

# INTRODUCTION AU DOMAINE DE RECHERCHE : FORME LIMITE DE PERMUTATIONS ALÉATOIRES ET GRAPHES ALÉATOIRES

THÉO LENOIR

## 1. RÉSUMÉ

Les permutations et graphes aléatoires constituent deux vastes objets d'étude bien connus en probabilités. Pour les permutations, beaucoup de phénomènes sont déjà bien compris : nombre de points fixes, nombre de records, nombre de cycles, etc. Pour les graphes aléatoires, beaucoup de modèles ont déjà été étudiés (graphe d'Erdős-Renyi, graphes d'attachement préférentiel, etc) et de nombreuses propriétés sont connues : nombre d'arêtes, nombre de sommets isolés, connexité, taille de la plus grande composante connexe.

Néanmoins, ici on présente un nouveau point de vue sur les permutations : l'objectif est de voir une permutation comme une mesure aléatoire sur  $[0, 1]^2$  et de regarder comment se comporte une telle mesure : a-t-on convergence en loi ? Si oui, quel est l'objet limite ?

Nous introduisons d'abord les différentes connaissances dans le domaine des permutations aléatoires, puis nous évoquons trois outils importants : l'espace des permutons, la décomposition d'une permutation, et la combinatoire analytique. Ces outils permettent dans [1] de prouver la convergence de certaines permutations aléatoires vers un objet appelé le permuton brownien

## TABLE DES MATIÈRES

1. Résumé	1
2. Permutation aléatoire	3
3. Permutons aléatoire	4
3.1. Convergence : notion de permuton	4
3.2. Nombre d'occurrences	5
3.3. Nombre d'occurrences pour un permuton	5
4. Décomposition d'une permutation et arbre associé	7
5. Combinatoire analytique	8
6. Quelques questions ouvertes	11
Références	12

## 2. PERMUTATION ALÉATOIRE

Ici, on va s'intéresser aux permutations d'un ensemble de cardinal fini. Une permutation est tout simplement une bijection d'un ensemble dans lui-même.

**Définition 2.1.** On note  $\mathfrak{S}_n$  l'ensemble des bijections de  $\{1, \dots, n\}$  dans lui-même. Une permutation aléatoire est une variable aléatoire à valeur dans  $\mathfrak{S}_n$ .

Il existe plusieurs modèles de permutations aléatoires, mais le plus connu est le modèle uniforme, où  $\sigma_n$  est une variable aléatoire uniforme dans  $\mathfrak{S}_n$ . On s'intéresse alors à avoir des propriétés exactes, ou des propriétés asymptotiques, i.e. des convergences ou limites lorsque  $n$  tend vers  $+\infty$ .

De nombreuses études ont déjà été effectuées sur le fonctionnement d'une permutation aléatoire uniforme. Par exemple :

**Proposition 2.1.** *Soit  $\sigma$  une permutation aléatoire uniforme,  $\sigma$  a en moyenne 1 points fixes.*

On connaît aussi des moyens de simuler des permutations, comme par exemple l'algorithme suivant :

**Proposition 2.2.** *Pour obtenir une permutation aléatoire uniforme, on applique l'algorithme suivant, initialisé à  $E = \{1, \dots, n\}$  et  $a = 1$ .*

- *On tire un point  $a_1$  uniformément, indépendamment du reste, dans  $E$ . Si  $a_1$  ne vaut pas  $a$ , on enlève  $a_1$  de  $E$ . On itère le procédé pour obtenir  $a_1, \dots, a_k$  jusqu'à avoir  $a_k = a$ . On ajoute alors le cycle  $(a_1, \dots, a_k)$  à la décomposition de  $\sigma$  en cycles disjoints, et on remplace  $E$  par  $E$  privé de  $\{a_1, \dots, a_k\}$  puis on pose  $a$  le minimum de ce nouveau  $E$  s'il est non vide.*
- *Si  $E$  est vide, on retourne  $\sigma$*

L'algorithme précédent permet de simuler des permutations aléatoirement et d'obtenir des informations intéressantes sur le nombre de cycle.

