Random permutations and representations of symmetric groups

Pierre-Loïc Méliot

Contents

CHAPTER 1

From random walks on symmetric groups to the theory of representations

Given an integer $N \geq 1$, we denote $[\![1, N]\!] = \{1, 2, \ldots, N\}$ the set of integers between 1 and N. We recall that the *symmetric group* of size N is the set $\mathfrak{S}(N)$ of all the bijections $\sigma : [\![1, N]\!] \mapsto [\![1, N]\!]$. The elements of $\mathfrak{S}(N)$ are called the *permutations* of size N; there are

$$N! = 1 \times 2 \times 3 \times \cdots \times N$$

permutations in $\mathfrak{S}(N)$. For instance, $\mathfrak{S}(3)$ contains 6 elements that can be described by the following *words*:

$$123,\ 132,\ 213,\ 231,\ 312,\ 321.$$

The word of a permutation σ of size N is the sequence $\sigma(1)\,\sigma(2)\,\sigma(3)\ldots\sigma(N)$.

In this first chapter, we consider random processes $(\sigma_n)_{n\in\mathbb{N}}$ with values in a fixed symmetric group $\mathfrak{S}(N)$. Our objective is to evaluate the mixing times of these Markov processes. The calculation of the powers of the generators of these Markov chains leads one to understand the structure of the group algebra $\mathbb{C}[\mathfrak{S}(N)]$; this is provided by the *representation theory of the finite groups*. Chapter 2 will focus on the representations of the symmetric groups, and in Chapter 3 we shall answer the question of the mixing times for two specific random walks on $\mathfrak{S}(N)$.

1. Several random walks on the symmetric groups

In order to walk randomly on the symmetric group $\mathfrak{S}(N)$, one can use its group structure and take products of « elementary » random permutations. Recall that given two permutations σ, τ in $\mathfrak{S}(N)$, their product $\sigma\tau$ is the bijection $[\![1, N]\!] \to [\![1, N]\!]$ which is the composed map $\sigma \circ \tau$. One way to understand this product is to represent permutations by *braid diagrams*. If $\sigma \in \mathfrak{S}(N)$, its braid diagram connects two rows of N numbered points by N braids, with the number i on the first row being connected to the number $\sigma(i)$ on the second row. Then, the braid diagram of $\sigma \circ \tau$ is obtained

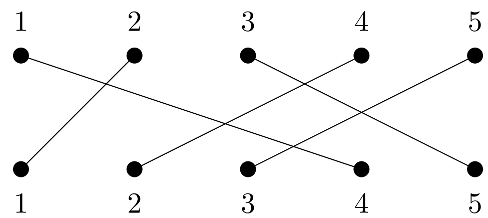


FIGURE 1. The braid diagram of the permutation $\sigma = 41523$.

by placing the braid diagram of σ *under* the braid diagram of τ. The product of permutations endows $\mathfrak{S}(N)$ with a group structure, the neutral element being the identity permutation $\mathrm{id} : i \mapsto i$, and the inverse σ^{-1} of a permutation σ being obtained by reversing its braid diagram (taking the symmetric diagram with respect to the axis of abscissa). Beware that the group structure is non-commutative: given two permutations σ and τ, in general we do not have $\sigma\tau = \tau\sigma$.

DEFINITION 1.1 (Random walk on $\mathfrak{S}(N)$). *Let μ be a probability distribution on $\mathfrak{S}(N)$. The random walk on $\mathfrak{S}(N)$ with generator μ is the sequence of random permutations $(\sigma_n)_{n\in\mathbb{N}}$ defined by*

$$\sigma_0 = \mathrm{id}_{[\![1,N]\!]} \qquad ; \qquad \sigma_{n+1} = \sigma_n \circ \tau_{n+1}$$

where $(\tau_n)_{n\geq 1}$ is a sequence of independent and identically distributed permutations chosen according to the distribution μ.

In a moment we shall present several interesting examples of generators μ. Let us first make two remarks. Since the τ_n's are independent and identically distributed, a random walk $(\sigma_n)_{n\in\mathbb{N}}$ defined in $\mathfrak{S}(N)$ as above is always a *Markov chain*, with state space $\mathfrak{S}(N)$, initial element $\sigma_0 = \mathrm{id}_{[\![1,N]\!]}$, and transition matrix $P(\sigma,\sigma') = \mu(\sigma^{-1}\sigma')$. On the other hand, the n-th element σ_n writes as the product $\tau_1\tau_2\cdots\tau_n$, so we would obtain the same distribution at each time n if we were taking products on the left, with a random process $(\sigma'_n)_{n\in\mathbb{N}}$ defined by $\sigma'_{n+1} = \tau_{n+1}\circ\sigma'_n$. Thus, for each n,

$$\sigma_n =_{(\text{distribution})} \sigma'_n.$$

However, the whole trajectories $(\sigma_n)_{n\in\mathbb{N}}$ and $(\sigma'_n)_{n\in\mathbb{N}}$ do not have the same distribution in $(\mathfrak{S}(N))^{\mathbb{N}}$. All the examples considered below are random walks according to Definition 1.1, with products on the right.

Example 1.2 (Products of random transpositions). Consider a deck of cards numbered by the integers in $[\![1,N]\!]$; when enumerating the cards from the top of the deck to the bottom, we obtain the word $\sigma(1)\,\sigma(2)\cdots\sigma(N)$ of a permutation $\sigma\in\mathfrak{S}(N)$. We consider the random walk on $\mathfrak{S}(N)$ where at each step, we take two random positions i and j in $[\![1,N]\!]$, and we exchange the positions of the i-th card of the deck and the j-th card of the deck. The two indices i and j are chosen uniformly and independently in $[\![1,N]\!]$; if $i=j$, we convene that the deck of cards stays unchanged. This construction corresponds to the following recurrence equation:

$$\sigma_{n+1} = \sigma_n\circ\tau_{n+1}\quad \text{with } \tau_{n+1} = \begin{cases}(i,j) & \text{with probability } \frac{2}{N^2} \text{ for any pair } 1\leq i<j\leq N,\\ \mathrm{id} & \text{with probability } \frac{1}{N}.\end{cases}$$

Here we take the usual notations for cycles in $\mathfrak{S}(N)$: $(c_1,c_2,\ldots,c_r)$ is the permutation which sends c_1 to c_2, c_2 to c_3, *etc.*, c_r to c_1 and which leaves the other elements invariant. In particular, (i,j) is the *transposition* which exchanges i and j and leaves the other elements invariant.

For reasons that will become evident later, it is convenient to represent the generator μ of this random walk as an element of the *group algebra* of $\mathfrak{S}(N)$. Consider more generally a finite group G, and denote $\mathbb{C}[G]$ or $\mathbb{C}G$ the complex vector space with dimension

$$\dim(\mathbb{C}G) = \mathrm{card}(G),$$

and with a basis $(e_g)_{g\in G}$ labelled by the elements of the group G. We say that $\mathbb{C}G$ is the group algebra of G; by definition this is already a vector space, and the algebra structure will be detailed in a moment. An element of $\mathbb{C}G$ is thus a linear combination with complex coefficients $\sum_{g\in G} c_g\, e_g$ of elements (labelled by the elements) of the group G. We shall simplify a bit these notations and denote the basis elements g instead of e_g; in general this does not create any confusion. So, for instance,

$$3\,[213] + (1+\mathrm{i})\,[312] - \mathrm{e}^{\pi}\,[132]$$

is an element of $\mathbb{C}\mathfrak{S}(3)$. A probability distribution μ on a finite group G can then be represented by the element

$$\sum_{g\in G}\mu(g)\,g$$

of the group algebra $\mathbb{C}G$. With these conventions, the random walk on $\mathfrak{S}(N)$ associated to the product of random transpositions has for generator

$$\mu = \frac{1}{N}\,\mathrm{id} + \frac{2}{N^2}\sum_{1\leq i<j\leq N}(i,j).$$

Example 1.3 (Top-with-random transpositions). Instead of exchanging two arbitrary cards at random, one can choose to exchange at each step of the random walk a random card i chosen uniformly in $[\![1,N]\!]$ with the top card 1. Again, we convene that if the random card i is $i=1$, then we leave the deck of cards invariant. The corresponding random process $(\sigma_n)_{n\in\mathbb{N}}$ is the random walk on $\mathfrak{S}(N)$ with generator

$$\mu = \frac{1}{N}\,\mathrm{id} + \frac{1}{N}\sum_{i=2}^{N}(1,i).$$

It is natural to expect that this random walk takes a bit more time to mix the cards (there is less randomness in each exchange). We shall see that in order to mix the deck of cards, one needs a time approximatively twice larger with the top-with-random shuffle than with the random transposition shuffle.

Example 1.4 (Top-to-random cycles). Another simple way to mix cards is to take at each step the card on top of the deck, and to place it at a random position i chosen uniformly in $[\![1,N]\!]$. Here we do not exchange positions: instead, we apply a cyclic permutation, sending the first card to the i-th position, and each card with position $j\in[\![2,i]\!]$ to the position $j-1$. This corresponds to the following generator:

$$\mu = \frac{1}{N}\,\mathrm{id} + \frac{1}{N}\sum_{i=2}^{N}(i,i-1,\ldots,2,1).$$

Example 1.5 (Adjacent transpositions). Yet another interesting way to mix the deck of cards is to exchange at each step the positions of two adjacent cards i and $i+1$. We convene to take i at random in $[\![1,N]\!]$, and to leave the deck invariant if $i=N$. The generator of the corresponding random walk on $\mathfrak{S}(N)$ is:

$$\mu = \frac{1}{N}\,\mathrm{id} + \frac{1}{N}\sum_{i=1}^{N-1}(i,i+1).$$

This generator is natural if we think of $\mathfrak{S}(N)$ as a group of reflections (Coxeter group). Denote $s_i = (i,i+1)$ the i-th adjacent transposition, also called *elementary transposition*. It is not very difficult to see that $(s_i)_{i\in[\![1,N-1]\!]}$ spans the group $\mathfrak{S}(N)$, and that we have the following *braid relations*:

$$\begin{aligned}(s_i)^2 &= \mathrm{id};\\ s_i s_j &= s_j s_i \text{ if } |j-i|\geq 2;\\ s_i s_{i+1} s_i &= s_{i+1} s_i s_{i+1}.\end{aligned}$$

It can be shown (this is not easy) that $\mathfrak{S}(N)$ is the largest group spanned by elements $s_{i\in[\![1,N-1]\!]}$ which satisfy the relations above: any group $H=\langle s_1',s_2',\ldots,s_{N-1}'\rangle$ with these properties is the surjective image of a morphism of groups $\mathfrak{S}(N)\to H$ sending s_i to s_i' for any i.

Example 1.6 (Riffle shuffle). As a last example, let us consider the way that card dealers in a casino use in order to mix the decks of cards. Given a deck of N cards, one first splits it in two halves (with a size of the two halves which is random), and then one interlaces randomly the cards of the left deck with the cards of the right deck. This algorithm of mixing is called the *riffle shuffle*. Let us be a bit more precise in its description.

(RS1) For the sizes of the left and right halves, a natural choice is to take i and $N-i$ with i following a binomial distribution $\mathcal{B}(N,\frac{1}{2})$; thus, the two halves have an expected number of cards equal to $\frac{N}{2}$.

(RS2) When interlacing the two halves, it is reasonable to make the following choice: if there are j_{left} and j_{right} remaining cards to interlace, then the next card comes from the left deck (respectively, the right deck) with probability $\frac{j_{\mathrm{left}}}{j_{\mathrm{left}}+j_{\mathrm{right}}}$ (respectively, $\frac{j_{\mathrm{right}}}{j_{\mathrm{left}}+j_{\mathrm{right}}}$).

If the sizes i and $N-i$ are fixed, then the number of possibilities for the interlacings of the two halves is $\binom{N}{i}$: one has to choose which are the i positions among N which will contain the cards of the left deck, and the other positions will necessarily contain the cards of the right deck. Then, it is easily seen by induction that, conditionnally to the value of i, any possible interlacing chosen according to (RS2) has probability

$$\frac{i!\,(N-i)!}{N!} = \frac{1}{\binom{N}{i}}.$$

Therefore, the generator of the random walk of the riffle shuffle is

$$\begin{aligned}\mu &= \sum_{i=0}^{N} \frac{1}{2^N}\binom{N}{i} \sum_{\substack{\sigma \text{ interlacing permutation of} \\ \text{two decks of size } i \text{ and } N-i}} \frac{1}{\binom{N}{i}}\,\sigma \\ &= \frac{1}{2^N}\sum_{i=0}^{N} \sum_{\substack{\sigma \text{ interlacing permutation of} \\ \text{two decks of size } i \text{ and } N-i}} \sigma.\end{aligned}$$

In order to explain more concretely what is meant by « interlacing permutation of two decks of size i and $N-i$ », we introduce the notion of *shuffle product*. The shuffle product $v \sqcup\mathchoice{\mkern-7mu}{\mkern-7mu}{\mkern-3.2mu}{\mkern-3.8mu}\sqcup w$ of two words $v = v_1v_2\ldots v_s$ and $w = w_1w_2\ldots w_t$ with disjoint sets of letters is the set of $\binom{s+t}{s}$ words z with length $s+t$ and such that, when reading from left to right the letters of z which are letters of v (respectively, which are letters of w), one obtains v (respectively, w). For instance,

$$123 \sqcup\mathchoice{\mkern-7mu}{\mkern-7mu}{\mkern-3.2mu}{\mkern-3.8mu}\sqcup 45 = \{12345, 12435, 14235, 41235, 12453, 14253, 41253, 14523, 41523, 45123\}.$$

The set of permutations of size 5 above can be represented by a formal sum of permutations in $\mathbb{C}\mathfrak{S}(5)$:

$$123 \sqcup\mathchoice{\mkern-7mu}{\mkern-7mu}{\mkern-3.2mu}{\mkern-3.8mu}\sqcup 45 = 12345 + 12435 + 14235 + 41235 + 12453 + 14253 + 41253 + 14523 + 41523 + 45123.$$

Then, with the same conventions in any size N, we have:

$$\left(\sum_{\substack{\sigma \text{ interlacing permutation of} \\ \text{two decks of size } i \text{ and } N-i}} \sigma\right) = \big(12\ldots i\big) \sqcup\mathchoice{\mkern-7mu}{\mkern-7mu}{\mkern-3.2mu}{\mkern-3.8mu}\sqcup \big((i+1)(i+2)\ldots N\big).$$

So, the generator of the riffle shuffle is

$$\mu = \frac{1}{2^N}\sum_{i=0}^{N} \big(12\ldots i\big) \sqcup\mathchoice{\mkern-7mu}{\mkern-7mu}{\mkern-3.2mu}{\mkern-3.8mu}\sqcup \big((i+1)(i+2)\ldots N\big).$$

In the sequel, we fix a random walk $(\sigma_n)_{n\in\mathbb{N}}$ with generator $\mu \in \mathbb{C}\mathfrak{S}(N)$, and we denote μ_n the distribution of σ_n, which we also represent by an element of $\mathbb{C}\mathfrak{S}(N)$. Since $\sigma_1 = \tau_1$, we have by definition $\mu_1 = \mu$. We can generalize this relation by making $\mathbb{C}\mathfrak{S}(N)$ into an algebra. More generally, given a finite group G, we define an internal product in $\mathbb{C}G$ as follows:

$$\left(\sum_{g\in G} c_g\, g\right)\left(\sum_{h\in G} d_h\, h\right) = \sum_{(g,h)\in G^2} c_g d_h\, gh,$$

where the product gh is the one of the group G. In other words, we expand the product from G to $\mathbb{C}G$ by bilinearity. This is the so-called *convolution product* of the group algebra. By gathering the terms in the basis of group elements of $\mathbb{C}G$, we can rewrite:

$$\left(\sum_{g\in G} c_g\, g\right)\left(\sum_{h\in G} d_h\, h\right) = \sum_{k\in G}\left(\sum_{(g,h)\,|\,gh=k} c_g d_h\right) k.$$

We leave the reader check that one obtains in this way a structure of algebra over $\mathbb{C}$ for $\mathbb{C}G$. Moreover, G is a commutative group if and only if $\mathbb{C}G$ is a commutative algebra (this is never the case for $G = \mathfrak{S}(N)$ with $N \geq 3$).

Proposition 1.7 (Marginales of a random walk on $\mathfrak{S}(N)$). *Given a random walk $(\sigma_n)_{n\in\mathbb{N}}$ in $\mathfrak{S}(N)$ with generator μ, the distribution μ_n of σ_n is given by*

$$\mu_n = (\mu)^n,$$

the power on the right-hand side being taken in $\mathbb{C}\mathfrak{S}(N)$.

Proof. We have by conditioning

$$\mathbb{P}[\sigma_{n+1} = \sigma] = \sum_{\rho\in\mathfrak{S}(N)} \mathbb{P}[\sigma_n = \rho]\, \mu(\rho^{-1}\sigma),$$

therefore,

$$\begin{aligned}
\mu_{n+1} &= \sum_{\sigma\in\mathfrak{S}(N)} \mathbb{P}[\sigma_{n+1} = \sigma]\, \sigma = \sum_{(\rho,\sigma)\in(\mathfrak{S}(N))^2} \mathbb{P}[\sigma_n = \rho]\, \mu(\rho^{-1}\sigma)\, \sigma \\
&= \sum_{(\rho,\sigma)\in(\mathfrak{S}(N))^2} \big(\mathbb{P}[\sigma_n = \rho]\, \rho\big)\, \big(\mu(\rho^{-1}\sigma)\, \rho^{-1}\sigma\big) \\
&= \sum_{(\rho,\sigma')\in(\mathfrak{S}(N))^2} \big(\mathbb{P}[\sigma_n = \rho]\, \rho\big)\, \big(\mu(\sigma')\, \sigma'\big) = \mu_n\, \mu
\end{aligned}$$

by making the change of variables $\sigma' = \rho^{-1}\sigma$, which is valid since $\sigma \mapsto \rho^{-1}\sigma$ is a permutation of the $N!$ elements of $\mathfrak{S}(N)$. The relation $\mu_{n+1} = \mu_n\, \mu$ in $\mathbb{C}\mathfrak{S}(N)$ enables one to prove the result by induction. □

Here we see clearly the interest of the notion of group algebra, and this object will be one of the main tool of our analysis of the random walks. Let us now state the main question around random walks on the symmetric groups. As we consider a Markov chain on the finite state space $\mathfrak{S}(N)$, it is natural to study the asymptotic behavior of the marginal distributions μ_n, and the speed of convergence of these distributions towards an invariant measure. The invariant measure is almost always the uniform distribution on $\mathfrak{S}(N)$:

Proposition 1.8 (Ergodicity of a random walk on $\mathfrak{S}(N)$). *Given a random walk $(\sigma_n)_{n\in\mathbb{N}}$ in $\mathfrak{S}(N)$ with generator μ, suppose that the two following conditions are satisfied:*

(1) The identity permutation $\mathrm{id}_{[\![1,N]\!]}$ has positive probability under μ.

(2) The set of permutations σ which have positive probability under μ spans the group $\mathfrak{S}(N)$.

Then, $\mu_n \rightharpoonup_{n\to\infty}$ uniform measure.

Proof. The first item ensures that the Markov chain is aperiodic, and the second item that it is irreducible over the state space $\mathfrak{S}(N)$. Therefore, there exists a unique invariant probability measure ν for the Markov chain, and $\mu_n \rightharpoonup_{n\to\infty} \nu$. The recurrence relation $\mu_{n+1} = \mu_n\mu$ in $\mathbb{C}\mathfrak{S}(N)$ leads to the equation:

$$\nu = \nu\mu.$$

Here it is legit to take limits, since we are working in a finite-dimensional algebra, with a product which is obviously continuous. However, $\nu = \frac{1}{N!}\sum_{\sigma\in\mathfrak{S}(N)} \sigma$ clearly satisfies the equation above:

$$\nu\mu = \sum_{(\sigma,\tau)\in(\mathfrak{S}(N))^2} \frac{\mu(\tau)}{N!}\,\sigma\tau = \sum_{\rho\in\mathfrak{S}(N)} \left(\sum_{\tau\in\mathfrak{S}(N)} \frac{\mu(\tau)}{N!}\right)\rho = \frac{1}{N!}\sum_{\rho\in\mathfrak{S}(N)} \rho = \nu.$$

By unicity, we conclude that the invariant probability measure is the uniform measure. □

The uniform measure $\frac{1}{|G|}\sum_{g\in G} g$ on a finite group G is also called the *Haar measure* of the group; the notion expands to any *compact topological group* (not necessarily finite), and many arguments hereafter can be expanded to this setting. All the examples of random walks on $\mathfrak{S}(N)$ presented before satisfy the hypotheses of Proposition 1.8 (exercise: the first condition is trivial, and for the second condition one has to prove elementary facts such as: every permutation in $\mathfrak{S}(N)$ is the product of transpositions / of transpositions of the form $(1,i)$ / of elementary transpositions). Therefore, they all admit for invariant measure $\mu_\infty = \mathrm{Haar}$.

Problem. *Estimate the speed of convergence $\mu_n \rightharpoonup_{n\to\infty} \mu_\infty$ for each of the random walks previously introduced (the answer will depend on the size N of the symmetric group).*

In order to speak of the speed of convergence, we need to introduce a distance between probability measures on $\mathfrak{S}(N)$, or more generally on a finite group G. There are many possible choices, but one of the most useful distance is the *total variation distance*:

$$d_{\mathrm{TV}}(\mu,\nu) = \sup_{A\subset G} |\mu(A) - \nu(A)| = \frac{1}{2}\sum_{g\in G} |\mu(g) - \nu(g)|.$$

When one of the two probability measures is the Haar measure, this distance can be related to the norm of the Banach space $\mathscr{L}^1(G,\mathrm{Haar})$ (we shall also denote the Haar measure dg). Indeed,

$$d_{\mathrm{TV}}(\mu,\mathrm{Haar}) = \frac{1}{2}\sum_{g\in G}\left|\mu(g) - \frac{1}{|G|}\right| = \frac{1}{2}\int_G |f-1|\;dg = \frac{1}{2}\,\|f-1\|_{\mathscr{L}^1(G,\,dg)}$$

where $f = \frac{d\mu}{dg}$ is the density of μ with respect to the Haar measure. In particular, if we are interested in an upper bound on the total variation distance, then we can use the Cauchy–Schwarz inequality:

$$d_{\mathrm{TV}}(\mu,\mathrm{Haar}) \leq \frac{1}{2}\|f-1\|_{\mathscr{L}^2(G,\,dg)}.$$

In the following, we shall introduce tools which allow the computation of $\mathscr{L}^2$-norms of functions on a finite group, and which are particularly adapted to the case where $f = f_n$ is the density of a measure $\mu_n = \mu^n$ which is a convolution power of a generator. A related problem is:

Problem. *Consider a finite group G, and an element μ of its group algebra $\mathbb{C}G$ (*e.g.*, a probability measure). How can one compute efficiently the convolution powers μ^n of this element?*

In the setting of finite Markov chains, it is well known that the computation of the distribution at time n is equivalent to the computation of the powers of the transition matrix, and therefore to the diagonalisation of this matrix. One can guess that the problem above is of similar nature, and we shall see in the sequel that actually, it is entirely equivalent to the computation of powers of a finite collection of matrices.

2. The category of representations of a finite group

In this section, we fix a finite group G; if V is a complex finite-dimensional vector space, we denote $\mathrm{End}(V)$ the complex algebra of linear endomorphisms of V, and $\mathrm{GL}(V)$ the group of linear isomorphisms of V.

DEFINITION 1.9 (Representation of a group). *A (complex, linear, finite-dimensional) representation of a finite group G is a pair (V, ρ), where V is a complex finite-dimensional vector space, and $\rho : G \to \mathrm{GL}(V)$ is a morphism of groups. This defines an action of G on V:*

$$g \cdot v = \rho(g)(v).$$

Given a representation of G on a space V, we have a natural bilinear map

$$\begin{aligned} \mathbb{C}G \times V &\to V \\ \sum_{g \in G} c_g\, g,\ v &\mapsto \sum_{g \in G} c_g\, (g \cdot v), \end{aligned}$$

and one checks readily that for any $v \in V$ and any elements $x, y \in \mathbb{C}G$, we have $x \cdot (y \cdot v) = (xy) \cdot v$, the product xy being taken in the group algebra $\mathbb{C}G$. Therefore, a representation of G on a space V is equivalent to a structure of $\mathbb{C}G$-module for V.

Given two representations (V_1, ρ_1) and (V_2, ρ_2) of a group G, a morphism of representations (respectively, an isomorphism of representations) $\phi : V_1 \to V_2$ is a linear map (respectively, a bijective linear map) which is compatible with the action of the group: for any $g \in G$, the following diagram of linear maps is commutative:

$$\begin{array}{ccc} V_1 & \xrightarrow{\rho_1(g)} & V_1 \\ \phi \big\downarrow & & \big\downarrow \phi \\ V_2 & \xrightarrow[\rho_2(g)]{} & V_2. \end{array}$$

With the language of modules, a morphism of representations between (V_1, ρ_1) and (V_2, ρ_2) is simply of morphism of $\mathbb{C}G$-modules. Our objective is to describe all the representations of a finite group G *up to isomorphisms*. The following trivial example makes this objective more precise:

EXAMPLE 1.10 (Category of representations of the trivial group). Suppose that $G = \{1\}$ is the trivial group with one element. Then, a representation of G is simply a finite-dimensional complex vector space V, the action of the group being $1 \cdot v = v$ for any $v \in V$. So,

$$\mathbf{Representations}(\{1\}) = \mathbf{Vector\,spaces}_{\mathbb{C}}.$$

Any finite-dimensional vector space can be decomposed as a direct sum of lines: $V = \bigoplus_{i=1}^{\dim V} \mathbb{C}$, the symbol $=$ meaning here that there exists an isomorphism. Therefore, there is a unique « building block » for the representations of the trivial group, which is the one-dimensional complex line; and any representation of $\{1\}$ is a direct sum of such building blocks. We shall see that this result expands to the category of representations of any finite group G, the main difference being that there exists in the general case a finite number of possible building blocks (as many as the number of conjugacy classes in G).

Given two representations (V_1, ρ_1) and (V_2, ρ_2) of G, their *direct sum* is the representation of G whose underlying space is $V = V_1 \oplus V_2$, and whose defining morphism $\rho : G \to \mathrm{GL}(V)$ is given by:

$$(\rho(g))(v_1 + v_2) = (\rho_1(g))(v_1) + (\rho_2(g))(v_2).$$

Hence, in a basis of V adapted to the decomposition $V = V_1 \oplus V_2$, the matrix of $\rho(g)$ is the block-diagonal matrix

$$\begin{pmatrix} \rho_1(g) & 0 \\ 0 & \rho_2(g) \end{pmatrix}.$$

On the other hand, given a representation (V, ρ) of G, a *subrepresentation* of G is a vector subspace $W \subset V$ which is left invariant by the action of G: for any $w \in W$ and $g \in G$, $g \cdot w \in W$. In other words, W is a $\mathbb{C}G$-submodule of V. One says that (V, ρ) is an *irreducible* representation of G if $\dim V > 0$ and if the only subrepresentations of V are 0 and V itself.

THEOREM 1.11 (Maschke). *If $W \subset V$ is a subrepresentation of a representation of G, then there exists another subrepresentation $W' \subset V$ such that $V = W \oplus W'$. Any representation of a finite group G can be split in a direct sum of irreducible representations:*

$$V = \bigoplus_{i=1}^{r} V_i, \quad \text{with the } V_i\text{'s irreducible representations of } G.$$

PROOF. The theorem relies on the following important fact: for any representation (V, ρ) of a finite group G, there exists a Hermitian scalar product $\langle \cdot \mid \cdot \rangle_V$ on V such that $\rho(G)$ is included in the unitary group $\mathrm{U}(V)$ with respect to this scalar product. In other words, any element $g \in G$ acts on V by isometry with respect to this scalar product. The construction of $\langle \cdot \mid \cdot \rangle_V$ is as follows: we start from an arbitrary scalar product $(\cdot \mid \cdot)$ on V, and we define $\langle \cdot \mid \cdot \rangle_V$ by

$$\langle v_1 \mid v_2 \rangle_V = \frac{1}{|G|} \sum_{g \in G} (g \cdot v_1 \mid g \cdot v_2).$$

Then, any $h \in G$ acts indeed by isometry with respect to $\langle \cdot \mid \cdot \rangle_V$:

$$\langle h \cdot v_1 \mid h \cdot v_2 \rangle_V = \frac{1}{|G|} \sum_{g \in G} (gh \cdot v_1 \mid gh \cdot v_2) = \frac{1}{|G|} \sum_{g' \in G} (g' \cdot v_1 \mid g' \cdot v_2) = \langle v_1 \mid v_2 \rangle_V.$$

the map $g \mapsto gh$ being a permutation of the elements of G.

Suppose given a subrepresentation W of V. Then, its orthogonal $W^\perp$ with respect to $\langle \cdot \mid \cdot \rangle_V$ is also a subrepresentation of V: if $z \in W^\perp$ and $w \in W$, then for any $g \in G$,

$$\langle g \cdot z \mid w \rangle_V = \left\langle z \mid g^{-1} \cdot w \right\rangle_V = 0 \quad \text{since } z \in W^\perp \text{ and } g^{-1} \cdot w \in W.$$

As this is true for any $w \in W$, $g \cdot z \in W^\perp$. Thus, we have shown that $W^\perp$ is a subrepresentation of V, and we have of course $V = W \oplus W^\perp$. This proves the first part of the theorem, and the second part follows immediately by induction on the dimension of V (if V is not irreducible, then we can split it in smaller subrepresentations for which the induction hypothesis applies). ☐

DEFINITION 1.12 (Dual of a finite group). *The dual $\widehat{G}$ of a finite group G is the set*

$$\widehat{G} = \{\text{isomorphism classes of irreducible representations } \lambda = (V^\lambda, \rho^\lambda) \text{ of } G\}.$$

Maschke's theorem 1.11 ensures that the elements of $\widehat{G}$ are the building blocks of the category of representations of G: any representation of G is a direct sum of a finite number of elements of $\widehat{G}$. To complete our understanding of this, we are going to show that $\widehat{G}$ is a finite set, and that a decomposition of a representation V in irreducibles is unique (up to isomorphisms). To this purpose, we shall study the space of morphisms between representations of G:

$$\hom_G((V_1, \rho_1), (V_2, \rho_2)) = \{\text{morphisms of representations } \phi : V_1 \to V_2\}.$$

To simplify a bit the notations, in the following we omit the defining morphisms ρ_1 and ρ_2 and we simply denote $\hom_G(V_1, V_2)$. This set is clearly a complex vector space: any linear combination of morphisms of representations is again a morphism of representations.

PROPOSITION 1.13 (Schur's lemma). *The map* $\hom_G(\cdot,\cdot)$ *is additive on the left: for any representations* V, W *and* Z, $\hom_G(V\oplus W,Z)=\hom_G(V,Z)\oplus\hom_G(W,Z)$. *On the other hand, given two irreducible representations* V^λ *and* V^μ, *we have:*

$$\dim\big(\hom_G(V^\lambda,V^\mu)\big)=\begin{cases}1 & \text{if the representation } V^\lambda \text{ is isomorphic to } V^\mu,\\ 0 & \text{otherwise.}\end{cases}$$

PROOF. The additivity on the left is obvious: if $\phi:V\oplus W\to Z$ is a morphism of representations, then its restrictions $\phi_{|V}:V\to Z$ and $\phi_{|W}:W\to Z$ are also morphisms of representations, and $\phi=\phi_{|V}+\phi_{|W}$, this decomposition being unique. Let us now consider the space of morphisms between two irreducible representations V^λ and V^μ. If $\phi\in\hom_G(V^\lambda,V^\mu)$, then its kernel and its image are easily seen to be subrepresentations respectively of V^λ and of V^μ. As V^λ and V^μ are irreducible, we conclude that if $\phi\neq 0$, then $\ker\phi=0$ and $\operatorname{Im}\phi=V^\mu$, so ϕ is an isomorphism. In particular,

$$\dim\big(\hom_G(V^\lambda,V^\mu)\big)=0\ \text{ if the representation } V^\lambda \text{ is not isomorphic to } V^\mu.$$

Suppose now that V^λ and V^μ are isomorphic by some morphism of representations ψ. We then have an isomorphism of vector spaces:

$$\begin{aligned}\hom_G(V^\lambda,V^\lambda)&\to\hom_G(V^\lambda,V^\mu)\\ \phi&\mapsto\psi\circ\phi.\end{aligned}$$

Therefore, in order to compute $\dim(\hom_G(V^\lambda,V^\mu))$, we can assume that $V^\lambda=V^\mu$. Consider then an arbitrary morphism $\phi\in\hom_G(V^\lambda,V^\lambda)$; since we are working with complex vector spaces, there exists a complex eigenvalue $z\in\mathbb{C}$ such that $\phi-z\,\mathrm{id}_{V^\lambda}$ is not invertible. Then, $\phi-z\,\mathrm{id}_{V^\lambda}$ is an endomorphism of representations of V^λ which is not an isomorphism, so by the previous argument it is equal to 0, and

$$\phi=z\,\mathrm{id}_{V^\lambda}.$$

We conclude that $\hom_G(V^\lambda,V^\lambda)=\mathbb{C}\,\mathrm{id}_{V^\lambda}$ and that the dimension of the space of morphisms is in this case equal to 1. □

THEOREM 1.14 (Uniqueness of the decomposition in irreducibles). *Consider a representation* V *of a finite group* G, *which we split in irreducibles:*

$$V=\bigoplus_{\lambda\in\widehat{G}}m_\lambda\,V^\lambda,$$

The m_λ*'s are multiplicities in* $\mathbb{N}$*: by* $m_\lambda\,V^\lambda$, *we mean a direct sum of* m_λ *copies of the irreducible representation* V^λ. *The decomposition of* V *is unique: for each isomorphism class* $\lambda=(V^\lambda,\rho^\lambda)$ *in* $\widehat{G}$, *the multiplicity* m_λ *is entirely determined by* V.

PROOF. We have $m_\lambda=\dim(\hom_G(V,V^\lambda))$. □

THEOREM 1.15 (Finiteness of the dual $\widehat{G}$). *For any finite group* G, *its dual* $\widehat{G}$ *is finite: there only exists a finite number of isomorphism classes of irreducible representations of* G.

PROOF. Consider the vector space $V=\mathbb{C}G$. It is a representation of G for the action:

$$g\cdot\left(\sum_{h\in G}c_h\,h\right)=\sum_{h\in G}c_h\,(gh).$$

Given another representation W, we introduce the following linear map:

$$\begin{aligned}\Psi : \hom_G(\mathbb{C}G, W) &\to W\\ \phi &\mapsto \phi(e_G),\end{aligned}$$

where e_G denotes the neutral element of the group G. We claim that Ψ is an isomorphism of vector spaces. Note that for any element $x = \sum_{g\in G} c_g\, g$ in $\mathbb{C}G$, we have

$$\phi(x) = \sum_{g\in G} c_g\, \phi(g) = \left(\sum_{g\in G} c_g\, g\right) \cdot \phi(e_G) = x \cdot \Psi(\phi).$$

Consequently, if ϕ belongs to $\ker \Psi$, then $\phi(x) = 0$ for any $x \in \mathbb{C}G$, so $\phi = 0$; thus, Ψ is injective. Now, if $v \in V$, then the map

$$\begin{aligned}\phi : \mathbb{C}G &\to v\\ x &\mapsto x \cdot v\end{aligned}$$

defines a morphism of representations, and we have $\Psi(\phi) = \phi(e_G) = e_G \cdot v = v$; so Ψ is also surjective.

We conclude that for any representation W of G, we have $\hom_G(\mathbb{C}G, W) = W$. In particular, given any element $\lambda \in \widehat{G}$, $\dim(\hom_G(\mathbb{C}G, V^\lambda)) = \dim V^\lambda$. By Theorem 1.14 and its proof, we see that any irreducible representation of G occurs as a component of the *regular representation* $\mathbb{C}G$, and that we have the decomposition:

$$\mathbb{C}G = \bigoplus_{\lambda\in\widehat{G}} (\dim \lambda)\, V^\lambda,$$

where $\dim \lambda = \dim(V^\lambda)$. If we take the dimensions, we obtain the important formula

$$\mathrm{card}(G) = \sum_{\lambda\in\widehat{G}} (\dim \lambda)^2$$

and this implies in particular that $\widehat{G}$ is a finite set. □

So, to summarise the discussion above: given a finite group G, there exists a finite set $\widehat{G}$ of non-isomorphic irreducible representations $\lambda = (V^\lambda, \rho^\lambda)$, such that any representation V of G writes uniquely as

$$V = \bigoplus_{\lambda\in\widehat{G}} m_\lambda\, V^\lambda$$

with the multiplicities $m_\lambda \in \mathbb{N}$. Moreover, the regular representation $\mathbb{C}G$ « contains » all the irreducible representations, each V^λ occurring $\dim \lambda$ times as a subrepresentation of $\mathbb{C}G$.

Example 1.16 (Representations of $\mathfrak{S}(3)$). Let us describe the representations of the symmetric group of size 3. We first have two irreducible representations with dimension 1:

- the *trivial representation* of $\mathfrak{S}(3)$ on $\mathbb{C}$, with $\sigma \cdot v = v$ for any $v \in \mathbb{C}$. We shall denote this representation V_{trivial}.
- the *signature representation* of $\mathfrak{S}(3)$ on $\mathbb{C}$, with $\sigma \cdot v = \varepsilon(\sigma)\, v$, $\varepsilon(\sigma)$ being the signature of a permutation (the parity of the number of inversions of the permutation). We shall denote this representation V_{sign}.

It is clear that V_{trivial} and V_{sign} are not isomorphic: a morphism of representations $\phi : V_{\mathrm{sign}} \to V_{\mathrm{trivial}}$ must satisfy

$$-\phi(v) = \phi((1,2) \cdot v) = (1,2) \cdot \phi(v) = \phi(v),$$

so $\phi(v) = 0$ for any v, and $\phi = 0$. We have thus identified two elements of $\widehat{\mathfrak{S}(3)}$. Let us exhibit a third irreducible representation, this time with dimension 2. The *geometric representation* of $\mathfrak{S}(3)$ on $\mathbb{C}^3$ is given by

$$\sigma \cdot (x_1, x_2, x_3) = (x_{\sigma^{-1}(1)}, x_{\sigma^{-1}(2)}, x_{\sigma^{-1}(3)}).$$

This 3-dimensional representation is not irreducible, because it admits for subrepresentation the one-dimensional vector space $\mathbb{C}(1, 1, 1)$, on which permutations act trivially. The standard scalar product $\langle x \mid y\rangle = \overline{x_1}y_1 + \overline{x_2}y_2 + \overline{x_3}y_3$ is compatible with the action of the symmetric group on $\mathbb{C}^3$, and the orthogonal of $\mathbb{C}(1, 1, 1)$ with respect to this scalar product is the space

$$W = \{(x_1, x_2, x_3) \in \mathbb{C}^3 \mid x_1 + x_2 + x_3 = 0\}.$$

So, we have a decomposition in subrepresentations:

$$\mathbb{C}^3 = \mathbb{C}(1, 1, 1) \oplus W,$$

and the first component is an irreducible representation isomorphic to V_{trivial}. To see that the other component is irreducible, one can use the following geometric argument (another argument and a generalisation of this fact will be given much later by Proposition 3.18). Given the non-zero vector $\alpha = (1, -1, 0) \in W$, its images by the six permutations in $\mathfrak{S}(3)$ are arranged according to a hexagon in the plane W. The elementary transpositions $s_1 = (1, 2)$ and $s_2 = (2, 3)$ act on the plane by reflections, see Figure 2. Let W' be a non-zero subrepresentation of W, and $w \neq 0$ a vector of

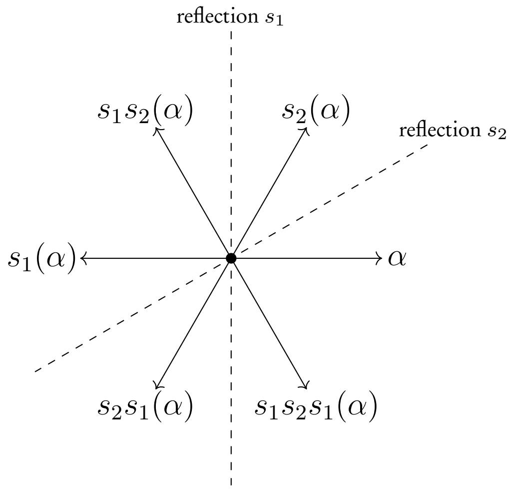


FIGURE 2. Action of $\mathfrak{S}(3)$ on the 2-dimensional irreducible representation W; $s_1 = (1, 2)$ and $s_2 = (2, 3)$ are the elementary transpositions which span the group.

this subrepresentation.

- If w does not belong to the axis of the reflection s_1, then w and $s_1(w)$ are not colinear, so W' contains a linear basis of W, and $W' = W$.
- If w belongs to the axis of s_1, then it does not belong to the axis of s_2, so we can use the same argument as above with w and $s_2(w)$.

We have thus proved that W is an irreducible representation of $\mathfrak{S}(3)$ with dimension 2, and for dimension reasons, it is not isomorphic to V_{trivial} or V_{sign}. Finally, since

$$\mathrm{card}(\mathfrak{S}(3)) = 6 = 1^2 + 1^2 + 2^2,$$

the formula for the sum of the squares of dimensions shows that we have found all the irreducible representations of $\mathfrak{S}(3)$:

$$\widehat{\mathfrak{S}(3)} = \{V_{\mathrm{trivial}}, V_{\mathrm{sign}}, W\}.$$

Any representation of the group $\mathfrak{S}(3)$ writes as a direct sum $aV_{\mathrm{trivial}} \oplus bV_{\mathrm{sign}} \oplus cW$ with multiplicities $(a, b, c) \in \mathbb{N}^3$.

3. The non-commutative Fourier transform

Let us restate a result which we obtained during the proof of Theorem 1.15:

THEOREM 1.17 (Plancherel formula of the squares). *For any finite group G,*

$$\mathrm{card}(G) = \sum_{\lambda \in \widehat{G}} (\dim \lambda)^2.$$

On the left-hand side, $\mathrm{card}(G)$ is the dimension of the group algebra $\mathbb{C}G$, whereas on the right-hand side, each term $(\dim \lambda)^2$ is the dimension of the endomorphism algebra $\mathrm{End}(V^\lambda)$. Therefore, it is tempting to compare the two complex algebras

$$\mathbb{C}G \quad \text{and} \quad \bigoplus_{\lambda \in \widehat{G}} \mathrm{End}(V^\lambda).$$

We are going to show that these two algebras are isomorphic, and also isometric for two adequate scalar products defined on them. Before stating precisely this result, let us recall a similar result which stems from the classical Fourier analysis on the circle or on the torus.

EXAMPLE 1.18 (Fourier analysis on the torus). Denote $\mathbb{T}^d = \mathbb{R}^d/(2\pi\mathbb{Z})^d$ the torus with dimension d; its elements consists in vectors $\theta = (\theta_1, \theta_2, \ldots, \theta_d)$ of real numbers modulo 2π. The torus $\mathbb{T}^d$ is a group for the addition of vectors modulo 2π on each coordinate. It is endowed with the probability measure

$$\frac{d\theta}{(2\pi)^d} = \frac{d\theta_1}{2\pi} \frac{d\theta_2}{2\pi} \cdots \frac{d\theta_d}{2\pi}.$$

Given a function $f \in \mathscr{L}^2(\mathbb{T}^d, \frac{d\theta}{(2\pi)^d})$ and a vector $k = (k_1, \ldots, k_d) \in \mathbb{Z}^d$, the Fourier coefficient $\widehat{f}(k)$ is defined by the integral:

$$\widehat{f}(k) = \frac{1}{(2\pi)^d} \int_{[0,2\pi]^d} f(\theta_1, \ldots, \theta_d)\, \mathrm{e}^{\mathrm{i}(k_1\theta_1 + \cdots + k_d\theta_d)}\, d\theta.$$

The well-known Parseval formula ensures that:

$$\|f\|^2_{\mathscr{L}^2(\mathbb{T}^d)} = \int_{\mathbb{T}^d} |f(\theta)|^2 \frac{d\theta}{(2\pi^d)} = \sum_{k \in \mathbb{Z}^d} |\widehat{f}(k)|^2 = \|\widehat{f}\|^2_{\ell^2(\mathbb{Z}^d)}.$$

Thus, $f \mapsto \widehat{f}$ is an isometry between the two Hilbert spaces $\mathscr{L}^2(\mathbb{T}^d)$ and $\ell^2(\mathbb{Z}^d)$. It is also an isomorphism of algebras with respect to the following products:

- the convolution product in $\mathscr{L}^2(\mathbb{T}^d)$:

$$(f * g)(\theta) = \int_{\mathbb{T}^d} f(\theta - \phi)\, g(\phi)\, \frac{d\phi}{(2\pi)^d}.$$

- the pointwise product in $\ell^2(\mathbb{Z}^d)$.

Indeed, it is well known (and easy to prove) that $\widehat{(f * g)}(k) = \widehat{f}(k)\, \widehat{g}(k)$. So, in a very strong sense, $\mathscr{L}^2(\mathbb{T}^d) = \ell^2(\mathbb{Z}^d)$. The goal of this last section of the chapter is to establish an analogous result with instead of $\mathbb{T}^d$ a finite group G, and instead of $\mathbb{Z}^d$ the dual $\widehat{G}$.

Let us introduce a few notations. First, if f is an element of $\mathbb{C}G$, then we can consider it as a function $G \to \mathbb{C}$, by denoting

$$f = \sum_{g \in G} f(g)\, g.$$

So, complex-valued functions on G and formal complex linear combinations of elements of the group G are the same objects. The functional viewpoint provides $\mathbb{C}G$ with a natural scalar product:

$$\langle f_1 \mid f_2 \rangle = \int_G \overline{f_1(g)}\, f_2(g)\, dg = \frac{1}{|G|} \sum_{g \in G} \overline{f_1(g)}\, f_2(g).$$

So, as a Hilbert space, $\mathbb{C}G = \mathscr{L}^2(G,\, dg)$. On the dual side, we denote $\mathbb{C}\widehat{G} = \bigoplus_{\lambda \in \widehat{G}} \mathrm{End}(V^\lambda)$; this is a direct sum of complex algebras, hence also a complex algebra. We endow each space $\mathrm{End}(V^\lambda)$ with the following Hermitian scalar product:

$$\langle u_1 \mid u_2 \rangle_{\mathrm{End}(V^\lambda)} = \frac{(\dim \lambda)}{|G|^2}\, \mathrm{tr}(u_1^*\, u_2),$$

where the adjoint u^* of an endomorphism u is taken with respect to a G-invariant scalar product on V^λ (*cf.* the proof of Theorem 1.11). We then equip $\mathbb{C}\widehat{G}$ with the unique scalar product whose restriction to an endomorphism space $\mathrm{End}(V^\lambda)$ is $\langle \cdot \mid \cdot \rangle_{\mathrm{End}(V^\lambda)}$, and which makes all the endomorphism spaces $\mathrm{End}(V^\lambda)$ orthogonal. For the moment, one might think that there are several possibilities for the G-invariant scalar products $\langle \cdot \mid \cdot \rangle_{V^\lambda}$ on V^λ, and therefore for the scalar product on $\mathbb{C}\widehat{G}$. In fact, Theorem 1.21 below will imply that the scalar products $\langle \cdot \mid \cdot \rangle_{\mathrm{End}(V^\lambda)}$ are unique.

DEFINITION 1.19 (Non-commutative Fourier transform). *For $f \in \mathbb{C}G$ and $\lambda = (V^\lambda, \rho^\lambda) \in \widehat{G}$, we define the non-commutative Fourier transform of f at λ by:*

$$\widehat{f}(\lambda) = \sum_{g \in G} f(g)\, \rho^\lambda(g).$$

This is a linear combination of elements of $\mathrm{GL}(V^\lambda)$*, hence an endomorphism in* $\mathrm{End}(V^\lambda)$*. The collection of endomorphisms $\widehat{f} = (\widehat{f}(\lambda))_{\lambda \in \widehat{G}}$ can be considered as an element of* $\bigoplus_{\lambda \in \widehat{G}} \mathrm{End}(V^\lambda) = \mathbb{C}\widehat{G}$*.*

REMARK 1.20. The definition of $\widehat{f}(\lambda)$ is natural if we think of representations as $\mathbb{C}G$-modules; then, $\widehat{f}(\lambda)$ is simply the endomorphism of V^λ induced by f. This leads one to wonder whether the theorem below can be extended to more general complex algebras than the finite group algebras $\mathbb{C}G$. The answer is positive for the class of *semisimple complex algebras*, which are the algebras isomorphic to direct sums of matrix algebras.

THEOREM 1.21 (Peter–Weyl theorem for finite groups). *The map $f \in \mathbb{C}G \mapsto \widehat{f} \in \mathbb{C}\widehat{G}$ is an isomorphism of complex algebras and an isometry of Hilbert spaces.*

The easy part of the theorem is the compatibility of the Fourier transform with products. Indeed, since ρ^λ is a morphism of groups, one computes readily:

$$\begin{aligned}
(\widehat{f_1 * f_2})(\lambda) &= \sum_{k \in G} (f_1 * f_2)(k)\, \rho^\lambda(k) = \sum_{(g,h) \in G^2} f_1(g) f_2(h)\, \rho^\lambda(gh) \\
&= \left(\sum_{g \in G} f_1(g)\, \rho^\lambda(g) \right) \left(\sum_{h \in G} f_2(h)\, \rho^\lambda(h) \right) = \widehat{f_1}(\lambda)\, \widehat{f_2}(\lambda).
\end{aligned}$$

Therefore, it remains to prove that the Fourier transform is an isometry. Given $\lambda \in \widehat{G}$, we fix an orthonormal basis $(e_1^\lambda, \ldots, e_{\dim \lambda}^\lambda)$ of the representation space V^λ (with respect to a G-invariant scalar product), and we write the matrices of the maps $\rho^\lambda(g)$ in this basis:

$$(\rho^\lambda(g))(e_j^\lambda) = \sum_{i=1}^{\dim \lambda} \rho_{ij}^\lambda(g)\, e_i^\lambda.$$

LEMMA 1.22 (Orthogonality of the matrix coefficients). *The matrix coefficients ρ^λ_{ij} with λ running over $\widehat{G}$ and $(i,j) \in [\![1, \dim\lambda]\!]^2$ form an orthogonal basis of the space $\mathbb{C}G$, with*

$$(\|\rho^\lambda_{ij}\|_{\mathbb{C}G})^2 = \frac{1}{\dim\lambda}.$$

PROOF. By Theorem 1.17, the family of functions $(\rho^\lambda_{ij})_{\lambda\in\widehat{G},\,1\leq i,j\leq \dim\lambda}$ has the correct cardinality for a basis of $\mathbb{C}G$. We have $\rho^\lambda_{ij}(g) = \langle e^\lambda_i \mid g\cdot e^\lambda_j\rangle_{V^\lambda}$, and

$$\begin{aligned}\langle \rho^\lambda_{ij} \mid \rho^\mu_{kl}\rangle_{\mathbb{C}G} &= \frac{1}{|G|}\sum_{g\in G} \overline{\langle e^\lambda_i \mid g\cdot e^\lambda_j\rangle_{V^\lambda}}\,\langle e^\mu_k \mid g\cdot e^\mu_l\rangle_{V^\mu} \\ &= \frac{1}{|G|}\sum_{g\in G} \langle e^\mu_k \mid g\cdot e^\mu_l\rangle_{V^\mu}\,\langle g\cdot e^\lambda_j \mid e^\lambda_i\rangle_{V^\lambda}.\end{aligned}$$

The indices j and l being fixed, we consider the linear map $\phi : V^\lambda \to V^\mu$ defined by

$$\phi(v) = \frac{1}{|G|}\sum_{g\in G}(g\cdot e^\mu_l)\,\langle g\cdot e^\lambda_j \mid v\rangle_{V^\lambda}.$$

We claim that ϕ is a morphism of representations. Indeed, we have:

$$\begin{aligned}\phi(h\cdot v) &= \frac{1}{|G|}\sum_{g\in G}(g\cdot e^\mu_l)\,\langle g\cdot e^\lambda_j \mid h\cdot v\rangle_{V^\lambda} = \frac{1}{|G|}\sum_{g\in G}(g\cdot e^\mu_l)\,\langle h^{-1}g\cdot e^\lambda_j \mid v\rangle_{V^\lambda} \\ &= h\cdot\left(\frac{1}{|G|}\sum_{g\in G}(h^{-1}g\cdot e^\mu_l)\,\langle h^{-1}g\cdot e^\lambda_j \mid v\rangle_{V^\lambda}\right) = h\cdot(\phi(v)).\end{aligned}$$

By Schur's lemma, if $\lambda\neq\mu$, then $\phi = 0$, and therefore $\langle \rho^\lambda_{ij} \mid \rho^\mu_{kl}\rangle_{\mathbb{C}G} = \langle e^\mu_k \mid \phi(e^\lambda_i)\rangle_{V^\mu} = 0$. Suppose now that $\lambda = \mu$; in this case, Schur's lemma ensures that ϕ is a scalar multiple of the identity map. To find which scalar multiple of the identity ϕ is, we compute its trace:

$$\begin{aligned}\mathrm{tr}(\phi) &= \sum_{i=1}^{\dim\lambda}\langle e^\lambda_i \mid \phi(e^\lambda_i)\rangle_{V^\lambda} = \frac{1}{|G|}\sum_{g\in G}\sum_{i=1}^{\dim\lambda}\langle g\cdot e^\lambda_j \mid e^\lambda_i\rangle\,\langle e^\lambda_i \mid g\cdot e^\lambda_l\rangle_{V^\lambda} \\ &= \frac{1}{|G|}\sum_{g\in G}\langle g\cdot e^\lambda_j \mid g\cdot e^\lambda_l\rangle_{V^\lambda} = \langle e^\lambda_j \mid e^\lambda_l\rangle_{V^\lambda} = 1_{(j=l)},\end{aligned}$$

by using the fact that $(e^\lambda_i)_{1\leq i\leq\dim\lambda}$ is an orthonormal basis of V^λ, and the fact that G acts by isometries on V^λ. Therefore, if $\lambda=\mu$, then $\phi = 1_{(j=l)}\,\frac{\mathrm{id}_{V^\lambda}}{\dim\lambda}$, and finally,

$$\langle \rho^\lambda_{ij} \mid \rho^\lambda_{kl}\rangle_{\mathbb{C}G} = \langle e^\lambda_i \mid \phi(e^\lambda_k)\rangle_{V^\lambda} = \frac{1_{(i=k,j=l)}}{\dim\lambda}.$$ ☐

PROOF OF THEOREM 1.21. For any representation (V,ρ), the *conjugate representation* (V^*,ρ^*) is defined as follows: $V^* = \hom(V,\mathbb{C})$ is the space of linear forms $V\to\mathbb{C}$, and for $\phi\in V^*$, we set

$$(\rho^*(g))(\phi) = \phi\circ\rho(g^{-1}).$$

Suppose given an orthonormal basis $(e_i)_{1\leq i\leq\dim V}$ of V. Then, a basis of V^* consists in the forms e^*_i defined by $e^*_i = \langle e_i \mid \cdot\rangle$. The representation matrices of (V^*,ρ^*) in this basis are related to those of (V,ρ) by taking the complex conjugates. Indeed,

$$\rho^*_{ij}(g) = (\rho^*(g)(e^*_j))(e_i) = (e^*_j\circ\rho(g^{-1}))(e_i) = \langle e_j \mid g^{-1}e_i\rangle = \overline{\langle e_i \mid g\cdot e_j\rangle} = \overline{\rho_{ij}(g)}.$$

If $\lambda\in\widehat{G}$, we shall denote λ^* the label of the corresponding conjugate representation; it is easy to see that it is again irreducible, and that $(\lambda^*)^* = \lambda$; thus, the conjugation of representation induces an involution of $\widehat{G}$.

We know that the two algebras $\mathbb{C}G$ and $\mathbb{C}\widehat{G}$ have the same dimension; therefore, it suffices to show that the Fourier transforms $\widehat{\rho^\lambda_{ij}}$ of the matrix coefficients are orthogonal elements of $\mathbb{C}\widehat{G}$, with the correct norms. We start from the orthogonality of matrix coefficients in $\mathbb{C}G$, and we rewrite it as a formula for $\widehat{\rho^\lambda_{ij}}$:

$$\frac{|G|}{\dim\lambda}\,1_{(\lambda=\mu,i=j,k=l)} = \sum_{g\in G}\rho^\lambda_{ij}(g)\,\overline{\rho^\mu_{kl}(g)} = \left(\sum_{g\in G}\rho^\lambda_{ij}(g)\,\rho^{\mu^*}(g)\right)_{kl} = \left(\widehat{\rho^\lambda_{ij}}(\mu^*)\right)_{kl}.$$

So, the Fourier transform $\widehat{\rho^\lambda_{ij}}$ is the element of $\mathrm{End}(V^{\lambda^*})\subset\mathbb{C}\widehat{G}$ which is proportional to the elementary matrix E_{ij}, with a coefficient $\frac{|G|}{\dim\lambda}$. In $\mathbb{C}\widehat{G}$, all these elements are orthogonal, and

$$(\|\widehat{\rho^\lambda_{ij}}\|_{\mathbb{C}\widehat{G}})^2 = \frac{|G|^2}{(\dim\lambda)^2}\frac{\dim\lambda}{|G|^2} = \frac{1}{\dim\lambda}.$$

This ends the proof of the theorem. □

Remark 1.23. Theorem 1.21 extends readily to the setting of compact topological groups, the main difference being that for an infinite compact topological group G, the set $\widehat{G}$ of isomorphism classes of irreducible representations is countable instead of finite. For compact topological groups, the algebra $\mathbb{C}G$ is replaced by $\mathscr{L}^2(G,\,dg)$, and the product of this algebra is the convolution product

$$(f_1 * f_2)(h) = \int_G f_1(hg^{-1})\,f_2(h)\,dg.$$

In the case of finite groups, this normalisation of the products differs from the one used throughout the chapter by a factor $\frac{1}{|G|}$. We then still have an isomorphism and isometry between $\mathscr{L}^2(G,\,dg)$ and the dual space

$$\mathscr{L}^2(\widehat{G}) = \bigoplus_{\lambda\in\widehat{G}}^{\perp}\mathrm{End}(V^\lambda),$$

provided that we renormalise the scalar products on the endomorphism spaces $\mathrm{End}(V^\lambda)$ as follows: $\langle u_1\mid u_2\rangle_{\mathrm{End}(V^\lambda)} = (\dim\lambda)\,\mathrm{tr}(u_1^*u_2)$. The isomorphism is given by the non-commutative Fourier transform

$$\widehat{f}(\lambda) = \int_G f(g)\,\rho^\lambda(g)\,dg;$$

in the case of finite groups, this definition differs from Definition 1.19 by a factor $\frac{1}{|G|}$.

As an application of the isomorphism theorem 1.21, we can now prove the following:

Proposition 1.24 (Cardinality of the dual). *Given a finite group G, the cardinality of $\widehat{G}$ equals the number of conjugacy classes of G.*

Proof. Consider the center $\mathrm{Z}(\mathbb{C}G)$ of the group algebra of G: this is the vector subspace of $\mathbb{C}G$ which consists in elements f such that $f*e=e*f$ for any $e\in\mathbb{C}G$. It is easy to see that $f=\sum_{g\in G}f(g)\,g$ belongs to $\mathrm{Z}(\mathbb{C}G)$ if and only if, for any $g,h\in G$,

$$f(hgh^{-1}) = f(g).$$

So, f viewed as a function is constant on any conjugacy class of G (an equivalence class for the relation $g_1\sim g_2 \iff g_2=hg_1h^{-1}$ for some $h\in G$). Therefore, a linear basis of $\mathrm{Z}(\mathbb{C}G)$ consists in the formal sums associated to the conjugacy classes. In particular,

$$\dim \mathrm{Z}(\mathbb{C}G) = \text{number of conjugacy classes of } G.$$

Now, by Theorem 1.21, this dimension is also the dimension of $\mathrm{Z}(\mathbb{C}\widehat{G}) = \bigoplus_{\lambda\in\widehat{G}} \mathrm{Z}(\mathrm{End}(V^\lambda))$. However, it is well known that the center of an algebra of matrices $\mathrm{End}(V)$ with V finite-dimensional vector space consists in the scalar matrices $z\,\mathrm{id}_V$ with $z \in \mathbb{C}$. Therefore,

$$\dim \mathrm{Z}(\mathbb{C}\widehat{G}) = \sum_{\lambda\in\widehat{G}} \dim \mathrm{Z}(\mathrm{End}(V^\lambda)) = \sum_{\lambda\in\widehat{G}} 1 = \mathrm{card}(\widehat{G}). \qquad \square$$

References

The use of representation theory for the study of random walks on groups goes back to the works of Poincaré [Poi12]. For random walks on the symmetric groups, it became popular in the eighties and nineties, in particular through the works of Diaconis. The mixing time of the random transposition model is computed with these techniques in [DS81]; around the same time, the result was also obtained by Aldous and by Diaconis, Flatto and Shepp by using stopping time techniques. For the top-with-random transposition model, we refer to [FOW85; Dia88; DG89]. The top-to-random cycle model is dealt with in [AD86; DFP92]. The problem of the mixing time of the adjacent transposition shuffle stayed unsolved for a long time, until it was solved by Lacoin in [Lac16]. Finally, the riffle shuffle model has been introduced by Gilbert, Shannon and Reeds, and its mixing time has been considered in [Ald83; BD92]; an excellent survey of the mathematics of this model can be found in [Dia03]. In particular, a representation theoretic approach for the riffle shuffle is provided by the so-called Solomon descent algebra. In Chapter 3, we shall give for each model the asymptotics of the mixing time, and we shall present the proofs of these results for the random transposition and the top-with-random transposition models.

For the representation theory of finite groups, there are many textbooks available, one of the most famous being the one by Serre [Ser77]. With the viewpoint of harmonic analysis, a more recent treatment is provided by [CST08]; see also [CST10] for the case of the symmetric groups, which we shall detail in Chapter 2. We strongly insist on the fact that many definitions and constructions around representations of groups are natural if we consider them as modules over the group algebras: morphisms of representation, induction and restriction of representations, *etc.* For the extension of the Fourier isomorphism to the setting of compact groups, see for instance the first part of [Bum13].

Exercises

(1) **Representations of abelian groups.** Let G be a finite abelian group: $gh = hg$ for any $g, h \in G$. Show that any irreducible representation of G has dimension 1 (hint: given a representation (V, ρ), consider a common eigenvector of the endomorphisms $\rho(g)$, $g \in G$). In the abelian case, what is the cardinality of the dual $\widehat{G}$?

(2) **One-dimensional representations of the symmetric groups.** Find all the (isomorphism classes of) representations of dimension 1 of the symmetric group $\mathfrak{S}(N)$, $N \geq 2$ (hint: show that all the transpositions must be sent by the defining morphism $\rho : \mathfrak{S}(N) \to \mathbb{C}^* = \mathrm{GL}(\mathbb{C})$ to the same element).

(3) **The character of a representation.** Given a representation (V, ρ) of a finite group G, its *character* is the map

$$\begin{aligned} \mathrm{ch}^V : G &\to \mathbb{C} \\ g &\mapsto \mathrm{tr}(\rho(g)). \end{aligned}$$

(a) Show that any character satisfies $\mathrm{ch}^V(gh) = \mathrm{ch}^V(hg)$ for $g, h \in G$. Hence, the characters viewed as elements of $\mathbb{C}G$ belong to $\mathrm{Z}(\mathbb{C}G)$.

(b) Show that the irreducible characters $\mathrm{ch}^\lambda = \mathrm{tr}\,\rho^\lambda$, which are labelled by the irreducible representations $\lambda = (V^\lambda, \rho^\lambda)$ in $\widehat{G}$, form an orthonormal basis of $\mathrm{Z}(\mathbb{C}G)$ (hint: use the orthogonality of the matrix coefficients).

(c) Show that a representation V of a finite group G is entirely determined (up to isomorphisms) by its character ch^V.

(d) Establish the following Fourier inversion formula for functions in $\mathrm{Z}(\mathbb{C}G)$: if $f \in \mathrm{Z}(\mathbb{C}G)$, then

$$f = \sum_{\lambda \in \widehat{G}} \widetilde{f}(\lambda)\,\mathrm{ch}^\lambda,$$

where $\widetilde{f}(\lambda) = \frac{1}{|G|}\sum_{g\in G}\mathrm{ch}^\lambda(g^{-1})\,f(g)$.

(4) **The Fourier inversion formula.** By using the orthogonality of characters, show the following general Fourier inversion formula: if $f \in \mathbb{C}G$ is viewed as a function, then for any $g \in G$,

$$f(g) = \sum_{\lambda\in\widehat{G}} \frac{\dim\lambda}{|G|}\,\mathrm{tr}\left(\widehat{f}(\lambda)\,\rho^\lambda(g^{-1})\right).$$

(5) **The discrete Fourier transform.** Suppose that $G = \mathbb{Z}/n\mathbb{Z}$. Show that the irreducible representations of G are as follows: they are all one-dimensional, they are also labelled by the elements of $\mathbb{Z}/n\mathbb{Z}$, and a class of integers k modulo n gives rise to the irreducible representation

$$\begin{aligned}\mathbb{Z}/n\mathbb{Z} &\to \mathbb{C}^* = \mathrm{GL}(\mathbb{C})\\ l &\mapsto \mathrm{e}^{\frac{2\mathrm{i}\pi kl}{n}}.\end{aligned}$$

Write the Fourier inversion formula for functions $f : \mathbb{Z}/n\mathbb{Z} \to \mathbb{C}$. Reprove it without using representation theory.

(6) **The second orthogonality formula.** Let G be a finite group. The objective of the exercise is to show the following orthogonality formula: for any $g, h \in G$, we have

$$\sum_{\lambda\in\widehat{G}} \mathrm{ch}^\lambda(g^{-1})\,\mathrm{ch}^\lambda(h) = \begin{cases}\frac{|G|}{|C_g|} & \text{if } g \text{ and } h \text{ are conjugated in } G,\\ 0 & \text{otherwise,}\end{cases}$$

where C_g denotes the conjugacy class $\{g' = hgh^{-1},\ h \in G\}$ of g.

(a) Let G and H be two finite groups. Given two representations (V, ρ^V) and (W, ρ^W) of respectively G and H, their *exterior tensor product* is the representation of the product group $G \times H$ with underlying space $V \otimes W$, and underlying morphism

$$\begin{aligned}\rho^V \otimes \rho^W : G \times H &\to \mathrm{GL}(V \otimes W)\\ (g,h) &\mapsto \rho^V(g) \otimes \rho^W(h).\end{aligned}$$

Show that the exterior tensor product induces a bijection

$$\begin{aligned}\widehat{G} \times \widehat{H} &\to \widehat{G \times H}\\ (V^\lambda, \rho^\lambda), (V^\mu, \rho^\mu) &\mapsto (V^\lambda \otimes V^\mu, \rho^\lambda \otimes \rho^\mu).\end{aligned}$$

(Hint: a representation V is irreducible if and only if its character has norm 1, and two irreducible representations are non-isomorphic if and only if their characters are orthogonal. Show that the character of the exterior tensor product of two representations is the product of the characters, and then prove that the exterior tensor products $\lambda \otimes \mu$ are irreducible and non-isomorphic representations of $G \times H$).

(b) Given a finite group G and an irreducible representation V^λ, we consider the following representation of $G \times G$ on $\mathrm{End}(V^\lambda)$:

$$(g,h) \cdot u = \rho^\lambda(g) \circ u \circ \rho^\lambda(h^{-1}).$$

Show that this representation of $G \times G$ is irreducible, and give its character.

(c) Construct a representation of $G \times G$ on $\mathbb{C}G$ such that the Fourier isomorphism $\mathbb{C}G \to \bigoplus_{\lambda \in \widehat{G}} \mathrm{End}(V^\lambda)$ is a morphism of representations. Compute the character of this representation by two different ways, and deduce from this the orthogonality relation stated at the beginning of the exercise.

(7) **Analysis of the random walk on the hypercube.** We consider the following random walk on $G = (\mathbb{Z}/2\mathbb{Z})^N$: at each step of the random walk, one stays at the same vector $(\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_N)$ with probability $\frac{1}{N+1}$, and one changes the parity of one of the N coordinates with probability $\frac{1}{N+1}$ for each possibility. Hence, one considers the random walk on G with generator:

$$\mu = \frac{1}{N+1}\,(0,0,\ldots,0) + \frac{1}{N+1}\sum_{i=1}^{N}(0,\ldots,1_i,\ldots,0) \quad \in \mathbb{C}G.$$

(a) Prove that G admits 2^N irreducible representations, which are all one-dimensional and write as follows:

$$\rho^t(\varepsilon_1,\ldots,\varepsilon_N) = (-1)^{t_1\varepsilon_1+\cdots+t_N\varepsilon_N} \in \mathbb{C}^* = \mathrm{GL}(\mathbb{C}).$$

for $t = (t_1,\ldots,t_N) \in (\mathbb{Z}/2\mathbb{Z})^N$.

(b) For $t \in (\mathbb{Z}/2\mathbb{Z})^N$, compute the Fourier transform $\widehat{\mu}(t)$ in terms of $|t| = \sum_{i=1}^N |t_i|$ (the symbols $|\cdot|$ mean that the sum is taken in $\mathbb{N}$ instead of $\mathbb{Z}/2\mathbb{Z}$).

(c) Show that for any $n \in \mathbb{N}$ and any $\varepsilon \in (\mathbb{Z}/2\mathbb{Z})^N$, one has:

$$\mu^n(\varepsilon) = \frac{1}{2^N}\sum_{t \in (\mathbb{Z}/2\mathbb{Z})^N} (\widehat{\mu}(t))^n\,\rho^t(\varepsilon).$$

We denote μ_∞ the uniform measure on G. Show that $\mu^n \to_{n\to\infty} \mu_\infty$, and give an exact formula for $(\|\mu^n - \mu_\infty\|_{\mathbb{C}G})^2$. Interpret this in terms of asymptotic properties of the random walk on the hypercube.

CHAPTER 2

The Frobenius–Schur isomorphism and the Jucys–Murphy elements

In the previous chapter, we have introduced several random walks on $\mathfrak{S}(N)$, all of them being described by a generator μ in the group algebra $\mathbb{C}\mathfrak{S}(N)$. The Fourier isomorphism ensures that the law at time n of the random walk has for Fourier transform

$$\sum_{\lambda \in \widehat{\mathfrak{S}(N)}} (\widehat{\mu}(\lambda))^n,$$

so by computing the powers of the matrices $\widehat{\mu}(\lambda)$ and by reversing the Fourier isomorphism, one has access to the marginal laws of the random walk, and one will be able to compute for instance the total variation distance between μ^n and the uniform measure. The obvious prerequisites of these computations are the following:

- we need a description of the set $\widehat{\mathfrak{S}(N)}$: can one give a complete list of the irreducible representations of $\mathfrak{S}(N)$, and compute for instance their dimensions?
- we also need to be able to compute the Fourier transforms $\widehat{\mu}(\lambda)$.

It turns out that many computations on the representations of the symmetric groups can be performed *without knowing precisely the representation matrices*. The tool which enables this is the algebra of symmetric functions, and the Frobenius–Schur isomorphism; this is the main focus of this chapter. We shall also present without proof an alternative approach to the representation theory of the symmetric groups, which is due to Okounkov and Vershik and which involves certain special elements of the group algebra $\mathbb{C}\mathfrak{S}(N)$ called the Jucys–Murphy elements; they will play an important role in certain computations of Chapter 3.

1. The five bases of the algebra of symmetric functions

Let $X = \{x_1, x_2, \ldots, x_n, \ldots\}$ be an infinite sequence of commutative variables. A *monomial* is a finite product of such variables: for instance, x_1x_3, $(x_2)^3$, $x_1(x_3)^4x_4$ are monomials. The *degree* of a monomial is the number of terms of the product: for instance, $\deg(x_1(x_3)^4x_4) = 6$. A *polynomial* in an infinite number of variables is a formal linear combination of monomials

$$p(X) = \sum_{I \in \mathcal{I}} c_I \, x_I,$$

with coefficients c_I for instance in the field of real numbers $\mathbb{R}$, and with each monomial x_I bounded in degree by a common upper bound $\deg p = \sup_{I \in \mathcal{I}}(\deg x_I) < +\infty$. We allow infinite sums such as

$$\sum_{i=1}^{\infty} (x_i)^2 x_{i+1},$$

but we do not allow sums unbounded in degree such as $\sum_{k=1}^{\infty} (x_1)^k$. Given a monomial $x_I = (x_{i_1})^{m_1} \cdots (x_{i_r})^{m_r}$, the number of pairs of monomials (x_J, x_K) such that $x_I = x_J x_K$ is finite, so we can define without ambiguity the product of two polynomials $p(X) = \sum_{J \in \mathcal{J}} c_J \, x_J$ and $q(X) = \sum_{K \in \mathcal{K}} d_K \, x_K$ by

$$(pq)(X) = \sum_{I \,|\, \deg(x_I) \leq \deg(p) + \deg(q)} \left(\sum_{(J,K) \in \mathcal{J} \times \mathcal{K} \,|\, x_J x_K = x_I} c_J d_K \right) x_I.$$

This definition obviously extends the usual rules for products of polynomials in a finite number of variables. It yields a structure of graded $\mathbb{R}$-algebra on the space $\mathbb{R}[X]$ of polynomials in an infinite number of variables.

