4 - The longest increasing subsequence

4.1 - The longest increasing subsequence of uniform permutations

Given a permutation σ, an increasing subsequence is a subsequence of the values of σ which is increasing.

Example: $\sigma = 4\,6\,3\,5\,2\,1\,7$ 3,5,7 4,5,7 form an increasing subsequence of σ

Note that it might be NOT unique.

We denote by $LIS(\sigma)$ the length of the longest increasing subsequence of σ.

Question: (Ulam '60s) How large is $LIS(\sigma_n)$ if σ_n is a uniform random perm of size n?

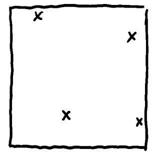
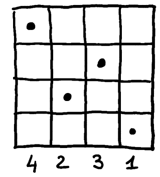
- Hammersley (1972): $LIS(\sigma_n) \sim c \cdot n^{1/2}$ for some $c>0$.
- Veršik & Kerov (1977), Logan & Shepp (1977): $c=2$
- Baik, Deift & Johansson (1999): $LIS(\sigma_n) \sim 2\, n^{1/2} + TW \cdot n^{1/6}$
  ↳ Random Variable with Tracy-Widom distribution.
- Dauvergne & Virág (2022): Full description of the scaling limit (= geodesic of the directed landscape)

The problem of studying LIS for unif. permutations has beautiful connections with many different areas of mathematics: see the book written by D. Romik (2015).
"The Surprising Mathematics of Longest Increasing Subsequences"

4.2. - The length of the longest increasing subsequence in μ_p^S-permutations.

Let μ_p^S be the BSP of parameter $p \in (0,1)$.

We can easily contruct a μ_p^S-random permutation sampled from μ_p^S:

Sample μ_p^S $\rightsquigarrow$  $\rightsquigarrow$ 

$= \sigma_n$

(Recall that μ_p^S is a random perm.)

(Sample n iid points using μ_p^S)

(Consider the unique permutation σ_n that has the same rel order as the points on the left)

Question: If σ_n is sampled from μ_p^S, how large is $LIS(\sigma_n)$?

Motivations:

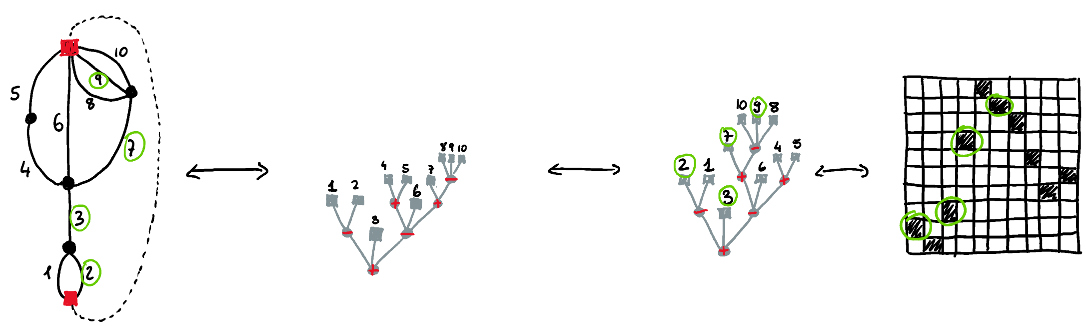
Not formally proved (but we know how to do it + simulations)
Note that this fact is NOT obvious since $LIS(\cdot)$ is NOT continuous for permutons.

- This question should be equivalent to study the LIS of a unif. sep. permutation.
- The length of the longest increasing subsequence of a separable permutation is equal to:

  ① the length of the longest directed path in the associated series-parallel map

  ② the size (# of leaves) of the largest subtree (spanned by leaves) of the corresponding Schröder tree containing only $\oplus$ at the branching points. [THIS CAN BE THOUGHT AS A SORT OF PERCOLATION MODEL]

  [③ the size of the largest clique in a certain model of random graphs (uniform cographs)] Not discussed here!



