Lecture notes of *Introduction to random permutations*

Mattia Cafasso and Sofia Tarricone

Masterclass Angers, 17 - 19 December 2024

Contents

Abstract

These are lecture notes for the Masterclass cours held in Angers (December 17-19, 2024) of title *Introduction to random permutations*. The main reference we used is the book by Baik, Deift and Suidan [4], with some modifications in the presentation of the material (for instance, we gave a different proof of the Borodin-Okounkov-Geronimo-Case formula, following [5]).

1 Introduction to the Ulam problem

The main subject of this course is the analysis of the so-called *Ulam's problem*. In order to state it, let us consider the symmetric group S_N of the permutations of the finite set $\{1, \ldots, N\}$.

Definition 1.1. *Let $\pi \in S_N$ be a permutation. Given k integers*

$$1 \leq i_1 < i_2 < \cdots < i_k \leq N$$

we say that $\pi(i_1), \ldots, \pi(i_k)$ is an increasing subsequence of π of length k *if π preserves the order of the integers i_j, $j = 1, \ldots, k$, i.e.*

$$\pi(i_1) < \pi(i_2) < \cdots < \pi(i_k).$$

We will denote with $\ell_N(\pi)$ the maximal length of all the increasing subsequences of π.

Remark 1.2. *Given a permutation $\pi \in S_N$, there might be several increasing subsequence of maximal length.*

Example 1.3. *Let $\pi = 54172386$. Then (by inspection) $\ell_N(\pi) = 4$. $1, 2, 3, 6$ and $1, 2, 3, 8$ are both increasing subsequences of maximal length.*

Now suppose that we equip S_N with the uniform measure, i.e.

$$\mathbb{P}\{\pi\} = \frac{1}{N!}, \quad \forall \pi \in S_N.$$

In this course, we will study the statistical behavior of the random variable ℓ_N, for N large. This is a problem with a long history. Let us recall here some of the most important milestones.

1. Stanisław Ulam, in an article published in 1961 [26], discussed the idea of using Monte Carlo approaches to study the asymptotic behavior of $\mathbb{E}[\ell_N]$, when $N \to \infty$. Based on his simulations, Ulam conjectured that it exists a constant c such that

$$\lim_{N\to\infty} \frac{\mathbb{E}[\ell_N]}{\sqrt{N}} = c.$$

   The term Ulam's problem, originally, referred to the proof of this conjecture and the identification of the constant c.

2. Hammersley, in 1972, introduced a *Poissonized* version of this problem [18] (see below for more details) and proved that the constant c exists.

3. In 1977, Logan and Shepp [21] and, indipendently, Vershik and Kherov [27] showed that $c = 2$.

4. Odlyzko and Rains, in the nineties [22], performed analytical simulations indicating that

$$\lim_{N\to\infty} \frac{\mathrm{Var}(\ell_N)}{N^{1/3}} = c_0 \equiv 0.819..., \tag{1.1}$$

   and also

$$\lim_{N\to\infty} \frac{\mathbb{E}[\ell_N] - 2\sqrt{N}}{N^{1/6}} = c_1 \equiv -1.758.... \tag{1.2}$$

Nowadays, using a simple laptop, one can easily run simulations giving a clear visualization of the results stated above, see Figure 1.

The final word on Ulam's problem was given by Baik, Deift and Johansson in 1999, who studied the convergence in distribution of the rescaled variable

$$\chi_N := \frac{\ell_N - 2\sqrt{N}}{N^{1/6}}. \tag{1.3}$$

In order to state their theorem, we need some results from the theory of Painlevé equations.

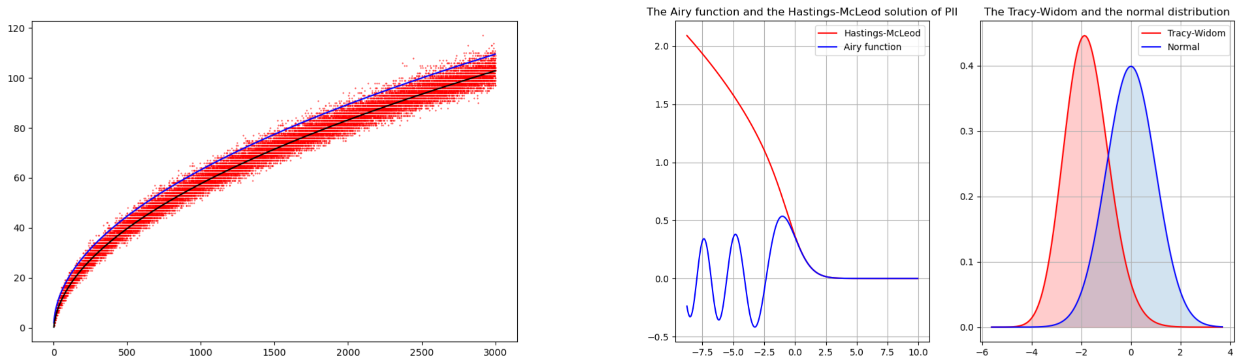


Figure 1: On the left: For each $N = 1, \dots, 3 \times 10^3$, we took 10 random permutations of size N and compute the associated value $\ell_N(\pi)$. In blue, the curve $y(x) = 2\sqrt{x}$ and in black the curve $y(x) = 2\sqrt{x} - 1.7711x^{1/6}$. On the right: The Airy function, solution to the equation $y''(x) = xy(x)$, is a special function with the same asymptotic, as $x \to +\infty$, of the Hasting-McLeod solution of the Painlevé II equation.

Theorem 1.4 (Hastings-McLeod, 1980 [19])**.** *The Painlevé II equation*

$$u''(x) = 2u^3(x) + xu(x) \tag{1.4}$$

has a unique solution u such that

$$u(x) \sim \frac{e^{-\frac{2}{3}x^{3/2}}}{2\sqrt{\pi}x^{1/4}} \quad \text{as} \quad x \to +\infty. \tag{1.5}$$

Moreover, this solution has the following asymptotics

$$u(x) \sim \sqrt{\frac{-x}{2}}\left(1 + \mathcal{O}(x^{-2})\right) \quad \text{as} \quad x \to -\infty. \tag{1.6}$$

This particular solution of the Painlevé II equation is commonly referred as *the Hastings-McLeod solution of the Painlevé II equation*. The theorem above, in particular, implies that the function

$$F_{GUE}(x) := \exp\left(-\int_x^{+\infty} (y - x)u^2(y)\mathrm{d}y\right) \tag{1.7}$$

is a cumulative distribution function.

Theorem 1.5 (Baik-Deift-Johansson, 1999)**.**
Let χ be a random variable with distribution function $F_{GUE}(x)$. Then χ_N, defined in (1.3)*, converges in distribution to χ:*

$$\lim_{N\to\infty} \mathbb{P}\left\{\frac{\ell_N - 2\sqrt{N}}{N^{1/6}} \leq x\right\} \to F_{GUE}(x). \tag{1.8}$$

We have moreover, convergence of moments:

$$\lim_{N\to+\infty} \mathbb{E}\left[\chi_N^m\right] = \mathbb{E}[\chi^m], \quad \forall m \geq 1. \tag{1.9}$$

Remark 1.6. *One should think about the Hastings-McLeod solution of Painlevé II as a sort of* nonlinear special function. *Tables are available for such functions, which allows us to plot it, and also plot the distribution function F_{GUE} and to compute its moments, see Fig. 1. In particular, the theorem above implies that*

$$\lim_{N\to+\infty} \frac{\mathrm{Var}(\ell_N)}{N^{1/3}} = \int_{-\infty}^{+\infty} t^2 dF_{GUE}(t) - \left(\int_{-\infty}^{+\infty} t dF_{GUE}(t)\right)^2 = 0.81132... \tag{1.10}$$

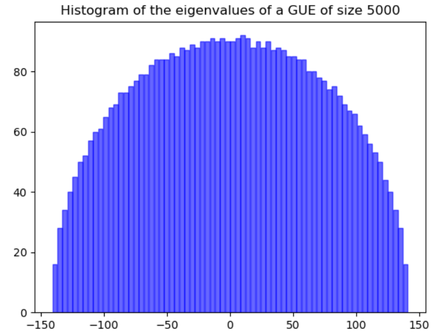


Figure 2: Histogram of the eigenvalues of a GUE of size 5×10^3. The semi-circle Wigner's law is clearly visible.

$$\lim_{N \to +\infty} \frac{\mathbb{E}\left[\ell_N\right] - 2\sqrt{N}}{N^{1/6}} = \int_{-\infty}^{+\infty} t dF_{GUE}(t) = -1.711... \tag{1.11}$$

which are in good agreement with (1.1), (1.2).

Remark 1.7. *Back in 1999, the Tracy-Widom distribution was already known as the limiting distribution of the largest eigenvalue on a large class of random matrices. More specifically, consider for instance the space of Hermitian matrices of size* N*, denoted by* $M = (M_{\ell,m})_{\ell,m=1}^N$ *with* $M_{\ell,m} = M_{\ell,m}^R + \mathrm{i} M_{\ell,m}^I$*. One can endow such a space with the probability measure*

$$\frac{1}{\mathcal{Z}_N} \mathrm{e}^{-\mathrm{Tr} M^2} \mathrm{d}M \equiv \frac{1}{\mathcal{Z}_N} \mathrm{e}^{-\mathrm{Tr} M^2} \prod_{\ell=1}^{N} \mathrm{d}M_{\ell,\ell}^R \prod_{1 \le \ell < m \le N} \mathrm{d}M_{\ell,m}^R \mathrm{d}M_{\ell,m}^I, \tag{1.12}$$

and consider the distribution of the corresponding eigenvalues (see Figure 2). Let us denote $\xi_1(M)$ *the biggest one. Then, Tracy and Widom, in [25] proved that*

$$\lim_{N \to \infty} \mathbb{P} \left\{ \frac{\xi_1(M) - \sqrt{2N}}{2^{-1/2} N^{-1/6}} \le x \right\} = F_{GUE}(x). \tag{1.13}$$

Note that, a priori, *there are no physical reasons to explain why the longest increasing subsequence of a random permutation behaves like the largest eigenvalue of a random matrix.*

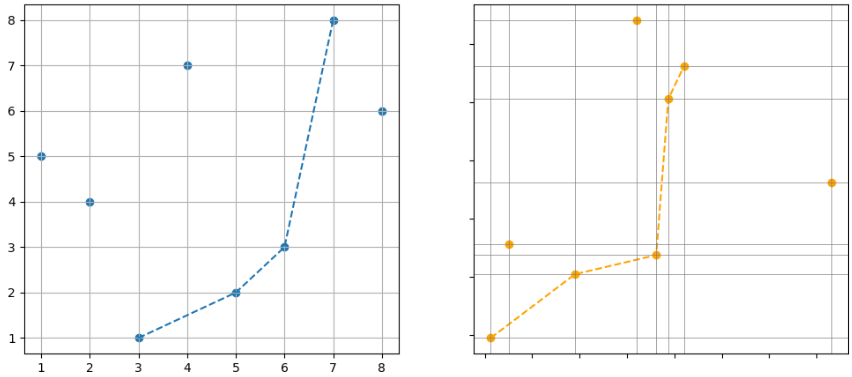


Figure 3: On the left, graph of the permutation (54172386) and visualization of one the longest increasing subsequence. On the right, sample of 8 points chosen uniformly. The corresponding permutation is (14283675). The increasing subsequence of maximal length which is visualized is $1, 2, 3, 6, 7$.

2 Poissonization and de-Poissonization

Giving a permutation $\pi \in S_N$, one can consider its graph as a collection of integer points in the square $D_N := [0, N] \times [0, N]$. Let us then consider Γ_π as the set of piecewise linear up/right paths with nodes at the points $\{(i, \pi(i)),\ i = 1, \dots, N\}$. It is easy to realize that

$$\ell_N(\pi) = \max_{p \in \Gamma_\pi} \{\#\text{ nodes on } p\}.$$

(see Fig.3).

Now suppose to take, at random and uniformly, N points $\{P_i = (x_i, y_i),\ i = 1, \dots, N\}$ in the square D_N. With probability one, all the x and y coordinates are distinct. Hence, one can associate to such a configuration of points a permutation π, and prove that such random permutation is uniform in S_N (see exercise 1, and Fig.3).

We want now to define a new random variable $L(t)$, depending on a parameter $t > 0$, which will be a "Poissonized" version of the random variable ℓ_N. In order to do so, we consider the square $D_t := [0, t] \times [0, t]$, and take $\mathcal{N}(t)$ uniform random points on D_t, where $\mathcal{N}(t)$ is itself a Poisson random variable with parameter t^2:

$$\mathbb{P}(\mathcal{N}(t) = N) = \mathrm{e}^{-t^2} \frac{t^{2N}}{N!}. \tag{2.1}$$

We will then define $\Gamma(t)$ as the random set of piecewise linear up/right paths in the square D_t, with nodes at the points $\{P_1, \dots, P_{\mathcal{N}(t)}\}$, and define the random variable

$$L(t) = \max_{p \in \Gamma(t)} \{\#\text{ nodes on } p\}. \tag{2.2}$$

Proposition 2.1. *For any $n \geq 0$,*

$$\mathbb{P}\{L(t) \leq n\} = \sum_{N=0}^{\infty} \mathrm{e}^{-t^2} \frac{t^{2N}}{N!} \mathbb{P}\{\ell_N \leq n\}. \tag{2.3}$$

Proof. This is simply because, if we condition $\mathcal{N}(t)$ to be equal to N, then $L(t)$ has the same distribution as ℓ_N. Hence,

$$\mathbb{P}\{L(t) \leq n\} = \sum_{N=0}^{\infty} \mathbb{P}\{L(t) \leq n | \mathcal{N}(t) = N\} \mathbb{P}\{\mathcal{N}(t) = N\} = \sum_{N=0}^{\infty} \mathrm{e}^{-t^2} \frac{t^{2N}}{N!} \mathbb{P}\{\ell_N \leq n\}.$$

□

Remark 2.2. *The arguments above can be made more precise introducing a* rate 1 Poisson point process *on the quarter plane* $\mathbb{R}_+^2$, *and defining* $\Gamma(t)$ *as the set of piecewise linear up/right paths from* $(0,0)$ *to* (t,t), *with nodes at the points of the Poisson point process contained in the square* D_t. *Roughly speaking, a Poisson point process is a measure on the space of point configurations on the underlying space* ($\mathbb{R}_+^2$, *in our case). It is uniquely defined by two properties:*

- *For any bounded Borel set* $A \subseteq \mathbb{R}_+^2$, *the number of points of the process contained in* A *is a Poisson random variable* $\mathcal{N}_A$ *of parameter* $\mu(A)$, *where* μ *is the Lebesgue measure.*
- *If* A *and* B *are disjoint, then* $\mathcal{N}_A$ *and* $\mathcal{N}_B$ *are independent.*

For more details, see for instance [12].

2.1 De-Poissonization of $L(t)$

It will be clearer from the next sections that $L(t)$ has a much richer structure then ℓ_N, and this is the reason why we introduced it. On the other hand, the Poisson variable $\mathcal{N}(\sqrt{N})$ has expected value equal to N and standard deviation equal to $\sqrt{N}$. Hence, it is reasonable to expect that $L(\sqrt{N})$ is "close to" ℓ_N, when $N \to \infty$. The following proposition makes this heuristic argument clearer.

Proposition 2.3. *Let us fix* $s > 1$ *and define* $\mu_N := N + 2\sqrt{sN \log N}$, $\nu_N := N - 2\sqrt{sN \log N}$ *for* $N \geq 2$. *Then, for* $4s \log N < N$ *and for any* $n \geq 0$,

$$\mathbb{P}\left\{L(\sqrt{\mu_N}) \leq n\right\} - \frac{1}{N^s} \leq \mathbb{P}\left\{\ell_N \leq n\right\} \leq \mathbb{P}\left\{L(\sqrt{\nu_N} \leq n)\right\} + \frac{1}{N^s}. \tag{2.4}$$

Remark 2.4. *To appreciate the importance of the proposition above, suppose that we prove that*

$$\lim_{t\to\infty} \mathbb{P}\left\{\frac{L(t) - 2t}{t^{1/3}} \leq x\right\} = F_{GUE}(x), \tag{2.5}$$

i.e., convergence in law of $L(t)$, *up to normalization. We can use Proposition 2.3 to easily prove that*

$$\lim_{N\to\infty} \mathbb{P}\left\{\frac{\ell_N - 2\sqrt{N}}{N^{1/6}} \leq x\right\} = F_{GUE}(x).$$

Indeed, as $N \to \infty$,

$$\mathbb{P}\left\{L(\sqrt{\nu_N}) \leq 2\sqrt{N} + xN^{1/6}\right\} = \tag{2.6}$$

$$= \mathbb{P}\left\{\frac{L(\sqrt{\nu_N}) - 2\sqrt{\nu_N}}{\nu_N^{1/6}} \leq \frac{2\sqrt{N} + xN^{1/6} - 2\sqrt{\nu_N}}{\nu_N^{1/6}}\right\} \to F_{GUE}(x), \tag{2.7}$$

and then, using the Proposition above, one has that

$$\limsup_{N\to\infty} \mathbb{P}\left\{\frac{\ell_N - 2\sqrt{N}}{N^{1/6}} \leq x\right\} \leq F_{GUE}(x).$$

Analogously, studying $\mathbb{P}\left\{L(\sqrt{\mu_N}) \leq 2\sqrt{N} + xN^{1/6}\right\}$, *we prove that*

$$F_{GUE} \leq \liminf_{N\to\infty} \mathbb{P}\left\{\frac{\ell_N - 2\sqrt{N}}{N^{1/6}} \leq x\right\}.$$

The proposition above is, actually, a particular case of the lemma and corollary below. Suppose that we are given a sequence of real numbers $(a_N)_{N\geq 0}$ such that

$$0 \leq a_N \leq 1 \quad \text{and} \quad a_{N+1} \leq a_N. \tag{2.8}$$

We introduce the generating function ϕ such that

$$\phi(t) = \phi(t, (a_N)_{N\geq 0}) := \sum_{N\geq 0} \mathrm{e}^{-t^2} \frac{t^{2N}}{N!} a_N. \tag{2.9}$$

