

Mélange de jeu de cartes et phénomène de cut-off

Louis Perot

Yann Adorno

Juin 2024

Sous la direction de Ons Rameh et Alexis Metz-Donnadieu.

Contents

1 Généralités	2
1.1 Introduction	2
2 Chaîne de Markov	3
2.1 Définitions	3
2.2 Théorèmes et propositions sur les chaînes de Markov	5
2.3 Chaînes de Markov et jeu de carte	7
3 Eléments de théorie des représentations	9
3.1 Définitions	9
3.2 Formule de Plancherel	10
3.3 Représentations irréductibles de \mathbb{S}_n	13
3.4 Lemmes utiles pour le théorème	15
4 Preuve des théorèmes	18
4.1 Théorème sur la majoration	18
4.2 Théorème sur la minoration	19



1 Généralités

1.1 Introduction

Notre projet de licence, intitulé "Mélange de jeu de cartes et phénomène de cut-off" est basé sur le travail de Persi Diaconis, dont les premiers travaux remontent à 1974. Fort de son expérience d'ancien magicien et de sa passion pour les mathématiques, il s'est naturellement penché sur l'étude des mélanges de cartes. Ses principaux travaux traitent en particulier du nombre d'itération requis pour obtenir un jeu *bien mélangé*, son résultat le plus connu est que le mélange américain, le plus répandu dans les casinos, nécessite 5 itérations du mélange pour que le jeu soit *bien mélangé*.

Le premier enjeu est d'abord la modélisation mathématique de ce problème. Pour cela, nous numérotions les cartes de un à cinquante deux. De sorte qu'un jeu de cartes sera représenté par une permutation de \mathbb{S}_{52} . La notion de mélange sera traduite par une distribution de probabilités sur l'ensemble des permutations de \mathbb{S}_{52} . La mesure de probabilité associera à chaque permutation sa probabilité d'apparaître, pour chaque itération du mélange. Un paquet sera considéré comme bien mélangé lorsque la mesure de probabilité tout juste décrite sera suffisamment proche de la mesure de probabilité uniforme sur \mathbb{S}_{52} . Ainsi, les mélanges acceptables pour cette définition seront ceux dont la distribution de probabilité tend vers la distribution uniforme lorsque le nombre d'itérations du mélange augmente.

Pour modéliser cet écart entre les deux distributions de probabilité nous devons introduire une distance sur $\mathcal{P}(\mathbb{S}_{52})$. Pour cela nous utiliserons la distance en variation totale :

$$\forall P, Q \in \mathcal{P}(\mathbb{S}_n), \quad \|P - Q\|_{VT} = \sum_{\pi \in \mathbb{S}_n} |P(\pi) - Q(\pi)|,$$

où $\mathcal{P}(\mathbb{S}_n)$ désigne l'ensemble des distributions de probabilité sur \mathbb{S}_n .

L'enjeu est donc ici de déterminer le nombre d'étapes nécessaires pour que la distance de la distribution de probabilités du mélange à la distribution uniforme soit nulle à ϵ près. Afin de ne pas restreindre notre étude seulement aux mélanges de paquets de cinquante deux cartes, nous étudierons par la suite les permutations de \mathbb{S}_n .

Nous nous concentrerons particulièrement sur le mélange suivant : à chaque étape, on tire au hasard deux entiers p_1 et p_2 entre 1 et n avec remise. On échange les cartes numérotées p_1 et p_2 dans la paquet, c'est-à-dire qu'on réalise la transposition $(p_1 p_2)$ dans \mathbb{S}_n . Ainsi, on obtient la distribution de probabilité suivante sur \mathbb{S}_n :

$$T(\sigma) = \begin{cases} \frac{1}{n} & \text{si } \sigma = \text{Id} \\ \frac{2}{n^2} & \text{si } \sigma \text{ est une transposition} \end{cases}$$

La distribution du mélange à l'ordre k correspond à la convolution itérée k fois de cette distribution. Plus précisément, on s'intéresse à la suite de variables aléatoires $(X_k^{(n)})_{k \geq 1}$:

$$\begin{cases} X_1^{(n)} = \text{Id}_{\mathbb{S}_n}, \\ X_{k+1}^{(n)} = \sigma^{(k)} \circ X_k^{(n)}. \end{cases}$$

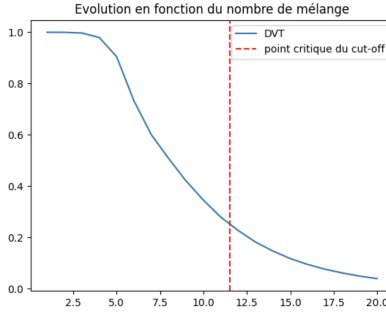
Où $\sigma^{(k)}$ à la loi T et est indépendant de $X_k^{(n)}$.

Dans le cas de ce mélange nous étudierons donc : $C_k = \|T^{*k} - U\|_{VT}$, où T^{*k} est la loi de $X_k^{(n)}$.

L'étude de cette distance fait alors apparaître un phénomène propre à certaines chaînes de Markov, le *cut-off* :

Les premières étapes n'ont presque aucun impact sur la distance de variation totale à la distribution uniforme, puis, en quelques étapes la distance de variation totale diminue drastiquement, pour ensuite stagner proche de 0. Le point d'inflexion de la courbe est appelé point critique.

Appliqué à notre sujet d'étude on observe que le paquet de cartes reste *mal mélangé*, pendant les premières étapes, puis en quelques étapes, devient "presque parfaitement mélangé". Ainsi, l'enjeu de ce mémoire est d'abord de mettre en évidence l'existence de ce phénomène de *cut-off* pour le mélange que l'on étudie, puis, de déterminer le point critique. Dans un premier temps, nous étudierons la distribution de probabilité sur \mathbb{S}_n du point de vue des chaînes de Markov dans la section 2 et nous obtiendrons une approximation grossière du point critique en $O(n^2)$. Dans un deuxième temps, avec une étude plus fine, utilisant la structure algébrique de \mathbb{S}_n et de ses représentations irréductibles, nous obtiendrons une estimation plus précise du point critique en $\frac{1}{2}n \log(n)$. Cette estimation constitue les théorèmes 1 et 2 dont les preuves, réalisées par Diaconis et Shahshahani en 1981, font l'objet des section 3 et 4.



Evolution de la distance en variation totale en fonction du nombre de mélanges

Ainsi, le premier théorème ci-dessous nous donne une majoration sur C_k pour k assez grand, le deuxième théorème nous donne une minoration sur C_k si k est assez petit.

Théorème I : Majoration

Supposons $n \geq 10$ et posons $c = \frac{k - \frac{n \log(n)}{2}}{n}$, si $c > 0$, alors il existe b tel que $C_k \leq be^{-2c}$

Comme première application de ce théorème, on peut donner le nombre de fois qu'il faut transposer deux cartes d'un jeu initialement connu pour que les cartes soient bien mélangées. En considérant un jeu classique à 52 cartes, on obtient $k = \lceil \frac{52 \log(52)}{2} \rceil = 103$, il faut donc au moins 103 transpositions pour que le jeu soit suffisamment mélangé.

Théorème II: Minoration

Supposons que n tende vers l'infini alors $C_k \geq 2(\frac{1}{e} - e^{-e^{-2c}}) + o(1)$.

On verra plus tard que ces deux théorèmes impliquent un phénomène de cut-off qui n'est pas totalement complet. En effet, ils induisent l'équivalence à l'infini des temps de mélange pour des valeurs comprises entre 0 et $\frac{2}{e}$, alors que le phénomène de cut-off habituel concerne toutes les valeurs comprises entre 0 et 1.

2 Chaîne de Markov

Dans cette partie on s'intéresse à l'étude probabiliste de notre problème en développant des outils nouveaux pour mener à bien notre étude. L'objectif est de démontrer la convergence de notre distribution convolée vers la distribution uniforme puis d'estimer cette vitesse de convergence.

2.1 Définitions

2.1.1 Chaîne de Markov

Une chaîne de Markov est la donnée d'une suite de variable aléatoire \mathbf{X} à valeurs dans \mathcal{X} sur un espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$ et d'un triplet (\mathcal{X}, P, ν) défini de la manière suivante :

- \mathcal{X} est l'espace d'états dans notre cas il est fini et est égale à \mathbb{S}_n ,
- P est un noyau de transition, $P : \mathcal{X} \times \mathcal{X} \rightarrow [0, 1]$ et vérifie

$$\forall x \in \mathcal{X}, \quad \sum_{y \in \mathcal{X}} P(x, y) = 1.$$

,

- ν est la loi initiale, $\nu : \mathcal{X} \rightarrow [0, 1]$ et vérifie :

$$\sum_{x \in \mathcal{X}} \nu(x) = 1.$$

,

De plus, la suite de variables aléatoires doit vérifier la propriété suivante :

$$\forall t \in \mathbb{N}, \forall (x_0, \dots, x_t) \in \mathcal{X}^{t+1}, \quad \mathbb{P}(X_0 = x_0, \dots, X_t = x_t) = \nu(x_0)P(x_0, x_1) \dots P(x_{t-1}, x_t).$$

(On notera, $\mathbf{X} \sim MC(\mathcal{X}, P, \nu)$.)