**Proposition 2.3.** *Soit  $\sigma$  une permutation aléatoire uniforme. Le nombre de cycle de sigma a la même loi que  $\sum_{i=1}^n X_i$  où les  $(X_i)_{1 \leq i \leq n}$  forment une famille de variables indépendantes de Bernoulli de paramètres  $(\frac{1}{i})_{1 \leq i \leq n}$ .*

En particulier, on peut comprendre le comportement de la loi du nombre de cycle.

Globalement, les permutations uniformes sont bien connues, et leurs propriétés liées au nombre de points fixes, à la décomposition en cycle disjoints sont assez connues et bien maîtrisées.

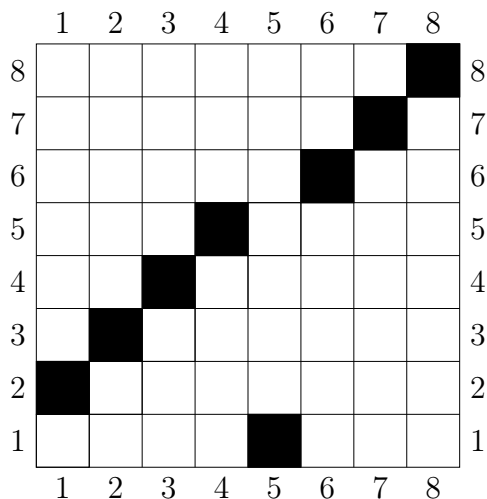
Néanmoins, même si on comprend les différents aspects d'une permutation aléatoire, il n'y a pas de notions de limite "géométrique" d'une permutation aléatoire. On aimerait pouvoir définir une notion ressemblant aux limites locales ou limites d'échelles pour des permutations aléatoires.

### 3. PERMUTONS ALÉATOIRE

**3.1. Convergence : notion de permuton.** On souhaite obtenir une forme de convergence locale d'une permutation. Pour cela, on peut regarder le graphe d'une permutation.

Une permutation aléatoire donne une courbe aléatoire. Mais il est difficile d'envisager un objet limite sous cette forme : il faut renormaliser le graphique pour obtenir une courbe interne au carré  $[0, 1]^2$ .

Une idée naturelle est d'essayer d'obtenir un objet géométrique un peu plus consistant : ici on peut noircir un carré de côté  $\frac{1}{n}$  au dessus de chaque point, pour obtenir un ensemble géométrique.



Néanmoins, cette figure géométrique a une aire valant  $\frac{1}{n}$ . On cherche donc à effectuer une renormalisation convenable, afin d'avoir une limite intéressante.

À une permutation on peut associer une mesure qui consiste à tirer aléatoirement un point parmi ces carrés. On obtient alors une mesure sur  $[0, 1]^2$  dont chaque marginale est uniforme.

On a alors les deux définitions suivantes :

**Définition 3.1.** Un permuton est une mesure de probabilité  $\mu$  sur  $[0, 1] \times [0, 1]$  dont les marginales sont uniformes. On définit  $\mathcal{L}$  l'ensemble des permutons qu'on munit de la topologie étroite.

**Définition 3.2.** Soit  $n \geq 2$ ,  $\sigma \in \mathfrak{S}_n$ , on définit  $\mu_\sigma$  le permuton associé à  $\sigma$  par  $\mu_\sigma(dx dy) = n \mathbb{1}_{\sigma(\lceil nx \rceil) = \lceil ny \rceil} dx dy$ .

Il y a donc naturellement une notion de convergence étroite d'un permuton aléatoire, mais peu pratique et peu calculable. Néanmoins il existe un outil, le nombre d'occurrences, qui permet de mesurer assez facilement la convergence au sens des permutons de permutations aléatoires.

### 3.2. Nombre d'occurrences.

**Définition 3.3.** Soit  $n \geq 1$ ,  $\sigma \in \mathfrak{S}_n$ ,  $I \subset \llbracket 1, n \rrbracket$ . On définit  $\text{pat}_I(\sigma)$  la permutation induite par la restriction de  $\sigma$  à  $I$ .