If $p(x_1,\ldots,x_n)$ is a polynomial in n variables, there is a natural action of $\mathfrak{S}(n)$ on $\mathbb{R}[x_1,\ldots,x_n]$ given by:

$$(\sigma\cdot p)(x_1,\ldots,x_n)=p(x_{\sigma(1)},\ldots,x_{\sigma(n)}).$$

For instance, if $p(x_1,x_2,x_3)=(x_1)^2+x_3$ and $\sigma=(1,2)$, then $(\sigma\cdot p)(x_1,x_2,x_3)=(x_2)^2+x_3$. A polynomial in n variables is called symmetric if $\sigma\cdot p=p$ for any $\sigma\in\mathfrak{S}(n)$. For instance, $p(x)=x_1x_2+x_2x_3+x_3x_1$ is a symmetric polynomial in 3 variables. If we want the same notion for polynomials $p\in\mathbb{R}[X]$ in an infinite number of variables, then we need to introduce the infinite symmetric group

$$\mathfrak{S}(\infty)=\bigcup_{n\geq 1}\uparrow\mathfrak{S}(n).$$

For every $n\in\mathbb{N}^*$, the group $\mathfrak{S}(n)$ identifies with the subgroup of $\mathfrak{S}(n+1)$ which consists in permutations σ such that $\sigma(n+1)=n+1$. Therefore, we have an increasing sequence of groups $(\mathfrak{S}(n))_{n\geq 1}$, all of them acting on $\mathbb{N}^*$; the group $\mathfrak{S}(n)$ leaves invariant the elements of $[\![n+1,+\infty]\!]$. The infinite symmetric group $\mathfrak{S}(\infty)$ is the increasing union of these groups: an element of $\mathfrak{S}(\infty)$ is a permutation of $\mathbb{N}^*$ which moves only a finite number of integers (in some interval $[\![1,n]\!]$). Then, we have an action of $\mathfrak{S}(\infty)$ on $\mathbb{R}[X]$, which is analogous to the action of $\mathfrak{S}(n)$ on $\mathbb{R}[x_1,\ldots,x_n]$:

$$(\sigma\cdot p)(x_1,\ldots,x_n,\ldots)=p(x_{\sigma(1)},\ldots,x_{\sigma(n)},\ldots).$$

Definition 2.1 (Symmetric function). *A symmetric function is an element $f\in\mathbb{R}[X]$ (polynomial in an infinite number of variables) such that, for any permutation $\sigma\in\mathfrak{S}(\infty)$, we have $\sigma\cdot f=f$.*

It is easy to see that the action of the infinite symmetric group on $\mathbb{R}[X]$ is compatible with the grading and with the product of polynomials, and therefore that the symmetric functions form a graded subalgebra Sym of $\mathbb{R}[X]$. The remainder of this section is devoted to a description of five graded linear bases of the algebra Sym.

Example 2.2 (Power sums). For $k\geq 1$, the k-th *power sum* is the symmetric function

$$p_k(X)=\sum_{i=1}^{\infty}(x_i)^k;$$

it is homogeneous with degree k (all the monomials appearing in p_k have the same degree k). More generally, consider a non-increasing sequence $\lambda=(\lambda_1\geq\lambda_2\geq\cdots\geq\lambda_\ell)$ of positive integers. Such a sequence is also called an *partition* of the integer $|\lambda|=\sum_{i=1}^{\ell}\lambda_i$. For instance, $(4,3,3)$ is a partition of the integer 10. Given an integer partition λ, we set $p_\lambda(X)=p_{\lambda_1}(X)\,p_{\lambda_2}(X)\cdots p_{\lambda_\ell}(X)$. This is again a symmetric function, with degree $\deg p_\lambda=|\lambda|$.

Example 2.3 (Elementary functions). For $k\geq 1$, the k-th *elementary symmetric function* is the symmetric function

$$e_k(X)=\sum_{1\leq i_1<i_2<\cdots<i_k}x_{i_1}x_{i_2}\cdots x_{i_k}.$$

For instance, $e_2(X)=\sum_{1\leq i<j}x_ix_j$. Each e_k is symmetric with degree k, and as above, we can construct more complicated symmetric functions by taking products: thus, for $\lambda=(\lambda_1,\ldots,\lambda_\ell)$ integer partition, we set $e_\lambda(X)=e_{\lambda_1}(X)\,e_{\lambda_2}(X)\cdots e_{\lambda_\ell}(X)$. This is again a symmetric function, with degree $\deg e_\lambda=|\lambda|$.

EXAMPLE 2.4 (Homogeneous functions). For $k \geq 1$, the k-th *homogeneous symmetric function* is the symmetric function

$$h_k(X) = \sum_{1 \leq i_1 \leq i_2 \leq \cdots \leq i_k} x_{i_1} x_{i_2} \cdots x_{i_k}.$$

For instance, $h_2(X) = \sum_{1 \leq i \leq j} x_i x_j = \sum_{1 \leq i < j} x_i x_j + \sum_{i=1}^{\infty} (x_i)^2 = e_2(X) + p_2(X)$. Each h_k is symmetric with degree k, and as before, given $\lambda = (\lambda_1, \ldots, \lambda_\ell)$ integer partition, we set $h_\lambda(X) = h_{\lambda_1}(X)\, h_{\lambda_2}(X) \cdots h_{\lambda_\ell}(X)$. This is again a symmetric function, with degree $\deg h_\lambda = |\lambda|$.

LEMMA 2.5 (Generating series of the three bases p, e and h). *Any power sum p_λ can be expressed as a linear combination with real coefficients of elementary functions e_μ with $|\mu| = |\lambda|$. The same statement holds with any pair among*

$$(p, e),\ (e, p),\ (p, h),\ (h, p),\ (h, e),\ (e, h).$$

PROOF. We introduce a variable z commuting with all the variables x_i, and the three generating series:

$$P(X, z) = \sum_{k=1}^{\infty} \frac{p_k(X)}{k}\, z^k;$$

$$E(X, z) = 1 + \sum_{k=1}^{\infty} e_k(X)\, z^k;$$

$$H(X, z) = 1 + \sum_{k=1}^{\infty} h_k(X)\, z^k.$$

These functions belong to the algebra of formal power series in the variables $\{z, x_1, \ldots, x_n, \ldots\}$. We can compute them all. First,

$$P(X, z) = \sum_{i=1}^{\infty} \left(\sum_{k=1}^{\infty} \frac{(x_i z)^k}{k} \right) = \sum_{i=1}^{\infty} \log\left(\frac{1}{1 - x_i z} \right) = \log\left(\prod_{i=1}^{\infty} \frac{1}{1 - x_i z} \right).$$

If we consider the exponential of $P(X, z)$, we obtain

$$\prod_{i=1}^{\infty} \frac{1}{1 - x_i z} = \prod_{i=1}^{\infty} \left(\sum_{d_i=1}^{\infty} (x_i z)^{d_i} \right) = \sum_{k=0}^{\infty} z^k \left(\sum_{d_{i_1} + \cdots + d_{i_r} = k} (x_{i_1})^{d_{i_1}} \cdots (x_{i_r})^{d_{i_r}} \right).$$

The sum associated to the term z^k is the sum over all possible ways to choose indices $i_1, \ldots, i_r \in \mathbb{N}^*$ and positive multiplicities $d_{i_1}, \ldots, d_{i_r}$ such that $d_{i_1} + \cdots + d_{i_r} = k$. By gathering the equal indices in the definition of $h_k(X)$, one sees that this sum is equal to the k-th homogeneous symmetric function (with by convention $h_0(X) = 1$). Thus,

$$\exp(P(X, z)) = H(X, z).$$

This implies that every homogeneous symmetric function $h_{k \geq 1}$ can be expressed as a polynomial in the power sums p_l, and conversely. Indeed, if one wants for instance to express h_k in terms of the power sums, one writes:

$$h_k(X) = [z^k](\exp(P(X, z))) = \sum_{\ell \geq 1} \frac{1}{\ell!} \sum_{k_1 + \cdots + k_\ell = k} \frac{p_{k_1}(X) \cdots p_{k_\ell}(X)}{k_1 \cdots k_\ell}.$$

By reordering the integers k_i in order to obtain integer partitions, one gets:

$$h_k(X) = \sum_{\lambda \vdash k} \frac{p_\lambda(X)}{z_\lambda},$$

where the sum runs over all integer partitions of size k, and where if λ is an integer partition with $m_1(\lambda)$ parts equal to 1, $m_2(\lambda)$ parts equal to 2, *etc.*, then

$$z_\lambda = \prod_{s\geq 1} (s^{m_s(\lambda)}\, m_s(\lambda)!).$$

Since $h_\mu(X) = h_{\mu_1}(X)\cdots h_{\mu_\ell}(X)$ for any integer partition $\mu = (\mu_1,\ldots,\mu_\ell)$, we have thus proved that any homogeneous symmetric function $h_\mu(X)$ is a linear combination (with positive rational coefficients) of functions $p_\lambda(X)$ with the same degree. The converse change of bases is obtained by using the relation $P(X,z) = \log(H(X,z))$.

The passage from the basis h to the basis e is analogous. Indeed, let us remark that

$$\prod_{i=1}^{\infty}(1+zx_i) = 1 + \sum_{k=1}^{\infty} z^k \left(\sum_{1\leq i_1<i_2<\cdots<i_k} x_{i_1}x_{i_2}\cdots x_{i_k}\right) = 1+\sum_{k=1}^{\infty} e_k(X)\, z^k = E(X,z).$$

Therefore, $E(X,z) = (H(X,-z))^{-1}$, and this relation between generating series allows one to express elementary functions in terms of homogeneous functions and conversely (this is the same kind of argument as above). By composition, we can also go from the basis p to the basis e and *vice versa*. □

EXAMPLE 2.6. For $k = 4$, the change of basis formula between homogeneous functions and power sums yields:

$$h_4(X) = \frac{p_4(X)}{4} + \frac{p_{(3,1)}(X)}{3} + \frac{p_{(2,2)}(X)}{8} + \frac{p_{(2,1,1)}(X)}{4} + \frac{p_{(1,1,1,1)}(X)}{24}.$$

We can also express $h_4(X)$ in terms of elementary functions:

$$h_4(X) = -\,e_4(X) + 2\,e_{(3,1)}(X) + e_{(2,2)}(X) - 3\,e_{(2,1,1)}(X) + e_{(1,1,1,1)}(X).$$

We have used above the terminology of basis, but we have not yet proved that any symmetric function can be expressed as a linear combination of power sums (or, equivalently, of elementary symmetric functions, or of homogeneous symmetric functions). It is also not clear that the family of functions $(p_\lambda)_\lambda$ with λ running over the set of all integer partitions consists in linearly independent functions. The following results will show that this is indeed the case. In the sequel, we denote $\mathfrak{Y}(n)$ the set of integer partitions of size n, and $\mathfrak{Y} = \bigsqcup_{n\in\mathbb{N}} \mathfrak{Y}(n)$. The *length* of an integer partition $\lambda = (\lambda_1,\lambda_2,\ldots,\lambda_\ell)$ is its number of parts $\ell = \ell(\lambda)$.

PROPOSITION 2.7 (Monomial basis of the algebra of symmetric functions). *Denote* Sym_d *the vector subspace of* Sym *which consists in symmetric functions which are homogeneous of degree* d *(if* $f = \sum_{I\in\mathcal{I}} c_I\, x_I$ *belongs to* Sym_d*, then* $\deg(x_I) = d$ *for every* $I \in \mathcal{I}$*). We have* $\mathrm{Sym} = \bigoplus_{d\in\mathbb{N}} \mathrm{Sym}_d$*, and for every* $d \in \mathbb{N}$*, there exists a linear basis* $(m_\lambda(X))_{\lambda\in\mathfrak{Y}(d)}$ *of the vector space* Sym_d *labelled by the integer partitions* $\lambda \in \mathfrak{Y}(d)$*.*

PROOF. Given a monomial $x_I = (x_{i_1})^{d_1}(x_{i_2})^{d_2}\cdots(x_{i_\ell})^{d_\ell}$ with degree $d = d_1+d_2+\cdots+d_\ell$, there exists a unique non-increasing reordering of the multiplicities d_j which gives an integer partition $\lambda \in \mathfrak{Y}(d)$. By invariance by action of the symmetric group $\mathfrak{S}(\infty)$, if x_I occurs with a coefficient c in a symmetric function f, then

$$x^\lambda = (x_1)^{\lambda_1}(x_2)^{\lambda_2}\cdots(x_\ell)^{\lambda_\ell}$$

occurs in f with the same coefficient $c \in \mathbb{R}$, and moreover, any other monomial x_J with the same reordering λ of the multiplicities must also have the same coefficient c. Conversely, given $\lambda \in \mathfrak{Y}(d)$, the list of the monomials x_I corresponding to λ can be obtained as follows:

- first, choose a permutation $(d_1, d_2, \ldots, d_\ell)$ of the parts of λ; the number of distinct such permutations is
$$\frac{\ell(\lambda)!}{\prod_{s\geq 1} m_s(\lambda)!}.$$
- then, choose an increasing sequence of indices $i_1 < i_2 < \cdots < i_\ell$; the monomial thus obtained is
$$x_I = \prod_{j=1}^{\ell} (x_{i_j})^{d_j}.$$

Denote $m_\lambda(X) = \sum x_I$ the sum of all monomials corresponding to the integer partition λ; this is the *monomial symmetric function* associated to the partition λ. The reasoning above proves that any symmetric function f is a linear combination of functions $m_\lambda(X)$, with λ integer partition. If we restrict ourselves to functions which are homogeneous of degree d, then the possible integer partitions λ are those in $\mathfrak{Y}(d)$. Finally, the $m_\lambda(X)$'s are linearly independent, because if $f = \sum_\lambda c_\lambda\, m_\lambda(X)$, then the coefficient c_λ can be recovered as the coefficient of the monomial x^λ in the polynomial f. ☐

Proposition 2.8 (Three bases of the algebra of symmetric functions). *The families* $(p_\lambda)_{\lambda\in\mathfrak{Y}(d)}$, $(e_\lambda)_{\lambda\in\mathfrak{Y}(d)}$ *and* $(h_\lambda)_{\lambda\in\mathfrak{Y}(d)}$ *are three linear basis of* Sym_d*. Therefore,* $(p_\lambda)_{\lambda\in\mathfrak{Y}}$, $(e_\lambda)_{\lambda\in\mathfrak{Y}}$ *and* $(h_\lambda)_{\lambda\in\mathfrak{Y}}$ *are three graded linear basis of* Sym.

Proof. We are going to introduce a partial order $\preceq$ on $\mathfrak{Y}(d)$ such that, for any integer partition $\lambda \in \mathfrak{Y}(d)$,
$$p_\lambda(X) = \left(\prod_{s\geq 1} m_s(\lambda)!\right) m_\lambda(X) + \sum_{\substack{\mu\in\mathfrak{Y}(d)\\ \mu\prec\lambda}} c_{\lambda\mu}\, m_\mu(X)$$
for some non-negative integer coefficients $c_{\lambda\mu}$. This will imply that $(p_\lambda)_{\lambda\in\mathfrak{Y}(d)}$ is a basis of Sym_d, with a matrix of change of basis from $(p_\lambda)_{\lambda\in\mathfrak{Y}(d)}$ to $(m_\lambda)_{\lambda\in\mathfrak{Y}(d)}$ which is upper-triangular with positive coefficients on the diagonal. The two other bases will then be covered by Lemma 2.5.

Given two integer partitions λ and μ with the same size, we say that λ is finer than μ (notation: $\lambda \preceq \mu$) if one can obtain μ from λ by summing some of its parts, and then reordering them in non-increasing order. For instance, starting from the integer partition $\lambda = (5, 3, 2)$, one can construct the following partitions μ with $\lambda \prec \mu$:

$$\begin{tikzcd}[column sep=small]
 & (10) & \\
(8,2) \arrow[ur, dash] & (7,3) \arrow[u, dash] & (5,5) \arrow[ul, dash] \\
 & (5,3,2) \arrow[ul, dash] \arrow[u, dash] \arrow[ur, dash] &
\end{tikzcd}$$

Let us now consider the symmetric function
$$p_\lambda(X) = \sum_{i_1, i_2, \ldots, i_\ell \geq 1} (x_{i_1})^{\lambda_1} (x_{i_2})^{\lambda_2} \cdots (x_{i_\ell})^{\lambda_\ell}.$$

By Proposition 2.7, $p_\lambda(X)$ can be written as a linear combination of functions $m_\mu(X)$. To find which monomial symmetric functions are involved in this expansion, we associate to each ℓ-tuple of indices $(i_1, \ldots, i_\ell)$ the set partition π of $[\![1, \ell]\!]$ whose parts correspond to the identities of indices.

For instance, if the ℓ-tuple of indices is $(1,3,1,4,5,3,1,2)$, then it corresponds to the set partition of $[\![1,8]\!]$:

$$\pi = \{1,3,7\} \sqcup \{2,6\} \sqcup \{4\} \sqcup \{5\} \sqcup \{8\}.$$

Given an integer partition λ with length ℓ and a set partition $\pi = \pi_1 \sqcup \pi_2 \sqcup \cdots \sqcup \pi_s$ of $[\![1,\ell]\!]$, we can associate to it an integer partition $\mu \succeq \lambda$ by taking for each part π_j the sum of parts $\sum_{i\in\pi_j} \lambda_i$, and by reordering in non-increasing order the integers thus obtained. For instance, if π is as above and if $\lambda = (8,6,6,3,2,2,2,1)$, then $\mu = (16,8,3,2,1)$. By construction, it is clear that $\lambda \preceq \mu$ regardless of the value of π. On the other hand, the monomial $(x_{i_1})^{\lambda_1}\cdots(x_{i_r})^{\lambda_\ell}$ is conjugated by the action of $\mathfrak{S}(\infty)$ to x^μ, so it contributes to a term $m_\mu(X)$ in the expansion of $p_\lambda(X)$. Note that each monomial symmetric function can appear numerous times in this expansion, in particular because there can be several distinct set partitions which yield the same μ by summing parts of λ. However, our argument proves that if $m_\mu(X)$ occurs with a non-zero coefficient in the expansion of $p_\lambda(X)$, then $\lambda \preceq \mu$. To conclude the proof, we have to show that $m_\lambda(X)$ occurs with positive multiplicity. The monomials x_I occurring in $p_\lambda(X)$ and conjugated by $\mathfrak{S}(\infty)$ to x^λ come from the ℓ-tuples of indices $(i_1,\ldots,i_\ell)$ which are all distinct. Moreover, there are $\prod_{s\geq 1} m_s(\lambda)!$ such monomials equal to x^λ; indeed, this happens when $(i_1,\ldots,i_\ell)$ is a permutation of $(1,2,\ldots,\ell)$ which permutes the blocks $(1,2,\ldots,m_1(\lambda))$, $(m_1(\lambda)+1,\ldots,m_1(\lambda)+m_2(\lambda))$, *etc.* This ends the proof of the proposition. □

EXAMPLE 2.9. The expansion of the power sum $p_{(5,3,2)}(X)$ in the basis of monomial symmetric functions is:

$$p_{(5,3,2)}(X) = m_{(5,3,2)}(X) + 2\,m_{(5,5)}(X) + m_{(7,3)}(X) + m_{(8,2)}(X) + m_{(10)}(X).$$

REMARK 2.10. Proposition 2.8 can be restated as follows: $(p_k)_{k\geq 1}$, $(e_k)_{k\geq 1}$ and $(h_k)_{k\geq 1}$ are algebraic bases of Sym, so any symmetric function can be written uniquely as a multivariate polynomial in the elements of one of these three families. As a consequence, if A is a real algebra and if one wants to define a morphism of algebras $\Phi : \mathrm{Sym} \to A$, then it suffices to specify the images $\Phi(p_k)$ for $k \geq 1$ (or the images $\Phi(e_{k\geq 1})$, or the images $\Phi(h_{k\geq 1})$). This remark will play an important role in Chapter 5, where we shall study a class of morphisms $\Phi : \mathrm{Sym} \to \mathbb{R}$.

In order to introduce the last important basis of Sym, we need another viewpoint on this algebra, namely, as a projective limit. For any $n \geq 1$, there is a morphism of algebras from $\mathbb{R}[X]$ to the algebra $\mathbb{R}[x_1,\ldots,x_n]$ of polynomials in n variables:

$$\pi_n(f(X)) = f(x_1,\ldots,x_n,0,0,\ldots).$$

One says that $\pi_n(f)$ is the *specialisation* of f with the variables x_{n+1}, x_{n+2}, *etc.* set to 0; we shall also denote $\pi_n(f) = f(x_1,\ldots,x_n)$. Given $N \geq n$, we have similar morphisms of algebras π_n^N : $\mathbb{R}[x_1,\ldots,x_N] \to \mathbb{R}[x_1,\ldots,x_n]$ corresponding again to setting $x_{n+1} = x_{n+2} = \cdots = x_N = 0$. These morphisms are compatible with one another: the following diagram is commutative

$$\begin{tikzcd}
\mathbb{R}[X] \arrow[dr, "\pi_{n+1}"'] \arrow[drr, "\pi_n"] \arrow[drrr, "\pi_{n-1}"] & & & & \\
\cdots \arrow[r] & \mathbb{R}[x_1,\ldots,x_{n+1}] \arrow[r, "\pi_n^{n+1}"'] & \mathbb{R}[x_1,\ldots,x_n] \arrow[r, "\pi_{n-1}^n"'] & \mathbb{R}[x_1,\ldots,x_{n-1}] \arrow[r] & \cdots
\end{tikzcd}$$

Since $\mathfrak{S}(\infty) = \bigcup_{n\geq 1} \uparrow \mathfrak{S}(n)$, if $f \in \mathrm{Sym}$, then its images $\pi_n(f)$ are invariant under the actions of the corresponding symmetric groups $\mathfrak{S}(n)$. Therefore, if $\mathrm{Sym}(n)$ denotes the subalgebra of $\mathbb{R}[x_1,\ldots,x_n]$ of symmetric polynomials in n variables, then the previous diagram restricts to:

$$\begin{tikzcd}
 & \mathrm{Sym} \arrow[d, "\pi_{n+1}"'] \arrow[rd, "\pi_n"] \arrow[rrd, "\pi_{n-1}"] & & & \\
\cdots \arrow[r] & \mathrm{Sym}(n+1) \arrow[r, "\pi_n^{n+1}"'] & \mathrm{Sym}(n) \arrow[r, "\pi_{n-1}^n"'] & \mathrm{Sym}(n-1) \arrow[r] & \cdots
\end{tikzcd}$$

As before, we denote with an index d the vector subspace $\mathrm{Sym}(n)_d$ of $\mathrm{Sym}(n)$ which consists in symmetric polynomials homogeneous of degree d. The same argument as in the proof of Proposition 2.7 shows that a linear basis of $\mathrm{Sym}(n)_d$ consists in monomial functions $m_\lambda(x_1,\ldots,x_n)$, with the following restrictions: we need $|\lambda| = d$ (degree condition) and also $\ell(\lambda) \leq n$ (otherwise, the number of distinct variables is not large enough to ensure that one can distribute the exponents $\lambda_1,\ldots,\lambda_\ell$). Thus,

$$\mathrm{Sym}_d = \bigoplus_{\lambda\in\mathfrak{Y}(d)} \mathbb{C}m_\lambda(X);$$
$$\mathrm{Sym}(n)_d = \bigoplus_{\substack{\lambda\in\mathfrak{Y}(d)\\ \ell(\lambda)\leq n}} \mathbb{C}m_\lambda(x_1,\ldots,x_n).$$

If $d \leq n$, then the restriction $\ell(\lambda) \leq n$ is superfluous, because $\ell(\lambda) \leq |\lambda|$ for any partition λ. It follows that the restriction of π_n to symmetric functions with degree smaller than or equal to n is an isomorphism of vector spaces. In particular, in order to define a symmetric function f with degree d, it suffices to define symmetric polynomials $f_n(x_1,\ldots,x_n)$ for n large enough (larger than $n_0 \geq d$), so that these polynomials are compatible with one another: $\pi_n^N(f_N) = f_n$ for any $N \geq n \geq n_0$. Then, there is a unique symmetric function f such that $\pi_n(f) = f_n$ for any n.

Remark 2.11. The previous discussion shows that Sym is the *projective limit in the category of graded algebras* of the sequence of graded algebras $(\mathrm{Sym}(n))_{n\in\mathbb{N}}$.

A polynomial in n variables $f(x_1,\ldots,x_n)$ is called *antisymmetric* if, for any permutation $\sigma \in \mathfrak{S}(n)$, we have

$$f(x_{\sigma(1)},\ldots,x_{\sigma(n)}) = \varepsilon(\sigma)\, f(x_1,\ldots,x_n),$$

where $\varepsilon(\sigma)$ is the signature of the permutation σ. For instance, $(x_1-x_2)(x_2-x_3)(x_1-x_3)$ is an antisymmetric polynomial with degree 3. We denote $\mathrm{AntiSym}(n)$ the vector space of antisymmetric polynomials in n variables. Consider the antisymmetrisation operator

$$\begin{aligned}\mathcal{A}_n : \mathbb{R}[x_1,\ldots,x_n] &\to \mathrm{AntiSym}(n)\\ f &\mapsto \sum_{\sigma\in\mathfrak{S}(n)} \varepsilon(\sigma)\,(\sigma\cdot f).\end{aligned}$$

Set $\rho_n = (n-1, n-2, \ldots, 1)$. Note that if $f(x_1,\ldots,x_n) = x^{\rho_n} = (x_1)^{n-1}(x_2)^{n-2}\cdots(x_n)^0$, then $(\mathcal{A}_n f)(x_1,\ldots,x_n)$ is the *Vandermonde determinant*:

$$(\mathcal{A}_n f)(x_1,\ldots,x_n) = \Delta(x_1,\ldots,x_n) = \prod_{1\leq i<j\leq n} (x_i - x_j).$$

Theorem 2.12 (Antisymmetric polynomials). *The map $f \mapsto \Delta \times f$ is an isomorphism of vector spaces from $\mathrm{Sym}(n)$ to $\mathrm{AntiSym}(n)$. On the other hand, a linear basis of $\mathrm{AntiSym}(n)$ consists in the polynomials $\mathcal{A}_n(x^{\lambda+\rho_n})$ with λ integer partition with length smaller than n.*

Proof. The ring of polynomials $\mathbb{R}[x_1,x_2,\ldots,x_n]$ is a unique factorisation domain, and the polynomials $x_i - x_j$ are irreducible. Let us prove that, for any pair $1 \leq i < j \leq n$ and for any antisymmetric polynomial $g(x_1,\ldots,x_n)$, g is divisible by $x_i - x_j$; by the unique factorisation property, this will imply that Δ divides g. By symmetry, it suffices to treat the case $(i,j)=(1,2)$. If $x_I=(x_1)^{n_1}(x_2)^{n_2}\,x_{I'}$ with $x_{I'}$ involving only variables $x_{k\geq 3}$, then the coefficient c of x_I in an antisymmetric polynomial $g(x_1,\ldots,x_N)$ is the opposite of the coefficient of $x_J=(x_1)^{n_2}(x_2)^{n_1}\,x_{I'}$. Therefore, if $n_1=n_2$, then $c=0$. Otherwise, we have a pairing of the monomials x_I and x_J, and it suffices to show that x_1-x_2 divides x_I-x_J. This is obvious: for instance, if $n_1>n_2$, then

$$(x_1)^{n_1}(x_2)^{n_2}-(x_1)^{n_2}(x_2)^{n_1}=(x_1-x_2)\left(\sum_{k=0}^{n_1-n_2-1}(x_1)^{n_2+k}(x_2)^{n_1-1-k}\right).$$

Now, the map $f\mapsto \Delta f$ is injective ($\mathbb{R}[x_1,\ldots,x_n]$ is an integral ring), and if f is symmetric, then Δf is antisymmetric:

$$\sigma\cdot(\Delta f)=(\sigma\cdot\Delta)(\sigma\cdot f)=\varepsilon(\sigma)\,\Delta f.$$

Indeed, the action of σ on Δ captures the parity of the number of inversions of σ, hence its signature. Conversely, if g is antisymmetric, then by the discussion above $g=\Delta f$ for some polynomial f, and we have

$$\varepsilon(\sigma)\,\Delta f=\varepsilon(\sigma)g=\sigma\cdot g=(\sigma\cdot\Delta)(\sigma\cdot f)=\varepsilon(\sigma)\,\Delta(\sigma\cdot f),$$

so $f=\sigma\cdot f$ and f is symmetric. The first part of the theorem is thus proved.

For the second part, note that if x_I is a monomial in $\mathbb{R}[x_1,\ldots,x_n]$, then $\mathcal{A}_n(x_I)=\pm\,\mathcal{A}_n(x^\lambda)$, where λ is the integer partition obtained by reordering the exponents in x_I. Therefore, $\mathrm{AntiSym}(n)$ is spanned linearly by the polynomials $\mathcal{A}_n(x^\lambda)$ with λ integer partition with length smaller than n. If λ has two parts $\lambda_i=\lambda_{i+1}$ (including the case where $\ell(\lambda)\leq n-2$ and $\lambda_{n-1}=\lambda_n=0$), then with $\tau=(i,i+1)$, we have

$$-\,\mathcal{A}_n(x^\lambda)=\tau\cdot\mathcal{A}_n(x^\lambda)=\mathcal{A}_n(\tau\cdot x^\lambda)=\mathcal{A}_n(x^\lambda),$$

so $\mathcal{A}_n(x^\lambda)=0$. Therefore,

$$\mathrm{AntiSym}(n)=\mathrm{Span}(\{\mathcal{A}_n(x^\lambda),\ \lambda\text{ strict integer partition with length }n-1\text{ or }n\}),$$

where by strict integer partition we mean an integer partition whose parts are in strict decreasing order (hence, all distinct). These integer partitions are those that write as

$$\lambda+\rho_n=(\lambda_1+\rho_{n,1},\lambda_2+\rho_{n,2},\ldots,\lambda_n+\rho_{n,n}),$$

with λ arbitrary integer partition with length smaller than n, and $\rho_n=(n-1,n-2,\ldots,1)$ (with by convention $\lambda_k=0$ if $k>\ell(\lambda)$, and $\rho_{n,n}=0$). To end the proof, we have to show that the polynomials $\mathcal{A}_n(x^{\lambda+\rho_n})$ with λ running over the set of partitions with length smaller than n are linearly independent. By using the degree, it suffices to prove that given a sum

$$g(x_1,\ldots,x_n)=\sum_{\lambda\in\mathfrak{Y}(d)}c_\lambda\,\mathcal{A}_n(x^{\lambda+\rho_n}),$$

one can recover each coefficient c_λ from g. However, it is easy to see that c_λ is the coefficient of $x^{\lambda+\rho_n}$ in g; this ends the proof. □

Given an integer partition λ with $\ell(\lambda)\leq n$, the previous theorem ensures that the polynomial $\mathcal{A}_n(x^{\lambda+\rho_n})$ is divisible by $\Delta=\mathcal{A}_n(x^{\rho_n})$, and that the *Schur polynomials*

$$s_\lambda(x_1,\ldots,x_n)=\frac{\mathcal{A}_n(x^{\lambda+\rho_n})}{\mathcal{A}_n(x^{\rho_n})}=\frac{\det((x_i)^{\lambda_j+n-j})_{1\leq i,j\leq n}}{\det((x_i)^{n-j})_{1\leq i,j\leq n}}$$

with $\ell(\lambda)\leq n$ form a linear basis of $\mathrm{Sym}(n)$. Moreover, the definition of s_λ does not depend on n, in the following sense: for any $n\geq\ell(\lambda)$,

$$s_\lambda(x_1,\ldots,x_n,0)=s_\lambda(x_1,\ldots,x_n),$$

where on the left-hand side we used the definition of s_λ with operators $\mathcal{A}_{n+1}$, and on the right-hand side with operators $\mathcal{A}_n$. This identity is obvious by using the formula defining Schur polynomials as ratios of determinants (we then develop the determinants along the last column). So, we have a sequence of symmetric polynomials compatible with one another, and there is for any integer partition $\lambda \in \mathfrak{Y}$ a unique symmetric function $s_\lambda(X)$ with degree $|\lambda|$ such that

$$\pi_n(s_\lambda) = \frac{\mathcal{A}_n(x^{\lambda+\rho_n})}{\mathcal{A}_n(x^{\rho_n})}$$

for any $n \geq \ell(\lambda)$. We call s_λ the *Schur function* with label λ; the Schur functions form another graded basis $(s_\lambda)_{\lambda\in\mathfrak{Y}}$ of Sym.

Example 2.13. Consider the integer partition $\lambda = (2,1)$. The corresponding Schur function has degree 3, so it can be identified by looking at the case of 3 variables x_1, x_2, x_3. A simple computation of determinants shows that

$$\begin{aligned} s_{(2,1)}(x_1,x_2,x_3) &= 2\,x_1x_2x_3 + (x_1)^2x_2 + (x_1)^2x_3 + (x_2)^2x_1 + (x_2)^2x_3 + (x_3)^2x_1 + (x_3)^2x_2 \\ &= 2\,m_{(1,1,1)}(x_1,x_2,x_3) + m_{(2,1)}(x_1,x_2,x_3); \end{aligned}$$

therefore, $s_{(2,1)}(X) = 2\,m_{(1,1,1)}(X) + m_{(2,1)}(X)$ in the algebra of symmetric functions Sym.

Remark 2.14. The five bases of the algebra of symmetric functions can easily be manipulated with a computer algebra system. We recommend in particular `SageMath`, which is built upon `Python` and contains numerous packages which enable computations with symmetric functions. For instance, the conversion of the Schur function $s_{(3,1)}$ in the other bases of Sym is performed as follows:

```
Sym = SymmetricFunctions(QQ)
s = Sym.schur()
p = Sym.power()
m = Sym.monomial()
e = Sym.elementary()
h = Sym.homogeneous()
p(s([3,1]))
>> 1/8*p[1, 1, 1, 1] + 1/4*p[2, 1, 1] - 1/8*p[2, 2] - 1/4*p[4]

m(s([3,1]))
>> 3*m[1, 1, 1, 1] + 2*m[2, 1, 1] + m[2, 2] + m[3, 1]

e(s([3,1]))
>> e[2, 1, 1] - e[2, 2] - e[3, 1] + e[4]

h(s([3,1]))
>> h[3, 1] - h[4]
```

There are three important properties of Schur functions which we shall use hereafter in order to relate Sym to the representations of the symmetric groups. The first identity is the change of basis relation between Schur functions and homogeneous functions:

Proposition 2.15 (Jacobi–Trudy formula). *For any integer partition λ with length $\ell(\lambda) \leq n$,*

$$s_\lambda = \det((h_{\lambda_i - i + j})_{1\leq i,j\leq n})$$

where by convention $h_{-k} = 0$ if $k \geq 1$.

Proof. We prove the formula with a finite number of variables $n \geq \ell(\lambda)$; this is sufficient, because the polynomials $s_\lambda(x_1,\ldots,x_n)$ with $n \geq \ell(\lambda)$ determine the symmetric function s_λ. For

$i,j \in [\![1,N]\!]$, let us introduce the symmetric polynomials

$$\begin{aligned} e_{i,j} &= e_j(x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n); \\ h_j &= h_j(x_1,\ldots,x_n). \end{aligned}$$

If $H(z) = \sum_{j=0}^{\infty} h_j\, z^j$ and $E_i(z) = \sum_{j=0}^{n-1} e_{i,j}\, z^j$, then

$$E_i(-z)\, H(z) = \left(\prod_{i' \neq i} 1 - x_{i'}\, z\right)\left(\prod_{i'} \frac{1}{1 - x_{i'}\, z}\right) = \frac{1}{1 - x_i z} = \sum_{k=0}^{\infty} (x_i z)^k.$$

Fix a sequence of non-negative integers $(k_1, k_2, \ldots, k_n)$. If we introduce the two matrices $M = ((-1)^{n-j}\, e_{i,n-j})_{1 \leq i,j \leq n}$ and $H = (h_{k_j - i + n})_{1 \leq i,j \leq n}$, then the identity above rewrites as:

$$(x_i)^{k_j} = [z^{k_j}] \left(\sum_{k=0}^{\infty} (x_i z)^k\right) = \sum_{a=1}^{n} (-1)^{n-a} e_{i,n-a}\, h_{k_j + a - n} = (M\, H)_{i,j}.$$

In particular, with $k_j = \lambda_j + n - j = (\lambda + \rho_n)_j$,

$$\mathcal{A}_n(x^{\lambda + \rho_n}) = \det((x_i^{\lambda_j + n - j})_{1 \leq i,j \leq n}) = (\det M)(\det H) = (\det M)\, \det((h_{\lambda_j - j + i})_{1 \leq i,j \leq n}).$$

The case $\lambda = 0$ gives $\mathcal{A}_n(x^{\rho_n}) = \det M$, so by taking the ratio of the two antisymmetric functions we obtain the claimed identity. □

Proposition 2.16 (Cauchy identity). *Consider two independent alphabets of commuting variables $X = \{x_1, x_2, \ldots\}$ and $Y = \{y_1, y_2, \ldots\}$. We have the following identity of formal power series:*

$$\sum_{\lambda \in \mathfrak{Y}} s_\lambda(X)\, s_\lambda(Y) = \prod_{i,j \geq 1} \frac{1}{1 - x_i y_j}.$$

Proof. It suffices to prove the same formula with two finite sets of variables $\{x_1, x_2, \ldots, x_n\}$ and $\{y_1, y_2, \ldots, y_n\}$, for any $n \geq 1$; this allows us to use the definition of Schur polynomials as ratios of antisymmetric polynomials. We have

$$\begin{aligned} \mathcal{A}_n(x^{\rho_n})\, \mathcal{A}_n(y^{\rho_n}) \prod_{1 \leq i,j \leq n} \frac{1}{1 - x_i y_j} &= \mathcal{A}_n(x^{\rho_n})\, \mathcal{A}_n(y^{\rho_n}) \prod_{1 \leq j \leq n} \left(\sum_{k_j = 0}^{\infty} h_{k_j}(x_1, \ldots, x_n)\, (y_j)^{k_j}\right) \\ &= \mathcal{A}_n(x^{\rho_n}) \sum_{\substack{(k_1,\ldots,k_n) \in \mathbb{N}^n \\ \sigma \in \mathfrak{S}(n)}} \varepsilon(\sigma)\, h_k(x_1, \ldots, x_n)\, y^{k + \sigma \cdot \rho_n} \end{aligned}$$

where $h_k = h_{k_1} h_{k_2} \cdots h_{k_n}$, and $y^{k + \sigma \cdot \rho_n} = (y_1)^{k_1 + \rho_{n,\sigma(1)}} (y_2)^{k_2 + \rho_{n,\sigma(2)}} \cdots (y_n)^{k_n + \rho_{n,\sigma(n)}}$. We set $l = k + \sigma \cdot \rho_n = (k_j + n - \sigma(j))_{1 \leq j \leq n}$. With the convention that $h_{-k}(X) = 0$ for $k \geq 1$, we can rewrite the quantity above as:

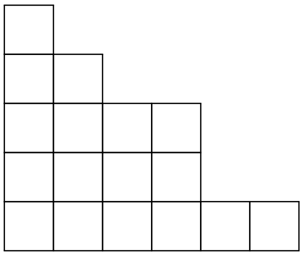
$$\mathcal{A}_n(x^{\rho_n}) \sum_{\substack{(l_1,\ldots,l_n) \in \mathbb{N}^n \\ \sigma \in \mathfrak{S}(n)}} \varepsilon(\sigma)\, h_{l - \sigma \cdot \rho_n}(x_1, \ldots, x_n)\, y^l = \sum_{(l_1,\ldots,l_n) \in \mathbb{N}^n} \mathcal{A}_n(x^l)\, y^l$$

by using the Jacobi–Trudy formula. If $l = (l_1, \ldots, l_n)$ is a sequence of integers with two terms $l_i = l_j$, then $\mathcal{A}_n(x^l) = 0$. Otherwise, there exists a unique permutation $\sigma \in \mathfrak{S}(n)$ and a unique integer partition $\lambda \in \mathfrak{Y}$ such that $l = \sigma \cdot (\lambda + \rho_n)$, so

$$\begin{aligned}\mathcal{A}_n(x^{\rho_n})\,\mathcal{A}_n(y^{\rho_n}) \prod_{1 \leq i,j \leq n} \frac{1}{1 - x_i y_j} &= \sum_{\lambda \in \mathfrak{Y}} \sum_{\sigma \in \mathfrak{S}(n)} \mathcal{A}_n(x^{\sigma(\lambda+\rho_n)})\, y^{\sigma(\lambda+\rho_n)} \\ &= \sum_{\lambda \in \mathfrak{Y}} \sum_{\sigma \in \mathfrak{S}(n)} \varepsilon(\sigma)\,\mathcal{A}_n(x^{\lambda+\rho_n})\, y^{\sigma(\lambda+\rho_n)} \\ &= \sum_{\lambda \in \mathfrak{Y}} \mathcal{A}_n(x^{\lambda+\rho_n})\,\mathcal{A}_n(y^{\lambda+\rho_n})\end{aligned}$$

and we conclude by dividing the two sides by $\mathcal{A}_n(x^{\rho_n})\,\mathcal{A}_n(y^{\rho_n})$. $\square$

The last important property of Schur functions is a special case of the so-called *Pieri rules*. It will also allow us to introduce the notion of *Young diagram* of an integer partition. Given a partition $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_\ell)$, its Young diagram is the array of boxes with λ_1 cells on the first bottom row, λ_2 cells on the second row just above, *etc.* (with $\ell = \ell(\lambda)$ rows of boxes). This is better understood with a drawing:

$(6,4,4,3,1) =$  .

If $\lambda \in \mathfrak{Y}(n)$ and $\Lambda \in \mathfrak{Y}(n+1)$, we denote $\lambda \nearrow \Lambda$ if the Young diagram of Λ can be obtained from the Young diagram of λ by adding one cell. For instance, the previous integer partition of size 17 satisfies $\lambda \nearrow \Lambda$ for the following integer partitions $\Lambda \in \mathfrak{Y}(18)$:

$$(6,4,4,2,1,1),\ (6,4,4,2,2),\ (6,4,4,3,1),\ (6,5,4,2,1),\ (7,4,4,2,1).$$

Proposition 2.17 (Pieri rule). *For any integer partition λ,*

$$s_\lambda(X)\, s_{(1)}(X) = \sum_{\Lambda \,|\, \lambda \nearrow \Lambda} s_\Lambda(X).$$

Proof. Note that in degree 1, all the functions of the five bases of Sym correspond:

$$p_1(X) = e_1(X) = h_1(X) = m_1(X) = s_1(X) = \sum_{i=1}^{\infty} x_i.$$

We prove the formula of the proposition with a large enough number of variables $n > \ell(\lambda)$. We have

$$\begin{aligned}\mathcal{A}_n(x^{\rho_n})\, s_\lambda(x_1,\ldots,x_n)\, s_1(x_1,\ldots,x_n) &= \sum_{\sigma \in \mathfrak{S}(n)} \sum_{j=1}^{n} \varepsilon(\sigma)\, x_j\ \sigma \cdot \left((x_1)^{\lambda_1+n-1}(x_2)^{\lambda_2+n-2}\cdots(x_n)^{\lambda_n}\right) \\ &= \sum_{\sigma \in \mathfrak{S}(n)} \sum_{j=1}^{n} \varepsilon(\sigma)\ \sigma \cdot \left(x_j\,(x_1)^{\lambda_1+n-1}(x_2)^{\lambda_2+n-2}\cdots(x_n)^{\lambda_n}\right) \\ &= \sum_{j=1}^{n} \mathcal{A}_n(x^{\rho_n+\lambda+(0,\ldots,1_j,\ldots,0)}).\end{aligned}$$

If $\lambda_{j-1} = \lambda_j$, then $(\rho_n+\lambda+(0,\ldots,1_j,\ldots,0))_{j-1} = (\rho_n+\lambda+(0,\ldots,1_j,\ldots,0))_j$ and the antisymmetrisation of the monomial vanishes. Otherwise, $\rho_n + \lambda + (0,\ldots,1_j,\ldots,0) = \rho_n + \Lambda$, where Λ is the integer partition obtained from λ by adding a cell on the j-th row. We conclude by dividing both sides of the equation by $\mathcal{A}_n(x^{\rho_n})$. $\square$

2. The representation ring of the symmetric groups

In order to establish a link between the ring of symmetric functions Sym and the representations of the symmetric groups, we start with the following basic observation:

LEMMA 2.18. *For any $n \in \mathbb{N}$, there is a bijection between the set $\widehat{\mathfrak{S}(n)}$ of irreducible representations of $\mathfrak{S}(n)$ and the set $\mathfrak{Y}(n)$ of integer partitions of size n.*

PROOF. By Proposition 1.24, it suffices to show that there is a bijection between the conjugacy classes in $\mathfrak{S}(n)$ and the integer partitions in $\mathfrak{Y}(n)$. Given a permutation $\sigma \in \mathfrak{S}(n)$, its *cycle-type* $t(\sigma)$ is the integer partition whose parts are the sizes of the cycles in the decomposition of σ as a product of disjoint cycles (with fixed points counted as cycles of size 1). For instance, if $\sigma = (1, 5, 2, 9)(3, 6)(4, 8)(7)$ in $\mathfrak{S}(9)$, then $t(\sigma) = (4, 2, 2, 1)$. Notice then that given a permutation τ and a cycle $(c_1, c_2, \ldots, c_r)$, one has

$$\tau\,(c_1, c_2, \ldots, c_r)\,\tau^{-1} = (\tau(c_1), \tau(c_2), \ldots, \tau(c_r)).$$

Moreover, the conjugate of a product is the product of conjugates; therefore, two permutations σ and σ' are conjugated by some permutation τ if and only if they have the same cycle-type. So, the map $\sigma \mapsto t(\sigma)$ induces a bijection between conjugacy classes in $\mathfrak{S}(n)$ and integer partitions in $\mathfrak{Y}(n)$. □

Denote $(S^\lambda)_{\lambda \in \mathfrak{Y}(n)}$ the collection of non-isomorphic irreducible representations of $\mathfrak{S}(n)$; for the moment we only know that there is a way to label the irreducibles by integer partitions, but later on we shall have a more precise description of the bijection. The discussion of Chapter 1 shows that any representation of $\mathfrak{S}(n)$ writes uniquely as a direct sum

$$\bigoplus_{\lambda \in \mathfrak{Y}(n)} m_\lambda\, S^\lambda, \quad m_\lambda \in \mathbb{N}.$$

We can see the direct sum above as an element of a $\mathbb{R}$-vector space with dimension $\mathrm{card}(\mathfrak{Y}(n))$, and with a linear basis denoted S^λ, $\lambda \in \mathfrak{Y}(n)$. Let us denote this vector space $\mathrm{R}(\mathfrak{S}(n))$; this is the *Grothendieck group* of the category of representations of the symmetric group $\mathfrak{S}(n)$ (with coefficients taken in the field of real numbers $\mathbb{R}$). The Grothendieck group of the representations of the symmetric groups is the direct sum of these spaces:

$$\mathrm{R}(\mathfrak{S}) = \bigoplus_{n \in \mathbb{N}} \mathrm{R}(\mathfrak{S}(n)).$$

This is a graded vector space, with a graded linear basis $(S^\lambda)_{\lambda \in \mathfrak{Y}}$ labelled by the set of all integer partitions. We are going to show that there is a nice isomorphism of graded vector spaces between Sym and $\mathrm{R}(\mathfrak{S})$, which enables many computations around the representations of the symmetric groups. This isomorphism will be:

- an isomorphism of rings with respect to a certain product on $\mathrm{R}(\mathfrak{S})$;
- and an isometry with respect to two scalar products on Sym and $\mathrm{R}(\mathfrak{S})$.

The next paragraphs are devoted to the construction of these structures.

Given a group G and a subgroup H, there is a natural way to *restrict* a representation of G in order to get a representation of H. Indeed, given (V, ρ) with $\rho : G \to \mathrm{GL}(V)$ morphism of groups, the restriction $\rho_{|H} : H \to \mathrm{GL}(V)$ yields a representation of the subgroup. This new representation is usually denoted $\mathrm{Res}_H^G(V)$; it has the same underlying vector space V as the original representation. Conversely, given a representation W of the subgroup H, there is a natural way to *induce* from it a representation of G. Consider W as a $\mathbb{C}H$-module; then, $\mathbb{C}H$ is a subalgebra of $\mathbb{C}G$, and we can take the tensor product

$$V = \mathbb{C}G \otimes_{\mathbb{C}H} W,$$

which is a vector space with dimension $\frac{\operatorname{card} G}{\operatorname{card} H} \times \dim W$. This new vector space is endowed with a natural structure of $\mathbb{C}G$-module (for the multiplication on the left), so it is a representation of G. We denote it by $\operatorname{Ind}_H^G(W)$. Note that the induction of representations is not the opposite operation of restriction; the two functors are rather related by a relation of duality, see in particular the proof of Proposition 2.31.

Lemma 2.19 (Character of an induced representation). *Let $H \subset G$ be two groups and W be a representation of H. The character of the induced representation $V = \operatorname{Ind}_H^G(W)$ is given by the following formula:*

$$\operatorname{ch}^V(g) = \sum_{k \in G/H} \operatorname{ch}^W(k^{-1}gk),$$

where the sum runs over a family of representatives of the left cosets kH of H in G, and where by convention $\operatorname{ch}^W(g') = 0$ if g' is not in H.

Proof. Fix a linear basis $(w_j)_{w \in [\![1,d]\!]}$ of W over $\mathbb{C}$. Given a family $\{k_1, \ldots, k_r\}$ of representatives of the left cosets of G/H, any element of G writes uniquely as $g = k_i h$ for some $i \in [\![1,r]\!]$ and some $h \in H$. Then, a linear basis of $V = \mathbb{C}G \otimes_{\mathbb{C}H} W$ over $\mathbb{C}$ consists in the tensors $k_i \otimes w_j$ with $i \in [\![1,r]\!]$ and $j \in [\![1,d]\!]$. Let us use this basis in order to compute the trace of the action of an element g. Given $i \in [\![1,r]\!]$ and $g \in G$, a product gk_i writes uniquely as $k_{i'(i,g)}\, h(i,g)$ with $i'(i,g) \in [\![1,r]\!]$ and $h(i,g) \in H$. We then have:

$$\operatorname{ch}^V(g) = \sum_{i=1}^{r} \sum_{j=1}^{d} [k_i \otimes w_j](gk_i \otimes w_j),$$

where $[k_i \otimes w_j](x)$ denotes the coefficient of $k_i \otimes w_j$ in $x \in V$. We rewrite the quantity above as follows:

$$\begin{aligned}\operatorname{ch}^V(g) &= \sum_{i=1}^{r} \sum_{j=1}^{d} [k_i \otimes w_j](k_{i'(i,g)} \otimes h(i,g)w_j) = \sum_{i=1}^{r} 1_{(i=i'(i,g))} \left(\sum_{j=1}^{d} [w_j](h(i,g)w_j) \right) \\ &= \sum_{i=1}^{r} 1_{(i=i'(i,g))} \operatorname{ch}^W(h(i,g)).\end{aligned}$$

Finally, $i = i'(h,g)$ if and only if $gk_i = k_i h(i,g)$, or equivalently if and only if $h(i,g) = k_i^{-1} g k_i$. So, when taking the sum over all the representatives $k_1, \ldots, k_r$ of G/H, the non-zero terms are those such that $k_i^{-1} g k_i$ belongs to H. ☐

Consider two irreducible representations S^λ and S^μ respectively of $\mathfrak{S}(m)$ and $\mathfrak{S}(n)$. Let us remark that the product group $\mathfrak{S}(m) \times \mathfrak{S}(n)$ can be considered as a subgroup of $\mathfrak{S}(m+n)$: the first group $\mathfrak{S}(m)$ acts on $[\![1,m]\!]$ and the second group $\mathfrak{S}(n)$ acts on $[\![m+1,m+n]\!]$.

Definition 2.20 (Product of representations of symmetric groups). *The product $S^\lambda \times S^\mu$ is defined as*

$$\operatorname{Ind}_{\mathfrak{S}(m)\times\mathfrak{S}(n)}^{\mathfrak{S}(m+n)}(S^\lambda \otimes S^\mu),$$

where $S^\lambda \otimes S^\mu$ is the exterior tensor product of the two representations S^λ and S^μ. The product $S^\lambda \times S^\mu$ can be considered as an element of $\mathrm{R}(\mathfrak{S}(m+n))$, and by extending the formula above by bilinearity, we get a graded product:

$$\mathrm{R}(\mathfrak{S}) \times \mathrm{R}(\mathfrak{S}) \to \mathrm{R}(\mathfrak{S})$$

which is associative and which makes $\mathrm{R}(\mathfrak{S})$ into a graded real algebra.

The only point which is not entirely trivial is the associativity of the product; it is an easy consequence of the relation $\mathrm{Ind}_H^G \circ \mathrm{Ind}_I^H = \mathrm{Ind}_I^G$ for a sequence of subgroups $I \subset H \subset G$. This relation is quite easy to prove by using the definition of induced representations with tensor products.

Let us now endow Sym and $\mathrm{R}(\mathfrak{S})$ with two inner products. On the side of the representations of the symmetric groups, we have a natural pairing in $\mathrm{R}(\mathfrak{S}(n))$:

$$\left\langle S^\lambda \,\middle|\, S^\mu \right\rangle = \dim(\hom_{\mathfrak{S}(n)}(S^\lambda, S^\mu)) = 1_{(\lambda=\mu)}.$$

If we extend this rule by linearity to the vector space $\mathrm{R}(\mathfrak{S}(n))$ and if we convene that the spaces $\mathrm{R}(\mathfrak{S}(n))$ are orthogonal in $\mathrm{R}(\mathfrak{S}) = \bigoplus_{n\in\mathbb{N}} \mathrm{R}(\mathfrak{S}(n))$, then we get a scalar product on the representation ring $\mathrm{R}(\mathfrak{S})$, such that the irreducible representations $(S^\lambda)_{\lambda\in\mathfrak{Y}}$ form an orthonormal basis. The corresponding scalar product on Sym will be related to the Cauchy identity 2.16.

LEMMA 2.21 (Hall scalar product). *We endow* Sym *with the unique scalar product which makes* $(s_\lambda)_{\lambda\in\mathfrak{Y}}$ *an orthonormal basis (it is called the* Hall scalar product*). Then, given another graded family* $(u_\lambda)_{\lambda\in\mathfrak{Y}}$ *of symmetric functions (with each* u_λ *homogeneous of degree* $|\lambda|$*), the following assertions are equivalent:*

(1) The family $(u_\lambda)_{\lambda\in\mathfrak{Y}}$ *is a graded orthonormal basis of* Sym.

(2) The family $(u_\lambda)_{\lambda\in\mathfrak{Y}}$ *satisfies the Cauchy identity:*

$$\sum_{\lambda\in\mathfrak{Y}} u_\lambda(X)\, u_\lambda(Y) = \prod_{i,j} \frac{1}{1-x_i y_j}.$$

PROOF. Denote $(U_{\lambda\mu})_{\lambda,\mu}$ the matrix of the family $(u_\lambda)_{\lambda\in\mathfrak{Y}}$ in the basis of Schur functions:

$$u_\lambda(X) = \sum_\mu U_{\lambda\mu}\, s_\mu(X).$$

The first assertion is equivalent to the fact that, for any $n \in \mathbb{N}$, the matrix $(U_{\lambda\mu})_{\lambda,\mu\in\mathfrak{Y}(n)}$ is orthogonal. If this is the case, then

$$\begin{aligned}\sum_{\lambda\in\mathfrak{Y}} u_\lambda(X)\, u_\lambda(Y) &= \sum_{\lambda,\mu,\nu} U_{\lambda\mu}U_{\lambda\nu}\, s_\mu(X)\, s_\nu(Y)\\ &= \sum_{\mu,\nu} 1_{(\mu=\nu)}\, s_\mu(X)\, s_\nu(Y) = \sum_\mu s_\mu(X)\, s_\mu(Y) = \prod_{i,j}\frac{1}{1-x_i y_j}.\end{aligned}$$

Conversely, if $(u_\lambda)_{\lambda\in\mathfrak{Y}}$ satisfies the Cauchy identity, then $\sum_{\lambda\in\mathfrak{Y}} u_\lambda(X)\, u_\lambda(Y) = \sum_{\lambda\in\mathfrak{Y}} s_\lambda(X)\, s_\lambda(Y)$, and we recover the orthogonality of the matrices $(U_{\lambda\mu})_{\lambda,\mu\in\mathfrak{Y}(n)}$ by taking the scalar product of this relation with $s_\lambda(X) \otimes s_\lambda(Y)$ in $\mathrm{Sym}(X) \otimes_{\mathbb{R}} \mathrm{Sym}(Y)$. □

COROLLARY 2.22 (Orthogonality of the power sums). *The basis* $(p_\lambda)_{\lambda\in\mathfrak{Y}}$ *of power sums is an orthogonal basis of* Sym *for the scalar product previously defined, with*

$$\langle p_\lambda \mid p_\lambda\rangle = z_\lambda.$$

PROOF. Recall the relation $h_k(X) = \sum_{\lambda\in\mathfrak{Y}(k)} \frac{p_\lambda(X)}{z_\lambda}$; it implies that

$$\prod_i \frac{1}{1-x_i} = \sum_{k=0}^{\infty} h_k(X) = \sum_{\lambda\in\mathfrak{Y}} \frac{p_\lambda(X)}{z_\lambda}.$$

In the formula above, we replace the set of variables $X = \{x_1, x_2, \ldots\}$ by $XY = \{x_i y_j,\, i, j \geq 1\}$. On the left-hand side, we obtain the Cauchy product $\prod_{i,j} \frac{1}{1-x_i y_j}$. On the right-hand side, note that

$$p_k(XY) = \sum_{i,j} (x_i y_j)^k = \left(\sum_i (x_i)^k\right)\left(\sum_j (y_j)^k\right) = p_k(X)\, p_k(Y).$$

Therefore, $p_\lambda(XY) = p_\lambda(X)\, p_\lambda(Y)$, and we have proved the Cauchy identity for the family of functions $(\frac{p_\lambda}{\sqrt{z_\lambda}})_{\lambda \in \mathfrak{Y}}$. ☐

We can now state the main result of this chapter:

Theorem 2.23 (Frobenius–Schur). *There is an isomorphism of graded real algebras between* Sym *and* $\mathrm{R}(\mathfrak{S})$, *which sends the basis of Schur functions* $(s_\lambda)_{\lambda \in \mathfrak{Y}}$ *to the basis of irreducible representations of symmetric groups* $(S^\lambda)_{\lambda \in \mathfrak{Y}}$. *This isomorphism is also an isometry for the scalar products previously introduced.*

This theorem ensures that many computations on the irreducible representations of symmetric groups can be performed by using the combinatorial properties of the Schur functions, which are essentially polynomials. In particular, we shall use it in order to compute the dimensions $\dim \lambda = \dim S^\lambda$ of the irreducible representations, and the values of the irreducible characters. There is a slightly stronger version of Theorem 2.23 involving the structures of *positive self-adjoint Hopf algebras* on Sym and $\mathrm{R}(\mathfrak{S})$; see the exercises at the end of the chapter.

The proof of Theorem 2.23 starts with a reinterpretation of the Grothendieck groups of representations $\mathrm{R}(\mathfrak{S}(n))$. Given a finite group G and the Grothendieck group $\mathrm{R}(G)$ of its category of representations, the character map

$$\begin{aligned} \mathrm{ch} : \mathrm{R}(G) &\to \mathrm{Z}(\mathbb{C}G) \\ V &\mapsto \mathrm{ch}^V \end{aligned}$$

can be extended by linearity in order to get a linear isomorphism between $\mathbb{C} \otimes_{\mathbb{R}} \mathrm{R}(G)$ and $\mathrm{Z}(\mathbb{C}G)$; indeed, as explained in Chapter 1 and its exercises, the two vector spaces have the same dimension, and a representation is entirely determined by its character. As we are working with symmetric functions with real coefficients and with the real Grothendieck groups of the symmetric groups $\mathfrak{S}(n)$, it would be nice if this isomorphism could be restricted to an isomorphism of real vector spaces. The tool which enables this restriction is the following:

Lemma 2.24. *A character* ch^V *of a representation* (V, ρ) *of a symmetric group* $\mathfrak{S}(n)$ *takes all its values in the ring of integers* $\mathbb{Z}$.

We admit this lemma; its proof relies on a bit of Galois theory and is presented in the exercises at the end of the chapter. Later, we shall give a formula for the irreducible characters of the symmetric groups (Corollary 2.27) which clearly implies that they are rational numbers in $\mathbb{Q}$. As a consequence of the lemma above, we see that $\mathrm{ch} : \mathrm{R}(\mathfrak{S}(n)) \to \mathrm{Z}(\mathbb{C}\mathfrak{S}(n))$ takes in fact its values in the real subalgebra $\mathrm{Z}(\mathbb{R}\mathfrak{S}(n))$, and is a linear isomorphism of real vector spaces.

A linear basis of the center of the real group algebra $\mathbb{R}G$ of a finite group G consists in the conjugacy classes

$$c_g = \sum_{h = g' g (g')^{-1}} h.$$

In the case of the symmetric group $\mathfrak{S}(n)$, these conjugacy classes correspond to the cycle-types $\mu \in \mathfrak{Y}(n)$, so a linear basis of $\mathrm{Z}(\mathbb{R}\mathfrak{S}(n))$ consists in the classes

$$C_\mu = \sum_{\substack{\sigma \in \mathfrak{S}(n) \\ t(\sigma)=\mu}} \sigma$$

with μ running over the set of integer partitions $\mathfrak{Y}(n)$. Thus, $\mathrm{Z}(\mathbb{R}\mathfrak{S}(n)) = \bigoplus_{\mu \in \mathfrak{Y}(n)} \mathbb{R}C_\mu$ (we have for the same reasons $\mathrm{Z}(\mathbb{C}\mathfrak{S}(n)) = \bigoplus_{\mu \in \mathfrak{Y}(n)} \mathbb{C}C_\mu$).

LEMMA 2.25. *For $n \in \mathbb{N}$, consider the linear map:*

$$\Phi_n : \mathrm{Z}(\mathbb{R}\mathfrak{S}(n)) \to \mathrm{Sym}_n$$
$$C_\mu \mapsto \frac{p_\mu}{z_\mu}$$

This map is an isometry with respect to the restriction of the scalar product of $\mathbb{C}\mathfrak{S}(n)$ to $\mathrm{Z}(\mathbb{R}\mathfrak{S}(n))$, and to the restriction of Hall scalar product of Sym to Sym_n.

PROOF. Recall that the scalar product on a group algebra $\mathbb{C}G$ is defined by $\langle g \mid g' \rangle = \frac{1_{(g=g')}}{|G|}$. Two conjugacy classes C_μ and $C_{\mu'}$ with $\mu \neq \mu'$ involve disjoint sets of permutations, so they are orthogonal in $\mathrm{Z}(\mathbb{R}\mathfrak{S}(n))$; the same is true for $\frac{p_\mu}{z_\mu}$ and $\frac{p_{\mu'}}{z_{\mu'}}$ by Corollary 2.22. It suffices now to prove that

$$\langle C_\mu \mid C_\mu \rangle_{\mathrm{Z}(\mathbb{R}\mathfrak{S}(n))} = \left\langle \frac{p_\mu}{z_\mu} \,\middle|\, \frac{p_\mu}{z_\mu} \right\rangle_{\mathrm{Sym}}.$$

On the right-hand side, by Corollary 2.22, we have $\frac{1}{z_\mu}$. On the left-hand side, we have $\frac{\mathrm{card}\, C_\mu}{n!}$. The integer partition μ being fixed, let us consider the map

$$\mathfrak{S}(n) \to C_\mu$$
$$a_1a_2\ldots a_n \mapsto (a_1,\ldots,a_{\mu_1})(a_{\mu_1+1},\ldots,a_{\mu_1+\mu_2})\cdots(a_{\mu_1+\cdots+\mu_{\ell-1}+1},\ldots,a_n).$$

This map is surjective, and every permutation $\sigma \in C_\mu$ is obtained z_μ times: without changing the permutation σ, we can permute cyclically the letters of each cycle (this contributes to a factor $\prod_{i\geq 1} i^{m_i(\mu)}$), and we can also permute the cycles with the same length (this contributes to a factor $\prod_{i \geq 1}(m_i(\mu))!$). Therefore, we have indeed:

$$\mathrm{card}\, C_\mu = \frac{n!}{z_\mu} \quad ; \quad \langle C_\mu \mid C_\mu \rangle_{\mathrm{Z}(\mathbb{R}\mathfrak{S}(n))} = \frac{1}{z_\mu}. \qquad \square$$

We can now define the isomorphism of Theorem 2.23. Let $\mathrm{ch}_n : \mathrm{R}(\mathfrak{S}(n)) \to \mathrm{Z}(\mathbb{R}\mathfrak{S}(n))$ be the character map, and $\Psi_n = \Phi_n \circ \mathrm{ch}_n : \mathrm{R}(\mathfrak{S}(n)) \to \mathrm{Sym}_n$. We aggregate these maps in a graded linear map $\Psi : \mathrm{R}(\mathfrak{S}) \to \mathrm{Sym}$, such that $\Psi_{|\mathrm{R}(\mathfrak{S}(n))} = \Psi_n$. Let us check that Ψ satisfies the hypotheses of the Frobenius–Schur theorem. By the previous lemma, we already know that it is an isometry. In order to check that it is compatible with the product, let us fix two representations V and W of the symmetric groups $\mathfrak{S}(m)$ and $\mathfrak{S}(n)$. We have:

$$\Psi_m(V) = \Phi_m(\mathrm{ch}^V) = \Phi_m\left(\sum_{\sigma \in \mathfrak{S}(m)} \mathrm{ch}^V(\sigma)\,\sigma\right)$$
$$= \Phi_m\left(\sum_{\mu \in \mathfrak{Y}(m)} \mathrm{ch}^V(\mu)\, C_\mu\right) = \sum_{\mu \in \mathfrak{Y}(m)} \mathrm{ch}^V(\mu)\,\frac{p_\mu}{z_\mu},$$

where $\mathrm{ch}^V(\mu)$ denotes the character of the representation V evaluated on any permutation with cycle-type μ. It is convenient to rewrite the formula above as:

$$\Psi_m(V) = \frac{1}{m!} \sum_{\sigma \in \mathfrak{S}(m)} \mathrm{ch}^V(\sigma)\, p_{t(\sigma)}.$$

Then, by using the formula for the character of an induced representation, and the easy fact that $\mathrm{ch}^{V \otimes W}(\sigma_1, \sigma_2) = \mathrm{ch}^V(\sigma_1)\, \mathrm{ch}^W(\sigma_2)$, we obtain:

$$\begin{aligned}\Psi_{m+n}(V \times W) &= \frac{1}{(m+n)!} \sum_{\sigma \in \mathfrak{S}(m+n)} \mathrm{ch}^{V \times W}(\sigma)\, p_{t(\sigma)} \\ &= \frac{1}{(m+n)!} \sum_{\sigma \in \mathfrak{S}(m+n)} \sum_{\substack{\tau \in \mathfrak{S}(m+n)/(\mathfrak{S}(m) \times \mathfrak{S}(n)) \\ \tau^{-1}\sigma\tau = (\sigma_1, \sigma_2) \in \mathfrak{S}(m) \times \mathfrak{S}(n)}} \mathrm{ch}^V(\sigma_1)\, \mathrm{ch}^W(\sigma_2)\, p_{t(\sigma)}\end{aligned}$$

If we fix $(\sigma_1, \sigma_2) \in \mathfrak{S}(m) \times \mathfrak{S}(n)$, then the permutation $\sigma = \tau(\sigma_1, \sigma_2)\tau^{-1}$ is entirely determined by τ and satisfies $t(\sigma) = t(\sigma_1) \sqcup t(\sigma_2)$. There are $\frac{(m+n)!}{m!\,n!}$ possibilities for $\tau \in \mathfrak{S}(m+n)/(\mathfrak{S}(m) \times \mathfrak{S}(n))$, therefore,

$$\Psi_{m+n}(V \times W) = \frac{1}{m!\,n!} \sum_{(\sigma_1, \sigma_2) \in \mathfrak{S}(m) \times \mathfrak{S}(n)} \mathrm{ch}^V(\sigma_1)\, \mathrm{ch}^W(\sigma_2)\, p_{t(\sigma_1)}\, p_{t(\sigma_2)} = \Psi_m(V)\, \Psi_n(W).$$

To end the proof of Theorem 2.23, we have to show that Ψ sends the irreducible representations of the symmetric groups to the Schur functions.

LEMMA 2.26. *For any integer partition $\lambda \in \mathfrak{Y}(n)$, s_λ is the image by the map Ψ of an integral linear combination of irreducible representations of $\mathfrak{S}(n)$.*

PROOF. Denote 1_n the trivial representation of $\mathfrak{S}(n)$ on $\mathbb{C}$, where every permutation acts by the identity. We have

$$\Psi_n(1_n) = \frac{1}{n!} \sum_{\sigma \in \mathfrak{S}(n)} p_{t(\sigma)} = \sum_{\mu \in \mathfrak{Y}(n)} \frac{p_\mu}{z_\mu} = h_n.$$

Since Ψ is compatible with products, for any product of homogeneous functions $h_{n_1} h_{n_2} \cdots h_{n_r}$, there is a representation sent to this product by Ψ, namely, $1_{n_1} \times 1_{n_2} \times \cdots \times 1_{n_r}$. Now, by the Jacobi–Trudy formula 2.15, $s_\lambda = \det((h_{\lambda_i - i + j})_{1 \leq i,j \leq \ell(\lambda)})$, so s_λ is the image by Ψ of a $\mathbb{Z}$-linear combination of representations of symmetric groups. $\square$

END OF THE PROOF OF THEOREM 2.23. Denote $\Psi^{-1}(s_\lambda) = \sum_{|\mu|=|\lambda|} c_{\lambda\mu}\, S^\mu$; the coefficients $c_{\lambda\mu}$ belong to $\mathbb{Z}$. Since Ψ is an isometry and the two bases $(s_\lambda)_{\lambda \in \mathfrak{Y}}$ and $(S^\lambda)_{\lambda \in \mathfrak{Y}}$ are orthonormal in their respective spaces,

$$1 = \|s_\lambda\|^2 = \left\| \sum_{|\mu|=|\lambda|} c_{\lambda\mu}\, S^\mu \right\|^2 = \sum_{|\mu|=|\lambda|} (c_{\lambda\mu})^2.$$

With integral coefficients, this is only possible if, for any λ, there is a unique integer partition $\mu = \mu(\lambda)$ such that $c_{\lambda\mu} \neq 0$; moreover, $c_{\lambda\mu} = \pm 1$. As the Schur functions (respectively, the irreducible representations) are mutually orthogonal, we also have that $\mu(\lambda) \neq \mu(\lambda')$ if $\lambda \neq \lambda'$. Therefore, up to a reindexation of the irreducible representations of $\mathfrak{S}(n)$, we can assume that $c_{\lambda\lambda} = \pm 1$ for any integer partition λ, and that $c_{\lambda\mu} = 0$ if $\lambda \neq \mu$. So, Ψ sends the Schur functions to the irreducible representations of the symmetric groups *possibly up to a sign*. To show that this sign is always $+1$, we use finally the Pieri rule 2.17. Fix an integer partition $\lambda \in \mathfrak{Y}(n)$, and set $\Psi(S^\lambda) = \varepsilon(\lambda)\, s_\lambda$. We have:

$$\varepsilon(\lambda)\, s_\lambda = \Psi_n(S^\lambda) = \Phi_n(\mathrm{ch}^\lambda) = \Phi_n\left(\sum_{\mu \in \mathfrak{Y}(n)} \mathrm{ch}^\lambda(\mu)\, C_\mu\right) = \sum_{\mu \in \mathfrak{Y}(n)} \mathrm{ch}^\lambda(\mu)\, \frac{p_\mu}{z_\mu},$$

where $\mathrm{ch}^\lambda(\mu)$ denotes the value of the character of S^λ on a permutation with cycle-type μ. Therefore,

$$\varepsilon(\lambda)\, \langle s_\lambda \,|\, (p_1)^n \rangle = \mathrm{ch}^\lambda(\mathrm{id}_{[\![1,n]\!]}) = \dim(S^\lambda) > 0.$$

However, by recursive use of the Pieri rule, $(p_1)^n$ is a positive linear combination of Schur functions; therefore, $\varepsilon(\lambda) = +1$. $\square$

3. Dimensions and irreducible characters

Let us give several important applications of the Frobenius–Schur theorem. The first application is a formula for the values of the irreducible characters of the symmetric groups:

Corollary 2.27 (From Schur functions to power sums). *Denote* $\mathrm{ch}^\lambda(\mu)$ *the value of the character of the irreducible representation* S^λ *of* $\mathfrak{S}(n)$ *on a permutation with cycle-type* μ*. We have*

$$\mathrm{ch}^\lambda(\mu) = \langle s_\lambda \,|\, p_\mu \rangle\,.$$

Therefore,

$$p_\mu(X) = \sum_{\lambda \,|\, |\lambda| = |\mu|} \mathrm{ch}^\lambda(\mu)\, s_\lambda(X);$$

$$s_\lambda(X) = \sum_{\mu \,|\, |\lambda| = |\mu|} \mathrm{ch}^\lambda(\mu)\, \frac{p_\mu(X)}{z_\mu}.$$

Proof. Since Ψ is an isometry,

$$\langle s_\lambda \,|\, p_\mu \rangle = \big\langle S^\lambda \,\big|\, \Psi(p_\mu) \big\rangle = \big\langle \mathrm{ch}^\lambda \,\big|\, z_\mu\, C_\mu \big\rangle = \frac{z_\mu \,\mathrm{card}\, C_\mu}{n!}\, \mathrm{ch}^\lambda(\mu) = \mathrm{ch}^\lambda(\mu).$$

The change of basis formulæ follow immediately from this, since $(s_\lambda)_{\lambda \in \mathfrak{Y}}$ and $(p_\mu)_{\mu \in \mathfrak{Y}}$ are orthogonal bases of Sym. $\square$

Since the beginning of this chapter, we have worked with polynomials and symmetric functions with coefficients in $\mathbb{R}$, but clearly we could have worked with $\mathbb{Q}$; in particular, all the coefficients of the matrices of change of basis are rational numbers. Therefore, the previous corollary shows that the character values $\mathrm{ch}^\lambda(\mu)$ are always rational numbers (so maybe there is a way to prove a "complex" Frobenius–Schur theorem without Lemma 2.24, and then come back to the real version by using the argument above).