Lemma 2.5. *For all $\epsilon \in (0,1)$ and $N \geq 1$*

$$\phi\left(\sqrt{(1+\epsilon)N}\right) - \mathrm{e}^{-\frac{\epsilon^2}{4}N} \leq a_N \leq \phi\left(\sqrt{(1-\epsilon)N}\right) + \frac{1-\epsilon}{\epsilon\sqrt{2\pi N}}\mathrm{e}^{-\frac{\epsilon^2}{2}N} \tag{2.10}$$

Proof. Using the monotonicity of the sequence (a_N) we find that

$$\phi(\sqrt{(1+\epsilon)N}) - a_N = \sum_{k=0}^{\infty}\left(\frac{\mathrm{e}^{-(1+\epsilon)N}((1+\epsilon)N)^k}{k!}a_k\right) - a_N = \sum_{k=0}^{\infty}\frac{\mathrm{e}^{-(1+\epsilon)N}((1+\epsilon)N)^k}{k!}(a_k - a_N) \leq$$
$$\leq \sum_{k=0}^{N-1}\frac{\mathrm{e}^{-(1+\epsilon)N}((1+\epsilon)N)^k}{k!}(a_k - a_N).$$

Using $0 \leq a_k - a_N \leq 1$ we can continue our chain of inequalities:

$$\sum_{k=0}^{N-1}\frac{\mathrm{e}^{-(1+\epsilon)N}((1+\epsilon)N)^k}{k!}(a_k - a_N) \leq \sum_{k=0}^{N-1}\frac{\mathrm{e}^{-(1+\epsilon)N}((1+\epsilon)N)^k}{k!} \leq$$
$$\leq \mathrm{e}^{-(1+\epsilon)N}(1+\epsilon)^N\sum_{k=0}^{N-1}\frac{N^k}{k!} \leq \mathrm{e}^{-\epsilon N}(1+\epsilon)^N. \tag{2.11}$$

This is enough to prove the first inequality in (2.10), since $\log(1+\epsilon) \leq \epsilon - \frac{\epsilon^2}{4}$. As for the second inequality, we start with similar computations:

$$a_N - \phi(\sqrt{(1+\epsilon)N}) \leq \sum_{k=N+1}^{\infty}\frac{\mathrm{e}^{-(1-\epsilon)N}((1-\epsilon)N)^k}{k!}(a_N - a_k) \leq \mathrm{e}^{-(1-\epsilon)N}\sum_{k=N+1}^{\infty}\frac{((1-\epsilon)N)^k}{k!} =$$
$$= \mathrm{e}^{-(1-\epsilon)N}((1-\epsilon)N)^N\sum_{m=1}^{\infty}\frac{((1-\epsilon)N)^m}{(m+N)!} \leq \mathrm{e}^{-(1-\epsilon)N}\frac{((1-\epsilon)N)^N}{N!}\sum_{m\geq 1}(1-\epsilon)^m =$$
$$= \mathrm{e}^{-(1-\epsilon)N}\frac{((1-\epsilon)N)^N}{N!}\frac{1-\epsilon}{\epsilon}.$$

Now, Stirling approximation tells us that

$$\frac{1}{N!} < \frac{1}{\sqrt{2\pi N}N^N}\mathrm{e}^{N-\frac{1}{12N}+\frac{1}{360N^3}},$$

and plugging the latter in the chain of inequalities above we get

$$a_N - \phi(\sqrt{(1+\epsilon)N}) \leq \mathrm{e}^{\epsilon N}\frac{(1-\epsilon)^N}{\sqrt{2\pi N}}\frac{1-\epsilon}{\epsilon}.$$

Finally, using $\log(1-\epsilon) \leq -\epsilon - \frac{\epsilon^2}{2}$, $\quad 0 \leq \epsilon < 1$, we finish the proof of the second inequality in (2.10). □

Corollary 2.6. *Let us fix $s > 1$ and define $\mu_N := N + 2\sqrt{sN\log N}$, $\nu_N := N - 2\sqrt{sN\log N}$ for $N \geq 2$. Then, for $4s\log N < N$,*

$$\phi(\sqrt{\mu_N}) - \frac{1}{N^s} \leq a_N \leq \phi(\sqrt{\nu_N}) + \frac{1}{N^s}. \tag{2.12}$$

Proof. We use the previous lemma setting $\epsilon := 2\sqrt{\frac{s\log N}{N}}$.The first inequality in (2.12) is immediately proven. As for the second one, we use the fact that

$$\frac{1-\epsilon}{\epsilon\sqrt{2\pi N}}\mathrm{e}^{-\frac{\epsilon^2}{2}N} \leq \frac{1}{\epsilon\sqrt{N}}\mathrm{e}^{-\frac{1}{2}\epsilon^2 N} = \frac{1}{2\sqrt{s\log N}N^{2s}} \leq \frac{1}{N^s}.$$

□

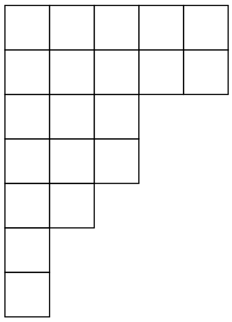
It is clear that Proposition 2.3 is a particular case of Corollary 2.6, applied to the sequence $a_N := \mathbb{P}\{\ell_N \leq n\}$, $\quad N \geq 1$. The only thing that one should verify is that $\mathbb{P}\{\ell_N \leq n\} \geq \mathbb{P}\{\ell_{N+1} \leq n\}$. We leave this as an exercise (Exercise 2).

3 Permutations and Young tableaux

Permutations are deeply related to integer partitions and Young tableaux. Let us introduce them.

Definition 3.1. *A partition $\lambda = (\lambda_1, \ldots, \lambda_r)$ of N is a set of positive integers such that $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_r$ and $|\lambda| := \sum_i \lambda_i = N$. We call $|\lambda|$ the size of the partition, and r its length. We will sometimes write $\lambda \vdash N$ to indicate that λ has size N.*

A convenient graphic representation of integer partitions is given by Young diagrams, which are constructing stacking r rows of boxes, each of length λ_i, in decreasing order. For instance, the partition $(5, 5, 3, 3, 2, 1, 1) \vdash 20$ corresponds to the Young diagram below:



The set of Young diagrams of size N will be denoted $\mathbb{Y}_N$, and the set of all Young diagrams $\mathbb{Y}$. From now on, we will implicitly identify a partition and the corresponding Young diagram.

Definition 3.2. *A* Young tableau *of size N is a Young diagram $(\lambda_1, \ldots, \lambda_r) \vdash N$ together with a bijection between its boxes and the integers $\{1, \ldots, N\}$. A* standard Young tableau *is a Young tableau whose integers increase along the lines and the columns. We will denote with SYT_N the set of standard Young tableau of size N. The Young diagram of a given Young tableau P, denoted $sh(P)$, is called the* shape *of the tableau.*

Example 3.3. *The two Young tableaux below are of shape $(5, 5, 4, 2, 2, 1) \vdash 19$, the first one is standard and the second one is not.*

1	3	4	8	10
2	5	9	13	14
6	11	12	16	
7	15			
17	18			
19				

5	3	4	8	10
2	1	12	13	14
6	11	9	16	
15	7			
17	18			
19				

Definition 3.4. *Given a Young diagram $\lambda \vdash N$, we will denote*

$$F_\lambda := \#\{P \in SYT_N \,:\, \mathrm{sh}(P) = \lambda\}. \tag{3.1}$$

An elegant and remarkable identity, due to Frobenius and Young, states that

$$\sum_{\lambda \vdash N} F_\lambda^2 = N! \tag{3.2}$$

We will now give a proof of this identity as a corollary of the so-called Robinson-Schensted correspondence, which is a bijection between permutations and couples of Young tableaux of the same shape:

$$\mathrm{RS} : S_N \longrightarrow \{(P, Q) \in SYT_N \times SYT_N \,:\, \mathrm{sh}(P) = \mathrm{sh}(Q)\} \tag{3.3}$$

3.1 The Robinson-Schensted algorithm

We will call a *partial tableau* (PT) a Young diagram whose boxes contain integers increasing along the lines and the columns. A partial tableau P belongs to SYT_N iff the integers in its boxes are exactly $\{1, \dots, N\}$. Now, suppose that $\pi \in S_N$. We will construct $RS(\pi) = (P, Q)$ through a sequence of couples of partial tableaux $(P_j, Q_j)_{j=1}^N$ such that $(P_0, Q_0) = (\emptyset, \emptyset)$, $(P_N, Q_N) = (P, Q)$ and, at every step, $\mathrm{sh}(P_i) = \mathrm{sh}(Q_i)$. In details, for every $i = 1, \dots, N$, we *row-insert* $\pi(i)$ into P_{i-1} and *place* i into Q_{i-1}. Q is sometimes called the *recording tableau*, the reason will be clear in a moment.

We now describe how to row-insert an integer x inside a partial tableau P:

- If x is greater than any element in the first row, just place it in a new box at the end.
- Otherwise, put it at the place of the first integer y greater than x.
- Repeat the procedure for y, going one row down.

Example 3.5. *Suppose that* $x = 4$ *and*

$$P = \begin{matrix} 1 & 3 & 6 \\ 2 & 9 & \\ 8 & & \end{matrix}$$

Then the row insertion will work as follows:

$$\begin{matrix} 1 & 3 & 4 \\ 2 & 9 & \\ 8 & & \end{matrix} \to \begin{matrix} 1 & 3 & 4 \\ 2 & 6 & \\ 8 & & \end{matrix} \to \begin{matrix} 1 & 3 & 4 \\ 2 & 6 & \\ 8 & 9 & \end{matrix}$$

We leave as an exercise to prove that, if P is a partial tableau and $x \notin P$, then the result of row-insertion is still a partial tableau. Now, suppose that you are given a permutation $\pi \in S_N$ and you have already constructed the couple (P_i, Q_i). In order to construct (P_{i+1}, Q_{i+1}), you first row-insert $\pi(i+1)$ into P_i, and then a new box will be created at the place (ℓ, m). Then, you add a box to Q_i, at the same position (ℓ, m), and write $i + 1$ inside the box. At the end, Q keeps trace of the order in which the boxes have been added, hence the name of *recording tableau*.

Example 3.6. *Suppose you are given the permutation* $(3541276) \in S_7$. *The steps of the RS algorithm are as follows:*

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

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

In order to prove that RS is a bijection, we will explicitly construct the inverse. To do so, recall that, starting from a permutation $\pi = (\pi(1), \dots, \pi(N)) \in S_N$, we constructed a sequence of couples of SYT_N

$$(\emptyset, \emptyset) \xrightarrow{\pi(1)} (P_1, Q_1) \xrightarrow{\pi(2)} \cdots \xrightarrow{\pi(n)} (P_N, Q_N) = (P, Q)$$

where, at each entry, the "input" from the permutation was given by the integer $\pi(i)$ (hence, the notation above). We now reverse these arrows and, at each step, starting from (P_j, Q_j), we recover (P_{j-1}, Q_{j-1}) and $\pi(j-1)$. The process is as follows:

- Select, from Q_j, the position (ℓ, m) of the box with the greatest integer. Remove it from Q_j. You obtained Q_{j-1}.
- Remove the same box from P_j. Now suppose that x was the integer contained in the box (ℓ, m) of P. Inverse the row insertion as follows:
  - If you are at the first row, just keep x as $\pi(j)$.
  - Otherwise, go one row above, put x at the place of the largest y smaller then x.
  - Repeat the procedure with x replaced by y, and so on till you reach the first row.

It might be useful, to get a grip on this algorithm, to check on example 3.6 that, indeed, starting from the last column, one gets backward and recover the permutation (3541276).

The RS correspondence enjoy many interesting properties. One in particular will be central for the sequel of these lectures. It is known under the name of *Schensted's Theorem.*

Theorem 3.7 (Schensted)**.** *Given a permutation $\pi \in S_N$, the length of its longest increasing subsequence is equal to length of the first row of the corresponding Young diagram, under the RS correspondence.*

We will actually prove a stronger result, which is the following

Theorem 3.8. *Let $\pi \in S_N$ and consider the corresponding sequence*

$$(\emptyset, \emptyset) \to (P_1, Q_1) \to \cdots \to (P_N, Q_N)$$

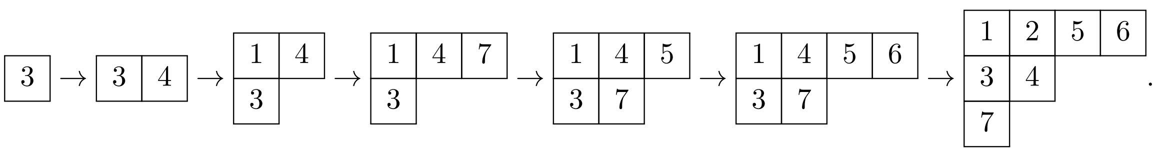
of couples of PT under the RS correspondence. If $\pi(k)$ enter into P_{k-1} at the column j, then the longest increasing subsequence ending in $\pi(k)$ has length j.

Proof. We prove it by induction on k. For $k = 1$, there is nothing to prove. Now suppose that $\pi(k)$ entered P_{k-1} at column j. We will prove that:
a) there is an increasing subsequence of length j ending at $\pi(k)$.
b) there isn't any longer subsequence of such a type.
As for *a*), suppose that $\pi(r)$, $r < k$, is at position $(1, j-1)$ in P_{k-1}. Then, by induction, there is a subsequence of length $(j-1)$ ending at $\pi(r)$, and we can add $\pi(k)$ to this sequence (since $\pi(r) > \pi(k)$) to obtain a subsequence of length j.
As for *b*), suppose that $\pi(\ell), \cdots, \pi(i), \pi(k)$ has length greater or equal than $j+1$, then $\pi(\ell), \cdots, \pi(i)$ has length greater or equal than j, which means that $\pi(i)$, in the tableau P_i, is (weakly) to the right of the column j. Now we take $\pi(q)$ element in the box $(1, j)$ of P_i, and clearly we have

$$\pi(q) \leq \pi(i) < \pi(k).$$

Take also $\pi(r)$ in the same box $(1, j)$ of P_{k-1}, clearly we have $\pi(r) \leq \pi(q) \leq \pi(i) < \pi(k)$. But it is not possible that $\pi(r) < \pi(k)$, since $\pi(k)$ "bumped" $\pi(r)$. □

Example 3.9. *Suppose to take $\pi = (3417562)$. The sequence of the (P_i), in the RS algorithm, is given by*



Accordingly, the length of the longest increasing subsequence ending at

3 4 1 7 5 6 2

is equal respectively to

1 2 1 3 3 4 2.

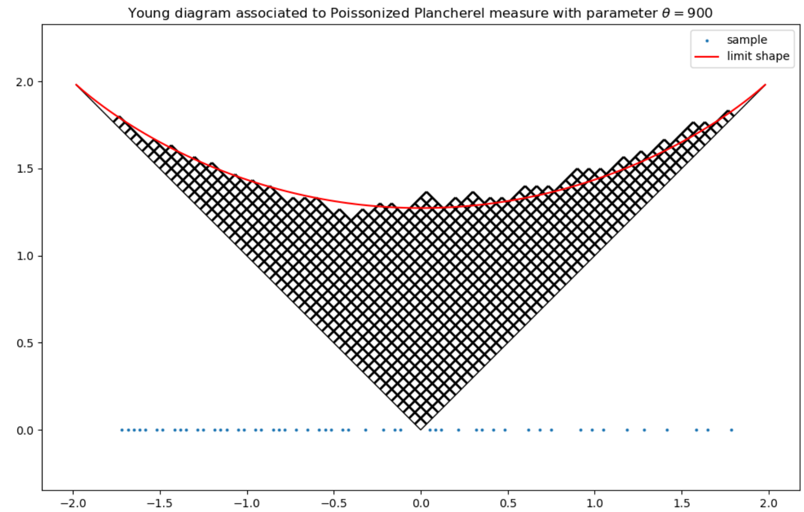


Figure 4

3.2 The Plancherel and Poissonized Plancherel measure

The RS correspondence induces a surjective map $S_N \twoheadrightarrow \mathbb{Y}_N$, which associated to each $\pi \in S_N$ the shape of the two standard Young tableaux $(P, Q) = RS(\pi)$. The push-forward of the uniform measure on S_N induces a measure on Y_N, given by

$$\mathbb{P}_{Pl_N}\{\lambda\} = \frac{F_\lambda^2}{|\lambda|!} = \frac{F_\lambda^2}{N!}, \quad \forall\, \lambda \vdash N. \tag{3.4}$$

Analogously, one has the Poissonized uniform measure on the space of permutations of arbitrary size, and its push-forward gives the Poissonized Plancherel measure

$$\mathbb{P}_{Pl(t)}\{\lambda\} = \mathrm{e}^{-t^2} \frac{t^{2|\lambda|} F_\lambda^2}{(|\lambda|!)^2}, \quad \forall\, \lambda \in \mathbb{Y}. \tag{3.5}$$

In view of the Schensted's theorem, it follows immediately that

$$\mathbb{P}(\ell_N \le n) = \mathbb{P}_{Pl_N}\{\lambda_1 \le n\} \tag{3.6}$$

and

$$\mathbb{P}(L(t) \le n) = \mathbb{P}_{Pl(t)}\{\lambda_1 \le n\}. \tag{3.7}$$

In what follows, we will give a more explicit formula for F_λ, and use it to express the right-hand side of (3.7) as the determinant of a $(n \times n)$ matrix (in fact, a Toeplitz matrix). For the moment being, we just mention that, sampling a (Poissonized) uniform permutation, and then using the RS correspondence, it is easy to sample a Young diagram distributed according to the (Poissonized) Plancherel measure. The picture in Fig. 4, obtained using [16], is the sample of a Poissonized Plancherel Young diagram, for $t = 30$. It is rescaled by a factor t, and it is clearly seen that the lenght of λ_1 is very close to 2, as proven by Logan and Shepp, and Vershik and Kherov. Actually, Logan and Shepp, and Vershik and Kherov, indipendently, found the equation of the limiting curve in red as the solution of a certain minimization problem, and obtained the solution of the Ulam's problem as a corollary of this result. The (very simple) expression of the curve is

$$\Omega(u) = \frac{2}{\pi}\left(u \arcsin(u/2) + \sqrt{4 - u^2}\right), \quad |u| \le 2. \tag{3.8}$$

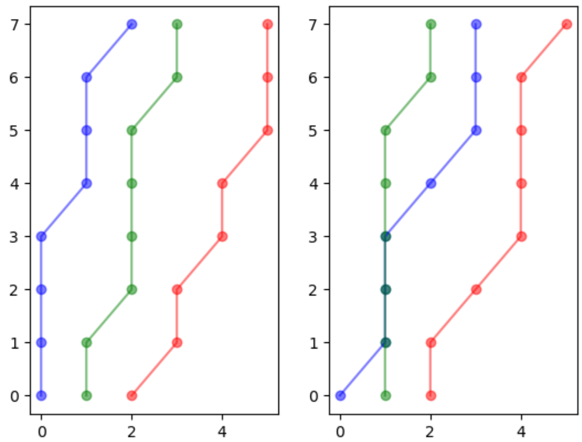


Figure 5: A visualization of two different ballot sequences. The first one is (strictly) order preserving, the second in not.