De façon plus détaillée, si  $i_1 < \dots < i_k$  sont les éléments de  $I$ ,  $\text{pat}_I(\sigma)$  est la permutation de  $\llbracket 1, k \rrbracket$  telle que si  $1 \leq l \neq j \leq k$ , on a  $\text{pat}_I(\sigma)(l) < \text{pat}_I(\sigma)(j)$  si et seulement si  $\sigma(i_l) < \sigma(i_j)$ .

**Exemple 3.1.** Pour  $\sigma = (1, 3, 2, 5, 4)$ ,  $\text{pat}_{\{1,3,5\}}(\sigma) = (1, 2)$

On peut ainsi définir le nombre d'occurrences d'un motif au sein d'une permutation.

**Définition 3.4.** Soit  $n, m \geq 1$  deux entiers,  $\sigma \in \mathfrak{S}_n$  et  $\pi \in \mathfrak{S}_m$ , on définit  $\text{occ}(\pi, \sigma)$  le nombre d'ensembles  $I$  inclus dans  $\llbracket 1, n \rrbracket$  de cardinal  $m$  tels que  $\text{pat}_I(\sigma) = \pi$ .

On définit également

$$\tilde{\text{occ}}(\pi, \sigma) = \frac{\text{occ}(\pi, \sigma)}{\binom{n}{m}}$$

En fait ce nombre moyen d'apparition d'un motif a une interprétation probabiliste simple. En effet, on a assez clairement :

$$\tilde{\text{occ}}(\pi, \sigma) = \mathbb{P}[\text{pat}_I(\sigma) = \pi]$$

où  $I$  est une variable aléatoire uniforme parmi les sous-ensembles à  $k$  éléments de  $\llbracket 1, n \rrbracket$ .

**3.3. Nombre d'occurrences pour un permuton.** L'objectif est de construire un équivalent de ce nombre d'occurrences pour un permuton. Naturellement, on aimerait tirer  $k$  points dans  $k$  carrés différents pour un permuton associé à une permutation et regarder la permutation induite.

Néanmoins ce procédé est difficilement définissable : pour un permuton quelconque, la notion de "carrés" n'a pas de sens. On peut donc décider de tirer  $k$  points de façon indépendante, selon la mesure associée à  $\mu$ . Lorsque  $n$  est grand, la probabilité d'avoir deux points dans le même carré est majorée par  $\frac{\binom{k}{2}}{n}$  donc petite à  $k$  fixé,  $n$  grand.

On obtient donc quasiment le nombre moyen d'occurrence associé à la permutation aléatoire.

**Définition 3.5.** Soit  $k \geq 1$  et  $\mu$  un permuton et  $\pi \in \mathfrak{S}_k$ , on définit le nombre moyen d'occurrences de  $\pi$  dans  $\mu$  par :

$$(1) \quad \tilde{\text{occ}}(\pi, \mu) = \mathbb{P}(\text{ord}(Y_1, \dots, Y_k) \circ \text{ord}(X_1, \dots, X_k)^{-1} = \pi)$$

où les  $(X_i, Y_i)_{1 \leq i \leq k}$  sont des variables i.i.d. suivant la loi de  $\mu$  et si les  $(x_i)_{1 \leq i \leq k}$  sont des réels deux à deux distincts  $\text{ord}(x_1, \dots, x_k)$  est la permutation de  $\llbracket 1, k \rrbracket$  telle que pour tout  $1 \leq i \neq j \leq k$ ,  $\sigma(i) < \sigma(j)$  si et seulement si  $x_i < x_j$ .

Si  $\mu$  est un permuton aléatoire, on étend la définition en disant que conditionnellement à  $\mu$  on a l'équation 1

**Théorème 3.1.** Soit  $(\sigma_n)_{n \geq 1}$  une suite de permutations aléatoires telle que pour tout entier strictement positif  $n$ ,  $\sigma_n$  est une permutation aléatoire de  $\llbracket 1, n \rrbracket$ . Soit  $(I_{n,k})_{n,k \geq 1}$  des variables aléatoires indépendantes et indépendantes de  $(\sigma_n)_{n \geq 1}$  telles que  $I_{n,k}$  est un sous-ensemble aléatoire uniforme à  $k$  éléments de  $\llbracket 1, n \rrbracket$ .