Another consequence of the Frobenius–Schur theorem is a formula for the dimensions of the irreducible representations S^λ. Given an integer partition $\lambda \in \mathfrak{Y}(n)$, we call *standard tableau* with shape λ a numbering of the cells of the Young diagram of λ by the integers in $[\![1, n]\!]$, such that every row and every column is strictly increasing. For instance,

10					
7	12				
6	11	15	16		
2	5	9	14		
1	3	4	8	13	17

is a standard tableau with shape $\lambda = (6, 4, 4, 2, 1)$. Let us remark that a standard tableau with shape λ corresponds to a sequence $\emptyset = \lambda^{(0)} \nearrow \lambda^{(1)} \nearrow \lambda^{(2)} \nearrow \cdots \nearrow \lambda^{(n)} = \lambda$, each pair $\lambda^{(i-1)} \nearrow \lambda^{(i)}$ of successive Young diagrams corresponding to the addition of the cell marked by the integer i. Denote $\mathrm{ST}(\lambda)$ the set of all standard tableaux with shape λ.

PROPOSITION 2.28 (Dimension of the irreducibles). *The dimension of the irreducible representation S^λ of $\mathfrak{S}(n)$ is the cardinality of $\mathrm{ST}(\lambda)$.*

PROOF. The dimension of S^λ is the trace of the action of $\mathrm{id}_{[\![1,n]\!]} \in \mathfrak{S}(n)$ on S^λ, so it is equal to

$$\mathrm{ch}^\lambda(1^n) = \langle s_\lambda \mid (p_1)^n \rangle .$$

We are therefore looking for the expansion in Schur functions of the symmetric function $(p_1(X))^n$. We claim that for any n,

$$(p_1(X))^n = \sum_{\lambda \in \mathfrak{Y}(n)} (\mathrm{card}\, \mathrm{ST}(\lambda))\, s_\lambda(X).$$

If we interpret the standard tableaux as increasing chains of Young diagrams, then the right-hand side of the formula above rewrites as a sum over all possible chains of length n:

$$\sum_{\emptyset = \lambda^{(0)} \nearrow \lambda^{(1)} \nearrow \cdots \nearrow \lambda^{(n)}} s_{\lambda^{(n)}}(X).$$

The identity can then be proved by induction on n, by an immediate application of the Pieri rule 2.17. □

COROLLARY 2.29 (Plancherel formula). *For any $n \in \mathbb{N}$,*

$$n! = \sum_{\lambda \in \mathfrak{Y}(n)} (\mathrm{card}\, \mathrm{ST}(\lambda))^2.$$

PROOF. This follows immediately from Theorem 1.17 and from the previous proposition. □

This remarkable formula ensures that there is a bijection between the set of permutations of size n, and the set of pairs (P, Q) of standard tableaux of size n and with the same shape. We shall make this bijection explicit in Chapter 4, which will be devoted to the study of the natural probability measure on $\mathfrak{Y}(n)$ stemming from the Plancherel formula.

EXAMPLE 2.30 (Representations of $\mathfrak{S}(3)$, revisited). In Example 1.16, we have identified the three irreducible representations of $\mathfrak{S}(3)$. They correspond to the integer partitions of size 3, which are (3), $(2, 1)$ and $(1, 1, 1)$. Let us make the bijection precise.

- For any n, $s_{(n)} = h_n$ (this is the Jacobi–Trudy formula for an integer partition with length n), and we saw during the proof of the Frobenius–Schur theorem that h_n is the image of the trivial representation of $\mathfrak{S}(n)$ by the isomorphism Ψ. Therefore, for any n, the integer partition which labels the one-dimensional trivial representation of $\mathfrak{S}(n)$ on $\mathbb{C}$ is the integer partition (n) with one part. In particular, the representation denoted V_{trivial} in Example 1.16 is $S^{(3)}$.
- The geometric representation W has dimension 2, and this is also the case of $S^{(2,1)}$, since there are two standard tableaux with shape $(2, 1)$:

$$\begin{array}{|c|c} \hline 3 & \\ \hline 1 & \multicolumn{1}{|c|}{2} \\ \hline \end{array} \quad \text{and} \quad \begin{array}{|c|c} \hline 2 & \\ \hline 1 & \multicolumn{1}{|c|}{3} \\ \hline \end{array} .$$

On the other hand, the sign representation V_{sign} has dimension 1, and there is a unique standard tableau with shape $(1,1,1)$. By identification, we therefore have:

$$V_{\mathrm{sign}} = S^{(1,1,1)} \quad ; \quad W = S^{(2,1)}.$$

We have therefore identified the partitions which label the three kinds of irreducible representations of $\mathfrak{S}(3)$. The formula for characters allows us to compute the table of characters of W: since

$$s_{(2,1)}(X) = \frac{p_{(1,1,1)}(X)}{3} - \frac{p_3(X)}{3},$$

we obtain $\mathrm{ch}^{(2,1)}(1,1,1) = 2$, $\mathrm{ch}^{(2,1)}(2,1) = 0$ and $\mathrm{ch}^{(2,1)}(3) = -1$.

During the construction of the product operation on $\mathrm{R}(\mathfrak{S})$, we have introduced the notions of induced and restricted representations. A special case which becomes explicitly computable thanks to Theorem 2.23 is with the two groups $H = \mathfrak{S}(n)$ and $G = \mathfrak{S}(n+1)$, where H is viewed as the subgroup of G acting on $[\![1,n]\!] \subset [\![1,n+1]\!]$.

PROPOSITION 2.31 (Branching rules). *For any integer partition $\lambda \in \mathfrak{Y}(n)$,*

$$\mathrm{Ind}_{\mathfrak{S}(n)}^{\mathfrak{S}(n+1)}(S^\lambda) = \bigoplus_{\Lambda \,|\, \lambda \nearrow \Lambda} S^\Lambda.$$

For any integer partition $\Lambda \in \mathfrak{Y}(n+1)$,

$$\mathrm{Res}_{\mathfrak{S}(n)}^{\mathfrak{S}(n+1)}(S^\Lambda) = \bigoplus_{\lambda \,|\, \lambda \nearrow \Lambda} S^\lambda.$$

PROOF. The first part of the proposition comes from the fact that $\mathfrak{S}(n)$ can be identified with $\mathfrak{S}(n) \times \mathfrak{S}(1) \subset \mathfrak{S}(n+1)$, since $\mathfrak{S}(1) = \{1\}$ is the trivial group with one element. Under this identification, S^λ corresponds to the external tensor product $S^\lambda \otimes S^{(1)}$, therefore,

$$\mathrm{Ind}_{\mathfrak{S}(n)}^{\mathfrak{S}(n+1)}(S^\lambda) = S^\lambda \times S^{(1)}.$$

The first formula is then the Pieri rule translated by the Frobenius–Schur isomorphism. For the second formula, we shall prove a general fact in representation theory: given two finite groups $H \subset G$ and two irreducible representations V of H and W of G, we have:

$$\text{multiplicity of } W \text{ in } \mathrm{Ind}_H^G(V) = \text{multiplicity of } V \text{ in } \mathrm{Res}_H^G(W).$$

This is the *Frobenius reciprocity formula*. These multiplicities can be computed by taking scalar products of characters; let us start from the left-hand side. We have

$$\begin{aligned}
\left\langle \mathrm{ch}^W \,\middle|\, \mathrm{ch}^{\mathrm{Ind}_H^G(V)} \right\rangle_{\mathbb{C}G} &= \frac{1}{|G|} \sum_{g\in G} \overline{\mathrm{ch}^W(g)}\, \mathrm{ch}^{\mathrm{Ind}_H^G(V)} \\
&= \frac{1}{|G|} \sum_{\substack{g\in G\\ k\in G/H}} 1_{(k^{-1}gk\in H)}\, \overline{\mathrm{ch}^W(g)}\, \mathrm{ch}^V(k^{-1}gk) \\
&= \frac{1}{|G|} \sum_{\substack{h\in H\\ k\in G/H}} \overline{\mathrm{ch}^W(h)}\, \mathrm{ch}^V(h) \\
&= \frac{1}{|H|} \sum_{h\in H} \overline{\mathrm{ch}^W(h)}\, \mathrm{ch}^V(h) = \left\langle \mathrm{ch}^{\mathrm{Res}_H^G(W)} \,\middle|\, \mathrm{ch}^V \right\rangle_{\mathbb{C}H}.
\end{aligned}$$

Therefore, S^λ occurs in $\mathrm{Res}_{\mathfrak{S}(n)}^{\mathfrak{S}(n+1)}(S^\Lambda)$ if and only if S^Λ occurs in $\mathrm{Ind}_{\mathfrak{S}(n)}^{\mathfrak{S}(n+1)}(S^\lambda)$; this is the case if and only if $\lambda \nearrow \Lambda$, and the multiplicity is then equal to one. $\square$

4. The Gelfand–Tsetlin algebras

We should remark that we have been able to prove numerous results on the irreducible representations of the symmetric groups $\mathfrak{S}(n)$ *without describing explicitly* these representations and their matrices. For instance, the dimension formula 2.28 shows that there exists for each irreducible representation S^λ a linear basis $(e_T)_T$ of the representation space labelled by the standard tableaux $T \in \mathrm{ST}(\lambda)$, but so far we have no idea of the value of $\sigma \cdot e_T$ for $\sigma \in \mathfrak{S}(n)$ (of course, the result depends on the choice of the basis). This is actually a common phenomenon in representation theory; the same thing happens for the representations of the classical groups of matrices and their Lie algebras, where the theory of weights enables similarly the computation of the dimensions and of the characters, but does not enable the computation of the matrices of representations. To close this chapter, we shall state without proof several results about *nice* linear bases of the irreducible representations S^λ. A part of these results could actually be used to recover most of the representation theory of the symmetric groups, see the references at the end of the chapter.

We start by providing an explicit but mostly useless (at least for our objectives) description of the representations S^λ. Given a standard tableau T with shape $\lambda \in \mathfrak{Y}(n)$, we associate to it the polynomial

$$\Delta_T(x_1, \ldots, x_n) = \prod_{\text{columns } c \text{ of the diagram}} \left(\prod_{i<j \text{ appearing in } c} (x_i - x_j) \right).$$

For instance, if

$$T = \begin{array}{|c|c|c|c|} \hline 6 & \multicolumn{3}{c}{} \\ \cline{1-3} 3 & 5 & 8 & \multicolumn{1}{c}{} \\ \hline 1 & 2 & 4 & 7 \\ \hline \end{array}\;,$$

then $\Delta_T(x_1, \ldots, x_8) = (x_1 - x_3)(x_1 - x_6)(x_3 - x_6)(x_2 - x_5)(x_4 - x_8)$. Denote V^λ the subspace of $\mathbb{C}[x_1, \ldots, x_n]$ spanned linearly by all the polynomials $\Delta_T(x_1, \ldots, x_n)$ with $T \in \mathrm{ST}(\lambda)$. The symmetric group $\mathfrak{S}(n)$ acts on the space of polynomials $\mathbb{C}[x_1, \ldots, x_n]$ by permutation of the variables.

Theorem 2.32 (Frobenius, Specht, Young). *The space V^λ is stable by the action of the symmetric group $\mathfrak{S}(n)$, and it is an irreducible representation of $\mathfrak{S}(n)$ isomorphic to S^λ (the image of the Schur function $s_\lambda(X)$ by the Frobenius–Schur isomorphism). The family of polynomials $(\Delta_T)_{T \in \mathrm{ST}(\lambda)}$ is a linear basis of V^λ.*

This theorem provides an easy answer to the question: can one construct explicitly the irreducible representations of $\mathfrak{S}(n)$? Unfortunately, its proof has essentially nothing in common with the previous discussion, and it is also quite difficult. In the sequel, we construct another linear basis of the representation S^λ, which is a bit more connected to the previous discussion, and in particular to the branching rules: the so-called *Gelfand–Tsetlin basis*.

Fix an integer n and an integer partition $\lambda \in \mathfrak{Y}(n)$. The branching rule ensures that, if we restrict the action to $\mathfrak{S}(n-1)$, then we have an expansion in direct sum:

$$S^\lambda = \bigoplus_{\lambda^{(n-1)} \,|\, \lambda^{(n-1)} \nearrow \lambda} S^{\lambda^{(n-1)}}.$$

We claim that this decomposition in direct sum is unique, in the sense that each linear projector $\pi_{\lambda^{(n-1)}} : S^\lambda \to S^{\lambda^{(n-1)}}$ is entirely determined by the representation S^λ of $\mathfrak{S}(n)$. Indeed, consider more generally a finite group G and a representation $V = \bigoplus_{\lambda \in \widehat{G}} m_\lambda V^\lambda$, possibly with multiplicities $m_\lambda \geq 1$ for the irreducibles. Through the non-commutative Fourier transform, the structure of $\mathbb{C}G$-module on V translates into a structure of $\mathbb{C}\widehat{G} = \bigoplus_{\lambda \in \widehat{G}} \mathrm{End}(V^\lambda)$-module, each matrix space $\mathrm{End}(V^\lambda)$ acting on a subrepresentation V^μ by 0 if $\mu \neq \lambda$, and by the usual matrix action if $\mu = \lambda$.

In particular, the projection on the *isotypic component* $m_\lambda V^\lambda$ of V is obtained by using the inverse Fourier transform of id_{V^λ}, which is

$$\pi_\lambda = \sum_{g\in G} \frac{(\dim\lambda)\,\mathrm{ch}^\lambda(g^{-1})}{|G|}\, g.$$

As we have a multiplicity-free decomposition of S^λ in $G = \mathfrak{S}(n-1)$-representations, we conclude that the space $S^{\lambda^{(n-1)}} = \pi_{\lambda^{(n-1)}}(S^\lambda)$ is well-defined in S^λ. Now, we can use this argument recursively to expand the representations $S^{\lambda^{(n-1)}}$ in $\mathfrak{S}(n-2)$-irreducible representations, and so on. By induction, we can therefore prove the following:

PROPOSITION 2.33 (Gelfand–Tsetlin decomposition). *Given an irreducible representation S^λ of $\mathfrak{S}(n)$, there exists a unique decomposition in direct sum $S^\lambda = \bigoplus_{T\in\mathrm{ST}(\lambda)} V_T$ with the following property: each space V_T has dimension* 1, *and for any standard tableau T with shape λ corresponding to an increasing sequence of partitions*

$$\emptyset = \lambda^{(0)} \nearrow \lambda^{(1)} \nearrow \cdots \nearrow \lambda^{(n)} = \lambda$$

and any $k \in [\![1,n]\!]$, the space $\mathbb{C}\mathfrak{S}(k)\cdot V_T$ spanned by the action of $\mathfrak{S}(k)$ on V_T is an irreducible representation of $\mathfrak{S}(k)$ isomorphic to $S^{\lambda^{(k)}}$.

Since each V_T has dimension 1, there is up to a scalar a unique vector e_T such that $V_T = \mathbb{C}e_T$; we call the family $(e_T)_{T\in\mathrm{ST}(\lambda)}$ a Gelfand–Tsetlin basis of S^λ, and it is unique up to action of a diagonal matrix. The interest of this basis is that one can compute on it the action of a large family of elements of the group algebra $\mathbb{C}\mathfrak{S}(n)$.

DEFINITION 2.34 (Jucys–Murphy element). *For $k \in [\![1,n]\!]$, the k-th Jucys–Murphy element of $\mathbb{C}\mathfrak{S}(n)$ is the group algebra element*

$$J_k = \sum_{j=1}^{k-1} (j,k),$$

with by convention $J_1 = 0$. It is easy to check that the Jucys–Murphy elements commute with one another; we call Gelfand–Tsetlin algebra $\mathrm{GZ}(n)$ of order n the commutative subalgebra of $\mathbb{C}\mathfrak{S}(n)$ spanned by $J_1, J_2, \ldots, J_n$.

The following theorem, which we state without proof, summarises the important properties of the Gelfand–Tsetlin algebra and of the Jucys–Murphy elements. We shall need the notion of *content* of a cell of a Young diagram. If $\square$ is a cell with abscissa i and ordinate j in a Young diagram, its content $c(\square)$ is the difference $i-j$. Then, given a standard tableau T of size n, we denote $c(k,T)$ the content of the cell numbered k in the tableau T. For instance, if T is the standard tableau drawn before the statement of Theorem 2.32, then $c(6,T) = 1-3 = -2$.

THEOREM 2.35 (Jucys, Murphy, Okounkov–Vershik).

(1) The Gelfand–Tsetlin algebra $\mathrm{GZ}(n)$ is a maximal commutative subalgebra of $\mathbb{C}\mathfrak{S}(n)$.

(2) For any irreducible representation S^λ of $\mathfrak{S}(n)$, the elements of $\mathrm{GZ}(n)$ act diagonally on the Gelfand–Tsetlin basis $(e_T)_{T\in\mathrm{ST}(\lambda)}$. More precisely,

$$J_k \cdot e_T = c(k,T)\, e_T$$

for any $k \in [\![1,n]\!]$ and any $T \in \mathrm{ST}(\lambda)$.

(3) The image of $\mathrm{GZ}(n)$ by the non-commutative Fourier transform is $\bigoplus_{\lambda\in\mathfrak{Y}(n)} \mathrm{Diag}(S^\lambda)$, where $\mathrm{Diag}(S^\lambda)$ is the subalgebra of $\mathrm{End}(S^\lambda)$ which consists in endomorphisms acting diagonally in

the Gelfand–Tsetlin basis of S^λ. In particular,

$$\dim \mathrm{GZ}(n) = \sum_{\lambda \in \mathfrak{Y}(n)} \mathrm{card}\,\mathrm{ST}(\lambda).$$

(4) *The center $\mathrm{Z}(n) = \mathrm{Z}(\mathbb{C}\mathfrak{S}(n))$ is spanned by the symmetric polynomials in the Jucys–Murphy elements:*

$$\mathrm{Z}(n) = \mathrm{Span}(\{f(J_1, J_2, \ldots, J_n),\ f \in \mathrm{Sym}\}).$$

The Gelfand–Tsetlin algebra $\mathrm{GZ}(n)$ plays for $\mathbb{C}\mathfrak{S}(n)$ a role analogous to the role of a Cartan algebra in a Lie algebra, and of a maximal torus in a compact Lie group. In particular, many properties of the irreducible representations of the symmetric groups can be recovered by proving the theorem above, independently from the theory of symmetric functions. The starting point of the theory is the multiplicity-free property of the restriction functors $\mathrm{Res}^{\mathfrak{S}(n)}_{\mathfrak{S}(n-1)}$, which can be established independently from the Pieri rule by using the theory of Gelfand pairs. In Chapter 3, we shall only use the fact that the Jucys–Murphy elements act on the Gelfand–Tsetlin bases through the contents of the cells of the tableaux.

References

The classical reference for the theory of symmetric functions is [Mac95]; it is really the bible on the subject, and it has been read and cited by almost every combinatorist. The isomorphism between $\mathrm{R}(\mathfrak{S})$ and Sym is also explained in [Zel81], with a generalisation of this approach to the case of the Grothendieck ring of representations $\mathrm{R}(\mathrm{GL}(\mathbb{F}_q))$ of the finite general linear groups. For the representation theory of the symmetric groups, a classical textbook is [Sag01]. Let us also mention the monography by the author of these notes [Mél17], which contains the missing proofs of certain statements in this chapter (in particular, all the results from the last section).

The main properties of the Jucys–Murphy elements have been established in [Juc74; Mur81]. In a famous paper (see [OV04]), Okounkov and Vershik used these elements in order to reprove most of the properties of the representations of the symmetric groups; more details on this approach are provided by [CST10]. The Jucys–Murphy elements have been used in [Bia98] in order to relate the representations of the symmetric groups to certain ideas from the theory of free probability; they play an important role in many recent developments regarding the combinatorial properties of the irreducible characters ch^λ.

Exercises

(1) **The third Cauchy identity.** Prove the third Cauchy identity:

$$\prod_{i,j} \frac{1}{1 - x_i y_j} = \sum_{\lambda \in \mathfrak{Y}} h_\lambda(X)\, m_\lambda(Y).$$

What is the value of a scalar product $\langle h_\lambda \mid m_\mu \rangle$ in the algebra of symmetric functions Sym?

(2) **The Newton formulæ.** Consider the other generating function $Q(X, z)$ of power sums defined by

$$Q(X, z) = \sum_{k=1}^{\infty} p_k(X)\, z^k.$$

Show that

$$Q(X, z) = z\, P'(X, z) = z\, \frac{H'(X, z)}{H(X, z)} = z\, \frac{E'(X, -z)}{E(X, -z)},$$

where the $'$ denotes the derivation with respect to the variable z. Deduce from this the *Newton identities*: for any $k \geq 1$,

$$k\,h_k(X) = \sum_{j=1}^{k} p_j(X)\,h_{k-j}(X);$$
$$k\,e_k(X) = \sum_{j=1}^{k} (-1)^{j-1}\,p_j(X)\,e_{k-j}(X).$$

(3) **The characters of the symmetric groups take integer values.** This exercise is devoted to the proof of Lemma 2.24.

(a) Fix a finite group G and an element $g \in G$ with order m. We denote $\xi = \mathrm{e}^{\frac{2i\pi}{m}}$. Show that for any representation (V, ρ) of G, $\mathrm{ch}^V(g) = \mathrm{tr}(\rho(g))$ belongs to the field of numbers $\mathbb{Q}[\xi]$, and that it is an algebraic integer (solution of a monic polynomial equation $X^r + c_{r-1}X^{r-1} + \cdots + c_0 = 0$ with coefficients c_i in $\mathbb{Z}$).

We recall that the cyclotomic extension $\mathbb{Q}[\xi]$ admits for Galois group

$$\mathrm{Gal}(\mathbb{Q}[\xi], \mathbb{Q}) = (\mathbb{Z}/m\mathbb{Z})^*,$$

where the group of invertible elements $(\mathbb{Z}/m\mathbb{Z})^*$ of the ring $\mathbb{Z}/m\mathbb{Z}$ acts on elements of $\mathbb{Q}[\xi]$ as follows: a class of integers $k \bmod m$ with $k \wedge m = 1$ acts by the unique automorphism of fields of $\mathbb{Q}[\xi]$ which fixes $\mathbb{Q}$ and which sends ξ to ξ^k.

(b) Consider an element $g \in G$ with order m and such that, for any $k \in (\mathbb{Z}/m\mathbb{Z})^*$, g and g^k are conjugated. Show then that $\mathrm{ch}^V(g)$ belongs to $\mathbb{Q}$, and even to $\mathbb{Z}$.

(c) Use the criterion of the previous question in order to prove that the characters of the symmetric groups take their values in $\mathbb{Z}$.

(4) **The Grothendieck group of representations of a product of group.** Let G and H be two finite groups. Use the bijection $\widehat{G} \times \widehat{H} \to \widehat{G \times H}$ in order to construct a canonical isomorphism of vector spaces between $\mathrm{R}(G \times H)$ and the tensor product $\mathrm{R}(G) \otimes_{\mathbb{R}} \mathrm{R}(H)$.

(5) **The coproduct of symmetric functions and of representations of symmetric groups.** This exercise introduces two isomorphic structures of self-adjoint graded bialgebra on Sym and on $\mathrm{R}(\mathfrak{S})$.

(a) Let $X = \{x_1, \ldots, x_n, \ldots\}$ and $Y = \{y_1, \ldots, y_n, \ldots\}$ be two independent alphabets of variables. Given a symmetric function $f \in \mathrm{Sym}$, we set $\Delta(f) = f(X+Y)$, where $X + Y$ denotes the disjoint union of these sets of variables. For instance,

$$\Delta(p_k) = \sum_{i\geq 1} (x_i)^k + \sum_{i \geq 1} (y_i)^k.$$

Show that Δ yields a morphism of graded algebras from Sym to $\mathrm{Sym} \otimes_{\mathbb{R}} \mathrm{Sym}$. With this viewpoint, prove the following identities:

$$\Delta(p_k) = p_k \otimes 1 + 1 \otimes p_k;$$
$$\Delta(h_k) = h_k \otimes 1 + h_{k-1} \otimes h_1 + h_{k-2} \otimes h_2 + \cdots + 1 \otimes h_k;$$
$$\Delta(e_k) = e_k \otimes 1 + e_{k-1} \otimes e_1 + e_{k-2} \otimes e_2 + \cdots + 1 \otimes e_k.$$

We say that Δ is a *coproduct* on Sym.

(b) We endow Sym with the Hall scalar product, and $\mathrm{Sym} \otimes_{\mathbb{R}} \mathrm{Sym}$ with the tensor product of two Hall scalar products:

$$\langle f_1 \otimes g_1 \mid f_2 \otimes g_2 \rangle_{\mathrm{Sym} \otimes_{\mathbb{R}} \mathrm{Sym}} = \langle f_1 \mid f_2 \rangle_{\mathrm{Sym}} \, \langle g_1 \mid g_2 \rangle_{\mathrm{Sym}}$$

for $f_1, f_2, g_1, g_2 \in \mathrm{Sym}$. Show that Δ is the dual map of the product of the algebra Sym: for any $f, g, h \in \mathrm{Sym}$,

$$\langle \Delta(f) \mid g \otimes h \rangle_{\mathrm{Sym} \otimes_{\mathbb{R}} \mathrm{Sym}} = \langle f \mid gh \rangle_{\mathrm{Sym}} \, .$$

(c) Given a representation V of $\mathfrak{S}(n)$, we denote $\Delta(V)$ the element of $\mathrm{R}(\mathfrak{S}) \otimes \mathrm{R}(\mathfrak{S})$ defined as:

$$\sum_{k=0}^{n} \mathrm{Res}^{\mathfrak{S}(n)}_{\mathfrak{S}(k) \times \mathfrak{S}(n-k)}(V),$$

where $\mathfrak{S}(k) \times \mathfrak{S}(n-k)$ is considered as the subgroup of $\mathfrak{S}(n)$ which consists in pairs of permutations (σ_1, σ_2) such that σ_1 acts on $[\![1, k]\!]$ and σ_2 acts on $[\![k+1, n]\!]$; and where a representation of this group can be seen as an element of $\mathrm{R}(\mathfrak{S}(k)) \otimes_{\mathbb{R}} \mathrm{R}(\mathfrak{S}(n-k))$. Use Frobenius reciprocity to prove that Δ is the dual map of the product of $\mathrm{R}(\mathfrak{S})$: for any representations V, W, Z of $\mathfrak{S}(m)$, $\mathfrak{S}(n)$ and $\mathfrak{S}(m+n)$,

$$\langle \Delta(V) \mid W \otimes Z \rangle_{\mathrm{R}(\mathfrak{S}) \otimes_{\mathbb{R}} \mathrm{R}(\mathfrak{S})} = \langle V \mid W \times Z \rangle_{\mathrm{R}(\mathfrak{S})} \, .$$

(d) Show that the Frobenius–Schur isomorphism $\Psi : \mathrm{R}(\mathfrak{S}) \to \mathrm{Sym}$ is compatible with the coproducts Δ defined above: for any representation V of a symmetric group $\mathfrak{S}(n)$,

$$\Delta(\Psi(V)) = (\Psi \otimes \Psi)(\Delta(V));$$

(hint: use the two duality properties).

(6) **The antipode of symmetric functions and of representations of symmetric groups.** This exercise introduces two involutions of Sym and $\mathrm{R}(\mathfrak{S})$ which correspond through the Frobenius–Schur isomorphism.

(a) Consider the morphism of algebras $\varepsilon : \mathrm{Sym} \to \mathrm{Sym}$ defined by $\varepsilon(p_k) = -p_k$ for any $k \geq 1$. Show that $\varepsilon(h_k) = (-1)^k \, e_k$ and $\varepsilon(e_k) = (-1)^k \, h_k$. Show also that ε satisfies the Hopf identity: if Δ is the coproduct $\mathrm{Sym} \to \mathrm{Sym} \otimes_{\mathbb{R}} \mathrm{Sym}$ introduced in the previous exercise and $\nabla : \mathrm{Sym} \otimes_{\mathbb{R}} \mathrm{Sym} \to \mathrm{Sym}$ is the product of the algebra Sym, then

$$\nabla \circ (\mathrm{id} \otimes \varepsilon) \circ \Delta = \nabla \circ (\varepsilon \otimes \mathrm{id}) \circ \Delta = 0.$$

(b) For each $n \in \mathbb{N}$, denote ε_n the signature representation of $\mathfrak{S}(n)$, which has dimension 1. The *internal tensor product* of two representations (V, ρ^V) and (W, ρ^W) of the same finite group G is the representation with space $V \otimes W$ and with defining morphism

$$\begin{aligned} \rho^V \otimes \rho^W : G &\to \mathrm{GL}(V \otimes W) \\ g &\mapsto \rho^V(g) \otimes \rho^W(g). \end{aligned}$$

We denote $V \boxtimes W$ this representation. Show that the linear map $\varepsilon : \mathrm{R}(\mathfrak{S}) \to \mathrm{R}(\mathfrak{S})$ defined on a representation V of $\mathfrak{S}(n)$ by $\varepsilon(V) = (-1)^n \, \varepsilon_n \boxtimes V$ is an involutive isomorphism of the algebra $\mathrm{R}(\mathfrak{S})$.

(c) We identify $\mathrm{R}(\mathfrak{S}(n))$ and $\mathrm{Z}(\mathbb{C}\mathfrak{S}(n))$ by using the character map ch_n, and we continue to denote ε the involution of $\mathrm{Z}(\mathbb{C}\mathfrak{S}(n))$ obtained by transport by the map ch_n. Compute $\varepsilon(C_\mu)$ for $\mu \in \mathfrak{Y}(n)$.

(d) Show that the Frobenius–Schur isomorphism $\Psi : \mathrm{R}(\mathfrak{S}) \to \mathrm{Sym}$ is compatible with the involutions ε defined above: for any representation V of a symmetric group $\mathfrak{S}(n)$,
$$\varepsilon(\Psi(V)) = \Psi(\varepsilon(V)).$$

(e) We admit the following dual Jacobi–Trudy formula: for any integer partition λ with $\lambda_1 \leq m$,
$$s_\lambda = \det((e_{\lambda'_i - i + j})_{1 \leq i,j \leq m}),$$
where λ' denotes the integer partition with the same size as λ, and with a Young diagram obtained from the one of λ by symmetrising with respect to the diagonal axis (for instance, if $\lambda = (6,4,4,2,1)$, then $\lambda' = (5,4,3,3,1,1)$). Show that $\varepsilon(s_\lambda) = (-1)^{|\lambda|}\, s_{\lambda'}$ in the algebra of symmetric functions.

(f) Let λ be an integer partition with size n. Show that the internal tensor product of representations $\varepsilon_n \boxtimes S^\lambda$ is again an irreducible representation of $\mathfrak{S}(n)$, and identify the corresponding integer partition.

(7) **Product of Jucys–Murphy elements**. Check that the Jucys–Murphy elements commute with one another, by giving an explicit formula for the product $J_i J_j = J_j J_i$.

(8) **Symmetric functions of Jucys–Murphy elements**. We place ourselves in the algebra $\mathbb{C}[z](\mathfrak{S}(n))$ of formal linear combinations of permutations of size n, with coefficients in the algebra of polynomials $\mathbb{C}[z]$. Prove the following identity:
$$\prod_{i=1}^{n}(z + J_i) = \sum_{\sigma \in \mathfrak{S}(n)} z^{\text{number of cycles of } \sigma}\, \sigma.$$
Deduce from this a formula for the k-th elementary symmetric function of the Jucys–Murphy elements:
$$e_k(J_1, \ldots, J_n) = \sum_{\substack{\mu \in \mathfrak{Y}(n) \\ \ell(\mu) = n-k}} C_\mu.$$

CHAPTER 3

Computation of the mixing times

We can now combine the results from the two first chapters in order to evaluate the mixing times of two random walks on the symmetric group $\mathfrak{S}(n)$: the random walk of random transpositions with generator

$$\mu_{\mathrm{RT}} = \frac{1}{N}\,\mathrm{id} + \frac{2}{N^2} \sum_{1 \leq i < j \leq N} (i,j),$$

and the random walk of top-with-random transpositions with generator

$$\mu_{\mathrm{TWRT}} = \frac{1}{N}\,\mathrm{id} + \frac{1}{N} \sum_{i=2}^{N} (1,i).$$

The other kinds of random walks introduced in Chapter 1 are examined in the exercises at the end of this chapter.

1. The Diaconis upper-bound lemma

Let μ be a probability measure on $\mathfrak{S}(N)$ which satisfies the hypotheses of Proposition 1.8; we know that the law at time n of the random walk with generator μ is $\mu_n = \mu^n$, and that $\mu_n(\sigma) \to \frac{1}{N!}$ for any permutation $\sigma \in \mathfrak{S}(N)$. We want to evaluate the total variation distance

$$d_{\mathrm{TV}}(\mu_n, \mathrm{Haar}) = \sup_{A \subset \mathfrak{S}(N)} |\mu_n(A) - \mathrm{Haar}(A)| = \frac{1}{2} \sum_{\sigma \in \mathfrak{S}(N)} \left| \mu_n(\sigma) - \frac{1}{N!} \right|.$$

The quantity above can be rewritten as a $\mathscr{L}^1$-norm with respect to the Haar measure, which we denote μ_∞ in the sequel:

$$d_{\mathrm{TV}}(\mu_n, \mu_\infty) = \frac{1}{2}\, \|f_n - 1\|_{\mathscr{L}^1(\mathfrak{S}(N),\,\mu_\infty)}, \quad \text{where } f_n(\sigma) = \frac{d\mu_n}{d\mu_\infty}(\sigma) = N!\,\mu_n(\sigma).$$

By the Cauchy–Schwarz inequality, we therefore have:

$$4\,(d_{\mathrm{TV}}(\mu_n,\mu_\infty))^2 = \left(\|f_n - 1\|_{\mathscr{L}^1(\mathfrak{S}(N),\,\mu_\infty)}\right)^2 \leq \left(\|f_n - 1\|_{\mathscr{L}^2(\mathfrak{S}(N),\,\mu_\infty)}\right)^2.$$

The $\mathscr{L}^2$-norm on the right-hand side is the norm of the group algebra $\mathbb{C}\mathfrak{S}(N)$, so by Theorem 1.21,

$$\begin{aligned}(\|f_n - 1\|_{\mathbb{C}\mathfrak{S}(N)})^2 &= \sum_{\lambda \in \mathfrak{Y}(N)} \frac{\dim \lambda}{(N!)^2}\, \mathrm{tr}\left(((\widehat{f_n - 1})(\lambda))^* \, (\widehat{f_n - 1})(\lambda) \right) \\ &= \sum_{\lambda \in \mathfrak{Y}(N)} (\dim \lambda)\, \mathrm{tr}\left(((\widehat{\mu_n - \mu_\infty})(\lambda))^* \, (\widehat{\mu_n - \mu_\infty})(\lambda) \right).\end{aligned}$$

Here we used the fact that the dual of the finite group $\mathfrak{S}(N)$ is the set of integer partitions $\mathfrak{Y}(N)$. Now, the expression above can be simplified a bit. First, note that the trivial representation 1_N of $\mathfrak{S}(N)$ on $\mathbb{C}$ corresponds to the integer partition (N) with one part; indeed, we have seen during the proof of Theorem 2.23 that $\Psi(1_N) = h_N$, and the homogeneous symmetric function h_N is a particular case of Schur functions: $h_N = s_{(N)}$ by the Jacobi–Trudy formula for a Schur function of an integer partition with one part. Therefore, $1_N = S^{(N)}$. Now, we claim that for any other integer partition $\lambda \in \mathfrak{Y}(N)$,

$$\lambda \neq (N) \quad \Rightarrow \quad \widehat{\mu_\infty}(\lambda) = 0.$$

Indeed, this is a particular case of the orthogonality of the matrix coefficients stated in Lemma 1.22: if ρ^λ_{ij} denotes a matrix coefficient of the irreducible representation $(S^\lambda, \rho^\lambda)$ with respect to an arbitrary basis of S^λ, then

$$(\widehat{\mu_\infty}(\lambda))_{ij} = \frac{1}{N!} \sum_{\sigma \in \mathfrak{S}(N)} \rho^\lambda_{ij}(\sigma) = \left\langle \rho^{(N)}_{11} \,\middle|\, \rho^\lambda_{ij} \right\rangle_{\mathbb{C}\mathfrak{S}(N)} = 0.$$

On the other hand, for any probability measure ν on $\mathfrak{S}(N)$, since the image of any permutation σ by the trivial representation is the identity of $\mathbb{C}$,

$$\widehat{\nu}((N)) = \sum_{\sigma \in \mathfrak{S}(N)} \nu(\sigma)\, \mathrm{id}_C = \mathrm{id}_{\mathbb{C}}.$$

Consequently, $(\widehat{\mu_n - \mu_\infty})((N)) = \mathrm{id}_{\mathbb{C}} - \mathrm{id}_{\mathbb{C}} = 0$. So, in the sum over partitions corresponding to the $\mathscr{L}^2$-norm $(\|f_n - 1\|_{\mathbb{C}\mathfrak{S}(N)})^2$, one can remove the term corresponding to the integer partition (N), and in the other terms, one can remove the quantities $\widehat{\mu_\infty}(\lambda)$. Thus,

$$(\|f_n - 1\|_{\mathbb{C}\mathfrak{S}(N)})^2 = \sum_{\substack{\lambda \in \mathfrak{Y}(N) \\ \lambda \neq (N)}} (\dim \lambda)\, \mathrm{tr}\big((\widehat{\mu_n}(\lambda))^* \,\widehat{\mu_n}(\lambda)\big).$$

Finally, since the non-commutative Fourier transform is compatible with the convolution products, $\widehat{\mu_n}(\lambda) = (\widehat{\mu}(\lambda))^n$ for any $\lambda \in \mathfrak{Y}(N)$. We have therefore shown:

LEMMA 3.1 (Diaconis). *Consider a random walk on $\mathfrak{S}(N)$ with generator μ. For any $n \in \mathbb{N}$,*

$$4\,(d_{\mathrm{TV}}(\mu_n, \mu_\infty))^2 \leq \sum_{\substack{\lambda \in \mathfrak{Y}(N) \\ \lambda \neq (N)}} (\dim \lambda)\, \mathrm{tr}\big(((\widehat{\mu}(\lambda))^n)^* \,(\widehat{\mu}(\lambda))^n\big).$$

REMARK 3.2. In the computations above, the fact that the underlying group was $G = \mathfrak{S}(N)$ was only used in order to label the irreducible representations, and to identify the integer partition which corresponds to the trivial representation. Therefore, Lemma 3.1 generalises readily to any random walk on a finite group G, with a sum running over the set

$$\widehat{G}^* = \{\text{(isomorphism classes of) non-trivial irreducible representations of G}\}.$$

In the sequel we shall compute explicitly the terms of the sum of Lemma 3.1 when μ is one of the two generators described at the beginning of this chapter. The upper-bound lemma of Diaconis takes a simpler form when the generator μ is a probability measure constant on each conjugacy class, or equivalently when $\mu \in \mathrm{Z}(\mathbb{C}\mathfrak{S}(N))$. Indeed, as explained during the proof of Proposition 1.24, the image of the center of the group algebra by the non-commutative Fourier transform is the direct sum of the one-dimensional algebras $\mathbb{C}\mathrm{id}_{S^\lambda}$, $\lambda \in \mathfrak{Y}(N)$. Therefore, if $\mu \in \mathrm{Z}(\mathbb{C}\mathfrak{S}(N))$, then for any $\lambda \in \mathfrak{Y}(N)$,

$$\widehat{\mu}(\lambda) = \frac{\mathrm{tr}(\widehat{\mu}(\lambda))}{\dim \lambda}\, \mathrm{id}_{S^\lambda}.$$

The trace of the Fourier transform can be rewritten in terms of characters:

$$\mathrm{tr}(\widehat{\mu}(\lambda)) = \sum_{\sigma \in \mathfrak{S}(N)} \mu(\sigma)\, \mathrm{tr}(\rho^\lambda(\sigma)) = \mathrm{tr}(\widehat{\mu}(\lambda)) = \sum_{\sigma \in \mathfrak{S}(N)} \mu(\sigma)\, \mathrm{ch}^\lambda(\sigma).$$

We shall denote this quantity $\widetilde{\mu}(\lambda)$; this version of the non-commutative Fourier transform for elements of the center of the group algebra yields a collection of complex numbers instead of a collection of complex matrices. If we rewrite the upper bound in terms of these quantities, we get:

COROLLARY 3.3 (Upper-bound for central generators). *Consider a random walk on $\mathfrak{S}(N)$ with generator $\mu \in \mathrm{Z}(\mathbb{C}\mathfrak{S}(N))$. For any $n \geq 1$,*

$$4\,(d_{\mathrm{TV}}(\mu_n, \mu_\infty))^2 \leq \sum_{\substack{\lambda \in \mathfrak{Y}(N) \\ \lambda \neq (N)}} \frac{|\widetilde{\mu}(\lambda)|^{2n}}{(\dim \lambda)^{2n-2}}.$$

This version of the upper-bound lemma will be used when μ is the generator of the products of random transpositions, whereas the general version will be used when μ is the generator of the top-with-random transposition random walk.

2. The hook-length formula

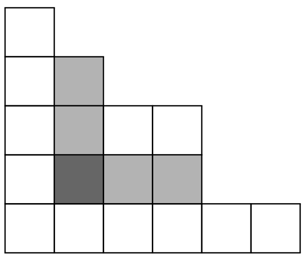
In order to use Lemma 3.1 and Corollary 3.3 with the two generators μ_{RT} and μ_{TWRT}, we need to make explicit the following quantities:

$$\dim \lambda,\ \ \widetilde{\mu_{\mathrm{RT}}}(\lambda),\ \ \widehat{\mu_{\mathrm{TWRT}}}(\lambda).$$

The dimension of the irreducible representation S^λ of $\mathfrak{S}(N)$ with label λ is given by the celebrated *hook-length formula*. Given an integer partition λ and a cell $\square$ of its Young diagram, we call *leg-length* (respectively, *arm-length*) of $\square$ in λ the number $l(\square, \lambda)$ of cells on top of $\square$ in the Young diagram (respectively, the number $a(\square, \lambda)$ of cells to the right of $\square$). The hook-length of $\square$ is given by the formula:

$$h(\square, \lambda) = 1 + l(\square, \lambda) + a(\square, \lambda).$$

It is the size of the largest hook which one can draw by starting from $\square$ and by trying to reach the border of the Young diagram. For instance, in $\lambda = (6, 4, 4, 2, 1)$, the cell with coordinates $(2, 2)$ has for hook-length $1 + 2 + 2 = 5$.



THEOREM 3.4 (Frame–Robinson–Thrall). *For any integer partition $\lambda \in \mathfrak{Y}(N)$, the dimension of the corresponding irreducible representation S^λ of $\mathfrak{S}(N)$ is*

$$\dim \lambda = \frac{N!}{\prod_{\square \in \lambda} h(\square, \lambda)}.$$

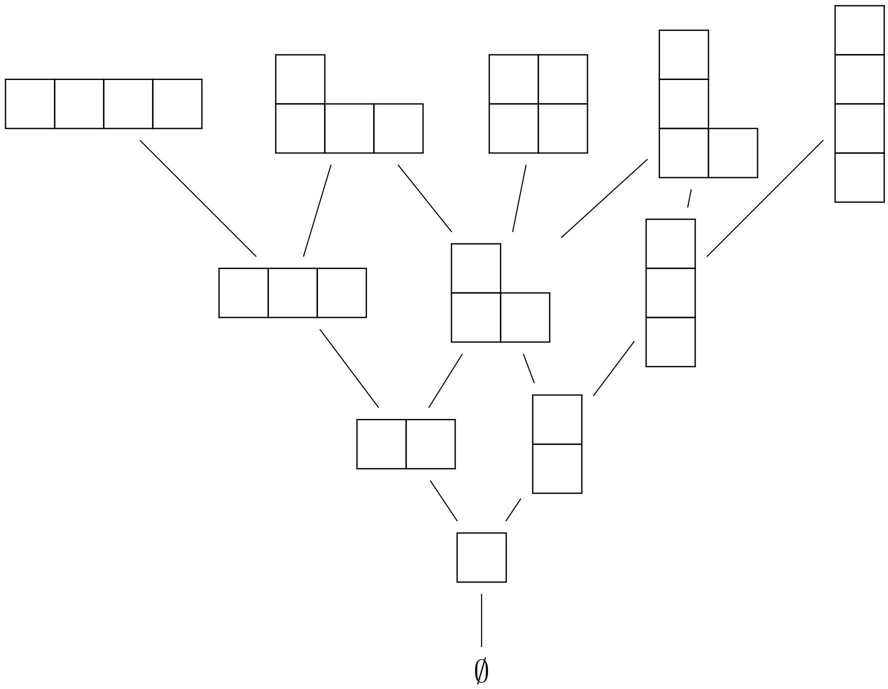
One of the easiest proof of this result is of probabilistic nature, as it involves a random walk on the so-called *Young graph*. The Young graph is the infinite graph whose vertices are all the Young diagrams of the integer partitions $\lambda \in \mathfrak{Y}$, and with an edge from λ to Λ if $|\Lambda| = |\lambda| + 1$ and $\lambda \nearrow \Lambda$. The first levels of this graph are drawn hereafter.

Consider an integer partition Λ with size $N + 1$. We choose a random integer partition λ of size N among those such that $\lambda \nearrow \Lambda$, according to the following procedure:

- We start by taking at random a cell $\square_1$ uniformly in the Young diagram of Λ; each cell has probability $\frac{1}{N+1}$.
- If $\square_1$ is a corner in the upper-right border of Λ, we can remove it in order to obtain an integer partition λ with $\lambda \nearrow \Lambda$.

- Otherwise, the hook-length of $\square_1$ is greater than 2, and we can take a new random cell $\square_2$ uniformly in the hook of $\square_1$, $\square_1$ being excluded. Each cell of the hook of $\square_1$ has uniform probability $\frac{1}{h(\square_1,\lambda)-1}$.
- If $\square_2$ is a corner, we remove it and get $\lambda \in \mathfrak{Y}(N)$ such that $\lambda \nearrow \Lambda$. Otherwise, we continue the procedure by choosing a uniform random cell $\square_3$ in the hook of $\square_2$, then a uniform random cell $\square_4$ in the hook of $\square_3$, *etc.* until we get a removable corner.

At each step, the hook-length of the new cell $\square_i$ is strictly smaller than the hook-length of $\square_{i-1}$, so the procedure terminates almost surely.



Lemma 3.5. *Denote $p(\Lambda \to \lambda)$ the probability to obtain λ by removing a random corner from Λ according to the procedure described above; and $f(\lambda)$ the right-hand side of the formula of Theorem 3.4 (the ratio with hook-lengths). We have*

$$p(\Lambda \to \lambda) = \frac{f(\lambda)}{f(\Lambda)}.$$

Proof. Denote (x, y) the coordinates of the cell of the Young diagram of Λ which is removed in order to get λ. We have

$$\frac{f(\lambda)}{f(\Lambda)} = \frac{1}{N+1} \frac{\prod_{\square\in\Lambda} h(\square,\Lambda)}{\prod_{\square\in\lambda} h(\square,\lambda)},$$

and the only cells which have a different hook-length in λ and in Λ are those in the same row or the same column as (x, y). Thus,

$$\begin{aligned}
\frac{f(\lambda)}{f(\Lambda)} &= \frac{1}{N+1} \prod_{i=1}^{x-1} \frac{h(i,y)}{h(i,y)-1} \prod_{j=1}^{y-1} \frac{h(x,j)}{h(x,j)-1} \\
&= \frac{1}{N+1} \left(\prod_{i=1}^{x-1} 1 + \frac{1}{h(i,y)-1} \right) \left(\prod_{j=1}^{y-1} 1 + \frac{1}{h(x,j)-1} \right) \\
&= \frac{1}{N+1} \sum_{\substack{I \subset [\![1,x-1]\!] \\ J \subset [\![1,y-1]\!]}} \frac{1}{\prod_{i\in I}(h(i,y)-1) \prod_{j\in J}(h(x,j)-1)},
\end{aligned}$$

where $h(a,b) = h((a,b),\Lambda)$.

Let us now prove that one obtains the same formula when following the random procedure of deletion of a corner. Let

$$(\square_1, \square_2, \ldots, \square_r = (x,y)) = ((x_1,y_1),(x_2,y_2),\ldots,(x_r,y_r) = (x,y))$$

be a sequence of cells in the Young diagram of Λ, such that each $\square_i$ is in the hook of $\square_{i-1}$, and such that $\square_r = (x,y)$ is a corner in the top-right border of the Young diagram. Thus, we are looking at a path of cells which can be followed during the procedure of random deletion of a corner in Λ. Let $I = \{x_1, x_2, \ldots, x_r\} \setminus \{x\}$ and $J = \{y_1, y_2, \ldots, y_r\} \setminus \{y\}$. Note that repetitions are possible in the sequences $(x_1, \ldots, x_r)$ and $(y_1, \ldots, y_r)$; therefore, the sets I and J can have different cardinalities, and usually they have cardinality smaller than $r-1$. Let us denote $p(\Lambda \to \lambda, I, J)$ the probability to obtain $\lambda = \Lambda \setminus (x,y)$ and the two traversal sets I and J. We claim that

$$p(\Lambda \to \lambda, I, J) = \frac{1}{N+1} \frac{1}{\prod_{i\in I}(h(i,y)-1) \prod_{j \in J}(h(x,j)-1)}.$$

This will end the proof of the lemma, by summing over all the pairs (I,J) of possible traversal sets. We now reason by induction on $s = |I| + |J|$. If $s = 0$, then the sequence of cells only contains $\square_1 = (x,y)$, and the probability of choosing right away this cell is indeed $\frac{1}{N+1}$, so the result is true in this case. Suppose the result true up to order $s-1 \geq 0$. Note that the first cell $\square_1 = (x_1,y_1)$ of a path counted by the probability $p(\Lambda \to \lambda, I, J)$ is determined by the two sets I and J:

$$x_1 = \min(I \sqcup \{x\}) \quad ; \quad y_1 = \min(J \sqcup \{y\}).$$

Then, there are two possibilities for $\square_2$: either $\square_2 = (x_2, y_1)$ with $x_2 = \min(I \sqcup \{x\} \setminus \{x_1\})$, or $\square_2 = (x_1, y_2)$ with $y_2 = \min(J \sqcup \{y\} \setminus \{y_1\})$. Consequently,

$$p(\Lambda \to \lambda, I, J) = \frac{1}{h(x_1,y_1)-1} \big(p(\Lambda \to \lambda, I \setminus \{x_1\}, J) + p(\Lambda \to \lambda, I, J \setminus \{y_1\})\big).$$

By the induction hypothesis,

$$\begin{aligned} p(\Lambda \to \lambda, I \setminus \{x_1\}, J) &= \frac{1}{N+1} \frac{1}{\prod_{i\in I\setminus\{x_1\}}(h(i,y)-1) \prod_{j \in J}(h(x,j)-1)} \\ &= \frac{h(x_1,y)-1}{N+1} \frac{1}{\prod_{i\in I}(h(i,y)-1) \prod_{j \in J}(h(x,j)-1)}, \end{aligned}$$

and similarly,

$$p(\Lambda \to \lambda, I, J \setminus \{y_1\}) = \frac{h(x,y_1)-1}{N+1} \frac{1}{\prod_{i\in I}(h(i,y)-1) \prod_{j \in J}(h(x,j)-1)}.$$

To get the result with $|I| + |J| = s$, we therefore have to prove that $h(x,y_1) + h(x_1,y) - 2 = h(x_1,y_1) - 1$. This is an easy computation if one introduces the leg- and arm-lengths:

$$\begin{aligned} h(x,y_1) + h(x_1,y) - 2 &= l(x,y_1) + a(x,y_1) + l(x_1,y) + a(x_1,y) \\ &= (l(x,y_1) + l(x_1,y)) + (a(x,y_1) + a(x_1,y)) \\ &= l(x_1,y_1) + a(x_1,y_1) = h(x_1,y_1) - 1. \end{aligned}$$ ☐

PROOF OF THEOREM 3.4. Since $p(\Lambda \to \cdot)$ is a probability measure on the set of integer partitions λ such that $\lambda \nearrow \Lambda$, we have

$$1 = \sum_{\lambda \,|\, \lambda \nearrow \Lambda} \frac{f(\lambda)}{f(\Lambda)}.$$

In other words, $f(\cdot)$ satisfies the recurrence relation: $f(\Lambda) = \sum_{\lambda \,|\, \lambda \nearrow \Lambda} f(\lambda)$. However, the same recurrence relation is satisfied by $\dim \lambda = \mathrm{card}(\mathrm{ST}(\lambda))$; therefore, $\dim \lambda = f(\lambda)$. ☐

EXAMPLE 3.6. The integer partition $\lambda = (6,4,4,2,1)$ admits the following hook-lengths:

1					
3	1				
6	4	2	1		
7	5	3	2		
10	8	6	5	2	1

Dividing 17! by the product of these numbers, one obtains:

$$\dim(6,4,4,2,1) = 2450448.$$

The two other unknown quantities appearing in the Diaconis upper-bound lemma are much easier to compute. We start with the computation of the coefficients $\widetilde{\mu_{\mathrm{RT}}}(\lambda)$:

PROPOSITION 3.7 (Fourier transform for the random transposition shuffle). *For any integer partition $\lambda \in \mathfrak{Y}(N)$,*

$$\widetilde{\mu_{\mathrm{RT}}}(\lambda) = \frac{\dim\lambda}{N^2} \sum_{i=1}^{\ell(\lambda)} \lambda_i(\lambda_i - 2i + 2).$$

PROOF. Notice that $x_N = \sum_{1\leq i<j\leq N}(i,j) = \sum_{i=1}^{N} J_i$ is the sum of all the Jucys–Murphy elements. Therefore, the action of x_N on any element e_T of the Gelfand–Tsetlin basis of the irreducible representation S^λ is by multiplication by the sum of all the contents of the cells of the Young diagram of λ; this does not depend on the standard tableau T. The trace of the action of x_N on S^λ is therefore:

$$\sum_{T\in\mathrm{ST}(\lambda)} \left(\sum_{\square\in\lambda} c(\square,\lambda)\right) = (\dim\lambda)\left(\sum_{\square\in\lambda} c(\square,\lambda)\right).$$

This is also $\mathrm{tr}(\rho^\lambda(x_N)) = \sum_{1\leq i<j\leq N} \mathrm{ch}^\lambda(i,j)$. Consequently,

$$\widetilde{\mu_{\mathrm{RT}}}(\lambda) = \frac{1}{N}\,\mathrm{ch}^\lambda(\mathrm{id}) + \frac{2}{N^2}\sum_{1\leq i<j\leq N} \mathrm{ch}^\lambda(i,j) = (\dim\lambda)\left(\frac{1}{N} + \frac{2}{N^2}\sum_{\square\in\lambda} c(\square,\lambda)\right).$$

To compute the sum of the contents, note that on the i-th row of the Young diagram of λ, one sums

$$(\lambda_i - i) + (\lambda_i - i - 1) + \cdots + (1-i) = \frac{\lambda_i(\lambda_i+1)}{2} - i\,\lambda_i = \frac{\lambda_i(\lambda_i - 2i + 1)}{2}.$$

So,

$$\sum_{\square\in\lambda} c(\square,\lambda) = \frac{1}{2}\sum_{i=1}^{\ell(\lambda)} \lambda_i(\lambda_i - 2i + 1).$$

We conclude by using the fact that $N = \sum_{i=1}^{\ell(\lambda)} \lambda_i$. ☐

For the non-commutative Fourier transform of μ_{TWRT}, notice that the element $y_N = \sum_{i=2}^{N}(1,i)$ is conjugated by $\mathfrak{S}(N)$ to the N-th Jucys–Murphy element $J_N = \sum_{i=1}^{N-1}(i,N)$:

$$y_N = (1,N)\,J_N\,(1,N).$$

Therefore, if $(e_T)_{T\in\mathrm{ST}(\lambda)}$ is the Gelfand–Tsetlin basis of S^λ and if $f_T = (\rho^\lambda(1,N))(e_T)$, then

$$y_N\,f_T = (1,N)\,J_N\,e_T = c(N,T)\,(1,N)\,e_T = c(N,T)\,f_T.$$

So, y_N acts diagonally on the modified Gelfand–Tsetlin basis $(f_T)_{T\in\mathrm{ST}(\lambda)}$, and its action is given by the family of contents $(c(N,T))_{T\in\mathrm{ST}(\lambda)}$. Note that a cell $\square$ of a Young diagram can be numbered by N in a standard tableau T only if $\square$ is a corner at the top-right edge of the diagram. Moreover, the number of standard tableaux with shape λ and with the corner cell $\square$ numbered by N is the number of standard tableaux with size $N-1$ and shape $\mu = \lambda \setminus \square$. If $\square$ is in the i-th row of λ, then $c(N,T) = \lambda_i - i$ in this case. So:

PROPOSITION 3.8 (Fourier transform for the top-with-random transposition shuffle). *For any integer partition $\lambda \in \mathfrak{Y}(N)$, the eigenvalues of the matrix $\widehat{\mu_{\mathrm{TWRT}}}(\lambda)$ are the numbers $\frac{\lambda_i - i + 1}{N}$, where i is the number of a row such that $\lambda_i > \lambda_{i+1}$. The multiplicity of $\lambda_i - i$ as an eigenvalue is $\dim(\lambda \setminus \square_i)$, where $\square_i$ is the last cell of the i-th row of the Young diagram of λ.*

We can now rewrite more explicitly the upper bound for the two random walks on $\mathfrak{S}(N)$ that we are considering in this chapter. For the random transposition shuffle with generator μ_{RT}, we have by Corollary 3.3 and Proposition 3.7:

$$4\,(d_{\mathrm{TV}}(\mu_n, \mu_\infty))^2 \leq \sum_{\substack{\lambda \in \mathfrak{Y}(N) \\ \lambda \neq (N)}} (\dim \lambda)^2 \left(\frac{1}{N^2} \sum_{i=1}^{\ell(\lambda)} \lambda_i(\lambda_i - 2i + 2) \right)^{2n}.$$

For the top-with-random transposition shuffle, we have by Lemma 3.1 and Proposition 3.8:

$$4\,(d_{\mathrm{TV}}(\mu_n, \mu_\infty))^2 \leq \sum_{\substack{\lambda \in \mathfrak{Y}(N),\, \lambda \neq (N) \\ i \,|\, \square_i = (\lambda_i, i) \text{ is a corner of } \lambda}} (\dim \lambda)(\dim(\lambda \setminus \square_i)) \left(\frac{\lambda_i - i + 1}{N} \right)^{2n}.$$

3. Analysis of the bounding series

We want to prove that for n large enough, the upper bounds computed above are small. The first question to address is: how to choose n so that the total variation distance $d_{\mathrm{TV}}(\mu_n, \mu_\infty)$ becomes small? The following rule gives an intuitive answer, which will be made rigorous by the arguments of representation theory developed so far.

INTUITION. *Consider an ergodic Markov chain $(X_n)_{n \in \mathbb{N}}$ on a finite set $\mathfrak{X}$ of « configurations » on a finite space A. We assume that the transitions of the Markov chain are local modifications of the configurations, which only concern a bounded number of elements $a \in A$. Then, in order to reach the stationary measure, one needs to have modified at least once each element of the space A. In some (but not all) cases, this intuitive rule yields the correct order of magnitude for the mixing time of the Markov chain.*

Let us apply this intuitive rule to the two random walks on $\mathfrak{S}(N)$. For the random transposition shuffle, the Markov chain considered lives in the set $\mathfrak{X} = \mathfrak{S}(N)$ of permutations of the finite space $A = [\![1, N]\!]$. Each transposition τ_n modifies the images of two randomly chosen elements i_n and j_n in A. Therefore, one expects that the mixing time of the random walk $\sigma_n = \tau_1\tau_2\cdots\tau_n$ is of the same order as the expectation of the following stopping time:

$$(\text{mixing time})_{\mathrm{RT}} \sim \mathbb{E}[\tau_N], \;\; \text{with } \tau_N = \inf\{n \in \mathbb{N},\; \{i_1, i_2, \ldots, i_n, j_1, j_2, \ldots, j_n\} = [\![1, N]\!]\}.$$

Similarly, for the top-with-random transposition shuffle, each transposition τ_n modifies 1 and a randomly chosen element $i_n \in [\![1, N]\!]$. Therefore, one expects:

$$(\text{mixing time})_{\mathrm{TWRT}} \sim \mathbb{E}[\tau'_N], \;\; \text{with } \tau'_N = \inf\{n \in \mathbb{N},\; \{i_1, i_2, \ldots, i_n\} = [\![1, N]\!]\}.$$

The expectations of the stopping times are easy to compute:

LEMMA 3.9 (Problem of the collector). *Consider a sequence of independent random variables $(U_n)_{n\in\mathbb{N}}$ uniformly distributed in a finite set $[\![1,N]\!]$. Let $\tau_N = \inf\{n \in \mathbb{N},\ \{U_1,\ldots,U_n\} = [\![1,N]\!]\}$. We have:*

$$\mathbb{E}[\tau_N] = N\sum_{i=1}^{N}\frac{1}{i} = N\log N + O(N).$$

PROOF. If $k \in [\![0, N-1]\!]$ elements have already been attained by $\{U_1,\ldots,U_n\}$, then in order to attain the $(k+1)$-th element, one needs to wait a random time following a geometric distribution with parameter $p = \frac{N-k}{N}$. Therefore,

$$\tau_N =_{(\mathrm{law})} \sum_{k=0}^{N-1}\mathcal{G}\left(\frac{N-k}{N}\right),$$

where the sum on the right-hand side is a sum of independent geometric random variables. Consequently, $\mathbb{E}[\tau_N] = \sum_{k=0}^{N-1}\frac{N}{N-k} = N\sum_{i=1}^{N}\frac{1}{i}$. □

As a consequence of Lemma 3.9, a reasonable conjecture for the mixing time of the top-with-random transposition shuffle is

$$(\text{mixing time})_{\mathrm{TWRT}} \sim N\log N.$$

For the random transposition shuffle, one chooses at each step two random elements i_n and j_n instead of one element i_n, so one can expect that the mixing time will be divided by two in comparison to the previous model:

$$(\text{mixing time})_{\mathrm{RT}} \sim \frac{N\log N}{2}.$$

We shall see in the sequel that these conjectures hold: the intuitive rule yields the correct estimate of the mixing time for the two random walks on $\mathfrak{S}(N)$ that we are considering. It should be noticed that for certain Markov chains, the intuitive rule fails: for instance, when looking at the Glauber dynamics on the Ising model, if the temperature parameter β^{-1} is too small, then reaching every site is not sufficient in order to mix correctly the configurations, as certain areas of positive or negative spins stay frozen with large probability. Therefore, in this case, the mixing time is much larger than the hitting time of the whole set of sites; see the references at the end of the chapter.

Let us now proceed with the analysis of the bounding series. For the random transposition shuffle, we set

$$n = \frac{N(\log N + c)}{2},$$

with $c \in \mathbb{R}_+$ (in the sequel, it is always assumed that c is chosen so that the quantity above is an integer). We denote $S_{N,n} = \sum_{\lambda\neq(N)}(\dim\lambda)^2\,(r(\lambda))^{2n}$, where $r(\lambda) = \frac{1}{N^2}\sum_{i=1}^{\ell(\lambda)}\lambda_i(\lambda_i - 2i + 2)$.

Step 1: getting rid of the signs of the eigenvalues. The main idea is that the behavior of $S_{N,n}$ is determined *after the mixing time* by the integer partitions λ such that $|r(\lambda)|$ is maximal. To get an idea of what happens, let us compute $N^2\,r(\lambda)$ for $\lambda \in \mathfrak{Y}(8)$:

(8)	(71)	(62)	(61^2)	(53)	(521)	(51^3)	(4^2)	(431)	(42^2)	(421^2)
64	48	36	32	28	22	16	24	16	12	8
(41^4)	(3^22)	(3^21^2)	(32^21)	(321^3)	(31^5)	(2^4)	(2^31^2)	(2^21^4)	(21^6)	(1^8)
0	8	4	0	-6	-16	-8	-12	-20	-32	-48

We observe that the largest coefficients $r(\lambda)$ are obtained when the first parts $\lambda_1, \lambda_2, \ldots$ of λ are large. On the contrary, if λ has many small parts, then $r(\lambda)$ is small, and even possibly negative. The negative coefficients $r(\lambda)$ tend to be smaller in absolute value than the positive ones; the following lemma makes this statement more precise.

LEMMA 3.10. *For any $N \geq 4$ and any $n \geq 1$, we have*

$$S_{N,n} \leq 2 \sum_{\substack{\lambda \in \mathfrak{Y}(N) \\ \lambda \neq (N),\, r(\lambda) \geq 0}} (\dim \lambda)^2 \, (r(\lambda))^{2n}.$$

PROOF. Let us first consider the integer partitions λ which are not in

$$E_N = \left\{(N),\, (N-1,1),\, (2,1^{N-2}), (1^N)\right\}.$$

We have $r(\lambda) = \frac{1}{N} + \frac{2}{N^2}\, c(\lambda)$, where $c(\lambda)$ is the sum of all the contents of the cells of the Young diagram of λ. Given an integer partition λ, we denote λ' the integer partition with the same size and whose Young diagram is obtained by symmetrising the Young diagram of λ with respect to the first diagonal. For instance, if $\lambda = (6,4,2,2,1)$, then $\lambda' = (5,4,3,3,1,1)$. This operation is involutive, and it leaves $\dim \lambda$ invariant; indeed, the symmetry with respect to the first diagonal also yields a bijection between the standard tableaux of shape λ and those of shape λ'. On the other hand, we have obviously $c(\lambda') = -c(\lambda)$. Suppose $r(\lambda) < 0$. This is only possible if $c(\lambda) < 0$, and then,

$$r(\lambda') = \frac{1}{N} + \frac{2}{N^2}\, c(\lambda') = \frac{1}{N} - \frac{2}{N^2}\, c(\lambda) > -r(\lambda).$$

In particular, $r(\lambda') > 0$. Therefore,

$$\sum_{\substack{\lambda \in \mathfrak{Y}(N) \\ \lambda \notin E_N,\, r(\lambda) < 0}} (\dim \lambda^2)\, (r(\lambda))^{2n} \leq \sum_{\substack{\lambda \in \mathfrak{Y}(N) \\ \lambda' \notin E_N,\, r(\lambda') > 0}} (\dim \lambda')^2\, (r(\lambda'))^{2n} \leq \sum_{\substack{\lambda \in \mathfrak{Y}(N) \\ \lambda \notin E_N,\, r(\lambda) \geq 0}} (\dim \lambda)^2\, (r(\lambda))^{2n}.$$

This implies that

$$S_{N,n} \leq 2 \sum_{\substack{\lambda \in \mathfrak{Y}(N) \\ \lambda \notin E_N,\, r(\lambda) \geq 0}} (\dim \lambda)^2\, (r(\lambda))^{2n} + \sum_{\lambda \in \{(N-1,1),(2,1^{N-2}),(1^N)\}} (\dim \lambda)^2\, (r(\lambda))^{2n}.$$

To end the proof, let us show that the sum of the contributions of the two integer partitions $(2,1^{N-2})$ and (1^N) is smaller than the contribution of $(N-1,1)$. By using the hook-length formula, we get:

$$(\dim(N-1,1))^2\, (r(N-1,1))^{2n} = (N-1)^2 \left(1 - \frac{2}{N}\right)^{2n};$$
$$(\dim(2,1^{N-2}))^2\, (r(2,1^{N-2}))^{2n} = (N-1)^2 \left(1 - \frac{4}{N}\right)^{2n};$$
$$(\dim(1^N))^2\, (r(1^N))^{2n} = \left(1 - \frac{2}{N}\right)^{2n}$$

The sum of the two last terms is smaller than

$$\begin{aligned}(N-1)^2 \left(1-\frac{2}{N}\right)^{2n} \left(\frac{1}{(N-1)^2} + \frac{(N-4)^2}{(N-2)^2}\right) &\leq (N-1)^2 \left(1-\frac{2}{N}\right)^{2n} \left(\frac{(N-4)^2+1}{(N-2)^2}\right) \\ &\leq (N-1)^2 \left(1-\frac{2}{N}\right)^{2n}.\end{aligned}$$

□

Step 2: estimating $\dim \lambda$ and $r(\lambda)$ knowing the value of λ_1. We now try to take into account our other observation: the integer partitions whose first parts are large are those which yield the largest coefficients $r(\lambda)$.