- ① can be interpreted as a model of LAST-PASSAGE-PERCOLATION on planar maps.

The first interesting result on $LIS(\sigma_n)$ when σ_n is sampled from μ_p^S is the following one:

THEOREM (Bassino, Bouvel, Drmota, Feray, Gerin, Maazoun, Pierrot, 2022)

For all $p \in (0,1)$, it holds that

$$\frac{LIS(\sigma_n)}{n} \xrightarrow[n\to\infty]{P} 0.$$

The above theorem tells us that $LIS(\sigma_n)$ is always $(\forall p \in (0,1))$ sublinear. This seems a quite weak result, but we will see later that this is the best you can do, if you want a result $\forall p \in (0,1)$.

Observation: When $p = 1/2$, using the symmetries of $\mathcal{M}^S_p$ and the Erdös–Szekeres thm. we can immediately deduce that

$$\mathbb{P}\left(LIS(\sigma_n) \geq n^{1/2} \right) \xrightarrow[n\to\infty]{} 1.$$

This work motivated the question of determining more precise estimates for $LIS(\sigma_n)$:
For instance, for $p \in (0,1)$ fixed, $LIS(\sigma_n) \approx n^d$ for some $0<d<1$
or $LIS(\sigma_n) \approx n/\log(n)$ or something similar?

4.2 - Power-law bounds for the LIS of $\mathcal{M}^S_p$-permutations.

THEOREM [B., Da Silva, Gwynne, 2023]

There exists explicit functions

$$\alpha_* : (0,1) \longrightarrow (1/2, 1)$$

$$\beta^* : (0,1) \longrightarrow (1/2, 1)$$

such that for all $p \in (0,1)$:

- $1/2 < \alpha_*(p) < \beta^*(p) < 1$
- For each $\alpha < \alpha_*(p)$ and $\beta > \beta^*(p)$, it holds with probability tending to one as $n \to \infty$ that

$$n^\alpha \leq LIS(\sigma_n) \leq n^\beta.$$

Equivalently, $$P\left(n^{\alpha_*(p)+o(1)} \leq LIS(\sigma_n) \leq n^{\beta^*(p)+o(1)}\right) \xrightarrow[n\to\infty]{} 1.$$

Remark: For all $p \in (0,1)$,

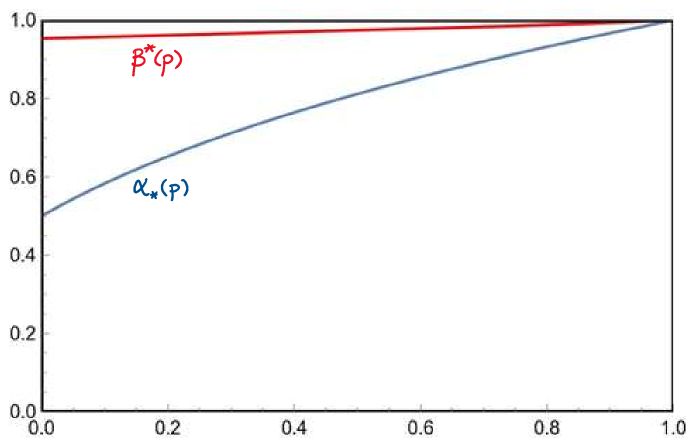
$$\alpha_*(p) = 1 - \lambda_*(p)$$

where $\lambda_*(p)$ is the only positive solution to the equation

$$\phi^S(-\lambda_*(p)) = -2(1-p)\sqrt{\frac{2}{\pi}}$$

with

$$\phi^S(q) = \int_0^\infty (1-e^{-qx}) \left(\mathbb{1}_{\{x \in (0, \log(2)]\}} + p\, \mathbb{1}_{\{x \in \log(2), +\infty\}} \right) \frac{2e^x\, dx}{\sqrt{2\pi(e^x-1)^3}}$$



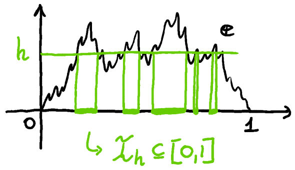
Also $\beta^*(p)$ has an explicit expression but it is even less nice than the one for $\alpha_*(p)$.