Les propriétés suivantes sont équivalentes :

- (1)  $(\mu_{\sigma_n})_{n \geq 1}$  converge étroitement vers un permuton aléatoire  $\mu$
- (2) Pour tout  $\pi \in \mathfrak{S}$ , il existe  $\Delta_\pi \geq 0$  tel que  $\mathbb{E}[\tilde{\text{occ}}(\pi, \sigma_n)]$  converge vers  $\Delta_\pi$ .

Globalement que veut-dire ce théorème ? Il suffit à priori de connaître l'asymptotique du nombre d'occurrences pour chaque motif, afin de voir s'il y a un objet limite et de le comprendre.

4. DÉCOMPOSITION D'UNE PERMUTATION ET ARBRE ASSOCIÉ

Géométriquement, le permuton associé à une permutations est composé de carrés. Cependant, on peut parfois voir apparaître des blocs dans le dessins de ce permuton, i.e. des intervalles dont l'image est un intervalle.

Sur le premier exemple, on voit que les intervalles  $[1, 4]$  et  $[6, 8]$  ont pour image un intervalle. Certaines permutations jouent alors un rôle spécifique : les permutations dont les seuls intervalles stables sont triviaux, i.e. égaux à  $\{1, n\}$  ou de cardinal au plus 1.

**Définition 4.1.** Soit  $n > 2$ ,  $\sigma \in \mathfrak{S}_n$  : on dit que  $\sigma$  est simple s'il existe par d'intervalle  $I \subset \llbracket 1, n \rrbracket$ , dont l'image par  $\sigma$  est un intervalle, vérifiant  $2 \leq |I| \leq n - 1$

On peut alors essayer de formaliser cette notion de "blocs" par la définition suivante :

**Définition 4.2.** Soit  $d \geq 1$ ,  $\sigma \in \mathfrak{S}_d$  et  $\theta_1, \dots, \theta_d$   $d$  permutations de respectivement  $\llbracket 1, i_1 \rrbracket, \dots, \llbracket 1, i_d \rrbracket$ , on note  $s_p := i_1 + \dots + i_p$  pour  $1 \leq p \leq d$  avec  $s_0 = 0$ , et  $I_p := [s_p + 1, s_{p+1}]$  pour  $0 \leq p \leq d - 1$ . On définit  $\sigma[\theta_1, \dots, \theta_d]$  l'unique permutation telle que la permutation induite sur  $I_p$  est  $\theta_{p+1}$  pour  $0 \leq p \leq d - 1$ , l'image de chaque  $I_p$  est un intervalle, et la permutation induite sur  $s_1, \dots, s_d$  est  $\sigma$

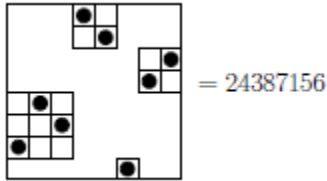


FIGURE 1. Ici on a la décomposition  $24387156 = 2413[132, 21, 1, 12]$

Ainsi, l'observation précédente permet de décomposer itérativement une permutation :

**Théorème 4.1.** Soit  $n \geq 2$ ,  $\sigma \in \mathfrak{S}_n$ ,  $\sigma$  est décomposable de façon unique d'une des trois façons suivantes :

- $\alpha[\theta_1, \dots, \theta_d]$  avec  $\alpha \in \mathfrak{S}_d$  une permutation simple et  $\theta_1, \dots, \theta_d$  des permutations
- $\oplus[\theta_1, \dots, \theta_d]$  avec  $\theta_1, \dots, \theta_d \oplus$  indécomposable
- $\ominus[\theta_1, \dots, \theta_d]$  avec  $\theta_1, \dots, \theta_d \ominus$  indécomposable

En particulier, on peut construire un arbre planaire à partir d'une permutation, en mettant pour noeud  $\oplus, \ominus$  ou une permutation simple et des feuilles non numérotées.