4 Toeplitz formulas and asymptotics

We start giving a formula for F_λ, which is known under the name of Frobenius-Young determinantal formula.

4.1 The Frobenius-Young determinantal formula

Theorem 4.1. *For any $\lambda = (\lambda_1, \dots, \lambda_r) \vdash N$,*

$$F_\lambda = N! \det \left(\frac{1}{(\lambda_i - i + j)!} \right)_{i,j=1}^{r} . \tag{4.1}$$

Remark 4.2. *For reasons that will be clearer in the next section, imagine that we represent $\lambda = (\lambda_1, \dots, \lambda_r) \vdash N$ via a longer sequence of integers $(\lambda_1, \dots, \lambda_r, \lambda_{r+1}, \lambda_n)$, with $\lambda_j = 0$ for any $j \geq r+1$. The formula* (4.1) *still holds, imposing $k! = 0$ for any negative integer k.*

The proof we are giving is by elementary arguments in combinatorics. We start giving a bijection between SYT and another type of object.

Definition 4.3. *A* ballot sequence *with N voters, r candidates, standing u and outcome v is a sequence of vectors $(u_0 = u, u_1, \dots, u_N = v)$ such that $u_j = (u_j^{(1)}, \dots, u_j^{(r)}) \in \mathbb{Z}_+^r$ for every j and, at each step, u_{j+1} is obtained from u_j adding one to a single coordinate, and leaving the others unchanged. Moreover, we say that a ballot sequence is* strictly order preserving *(SOPBS) if, for every j, the coordinates of u_j form a strictly decreasing sequence $u_j^{(1)} > u_j^{(2)} > \dots > u_j^{(r)}$.*

Remark 4.4. *One can visualize a ballot sequence as a collection of r paths, each going through the points $(k, u_k^{(j)})$, $j = 1, \dots, r$, that are such that, at each step, all of them goes north but one, which goes north-east. Moreover, a ballot sequence is strictly order preserving if and only if those paths never intersect, see Fig. 4.*

Lemma 4.5. *Let $\lambda = (\lambda_1, \dots, \lambda_r) \vdash N$ and denote $v = (\ell_1, \dots, \ell_r) = (\lambda_1 + r - 1, \lambda_2 + r - 2, \dots, \lambda_r)$. We also denote $u := (r-1, r-2, \dots, 0)$. Then*

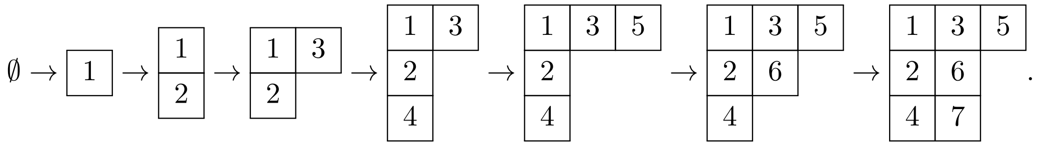
$$F_\lambda = \#\, SOPBS_{(u,v)} = \#\{\text{SOPBS } \textit{with } N \textit{ voters, standing } u \textit{ and outcome } v\}. \tag{4.2}$$

The proof of this lemma is left as an exercise. It shows an explicit map from $SOPBS_{(u,v)}$ to SYT of shape λ as follows: if, going from u_i to u_{i+1}, the candidate j obtained one vote, then add a box to the j-th row, and put $i+1$ in the box.

Example 4.6. *Consider the SOPBS*

$$(2,1,0) \to (3,1,0) \to (3,2,0) \to (4,2,0) \to (4,2,1) \to (5,2,1) \to (5,3,1) \to (5,3,2).$$

The corresponding SYT is obtained through the following chain:



In the Fig 4, the system of paths on the left hand side is associated to this SOPBS.

We will now compute the cardinality of $SOPBS_{u,v}$ as a determinant. Now, for any $\sigma \in S_r$, let us denote

$$v_\sigma := (\ell_{\sigma(1)}, \dots, \ell_{\sigma(r)}),$$

and

$$\mathcal{P}_v := \{P \ : \ P \in BS_{(u,v_\sigma)} \text{ for some } \sigma \in S_r\}.$$

We will compute

$$A := \sum_{P \in \mathcal{P}_v} \text{sgn} P,$$

where, if $P \in BS_{(u,v_\sigma)}$, then $\text{sgn}\, P = \text{sgn}\, \sigma$. To do so, we remark that $\mathcal{P}_v = \mathcal{P}_v' \cup \mathcal{P}_v''$, where $\mathcal{P}_v''$ is the set of P which are strictly order preserving, and $\mathcal{P}_v'$ is the complement. Clearly, $A = A' + A'' = \sum_{P \in \mathcal{P}_v'} \text{sgn}\, P + \sum_{P \in \mathcal{P}_v''} \text{sgn}\, P$. We now claim that $A' = 0$ and $A'' = \#SOPBS_{(u_0,v)}$. As for A', it is enough to construct an involution $\iota : \mathcal{P}_v' \longrightarrow \mathcal{P}_v'$ such that $\text{sgn}(\iota(P)) = -\text{sgn} P$ for any $P \in \mathcal{P}_v'$. Let $(u_0, u_1, \dots, u_{N-1}, u_N) \in \mathcal{P}_v'$, and denote $u_k = (u_k^{(1)}, \dots, u_k^{(r)}), k = 1, \dots, N$. It exists a $j < N$ and two integers $s_1, s_2 < r$ such that $u_j^{(s_1)} = u_j^{(s_2)}$. We take the biggest of such a j, and then s_1, s_2 will be unique. We define $\tilde{P} = (u_0, u_1, \dots, u_j, \tilde{u}_{j+1}, \tilde{u}_{N-1}, \tilde{u}_N) \in \mathcal{P}_v'$ exchanging, starting from j, the coordinate s_1 with the coordinate s_2 (for instance, in Fig. 4, the involution exchanges green with blue starting from $y = 3$).ι has clearly the requested properties, so that $A' = 0$. As for A'', if $P \in \mathcal{P}_v''$, then the associated permutation is the identity, and this means that $\mathcal{P}_v'' = SOPBS_{u,v}$ and that $A'' = \#SOPBS_{(u,v)}$.

On the other hands, by elementary combinatorics,

$$\#BS_{u,v_\sigma} = \frac{N!}{(\ell_{\sigma(1)} - r + 1)!(\ell_{\sigma(2)} - r + 2)! \cdots \ell_{\sigma(r)}!},$$

so that by definition

$$A = \sum_{\sigma \in S_r} \sum_{P \in BS_{u,v_\sigma}} \text{sgn}\ P = \sum_{\sigma \in S_r} \text{sgn}\ \sigma \frac{N!}{(\ell_{\sigma(1)} - r + 1)!(\ell_{\sigma(2)} - r + 2)! \cdots \ell_{\sigma(r)}!} = N! \det \left(\frac{1}{(\ell_i - r + j)!} \right)_{i,j=1}^r$$

and, since $\ell_i = \lambda_i + r - i$, we finally obtain

$$F_\lambda = \#SOPBS_{u,v} = A = N! \det \left(\frac{1}{(\lambda_i - i + j)!} \right)_{i,j=1}^r .$$

4.2 Gessel's formula

We now go back to our main issue, which is to find a formula for $\mathbb{P}\{L(t) \le n\}$. Using the RS correspondence, combined with the Young-Frobenius formula, we can write it as

$$\mathbb{P}\{L(t) \le n\} = \mathrm{e}^{-t^2} \sum_{\lambda | \lambda_1 \le n} \frac{t^{2|\lambda|}}{(|\lambda|!)^2} F_\lambda^2 = \mathrm{e}^{-t^2} \sum_{\lambda | r(\lambda) \le n} t^{2|\lambda|} \left[\det \left(\frac{1}{(\lambda_i - i + j)!} \right)_{i,j=1}^n \right]^2 . \qquad (4.3)$$

To establish the second identity, where $r(\lambda)$ is equal to the number of row of λ, we used the fact that $F_\lambda = F_{\lambda^{\mathrm{T}}}$, and that the length of the first row of λ is equal to the number of row of λ^{T}.

Gessel's formula gives $\mathbb{P}\{L(t) \le n\}$, in terms of a Toeplitz determinant, which we now define.

Definition 4.7. *Let $a : S^1 \longrightarrow \mathbb{C}$ an integrable function on the unit circle, and let*

$$a_k := \int_{|z|=1} z^{-k} a(z) \frac{\mathrm{d}z}{2\pi \mathrm{i} z}. \tag{4.4}$$

its Fourier coefficients. We define the $(n+1)\times(n+1)$ Toeplitz matrix and its determinant as

$$T_n(a) := (a_{j-k})_{j,k=0}^n, \qquad D_n(a) := \det(T_n(a)). \tag{4.5}$$

Theorem 4.8. *Let*

$$\varphi(t;z) \equiv \varphi(z) = \mathrm{e}^{t(z+z^{-1})}.$$

For any $n \ge 0$,

$$\mathbb{P}\{L(t) \le n\} = \mathrm{e}^{-t^2} D_{n-1}(\varphi). \tag{4.6}$$

For the proof of Gessel's formula, we will use another well known formula, which is known under the name of *Andréief's identity.*

Lemma 4.9. *Let (X,μ) be a measure space, and $(f_i)_{i=1}^n, (g_i)_{i=1}^n$ two sequences of square-integrable functions. Then*

$$\int_X \cdots \int_X \det\Big(f_i(x_j)\Big)_{i,j=1}^n \det\Big(g_i(x_j)\Big)_{i,j=1}^n \prod_{k=1}^n \mathrm{d}\mu(x_k) = n! \det\left(\int_X f_i(x) g_j(x) \mathrm{d}\mu(x)\right)_{i,j=1}^n. \tag{4.7}$$

Proof. We apply the definition of det for $\det(g_i(x_j))$ and rewrite the left-hand side of (4.7) as

$$\sum_{\pi\in S_n} \operatorname{sgn}\pi \int_{X^n} \det\Big(f_i(x_j)\Big) \prod_{j=1}^n g_{\pi(j)}(x_j) \prod_{k=1}^n \mathrm{d}\mu(x_k) = \sum_{\pi\in S_n} \operatorname{sgn}\pi \int_{X^n} \det\Big(f_i(x_j) g_{\pi(j)}(x_j)\Big) \prod_{k=1}^n \mathrm{d}\mu(x_k)$$

$$\sum_{\pi\in S_n} \operatorname{sgn}\pi \det\left(\int_X f_i(x) g_{\pi(j)}(x) \mathrm{d}\mu(x)\right) = \sum_{\pi\in S_n} \det\left(\int_X f_i(x) g_j(x) \mathrm{d}\mu(x)\right) = n! \det\left(\int_X f_i(x) g_j(x) \mathrm{d}\mu(x)\right).$$

□

Proof of Theorem 4.8. We start from equation (4.3) and write it back in function of the variables

$$\ell_j := \lambda_j + n - j.$$

Since $\sum_{j=1}^n \ell_j = |\lambda| + n^2 - \frac{n(n-1)}{2}$, we obtain that

$$\mathbb{P}\{L(t) \le n\} = \mathrm{e}^{-t^2} t^{-n(n-1)} \sum_{\ell_1 > \ell_2 \cdots > \ell_n \ge 0} \left[t^{\sum_j \ell_j} \det\left(\frac{1}{(\ell_i - n + j)!}\right)_{i,j=1}^n \right]^2$$

$$= \mathrm{e}^{-t^2} \sum_{\ell_1 > \ell_2 \cdots > \ell_n \ge 0} \left[\det\left(\frac{t^{\ell_i - n + j}}{(\ell_i - n + j)!}\right)_{i,j=1}^n\right]^2 = \mathrm{e}^{-t^2} \sum_{\ell_1 > \ell_2 \cdots > \ell_n \ge 0} \left[\det\left(\frac{t^{\ell_i - j}}{(\ell_i - j)!}\right)_{i,j=0}^{n-1}\right]^2$$

The determinant in the last sum is symmetric in $\ell_1, \ldots, \ell_n$, and vanishes when $\ell_i = \ell_j$. This means that, after taking a transpose, we can rewrite

$$\mathbb{P}\{L(t) \le n\} = \frac{\mathrm{e}^{-t^2}}{n!} \sum_{\ell = (\ell_1, \ldots, \ell_n) \in \mathbb{Z}_{\ge 0}^n} \left[\det\left(\frac{t^{\ell_j - i}}{(\ell_j - i)!}\right)_{i,j=0}^{n-1}\right]^2.$$

At this stage, we can apply Andreief's formula, using the counting measure of $\mathbb{Z}_{\geq 0}$, and so we obtain

$$\mathbb{P}\{L(t) \leq n\} = \mathrm{e}^{-t^2} \left[\det \left(\sum_{m \geq 0} \frac{t^{2m-i-j}}{(m-i)!(m-j)!} \right)_{i,j=1}^{n} \right].$$

We now analyze the term inside the determinant. Using the fact that $\frac{1}{a!} = 0$ for $a < 0$ we rewrite

$$\sum_{m \geq 0} \frac{t^{2m-i-j}}{(m-i)!(m-j)!} = \sum_{m=-\infty}^{\infty} \frac{t^{2m-i-j}}{(m-i)!(m-j)!} = \sum_{m=-\infty}^{\infty} \frac{t^{2m+i-j}}{m!(m+i-j)!} =: d_{i-j}.$$

Hence, we obtained that

$$\mathbb{P}\{L(t) \leq n\} = \mathrm{e}^{-t^2} \det(d_{i-j})_{i,j=0}^{n-1}.$$

To conclude, we now compute the generating function of $\sum_{k \in \mathbb{Z}} d_k z^k$:

$$\sum_{k \in \mathbb{Z}} d_k z^k = \sum_{k \in \mathbb{Z}} \sum_{m \in \mathbb{Z}} \frac{t^{2m+k}}{m!(m+k)!} z^k = \sum_{m,s \in \mathbb{Z}} \frac{t^{m+s}}{m!s!} z^{s-m} = \mathrm{e}^{t(z+z^{-1})} = \varphi(z).$$

□

4.3 The Borodin-Okounkov formula

Once the Gessel formula

$$\mathbb{P}(L(t) \leq n) = \mathrm{e}^{-t^2} D_{n-1}(\varphi), \qquad \varphi(z) = \mathrm{e}^{t(z+z^{-1})}$$

is established, in order to obtain the Baik-Deift-Johansson theorem one has to study large n asymptotics of the Toeplitz determinants above. Two approaches are possible: one is to study the related orthogonal polynomial on the unit circle, and this what the next chapters are about. Without giving much details, here we discuss another approach, based on a formula connecting Toeplitz determinants with determinants of operators (namely, Fredholm determinants). In what follows, we consider H as the Hilbert space $L^2(S^1)$, equipped with the standard scalar product

$$< f, g > := \int_{-\pi}^{\pi} f(\theta)\overline{g(\theta)} \frac{\mathrm{d}\theta}{2\pi} = \int_{S^1} f(z)\overline{g(z)} \frac{\mathrm{d}z}{2\pi \mathrm{i} z}. \tag{4.8}$$

It will be convenient, sometimes, to identify H with $\ell^2(\mathbb{Z})$, via the map

$$f(\theta) = \sum_{k \in \mathbb{Z}} f_k \mathrm{e}^{\mathrm{i}k\theta} \mapsto (f_k)_{k \in \mathbb{Z}}.$$

We will also denote with $H_\pm$ the subspaces of fonctions in which just the non–negative (or negative) Fourier modes are non-zero. Of course $H = H_+ \oplus H_-$. We will denote with $\Pi_\pm$ the orthogonal projections onto $H_\pm$, and with ι the involution $(\iota f)(z) = z^{-1} f(z^{-1})$.

Now, suppose that f is a (essentially) bounded function on S^1, such that $\sum_{k \in \mathbb{Z}} |k||f_k|^2 < \infty$. We introduce three different operators, acting on H_+ associated to f:

$$T(f) = \Pi_+ \circ f_{|H_+}, \qquad H(f) = \Pi_+ \circ f \circ \iota_{|H_+}, \qquad \widetilde{H}(f) = \iota \circ \Pi_- \circ f_{|H_+}. \tag{4.9}$$

One can easily verify that, with respect to the standard base for H_+, the three operators above are associated to the following $\mathbb{N} \times \mathbb{N}$ matrices

$$T(f) = (f_{\ell-m})_{\ell,m \in \mathbb{N}}, \qquad H(f) = (f_{1+\ell+m})_{\ell,m \in \mathbb{N}}, \qquad \widetilde{H}(f) = (f_{-1-\ell-m})_{\ell,m \in \mathbb{N}}. \tag{4.10}$$

Another important (and easy to prove) identity is

$$T(f)T(g) = T(fg) - H(f)\widetilde{H}(g) \tag{4.11}$$

which tells you, in particular, that $T(f)T(g) = T(fg)$ whenever $T(f)$ is upper triangular or $T(g)$ is lower triangular. We are ready to prove the Borodin-Okounkov formula, as stated (and proved) in [5].

