1. Ulam's problem and the RSK algorithm.

schedule: every thursday, 1:30 → 3:30 pm, except 02-27? (or 02-20).
9 sessions + 1 session with presentations (04-03)

1. Increasing subsequences in random words.

Let us fix $N \geq 1$. A word with length n and letters in $[\![1, N]\!]$ is just a sequence $w = (w_1, w_2, \dots, w_n)$ with each $w_i \in [\![1, N]\!]$.
$= w_1 w_2 \cdots w_n$

example: $N = 6$, $n = 9$: $w = 431631263$
If $N = n$ and every letter $i \in [\![1, n]\!]$ appears exactly once, one obtains a permutation $\sigma \in \mathfrak{S}(n)$.
ex: $n = 9$, $\sigma = 741359268$

$|[\![1, N]\!]^n| = N^n$, $|\mathfrak{S}(n)| = n! = \prod_{i=1}^{n} i$.
One endows these set of words with their uniform probability measure.

Definition: An increasing subword or subsequence in a word w with length n is a subword $w_{i_1} w_{i_2} \cdots w_{i_\ell}$ with $i_1 < i_2 < i_3 < \cdots < i_\ell$ and $w_{i_1} \leq w_{i_2} \leq \cdots \leq w_{i_\ell}$.

ex: in 431631263, 1123 is an increasing subword.

We denote $\ell(w) = \max \{ \text{length of an increasing subword of } w \}$.

Ulam's problem: (1961) what is the distribution of $\ell_n = \ell(\sigma_n)$ with $\sigma_n \sim \mathrm{Unif}(\mathfrak{S}(n))$, $n \to +\infty$?

We can ask the same question with a large random word w_n in $[\![1, N]\!]^n$, $n \to +\infty$ (and possibly $N \to +\infty$ as well)
This question led Ulam to develop the Monte Carlo simulation method.

Hammersley, 1972: $\exists c > 0 \mid \frac{\ell_n}{\sqrt{n}} \xrightarrow[\mathbb{P}]{n \to +\infty} c$.

Logan-Shepp : $c > 2$ 1977

Kerov-Vershik : $c = 2$ 1977 (with the proof of $c \leq 2$ hidden in a 1985 paper)

Baik-Deift-Johansson 1999 $\left(\frac{\ell_n}{2\sqrt{n}} - 1\right) n^{1/3} \xrightarrow[\text{distribution}]{n \to +\infty}$ some limiting distribution TW.

The limiting Tracy-Widom distribution plays an important role in random matrix theory...

2. Greene invariants and random partitions.

To explain the whole story, we need to introduce the Greene invariants of a word.

__Definition:__ The k-th Greene invariant of a word w is

$$I_k(w) = \max \{ \ell(v_1) + \ell(v_2) + \dots + \ell(v_k),\ v_1, \dots, v_k \text{ disjoint increasing subwords of } w \}$$

Obviously, $(I_k(w))_{k \geq 0}$ is non-decreasing, so one can consider the sequence of non-negative increments

$$\lambda_k(w) = I_k(w) - I_{k-1}(w),\ k \geq 1.$$

__example__ : in $\sigma = 741359268$

$\lambda_1(\sigma) = \ell(\sigma) = 5$ $(v_1' = 13568)$

$\lambda_2(\sigma) = 2$ $(v_2' = 49)$

$\lambda_3(\sigma) = \lambda_4(\sigma) = 1$

$\lambda_{r \geq 5}(\sigma) = 0$

⚠ to compute $I_k(w)$, one cannot always use the subwords involved in the calculation of $I_{k-1}(w)$.

$$\lambda(w) = (\lambda_1(w), \dots, \lambda_r(w))$$

λ_r: last non-zero increment ; $r \leq \text{length}(w)$.

Non-trivial fact: $\lambda(w)$ is a non-increasing sequence ($(5,2,1,1)$ in the previous example).

Definition An integer partition with size n is a non-increasing sequence $\lambda = (\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_r)$ of positive integers with $|\lambda| = \sum_{i=1}^{r} \lambda_i = n$.

ex: $(5,2,1,1)$ is an integer partition with size 9 and length 4.