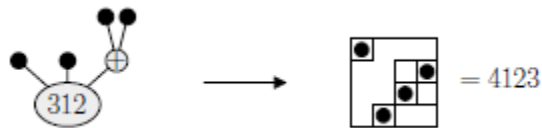


FIGURE 2. Arbre canonique associé à 4123

## 5. COMBINATOIRE ANALYTIQUE

Notons  $\mathcal{T}$  l'ensemble des arbres binaires finis. Compter le nombre d'arbre planaire possédant  $n$  feuilles directement est faisable via des bijections ad hoc, mais rarement adaptable facilement dans des cas plus compliqués (par exemple avec des numérotations des feuilles ou des noeuds ou des contraintes supplémentaires).

Néanmoins, il y a une manière plus robuste d'obtenir de l'information sur le nombre d'arbre de taille  $n$  vérifiant une certaine propriété : les séries génératrices.

**Définition 5.1.** On définit  $T := \sum_{t \in \mathcal{T}} z^{|t|}$  où  $|t|$  est la nombre de feuille de  $t$ .

Or un arbre binaire est très simple à décrire : soit il est constitué d'une feuille, soit d'un sous-arbre binaire gauche et d'un sous-arbre binaire droit.

Ceci se traduit par l'équation suivante :

$$(2) \quad T = z + T^2$$

Couplé à  $T(0) = 0$ , cette équation n'a qu'une solution dans l'anneau des séries formelles.

Il s'avère que cette équation se résout et permet d'obtenir les deux résultats suivants :

**Théorème 5.1.** *La fonction  $T$  a un rayon de convergence  $\frac{1}{4}$ , se prolonge sur  $\mathbb{C}$  privé de  $[\frac{1}{4}, +\infty[$  et vérifie :*

$$T = \frac{1 - \sqrt{1 - 4z}}{2z}$$



pour tout  $z$  dans le domaine défini précédemment.

**Corollaire.** Il y a  $\frac{\binom{2n}{n}}{n+1}$  arbres binaires avec  $n+1$  feuilles pour tout  $n \in \mathbb{N}$ .

Ici, tout est calculable puisqu'on a une équation quadratique en  $T$ . Néanmoins dans de très nombreux cas,  $T$  n'est pas du tout calculable. Mais il est souvent possible de trouver une équation fonctionnelle vérifiée par  $T$  de la forme :

$$(3) \quad T = z + \Lambda(T)$$

avec  $\gamma$  une fonction entière dont le développement est à coefficient positif.

Dans la littérature, il y a déjà beaucoup de théorie sur la résolution d'une telle équation (voir [4] pour des détails). En effet, celle-ci se factorise en une équation de la forme  $\Phi(T) = z$  où  $\Phi(z) = z - \Lambda(z)$  pour tout  $z \in \mathbb{C}$  : c'est donc un problème qui se résume à inverser localement  $\Phi$ .

**Théorème 5.2.** Soit  $\Lambda$  une fonction holomorphe non nulle sur  $\mathbb{C}$ , dont le développement en série entière en 0 n'a que des termes positifs, avec  $\Lambda(0) = \Lambda'(0) = 0$ . Alors l'équation

$$(4) \quad T = \Lambda(T) + z$$

a une unique solution dans l'anneau des séries formelles vérifiant  $T(0) = 0$  notée  $T$ . Ses coefficients sont positifs et  $T$  est holomorphe sur un ouvert connexe contenant 0.

Il s'avère qu'une telle équation permet d'obtenir, sous des hypothèses raisonnables sur  $\Lambda$  qui ne seront pas détaillées ici, le comportement de  $T$ .  $T$  va avoir un rayon fini noté  $R$ , calculable à partir des informations sur  $\Lambda$ .

Son comportement au bord est aussi compréhensible : sous des bonnes hypothèses sur  $\Lambda$ ,  $T$  se prolonge holomorphiquement en tout point du bord, sauf en  $R$ , qui est une singularité d'après le théorème de Pringsheim.

On a même le développement asymptotique suivant :

**Théorème 5.3.** Il existe  $\beta > 0$ ,  $A, B$  deux complexes tel que  $T$  se prolonge holomorphiquement sur  $B(R, \beta)$  privé de  $[R, R + \beta[$  et vérifie le développement asymptotique suivant au voisinage de  $R$  :

$$T(z) = A + B\sqrt{1 - \frac{z}{R}} + \mathcal{O}\left(1 - \frac{z}{R}\right)$$

De plus,  $R$  est calculable explicitement à partir des données du problème.