LEMMA 3.11. *Let $\lambda \in \mathfrak{Y}(N)$ be such that $\lambda_1 = N - k$.*

(1) If $1 \leq k \leq \frac{N}{2}$, then

$$r(\lambda) \leq 1 - \frac{2k(N+1-k)}{N^2}.$$

(2) If $k \geq \frac{N}{2}$, then

$$r(\lambda) \leq 1 - \frac{k}{N}.$$

Proof. Suppose first $\lambda_1 \geq \frac{N}{2}$. We have:

$$N^2\, r(\lambda) = \sum_{i=1}^{\ell(\lambda)} \lambda_i(\lambda_i - 2i + 2) = (N-k)^2 + \sum_{i=2}^{\ell(\lambda)} \lambda_i(\lambda_i - 2i + 2).$$

If $\lambda_2 < k = N - \lambda_1$, then we can raise the value of $r(\lambda)$ by transfering the last cell of a row $\lambda_{i \geq 3}$ to the end of λ_2. Indeed, $N^2\, r(\lambda)$ is modified by this shift of cells by

$$(\lambda_2+1)(\lambda_2-1) - \lambda_2(\lambda_2-2) + (\lambda_i-1)(\lambda_i-2i+1) - \lambda_i(\lambda_i-2i+2) = 2(\lambda_2 - \lambda_i + (i-1)) > 0.$$

So, the maximal possible value of $N^2\, r(\lambda)$ if $\lambda_1 = N - k \geq \frac{N}{2}$ is when $\ell(\lambda) = 2$ and $\lambda_2 = k$. This case yields the first upper bound. Suppose now $\lambda_1 \leq \frac{N}{2}$. Then,

$$\begin{aligned} N^2\, r(\lambda) = (N-k)^2 + \sum_{i=2}^{\ell(\lambda)} \lambda_i(\lambda_i - 2i + 2) &\leq (N-k)^2 + (N-k) \sum_{i=2}^{\ell(\lambda)} \lambda_i \\ &\leq (N-k)^2 + k(N-k) = N(N-k), \end{aligned}$$

whence the second upper bound. ☐

Lemma 3.12. *We suppose that $\lambda = (\lambda_1, \ldots, \lambda_r) \in \mathfrak{Y}(N)$ has its first part $\lambda_1 = N - k$ with $k \geq 1$. Then,*

$$\dim \lambda \leq \binom{N}{k} \dim(\lambda_2, \ldots, \lambda_r).$$

Proof. In order to construct a standard tableau with shape λ, we can first choose the $N-k$ entries of the first row, and there are $\binom{N}{k}$ possibilities. Then, once this choice is made, there remains k integers $i_1 < i_2 < \cdots < i_k$ to place in the rows $\lambda_2, \ldots, \lambda_r$; one can interpret this part of the tableau as a standard tableau with size k and shape $(\lambda_2, \ldots, \lambda_r)$. Note that all the choices of the first row are not compatible with all the choices of a standard tableau on the remaining rows; nonetheless, this enumeration leads to the inegality

$$\mathrm{card}(\mathrm{ST}(\lambda)) \leq \binom{N}{k} \mathrm{card}(\mathrm{ST}(\lambda_2, \ldots, \lambda_r)).$$ ☐

Step 3: computing an upper bound. We split the upper bound on $\frac{S_{N,n}}{2}$ in two parts:

$$T_{N,n} = \sum_{\substack{\lambda \in \mathfrak{Y}(N) \\ r(\lambda) \geq 0 \\ \frac{N}{2} \leq \lambda_1 < N}} (\dim \lambda)^2\, (r(\lambda))^{2n} \qquad ; \qquad U_{N,n} = \sum_{\substack{\lambda \in \mathfrak{Y}(N) \\ r(\lambda) \geq 0 \\ \lambda_1 < \frac{N}{2}}} (\dim \lambda)^2\, (r(\lambda))^{2n}.$$

By the first part of Lemma 3.11,

$$T_{N,n} \leq \sum_{k=1}^{\lfloor \frac{N}{2} \rfloor} \sum_{\substack{\lambda \in \mathfrak{Y}(N) \\ \lambda_1 = N-k}} (\dim \lambda)^2 \left(1 - \frac{2k(N+1-k)}{N^2}\right)^{2n}$$

$$\leq \sum_{k=1}^{\lfloor \frac{N}{2} \rfloor} \binom{N}{k}^2 \left(\sum_{\mu \in \mathfrak{Y}(k)} (\dim \mu)^2\right) \mathrm{e}^{-\frac{(\log N + c)\, 2k(N+1-k)}{N}} = \sum_{k=1}^{\lfloor \frac{N}{2} \rfloor} \frac{1}{k!} \left(N^{\downarrow k}\right)^2 \mathrm{e}^{-\frac{(\log N+c)\, 2k(N+1-k)}{N}}.$$

The last identity relies on the Parseval identity $k! = \sum_{\mu \in \mathfrak{Y}(k)} (\dim \mu)^2$; $N^{\downarrow k}$ is a notation for the falling factorial $N(N-1)\cdots(N-k+1)$. Since $N^{\downarrow k} \leq N^k$, we have

$$N^{\downarrow k}\, \mathrm{e}^{-\frac{(\log N) k(N+1-k)}{N}} \leq N^{\frac{k(k-1)}{N}};$$

$$T_{N,n} \leq \sum_{k=1}^{\lfloor \frac{N}{2} \rfloor} \frac{\mathrm{e}^{-kc}}{k!}\, N^{\frac{2k(k-1)}{N}} \leq \frac{1}{\sqrt{2\pi}} \sum_{k=1}^{\lfloor \frac{N}{2} \rfloor} \left(\mathrm{e}^{1-c}\, k^{-1}\, N^{\frac{2(k-1)}{N}}\right)^k,$$

the last inequality following from a weak form of the Stirling estimates: $k! \geq \sqrt{2\pi}(\frac{k}{\mathrm{e}})^k$ for any $k \geq 1$. If the integer N is fixed, then the map $f_N(x) = N^{\frac{2(x-1)}{N}}\, x^{-1}$ is decreasing on $[1, \frac{N}{2\log N}]$ and then increasing; therefore, its maximum on $[1, \frac{N}{2}]$ is

$$\max\left(f_N(1), f_N\left(\frac{N}{2}\right)\right) = \max\left(1, 2N^{-\frac{2}{N}}\right) \leq 2.$$

Suppose $c \geq 2$; then, $2\mathrm{e}^{1-c} < 1$ and

$$T_{N,n} \leq \frac{1}{\sqrt{2\pi}} \sum_{k=1}^{\infty} (2\mathrm{e}^{1-c})^k = \frac{2\mathrm{e}^{1-c}}{\sqrt{2\pi}(1-2\mathrm{e}^{1-c})} \leq \frac{2\mathrm{e}}{\sqrt{2\pi}(1-2\mathrm{e}^{-1})}\, \mathrm{e}^{-c} \leq 8.208\, \mathrm{e}^{-c}.$$

The other sum $U_{N,n}$ is treated by similar arguments. Indeed, the second part of Lemma 3.11 implies that

$$U_{N,n} \leq \sum_{\frac{N}{2} < k \leq N} \frac{1}{k!} \left(N^{\downarrow k}\right)^2 \left(1 - \frac{k}{N}\right)^{N(\log N + c)} \leq \sum_{\frac{N}{2} < k \leq N} \frac{1}{k!} \left(N^{\downarrow k}\right)^2 \left(1 - \frac{k}{N}\right)^{N \log N} \mathrm{e}^{-kc}.$$

Set $k = xN$ with $x \in (\frac{1}{2}, 1]$. The Stirling estimates yields:

$$N^{\downarrow k} = \frac{N!}{(N-k)!} \leq \frac{2}{\sqrt{1-x}}\, \mathrm{e}^{Nx\log N - N(1-x)\log(1-x) - Nx};$$

$$\frac{1}{k!} \leq \frac{1}{\sqrt{2\pi x N}}\, \mathrm{e}^{Nx - Nx\log N - Nx\log x}.$$

Gathering these upper bounds, we obtain

$$\frac{1}{k!} \left(N^{\downarrow k}\right)^2 \left(1 - \frac{k}{N}\right)^{N\log N} \leq \frac{4}{(1-x)\sqrt{2\pi x N}}\, \mathrm{e}^{N\log N\, f(x) + N\, g(x)} \leq \frac{4\mathrm{e}}{\sqrt{\pi N}}\, \mathrm{e}^{(N\log N - 1)\, f(x) + N\, g(x)},$$

with $f(x) = x + \log(1-x)$ and $g(x) = -x - x\log x - 2(1-x)\log(1-x)$. Now, both functions f and g are decreasing on $(\frac{1}{2}, 1]$, so we can use their values at $x = \frac{1}{2}$ to get:

$$\frac{1}{k!} \left(N^{\downarrow k}\right)^2 \left(1 - \frac{k}{N}\right)^{N\log N} \leq \frac{4\mathrm{e}}{\sqrt{\pi N}}\, \mathrm{e}^{(N\log N - N - 1)\,(\frac{1}{2} + \log(\frac{1}{2})) - N\,(\frac{1}{2}\log(\frac{1}{2}))}.$$

Since $\frac{1}{2}+\log(\frac{1}{2})<0$, this sequence goes to 0 as N goes to infinity, and one can even check that it is decreasing with N, and smaller than 11.05 for $N\geq 4$. So, for $c\geq 3$,

$$U_{N,n}\leq 11.05\sum_{k>\frac{N}{2}}\mathrm{e}^{-kc}\leq\frac{11.05\,\mathrm{e}^{-3c}}{1-\mathrm{e}^{-c}}\leq\frac{11.05}{\mathrm{e}^4-\mathrm{e}^2}\,\mathrm{e}^{-c}\leq 0.235\,\mathrm{e}^{-c}.$$

Putting everything together, we get:

PROPOSITION 3.13 (Upper bound for the random transposition model). *Consider the random transposition shuffle on $\mathfrak{S}(N)$ at time $n=\frac{N(\log N+c)}{2}$. We assume $N\geq 4$ and $c\geq 2$. Then, the sum $S_{N,n}$ which is an upper bound for $4\,(d_{\mathrm{TV}}(\mu_n,\mu_\infty))^2$ is smaller than $17\,\mathrm{e}^{-c}$. Therefore, for any $N\geq 4$ and any $c\geq 0$,*

$$d_{\mathrm{TV}}\left(\mu_{\frac{N(\log N+c)}{2}},\mu_\infty\right)\leq\mathrm{e}^{1-\frac{c}{2}}.$$

PROOF. The discussion above yields an upper bound on the total variation distance equal to $K\,\mathrm{e}^{-\frac{c}{2}}$ when $c\geq 2$, with a constant $K=\frac{\sqrt{17}}{2}$. Since the total variation distance is always smaller than 1, we can incorporate the case $c\in[0,2]$ in the upper bound by raising the value of K (taking for instance $K=\mathrm{e}$). □

We can follow the same steps to prove an upper bound for the top-with-random transposition shuffle. Set $n=N(\log N+c)$ with $c\in\mathbb{R}_+$, and

$$S^*_{N,n}=\sum_{\lambda\neq(N)}(\dim\lambda)\sum_{i\,|\,(\lambda_i,i)\text{ corner of }\lambda}\dim(\lambda\setminus(\lambda_i,i))\,(r^*(\lambda,i))^{2n},$$

where $r^*(\lambda,i)=\frac{\lambda_i-i+1}{N}$. In the following, we use notations similar to those used for the random transposition shuffle, with an exponent $*$ in order to make clear that we are now considering the top-with-random transposition model.

Step 1: getting rid of the signs of the eigenvalues. We have the exact analogue of Lemma 3.10:

LEMMA 3.14. *For any $N\geq 4$ and any $n\geq 1$, we have:*

$$S^*_{N,n}\leq\sum_{\lambda\neq(N)}(\dim\lambda)\sum_{\substack{i\,|\,(\lambda_i,i)\text{ corner of }\lambda\\ r^*(\lambda,i)\geq 0}}\dim(\lambda\setminus(\lambda_i,i))\,(r^*(\lambda,i))^{2n}.$$

PROOF. As in the proof of Lemma 3.10, we first consider partitions $\lambda\notin E_N$. The symmetrisation with respect to the first diagonal yields an involution of the pairs (λ,i), where i is the index of a row of a corner of λ: one sends λ to λ' and i to λ_i. This operation:

(1) yields by restriction an involution of the pairs (λ,i) with $\lambda\in E_N$;

(2) leaves $\dim\lambda$ and $\dim(\lambda\setminus(i,\lambda_i))$ invariant, as is clear by considering these quantities as the cardinalities of certain sets of standard tableaux.

Moreover,

$$r^*(\lambda',\lambda_i)=\frac{i-\lambda_i+1}{N}=\frac{2}{N}-r^*(\lambda,i)>-r^*(\lambda,i).$$

Therefore, we obtain by the same argument as in Lemma 3.10:

$$S^*_{N,n} \leq 2 \leq \sum_{\lambda \notin E_N} (\dim \lambda) \sum_{\substack{i \,|\, (\lambda_i,i) \text{ corner of } \lambda \\ r^*(\lambda,i) \geq 0}} \dim(\lambda \setminus (\lambda_i,i))\, (r^*(\lambda,i))^{2n}$$
$$+ \sum_{\lambda \in \{(N-1,1),(2,1^{N-2}),(1^N)\}} (\dim \lambda) \sum_{i \,|\, (\lambda_i,i) \text{ corner of } \lambda} \dim(\lambda \setminus (\lambda_i,i))\, (r^*(\lambda,i))^{2n}.$$

To conclude the proof, let us analyse the three integer partitions $(N-1,1)$, $(2,1^{N-2})$ and (1^N):

- the integer partition $(N-1,1)$ has two positive corners: $(N-1,1)$ which gives a contribution $(N-1)(N-2)\,(1-\frac{1}{N})^{2n}$, and $(1,2)$ which gives a contribution $(N-1)\,0^{2n}=0$.
- the integer partition $(2,1^{N-2})$ has one negative corner $(1,N-1)$ which gives a contribution $(N-1)(N-2)\,(\frac{3}{N}-1)^{2n}$, and one positive corner $(2,1)$ with a contribution $(N-1)\,(\frac{2}{N})^{2n}$.
- the integer partition (1^N) has one negative corner $(1,N)$ with a contribution $(\frac{2}{N}-1)^{2n}$.

It is then easy to prove that for any $n \geq 1$,

$$(N-1)(N-2)\left(1-\frac{1}{N}\right)^{2n} \geq (N-1)(N-2)\left(1-\frac{3}{N}\right)^{2n} + \left(1-\frac{2}{N}\right)^{2n}$$

assuming $N \geq 4$. □

Step 2: only using the corners with the largest contents; surprisingly, the computation is then much easier than before. We want to use the following fact, which is a consequence of the branching rules (Proposition 2.31): for any integer partition $\lambda \in \mathfrak{Y}(N)$,

$$\sum_{i \,|\, (\lambda_i,i) \text{ corner of } \lambda} \dim(\lambda \setminus (\lambda_i,i)) = \dim \mathrm{Res}^{\mathfrak{S}(N)}_{\mathfrak{S}(N-1)}(S^\lambda) = \dim \lambda.$$

To this purpose, we need to have a common upper bound on all the coefficients $r^*(\lambda,i)$ for λ fixed integer partition and i among the indices of corners of λ. However, if $\lambda_1 = N-k$ with $k \geq 1$, then we have obviously for any corner (λ_i,i) of λ:

$$r^*(\lambda,i) = \frac{\lambda_i - i + 1}{N} \leq \frac{\lambda_1}{N} = 1 - \frac{k}{N}.$$

Therefore, with computations identical to those of the random transposition model,

$$S^*_{N,n} \leq 2 \sum_{k=1}^{N} \sum_{\substack{\lambda \in \mathfrak{Y}(N) \\ \lambda_1 = N-k}} (\dim \lambda)^2 \left(1-\frac{k}{N}\right)^{2n} \leq 2 \sum_{k=1}^{N} \frac{1}{k!}\,(N^{\downarrow k})^2 \left(1-\frac{k}{N}\right)^{2n}.$$

Now, n is equal to *twice* the quantity considered for the other model, so we have:

$$S^*_{N,n} \leq 2 \sum_{k=1}^{N} \frac{1}{k!}\,(N^{\downarrow k})^2\, \mathrm{e}^{-2k(\log N + c)} \leq 2 \sum_{k=1}^{N} \frac{1}{k!}\, \mathrm{e}^{-2kc} \leq 2\left(\mathrm{e}^{\mathrm{e}^{-2c}} - 1\right).$$

If c is larger than 1, then $\mathrm{e}^{\mathrm{e}^{-2c}} - 1 \leq \frac{3}{2}\,\mathrm{e}^{-2c}$, so we obtain the following:

PROPOSITION 3.15 (Upper bound for the top-with-random transposition model). *Consider the top-with-random transposition shuffle on $\mathfrak{S}(N)$ at time $n = N(\log N + c)$. We assume $N \geq 4$ and $c \geq 1$. Then, the sum $S^*_{N,n}$ which is an upper bound for $4\,(d_{\mathrm{TV}}(\mu_n,\mu_\infty))^2$ is smaller than $3\,\mathrm{e}^{-2c}$. Therefore, for any $N \geq 4$ and any $c \geq 0$,*

$$d_{\mathrm{TV}}\left(\mu_{N(\log N + c)}, \mu_\infty\right) \leq \mathrm{e}^{1-c}.$$

REMARK 3.16. The inequalities stated in Propositions 3.13 and 3.15 are not optimal, in particular with respect to the dependency in the parameter c. For the random transposition shuffle, one can show that

$$d_{\mathrm{TV}}(\mu_{\frac{N(\log N+c)}{2}}, \mu_\infty) = O(\mathrm{e}^{-c})$$

instead of $O(\mathrm{e}^{-\frac{c}{2}})$; see the references at the end of the chapter.

4. Discriminating random variables and the cut-off phenomenon

So far we have shown that, after a time larger than $\frac{1}{2}(N \log N)$ or $N \log N$, the total variation distance to the stationary (uniform) distribution is small. We are now going to prove that on the contrary, if one looks at the distribution at a time n smaller than $\frac{1}{2}(N \log N)$ or $N \log N$, then the total variation distance $d_{\mathrm{TV}}(\mu_n, \mu_\infty)$ stays large (close to 1). This phase transition is called a *cut-off phenomenon* and it is a common feature of many Markov chains living on a large finite space state. It is established thanks to the following general lemma:

LEMMA 3.17 (Discriminating random variable). *Let μ and ν be two probability distributions on a measured space $(\mathfrak{X}, \mathscr{X})$. We suppose given a real-valued random variable X on this space such that:*

$$\mathbb{E}_\mu[X] = a \qquad ; \qquad \mathbb{E}_\nu[X] = b \qquad ; \qquad \max(\mathrm{Var}_\mu(X), \mathrm{Var}_\nu(X)) \leq v.$$

Then,

$$d_{\mathrm{TV}}(\mu, \nu) \geq 1 - \frac{8v}{(b-a)^2}.$$

PROOF. By symmetry, we can assume $b > a$; we consider the event $E = \{X \geq \frac{a+b}{2}\}$. By the Bienaymé–Chebyshev inequality,

$$\mu(E) \leq \mu\left(\left\{|X - \mathbb{E}_\mu[X]| \geq \frac{b-a}{2}\right\}\right) \leq \frac{4v}{(b-a)^2}.$$

Similarly,

$$1 - \nu(E) \leq \nu\left(\left\{|X - \mathbb{E}_\nu[X]| \geq \frac{b-a}{2}\right\}\right) \leq \frac{4v}{(b-a)^2}.$$

Therefore, $|\mu(E) - \nu(E)| \geq 1 - \frac{8v}{(b-a)^2}$, and *a fortiori* the total variation distance between μ and ν is larger than this quantity. □

Now the question is: how to choose a random variable X which discriminates the distribution μ_n before cut-off time and the uniform distribution μ_∞? If one looks at the proofs of the upper-bounds in the previous section, one sees that the behavior of the two series $S_{N,n}$ and $S^*_{N,n}$ is dictated by the term $\lambda = (N-1, 1)$, which yields the largest coefficient $r(\lambda)$ or $r^*(\lambda, i)$. Therefore, it is tempting to use discriminating functions on $G = \mathfrak{S}(N)$ which are related to the irreducible representation $S^{(N-1,1)}$; for instance, the character of this representation. This approach is successful for many random walks on groups or related objects, and here it will indeed enable the proof of the cut-off phenomenon. We start by identifying the representation $S^{(N-1,1)}$.

PROPOSITION 3.18 (Geometric representation). *Consider the representation of $\mathfrak{S}(N)$ by permutation of the coordinates of the vectors in $\mathbb{C}^{N \geq 2}$. It admits as a subrepresentation*

$$W = \{(x_1, \ldots, x_N) \in \mathbb{C}^N \,|\, x_1 + \cdots + x_N = 0\},$$

which has dimension $N-1$. The space W is an irreducible representation of $\mathfrak{S}(N)$ with label $\lambda = (N-1, 1)$.

PROOF. Note that the standard scalar product on $\mathbb{C}^N$ is invariant by the action of $\mathfrak{S}(N)$, and that the orthogonal of W is the one-dimensional space $V = \{(x,x,\ldots,x)\,|\,x \in \mathbb{C}\}$, on which $\mathfrak{S}(N)$ acts trivially. Thus,

$$\mathbb{C}^N = S^{(N)} \oplus W$$

and this decomposition proves that

$$\forall \sigma \in \mathfrak{S}(N),\ \ \mathrm{ch}^{\mathbb{C}^N}(\sigma) = 1 + \mathrm{ch}^W(\sigma),$$

since the character of the trivial one-dimensional representation $S^{(N)}$ of $\mathfrak{S}(N)$ is the constant function equal to 1. However, the character of $\mathbb{C}^N$ is:

$$\mathrm{ch}^{\mathbb{C}^N}(\sigma) = \sum_{i=1}^N \langle e_i \,|\, \sigma \cdot e_i\rangle = \text{number of fixed points of } \sigma.$$

Thus, $\mathrm{ch}^W(\sigma) = \mathrm{Fix}(\sigma) - 1$, where $\mathrm{Fix}(\sigma)$ denotes the number of fixed points of σ in $[\![1,N]\!]$. Now, in order to prove that W is isomorphic to $S^{(N-1,1)}$ (hence irreducible), it suffices to compute separately $\mathrm{ch}^{(N-1,1)}(\sigma)$ and to prove that it is equal to $\mathrm{Fix}(\sigma) - 1$. Note that by the Jacobi–Trudy formula,

$$s_{(N-1,1)} = \det\begin{pmatrix} h_{N-1} & h_N \\ 0 & h_1 \end{pmatrix} = h_{N-1}\,h_1 - h_N = h_{N-1}\,p_1 - h_N.$$

By Corollary 2.27, for any $\mu \in \mathfrak{Y}(N)$, the coefficient of the power sum p_μ in $s_{(N-1,1)}$ equals $(z_\mu)^{-1}$ times the value of the character $\mathrm{ch}^{(N-1,1)}$ on a permutation with cycle-type μ. On the other hand, the expansion of h_{N-1} and h_N in power sums is known:

$$h_{N-1} = \sum_{\nu \in \mathfrak{Y}(N-1)} \frac{p_\nu}{z_\nu} \qquad ; \qquad h_N = \sum_{\mu \in \mathfrak{Y}(N)} \frac{p_\mu}{z_\mu}.$$

We now remark that

$$h_{N-1}\,p_1 = \sum_{\substack{\mu \in \mathfrak{Y}(N) \\ \mu = (\nu,1) \text{ with } \nu \in \mathfrak{Y}(N-1)}} \frac{p_\mu}{z_\nu} = \sum_{\mu \in \mathfrak{Y}(N)} m_1(\mu)\,\frac{p_\mu}{z_\mu},$$

where $m_1(\mu)$ is the multiplicity of 1 as a part of μ, and the number of fixed points of a permutation $\sigma \in \mathfrak{S}(N)$ with cycle-type μ. Therefore,

$$s_{(N-1,1)} = h_{N-1}\,p_1 - h_N = \sum_{\mu \in \mathfrak{Y}(N)} (m_1(\mu) - 1)\,\frac{p_\mu}{z_\mu},$$

so $\mathrm{ch}^{(N-1,1)}(\mu) = m_1(\mu) - 1$ as wanted. □

In the sequel we set $X = \mathrm{ch}^{(N-1,1)} : \mathfrak{S}(N) \to \mathbb{R}$. If μ is a probability distribution on $\mathfrak{S}(N)$ and σ is chosen according to μ, then $X(\sigma)$ is a real-valued random variable. Let us compute its expectation under μ_n and μ_∞ for the two random walks considered in this chapter.

PROPOSITION 3.19 (Expectation of the discriminating variable). *We have for any $n \geq 1$ and any $N \geq 3$:*

$$\mathbb{E}_{\mu_n,\mathrm{RT}}[X] = (N-1)\left(1 - \frac{2}{N}\right)^n;$$

$$\mathbb{E}_{\mu_n,\mathrm{TWRT}}[X] = (N-2)\left(1 - \frac{1}{N}\right)^n;$$

$$\mathbb{E}_{\mu_\infty}[X] = 0.$$

Proof. The third identity comes from the orthogonality of characters:

$$\mathbb{E}_{\mu_\infty}[X] = \frac{1}{N!} \sum_{\sigma \in \mathfrak{S}(N)} \mathrm{ch}^{(N-1,1)}(\sigma) = \left\langle \mathrm{ch}^{(N)} \,\middle|\, \mathrm{ch}^{(N-1,1)} \right\rangle = 0.$$

Now, note that for any irreducible character ch^λ and any random walk on $\mathfrak{S}(N)$ with generator μ, we have:

$$\mathbb{E}_{\mu_n}[\mathrm{ch}^\lambda(\sigma)] = \sum_{\sigma \in \mathfrak{S}(N)} \mu_n(\sigma)\, \mathrm{tr}(\rho^\lambda(\sigma)) = \mathrm{tr}\left(\widehat{\mu_n(\lambda)}\right) = \mathrm{tr}\left((\widehat{\mu}(\lambda))^n\right).$$

Consider first the random transposition model. Each matrix $\widehat{\mu_{\mathrm{RT}}}(\lambda)$ is proportional to id_{S^λ}, and equal to $\widetilde{\mu_{\mathrm{RT}}}(\lambda)\, \frac{\mathrm{id}_{S^\lambda}}{\dim \lambda}$. Therefore, by Proposition 3.7,

$$\mathbb{E}_{\mu_n,\mathrm{RT}}[\mathrm{ch}^\lambda(\sigma)] = \frac{1}{(\dim \lambda)^{n-1}}\, (\widetilde{\mu_{\mathrm{RT}}}(\lambda))^n = (\dim \lambda) \left(\frac{1}{N^2} \sum_{i=1}^{\ell(\lambda)} \lambda_i(\lambda_i - 2i + 2) \right)^n.$$

The particular case $\lambda = (N-1,1)$ gives the first formula. For the top-with-random transposition model, we have similarly by Proposition 3.8

$$\mathbb{E}_{\mu_n,\mathrm{TWRT}}[\mathrm{ch}^\lambda(\sigma)] = \sum_{i \,|\, (\lambda_i,i) \text{ corner of } \lambda} \dim(\lambda \setminus (\lambda_i,i)) \left(\frac{\lambda_i - i + 1}{N} \right)^n.$$

The case $\lambda = (N-1,1)$ yields the second identity. □

In order to apply Lemma 3.17, we also need to compute the second moment of X under the probability measures of the previous proposition. This computation is related to the square of the irreducible character $\mathrm{ch}^{(N-1,1)}$, and to the internal tensor product of representations $S^{(N-1,1)} \boxtimes S^{(N-1,1)}$. For our purpose, we shall only require the following result:

Proposition 3.20 (Square of the geometric representation). *For any $N \geq 4$ and any $\sigma \in \mathfrak{S}(N)$, we have*

$$(\mathrm{ch}^{(N-1,1)}(\sigma))^2 = \mathrm{ch}^{(N)}(\sigma) + \mathrm{ch}^{(N-1,1)}(\sigma) + \mathrm{ch}^{(N-2,2)}(\sigma) + \mathrm{ch}^{(N-2,1,1)}(\sigma).$$

Proof. In the sequel we fix a permutation $\sigma \in \mathfrak{S}(N)$ with cycle-type μ. The left-hand side of the formula equals $(\mathrm{Fix}(\sigma)-1)^2 = (m_1(\mu)-1)^2$. To compute the right-hand side, we combine as before the Frobenius–Schur formula and the Jacobi–Trudy formula:

$$\begin{aligned}
&\mathrm{ch}^{(N)}(\sigma) + \mathrm{ch}^{(N-1,1)}(\sigma) + \mathrm{ch}^{(N-2,2)}(\sigma) + \mathrm{ch}^{(N-2,1,1)}(\sigma)\\
&= \left\langle s_{(N)} + s_{(N-1,1)} + s_{(N-2,2)} + s_{(N-2,1,1)} \,\middle|\, p_\mu \right\rangle\\
&= \left\langle h_N + (h_{N-1}\, h_1 - h_N) + (h_{N-2}\, h_2 - h_{N-1}\, h_1) + (h_{N-2}\,(h_1)^2 + h_N - h_{N-2}\, h_2 - h_{N-1}\, h_1) \,\middle|\, p_\mu \right\rangle\\
&= \left\langle h_N - h_{N-1}\, p_1 + h_{N-2}\,(p_1)^2 \,\middle|\, p_\mu \right\rangle.
\end{aligned}$$

We now reason as in the proof of Proposition 3.18.

- If $m_1(\mu) = 0$, then p_μ appears in $h_N - h_{N-1}\, p_1 + h_{N-2}\,(p_1)^2$ with a coefficient $\frac{1}{z_\mu}$ (by using the formula of expansion of h_N in power sums), so the scalar product is equal to $1 = (0-1)^2$.
- If $m_1(\mu) = 1$, then p_μ appears in $h_N - h_{N-1}\, p_1 + h_{N-2}\,(p_1)^2$ with a coefficient $\frac{1}{z_\mu} - \frac{1}{z_\nu}$ with $\mu = (\nu, 1)$. However, as $m_1(\nu) = 0$, $z_\mu = z_\nu$ and thus, the scalar product equals $0 = (1-1)^2$.

- Finally, if $m_1(\mu) = m \geq 2$, let θ be the partition obtained from μ by removing all the parts with size 1. Then p_μ appears in $h_N - h_{N-1}\, p_1 + h_{N-2}\, (p_1)^2$ with a coefficient
$$\frac{1}{z_{(\theta,1^m)}} - \frac{1}{z_{(\theta,1^{m-1})}} + \frac{1}{z_{(\theta,1^{m-2})}} = \frac{1}{z_\theta}\left(\frac{1}{m!} - \frac{1}{(m-1)!} + \frac{1}{(m-2)!}\right) = \frac{(m-1)^2}{z_\mu}.$$
Therefore, the scalar product equals again $(m_1(\mu)-1)^2$. □

Corollary 3.21 (Variance of the discriminating variable). *We have for any $n \geq 1$ and any $N \geq 5$:*
$$\begin{aligned}
\mathrm{Var}_{\mu_n,\mathrm{RT}}[X] &= 1 + (N-1)\left(1-\frac{2}{N}\right)^n + \frac{N^2-3N+2}{2}\left(1-\frac{4}{N}\right)^n - \frac{N^2-N+2}{2}\left(1-\frac{2}{N}\right)^{2n};\\
\mathrm{Var}_{\mu_n,\mathrm{TWRT}}[X] &= 1 + (N-2)\left[\left(1-\frac{1}{N}\right)^n + \left(\frac{1}{N}\right)^n + \left(-\frac{1}{N}\right)^n\right] + (N^2-5N+5)\left(1-\frac{2}{N}\right)^n\\
&\quad - (N^2-4N+4)\left(1-\frac{1}{N}\right)^{2n};\\
\mathrm{Var}_{\mu_\infty}[X] &= 1.
\end{aligned}$$

Proof. The previous proposition and the general formula for the expectation of a random character value $\mathrm{ch}^\lambda(\sigma)$ with $\lambda \in \mathfrak{Y}(N)$ and $\sigma \sim \mu_n$ or μ_∞ enables the computation of $\mathbb{E}[X^2]$ under the three probability measures. Then, one substacts $\mathbb{E}[X]^2$ to obtain the variance; the calculations are straightforward. □

We are now ready to prove lower bounds on the total variation distance before the cut-off time:

Proposition 3.22 (Lower bound for the random transposition model). *Consider the random transposition shuffle on $\mathfrak{S}(N)$ at time $n = \frac{N(\log N - c)}{2}$. We assume $N \geq 5$ and $c \geq 0$. Then,*
$$d_{\mathrm{TV}}\left(\mu_{\frac{N(\log N-c)}{2}}, \mu_\infty\right) \geq 1 - \mathrm{e}^{3-c}.$$

Proof. Let us remark that in the formula for the variance of X under μ_n, the last term is larger than the second last term. Therefore,
$$\mathrm{Var}_{\mu_n,\mathrm{RT}}[X] \leq v = 1 + (N-1)\left(1-\frac{2}{N}\right)^n,$$
and we can apply Lemma 3.17 with $a = 0$ and $b = v-1$. We get:
$$d_{\mathrm{TV}}\left(\mu_{\frac{N(\log N-c)}{2}}, \mu_\infty\right) \geq 1 - 8\left(\frac{1}{b} + \frac{1}{b^2}\right).$$
Let us then evaluate b. For $N \geq 5$,
$$\log\left(1-\frac{2}{N}\right) \geq -\frac{2}{N} - \frac{1}{N\log N} \qquad ; \qquad b \geq \frac{N-1}{N}\,\mathrm{e}^{c-\frac{1}{2}} \geq \frac{4}{5}\,\mathrm{e}^{c-\frac{1}{2}}.$$
Consequently,
$$d_{\mathrm{TV}}\left(\mu_{\frac{N(\log N-c)}{2}}, \mu_\infty\right) \geq 1 - 10\,\mathrm{e}^{\frac{1}{2}-c} - \frac{50}{4}\,\mathrm{e}^{1-2c}.$$
If $c \geq 3$, then the lower bound obtained is larger than $1-\mathrm{e}^{3-c}$. The case $c \leq 3$ is also true since the total variation distance is always positive. □

Proposition 3.23 (Lower bound for the top-with-random transposition model). *Consider the random transposition shuffle on $\mathfrak{S}(N)$ at time $n = N(\log N - c)$. We assume $N \geq 5$ and $c \geq 0$. Then,*
$$d_{\mathrm{TV}}\left(\mu_{N(\log N-c)}, \mu_\infty\right) \geq 1 - \mathrm{e}^{3-c}.$$

Proof. As before, we can bound the variance of X under μ_n by

$$\mathrm{Var}_{\mu_n,\mathrm{RT}}[X] \leq v = 1 + (N-2)\left(1-\frac{1}{N}\right)^n,$$

and we can apply Lemma 3.17 with $a = 0$ and $b = v - 1$:

$$d_{\mathrm{TV}}\left(\mu_{N(\log N - c)}, \mu_\infty\right) \geq 1 - 8\left(\frac{1}{b} + \frac{1}{b^2}\right).$$

We can evaluate b as follows:

$$\log\left(1-\frac{1}{N}\right) \geq -\frac{1}{N} - \frac{1}{5N\log N} \qquad ; \qquad b \geq \frac{N-2}{N}\,\mathrm{e}^{c-\frac{1}{5}} \geq \frac{3}{5}\,\mathrm{e}^{c-\frac{1}{5}}.$$

Therefore,

$$d_{\mathrm{TV}}\left(\mu_{N(\log N - c)}, \mu_\infty\right) \geq 1 - \frac{40}{3}\,\mathrm{e}^{\frac{1}{5}-c} - \frac{200}{9}\,\mathrm{e}^{\frac{2}{5}-2c}.$$

Again, one can check that if $c \geq 3$, then the lower bound is larger than $1 - \mathrm{e}^{3-c}$; and in the case $c \leq 3$, the inequality is trivially satisfied. □

Let us summarise the results obtained in this chapter.

Theorem 3.24 (Cut-off phenomenon). *Consider the random transposition shuffle or the top-with-random transposition shuffle on $\mathfrak{S}(N)$, with $N \geq 5$. Set*

$$n_{\mathrm{mixing}} = \begin{cases} \frac{N\log N}{2} & \textit{for the random transposition model}, \\ N\log N & \textit{for the top-with-random transposition model}. \end{cases}$$

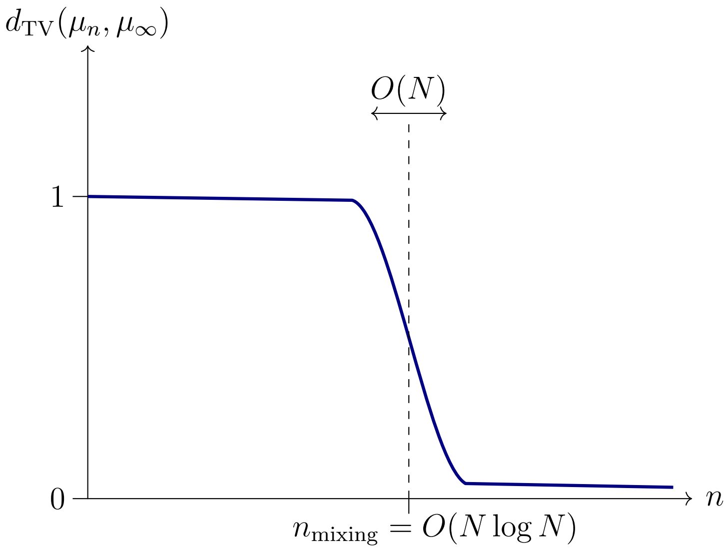
(1) For any $\varepsilon > 0$,

$$d_{\mathrm{TV}}(\mu_{(1+\varepsilon)n_{\mathrm{mixing}}}, \mu_\infty) = \begin{cases} O(N^{-\frac{\varepsilon}{2}}) & \textit{for the random transposition model}, \\ O(N^{-\varepsilon}) & \textit{for the top-with-random transposition model}. \end{cases}$$

(2) For any $\varepsilon > 0$,

$$1 - d_{\mathrm{TV}}(\mu_{(1-\varepsilon)n_{\mathrm{mixing}}}, \mu_\infty) = O(N^{-\varepsilon}).$$

In fact, one could improve the first item of the theorem and prove an upper bound of the form $O(N^{-\varepsilon})$ for both cases. The drawing below illustrates the phase transition:



It turns out that the three other examples of random walks on $\mathfrak{S}(N)$ introduced in Chapter 1 also exhibit a cut-off, but with different mixing times; see the references and exercises hereafter.

References

The upper bound on the total variation distance between the law of a random walk on group at time n and the uniform measure appears in [DS81, Lemma 14], in the setting of the random transposition model; it might have been stated before by other authors. The arguments which we gave for the computation of an upper bound for this model differ a bit from those of [DS81] or of [CST08, Chapter 10]; as far as we know, all the representation-theoretic proofs of the cut-off phenomenon for this model gather the integer partitions according to their first part, and then study several zones for this first part (in our case, $\lambda_1 \geq \frac{N}{2}$ and $\lambda_1 < \frac{N}{2}$). More recently, the *profile* of the convergence to stationarity has been computed by Teyssier in [Tey20]. Hence, it has been shown that

$$d_{\mathrm{TV}}\big(\mu_{\frac{N(\log N+c)}{2}}, \mu_\infty\big) \to_{N\to\infty} d_{\mathrm{TV}}\big(\mathcal{P}(1+\mathrm{e}^{-c}), \mathcal{P}(1)\big)$$

with Poisson laws on the right-hand side.

For the top-with-random transposition model, most of the arguments appear in [Dia88] (in particular the proof of the lower bound), but as far as we know a complete proof of the upper bound has never been published before (although it is much simpler than for the random transposition model). We refer the reader to the paragraph at the end of Chapter 1 for a list of articles in which are computed the mixing times of the five models 1.2–1.6. These mixing times are the following:

model	RT (Ex. 1.2)	TWRT (Ex. 1.3)	TTRC (Ex. 1.4)	AT (Ex. 1.5)	RS (Ex. 1.6)
mixing time	$\frac{N\log N}{2}$	$N\log N$	$N\log N$	$\frac{N^3\log N}{2\pi^2}$	$\frac{3\log N}{2\log 2}$

Each model exhibits a cut-off in a small window around the mixing time. The top-to-random cycle model is actually quite easy, if one uses the technique of uniform stopping times; the most difficult case is the model of the adjacent transpositions. The riffle shuffle model is related to two subalgebras of the group algebra $\mathbb{C}\mathfrak{S}(N)$: the descent algebra $\mathfrak{D}(N)$ of Solomon, and its commutative subalgebra spanned by the classes of permutations with a fixed number of descents. For this model, the profile of the convergence to stationarity is also known, and it is related to the Gaussian distribution:

$$d_{\mathrm{TV}}\big(\mu_{\frac{3\log(Nt)}{2\log 2}}, \mu_\infty\big) \to_{N\to\infty} \mathbb{P}\left[|\mathcal{N}(0,1)| \leq \frac{1}{4\sqrt{3t^3}}\right].$$

For a general treatment of the problem of mixing times of Markov chains, we refer to [LPW17].

Exercises

(1) **Submultiplicativity of the distance to equilibrium.** Remark that for any probability measures $\mu = \sum_{\sigma\in\mathfrak{S}(N)} \mu(\sigma)\,\sigma$ and $\nu = \sum_{\sigma\in\mathfrak{S}(N)} \nu(\sigma)\,\sigma$ on the symmetric group $\mathfrak{S}(N)$, one has

$$(\mu - \mathrm{Haar})(\nu - \mathrm{Haar}) = \mu\nu - \mathrm{Haar}$$

in the group algebra $\mathbb{C}\mathfrak{S}(N)$, where $\mathrm{Haar} = \frac{1}{N!}\sum_{\sigma\in\mathfrak{S}(N)} \sigma$. Deduce from this observation that, given a random walk on $\mathfrak{S}(N)$ with generator μ and marginal laws $\mu_n = \mu^n$, one has

$$d_{\mathrm{TV}}(\mu_{n+m}, \mathrm{Haar}) \leq 2\, d_{\mathrm{TV}}(\mu_n, \mathrm{Haar})\, d_{\mathrm{TV}}(\mu_m, \mathrm{Haar})$$

for any $n, m \in \mathbb{N}$.

(2) **A proof of the formula for the character of a transposition.** For $\lambda \in \mathfrak{Y}(N)$, we denote $\mu = \lambda + \rho$, where $\rho = (N-1, N-2, \ldots, 1, 0)$; μ is therefore a decreasing sequence of N non-negative integers. Recall that $s_\lambda(x_1, \ldots, x_N) = \frac{\det(x_i^{\mu_j})_{1\leq i,j\leq N}}{\det(x_i^{\rho_j})_{1\leq i,j\leq N}}$.

(a) Show that the dimension $\dim \lambda$ is the coefficient of $s_\lambda(x_1, \ldots, x_N)$ in the expansion in Schur polynomials of $(x_1 + \cdots + x_N)^N$. Deduce from this observation the following formula:

$$\dim \lambda = [x^\mu] \left((x_1 + \cdots + x_N)^N \sum_{\sigma \in \mathfrak{S}(N)} \varepsilon(\sigma)\, \sigma \cdot (x_1^{\rho_1} \cdots x_N^{\rho_N}) \right).$$

(b) In the previous formula, show that the term corresponding to a permutation σ is

$$\varepsilon(\sigma) \frac{N!}{\prod_{i=1}^N (\mu_i - N + \sigma(i))!}.$$

Prove then the following formula:

$$\dim \lambda = N! \det \left(\frac{1}{(\mu_i - N + j)!} \right)_{1 \leq i,j \leq N}.$$

(c) Use the standard transformation properties of determinants to rewrite the previous formula as:

$$\dim \lambda = \frac{N!}{\prod_{i=1}^N (\mu_i)!} \Delta(\mu_1, \ldots, \mu_N),$$

where $\Delta(\mu_1, \ldots, \mu_N)$ is the Vandermonde determinant $\prod_{1 \leq i < j \leq N} (\mu_i - \mu_j)$.

We can adapt the previous reasoning in order to compute $\mathrm{ch}^\lambda(21^{N-2})$, thereby proving Proposition 3.7 without using Jucys–Murphy elements.

(d) Show that

$$\mathrm{ch}^\lambda(21^{N-2}) = \sum_i [x^{\mu - 2e_i}] \left((x_1 + \cdots + x_N)^{N-2} \sum_{\sigma \in \mathfrak{S}(N)} \varepsilon(\sigma)\, \sigma \cdot (x_1^{\rho_1} \cdots x_N^{\rho_N}) \right)$$

where $e_i = (0, \ldots, 1_i, \ldots, 0)$, and where the sum runs over indices i such that $\mu_i \geq 2$. Adapt the reasoning of the previous questions to prove that

$$\mathrm{ch}^\lambda(21^{N-2}) = \sum_i \frac{(N-2)!}{(\mu_1)!(\mu_2!) \cdots (\mu_i - 2)! \cdots (\mu_N)!} \Delta(\mu_1, \ldots, \mu_i - 2, \ldots, \mu_N).$$

(e) Prove that

$$N(N-1)\, \chi^\lambda(21^{N-2}) = \sum_{i=1}^N \mu_i(\mu_i - 1) \prod_{j \neq i} \frac{\mu_i - \mu_j - 2}{\mu_i - \mu_j}.$$

(f) We introduce the polynomial $\phi_\lambda(z) = \prod_{i=1}^N (z - \mu_i)$, and the rational function $F(z) = -\frac{z(z-1)}{2} \frac{\phi_\lambda(z-2)}{\phi_\lambda(z)}$. Show that the quantity of the previous question is the sum of the residues of $F(z)$ at its (simple) poles $\mu_1, \ldots, \mu_N \in \mathbb{C}$. By considering the behavior of $F(z)$ when z grows to infinity, conclude that

$$N(N-1)\, \chi^\lambda(21^{N-2}) = [z^{-1}] \left(-\frac{z(z-1)}{2} \prod_{i=1}^N \left(1 - \frac{2}{z - \mu_i}\right) \right),$$

where the right-hand side is expanded as a Laurent series in z.

(g) Recover finally the formula $\chi^\lambda(21^{N-2}) = \frac{1}{N(N-1)} \sum_{i=1}^{\ell(\lambda)} \lambda_i(\lambda_i - 2i + 1)$.

(3) **Analysis of the top-to-random cycle model.** This exercise introduces the method of *strong uniform times* in order to estimate the mixing time of random walks on $\mathfrak{S}(N)$. Given a Markov chain $(\sigma_n)_{n\in\mathbb{N}}$ on $\mathfrak{S}(N)$, recall that a *stopping time* is a random variable T with values in $\mathbb{N}\sqcup\{+\infty\}$ such that $\{T=n\}\in\mathscr{F}_n$ for any n, $(\mathscr{F}_n)_{n\in\mathbb{N}}$ being the filtration spanned by the Markov chain. A strong uniform time is a stopping time T such that, for any $n\in\mathbb{N}$,

$$\mathbb{P}[\sigma_n=\sigma\,|\,T=n]=\frac{1}{N!}.$$

If the stopping time T is almost surely finite, then the condition above states that σ_T is independent from T and uniformly distributed on $\mathfrak{S}(N)$.

(a) Show that given a strong uniform time T for a random walk $(\sigma_n)_{n\in\mathbb{N}}$ with generator μ, we have

$$d_{\mathrm{TV}}(\mu_n,\mathrm{Haar})\leq\mathbb{P}[T>n].$$

One can first show that, given $n\geq m$, one has

$$\mathbb{P}[T=m\text{ and }\sigma_n\in A]=\mathbb{P}[T=m]\,\frac{|A|}{N!}$$

for any subset $A\subset\mathfrak{S}(N)$.

(b) We consider the top-to-random cycle model, with generator

$$\mu=\frac{1}{N}\,\mathrm{id}+\frac{1}{N}\sum_{i=2}^{N}(i,i-1,\ldots,2,1).$$

It is convenient to think of the corresponding random permutations $(\sigma_n)_{n\in\mathbb{N}}$ as the configurations of a deck of cards. At each time n of the random walk, one takes the card on top of the deck, and one inserts it at a random place $i\in[\![1,N]\!]$ chosen uniformly. Denote $T-1=\inf\{n\in\mathbb{N}\,|\,\sigma_n(1)=N\}$. Show that the positions $(\sigma_0^{-1})(N),(\sigma_1^{-1})(N),\ldots,(\sigma_{T-1})^{-1}(N)$ of the card N form a non-increasing sequence with $\sigma_{i+1}^{-1}(N)-\sigma_i^{-1}(N)\in\{0,-1\}$ for all $i<T-1$.

(c) Let $m\leq T-1$. Show that conditionally to the event

$$\{\sigma_m^{-1}(N)=k,\ \sigma_m([\![k+1,N]\!])=S\},$$

the sequence $(\sigma_m(k+1),\ldots,\sigma_m(N))$ is a random permutation of the set S with uniform distribution. Deduce from this observation that T is a strong uniform time for the top-to-random model.

(d) Show that T is distributed as the sum $G_1+G_2+\cdots+G_N$ of independent geometric variables, with each variable G_i geometric with parameter $\frac{i}{N}$:

$$\mathbb{P}[G_i=k]=\frac{i}{N}\left(1-\frac{i}{N}\right)^{k-1}.$$

Show that $\mathbb{E}[T]=\sum_{i=1}^{N}\frac{N}{i}\leq N\log N+N$, where γ is the Euler–Mascheroni constant; and that $\mathrm{var}(T)=\sum_{i=1}^{N}\frac{N}{i}(\frac{N}{i}-1)\leq\frac{N^2\pi^2}{6}$.

(e) Show that, with $n=N(\log N+c)$, one has $d_{\mathrm{TV}}(\mu_n,\mathrm{Haar})=O(c^{-2})$.

(f) Denote $T_j-1=\inf\{n\in\mathbb{N}\,|\,\sigma_n(1)=j\}$. Show that if $n<T_j$, then the cards $j,j+1,\ldots,N$ are in this order in the word of σ_n. Deduce from this observation that

$$d_{\mathrm{TV}}(\mu_n,\mathrm{Haar})\geq 1-\mathbb{P}[T_j\leq n]-\frac{1}{(N-j+1)!}$$

Show that T_j is distributed as the sum $G_{N-j+1} + G_{N-j+2} + \cdots + G_N$. Prove the estimates $\mathbb{E}[T_j] \geq N \log(\frac{N+1}{N-j+1})$ and $\mathrm{var}(T) \leq \frac{N^2}{N-j}$.

(g) Show that, with $n = N(\log N - c)$, one has $d_{\mathrm{TV}}(\mu_n, \mathrm{Haar}) \geq 1 - O(\mathrm{e}^{-c})$. Conclude that the top-to-random cycle model exhibits a cut-off at time $n_{\text{mixing}} = N \log N$.

(4) **An elementary proof of the lower bounds on mixing times.** This exercise proposes an alternative proof of Propositions 3.22 and 3.23, which does not rely on representation theory.

(a) By using an inclusion-exclusion argument, show that the number of permutations $\sigma \in \mathfrak{S}(N)$ which have no fixed point is

$$D_N = N! \sum_{r=0}^{N} \frac{(-1)^r}{r!}.$$

(b) Show that if σ is chosen uniformly in $\mathfrak{S}(N)$, then its number of fixed points $\mathrm{Fix}(\sigma)$ satisfies:

$$\mathbb{P}[\mathrm{Fix}(\sigma) = k] = \frac{1}{k!\,\mathrm{e}} + O\left(\frac{1}{k!\,(N-k)!}\right) \to_{N\to\infty} \frac{1}{k!\,\mathrm{e}}.$$

In other words, the number of fixed points converges in law to a standard Poisson distribution.

(c) Deduce from the previous question that for any $k \in [\![0, N]\!]$,

$$\mathbb{P}[\mathrm{Fix}(\sigma) \leq k] = 1 - O\left(\frac{1}{k!} + \frac{2^N}{N!}\right)$$

for a uniform random permutation $\sigma \in \mathfrak{S}(N)$.

(d) For the random transposition model, denote $S_n = \{i_m, j_m,\ 1 \leq m \leq n\}$, where the (i_m, j_m)'s are the random transpositions used to increment the random walk. Show that

$$d_{\mathrm{TV}}(\mu_n, \mathrm{Haar}) \geq \mathbb{P}[\mathrm{Fix}(\sigma) \leq k] - \mathbb{P}[|S_n| \geq N - k].$$

With $n = \frac{N(\log N - c)}{2}$, use this inequality to prove that the total variation distance is close to 1.

(e) Use an analogous argument in order to prove the lower bound for the top-with-random transposition model.

(5) **The problem with 26 questions.** This exercise proposes a representation-theoretic analysis of the cut-off phenomenon for the riffle shuffle model introduced in Example 1.6. We keep the same notations as in Chapter 1; in particular, the generator μ of the random walk is

$$\mu = \frac{1}{2^N} \sum_{i=0}^{N} (12\ldots i) \sqcup ((i+1)(i+2)\ldots N).$$

A *return* of a permutation $\sigma \in \mathfrak{S}(N)$ is an integer $i \in [\![1, N-1]\!]$ such that the letter i appears before the letter $i+1$ in the word of σ. For instance, the set of returns of $\sigma = 861734952$ is $R(\sigma) = \{2, 5, 7\}$. For $r \in [\![0, N-1]\!]$, we denote

$$U_r = \sum_{\substack{\sigma \in \mathfrak{S}(N) \\ \mathrm{card}(R(\sigma)) = r}} \sigma.$$

(a) Show that $\mu = \frac{N+1}{2^N}\, U_0 + \frac{1}{2^N}\, U_1$.

A *descent* (respectively, a *rise*) of a permutation $\sigma \in \mathfrak{S}(N)$ is an integer $i \in [\![1, N-1]\!]$ such that $\sigma(i) > \sigma(i+1)$ (respectively, such that $\sigma(i) < \sigma(i+1)$). For instance, the set of descents of $\sigma = 861734952$ is $D(\sigma) = \{1, 2, 4, 7, 8\}$.

(b) Show that for any permutation σ, $D(\sigma) = R(\sigma^{-1})$.

(c) For $m \in [\![0, N-1]\!]$, the *Eulerian number*$A_{N,m}$ is the number of permutations $\sigma \in \mathfrak{S}(N)$ with m rises. Show that $A_{N,m}$ is also the number of permutations with m descents, or with m returns. Show the symmetry: $A_{N,m} = A_{N,N-1-m}$.

We consider for $m \in [\![0, N-1]\!]$ the following regions of the hypercube $[0,1)^N$:

$$M_{N,m} = \{(x_1, \ldots, x_N) \in [0,1)^N \mid \text{there are } m \text{ indices } i \in [\![1, N-1]\!] \text{ such that } x_i < x_{i+1}\};$$
$$H_{N,m} = \{(x_1, \ldots, x_N) \in [0,1)^N \mid m \leq x_1 + x_2 + \cdots + x_N < m+1\},$$

and the map

$$\Psi : [0,1)^N \to [0,1)^N$$
$$(x_1, \ldots, x_N) \mapsto (y_1, \ldots, y_N), \quad \text{with } y_i = \begin{cases} x_{i-1} - x_i & \text{if } x_{i-1} > x_i, \\ 1 + x_{i-1} - x_i & \text{if } x_{i-1} < x_i, \end{cases}$$

with by convention $x_0 = 0$. We do not define Ψ on the subset with zero Lebesgue measure where some coordinates x_i are equal.

(d) Show that the Lebesgue volume of $M_{N,m}$ is $\frac{A_{N,m}}{N!}$.

(e) Show that Ψ is inversed by the formula $x_i = 1 + \lfloor y_1 + \cdots + y_i \rfloor - y_1 + \cdots + y_i$ (outside a negligible subset of the hypercube). Show that Ψ is a Lebesgue isomorphism of the hypercube, and that $\Psi(M_{N,m}) = H_{N,m}$. Deduce from this observation that:

$$A_{N,m} = N!\,\mathbb{P}[m \leq X_1 + X_2 + \cdots + X_N < m+1],$$

where the X_i's are independent uniform variables on $[0,1)$.

Given $N \geq 1$, $a \geq 2$ and $\sigma \in \mathfrak{S}(N)$, the *$a$-shuffle* of σ is the random permutation $\tau = \overline{\sqcup\!\sqcup}^{(a)}(\sigma) \in \mathfrak{S}(N)$ computed as follows:

- We pick at random N independent and uniform points in $[0,1)$, and we denote $0 \leq x_{\sigma(1)} < x_{\sigma(2)} < \cdots < x_{\sigma(N)} < 1$ their increasing reordering.
- We set $y_i = a x_i \bmod 1$; the increasing reordering $y_{\tau(1)} < y_{\tau(2)} < \cdots < y_{\tau(N)}$ of the y_i's yields the permutation $\tau \in \mathfrak{S}(N)$.

For instance, if $N = 4$, $a = 2$ and $\sigma = 4312$, a possible set of x_i's is $\{0.2, 0.4, 0.55, 0.75\}$, and we set $x_4 = 0.2$, $x_3 = 0.4$, $x_1 = 0.55$ and $x_2 = 0.75$. The y_i's are $y_4 = 0.4$, $y_3 = 0.8$, $y_1 = 0.1$ and $y_2 = 0.5$, and their increasing reodering is $y_1 < y_4 < y_2 < y_3$, so we set in this case $\tau = 1423$.

(f) Show that the sequence $(y_{\tau(1)}, \ldots, y_{\tau(N)})$ used during the construction of $\tau = \overline{\sqcup\!\sqcup}^{(a)}(\sigma)$ is independent from τ, and is uniformly distributed on the simplex

$$\Delta^N[0,1) = \{(x_1, x_2, \ldots, x_N) \,|\, 0 < x_1 < \cdots < x_N < 1\}$$

endowed with the Lebesgue measure.

(g) Show that for any $a \geq 2$ and any $\sigma \in \mathfrak{S}(N)$, the two random variables $\overline{\sqcup\!\sqcup}^{(a)}(\sigma)$ and $\sigma \circ \overline{\sqcup\!\sqcup}^{(a)}(\mathrm{id}_{[\![1,N]\!]})$ have the same distribution.

(h) We denote $(\sigma_n)_{n\in\mathbb{N}}$ the random walk on $\mathfrak{S}(N)$ with generator μ. Show that the transitions of this random walk are given by a 2-shuffle: conditionally to $\{\sigma_n = \sigma\}$, the law of σ_{n+1} is the law of the random permutation $\overline{\sqcup\!\sqcup}^{(2)}(\sigma)$.

(i) We denote V_a the distribution of $\overline{\sqcup\!\sqcup}^{(a\geq 2)}(\mathrm{id}_{[\![1,N]\!]})$, viewed as an element of $\mathbb{C}\mathfrak{S}(N)$. For instance, by the previous question,

$$V_2 = \mu = \frac{1}{2^N}\sum_{k=0}^{N}(12\cdots k) \sqcup\!\sqcup ((k+1)(k+2)\cdots N).$$

Give a similar formula for V_a, for any $a \geq 2$.

(j) Show that for any $a, b \geq 2$ and any $\sigma \in \mathfrak{S}(N)$, we have the identity of distributions:

$$\overline{\sqcup\!\sqcup}^{(a)}(\overline{\sqcup\!\sqcup}^{(b)}(\sigma)) =_{(\mathrm{law})} \overline{\sqcup\!\sqcup}^{(ab)}(\sigma),$$

the a-shuflle and the b-shuffle in the left-hand side being performed independently.

(k) Show that the number of returns of $\overline{\sqcup\!\sqcup}^{(a)}(\mathrm{id}_{[\![1,N]\!]})$ belongs to $[\![0, a-1]\!]$, and that if σ is a permutation with $0 \leq r = \mathrm{card}(R(\sigma)) \leq a-1$, then

$$\mathbb{P}[\overline{\sqcup\!\sqcup}^{(a)}(\mathrm{id}_{[\![1,N]\!]}) = \sigma] = \frac{\binom{N+a-r-1}{N}}{a^N}.$$

Deduce from this formula an expansion of $V_{a\geq 2} \in \mathbb{C}\mathfrak{S}(N)$ as a linear combination of the classes U_r, $r \in [\![0, N-1]\!]$.

(l) Show that the distribution μ_n of σ_n is the following element of $\mathbb{C}\mathfrak{S}(N)$:

$$\mu_n = \sum_{r=0}^{N-1} \frac{\binom{2^n+N-r-1}{N}}{2^{nN}}\, U_r.$$

(m) Show that the space $\mathfrak{B}\mathfrak{D}(N)$ spanned linearly by the $U_{r\in[\![0,N-1]\!]}$'s is a commutative subalgebra of $\mathbb{C}\mathfrak{S}(N)$ (hint: perform a change of basis between the U_r's and the V_a's). This algebra has been introduced by Bayer and Diaconis in order to compute the mixing time of the riffle shuffle.

(n) Show that for any $a \geq 2$,

$$V_a = \frac{1}{N!}\sum_{l=0}^{N-1} a^{-l}\left(\sum_{r=0}^{N-1} e_l(N-1-r, N-2-r, \ldots, 1-r, -r)\, U_r\right).$$

For $l \in [\![0, N-1]\!]$, we set $E_l = \sum_{r=0}^{N-1} e_l(N-1-r, N-2-r, \ldots, 1-r, -r)\, U_r$. Show that $\mathbb{C}[V_2] = \mathrm{Span}(1, V_2, (V_2)^2, (V_2)^3, \ldots)$ contains all the E_l's, and is equal to $\mathfrak{B}\mathfrak{D}(N)$.

(o) Show that $(E_l)_{l\in[\![0,N-1]\!]}$ is a linear basis of $\mathfrak{B}\mathfrak{D}(N)$, and that for $l, m \in [\![0, N-1]\!]$, we have $E_l\, E_m = 1_{(l=m)}\, N!\, E_l$ (hint: consider the matrix of the linear map $x \in \mathfrak{B}\mathfrak{D}(N) \mapsto P(V_2)\cdot x$, P being a polynomial).

(p) Prove that the map

$$\mathfrak{B}\mathfrak{D}(N) \to \mathbb{C}^N$$

$$x = \frac{1}{N!}\sum_{l=0}^{N-1} x_l\, E_l \mapsto (x_0, x_1, \ldots, x_{N-1})$$

is an isomorphism of algebras.

(q) Show that

$$d_{\mathrm{TV}}(\sigma_n, \mathrm{Haar}) = \frac{1}{2} \sum_{m=0}^{N-1} \frac{A_{N,m}}{N!} \left| \frac{(2^n + N - 1 - m)!}{(2^n - 1 - m)!\, 2^{nN}} - 1 \right|.$$

(r) We set $n = \frac{\log(N^{3/2}c)}{\log 2}$, and we make the change of variables $m = \frac{N}{2} + \sqrt{\frac{N}{12}}\, x$; the parameters $c \in \mathbb{R}_+^*$ and $x \in \mathbb{R}$ are chosen so that n and m are integers. Estimate a sum

$$\sum_{m=0}^{N-1} \frac{A_{N,m}}{N!}\, f(x)$$

by using the central limit theorem, f being a bounded Lipschitz function. Show on the other hand that

$$\frac{(2^n + N - 1 - m)!}{(2^n - 1 - m)!\, 2^{nN}} = \exp\left(-\frac{x}{c\sqrt{12}} - \frac{1}{24c^2} + O_x\left(\frac{1}{\sqrt{N}}\right)\right).$$

We admit that we can use the estimate from the central limit theorem even though the function above is not bounded Lipschitz. Show then that

$$\begin{aligned}\lim_{N \to +\infty} d_{\mathrm{TV}}(\sigma_n, \mathrm{Haar}) &= \frac{1}{2} \int_{\mathbb{R}} \left| \mathrm{e}^{-\frac{x}{c\sqrt{12}} - \frac{1}{24c^2}} - 1 \right| \frac{\mathrm{e}^{-\frac{x^2}{2}}}{\sqrt{2\pi}}\, dx \\ &= 1 - 2F\left(-\frac{1}{4\sqrt{3}c}\right),\end{aligned}$$

where F is the cumulative distribution function of the standard normal distribution.

(s) Establish a cut-off for the riffle shuffle model, at time $n_{\mathrm{mixing}} = \frac{3 \log N}{2 \log 2}$.

The image of the Bayer–Diaconis algebra $\mathfrak{BD}(N)$ by the anti-isomorphism $\sigma \mapsto \sigma^{-1}$ is included in another interesting subalgebra of $\mathbb{C}\mathfrak{S}(N)$ called the descent subalgebra; the next questions detail the construction of this object. For $i \in [\![1, N-1]\!]$, we denote $s_i = (i, i+1)$ the elementary transposition which exchanges i and $i+1$; the $s_{i \in [\![1,N-1]\!]}$ form a generator subset of the group $\mathfrak{S}(N)$.

(t) If $\sigma \in \mathfrak{S}(N)$ and s_i is fixed, show that $\sigma^{-1} s_i \sigma \in \{s_1, s_2, \ldots, s_{N-1}\}$ if and only if $D(\sigma) \neq D(s_i\sigma)$.

(u) We fix integers $i_1, \ldots, i_k \in [\![1, N-1]\!]$, and two permutations τ and $s_{i_k} s_{i_{k-1}} \cdots s_{i_1} \tau$ with the same descent set. Show the following alternative:

- either $D(\tau) = D(s_{i_{k-1}} \cdots s_{i_1} \tau)$;
- or, there exists $j < k$ such that $s_{i_k} \cdots s_{i_{j+1}} = s_{i_{k-1}} \cdots s_{i_j}$.

Indication: if the first hypothesis is not satisfied, consider a which is for instance not in $D(s_{i_k} \cdots s_{i_1}\tau) = D(\tau)$ and which is in $D(s_{i_{k-1}} \cdots s_{i_1}\tau)$ (the other case is symmetric), and the first integer $j \geq 1$ such that $a \in D(s_{i_j} \cdots s_{i_1}\tau)$.

(v) Consider $\sigma, \tau \in \mathfrak{S}(N)$ two permutations with the same descent set. If $s_{i_k} s_{i_{k-1}} \cdots s_{i_1} = \tau\sigma^{-1}$ is a decomposition with minimal length k, show that the permutations $s_{i_1}\sigma$, $s_{i_2}s_{i_1}\sigma$, *etc.*, $s_{i_k}s_{i_{k-1}} \cdots s_{i_1}\sigma = \tau$ have the same descent set (hint: prove it by induction on k). This implies that the descent classes $D_A = \{\sigma \in \mathfrak{S}(N),\ D(\sigma) = A\}$ with $A \subset [\![1, N-1]\!]$ are connex subsets of the left Cayley graph of $(\mathfrak{S}(N), \{s_1, \ldots, s_{N-1}\})$.

Given $A, B \subset [\![1, N-1]\!]$ et $\rho \in \mathfrak{S}(N)$, we denote $N_{A,B,\rho}$ the set of pairs (σ, τ) such that $\sigma\tau = \rho$, $D(\sigma) = A$ and $D(\tau) = B$. We fix in the following an elementary transposition s_i such that $D(\rho) = D(s_i\rho)$.

(w) For $(\sigma, \tau) \in N_{A,B,\rho}$, prove the following alternative:
- either $D(s_i\sigma) = D(\sigma)$,
- or, $s_i\sigma = \sigma s_j$ for some s_j, which satisfies then $D(s_j\tau) = D(\tau)$.

(x) We set $N_{A,B,\rho} = N_0 \sqcup \bigsqcup_{j=1}^{N-1} N_j$, with

$$N_0 = \{(\sigma, \tau) \in N_{A,B,\rho} \,|\, \sigma^{-1}s_i\sigma \notin \{s_1, \ldots, s_{N-1}\}\};$$
$$N_j = \{(\sigma, \tau) \in N_{A,B,\rho} \,|\, \sigma^{-1}s_i\sigma = s_j\}.$$

For $j \in [\![0, N-1]\!]$, we also set:

$$N_0^* = \{(s_i\sigma, \tau) \,|\, (\sigma, \tau) \in N_0\};$$
$$N_j^* = \{(\sigma, s_j\tau) \,|\, (\sigma, \tau) \in N_j\}.$$

Show that $N_{A,B,s_i\rho} = N_0^* \sqcup \bigsqcup_{j=1}^{N-1} N_j^*$.

(y) Deduce from the previous questions the following fact: if ρ and ρ' have the same descent set, then $\mathrm{card}(N_{A,B,\rho}) = \mathrm{card}(N_{A,B,\rho'})$ for any $A, B \subset [\![1, N-1]\!]$.

(z) For $A \subset [\![1, N-1]\!]$, set

$$D_A = \sum_{\substack{\sigma \in \mathfrak{S}(N) \\ D(\sigma) = A}} \sigma.$$

Show that the subspace of $\mathbb{C}\mathfrak{S}(N)$ spanned linearly by the 2^{N-1} descent classes D_A is a (unitary) subalgebra. This is the *Solomon descent algebra*.

CHAPTER 4

Plancherel measures and their asymptotics

In Chapter 2, we observed that the Plancherel formula 1.17 yields in the setting of symmetric groups the following combinatorial formula:

$$n! = \sum_{\lambda \in \mathfrak{Y}(n)} (\dim \lambda)^2 = \sum_{\lambda \in \mathfrak{Y}(n)} (\mathrm{card}\, \mathrm{ST}(\lambda))^2.$$

We shall now see that this formula can be proved directly, by exhibiting a bijection between the set of permutations of size n and the set of pairs of standard tableaux (P, Q) with the same shape $\lambda \in \mathfrak{Y}(n)$. On the other hand, the formula above yields a natural probability measure on integer partitions of size n:

Definition 4.1 (Plancherel measure). *Let G be a finite group. The Plancherel measure of G is the probability measure PL_G on $\widehat{G}$ defined by:*

$$\mathrm{PL}_G(\lambda) = \frac{(\dim \lambda)^2}{\mathrm{card}\, G}.$$

If $G = \mathfrak{S}(n)$, this Plancherel measure is the probability measure on $\mathfrak{Y}(n)$ given by

$$\mathrm{PL}_n(\lambda) = \frac{(\mathrm{card}\, \mathrm{ST}(\lambda))^2}{n!}.$$

Several computations of the previous chapters involve sums which can be rewritten as integrals against the Plancherel measure of $\mathfrak{S}(n)$. We shall see that the Young diagram of a random integer partition $\lambda \in \mathfrak{Y}(n)$ chosen according to the Plancherel measure has a typical shape. This observation leads to the *Logan–Shepp–Kerov–Vershik law of large numbers* (Theorem 4.28), and to a solution of Ulam's problem of the longest increasing subsequence of a random permutation.

1. The Robinson–Schensted correspondence

The *Robinson–Schensted correspondence* (later generalized by Knuth) associates to any permutation of size n a pair of standard tableaux (P, Q) with the same shape $\lambda \in \mathfrak{Y}(n)$. It is actually a special case of an algorithm which can be applied to any *word*, that is a sequence of integers possibly with repetitions. In the sequel, we call *tableau* with shape $\lambda \in \mathfrak{Y}(n)$ an arbitrary numbering of the cells of the Young diagram of λ by numbers in $\mathbb{N}^*$ (possibly with repetitions, and without any condition of order). We say that a tableau is *semi-standard* if it is *strictly* increasing along the columns and *weakly* increasing along the rows. The standard tableaux correspond to the particular case where one has also strictly increasing rows, and where the numbers are chosen in $[\![1, n]\!]$, with $n = |\lambda|$.

Consider now an alphabet $A = [\![1, N]\!]$, and a word $w = w_1 w_2 \cdots w_n$ with letters in A. We construct inductively two tableaux $P(w)$ and $Q(w)$ according to the following rules.

(1) We start with $P_0 = Q_0 = \emptyset$, and we read the letters of the word w from left to right. At the i-th step of the algorithm, P_i will be a semi-standard tableau whose entries are the letters $w_1, w_2, \ldots, w_i$, and Q_i will be a standard tableau of size i.

(2) Suppose that P_{i-1} and Q_{i-1} are already constructed. To obtain P_i from P_{i-1}, we insert w_i in P_{i-1} as follows:

- If w_i is larger than all the entries of the first row of P_{i-1}, we put it at the end of this first row.
- Otherwise, there exists a first entry l in the first row of P_{i-1} which is strictly larger than w_i. We replace l by w_i in this first row, and we bump l to the next row, trying to insert it according to the same rules. Namely, if l is larger than all the entries of the second row, we put it at the end of this second row, and otherwise, we replace $l' > l$ by l and we bump l' to the next row.
- One might need to bump several letters to higher rows, but the process ends after a finite number of steps (possibly, by adding a new row to P_{i-1} if a letter of the last row of P_{i-1} has to be bumped to an higher row).

The shape of the new tableau P_i differs from P_{i-1} by one cell on the right-top border. We construct Q_i from Q_{i-1} by adding the same cell and numbering it by i. By construction, Q_i is a standard tableau of size i for any $i \in [\![0, n]\!]$; we shall prove in a moment that P_i is semi-standard for any $i \in [\![0, n]\!]$.

(3) We set $P(w) = P_n$ and $Q(w) = Q_n$. Thus, $P(w)$ is a tableau whose entries are the letters of w, and $Q(w)$ is a standard tableau of size n.