We denote $\mathcal{Y}(n) = \{$ integer partitions with size $n \}$

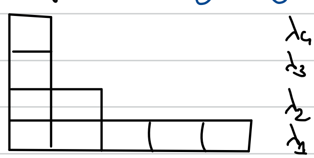
$\mathcal{Y}(1) = \{(1)\}$; $\mathcal{Y}(2) = \{(2), (1,1)\}$; $\mathcal{Y}(3) = \{(3), (2,1), (1,1,1)\}$

$\mathcal{Y}(4) = \{(4), (3,1), (2,2), (2,1,1), (1^4)\}$.

$$\mathcal{Y} = \bigsqcup_{n=0}^{\infty} \mathcal{Y}(n)$$

An integer partition is usually represented by its Young diagram:

$\lambda = (5,2,1,1) \longrightarrow$



$|\mathcal{Y}(n)|$? We can rewrite multiplicatively an integer partition:

$\lambda = 1^{m_1} 2^{m_2} \dots s^{m_s}$ with $m_i(\lambda)$ = number of parts of λ equal to i.

$$\prod_{n \geq 1} \frac{1}{1-z^n} = \prod_{n \geq 1} \sum_{m_n = 0}^{\infty} z^{n m_n}$$

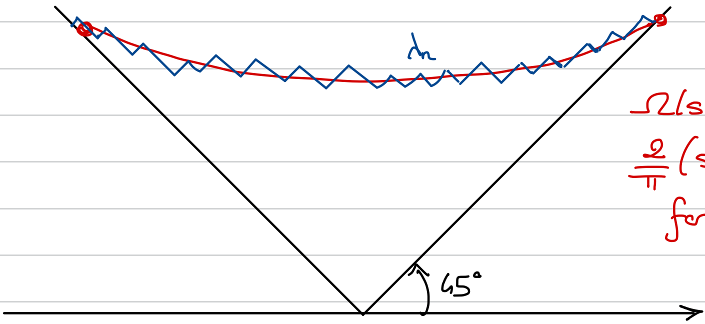
$$= \sum_{m_1, m_2, \dots \geq 0} z^{1 m_1 + 2 m_2 + \dots + s m_s}$$

$$= \sum_{k=0}^{\infty} |\mathcal{Y}(k)| z^k \longrightarrow \text{allows one to compute } |\mathcal{Y}(n)|.$$

$$|\mathcal{Y}(n)| \underset{n \to +\infty}{\simeq} \frac{1}{4n\sqrt{3}} \exp\left(\pi \sqrt{\frac{2n}{3}}\right).$$

Consider $\sigma_n \sim \mathrm{Unif}(\mathfrak{S}(n))$, and $\lambda_n = \lambda(\sigma_n)$. $\lambda_{n,1} = \ell_n$.

Logan-Shepp, Kerov-Vershik 1977: if one draws the Young diagram of λ_n with $\frac{1}{\sqrt{n}} \times \frac{1}{\sqrt{n}}$ cells, its boundary converges in probability to a continuous limit shape:



Kerov, 1993: $(\omega_n(s) - \Omega(s))\sqrt{n} \xrightarrow[n \to +\infty]{\text{law}}$ Gaussian distribution on $[-2,2]$

(random) continuous function associated to λ_n

$$\frac{2}{\pi} \sum_{k=2}^{+\infty} \frac{\xi_k}{\sqrt{k}} \sin\left(k \arccos \frac{s}{2}\right).$$

$\xi_k \sim \mathcal{N}(0,1)$ iid.

Borodin-Okounkov-Olshanski
Okounkov
Johansson
) 2000

$$n^{1/3}\left(\frac{\lambda_{n,1}}{2\sqrt{n}} - 1, \frac{\lambda_{n,2}}{2\sqrt{n}} - 1, \dots, \frac{\lambda_{n,k}}{2\sqrt{n}} - 1\right) \xrightarrow[n \to +\infty]{\text{law}}$$ k first particles of the Airy process. $\forall k \geq 1$.