4.3 - Ideas for the proof of the lower bound.

- Recall that σ_n is sampled from μ_p^S.
- We want to construct a long increasing subsequence of σ_n.
  ↳ (Not necessarily the longest one)
- From the construction of μ_p^S from $\lhd$, it is not too hard to realize that this is equivalent to find a large ordered subset O of $[0,1]$.
  in some sense (for instance some fractal dimension)
  ↳ w.r.t. $\lhd$, i.e. $\forall\, x<y \in O$ we want $x \lhd y \iff S_{xy} = \oplus$
- Hence our goal is the following one:

  GOAL: Construct a subset $O \subseteq [0,1]$ such that $S_{xy} = \oplus$ for all $x, y \in O$.

- IDEA: Explore the excursion from bottom to top:



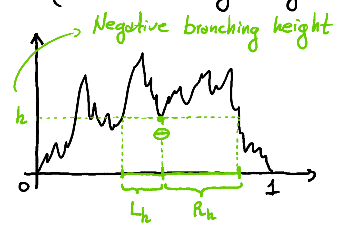
$\mathcal{I}_h$ = Union of "excursion intervals" above height h (see picture)

Obviously, if $h < h'$ then $\mathcal{I}_h \subseteq \mathcal{I}_{h'}$

[Bertoin '02] & [Aldous-Pitman '98] proved that $(\mathcal{I}_h)_{h \geq 0}$ is a SELF-SIMILAR FRAGMENTATION PROCESS

- Note that if h is a NEGATIVE-BRANCHING HEIGTH (i.e. the height of a local minimum of e with a negative sign), then we need to exclude either L_h or R_h from O.

(Note also that if h is a positive branching heigth then there are NO restrictions.)



One among $O \cap L_h$ and $O \cap R_h$ must be empty, otherwise I can find $x, y \in O$ s.t. $S_{xy} = \ominus$ ☹

- The idea is that we start with $h = 0$ and $O_h = [0,1]$, then we let h increase, and each time we hit a negative branching height then we remove from O_h either L_h or R_h.

  Note that by construction then O_∞ is a ordered set with respect to $\triangleleft$.

- QUESTION: How do we select which one among L_h and R_h we discard from O_h?

  - OPTIMAL STRATEGY: Discard the side containing the smallest ordered subset.
    (For some reasonable notion of "size" (for instance Hausdorff dimension).)

    ⚠ This is a terribly complicated strategy to analyze! It is not Markovian!
    ↳ Recall that at the end of the day we need to prove that O_∞ is large, so we will need to do some estimates on our "selection rule".

  - A MORE REASONABLE STRATEGY: Use the following SELECTION RULE S:

$$S := \text{"At every negative branching heigth, discard the smallest between } L_h \text{ \& } R_h\text{"}$$

(in terms of Lebesgue measure)

- Using this strategy $\mathcal{S}$ we get a set $O_\infty^{\mathcal{S}} \subseteq [0,1]$ which is ordered with respect to $\triangleleft$ by construction!

- We are now left with the following task:

  Show that $O_\infty^{\mathcal{S}}$ is large [$O_\infty^{\mathcal{S}}$ is a fractal subset of $[0,1]$ so the correct notion for "large" is large fractal dimension.]

- It is quite standard (but not easy to prove) that in order to prove that a set $A \subseteq [0,1]$ has fractal dimension $\alpha \in [0,1]$, one needs to prove that

$$\mathbb{P}(U \in A_\varepsilon) \underset{\varepsilon\to 0}{\sim} C \cdot \varepsilon^{\lambda} \quad \text{with} \quad \lambda = 1-\alpha \quad \left[\begin{array}{l} \text{Plus a second moment estimate:} \\ \mathbb{P}(U \in A_\varepsilon, V \in A_\varepsilon) \underset{\varepsilon\to 0}{\sim} C' \cdot \varepsilon^{2\lambda} \end{array} \right]$$

where U is uniform in $[0,1]$ and $A_\varepsilon = \{ y \in [0,1] \mid d(y, A) < \varepsilon \}$.