This is better understood on a example. Suppose that $w = 174723521846$. The two first pairs of tableaux obtained by the RS algorithm are:

$$P_1 = \begin{ytableau} 1 \end{ytableau}\,,\ Q_1 = \begin{ytableau} 1 \end{ytableau} \qquad ; \qquad P_2 = \begin{ytableau} 1 & 7 \end{ytableau}\,,\ Q_2 = \begin{ytableau} 1 & 2 \end{ytableau}\,.$$

Indeed, since $7 \geq 1$, we can insert it at the end of the first row when going from P_1 to P_2; and the tableaux Q_1 and Q_2 record the growing of the shape of the tableaux P_1 and P_2. Now, when inserting $w_3 = 4$ in P_2, we have to replace 7 by 4 in the first row, and to bump 7 to a second row:

$$P_3 = \begin{ytableau} 7 \\ 1 & 4 \end{ytableau}\,,\ Q_3 = \begin{ytableau} 3 \\ 1 & 2 \end{ytableau}\,.$$

The next tableaux P_i are the following:

$$P_4 = \begin{ytableau} 7 \\ 1 & 4 & 7 \end{ytableau} \quad ; \quad P_5 = \begin{ytableau} 7 \\ 4 \\ 1 & 2 & 7 \end{ytableau} \quad ; \quad P_6 = \begin{ytableau} 7 \\ 4 & 7 \\ 1 & 2 & 3 \end{ytableau} \quad ; \quad P_7 = \begin{ytableau} 7 \\ 4 & 7 \\ 1 & 2 & 3 & 5 \end{ytableau} \quad ;$$

$$P_8 = \begin{ytableau} 7 \\ 4 \\ 3 & 7 \\ 1 & 2 & 2 & 5 \end{ytableau} \quad ; \quad P_9 = \begin{ytableau} 7 \\ 4 \\ 3 \\ 2 & 7 \\ 1 & 1 & 2 & 5 \end{ytableau} \quad ; \quad P_{10} = \begin{ytableau} 7 \\ 4 \\ 3 \\ 2 & 7 \\ 1 & 1 & 2 & 5 & 8 \end{ytableau} \quad ;$$

$$P_{11} = \begin{ytableau} 7 \\ 4 \\ 3 & 7 \\ 2 & 5 \\ 1 & 1 & 2 & 4 & 8 \end{ytableau} \quad ; \quad P_{12} = \begin{ytableau} 7 \\ 4 \\ 3 & 7 \\ 2 & 5 & 8 \\ 1 & 1 & 2 & 4 & 6 \end{ytableau}\,.$$

The last tableau is $P_{12} = P(w)$. Since $Q(w)$ records the growing of the shape of the tableaux P_i, we have

$$Q(w) = \begin{array}{|c|c|c|c|c|}
\cline{1-1} 9 \\
\cline{1-1} 8 \\
\cline{1-2} 5 & 11 \\
\cline{1-3} 3 & 6 & 12 \\
\hline 1 & 2 & 4 & 7 & 10 \\
\hline
\end{array}\ .$$

Theorem 4.2 (Robinson–Schensted). *The Robinson–Schensted algorithm yields a bijection between the set of words with length n and letters in $A = [\![1, N]\!]$, and the set of pairs (P, Q) of tableaux with size n and same shape $\lambda \in \mathfrak{Y}(n)$, P being semi-standard with entries in A and Q being standard. This bijection is such that the entries of $P(w)$ are the letters of w, counting multiplicities.*

The restriction of the algorithm to the set of words of the permutations in $\mathfrak{S}(n)$ yields a bijection between the symmetric group of order n, and the set of pairs (P, Q) of standard tableaux with size n and same shape $\lambda \in \mathfrak{Y}(n)$.

Proof. We first prove that the insertion of a letter l in a semi-standard tableau P yields a tableau P' which is still semi-standard. By construction, the letter l is either inserted at the end of the first row if it is larger than all the entries of the first row of P; or, it replaces the first letter m which is strictly larger than l in this first row. In these two cases, it is clear than the new first row of P' is still weakly increasing. Let us now check that if a letter m from the $(i-1)$-th row of P is bumped to the i-th row, then the following relations in P' still hold:

$$\begin{array}{lccccc}
i\text{-th row} & \boxed{b} & \leq & \boxed{m} & \leq & \boxed{c} \\
 & & & \vee & & \\
(i-1)\text{-th row} & & & \boxed{a} & &
\end{array}$$

The relations $b \leq m \leq c$ follow from the same argument as before; actually, if c exists (that is, if we do not place m at the end of a row of P'), then by construction $m < c$. Therefore, the only difficult thing to check is that $a < m$. If we had $a \geq m$, then m would be in P a letter of the $(i-1)$-th row which came before a, and which needed to be bumped during the construction of P'. But in this case, we have the following configuration in P:

$$\begin{array}{lccc}
i\text{-th row} & \boxed{d} & & \\
 & \vee & & \\
(i-1)\text{-th row} & \boxed{m} & \cdots & \boxed{a}
\end{array}$$

When bumping m from the $(i-1)$-th to the i-th row, we can insert m before d, because $m < d$. In particular, we cannot insert it on top of a; whence a contradiction. So, P' is still a semi-standard tableau.

The remainder of the theorem comes from the following observation: given the two tableaux $P(w)$ and $Q(w)$ with P semi-standard and Q standard, we can recursively uninsert the letters $w_n, w_{n-1}, \ldots, w_1$ in order to reconstruct the word w. For instance, let us consider the previous example and explain how to go from the pair (P_{12}, Q_{12}) to the pair (P_{11}, Q_{11}) and the letter $w_{12} = 6$. We observe that 12 is on the second row of Q_{12}, so the corresponding entry 8 in P_{12} was in the first row in P_{11} and was bumped by the insertion of a letter, namely, the largest letter strictly smaller than 8 in the first row of P_{12}; this is $w_{12} = 6$. So, P_{11} is obtained from P_{12} by moving the entry 8 from the end of the second row to the position occupied by 6 in the first row, and by removing this 6; whereas Q_{11} is obtained from Q_{12} by deleting the entry 12. This algorithm is general: each time, to go from Q_i to Q_{i-1}, we delete the entry i, and to go from P_i to P_{i-1}, we slide the *bumping* *route* starting from the entry corresponding to i in Q_i; this bumping route ends on the first row of P_i, and the corresponding entry is w_i, which we remove from P_i to obtain P_{i-1}.

Finally, for the statement on permutations, it suffice to observe that if $w = \sigma$ is a permutation word, then its tableau $P(\sigma)$ is semi-standard with n distinct entries in $[\![1, n]\!]$ (the letters of w): this is equivalent to being a standard tableau. □

Given a word w, we denote $\lambda(w)$ the common shape of the two tableaux $P(w)$ and $Q(w)$ associated to w. The parts of this integer partition can be related to the lengths of the longest increasing subwords of w:

Proposition 4.3 (Longest increasing subword). *Given a word $w = w_1 w_2 \cdots w_n$, the first part $\lambda_1(w)$ is equal to the maximal length $\ell(w)$ of a weakly increasing subword $w_{i_1} w_{i_2} \cdots w_{i_l}$ with*

$$1 \leq i_1 < i_2 < \cdots < i_l \leq n \quad \text{and} \quad w_{i_1} \leq w_{i_2} \leq \cdots \leq w_{i_l}.$$

Proof. We reproduce here the original proof of this statement, due to Schensted. In the sequel, if a letter l appears several time in a word w, we index its consecutive occurrences $l_1, l_2, \ldots, l_s$, so for instance $w = 174723521846$ becomes $\underline{w} = 1_1 7_1 4_1 7_2 2_1 3 5 2_2 1_2 8 4_2 6$. We can apply the Robinson–Schensted algorithm to this modified word $\underline{w}$, and forget at the end the indices in order to recover the tableaux $P(w)$ and $Q(w)$.

Call *basic j-th subword* of w the list $\underline{B}_j(w)$ of the indiced letters of $\underline{w}$ which are inserted in the j-th position of the first row of the P-tableaux during the Robinson–Schensted algorithm (these letters can then be bumped to higher rows). For instance, the basic subwords of $w = 174723521846$ are:

$$\underline{B}_1(w) = 1_1 \quad ; \quad \underline{B}_2(w) = 7_1 4_1 2_1 1_2 \quad ; \quad \underline{B}_3(w) = 7_2 3 2_2$$
$$\underline{B}_4(w) = 5 4_2 \quad ; \quad \underline{B}_5(w) = 86.$$

By construction, the words $\underline{B}_j(w)$ form a partition of $\underline{w}$ into disjoint subwords. Moreover, the unindiced version $B_j(w)$ of each basic j-th subword $\underline{B}_j(w)$ is strictly decreasing: each time one inserts a new letter l in the j-th position of the first row of the P-tableau, it bumps the previous entry m of $B_j(w)$ to a higher row, so $m > l$. Now, observe that if $m = \underline{w}_b$ appears in $\underline{B}_j(w)$ with $j \geq 2$, then one can find $l = \underline{w}_a$ in $\underline{B}_{j-1}(w)$ such that $l \leq m$ and $a < b$. Indeed, when m is inserted in $P_b(\underline{w})$, it is at the j-th position of the first row, so it is larger than the entry l of the $(j-1)$-th position of the first row of $P_b(\underline{w})$. This entry l belongs to $\underline{B}_{j-1}(w)$ and is a letter $\underline{w}_a$ with $a < b$.

If $z = w_{i_1} \ldots w_{i_l}$ is a weakly increasing subword of w, then its version $\underline{z}$ with indiced letters can intersect each basic subword $\underline{B}_j(w)$ only once, because these basic subwords are strictly decreasing. Therefore, l is smaller than the number of basic subwords, which is $\lambda_1(w)$. Moreover, by taking an entry of $\underline{B}_{\lambda_1(w)}(w)$ and finding smaller entries in $\underline{B}_{\lambda_1(w)-1}(w)$, $\underline{B}_{\lambda_1(w)-2}(w)$, *etc.*, we can indeed construct a weakly increasing subword of $\underline{w}$ with length $l = \lambda_1(w)$. So, we have indeed proved the identity $\ell(w) = \lambda_1(w)$. □

Remark 4.4. The proof is easier to understand in the special case where all the letters of w are distinct: we then do not need to introduce labels for the occurrences of the letters. This argument is required in the general case, because we want the basic subwords $\underline{B}_j(w)$ to form a partition of w in subwords.

Remark 4.5 (Greene invariants). A generalisation of Proposition 4.3 due to Greene states that, for any $k \in [\![1, \ell(\lambda)]\!]$, $\lambda_1(w) + \lambda_2(w) + \cdots + \lambda_k(w)$ is the maximal sum of the lengths of k disjoint weakly increasing subwords of w. Let us highlight a small subtlety in this statement: for instance with $k = 2$, $\lambda_1(w) + \lambda_2(w)$ is the maximal sum of the lengths $l \geq m$ of two disjoint weakly increasing subwords of w, but we do not have necessarily $l = \lambda_1(w)$ and $m = \lambda_2(w)$. The proof of this generalisation is much harder than the previous proof, and it is related to the theory of the *plactic monoid*, which leads to other surprising results detailed in the two next remarks.

Remark 4.6 (Robinson–Schensted correspondence and inversion of permutations). If σ is a permutation, then we have the following surprising relation between its P and Q tableaux (this is due to Schützenberger):

$$P(\sigma) = Q(\sigma^{-1}).$$

Remark 4.7 (Jeu de taquin). Given a word w, there is a way to compute $P(w)$ which relies on sliding numbered cells as in the *jeu de taquin*. We associate to w the *ribbon* $R(w)$, which is the diagram obtained by reading the word w from left to right, and, when reading two consecutive letters a and b:

- by adding a cell $\boxed{b}$ to the right of the cell $\boxed{a}$ if $a \leq b$;
- by adding a cell $\boxed{b}$ below the cell $\boxed{a}$ if $a > b$ (descent).

For instance, the ribbon associated to the word $w = 174723521846$ is

$$\begin{array}{ccccccc}
1 & 7 & & & & & \\
 & 4 & 7 & & & & \\
 & & 2 & 3 & 5 & & \\
 & & & & 2 & & \\
 & & & & 1 & 8 & \\
 & & & & & 4 & 6
\end{array}.$$

Starting from this ribbon, let us slide its numbered cells towards the bottom left corner, while still keeping the rows weakly increasing and the columns strictly increasing. For instance, one can slide the left-most cell $\boxed{1}$ one step down, and then the cell $\boxed{7}$ of the first row to the left, thereby obtaining:

$$\begin{array}{ccccccc}
7 & & & & & & \\
1 & 4 & 7 & & & & \\
 & & 2 & 3 & 5 & & \\
 & & & & 2 & & \\
 & & & & 1 & 8 & \\
 & & & & & 4 & 6
\end{array}.$$

If one completes the bottom left border of the ribbon $R(w)$ by inner empty cells so as to have a Young diagram, then the two sliding operations described above have reduced by one the number of inner empty cells. We continue the sliding of cells towards the left or the bottom until all the inner empty cells have disappeared; the order chosen for the slidings is arbitrary. Let us give a few intermediary steps of this algorithm:

$$\begin{array}{ccccccc}
7 & & & & & & \\
1 & 4 & 7 & & & & \\
 & & 2 & 3 & 5 & & \\
 & & & & 2 & 8 & \\
 & & & & 1 & 4 & 6
\end{array}\ ;\quad
\begin{array}{cccccc}
7 & & & & & \\
1 & 4 & 7 & & & \\
 & 2 & 3 & 5 & & \\
 & & & 2 & 8 & \\
 & & & 1 & 4 & 6
\end{array}\ ;\quad
\begin{array}{cccccc}
7 & & & & & \\
4 & 7 & & & & \\
1 & 2 & 3 & 5 & & \\
 & & & 2 & 8 & \\
 & & & 1 & 4 & 6
\end{array}\ ;\quad
\begin{array}{ccccc}
7 & & & & \\
4 & 7 & & & \\
1 & 2 & 3 & & \\
 & & 2 & 5 & 8 \\
 & & 1 & 4 & 6
\end{array}$$

$$\begin{array}{ccccc}
7 & & & & \\
4 & 7 & & & \\
1 & 3 & & & \\
 & 2 & 2 & 5 & 8 \\
 & & 1 & 4 & 6
\end{array}\ ;\quad
\begin{array}{ccccc}
7 & & & & \\
4 & 7 & & & \\
1 & 3 & & & \\
 & 2 & 5 & 8 & \\
 & 1 & 2 & 4 & 6
\end{array}\ ;\quad
\begin{array}{ccccc}
7 & & & & \\
4 & & & & \\
3 & 7 & & & \\
1 & 2 & 5 & 8 & \\
 & 1 & 2 & 4 & 6
\end{array}\ ;\quad
\begin{array}{ccccc}
7 & & & & \\
4 & & & & \\
3 & 7 & & & \\
2 & 5 & 8 & & \\
1 & 1 & 2 & 4 & 6
\end{array}.$$

We recognise at the end the tableau $P(w)$, and it turns out that this is a general phenomenon: for any choice of slidings, starting from the ribbon $R(w)$ and deleting the inner empty cells, one always obtain in the end $P(w)$. This deep result is due to Lascoux and Schützenberger.

Proposition 4.3 is the starting point of the solution of an important combinatorial problem studied during the second part of the twentieth century. Consider a random permutation σ_n chosen uniformly in $\mathfrak{S}(n)$: $\mathbb{P}[\sigma_n = \sigma] = \frac{1}{n!}$ for any $\sigma \in \mathfrak{S}(n)$. We are interested in the random variable:

$$\ell_n = \ell(\sigma_n) = \text{length of the longest increasing subsequence of a random permutation.}$$

We shall prove that $\frac{\ell_n}{2\sqrt{n}}$ converges in probability to the constant 1, and in Chapter 5, we shall also study the rescaled fluctuations

$$X_n = n^{\frac{1}{3}} \left(\frac{\ell_n}{2\sqrt{n}} - 1 \right).$$

PROPOSITION 4.8. *The law of the random variable ℓ_n is the law of the first part of an integer partition λ chosen according to the Plancherel measure PL_n on $\mathfrak{Y}(n)$.*

PROOF. We have

$$\mathbb{P}[\ell_n = k] = \mathbb{P}[\lambda_1(\sigma_n) = k] = \frac{1}{n!} \sum_{\substack{(P,Q) \text{ standard tableaux} \\ \text{with same shape } \lambda \text{ in } \mathfrak{Y}(n)}} 1_{(\lambda_1 = k)} = \frac{1}{n!} \sum_{\substack{\lambda \in \mathfrak{Y}(n) \\ \lambda_1 = k}} (\mathrm{card}\, \mathrm{ST}(\lambda))^2 = \mathrm{PL}_n[\lambda_1 = k]$$

by using Proposition 4.3 for the first identity, the Robinson–Schensted correspondence for the second identity. □

As a consequence, if we can describe the shape of a large random Young diagram chosen according to the Plancherel measure, then we shall be able to solve the problem of the longest increasing subsequence of a random permutation. These limiting results will be stated at the end of this chapter, see Theorem 4.28 and Corollary 4.29. As a first consequence of the previous observation, let us prove the following non trivial estimate:

PROPOSITION 4.9 (Upper bound on the expectation of the longest increasing subsequence). *For any $n \geq 1$, $\mathbb{E}[\ell_n] \leq 2\sqrt{n}$.*

PROOF. It suffices to prove that $\mathbb{E}[\ell_n - \ell_{n-1}] \leq \frac{1}{\sqrt{n}}$, because this entails:

$$\mathbb{E}[\ell_n] \leq \sum_{k=1}^{n} \frac{1}{\sqrt{k}} \leq \int_0^n \frac{1}{\sqrt{x}}\, dx = 2\sqrt{n}.$$

Consider a random permutation σ_n uniformly chosen in $\mathfrak{S}(n)$. By the previous discussion, $\lambda_1(\sigma_n)$ has the law of ℓ_n. Moreover, it is easy to see that conditionally to the value $\sigma_n(n)$, the subword $\sigma_{n,n-1}$ of σ_n which consists in its $n-1$ first letters is uniformly distributed among the permutations of the set $[\![1, n]\!] \setminus \{\sigma_n(n)\}$. As a consequence:

$$\lambda_1(\sigma_{n,n-1}) =_{\text{law}} \ell_{n-1}.$$

Indeed, this is true conditionally to the value of $\sigma_n(n)$, and we obtain the same law for every possible value. We conclude that

$$\mathbb{E}[\ell_n - \ell_{n-1}] = \mathbb{E}[\lambda_1(\sigma_n) - \lambda_1(\sigma_{n,n-1})].$$

However, we increase the length of the first row when going from $\sigma_{n,n-1}$ to σ_n if and only if n is placed on the first row of the recording tableau $Q(\sigma_n)$. So,

$$\mathbb{E}[\lambda_1(\sigma_n) - \lambda_1(\sigma_{n,n-1})] = \mathbb{P}[n \text{ is in the first row of } Q(\sigma_n)].$$

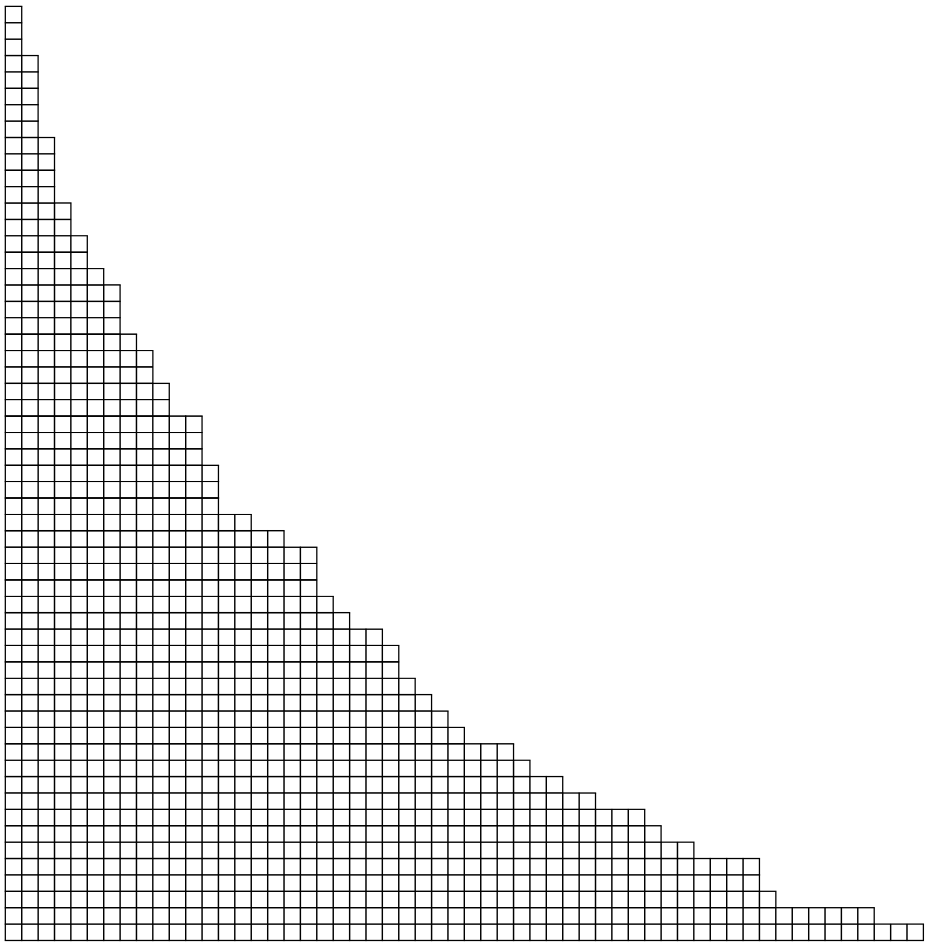
We now remark that n is at the end of the first rows of both tableaux $P(\sigma)$ and $Q(\sigma)$ if and only if $\sigma(n) = n$; this happens with probability $\frac{1}{n}$ if $\sigma = \sigma_n$ is chosen uniformly in $\mathfrak{S}(n)$. Then:

$$\begin{aligned}\frac{1}{n} &= \mathbb{P}[n \text{ is in the first rows of } P(\sigma_n) \text{ and } Q(\sigma_n)] \\ &= \sum_{\lambda \in \mathfrak{Y}(n)} \mathrm{PL}_n[\lambda]\, \mathbb{P}[n \text{ is in the first rows of } P(\sigma_n) \text{ and } Q(\sigma_n) \,|\, \lambda(\sigma_n) = \lambda] \\ &= \sum_{\lambda \in \mathfrak{Y}(n)} \mathrm{PL}_n[\lambda]\, \left(\mathbb{P}[n \text{ is in the first row of } Q(\sigma_n) \,|\, \lambda(\sigma_n) = \lambda]\right)^2.\end{aligned}$$

Indeed, conditionnally to the shape $\lambda(\sigma_n)$, the two tableaux $P(\sigma_n)$ and $Q(\sigma_n)$ are independent, and they are both uniformly distributed in $\mathrm{ST}(\lambda(\sigma_n))$. Finally, by the Cauchy–Schwarz inequality, the last line of the equations above is greater than the square of

$$\sum_{\lambda \in \mathfrak{Y}(n)} \mathrm{PL}_n[\lambda]\, \mathbb{P}[n \text{ is in the first row of } Q(\sigma_n) \,|\, \lambda(\sigma_n) = \lambda] = \mathbb{P}[n \text{ is in the first row of } Q(\sigma_n)].$$

This ends the proof. $\square$



2. Asymptotics of the random character values

We have drawn above the Young diagram of a random integer partition with size $n = 1000$ chosen according to the Plancherel measure PL_{1000}. Although the result is random, it seems that the upper-right boundary of this Young diagram stays close to some smooth curve if one rescales the cells in an appropriate way. In order to prove this *geometric* law of large numbers, we shall first study *algebraic* observables of the random partitions $\lambda_n \sim \mathrm{PL}_n$, which are related to the associated irreducible characters of $\mathfrak{S}(n)$.

DEFINITION 4.10 (Rescaled character value). *Fix an integer partition* $\mu = (\mu_1, \ldots, \mu_\ell)$ *with size* k*, and denote*

$$\sigma_\mu = (1, 2, \ldots, \mu_1)(\mu_1 + 1, \ldots, \mu_1 + \mu_2) \cdots (\mu_1 + \cdots + \mu_{\ell-1} + 1, \ldots, \mu_1 + \cdots + \mu_\ell);$$

this is a permutation in $\mathfrak{S}(k)$ *with cycle-type* μ*. If* $n \geq k$*, we can consider* σ_μ *as an element of* $\mathfrak{S}(n)$ *thanks to the natural inclusion* $\mathfrak{S}(k) \subset \mathfrak{S}(n)$ *coming from* $[\![1, k]\!] \subset [\![1, n]\!]$*. The rescaled character value* Σ_μ *is the function* $\mathfrak{Y} \to \mathbb{R}$ *defined by:*

$$\Sigma_\mu(\lambda) = \begin{cases} n^{\downarrow k}\, \frac{\mathrm{ch}^\lambda(\sigma_\mu)}{\dim \lambda} & \text{if } |\lambda| = n \geq k = |\mu|, \\ 0 & \text{if } |\lambda| = n < k = |\mu|. \end{cases}$$

The notation $n^{\downarrow k}$ *is the falling factorial* $n(n-1)(n-2)\cdots(n-k+1)$.

The functions Σ_μ are observables of the integer partition λ, and given $\lambda = \lambda_n \sim \mathrm{PL}_n$, they yield a family of random variables whose joint distribution will enable us to understand the typical geometry of the Young diagram of λ_n. The method that we shall use in the sequel works in theory for any probability measure on $\mathfrak{Y}(n)$ which is the *spectral measure* of a representation.

DEFINITION 4.11 (Spectral measure). *Let* G *be a finite group, and* V *be a non-zero representation of* G*. The spectral measure of the representation* V *is the probability measure* $\mathbb{P}_V$ *on* $\widehat{G}$ *defined by*

$$\mathbb{P}_V[\lambda] = \frac{m_\lambda\,(\dim \lambda)}{\dim V},$$

where the m_λ*'s are the multiplicities of the irreducible representations of* G *as components of* V*:* $V = \bigoplus_{\lambda \in \widehat{G}} m_\lambda\, V^\lambda$.

PROPOSITION 4.12. *The Plancherel measure of a finite group* G *is the spectral measure of its regular representation on* $\mathbb{C}G$*. On the other hand, for any representation* V *of* G *and any element* $g \in G$*, if we consider the normalised character*

$$\chi^\lambda(g) = \frac{\mathrm{ch}^\lambda(g)}{\dim \lambda}$$

as a random variable with $\lambda \sim \mathbb{P}_V$*, then*

$$\mathbb{E}_V[\chi^\lambda(g)] = \frac{\mathrm{ch}^V(g)}{\dim V} = \chi^V(g)$$

is the value of the normalised character of the representation V *on the element* g.

PROOF. We have seen in Chapter 1 that the decomposition of $\mathbb{C}G$ in irreducible representations is $\mathbb{C}G = \bigoplus_{\lambda \in \widehat{G}} (\dim \lambda) V^\lambda$; *cf.* the proof of Theorem 1.15. Consequently,

$$\mathbb{P}_{\mathbb{C}G}[V] = \frac{(\dim \lambda)^2}{\dim \mathbb{C}G} = \frac{(\dim \lambda)^2}{\mathrm{card}\, G} = \mathrm{PL}_G[V].$$

This proves the first part of the proposition, and the second part is an easy computation: if $V = \bigoplus_{\lambda \in \widehat{G}} m_\lambda\, V^\lambda$, then

$$\mathbb{E}_V[\chi^\lambda(g)] = \sum_{\lambda \in \widehat{G}} \mathbb{P}_V[\lambda]\, \chi^\lambda(g) = \frac{1}{\dim V} \sum_{\lambda \in \widehat{G}} m_\lambda\, \mathrm{ch}^\lambda(g) = \frac{\mathrm{ch}^V(g)}{\dim V}. \qquad \square$$

COROLLARY 4.13. *If* $\mu \in \mathfrak{Y}(k)$*, then for any* $n \geq k$*, the random partition* λ_n *chosen according to the Plancherel measure* PL_n *of* $\mathfrak{S}(n)$ *satisfies:*

$$\mathbb{E}_n[\Sigma_\mu(\lambda_n)] = \begin{cases} n^{\downarrow k} & \text{if } \mu = 1^k, \\ 0 & \text{otherwise.} \end{cases}$$

PROOF. The normalised character of the regular representation of a finite group G is $\chi^{\mathbb{C}G}(g) = 1_{(g=e_G)}$. Indeed, by taking the linear basis $(h)_{h\in G}$ of $\mathbb{C}G$ in order to compute the traces, we obtain:

$$\mathrm{ch}^{\mathbb{C}G}(g) = \sum_{h\in G}[\text{coefficient of } h \text{ in } gh] = \begin{cases} \mathrm{card}\, G & \text{if } g = e_G, \\ 0 & \text{if } g \neq e_G. \end{cases}$$

As a consequence,

$$\mathbb{E}_n[\varSigma_\mu(\lambda_n)] = n^{\downarrow k}\, \mathbb{E}_n[\chi^{\lambda_n}(\sigma_\mu)] = n^{\downarrow k}\, \chi^{\mathbb{C}\mathfrak{S}(n)}(\sigma_\mu) = n^{\downarrow k}\, 1_{(\sigma_\mu = \mathrm{id}_{[\![1,n]\!]})} = n^{\downarrow k}\, 1_{(\mu=1^k)}. \qquad \square$$

If we want to understand the distribution of the random variables $\varSigma_\mu(\lambda_n)$, then we need to compute more generally their (joint) moments. It turns out that the functions $\varSigma_\mu$ form a graded algebra of functions on $\mathfrak{Y}$ with explicit structure coefficients. The construction of this algebra is due to Kerov and Olshanski, and it can be performed thanks to the following extremely nice argument. Call *partial permutation* a pair (σ, A), where $A \subset \mathbb{N}^*$ is a finite subset (possibly empty), and $\sigma \in \mathfrak{S}(A)$. The cycle-type of a partial permutation (σ, A) is the cycle-type of σ viewed as an element of $\mathfrak{S}(A)$; hence, it is an integer partition with size $|A|$. The product of two partial permutations (σ, A) and (τ, B) is defined by:

$$(\sigma, A) \times (\tau, B) = (\sigma \circ \tau, A \cup B).$$

If $\deg(\sigma, A) = \mathrm{card}\, A$, then the product above makes the space of formal linear combinations of partial permutations into a graded algebra. We can even allow formal linear combinations with an infinite number of terms, as long as the degree of a sum stays bounded: indeed, given a partial permutation (ρ, C), there is a finite number of ways of decomposing it as a product $(\sigma \circ \tau, A \cup B)$ with $\sigma \in \mathfrak{S}(A)$ and $\tau \in \mathfrak{S}(B)$. Thus, the *algebra of partial permutations* is the algebra $\mathscr{P}$ whose elements are the formal series

$$\sum_{(\sigma,A)\in P} c_{(\sigma,A)}\, (\sigma, A),$$

where the coefficients $c_{(\sigma,A)}$ are complex numbers, and P is a possibly infinite set of partial permutations with $\sup\{\deg(\sigma, A) \,|\, (\sigma, A) \in P\} < +\infty$. The algebra $\mathscr{P}$ is graded by

$$\deg\left(\sum_{(\sigma,A)\in P} c_{(\sigma,A)}\, (\sigma, A)\right) = \sup\{\deg(\sigma, A) \,|\, (\sigma, A) \in P \text{ and } c_{(\sigma,A)} \neq 0\},$$

and $\deg(fg) \leq \deg(f) + \deg(g)$ for any $f, g \in \mathscr{P}$. In the following, we shall be interested in a commutative subalgebra of $\mathscr{P}$ which will turn out to be related to the functions $\varSigma_\mu$. Let us introduce the *infinite symmetric group* $\mathfrak{S}(\infty) = \bigcup_{n=0}^{\infty} \uparrow \mathfrak{S}(n)$, which consists in the permutations of the integers which fix every integer except the elements of a finite subset of $\mathbb{N}^*$. The group $\mathfrak{S}(\infty)$ acts by conjugation on the partial permutations:

$$\tau \cdot (\sigma, A) = (\tau\sigma\tau^{-1}, \tau(A)).$$

This action is compatible with the products and with the gradation; therefore, the elements of $\mathscr{P}$ which are invariant by conjugation by any permutation $\sigma \in \mathfrak{S}(\infty)$ form a graded subalgebra of $\mathscr{P}$. We denote

$$\mathscr{O} = \mathscr{P}^{\mathfrak{S}(\infty)} = \{f \in \mathscr{P} \,|\, \forall \sigma \in \mathfrak{S}(\infty), \sigma \cdot f = f\},$$

and we call it the *algebra of observables* (of Young diagrams). This terminology will be explained in a moment. In the sequel, given an integer partition μ of size k and an arrangement of integers $a = (a_1 \neq a_2 \neq \cdots \neq a_k)$, we denote

$$\sigma_\mu(a) = \big((a_1, \ldots, a_{\mu_1})(a_{\mu_1+1}, \ldots, a_{\mu_1+\mu_2}) \cdots (a_{\mu_1+\cdots+\mu_{\ell-1}+1}, \ldots, a_k)\,,\, \{a_1, \ldots, a_k\}\big);$$

this partial permutation has cycle-type μ.

THEOREM 4.14 (Algebra of observables). *The algebras $\mathscr{P}$ and $\mathscr{O}$ have the following properties:*

(1) *The algebra $\mathscr{O}$ is commutative, and a graded linear basis of it consists in the symbols Σ_μ with $\mu \in \mathfrak{Y}$:*
$$\Sigma_\mu = \sum_{\substack{a:[\![1,|\mu|]\!]\to\mathbb{N}^* \\ \text{arrangement}}} \sigma_\mu(a).$$

(2) *For any $n \geq 1$, we have a surjective morphism of algebras:*
$$\pi_n : \mathscr{P} \to \mathbb{C}\mathfrak{S}(n)$$
$$(\sigma, A) \mapsto \begin{cases} \sigma & \text{if } A \subset [\![1,n]\!]\,, \\ 0 & \text{otherwise.} \end{cases}$$
This morphism restricts to a surjective morphism of commutative algebras from $\mathscr{O}$ to the center $\mathrm{Z}(\mathbb{C}\mathfrak{S}(n))$.

(3) *Fix an integer partition $\lambda \in \mathfrak{Y}(n)$ with arbitrary size $n \geq 1$, and consider the linear map on $\mathscr{O}$ defined by*
$$\Sigma_\mu \mapsto \Sigma_\mu(\lambda),$$
the image being the rescaled character value from Definition 4.10. This map is a morphism of algebras from $\mathscr{O}$ to $\mathbb{C}$.

Proof. Consider the linear map π_n defined in the second item of the theorem. It is indeed a morphism of algebras, because we have in $\mathbb{C}\mathfrak{S}(n)$:
$$\begin{aligned}\pi_n((\sigma,A)\times(\tau,B) &= \pi_n(\sigma\circ\tau, A\cup B) = 1_{A\cup B\subset[\![1,n]\!]}\,(\sigma\circ\tau) \\ &= \big((1_{A\subset[\![1,n]\!]})\,\sigma\big)\big((1_{B\subset[\![1,n]\!]})\,\tau\big) = \pi_n(\sigma,A)\,\pi_n(\tau,B).\end{aligned}$$
It is not true that the collection of maps $(\pi_n)_{n\geq 1}$ separate the elements of $\mathscr{P}$: for instance, the two partial permutations $(\mathrm{id}_{[\![1,2]\!]}, [\![1,2]\!])$ and $(\mathrm{id}_{\{2\}}, \{2\})$ have the same image by all the maps π_n, but they are distinct. However, the collection of maps $(\pi_n)_{n\geq 1}$ separate the elements of $\mathscr{O}$. To prove this, let us first determine a linear basis of the graded subalgebra $\mathscr{O}$. If a partial permutation (σ, A) with cycle-type μ appears with a coefficient c in $f \in \mathscr{O}$, then by invariance by conjugation by $\mathfrak{S}(\infty)$, the partial permutation $\sigma_\mu(a)$ also appears with the same coefficient in f for any arrangement a of size $k = \operatorname{card} A$. Moreover, the symbols Σ_μ are linearly independent, since they involve partial permutations in disjoint sets. Thus, we have shown that
$$\mathscr{O} = \mathrm{Span}(\{\Sigma_\mu \,|\, \mu \in \sqcup_{n\geq 0}\mathfrak{Y}(n)\}),$$
and that the symbols Σ_μ form a linear basis of $\mathscr{O}$. We obviously have $\deg \Sigma_\mu = |\mu|$ for any integer partition μ. Let us now determine the image of Σ_μ by π_n. Notice that
$$\pi_n(\tau\cdot(\sigma,A)) = \tau(\pi_n(\sigma,A))\tau^{-1}$$
for any permutation $\tau \in \mathfrak{S}(n)$. Therefore, if $f \in \mathscr{O}$, then $\pi_n(f)$ is invariant by conjugation in $\mathbb{C}\mathfrak{S}(n)$, so it belongs to the center of the group algebra $\mathbb{C}\mathfrak{S}(n)$. This center is spanned by the conjugacy classes C_ν with $\nu \in \mathfrak{Y}(n)$. Denote $\kappa_\nu = \frac{C_\nu}{\operatorname{card} C_\nu}$. We have in fact:
$$\pi_n(\Sigma_\mu) = \begin{cases} n^{\downarrow k}\,\kappa_{\mu\sqcup 1^{n-k}} & \text{if } n \geq k, \\ 0 & \text{if } n < k, \end{cases}$$
if $k = |\mu| = \deg \Sigma_\mu$. Indeed, the image by π_n of a partial permutation (σ, A) with cycle-type μ is either 0 if $A \not\subset [\![1,n]\!]$, or a permutation with completed cycle-type $\mu \sqcup ((n-k)$ fixed points) if $A \subset [\![1,n]\!]$. The formula above is then a consequence of the following facts:

- there are $n^{\downarrow k}$ arrangements a with size k and values in $[\![1,n]\!]$;
- the image of $\pi_n(\Sigma_\mu)$ belongs to the center of $\mathbb{C}\mathfrak{S}(n)$, hence is a multiple of the conjugacy class $C_{\mu\sqcup 1^{n-k}}$ when $n \geq k$.

This implies that an element $f \in \mathscr{O}$ is entirely determined by its images $\pi_n(f)$, $n \geq 1$. Indeed, consider an integer partition μ without parts of size 1, and all the completed versions $\mu \sqcup 1, \ldots, \mu \sqcup 1^d$ whose symbols appear with a non-zero coefficient in f. There are only a finite number of them since f is of finite degree. So,

$$f = \sum_{k=0}^{d} c_k\, \Sigma_{\mu \sqcup 1^k} + \text{other terms (not of the form } \Sigma_{\mu \sqcup 1^j}).$$

When looking at $\pi_n(f)$ with $n \geq |\mu| + d$, we obtain:

$$\pi_n(f) = \left(\sum_{k=0}^{d} c_k\, n^{\downarrow n-k-|\mu|}\right) \kappa_{\mu \sqcup 1^{n-|\mu|}} + \text{linear combination of other conjugacy classes.}$$

So, from the knowledge of the $\pi_n(f)$'s, we deduce the polynomials $\sum_{k=0}^{d} c_k\, n^{\downarrow n-k-|\mu|}$, whence all the coefficients c_k; this proves our claim. It implies that $\mathscr{O}$ is a commutative algebra: indeed, given f and g in $\mathscr{O}$, since $\pi_n(fg) = \pi_n(f)\, \pi_n(g) = \pi_n(g)\, \pi_n(f) = \pi_n(gf)$ for any n, $fg = gf$.

Let us finally establish the third item of the proposition. It follows from a rewriting of the rescaled character value: if $|\lambda| = n$, then

$$\Sigma_\mu(\lambda) = \chi^\lambda(\pi_n(\Sigma_\mu)),$$

where $\chi^\lambda(\sum_{\sigma \in \mathfrak{S}(n)} c_\sigma\, \sigma) = \sum_{\sigma \in \mathfrak{S}(n)} c_\sigma\, \chi^\lambda(\sigma)$. However, π_n is a morphism of algebras from $\mathscr{O}$ to $\mathrm{Z}(\mathbb{C}\mathfrak{S}(n))$, and the normalised character value χ^λ is a morphism of algebras from $\mathrm{Z}(\mathbb{C}\mathfrak{S}(n))$ to $\mathbb{C}$. For this second claim, notice that any element x of $\mathrm{Z}(\mathbb{C}\mathfrak{S}(n))$ is sent by the Fourier transform to a linear combination of identity matrices id_{S^λ} (this is because $\mathrm{Z}(\mathrm{End}(S^\lambda)) = \mathbb{C}\mathrm{id}_{S^\lambda}$); and $\chi^\lambda(x)$ is the coefficient of the matrix id_{S^λ} in $\widehat{x}$. Therefore,

$$\begin{aligned}\chi^\lambda(xy) &= \text{coefficient of } \mathrm{id}_{S^\lambda} \text{ in } \widehat{xy} = \widehat{x}\,\widehat{y} \\ &= (\text{coefficient of } \mathrm{id}_{S^\lambda} \text{ in } \widehat{x})(\text{coefficient of } \mathrm{id}_{S^\lambda} \text{ in } \widehat{y}) = \chi^\lambda(x)\,\chi^\lambda(y)\end{aligned}$$

for any x, y in the center of the group algebra. Now, by composition, we conclude that $\Sigma_\mu \mapsto \Sigma_\mu(\lambda)$ is indeed a morphism of algebras from $\mathscr{O}$ to $\mathbb{C}$. □

The last part of Theorem 4.14 enables us to compute products of observables by using the formalism of partial permutations, and then to use these products in order to evaluate the moments of $\Sigma_\mu(\lambda_n)$ with $\lambda_n \sim \mathrm{PL}_n$. This method leads to the following important result:

THEOREM 4.15 (Kerov's central limit theorem). *For $k \geq 2$, denote $c_k = (1, 2, \ldots, k)$. Under the Plancherel measures PL_n, the random character values*

$$n^{\frac{k}{2}}\, \frac{\chi^{\lambda_n}(c_k)}{\sqrt{k}}$$

converge in law towards independent standard normal variables with law $\mathcal{N}_{\mathbb{R}}(0, 1)$.

The end of this section is devoted to a proof of this central limit theorem. Let us start by explaining how to compute a general product $\Sigma_\mu\, \Sigma_\nu$. In the following, we fix two integer partitions μ and ν with respective sizes m and n, and we shall consider a disjoint copy $\mathbb{N}'$ of $\mathbb{N}$, whose elements will be denoted $0'$, $1'$, *etc.* A *partial pairing* between $[\![1, m]\!]$ and $[\![1', n']\!]$ is a (possibly empty) set P of disjoint pairs (x, y') with $x \in [\![1, m]\!]$ and $y' \in [\![1', n']\!]$. For instance, $P = \{(1, 3'), (3, 2')\}$ is a partial pairing between $[\![1, 3]\!]$ and $[\![1', 4']\!]$. We denote $\mathfrak{P}(m, n)$ the finite set of partial pairings between $[\![1, m]\!]$ and $[\![1', n']\!]$. This is a finite set with cardinality

$$\sum_{k=0}^{\min(m,n)} k! \binom{m}{k} \binom{n}{k} = \sum_{k=0}^{\min(m,n)} \frac{m!\, n!}{(m-k)!\, (n-k)!\, k!}.$$

Given two cycle-types $\mu \in \mathfrak{Y}(m)$ and $\nu \in \mathfrak{Y}(n)$ and $P \in \mathfrak{P}(m,n)$, we denote $\rho(\mu,\nu,P)$ the cycle-type with size $|\mu|+|\nu|-|P|$ of the partial permutation

$$\sigma_\mu(1,2,\ldots,n) \times \sigma_\nu(a'),$$

where a' is the arrangement obtained from $1',2',\ldots,n'$ by replacing each element y' of a pair $(x,y') \in P$ by x. For instance, if $\mu = (3)$, $\nu = (2,2)$ and $P = \{(1,3'),(3,2')\}$, then $\rho(\mu,\nu,P)$ is the cycle-type of

$$(1,2,3) \times (1',3)(1,4') = (1,4',2,3,1')$$

viewed as a permutation of $\{1,2,3,1',4'\}$, hence (5).

Proposition 4.16 (Product of symbols Σ_μ). *For any integer partitions μ and ν with respective sizes m and n,*

$$\Sigma_\mu\, \Sigma_\nu = \sum_{P \in \mathfrak{P}(m,n)} \Sigma_{\rho(\mu,\nu,P)}.$$

Proof. Given two arrangements $a : [\![1,m]\!] \to \mathbb{N}^*$ and $b : [\![1,n]\!] \to \mathbb{N}^*$, we associate to them a partial pairing $P(a,b)$ whose pairs are the (x,y') with $a(x) = b(y)$. Then, it is easily seen that the cycle-type of $\sigma_\mu(a)\,\sigma_\nu(b)$ only depends on $P(a,b)$, and is $\rho(\mu,\nu,P(a,b))$. Moreover, if $P \in \mathfrak{P}(m,n)$ is fixed, then there is a natural bijection between

$$\{\text{pairs of arrangements } (a,b) \text{ with } |a| = m,\ |b| = n,\ P(a,b) = P\}$$
$$\text{and} \quad \{\text{arrangements } c \text{ with } |c| = m+n-|P|\},$$

such that $\sigma_\mu(a)\,\sigma_\nu(b) = \sigma_{\rho(\mu,\nu,P(a,b))}(c)$. This natural bijection correspond to the rewriting as $m+1, m+2, \ldots, m+n-|P|$ of those elements y' in $[\![1',n']\!]$ which are not in a pair $(x,y') \in P(a,b)$. Then, $c(x) = a(x)$ for $x \in [\![1,m]\!]$, and $c(x) = b(y)$ for $x \in [\![m+1, m+n-|P|]\!]$ corresponding to a $y' \in [\![1',n']\!]$ as described above. Now, thanks to these bijections, we get:

$$\Sigma_\mu\,\Sigma_\nu = \sum_{\substack{a \text{ arrangement with size } m \\ b \text{ arrangement with size } n}} \sigma_\mu(a)\,\sigma_\nu(b) = \sum_{P \in \mathfrak{P}(m,n)} \left(\sum_{\substack{a \text{ arrangement with size } m \\ b \text{ arrangement with size } n \\ P(a,b) = P}} \sigma_\mu(a)\,\sigma_\nu(b) \right)$$
$$= \sum_{P \in \mathfrak{P}(m,n)} \left(\sum_{c \text{ arrangement with size } m+n-|P|} \sigma_{\rho(\mu,\nu,P)}(c) \right) = \sum_{P \in \mathfrak{P}(m,n)} \Sigma_{\rho(\mu,\nu,P)}. \qquad \square$$

Example 4.17. Proposition 4.16 is probably better understood with an example. If we take $\lambda = (3)$ and $\mu = (2)$, then we need to consider the following partial pairings:

- the empty partial pairing: a product of a 3-cycle with a 2-cycle with disjoint support is a partial permutation with cycle-type $(3,2)$, whence a term $\Sigma_{(3,2)}$ in the product $\Sigma_3\,\Sigma_2$.
- the 6 partial pairings with size 1: if a 3-cycle and a 2-cycle have their support that meet at exactly one point, then their product is a 4-cycle; whence a term $6\,\Sigma_4$ in the product $\Sigma_3\,\Sigma_2$.
- finally, the 6 partial pairings with size 2: it is easy to check that the corresponding cycle-type is always the same, and is the cycle-type of the partial permutation $\sigma_3(1,2,3)\sigma_2(1,2) = ((1,3),\{1,2,3\})$, that is $(2,1)$. Beware that since we are working with partial permutations, we have to take into account the fixed points of the permutation on its support. So, these pairings give a contribution $6\,\Sigma_{(2,1)}$ to the product.

We conclude that

$$\Sigma_3\,\Sigma_2 = \Sigma_{(3,2)} + 6\,\Sigma_4 + 6\,\Sigma_{(2,1)}.$$

REMARK 4.18. Since $\deg \Sigma_{\rho(\mu,\nu,P)} = |\mu| + |\nu| - |P|$, the leading term of a product $\Sigma_\mu\, \Sigma_\nu$ with respect to the degree corresponds to the empty partial pairing:

$$\Sigma_\mu\, \Sigma_\nu = \Sigma_{\mu\sqcup\nu} + \text{terms with strictly lower degree}.$$

Some analogues of this factorisation property with respect to other gradations on $\mathscr{O}$ will play an important role in the sequel.

LEMMA 4.19 (Counting twice the fixed points). *We introduce another gradation* rank *on* $\mathscr{O}$*:* $\mathrm{rank}(\Sigma_\mu) = |\mu| + m_1(\mu)$, *where* $m_1(\mu)$ *is the number of parts of size* 1 *in* μ. *Then,* rank *is a gradation of algebras:* $\mathrm{rank}(fg) \leq \mathrm{rank}(f) + \mathrm{rank}(g)$ *for any* $f, g \in \mathscr{O}$.

PROOF. The gradation rank is the restriction to $\mathscr{O}$ of the following gradation on partial permutations:

$$\mathrm{rank}(\sigma, A) = \mathrm{card}\, A + \text{number of fixed points of } \sigma \text{ in } A.$$

Let us then prove that we always have $\mathrm{rank}(\sigma\tau, A \cup B) \leq \mathrm{rank}(\sigma, A) + \mathrm{rank}(\tau, B)$. Given a permutation σ and a set A, we denote $\mathrm{Fix}(\sigma, A)$ the set of fixed points of σ in A. We decompose $\mathrm{Fix}(\sigma\tau, A \cup B)$ in three parts:

- the set of points of $B \setminus A$ which are fixed by τ;
- the set of points of $A \setminus B$ which are fixed by σ;
- the set of points of $A \cap B$ which are fixed by $\sigma\tau$.

This decomposition yields:

$$\begin{aligned}|\mathrm{Fix}(\sigma\tau, A \cup B)| &= |\mathrm{Fix}(\sigma, A \setminus B)| + |\mathrm{Fix}(\tau, B \setminus A)| + |\mathrm{Fix}(\sigma\tau, A \cap B)| \\ &\leq |\mathrm{Fix}(\sigma, A \setminus B)| + |\mathrm{Fix}(\tau, B \setminus A)| + |A \cap B|; \\ |\mathrm{Fix}(\sigma\tau, A \cup B)| + |A \cup B| &\leq |\mathrm{Fix}(\sigma, A \setminus B)| + |A| + |\mathrm{Fix}(\tau, B \setminus A)| + |B| \\ &\leq |\mathrm{Fix}(\sigma, A)| + |A| + |\mathrm{Fix}(\tau, B)| + |B|,\end{aligned}$$

whence the result. □

LEMMA 4.20. *Let* μ *and* ν *be two integer partitions without part in common: if* $m_k(\mu) > 0$ *for some* $k \geq 1$, *then* $m_k(\nu) = 0$. *Note that under this hypothesis, if* $\mu \sqcup \nu$ *is the integer partition whose multiset of parts is the disjoint union of the multiset of the parts of* μ *and of the multiset of the parts of* ν, *then we have* $\mathrm{rank}(\Sigma_\mu) + \mathrm{rank}(\Sigma_\nu) = \mathrm{rank}(\Sigma_{\mu\sqcup\nu})$. *We have in* $\mathscr{O}$*:*

$$\Sigma_\mu\, \Sigma_\nu = \Sigma_{\mu\sqcup\nu} + \textit{terms of strictly lower rank}.$$

PROOF. Let us examine under which conditions we have $\mathrm{rank}(\sigma\tau, A \cup B) = \mathrm{rank}(\sigma, A) + \mathrm{rank}(\tau, B)$. We need all the inequalities of the previous proof to be identities. Therefore:

- any point in $A \cap B$ has to be fixed by $\sigma\tau$;
- there are no fixed point of σ or of τ in $A \cap B$.

This can only happen if σ and τ leave $A \cap B$ invariant and if $\sigma_{|A\cap B} = (\tau_{|A\cap B})^{-1}$. However, σ and τ do not have cycles with the same length, so the only possibility is that $A \cap B = \emptyset$. We conclude that the term of higher rank in the product $\Sigma_\mu\, \Sigma_\nu$ corresponds to the empty partial pairing P, and is indeed $\Sigma_{\mu\sqcup\nu}$. □

REMARK 4.21. As a consequence of this lemma, we see that a linear basis of $\mathscr{O}$ which is graded with respect to the rank consists in the products $\Sigma_\mu\, (\Sigma_1)^k$ with $m_1(\mu) = 0$ and $k \in \mathbb{N}$. Indeed, given an arbitrary integer partition ν, we can always write it $\nu = \mu \sqcup 1^k$ with $m_1(\mu) = 0$ and $m_1(\nu) = k$, and the lemma states that Σ_ν and $\Sigma_\mu\, \Sigma_{1^k}$ differ by terms of strictly lower rank. On the other hand, one has also

$$\Sigma_{1^k} = (\Sigma_1)^k + \text{terms of strictly lower rank}.$$

Indeed, the subalgebra $\mathbb{C}[\Sigma_1]$ of $\mathscr{O}$ spanned by Σ_1 is an algebra of polynomials in one variable, and it is spanned by the symbols Σ_{1^k} with $k \geq 0$. This claim follows from the identity

$$\Sigma_{1^k}(\lambda) = n^{\downarrow k} \quad \text{if } |\lambda| = n,$$

which holds for any $\lambda \in \mathfrak{Y}$ and any $k \geq 0$. As a consequence, the two families $(\Sigma_\nu)_{\nu \in \mathfrak{Y}}$ and $(\Sigma_\mu\, (\Sigma_1)^k)_{m_1(\mu)=0,\, k\geq 0}$ differ by a matrix which is triangular with respect to the rank. So, they are both graded linear bases of $\mathscr{O}$ with respect to the rank. The rank of an observable $\Sigma_\mu\, (\Sigma_1)^k$ with $m_1(\mu) = 0$ is $|\mu| + 2k$.

LEMMA 4.22. *If $k \geq 2$ and $m \geq 1$, then*

$$\Sigma_k\, \Sigma_{k^m} = \Sigma_{k^{m+1}} + km\, \Sigma_{k^{m-1}} \Sigma_{1^k} + \textit{terms of strictly lower rank}.$$

PROOF. The main terms of the formula are of rank $k(m+1)$, so we need again to examine under which conditions we have $\mathrm{rank}(\sigma\tau, A \cup B) = \mathrm{rank}(\sigma, A) + \mathrm{rank}(\tau, B)$ when (σ, A) has cycle-type k and (τ, B) has cycle-type k^m. Again, this can only happen if σ and τ leave $A \cap B$ invariant and if $\sigma_{|A\cap B} = (\tau_{|A\cap B})^{-1}$. Now, there are two possibilities:

(1) $A \cap B = \emptyset$. This corresponds to the empty partial pairing in Proposition 4.16, and the corresponding contribution in $\Sigma_k\, \Sigma_{k^m}$ is $\Sigma_{k^{m+1}}$.

(2) $A \cap B = A$ has size k, and the restriction of τ to A is the inverse of the cycle σ. There are km corresponding partial pairings in $\mathfrak{P}(k, k^m)$:

   (a) we choose one of the interval $[\![(j-1)k + 1', jk']\!]$ with $j \in [\![1, m]\!]$.

   (b) we choose $t \in [\![1, k]\!]$ and we pair 1 with $(j-1)k + t'$, 2 with $(j-1)k + t - 1'$, *etc.*, t with $(j-1)k + 1'$, $t+1$ with jk', *etc.*, and k with $(j-1)k + t + 1'$.

   The second item comes from the fact that there are k ways to write in cyclic order the elements of a k-cycle. Now, for any of these pairings, the cycle-type of the product of partial permutations $(\sigma, A)(\tau, B) = (\sigma\tau, B)$ is $k^{m-1}\, 1^m$, whence a contribution $km\, \Sigma_{k^{m-1}\, 1^k}$ in the product $\Sigma_k\, \Sigma_{k^m}$.

Finally we can replace $\Sigma_{k^{m-1}\, 1^k}$ by $\Sigma_{k^{m-1}}\, \Sigma_{1^k}$ by the previous remark. ☐

The renormalisation of the random character values in Kerov's central limit theorem 4.15 makes it natural to extend a bit the algebra of observables $\mathscr{O}$, by adding the (possibly negative) half-powers of the observable Σ_1. Indeed, since $\Sigma_1(\lambda) = n$ if $|\lambda| = n$, and since $n^{\downarrow k} = n^k + O(n^{k-1})$, we have

$$n^{\frac{k}{2}}\, \frac{\chi^\lambda(c_k)}{\sqrt{k}} = \frac{\Sigma_k(\lambda)}{\sqrt{k}\, (\Sigma_1(\lambda))^{\frac{k}{2}}}(1 + O(n^{-1})),$$

so it is natural to work with the *extended observable*

$$X_k = \frac{\Sigma_k}{\sqrt{k}\, (\Sigma_1)^{\frac{k}{2}}},$$

which belongs to the commutative $\mathscr{O} \otimes_{\mathbb{C}[\Sigma_1]} \mathbb{C}[(\Sigma_1)^{\frac{1}{2}}, (\Sigma_1)^{-\frac{1}{2}}]$. Denote $\mathscr{O}^+$ this extended algebra; a linear basis of $\mathscr{O}^+$ consists in the products

$$\Sigma_\mu\, (\Sigma_1)^{\frac{k}{2}} \quad \text{with } m_1(\mu) = 0 \text{ and } k \in \mathbb{Z},$$

and the rank gradation extends to $\mathscr{O}^+$ by setting $\mathrm{rank}(\Sigma_\mu\,(\Sigma_1)^{\frac{k}{2}}) = |\mu| + k$ if $m_1(\mu) = 0$. Given an integer partition μ without parts of size 1, we set

$$X_\mu = \frac{1}{\sqrt{\prod_{i=1}^{\ell(\mu)} \mu_i}}\,\frac{\Sigma_\mu}{(\Sigma_1)^{\frac{|\mu|}{2}}};$$

it is an extended observable with rank 0.

Proposition 4.23. *For any integer partition μ with $m_1(\mu) = 0$, we have the following relation in the extended algebra $\mathscr{O}^+$:*

$$X_\mu = \prod_{k\geq 2} H_{m_k(\mu)}(X_k) + \textit{terms with negative rank},$$

where $(H_m)_{m\geq 0}$ is the family of orthogonal polynomials associated to the standard Gaussian distribution $(2\pi)^{-\frac{1}{2}}\,\mathrm{e}^{-\frac{x^2}{2}}\,dx$.

Proof. The family of *Hermite polynomials* is obtained by Gram–Schmidt orthogonalisation of the family $(x^m)_{m\geq 0}$ in $\mathscr{L}^2(\mathbb{R},(2\pi)^{-\frac{1}{2}}\,\mathrm{e}^{-\frac{x^2}{2}}\,dx)$. Thus, $H_0 = 1$, and H_{m+1} is characterised by the following relations:

- $H_{m+1}(x) = x^{m+1} +$ polynomial with degree smaller than m.
- $\langle H_{m+1} \mid H_{l\leq m}\rangle = 0$, the scalar product being taken with respect to the Gaussian distribution.

It is easily seen that a solution of this recurrence is provided by the Rodrigues' formula:

$$H_m(x) = (-1)^m\,\mathrm{e}^{\frac{x^2}{2}}\,\frac{d^m}{dx^m}\left(\mathrm{e}^{-\frac{x^2}{2}}\right).$$

The first item is trivial by recurrence on m, and for the second item, we use an integration by parts:

$$\langle H_{m+1} \mid H_{l\leq m}\rangle = \frac{(-1)^{m+1}}{\sqrt{2\pi}}\int_{\mathbb{R}} \frac{d^{m+1}}{dx^{m+1}}\left(\mathrm{e}^{-\frac{x^2}{2}}\right) H_l(x)\,dx = \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}}\left(\mathrm{e}^{-\frac{x^2}{2}}\right)\frac{d^{m+1}}{dx^{m+1}}(H_l(x))\,dx = 0;$$

indeed, we derivate $(m+1)$ times a polynomial with degree $l \leq m$. As a consequence of the Rodrigues' formula and of the Taylor expansion at x of the entire function $z \mapsto \mathrm{e}^{-\frac{z^2}{2}}$, we get:

$$\mathrm{e}^{-\frac{(x-t)^2}{2}} = \sum_{m=0}^{\infty} \frac{(-t)^m}{m!}\frac{d^m}{dx^m}\left(\mathrm{e}^{-\frac{x^2}{2}}\right) = \mathrm{e}^{-\frac{x^2}{2}}\sum_{m=0}^{\infty} H_m(x)\,\frac{t^m}{m!};$$

$$\mathrm{e}^{xt-\frac{t^2}{2}} = \sum_{m=0}^{\infty} H_m(x)\,\frac{t^m}{m!}.$$

If we derivate the last formula with respect to the variable t, we get on the left-hand side

$$(x-t)\,\mathrm{e}^{xt-\frac{t^2}{2}} = \sum_{m=0}^{\infty}(x\,H_m(x) - m\,H_{m-1}(x))\,\frac{t^m}{m!},$$

and on the right-hand side $\sum_{m=0}^{\infty} H_{m+1}(x)\,\frac{t^m}{m!}$. This proves the recurrence relation $H_{m+1}(x) = x\,H_m(x) - m\,H_{m-1}(x)$ for any $m \geq 1$. Let us now prove the formula for the extended observable X_μ. It follows from Lemma 4.20 that

$$\Sigma_\mu = \prod_{s\geq 2} \Sigma_{s^{m_s(\mu)}} + \text{terms with strictly lower rank};$$

$$X_\mu = \prod_{s\geq 2} X_{s^{m_s(\mu)}} + \text{terms with negative rank}.$$

Therefore, it suffices to prove that $X_{k^m} = H_m(X_k) +$ terms with negative rank for any $k \geq 2$ and any $m \geq 1$. Suppose the result true up to order $m \geq 1$. Then, by Lemma 4.22,

$$\begin{aligned} X_{k^{m+1}} &= \frac{\Sigma_{k^{m+1}}}{k^{\frac{m+1}{2}}\,(\Sigma_1)^{\frac{k(m+1)}{2}}} = \frac{\Sigma_k\,\Sigma_{k^m}}{k^{\frac{m+1}{2}}\,(\Sigma_1)^{\frac{k(m+1)}{2}}} - m\,\frac{\Sigma_{k^{m-1}}}{k^{\frac{m-1}{2}}\,(\Sigma_1)^{\frac{k(m-1)}{2}}} + \cdots \\ &= X_k\,X_{k^m} - m\,X_{k^{m-1}} + \cdots \\ &= X_k\,H_m(X_k) - m\,H_{m-1}(X_k) + \cdots, \end{aligned}$$

where on each line the dots indicate terms with negative rank. By the previous discussion, the combination of Hermite polynomials is $H_{m+1}(X_k)$, whence the result at order $m+1$. □

Proof of Theorem 4.15. Note that for any extended observable $f \in \mathscr{O}^+$, we have

$$\mathbb{E}_n[f(\lambda_n)] = O\left(n^{\frac{\mathrm{rank}(f)}{2}}\right),$$

with a constant in the $O(\cdot)$ which depends on f. Indeed, it suffices to prove it for the linear basis $\Sigma_\mu\,(\Sigma_1)^{\frac{k}{2}}$. If $\mu \neq \emptyset$, then $\mathbb{E}_n[\Sigma_\mu(\lambda_n)\,(\Sigma_1(\lambda_n))^{\frac{k}{2}}] = n^{\frac{k}{2}}\,\mathbb{E}_n[\Sigma_\mu(\lambda_n)] = 0$ by Corollary 4.13; and if $\mu = \emptyset$, then Σ_1 is the constant function equal to n on $\mathfrak{Y}(n)$, so $\mathbb{E}_n[(\Sigma_1(\lambda_n))^{\frac{k}{2}}] = n^{\frac{k}{2}}$ and the estimate is also true in this case.

We now compute the joint moments of a collection of renormalised character values. We denote $Y_{k\geq 2,n}$ the random variables involved in the statement of Theorem 4.15, and we consider a non-zero integer partition $\mu = 2^{m_2}3^{m_3}\cdots s^{m_s}$. Note that we have

$$\begin{aligned} \prod_{k=2}^{s} H_{m_k}(Y_{k,n}) &= \prod_{k=2}^{s} H_{m_k}\big(X_k(\lambda_n)(1+O(n^{-1}))\big) \\ &= \prod_{k=2}^{s} H_{m_k}(X_k(\lambda_n)) + \sum_{i\in I} P_i(n^{-1})\,Q_i(X_2,X_3,\ldots,X_s), \end{aligned}$$

where the P_i's are polynomials without constant term, and the $Q_i(X_2(\lambda_n),X_3(\lambda_n),\ldots,X_s(\lambda_n))$ are polynomials in the extended observables $X_k(\lambda)$, and therefore have rank smaller than or equal to 0. Taking expectations, we get:

$$\begin{aligned} \mathbb{E}_n\left[\prod_{k=2}^{s} H_{m_k}(Y_{k,n})\right] &= \mathbb{E}_n\left[\prod_{k=2}^{s} H_{m_k}(X_k(\lambda_n))\right] + O(n^{-1}) \\ &= \mathbb{E}_n[X_\mu(\lambda_n)] + O\left(n^{-\frac{1}{2}}\right) = O\left(n^{-\frac{1}{2}}\right) \end{aligned}$$

by using on the second line the content of Proposition 4.23, and the fact $\mathbb{E}_n[\Sigma_\mu(\lambda_n)] = 0$ if μ has parts of size greater than 2. On the other hand, if $(Y_k)_{k\geq 1}$ is a family of independent standard Gaussian variables, then

$$\begin{aligned} \mathbb{E}\left[\prod_{k=2}^{s} H_{m_k}(Y_k)\right] &= \prod_{k=2}^{s} \langle H_{m_k} \mid 1\rangle_{\mathscr{L}^2\left(\mathbb{R},(2\pi)^{-\frac{1}{2}}\mathrm{e}^{-\frac{x^2}{2}}\,dx\right)} \\ &= 0 \quad \text{if } \mu \neq \emptyset. \end{aligned}$$

Since the Hermite polynomials form a linear basis of the space of polynomials, we conclude that the joint moments of the variables $Y_{k,n}$ all converge towards those of a family of independent standard Gaussian variables. The distribution of these limiting variables is characterised by its joint moments (because we have convergent joint Laplace transforms), and it is well known that this implies the convergence in law. □

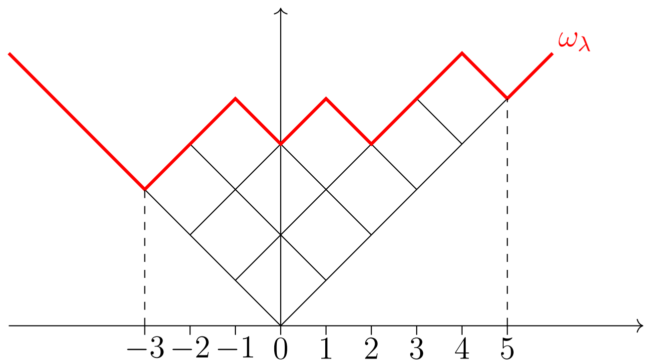
3. From character values to geometric observables

The previous section enabled us to understand the behavior of the observables $\Sigma_\mu(\lambda_n)$ when $\lambda_n \sim \mathrm{PL}_n$. However, we were initially interested in the geometry of the random integer partitions λ_n, so we now need to explain how the character values can be related to the geometry of Young diagrams. As a preliminary step, let us construct a functional space which is convenient in order to speak of convergence of the shape of a sequence of Young diagrams.

Definition 4.24 (Continuous Young diagram). *A continuous Young diagram ω is a function $\omega : \mathbb{R} \to \mathbb{R}_+$ which is Lipschitz with constant 1, and such that $\omega(s) = |s|$ for $|s|$ large enough. The area of a continuous Young diagram ω is defined as*

$$\frac{1}{2}\,\widetilde{p}_2(\omega) = \int_{\mathbb{R}} \frac{\omega(s) - |s|}{2}\, ds.$$

Example 4.25 (Continuous Young diagram of an integer partition). Given an integer partition λ, we can associate to it a continuous Young diagram ω_λ with area $\frac{1}{2}\,\widetilde{p}_2(\omega_\lambda) = |\lambda|$. Let us draw the cells of the Young diagram of λ as squares with area 2, and rotate this Young diagram by 45 degrees. The upper boundary of this drawing can be considered as an affine by parts function ω_λ on $[-\ell(\lambda), \lambda_1]$, and we extend this function to $\mathbb{R}$ by setting $\omega_\lambda(s) = |s|$ outside the interval $[-\ell(\lambda), \lambda_1]$. By construction, the function ω_λ is then a continuous Young diagram with area $|\lambda|$. For instance, if $\lambda = (5, 3, 2)$, then we obtain the following continuous Young diagram:



Example 4.26 (Rescaling of continuous Young diagrams). For $t > 0$, the space $\mathscr{Y}$ of all the continuous Young diagrams is stable by the operation of rescaling $\omega \mapsto \omega_{(t)}$, where

$$\omega_{(t)}(s) = \sqrt{t}\,\omega\left(\frac{s}{\sqrt{t}}\right).$$

This operation scales both axes by a factor $\sqrt{t}$, so it multiplies the areas by t: $\frac{1}{2}\,\widetilde{p}_2(\omega_{(t)}) = \frac{t}{2}\,\widetilde{p}_2(\omega)$. By rescaling the continuous Young diagrams of integer partitions, we can define a notion of *convergence of shape* for a sequence of integer partitions $(\lambda_n)_{n\in\mathbb{N}}$ with $|\lambda_n| = n$ for every n. Thus, we shall say that the shapes of these integer partitions converge if the sequence of rescaled continuous Young diagrams

$$\omega_n = (\omega_{\lambda_n})_{(\frac{1}{n})}$$

converges uniformly on $\mathbb{R}$ to some *limit shape* $\omega_\infty \in \mathscr{Y}$. We shall see in a moment that the shapes of the random integer partitions $\lambda_n \sim \mathrm{PL}_n$ converge in probability towards some explicit limit shape Ω.

Example 4.27 (Logan–Shepp–Kerov–Vershik curve). The *Logan–Shepp–Kerov–Vershik* curve is the continuous Young diagram Ω defined by

$$\Omega(s) = \begin{cases} \frac{2}{\pi}\left(s \arcsin \frac{s}{2} + \sqrt{4 - s^2}\right) & \text{if } |s| < 2, \\ |s| & \text{if } |s| \geq 2. \end{cases}$$

Its area is equal to 1, and more generally, we can compute the moments

$$\widetilde{p}_k(\Omega) = k(k-1)\int_{\mathbb{R}} \frac{\Omega(s)-|s|}{2}\, s^{k-2}\,ds = \int_{\mathbb{R}} \left(\frac{\Omega(s)-|s|}{2}\right)''\, s^k\,ds$$

for $k \geq 2$ (in the second integral, the second derivative might have to be considered as a distribution). Indeed, by integration by parts on $[-2,2]$,

$$\widetilde{p}_k(\Omega) = -\frac{k}{2}\int_{-2}^{2} \left(\frac{2}{\pi}\,\arcsin\frac{s}{2} - \mathrm{sgn}(s)\right) s^{k-1}\,ds$$

vanishes if k is odd (by parity), and it writes otherwise as

$$\begin{aligned}\widetilde{p}_{2j}(\Omega) &= \int_0^2 \left(1-\frac{2}{\pi}\,\arcsin\frac{s}{2}\right) d(s^{2j}) = 2^{2j+1}j\int_0^{\frac{\pi}{2}} \left(1-\frac{2\theta}{\pi}\right)(\sin\theta)^{2j-1}\,\cos\theta\,d\theta \\ &= \frac{2^{2j+1}}{\pi}\int_0^{\frac{\pi}{2}} (\sin\theta)^{2j}\,d\theta = \binom{2j}{j}.\end{aligned}$$

Here we have used the classical recurrence relation $I_{2j} = \frac{2j-1}{2j}\,I_{2j-2}$ satisfied by the Wallis integrals $I_{2j} = \int_0^{\frac{\pi}{2}}(\sin\theta)^{2j}\,d\theta$ in order to compute these integrals.

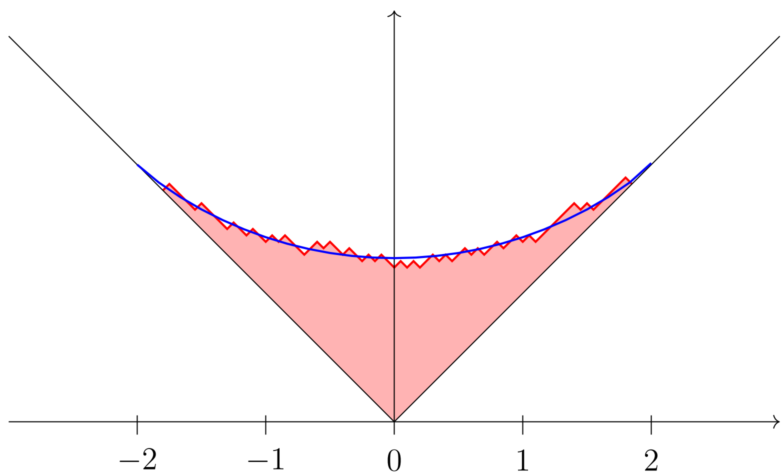
We are now ready to state the main result of this chapter, which is the law of large numbers satisfied by the Plancherel measures:

THEOREM 4.28 (Logan–Shepp, Kerov–Vershik). *Let $(\lambda_n)_{n\in\mathbb{N}}$ be a sequence of random integer partitions distributed according to the Plancherel measures PL_n of the symmetric groups. We denote $(\omega_n)_{n\in\mathbb{N}}$ the corresponding sequence of (random) rescaled continuous Young diagrams: $\omega_n = (\omega_{\lambda_n})_{(\frac{1}{n})}$. The sequence $(\omega_n)_{n\in\mathbb{N}}$ converges in probability to the LSKV curve Ω: for any $\varepsilon > 0$,*

$$\lim_{n\to\infty} \mathbb{P}\left[\sup_{s\in\mathbb{R}} |\omega_n(s)-\Omega(s)| \geq \varepsilon\right] = 0.$$

Thus, the shapes of the random integer partitions λ_n converge in probability to the LSKV curve.