2.1.2 Irréductibilité et Ergodicité

On dit qu'une chaîne de Markov \mathbf{X} est irréductible si elle vérifie :

$$\forall (x, y) \in \mathcal{X} \times \mathcal{X}, \quad \exists t \in \mathbb{N}, \quad P^t(x, y) > 0.$$

On dit que X est ergodique si elle vérifie :

$$\exists t \in \mathbb{N}, \quad \forall (x, y) \in \mathcal{X} \times \mathcal{X}, \quad P^t(x, y) > 0.$$

L'irréductibilité affirme qu'en partant de n'importe quel point on peut arriver à n'importe quel autre point en un nombre fini de pas. L'ergodicité, quant elle, donne qu'il existe un nombre de pas fini tels qu'en partant de n'importe quel point on peut arriver à tous les autres points.

2.1.3 Loi stationnaire

On dit qu'une mesure μ sur \mathcal{X} est une loi stationnaire pour le noyau de transition P si :

- $\mu = \mu P$,
- $\sum_{x \in \mathcal{X}} \mu(x) = 1$.

2.1.4 Temps de mélange

On appelle le temps de mélange de précision $\epsilon \in (0, 1)$ de P :

$$t_{MIX}^{(\epsilon)}(P) = \min\{t \in \mathbb{N} : \max_{x \in \mathcal{X}} \|P^t(x, \cdot) - U\|_{VT} \leq \epsilon\}.$$

2.1.5 Phénomène de *Cut-off*

On dit qu'une suite de noyau de transition (P_n) exhibe un phénomène de *cut-off* si :

$$\forall \epsilon \in (0, 1), \quad t_{MIX}^{(\epsilon)}(P_n) \underset{n}{\sim} t_{MIX}^{(\frac{1}{4})}(P_n).$$

2.1.6 Lien entre le phénomène de *cut-off* et Théorèmes I et II

Maintenant que le temps de mélange a été défini, on peut expliciter le lien entre nos théorèmes et le *cut-off*.

Les théorèmes nous donnent des bornes sur la distance en variation totale, quand n tend vers l'infini on a $be^{-2c} \geq C_k \geq 2(\frac{1}{e} - e^{-e^{-2c}}) + o(1)$ où b est une certaine constante et c est défini dans l'introduction (et dépend de n et k). Ainsi pour n assez grand,

$$\forall k < k_\epsilon^- \quad C_k \geq \frac{2}{e} \quad \text{et} \quad \forall k \geq k_\epsilon^+ \quad C_k \leq be^{-\epsilon \log(n)},$$

$$\text{où } k_\epsilon^\pm = \frac{n \log(n) \pm \epsilon n \log(n)}{2}.$$

Donc si on prend un temps de mélange avec $\epsilon < \frac{2}{e}$, on choisit n assez grand tel que $be^{-\epsilon \log(n)} \leq \min(\frac{1}{4}, \epsilon)$ alors on sait que les temps de mélange seront atteints dans la fenêtre, or les deux bornes de la fenêtre sont des équivalents de $n \log(n)$ donc on a bien que les temps de mélange sont des équivalents à l'infini de $n \log(n)$. Ainsi on a bien un phénomène de *cut-off* incomplet avec ces bornes. Dans le cas où $\epsilon \geq \frac{2}{e}$ on ne peut pas garantir que le temps de mélange associé à ϵ est dans la fenêtre k_ϵ^\pm , donc le même raisonnement n'est pas faisable. Toutefois une forme plus précise du théorème de minoration existe, dans laquelle la constante devant la parenthèse est e et non 2 ainsi on peut répéter l'argument précédent pour ce nouveau résultat.

2.1.7 Couplage Markovien

Soit $(X_t)_{t=0}^\infty, (Y_t)_{t=0}^\infty$ deux chaînes de Markov avec un noyau de transition P , on dit que $(X_t, Y_t)_{t=0}^\infty$ est un couplage markovien si c'est une chaîne de Markov avec l'espace d'états $\mathcal{X} \times \mathcal{X}$ et si elle vérifie la propriété suivante :

$\forall x, x', y, y'$:

$$\mathbb{P}(X_{t+1} = x' | X_t = x, Y_t = y) = P(x, x') \quad \text{et} \quad \mathbb{P}(Y_{t+1} = y' | X_t = x, Y_t = y) = P(y, y').$$

2.2 Théorèmes et propositions sur les chaînes de Markov

Dans cette partie on joue avec les concepts mathématiques introduits dans la section précédente dans le but de se familiariser avec ces nouvelles notions et de développer des outils d'analyse utiles pour notre problème.

2.2.1 Proposition

Tout noyau de transition P a une loi stationnaire et si P est irréductible alors cette loi stationnaire est unique.

Preuve :

Avant de plonger dans les détails, on note μ_t la loi de chaque variable $X_t, t \geq 1$, on sait par définition des chaînes de Markov que :

$$\mu_t(y) = \mathbb{P}(X_t = y) = \sum_{x \in \mathcal{X}} \mathbb{P}(X_{t-1} = x)P(x, y) = \sum_{x \in \mathcal{X}} \mu_{t-1}(x)P(x, y).$$

Pour le cas $t = 0$ on a $\mu_0(y) = \nu(y)$. Soit $t \geq 1$, on pose $\pi_t = \frac{1}{t} \sum_{i=0}^{t-1} \mu_i$, π_t est une suite de $\mathcal{P}(\mathcal{X})$ qui est compact. On sait qu'il existe une sous-suite qui converge vers un certain élément $\pi \in \mathcal{P}(\mathcal{X})$.

De plus, $\forall x \in \mathcal{X}$ on a :

$$(\pi_t P)(x) - \pi_t(x) = \frac{1}{t} \sum_{i=0}^{t-1} (\mu_i P)(x) - \frac{1}{t} \sum_{i=0}^{t-1} \mu_i(x) = \frac{\mu_{t+1}(x) - \mu_0(x)}{t} \xrightarrow{t \rightarrow +\infty} 0.$$

Ainsi, la mesure de probabilité π est stationnaire.

Prouvons maintenant l'unicité en cas d'irréductibilité de P .

Tout d'abord, on montre que chaque loi stationnaire a un support égal à l'espace d'états. Soit $x \in \mathcal{X}$, on sait qu'il existe x_0 tel que $\pi(x_0) > 0$, car π est une mesure de probabilité discrète.

Par irréductibilité de P , il existe un certain t entier tel que $P^t(x_0, x) > 0$. Ainsi $\pi(x) = (\pi P)(x) = (\pi P^t)(x) = \sum_{y \in \mathcal{X}} \pi(y)P^t(y, x) \geq \pi(x_0)P^t(x_0, x) > 0$. Donc x est dans le support de π . Comme le choix de x est arbitraire le support de π est tout \mathcal{X} .

Supposons par l'absurde qu'il existe deux loi stationnaire distinctes π_1 et π_2 , on pose alors la fonction :

$$m : [0, 1] \rightarrow [-1, 1]$$

$$\theta \rightarrow \min_{x \in \mathcal{X}} (\pi_1(x) - \theta \pi_2(x)).$$

On sait que m est continue, $m(0) = \min_{x \in \mathcal{X}} (\pi_1(x)) > 0$ car $\text{supp}(\pi_1) = \mathcal{X}$ et $m(1) = \min_{x \in \mathcal{X}} (\pi_1(x) - \pi_2(x)) < 0$, car $\pi_1 \neq \pi_2$, donc par le TVI il existe $\theta_0 \in (0, 1)$ tel que $m(\theta_0) = 0$.

Ainsi $m(\theta_0)$ donne lieu a une nouvelle loi stationnaire, $\pi^* = \frac{\pi_1(x) - \theta_0 \pi_2(x)}{1 - \theta_0}$ qui vérifie bien l'égalité $\pi = \pi P$ car les mesures π_1 et π_2 la vérifie. De plus, comme le minimum est atteint on sait qu'il existe un certain x_0 tel que $\pi^*(x_0) = 0$ or on a montré que toutes les lois stationnaires de P avait forcément un support complet si P était irréductible on obtient donc une contradiction et la loi stationnaire est unique. \square

2.2.2 Théorème

Si P est un noyau de transition ergodique alors,

$$\forall y \in \mathcal{X}, \quad \mathbb{P}(X_t = y) \xrightarrow{t \rightarrow +\infty} \pi(y).$$

Ce théorème nous donne une convergence pour la suite de variable aléatoire de la chaîne de Markov qui est très pertinente dans le cadre de notre problème. On va le découvrir dans la section suivante. Nous avons décidé de ne pas mettre de preuve, toutefois dans [4] Salez le montre.