Intuition: One possible definition of fractal dimension [Minkowski dimension] for $A \subseteq [0,1]$ is:

$$\dim_M(A) := \lim_{\varepsilon \to 0} \frac{\log(N_A(\varepsilon))}{\log(\varepsilon^{-1})} \quad \text{, where} \quad N_A(\varepsilon) := \text{minimal \# of intervals of size } \varepsilon \text{ needed to cover } A$$

Why? If you take a d-dimensional box of side length 1 you need $\approx \varepsilon^{-d}$ boxes of side length ε to cover it.

Hence if A has fractal dimension $\dim(A)$ we expect that

$$N_A(\varepsilon) \approx \varepsilon^{-\dim(A)}$$

$$\Rightarrow \log(N_A(\varepsilon)) \approx \dim(A) \cdot \log(\varepsilon^{-1}) \Rightarrow \frac{\log(N_A(\varepsilon))}{\log(\varepsilon^{-1})} \xrightarrow[\varepsilon \downarrow 0]{} \dim(A)$$

Hence, if we want to prove that $\dim(A) = \alpha$, we need to prove that

$$N_A(\varepsilon) \approx \varepsilon^{-\alpha}$$

But if U is uniform in $[0,1]$ then

$$P(U \in A_\varepsilon) \approx \frac{\varepsilon^{-\alpha}}{\varepsilon^{-1}} = \varepsilon^{1-\alpha}$$

good intervals

all possible intervals

Chapter 4 "Hausdorff dim."

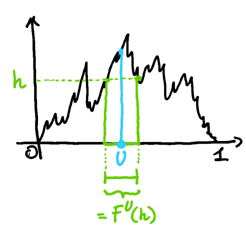
A good reference for fractal dimensions (of random sets) is "Brownian Motion" by Mörters Peres

Hence we now want to estimate

$$\mathbb{P}(U \in O_\varepsilon)$$

where here $O = O_\infty^{S}$. One can show that the above probability is close to:

$$\mathbb{P}(U \text{ is } \underline{\text{not}} \text{ discarded by } S \text{ before } F^U(h) < \varepsilon)$$



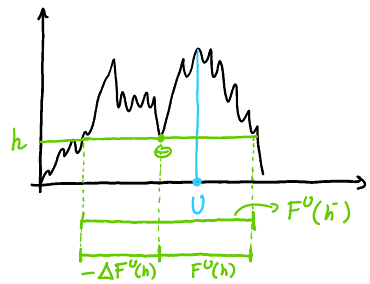
where $F^U(h) :=$ the length of the excursion containing U at height h

- We are left with this new GOAL:

> Show that
>
> $$\mathbb{P}(U \text{ is not discarded by } S \text{ before } F^U(h) < \varepsilon) \underset{\varepsilon \to 0}{\sim} \varepsilon^{\lambda_*(p)}$$

4.4 - Estimates for the above probability

When do $\mathcal{S}$ discards U ?



$\Delta F^U(h) = F^U(t) - F^U(t^-) < 0$

$\mathcal{S}$ discards U at the negative height h $\iff$ $F^U(h) < -\Delta F^U(h)$

Set

$$H^U_{\mathcal{S}} := \inf \{ h>0 \mid h \text{ is a neg-branching height } \& \ F^U(h) < -\Delta F^U(h) \}$$ "1st time when $\mathcal{S}$ discards U"

$$H^U_{\varepsilon} := \inf \{ h>0 \mid F^U(h) < \varepsilon \}$$ "1st time F^U drops below ε"

Then we want to compute

$$\mathbb{P}\left(H^U_{\varepsilon} < H^U_{\mathcal{S}} \right)$$

Proposition: [Bertoin, 2002]