We have drawn below a random integer partition $\lambda \sim \mathrm{PL}_{400}$ (in red), and its limit shape Ω (in blue); the Logan–Shepp–Kerov–Vershik law of large numbers states that the boundary of the Young diagram λ is close to the LSKV curve with high probability.



COROLLARY 4.29 (Solution of Ulam's problem). *The length ℓ_n of a random permutation σ_n chosen uniformly in $\mathfrak{S}(n)$ satisfies the following law of large numbers:*

$$\frac{\ell_n}{2\sqrt{n}} \to_{\mathbb{P}} 1,$$

the arrow $\to_{\mathbb{P}}$ meaning that we have convergence in probability.

PROOF. Fix $\varepsilon > 0$, and let us evaluate the probability $\mathbb{P}[\ell_n \geq 2(1-\varepsilon)\sqrt{n}+1]$. By Proposition 4.8, this probability is also the probability under PL_n that the first part $\lambda_{n,1}$ of a random integer partition is larger than $2(1-\varepsilon)\sqrt{n}+1$. In terms of continuous Young diagrams, this implies that

$$\omega_{\lambda_n}(2(1-\varepsilon)\sqrt{n}) \geq 2(1-\varepsilon)\sqrt{n}+2;$$
$$\omega_n(2(1-\varepsilon)) \geq 2(1-\varepsilon)+\frac{2}{\sqrt{n}}.$$

However, at $s = 2(1-\varepsilon)$, $\Omega(s) > s$, so for n large enough, $\Omega(2(1-\varepsilon)) \geq 2(1-\varepsilon)+\frac{2}{\sqrt{n}}+\eta$ for some $\eta > 0$ (the LSKV curve $\Omega(s)$ is strictly larger than $|s|$ on $(-2,2)$). Then, if $\|\omega_n - \Omega\|_\infty \leq \eta$, we obtain

$$\omega_n(2(1-\varepsilon)) \geq \Omega(2(1-\varepsilon)) - \eta \geq 2(1-\varepsilon)+\frac{2}{\sqrt{n}}.$$

We conclude that

$$\mathbb{P}[\ell_n \geq 2(1-\varepsilon)\sqrt{n}+1] \geq \mathbb{P}[\|\omega_n - \Omega\|_\infty \leq \eta]$$

for n large enough. By Theorem 4.28, the probability on the right-hand side goes to 1, so the same is true for the probability on the left-hand side. In order to prove the convergence in probability, it remains to see that we also have

$$\mathbb{P}[\ell_n \leq 2(1+\varepsilon)\sqrt{n}] \to 1.$$

To this purpose, we use the upper bound on the expectation of ℓ_n (Proposition 4.9):

$$\begin{aligned} 1 \geq \frac{\mathbb{E}[\ell_n]}{2\sqrt{n}} &\geq (1+\varepsilon)\,\mathbb{P}\left[\frac{\ell_n}{2\sqrt{n}} \geq 1+\varepsilon\right] + \left(1-\frac{\varepsilon}{t}\right)\,\mathbb{P}\left[1+\varepsilon \geq \frac{\ell_n}{2\sqrt{n}} \geq 1-\frac{\varepsilon}{t}\right] \\ &\geq \varepsilon\left(1-\frac{1}{t}\right)\,\mathbb{P}\left[\frac{\ell_n}{2\sqrt{n}} \geq 1+\varepsilon\right] + \left(1-\frac{\varepsilon}{t}\right)\,\mathbb{P}\left[\frac{\ell_n}{2\sqrt{n}} \geq 1-\frac{\varepsilon}{t}\right] \end{aligned}$$

for any parameter $t > 1$. Taking the limsup of the right-hand side and using the first part of the proof, we obtain after rearrangement of the terms:

$$\frac{1}{t-1} \geq \limsup_{n\to\infty}\left(\mathbb{P}\left[\frac{\ell_n}{2\sqrt{n}} \geq 1+\varepsilon\right]\right)$$

for any $t > 1$. Hence, $\mathbb{P}[\ell_n \geq 2(1+\varepsilon)\sqrt{n}]$ goes to 0, and this ends the proof. □

The end of this chapter is devoted to the proof of the LSKV law of large numbers 4.28. We start by giving a criterion for the uniform convergence of a sequence of continuous Young diagrams.

LEMMA 4.30 (Convergence of the geometric observables). *Let $(\omega_n)_{n\in\mathbb{N}}$ and ω_∞ a sequence of continuous Young diagrams with a common support: there exists $C > 0$ such that $\omega_n(s) = \omega_\infty(s) = |s|$ for any $|s| > C$ and any $n \in \mathbb{N}$. We suppose that*

$$\widetilde{p}_k(\omega_n) = k(k-1)\int_{\mathbb{R}} \frac{\omega_n(s)-|s|}{2}\, s^{k-2}\,ds$$

converges to $\widetilde{p}_k(\omega_\infty)$ for any $k \geq 2$. Then, $\|\omega_n - \omega_\infty\|_\infty$ goes to 0.

PROOF. Obviously, it suffices to prove the uniform convergence on $[-C,C]$. Set $\sigma_n(s) = \frac{\omega_n(s)-|s|}{2}$; this is a Lipschitz function with constant 1. Its derivative $\sigma_n'(s)$ gives a signed measure $\sigma_n'(s)\,ds$ on $[-C,C]$. For any polynomial $P \in \mathbb{C}[s]$, we have

$$\int_{\mathbb{R}} \sigma_n'(s)\,P(s)\,ds \to \int_{\mathbb{R}} \sigma_\infty'(s)\,P(s)\,ds;$$

indeed, by integration by parts, $\int_{\mathbb{R}} \sigma_n'(s)\,P(s)\,ds = -\int_{\mathbb{R}} \sigma_n(s)\,P'(s)\,ds$ is a linear combination of the observables $\widetilde{p}_k(\omega_n)$, so by hypothesis it converges to the same observable of the continuous Young diagram ω_∞. Consider now a continuous function F on $[-C,C]$. By the Stone–Weierstrass theorem, $\|F-P\|_{\infty,[-C,C]} \leq \varepsilon$ for some polynomial P, therefore,

$$\begin{aligned}\limsup_{n\to\infty}\left(\int_{\mathbb{R}} \sigma_n'(s)\,F(s)\,ds\right) &\leq 2C\varepsilon + \limsup_{n\to\infty}\left(\int_{\mathbb{R}} \sigma_n'(s)\,P(s)\,ds\right) = 2C\varepsilon + \int_{\mathbb{R}} \sigma_\infty'(s)\,P(s)\,ds \\ &\leq 4C\varepsilon + \int_{\mathbb{R}} \sigma_\infty'(s)\,F(s)\,ds\end{aligned}$$

since the derivatives $\sigma_n'(s)$ and $\sigma_\infty'(s)$ are bounded in absolute value by 1. As this is true for any ε, we obtain that $\limsup_{n\to\infty}(\int_{\mathbb{R}} \sigma_n'(s)\,F(s)\,ds) \leq \int_{\mathbb{R}} \sigma_\infty'(s)\,F(s)\,ds$, and we prove similarly that $\liminf_{n\to\infty}(\int_{\mathbb{R}} \sigma_n'(s)\,F(s)\,ds) \geq \int_{\mathbb{R}} \sigma_\infty'(s)\,F(s)\,ds$. So,

$$\int_{\mathbb{R}} \sigma_n'(s)\,F(s)\,ds \to_{n\to\infty} \int_{\mathbb{R}} \sigma_\infty'(s)\,F(s)\,ds$$

for any continuous function F on $[-C,C]$.

Given $x \in (-C,C)$ and $\varepsilon > 0$, let $F : [-C,C] \to [0,1]$ be a continuous test function equal to 1 before $x-\varepsilon$, and to 0 after x. We have the following estimates:

$$\begin{aligned}\int_{\mathbb{R}} \sigma_n'(s)\,F(s)\,ds &\geq \int_{-C}^{x-\varepsilon} \sigma_n'(s)\,ds - \int_{x-\varepsilon}^{x} 1\,ds \geq \sigma_n(x-\varepsilon) - \varepsilon \geq \sigma_n(x) - 2\varepsilon; \\ \int_{\mathbb{R}} \sigma_n'(s)\,F(s)\,ds &\leq \int_{-C}^{x-\varepsilon} \sigma_n'(s)\,ds + \int_{x-\varepsilon}^{x} 1\,ds \leq \sigma_n(x-\varepsilon) + \varepsilon \leq \sigma_n(x) + 2\varepsilon.\end{aligned}$$

As a consequence, $\limsup_{n\to\infty} \sigma_n(x) - 4\varepsilon \leq \sigma_\infty(x) \leq \liminf_{n\to\infty} \sigma_n(x) + 4\varepsilon$, and since ε is arbitrary, we conclude that $\sigma_n(x) \to \sigma_\infty(x)$. Thus, $(\sigma_n)_{n\in\mathbb{N}}$ converges everywhere to σ_∞, and $(\omega_n)_{n\in\mathbb{N}}$ converges everywhere on $[-C,C]$ to ω_∞. Finally, it is well known that for Lipschitz functions with a uniform constant, the convergence everywhere implies the uniform convergence on compact sets (this is for instance a particular case of the Arzela–Ascoli theorem). □

COROLLARY 4.31 (Criterion of convergence of the shapes of a random sequence of partitions). *Let $(\lambda_n)_{n\in\mathbb{N}}$ be a sequence of random integer partitions with $|\lambda_n| = n$ for every $n \in \mathbb{N}$, and such that the following holds:*

(1) There exists a constant C such that $\mathbb{P}[\lambda_{n,1} \leq C\sqrt{n}] \to 1$ and $\mathbb{P}[\ell(\lambda_n) \leq C\sqrt{n}] \to 1$.

(2) The observables $\widetilde{p}_k(\omega_{\lambda_n})$ satisfy:

$$\forall k \geq 2,\ \frac{\widetilde{p}_k(\omega_{\lambda_n})}{n^{\frac{k}{2}}} \to_{\mathbb{P}} \widetilde{p}_k(\omega)$$

for some shape $\omega \in \mathscr{Y}$.

Then, the shapes of the random integer partitions $(\lambda_n)_{n\in\mathbb{N}}$ converge in probability to ω (in the same sense as in Theorem 4.28).

PROOF. Note that by a change of variables, if $\omega_n = (\omega_{\lambda_n})_{(\frac{1}{n})}$, then

$$\widetilde{p}_k(\omega_n) = \frac{\widetilde{p}_k(\omega_{\lambda_n})}{n^{\frac{k}{2}}}.$$

Let us recall the classical link between convergence in probability and almost sure convergence: given a sequence of random variables $(\omega_n)_{n\in\mathbb{N}}$ in some metric space $\mathscr{Y}$, $(\omega_n)_{n\in\mathbb{N}}$ converges in probability to an element ω if and only if, for any subsequence $(\omega_{n_k})_{k\in\mathbb{N}}$, there is a further subsequence $(\omega_{n_{k_l}})_{l\in\mathbb{N}}$ which converges almost surely to ω. As a consequence, it suffices to prove the following criterion: if $(\omega_n)_{n\in\mathbb{N}}$ is a sequence of random continuous Young diagrams defined on a common probability space and such that

(1') $\omega_n(s) = |s|$ outside $[-C, C]$ with probability 1;

(2') for any $k \geq 2$, $\widetilde{p}_k(\omega_n)$ converges almost surely to $\widetilde{p}_k(\omega)$;

then $(\omega_n)_{n\in\mathbb{N}}$ converges almost surely to ω in the space of continuous Young diagram. However, this follows immediately from the deterministic statement of Lemma 4.30. □

Let us check that the first assumption of the previous corollary is satisfied by the random integer partitions chosen under the Plancherel measures. Note first that the Plancherel measure gives the same weight to an integer partition λ and to its conjugate λ'; this follows immediately from the hook-length formula for $\dim\lambda$. As a consequence, the length $\ell(\lambda_n)$ and the first part $\lambda_{n,1}$ of a random integer partition λ_n chosen according to the Plancherel measure PL_n have the same distribution. So, it suffices to prove that $\lambda_{n,1} \leq C\sqrt{n}$ with very high probability for some constant C. By Proposition 4.8, this is equivalent to the following statement: given σ_n random uniform permutation in $\mathfrak{S}(n)$, the length ℓ_n of a longest increasing subsequence in σ_n is smaller than $C\sqrt{n}$ with very high probability for some constant C. Let $C = 2\mathrm{e}$. Note that given $l \geq 1$, we always have

$$\mathbb{P}[\ell_n = l] \leq \frac{1}{n!}\binom{n}{l}^2 (n-l)! = \frac{1}{l!}\binom{n}{l}.$$

Indeed, in order to construct a permutation $\sigma \in \mathfrak{S}(n)$ with $\ell(\sigma) = l$, one can first choose the l elements of a longest increasing subsequence with length l and their positions in the word of σ ($\binom{n}{l}^2$ possibilities), and then distribute the other elements (less than $(n-l)!$ possibilities). Therefore,

$$\begin{aligned}\mathbb{P}[\ell_n \geq 2\mathrm{e}\sqrt{n}] &\leq \sum_{l=\lfloor 2\mathrm{e}\sqrt{n}\rfloor}^{n} \frac{1}{l!}\binom{n}{l} \leq \sum_{l=\lfloor 2\mathrm{e}\sqrt{n}\rfloor}^{n} \frac{n^l}{(l!)^2} \\ &\leq \sum_{l=\lfloor 2\mathrm{e}\sqrt{n}\rfloor}^{n} (2\pi l)\left(\frac{n\mathrm{e}^2}{l^2}\right)^l \leq (2\pi n)\sum_{l=\lfloor 2\mathrm{e}\sqrt{n}\rfloor}^{n} \frac{1}{2^l} \leq \frac{4\pi n}{2^{\lfloor 2\mathrm{e}\sqrt{n}\rfloor}}\end{aligned}$$

by using Stirling's estimate. The last upper bound clearly goes to 0, whence the result.

In the sequel, given an integer partition λ, we denote $\widetilde{p}_k(\lambda) = \widetilde{p}_k(\omega_\lambda)$. The end of the proof of Theorem 4.28 relies on the following deep result:

THEOREM 4.32 (Geometric and algebraic observables correspond). *We endow the algebra $\mathscr{O}$ with the weight grading:* $\mathrm{wt}(\Sigma_\mu) = |\mu| + \ell(\mu)$.

(1) The weight is a gradation of algebra: $\mathrm{wt}(fg) \leq \mathrm{wt}(f) + \mathrm{wt}(g)$ *for any* $f, g \in \mathscr{O}$.

(2) For any $k \geq 2$, the function $\widetilde{p}_k$ on $\mathfrak{Y}$ can be written as a linear combination of the functions Σ_μ (renormalised character values).

(3) More precisely, $\widetilde{p}_k$ is an observable with total weight k, and its leading term is

$$\widetilde{p}_k = \sum_{\mu \,|\, |\mu|+\ell(\mu)=k} \frac{k^{\downarrow \ell(\mu)}}{\prod_{i\geq 1}(m_i(\mu))!}\, \Sigma_\mu + \textit{terms with strictly lower weight.}$$

We admit this result, which is due to Ivanov and Olshanski; see the references at the end of the chapter for a sketch of proof. The only part which is easy is the first item; its proof is analogous to the proof of Lemma 4.19. Indeed, we can extend the weight to the algebra of partial permutations $\mathscr{P}$ by setting

$$\mathrm{wt}(\sigma, A) = |A| + \text{number of cycles of } \sigma \text{ in } A.$$

Then, it suffices to prove that $\mathrm{wt}(\sigma\tau, A \cup B) \leq \mathrm{wt}(\sigma, A) + \mathrm{wt}(\tau, B)$ for any partial permutations (σ, A) and (τ, B). We decompose the set of cycles $\mathrm{Cyc}(\sigma\tau, A \cup B)$ in three parts:

- the cycles of σ which are entirely contained in $A \setminus B$;
- the cycles of τ which are entirely contained in $B \setminus A$;
- the other cycles, which are not contained in $A \setminus B$ or in $B \setminus A$.

Denote a, b, c the cardinalities of these three sets, so that $|\mathrm{Cyc}(\sigma\tau, A \cup B)| = a + b + c$. The inequality comes from the following claim: given a cycle of $\sigma\tau$ which is counted by c, its support intersects $A \cap B$. Indeed, such a cycle ρ meets by definition both A and B, so there exists an element $x \in A \cap (\text{support of } \rho)$ such that $\sigma\tau(x) \in B \cap (\text{support of } \rho)$. If $x \in A \cap B$, we are done, and otherwise, $x \notin B$, so $\sigma\tau(x) = \sigma(x)$ belongs to $A \cap B$. As a consequence of this discussion, $c \leq |A \cap B|$, and therefore,

$$\begin{aligned}|\mathrm{Cyc}(\sigma\tau, A \cup B)| + |A \cup B| &= a + b + c + |A \cup B| \leq a + |A| + b + |B| \\ &\leq |\mathrm{Cyc}(\sigma, A)| + |A| + |\mathrm{Cyc}(\tau, B)| + |B|.\end{aligned}$$

Let us examine the case where one has equality. We then need that:

- $c = |A \cap B|$: this implies that each cycle of $\sigma\tau$ which is not contained in $A \setminus B$ or in $B \setminus A$ is a cycle of length 1, hence a fixed point of $\sigma\tau$ in $A \cap B$.
- $a = |\mathrm{Cyc}(\sigma, A)|$ and $b = |\mathrm{Cyc}(\tau, B)|$; this is only possible if $A \cap B = \emptyset$.

So, $\mathrm{wt}(\sigma\tau, A \cap B) = \mathrm{wt}(\sigma, A) + \mathrm{wt}(\tau, B)$ if and only if the supports A and B are disjoint. This leads to the following property for the symbols Σ_μ, analogous to Lemma 4.20:

$$\Sigma_\mu\, \Sigma_\nu = \Sigma_{\mu \sqcup \nu} + \text{terms with strictly lower weight.}$$

Indeed, by the previous discussion, the leading term of the product with respect to the weight corresponds to the empty partial pairing.

PROOF OF THEOREM 4.28. We can now use arguments analogous to those used during the proof of Kerov's central limit theorem. Note first that for any observable $f \in \mathscr{O}$,

$$\mathbb{E}_n[f] = O\Big(n^{\frac{\mathrm{wt}(f)}{2}}\Big).$$

Indeed, it suffices to prove this estimate for an observable $f = \Sigma_\mu$: this is trivial if $\mu \neq 1^k$, and if $\mu = 1^k$, then $\mathrm{wt}(\Sigma_{1^k}) = 2k$ and $\mathbb{E}_n[\Sigma_{1^k}] = n^{\downarrow k} = O(n^k)$. We now take the expectation of the expansion of $\widetilde{p}_k$ provided by Theorem 4.32. We get:

$$\mathbb{E}_n[\widetilde{p}_k(\lambda_n)] = \sum_{\mu \,|\, |\mu|+\ell(\mu)=k} \frac{k^{\downarrow \ell(\mu)}}{\prod_{i\geq 1}(m_i(\mu))!}\, \mathbb{E}_n[\Sigma_\mu] + O\Big(n^{\frac{k-1}{2}}\Big).$$

If k is odd, there is no term Σ_{1^j} in the right-hand side, so all the remaining expectations vanish. Thus,

$$\mathbb{E}_n[\widetilde{p}_k(\lambda_n)] = O\Big(n^{\frac{k-1}{2}}\Big) \quad \text{if } k \text{ is odd.}$$

If $k=2j$ is even, the only non-zero expectation is $\mathbb{E}_n[\varSigma_{1^j}]=n^{\downarrow j}$, so

$$\mathbb{E}_n[\widetilde{p}_{2j}(\lambda_n)]=\binom{2j}{j}\,n^{\downarrow j}+O\!\left(n^{j-\frac{1}{2}}\right)=\binom{2j}{j}\,n^{j}+O\!\left(n^{j-\frac{1}{2}}\right).$$

If ω_n is the scaled continuous Young diagram obtained from λ_n, then $\widetilde{p}_k(\omega_n)=n^{-\frac{k}{2}}\,\widetilde{p}_k(\lambda_n)$, so the previous estimates become:

$$\mathbb{E}_n[\widetilde{p}_k(\omega_n)]=\begin{cases}O\!\left(n^{-\frac{1}{2}}\right) & \text{if } k \text{ is odd,}\\ \binom{2j}{j}+O\!\left(n^{-\frac{1}{2}}\right) & \text{if } k=2j \text{ is even.}\end{cases}$$

We now remark that these limiting values are the observables $\widetilde{p}_k(\Omega)$, where Ω is the LSKV curve (see Example 4.27). To end the proof of the theorem, we have to prove the convergence in probability $\widetilde{p}_k(\omega_n)\to_{\mathbb{P}}\widetilde{p}_k(\Omega)$, which is stronger than the convergence in expectation. By the Bienaymé–Chebyshev inequality with the variances, it suffices to prove that we also have $\mathbb{E}_n[(\widetilde{p}_k(\omega_n))^2]\to(\widetilde{p}_k(\Omega))^2$. This follows readily from the fact that the weight is a gradation of algebra, with $\varSigma_\mu\,\varSigma_\nu=\varSigma_{\mu\sqcup\nu}+$ terms with strictly lower weight. Indeed, we have

$$(\widetilde{p}_k)^2=\sum_{\substack{\mu\,|\,|\mu|+\ell(\mu)=k\\ \nu\,|\,|\nu|+\ell(\nu)=k}}\frac{k^{\downarrow\ell(\mu)}\,k^{\downarrow\ell(\nu)}}{\prod_{i\geq 1}(m_i(\mu))!\,(m_i(\nu))!}\,\varSigma_{\mu\sqcup\nu}+\text{terms with weight smaller than } 2k-1,$$

from which it follows that

$$\mathbb{E}_n[(\widetilde{p}_k(\omega_n))^2]=1_{(k=2j\text{ is even})}\binom{2j}{j}^{2}+O\!\left(n^{-\frac{1}{2}}\right)=(\widetilde{p}_k(\Omega))^2+O\!\left(n^{-\frac{1}{2}}\right).$$

The criterion explained in Corollary 4.31 ends then the proof of the law of large numbers. □

References

The Robinson–Schensted algorithm appeared first in [Rob38]; its connections with the longest increasing and decreasing subsequences have been proved by Schensted and Greene in [Sch61; Gre74]. A generalisation of the Robinson–Schensted algorithm to two-line arrays and matrices with non-negative integer entries has been proposed by Knuth in [Knu70]; in this setting, the pair (P,Q) consists in two semistandard tableaux with the same shape. The problem of the longest increasing subsequences in random permutations goes back to the 1960s; see [Ula61], which actually developed the Monte–Carlo method of simulation in order to tackle this problem. In 1972, Hammersley proved that there exists a constant C such that $\ell(\sigma_n)/\sqrt{n}$ converges in probability to C; *cf.* [Ham72]. The value $C=2$ was computed by Logan and Shepp [LS77] and separately by Kerov and Vershik [KV77] in 1977; the two papers are remarkably similar, except that [LS77] only proves that $C\geq 2$, whereas [KV77] claims without proof that $C\leq 2$; the proof based on the Robinson–Schensted algorithm of this second inequality was only published in 1985 in [KV85].

The method of (moments of) observables of Young diagrams is due to Kerov, Olshanski and Vershik; see in particular [KV81; KV86; Ker93; KO94]. The central limit theorem for random character values appeared first in [Ker93]; a full proof is provided by [IO02]. The constructions with partial permutations are due to Ivanov and Kerov, see [IK99]; they make the manipulations of the symbols $\varSigma_\mu$ much easier. A far reaching generalisation of Kerov's central limit theorem is due to Śniady; see [Śni06a; Śni06b], which give general conditions on the spectral measures of representations of symmetric groups that ensure the asymptotic normality of the character values. We should mention that there is a geometric analogue of Theorem 4.15 for the difference of diagrams $\omega_n(s)-\Omega(s)$, where ω_n is the rescaled continuous Young diagram obtained from a random partition $\lambda_n \sim \mathrm{PL}_n$, and where Ω is the LSKV curve. Thus, it can be shown that

$$\sqrt{n}\,(\omega_n(s) - \Omega(s)) \rightharpoonup_{n\to\infty} 1_{s\in(-2,2)} \sum_{k\geq 2} \frac{\xi_k}{\sqrt{k}} \sin\left(k \arccos\left(\frac{s}{2}\right)\right),$$

where the $\xi_{k\geq 2}$ are independent standard Gaussian variables, and where the random series converges in the sense of distributions.

Let us sketch briefly the proof of Theorem 4.32, which we admitted in our discussion. By using the Frobenius–Schur formula for the irreducible characters of the symmetric groups (Theorem 2.23), we can relate the symbols $\varSigma_k(\lambda)$ to the so-called *Frobenius moments* $p_k(\lambda)$, and prove the Wassermann formula:

$$\varSigma_k(\lambda) = [t^{k+1}]\left\{-\frac{1}{k}\left(\prod_{i=1}^{k}\left(1-\left(i-\frac{1}{2}\right)t\right)\right)\exp\left(\sum_{j=1}^{\infty}\frac{p_j(\lambda)\,(1-(1-kt)^{-j})t^j}{j}\right)\right\}.$$

Here, the $p_j(\lambda) = \sum_{i=1}^{s}(a_i)^j - (-b_i)^j$ are the moments of the Frobenius coordinates

$$F_\lambda = (a_1, \ldots, a_s; -b_1, \ldots, -b_s)$$

of the Young diagram of λ; see the beginning of the next chapter for the definition of these coordinates. These Frobenius moments p_j are themselves related to the moments $\widetilde{p}_j$ by a triangular change of basis:

$$\widetilde{p}_k(\lambda) = \sum_{j=0}^{\lfloor\frac{k-1}{2}\rfloor}\binom{k}{2j+1}\frac{p_{k-2j-1}(\lambda)}{2^{2j}}.$$

This relation comes mostly from the combinatorial relation between Frobenius and descent coordinates (see again the beginning of Chapter 5), and from the manipulation of the generating series of these coordinates. The case $k = 2$ of this formula is essentially explained in Chapter 3, Exercise 2. By combining the two formulas above, we can express each symbol $\varSigma_k$ in terms of the functions $\widetilde{p}_j$, and conversely. Finally, we can use the weight grading on $\mathscr{O}$ and the formula from Proposition 4.16 in order to obtain the expansion of $\widetilde{p}_k$ stated in Theorem 4.32. A complete proof appears in [IO02]; see Proposition 3.7 in *loc. cit.*

Exercises

(1) **Longest decreasing subsequences.** Given a semistandard tableau S, we denote $S \leftarrow x$ the semistandard tableau obtained after insertion of x according to the rules of the Robinson–Schensted algorithm. Suppose that the entries of S are all distinct, and that x is not an entry of S. The *reverse Robinson–Schensted insertion* $x \rightarrow S$ is the same algorithm, but where one inserts the letter x column by column instead of row by row.

(a) Suppose given S with distinct entries, and $x \neq y$ letters which are not in S. Show that

$$x \rightarrow (S \leftarrow y) = (x \rightarrow S) \leftarrow y.$$

To this purpose, it will be convenient to distinguish three cases: (i) x is larger than y and than all the entries of S; (ii) y is larger than x and than all all the entries of S; (iii) the largest entry M of S is greater than x and y. In the third case, one can compare $x \rightarrow (S \leftarrow y)$ and $(x \rightarrow S) \leftarrow y$ with the tableaux

$$x \rightarrow (S' \leftarrow y) \quad \text{and} \quad (x \rightarrow S') \leftarrow y,$$

where $S' = S \setminus \{M\}$; and use an induction on the size of S.

(b) Given a permutation $\sigma = \sigma(1)\sigma(2)\cdots\sigma(N)$ in $\mathfrak{S}(N)$, we introduce the reversed permutation $\sigma' = \sigma(N)\sigma(N-1)\cdots\sigma(1)$. Show by induction on N that the two standard tableaux $P(\sigma) = \emptyset \leftarrow \sigma(1) \leftarrow \sigma(2) \leftarrow \cdots \leftarrow \sigma(N)$ and $P(\sigma') = \emptyset \leftarrow \sigma'(1) \leftarrow \sigma'(2) \leftarrow \cdots \leftarrow \sigma'(N)$ are conjugated (obtained from one another by symmetry with respect to the diagonal).

(c) Deduce from the previous question that the number of rows of $P(\sigma)$ is equal to the length of the longest *decreasing* subsequence in σ.

(d) Extend the result of the previous question to the tableaux associated to words: given a word w, the number of rows of $P(w)$ is equal to the length of the longest strictly decreasing subword of w.

(2) **The Schur–Weyl measures.** Denote $\mathrm{SST}(\lambda, N)$ the set of semistandard tableaux with shape λ and letters in $[\![1, N]\!]$.

(a) Prove the identity:

$$N^n = \sum_{\lambda \in \mathfrak{Y}(n)} \mathrm{card}(\mathrm{SST}(\lambda, N))\,\mathrm{card}(\mathrm{ST}(\lambda))$$

for any $N, n \geq 1$.

(b) Consider a vector space $U = \mathbb{C}^N$ with dimension N, and its n-th tensor power $V = (\mathbb{C}^N)^{\otimes n}$: it is the complex vector space which has dimension N^n and which is spanned linearly by the tensors $u_1 \otimes u_2 \otimes \cdots \otimes u_n$ with $u_1, \ldots, u_n \in U$. It is a representation of $\mathfrak{S}(n)$ for the action

$$\sigma \cdot (u_1 \otimes u_2 \otimes \cdots \otimes u_n) = u_{\sigma^{-1}(1)} \otimes u_{\sigma^{-1}(2)} \otimes \cdots \otimes u_{\sigma^{-1}(n)}.$$

We admit that the decomposition in irreducibles of this representation is:

$$V = \sum_{\lambda \in \mathfrak{Y}(n)} \mathrm{card}(\mathrm{SST}(\lambda, N))\, S^\lambda.$$

Reprove the identity of the first question by using this representation theoretic argument.

It is then natural to consider the spectral measure of the representation $V = (\mathbb{C}^N)^{\otimes n}$; this is the so-called *Schur–Weyl measure* with parameters N and n, and we shall denote it $\mathrm{SW}_{N,n}$. Thus,

$$\mathrm{SW}_{N,n}[\lambda] = \frac{\mathrm{card}(\mathrm{SST}(\lambda, N))\,\mathrm{card}(\mathrm{ST}(\lambda))}{N^n}.$$

(c) Prove that the character of the representation V is given by:

$$\mathrm{ch}^V(\sigma_\mu) = N^{\ell(\mu)}$$

for any $\mu \in \mathfrak{Y}(n)$. Deduce from this the following formula for the expectation of the algebraic observables $\varSigma_\mu$:

$$\mathbb{E}_V[\varSigma_\mu(\lambda)] = n^{\downarrow|\mu|}\, N^{\ell(\mu)-|\mu|}$$

for any $\mu \in \mathfrak{Y}$.

(3) **Asymptotic behavior of the Schur–Weyl measures.** We use the same notations as in the previous exercise, and we suppose that $N = \frac{\sqrt{n}}{c}$ with $c > 0$. Then, the asymptotic behavior of $\lambda_n \sim \mathrm{SW}_{N,n}$ is close to the one observed for random partitions under the Plancherel measures; the main difference is that the limiting shape is not anymore the Logan–Shepp–Kerov–Vershik curve. In the sequel, we admit that under $\mathrm{SW}_{N,n}$ with $N = \frac{\sqrt{n}}{c}$, there exists

a constant $C = C(c)$ such that $\mathbb{P}_n[\lambda_{n,1} \leq C\sqrt{n}] \to 1$ and such that $\mathbb{P}[\ell(\lambda_n) \leq C\sqrt{n}] \to 1$ when n grows to infinity (the proof is similar to the one given for Plancherel measures, by replacing permutation words in $\mathfrak{S}(n)$ by words in $[\![1, N]\!]^n$).

(a) Show that with the normalisation $N = \frac{\sqrt{n}}{c}$, we have $\mathbb{E}_n[f(\lambda)] = O(n^{\frac{\mathrm{wt}(f)}{2}})$ for any observable $f \in \mathscr{O}$.

(b) Prove that the rescaled Young diagrams $\omega_n = (\omega_{\lambda_n})_{\frac{1}{n}}$ satisfy:

$$\mathbb{E}_n[\widetilde{p}_k(\omega_n)] \to_{n\to\infty} \sum_{\mu \,|\, |\mu|+\ell(\mu)=k} \frac{k^{\downarrow \ell(\mu)}}{\prod_{i\geq 1}(m_i(\mu))!}\, c^{|\mu|-\ell(\mu)}.$$

(c) We fix $s \geq l \geq 1$. Prove that:

$$\sum_{\substack{|\mu|=s \\ \ell(\mu)=l}} \frac{l!}{\prod_{i\geq 1}(m_i(\mu))!} = \binom{s-1}{l-1}.$$

(Hint: count the number of compositions $s = s_1 + s_2 + \cdots + s_l$ with length l and with the s_i positive integers). Deduce from this combinatorial formula the following expression for the limits of the expected geometric moments:

$$\lim_{n\to\infty} \mathbb{E}_n[\widetilde{p}_k(\omega_n)] = \sum_{i=1}^{\lfloor \frac{k}{2} \rfloor} \frac{k^{\downarrow 2i}}{(k-i)\, i!\, (i-1)!}\, c^{k-2i}.$$

(d) Suppose that there exists a continuous Young diagram Ω_c such that $\widetilde{p}_k(\Omega_c)$ is given by the formula above. Show then that the rescaled Young diagram ω_n obtained from a random integer partition $\lambda_n \sim \mathrm{SW}_{N,n}$ converges in probability to Ω_c.

(e) Suppose $c \in (0, 1)$. Show that the unique continuous Young diagram $\Omega_c(s)$ equal to $|s|$ outside $[c-2, c+2]$, and such that

$$\Omega_c'(s) = \frac{2}{\pi}\arcsin\left(\frac{s+c}{\sqrt{1+sc}}\right)$$

on $[c-2, c+2]$, satisfies the hypothesis of the previous question.

Thus, the random partitions chosen under the Schur–Weyl measures $\mathrm{SW}_{c^{-1}\sqrt{n},n}$ admit a limiting shape Ω_c at least for $c < 1$. The Schur–Weyl measures with parameters $c \geq 1$ also admit limiting shapes, but the exact formula for Ω_c is then a bit more complicated.

(f) Deduce from the previous result an analogue of the law of large numbers $\ell(\sigma_n) \sim 2\sqrt{n}$ when the permutation σ_n is replaced by a random uniform word w_n chosen in $[\![1, N]\!]^n$.

(g) What happens when c goes to 0? Explain intuitively this result.

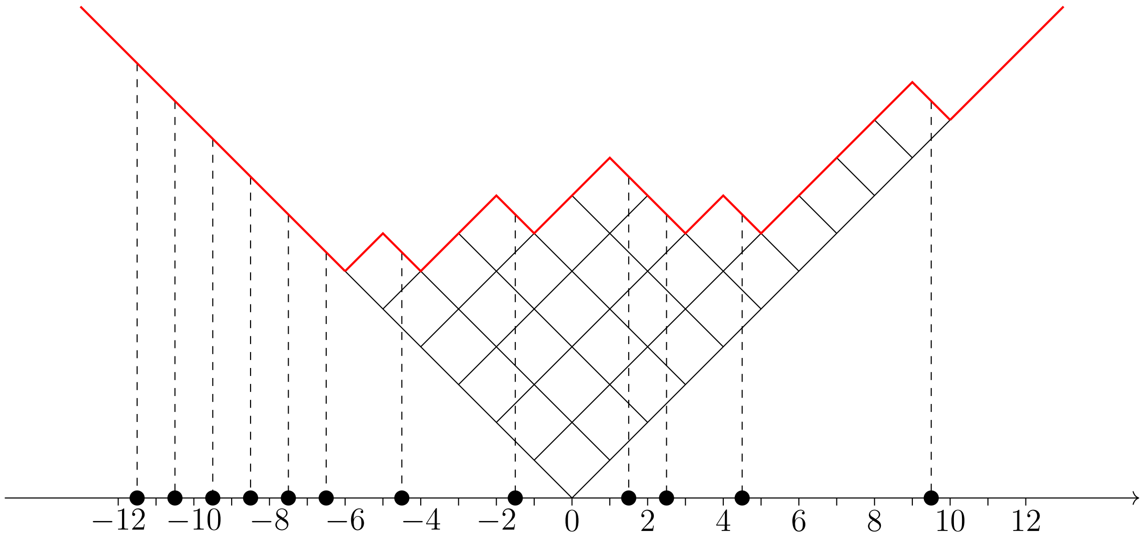
CHAPTER 5

Schur measures and the Tracy–Widom distribution

In this last chapter, we examine the random integer partitions $\lambda_n \sim \mathrm{PL}_n$ with another viewpoint: we are now interested in their *local* aspect, and in the correlations between the sizes of their parts. This problem is better understood by associating to each integer partition a *system of particles* on the real line. Given $\lambda \in \mathfrak{Y}$, we denote $M_\lambda = (\lambda_i - i + \frac{1}{2})_{i \geq 1}$, with by convention $\lambda_r = 0$ if $r > \ell(\lambda)$. This is a decreasing sequence of half-integers in $\mathbb{Z}' = \mathbb{Z} + \frac{1}{2}$, such that

$$M_\lambda \cap \mathbb{Z}'_+ \quad \text{and} \quad \mathbb{Z}'_- \setminus (M_\lambda \cap \mathbb{Z}'_-) \quad \text{are both finite sets with the same cardinality.}$$

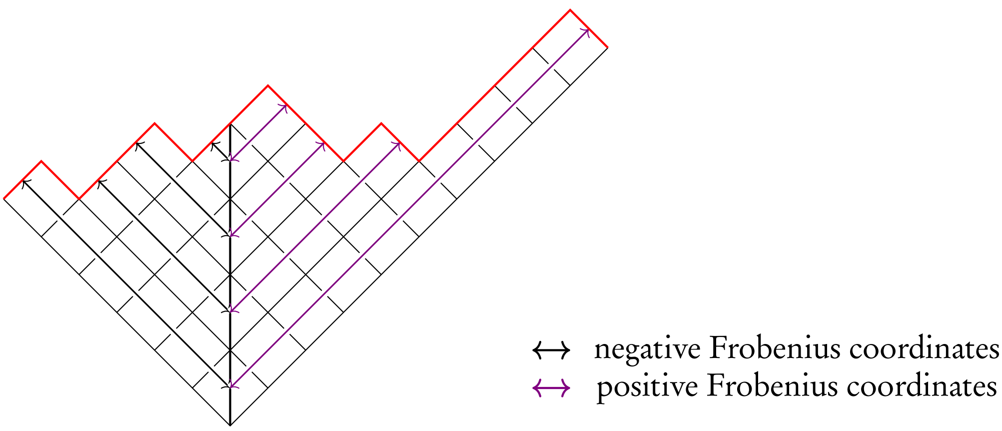
Graphically, the system of particles M_λ corresponds to the descents of the continuous Young diagram of λ: when drawing ω_λ, we can consider it as the union of segments ⟍ and ⟋ with endpoints in $\mathbb{Z} \times \mathbb{N}$, and the *descent coordinates* are the abscissa of the centers of the descending segments.



We have drawn above the continuous Young diagram of $\lambda = (10, 6, 5, 5, 3, 1)$ and its associated system of particles M_λ. The *Frobenius coordinates* of λ are the half-integers in

$$F_\lambda = (\mathbb{Z}'_- \setminus (M_\lambda \cap \mathbb{Z}'_-)) \sqcup (M_\lambda \cap \mathbb{Z}'_+).$$

Graphically, they are obtained by measuring the sizes of the rows and columns of λ *starting from the diagonal*, see the drawing below.



This interpretation explains in particular why $|M_\lambda \cap \mathbb{Z}'_+| = |\mathbb{Z}'_- \setminus (M_\lambda \cap \mathbb{Z}'_-)|$. With $\lambda = (10, 6, 5, 5, 3, 1)$, we obtain

$$F_\lambda = \left(-\frac{11}{2}, -\frac{7}{2}, -\frac{5}{2}, -\frac{1}{2}; \frac{3}{2}, \frac{5}{2}, \frac{9}{2}, \frac{19}{2}\right).$$

If λ is random, then the two random sets M_λ and F_λ are random configurations of elements of $\mathbb{Z}'$, also called *random point processes*. It turns out that these random point processes have a special structure when the law of λ is a Plancherel measure (actually, we shall need to consider Plancherel measures with a Poissonised size n, but this is a small technical detail): we get *determinantal* point processes, whose asymptotic properties when n goes to infinity can be studied by using classical analytical tools (mostly, the saddle-point analysis of certain path integrals in the complex plane). This structure is actually shared by many other probability measures on $\mathfrak{Y} = \bigsqcup_{n\in\mathbb{N}} \mathfrak{Y}(n)$: the so-called *Schur measures*, which are closely related to the representation theory of the infinite symmetric group. In the particular case of the Plancherel measures, this theory allows one to recover the Logan–Shepp–Kerov–Vershik law of large numbers, and to make more precise the estimate $\ell_n \sim 2\sqrt{n}$ of the length of a longest increasing subsequence of a random uniform permutation. Thus, we shall prove that $\ell_n - 2\sqrt{n} \sim n^{1/6}X$, where X is a random variable which follows the so-called *Tracy–Widom distribution*; this limiting law also appears when studying the largest eigenvalues of large random Hermitian matrices

1. An introduction to determinantal point processes

We start by giving a brief survey of the theory of random point processes, which will be the right framework in order to study the local aspects of certain models of random partitions. We shall skip most of the proofs of the results, but give precise references for them at the end of the chapter. Some results around large integer partitions will be stated as the convergence of a rescaled version of the random point process M_λ in $\mathbb{Z}'$ towards a random point process with points in $\mathbb{R}$; therefore, it is convenient to study right away the random point processes with values in a general space $\mathfrak{X}$. In the following, we thus fix a locally compact, separable, complete metrisable space $\mathfrak{X}$; in all the applications, $\mathfrak{X}$ will be a subset of $\mathbb{R}^d$. We denote $\mathscr{C}_{c,+}(\mathfrak{X})$ the set of positive continuous functions on $\mathfrak{X}$ with a compact support. We endow $\mathfrak{X}$ with its Borel σ-field $\mathscr{B}(\mathfrak{X})$, and we consider the set $\mathscr{M}(\mathfrak{X})$ of locally finite positive measures on $\mathfrak{X}$: they are the measures $\mu : \mathscr{B}(\mathfrak{X}) \to \mathbb{R}_+ \sqcup \{+\infty\}$ with $\mu(B) < +\infty$ for any relatively compact Borel subset B. A natural topology on the set $\mathscr{M}(\mathfrak{X})$ is the smallest topology which makes continuous the maps

$$\begin{aligned} \mathscr{M}(\mathfrak{X}) &\to \mathbb{R}_+ \\ \mu &\mapsto \mu(f) = \int_{\mathfrak{X}} f(x)\, \mu(dx) \end{aligned}$$

with $f \in \mathscr{C}_{c,+}(\mathfrak{X})$. Then, the corresponding σ-field of Borel subsets of $\mathscr{M}(\mathfrak{X})$ happens to be the same as the smallest σ-field which makes measurable the maps $\mu \mapsto \mu(B)$ with B relatively compact Borel subset of $\mathfrak{X}$. A *random measure* on $\mathfrak{X}$ is a random variable with values in $\mathscr{M}(\mathfrak{X})$; so, it is a measurable map M from a probability space $(\Omega, \mathscr{F}, \mathbb{P})$ to $\mathscr{M}^{\mathrm{atom}}(\mathfrak{X})$. By definition, all the quantities $M(B)$ with $B \in \mathscr{B}(\mathfrak{X})$ are then random variables with values in $\mathbb{R}_+ \sqcup \{+\infty\}$. A sequence of random measures $(M_n)_{n\in\mathbb{N}}$ converges in distribution to a random measure M if and only if one of the following equivalent assertions is true:

(1) For any $f \in \mathscr{C}_{c,+}(\mathfrak{X})$, $M_n(f)$ converges in law towards $M(f)$.

(2) For any $f \in \mathscr{C}_{c,+}(\mathfrak{X})$, $\mathbb{E}[\mathrm{e}^{-M_n(f)}] \to_{n\to\infty} \mathbb{E}[\mathrm{e}^{-M(f)}]$.

(3) For any relatively compact Borel subsets $B_1, \ldots, B_k$ such that $M(\partial B_1) = \cdots = M(\partial B_k) = 0$ almost surely, the random vectors $(M_n(B_1), \ldots, M_n(B_k))$ converge in distribution towards $(M(B_1), \ldots, M(B_k))$.

In the sequel, we shall only work with *atomic measures* on $\mathfrak{X}$, which are the elements of $\mathscr{M}(\mathfrak{X})$ which take integer values. If μ is an atomic measure on $\mathfrak{X}$, there exists a countable family $(x_i)_{i\in I}$ of points in $\mathfrak{X}$ such that

$$\mu = \sum_{i\in I} \delta_{x_i},$$

and such that for any compact subset K, $\{i \,|\, x_i \in K\}$ is finite. We denote $\mathscr{M}^{\mathrm{atom}}(\mathfrak{X})$ the set of atomic measures on $\mathfrak{X}$; it is a closed subset of $\mathscr{M}(\mathfrak{X})$. Then, a *random point process* on $\mathfrak{X}$ is a random element in $\mathscr{M}^{\mathrm{atom}}(\mathfrak{X})$, so a random measure with values in $\mathscr{M}^{\mathrm{atom}}(\mathfrak{X}) \subset \mathscr{M}(\mathfrak{X})$. The aforementioned criteria of convergence in distribution stay true for a sequence of random point processes, and moreover, if M is a *simple* point process (i.e., $M(\{x\}) \in \{0,1\}$ almost surely for any $x \in \mathfrak{X}$), then one can replace the third item by:

(3') For any relatively compact Borel subsets B such that $M(\partial B) = 0$ almost surely, the random variables $M_n(B)$ converge in distribution towards $M(B)$.

Example 5.1 (Poisson point processes). Consider a space $\mathfrak{X}$ and a measure $\mu \in \mathscr{M}(\mathfrak{X})$. A *Poisson point process* with intensity μ on $\mathfrak{X}$ is a random point process P_μ such that, for any family $(B_a)_{a\in A}$ of disjoint Borel subsets of $\mathfrak{X}$, $(P_\mu(B_a))_{a\in A}$ is a family of independent Poisson variables with parameters $\mu(B_a)$. Any locally finite positive Borel measure on $\mathfrak{X}$ gives rise to a Poisson point process, which is unique in law in $\mathscr{M}^{\mathrm{atom}}(\mathfrak{X})$.

In order to study a random point process M on a space $\mathfrak{X}$, it is natural to consider the joint moments of the associated positive random variables $M(B)$: they describe how many points fall in a given Borel subset B, and the correlations of these cardinalities for distinct Borel subsets $B_1, B_2, \ldots, B_n$. The *factorial moment measures* of M will enable one to encode all these joint moments in a convenient way. For $n \geq 1$, we first define the n-th *factorial power* $M^{\downarrow n}$ of the point process M. Although this is not trivial, given a random point process $M : (\Omega, \mathscr{F}, \mathbb{P}) \to \mathscr{M}^{\mathrm{atom}}(\mathfrak{X})$ on a locally compact polish space, we can actually define on the same probability space some random variables $X_i : (\Omega, \mathscr{F}, \mathbb{P}) \to \mathfrak{X} \sqcup \{\dagger\}$ for $i \geq 1$, such that:

- $X_i = \dagger$ if and only if $M(\mathfrak{X}) < +\infty$ and $i > M(\mathfrak{X})$;
- $M = \sum_{i=1}^{M(\mathfrak{X})} \delta_{X_i}$.

We then define $M^{\downarrow n}$ as the random point process on $\mathfrak{X}^n$ given by

$$M^{\downarrow n} = \sum_{\substack{i_1 \neq i_2 \neq \cdots \neq i_n \\ 1 \leq i_a \leq M(\mathfrak{X})}} \delta_{(X_{i_1}, X_{i_2}, \ldots, X_{i_n})}.$$

This random point process count the sequences $(X_{i_1}, X_{i_2}, \ldots, X_{i_n})$ of length n consisting in distinct points of the random point process M. Here the word "distinct" is a bit misleading, because if $X_i = X_j$ for $i \neq j$ (so, if M is not a *simple* random point process), then the pair (X_i, X_j) is allowed to appear in a sequence $(X_{i_1}, X_{i_2}, \ldots, X_{i_n})$ counted by the factorial power $M^{\downarrow n}$. It is obvious by construction that for any permutation $\sigma \in \mathfrak{S}(n)$, $M^{\downarrow n} = M^{\downarrow n} \circ \sigma$, where $\mathfrak{S}(n)$ acts on $\mathfrak{X}^n$ by permutation of the coordinates. The terminology of factorial power is justified by the following computation: if B is a (relatively) compact subset of $\mathfrak{X}$, then $M(B)$ is almost surely finite, and

$$\begin{aligned} M^{\downarrow n}(B^n) &= \text{number of } n\text{-sequences of distinct points in } B \\ &= M(B)(M(B)-1)\cdots(M(B)-n+1) = (M(B))^{\downarrow n}. \end{aligned}$$

The n-th factorial moment measure of M is the positive Borel measure $\mu_M^{\downarrow n}$ on $\mathfrak{X}^n$ defined by:

$$\mu_M^{\downarrow n}(B_1 \times B_2 \times \cdots \times B_n) = \mathbb{E}\big[M^{\downarrow n}(B_1 \times B_2 \times \cdots \times B_n)\big]\,.$$

In particular, for $n = 1$, $\mu_M^{\downarrow 1} = \mu_M$ is the intensity of the point process M, defined by $\mu_M(B) = \mathbb{E}[M(B)]$.

EXAMPLE 5.2 (Factorial moment measures of a Poisson point process). Consider a Poisson point process P on $\mathfrak{X}$ with intensity μ, and some locally compact subsets $B_1, \ldots, B_n$ in $\mathfrak{X}$. A way to construct the restriction of the Poisson point process P to $B = \bigcup_{a=1}^n B_a$ is as follows: we first take a Poisson random variable N with parameter $\mu(B)$, and we then set

$$P_{|B} = \sum_{i=1}^{N} \delta_{X_i},$$

where the X_i's are independent random variables in B with law $\frac{\mu(\cdot)}{\mu(B)}$, and are independent of N. Let us then compute the n-th factorial moment measure of P. By invariance by action of $\mathfrak{S}(n)$ and by additivity, it suffices to compute $\mu_P^{\downarrow n}(B_1 \times \cdots \times B_n)$ when

$$B_1 \times B_2 \times \cdots \times B_n = (C_1)^{n_1} \times \cdots \times (C_k)^{n_k},$$

with $n = n_1 + \cdots + n_k$ and with the C_j disjoint. Suppose first that $k = 1$. We then have:

$$\mu_P^{\downarrow n}(B^n) = \mathbb{E}[P^{\downarrow n}(B^n)] = \mathbb{E}[N^{\downarrow n}] = \sum_{j=n}^{\infty} \frac{\mathrm{e}^{-\mu(B)}\,(\mu(B))^j\, j^{\downarrow n}}{j!} = (\mu(B))^n.$$

In the general case, we have $P^{\downarrow n}((C_1)^{n_1} \times \cdots \times (C_k)^{n_k}) = P^{\downarrow n_1}((C_1)^{n_1}) \times \cdots \times P^{\downarrow n_k}((C_k)^{n_k})$, and the random point processes $P_{|C_1}, \ldots, P_{|C_k}$ are independent, so

$$\begin{aligned}\mu_P^{\downarrow n}(B_1 \times B_2 \times \cdots \times B_n) &= \mathbb{E}\left[\prod_{j=1}^{k} P^{\downarrow n_j}((C_j)^{n_j})\right] \\ &= \prod_{j=1}^{k} \mathbb{E}\big[P^{\downarrow n_j}((C_j)^{n_j})\big] = \prod_{j=1}^{k} (\mu(C_j))^{n_j} = \prod_{a=1}^{n} \mu(B_a).\end{aligned}$$

We conclude that $\mu_P^{\downarrow n} = \mu^{\otimes n}$. This identity encodes the independence of the restrictions of the Poisson point process P to disjoint subsets.

The factorial powers and the factorial moment measures are related to the computation of (the expectation of) products

$$\prod_{i \in I} (1 + f(X_i)),$$

where $f : \mathfrak{X} \to \mathbb{C}$ is some bounded measurable function with compact support, and the random point process M is given by $M = \sum_{i \in I} \delta_{X_i}$, I being a random interval $[\![1, M(\mathfrak{X})]\!] \subset \mathbb{N}$. Indeed, since f is supported by a compact, we can assume without loss of generality that $M(\mathfrak{X})$ is finite almost surely, and then,

$$\prod_{i \in I} (1 + f(X_i)) = 1 + \sum_{n=1}^{\infty} \left(\sum_{1 \leq i_1 < i_2 < \cdots < i_n \leq M(\mathfrak{X})} \prod_{a=1}^{n} f(X_{i_a}) \right) = 1 + \sum_{n=1}^{\infty} \frac{1}{n!}\, M^{\downarrow n}(f^{\otimes n}).$$

In many cases, given a random point process M on a locally compact polish space $\mathfrak{X}$, there exists a reference Radon measure (locally finite Borel positive measure) λ on $\mathfrak{X}$ such that for any $n \geq 1$, the factorial moment measure $\mu_M^{\downarrow n}$ is absolutely continuous with respect to $\lambda^{\otimes n}$. In this situation, the density

$$\rho_n(x_1, \ldots, x_n) = \frac{d(\mu_M^{\downarrow n})}{d(\lambda^{\otimes n})}(x_1, \ldots, x_n)$$

is called the n-th *correlation function* of the random point process. In particular, if $\mathfrak{X} = \mathbb{Z}^d$ or $\mathfrak{X} = \mathbb{R}^d$, we shall take for reference measure λ the counting measure or the standard Lebesgue measure.

THEOREM 5.3 (Lenard). *Let M be a random point process on a locally compact polish space $\mathfrak{X}$. We suppose that there exists a reference Radon measure λ on $\mathfrak{X}$ such that the correlation functions ρ_n with respect to M and λ are all well-defined. Then:*

(1) The correlation functions are symmetric: for any $\sigma \in \mathfrak{S}(n)$,

$$\rho_n(x_1, \dots, x_n) = \rho_n(x_{\sigma(1)}, \dots, x_{\sigma(n)}).$$

(2) The correlation functions are positive, in the following sense: for any set $(\phi_0, \phi_1, \dots, \phi_N)$ of compactly supported measurable functions $\phi_k : \mathfrak{X}^k \to \mathbb{R}$ such that

$$\phi_0 + \sum_{k=1}^{N} \sum_{\substack{i_1 \neq i_2 \neq \cdots \neq i_k \\ 1 \leq i_a \leq N}} \phi_k(x_{i_1}, \dots, x_{i_k}) \geq 0,$$

we also have

$$\phi_0 + \sum_{k=1}^{N} \int_{\mathfrak{X}^k} \phi_k(x_1, \dots, x_k)\, \rho_k(x_1, \dots, x_k)\, \lambda^{\otimes k}(dx_1 \cdots dx_k) \geq 0.$$

Conversely, given a family of locally integrable positive functions $(\rho_n)_{n \in \mathbb{N}}$ which satisfy the two conditions above, one can define a random point process M on $\mathfrak{X}$ with these correlation functions. This random point process M is unique in law if and only if the random variables $M(B)$ with $B \in \mathscr{B}(\mathfrak{X})$ are determined by their moments.

The determinantal point processes are the random point processes whose correlation functions write as

$$\rho_n(x_1, \dots, x_n) = \det(K(x_i, x_j))_{1 \leq i,j \leq n}$$

for some adequate kernel K which does not depend on n. Although we could work in a more general setting, we shall only consider in the sequel the point processes associated to *locally trace class Hermitian kernels*. We consider as before a locally compact separable complete metric space $\mathfrak{X}$ endowed with a Radon measure λ, We denote $\mathscr{L}^2(\mathfrak{X}, \lambda)$ the Hilbert space of square-integrable functions on $\mathfrak{X}$, which is separable. We recall that a *trace class operator* on a separable Hilbert space H is a bounded linear operator $A : H \to H$ such that, given an orthonormal basis $(e_i)_{i \in I}$ of H, we have

$$\sum_{i \in I} \left\langle e_i \,\middle|\, (A^* A)^{1/2}(e_i) \right\rangle_H < +\infty$$

the index set I being finite if H is finite-dimensional, and infinite countable if H is infinite-dimensional. The *trace* of the operator A is then defined by the aboslutely convergent series

$$\mathrm{tr}(A) = \sum_{i \in I} \langle e_i \mid A(e_i) \rangle_H\,;$$

this does not depend on the choice of the orthonormal basis. In the same setting, a bounded linear operator $A : H \to H$ is called *Hilbert–Schmidt* if $A^* A$ is a trace class operator; equivalently,

$$(\|A\|_{\mathrm{HS}})^2 = \mathrm{tr}(A^* A) = \sum_{i \in I} \langle A(e_i) \mid A(e_i) \rangle_H < +\infty$$

for any orthonormal basis $(e_i)_{i \in I}$ of H. We have the following inclusions of ideals of the Banach algebra of bounded linear operators on H:

$$\{\text{finite rank}\} \subset \{\text{trace class}\} \subset \{\text{Hilbert–Schmidt}\} \subset \{\text{compact}\}.$$

If $\mathscr{K} : \mathscr{L}^2(\mathfrak{X}, \lambda) \to \mathscr{L}^2(\mathfrak{X}, \lambda)$ is a Hilbert–Schmidt operator, then one can show that there exists a unique kernel $K \in \mathscr{L}^2(\mathfrak{X}^2, \lambda^{\otimes 2})$ such that

$$(\mathscr{K}f)(x) = \int_{\mathfrak{X}} K(x,y)\, f(y)\, \lambda(dy).$$

The map $K \mapsto \mathscr{K}$ is an isometry between $\mathscr{L}^2(\mathfrak{X}^2, \lambda^{\otimes 2})$ and the space of Hilbert–Schmidt operators $\mathrm{HS}(\mathscr{L}^2(\mathfrak{X}, \lambda))$. Moreover, if $\mathfrak{X} = \mathbb{R}^d$, λ is the Lebesgue measure, $\mathscr{K}$ is a trace class operator and if K is continuous at (x,x) for λ-almost any x, then

$$\mathrm{tr}(\mathscr{K}) = \int_{\mathfrak{X}} K(x,x)\, \lambda(dx).$$

Given a trace class operator A on a separable Hilbert space H, we can also define the *Fredholm determinant* $\det(I+A)$. Denote $\bigwedge^k H$ the k-th exterior power of H, which is the Hilbert completion of the algebraic k-th exterior power for the scalar product

$$\langle v_1 \wedge v_2 \wedge \cdots \wedge v_k \mid w_1 \wedge w_2 \wedge \cdots \wedge w_k \rangle_{\bigwedge^k H} = \det(\langle v_i \mid w_j \rangle_H)_{1 \leq i,j \leq k}.$$

It is again a separable Hilbert space, and if $(e_i)_{i \in I}$ is an orthonormal basis of H with $I \subset \mathbb{N}$, then $(e_{i_1} \wedge e_{i_2} \wedge \cdots \wedge e_{i_k})_{i_1 < i_2 < \cdots < i_k \in I}$ is an orthonormal basis of $\bigwedge^k H$. Now, the k-th exterior power of A defined by extension of the rule

$$\left(\bigwedge\nolimits^k A\right)(v_1 \wedge v_2 \wedge \cdots \wedge v_k) = A(v_1) \wedge A(v_2) \wedge \cdots \wedge A(v_k)$$

is again a trace class operator. Indeed, given A of trace class, consider an orthonormal basis $(e_i)_{i \in I}$ of diagonalisation of the compact self-adjoint operator $|A| = (A^*A)^{1/2}$, with $|A|(e_i) = \lambda_i\, e_i$. Each λ_i is a non-negative real number, and $\|A\|_1 = \mathrm{tr}(|A|) = \sum_{i \in I} \lambda_i$. We then have:

$$\begin{aligned}
&\sum_{i_1 < i_2 < \cdots < i_k} \left\langle e_{i_1} \wedge e_{i_2} \wedge \cdots \wedge e_{i_k} \,\middle|\, \left|\bigwedge\nolimits^k A\right| (e_{i_1} \wedge e_{i_2} \wedge \cdots \wedge e_{i_k}) \right\rangle_{\bigwedge^k H} \\
&= \sum_{i_1 < i_2 < \cdots < i_k} \langle e_{i_1} \wedge e_{i_2} \wedge \cdots \wedge e_{i_k} \mid |A|(e_{i_1}) \wedge |A|(e_{i_2}) \wedge \cdots \wedge |A|(e_{i_k}) \rangle_{\bigwedge^k H} \\
&= \sum_{i_1 < i_2 < \cdots i_k} \lambda_{i_1} \lambda_{i_2} \cdots \lambda_{i_k} = \frac{1}{k!} \sum_{i_1 \neq i_2 \neq \cdots \neq i_k} \lambda_{i_1} \lambda_{i_2} \cdots \lambda_{i_k} \leq \frac{1}{k!} (\mathrm{tr}(|A|))^k,
\end{aligned}$$

so $\bigwedge^k A$ is of trace class, with $\left\| \bigwedge^k A \right\|_1 \leq \frac{(\|A\|_1)^k}{k!}$. The Fredholm determinant is defined by:

$$\det(I + A) = 1 + \sum_{k=1}^{\infty} \mathrm{tr}\left(\bigwedge\nolimits^k A\right);$$

by the previous calculation, the series is convergent and $|\det(I+A)| \leq \mathrm{e}^{\|A\|_1}$. We recover the traditional determinant when H is finite-dimensional. On the other hand, if $\mathscr{K}$ is a trace class operator on $\mathscr{L}^2(\mathfrak{X}, \lambda)$ associated to a kernel $K \in \mathscr{L}^2(\mathfrak{X}^2, \lambda^2)$, then its Fredholm determinant is given by:

$$\det(I + \mathscr{K}) = 1 + \sum_{k=1}^{\infty} \mathrm{tr}\left(\bigwedge\nolimits^k \mathscr{K}\right) = 1 + \sum_{k=1}^{\infty} \frac{1}{k!} \int_{\mathfrak{X}^k} \det(K(x_i, x_j))_{1 \leq i,j \leq k}\, \lambda(dx_1) \cdots \lambda(dx_k).$$

This is the *Fredholm formula*.

In the following, we consider a linear operator $\mathscr{K} : \mathscr{L}^2(\mathfrak{X}, \lambda) \to \mathscr{L}^2(\mathfrak{X}, \lambda)$ which is:

(1) *Hermitian and non-negative*: for $f, g \in \mathscr{L}^2(\mathfrak{X}, \lambda)$, $\langle f \mid \mathscr{K}(g) \rangle_{\mathscr{L}^2(\mathfrak{X},\lambda)} = \langle \mathscr{K}(f) \mid g \rangle_{\mathscr{L}^2(\mathfrak{X},\lambda)}$ and $\langle f \mid \mathscr{K}(f) \rangle_{\mathscr{L}^2(\mathfrak{X},\lambda)} \geq 0$.

(2) *locally of trace class*: for any relatively compact subset $B \subset \mathfrak{X}$, $\mathscr{K}_B = 1_B \mathscr{K} 1_B$ is a trace class operator on $\mathscr{L}^2(B, \lambda_{|B})$.

This implies the existence of a unique measurable function $K : \mathfrak{X}^2 \to \mathbb{C}$ such that:

- $K(x,y) = \overline{K(y,x)}$.
- $\det(K(x_i,x_j))_{1\leq i,j\leq n} \geq 0$ for $\lambda^{\otimes n}$-almost any family of points $x_1,\ldots,x_n$.
- $\int_{B^2} |K(x,y)|^2\,\lambda(dx)\,\lambda(dy) < +\infty$ for any relatively compact subset $B \subset \mathfrak{X}$.

THEOREM 5.4 (Determinantal point process associated to a Hermitian kernel). *Suppose that K is the kernel of a Hermitian non-negative locally trace class operator $\mathscr{K}$ on $\mathscr{L}^2(\mathfrak{X},\lambda)$.*

(1) The spectrum of $\mathscr{K}$ (set of complex numbers such that $zI - \mathscr{K}$ is not invertible) is included in $[0,1]$ if and only if, for any relatively compact subset $B \subset \mathfrak{X}$, the spectrum of the restricted operator $\mathscr{K}_B$ is included in $[0,1]$.

(2) If the condition $\mathrm{Spec}(\mathscr{K}) \subset [0,1]$ is satisfied, then there exists a random point process M on $\mathfrak{X}$ whose correlation functions with respect to λ are given by:

$$\rho_n(x_1,\ldots,x_n) = \det(K(x_i,x_j))_{1\leq i,j\leq n}.$$

This random point process is unique in law; equivalently, all the random variables $M(B)$ with B relatively compact are determined by their moments (actually, they have subexponential tails).

(3) Conversely, suppose given a determinantal point process M whose correlations are associated to the kernel K of a Hermitian non-negative locally trace class operator $\mathscr{K}$ on $\mathscr{L}^2(\mathfrak{X},\lambda)$. Then, $\mathrm{Spec}(\mathscr{K}) \subset [0,1]$.

This result is due to Soshnikov. The determinantal point processes can be defined under weaker assumptions (for instance, with non-Hermitian kernels), but in the sequel we shall stick to the setting of Theorem 5.4. Note that in the first item of the theorem, a restricted operator $\mathscr{K}_B$ is trace class hence compact, so

$$\mathrm{Spec}(\mathscr{K}_B) \cup \{0\} = \{\text{eigenvalues of } \mathscr{K}_B\} \cup \{0\}.$$

The non-zero eigenvalues of $\mathscr{K}_B$ are involved in a precise description of the law of the random variable $M(B)$, M being a determinantal point process with kernel K. More generally, let us consider a family $B_1,\ldots,B_n$ of disjoint relatively compact subsets in $\mathfrak{X}$, and let us compute the joint generating function of the random variables $M(B_1),\ldots,M(B_n)$. We have:

$$\begin{aligned}\mathbb{E}\left[\prod_{a=1}^n (z_a)^{M(B_a)}\right] &= \mathbb{E}\left[\prod_{a=1}^n \sum_{m_a=0}^{M(B_a)} \binom{M(B_a)}{m_j} (z_a-1)^{m_a}\right] \\ &= 1 + \sum_{m=1}^{\infty} \sum_{m_1+\cdots+m_n=m} \mathbb{E}\left[\prod_{a=1}^n (M(B_a))^{\downarrow m_a}\right] \prod_{a=1}^n \frac{(z_a-1)^{m_a}}{(m_a)!}.\end{aligned}$$

The convergence of these series is ensured by the following identity: for any composition $m = m_1+\cdots+m_n$, setting $B = \bigsqcup_{a=1}^n B_a$, we have

$$\begin{aligned}\mathbb{E}\left[\prod_{a=1}^n (M(B_a))^{\downarrow m_a}\right] &= \mathbb{E}[M^{\downarrow m}((B_1)^{m_1}\times\cdots\times(B_n)^{m_n})] \\ &= \int_{(B_1)^{m_1}\times\cdots\times(B_n)^{m_n}} \det(K(x_i,x_j))_{1\leq i,j\leq m}\,\lambda(dx_1)\cdots\lambda(dx_m) \\ &= m!\,\mathrm{tr}\left((1_B\mathscr{K}1_{B_1})^{\wedge m_1}\wedge(1_B\mathscr{K}1_{B_2})^{\wedge m_2}\wedge\cdots\wedge(1_B\mathscr{K}1_{B_n})^{\wedge m_n}\right) \\ &\leq (\mathrm{tr}(\mathscr{K}_B))^m.\end{aligned}$$

If $a = a_j$ is the index of the set B_a corresponding to the variable x_j, then the same computation shows that

$$\begin{aligned}
&\mathbb{E}\left[\prod_{a=1}^{n}(z_a)^{M(B_a)}\right] \\
&= 1 + \sum_{\substack{m_1+\cdots+m_n=m\\ m\geq 1}} \int_{\mathfrak{X}^m} \det(1_B(x_i)\,(z_{a_j}-1)\,K(x_i,x_j)\,1_{B_{a_j}}(x_j))_{1\leq i,j\leq m}\,\frac{\lambda(dx_1)\cdots\lambda(dx_m)}{(m_1)!\cdots(m_n)!} \\
&= 1 + \sum_{m=1}^{\infty}\frac{1}{m!}\int_{\mathfrak{X}^m}\det\left(\sum_{a=1}^{n} 1_B(x_i)\,(z_{a_j}-1)\,K(x_i,x_j)\,1_{B_{a_j}}(x_j)\right)_{1\leq i,j\leq m}\lambda(dx_1)\cdots\lambda(dx_m).
\end{aligned}$$

The quantity that one obtains is a Fredholm determinant:

$$\mathbb{E}\left[\prod_{a=1}^{n}(z_a)^{M(B_a)}\right] = \det\left(I + \sum_{a=1}^{n} 1_B\,(z_a-1)\,\mathscr{K}\,1_{B_a}\right).$$

Let us consider in particular the case where $n = 1$ and the trace class self-adjoint non-negative operator $\mathscr{K}_B$ has a countable family of eigenvalues $(\lambda_{B,i})_{i\in I}$ with $0 \leq \lambda_{B,i} \leq 1$ for any i. The Fredholm determinant is then given by:

$$\mathbb{E}[z^{M(B)}] = \det(I + (z-1)\mathscr{K}_B) = \prod_{i\in I}(1+(z-1)\lambda_{B,i}).$$

This is precisely the generating series of the random variable $X = \sum_{i\in I}\mathrm{Ber}(\lambda_{B,i})$, where all the Bernoulli variables are assumed to be independent. This random series converges almost surely by the two-series Kolmogorov criterion, since $\mathrm{tr}(\mathscr{K}_B) = \sum_{i\in I}\lambda_{B,i} < +\infty$ by hypothesis. We thus have:

PROPOSITION 5.5 (Marginales of a determinantal point process). *Consider a determinantal point process M associated to a Hermitian non-negative locally trace class operator with kernel K. For any relatively compact subset $B \subset \mathfrak{X}$, if $(\lambda_{B,i})_{i\in I}$ is the collection of eigenvalues of $\mathscr{K}_B$, then the law of $M(B)$ is the law of a random series of independent Bernoulli variables with parameter $\lambda_{B,i}$.*

The following remarks give further details on the connection between a determinantal point process M and the spectral properties of its associated locally trace class Hermitian kernel K.

REMARK 5.6 (Density of particles). The case $n = 1$ of Proposition 5.5 gives for any relatively compact subset B:

$$\mathbb{E}[M(B)] = \sum_{i\in I}\lambda_{B,i} = \mathrm{tr}(\mathscr{K}_B).$$

If $\mathfrak{X} \subset \mathbb{R}^d$ and the Hermitian kernel K is a continuous function, then this trace is given by

$$\mathbb{E}[M(B)] = \int_B K(x,x)\,\lambda(dx).$$

REMARK 5.7 (Total number of particles). If $\mathscr{K} : \mathscr{L}^2(\mathfrak{X},\lambda) \to \mathscr{L}^2(\mathfrak{X},\lambda)$ is the operator associated to K, then one can define the trace of the whole operator $\mathscr{K}$ by taking the supremum of the traces of the restricted operators $\mathscr{K}_B$:

$$\mathrm{tr}(\mathscr{K}) = \left(\int_{\mathfrak{X}} K(x,x)\,\lambda(dx)\right) \in \mathbb{R}_+ \sqcup \{+\infty\}.$$

Then, it is easy to deduce from Proposition 5.5 that $M(\mathfrak{X}) = +\infty$ with probability 0 if $\mathrm{tr}(\mathscr{K}) < +\infty$, and with probability 1 if $\mathrm{tr}(\mathscr{K}) = +\infty$. In the first case, we have:

- $\mathbb{P}[M(\mathfrak{X}) \leq n] = 1$ if and only if $\mathrm{rank}(\mathscr{K}) \leq n$. If $\mathrm{rank}(\mathscr{K}) = n$, then there exists a family of n orthonormal functions $\psi_1, \ldots, \psi_n$ in $\mathscr{L}^2(\mathfrak{X}, \lambda)$ and eigenvalues $\lambda_1, \ldots, \lambda_n$ in $(0, 1]$ such that
$$K(x,y) = \sum_{i=1}^{n} \lambda_i \, \overline{\psi_i(x)} \, \psi_i(y)$$
$\lambda^{\otimes 2}$ almost everywhere.
- $\mathbb{P}[M(\mathfrak{X}) = n] = 1$ if and only if $\mathrm{rank}(\mathscr{K}) = n$ and $\mathscr{K}$ is the orthogonal projection on a vector space with rank n. Then, in the decomposition above, all the λ_i's are equal to 1. Thus, the determinantal point processes with a fixed number of points are naturally associated to orthogonal projections, and in the applications this usually leads to computations with certain families of orthogonal polynomials.