2.2.3 Propriété (*)

Soit $(X_t, Y_t)_{t=1}^{\infty}$ un P -couplage Markovien on dit qu'il vérifie la propriété (*) si la matrice P vérifie la propriété suivante :

$$\text{si } X_s = Y_s, \quad \text{alors } X_t = Y_t \quad \forall t \geq s.$$

2.2.4 Lien entre couplage et temps de mélange

Soit P une chaîne de Markov irréductible et apériodique qui converge vers sa loi stationnaire π alors :

$$\bar{d}(t) \geq d(t),$$

où $d(t) := \max_{x \in \mathcal{X}} \|P^t(x, \cdot) - \pi\|_{VT}$ et $\bar{d}(t) := \max_{x \in \mathcal{X}} \|P^t(x, \cdot) - P^t(y, \cdot)\|_{VT}$.

Preuve :

Rappelons que: $\forall A \subset \mathcal{X} \quad \pi(A) = \sum_{y \in \mathcal{X}} \pi(y) P^t(y, A)$.

Alors, pour tout A inclus dans \mathcal{X} ,

$$|P^t(x, A) - \pi(A)| \leq \sum_{y \in \mathcal{X}} |\pi(y)[P^t(x, A) - P^t(y, A)]| \leq \sum_{y \in \mathcal{X}} \pi(y) \|P^t(x, \cdot) - P^t(y, \cdot)\|_{VT} \leq \bar{d}(t).$$

En prenant le max sur les A on trouve $d(t) \leq \bar{d}(t)$ □

2.2.5 Propriété d'un couplage Markovien (*)

Soit $(X_t, Y_t)_{t=1}^{\infty}$ un P -couplage Markovien vérifiant (*) et soit (x, y) le point de départ du processus alors :

$$\|P^t(x, \cdot) - P^t(y, \cdot)\|_{VT} \leq \mathbb{P}_{x,y}(\tau_* > t),$$

où $\tau_* := \min\{t \in \mathbb{N} : X_s = Y_s \quad \forall s \geq t\}$.

Preuve :

On sait que:

$$\mathbb{P}_{x,y}(X_t = z) = P^t(x, z) \quad \text{et} \quad \mathbb{P}_{x,y}(Y_t = z) = P^t(y, z),$$

Or $\forall A \subset \mathcal{X}$,

$$\mathbb{P}_{x,y}\{X_t \in A\} - \mathbb{P}_{x,y}\{Y_t \in A\} \leq \mathbb{P}_{x,y}\{X_t \in A, Y_t \notin A\} \leq \mathbb{P}_{x,y}\{X_t \neq Y_t\}.$$

Donc $\max_{A \subset \mathcal{X}} |\mathbb{P}_{x,y}(X_t \in A) - \mathbb{P}_{x,y}(Y_t \in A)| \leq \mathbb{P}_{x,y}\{X_t \neq Y_t\}$.

Ainsi $\|P^t(x, \cdot) - P^t(y, \cdot)\|_{VT} \leq \mathbb{P}_{x,y}\{X_t \neq Y_t\} \underset{\text{par (*)}}{\leq} \mathbb{P}_{x,y}(\tau_* > t)$. □

2.2.6 Proposition

Soit $(X_t, Y_t)_{t=1}^{\infty}$ un P -couplage Markovien vérifiant (*) et soit (x, y) le point de départ du processus et τ_* défini précédemment alors:

$$t_{MIX}^{(\frac{1}{4})} \leq 4 \max_{x,y} \mathbb{E}_{x,y}(\tau_*).$$

Preuve :

$$t_{MIX} = \min \left\{ t \in \mathbb{N} : d(t) \leq \frac{1}{4} \right\} \leq \min \left\{ t \in \mathbb{N} : \max_{x,y} \mathbb{P}_{x,y}(\tau_* > t) \leq \frac{1}{4} \right\} \leq \min \left\{ t \in \mathbb{N} : \max_{x,y} 4\mathbb{E}(\tau_*) \leq t \right\} = 4 \max_{x,y} \mathbb{E}(\tau_*),$$

où l'on a utilisé les deux propositions précédentes et l'inégalité de Markov. □

2.3 Chaînes de Markov et jeu de carte

On pose $\mathbf{X}^{(n)} = (X_1^{(n)}, X_2^{(n)}, \dots)$ où $X_i^{(n)}$ est la permutation obtenue après i mélanges avec un jeu de n cartes. Donc la distribution de $X_i^{(n)}$ est donnée par $T_{(n)}^{*i}$ qui est la convolution i fois de la distribution d'un seul mélange dans un jeu à n cartes.

Ainsi le noyau de transition est une matrice de dimension $n! \times n!$:

$$P(\sigma, \tau) = \begin{cases} 1/n & \text{si } \sigma = \tau \\ 2/n^2 & \text{si il existe une transposition } \pi \text{ tel que } \pi\sigma = \tau \\ 0 & \text{sinon} \end{cases}$$

Montrons que le noyau de transition P définit une chaîne de Markov ergodique:

On commence par montrer qu'elle est irréductible, soit τ et σ deux permutations dans \mathbb{S}_n alors on peut les décomposer en produit de transpositions :

$$\tau = \prod_{i=1}^r \tau_i^{m_i} \quad \text{et} \quad \sigma = \prod_{i=1}^{r'} \sigma_i^{m'_i}.$$

Il y a une probabilité positive de tirer chacune des transpositions présentes dans la décomposition des deux permutations.

$$P(\sum_{i=1}^r m_i)(\tau, Id) \geq P(\tau, \tau\tau_n)P(\tau\tau_n, \tau\tau_n^2) \dots P(\tau\tau_n^{m_n-1}, \tau\tau_n^{m_n}) \dots P(\tau\tau_n^{m_n} \dots \tau_1^{m_1-1}, \underbrace{\tau\tau_n^{m_n} \dots \tau_1^{m_1}}_{=Id}) > 0,$$

$$P(\sum_{i=1}^{r'} m'_i)(Id, \sigma) \geq P(Id, \sigma_1)P(\sigma_1, \sigma_1^2) \dots P(\sigma_1^{m'_1-1}, \sigma_1^{m'_1}) \dots P(\sigma_1^{m'_1} \dots \sigma_n^{m'_n-1}, \underbrace{\sigma_1^{m'_1} \dots \sigma_n^{m'_n}}_{=\sigma}) > 0,$$

Ainsi,

$$P(\sum_{i=1}^r m_i + \sum_{i=1}^{r'} m'_i)(\tau, \sigma) \geq P(\sum_{i=1}^r m_i)(\tau, Id)P(\sum_{i=1}^{r'} m'_i)(Id, \sigma) > 0.$$

Donc P est irréductible. On a un noyau de transition irréductible, on peut donc définir le temps $t(x, y)$ comme le temps minimal tel qu'on puisse aller de x à y avec une probabilité strictement positive en $t(x, y)$ étapes. De plus, pour tout état x il y a une probabilité positive de rester au même endroit. Soit $t^* := \max_{(x,y)} t(x, y)$.

$$\forall (x, y) \in \mathcal{X} \times \mathcal{X}, \quad P^{t^*}(x, y) \geq P^{t(x,y)}(x, y)P^{t^*-t(x,y)}(y, y) > 0.$$

Donc P est ergodique, et par le théorème 2.2.2 on peut établir :

$$P^t(x, \cdot) \xrightarrow[t \rightarrow \infty]{} \pi(\cdot),$$

où π est la loi stationnaire.

On montre que la mesure uniforme U est une loi stationnaire pour le noyau de transition P , on note $\forall x \in \mathbb{S}_n$:

$$U(x) = \frac{1}{n!} \quad \text{et} \quad UP(x) = \sum_{z \in \mathbb{S}_n} U(x)P(z, x) = \frac{1}{n!} \sum_{z \in \mathbb{S}_n} P(z, x) = \frac{1}{n!}.$$

Donc U est une loi stationnaire pour notre noyau de transition, or on sait que cette loi stationnaire est unique dans le cadre de noyaux irréductibles donc on a

$$P^t(x, \cdot) \xrightarrow[t \rightarrow \infty]{VT} U(\cdot).$$

Maintenant que l'on a montré la convergence vers la loi uniforme on s'intéresse à la vitesse de convergence, pour cela on utilise les liens faits précédemment entre les couplages et la distance en variation totale. Ici, on crée un couplage P -Markovien (*) de toute pièce qui va nous permettre de borner le temps de mélange.

On considère X_t une variable aléatoire qui tire uniformément une carte dans le paquet entre 1 et n , et une variable aléatoire Y_t qui tire une position de manière uniforme entre 1 et n .

Soit D1 et D2 deux paquets mélangés.

On souhaite coupler le mélange des deux paquets. On dénote par a_t le nombre de cartes qui sont à la même place dans les deux paquets à l'instant t . On échange simultanément et dans chaque paquet, la carte numéro X_t et la carte de position Y_t .