Il s'avère que pour des fonctions holomorphes, l'asymptotique des coefficients est très lié aux singularités au bord du disque de convergence (voir [5] pour plus de détails). Ainsi on peut obtenir asymptotiquement de l'information sur les coefficients de  $T$ , en exploitant par exemple le théorème suivant.

**Théorème 5.4** (Théorème de transfert). *Soit  $R > 0$  et  $f$  une fonction holomorphe sur un domaine  $\Delta(R)$  vérifiant  $f(z) = o\left(\left(1 - \frac{z}{R}\right)^{-a}\right)$  au voisinage de  $R$ . Alors  $[z^n]f = o\left(\frac{n^{a-1}}{R^n}\right)$ .*

Fixons  $S$  un ensemble de permutations simples, stable par permutation induite, qui ne contient pas toutes les permutations simples (à l'origine seul le cas  $S = \emptyset$  avait été traité dans [3]). On définit alors

$$L(z) = \sum_{n \geq 0} s_n z^n$$

où  $s_n$  est le nombre d'éléments de cardinal  $n$  dans  $S$ .

Il s'avère que les différentes données (nombre de permutations dont l'arbre associé à des noeuds à valeurs dans  $S \cup \{\oplus, \ominus\}$ , nombre d'occurrences moyen) se "calculent" à partir d'équations fonctionnelles avec  $S$  : on est donc capable de calculer la limite, et d'obtenir un résultat limite.

On a ainsi convergence d'une permutation aléatoire uniforme de taille  $n$  dont les noeuds sont dans  $S$  vers le permuton brownien (dont le paramètre dépend de  $S$ ), objet qu'on peut construire à partir de l'excursion brownienne. Cet objet est défini dans [6].

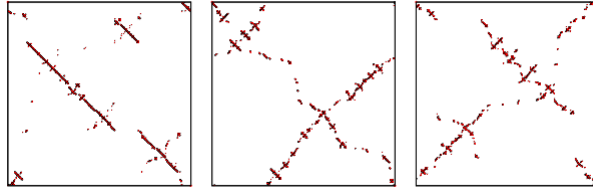


FIGURE 3. Différents permutons browniens pour différentes valeurs du paramètre

## 6. QUELQUES QUESTIONS OUVERTES

L'objectif de ma thèse sera d'étendre ces différentes techniques, pour des graphes aléatoires. La décomposition qui sera utilisée ne sera donc pas la même (ce sera la décomposition modulaire).

Pour l'instant les résultats connus sont uniquement dans le cas où l'ensemble des graphes premiers est vide (donc uniquement avec  $\oplus$  et  $\ominus$ ), ce qui est fait dans [2]. L'objectif est d'étendre cela à des classes plus compliquées, ce qui pourra donner une convergence vers un objet différent.

La combinatoire est différente, puisque comme le graphe est numéroté, les arbres associés deviennent non planaires, d'où l'emploi de séries exponentielles.

## RÉFÉRENCES

- [1] Frédérique Bassino, Mathilde Bouvel, Valentin Féray, Lucas Gerin, Mickaël Maazoun, and Adeline Pierrot. Universal limits of substitution-closed permutation classes, 2017.
- [2] Frédérique Bassino, Mathilde Bouvel, Valentin Féray, Lucas Gerin, Mickaël Maazoun, and Adeline Pierrot. Random cographs : Brownian graphon limit and asymptotic degree distribution, 2021.
- [3] Frédérique Bassino, Mathilde Bouvel, Valentin Féray, Lucas Gerin, and Adeline Pierrot. The brownian limit of separable permutations, 2017.
- [4] Mathilde Bouvel, Marni Mishna, and Cyril Nicaud. Some families of trees arising in permutation analysis, 2016.
- [5] P. Flajolet and R. Sedgewick. *Analytic Combinatorics*. Cambridge University Press, 2009.
- [6] Mickaël Maazoun. On the brownian separable permuton, 2017.