Theorem 4.10. *Suppose that φ is a function on S^1, such that*

$$\varphi(z) = \varphi_+(z)\varphi_-(z) = \mathrm{e}^{V_+(z)}\mathrm{e}^{V_-(z)}, \quad \textit{with} \quad V_+(z) = \sum_{k>0}(V_+)_k z^k \quad \textit{and} \quad V_-(z) = \sum_{k<0}(V_-)_k z^k.$$

Suppose moreover that $V_\pm$ are bounded and $\sum_k |k||(\log\varphi)_k|^2 < \infty$.
Then

$$D_{n-1}(\varphi) = Z(\varphi)\det(\mathrm{Id} - \mathcal{K}_{|\ell^2\{n,n+1,n+2,\dots\}}). \tag{4.12}$$

with $Z(\varphi) = \det(T(\varphi)T(\varphi^{-1})) = \det(\mathrm{Id} - H(\varphi)\widetilde{H}(\varphi^{-1}))$ and $\mathcal{K} = H(\varphi_-/\varphi_+)\widetilde{H}(\varphi_+/\varphi_-)$. All the determinants above are Fredholm determinant, meaning that

$$\det(\mathrm{Id} - \mathcal{K}_{|\ell^2\{n,n+1,n+2,\dots\}}) = \sum_{m=0}^{\infty}(-1)^m \sum_{n\le\ell_1<\dots<\ell_m} \det\Big[\mathcal{K}(\ell_i,\ell_j)\Big]_{i,j=1}^{m}, \tag{4.13}$$

and analogously for Z.

Proof. We will work with infinite matrices indexed on $\mathbb{N}$. We start observing that, denoting with P_n the infinite matrix with n ones on the main diagonal and all the other entries equal to 0, we have to compute the determinant on the $n\times n$ upper -left principal minor of $P_nT(\varphi)P_n$. We also observe that

$$P_nT(\varphi_+) = P_nT(\varphi_+)P_n \qquad \text{and} \qquad T(\varphi_-)P_n = P_nT(\varphi_-)P_n.$$

Hence, we can use the following chain of equalities

$$\begin{aligned} P_nT(\varphi)P_n &= P_nT(\varphi_+)T(\varphi_+^{-1})T(\varphi)T(\varphi_-^{-1})T(\varphi_-)P_n = \\ &P_nT(\varphi_+)P_n\Big[P_nT(\varphi_+^{-1})T(\varphi)T(\varphi_-^{-1})P_n\Big]P_nT(\varphi_-)P_n \end{aligned} \tag{4.14}$$

The $n\times n$ upper-left principal minors of the matrices outside the square brackets have determinant equal to 1, so that we are just interested in $A := T(\varphi_+^{-1})T(\varphi)T(\varphi_-^{-1})$. More precisely, we have to compute $\det(P_nAP_n + Q_n)$, where $Q_n := \mathrm{Id} - P_n$. Upon proving that A is invertible and differs from the identity by a trace-class operator (which we will do in a moment) we can write

$$\begin{aligned} \det(P_nAP_n + Q_n) &= \det(A)\det(A^{-1}P_nAP_n + A^{-1}Q_n) = \det(A)\det(A^{-1}(\mathrm{Id} - Q_n)AP_n + A^{-1}Q_n) = \\ &\det(A)\det(P_n - A^{-1}Q_nAP_n + A^{-1}Q_n) = \det(A)\det(P_n + A^{-1}Q_n)\det(\mathrm{Id} - Q_nAP_n). \end{aligned}$$

The last determinant is equal to 1, hence we are left with

$$\det(A)\det(\mathrm{Id} - (\mathrm{Id} - A^{-1})Q_n) = \det(A)\det(\mathrm{Id} - Q_n(\mathrm{Id} - A^{-1})Q_n) \tag{4.15}$$

We now observe that

$$\det(A) = \det(T(\varphi_+^{-1})T(\varphi)T(\varphi_-^{-1})) = \det(T(\varphi)T(\varphi_-^{-1})T(\varphi_+^{-1})) = \det(T(\varphi)T(\varphi^{-1})) = \det(\mathrm{Id} - H(\varphi)\widetilde{H}(\varphi^{-1})),$$

which also proves that $\det A$ is well defined (indeed, under the hypothesis above, both $H(\varphi)$ and $\widetilde{H}(\varphi^{-1})$ are Hilbert-Schmidt operators: see [5] and reference therein). As for the second part, we have

$$\begin{aligned} \mathrm{Id} - A^{-1} &= \mathrm{Id} - T(\varphi_-)(T(\varphi))^{-1}T(\varphi_+) = \mathrm{Id} - T(\varphi_-)T(\varphi_+^{-1})T(\varphi_-^{-1})T(\varphi_+) \\ &= \mathrm{Id} - T(\varphi_-/\varphi_+)T(\varphi_+/\varphi_-) = H(\varphi_-/\varphi_+)\widetilde{H}(\varphi_+/\varphi_-). \end{aligned}$$

□

Remark 4.11. *One can also prove a more explicit expression for Z; namely*

$$Z = \exp\left(\sum_{k\ge0} k(V_+)_k(V_-)_k\right).$$

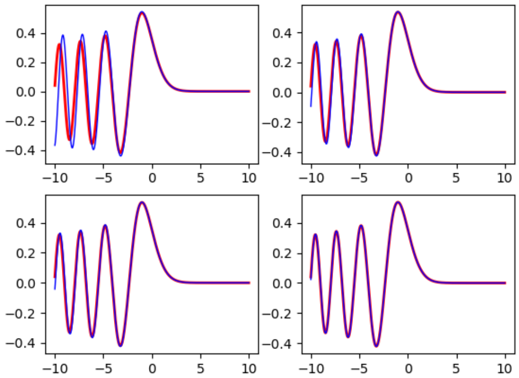


Figure 6: In blue the graph of the Airy function, and in red the graph of $t^{1/3}J_{2t+xt^{1/3}}(2t)$ for $t = 10, 50, 100, 1000$.

This constant $Z(\varphi)$, indeed, is the one appearing in the strong Szegö theorem, stating that

$$\lim_{n\to\infty} D_n(\varphi) = Z(\varphi).$$

In the case we are interested in, $Z(\varphi) = \exp(t^2)$, and hence

$$\mathbb{P}(L(t) \leq n) = \det(\mathrm{Id} - \mathcal{K}_{|\ell^2\{n,n+1,n+2,\ldots\}}).$$

4.4 Towards the Baik-Deift-Johansson Theorem

In this short paragraph, we will outline a proof of the Baik-Deift-Johansson which is different from the original one, and goes through the asymptotic analysis of the Fredholm determinant given by the Borodin-Okounkov formula. To start with, we give an alternative definition of the Tracy-Widom distribution (which is actually the first one that appeared in the literature). To construct it, we need to introduce the Airy kernel as the function of two real variables defined as follows

$$K_{\mathrm{Ai}}(x,y) = \int_0^{+\infty} \mathrm{Ai}(x+z)\mathrm{Ai}(y+z)dz = \frac{\mathrm{Ai}(x)\mathrm{Ai}'(y) - \mathrm{Ai}'(x)\mathrm{Ai}(y)}{x-y}, \quad x, y \in \mathbb{R} \tag{4.16}$$

where $\mathrm{Ai}(\cdot)$ denotes the Airy function, a real, fast decaying (at $+\infty$) solution of the Airy differential equation

$$f''(t) = tf(t) \tag{4.17}$$

that can be represented as

$$\mathrm{Ai}(t) = \frac{1}{\pi}\int_0^{+\infty} \cos\left(\frac{x^3}{3} + xt\right) dx. \tag{4.18}$$

The integral operator associated to the Airy kernel $\mathcal{K}_{\mathrm{Ai}}$ acting on the space $L^2((0,\infty))$-functions is defined as

$$\left(\mathcal{K}_{\mathrm{Ai}}h\right)(x) = \int_0^{+\infty} K_{\mathrm{Ai}}(x,y)h(y)dy. \tag{4.19}$$

It has the properties to be : Hermitian, locally trace-class (trace-class on every bounded Borel subset of $\mathbb{R}$) and such that $0 \leq \mathcal{K}_{\mathrm{Ai}} \leq 1$.

Remark 4.12. *We shortly remind some definitions and results for compact operator on Hilbert spaces. For more details, look at [15].*

- *A compact operator on a Hilbert space $\mathcal{H}$ (for us $L^2(J), J \subset \mathbb{R}$), is said to be of trace-class if it has finite trace, or*

$$Tr(\mathcal{K}) := \sum_{n\geq 1} \lambda_n(|\mathcal{K}|) < +\infty \tag{4.20}$$

where $\lambda_n(|\mathcal{K}|)$ denote the the singular values of $\mathcal{K}$ or the eigenvalues of $|\mathcal{K}| = (\mathcal{K}\mathcal{K}^*)^{\frac{1}{2}}$ ($\mathcal{K}^*$ being the adjoint operator of $\mathcal{K}$). For a (compact) integral operator with kernel K on $L^2((J))$ then its trace

$$Tr(\mathcal{K}) = \int_J K(x,x)dx. \tag{4.21}$$

- *For a trace-class operator $\mathcal{K}$ we define the Fredholm determinant of $1+\mathcal{K}$ as*

$$\det(1+\mathcal{K}) = \sum_{n\geq 0} Tr(\Lambda^n(\mathcal{K})) \tag{4.22}$$

where $\Lambda^n(\mathcal{K})$ denotes the operator obtained by tensoring n times $\mathcal{K}$ i.e. by taking $\mathcal{K}\otimes\cdots\otimes\mathcal{K}$ considered acting on the space $\Lambda^n\mathcal{H}$.

For a trace-class integral operator with kernel K on $L^2((J))$ then its Fredholm determinant coincide with the following Fredholm expansion

$$\det(1+\mathcal{K}) = \sum_{n=0} \frac{1}{n!} \int_J \cdots \int_J \det_{i,j=1}^{n} \left(K(x_i,x_j)\right) dx_1 \dots dx_n \tag{4.23}$$

The Tracy-Widom distribution is then defined as the function of the real variable s such that

$$F_{GUE}(s) = \det(1 - \mathcal{K}_{\mathrm{Ai},s}) \tag{4.24}$$

where in the right hand side we find the Fredholm determinant of the trace-class integral operator (over $L^2((0,\infty))$) acting through the s shifted Airy kernel kernel $K_{\mathrm{Ai},s}$ given by

$$K_{\mathrm{Ai},s}(x,y) = K_{\mathrm{Ai}}(x+s, y+s). \tag{4.25}$$

The equivalence between (1.7) and (4.24) is exactly the Tracy-Widom theorem [25].

Going back to the Borodin-Okounkov formula (4.12), we start noticing that, in our specific case,

$$\varphi_-(z)/\varphi_+(z) = \exp^{t(z^{-1}-z)} = \sum_{m\geq 0} (-1)^m z^m J_m(2t). \tag{4.26}$$

where J_m are the Bessel function of the first kind, solutions of the equation

$$t^2\frac{d^2w}{dt^2} + t\frac{dw}{dt} + (z^2 - m^2)w = 0$$

(see [11]). Hence, we obtain that

$$H(\varphi_-(z)/\varphi_+(z))_{\ell,m} = (-1)^{1+\ell+m} J_{1+\ell+m}(2t)$$

and

$$\widetilde{H}(\varphi_+(z)/\varphi_-(z))_{\ell,m} = J_{-1-\ell-m}(2t) = (-1)^{1+\ell+m} J_{1+m+\ell}(2t).$$

Hence,

$$\mathcal{K}(i,j) = \sum_{\ell\geq 0} (-1)^{i+j} J_{1+i+\ell}(2t) J_{1+\ell+j}(2t).$$

Up to the sign $(-1)^{i+j}$ (which does not change the determinants we are interested in) this is the so-called *discrete Bessel kernel*, which is nothing but the kernel of the determinantal point process associated to the Poissonized Plancherel measure (see Figure 4). Frequently, it is written as acting on the set of semi-integers $\mathbb{Z}'$, and in this case it reads

$$K^{\mathrm{Be}}(a,b) = \sum_{\ell\in\mathbb{Z}'} J_{a+\ell}(2t)J_{b+\ell}(2t).$$

Now, Bessel functions satisfy, with respect to the order, a discrete version of the Airy equation:

$$J_{n-1}(2t) + J_{n+1}(2t) = \frac{n}{t}J_n(2t)$$

and indeed, in a certain asymptotic regime, they converge to the latter. Namely

$$\lim_{t\to+\infty} t^{1/3}J_{2t+st^{1/3}}(2t) = \mathrm{Ai}(s)$$

uniformly on compact set (see [11] and Figure 6). Using this property, one can prove that

$$\lim_{t\to\infty} t^{1/3}K^{\mathrm{Be}}(\lfloor 2t+xt^{1/3}\rfloor, \lfloor 2t+yt^{1/3}\rfloor) \to K^{\mathrm{Ai}}(x,y) := \int_0^\infty \mathrm{Ai}(x+z)\mathrm{Ai}(y+z)\mathrm{d}z. \tag{4.27}$$

Taking care of the necessary analytical details (we did not explain, for instance, how the convergence of the kernels imply the convergence of the corresponding Fredholm determinants), as done in [7] and [20], one obtains an alternative proof of the Baik-Deift-Johasson theorem and much more, since this method also gives an interpretation of the theorem in term of convergence of point processes, see [7] for details.

5 Orthogonal polynomials on the unit circle

In this Section we define the orthogonal polynomials on the unit circle (OPUC), and we see their main properties namely: the formula for the OPUC in terms of Toeplitz determinants and OP, the Szego reccurence relations and their Riemann-Hilbert characterization. For the first two parts we follow Chapter XI of [24] and for the last part Chapter V of [4].

5.1 Introduction to OPUC

We proceed with general facts about the theory of orthogonal polynomials with respect to a measure defined on the unit circle $S^1 = \{z \in \mathbb{C}, |z| = 1\}$. For a given (well-defined) measure, the definition of the corresponding family of orthogonal polynomials is analogue to the one in the real case , i.e. for (well-defined) measures supported on the real line. The properties of such families of orthogonal polynomials are then slight generalizations of the ones holding for real orthogonal polynomials.

Let us consider a function $f(\theta)$ defined for real values of θ, positive and 2π periodic, integrable and such that

$$\int_{-\pi}^{\pi} f(\theta)d\theta > 0 \tag{5.1}$$

and we consider the measure for $z \in S^1, z = \mathrm{e}^{\mathrm{i}\theta}$ given by

$$d\mu(\theta) = f(\theta)\frac{d\theta}{2\pi}. \tag{5.2}$$

The family of orthonormal polynomials on the unit circle $p_n(z), n \geq 0$ defined with respect to the measure (5.2) are uniquely defined by

- $p_n(z) = \kappa_n z^n + \dots$ for every $n \geq 0$ with $\kappa_n \in \mathbb{R}_{>0}$;
- and the orthogonality condition

$$\int_{-\pi}^{\pi} p_n(\mathrm{e}^{\mathrm{i}\theta})\overline{p_m(\mathrm{e}^{\mathrm{i}\theta})}f(\theta)\frac{d\theta}{2\pi} = \delta_{n,m} \tag{5.3}$$

  for all $n, m \geq 0$.

The polynomials $p_n(z)$ are obtained by orthogonalizing the family

$$f(\theta)^{\frac{1}{2}} z^n, \; z = \mathrm{e}^{\mathrm{i}\theta}, \; n \geq 0 \tag{5.4}$$

or in other words by finding the coefficients of $p_n(z)$ imposing the orthogonality condition to be zero for any $z^k, k < n$ and then finding the normalization constant.

Remark 5.1. *The corresponding monic orthogonal polynomials are denoted by*

$$\pi_n(z) = \kappa_n^{-1} p_n(z), n \geq 0. \tag{5.5}$$

Remark 5.2. *For even functions f (such that $f(\theta) = f(-\theta)$ for all $\theta \in \mathbb{R}$) the coefficients of $p_n(z)$ are all real.*

5.2 Toeplitz determinant formula for OPUC

An useful explicit formula for any $p_n(z)$ is given in terms of the Toeplitz determinants associated to the measure (5.2).

Fo any such given measure, let us recall the definition of the Fourier coefficients $f_n, n \geq 0$ given in (4.4), equivalent to

$$f_n = \int_{-\pi}^{\pi} \mathrm{e}^{-\mathrm{i}n\theta} f(\theta) \frac{d\theta}{2\pi}. \tag{5.6}$$

Remark 5.3. *We have $\overline{f_n} = f_{-n}$ for all $n \geq 0$.*

Then for any $n \geq 0$, the Toeplitz matrix T_n associated to the function (symbol) f, of size $n+1$, and their determinants $D_n = \det(T_n)$ are defined as in equation (4.5).

Remark 5.4. *The matrix T_n is Hermitian for every n. Moreover, it is positive definite. Indeed if we consider the associated quadratic form τ_n, we can see that*

$$\begin{aligned} \tau_n(u) &= \sum_{k,l=0}^{n} (T_n)_{k,l} u_k \overline{u}_l = \sum_{k,l=0}^{n} f_{k-l} u_k \overline{u}_l = \frac{1}{2\pi} \int_{-\pi}^{\pi} \sum_{k,l=0}^{n} \mathrm{e}^{-i(k-l)\theta} f(\theta) u_k \overline{u}_l d\theta \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \sum_{k=0}^{n} \mathrm{e}^{-ik\theta} u_k \sum_{l=0}^{n} \mathrm{e}^{il\theta} \overline{u}_l f(\theta) d\theta = \frac{1}{2\pi} \int_{-\pi}^{\pi} \left| \sum_{k=0}^{n} \mathrm{e}^{-ik\theta} u_k \right|^2 f(\theta) d\theta > 0. \end{aligned} \tag{5.7}$$

In particular we have $D_n \in \mathbb{R}_{>0}$.