A l'instant $t + 1$ il y a plusieurs possibilités :

- Les cartes X_t sont à la même place dans les deux paquets. Dans ce cas, en tirant la position Y_t , on ne va pas créer de nouvelles liaisons entre les deux paquets, mais on ne va pas en supprimer non plus. Alors $a_t = a_{t+1}$. Par exemple,

$$\begin{array}{cccc} 1 & 2 & 3 & 4 \\ 2 & 4 & 3 & 1 \end{array} \implies \begin{array}{cccc} 3 & 2 & 1 & 4 \\ 3 & 4 & 2 & 1 \end{array}, \text{ où } Y_t = 1 \text{ et } X_t = 3.$$

- Si les cartes X_t sont à des places différentes et qu'on tire une position Y_t sur lesquelles les cartes coïncident dans les deux paquets, alors on va briser la connexion à la position Y_t mais elle va être remplacée par une connexion entre les cartes de valeur X_t . Donc $a_t = a_{t+1}$. Par exemple,

$$\begin{array}{cccc} 1 & 2 & 3 & 4 \\ 2 & 4 & 3 & 1 \end{array} \implies \begin{array}{cccc} 3 & 2 & 1 & 4 \\ 2 & 4 & 1 & 3 \end{array}, \text{ où } Y_t = 3 \text{ et } X_t = 1.$$

- Si les cartes X_t ne sont pas à la même place, mais que les cartes à la position Y_t ne coïncident pas non plus, alors on crée au minimum la connexion entre les cartes de valeur X_t . On peut créer jusqu'à deux connexions en plus de la première :

Cas où une seule connexion se forme $\begin{array}{cccc} 1 & 2 & 3 & 4 \\ 2 & 4 & 1 & 3 \end{array} \implies \begin{array}{cccc} 1 & 2 & 3 & 4 \\ 1 & 4 & 2 & 3 \end{array}, \text{ où } Y_t = 1 \text{ et } X_t = 1.$

– si la carte dans le paquet 1 à la position initiale Y_t coïncide avec la carte dans le paquet 2 à la position initiale de X_t dans le paquet 1

– et inversement :

Cas où deux connexions se forment $\begin{array}{cccc} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \end{array} \implies \begin{array}{cccc} 1 & 2 & 3 & 4 \\ 1 & 4 & 2 & 3 \end{array}, \text{ où } Y_t = 1 \text{ et } X_t = 1.$

Cas où trois connexions se forment $\begin{array}{cccc} 2 & 3 & 1 & 4 \\ 3 & 1 & 2 & 4 \end{array} \implies \begin{array}{cccc} 1 & 3 & 2 & 4 \\ 1 & 3 & 2 & 4 \end{array}, \text{ où } Y_t = 1 \text{ et } X_t = 1.$

2.3.1 Proposition

Soit τ le temps auquel les paquets coïncident alors :

$$\mathbb{E}(\tau) \leq \frac{(\pi n)^2}{6}.$$

Preuve :

On considère la décomposition de τ en une somme de τ_i où chaque τ_i représente le nombre de transpositions dont on a besoin pour passer de $i-1$ cartes qui coïncident à i cartes qui coïncident. En sachant qu'à l'instant 0, il est possible que des cartes coïncident déjà et que le nombre de cartes coïncidentes peut augmenter de deux ou trois par transpositions. Alors il est possible que certains τ_i soit nuls.

On commence par déterminer la loi des τ_{i+1} , lorsque $a_t = i$, alors pour que a_{t+1} soit supérieur à i il faut tirer une carte qui ne coïncide pas dans les deux paquets. Il y a $n - i$ possibilités, et une position de carte où il n'y a pas déjà de connexion, $n - i$ possibilités, donc avec probabilité $\frac{(n-i)^2}{n^2}$ et on recommence jusqu'à ce que la transposition vérifie une nouvelle connexion. Donc on a une variable aléatoire géométrique de paramètre $\frac{(n-i)^2}{n^2}$, donc $\mathbb{E}(\tau_{i+1}|a_t = i) = \frac{n^2}{(n-i)^2}$. On

remarque que $\tau_i = \tau_i \mathbf{1}_{a_{\tau_1+\dots+\tau_{i-1}}=i-1}$, ainsi :

$$\mathbb{E}(\tau) = \sum_{i=1}^n \mathbb{E}(\tau_i) = \sum_{i=1}^n \mathbb{E}(\tau_i \mathbf{1}_{a_{\tau_1+\dots+\tau_{i-1}}=i-1}) < \sum_{i=1}^n \mathbb{E}(\tau_i | a_{\tau_1+\dots+\tau_{i-1}} = i-1) = \sum_{i=1}^n \frac{n^2}{(n-i)^2} < \frac{(n\pi)^2}{6}.$$

En utilisant ce résultat, le fait que $(X_t, Y_t)_{t=0}^\infty$ défini un couplage Markovien et la propriété des couplages Markoviens * on obtient que le temps de mélange du jeu de carte vérifie :

$$t_{MIX} \leq 4 \max_{x,y} \mathbb{E}(\tau_*) \leq \frac{4(\pi n)^2}{6}.$$

3 Eléments de théorie des représentations

La partie 2 sur les chaînes de Markov nous a permis de comprendre plus finement les mécanismes qui sont en jeu dans le mélange de cartes présenté par Diaconis et Shahshahani. Toutefois, dans une volonté de mettre en lumière le phénomène de cut-off apparaissant durant le mélange, la borne en $O(n^2)$ obtenue par l'étude des chaînes de Markov n'est pas suffisante. Il est donc nécessaire d'utiliser un point de vue différent pour aborder ce problème.

Pour cela, nous allons étudier les représentations irréductibles de \hat{S}_n afin de donner une preuve algébrique du théorème de majoration. Nous nous servirons d'abord de la formule de Plancherel (partie 3.2) pour faire la bascule entre les deux points de vue. Puis grâce à l'équivalence entre la considération des tableaux d'Young et des représentations irréductibles de \hat{S}_n que nous prouverons (partie 3.3), nous pourrons établir des propriétés sur les représentations irréductibles (partie 3.4) pour enfin aboutir par la preuve du théorème de majoration (partie 4.1).

3.1 Définitions

Il s'agit dans un premier temps de définir les outils qui nous seront utiles pour la suite de notre étude et de rappeler quelques propriétés fondamentales.

3.1.1 Représentation d'un groupe

Une représentation d'un groupe G sur un Espace vectoriel V est un homomorphisme de groupe de G dans le groupe linéaire $GL(V)$. C'est-à-dire, c'est une application $\rho : G \rightarrow GL(V)$ qui vérifie les propriétés suivantes :

- Pour tout élément g de G , $\rho(g)$ est une transformation linéaire de V .
- Pour tous g, h dans G , $\rho(gh) = \rho(g)\rho(h)$.

3.1.2 Dimension d'une représentation

La dimension d'une représentation $\rho : G \rightarrow GL(V)$ est la dimension de l'espace vectoriel V . Si V est de dimension finie n , alors la représentation est dite de dimension n .

3.1.3 Propriété universelle de l'algèbre de groupe

Il est équivalent de se donner une représentation \mathbb{C} -linéaire du groupe G et un $\mathbb{C}[G]$ -module.

Si on note (ρ, V) une représentation \mathbb{C} -linéaire de G , il existe une unique structure de $\mathbb{C}[G]$ -module, appelée $\mathbb{C}[G]$ -module associé, vérifiant :

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

3.1.4 Remarque :

L'équivalence entre les deux points de vue est essentielle pour la preuve du théorème de majoration. Elle permet notamment d'utiliser les tableaux d'Young (cf partie 3.3), pour la suite on confondra régulièrement les deux visions.

3.1.5 Représentation irréductible

Une représentation $\rho : G \rightarrow GL(V)$ est dite irréductible si V n'admet aucun sous-espace vectoriel W propre stable par les $\rho(g)$, autre que 0 et V lui-même. Autrement dit, il n'existe aucun sous-espace vectoriel invariant non trivial de V sous l'action de ρ .

3.1.6 Représentations équivalentes

Deux représentations $\rho_1 : G \rightarrow GL(V_1)$ et $\rho_2 : G \rightarrow GL(V_2)$ d'un même groupe G sont dites équivalentes si elles sont "semblables" d'un point de vue structurel. Formellement, il existe un isomorphisme linéaire $T : V_1 \rightarrow V_2$ tel que pour tout g dans G , $\rho_2(g) = T \circ \rho_1(g) \circ T^{-1}$. Cela signifie que les actions des éléments du groupe sur les espaces vectoriels sont essentiellement les mêmes, à une transformation linéaire près.