Remark 5.8 (Gap probabilities). A particular case of the Fredholm formula for the generating series $\mathbb{E}[z^{M(B)}]$ is with $z = 0$. We are then evaluating the *gap probability* to not having any point of M in B:

$$\mathbb{P}[M(B) = 0] = \det(I - \mathscr{K}_B) = 1 + \sum_{m=1}^{\infty} \frac{(-1)^m}{m!} \int_{B^m} \det(K(x_i, x_j))_{1 \leq i,j \leq m} \, \lambda(dx_1) \cdots \lambda(dx_m).$$

The formula extends readily to any subset $B \subset \mathfrak{X}$ such that $\mathscr{K}_B$ is still a trace class operator. This formula will play an essential role in the study of extremal points of determinantal point processes. On the other hand, it implies the following elementary fact: if the reference measure λ has no atom, then a determinantal point process M with Hermitian locally trace class operator associated to a kernel K is always a simple random point process (atomic measure which is almost surely without multiple point). Indeed, consider a relatively compact subset B, and let us evaluate the expected number $X(B, \varepsilon)$ of ordered pairs of points of M that fall in B and are at distance smaller than ε. This is given by:

$$\mathbb{E}[X(B,\varepsilon)] = \int_{B^2} \det \begin{pmatrix} K(x,x) & K(x,y) \\ K(y,x) & K(y,y) \end{pmatrix} 1_{d(x,y) \leq \varepsilon} \, \lambda(dx) \, \lambda(dy).$$

As ε goes to zero, the locally integrable correlation function $\rho_2(x, y)$ yields an integral which goes to 0. Since $\mathbb{E}[X(B, \varepsilon)] \geq \mathbb{P}[X(B, \varepsilon) > 0] \geq \mathbb{P}[M$ has a multiple point in $B]$, we can conclude that M is a simple random point process.

To close this section, let us state and prove a limiting result regarding sequences of determinantal point processes.

Proposition 5.9 (Convergence of determinantal point processes). *Suppose that M is a determinantal point process on $(\mathfrak{X}, \lambda)$ with a locally bounded Hermitian kernel $K(x, y)$, and that $(M_N)_{N \in \mathbb{N}}$ is a sequence of determinantal point processes with Hermitian kernels $K_N(x, y)$. If $K_N(x, y) \to K(x, y)$ locally uniformly, then M_N converges in law to M as N goes to infinity.*

Proof. By the discussion at the beginning of this section, we have to prove that for any family of locally compact measurable subsets $B_1, \ldots, B_k \subset \mathfrak{X}$ with $\mu(\partial B_1) = \cdots = \mu(\partial B_k) = 0$ almost surely, we have the convergence in law

$$(M_N(B_1), \ldots, M_N(B_k)) \rightharpoonup_{N \to \infty} (M(B_1), \ldots, M(B_k)).$$

The convergence of the kernels implies the local uniform convergence of all the correlations functions. However, if the correlation functions converge locally uniformly, then by dominated convergence the joint moments of the vectors $(M_N(B_1), \ldots, M_N(B_n))$ converge towards the joint moments of the vector $(M(B_1), \ldots, M(B_n))$. As the law of this random vector is determined by its moments, this implies the convergence in law. $\square$

2. The Thoma simplex

We are now going to describe a large family of probability measures on $\mathfrak{Y} = \bigsqcup_{N\geq 0} \mathfrak{Y}(N)$ such that the associated random point processes M_λ and F_λ on $\mathbb{Z}'$ are determinantal point processes with explicit kernels. This family of *Schur measures* was identified by Okounkov in the 90's, and its definition is related to an older result of classification of the extremal characters of the infinite symmetric group $\mathfrak{S}(\infty)$. In this section, we present this classification result (Theorem 5.22) and we give a partial proof of it. We start by extending a bit our Definition 4.11 of spectral measure. Suppose given a finite group G, and a function $f : G \to \mathbb{C}$ which is invariant by conjugation: $f(hgh^{-1}) = f(g)$ for any $g, h \in G^2$ (equivalently, $f(gh) = f(hg)$). Then, viewed as an element of $\mathbb{C}G$, f is central, so it is a linear combination of the irreducible characters of G:

$$f = \sum_{\lambda \in \widehat{G}} c_\lambda\, \chi^\lambda, \quad \text{with } \chi^\lambda(\cdot) = \frac{\mathrm{ch}^\lambda(\cdot)}{\dim \lambda}, \quad \text{and the } c_\lambda \text{ in } \mathbb{C}.$$

LEMMA 5.10. *The coefficients c_λ of the expansion above form a probability measure on $\widehat{G}$ if and only if the following conditions are satisfied:*

(1) We have $f(e_G) = 1$.

(2) For any finite family $(g_1, \ldots, g_n)$ of elements of G, the matrix $(f(g_i g_j^{-1}))_{1\leq i,j\leq n}$ is Hermitian and non-negative definite.

We then say that f is a normalised trace on G, and that $(c_\lambda)_{\lambda\in\widehat{G}}$ is its spectral measure. We shall then rewrite $c_\lambda = \mathbb{P}_f[\lambda]$.

PROOF. The first condition is equivalent to the fact that the sum of the coefficients c_λ is equal to one: indeed,

$$f(e_G) = \sum_{\lambda\in\widehat{G}} c_\lambda = 1.$$

We claim that the second condition is equivalent to the fact that the coefficients c_λ are real and non-negative. If the c_λ are real, then any matrix $M = M(g_1, g_2, \ldots, g_n) = (f(g_i g_j^{-1}))_{1\leq i,j\leq n}$ is Hermitian, because

$$f(g_j g_i^{-1}) = \sum_{\lambda\in\widehat{G}} c_\lambda\, \chi^\lambda(g_j g_i^{-1}) = \sum_{\lambda\in\widehat{G}} c_\lambda\, \overline{\chi^\lambda(g_i g_j^{-1})} = \overline{f(g_i g_j^{-1})}.$$

Conversely, if all these matrices are Hermitian, then by taking $n = 2$, $g_1 = e_G$ and $g_2 = g$, we see that $f(g) = \overline{f(g^{-1})}$ for any $g \in G$. This is only possible if the coefficients of f in the basis $(\chi^\lambda)_{\lambda\in\widehat{G}}$ of $\mathrm{Z}(\mathbb{C}G)$ are real.

It remains to prove that all the matrices $M(g_1, \ldots, g_n)$ are non-negative definite if and only if all the coefficients c_λ are non-negative. However, a Hermitian matrix with size n is non-negative definite if and only if, for any vector $v \in \mathbb{C}^n$, $vMv^* \geq 0$. Given $v = (v_1, \ldots, v_n)$, one has

$$\begin{aligned} vMv^* &= \sum_{1\leq i,j\leq n} v_i\, f(g_i g_j^*)\, \overline{v_j} = \sum_{\lambda\in\widehat{G}} \frac{c_\lambda}{\dim \lambda} \left(\sum_{1\leq i,j\leq n} v_i \overline{v_j}\, \mathrm{tr}(\rho^\lambda(g_i g_j^{-1})) \right) \\ &= \sum_{\lambda\in\widehat{G}} \frac{c_\lambda}{\dim\lambda}\, \mathrm{tr}(\hat{v}(\lambda)\, (\hat{v}(\lambda))^*) \end{aligned}$$

where $v = \sum_{i=1}^n v_i g_i \in \mathbb{C}G$. Since n and v are arbitrary, by using the isomorphism $\mathbb{C}G = \bigoplus_{\lambda\in\widehat{G}} \mathrm{End}(V^\lambda)$, we get that for any $\lambda \in \widehat{G}$ and any matrix $P \in \mathrm{End}(V^\lambda)$, $\frac{c_\lambda}{\dim\lambda}\, \mathrm{tr}(PP^*) \geq 0$. Obviously, since $\mathrm{tr}(PP^*) \geq 0$, this happens if and only if all the coefficients c_λ are non-negative. $\square$

Example 5.11. If $f = \chi^V$ is the normalised character of a finite-dimensional representation V of G, then it is a non-negative normalised trace on G, with spectral measure $\mathbb{P}_V$ given by Definition 4.11. The set of normalised traces on a finite group G is actually the closure of the set of normalised characters in the finite-dimensional space $\mathbb{C}G$. Indeed, given a normalised trace $f = \sum_{\lambda \in \widehat{G}} c_\lambda \, \chi^\lambda$, we can approximate the coefficients c_λ by rational numbers of the form

$$c_{\lambda,\mathrm{approx}} = \frac{n_\lambda \, \dim \lambda}{\sum_{\rho \in \widehat{G}} n_\rho \, \dim \rho},$$

and then, f is approximated by the normalised character of the representation $V = \bigoplus_{\lambda \in \widehat{G}} n_\lambda \, V^\lambda$.

If G is an *infinite* group, then its representation theory is usually much more complicated than the finite case of Chapter 1. However, we can still consider the normalised traces on G: they are the functions $f : G \to \mathbb{C}$ such that $f(gh) = f(hg)$ and such that the two conditions from Lemma 5.10 are satisfied. In this more general setting, the role of the irreducible characters is played by the so-called *extremal characters*: they are the normalised traces τ that cannot be written as a linear combination $t\,\tau_1 + (1-t)\,\tau_2$ with $t \in (0,1)$ and $\tau_1 \neq \tau_2$ normalised traces.

Definition 5.12 (Extremal characters and central measures). *Let χ be an extremal character of the infinite symmetric group $\mathfrak{S}(\infty)$. The family of central measures associated to it is the family of spectral measure $(\mathbb{P}_{\chi_{|\mathfrak{S}(n)}})_{n \in \mathbb{N}}$. Thus, for any $n \in \mathbb{N}$ and any $\sigma \in \mathfrak{S}(n)$,*

$$\chi(\sigma) = \sum_{\lambda \in \mathfrak{Y}(n)} \mathbb{P}_{\chi,n}[\lambda] \, \chi^\lambda(\sigma),$$

with $\mathbb{P}_{\chi,n} = \mathbb{P}_{\chi_{|\mathfrak{S}(n)}}$.

Example 5.13 (Plancherel measures are central). Consider the map $\tau : \sigma \in \mathfrak{S}(\infty) \mapsto 1_{(\sigma = \mathrm{id}_{\mathbb{N}^*})}$. We have $\tau(gh) = 1_{(g=h^{-1})} = \tau(hg)$, and $\tau(\mathrm{id}_{\mathbb{N}^*}) = 1$. Moreover, for any finite permutations $g_1, g_2, \ldots, g_n$, the associated matrix is $M(g_1,\ldots,g_n) = (1_{(g_i = g_j)})_{1 \leq i,j \leq n}$, so it is conjugated by a permutation matrix to a block diagonal matrix whose blocks are of the form

$$J_r = \begin{pmatrix} 1 & \cdots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \cdots & 1 \end{pmatrix}.$$

These blocks are obviously non-negative definite (the spectrum of J_r is $\{0^{r-1}, r^1\}$), so τ is a normalised trace on $\mathfrak{S}(\infty)$. We shall prove in a moment that τ is an extremal character. The associated central measures are the Plancherel measures PL_n of the symmetric groups; indeed, the restriction of τ to $\mathfrak{S}(n)$ is the normalised character of its regular representation, see the proof of Corollary 4.13.

Given a family $(\mathbb{P}_n)_{n \in \mathbb{N}}$ of probability measures on the sets $\mathfrak{Y}(n)$, there is a simple criterion in order to know if they come from a normalised trace of $\mathfrak{S}(\infty)$ (by taking the spectral measures of the restrictions to the finite symmetric groups $\mathfrak{S}(n)$). Suppose that χ is a normalised trace on $\mathfrak{S}(\infty)$, and consider its spectral measures $\mathbb{P}_n = \mathbb{P}_{\chi_{|\mathfrak{S}(n)}}$. For any $n \in \mathbb{N}$ and any permutation $\sigma \in \mathfrak{S}(n)$, we have:

$$\chi(\sigma) = \sum_{\lambda \in \mathfrak{Y}(n)} \mathbb{P}_n[\lambda] \, \chi^\lambda(\sigma) = \sum_{\Lambda \in \mathfrak{Y}(n+1)} \mathbb{P}_{n+1}[\Lambda] \, \chi^\Lambda(\sigma).$$

However, by the branching rule 2.31, $(\mathrm{ch}^\Lambda)_{|\mathfrak{S}(n)} = \sum_{\lambda \,|\, \lambda \nearrow \Lambda} \mathrm{ch}^\lambda$, so:

$$\sum_{\lambda \in \mathfrak{Y}(n)} \mathbb{P}_n[\lambda]\, \chi^\lambda(\sigma) = \sum_{\substack{\lambda \in \mathfrak{Y}(n) \\ \Lambda \in \mathfrak{Y}(n+1) \\ \lambda \nearrow \Lambda}} \mathbb{P}_{n+1}[\Lambda]\, \frac{\dim \lambda}{\dim \Lambda}\, \chi^\lambda(\sigma).$$

Since the irreducible characters form a linear basis of the space of conjugacy-invariant functions on $\mathfrak{S}(n)$, we conclude that for any $n \in \mathbb{N}$ and any integer partition $\lambda \in \mathfrak{Y}(n)$,

$$\frac{\mathbb{P}_n[\lambda]}{\dim \lambda} = \sum_{\Lambda \,|\, \lambda \nearrow \Lambda} \frac{\mathbb{P}_{n+1}[\Lambda]}{\dim \Lambda}.$$

The equation above can be restated as follows. Consider the Young graph $\mathfrak{Y} = \bigsqcup_{n \in \mathbb{N}} \mathfrak{Y}(n)$, with a directed edge $\lambda \to \Lambda$ for each pair of integer partitions $(\lambda \nearrow \Lambda)$. A function $m : \mathfrak{Y} \to \mathbb{R}$ is said *harmonic* if $m(\lambda) = \sum_{\Lambda \,|\, \lambda \nearrow \Lambda} m(\Lambda)$ for any integer partition λ. Then:

Proposition 5.14 (Spectral measures and harmonic functions). *The spectral measures of normalised traces on $\mathfrak{S}(\infty)$ are in bijection with the non-negative harmonic functions on the Young graph normalised with $m(\emptyset) = 1$. The bijection is provided by the following formula:*

$$\forall n \in \mathbb{N},\ \forall \lambda \in \mathfrak{Y}(n),\ \mathbb{P}_n[\lambda] = (\dim \lambda)\, m(\lambda).$$

Proof. We have proved above that a family of spectral measures associated to a normalised trace of $\mathfrak{S}(\infty)$ gives rise to a (non-negative, normalised) harmonic function on the Young graph. Conversely, suppose given such a harmonic function m. Then, the formula

$$\chi(\sigma) = \sum_{\lambda \in \mathfrak{Y}(n)} m(\lambda)\, \mathrm{ch}^\lambda(\sigma) \quad \forall \sigma \in \mathfrak{S}(n) \subset \mathfrak{S}(\infty)$$

defines a function on $\mathfrak{S}(\infty)$, and the definition above does not depend on the choice of n, because of the property of harmonicity. In particular, given a family $(\sigma_1, \ldots, \sigma_r)$ of finite permutations, we can compute the matrix $(\chi(\sigma_i \sigma_j^{-1}))_{1 \leq i,j \leq r}$ by placing ourselves in a sufficiently large finite symmetric group $\mathfrak{S}(n)$, and the result is a non-negative definite Hermitian matrix, because $\chi_{|\mathfrak{S}(n)}$ is a positive linear combination of irreducible characters (Lemma 5.10). The property $\chi(\sigma\tau) = \chi(\tau\sigma)$ follows from the same argument, since it is true in any finite symmetric group $\mathfrak{S}(n)$. Finally, since $m(\emptyset) = 1$, $\chi(\mathrm{id}_{\mathbb{N}^*}) = 1$. So, from a non-negative harmonic function on $\mathfrak{Y}$, we can indeed recover a normalised trace on $\mathfrak{S}(\infty)$. □

Among the harmonic functions on $\mathfrak{Y}$, a particular role is played by the *specialisations* of the ring of symmetric functions. Consider any morphism of rings $\psi : \mathrm{Sym} \to \mathbb{R}$; note that $\psi(c) = c$ for any constant $c \in \mathbb{R}$. Instead of denoting the images $\psi(f)$, it is convenient to think of ψ as a *virtual alphabet* of variables X, and to write $\psi(f) = f(X)$. As a particular case, consider a summable family $X = \{x_1, x_2, \ldots\}$ of real numbers. Then,

$$p_k(X) = \sum_{i=1}^{\infty} (x_i)^k$$

converges for any $k \geq 1$, so $p_\lambda(X) = \prod_{k \geq 1} (p_k(X))^{m_k(\lambda)}$ is well defined for any integer partition λ. As $(p_\lambda)_{\lambda \in \mathfrak{Y}}$ is a linear basis of Sym, this enables us to define $f(X)$ for any $f \in \mathrm{Sym}$, thereby obtaining a morphism of rings from Sym to $\mathbb{R}$. Note that all the specialisations $f \in \mathrm{Sym} \mapsto f(X) \in \mathbb{R}$ are not obtained by this way. Indeed, for any choice of real numbers $(a_k)_{k \geq 1}$, if we set $p_k(X) = a_k$, then there is a unique morphism of rings from Sym to $\mathbb{R}$ which extends this definition: it suffices to set

$$p_\lambda(X) = \prod_{k \geq 1} (a_k)^{m_k(\lambda)},$$

and to extend this definition by linearity to the whole ring Sym. Consider for instance the *exponential alphabet* E, which is the specialisation of Sym defined by $p_1(E) = 1$ and $p_k(E) = 0$ for any $k \geq 2$. This morphism of rings does not come from a summable family $\{x_1, x_2, \ldots\}$: if it were the case, we would have

$$0 = p_2(E) = \sum_{i=1}^{\infty} (x_i)^2,$$

so $x_i = 0$ for any index i, and then $p_1(E) = \sum_{i=1}^{\infty} x_i = 0$.

A specialisation $f \mapsto f(X)$ of Sym is called non-negative normalised if $p_1(X) = 1$ and $s_\lambda(X) \geq 0$ for any $\lambda \in \mathfrak{Y}$. Note then that $m(\lambda) = s_\lambda(X)$ is a normalised non-negative harmonic function on the Young graph: indeed, by the Pieri rule,

$$m(\lambda) = m(\lambda)\, m(\Box) = (s_\lambda\, s_1)(X) = \sum_{\Lambda \,|\, \lambda \nearrow \Lambda} s_\Lambda(X) = \sum_{\Lambda \,|\, \lambda \nearrow \Lambda} m(\Lambda).$$

It turns out that the specialisations of the ring of symmetric functions are precisely the extremal harmonic functions.

THEOREM 5.15 (Kerov–Vershik ring theorem). *A harmonic function $m : \mathfrak{Y} \to \mathbb{R}$ which is non-negative and normalised is extremal in the convex set of such functions if and only if $m(\lambda) = s_\lambda(X)$ for some non-negative normalised specialisation X of the ring of symmetric functions.*

COROLLARY 5.16. *The extremal characters of $\mathfrak{S}(\infty)$ are in bijection with the non-negative normalised specialisations of the ring of symmetric functions.*

PROOF. Any function $m : \mathfrak{Y} \to \mathbb{R}$ gives rise to a linear function $\psi : \mathrm{Sym} \to \mathbb{R}$, by setting $\psi(\sum_{\lambda \in \mathfrak{Y}} c_\lambda\, s_\lambda) = \sum_{\lambda \in \mathfrak{Y}} c_\lambda\, m(\lambda)$. The function m is normalised non-negative harmonic if and only if:

(1) $\psi(1) = 1$;

(2) $\psi(s_\lambda) \geq 0$ for any Schur function s_λ;

(3) $\psi(f\, s_1) = \psi(f)$ for any $f \in \mathrm{Sym}$.

The ring theorem states that m is extremal if and only if $\psi(fg) = \psi(f)\psi(g)$ for any symmetric functions f and g (morphism of rings). Because of the third item above, if m is harmonic and ψ is the associated linear function, it is convenient to consider ψ as a linear map from the quotient ring $A = \mathrm{Sym}/(s_1 - 1)$ to $\mathbb{R}$. Then, the second item can be restated more abstractly as follows: if J is the cone in Sym spanned by the positive linear combinations of Schur functions, and K is its image by the projection map $\pi_A : \mathrm{Sym} \to A$, then ψ is non-negative on the cone K. In order to prove the theorem, we shall use the two following properties of the cone $K \subset A$:

(C1) The cone K is stable by multiplication. It suffices to prove that J is stable by multiplication, since $K = \pi_A(J)$ and π_A is a morphism of rings. Equivalently, we have to prove that a product of two Schur functions is a positive linear combination of Schur functions. However, by the Frobenius–Schur theorem, this is equivalent to the fact that in $\mathrm{R}(\mathfrak{S})$, the product of two irreducible representations S^λ of $\mathfrak{S}(m)$ and S^ν of $\mathfrak{S}(n)$ is a positive linear combination of irreducible representations of $\mathfrak{S}(m+n)$: this is obvious, since $S^\lambda \times S^\nu$ is a (reducible) representation of this group.

(C2) For any $b \in K$, there exists $\varepsilon > 0$ such that $1_A - \varepsilon b$ belongs to K. Notice first that the set K' of elements b with this property is a subcone of K. Indeed, it is obviously stable by multiplication by a positive scalar, and if $1_A - \varepsilon_1 b_1 \in K$ and $1_A - \varepsilon_2 b_2 \in K$, then we can assume without generality $\varepsilon_1 \geq \varepsilon_2$; in this case,

$$1_A - \frac{\varepsilon_2}{2}(b_1 + b_2) = \frac{1}{2}(1_A - \varepsilon_1 b_1) + \frac{1}{2}(1_A - \varepsilon_2 b_2) + \frac{\varepsilon_1 - \varepsilon_2}{2}\, b_1$$

belongs to K, so $b_1 + b_2 \in K'$. As a consequence, in order to prove that $K' = K$, it suffices to establish that for any Schur function s_λ, $b = \pi_A(s_\lambda)$ belongs to K'. However, if $n = |\lambda|$, then we saw during the proof of Proposition 2.28 that $(s_1)^n = (p_1)^n = \sum_{\mu \in \mathfrak{Y}(n)} (\dim \mu)\, s_\mu$, so by taking the image by π_A, we obtain:

$$1_A = \sum_{\mu \in \mathfrak{Y}(n)} (\dim \mu)\, \pi_A(s_\mu) = b + \pi_A(j) \quad \text{for some } j \in J.$$

This ends the proof of this second property. In terms of the partial order on A defined by the cone K, it can be restated as follows: for any $b \in K$, there exists $\varepsilon > 0$ such that $1_A \succeq_K \varepsilon b$.

Consider an extremal non-negative normalised harmonic function m, and the associated linear map $\psi : A \to \mathbb{R}$. We want to prove that $\psi(ab) = \psi(a)\psi(b)$ for any $a, b \in A$. As J spans linearly Sym, $K = \pi_A(J)$ spans linearly A, and we can assume without loss of generality that $a, b \in K$, and by (C2), that $1_A - b \in K$. Then,

$$\psi(a) = \psi(ab) + \psi(a(1_A - b)),$$

and by Property (C1), the right-hand side is the sum of the images by ψ of two elements of K. Suppose first that $\psi(b)\psi(1_A - b) \neq 0$. Then, we rewrite the decomposition above as:

$$\psi(a) = \psi(b)\, \psi_b(a) + \psi(1_A - b)\, \psi_{1_A - b}(a),$$

where ψ_b and $\psi_{1_A - b}$ are the two linear maps defined by $\psi_b(a) = \frac{\psi(ab)}{\psi(b)}$ and $\psi_b(a) = \frac{\psi(a(1_A - b))}{\psi(1_A - b)}$. These two linear maps correspond to normalised non-negative harmonic functions, so by extremality of m, $\psi = \psi_b = \psi_{1_A - b}$. The identity $\psi = \psi_b$ amounts to $\psi(ab) = \psi(a)\psi(b)$. It remains to treat the case where $\psi(b) = 0$ or $\psi(1_A - b) = 0$. Let us suppose for instance that $\psi(b) = 0$; then, $a \preceq_K \varepsilon\, 1_A$ for some $\varepsilon > 0$, so $ab \preceq_K \varepsilon\, b$ for some $\varepsilon > 0$. It follows that $\psi(ab) \leq \varepsilon\, \psi(b) = 0$, and that $\psi(ab) = 0$. In particular, we have again $\psi(ab) = \psi(a)\, \psi(b)$. We conclude that ψ is indeed a morphism of rings when m is extremal.

To establish the converse implication, we shall use a form of the Choquet theorem: any non-negative normalised harmonic function m can be written as an integral

$$m = \int_{\mathcal{E}} n\, \mu_m(dn),$$

where $\mathcal{E}$ is the set of extremal harmonic functions, and μ_m is a probability measure on this set (in general, μ_m is not unique). The measure μ_m is a Borel measure with respect to the (compact, metrisable) topology on the set $\mathcal{H}$ of non-negative normalised harmonic functions which comes from the embedding

$$m \in \mathcal{H} \mapsto (m(\lambda))_{\lambda \in \mathfrak{Y}} \in [0,1]^{\mathfrak{Y}}.$$

The inequality $m(\lambda) \leq 1$ for any $\lambda \in \mathfrak{Y}$ comes from the fact that $\pi_A(s_\lambda) \preceq_K 1_A$ for any integer partition. We are going to prove that when m is a specialisation of Sym, the probability measure μ_m is concentrated on a unique extremal point n. In the following, we write $\psi_m(s_\lambda) = m(\lambda)$, and similarly for ψ_n and n. Given $\lambda \in \mathfrak{Y}$ and m harmonic function associated to a specialisation ψ_m, we consider the random variable $X_\lambda \in [0,1]$ whose law is the image of the law μ_m by the map

$$n \in \mathcal{E} \mapsto n(\lambda) \in [0,1].$$

By construction, $\mathbb{E}[X_\lambda] = \int_{\mathcal{E}} n(\lambda)\mu_m(dn) = m(\lambda)$ (this is true for any harmonic function m). Now, if m comes from a specialisation and if we write $(s_\lambda)^2 = \sum_\mu c_\mu\, s_\mu$, then

$$\mathbb{E}[(X_\lambda)^2] = \int_{\mathcal{E}} (\psi_n(s_\lambda))^2\, \mu_m(dn) = \int_{\mathcal{E}} \psi_n((s_\lambda)^2)\, \mu_m(dn) = \sum_\mu c_\mu \int_{\mathcal{E}} \psi_n(s_\mu)\, \mu_m(dn)$$

since extremal harmonic functions are multiplicative. Reworking the right-hand side of this formula, we get:

$$\sum_\mu c_\mu \int_{\mathcal{E}} n(\mu)\, \mu_m(dn) = \sum_\mu c_\mu\, \mathbb{E}[X_\mu] = \sum_\mu c_\mu\, m(\mu) = \psi_m((s_\lambda)^2) = (\psi_m(s_\lambda))^2 = (m(\lambda))^2.$$

We conclude that the variance of X_λ is equal to 0, and therefore that X_λ is the random variable almost surely equal to $m(\lambda)$. As this is true for any λ, the random variable n chosen under μ_m is a constant function in $\mathcal{E}$, so the law μ_m is concentrated on one point and m belongs to $\mathcal{E}$. $\square$

By the previous discussion, the classification of the extremal characters of the infinite symmetric group is equivalent to the classification of all the non-negative specialisations of the ring of symmetric functions. This classification was obtained separately by Edrei and by Thoma in the 50's. To describe these non-negative specialisations, it is convenient to introduce the *Miwa parameters* of a specialisation $f \in \mathrm{Sym} \mapsto f(X) \in \mathbb{R}$: they are the real numbers

$$t_k = \frac{p_k(X)}{k},$$

and the collection $(t_k)_{k\geq 1}$ entirely determines the specialisation X.

LEMMA 5.17. *If $(t_k^X)_{k\geq 1}$ and $(t_k^{X'})_{k\geq 1}$ are two families of Miwa parameters associated to non-negative specialisations X and X', then the family of sums $(t_k^X + t_k^{X'})_{k\geq 1}$ is also associated to a non-negative specialisation of* Sym.

PROOF. This is a clever application of the Cauchy identity 2.16. Given two alphabets $X = \{x_1, x_2, \ldots\}$ and $Y = \{y_1, y_2, \ldots\}$, we have:

$$\begin{aligned}\sum_{\lambda \in \mathfrak{Y}} s_\lambda(X)\, s_\lambda(Y) &= \prod_{i,j\geq 1} \frac{1}{1 - x_i y_j} = \exp\left(\sum_{i,j\geq 1} -\log(1 - x_i y_j)\right)\\ &= \exp\left(\sum_{i,j,k\geq 1} \frac{(x_i y_j)^k}{k}\right) = \exp\left(\sum_{k\geq 1} \frac{p_k(X)\, p_k(Y)}{k}\right).\end{aligned}$$

As a consequence, if $(t_k^X)_{k\geq 1}$ is the family of Miwa parameters associated to a specialisation X of Sym, then

$$(X \text{ is non-negative}) \iff \left(\text{the expansion of } \exp\left(\sum_{k\geq 1} t_k^X\, p_k(Y)\right) \text{ is Schur non-negative}\right).$$

In this equivalence, the exponential of the right-hand side lives in an algebra slightly larger than $\mathrm{Sym}[Y]$: it belongs to the algebra of symmetric functions with possibly an infinite degree (formal infinite linear combinations of functions $p_\lambda(Y)$ or $s_\lambda(Y)$ without restriction on the degree). We can nonetheless consider products of such elements, because the product of symmetric functions is graded, and because for every degree d, the dimension of the space of symmetric functions homogeneous with degree d is finite. Now, suppose that $(t_k^X)_{k\geq 1}$ and $(t_k^{X'})_{k\geq 1}$ are Miwa parameters of non-negative specialisations. Then, the specialisation associated to the family $(t_k^X + t_k^{X'})_{k\geq 1}$, which we shall denote $X+X'$, gives a generalised symmetric function

$$\exp\left(\sum_{k\geq 1}(t_k^X+t_k^{X'})\,p_k(Y)\right)=\exp\left(\sum_{k\geq 1}t_k^X\,p_k(Y)\right)\exp\left(\sum_{k\geq 1}t_k^{X'}\,p_k(Y)\right)$$

which is a product of Schur non-negative series. But the product of two Schur functions is a positive linear combination of Schur functions, so the left-hand side is again a Schur non-negative series. Thus, $X+X'$ is again a non-negative specialisation of Sym. □

Let us now describe three natural families of non-negative specialisations of Sym.

Example 5.18 (Specialisations with non-negative variables). Suppose that $X=\{x_1,x_2,\ldots\}$ is a specialisation of Sym associated to a summable family of non-negative real numbers; up to permutation, we can assume $x_1\geq x_2\geq\cdots$, and set $x_n=0$ for n large enough if the sequence only contains a finite number of positive terms. Then, X yields a non-negative specialisation of Sym. Indeed, since non-negative specialisations are stable by addition of Miwa parameters, and also obviously by taking limits of sequences of Miwa parameters, it suffices to prove the result when $X=\{x_1\}$ consists of a single positive real number. However, in this case, $t_k^X=\frac{(x_1)^k}{k}$, and

$$\exp\left(\sum_{k\geq 1}t_k^X\,p_k(Y)\right)=\exp\left(\sum_{k=1}^{\infty}\frac{p_k(Y)}{k}\,(x_1)^k\right)=\sum_{k=0}^{\infty}h_k(Y)\,(x_1)^k,$$

as seen during the proof of Lemma 2.5. Since $h_k(Y)=s_k(Y)$, the right-hand side is a Schur non-negative series, so the specialisation X is non-negative. So, any non-increasing, non-negative and summable sequence $(x_n)_{n\geq 1}$ yields a non-negative specialisation of Sym; it is normalised if and only if $p_1(X)=1$.

Example 5.19 (Conjugates of non-negative specialisations). In the exercises of Chapter 2, we introduced an involutive isomorphism of Sym related *via* the Frobenius–Schur isomorphism to the operation

$$\theta : S\in \mathrm{R}(\mathfrak{S}(n))\mapsto \varepsilon_n\otimes S\in \mathrm{R}(\mathfrak{S}(n)),$$

ε_n being the sign representation of $\mathfrak{S}(n)$, and the tensor product being the internal tensor product of representations of $\mathfrak{S}(n)$. A consequence of the *dual* Jacobi–Trudy formula is the following explicit description of this map:

- The linear map $\theta:\mathrm{R}(\mathfrak{S})\to\mathrm{R}(\mathfrak{S})$ sends an irreducible representation S^λ to the irreducible representation $S^{\lambda'}$, λ' being the integer partition conjugate of λ.
- By using the Frobenius–Schur isomorphism, we transport θ to an involution of the algebra of symmetric functions. Then, θ is an isomorphism of algebras, and:

$$\theta(s_\lambda)=s_{\lambda'}\qquad;\qquad \theta(p_k)=(-1)^{k-1}\,p_k\quad\forall k\geq 1.$$

If X is a non-negative specialisation of Sym, then the specialisation $\overline{X}$ defined by $f(\overline{X})=(\theta(f))(X)$ is again non-negative: $s_\lambda(\overline{X})=s_{\lambda'}(X)\geq 0$. The Miwa parameters of this new specialisation are related to the Miwa parameters of X by:

$$t_k^{\overline{X}}=\frac{p_k(\overline{X})}{k}=(-1)^{k-1}\,\frac{p_k(X)}{k}=(-1)^{k-1}\,t_k^X.$$

In particular, given a non-increasing, non-negative and summable sequence $(x_n)_{n\geq 1}$, we can define a non-negative specialisation $\overline{X}$ of Sym by setting:

$$p_k(\overline{X})=(-1)^{k-1}\sum_{n=1}^{\infty}(x_n)^k\quad\text{for any } k\geq 1.$$

EXAMPLE 5.20 (Exponential alphabet). Consider the exponential alphabet E, which is the specialisation given by $p_k(E) = 1_{(k=1)}$. It is a non-negative specialisation: by the Frobenius formula, if $\lambda \in \mathfrak{Y}(n)$, then

$$s_\lambda(E) = \sum_{\mu \in \mathfrak{Y}(n)} \frac{p_\mu(E)}{z_\mu} \operatorname{ch}^\lambda(\mu) = \frac{p_{1^n}(E)}{z_{1^n}} \operatorname{ch}^\lambda(1^n) = \frac{\dim \lambda}{n!} \geq 0.$$

The associated extremal character of $\mathfrak{S}(\infty)$ admits for spectral measures:

$$\mathbb{P}_n[\lambda] = (\dim \lambda)\, s_\lambda(E) = \frac{(\dim \lambda)^2}{n!} = \mathrm{PL}_n[\lambda].$$

Hence, the Plancherel measures are indeed central measures; they are associated to the extremal character of $\mathfrak{S}(\infty)$ given by $\chi(\sigma) = 1_{(\sigma = \mathrm{id}_{\mathbb{N}^*})}$.

DEFINITION 5.21 (Thoma simplex). *The Thoma simplex is the (compact) set Ω of pairs $(A, B) = ((\alpha_n)_{n \geq 1}, (\beta_n)_{n \geq 1})$ of non-increasing sequences of non-negative real numbers, such that*

$$\sum_{n=1}^{\infty} \alpha_n + \sum_{n=1}^{\infty} \beta_n = 1 - \gamma \leq 1.$$

To any element $(A, B) \in \Omega$, we can associate a normalised non-negative specialisation of Sym, by considering the sum of specialisations (in the sense of Lemma 5.17) $A + \overline{B} + \gamma E$. We have:

$$p_1(A + \overline{B} + \gamma E) = p_1(A) + p_1(\overline{B}) + \gamma p_1(E) = \sum_{n=1}^{\infty} \alpha_n + \sum_{n=1}^{\infty} \beta_n + \gamma = 1$$

and for $k \geq 2$,

$$p_k(A + \overline{B} + \gamma E) = p_k(A) + p_k(\overline{B}) + \gamma p_k(E) = \sum_{n=1}^{\infty} (\alpha_n)^k + (-1)^{k-1} \sum_{n=1}^{\infty} (\beta_n)^k.$$

THEOREM 5.22 (Edrei, Thoma). *The only normalised non-negative specialisations of the algebra* Sym *are those associated to pairs (A, B) in the Thoma simplex. Moreover, two distinct pairs (A, B) and (A', B') give rise to distinct specialisations.*

COROLLARY 5.23. *The set of extremal characters of $\mathfrak{S}(\infty)$ is parametrised bijectively by the Thoma simplex.*

We admit this classification; a proof of it is presented in the exercises at the end of the Chapter, and it relies on the analogue of Theorem 4.32 with respect to the *degree* gradation of the algebra of observables $\mathscr{O}$ from Chapter 4 (instead of the *weight* gradation).

3. The Schur measures

We can now define the Schur measures, which are probability measures on the Young graph $\mathfrak{Y} = \bigsqcup_{n=0}^{\infty} \mathfrak{Y}(n)$. Let us remark that if X is a non-negative specialisation of Sym, then:

- either $p_1(X) = 0$, in which case the specialisation is the projection from Sym to the subalgebra of degree 0 symmetric function (constants). Indeed, for any $n \geq 1$, $0 = p_{1^n}(X) = \sum_{\lambda \in \mathfrak{Y}(n)} (\dim \lambda)\, s_\lambda(X)$; since $s_\lambda(X) \geq 0$, we get $s_\lambda(X) = 0$ for any λ with $|\lambda| = n \geq 1$.
- or, $p_1(X) > 0$. We can then renormalise the specialisation by setting

$$s_\lambda(X_{\mathrm{norm}}) = \frac{s_\lambda(X)}{(p_1(X))^{|\lambda|}};$$

  X_{norm} is then a non-negative normalised specialisation of Sym.

It follows readily from this and from the Thoma classification theorem 5.22 that a non-negative specialisation of Sym can be parametrised bijectively by a triple

$$((\alpha_n)_{n\geq 1}, (\beta_n)_{n\geq 1}, \gamma)$$

with $(\alpha_n)_{n\geq 1}$ and $(\beta_n)_{n\geq 1}$ non-negative non-increasing summable sequences, and γ non-negative real number. The corresponding specialisation is

$$p_1(X) = \sum_{n=1}^{\infty} \alpha_n + \sum_{n=1}^{\infty} \beta_n + \gamma;$$

$$p_{k\geq 2}(X) = \sum_{n=1}^{\infty} (\alpha_n)^k + (-1)^{k-1} \sum_{n=1}^{\infty} (\beta_n)^k.$$

Now, consider two non-negative specialisations X and Y of the algebra of symmetric functions Sym. We assume that the series $\sum_{k=1}^{\infty} \frac{p_k(X)\,p_k(Y)}{k}$ is absolutely convergent: beware that this is not always the case (for instance, if $X = Y = \{1\}$ are the specialisations associated to a single variable $x_1 = y_1 = 1$, then we get $\sum_{k=1}^{\infty} \frac{p_k(X)\,p_k(Y)}{k} = \sum_{k=1}^{\infty} \frac{1}{k} = +\infty$).

Definition 5.24 (Schur measure). *The Schur measure associated to two non-negative specialisations X and Y of* Sym *is the probability measure on $\mathfrak{Y}$ given by*

$$\mathbb{S}_{X,Y}[\lambda] = \frac{s_\lambda(X)\,s_\lambda(Y)}{\exp\left(\sum_{k=1}^{\infty} \frac{p_k(X)\,p_k(Y)}{k}\right)},$$

assuming that the denominator involves an absolutely convergent series.

The fact that one obtains a probability measure is an immediate consequence of the Cauchy identity, as already used during the proof of Lemma 5.17. Let us denote tX the non-negative specialisation with $p_k(tX) = t^k\,p_k(X)$. It is easy to see that $\mathbb{S}_{X,Y}[\lambda] = \mathbb{S}_{tX,t^{-1}Y}[\lambda]$ for any positive real number t; thus, there is always a choice of normalisation when defining a Schur measure. In many cases, it is convenient to choose the renormalisation of the virtual alphabets X and Y so that $p_1(X) = p_1(Y)$.

Example 5.25 (Poissonised Plancherel measures). Suppose that $X = Y = \sqrt{\theta}\,E$, where E is the exponential alphabet and θ is an arbitrary positive real number. Then,

$$\mathbb{S}[\lambda] = \frac{(s_\lambda(\sqrt{\theta}E))^2}{\exp(\theta)} = \mathrm{e}^{-\theta}\,\theta^{|\lambda|}\left(\frac{\dim\lambda}{|\lambda|!}\right)^2 = \frac{\mathrm{e}^{-\theta}\,\theta^{|\lambda|}}{|\lambda|!}\,\mathrm{PL}_{|\lambda|}[\lambda].$$

A random integer partition chosen under this Schur measure can be obtained as follows: we first choose a random integer n according to the Poisson law $\mathcal{P}(\theta)$ with parameter θ, and we then choose a random integer partition according to the Plancherel measure PL_n. This fact provides us with a new way to study large Plancherel-distributed integer partitions. If n is a large integer, then the Poisson law $\mathcal{P}(n)$ is concentrated around its mean n (with a standard deviation equal to $\sqrt{n}$), so for many asymptotic equations, the study of $\lambda \sim \mathrm{PL}_n$ is equivalent to the study of $\lambda \sim \mathbb{S}_{\sqrt{n}E,\sqrt{n}E}$. The latter distribution on $\mathfrak{Y}$ turns out to give rise to determinantal point processes with explicit kernels, by considering the descent or Frobenius coordinates M_λ or F_λ.

Theorem 5.26 (Okounkov). *Consider the descent coordinates M_λ of a random integer partition λ chosen according to a Schur measure $\mathbb{S}_{X,Y}$. We denote $t_k^X = \frac{p_k(X)}{k}$ and $t_k^Y = \frac{p_k(Y)}{k}$ the Miwa parameters of the non-negative specialisations X and Y, and we set*

$$\mathscr{J}_{X,Y}(z) = \exp\left(\sum_{k=1}^{\infty} t_k^X\,z^k - t_k^Y\,z^{-k}\right).$$

Then, $M_\lambda \subset \mathbb{Z}'$ is a determinantal point process. Its kernel $(K_{X,Y}(x,y))_{x,y\in\mathbb{Z}'}$ is encoded by the double generating series $\mathscr{K}_{X,Y}(z,w) = \sum_{x,y\in\mathbb{Z}'} z^x\, w^{-y}\, K_{X,Y}(x,y)$, which is given by the following explicit formula:

$$\mathscr{K}_{X,Y}(z,w) = \frac{\sqrt{zw}}{z-w}\,\frac{\mathscr{J}_{X,Y}(z)}{\mathscr{J}_{X,Y}(w)}.$$

REMARK 5.27. It can also be shown that the Frobenius coordinates F_λ of a random integer partition λ chosen according to a Schur measure $\mathbb{S}_{X,Y}$ form a determinantal point process; this follows directly from the set-theoretic relations between F_λ and M_λ, and from a manipulation of the determinants involved in the correlation functions.

The remainder of this section is devoted to a proof of Okounkov's theorem. In the sequel, we shall deal with formal series in the Miwa parameters t^X and t^Y and the variables z and w; we leave the reader check that each time one can replace these formal series by absolutely convergent series (at least for adequate values of the parameters z and w).

To begin with, notice that in the discrete setting, a determinantal point process with kernel K on $\mathbb{Z}'$ is a random subset $M \subset \mathbb{Z}'$ such that, for any finite family of points $A = \{a_1,\ldots,a_n\}$, we have:

$$\mathbb{P}[A \subset M] = \mathbb{E}[M^{\downarrow n}(\{a_1\}\times\{a_2\}\times\cdots\times\{a_n\})] = \rho_n(a_1,\ldots,a_n) = \det(K(a_i,a_j))_{1\leq i,j\leq n}.$$

The appearance of a determinant in the setting of Schur measures is related to the combinatorics of the so-called *infinite wedge space* (also called *fermionic Fock space*). This space $\bigwedge^\infty$ is the Hilbert space spanned by the orthonormal basis of vectors

$$v_M = x_0 \wedge x_1 \wedge \cdots \wedge x_n \wedge \cdots\;,$$

$M = \{x_0 > x_1 > \cdots > x_n > \cdots\}$ running over the set of infinite decreasing sequences in $\mathbb{Z}'$ which contain all the sufficiently large negative half-integers. Among these vectors, the *vacuum state* is

$$v_\emptyset = -\frac{1}{2} \wedge -\frac{3}{2} \wedge -\frac{5}{2} \wedge \cdots\;;$$

it corresponds to the empty integer partition, and more generally, we shall denote $v_\lambda = v_{M_\lambda}$. A basis vector v_M corresponds to an integer partition λ ($M = M_\lambda$) if and only if the number of positive coordinates x_i in v_M is equal to the number of negative coordinates x_j which are not in v_M. The half-integers $x \in \mathbb{Z}'$ act on $\bigwedge^\infty$ by the *free fermion operators* ψ_x and ψ_x^*:

$$\psi_x(v_M) = \begin{cases} x \wedge v_M & \text{if } x \notin M \\ 0 & \text{if } x \in M \end{cases} \quad;\quad \psi_x^*(v_M) = \begin{cases} \varepsilon_{x,M}\, v_{M\setminus\{x\}} & \text{if } x \in M \\ 0 & \text{if } x \notin M \end{cases}$$

with the usual rules of anticommutation for the $\wedge$ symbols (in order to replace x inside a decreasing sequence), and where $\varepsilon_{x,M}$ is the parity of the position of x in the decreasing sequence M. The operator ψ_x^* is the adjoint of ψ_x, and we have the anticommutation formula $\psi_x\psi_x^* + \psi_x^*\psi_x = \mathrm{id}_{\bigwedge^\infty}$; all the other anticommutators vanish. The free fermions enable us to define the *charge operator and the energy operator*:

$$C = \sum_{x>0} \psi_x\psi_x^* - \sum_{x<0} \psi_x^*\psi_x;$$
$$H = \sum_{x>0} x\,\psi_x\psi_x^* - \sum_{x<0} x\,\psi_x^*\psi_x.$$

LEMMA 5.28. *We have $C(v_\lambda) = 0$ and $H(v_\lambda) = |\lambda|\, v_\lambda$ for any integer partition λ. The kernel of C is the span of the vectors v_λ with $\lambda \in \mathfrak{Y}$.*

Proof. We remark that $\psi_x\psi_x^*(v_M) = 1_{(x\in M)}\, v_M$, and that $\psi_x^*\psi_x(v_M) = 1_{(x\notin M)} v_M$. Therefore, $C(v_M)$ is equal to $t\, v_M$, where t is the difference $|M\cap \mathbb{Z}'_+| - |\mathbb{Z}'_- \setminus (M\cap\mathbb{Z}'_-)|$. This quantity vanishes if and only if M is the set of descent coordinates of an integer partition, whence the results on the charge operator. For the energy operator, we get $H(v_M) = u\, v_M$, with

$$u = \sum_{x\in M\cap\mathbb{Z}'_+} x - \sum_{x\in\mathbb{Z}'_-\setminus(M\cap\mathbb{Z}'_-)} x = \sum_{x\in F_\lambda} |x|$$

if $M = M_\lambda$. It is then clear from the drawing at the beginning of this chapter that the sum of the absolute values of the Frobenius coordinates of a Young diagram λ is equal to the area $|\lambda|$ of this diagram. □

Because of the lemma above, we can embed isometrically the vector space of symmetric functions Sym (endowed with the Hall scalar product) in the subspace of $\bigwedge^\infty$ spanned by the vectors v_M with charge 0, by sending the Schur function s_λ to the vector v_λ. The *boson-fermion correspondence* relates the multiplication by p_k in Sym to some explicit linear operators on $\bigwedge^\infty$. For $k \in \mathbb{Z}\setminus\{0\}$, we set $\alpha_k = \sum_{l\in\mathbb{Z}'} \psi_l\psi^*_{l+k}$. Note that $\alpha_k^* = \alpha_{-k}$; one can also show that the operators α_k satisfy the Heisenberg commutation relations $[\alpha_m,\alpha_n] = m\, 1_{(m=-n)}$. Besides, if $k>0$, then $\alpha_k(v_\emptyset) = 0$: indeed, if $l\in\mathbb{Z}'$ is such that $\psi^*_{l+k}(v_\emptyset)$ does not vanish, then $l+k\in\mathbb{Z}'_-$, so $l\in\mathbb{Z}'_-$ and $\psi_l(\psi^*_{l+k}(v_\emptyset))$ vanishes.

Theorem 5.29 (Pieri rules and boson-fermion correspondence). *If* $p_k\, s_\lambda = \sum_{|\mu|=|\lambda|+k} c_{\lambda,\mu,k}\, s_\mu$ *in* Sym, *then* $\alpha_{-k}\, v_\lambda = \sum_{|\mu|=|\lambda|+k} c_{\lambda,\mu,k}\, v_\mu$.

Proof of Theorem 5.29 in the case $k=1$. We prove the case $k=1$; the general case follows from a generalisation of Proposition 2.17, which is essentially the Murnaghan–Nakayama rule. When $k=1$, we know that $p_1\, s_\lambda = \sum_{\Lambda\,|\,\lambda\nearrow\Lambda} s_\Lambda$. In terms of descent coordinates, the adjunction of a box to a Young diagram corresponds to the shift $x_i \to x_i+1$ of one descent coordinate $x_i\in M_\lambda$, assuming that x_i+1 is not in M_λ. So, if we denote $\sigma : \mathrm{Sym} \to \bigwedge^\infty$ the isometric embedding described above, then

$$\sigma(p_1\, s_\lambda) = \sum_{k\in\mathbb{Z}'} 1_{(k\in M_\lambda,\,(k+1)\notin M_\lambda)} v_{(M_\lambda\setminus\{k\})\cup\{k+1\}} = \left(\sum_{k\in\mathbb{Z}'} \psi_{k+1}\psi_k^*\right)(v_\lambda) = \alpha_{-1}(\sigma(s_\lambda)).$$

We can therefore represent the multiplication by p_1 in Sym by the operator α_{-1} on the infinite wedge space. □

As a corollary of Theorem 5.29, we can now write a Schur measure $\mathbb{S}_{X,Y}$ as a scalar product in $\bigwedge^\infty$. Given an infinite sequence of formal parameters $t=(t_1,t_2,\ldots)$, we set

$$\Gamma_\pm(t) = \exp\left(\sum_{k=1}^\infty t_k\,\alpha_{\pm k}\right).$$

The operators $\Gamma_+(t)$ and $\Gamma_-(t)$ are adjoint. We denote as above σ the embedding $\mathrm{Sym}\to\bigwedge^\infty$, and we consider the symmetric functions in Sym as formal series in the variables in an alphabet Y. Notice that if X is the specialisation of Sym with Miwa parameters $t_k = \frac{p_k(X)}{k}$, then

$$\begin{aligned}\Gamma_-(t)(v_\emptyset) &= \exp\left(\sum_{k=1}^\infty t_k\,\alpha_{-k}\right)(v_\emptyset) = \sigma\left(\exp\left(\sum_{k=1}^\infty t_k\, p_k(Y)\right)\right)\\ &= \sigma\left(\sum_{\lambda\in\mathfrak{Y}} s_\lambda(X)\, s_\lambda(Y)\right) = \sum_{\lambda\in\mathfrak{Y}} s_\lambda(X)\, v_\lambda,\end{aligned}$$

by using as before the Cauchy formula. On the other hand, as $\alpha_k(v_\emptyset) = 0$ for $k \geq 1$, $\Gamma_+(t)(v_\emptyset) = v_\emptyset$. Consider now two positive specialisations X and Y of the algebra Sym, and the associated sequences of Miwa parameters t_X and t_Y. We have, for any finite set $A = \{a_1, a_2, \ldots, a_n\} \subset \mathbb{Z}'$:

$$\begin{aligned}
\mathbb{S}_{X,Y}[A \subset M_\lambda] &= \exp\left(-\sum_{k=1}^{\infty} k\, t_k^X\, t_k^Y\right) \sum_{\lambda \,|\, A \subset M_\lambda} s_\lambda(X)\, s_\lambda(Y) \\
&= \exp\left(-\sum_{k=1}^{\infty} k\, t_k^X\, t_k^Y\right) \left\langle \Gamma_-(t_X)\,(v_\emptyset) \,\middle|\, \left(\prod_{a \in A} \psi_a \psi_a^*\right) \Gamma_-(t_Y)\,(v_\emptyset) \right\rangle \\
&= \exp\left(-\sum_{k=1}^{\infty} k\, t_k^X\, t_k^Y\right) \left\langle \Gamma_-(t_X)\,\Gamma_+(-t_Y)\,(v_\emptyset) \,\middle|\, \left(\prod_{a \in A} \psi_a \psi_a^*\right) \Gamma_-(t_Y)\,\Gamma_+(-t_X)\,(v_\emptyset) \right\rangle \\
&= \exp\left(-\sum_{k=1}^{\infty} k\, t_k^X\, t_k^Y\right) \left\langle v_\emptyset \,\middle|\, \Gamma_-(-t_Y)\,\Gamma_+(t_X) \left(\prod_{a \in A} \psi_a \psi_a^*\right) \Gamma_-(t_Y)\,\Gamma_+(-t_X)\,(v_\emptyset) \right\rangle.
\end{aligned}$$

LEMMA 5.30. *We have the commutation relation*

$$\Gamma_+(t_X)\,\Gamma_-(-t_Y) = \Gamma_-(-t_Y)\,\Gamma_+(t_X)\, \exp\left(-\sum_{k=1}^{\infty} k\, t_k^X\, t_k^Y\right).$$

PROOF. On a finite dimensional vector space, given two linear operators U, V such that $[U, V]$ commutes with U and with V, we have:

$$\mathrm{e}^U\, \mathrm{e}^V = \mathrm{e}^V\, \mathrm{e}^U\, \mathrm{e}^{[U,V]}.$$

Indeed, consider the matrix-valued map $M : t \mapsto \mathrm{e}^{-tV}\, \mathrm{e}^U\, \mathrm{e}^{tV}$. Its derivative with respect to t is $M'(t) = [M(t), V] = \mathrm{e}^{-tV}\, [\mathrm{e}^U, V]\, \mathrm{e}^{tV}$. However,

$$\begin{aligned}
[\mathrm{e}^U, V] = \sum_{n=1}^{\infty} \frac{U^n V - V U^n}{n!} &= \sum_{n=1}^{\infty} \frac{1}{n!} \sum_{k=0}^{n-1} (U^{n-k} V U^k - U^{n-k-1} V U^{k+1}) \\
&= \sum_{n=1}^{\infty} \frac{1}{n!} \sum_{k=0}^{n-1} U^{n-k-1} [U, V] V^k = \sum_{n=1}^{\infty} \frac{1}{n!} \sum_{k=0}^{n-1} U^{n-1}\, [U, V] \\
&= \sum_{n=1}^{\infty} \frac{U^{n-1}}{(n-1)!}\, [U, V] = \mathrm{e}^U\, [U, V].
\end{aligned}$$

Therefore, $M'(t) = \mathrm{e}^{-tV}\, \mathrm{e}^U\, [U, V]\, \mathrm{e}^{tV} = M(t)\, [U, V]$, and the solution of this differential equation is $M(t) = M(0)\, \mathrm{e}^{t[U,V]} = \mathrm{e}^U\, \mathrm{e}^{t[U,V]}$. The value at $t = 1$ yields the claimed formula. Now, it is not difficult to see that one can extend in certain cases the formula to an infinite-dimensional setting; in particular, this works with the operators $U = \sum_{k=1}^{\infty} t_k^X\, \alpha_k$ and $V = -\sum_{k=1}^{\infty} t_k^Y\, \alpha_{-k}$ acting on $\bigwedge^\infty$, whence the claimed formula, at least with formal Miwa parameters. ☐

Consequently, given a Schur measure $\mathbb{S}_{X,Y}$, we obtain the following representation formula in the infinite wedge space:

$$\begin{aligned}
\mathbb{S}_{X,Y}[A \subset M_\lambda] &= \left\langle v_\emptyset \,\middle|\, \Gamma_+(t_X)\,\Gamma_-(-t_Y) \left(\prod_{a \in A} \psi_a \psi_a^*\right) \Gamma_-(t_Y)\,\Gamma_+(-t_X)\,(v_\emptyset) \right\rangle \\
&= \left\langle v_\emptyset \,\middle|\, \left(\prod_{a \in A} \Psi_a \Psi'_a\right) (v_\emptyset) \right\rangle
\end{aligned}$$

where $\Psi_a = G\, \psi_a\, G^{-1}$, $\Psi'_a = G\, \psi_a^*\, G^{-1}$, and $G = \Gamma_+(t_X)\,\Gamma_-(-t_Y)$ (beware that since G is not unitary, Ψ'_a *is not* the adjoint of Ψ_a).

LEMMA 5.31. *The operators Ψ_a (respectively, Ψ'_a) are (infinite) linear combinations of the operators ψ_x (respectively, of the operators ψ^*_x), with coefficients depending on the two specialisations X and Y.*

PROOF. Consider the generating series $\psi(z) = \sum_{n\in\mathbb{Z}'} \psi_n\, z^n$ and $\psi^*(w) = \sum_{n\in\mathbb{Z}'} w^{-n}\, \psi^*_n$. We start by computing the commutators of the free bosons α_k with the free fermions ψ_m. Note that if $m \neq l+k$, then ψ_m anticommutes with ψ_l and with ψ^*_{l+k}, so

$$\psi_l\psi^*_{l+k}\psi_m = -\psi_l\psi_m\psi^*_{l+k} = \psi_m\psi_l\psi^*_{l+k} \qquad ; \qquad [\psi_l\psi^*_{l+k}, \psi_m] = 0.$$

Therefore,

$$[\alpha_k, \psi_m] = [\psi_{m-k}\psi^*_m, \psi_m] = \psi_{m-k}\psi^*_m\psi_m - \psi_m\psi_{m-k}\psi^*_m = \psi_{m-k}(\psi^*_m\psi_m + \psi_m\psi^*_m) = \psi_{m-k}.$$

We can encode these commutation relations by the generating series:

$$[\alpha_k, \psi(z)] = \sum_{m\in\mathbb{Z}} z^m\, \psi_{m-k} = z^k\, \psi(z).$$

Similarly, $[\alpha_k, \psi^*(w)] = -w^k\, \psi^*(w)$. Now, the elements α_k belong to an infinite-dimensional Lie algebra, and the $\Gamma_\pm(t)$ are group-like elements of a corresponding infinite-dimensional Lie group. The commutation relations established above lead to:

$$\Gamma_\pm(t)\, \psi(z)\, \Gamma_\pm(-t) = \exp\left(\sum_{k=1}^{\infty} t_k\, z^{\pm k}\right)\, \psi(z);$$
$$\Gamma_\pm(t)\, \psi^*(w)\, \Gamma_\pm(-t) = \exp\left(\sum_{k=1}^{\infty} -t_k\, w^{\pm k}\right)\, \psi^*(w).$$

Therefore,

$$G\, \psi(z)\, G^{-1} = \exp\left(\sum_{k=1}^{\infty} t^X_k\, z^k - \sum_{k=1}^{\infty} t^Y_k\, z^{-k}\right)\, \psi(z);$$
$$G\, \psi^*(w)\, G^{-1} = \exp\left(-\sum_{k=1}^{\infty} t^X_k\, w^k + \sum_{k=1}^{\infty} t^Y_k\, w^{-k}\right)\, \psi^*(w).$$

By taking the coefficient of z^n or of w^{-n} of these expressions, we obtain on the left-hand side Ψ_n or Ψ'_n, and on the right-hand side a linear combination of operators ψ_x or ψ^*_x. $\square$

We now want to show that, for any finite part $A = \{a_1, \ldots, a_n\}$, we have

$$\left\langle v_\emptyset \,\middle|\, \left(\prod_{i=1}^n \Psi_{a_i}\Psi'_{a_i}\right)\, (v_\emptyset) \right\rangle = \det\left(\left(\left\langle v_\emptyset \,\middle|\, \Psi_{a_i}\Psi'_{a_j}\, (v_\emptyset)\right\rangle\right)_{1\leq i,j\leq n}\right).$$

Denote $[[U,V]] = UV + VU$ the anticommutator of two operators. The anticommutation relations $[[\psi_k, \psi^*_l]] = 1_{(k=l)}$ and $[[\psi_k, \psi_l]] = [[\psi^*_k, \psi^*_l]] = 0$ also hold for the G-conjugated operators Ψ_k and Ψ'_l. Therefore,

$$\Psi_{a_1}\Psi'_{a_1}\Psi_{a_2}\Psi'_{a_2}\cdots\Psi_{a_n}\Psi'_{a_n} = \Psi_{a_1}\Psi_{a_2}\cdots\Psi_{a_n}\Psi'_{a_n}\cdots\Psi'_{a_2}\Psi'_{a_1}.$$

Denote $\Psi_a = \sum_{k\in\mathbb{Z}'} c(a,k)\, \psi_k$ and $\Psi'_a = \sum_{n\in\mathbb{Z}'} c'(a,k)\, \psi^*_k$, the coefficients $c(a,k)$ and $c'(a,k)$ coming from (the proof of) Lemma 5.31. By multilinearity,

$$\left\langle v_\emptyset \,\middle|\, \left(\prod_{i=1}^n \Psi_{a_i}\Psi'_{a_i}\right)\, (v_\emptyset) \right\rangle = \left\langle v_\emptyset \,\middle|\, \Psi_{a_1}\Psi_{a_2}\cdots\Psi_{a_n}\Psi'_{a_n}\cdots\Psi'_{a_2}\Psi'_{a_1}(v_\emptyset)\right\rangle$$
$$= \sum_{\substack{k_1,\ldots,k_n\in\mathbb{Z}'\\ l_1,\ldots,l_n\in\mathbb{Z}'}} c(a_1,k_1)\cdots c(a_n,k_n)\, c'(a_1,l_1)\cdots c'(a_n,l_n)\, \left\langle v_\emptyset \,\middle|\, \psi_{k_1}\cdots\psi_{k_n}\psi^*_{l_n}\cdots\psi^*_{l_1}(v_\emptyset)\right\rangle.$$

We then see that it suffices to prove:

$$\left\langle v_\emptyset \,\middle|\, \psi_{k_1}\cdots\psi_{k_n}\psi^*_{l_n}\cdots\psi^*_{l_1}(v_\emptyset)\right\rangle = \det\left(\left(\left\langle v_\emptyset \,\middle|\, \psi_{k_i}\psi^*_{l_j}(v_\emptyset)\right\rangle\right)_{1\leq i,j\leq n}\right)$$

for any $2n$-tuple of half-integers $(k_1,\ldots,k_n,l_1,\ldots,l_n)$ (note that the k_i's and the l_j's are not necessarily distinct). Indeed, resumming over the $2n$-tuples this relation, we shall obtain the formula with the operators Ψ_{a_i} and Ψ'_{a_j}. Now, this last determinantal formula is an instance of *Wick's theorem* for anticommutating operators; here it can be given an elementary proof. Let us first analyse the left-hand side, which can be rewritten as

$$\left\langle \psi^*_{k_n}\cdots\psi^*_{k_1}(v_\emptyset) \,\middle|\, \psi^*_{l_n}\cdots\psi^*_{l_1}(v_\emptyset)\right\rangle.$$

If $\psi^*_{l_n}\cdots\psi^*_{l_1}\neq 0$, then all the l_j's are distinct. Moreover, if $\psi^*_{l_n}\cdots\psi^*_{l_1}(v_\emptyset)\neq 0$, then all the l_j's are in $\mathbb{Z}'_-$. So,

$$\psi^*_{l_n}\cdots\psi^*_{l_1}(v_\emptyset) = \pm 1_{(\text{the } l_j\text{'s are distinct negative half-integers})}\, v_{\mathbb{Z}'\setminus\{l_1,\ldots,l_n\}}.$$

The scalar product $\left\langle \psi^*_{k_n}\cdots\psi^*_{k_1}(v_\emptyset) \,\middle|\, \psi^*_{l_n}\cdots\psi^*_{l_1}(v_\emptyset)\right\rangle$ does not vanish if the two terms do not vanish and are colinear vectors. This happens if and only if $\{k_1,\ldots,k_n\}=\{l_1,\ldots,l_n\}$ is the same set of distinct negative half-integers. So,

$$\left\langle \psi^*_{k_n}\cdots\psi^*_{k_1}(v_\emptyset) \,\middle|\, \psi^*_{l_n}\cdots\psi^*_{l_1}(v_\emptyset)\right\rangle = \pm 1_{(\{k_1,\ldots,k_n\}=\{l_1,\ldots,l_n\})}\, 1_{(\text{the } l_j\text{'s are distinct negative half-integers})}.$$

The sign $\pm$ is the signature of the permutation σ of $[\![1,n]\!]$ such that $k_i=l_{\sigma(i)}$ for any i: indeed, we can first check that the sign is $+1$ when $\sigma=\mathrm{id}_{[\![1,n]\!]}$ (by using the classical anticommutation relations with the $\wedge$ symbol), and the other cases follow by using the anticommutation relations satisfied by the operators ψ_{k_i}. Now, by using the formula above in the case $n=1$, we get:

$$\left\langle v_\emptyset \,\middle|\, \psi_{k_i}\psi^*_{l_j}(v_\emptyset)\right\rangle = 1_{(k_i=l_j)}\, 1_{(l_j\in\mathbb{Z}'_-)}.$$

So, the determinant on the right-hand side of the claimed formula is:

$$\sum_{\sigma\in\mathfrak{S}(n)} \varepsilon(\sigma) 1_{(k_1=l_{\sigma(1)},\ldots,k_n=l_{\sigma(n)})}\, 1_{(l_1,\ldots,l_n\in\mathbb{Z}'_-)}.$$

If the two multisets $\{k_1,\ldots,k_n\}$ and $\{l_1,\ldots,l_n\}$ are distinct, then no permutation σ is such that $k_i=l_{\sigma(i)}$, and the determinant vanishes. Suppose now that the two multisets $\{k_1,\ldots,k_n\}$ are $\{l_1,\ldots,l_n\}$ are equal, but that there are multiple occurrences of one element: for instance, $k_1=k_2$. Then, the involution $\sigma\mapsto\sigma\circ(1,2)$ reverses the signatures $\varepsilon(\sigma)$, but leaves invariant the set of permutations σ such that $k_i=l_{\sigma(i)}$ for $i\in[\![1,n]\!]$. So, there are as many such permutations σ with signature $+1$ and with signature -1, and the determinant vanishes again. We conclude that

$$\det = \left(\sum_{\sigma\in\mathfrak{S}(n)} \varepsilon(\sigma) 1_{(k_1=l_{\sigma(1)},\ldots,k_n=l_{\sigma(n)})}\right) 1_{(\text{the } l_j\text{'s are distinct negative half-integers})}.$$

Now, with two equal sets $\{k_1,\ldots,k_n\}=\{l_1,\ldots,l_n\}$ of distinct half-integers, there is only one permutation σ such that $k_i=l_{\sigma(i)}$, so we conclude that the determinant is again equal to the signature of this permutation, assuming the condition

$$1_{(\{k_1,\ldots,k_n\}=\{l_1,\ldots,l_n\})}\, 1_{(\text{the } l_j\text{'s are distinct negative half-integers})} = 1.$$

This ends the proof of the determinantal formula for $\mathbb{S}_{X,Y}[A\subset M_\lambda]$: for any finite subset $A=\{a_1,\ldots,a_n\}$ of $\mathbb{Z}'$,

$$\mathbb{S}_{X,Y}[A\subset M_\lambda] = \det\left(\left(\left\langle v_\emptyset \,\middle|\, \Psi_{a_i}\Psi'_{a_j}(v_\emptyset)\right\rangle\right)_{1\leq i,j\leq n}\right).$$

PROOF OF THEOREM 5.26. It remains to see that the kernel $K_{X,Y}(x,y)$ of our determinantal point process can be related to the generating series $\mathscr{J}_{X,Y}(z)$ as explained in the statement of the theorem. By the discussion above,

$$K_{X,Y}(x,y) = \left\langle v_\emptyset \;\middle|\; \Psi_x \Psi'_y(v_\emptyset) \right\rangle .$$

If one forms the generating series of this kernel, one gets:

$$\begin{aligned}
\mathscr{K}_{X,Y}(z,w) &= \sum_{n,m\in\mathbb{Z}'} z^n\, w^{-m}\, K_{X,Y}(n,m) \\
&= \sum_{n,m\in\mathbb{Z}'} \left\langle v_\emptyset \;\middle|\; G\,(z^n\psi_n)\,(w^{-m}\psi^*_m)\,G^{-1}(v_\emptyset) \right\rangle \\
&= \left\langle v_\emptyset \;\middle|\; G\,\psi(z)\,\psi^*(w)\,G^{-1}(v_\emptyset) \right\rangle \\
&= \frac{\mathscr{J}_{X,Y}(z)}{\mathscr{J}_{X,Y}(w)} \left\langle (v_\emptyset) \;\middle|\; \psi(z)\,\psi^*(w)\,(v_\emptyset) \right\rangle ,
\end{aligned}$$

by using on the last line the formulas for the conjugation action of G on $\psi(z)$ and $\psi^*(w)$, which have already been used during the proof of Lemma 5.31. Finally,

$$\psi^*(w)\,(v_\emptyset) = \sum_{n\in\mathbb{Z}'_-} (-1)^{n+\frac{1}{2}}\, w^{-n}\, v_{\mathbb{Z}'_-\setminus\{n\}}$$

$$\langle v_\emptyset \mid \psi(z)\,\psi^*(w)\,(v_\emptyset)\rangle = \sum_{n\in\mathbb{Z}'_-} \left(\frac{z}{w}\right)^n = \sqrt{\frac{w}{z}}\,\frac{1}{1-\frac{w}{z}} = \frac{\sqrt{wz}}{z-w}. \qquad \square$$

REMARK 5.32. Suppose that $X = Y = \sqrt{\theta}\,E$ (Poissonised Plancherel measures). Then, the group-like elements $\Gamma_\pm(t^X)$ and $\Gamma_\pm(t^Y)$ only involve the bosons $\alpha_{\pm 1}$, for which the boson-fermion correspondance (Theorem 5.29) has been proved. So, if we are only interested in this case, then there is no gap in our discussion.