Proposition 5.5. *For every $n \geq 0$, the n-th orthonormal polynomial for the given measure $d\mu(\theta)$ in (5.2) is given by*

$$p_n(z) = \frac{1}{\sqrt{D_n D_{n-1}}} \begin{vmatrix} f_0 & f_{-1} & f_{-2} & \cdots & f_{-n+1} & f_{-n} \\ f_1 & f_0 & f_{-1} & \cdots & f_{-n+2} & f_{-n+1} \\ f_2 & f_1 & f_0 & \cdots & f_{-n+3} & f_{-n+2} \\ \vdots & & & \ddots & & \vdots \\ f_{n-1} & f_{n-2} & & \cdots & f_0 & f_{-1} \\ 1 & z & z^2 & \cdots & z^{n-1} & z^n \end{vmatrix}. \tag{5.8}$$

In particular $\kappa_n = \sqrt{\frac{D_{n-1}}{D_n}}$.

Proof. Let us call $q_n(z)$ the right hand side of formula (5.8). First we show that $q_n(z)$ is orthogonal to every $z^k, k<n$. Indeed

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} q_n(\mathrm{e}^{i\theta})\overline{\mathrm{e}^{ik\theta}}f(\theta)d\theta = \frac{1}{2\pi}\int_{-\pi}^{\pi} q_n(\mathrm{e}^{i\theta})\mathrm{e}^{-ik\theta}f(\theta)d\theta$$

$$= \frac{1}{\sqrt{D_nD_{n-1}}}\begin{vmatrix} f_0 & f_{-1} & f_{-2} & \cdots & f_{-n+1} & f_{-n} \\ f_1 & f_0 & f_{-1} & \cdots & f_{-n+2} & f_{-n+1} \\ f_2 & f_1 & f_0 & \cdots & f_{-n+3} & f_{-n+2} \\ \vdots & & & \ddots & & \vdots \\ f_{n-1} & f_{n-2} & & \cdots & f_0 & f_{-1} \\ \frac{1}{2\pi}\int_{-\pi}^{\pi}\mathrm{e}^{-ik\theta}f(\theta)d\theta & \frac{1}{2\pi}\int_{-\pi}^{\pi}\mathrm{e}^{i(-k+1)\theta}f(\theta)d\theta & & \cdots & & \frac{1}{2\pi}\int_{-\pi}^{\pi}\mathrm{e}^{i(-k+n)\theta}f(\theta)d\theta \end{vmatrix} = 0 \tag{5.9}$$

since we notice that the last line in the matrix, for $k<n$, corresponds to one of the previous line and so the determinant of this matrix is zero. Thus the $q_n(z)$ satisfy the orthogonality condition, up to the normalization constant they should coincide with the $p_n(z)$. To see it, we consider now

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} q_n(\mathrm{e}^{\mathrm{i}\theta})\overline{q_n(\mathrm{e}^{\mathrm{i}\theta})}f(\theta)d\theta =$$

$$\frac{1}{2\pi D_nD_{n-1}}\int_{-\pi}^{\pi}\begin{vmatrix} f_0 & f_{-1} & f_{-2} & \cdots & f_{-n+1} & f_{-n} \\ f_1 & f_0 & f_{-1} & \cdots & f_{-n+2} & f_{-n+1} \\ f_2 & f_1 & f_0 & \cdots & f_{-n+3} & f_{-n+2} \\ \vdots & & & \ddots & & \vdots \\ f_{n-1} & f_{n-2} & & \cdots & f_0 & f_{-1} \\ 1 & \mathrm{e}^{\mathrm{i}\theta} & \mathrm{e}^{\mathrm{i}2\theta} & \cdots & \mathrm{e}^{\mathrm{i}(n-1)\theta} & \mathrm{e}^{\mathrm{i}n\theta} \end{vmatrix} \underbrace{\overline{\begin{vmatrix} f_0 & f_{-1} & f_{-2} & \cdots & f_{-n+1} & f_{-n} \\ f_1 & f_0 & f_{-1} & \cdots & f_{-n+2} & f_{-n+1} \\ f_2 & f_1 & f_0 & \cdots & f_{-n+3} & f_{-n+2} \\ \vdots & & & \ddots & & \vdots \\ f_{n-1} & f_{n-2} & & \cdots & f_0 & f_{-1} \\ 1 & \mathrm{e}^{\mathrm{i}\theta} & \mathrm{e}^{\mathrm{i}2\theta} & \cdots & \mathrm{e}^{\mathrm{i}(n-1)\theta} & \mathrm{e}^{\mathrm{i}n\theta} \end{vmatrix}}}_{=\mathrm{e}^{-\mathrm{i}n\theta}D_{n-1}+\sum_{k=0}^{n-1}\mathrm{e}^{-\mathrm{i}k\theta}\overline{D_n^{n+1,k}}(-1)^{n-k}} f(\theta)d\theta$$

$$= \frac{1}{2\pi D_n}\int_{-\pi}^{\pi}\begin{vmatrix} f_0 & f_{-1} & f_{-2} & \cdots & f_{-n+1} & f_{-n} \\ f_1 & f_0 & f_{-1} & \cdots & f_{-n+2} & f_{-n+1} \\ f_2 & f_1 & f_0 & \cdots & f_{-n+3} & f_{-n+2} \\ \vdots & & & \ddots & & \vdots \\ f_{n-1} & f_{n-2} & & \cdots & f_0 & f_{-1} \\ 1 & \mathrm{e}^{\mathrm{i}\theta} & \mathrm{e}^{\mathrm{i}2\theta} & \cdots & \mathrm{e}^{\mathrm{i}(n-1)\theta} & \mathrm{e}^{\mathrm{i}n\theta} \end{vmatrix}\mathrm{e}^{-\mathrm{i}n\theta}f(\theta)d\theta$$

$$= \frac{1}{D_n}\begin{vmatrix} f_0 & f_{-1} & f_{-2} & \cdots & f_{-n+1} & f_{-n} \\ f_1 & f_0 & f_{-1} & \cdots & f_{-n+2} & f_{-n+1} \\ f_2 & f_1 & f_0 & \cdots & f_{-n+3} & f_{-n+2} \\ \vdots & & & \ddots & & \vdots \\ f_{n-1} & f_{n-2} & & \cdots & f_0 & f_{-1} \\ \frac{1}{2\pi}\int_{-\pi}^{\pi}\mathrm{e}^{-\mathrm{i}n\theta}f(\theta)d\theta & \frac{1}{2\pi}\int_{-\pi}^{\pi}\mathrm{e}^{-\mathrm{i}(n-1)\theta}f(\theta)d\theta & & \cdots & & \frac{1}{2\pi}\int_{-\pi}^{\pi}f(\theta)d\theta \end{vmatrix} = 1 \tag{5.10}$$

where in the second identity we used the orthogonality we just proved and in the last line we recognize the last row as the one given by $(f_n, \dots f_0)$ and so the determinant as exactly D_n. This concludes the proof of formula (5.8). □

Remark 5.6. *This is the analogue of the formula for real orthogonal polynomials in terms of Hanekl type determinants, instead of Toeplitz (see e.g. [24, 10]).*

Remark 5.7. *There are also other explicit formulas for the OPUC, see the exercises section.*

5.3 Szego relations

General orthogonal polynomials on the real line are known to be equivalently defined by the so called three terms recurrence relation (see [24], Chapter 3). For the OPUC this same property does not hold, but still the family of $p_n(z)$ and their reciprocal polynomials $p_n^*(z)$, defined as polynomials of the same degree given by

$$p_n^*(z) = z^n \overline{p_n(\bar{z}^{-1})} = z^n \overline{p_n}(z^{-1}), \tag{5.11}$$

they satisfy together a coupled system of recurrence relations called the Szego relations.

Theorem 5.8 (e.g. [24], Theorem 11.4.2). *The following identity holds for all z, a,*

$$\sum_{k=0}^{n} \overline{p_k(a)} p_k(z) = \frac{\overline{p_{n+1}^*(a)} p_{n+1}^*(z) - \overline{p_{n+1}(a)} p_{n+1}(z)}{1 - \bar{a}z}, \tag{5.12}$$

and so the recurrence formulas

$$\begin{aligned} \kappa_n z p_n(z) &= \kappa_{n+1} p_{n+1}(z) - p_{n+1}(0) p_{n+1}^*(z) \\ \kappa_n p_{n+1}(z) &= \kappa_{n+1} z p_n(z) + p_{n+1}(0) p_n^*(z). \end{aligned} \tag{5.13}$$

Proof. Details of the proof of the first formula (5.12) can be found in [24], pg. 293. Here we only explain how to derive the recurrence relations (5.13) starting from (5.12). For the first one, it is enough to look at the coefficients of the term $\bar{a}^{n+1}$ both sides in

$$(1 - \bar{a}z) \sum_{k=0}^{n} \overline{p_k(a)} p_k(z) = \overline{p_{n+1}^*(a)} p_{n+1}^*(z) - \overline{p_{n+1}(a)} p_{n+1}(z) \tag{5.14}$$

i.e. (5.12) after multiplying both sides by $1 - \bar{a}z$.

On the left hand side we get only one term containing $\bar{a}^{n+1}$ (coming from $-\bar{a}z\overline{p_n(a)}p_n(z) = -\bar{a}z(\kappa_n \bar{a}^n + \dots)p_n(z)$) which is

$$-z\kappa_n p_n(z). \tag{5.15}$$

On the right hand side instead we get two terms (one for each of the addends), they are

$$p_{n+1}(0) p_{n+1}^*(z) - \kappa_{n+1} p_{n+1}(z). \tag{5.16}$$

In particular the first term is obtained by expanding

$$\overline{a^{n+1} \overline{p_{n+1}}(a^{-1})} p_{n+1}^*(z) = \bar{a}^{n+1} \left(\bar{a}^{-n-1} \kappa_{n+1} + \bar{a}^{-n} \gamma_{n+1,n} + \dots + \underbrace{\gamma_{n+1,0}}_{= p_{n+1}(0)} \right) p_{n+1}^*(z). \tag{5.17}$$

Thus equating these coefficients give

$$-z\kappa_n p_n(z) = p_{n+1}(0) p_{n+1}^*(z) - \kappa_{n+1} p_{n+1}(z) \tag{5.18}$$

which is the first equation in (5.13). The second equation can be obtained from (5.12) by looking both sides at the reciprocal polynomials, we leave as an exercise. □

Remark 5.9. *We will come back to these recurrence relation in the next Section, where we will reformulate it as a second matrix equation of a discrete Lax pair.*

Remark 5.10. *The equivalent version of the recurrence relation* (5.13) *for the monic orthogonal polynomials $\pi_n(z)$ is*

$$\begin{aligned} \pi_{n+1}(z) &= z\pi_n(z) - \bar{\alpha}_n \pi_n^*(z) \\ \pi_{n+1}^*(z) &= \pi_n^*(z) - \alpha_n z \pi_n(z) \end{aligned} \tag{5.19}$$

and the values $\alpha_n = -\overline{\pi_{n+1}(0)}$ are usually called the Verblusky coefficients.

5.4 Riemann-Hilbert approach

We now show an alternative characterization, via Riemann-Hilbert (matrix factorization) problems, of the family of orthogonal polynomials on the unit circle for the given measure $d\mu(\theta)$ (5.2) with the convention

$$f(\theta) = w(e^{i\theta}). \tag{5.20}$$

This characterization was first proven in the paper [3], inspired by the real counterpart previously discovered in [13]. Here we follow the presentation of the results given in Chapter V of [4].

Remark 5.11. *Let us consider a pair (Σ, J) where Σ is a smooth (possible with a finite number of endpoints or singular points) oriented contour in the complex plane and $J = J(z)$ is a matrix-valued sufficiently well-behaved function defined for all $z \in \Sigma$. Generally speaking, a Riemann-Hilbert (matrix factorization) problem for the given pair (Σ, J) is intended as the problem of finding a matrix-valued function $Y : \mathbb{C} \to GL(n, \mathbb{C}), n \geq 2$ such that*

- *$Y(z)$ is analytic for $z \in \mathbb{C} \setminus \Sigma$;*
- *$Y(z)$ have continuous boundary values when approaching Σ from the left (+ side) and the right (− side) non-tangentially $Y_\pm(z), z \in \Sigma$, namely*

$$Y_\pm(z) = \lim_{\substack{z' \to z \\ z' \in \pm \text{ side of } \Sigma}} Y(z') \tag{5.21}$$

  and they are related by $Y_+(z) = Y_-(z)J(z), z \in \Sigma$;

- *$Y(z)$ has a prescribed behaviour for $z \to \infty$ (and eventually also at the discontinuities points).*

Let us fix $n \geq 1$. The OPUC Riemann-Hilbert problem [3] is the following Riemann-Hilbert 2×2 matrix factorization problem, for the pair

$$(\Sigma = S^1, J = J_Y(z) = \begin{pmatrix} 1 & z^{-n}w(z) \\ 0 & 1 \end{pmatrix}, z \in S^1) \tag{5.22}$$

where S^1 is oriented counter-clockwise.

Riemann–Hilbert Problem 1. *The function $Y(z) := Y(n; z) : \mathbb{C} \to \mathrm{GL}(2, \mathbb{C})$, has the following properties:*

(1) $Y(z)$ is analytic for every $z \in \mathbb{C} \setminus S^1$;

(2) $Y(z)$ has continuous boundary values $Y_\pm(z)$ while approaching non-tangentially S^1 either from the left or from the right, and they are related for all $z \in S^1$ through

$$Y_+(z) = Y_-(z)J_Y(z), \quad \textit{with}; \tag{5.23}$$

(3) $Y(z)$ is normalized at ∞ as

$$Y(z) \sim \left(I + \sum_{j=1}^{\infty} \frac{Y_j(n)}{z^j} \right) z^{n\sigma_3}, \quad z \to \infty, \tag{5.24}$$

where σ_3 denotes the Pauli's matrix $\sigma_3 := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$.

It is known from [3] that the above Riemann-Hilbert problem, admits a unique solution which is explicitly written in terms of the family $\{\pi_n(z)\}, n \geq 0$. Before stating the result, we need to introduce the Cauchy transform and its well-known properties. For every ($L^p(S^1)$) function $f(y)$, its Cauchy transform $\mathcal{C}f(z)$ is defined for any $z \notin S^1$ as

$$(\mathcal{C}f)(z) := \frac{1}{2\pi \mathrm{i}} \int_{S^1} \frac{f(y)}{y-z} \mathrm{d}y. \tag{5.25}$$

A fundamental property for us, is that for the boundary operators

$$(\mathcal{C}_{\pm} f)(z) = \lim_{\substack{z' \to z \\ z' \in \pm \text{ side of } S^1}} (\mathcal{C}f)(z') \tag{5.26}$$

satisfy the operator equation

$$\mathcal{C}_+ - \mathcal{C}_- = I \tag{5.27}$$

for I the identity operator of $L^p(S^1)$ (Sokhotski-Plemelji formula) and we have the norm estimate, for a constant c_p

$$||\mathcal{C}_{\pm} f||_{L^p(S^1)} \leq c_p ||f||_{L^p(S^1)}. \tag{5.28}$$

A classical reference for this kind of results is [9, 14].

Theorem 5.12 ([3], Lemma 4.1). *For any $n \geq 1$ the Riemann-Hilbert problem* (1) *admits a unique solution*

$$Y(z) = \begin{pmatrix} \pi_n(z) & \mathcal{C}\left(y^{-n}\pi_n(y)w(y)\right)(z) \\ -\kappa_{n-1}^2 \pi_{n-1}^*(z) & -\kappa_{n-1}^2 \mathcal{C}\left(y^{-n}\pi_{n-1}^*(y)w(y)\right)(z) \end{pmatrix}. \tag{5.29}$$

Moreover, $\det(Y(z)) \equiv 1$.

Proof. We start by rewriting the jump condition (5.23) in terms of scalar equations for the entries of Y, denoted by $Y^{ij}, i,j = 1,2$. From the first column we have

$$\begin{aligned} Y_+^{11}(z) &= Y_-^{11}(z), \\ Y_+^{21}(z) &= Y_-^{21}(z) \end{aligned} \tag{5.30}$$

while from the second column we have

$$\begin{aligned} Y_+^{12}(z) &= Y_-^{11}(z) z^{-n} w(z) + Y_-^{12}(z), \\ Y_+^{22}(z) &= Y_-^{21}(z) z^{-n} w(z) + Y_-^{22}(z). \end{aligned} \tag{5.31}$$

We rewrite as well the asymptotic condition for $z \to \infty$ (5.10) entry by entry, for the first column

$$\begin{aligned} Y^{11}(z) &= z^n + \mathcal{O}(z^{n-1}), \\ Y^{21}(z) &= \mathcal{O}(z^{n-1}) \end{aligned} \tag{5.32}$$

and for the second column

$$\begin{aligned} Y^{12}(z) &= \mathcal{O}(z^{-n-1}), \\ Y^{22}(z) &= z^{-n} + \mathcal{O}(z^{-n-1}). \end{aligned} \tag{5.33}$$

For the entry Y^{11}, the jump condition tells us that it should be an entire function and the asymptotic condition that it should go as a polynomial of degree n at ∞. By the generalized Liouville theorem[1], we can then conclude that $Y^{11}(z) = Q_n(z)$ a monic polynomial of degree n.

For the entry Y^{21} we have a similar result, since the jump condition tells us again that it should be an entire function and the asymptotic condition that it should behave as a polynomial of degree at most $n-1$ at ∞. Thus we conclude that $Y^{21}(z) = Q_{n-1}(z)$ a polynomial of degree at most $n-1$.

To show that these two polynomials Q_n, Q_{n-1} should be exactly the ones in the statement we need to use the asymptotic condition on the second column together with the jump condition.

[1](Generalized) Liouville theorem. If f is an holomorphic function $|f(z)| \leq M|z|^n$ for $|z| \to \infty$, then f is a polynomial of degree at most n.