3.1.7 Caractère d'une représentation d'un groupe fini

On appelle χ_ρ le caractère de la représentation (V, ρ) de G un groupe fini et V un \mathbb{C} espace vectoriel, l'application suivante :

$$\begin{aligned} \chi_\rho : G &\rightarrow \mathbb{C} \\ \gamma &\rightarrow \text{Tr}(\rho(\gamma)). \end{aligned}$$

3.1.8 Formule de Frobenius

Soit k le nombre de classes de conjugaisons du groupe fini G , on a :

- (i) A isomorphisme près il existe exactement k représentations irréductibles de G ,
- (ii) Leurs dimensions d_1, \dots, d_k vérifient :

$$|G| = \sum_{i=1}^n d_i^2.$$

3.2 Formule de Plancherel

L'égalité de Plancherel est essentielle dans notre preuve du théorème de majoration car elle permet de faire le lien entre la sommation probabiliste et la sommation sur les représentations de \mathbb{S}_n .

3.2.1 Formule de Plancherel

Soit G un groupe fini et P une fonction sur G ,

$$\sum_{\eta \in G} |P(\eta)|^2 = \frac{1}{|G|} \sum_{\rho \in \hat{G}} d_\rho \text{Tr}[\rho(P)\rho(P)^*].$$

Où la somme de droite est prise sur toutes les classes d'équivalence de représentations irréductibles et où l'adjonction $\rho(P)^*$ est définie pour un produit scalaire sur V_ρ rendant les ρ_g orthogonaux

3.2.2 Remarque :

En théorie des probabilités, cette égalité peut être interprétée comme suit : si G est le groupe symétrique à n éléments, alors on peut prendre une fonction P qui est une mesure de probabilités sur \mathbb{S}_n , et l'égalité de Plancherel permet de relier la norme L^2 de P à une quantité purement algébrique.

Avant d'aborder la preuve de la formule, il est nécessaire d'introduire certaines notions et de démontrer certaines propriétés.

3.2.3 Représentation de permutations, représentation régulière

Soit G un groupe fini. On suppose que G agit linéairement sur un ensemble fini X . On appelle $V = \mathbb{C}X$ la représentation de permutation associée telle que :

$$\forall g \in G, \quad \forall x \in X, \quad \rho_g(e_x) = e_{g \cdot x}.$$

La représentation régulière est la représentation de permutation associée au cas $X = G$.

3.2.4 Lemme sur le caractère d'une représentation régulière

Soit χ_{reg} la représentation de la permutation régulière, elle vérifie :

- $\chi_{reg}(1) = |G|$,
- $\forall s \neq 1, \chi_{reg}(s) = 0$.

Ce résultat est montré par Diaconis dans [6].

3.2.5 Corollaire

Toute représentation irréductible W est incluse dans une sous-représentation régulière.

De plus, on a :

$$d_W = (\chi_{reg} | \chi_W).$$

On a également, $\forall g \in G, g \neq 1$:

$$\sum_{\rho \in \hat{G}} d_\rho \chi_\rho(g) = 0.$$

3.2.6 Remarque :

La deuxième partie de la formule de Frobenius est également un corollaire de cette propriété.

On peut désormais définir l'analogie de la transformation de Fourier, qui apparaît dans la formule de Plancherel.

3.2.7 Analogie de la transformation de Fourier

Soit f une fonction sur G et ρ une représentation, $\rho : G \rightarrow V$

On note $\rho(f)$ l'élément de $GL(V)$:

$$\rho(f) = \sum_{\eta \in G} f(\eta) \rho(\eta).$$

3.2.8 Transformée de Fourier d'une convolution

Soit f_1 et f_2 deux fonctions sur G et ρ une représentation. On a:

$$\rho(f_1 * f_2) = \rho(f_1) \rho(f_2).$$

Avec "*" la convolée de f_1 et f_2 :

$$\forall g \in G, f_1 * f_2(g) = \sum_{h \in G} f_2(gh^{-1}) f_1(h).$$

Maintenant que l'on a posé le cadre théorique pour comprendre les outils utiles à la formule de Plancherel, on peut énoncer le lemme qui nous permet de montrer la formule.

3.2.9 Formule d'inversion de Fourier

Soit f une fonction sur G . On a :

$$\forall g \in G, f(g) = \frac{1}{|G|} \sum_{\rho \in \hat{G}} d_\rho \text{Tr}(\rho(g^{-1}) \rho(f)).$$

Avec d_ρ la dimension de la représentation ρ .

Preuve :

Les deux côtés de l'égalité sont linéaires en f . On peut donc vérifier l'égalité seulement avec les fonctions $f(g) = \delta_{gt}$, vecteurs de la base. La fonction δ_{gt} vaut 1 si $t = g$ et est nulle sinon.

Avec un tel f on a :

$$\forall \rho \in \hat{G} : \rho(f) = \rho(t).$$

Le côté droit de l'égalité vaut alors :

$$\frac{1}{|G|} \sum_{\rho \in \hat{G}} d_\rho \chi_\rho(g^{-1}t).$$

En utilisant le corollaire 3.2.5, la somme vaut 1 si $g = t$ et 0 sinon, ainsi :

$$\frac{1}{|G|} \sum_{\rho \in \hat{G}} d_\rho \chi_\rho(gt^{-1}) = \delta_{st}.$$

□

Nous pouvons désormais faire la preuve de la formule de Plancherel.

Preuve de la formule de Plancherel

Montrons de manière plus générale la formule suivante, dont la formule de Plancherel est un cas particulier :

$$\sum_{g \in G} f(g^{-1})h(g) = \frac{1}{|G|} \sum_{\rho \in \hat{G}} d_\rho \text{Tr}(\rho(f)\rho(h))$$

Les deux côtés de l'égalité sont linéaires en f . On peut donc vérifier l'égalité seulement avec les fonctions $f(t) = \delta_{gt}$, vecteurs de la base. La fonction δ_{gt} vaut 1 si $t = g$ et est nulle sinon.

Le théorème d'inversion de Fourier donne que :

$$h(t^{-1}) = \frac{1}{|G|} \sum_{\rho \in \hat{G}} d_\rho \text{Tr}(\rho(t)\rho(h)).$$

Or :

$$\sum_{g \in G} f(g^{-1})h(g) = h(t^{-1}).$$

On a donc :

$$\sum_{g \in G} \bar{f}(g^{-1})h(g) = \frac{1}{|G|} \sum_{\rho \in \hat{G}} d_\rho \text{Tr}(\rho(f)\rho(h)).$$

Car $\rho(f) = \rho(t)$.

Ainsi, on conclut en remplaçant h par la fonction :

$$h = f^*, \forall g \in G f^*(g) = \bar{f}(g^{-1}).$$

□

3.2.10 Remarque :

Une telle preuve n'est possible que si tous les caractères sont orthogonaux. Pour cela on définit le produit scalaire hermitien suivant :

$$(\chi_\rho | \chi_{\rho'}) = \frac{1}{|G|} \sum_{g \in G} \text{Tr}(\rho(g) \bar{\rho'}(g^{-1})).$$

Les caractères sont orthogonaux 2 à 2 car on remarque en réindexant que si $\rho \neq \rho'$:

$$(\chi_\rho | \chi_{\rho'}) = (\chi_{\rho'} | \chi_\rho) = 0.$$

De plus :

$$(\chi_\rho | \chi_\rho) = \frac{1}{|G|} \sum_{g \in G} \text{Tr}(\rho(g) \bar{\rho}(g^{-1})) = \frac{1}{|G|} \sum_{g \in G} \text{Tr}(\rho(gg^{-1})) = 1.$$

3.2.11 Remarque essentielle pour la majoration : Tout d'abord, en utilisant Cauchy-Schwarz sur $\|T^{*k} - U\|_{VT}$ on obtient :

$$\|T^{*k} - U\|_{VT} = \sum_{\pi \in \mathbb{S}_n} 1 \times |T^{*k}(\pi) - U(\pi)| \leq \sqrt{n! \sum_{\pi \in \mathbb{S}_n} |(T^{*k} - U)(\pi)|^2}.$$

De plus, si on considère P comme étant la fonction $T^{*k} - U$ sur \mathbb{S}_n alors on obtient via Plancherel :

$$\sum_{\pi \in \mathbb{S}_n} |(T^{*k} - U)(\pi)|^2 = \sum_{\pi \in \hat{\mathbb{S}}_n} d_\rho \text{Tr}[\rho(T^{*k} - U) \rho(T^{*k} - U)^*].$$

Ainsi on a l'inégalité :

$$\|T^{*k} - U\|_{VT} \leq \sum_{\pi \in \hat{\mathbb{S}}_n} d_\rho \text{Tr}[\rho(T^{*k} - U) \rho(T^{*k} - U)^*].$$

3.3 Représentations irréductibles de \mathbb{S}_n

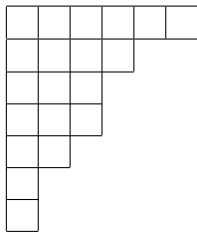
L'inégalité de Plancherel fait apparaître une sommation sur les représentations irréductibles de \mathbb{S}_n ($\hat{\mathbb{S}}_n$). L'enjeu de cette partie est donc d'étudier ces représentations irréductibles sous le point de vue des tableaux d'Young. Ensuite, en utilisant les propriétés des tableaux nous pourrions établir des lemmes sur les représentations dans la perspective du théorème de majoration.