Consider a random Hermitian matrix with size $n \times n$ and Gaussian entries:

$$M_n = \begin{pmatrix} X_1 & Y_{12} + i Z_{12} & & \\ Y_{12} - i Z_{12} & \ddots & & \\ & & \ddots & \\ & & & X_n \end{pmatrix}$$

with independent variables X_i, Y_{jk}, Z_{jk}

$$\mathbb{E}[X_i^2] = 1, \quad \mathbb{E}[Y_{jk}^2] = \mathbb{E}[Z_{jk}^2] = \frac{1}{2}.$$

We denote $x_{n,1} \geq x_{n,2} \geq \dots \geq x_{n,n}$ the spectrum of M_n.

$$n^{2/3}\left(\frac{x_{n,1}}{2\sqrt{n}} - 1, \frac{x_{n,2}}{2\sqrt{n}} - 1, \dots, \frac{x_{n,k}}{2\sqrt{n}} - 1\right) \xrightarrow[n \to +\infty]{} \text{a random vector with the same distribution as before!}$$
(law)

→ surprising connection between random matrix theory, interacting particle systems, and combinatorics.

3. The Robinson Schensted-Knuth algorithm.

If $\sigma_n \sim \mathrm{Unif}(\mathfrak{S}(n))$, then $\lambda_n = \lambda(\sigma_n)$ is not distributed uniformly on $\mathfrak{Y}(n)$...

related question: how to compute $\lambda(\sigma)$ or $\lambda(\omega)$?

Definition: A semi standard tableau with shape $\lambda \in \mathfrak{Y}(n)$ is a filling of the cells of the Young diagram of λ by numbers, which is
- non-decreasing along the rows →
- strictly increasing along the columns ↑.

If $n = |\lambda|$ and each number $i \in [\![1, n]\!]$ appears once, rows and columns are both strictly increasing and one says that the tableau is standard.

ex:

4	6		
3	3	6	
1	1	2	3

is a SST with shape $(4,3,2)$

3	6		
2	5	9	
1	4	7	8

is a ST ——— $(4,3,2)$.

$\mathrm{SST}(\lambda, N) = \{$ semistandard tableaux with shape λ and entries in $[\![1, N]\!]\}$

$\mathrm{ST}(\lambda) = \{$ standard tableaux with shape $\lambda \}$.

(difficult question: compute $|ST(\lambda)|$ and $|SST(\lambda, N)|$ ...).

Theorem Robinson 1938, Schensted 1960, Knuth 1970. There is a combinatorial bijection

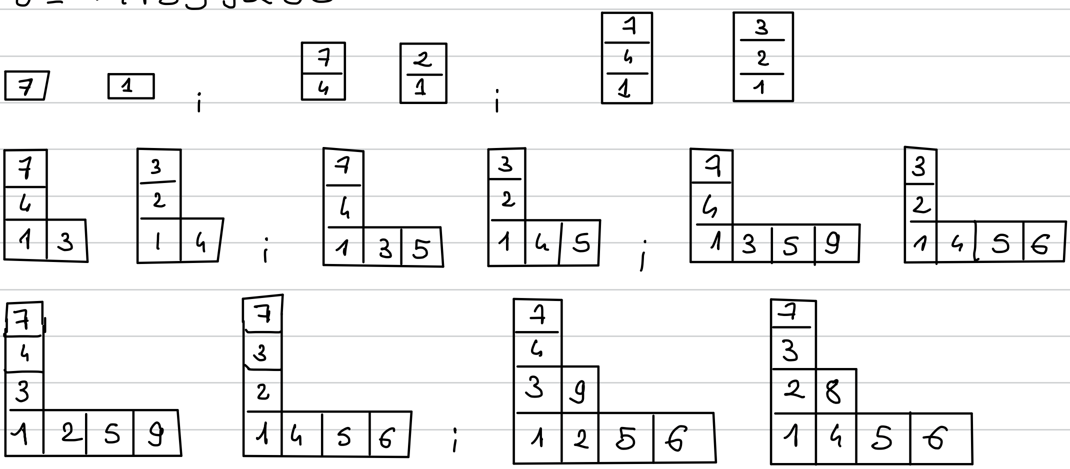
$$RSK : [\![1, N]\!]^n \longrightarrow \bigsqcup_{\lambda \in \mathcal{Y}(n)} SST(\lambda, N) \times ST(\lambda).$$
$$w \longmapsto (P(w), Q(w))$$

If $N = n$, the restriction of RSK to $\mathfrak{S}(n)$ yields a bijection

$$\mathfrak{S}(n) \longrightarrow \bigsqcup_{\lambda \in \mathcal{Y}(n)} ST(\lambda) \times ST(\lambda).$$

example:

$\sigma = 741359268$



and finally:

$P(\sigma) =$

7				
4				
3	9			
1	2	5	6	8

$Q(\sigma) =$

7				
3				
2	8			
1	4	5	6	9

The second tableau $Q(w)$ records the growth of the two tableaux P and Q, therefore it is standard.