Starting from Y^{12}, we see that its jump condition is satisfied by taking

$$Y^{12}(z) = \mathcal{C}\left(y^{-n}Q_n(y)w(y)\right)(z) \tag{5.34}$$

thanks to the Cauchy transform property (5.27). Now, we see that the required asymptotic condition for Y^{12} implies the specific choice of $Q_n(z) = \pi_n(z)$ the n-th monic orthogonal polynomial. Indeed

$$\begin{aligned} Y^{12}(z) = \mathcal{C}\left(y^{-n}Q_n(y)w(y)\right)(z) &= \frac{1}{2\pi \mathrm{i}}\int_{S^1}\frac{y^{-n}Q_n(y)w(y)}{y-z}dy \\ \sim_{z\to\infty} &-\frac{1}{2\pi \mathrm{i}}\sum_{k\geq 0}\frac{1}{z^{k+1}}\int_{S^1} y^{-n+k}Q_n(y)w(y)dy. \end{aligned} \tag{5.35}$$

For $Y^{12}(z) = \mathcal{O}(z^{-n-1})$ to hold, we then need to impose that

$$\frac{1}{2\pi \mathrm{i}}\int_{S^1} y^{-n+k}Q_n(y)w(y)dy = 0, \quad k = 0, \ldots, n-1 \tag{5.36}$$

or (by the change of variable $y = \mathrm{e}^{\mathrm{i}\theta}$)

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} Q_n(\mathrm{e}^{\mathrm{i}\theta})\mathrm{e}^{-\mathrm{i}(n-k-1)\theta}w(\mathrm{e}^{\mathrm{i}\theta})\mathrm{d}\theta = 0 \;\; k = 0, \ldots, n-1. \tag{5.37}$$

But this is exactly the orthogonality condition for the family defined with respect to the measure $d\mu(\theta)$, thus since we know that Q_n is also monic, we conclude that the only possibility is

$$Q_n(z) = \pi_n(z) \implies Y^{11}(z) = \pi_n(z), \; Y^{12}(z) = \mathcal{C}\left(y^{-n}\pi_n(y)w(y)\right)(z) \tag{5.38}$$

Finally, we look at the jump equation for the entry $Y^{22}(z)$, and we realize that again by defining

$$Y^{22}(z) = \mathcal{C}\left(y^{-n}Q_{n-1}(y)w(y)\right) \tag{5.39}$$

thanks to (5.27), the jump condition is satisfied. For the asymptotic condition to hold, given such we see that this translate again on a system of condition of Q_{n-1}. Using the same trick as before

$$\begin{aligned} Y^{22}(z) = \mathcal{C}\left(y^{-n}Q_{n-1}(y)w(y)\right) &= \frac{1}{2\pi \mathrm{i}}\int_{S^1}\frac{y^{-n}Q_{n-1}(y)w(y)}{y-z}dy \\ \sim_{z\to\infty} &-\frac{1}{2\pi \mathrm{i}}\sum_{k\geq 0}\frac{1}{z^{k+1}}\int_{S^1} y^{-n+k}Q_{n-1}(y)w(y)dy. \end{aligned} \tag{5.40}$$

This time, in order to have the asymptotic condition $Y^{22}(z) = z^{-n} + \mathcal{O}(z^{-n-1})$ to hold we need to impose

$$\begin{aligned} &-\frac{1}{2\pi \mathrm{i}}\int_{S^1} y^{-n+k}Q_{n-1}(y)w(y)dy = 0, \quad k = 0, \ldots, n-2 \\ &-\frac{1}{2\pi \mathrm{i}}\int_{S^1} y^{-1}Q_{n-1}(y)w(y)dy = 1. \end{aligned} \tag{5.41}$$

After change of variables ($y = \mathrm{e}^{\mathrm{i}\theta}$) the set of conditions corresponds to

$$\begin{aligned} &-\frac{1}{2\pi}\int_{-\pi}^{\pi} \mathrm{e}^{-\mathrm{i}(n-k-1)\theta}Q_{n-1}(\mathrm{e}^{\mathrm{i}\theta})w(\mathrm{e}^{\mathrm{i}\theta})d\theta = 0, \quad k = 0, \ldots, n-2 \\ &-\frac{1}{2\pi}\int_{-\pi}^{\pi} Q_{n-1}(\mathrm{e}^{\mathrm{i}\theta})w(\mathrm{e}^{\mathrm{i}\theta})d\theta = 1 \end{aligned} \tag{5.42}$$

and this is exactly the set of conditions satisfied by taking $Q_{n-1}(z) = -\kappa_{n-1}^2\pi_{n-1}^*(z)$.

Indeed for $k = 0, \dots, n-2$ we have

$$\begin{aligned}
&\frac{\kappa_{n-1}^2}{2\pi}\int_{-\pi}^{\pi} \mathrm{e}^{-\mathrm{i}(n-k-1)\theta}\pi_{n-1}^*(\mathrm{e}^{\mathrm{i}\theta})w(\mathrm{e}^{\mathrm{i}\theta})d\theta = \frac{\kappa_{n-1}^2}{2\pi}\int_{-\pi}^{\pi} \mathrm{e}^{-\mathrm{i}(n-k-1)\theta}(\mathrm{e}^{\mathrm{i}(n-1)\theta})\overline{\pi_{n-1}}(\mathrm{e}^{-\mathrm{i}\theta})w(\mathrm{e}^{\mathrm{i}\theta})d\theta \\
&= \frac{\kappa_{n-1}^2}{2\pi}\int_{-\pi}^{\pi} \mathrm{e}^{\mathrm{i}k\theta}\overline{\pi_{n-1}}(\mathrm{e}^{-\mathrm{i}\theta})w(\mathrm{e}^{\mathrm{i}\theta})d\theta = \frac{\kappa_{n-1}^2}{2\pi}\int_{-\pi}^{\pi} \overline{\mathrm{e}^{-\mathrm{i}k\theta}\pi_{n-1}(\mathrm{e}^{\mathrm{i}\theta})w(\mathrm{e}^{\mathrm{i}\theta})}d\theta \\
&= \frac{\kappa_{n-1}^2}{2\pi}\overline{\int_{-\pi}^{\pi} \mathrm{e}^{-\mathrm{i}k\theta}\pi_{n-1}(\mathrm{e}^{\mathrm{i}\theta})w(\mathrm{e}^{\mathrm{i}\theta})d\theta} = 0
\end{aligned} \tag{5.43}$$

by using the orthogonality condition satisfied by $\pi_{n-1}(z)$. And also

$$\begin{aligned}
&\frac{\kappa_{n-1}^2}{2\pi}\int_{-\pi}^{\pi} \pi_{n-1}^*(\mathrm{e}^{\mathrm{i}\theta})w(\mathrm{e}^{\mathrm{i}\theta})d\theta = \frac{\kappa_{n-1}^2}{2\pi}\int_{-\pi}^{\pi} \mathrm{e}^{\mathrm{i}(n-1)\theta}\overline{\pi_{n-1}}(\mathrm{e}^{-\mathrm{i}\theta})w(\mathrm{e}^{\mathrm{i}\theta})d\theta \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} p_{n-1}(\mathrm{e}^{\mathrm{i}\theta})\overline{p_{n-1}(\mathrm{e}^{\mathrm{i}\theta})}w(\mathrm{e}^{\mathrm{i}\theta})d\theta = 1
\end{aligned} \tag{5.44}$$

by using the orthonormality condition for p_{n-1}. Thus we conclude

$$Q_{n-1}(z) = -\kappa_{n-1}^2\pi_{n-1}^*(z) \implies Y^{21}(z) = -\kappa_{n-1}^2\pi_{n-1}^*(z),\ Y^{22}(z) = -\kappa_{n-1}^2\mathcal{C}(y^{-n}\pi_{n-1}^*(y)w(y)) \tag{5.45}$$

and this concludes the proof of the formula for the solution $Y(z)$.

Notice that the $\det(Y(z))$ is an entire function, indeed

$$\det(Y_+(z)) = \det(Y_-(z))\underbrace{\det(J_Y(z))}_{=1} = \det(Y_-(z)). \tag{5.46}$$

Moreover for $z \to \infty$ $\det(Y(z)) \sim 1$, thus by Liouville theorem necessarily $\det(Y(z)) = 1$. In particular $Y(z)$ is invertible. Now if there exist another solution $\tilde{Y}(z)$ of the same RH problem, we take $T(z) = \tilde{Y}(z)Y^{-1}(z)$ and this is also an entire function since $T_+(z) = T_-(z)$. And for $z \to \infty$ then $T(z) \sim I$, thus by Liouville theorem again $T(z) = I$ and so $\tilde{Y}(z) = Y(z)$. □

The solution $Y(z)$ has a symmetry which will be very useful in the following section.

Corollary 5.13 ([4],Proposition 5.12). *The unique solution $Y(z)$ of the Riemann-Hilbert problem 1 is such that*

$$Y(z) = \sigma_3 Y(0)^{-1}Y(z^{-1})z^{n\sigma_3}\sigma_3, \tag{5.47}$$

$$Y(z) = \overline{Y(\bar{z})}. \tag{5.48}$$

Proof. Both properties follow from the fact that the solution of Riemann-Hilbert problem 1 is unique.

For the first equation, we define the following matrix-valued function

$$H(z) = Y(z^{-1})\begin{pmatrix} z^n & 0 \\ 0 & z^{-n}\end{pmatrix}\sigma_3, \quad z \in \mathbb{C}\setminus(\Sigma\cup\{0\}) \tag{5.49}$$

and we show that it satisfies the same jump condition along Σ (5.23) of $Y(z)$, using that

$$\sigma_3^{-1}\begin{pmatrix} z^{-n} & 0 \\ 0 & z^n\end{pmatrix} J_Y(z^{-1})^{-1}\begin{pmatrix} z^n & 0 \\ 0 & z^{-n}\end{pmatrix}\sigma_3 = J_Y(z). \tag{5.50}$$

We can compute

$$
\begin{aligned}
H_+(z) &= \lim_{\substack{z'\to z\\ z' \text{ inside } S^1}} H(z') = \lim_{\substack{z'\to z\\ z' \text{ inside } S^1}} Y((z')^{-1}) \begin{pmatrix} (z')^n & 0 \\ 0 & (z')^{-n} \end{pmatrix} \sigma_3 \\
&= \underbrace{\lim_{\substack{(z')^{-1}\to z^{-1}\\ (z')^{-1} \text{ outside } S^1}} Y((z')^{-1})}_{=Y_-(z^{-1})} \begin{pmatrix} z^n & 0 \\ 0 & z^{-n} \end{pmatrix} \sigma_3 = \underbrace{\lim_{\substack{(z')^{-1}\to z^{-1}\\ (z')^{-1} \text{ inside } S^1}} Y((z')^{-1})}_{Y_+(z^{-1})} J_Y(z^{-1})^{-1} \begin{pmatrix} z^n & 0 \\ 0 & z^{-n} \end{pmatrix} \sigma_3 \\
&= \underbrace{\lim_{\substack{z'\to z\\ z' \text{ outside } S^1}} Y((z')^{-1}) \begin{pmatrix} z^n & 0 \\ 0 & z^{-n} \end{pmatrix} \sigma_3}_{=H_-(z)} \underbrace{\sigma_3^{-1} \begin{pmatrix} z^{-n} & 0 \\ 0 & z^{n} \end{pmatrix} J_Y(z^{-1})^{-1} \begin{pmatrix} z^n & 0 \\ 0 & z^{-n} \end{pmatrix} \sigma_3}_{=J_Y(z)} \\
&= H_-(z) J_Y(z).
\end{aligned} \tag{5.51}
$$

Notice that $H(z)$ has a removable singularity in $z = 0$, in particular $H(z) \to \sigma_3$. Moreover for $z \to \infty$ we have

$$
H(z) \begin{pmatrix} z^{-n} & 0 \\ 0 & z^{n} \end{pmatrix} \to Y(0)\sigma_3. \tag{5.52}
$$

Thus by taking

$$
L(z) = \sigma_3 Y(0)^{-1} H(z) \tag{5.53}
$$

we see that L satisfies the same Riemann-Hilbert problem 1 as Y and so by uniqueness of the solution we obtain the first identity.

For the second identity a similar proof hold, but defining $H(z) = \overline{Y(\bar{z})}$. We leave it as an exercise. □

Notice that the factor $Y(0) = Y(n;0)$ appearing in equation (5.47) has a very explicit form, by equation (5.29). This will be useful in the following sections.

Lemma 5.14 ([4], Corollary 5.13)**.** *For every $n \geq 1$ we have*

$$
Y(0) = Y(n;0) = \begin{pmatrix} x_n & \kappa_n^{-2} \\ -\kappa_{n-1}^2 & x_n \end{pmatrix}, \tag{5.54}
$$

where we denoted with $x_n := \pi_n(0) \in \mathbb{R}$.

Proof. The first column of $Y(n;0)$ directly follows from the evaluation in $z = 0$ of $Y(n;z)$ as given in equation (5.29). Indeed $Y^{11}(n;0) = \pi_n(0)$ and $Y^{21}(n;0) = -\kappa_{n-1}^2 \pi_{n-1}^*(0)$ but we observe that

$$
\pi_{n-1}^*(0) = z^{n-1}\overline{\pi_{n-1}(\bar{z}^{-1})}|_{z=0} = z^{n-1}(z^{-(n-1)} + \cdots + \overline{\pi_{n-1}(0)})|_{z=0} = 1. \tag{5.55}
$$

Thus we conclude that $Y^{21}(n;0) = -\kappa_{n-1}^2$.

For what concerns the second column of $Y(n;0)$, we first find the $(2,2)$-entry. This is indeed easily deduced from the symmetry given in (5.47). In the limit for $z \to 0$ it gives

$$
Y(n;0) = \sigma_3 Y^{-1}(n;0)\sigma_3, \tag{5.56}
$$

thus $Y^{22}(n;0) = Y^{11}(n;0) = \pi_n(0)$. Finally for the entry $(1,2)$ of $Y(n;0)$, we compute it explicitly using the orthonormality property of the polynomials $p_m(z)$

$$
\begin{aligned}
Y^{12}(n;0) &= \frac{1}{2\pi i}\int_{S^1} \frac{\pi_n(s)s^{-n}w(s)}{s} ds = \int_{-\pi}^{\pi} \pi_n(e^{i\theta})\overline{e^{in\theta}} w(e^{i\theta})\frac{d\theta}{2\pi} = \frac{1}{\kappa_n^2}\int_{-\pi}^{\pi} p_n(e^{i\theta})\overline{p_n(e^{i\theta})} w(e^{i\theta})\frac{d\theta}{2\pi} \\
&= \frac{1}{\kappa_n^2}.
\end{aligned} \tag{5.57}
$$

Finally, the fact that x_n is real follows from the entry $(1,1)$ of equation (5.48) together with equation (5.29). □

6 Toeplitz determinants and the discrete Painlevé II equation

In this Section we first proof a recurrence relation for the Toeplitz determinants in terms of orthogonal polynomials, which holds for any measure $d\mu(\theta)$ over the conditions above, and then we see how to specialize it in terms of the discrete Painlevé II equation, for the Toeplitz determinants of interest, already introduced in Section 4 Theorem 4.8, with weight (or symbol)

$$w(z) = w(z;t) = \mathrm{e}^{tz+tz^{-1}}, \quad t > 0. \tag{6.1}$$

In this Section, we follow the last part of Chapter VII of [4].

Remark 6.1. *Notice that this weight function can be expressed equivalently, for $z = \mathrm{e}^{i\theta}$ as*

$$w(\mathrm{e}^{i\theta}) = \mathrm{e}^{2t\cos\theta} = \mathrm{e}^{R(\theta)} \tag{6.2}$$

$R(\theta) = 2t\cos(\theta)$ *being a real analytic periodic function on the unit circle.*

From now on, we will denote by $Y(z) = Y(z;n,t)$ the solution of the Riemann-Hilbert problem 1 for this particular measure defined by the weight (6.1), depending on the parameter t.

6.1 Recurrence relation for Toeplitz determinants

We start by using the previous Lemma to obtain a recursive relation for the Toeplitz determinants.

Corollary 6.2. *We have for any $n \geq 1$*

$$\frac{D_n D_{n-2}}{D_{n-1}^2} = 1 - x_n^2 \tag{6.3}$$

Proof. Equation (6.3) comes from the fact that $\det(Y(n;z)) = 1$ identically in z and so in particular for $z = 0$ by writing $Y(n;0)$ as in equation (5.54), we have

$$\frac{\kappa_{n-1}^2}{\kappa_n^2} = 1 - x_n^2. \tag{6.4}$$

Now just recall that $\kappa_n^2 = \frac{D_{n-1}}{D_n}$ and (6.3) follows. □

We stress that the quantities D_n, x_n, κ_n now all depend parametrically on the parameter t and the equation (6.3) holds identically in t, but we use a simplified notation.

Remark 6.3. *Notice that this recurrence relation holds true for any family of Toeplitz determinant with a symbol or weight that respects the conditions given at the beginning. In particular this relation allows one to compute the n-th Toeplitz determinants by using the two previous one, together with the computation of $x_n = \pi_n(0)$. In the next Subsection we will see that this x_n can also be computed recursively, and for the specific weight* (6.1) *this recursion is the discrete Painlevé II equation.*

6.2 The Lax pair for the discrete Painlevé II equation

From now on, we fix the measure $d\mu(\theta) = w(\mathrm{e}^{\mathrm{i}\theta})\frac{d\theta}{2\pi}$ for w in (6.1), parametrically depending on t. In this case, we are going to show that $x_n = \pi_n(0)$ satisfies the discrete nonlinear equation

$$x_{n+1} + x_{n-1} = -\frac{nx_n}{t(1-x_n^2)} \tag{6.5}$$

known nowadays as the discrete Painlevé II equation [17].

In order to achieve this result, we are going to construct a Lax pair for this equation, using the solution of the Riemann-Hilbert problem 1 for the specified measure $d\mu(\theta;t)$ taken with (6.1).

We define the following matrix-valued function

$$\Psi(z;n,t) := \begin{pmatrix}1 & 0\\ 0 & \kappa_n^{-2}\end{pmatrix} Y(z;n,t) \begin{pmatrix}1 & 0\\ 0 & z^n\end{pmatrix} \mathrm{e}^{\ell(z;t)\frac{\sigma_3}{2}}, \tag{6.6}$$

where $\ell(z;t) = tz + tz^{-1}$ (recall that now the weight in the measure is now $w(z) = \mathrm{e}^{tz+tz^{-1}} = \mathrm{e}^{\ell(z;t)}$).