3.3.1 Diagramme d'Young

Un diagramme d'Young est un tableau en deux dimensions contenant n cases vérifiant que chaque ligne possède moins de colonnes que la précédente.

Remarque : Chaque diagramme peut être associé à une unique décomposition de l'entier n .

Exemple : Pour la partition $\lambda = (6, 4, 3, 3, 2, 1, 1)$ de 20, va correspondre le tableau :



3.3.2 Tabloïd

Un tabloïd est une classe d'équivalence de diagrammes d'Young pour la relation $T_1 \sim T_2$ si $\lambda_1 = \lambda_2$ et chaque ligne contient les mêmes entiers, sans considération d'ordre.

Exemple : Pour la partition $(6, 4, 3, 3, 2, 1, 1)$ de 20, le tabloïd suivant convient :

20	18	5	2	1	12
3	14	8	6		
7	13	4			
11	16	9			
10	15				
17					
19					

Notation : Si T est un tableau, on note $[T]$ sa classe d'équivalence de tabloïds. Il s'agit désormais d'établir une

3.3.3 Tableau d'Young

Etant donné un diagramme d'Young λ on construit un tableau en remplissant les cases du diagramme par des entiers en suivant deux règles :

- Les entiers croissent en avançant sur chaque ligne.
- Les entiers croissent strictement en avançant sur chaque colonne.

Remarque :

En particulier, nous étudierons les tableaux standards, pour lesquels les entiers du tableaux sont ceux de 1 à n, chacun n'apparaissant qu'une fois.

Exemple : Le tableau suivant est un tableau standard associé à la partition (6,4,3,3,2,1,1) de 20 et de la même classe d'équivalence de tabloïds que le précédent :

1	2	5	12	18	20
3	6	8	14		
4	7	13			
9	11	16			
10	15				
17					
19					

équivalence entre la considération des tableaux d'Young et des représentations irréductibles de \mathbb{S}_n . Pour cela nous allons construire un module, que nous appellerons module de Specht.

3.3.4 Dimension d'une représentation

La dimension d_λ d'une représentation de \mathbb{S}_n associée à la partition λ est le nombre de possibilité de remplissage d'un tableau d'Young standard de dimension λ .

Notation: Pour un tableau T on notera $R(T)$ le sous-groupe de \mathbb{S}_n des permutations qui permutent les éléments au sein d'une même ligne. On l'appellera "groupe des lignes".

De manière analogue on définit le groupe des colonnes.

Nous pouvons alors définir l'action de \mathbb{S}_n sur les tableaux standards d'Young.

3.3.5 Définition : action de \mathbb{S}_n sur les tableaux

Soit T un tableau de dimension λ et $\sigma \in \mathbb{S}_n$.

$\sigma * T$ est un tableau de dimension λ telle que l'élément i de chaque case est remplacé par son image $\sigma(i)$ par la permutation σ .

Remarque :

L'action de \mathbb{S}_n sur les classes d'équivalence donne, $\sigma * [T] = [\sigma * T]$.

Pour l'action définie au-dessus sur les lignes et les colonnes on a, $R(\sigma * T) = \sigma * R(T) * \sigma^{-1}$ et $C(\sigma * T) = \sigma * C(T) * \sigma^{-1}$.

Notation: Soit un tableau de Young T on notera a_T , b_T et c_T les éléments de $\mathbb{C}[\mathbb{S}_n]$ suivants :

- $a_T = \sum_{p \in R(T)} p$,
- $b_T = \sum_{q \in C(T)} \text{sgn}(q)q$,
- $c_T = a_T \cdot b_T$.

Ces éléments sont appelés symétriseurs de Young.

3.3.6 Cadre mathématique

On définit M^λ le \mathbb{C} -espace vectoriel ayant pour vecteurs de sa base les tabloïds $[T]$ de dimension λ . Etant donnée l'action de \mathbb{S}_n sur les tabloïds, M^λ est un $\mathbb{C}[\mathbb{S}_n]$ -module.

On associe alors à chaque tableau T de dimension λ l'élément :

$$v_T = b_T \cdot [T] = \sum_{q \in C(T)} \text{sgn}(q)[q.T] \in M^\lambda.$$

3.3.7 Module de Specht

On définit le module de Specht S_λ comme le sous-espace de M^λ engendré par les v_T , avec T variant sur l'ensemble des dimensions λ partitions de n .

Remarque : Une telle définition du module de Specht est rendue possible par la propriété suivante, qui assure que S^λ est préservé par l'action de \mathbb{S}_n .

Propriété

Pour tout tableau de Young T et toute permutation σ de \mathbb{S}_n on a,

$$\sigma.v_T = v_{\sigma.T}.$$

Ce résultat est prouvé par Fulton dans [7].

Remarque : Il s'agit désormais de montrer que le module de Specht tout juste défini convient bien. Pour cela nous allons établir une bijection entre l'ensemble des modules de Specht et des représentations irréductibles de \mathbb{S}_n . Pour cela nous devons nous appuyer sur le lemme suivant.

3.3.8 Type d'une permutation

Le type d'un élément de \mathbb{S}_n est la partition de n obtenue en considérant le cardinal de ses orbites dans sa décomposition en cycles.

3.3.9 Lemme 1

Deux éléments de \mathbb{S}_n sont conjugués si et seulement si ils ont même type.

3.3.10 Théorème assurant l'équivalence de point de vue des modules de Specht

Pour toute partition λ de n , S^λ est une représentation irréductible de \mathbb{S}_n .

Toute représentation irréductible de \mathbb{S}_n est isomorphe à un unique module de Specht.

Preuve :

Soit λ une partition de \mathbb{S}_n , montrons que S^λ est une représentation irréductible de \mathbb{S}_n .

Par équivalence entre le point de vue de $\mathbb{C}[\mathbb{S}_n]$ -module et de représentation de \mathbb{S}_n , S^λ est une représentation de \mathbb{S}_n .

Supposons que S^λ n'est pas irréductible, alors V et W deux sous-modules de S^λ tels que $S^\lambda = V \oplus W$. On a :

$$b_T * M_\lambda = b_T * S^\lambda = \mathbb{C} * v_T.$$

Ici, on aurait :

$$\mathbb{C} * v_T = b_T * V \oplus b_T * W.$$

Donc $v_T \in V$ ou $v_T \in W$.

Supposons v_T dans V , alors $\mathbb{C}[\mathbb{S}_n]v_T = V$, or $\mathbb{C}[\mathbb{S}_n]v_T = S^\lambda$, donc $V = S^\lambda$.

Montrons ensuite qu'il y a exactement le même nombre de Specht modules que de représentations irréductibles de \mathbb{S}_n .

D'après le lemme (cf ici), il y a le même nombre de partition λ de n que de classes de conjugaisons de \mathbb{S}_n .

Or, d'après le théorème de Frobenius, il y a le même nombre de classes de conjugaisons de \mathbb{S}_n que de $\mathbb{C}[\mathbb{S}_n]$ -modules irréductibles.

Ainsi, toute représentation irréductible de \mathbb{S}_n est isomorphe à un unique module de Specht. □

3.4 Lemmes utiles pour le théorème

Après avoir introduit l'ensemble des outils nécessaires, l'enjeu de cette partie est d'introduire l'ensemble des lemmes qui nous seront utiles à la preuve de la majoration développée dans la partie 4.1.

3.4.1 Lemme 2 En utilisant la formule de Frobenius on obtient :

$$\forall \rho \in \hat{\mathbb{S}}_n, d_\rho \leq \sqrt{n!}.$$

3.4.2 Lemme 3 (ou Lemme de Schur)

Soit ρ une représentation de G , ρ est irréductible si et seulement si pour toute matrice carré de taille d_ρ M , on a $M\rho(\gamma) = \rho(\gamma)M$ pour tout γ dans G .

3.4.2.1 Exemple et Remarque : Une preuve de ce théorème a été étudiée dans le cours d'algèbre I, donc il nous semblait peu utile de le prouver à nouveau. On peut par exemple l'utiliser pour montrer que si ρ est une représentation irréductible non triviale de \mathbb{S}_n alors $\rho(U) = 0$.

3.4.3 Lemme 4

Soit G un groupe fini, ρ une représentation irréductible de G et P une fonction de G dans \mathbb{C} constante sur les classes de conjugaison. On note G_i la i -ème classe de conjugaison, $P(G_i) = \{P_i\}$, $|G_i| = n_i$ et enfin $\chi_\rho(G_i) = \{\chi_i\}$. Alors,

$$\rho(P) = CI \text{ où } C = \frac{1}{d_\rho} \sum_i P_i \chi_i n_i.$$

Preuve :

$$\rho(P) = \sum_{\gamma \in G} P(\gamma) \rho(\gamma) = \sum_i P_i \underbrace{\sum_{\eta \in G_i} \rho(\eta)}_{M_i}.$$

Or M_i est invariant par conjugaison, $\forall \gamma \in G$, $\rho(\gamma)M_i\rho(\gamma^{-1}) = M_i$.