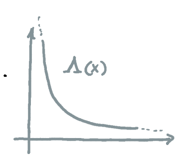
$F^U(h)$ is a positive self-similar Markov process and

$$F^U(h) = \exp\left(-\xi_{\rho(h)}\right)$$

where $\rho(h)$ is an explicit time-change and ξ_t is a subordinator with Laplace exponent

$$\phi(q) := -\log \mathbb{E}\left[e^{-q\xi_1}\right] = \int_0^\infty (1-e^{-qx}) \underbrace{\frac{2e^x}{\sqrt{2\pi(e^x-1)^3}}\,dx}_{:=\Lambda(dx)} \quad , \quad q > -\tfrac{1}{2}.$$

"Lévy measure"



Intuition: $\Lambda(A) := \mathbb{E}\left[\#\text{ of jumps of } \xi_t \text{ having step size in } A \text{ in an interval of size } 1\right]$

Note that

$$F^U(h) < -\Delta F^U(h) \iff F^U(h) < F^U(h^-) - F^U(h) \iff \underset{2e^{-\xi_{\rho(h)}} < e^{-\xi_{\rho(h^-)}}}{\underbrace{2F^U(h) < F^U(h^-)}} \iff e^{\log(2)} < e^{\Delta\xi_{\rho(h)}} \iff \Delta\xi_{\rho(h)} > \log(2)$$

So, we can rewrite

$$H_S^U = \inf\{h \geq 0 \mid h \text{ is a neg-branching height and } \Delta\xi_{\rho(h)} > \log(2)\}$$

We now introduce two new processes χ and $\bar{\xi}$.

- Conditioning on ξ, for every s s.t. $\Delta\xi_s > 0$ we independently set

$$\begin{cases} \chi_s = 1 & \text{w.p. } p \\ \chi_s = 0 & \text{w.p. } 1-p \end{cases}$$

- Now we introduce $\bar{\xi}$ which is a killed version of ξ.
  - If $\chi_s = 1$, we do nothing.
  - If $\chi_s = 0$, then we kill ξ (i.e. we set $\bar{\xi}_t = \infty$, $\forall t \geq s$) if $\Delta\xi_s > \log(2)$.

Note that one can write $\bar{\xi} = \xi^S + \xi^K$ where ξ^S is a subordinator with Levy measure

$$\Lambda^S(dx) := \Lambda(dx)\Big|_{x\in[0,\log(2)]} + p\cdot\Lambda(dx)\Big|_{x\in(\log(2),\infty)}$$

and ξ^K is an independent subordinator with with Levy measure

$$\Lambda^K(dx) = (1-p)\,\Lambda(dx)\Big|_{x\in(\log(2),\infty)}$$

With this description ξ has the law of ξ^S killed at the first time T when ξ^K takes a jump. T is an exponential random variable with parameter

$$\lambda(\rho) = \int_{\log(2)}^{\infty} \Lambda^K(dx) = 2(1-\rho)\sqrt{\frac{2}{\pi}}$$

Proposition: Let $\lambda_*(\rho)$ be the only positive solution to the equation

$$\phi^S(-\lambda_*(\rho)) = -\lambda(\rho)$$

where $\quad \phi^S(q) = \int_0^\infty (1-e^{-qx})\, \Lambda^S(dx)$. Then

$$\mathbb{P}(H_\varepsilon^U < H_S^U) \underset{\varepsilon\to 0}{\sim} c \cdot \varepsilon^{\lambda_*(\rho)}.$$

Proof: Note that

$$\mathbb{P}(H_\varepsilon^U < H_S^U) = \mathbb{P}(T_{-\log\varepsilon} < T)$$

where $T_x = \inf\{s>0,\ \xi_s^S > x\}$. Since T is an exponential random variable independend of $T_{-\log\varepsilon}$ (because ξ^K and ξ^S are independent):

$$\mathbb{P}(T_{-\log\varepsilon} < T) = \mathbb{E}\left[\mathbb{P}(T > T_{-\log\varepsilon} \mid T_{-\log\varepsilon})\right] \underbrace{=}_{\mathbb{P}(T>t) = e^{-\lambda t}} \mathbb{E}\left[e^{-\lambda(\rho)\, T_{-\log\varepsilon}}\right]$$

So we are left with computing the Laplace transform of $T_{-\log\varepsilon}$.
For simplicity set $a = \lambda(\rho)$ and $x = -\log\varepsilon$; we compute $\mathbb{E}[e^{-aT_x}]$.