Thanks to the properties of $Y(z;n,t)$ from the RH problem 1 one can prove that $\Psi(z;n,t)$ satisfies the following Riemann-Hilbert problem.

Riemann–Hilbert Problem 2. *The function* $\Psi(z) := \Psi(z;n,t) : \mathbb{C} \to \mathrm{GL}(2,\mathbb{C})$ *has the following properties:*

(1) $\Psi(z)$ *is analytic for every* $z \in \mathbb{C} \setminus \{S^1 \cup \{0\}\}$*;*

(2) $\Psi(z)$ *has continuous boundary values* $\Psi_\pm(z)$ *while approaching non-tangentially* S^1 *either from the left or from the right, and they are related for all* $z \in S^1$ *through*

$$\Psi_+(z) = \Psi_-(z) J_0, \quad J_0 = \begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix}; \tag{6.7}$$

(3) $\Psi(z)$ *has asymptotic behavior near* 0 *given by*

$$\Psi(z) \sim \begin{pmatrix}1 & 0\\ 0 & \kappa_n^{-2}\end{pmatrix} Y(0) \left(I + \sum_{j=1}^{\infty} z^j \widetilde{Y}_j(n)\right) \begin{pmatrix}1 & 0\\ 0 & z^n\end{pmatrix} \mathrm{e}^{\ell(z;t)\frac{\sigma_3}{2}}, \quad z \to 0; \tag{6.8}$$

(4) $\Psi(z)$ *has asymptotic behavior near* ∞ *given by*

$$\Psi(z) \sim \begin{pmatrix}1 & 0\\ 0 & \kappa_n^{-2}\end{pmatrix} \left(I + \sum_{j=1}^{\infty} \frac{Y_j(n)}{z^j}\right) \begin{pmatrix}z^n & 0\\ 0 & 1\end{pmatrix} \mathrm{e}^{\ell(z;t)\frac{\sigma_3}{2}}, \quad |z| \to \infty. \tag{6.9}$$

Proposition 6.4. *The function* $\Psi(z;n,t)$ *defined in* (6.6) *solves the Riemann–Hilbert problem 2.*

Proof. The analyticity condition and the asymptotic expansions at $0, \infty$ given in (6.8), (6.9) follows directly from the definition (6.6) and the fact that $Y(z)$ solves the RH problem 1. Condition (6.7) follows from direct computation

$$\Psi_+(z) = \begin{pmatrix}1 & 0\\ 0 & \kappa_n^{-2}\end{pmatrix} Y_+(z) \begin{pmatrix}1 & 0\\ 0 & z^n\end{pmatrix} \mathrm{e}^{\ell(z;t)\frac{\sigma_3}{2}} = \begin{pmatrix}1 & 0\\ 0 & \kappa_n^{-2}\end{pmatrix} Y_-(z) J_Y(z) \begin{pmatrix}1 & 0\\ 0 & z^n\end{pmatrix} \mathrm{e}^{\ell(z;t)\frac{\sigma_3}{2}} \tag{6.10}$$

$$= \Psi_-(z) \mathrm{e}^{-\ell(z;t)\frac{\sigma_3}{2}} \begin{pmatrix}1 & 0\\ 0 & z^{-n}\end{pmatrix} \begin{pmatrix}1 & z^{-n} e^{\ell(z;t)}\\ 0 & 1\end{pmatrix} \begin{pmatrix}1 & 0\\ 0 & z^n\end{pmatrix} \mathrm{e}^{\ell(z;t)\frac{\sigma_3}{2}} \tag{6.11}$$

$$= \Psi_-(z) \begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix}. \tag{6.12}$$

□

From the solution of the Riemann-Hilbert problem 2 we deduce a linear system that will be the Lax pair for the discrete Painlevé II equation (in the following we omit in Ψ the dependence on t that should be considered only as parameters and not actual variables like n, z).

Proposition 6.5. *We have*

$$\begin{aligned} \Psi(n+1;z) &= U(n;z)\Psi(n;z),\\ \partial_z \Psi(n;z) &= T(n;z)\Psi(n;z), \end{aligned} \tag{6.13}$$

with

$$U(n;z) := \begin{pmatrix} z + x_n x_{n+1} & -x_{n+1}\\ -(1 - x_{n+1}^2)x_n & 1 - x_{n+1}^2 \end{pmatrix} = \sigma_+ z + U_0(n), \tag{6.14}$$

where $\sigma_+ := \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ *and*

$$T(n;z) := T_0(n) + T_{-1}(n)z^{-1} + T_{-2}(n)z^{-2} \tag{6.15}$$

where

$$T_0(n) = \frac{t}{2}\sigma_3, \;\; T_{-1}(n) = \begin{pmatrix} n & -tx_{n+1}\kappa_n^{-2} \\ -t\kappa_{n-1}^2 x_{n-1} & 0 \end{pmatrix}, \;\; T_{-2}(n) = t\begin{pmatrix} \frac{1}{2} - x_n^2 & x_n \\ x_n(1-x_n)^2 & x_n^2 - \frac{1}{2} \end{pmatrix}. \tag{6.16}$$

Proof. We first prove the first equation. We start by defining the quantity $U(n;z) := \Psi(n+1;z)\Psi^{-1}(n;z)$. Since the jump condition for $\Psi(z)$ (6.7) is independent of n, $U(n;z)$ is analytic everywhere. Plugging in equation (6.9) we have the expansion at ∞

$$U(n;z) = \begin{pmatrix} 1 & 0 \\ 0 & \kappa_{n+1}^{-2} \end{pmatrix}\left(I + \frac{Y_1(n+1)}{z} + \mathcal{O}(z^{-2})\right)\underbrace{z^{(n+1)\sigma_3}\begin{pmatrix} 1 & 0 \\ 0 & z \end{pmatrix}z^{-n\sigma_3}}_{=\begin{pmatrix} z & 0 \\ 0 & 1 \end{pmatrix}}\left(I - \frac{Y_1(n)}{z} + \mathcal{O}(z^{-2})\right)\begin{pmatrix} 1 & 0 \\ 0 & \kappa_n^2 \end{pmatrix},$$

from which we deduce that $U(n;z)$ is a polynomial in z of degree 1, by Liouville theorem. Moreover its matrix-valued coefficient are written as

$$U(n;z) = z\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \underbrace{\begin{pmatrix} 1 & 0 \\ 0 & \kappa_{n+1}^{-2} \end{pmatrix}Y(n+1;0)\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}Y^{-1}(n;0)\begin{pmatrix} 1 & 0 \\ 0 & \kappa_n^2 \end{pmatrix}}_{=U_0(n)}, \tag{6.17}$$

where we computed the constant term $U_0(n)$ just by looking at the value at $z = 0$. Doing the computation and using equation (5.54) we obtain

$$U_0(n) = \begin{pmatrix} Y^{11}(n+1;0)Y^{22}(n;0) & -\kappa_n^2 Y^{11}(n+1;0)Y^{12}(n;0) \\ \kappa_{n+1}^{-2}Y^{21}(n+1;0)Y^{22}(n,0) & -Y^{21}(n+1;0)Y^{12}(n;0) \end{pmatrix} = \begin{pmatrix} x_{n+1}x_n & -x_{n+1} \\ -(1-x_{n+1}^2)x_n & 1 - x_{n+1}^2 \end{pmatrix}.$$

Notice that from this proof, we also obtained the following explicit formula for the asymptotic coefficient $Y_1(n)$

$$\begin{aligned} Y_1^{11}(n+1) - Y_1^{11}(n+1) &= x_n x_{n+1} \\ Y_1^{21}(n+1) &= -\kappa_n^2 x_n \\ Y_1^{12}(n) &= \frac{x_{n+1}}{\kappa_n^2} \end{aligned} \tag{6.18}$$

by comparing $U_0(n)$ with the constant term in the asymptotic expansion of $U(n,z)$ at ∞. This will be useful later in the second part of the proof.

For the second equation, we start by defining $T(z;n) = \partial_z\Psi(z;n)(\Psi(z;n))^{-1}$. Notice that the jump of Ψ does not depend on z, so T has no jump along S^1. Is is then a meromorphic function. Its behaviour at ∞, given (6.9), is like

$$\begin{aligned} T(z;n) &\sim \partial_z\Psi(z;n)(\Psi(z;n))^{-1} \\ &\sim \left(-\frac{Y_1(n)}{z^2} + \dots\right)\left(1 - \frac{Y_1(n)}{z} + \dots\right) + \left(1 + \frac{Y_1(n)}{z} + \dots\right)\begin{pmatrix} nz^{-1} & 0 \\ 0 & 0 \end{pmatrix}\left(1 - \frac{Y_1(n)}{z} + \dots\right) \\ &+ \underbrace{\frac{\ell'(z;t)}{2}}_{=\frac{t}{2} - \frac{t}{2z^2}}\left(1 + \frac{Y_1(n)}{z} + \dots\right)\sigma_3\left(1 - \frac{Y_1(n)}{z} + \dots\right) \end{aligned} \tag{6.19}$$

in particular the leading term is $\frac{t}{2}\sigma_3$. Its behaviour at 0 is instead dictated by (6.8)

$$
\begin{aligned}
T(z;n) \sim &\begin{pmatrix}1 & 0\\ 0 & \kappa_n^{-2}\end{pmatrix} Y(0)\left(\widetilde{Y}_1(n)+\dots\right)\left(I - z\widetilde{Y}_1(n)+\dots\right)Y(0)^{-1}\begin{pmatrix}1 & 0\\ 0 & \kappa_n^{2}\end{pmatrix}\\
&+\begin{pmatrix}1 & 0\\ 0 & \kappa_n^{-2}\end{pmatrix} Y(0)\left(I+z\widetilde{Y}_1(n)+\dots\right)\begin{pmatrix}0 & 0\\ 0 & nz^{-1}\end{pmatrix}\left(I - z\widetilde{Y}_1(n)+\dots\right)Y(0)^{-1}\begin{pmatrix}1 & 0\\ 0 & \kappa_n^{2}\end{pmatrix}\\
&+\begin{pmatrix}1 & 0\\ 0 & \kappa_n^{-2}\end{pmatrix} Y(0)\left(I+z\widetilde{Y}_1(n)+\dots\right)\underbrace{\frac{\ell'(z;t)}{2}}_{=\frac{t}{2}-\frac{t}{2z^2}}\sigma_3\left(I - z\widetilde{Y}_1(n)+\dots\right)Y(0)^{-1}\begin{pmatrix}1 & 0\\ 0 & \kappa_n^{2}\end{pmatrix}
\end{aligned}
\tag{6.20}
$$

from which we deduce that $U(z;n)$ has a pole of order 2 in $z=0$, with leading coefficient

$$
-\frac{t}{2}\begin{pmatrix}1 & 0\\ 0 & \kappa_n^{-2}\end{pmatrix} Y(0)\sigma_3 Y(0)^{-1}\begin{pmatrix}1 & 0\\ 0 & \kappa_n^{2}\end{pmatrix} = t\begin{pmatrix}\frac{1}{2}-x_n^2 & x_n\\ x_n(1-x_n)^2 & x_n^2-\frac{1}{2}\end{pmatrix}
\tag{6.21}
$$

where we used (6.4).

So we conclude that $T(z;n)$ is a meromorphic function with only one pole in 0 of order 2. In particular, we can write it down explicitly as

$$
T(z;n) = \frac{t}{2}\sigma_3 + \frac{1}{z}\begin{pmatrix} n & -tx_{n+1}\kappa_n^{-2}\\ -t\kappa_{n-1}^2 x_{n-1} & 0\end{pmatrix} + \frac{1}{z^2}t\begin{pmatrix}\frac{1}{2}-x_n^2 & x_n\\ x_n(1-x_n)^2 & x_n^2-\frac{1}{2}\end{pmatrix}.
\tag{6.22}
$$

Notice that the coefficient of z^{-1} is derived by e.g. looking at the coefficient in the expansion at ∞, given by

$$
\begin{pmatrix} n & 0\\ 0 & 0\end{pmatrix} + \frac{t}{2}\underbrace{\left(Y_1(n)\sigma_3 - \sigma_3 Y_1(n)\right)}_{=\begin{pmatrix}0 & -2Y_1^{12}(n)\\ 2Y_1^{21}(n) & 0\end{pmatrix}}
\tag{6.23}
$$

that is explicitly computed by equations (6.18). □

Now the system (6.13) gives a Lax pair for the discrete Painlevé II equation, in the sense that by looking at its compatibility condition, namely the fact that

$$
\begin{aligned}
\partial_z(\Psi(z;n+1)) &= T(z;n+1)\Psi(z;n+1) = T(z;n+1)U(z;n)\Psi(z;n)\\
\partial_z(\Psi(z;n+1)) &= \partial_z\left(U(z;n)\Psi(z;n)\right) = \left(\partial_z(U(z;n)) + U(z;n)T(z;n)\right)\Psi(z;n)
\end{aligned}
\tag{6.24}
$$

is equivalent to the following matrix equation for $U(z;n), T(z;n)$ that should hold for every z, n, t

$$
\partial_z U(z;n) = T(z;n+1)U(z;n) - U(z;n)T(z;n).
\tag{6.25}
$$

The only nontrivial scalar equation left from this matrix equation is

$$
t(1-x_n^2)\kappa_n^2 x_{n+1} + nx_n\kappa_n^2 + tx_{n-1}\kappa_{n-1}^2 = 0
\tag{6.26}
$$

that after dividing by κ_{n-1}^2 coincide exactly with (6.5).

We have finally proved the following result.

Theorem 6.6 ([2, 6, 1]). *The Toeplitz determinants D_n w.r.t. the symbol $w(\mathrm{e}^{\mathrm{i}\theta})$ in (6.1) on S^1, satisfy for every $n \geq 1$ the following recursion*

$$
\frac{D_n D_{n-2}}{D_{n-1}^2} = 1 - x_n^2
\tag{6.27}
$$

where x_n solves the discrete Painlevé II equation

$$
x_{n+1} + x_{n-1} = -\frac{nx_n}{t(1-x_n^2)}
\tag{6.28}
$$

with initial condition $x_0 = 1, x_1 = \pi_1(0,t)$.

Remark 6.7. *Different proofs, here we followed [2] (also explained in [4]), but another one using discrete Riemann-Hilbert problems associated to discrete integrable operators was given in [6] and yet another one, based on unitary matrix integrals was found in [1] independently and more or less at the same time.*

Remark 6.8. *Theorem 6.6 has been recently generalized for Toeplitz determinants associated with more general weights*

$$e^{\sum_{k=1}^{N}\frac{t_k}{k}(z^k+z^{-k})}$$

related to some specific instances of Okounkov Schur measures [23] on random partitions (generalization of the Poissonized Plancherel measure) in the work [8] (and references therein for the first values $N = 2, 3$). In this generalization the discrete Painlevé II equation is replaced by a $2N$ order nonlinear discrete equation, corresponding to the N-th equation of the discrete Painlevé II hierarchy.

7 Convergence to the Tracy-Widom GUE distribution

In this section we will see some heuristics of the Baik-Deift-Johansson theorem taking continuous limit of the discrete Painlevé II equation, and finally we will conclude with some comments on the rigorous proof [3].

7.1 Heuristics of BDJ theorem

In this paragraph we show an heuristic argument of the Baik Deift Johansson Theorem using the Tracy-Widom representation of $F_{GUE}(s)$ introduced in Theorem 1.5 and by taking a continuous limit of Theorem 6.6. We follow here the explanation given in Paragraph 7.4 of [4], another similar explanation is given in [6]. In particular, we will see first that under a certain scaling limit the discrete Painlevé II equation gives the Painlevé II equation, then that the recursion (6.3) gives the Tracy-Widom formula indeed.