Donc par le lemme de Schur on sait que M_i est une matrice scalaire, $M_i = C_i I$.

$$Tr(M_i) = \sum_{\eta \in G_i} Tr(\rho(\eta)) = \sum_{\eta \in G_i} \chi_\rho(\eta) = \sum_{\eta \in G_i} \chi_i = n_i \chi_i.$$

De plus, $Tr(M_i) = Tr(C_i I) = d_\rho C_i$ donc $C_i = \frac{n_i \chi_i}{d_\rho}$.

On obtient $\rho(P) = \sum_{\eta \in G_i} \frac{P_i n_i \chi_i}{d_\rho}$. □

3.4.3.1 Exemple Soit T défini ici est une fonction sur \mathbb{S}_n et soit ρ une représentation irréductible. On a bien que T est constante sur les classes de conjugaisons de \mathbb{S}_n car elle associe $\frac{1}{n}$ à l'identité, $\frac{2}{n^2}$ aux transpositions et 0 au reste, on peut donc utiliser le lemme 3. Ainsi

$$\begin{aligned} \rho(T) &= \left(\frac{1}{d_\rho} \sum_i T_i \chi_i n_i\right) I, \\ &= \frac{1}{d_\rho} (T(Id) \times Tr(\rho(Id)) \times 1 + T(\sigma) \times \chi_\rho(\sigma) \times \binom{n}{2}) I, \\ &= \left(\frac{1}{n} + \frac{(n-1)\chi_\rho(\sigma)}{nd_\rho}\right) I. \end{aligned}$$

où l'on a utilisé que le fait que le nombre de transposition s'élève à $\binom{n}{2}$ et que T est nulle sur le reste de la somme.

3.4.4 Lemme 5

Soit λ une partition de n alors

$$d_\lambda \leq \binom{n}{\lambda_1} \sqrt{(n - \lambda_1)!}.$$

Preuve :

On sait que la dimension de la représentation associée à la partition λ correspond aux nombres de manière de remplir le tableau de Young de λ en respectant la croissance dans les colonnes et dans les lignes. Il y a au plus $\binom{n}{\lambda_1}$ manières de remplir la première ligne, en effet on choisit λ_1 nombre parmi 1 à n et on les met dans la première ligne en ordre croissant. Il ne nous reste plus qu'à remplir un tableau de Young avec $(n - \lambda_1)$ cases, on peut associer à ce tableau une partition de $n - \lambda_1$ qui elle même peut être associée à une représentation irréductible de $S_{n-\lambda_1}$. Or on sait que la dimension d'une telle représentation est inférieure à $\sqrt{(n - \lambda_1)!}$ (cf lemme 2) ainsi on obtient $d_\lambda \leq \binom{n}{\lambda_1} \sqrt{(n - \lambda_1)!}$. □

3.4.5 Lemme 6

Soit χ_λ le caractère associé à la représentation irréductible de \mathbb{S}_n de la partition λ alors on a pour toute transposition σ :

$$\frac{\chi_\lambda(\sigma)}{d_\lambda} = \frac{1}{n(n-1)} \sum_{j=1}^m [(\lambda_j - j)(\lambda_j - j + 1) - j(j-1)].$$

Une version de identique de cette formule est prouvée dans [10], toutefois les notations utilisées sont assez anciennes. **3.4.6**

Lemme 7

Soit ρ et ρ' des representations irréductibles conjuguées alors :

$$d_\rho = d_{\rho'}, \quad \text{et} \quad \forall \sigma \in \text{Transp}(n), \quad \chi_\rho(\sigma) = -\chi_{\rho'}(\sigma).$$

James prouve ce lemme dans la section (6.6) de [8].

3.4.7 Lemme 8

Soit λ et λ' deux partitions de n . On a $\lambda \supseteq \lambda'$ si et seulement s'il existe un nombre fini p de partitions de n telles que :

$$\lambda = \lambda^0 \supseteq \lambda^1 \supseteq \dots \supseteq \lambda^p = \lambda'.$$

Marshall et Olkin montre ce lemme dans la section 5D de [9].

3.4.8 Lemme 9

Soit $\rho_\lambda = \rho$ et $\rho' = \rho_{\lambda'}$, si $\lambda \supseteq \lambda'$ alors $r(\rho) \leq r(\rho')$.

Ce lemme est montré intégralement dans l'article de Persi Diaconis, [1].

Il s'agit désormais d'utiliser certains de ces lemmes pour majorer $r(\rho)$ dans plusieurs cas de figures. Pour cela, nous étudierons les tableaux d'Young associés à ces représentations et les majorations viendront de leur forme, en utilisant le lemme 9. Pour la suite, on considérera un tableau associé à la partition $\lambda = (\lambda_1, \dots, \lambda_m)$ et ρ la représentation associée.

Pour les trois cas qui vont suivre, les preuves sont très similaires et suivent les mêmes étapes ; à savoir la définition de l'entier b , la définition d'un λ' , suivi de l'utilisation des lemme 9, puis 6, et de la conclusion (cf preuve du lemme 10 ci-dessous). Nous choisissons donc de ne présenter que la preuve du lemme 10.

3.4.9 Lemme 10 : cas $\lambda_1 \leq n/3$

Pour une telle représentation on a :

$$r(\rho) \leq \frac{1}{n(n-1)} \left[\frac{n^2}{3} + \frac{20}{3} - 3n \right].$$

Preuve :

On note b le plus petit entier supérieur ou égal à $\frac{n}{3}$, $b = \frac{n}{3} + \epsilon$, $\lambda' = (b, b, n-2b)$ (une partition de n) et ρ' la représentation associée.

On a $\lambda' \supseteq \lambda$, donc d'après le lemme 9, $r(\rho) \leq r(\rho')$.

On utilise le lemme 6 pour avoir :

$$r(\rho) \leq \frac{1}{n(n-1)} [b(b-1) + (b-1)(b-2) - 2 + (n-2b-2)(n-2b-3) - 6].$$

On remplace, b par $n/3 + \epsilon$ pour avoir :

$$r(\rho) \leq \frac{1}{n(n-1)} \left[\frac{n^2}{3} + 6(\epsilon^2 + \epsilon) - 3n \right].$$

Enfin, on trouve l'expression finale en utilisant, $\epsilon \leq 2/3$. Ainsi :

$$r(\rho) \leq \frac{1}{n(n-1)} \left[\frac{n^2}{3} + \frac{20}{3} - 3n \right].$$

□

3.4.10 Lemme 11 : cas $\lambda_1 \leq n/2$

Pour une telle représentation on a :

$$r(\rho) \leq \frac{1}{n(n-1)} \left[\frac{n^2}{2} - n \right].$$

3.4.11 Lemme 12 : cas $\lambda_1 > n/2$

Pour une telle représentation on a :

$$0 < r(\rho) \leq \frac{1}{n(n-1)} [\lambda_1(\lambda_1-1)(n-\lambda_1-1)(n-\lambda_1-2) - 2].$$

3.4.12 Lemme 13: Majoration du nombre de partitions d'un entier

On note $p(n)$ le nombre de partitions de l'entier n , on a :

$$p(n) \leq e^{\pi\sqrt{\frac{2n}{3}}}.$$

Ayoub prouve cette formule dans le chapitre 3, section 1 de [11] en 1963.

4 Preuve des théorèmes

4.1 Théorème sur la majoration

Il s'agit de démontrer le théorème sur la majoration de la distance quand n croit. Nous utiliserons pour cela les différents outils évoqués dans la partie 3 sur les propriétés des représentations de \mathbb{S}_n . Nous pourrions alors obtenir l'estimation du point critique en $\frac{n \log(n)}{2}$ annoncée.

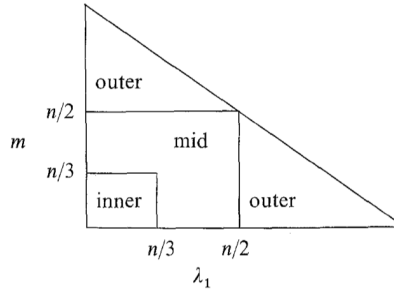
Dans un premier temps, nous allons majorer la distance de variation totale par une formule portant sur les représentations de \mathbb{S}_n . Par la suite, nous ferons une disjonction de cas en fonction des différentes formes possibles des tableaux d'Young, s'inspirant des lemmes 9 à 11. Enfin, nous pourrions exprimer la majoration.

4.1.1 Lemme 14

$$\|T^{*k} - U\|_{VT} \leq \left[\sum_{\rho \in \hat{G}} d_{\rho}^2 \left(\frac{1}{n} + \frac{n-1}{n} r(\rho) \right)^{2k} \right]^{1/2}$$

En combinant les remarques et exemples 3.4.3.1, 3.4.2.1 et 3.2.11 on obtient le résultat.