TRICK (classical in Lévy processes): Exponential tilting!

Let $a_* = \lambda_*(p)$, so that $\phi^S(-a_*) = -a$ and consider the process

$$M_s^a = e^{a_* \xi_s^S - as}, \quad s \geq 0.$$

This is a martingale:

$$\mathbb{E}\left[e^{a_* \xi_s^S - as} \,\middle|\, \mathcal{G}_t\right] = \mathbb{E}\left[e^{a_*(\xi_s^S - \xi_t^S) - a(s-t)} e^{a_* \xi_t^S - at} \,\middle|\, \mathcal{G}_t\right]$$

$$\stackrel{!}{=} e^{a_* \eta_t - at} \quad \text{by def of } \phi^S \text{ and } \phi^S(-a_*) = a.$$

Let $\mathbb{P}^a$ be the measure defined by $d\mathbb{P}^a = M_s^a \cdot d\mathbb{P}$. Then under $\mathbb{P}^a$

ξ_t^S is a Lévy process with Laplace exponent $\phi^{S,a}(q) := \phi^S(q - a_*) + a$

By an optional stopping type argument, we get:

$$\mathbb{E}^a\left[e^{-a_* \xi_{T_x}^S}\right] = \mathbb{E}\left[e^{-a_* \xi_{T_x}^S} \underbrace{e^{a_* \xi_{T_x}^S - a T_x}}_{= M_{T_x}^a}\right] = \mathbb{E}\left[e^{-a T_x}\right].$$

Expectation with respect to $\mathbb{P}^a$.

But, since ξ^S is a Levy process under $\mathbb{P}^a$, by standard theory on Lévy processes, one gets that

$$\mathbb{E}\left[e^{-a_* \xi_{T_x}^S}\right] \underset{x \to \infty}{\sim} c \cdot e^{-a_* \cdot x}. \qquad \left[\text{Since } \xi_{T_x}^S \approx x\right]$$

Substituting we get that $\mathbb{P}(H_\varepsilon^0 < H_S^0) \underset{\varepsilon \downarrow 0}{\sim} c \cdot \varepsilon^{\lambda_*(p)}$. $\square$

4.5 - Some additional comments.

① The estimate above is the key result to get the value $\alpha_*(p)$ in the theorem, but in order to complete the proof of the theorem there is quite a lot of technical work:

Roughly speaking one need to deduce from

$$(*) \quad \mathbb{P}(U \text{ is not discarded before } F^U(h) < \varepsilon) \underset{\varepsilon \downarrow 0}{\sim} c\, \varepsilon^{\lambda_*(p)}$$

that $\forall \alpha < \alpha_*(p) = 1 - \lambda_*(p)$

$$(**) \quad \mathbb{P}\left(\mathrm{LIS}(\sigma_n) \geq c \cdot n^{\alpha} \right) \xrightarrow[n\to\infty]{} 1$$

From $(*)$ and a similar two-point estimate one gets that $\exists\, q > 0$ s.t.

$$(***) \quad \mathbb{P}\left(\mathrm{LIS}(\sigma_n) \geq c \cdot n^{\alpha} \right) \geq q \quad \forall n \geq 1.$$

Then one concludes the proof of $(**)$, using $(***)$ and a (very delicate) zero-one law.

② The upper bound is using the idea that

"Whatever is the strategy S that I use, I must discard something at every negative-branching height, and so I must discard a lot in the end"

Enough so that the exponent is strictly less that 1.