Let us consider the following new function

$$q(t,s) := (-1)^{\lfloor 2t+st^{\frac{1}{3}}\rfloor} t^{\frac{1}{3}} x_{\lfloor 2t+st^{\frac{1}{3}}\rfloor} \tag{7.1}$$

where $\lfloor\cdot\rfloor$ denotes the integer part and t is the same parameter appearing in (6.5) and $x_\cdot$ is the discrete function in the same equation, $s \in \mathbb{R}$. Equivalently

$$x_{\lfloor 2t+st^{\frac{1}{3}}\rfloor} = q(t,s)(-1)^{-\lfloor 2t+st^{\frac{1}{3}}\rfloor} t^{-\frac{1}{3}} \tag{7.2}$$

We notice that, thanks to equation (6.5) and considering $n = \lfloor 2t + st^{\frac{1}{3}}\rfloor$, we have

$$\begin{aligned}
\frac{\lfloor 2t+st^{\frac{1}{3}}\rfloor}{t}q(t,s) &= \frac{\lfloor 2t+st^{\frac{1}{3}}\rfloor}{t}(-1)^{\lfloor 2t+st^{\frac{1}{3}}\rfloor} t^{\frac{1}{3}} x_{\lfloor 2t+st^{\frac{1}{3}}\rfloor}\\
&= -(-1)^{\lfloor 2t+st^{\frac{1}{3}}\rfloor} t^{\frac{1}{3}}\left(x_{\lfloor 2t+st^{\frac{1}{3}}+1\rfloor} + x_{\lfloor 2t+st^{\frac{1}{3}}-1\rfloor}\right)\left(1-x^2_{\lfloor 2t+st^{\frac{1}{3}}\rfloor}\right)\\
&= (-1)^{\lfloor 2t+st^{\frac{1}{3}}+1\rfloor} t^{\frac{1}{3}} q(t,s+t^{-\frac{1}{3}})(-1)^{-\lfloor 2t+st^{\frac{1}{3}}+1\rfloor} t^{-\frac{1}{3}}(1-t^{-\frac{2}{3}}q(t,s)^2)\\
&+ (-1)^{\lfloor 2t+st^{\frac{1}{3}}-1\rfloor} t^{\frac{1}{3}} q(t,s-t^{-\frac{1}{3}})(-1)^{-\lfloor 2t+st^{\frac{1}{3}}-1\rfloor} t^{-\frac{1}{3}}(1-t^{-\frac{2}{3}}q(t,s)^2)\\
&= \left(q(t,s+t^{\frac{1}{3}}) + q(t,s-t^{-\frac{1}{3}})\right)(1-t^{-\frac{2}{3}}q(t,s)^2)
\end{aligned} \tag{7.3}$$

where we used that

$$\begin{aligned}
q(t,s\pm t^{-\frac{1}{3}}) &= (-1)^{\lfloor 2t+(s\pm t^{-\frac{1}{3}})t^{\frac{1}{3}}\rfloor} t^{-\frac{1}{3}} x_{\lfloor 2t+(s\pm t^{-\frac{1}{3}})t^{\frac{1}{3}}\rfloor}\\
&= (-1)^{\lfloor 2t+st^{\frac{1}{3}}\pm 1\rfloor} t^{\frac{1}{3}} x_{\lfloor 2t+st^{\frac{1}{3}}\pm 1\rfloor}.
\end{aligned} \tag{7.4}$$

Then supposing $q(t,s)$ is smooth in s, we have its Taylor expansion

$$q(t, s \pm t^{-\frac{1}{3}}) = q(t,s) \pm q'(t,s)t^{-\frac{1}{3}} + q''(t,s)t^{-\frac{2}{3}} + O(t^{-1}) \tag{7.5}$$

(where the $'$ denotes the derivative with respect to s), that allows us to rewrite the previous computation as follows

$$\frac{\lfloor 2t + st^{\frac{1}{3}} \rfloor}{t} q(t,s) = \left(2q(t,s) + q''(t,s)t^{-\frac{2}{3}} + O(t^{-1})\right)\left(1 - t^{-\frac{2}{3}} q^2(t,s)\right) \tag{7.6}$$

and in particular, we obtain

$$q''(t,s) = 2q^3(t,s) + sq(t,s) + O(t^{-1}). \tag{7.7}$$

Taking $t \to +\infty$, supposing the limit $\lim_{t\to+\infty} q(t,s) = q(s)$ exists, we get from the previous equation that the limiting function $q(s)$ solves the Painlevé II equation in s

$$q''(s) = 2q^3(s) + sq(s). \tag{7.8}$$

Now we consider the continuous limit of the equation (6.3) describing the relation between the Toeplitz determinants D_n and the solution of the Painlevé II equation. We define

$$Q_n(t) = \log\left(\mathbb{P}\{L(t) \leq n\}\right) = \log\left(D_{n-1}\right) - t^2 \tag{7.9}$$

thanks to the Gessel formula, Theorem 4.8. Then the recursion relation (6.3) reads as

$$Q_{n-1}(t) - 2Q_n(t) + Q_{n+1}(t) = \log\left(1 - x_n^2\right). \tag{7.10}$$

Again, for given $s \in \mathbb{R}$ we consider now the continuous limit of the above equation where simultaneously both $n, t \to +\infty$ with $n = \lfloor 2t + st^{\frac{1}{3}} \rfloor$ the same scaling we used before for the discrete Painlevé II equation. Notice that we have

$$Q_{n\pm1}(t) = Q_{\lfloor 2t + st^{\frac{1}{3}} \pm 1 \rfloor}(t) = Q_{\lfloor 2t + t^{\frac{1}{3}}(s \pm t^{-\frac{1}{3}}) \rfloor}(t) \tag{7.11}$$

so supposing that the limit exist for $Q_n(t)$ in this scaling

$$\lim_{t\to+\infty} Q_{\lfloor 2t + st^{\frac{1}{3}} \rfloor}(t) = Q(s) \tag{7.12}$$

locally uniformly in s and for $Q(s)$ being a smooth function of s, the equation (7.10), together with assumption on the limiting behavior of x_n, becomes

$$Q(s - t^{-\frac{1}{3}}) - 2Q(s) + Q(s + t^{-\frac{1}{3}}) \sim \log\left(1 - \frac{q^2(s)}{t^{\frac{2}{3}}}\right). \tag{7.13}$$

Now, in the left hand side we recognize, up to a factor $t^{-\frac{2}{3}}$ the second derivative of Q in s in the limit $t \to \infty$ while in the right hand side just by expanding the logarithm near 1

$$\frac{Q''(s)}{t^{\frac{2}{3}}} + O(t^{-1}) = -\frac{q^2(s)}{t^{\frac{2}{3}}} + O(t^{-\frac{4}{3}}) \tag{7.14}$$

which gives at first order exactly

$$Q''(s) = -q^2(s). \tag{7.15}$$

Integrating this relation, supposing that $q(s)$ goes to zero fast enough for $s \to \infty$ we get

$$Q'(y) = \int_y^{+\infty} q^2(z)dz. \tag{7.16}$$

Then integrating a second time

$$
\begin{aligned}
Q(s) &= -\int_s^{+\infty}\int_y^{+\infty} q^2(z)dzdy \\
&= -\left(y\int_y^{+\infty} q^2(z)dz\Big|_{y=s}^{y=+\infty} - \int_s^{+\infty} y\left(-q^2(y)\right)dy\right) \\
&= s\int_s^{+\infty} q^2(z)dz - \int_s^{+\infty} yq^2(y)dy \\
&= -\int_s^{+\infty} (y-s)q^2(y)dy
\end{aligned}
\tag{7.17}
$$

where we first integrated by parts and then we used the assumption of *suffiently fast decay* to zero of the function q at infinity. Notice that the function q solves the Painlevé II equation (7.8).

The only thing that is left to do, in order to recognize the GUE Tracy-Widom distribution in the right hand side, as defined in the Introduction (1.7), is to show that the solution q is indeed the Hastings-McLeod solution of the Painlevé II equation, in other words we need to show $q(s) \sim \mathrm{Ai}(s), s \to \infty$. In order to see this, another final heuristic argument can be used. First, we notice that in the discrete Painlevé II equation (6.5), since the term x_n goes to zero for large n, the nonlinear part is negligible compared to the others. In this way, the remaining linear equation is exactly the discrete equation satisfied by the Bessel functions of first kind, that for integer parameter n are defined as

$$J_n(x) = \frac{1}{\pi}\int_0^{\pi} \cos(nt - x\sin t)dt \tag{7.18}$$

that can also be extended to real values of the parameter n. In particular we can identify, for $n \to \infty$

$$x_n \sim J_{-n}(2t) = (-1)^n J_n(2t). \tag{7.19}$$

Now, taking also $t \to \infty$ in the usual scaling limit where $n = 2t + st^{\frac{1}{3}}$, for any s, one can show that

$$t^{\frac{1}{3}} J_{2t+st^{\frac{1}{3}}}(2t) \sim \mathrm{Ai}(s). \tag{7.20}$$

see Lemma 4.4 in [7] for a detailed proof. All together we have, for all $s \in \mathbb{R}$

$$q(s) = \lim_{t\to\infty} q(t,s) = \lim_{t\to\infty} q(t,s) := (-1)^{\lfloor 2t+st^{\frac{1}{3}}\rfloor} t^{\frac{1}{3}} x_{\lfloor 2t+st^{\frac{1}{3}}\rfloor} \sim \lim_{t\to\infty} t^{\frac{1}{3}} J_{2t+st^{\frac{1}{3}}}(2t) \sim \mathrm{Ai}(s). \tag{7.21}$$

7.2 Final remarks

Main ideas of the actual proof of [3] The actual proof of Theorem 1.5 is based on the nonlinear steepest descent method applied to the Riemann-Hilbert problem 1 for the OPUC with the measure with weight (6.1).

The starting point for this proof is the formula for $Q_n(t)$, defined in (7.9) that we recall below

$$Q_n(t) = \log\left(\mathbb{P}\{L(t) \le n\}\right) = \log\left(D_{n-1}\right) - t^2 \tag{7.22}$$

together with the relation between D_n and κ_n given in (6.4). Using the Szego strong limiting theorem applied to the Toeplitz determinants D_n, one can see

$$Q_n(t) = \sum_{k=n}^{\infty} \log \kappa_k^2. \tag{7.23}$$

In fact, the Szego Strong limiting theorem allows to describe the asymptotic behavior of large Toeplitz determinants in the following way.

Theorem 7.1 ([24]). *Let $w(z) = \mathrm{e}^{R(\theta)}$ for $z = \mathrm{e}^{i\theta}$ and R a real analytic 2π-periodic function on the unit circle. Denote the Fourier coefficients of R by*

$$R_k = \int_{-\pi}^{\pi} \mathrm{e}^{-ik\theta} R(\theta) \frac{d\theta}{2\pi}. \tag{7.24}$$

Then $w(z)$ can be analytically continued and $|w(z)| \geq Re(w(z)) > 0$ in the annulus $\rho < |z| < \rho^{-1}$ for a certain $\rho \in (0,1)$ and the Toeplitz determinant associated to w have the following behaviour

$$\log D_n(w) \sim (n+1)R_0 + \sum_{k\geq 1} kR_k R_{-k}, \quad \textit{as } n \to \infty. \tag{7.25}$$

The Toeplitz determinants for the weight (6.1) satisfy the hypothesis, with

$$R(\theta) = 2t\cos(\theta). \tag{7.26}$$

Thus the large n behaviour of these Toeplitz determinants can be deduced by applying the Szego strong limiting theorem, computing

$$R_0 = \frac{t}{\pi}\int_{-\pi}^{\pi} \cos(\theta) d\theta = 0 \tag{7.27}$$

noticing that $R_{-k} = R_k$ for all $k \in \mathbb{Z}$ and in particular

$$R_{\pm 1} = \frac{t}{\pi}\int_{-\pi}^{\pi} \cos(\theta)^2 d\theta = \frac{t}{\pi}\left(\frac{1}{2}\left(\theta + \sin\theta\cos\theta\right)\right)\Big|_{\theta=-\pi}^{\theta=\pi} = t \tag{7.28}$$

and for all $k \geq 2$

$$R_{\pm k} = \frac{t}{\pi}\int_{-\pi}^{\pi} \cos(k\theta)\cos(\theta) d\theta = 0. \tag{7.29}$$

Thus we conclude

$$\lim_{n\to\infty} \log D_n = t^2. \tag{7.30}$$

which leads to (7.23) by writing

$$Q_n(t) = -t^2 + \lim_{M\to\infty}\left(\sum_{k=n}^{M} \log \kappa_k^2 + D_M\right). \tag{7.31}$$

Equation (7.23) tells us that in order to describe the behaviour of Q_n is enough to study the behavior of the $\kappa_k, k \geq n$. In fact these values are determined in terms of the solution of the Riemann-Hilbert problem 1 for the weight $w(z;t)$ (6.1), that we denoted $Y = Y(k,t;z)$. In particular, from the explicit solution (5.54) one can see that for any k

$$\kappa_k^2 = -Y_{21}(k+1,t;z=0). \tag{7.32}$$

The problem of studying the large k,t behaviour (in the appropriate scaling limit) of the coefficients $\kappa_k(t)$ then is encoded by the problem of studying the same asymptotic behaviour of the solution of the Riemann-Hilbert problem $Y(k,t;z)$ itself, that is exactly what the nonlinear steepest descend method does.

The main idea of this method, is to find a series of transformations starting from the Riemann-Hilbert problem 1 and leading to a Riemann-Hilbert problem normalized at infinity by the identity matrix and with a jump matrix which is, in the required scaling limit, and uniformly in z, asymptotically equivalent to the identity matrix. This requires many technicalities, which are explained from Section 4 to 6 of the original paper [3] and led in the end to the wanted asymptotic for the coefficients κ_k. This lead to the description of the asymptotic behaviour of $Q_n(t)$ (see Lemma 7.7 [3]). Finally, using the de-Poissonization lemma, the original original proof of the result 1.5 is achieved (Section 8 [3]).

8 Exercises

Exercise 1. *Suppose you are given n points* $\{(X_k, Y_k\}$ *distributed uniformly on* D_N *and consider it as the graph associated to a random permutation* σ*, as in Fig. 3. Prove that such a permutation is uniformly distributed in* S_N*.*
Hint: Given $X_1, \ldots, X_n$ uniformly distributed on $[0, N]$, there exists a unique $\sigma \in S_N$ such that $X_{\pi(1)} < \cdots < X_{\pi(N)}$. Analogously, there exists a unique $\eta \in S_N$ such that $Y_{\eta(1)} < \cdots < Y_{\eta(N)}$. Start proving that σ and η are uniformly distributed. What is the relation between σ, η and π?.

Exercise 2. *Prove that for any* $n, N \geq 1$

$$\mathbb{P}(\ell_N \leq n) \geq \mathbb{P}(\ell_{N+1} \leq n).$$

Exercise 3. *Given* $RS(\pi) = (P, Q)$*, is the first row of* P *an increasing subsequence of* π *of maximal length? Show how an increasing subsequence of* π *of maximal length can be constructed from the RS algorithm.*

Exercise 4. *Prove the following formula for the n-th orthogonal polynomial on the unit circle for a given measure*

$$p_n(z) = \frac{1}{n!\sqrt{D_n D_{n-1}}} \int_{|z_0|=1} \cdots \int_{|z_{n-1}|=1} \prod_{j=0}^{n-1} (z - z_j) \prod_{0 \leq j < k \leq n-1} |z_j - z_k|^2 \prod_{j=0}^{n-1} w(z_j) \frac{dz_j}{2\pi i z_j}$$

(it follows from the analogue of the Heine formula for Hankel determinants in the case of Toeplitz determinants)

Exercise 5. *Prove the symmetry relation for the solution of the Riemann-Hilbert problem 1 given by equation* (5.48)*.*

Exercise 6. *Prove the second relation in the Szego recurrence relations, system* (5.13) *and recover these recurrence relations from the Lax pair in Proposition 6.5.*

References

[1] M. Adler, P. van Moerbeke. “Integrals over classical groups, random permutations, Toda and Toeplitz lattices”. *Comm. Pure Appl. Math.* 54 no. 2, (2001).

[2] J. Baik. “Riemann-Hilbert problems for last passage percolation”. *ArXiv* (2001).

[3] J. Baik, P. Deift & K. Johansson. “On the distribution of the length of the longest increasing subsequence of random permutations”. *Journal of the Am. Math. Soc.* 12, no. 4, 1119–1178 (1999).

[4] J. Baik, P. Deift and T. Suidan. “Combinatorics and Random Matrix Theory”. *American Mathematical Society*, (2016).

[5] E. L. Basor, H. Widom. “ On a Toeplitz determinant identity of Borodin and Okounkov.” *Integral Equations and Operator Theory*, 37(4), 397-401 (2000).

[6] A. Borodin. “Discrete gap probabilities and iscrete Painlevé equations”. *Duke Math. J.* 117(3), (2003).

[7] A. Borodin, A. Okunkov and G. Olshanski. “Asymptotics of Plancherel measures for symmetric groups”. *Journal of the American Mathematical Society*, 13 no. 3, (2000).

[8] T. Chouteau, S. Tarricone. “Recursion relation for Toeplitz determinants and the discrete Painlevé II hierarchy”. *SIGMA, Special Edition in honour of Alexander Its ’70th birthday*, (2023).

[9] K. F. Clancey, I. Gohberg. “Factorization of Matrix Functions and Singular Integral Operators”. *Birkhäuser Basel*, (2014).

[10] P. Deift. “Orthogonal Polynonials and random matrix theory : a Riemann-Hilbert approach”. *American Mathematical Society*, vol. 3 (2000).

[11] NIST Digital Library of Mathematical Functions. https://dlmf.nist.gov/.

[12] R. Durrett. “Probability: theory and examples” *Duxbury Press*, 1996.

[13] A.S. Fokas, A. Its & A.V. Kitaev. “An isomonodromy approach to the theory of two-dimensional quantum gravity” *Russian Mathematical Surveys* 45 no. 6 (1990).

[14] F.D. Gakhov. “Boundary value problems”. *Pergamon Press* (1966).

[15] I. Gohberg, S. Goldberg and N. Krupnik. Traces and determinants of linear operators Vol. 116, Birkhäuser (2012).

[16] G. Gautier, G. Polito, R. Bardenet, and M. Valko. “DPPy: DPP Sampling with Python.” *Journal of Machine Learning Research - Machine Learning Open Source Software (JMLR- MLOSS)*, 20(180), 1-7, (2019).

[17] B. Grammaticos, F. W. Nijhoff, A. Ramani. “Discrete Painlevé equations, The Painlevé property”. *CRM Ser. Math. Phys., Springer*, (1999)

[18] J.M. Hammersley. “A few seedlings of research” In *Proc. Sixth Berkeley Symp. Math. Statist. and Probability* Vol. 1, pp. 345-394 (1972).

[19] S.P. Hastings , J. B. Mcleod. “A boundary value problem associated with the second Painlevé transcendent and the Korteweg-de Vries equation.” *Archive for Rational Mechanics and Analysis*, 73(1), 31-51 (1980).

[20] K. Johansson. “Discrete orthogonal polynomial ensembles and the Plancherel measure.” *Annals of Mathematics*, 259-296, 2001.

[21] B.F. Logan, L.A. Shepp. “A variational problem for random Young tableaux” *Advances in mathematics*, 26(2), 206-222 (1977).

[22] A.M. Odlyzko, E.M. Rains. “On longest increasing subsequences in random permutations.” *Contemporary Mathematics*, 251, 439-452, (2000).

[23] Okunkov. “Infinite wedge and random partitions”. *Selecta Mathematica*, 7, 1999.

[24] G. Szegö. “Orthogonal Polynomials”. *American Mathematical Society*, vol. 23 (1939).

[25] C.A. Tracy, H. Widom. “Level-spacing distributions and the Airy kernel.” *Communications in Mathematical Physics*, 159, 151-174 (1994).

[26] S. Ulam. “ Monte Carlo calculations in problems of mathematical physics.” *Modern Mathematics for the Engineers*, 261–281, (1961).

[27] A. M. Vershik, S.V.E. Kerov.(1977). “Asymptotics of the Plancherel measure of the symmetric group and the limiting form of Young tableaux.” In *Doklady akademii nauk*, Vol. 233, No. 6, pp. 1024-1027 (1977).