Why is $P(w)$ semistandard? non decreasing rows by construction.

One has to show that a configuration $\begin{array}{|c|c|}\hline b & c \\ \hline \multicolumn{1}{c|}{} & a \\ \hline\end{array}$ with $b \leq c \leq a$ is impossible.

by induction: if P is semistandard, then $P \leftarrow i$ also.

if $P \leftarrow i$ contains $\begin{array}{|c|c|}\hline b & c \\ \hline \multicolumn{1}{c|}{} & a \\ \hline\end{array}$ with $b \leq c \leq a$, then previously

c was on the lower row, before a, but strictly after the column of b

One can invert the insertion procedure, whence the bijective character. □.

<u>Theorem</u> (Greene, 1974). The shape of $P(w)$ and $Q(w)$ is $\lambda(w)$.

Let us prove that the length of the first row of $P(w)$ and $Q(w)$ is $\ell(w)$.

We define the i-th basic word of w as the subword which consists in letters inserted in position $(1, i)$ at the time of their insertion (later on, they might be pushed).

<u>ex:</u> with $w = \sigma = 741359268$,

$b_1 = 741$

$b_2 = 32 \quad \longleftarrow$ by construction, decreasing words.

$b_3 = 5$

$b_4 = 96$

$b_5 = 8$.

Any non-decreasing subword can intersect each b_i at most once: $\ell(w) \leq sh(P(w))_1$.

<u>lemma:</u> if $w_j \in b_{k+1}$, then there exists $w_i \in b_k$ with $w_i \leq w_j$, $i < j$.

Therefore, one can construct a non decreasing subword with length $sh(P(w))_1$ □.

4. Measures on integer partitions.

The image of the uniform measure by RSKshape is :

– the Schur-Weyl measure $SW_{N,n}$ if $w \sim \mathrm{Unif}([\![1,N]\!]^n)$:

$$SW_{N,n}(\lambda) = \frac{|SST(\lambda,N)|\ |ST(\lambda)|}{N^n}$$

– the Plancherel measure Pl_n if $\sigma \sim \mathrm{Unif}(\mathfrak{S}(n))$.

$$Pl_n(\lambda) = \frac{|ST(\lambda)|^2}{n!}.$$

problem: study $\lambda_n \sim Pl_n$. In particular, what is the asymptotic behavior of $\lambda_{n,1} \underset{\text{law}}{=} \ell(\sigma_n)$? .

Lemma: $\mathbb{E}[\ell_n] \leq 2\sqrt{n}$ $\forall n \geq 1$.

Proof: It suffices to prove that $\mathbb{E}[\ell_n - \ell_{n-1}] \leq \frac{1}{\sqrt{n}}$.

$\sigma_n \sim \mathrm{Unif}(\mathfrak{S}(n))$; $\sigma_{n,n-1} = n-1$ first letters.

$\ell_n = \ell(\sigma_n)$; $\ell_{n-1} \underset{(\text{law})}{=} \ell(\sigma_{n,n-1})$.

$$\mathbb{E}[\ell_n - \ell_{n-1}] = \mathbb{P}[\ell_n = 1 + \ell_{n-1}] = \mathbb{P}[n \in Q(\sigma_n)_1]$$

$$= \sum_{\lambda \in \mathcal{Y}(n)} Pl_n[\lambda]\ \mathbb{P}[n \in Q(\sigma_n)_1 \mid \lambda(\sigma_n) = \lambda]$$

$$\_\_\_\_^2 \underset{\text{Cauchy-Schwartz}}{\leq} \sum_{\lambda \in \mathcal{Y}(n)} Pl_n[\lambda]\ \mathbb{P}[n \in P(\sigma_n)_1 \text{ et } n \in Q(\sigma_n)_1 \mid \lambda(\sigma_n) = \lambda]$$

$$\leq \mathbb{P}[n \in P(\sigma_n) \text{ et } n \in Q(\sigma_n)] = \mathbb{P}[\sigma_n(n) = n] = \frac{1}{n}.$$

$\square$.