4. Asymptotics of the Plancherel kernel in the bulk and at the edge

To close this chapter, we explain how to use the determinantal structure of the random point process M_λ with $\lambda \sim \mathbb{S}_{\sqrt{\theta}E,\sqrt{\theta}E}$ in order to understand the asymptotic behavior of large Plancherel random partitions. Theorem 5.26 reads in this setting as follows. If PPL_θ is the Poissonised Plancherel measure with parameter θ and if $\lambda \sim \mathrm{PPL}_\theta$, then $M_{\{\theta\}} = M_\lambda$ is a determinantal point process on $\mathbb{Z}'$ with kernel $K_\theta(x,y)$, with

$$\sum_{x,y\in\mathbb{Z}'} z^x\, w^{-y}\, K_{\{\theta\}}(x,y) = \frac{\sqrt{zw}}{z-w}\,\exp\left(\sqrt{\theta}((z-z^{-1})-(w-w^{-1}))\right).$$

We recover the value of $K_{\{\theta\}}(x,y)$ by taking a double contour integral along two arbitrary non-crossing paths circling around 0:

$$K_{\{\theta\}}(x,y) = \frac{1}{(2\mathrm{i}\pi)^2} \oint\!\!\oint \frac{1}{(z-w)\sqrt{zw}}\,\exp\left(\sqrt{\theta}((z-z^{-1})-(w-w^{-1}))\right) z^{-x} w^{y}\,dz\,dw,$$

with the contours γ_w and γ_z chosen so that γ_w is in the interior of the domain determined by γ_z (this ensures that the series $\sum_{n\in\mathbb{Z}'_-}\left(\frac{z}{w}\right)^n$ considered at the very end of the proof of Theorem 5.26 converges absolutely). This is the so-called *discrete Bessel kernel*. The terminology comes from the following link with the Bessel functions of the first kind $J_n(x)$, which are the entire functions on the complex plane defined by:

$$J_n(x) = \sum_{m=0}^{\infty} \frac{(-1)^m}{m!\,\Gamma(m+n+1)} \left(\frac{x}{2}\right)^{2m+n} = \frac{1}{2\mathrm{i}\pi} \oint \mathrm{e}^{\frac{x}{2}(z-z^{-1})} \,\frac{dz}{z^{n+1}}.$$

PROPOSITION 5.33 (Discrete Bessel kernel). *We have*

$$K_{\{\theta\}}(x,y) = \sqrt{\theta}\,\frac{J_{x-\frac{1}{2}}(2\sqrt{\theta})\,J_{y+\frac{1}{2}}(2\sqrt{\theta}) - J_{x+\frac{1}{2}}(2\sqrt{\theta})\,J_{y-\frac{1}{2}}(2\sqrt{\theta})}{x-y}.$$

PROOF. The following argument is due to Okounkov; the appearance of a ratio as above is a classical phenomenon in the theory of determinantal point processes, and it is related to the Christoffel–Darboux formula for orthogonal polynomials. Note that:

$$\frac{1}{z^{x+\frac{1}{2}}\,w^{-y+\frac{1}{2}}} = -\frac{1}{x-y}\left(z\frac{\partial}{\partial z} + w\frac{\partial}{\partial w} + 1\right)\frac{1}{z^{x+\frac{1}{2}}\,w^{-y+\frac{1}{2}}}.$$

Suppose given two analytic functions f and g, defined in a neighborhood of the circle $\{z \,|\, |z| = r\}$. An integration by parts yields:

$$\begin{aligned}
\oint f(z)\left(z\frac{\partial}{\partial z}\right) g(z)\,dz &= \int_{\theta=0}^{2\pi} f(r\mathrm{e}^{\mathrm{i}\theta}) r\mathrm{e}^{\mathrm{i}\theta} g'(r\mathrm{e}^{\mathrm{i}\theta})\, r\mathrm{i}\mathrm{e}^{\mathrm{i}\theta}\,d\theta \\
&= \left[f(r\mathrm{e}^{\mathrm{i}\theta}) r\mathrm{e}^{\mathrm{i}\theta} g(r\mathrm{e}^{\mathrm{i}\theta})\right]_0^{2\pi} - \int_0^{2\pi} \left(f'(r\mathrm{e}^{\mathrm{i}\theta}) r^2 \mathrm{i}\mathrm{e}^{2\mathrm{i}\theta} + f(r\mathrm{e}^{\mathrm{i}\theta})\, r\mathrm{i}\mathrm{e}^{\theta}\right) g(r\mathrm{e}^{\mathrm{i}\theta})\,d\theta \\
&= -\oint (f'(z)z + f(z))\,g(z)\,dz = -\oint \frac{\partial}{\partial z}\left(zf(z)\right)\,g(z)\,dz.
\end{aligned}$$

In other words, the adjoint operator of $z\frac{\partial}{\partial z}(\cdot)$ is $-\frac{\partial}{\partial z}(z\cdot) = -1 - z\frac{\partial}{\partial z}(\cdot)$. It follows that the adjoint operator of $U = z\frac{\partial}{\partial z} + w\frac{\partial}{\partial w} + 1$ is $-U$. Therefore,

$$\begin{aligned}
K_{\{\theta\}}(x,y) &= \frac{1}{(2\mathrm{i}\pi)^2} \oint\!\!\oint \frac{\mathrm{e}^{\sqrt{\theta}((z-z^{-1})-(w-w^{-1}))}}{z-w}\,\frac{1}{z^{x+\frac{1}{2}}w^{-y+\frac{1}{2}}}\,dz\,dw \\
&= -\frac{1}{(2\mathrm{i}\pi)^2(x-y)} \oint\!\!\oint \frac{\mathrm{e}^{\sqrt{\theta}((z-z^{-1})-(w-w^{-1}))}}{z-w}\,U\!\left(\frac{1}{z^{x+\frac{1}{2}}w^{-y+\frac{1}{2}}}\right)dz\,dw \\
&= \frac{1}{(2\mathrm{i}\pi)^2(x-y)} \oint\!\!\oint U\!\left(\frac{\mathrm{e}^{\sqrt{\theta}((z-z^{-1})-(w-w^{-1}))}}{z-w}\right)\frac{1}{z^{x+\frac{1}{2}}w^{-y+\frac{1}{2}}}\,dz\,dw.
\end{aligned}$$

Finally,

$$\begin{aligned}
U\!\left(\frac{\mathrm{e}^{\sqrt{\theta}((z-z^{-1})-(w-w^{-1}))}}{z-w}\right) &= \frac{z+z^{-1}-w-w^{-1}}{z-w}\,\sqrt{\theta}\,\mathrm{e}^{\sqrt{\theta}((z-z^{-1})-(w-w^{-1}))} \\
&= \left(1-\frac{1}{zw}\right)\sqrt{\theta}\,\mathrm{e}^{\sqrt{\theta}((z-z^{-1})-(w-w^{-1}))},
\end{aligned}$$

and the formula follows now from the definition of $J_{x\pm\frac{1}{2}}(2\sqrt{\theta})$ as a path integral. □

In the following, we shall be interested in the asymptotics of $M_{\{\theta\}}$ in several regions:

- *in the bulk*: with $x_0 \in (-2,2)$, denote $M_{\{\theta\}}^{\mathrm{local},x_0}$ the random point process on $\mathbb{Z}'$ defined as the translation of $M_{\{\theta\}}$ by $-\lfloor x_0\sqrt{\theta}\rfloor$:

$$M_{\{\theta\}}^{\mathrm{local},x_0} = \{m \in \mathbb{Z}' \,|\, \lfloor x_0\sqrt{\theta}\rfloor + m \in M_{\{\theta\}}\}$$

Assuming that $x_0\sqrt{\theta} \in \mathbb{Z}$, $M_\lambda^{\mathrm{local},x_0}$ is a determinantal point process on $\mathbb{Z}'$ with kernel

$$K_{\{\theta\}}^{\mathrm{local},x_0}(x,y) = K_{\{\theta\}}(x_0\sqrt{\theta}+x, x_0\sqrt{\theta}+y).$$

We shall prove that this translated kernel admits a universal limit when θ goes to $+\infty$: the so-called *discrete sine kernel*, see Theorem 5.34.

- *outside the bulk*: with $x_0 \in \mathbb{R} \setminus [-2,2]$, we shall see that the kernel of the translated point process $M_{\{\theta\}}^{\mathrm{local},x_0}$ converges exponentially fast to 0 if $x_0 > 2$, and to $1_{(x=y)}$ if $x_0 < -2$.
- *at the edge*: in the neighborhood of the edge $\pm 2\sqrt{\theta}$ of our random Young diagram $\lambda \sim \mathrm{PPL}_\theta$, a limiting phenomenon for the random descent coordinates can be observed if we translate *and rescale* the random point process. Thus, it is convenient to introduce the random point process on $\mathbb{R}$ defined by:

$$M_{\{\theta\}}^{\mathrm{edge}} = \left\{x \in \mathbb{R} \,|\, 2\sqrt{\theta} + x\theta^{1/6} \in M_{\{\theta\}}\right\}.$$

We shall see that the corresponding renormalisation of the kernel $K_{\{\theta\}}$ admits a limit, which is the so-called *Airy kernel*; see Theorem 5.36. This limiting result implies a surprising link between large random integer partitions chosen according to the Plancherel measure, and the eigenvalues of large random Hermitian matrices; see Theorem 5.38.

Let us start by analysing the behavior of $M_{\{\theta\}}$ in the bulk $(-2\sqrt{\theta}, 2\sqrt{\theta})$. We consider two half-integers x_θ and y_θ in the neighborhood of $x_0\sqrt{\theta}$:

$$x_\theta = x_0\sqrt{\theta} + x \qquad ; \qquad y_\theta = x_0\sqrt{\theta} + y.$$

We can rewrite the translated kernel $K_{\{\theta\}}^{\mathrm{local},x_0}(x,y)$ as follows:

$$\begin{aligned}
K_{\{\theta\}}^{\mathrm{local},x_0}(x,y) &= K_{\{\theta\}}(x_\theta, y_\theta) \\
&= \frac{1}{(2\mathrm{i}\pi)^2} \oint\!\!\oint \frac{1}{(z-w)\sqrt{zw}} \exp\left(\sqrt{\theta}((z-z^{-1})-(w-w^{-1}))\right) z^{-x_\theta} w^{y_\theta}\, dz\, dw \\
&= \frac{1}{(2\mathrm{i}\pi)^2} \oint\!\!\oint \frac{1}{(z-w)\sqrt{zw}} \exp\left(\sqrt{\theta}\left(F\left(z, \frac{x_\theta}{\sqrt{\theta}}\right) - F\left(w, \frac{y_\theta}{\sqrt{\theta}}\right)\right)\right) dz\, dw
\end{aligned}$$

with $F(z,t) = z - z^{-1} - t\log z$, and where the double contour integral is taken for instance over two circles with radii $|z| > |w|$. This double contour integral will be estimated by a *saddle point analysis*. We start by finding the critical points of $F(\cdot, x)$, where $\frac{\partial F(z,x)}{\partial z} = 0$. They are given by

$$1 + \frac{1}{z^2} = \frac{x}{z} \qquad ; \qquad z = \frac{x \pm \mathrm{i}\sqrt{4-x^2}}{2} = \mathrm{e}^{\pm \mathrm{i}\phi},$$

with $\phi = \arccos(\frac{x}{2})$. Moreover, on the circle with radii 1, $\mathrm{Re}(F(z,x)) = 0$, and the gradient of $\mathrm{Re}(F(z,x))$ is given by

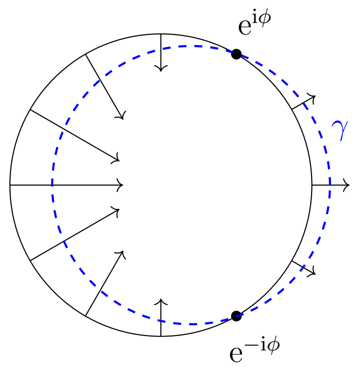
$$\nabla(\mathrm{Re}\, F(\cdot, x))(\mathrm{e}^{\mathrm{i}\theta}) = 2\,(\cos\theta - \cos\phi)\, u_\theta,$$

where u_θ is the unit vector $\mathrm{e}^{\mathrm{i}\theta}$ (perpendicular to the unit circle at $z = \mathrm{e}^{\mathrm{i}\theta}$).

Consider first a holomorphic function in two variables $g(z,w)$ defined in neighborhoods of the two circles $|z| = 1$ (path γ_z) and $|w| = 1 - \varepsilon$ (path γ_w). If we want to evaluate

$$\oint\!\!\oint_{\gamma_z,\,\gamma_w} g(z,w) \exp\left(\sqrt{\theta}\left(F\left(z, \frac{x_\theta}{\sqrt{\theta}}\right) - F\left(w, \frac{y_\theta}{\sqrt{\theta}}\right)\right)\right) dz\, dw,$$

then we can deform the path γ_w into the blue path γ drawn below:



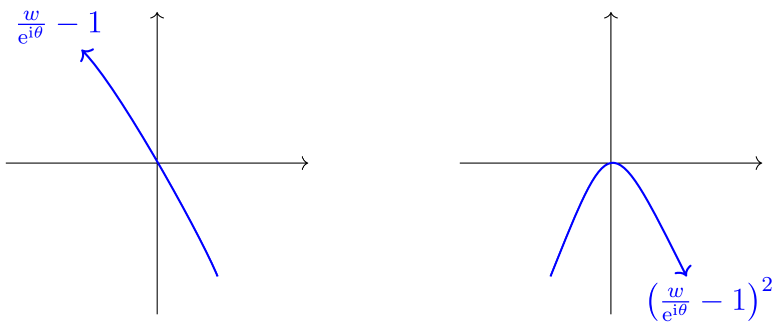
In this drawing, the black circle is the unit circle γ_z; the blue circle γ crosses γ_z at $e^{i\phi}$ and $e^{-i\phi}$, with $\phi = \arccos(\frac{y_\theta}{2\sqrt{\theta}})$; and the black arrows represent the gradient of $F(\cdot, \frac{y_\theta}{\sqrt{\theta}})$. Because of this gradient, $\mathrm{Re}(F(w, \frac{y_\theta}{\sqrt{\theta}})) > 0$ for any $w \in \gamma \setminus \{e^{\pm i\phi}\}$. It follows that the integral

$$\oint_\gamma \left| e^{-\sqrt{\theta}\,F(w,\frac{y_\theta}{\sqrt{\theta}})} \right| \,dw$$

goes to 0 as θ goes to infinity. Indeed, we can use the Laplace method to evaluate this integral: the main contribution to the integral is provided by the neighborhoods of the two critical points, where one can replace $F(w, \frac{y_\theta}{\sqrt{\theta}})$ by its second order Taylor approximation:

$$\begin{aligned} F\left(w, \frac{y_\theta}{\sqrt{\theta}}\right) &\simeq F\left(e^{i\phi}, \frac{y_\theta}{\sqrt{\theta}}\right) + F''\left(e^{i\phi}, \frac{y_\theta}{\sqrt{\theta}}\right) \frac{(w-e^{i\phi})^2}{2} \\ &\simeq 2i(\sin\phi - \phi\cos\phi) + i\sin\phi\,\left(\frac{w}{e^{i\phi}} - 1\right)^2; \\ \mathrm{Re}\left(F\left(w, \frac{y_\theta}{\sqrt{\theta}}\right)\right) &\simeq -(\sin\phi)\,\mathrm{Im}\left(\left(\frac{w}{e^{i\phi}} - 1\right)^2\right). \end{aligned}$$

We have drawn below the aspect of $\frac{w}{e^{i\phi}} - 1$ and of $(\frac{w}{e^{i\phi}} - 1)^2$ when γ crosses γ_z at $w = e^{i\phi}$.



It follows that the contribution to the integral $\oint_\gamma |e^{-\sqrt{\theta}\,F(w,\frac{y_\theta}{\sqrt{\theta}})}|\,dw$ corresponding to the parameters w in the neighborhood of $e^{i\phi}$ behaves as

$$\int e^{-\sqrt{\theta}\,Cs^2}\,ds = O\left(\theta^{-\frac{1}{4}}\right) \to_{\theta\to 0} 0$$

and the contribution of the neighborhood of $e^{-i\phi}$ is of the same order of magnitude. As a consequence,

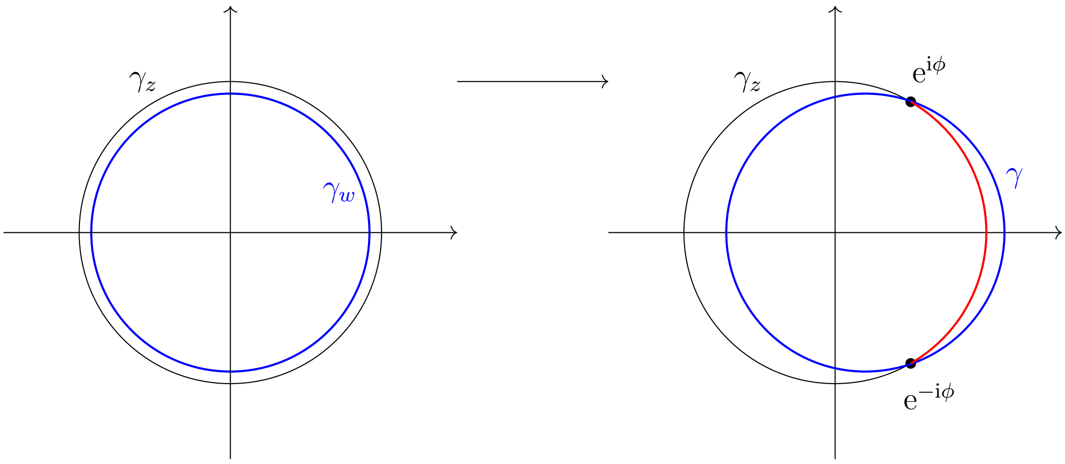
$$\begin{aligned} &\left| \oint\!\!\oint_{z\in\gamma_z,\, w\in\gamma_w} g(z,w)\,\exp\left(\sqrt{\theta}\left(F\left(z,\frac{x_\theta}{\sqrt{\theta}}\right) - F\left(w,\frac{y_\theta}{\sqrt{\theta}}\right)\right)\right)\,dz\,dw \right| \\ &= \left| \oint\!\!\oint_{z\in\gamma_z,\, w\in\gamma} g(z,w)\,\exp\left(\sqrt{\theta}\left(F\left(z,\frac{x_\theta}{\sqrt{\theta}}\right) - F\left(w,\frac{y_\theta}{\sqrt{\theta}}\right)\right)\right)\,dz\,dw \right| \\ &\leq C \oint_{w\in\gamma} \left| e^{-\sqrt{\theta}\,F(w,\frac{y_\theta}{\sqrt{\theta}})} \right|\,dw \to_{\theta\to+\infty} 0 \end{aligned}$$

since $|\exp(\sqrt{\theta}(F(z, \frac{x_\theta}{\sqrt{\theta}})))| = 1$ on γ_z, and since the holomorphic function $g(z, w)$ is bounded on $\gamma_z \times \gamma$.

Now, this is not what happens for the kernels $K^{\mathrm{local}, x_0}_{\{\theta\}}(x, y)$, because of the singularity of the ratio $\frac{1}{z-w}$ when $z = w$ (we do not have a holomorphic function). We can still use the invariance of contour integrals by deformation, but when we deform the path γ_w into γ, we pick up a residue equal to

$$\frac{1}{(2\mathrm{i}\pi)\, z^{x-y+1}}.$$

We then have to integrate this residue over the part of the path γ_z which is crossed by γ during the deformation $\gamma_w \to \gamma$ (the part in red in the figure below).



Notice that as θ goes to infinity, ϕ converges to $\phi_0 = \arccos(\frac{x_0}{2})$. So, we conclude that

$$\lim_{\theta \to \infty} K_{\{\theta\}}(x_\theta, y_\theta) = \frac{1}{2\mathrm{i}\pi} \oint_{\gamma_{\mathrm{red}} : \mathrm{e}^{-\mathrm{i}\phi_0} \to \mathrm{e}^{\mathrm{i}\phi_0}} \frac{1}{z^{x-y+1}}\, dz = \frac{\sin(\phi_0(x-y))}{\pi(x-y)}.$$

By applying Proposition 5.9, we obtain the following result due to Borodin–Okounkov–Olshanski:

THEOREM 5.34 (Convergence in the bulk towards the discrete sine kernel). *For $x_0 \in (-2, 2)$, we denote $\phi_0 = \arccos(\frac{x_0}{2}) \in (0, \pi)$. As θ goes to infinity, $M^{\mathrm{local}, x_0}_{\{\theta\}}$ converges towards the determinantal point process with discrete sine kernel*

$$K^{\mathrm{dsine}, x_0}(x, y) = \frac{\sin \phi_0(x-y)}{\pi(x-y)},$$

the reference measure being the counting measure on $\mathbb{Z}'$.

The limiting point process M^{x_0} which we obtained is translation-invariant on $\mathbb{Z}'$: for any $k \in \mathbb{Z}$, and any finite set of half-integers $A \subset \mathbb{Z}'$, $\mathbb{P}[A \subset M^{x_0}] = \mathbb{P}[A + k \subset M^{x_0}]$. The parameter $\phi_0 = \arccos(\frac{x_0}{2})$ drives the expected density

$$\mathbb{P}[x \subset M^{x_0}] = K^{\mathrm{dsine}, x_0}(x, x) = \frac{\phi_0}{\pi} \in (0, 1)$$

of the points in this point process. This allows us to recover formally Theorem 4.28 as a corollary of Theorem 5.34. Indeed, the formula above ensures that in the vicinity of $x_0\sqrt{\theta}$, each possible descent coordinate is a point of M_λ with probability close to $\frac{1}{\pi} \arccos(\frac{x_0}{2})$. Let us assume that the result translates to non-Poissonised random integer partitions: if $\lambda_n = \lambda \sim \mathrm{PL}_n$, then the random point process M_λ has almost the same properties as if we had $\lambda \sim \mathrm{PPL}_n$. An adequate *dePoissonisation procedure* actually enables such a transfer of results. Then, in terms of the random scaled diagram $\omega_n = (\omega_{\lambda_n})_{\frac{1}{n}}$, the previous estimate of density of descents implies:

$$(\omega_n)'(x_0) \simeq_{n\to\infty} 1 - \frac{2\phi_0}{\pi} = \frac{2}{\pi}\left(\frac{\pi}{2} - \arccos\left(\frac{x_0}{2}\right)\right) = \frac{2}{\pi}\,\arcsin\left(\frac{x_0}{2}\right),$$

which is exactly the derivative $\Omega'(x_0)$, where Ω is the LSKV curve.

In order to complete this approach to the Logan–Shepp–Kerov–Vershik law of large numbers, we also need to prove that the localised point process $M_{\{\theta\}}^{\mathrm{local},x_0}$ converges to a trivial limit if $x_0 \notin [-2,2]$. Let us explain what differs in this case from the previous discussion. We suppose first that $x > 2$. The two critical points of $F(z,x)$ are now:

$$z = \frac{x \pm \sqrt{x^2-4}}{2} = \mathrm{e}^{\pm h}, \quad \text{with } h = \operatorname{arccosh}\left(\frac{x}{2}\right).$$

Moreover, on the circle γ_z with radii 1, $\mathrm{Re}(F(z,x)) = 0$, and the gradient of $\mathrm{Re}(F(z,x))$ is given by

$$\nabla(\mathrm{Re}\, F(\cdot,x))(\mathrm{e}^{\mathrm{i}\theta}) = (2\cos\theta - x)\, u_\theta,$$

so it is always oriented towards the interior of the unit circle. Consequently, and as can be checked by a direct computation, $\mathrm{Re}\, F(\cdot, \frac{y_\theta}{\sqrt{\theta}})$ is bounded from below by a strictly positive constant on the path $\gamma_w = \{w \in \mathbb{C} \,|\, |w| = \mathrm{e}^{-h}\}$. As $\frac{1}{(z-w)\sqrt{zw}}$ is bounded uniformly on $\gamma_z \times \gamma_w$, this implies that

$$K_{\{\theta\}}^{\mathrm{local},x_0}(x,y) = K_{\{\theta\}}(x_\theta, y_\theta)$$

converges exponentially fast to 0, and that $M_{\{\theta\}}^{\mathrm{local},x_0}$ converges to the empty point process.

For $x < -2$, the critical points are $z = -\mathrm{e}^{\pm h}$, with $h = \operatorname{arccosh}(-\frac{x}{2})$; and on the circle γ_z with radii 1, we have the same formulæ as above, whence a real part $\mathrm{Re}(F(z,x))$ which vanishes and which has a gradient oriented towards the exterior of the unit circle. In order to make appear an integral going exponentially fast to 0, we then have to exchange the position of γ_z and γ_w, so that γ_w is the circle with radius e^h (exterior of γ_z). This exchange of position corresponds to a deformation of paths for γ_w, which picks up the residue

$$\lim_{\theta\to\infty} K_{\{\theta\}}(x_\theta, y_\theta) = \frac{1}{2\mathrm{i}\pi}\oint_{\gamma_z} \frac{1}{z^{x-y+1}}\,dz = 1_{(x=y)}.$$

This limiting kernel corresponds to the full point process $\mathbb{Z}'$: for any $A \subset \mathbb{Z}'$, $\mathbb{P}[A \subset M_{\{\theta\}}^{\mathrm{local},x_0}] \to \det(\mathrm{id}) = 1$.

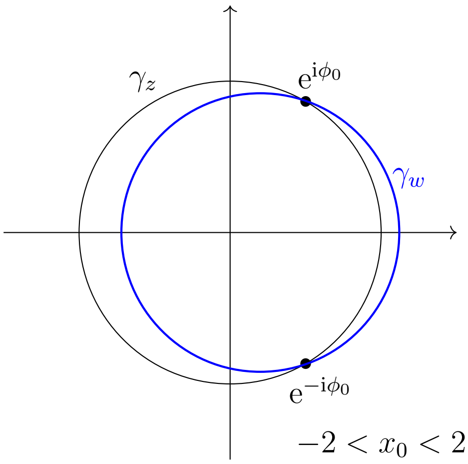
Proposition 5.35 (Trivial limit outside the bulk). *For $x_0 > 2$, $M_{\{\theta\}}^{\mathrm{local},x_0}$ converges towards the empty point process, and for $x_0 < -2$, $M_{\{\theta\}}^{\mathrm{local},x_0}$ converges towards the full point process $\mathbb{Z}'$.*

Again, this is formally related to the formulas $\Omega'(x_0) = 1$ for $x_0 > 2$ and $\Omega'(x_0) = -1$ for $x_0 < -2$, so it agrees with the LSKV theorem 4.28.

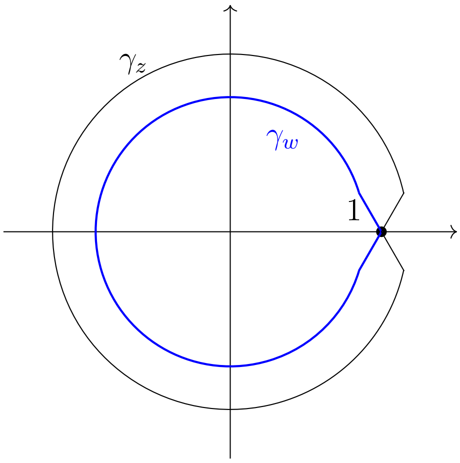
We now focus on the behavior of $M_{\{\theta\}}$ at the edge, that is to say in the vicinity of $2\sqrt{\theta}$ (by symmetry of the Plancherel-distributed random integer partitions, it suffices to study the right edge, the left edge in the vicinity of $-2\sqrt{\theta}$ exhibiting a symmetric behavior). The goal is to understand the asymptotics of the kernel $K_{\{\theta\}}(2\sqrt{\theta}+tx, 2\sqrt{\theta}+ty)$ when θ goes to infinity, and t is a suitable rescaling parameter, which later will be taken equal to $\theta^{1/6}$. We remark that the critical points of the action $F(z,x_0)$, which are complex conjugate numbers when $|x_0| < 2$, and which sit on the real line when $|x_0| > 2$, coalesce when $x_0 = \pm 2$, and yield a double critical point at $z = 1$ for $x_0 = 2$ (respectively, at $z = -1$ for $x_0 = -2$). In the neighborhood of $z = 1$, we therefore have:

$$F(z,2) = F(1,2) + \frac{F'''(1,2)}{6}\,(z-1)^3 + O(|z-1|^4) = \frac{(z-1)^3}{3} + O(|z-1|^4).$$

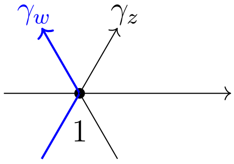
We have drawn below the adequate contours of integration γ_w and γ_z when $x_0 \in (-2, 2)$ and when $x_0 > 2$:



By adequate we mean that they yield integrals which converge exponentially fast to 0, by means of the Laplace method. For each case, $\mathrm{Re}(F(z, x_0))$ vanishes on γ_z, whereas $\mathrm{Re}(F(w, x_0))$ attains its maximum along γ_w at the critical point(s). In the first case, we had to perform a deformation of the initial path γ_w which led to the appearance of a non-trivial residue, whence a non-trivial limit for the localised kernel. Now, it is clear that when $x_0 = 2$, an adequate pair of paths of integration will be the geometric middle point between the two previous drawings, that is:



More precisely, let us make the following choices. The path γ_z is a circle with radius $1 + \varepsilon$, except in a small neighborhood of 1, where it follows the lines $1 + s\mathrm{e}^{\pm \frac{\mathrm{i}\pi}{3}}$, $s > 0$; whereas the path γ_w is a circle with radius $1 - \varepsilon$, except in a small neighborhood of 1, where it follows the lines $1 + s\mathrm{e}^{\pm \frac{2\mathrm{i}\pi}{3}}$, $s > 0$. Thus, if we zoom in on the neighborhood of 1, then we have the following configuration:



For $x_0 = 2$, the gradient of $\mathrm{Re}(F(z, x_0))$ on the unit circle is still oriented towards the interior, except at $z = 1$ where it vanishes. Therefore, one can show that on the part of γ_z which coincices with the circle with radius $1+\varepsilon$ (respectively, on the part of γ_w which coincides with the circle with radius $1-\varepsilon$), $\mathrm{Re}(F(z, 2))$ is bounded from above by a negative constant (respectively, $\mathrm{Re}(F(w, 2))$ is bounded from below by a positive constant). It is also not difficult to prove that the same phenomenon occurs when replacing $x = 2$ respectively by $2 + \frac{tx}{\sqrt{\theta}}$ and by $2 + \frac{ty}{\sqrt{\theta}}$, if $t \ll \sqrt{\theta}$. Therefore, the corresponding contributions to the double path integral will converge exponentially fast to 0. So, the main contribution comes from the parts of the paths which are lines stemming from $z = 1$

or $w = 1$, and

$$K_{\{\theta\}}(2\sqrt{\theta}+tx, 2\sqrt{\theta}+ty) \simeq \frac{1}{(2\mathrm{i}\pi)^2} \oint\oint_{1} \frac{1}{(z-w)\sqrt{zw}}\, \mathrm{e}^{\sqrt{\theta}\left(F\left(z,2+\frac{tx}{\sqrt{\theta}}\right)-F\left(w,2+\frac{ty}{\sqrt{\theta}}\right)\right)}\, dz\, dw.$$

If $w = 1 + t^{-1}w'$ and $t = \theta^{1/6}$, then

$$F\left(w, 2+\frac{ty}{\sqrt{\theta}}\right) = w - w^{-1} - (2+y\theta^{-1/3})\log w = \frac{-yw' + \frac{w'^3}{3}}{\sqrt{\theta}} + O(\theta^{-2/3}(w'^2 + w'^4))$$

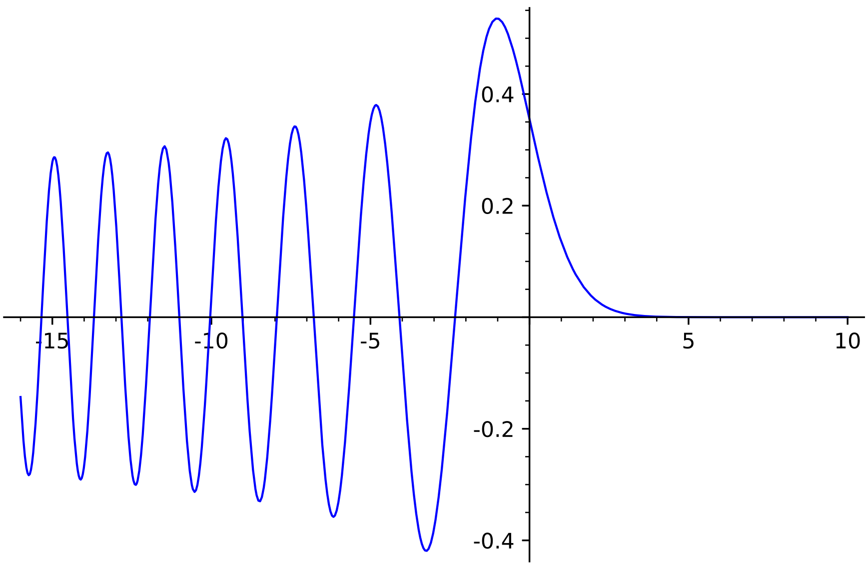
Making the same change of variables $z = 1+t^{-1}z'$, and assuming that $2\sqrt{\theta}+x\theta^{1/6}$ and $2\sqrt{\theta}+y\theta^{1/6}$ belong to $\mathbb{Z}'$, we get:

$$K_{\{\theta\}}(2\sqrt{\theta}+x\theta^{1/6}, 2\sqrt{\theta}+y\theta^{1/6}) \simeq \frac{\theta^{-1/6}}{(2\mathrm{i}\pi)^2} \oint\oint_{0} \frac{1}{(z'-w')}\, \mathrm{e}^{yw'-\frac{w'^3}{3}-xz'+\frac{z'^3}{3}}\, dz'\, dw',$$

where the two paths of integration now meet at 0 and go to infinity with directions $\mathrm{e}^{\pm\frac{\mathrm{i}\pi}{3}}, \mathrm{e}^{\pm\frac{2\mathrm{i}\pi}{3}}$. At this point, it is convenient to introduce the *Airy function*, which is the entire function defined by

$$\mathrm{Ai}(x) = \frac{1}{2\mathrm{i}\pi} \int \mathrm{e}^{xy-\frac{y^3}{3}}\, dy,$$

the path of integration being the union of the two half-lines which meet at 0 and which have directions $\mathrm{e}^{\pm\frac{2\mathrm{i}\pi}{3}}$. This special function satisfies the differential equation $\mathrm{Ai}''(x) - x\,\mathrm{Ai}(x) = 0$, and its graph as a function of a real parameter x is drawn below.



It can also be redefined as the real semi-convergent integral $\mathrm{Ai}(x) = \frac{1}{\pi}\int_0^\infty \cos(xy + \frac{y^3}{3})\, dy$, and its Fourier transform is

$$\widehat{\mathrm{Ai}}(\xi) = \mathrm{e}^{\frac{(\mathrm{i}\xi)^3}{3}}.$$

The reader will find more details on this special function in the references at the end of the chapter. Now, a small computation similar to the one performed in the proof of Proposition 5.33 and involving the operator $\frac{\partial}{\partial z'} + \frac{\partial}{\partial w'}$ enables one to transform the integral obtained above and rewrite it as:

$$K_{\{\theta\}}(2\sqrt{\theta}+x\theta^{1/6}, 2\sqrt{\theta}+y\theta^{1/6}) \simeq \theta^{-1/6}\, K^{\mathrm{Airy}}(x,y),$$

where K^{Airy} is the *Airy kernel* defined on real parameters by:

$$K^{\mathrm{Airy}}(x,y) = \frac{\mathrm{Ai}(x)\,\mathrm{Ai}'(y) - \mathrm{Ai}'(x)\,\mathrm{Ai}(y)}{x-y} = \int_0^\infty \mathrm{Ai}(s+x)\,\mathrm{Ai}(s+y)\,ds.$$

This leads to the following:

THEOREM 5.36 (Convergence at the edge towards the Airy kernel). *Consider the random point process $M^{\mathrm{edge}}_{\{\theta\}}$ on $\mathbb{R}$ defined by*

$$M^{\mathrm{edge}}_{\{\theta\}} = \{x \in \mathbb{R} \,|\, 2\sqrt{\theta} + x\theta^{1/6} \in M_{\{\theta\}}\},$$

where $M_{\{\theta\}}$ denotes the set of descent coordinates of a random integer partition $\lambda \sim \mathrm{PPL}_\theta$. As θ goes to infinity, $M^{\mathrm{edge}}_{\{\theta\}}$ converges towards the determinantal point process whose kernel is the Airy kernel $K^{\mathrm{Airy}}(x,y)$, *the reference measure being the Lebesgue measure on $\mathbb{R}$.*

PROOF. In the sequel, we denote $I(x,\theta)$ the indicator function of the condition $2\sqrt{\theta} + x\theta^{1/6} \in \mathbb{Z}'$. Given finite intervals $A_1,\ldots,A_n \subset \mathbb{R}$, we have:

$$\begin{aligned}
&\mathbb{E}[(M^{\mathrm{edge}}_{\{\theta\}})^{\downarrow n}(A_1 \times A_2 \times \cdots \times A_n)]\\
&= \mathbb{E}[(M_{\{\theta\}})^{\downarrow n}(2\sqrt{\theta} + \theta^{1/6}A_1 \times \cdots \times 2\sqrt{\theta} + \theta^{1/6}A_n)]\\
&= \sum_{\substack{k_1 \in (2\sqrt{\theta}+\theta^{1/6}A_1)\cap\mathbb{Z}' \\ \vdots \\ k_n \in (2\sqrt{\theta}+\theta^{1/6}A_n)\cap\mathbb{Z}'}} \det\left((K_{\{\theta\}}(k_i,k_j))_{1\leq i,j\leq n}\right)\\
&= \sum_{x_1\in A_1,\ldots,x_n\in A_n} I(x_1,\theta)\cdots I(x_n,\theta)\,\det\left((K_{\{\theta\}}(2\sqrt{\theta}+x_i\theta^{1/6}, 2\sqrt{\theta}+x_j\theta^{1/6}))_{1\leq i,j\leq n}\right)\\
&\simeq \sum_{x_1\in A_1,\ldots,x_n\in A_n} (\theta^{-1/6}I(x_1,\theta))\cdots(\theta^{-1/6}I(x_n,\theta))\,\det\left((K^{\mathrm{Airy}}(x_i,x_j))_{1\leq i,j\leq n}\right).
\end{aligned}$$

On the last line, we use the fact that the approximation of kernels discussed above can be shown to be locally uniform, so that we can sum the equivalences over points $x_1 \in A_1,\ldots,x_n \in A_n$. Now, for each index i, we have the Riemann sum approximation

$$\sum_{x_i \in A_i} \theta^{-1/6} I(x_i,\theta)(\cdot) \simeq \int_{A_i} (\cdot)\,dx_i,$$

where $(\cdot)$ can be replaced by any continuous bounded function of x_i. The same approximation holds for the multiple sum and the corresponding multiple integral, therefore,

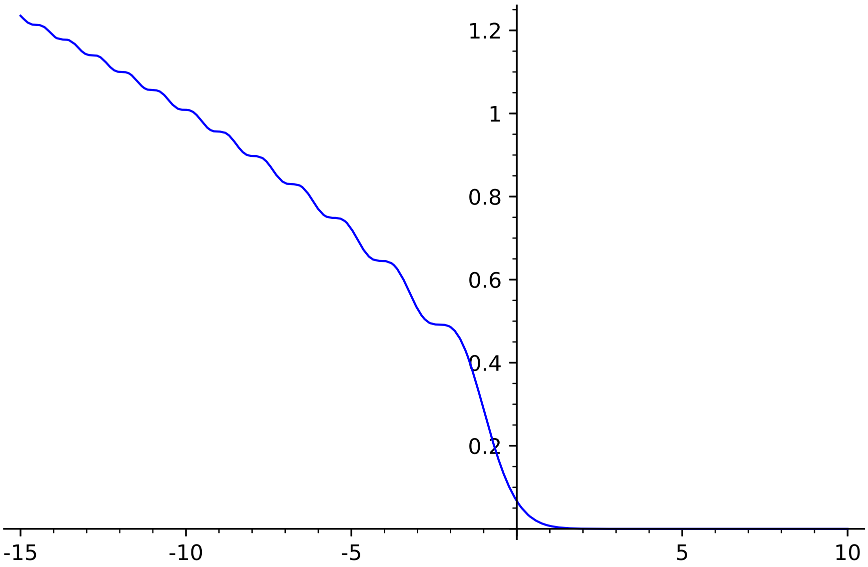
$$\lim_{\theta\to\infty} \mathbb{E}[(M^{\mathrm{edge}}_{\{\theta\}})^{\downarrow n}(A_1 \times A_2 \times \cdots \times A_n)] = \int_{A_1\times\cdots\times A_n} \det\left((K^{\mathrm{Airy}}(x_i,x_j))_{1\leq i,j\leq n}\right)\,dx_1\cdots dx_n.$$

Then, the same kind of argument as in the proof of Proposition 5.9 ensures the convergence of the random point process $M^{\mathrm{edge}}_{\{\theta\}}$ towards the Airy point process. □

To close this chapter, let us discuss an important consequence of Theorem 5.36 regarding the largest rows of a random partition chosen according to the Plancherel measure. The density (first correlation function) of the Airy point process is

$$\rho_1(t) = (\mathrm{Ai}'(t))^2 - \mathrm{Ai}(t)\,\mathrm{Ai}''(t) = (\mathrm{Ai}'(t))^2 - t\,(\mathrm{Ai}(t))^2,$$

and it is drawn below.



The asymptotics of the Airy function and its derivative as t goes to $\pm\infty$ can be derived by a saddle point analysis of the integral representation of $\mathrm{Ai}(t)$. Hence, one can show that

$$\rho_1(t) \simeq_{t\to+\infty} \frac{\mathrm{e}^{-\frac{4}{3}t^{\frac{3}{2}}}}{8\pi t} \qquad ; \qquad \rho_1(t) \simeq_{t\to-\infty} \frac{\sqrt{|t|}}{\pi}.$$

The exponential decay of $\rho_1(t)$ for t positive leads one to guess that with high probability, there are very few positive points of the Airy process (and in particular, only a finite number). On the other hand, we admit in the sequel that by using a dePoissonisation procedure, one can transfer the convergence result from Theorem 5.36 to the random integer partitions $\lambda \sim \mathrm{PL}_n$ when $n \to \infty$. Then, one expects that the largest part $\ell_n = \lambda_{n,1}$ of a random integer partition $\lambda_n \sim \mathrm{PL}_n$ is asymptotically of size

$$2\sqrt{n} + n^{1/6} X^{\mathrm{Airy}},$$

where X^{Airy} is *the largest point of the Airy point process*. This is a strong improvement over Corollary 4.29 for the size of the longest increasing subsequence in a random uniform permutation. Let us make this a bit more precise. The operator $\mathscr{K}^{\mathrm{Airy}}$ associated to the Airy kernel is trace class over any set $[t,+\infty)$, because of the representation $K^{\mathrm{Airy}}(x,y) = \int_0^\infty \mathrm{Ai}(x+s)\mathrm{Ai}(y+s)\,ds$. We can therefore use the formula for gap probabilities in order to compute the cumulative distribution function of X^{Airy}:

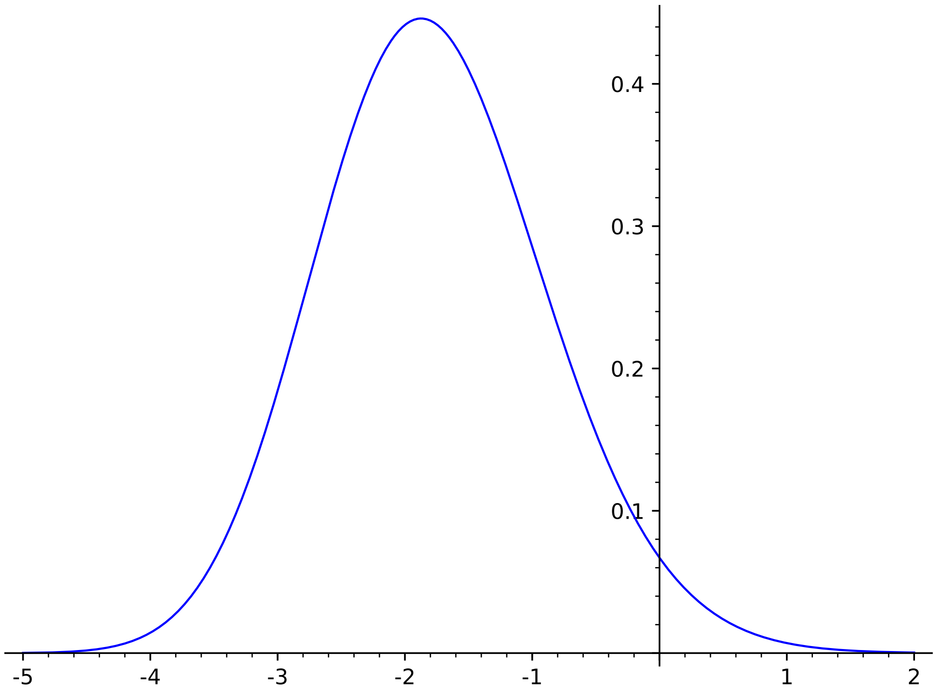
$$\begin{aligned}\mathbb{P}[X^{\mathrm{Airy}} \leq t] &= \mathbb{P}[M^{\mathrm{Airy}}(t,+\infty) = 0] = \det\Big(I - \mathscr{K}^{\mathrm{Airy}}_{|(t,+\infty)}\Big) \\ &= 1 + \sum_{m=1}^{\infty} \frac{(-1)^m}{m!} \int_{(t,+\infty)^m} \det(\mathrm{Ai}(x_i,x_j))_{1\leq i,j\leq m}\,dx_1\cdots dx_m.\end{aligned}$$

This distribution is known as the *Tracy–Widom distribution*, and it satisfies the following properties:

THEOREM 5.37 (Tracy–Widom). *Let M^{Airy} be a determinantal point process on $\mathbb{R}$ with kernel K^{Airy}. The largest point X^{Airy} of M^{Airy} exists almost surely, and it satisfies*

$$\mathbb{P}[X^{\mathrm{Airy}} \leq t] = F_2(t) = \exp\left(-\int_t^\infty (x-t)\,(q(x))^2\,dx\right),$$

where q is the unique solution of the Painlevé II equation $q''(x) = x\,q(x) + 2\,(q(x))^3$ such that $q(x) \simeq \mathrm{Ai}(x)$ as x goes to $+\infty$. The density of the F_2 distribution is drawn below.



We can finally state the following deep result, to be compared with Corollary 4.29:

THEOREM 5.38 (Baik–Deift–Johansson). *Let ℓ_n be the length of a longest increasing subsequence in a random permutation σ_n chosen uniformly in $\mathfrak{S}(n)$. We have the convergence in law:*

$$n^{1/3}\left(\frac{\ell_n}{2\sqrt{n}}-1\right) \rightharpoonup_{n\to\infty} \mathrm{TW},$$

where TW *denotes the Tracy–Widom F_2 distribution.*

It turns out that this Tracy–Widom distribution also describes the fluctuations of large random Hermitian matrices. Consider a Hermitian matrix H_N of size $N \times N$, with independent real diagonal entries $(H_N)_{i,i}$ distributed according to the standard real normal distribution $\mathcal{N}_{\mathbb{R}}(0,1)$, and with independent off-diagonal entries $(H_N)_{i,j} = \overline{(H_N)_{j,i}}$ distributed according to the standard complex normal distribution $\mathcal{N}_{\mathbb{C}} = \mathcal{N}_{\mathbb{R}}(0,\frac{1}{2}) + \mathrm{i}\mathcal{N}_{\mathbb{R}}(0,\frac{1}{2})$. We denote $x_{N,1} \geq x_{N,2} \geq \cdots \geq x_{N,N}$ the real eigenvalues of the random matrix H_N; they form a random point process L_N on $\mathbb{R}$, which can be shown to be determinantal. Then, we have the following matrix analogues of Theorems 5.34 and 5.36:

THEOREM 5.39 (Gaudin–Mehta). *For $x_0 \in (-2,2)$, we consider the rescaled point process*

$$L_N^{\mathrm{local},x_0} = \left\{ x \in \mathbb{R} \,\middle|\, x_0\sqrt{N} + \frac{2\pi x}{\sqrt{(4-(x_0)^2)N}} \in L_N \right\}.$$

As N goes to infinity, L_N^{local,x_0} converges to the determinantal point process with sine kernel

$$K^{\mathrm{sine}}(x,y) = \frac{\sin \pi(x-y)}{\pi(x-y)},$$

the reference measure being the Lebesgue measure on $\mathbb{R}$.

THEOREM 5.40 (Tracy–Widom). *We consider the rescaled point process of the largest eigenvalues of H_N:*

$$L_N^{\mathrm{edge}} = \{x \in \mathbb{R} \,|\, 2\sqrt{N} + N^{-1/6}x \in L_N\}.$$

As N goes to infinity, L_N^{edge} converges towards the Airy determinantal point process. In particular, the largest eigenvalue $x_{N,1}$ of H_N satisfies the following analogue of Theorem 5.38:

$$N^{2/3}\left(\frac{x_{N,1}}{2\sqrt{N}}-1\right) \rightharpoonup_{N\to\infty} \mathrm{TW}.$$

References

The theory of determinantal point processes is beautifully explained in [Hou+09], with applications to the study of zeroes of random analytic functions. Many important theoretical results on the subject are due to Soshnikov, see [Sos00a; Sos00b]. Our Theorem 5.3 is due to Lenard, see [Len73; Len75]: it allows one to focus on the correlation functions when studying determinantal point processes. For the theory of general random point processes (not necessarily determinantal), see [DV88; Kal17].

The classification of the positive specialisations of Sym is due to Thoma [Tho64]; an earlier result of Edrei (see [Ais+51]) is equivalent to it. A probabilistic proof of this classification, which uses random characters of large symmetric groups in order to approximate the extremal characters of $\mathfrak{S}(\infty)$, is due to Kerov and Vershik [KV81]. More recently, the irreducible admissible representations of the Gelfand pair $(\mathfrak{S}(\infty)\times\mathfrak{S}(\infty),\mathfrak{S}(\infty))$ have been classified by Olshanski and Okounkov, see [Ols90; Oko99]; the extremal characters of $\mathfrak{S}(\infty)$ correspond to the particular case of the spherical representations of this Gelfand pair.

The notion of Schur measure is due to Okounkov, and it appears in [Oko01a]; it explains several calculations which were performed in order to solve the Baik–Deift–Johansson conjecture on the largest parts of a random integer partition under the Plancherel measure. Baik, Deift and Johansson studied the first part λ_1 of a random partition in [BDJ99], and they conjectured the asymptotic behavior of the whole collection of the largest parts $\lambda_1 \geq \lambda_2 \geq \cdots$. This conjecture was solved by several methods in [BOO00; Oko00; Joh01], and the formalism of the infinite wedge space make clearer certain calculations of the solution. We also refer to [Oko01b], which explains the saddle point analysis of the renormalised kernels, and the appearance at the limit of the discrete sine and Airy kernels. For a general survey of the methods of asymptotic analysis of integrals (saddle point analysis, steepest descent method, *etc.*), see [Bru10, Chapters 5-6] and [Won01, Section II.4]. For the properties of the Bessel and Airy functions, we refer to [Olv97, Chapter 2, §8-9].

The Tracy–Widom distribution which is involved in Theorems 5.37-5.38 first appeared in the study of the largest eigenvalue of large random Hermitian matrices; see [TW94]. For a general study of the asymptotic behavior of the (determinantal) point process associated to the eigenvalues of random matrices, see [AGZ10, Chapter 3].

Exercises

(1) **Asymptotic independence of the localised point processes.** Theorem 5.34 ensures the convergence of the localised point process $M_{\{\theta\}}^{\mathrm{local},x_0}$ for $x_0 \in (-2,2)$. Adapt its proof in order to show that, if $x_0 \neq y_0$ are two points in $(-2,2)$, then the corresponding point processes $M_{\{\theta\}}^{\mathrm{local},x_0}$ and $M_{\{\theta\}}^{\mathrm{local},y_0}$ are asymptotically independent when θ goes to infinity.

(2) **The probabilistic proof of the Thoma theorem.** In this exercise, we follow the approach of Kerov and Vershik in order to prove the Thoma classification of the extremal characters of $\mathfrak{S}(\infty)$. We fix an extremal character χ of $\mathfrak{S}(\infty)$, and we denote $(\mathbb{P}_{\chi,n})_{n\in\mathbb{N}}$ the corresponding family of central measures. Let $\mathscr{T}$ be the set of all infinite paths $T = (\emptyset \nearrow \lambda^{(-1)} \nearrow \lambda^{(-2)} \nearrow \cdots \nearrow \lambda^{(-n)} \nearrow \lambda^{(-n-1)} \nearrow \cdots)$ which rise along the Young graph: for any $n \in \mathbb{N}$, $\lambda^{(-n)} \in \mathfrak{Y}(n)$, and $\lambda^{(-n)} \nearrow \lambda^{(-n-1)}$. The reason of the negative indices will

appear later. A path $T \in \mathscr{T}$ can be considered as an infinite standard tableau, and it is entirely determined by the maps:

$\square_{n\geq 1} : \mathscr{T} \to \mathbb{N}^2;$

$T \mapsto$ coordinates (x_n, y_n) of the bottom left corner of the cell $\lambda^{(-n-1)} \setminus \lambda^{(-n)}$.

In particular, there is a natural σ-field $\mathcal{F}$ on $\mathscr{T}$, which is the smallest σ-field making all the maps $\square_n$ measurable. A *backward* filtration $(\mathcal{F}_n)_{n\leq 0}$ on $\mathscr{T}$ is provided by the σ-subfields:

$$\mathcal{F}_{-n} = \sigma(\square_{n+1}, \square_{n+2}, \ldots).$$

The information contained in $\mathcal{F}_{-n}$ is the position of the cells $n+1, n+2, n+3, \ldots$ in an infinite tableau T, so in particular $\lambda^{(-n)}$ is a $\mathcal{F}_{-n}$-measurable random variable with values in $\mathfrak{Y}(n)$ (it is the complement of the aforementioned cells). What stays unknown at time $-n$ is the standard Young tableau $T_{|[\![1,n]\!]} \in \mathrm{ST}(\lambda^{(-n)})$.

(a) We consider the Markov chain $(\lambda^{(n)})_{n\leq 0}$ on $\mathfrak{Y} = \bigsqcup_{n\in\mathbb{N}} \mathfrak{Y}(n)$ which is the descending random walk with transition probabilities

$$p(\Lambda, \lambda) = \mathbb{P}[\lambda^{(-n)} = \lambda \,|\, \lambda^{(-n-1)} = \Lambda] = \frac{\dim \lambda}{\dim \Lambda}.$$

Recall why these formulas yield transition probabilities. Suppose that $\lambda^{(-N)}$ is distributed according to $\mathbb{P}_{\chi,N}$. Show then that $\lambda^{(-n)}$ is distributed according to $\mathbb{P}_{\chi,n}$ for any $n \leq N$.

By standard arguments from measure theory (Kolmogorov theorem), we can therefore construct for any extremal character (or more generally, for any normalised trace) χ on $\mathfrak{S}(\infty)$ a probability measure on $\mathscr{T}$ such that the random process $T = (\lambda^{(n)})_{n\leq 0}$ is a Markov chain with transition probabilities $p(\Lambda, \lambda) = \frac{\dim \lambda}{\dim \Lambda}$ and with marginal laws the central measures $\mathbb{P}_{\chi,n}$. In the following we fix this Markov chain indexed by the negative integers, and we associate to it backward martingales.

(b) Given μ integer partition with size k, we associate to it a permutation $\sigma_\mu \in \mathfrak{S}(k)$ with cycle-type μ, and we consider the random process $(\chi^{\lambda^{(n)}}(\sigma_\mu))_{n\leq -k}$, the χ^λ's being the normalised irreducible characters of the symmetric groups. Show that it is a bounded backward martingale with respect to the backward filtration $(\mathcal{F}_n)_{n\leq -k}$:

$$\mathbb{E}[\chi^{\lambda^{(-n)}}(\sigma_\mu) \,|\, \mathcal{F}_{-n-1}] = \chi^{\lambda^{(-n-1)}}(\sigma_\mu).$$

(c) Recall that a bounded backward martingale $(X_n)_{n\leq -k}$ converges almost surely and in moments towards $X_{-\infty} = \mathbb{E}[X_{-k}|\, \mathcal{F}_{-\infty}]$, where $\mathcal{F}_{-\infty} = \bigcap_{n\leq -k} \mathcal{F}_n$. We denote in the sequel

$$\chi^{\lambda^{(-\infty)}}(\sigma_\mu) = \lim_{n\to -\infty} \chi^{\lambda^{(-n)}}(\sigma_\mu).$$

We also introduce the non-negative normalised specialisation X of Sym associated by Corollary 5.16 to the extremal character χ. Prove that

$$\mathbb{E}[\chi^{\lambda^{(-\infty)}}(\sigma_\mu)] = \chi(\sigma_\mu) = p_\mu(X).$$

Recall from Chapter 4 the normalised character values: $\varSigma_\mu(\lambda) = n^{\downarrow|\mu|}\, \chi^\lambda(\sigma_\mu)$ if $n = |\lambda|$. These observables belong to the graded algebra $\mathscr{O}$, which is endowed with the degree $\deg \varSigma_\mu = |\mu|$. We admit the following consequence of the Frobenius–Schur formula: if $(a_1, \ldots, a_s; -b_1, \ldots, -b_s) = F_\lambda$ are the Frobenius coordinates of λ, and if

$$p_k(\lambda) = \sum_{i=1}^{s} (a_i)^k - \sum_{i=1}^{s} (-b_i)^k,$$

then $\mathscr{O} = \mathbb{R}[p_1, p_2, \ldots, p_k, \ldots]$, and $\Sigma_k - p_k$ is an observable of degree smaller than $k-1$ for any $k \geq 1$ (this follows from a two-page computation similar to the one of Chapter 3, Exercise 2). In particular, $\deg p_k = \deg \Sigma_k = k$. Notice that the Frobenius coordinates of λ write as:

$$a_i = \lambda_i - i + \frac{1}{2} \qquad ; \qquad b_i = \lambda'_i - i + \frac{1}{2}.$$

These relations are obvious if one draws the Young diagram of λ rotated by 45 degrees, as in the beginning of this chapter.

(d) Show that more generally, for any integer partition μ, if $p_\mu(\lambda) = \prod_{i=1}^{\ell(\mu)} p_{\mu_i}(\lambda)$, then $\Sigma_\mu(\lambda) - p_\mu(\lambda)$ is an observable of λ with degree strictly smaller than $|\mu|$. Deduce from this observation that for any integer partition $\lambda \in \mathfrak{Y}(n)$,

$$\chi^\lambda(\sigma_\mu) = \frac{p_\mu(\lambda)}{n^{|\mu|}} + O(n^{-1})$$

with a $O(n^{-1})$ which depends on μ but not on $\lambda \in \mathfrak{Y}(n)$.

(e) We associate to any integer partition λ with size larger than 1 a probability measure ν_λ on $[-1, 1]$:

$$\nu_\lambda = \frac{1}{|\lambda|} \left(\sum_{i=1}^{s} a_i\, \delta_{a_i} + b_i\, \delta_{-b_i} \right),$$

where $(a_1, \ldots, a_s; -b_1, \ldots, -b_s)$ are the Frobenius coordinates of λ. Given a sequence of integer partitions $(\lambda^{(n)})_{n \leq 0}$ with $|\lambda^{(-n)}| = n$ for any n, prove the equivalence between the following statements:

- There exists *limiting frequencies* $\alpha_{i \geq 1}$ and $\beta_{i \geq 1}$ such that $\sum_{i=1}^\infty \alpha_i + \beta_i \leq 1$, and such that $\frac{a_i(\lambda^{(-n)})}{n} \to \alpha_i$ and $\frac{b_i(\lambda^{(-n)})}{n} \to \beta_i$ for any $i \geq 1$. The pair (α, β) belongs then necessarily to the Thoma simplex Ω.
- The measures $\nu_{\lambda^{(-n)}}$ admit a weak limit ν, which is a probability measure on $[-1, 1]$.
- The observables $p_k(\lambda^{(-n)})$ satisfy $\frac{p_k(\lambda^{(-n)})}{n^k} \to_{n \to \infty} c_k$ for some parameters $c_{k \geq 1}$.

(f) Show that for the Markov chain $(\lambda^{(n)})_{n \leq 0}$ associated to an extremal character χ and a non-negative specialisation X of Sym, almost surely, the random observables $p_k(\lambda^{(-n)})$ satisfy the third criterion above with $c_k = p_k(X)$. Deduce from this that the specialisation X is the one associated to the Thoma parameter (α, β).

(g) Prove also that $\chi^{\lambda^{(-\infty)}}(\sigma) = \chi(\sigma)$ almost surely, and that the largest rows and columns of the random integer partitions $(\lambda^{(n)})_{n \leq 0}$ satisfy the law of large numbers:

$$\frac{(\lambda^{(-n)})_i}{n} \to_{n \to \infty,\, \text{a.s.}} \alpha_i \qquad ; \qquad \frac{(\lambda^{(-n)})'_i}{n} \to_{n \to \infty,\, \text{a.s.}} \beta_i$$

for any $i \geq 1$.

(3) **Schur–Weyl measures are central measures.** We recall from Chapter 4, Exercises 2 and 3 the following definition: the Schur–Weyl measure with parameters N and n is the probability measure on $\mathfrak{Y}(n)$ given by

$$\mathrm{SW}_{N,n}[\lambda] = \frac{\mathrm{card}(\mathrm{SST}(\lambda, N))\, \mathrm{card}(\mathrm{ST}(\lambda))}{N^n}.$$

This is also the spectral measure associated to the representation $V = (\mathbb{C}^N)^{\otimes n}$ of $\mathfrak{S}(n)$, which acts by permutation on the tensors.

(a) By combining the Frobenius–Schur formula and the formula $\mathrm{ch}^V(\sigma_\mu) = N^{\ell(\mu)}$, show that for any integer partition λ,

$$\mathrm{card}(\mathrm{SST}(\lambda, N)) = s_\lambda(\underbrace{1, 1, \ldots, 1}_{N \text{ terms}}) = s_\lambda(1^N).$$

(b) We consider the Poissonised version of the Schur–Weyl measures: $\mathrm{PSW}_{N,\theta}$ is the probability measure on $\mathfrak{Y}$ given by:

$$\mathrm{PSW}_{N,\theta}[\lambda] = \frac{\mathrm{e}^{-\theta}\,\theta^n}{n!}\,\mathrm{SW}_{N,n}[\lambda]$$

if $|\lambda| = n$. Show that this Poissonised Schur–Weyl measure is the Schur measure for the alphabets

$$A = \left(\underbrace{\frac{\sqrt{\theta}}{N}, \frac{\sqrt{\theta}}{N}, \ldots, \frac{\sqrt{\theta}}{N}}_{N \text{ terms}}\right) \qquad ; \qquad B = \sqrt{\theta}\,E.$$

(c) We take $N = \frac{\sqrt{\theta}}{c}$; this is the same normalisation as in Chapter 4, Exercise 3. Adapt to this case the saddle point analysis of the kernels performed for the Poissonised Plancherel measures, by replacing the function

$$F(z,t) = z - z^{-1} - t\log z$$

by

$$F_c(z,t) = -\frac{1}{c}\log(1-cz) - z^{-1} - t\log z.$$

Prove in particular the analogues of Theorems 5.34 and 5.36 for the Poissonised Schur–Weyl measures in the thermodynamic limit $N = \frac{\sqrt{\theta}}{c}$, $\theta \to +\infty$.

Bibliography

[Ais+51] M. Aissen, A. Edrei, I. Schoenberg, and A. Whitney. "On the generating functions of totally positive sequences". In: *Proc. Natl. Acad. Sci. USA* 37.5 (1951), pp. 303–307.

[Ald83] D. Aldous. "Random walks on finite groups and rapidly mixing Markov chains". In: *Seminar on Probability XVII*. Vol. 986. Lecture Notes in Mathematics. Springer Verlag, 1983, pp. 243–297.

[AD86] D. Aldous and P. Diaconis. "Shuffling cards and stopping times". In: *Amer. Math. Monthly* 93.5 (1986), pp. 333–348.

[AGZ10] G. W. Anderson, A. Guionnet, and O. Zeitouni. *An Introduction to Random Matrices*. Vol. 118. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 2010.

[BDJ99] J. Baik, P. Deift, and K. Johansson. "On the distribution of the length of the longest increasing subsequence of random permutations". In: *J. Amer. Math. Soc.* 12.4 (1999), pp. 1119–1178.

[BD92] D. Bayer and P. Diaconis. "Trailing the dovetail shuffle to its lair". In: *Ann. Appl. Prob.* 2.2 (1992), pp. 294–313.

[Bia98] P. Biane. "Representations of symmetric groups and free probability". In: *Adv. Math.* 138 (1998), pp. 126–181.

[BOO00] A. Borodin, A. Okounkov, and G. Olshanski. "Asymptotics of Plancherel measures for symmetric groups". In: *J. Amer. Math. Soc.* 13 (2000), pp. 491–515.

[Bru10] N. G. de Bruijn. *Asymptotic Methods in Analysis*. Dover Books on Mathematics. Dover Publications, 2010.

[Bum13] D. Bump. *Lie Groups*. 2nd edition. Vol. 225. Graduate Texts in Mathematics. Springer Verlag, 2013.

[CST08] T. Ceccherini-Silberstein, F. Scarabotti, and F. Tolli. *Harmonic Analysis on Finite Groups*. Vol. 108. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 2008.

[CST10] T. Ceccherini-Silberstein, F. Scarabotti, and F. Tolli. *Representation Theory of the Symmetric Groups. The Okounkov–Vershik Approach, Character Formulas, and Partition Algebras*. Vol. 121. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 2010.

[DV88] D. J. Daley and D. Vere–Jones. *An Introduction to the Theory of Point Processes*. Springer Series in Statistics. Springer–Verlag, 1988.

[Dia88] P. Diaconis. "Applications of non-commutative Fourier analysis to probability problems". In: *École d'été de Probabilités de Saint-Flour XV-XVII, 1985-1987*. Vol. 1362. Lecture Notes in Mathematics. Springer Verlag, 1988.

[Dia03] P. Diaconis. “Mathematical developments from the analysis of riffle-shuffling”. In: *Groups, Combinatorics and Geometry*. Ed. by A. Fuanou and M. Liebeck. World Scientific, 2003, pp. 73–97.

[DFP92] P. Diaconis, J. Fill, and J. Pitman. “Analysis of top to random shuffles”. In: *Combinatorics, Probability and Computing* 1 (1992), pp. 135–155.

[DG89] P. Diaconis and C. Greene. “Applications of Murphy’s elements”. In: *Stanford Technical Report* 335 (1989).

[DS81] P. Diaconis and M. Shahshahani. “Generating a random permutation with random transpositions”. In: *Z. Wahr. verw. Gebiete* 57.2 (1981), pp. 159–179.

[FOW85] L. Flatto, A. M. Odlyzko, and D. B. Wales. “Random shuffles and group representations”. In: *Ann. Probab.* 13.1 (1985), pp. 154–178.

[Gre74] C. Greene. “An extension of Schensted’s theorem”. In: *Adv. Math.* 14 (1974), pp. 254–265.

[Ham72] J. M. Hammersley. “A few seedlings of research”. In: *Proc. Sixth Berkeley Symp. Math. Statist. and Probability*. Vol. 1. University of California Press, 1972, pp. 345–394.

[Hou+09] J. Ben Hough, M. Krishnapur, Y. Peres, and B. Virág. *Zeros of Gaussian Analytic Functions and Determinantal Point Processes*. Vol. 51. University Lecture Series. AMS, 2009.

[IK99] V. Ivanov and S. Kerov. “The algebra of conjugacy classes in symmetric groups, and partial permutations”. In: *Representation Theory, Dynamical Systems, Combinatorial and Algorithmical Methods III*. Vol. 256. Zapiski Nauchnyh Seminarov (POMI). 1999, pp. 95–120.

[IO02] V. Ivanov and G. Olshanski. “Kerov’s central limit theorem for the Plancherel measure on Young diagrams”. In: *Symmetric Functions 2001: Surveys of Developments and Perspectives*. Ed. by S. Fomin. Vol. 74. NATO Science Series II. Mathematics, Physics and Chemistry. 2002, pp. 93–151.

[Joh01] K. Johansson. “Discrete orthogonal polynomial ensembles and the Plancherel measure”. In: *Ann. Math.* 153 (2001), pp. 259–296.

[Juc74] A. A. A. Jucys. “Symmetric polynomials and the center of the symmetric group ring”. In: *Rep. Math. Phys.* 5.1 (1974), pp. 107–112.

[Kal17] O. Kallenberg. *Random Measures, Theory and Applications*. Vol. 77. Probability Theory and Stochastic Modelling. Springer Verlag, 2017.

[Ker93] S. V. Kerov. “Gaussian limit for the Plancherel measure of the symmetric group”. In: *Comptes Rendus Acad. Sci. Paris Série I* 316 (1993), pp. 303–308.

[KO94] S. V. Kerov and G. Olshanski. “Polynomial functions on the set of Young diagrams”. In: *Comptes Rendus Acad. Sci. Paris Série I* 319 (1994), pp. 121–126.

[KV77] S. V. Kerov and A. M. Vershik. “Asymptotics of the Plancherel measure of the symmetric group and the limiting form of Young tableaux”. In: *Soviet Mathematics Doklady* 18 (1977), pp. 527–531.

[KV81] S. V. Kerov and A. M. Vershik. “Asymptotic theory of characters of the symmetric group”. In: *Funct. Anal. Appl.* 15.4 (1981), pp. 246–255.

[KV85] S. V. Kerov and A. M. Vershik. “Asymptotics of maximal and typical dimensions of irreducible representations of the symmetric group”. In: *Funct. Anal. Appl.* 19.1 (1985), pp. 25–36.

[KV86] S. V. Kerov and A. M. Vershik. “The characters of the infinite symmetric group and probability properties of the Robinson–Schensted–Knuth algorithm”. In: *J. Alg. Disc. Meth.* 7.1 (1986), pp. 116–124.

[Knu70] D. E. Knuth. “Permutations, matrices, and generalized Young tableaux”. In: *Pacific Journal of Mathematics* 34.3 (1970), pp. 709–727.

[Lac16] H. Lacoin. “Mixing time and cutoff for the adjacent transposition shuffle and the simple exclusion”. In: *Ann. Probab.* 44.2 (2016), pp. 1426–1487.

[Len73] A. Lenard. “Correlation functions and the uniqueness of the state in classical statistical mechanics”. In: *Comm. Math. Phys. 30* (1973), pp. 35–44.

[Len75] A. Lenard. “States of classical statistical mechanical systems of infinitely many particles”. In: *Arch. Rational Mech. Anal.* 59 (1975), pp. 219–256.

[LPW17] D. A. Levin, Y. Peres, and E. L. Wilmer. *Markov Chains and Mixing Times*. 2nd edition. AMS, 2017.

[LS77] B. F. Logan and L. A. Shepp. “A variational problem for random Young tableaux”. In: *Adv. Math.* 26 (1977), pp. 206–222.

[Mac95] I. G. Macdonald. *Symmetric Functions and Hall Polynomials*. 2nd edition. Oxford Mathematical Monographs. Oxford University Press, 1995.

[Mél17] P.-L. Méliot. *Representation Theory of Symmetric Groups*. CRC Press, 2017.

[Mur81] G. E. Murphy. “A new construction of Young’s seminormal representation of the symmetric groups”. In: *J. Algebra* 69.2 (1981), pp. 287–297.

[Oko99] A. Okounkov. “On representations of the infinite symmetric group”. In: *Journal of Mathematical Sciences* 96.5 (1999), pp. 3550–3589.

[Oko00] A. Okounkov. “Random matrices and random permutations”. In: *Intern. Math. Res. Notices* 2000.20 (2000), pp. 1043–1095.

[Oko01a] A. Okounkov. “Infinite wedge and random partitions”. In: *Selecta Mathematica, New Series* 7 (2001), pp. 57–81.

[Oko01b] A. Okounkov. “Symmetric Functions and Random Partitions”. In: *Symmetric Functions 2001: Surveys of Developments and Perspectives*. Ed. by S. Fomin. Vol. 74. NATO Science Series. 2001, pp. 223–252.

[OV04] A. Okounkov and A. M. Vershik. “A new approach to the representation theory of the symmetric groups. II”. In: *Zapiski Nauchnyh Seminarov (POMI)* 307 (2004), pp. 57–98.

[Ols90] G. Olshanski. “Unitary representations of (G, K)-pairs connected with the infinite symmetric group $S(\infty)$”. In: *Leningrad Math. J.* 1 (1990), pp. 983–1014.

[Olv97] F. W. J. Olver. *Asymptotics and Special Functions*. AKP Classics. A K Peters, 1997.

[Poi12] H. Poincaré. *Calcul des probabilités*. 2nd edition. Gauthier–Villars, 1912.

[Rob38] G. Robinson. “On the representations of the symmetric group”. In: *Amer. J. Math.* 60.3 (1938), pp. 745–760.

[Sag01] B. E. Sagan. *The Symmetric Group: Representations, Combinatorial Algorithms and Symmetric Functions*. 2nd edition. Vol. 203. Graduate Texts in Mathematics. Springer Verlag, 2001.

[Sch61] C. Schensted. "Longest increasing and decreasing subsequences". In: *Canadian Journal of Mathematics* 13 (1961), pp. 179–191.

[Ser77] J.-P. Serre. *Linear Representations of Finite Groups*. Vol. 42. Graduate Texts in Mathematics. Springer Verlag, 1977.

[Śni06a] P. Śniady. "Asymptotics of characters of symmetric groups, genus expansion and free probability". In: *Discrete Math.* 306.7 (2006), pp. 624–665.

[Śni06b] P. Śniady. "Gaussian fluctuations of characters of symmetric groups and of Young diagrams". In: *Prob. Th. Rel. Fields* 136.2 (2006), pp. 263–297.

[Sos00a] A. Soshnikov. "Determinantal random point fields". In: *Russian Math. Surveys* 55 (2000), pp. 923–975.

[Sos00b] A. Soshnikov. "Gaussian fluctuation for the number of particles in Airy, Bessel, sine, and other determinantal random point fields". In: *Journal of Statistical Physics* 100.3-4 (2000), pp. 491–522.

[Tey20] L. Teyssier. "Limit profile for random transpositions". In: *Ann. Probab.* 48.5 (2020), pp. 2323–2343.

[Tho64] E. Thoma. "Die unzerlegbaren, positive-definiten Klassenfunktionen der abzählbar unendlichen symmetrischen Gruppe". In: *Math. Zeit.* 85 (1964), pp. 40–61.

[TW94] C. A. Tracy and H. Widom. "Level spacing distributions and the Airy kernel". In: *Commun. Math. Phys.* 159 (1994), pp. 151–174.

[Ula61] S. M. Ulam. "Monte–Carlo calculations in problems of mathematical physics". In: *Modern Mathematics for the Engineers*. McGraw–Hill, 1961, pp. 261–281.

[Won01] R. Wong. *Asymptotic Approximations of Integrals*. Vol. 34. Classics in Applied Mathematics. Society for Industrial and Applied Mathematics, 2001.

[Zel81] A. V. Zelevinsky. *Representations of finite classical groups: a Hopf algebra approach*. Vol. 869. Lecture Notes in Mathematics. Springer Verlag, 1981.