On peut désormais étudier la majoration en termes de tableaux d'Young. Pour cela, nous distinguerons les tableaux en fonction du cardinal m de leur dimension $\lambda = (\lambda_1, \dots, \lambda_m)$ et de la taille de λ_1 . On distingue alors trois groupes que l'on peut résumer par le graphique suivant :



Zone (1) $\lambda_1 \leq \frac{n}{3}$ et $m \leq \frac{n}{3}$, Zone (2) $[\frac{n}{3} \leq \lambda_1 \leq \frac{n}{2}$ et $m \leq \frac{n}{2}]$ ou $[\frac{n}{3} \leq m \leq \frac{n}{2}$ et $\lambda_1 \leq \frac{n}{2}]$, Zone (3) $\lambda_1 > \frac{n}{2}$ ou $m > \frac{n}{2}$. Au sein de la zone 3, on distinguera également 3 zones : Zone (3.A) : $\frac{n}{2} < \lambda_1 \leq 0.7n$ ou $\frac{n}{2} < m \leq 0.7n$, Zone (3.B) : $0.7n < m \leq n$, Zone(3.C) $0.7n < \lambda_1 \leq n - 1$.

4.1.2 Majoration pour chacune des zones

Zone 1 :

$$\sum_{\rho \in (Zone(1))} d_{\rho}^2 \left| \frac{1}{n} + \frac{n-1}{n} r(\rho) \right|^{2k} \leq \left(\frac{1}{3} \right)^{2k} n!.$$

Zone 2 :

$$\sum_{\rho \in (Zone(2))} d_{\rho}^2 \left| \frac{1}{n} + \frac{n-1}{n} r(\rho) \right|^{2k} \leq e^{\pi\sqrt{\frac{2n}{3}}} 4^n \frac{1}{2} n^{2n/3}.$$

Zone 3.A

$$\sum_{\rho \in (Zone(3.A))} d_{\rho}^2 \left| \frac{1}{n} + \frac{n-1}{n} r(\rho) \right|^{2k} \leq e^{\pi\sqrt{\frac{2n}{3}}} 4^n \left[\frac{n}{2} \right]! \left(0.58 + \frac{1}{5n} + \frac{4}{n^2} \right)^{2k}$$

Zone 3.B

$$\sum_{\rho \in (\text{Zone}(3.B))} d_\rho^2 \left| \frac{1}{n} + \frac{n-1}{n} r(\rho) \right|^{2k} = B_{3.B} e^{-2c}.$$

Zone 3.C

$$\sum_{\rho \in (\text{Zone}(3.C))} d_\rho^2 \left| \frac{1}{n} + \frac{n-1}{n} r(\rho) \right|^{2k} = B_{3.C} e^{-2c}.$$

Les preuves de ces majorations sont principalement calculatoires et présentent une faible application réelle à cause de leur aspect calculatoire. Elles consistent principalement en une majoration de la dimension d_ρ utilisant les lemmes 5 et 7 et une majoration du terme $r(\rho)$ qui utilise les lemmes 6,9,10,11,12 et parfois le lemme 13 pour majorer le nombre de termes de la somme sur la zone i. Montrons désormais que la somme de tous ces termes est majorée par $B e^{-4c}$, avec B une constante que l'on ne cherchera pas à déterminer.

Preuve : Majorons chacun des termes par un terme de la forme $B e^{-4c}$, pour $n \geq 10$.

Zone 1 :

$\frac{1}{3}^{2k} n! \leq \frac{1}{3}^{2cn} \left(\frac{e}{3}\right)^{n \log(n)}$, en utilisant Stirling on a, $\frac{1}{3}^{2k} n! \leq e^{-2cn} \leq e^{-4c}$.

Zone 2 :

On réécrit le terme sous forme exponentielle, $e^{\pi \sqrt{\frac{2n}{3}}} 4^n \frac{1}{2}^{2k} n^{2n/3} = e^{[\pi \sqrt{\frac{2n}{3}} + n \log(4) + (2/3 - \log(2))n \log(n)]} e^{-2cn \log(2)}$.

Dans la première exponentielle, le terme de plus grand ordre est celui en $n \log(n)$, or $2/3 < \log(2)$, donc le crochet dans l'exponentielle tend vers $-\infty$, quand n tend vers l'infini.

On a donc une constante B_2 telle que, $e^{\pi \sqrt{\frac{2n}{3}}} 4^n \frac{1}{2}^{2k} n^{2n/3} \leq B_2 e^{-2 \log(2)cn} \leq B_2 e^{-4c}$.

Zone 3.A :

$e^{\pi \sqrt{\frac{2n}{3}}} 4^n \lfloor \frac{n}{4} \rfloor! (0.58 + \frac{1}{5n} + \frac{4}{n^2})^{n \log}$ est majoré et tend vers 0. On en déduit donc l'existence d'une constante $B_{3.A}$ telle que la somme pour la zone 3.1 soit majorée par $B_{3.A} e^{-4c}$.

Zone 3.B :

La majoration par $B_{3.B} e^{-4c}$ s'obtient directement. Et pour la zone 3.C la majoration est déjà obtenue dans la preuve précédente.

Ainsi, en rassemblant tous les résultats on a :

$$\|T^{*k} - U\|_{VT} \leq (B_1 + B_2 + B_{3.A} + B_{3.B} + B_{3.C}) e^{-4c}$$

□

4.2 Théorème sur la minoration

Nous cherchons dans cette partie à démontrer le théorème de minoration, supposons que n tende vers l'infini alors $C_k \geq 2\left(\frac{1}{e} - e^{-e^{-2c}}\right) + o(1)$.

Nous utiliserons pour cela les deux lemmes suivants venant de considérations probabilistes :

4.2.1 Lemme 15

Soit A_n l'ensemble des permutations avec au moins un point fixe.

On note U_n la distribution uniforme évoquée dans la définition de la distance de variation totale.

Quand n tend vers $+\infty$ on a :

$$U(A_n) = 1 - \frac{1}{e} + O\left(\frac{1}{n!}\right).$$

4.2.2 Lemme 16

On place avec une probabilité uniforme 2k boules dans n boîtes.

On note $B_{n,k}$ l'évènement et $B_k = \lim_{n \rightarrow +\infty} B_{n,k}$:

Au moins une boîte est vide. On a :

$$\lim_{n \rightarrow +\infty} P(B_{n,k}) = 1 - e^{-ne^{-2c}} + o(1)$$

Nous pouvons désormais démontrer le théorème 2, de minoration.

Preuve :

Nous utiliserons:

$$\|T^{*k} - U\|_{VT} = 2 \max_{E \subset \mathbb{S}_n} |T^{*k}(E) - U(E)|.$$

Le A considéré dans le lemme 15 est dans \mathbb{S}_n de même que $B_{n,k}$ pour tout n, k .

Donc : $\lim_{n \rightarrow +\infty} B_{n,k} \in \mathbb{S}_n$.

Or, $B_{n,k}$ a la même probabilité que l'évènement $C_{n,k}$:

Il y a au moins une des n cartes qui n'a pas été affectée après k étapes de mélange, sachant qu'une étape est ici l'échange des positions de deux cartes tirées aléatoirement avec remise.

On a donc aussi, à partir d'un certain rang : $C_{n,k} \subset A$.

Donc :

$$T^{*k}(A) \geq T^{*k}(B_k) \geq 1 - 2e^{-e^{-2c}} + o(1).$$

Ainsi :

$$\begin{aligned} \|T^{*k} - U\|_{VT} &\geq 2|T^{*k}(A) - U(A)|, \\ &\geq 2\left(\frac{1}{e} - e^{-e^{-2c}}\right) + o(1). \end{aligned}$$

□

References

- [1] Persi Diaconis et Mehrdad Shahshahani. Generating a Random Permutation with Random Transpositions. Stanford University, Departement of Statistics, 1981.
- [2] William Fulton. Young Tableaux. University of Chicago, 1996.
- [3] Persi Diaconis. Group Representations in Probability and Statistics. 1988.
- [4] Justin Salez. Mixing times of Markov chains. Lecture notes.
- [5] David A. Levin Yuval Peres. Markov Chains and Mixing Times, second edition.
- [6] Diaconis. Group Representations in Probability and Statistics, 1988.
- [7] William Fulton. Young Tableaux with Applications to Representation Theory and Geometry, 1997.
- [8] James. The Representation Theory of the Symetric Groups. Springer-Verlag 1978.
- [9] Marshall, A. W., Olkin. Inequalities: Theory of Majorization and its Application. New York : Academic Press 1974
- [10] Ingram. Some characters of the symetric group. page 358-369. 1950.
- [11] Ayoub. An Introduction to the Analytic Theory of Numbers Rhode Island: Providence 1963.