

Permutations aléatoires et matrices aléatoires

Ce dernier chapitre d'introduction est consacré au lien entre la théorie asymptotique des représentations des groupes symétriques et la théorie des **grandes matrices aléatoires**. La ressemblance entre ces deux thèmes concerne en premier lieu les outils : ainsi, les espérances de polynômes traciaux jouent un rôle sensiblement identique aux observables de diagrammes du chapitre 2, et d'autre part, les valeurs propres d'une matrice aléatoire constituent un processus ponctuel, qui est déterminantal pour le modèle du GUE (c'est-à-dire lorsque les entrées de la matrice sont des gaussiennes centrées indépendantes). Ainsi, des raisonnements semblables à ceux des paragraphes 3.3 et 4.2 permettent de déterminer la limite des lois empiriques renormalisées des valeurs propres (loi de Wigner), ainsi que l'asymptotique « locale » au centre et au bord du spectre (noyau sinus et noyau d'Airy). Nous rappelons ceci dans la section 5.1, en suivant pour l'essentiel [Kö05] — on renvoie à [Meh04, AGZ09] pour un traitement plus exhaustif. Le noyau d'Airy intervient également comme limite des noyaux mis en jeu dans la description déterminantale des processus ponctuels associés aux mesures de Plancherel poissonisées. Ce fait est à l'origine de l'équivalence de Baik-Deift-Johansson ([BDJ99, BDJ00]), qui peut être énoncée comme suit :

Théorème 5.1 (Baik-Deift-Johansson, [BOO00]). Soient $\lambda_1 \geq \lambda_2 \geq \dots$ les plus grandes valeurs propres d'une matrice hermitienne du GUE, et

$$Y_i = n^{2/3} \left(\frac{\lambda_i}{2\sqrt{n}} - 1 \right)$$

leurs déviations renormalisées. On note d'autre part $\mu_1 \geq \mu_2 \geq \dots$ les premières colonnes d'une partition aléatoire μ sous la mesure de Plancherel M_n , et

$$X_i = n^{1/3} \left(\frac{\mu_i}{2\sqrt{n}} - 1 \right)$$

leurs déviations renormalisées. Par la correspondance RSK, il s'agit aussi des déviations renormalisées des tailles des plus longs sous-mots croissants d'une permutation aléatoire de taille n . Lorsque n tend vers l'infini, les lois jointes de (X_1, \dots, X_k) et de (Y_1, \dots, Y_k) ont la même limite, et ce pour tout k .

Pour $k = 1$, la loi limite est la **loi de Tracy-Widom**, dont la densité est liée à l'équation différentielle de Painlevé II (voir [TW94]). Les sections 5.2 et 5.3 sont consacrées à deux ébauches¹

1. K. Johansson a donné une troisième preuve qui utilise des polynômes orthogonaux, cf. [Joh01].

de preuve du théorème 5.1 : la première s'appuie sur les processus ponctuels déterminantaux, et la seconde est de nature géométrique — cette approche due à Okounkov ([Ok000]) a fait l'objet d'un exposé à l'Institut Henri Poincaré en décembre 2009. Comme mentionné dans l'introduction de cette partie, les résultats de ce chapitre sont présentés essentiellement à titre culturel² ; néanmoins, la connection avec la théorie des matrices aléatoires est l'une des motivations principales de notre étude.

5.1 Théorie asymptotique des matrices aléatoires hermitiennes

Soit $H(n, \mathbb{C})$ l'ensemble des matrices complexes de taille $n \times n$ et hermitiennes, i.e., $M_{ji} = \overline{M_{ij}}$ pour tout couple d'indices (i, j) . On note $M_{ii} = x_{ii} \in \mathbb{R}$, et $M_{i < j} = x_{ij} + iy_{ij}$ avec $x_{ij} \in \mathbb{R}$, $y_{ij} \in \mathbb{R}$. La loi du *Gaussian Unitary Ensemble* est la mesure de probabilité gaussienne sur $H(n, \mathbb{C})$ donnée par la densité

$$d\mathbb{P}[M] = \prod_{i=1}^n \mathcal{N}_{(0,1)}(dx_{ii}) \prod_{i < j} \mathcal{N}_{(0,1/2)}(dx_{ij}) \prod_{i < j} \mathcal{N}_{(0,1/2)}(dy_{ij}) = \frac{1}{\tilde{Z}_n} e^{-\frac{\text{tr} M^2}{2}} dM,$$

où \tilde{Z}_n est une certaine constante de normalisation. Cette loi gaussienne est invariante pour l'action par conjugaison³ de $U(n, \mathbb{C})$ sur $H(n, \mathbb{C})$. Notons $\lambda_1 \geq \dots \geq \lambda_n$ les valeurs propres d'une matrice du GUE ; c'est un élément de la chambre de Weyl Δ_n . En calculant le jacobien de l'application

$$\begin{aligned} U(n, \mathbb{C}) / (U(1, \mathbb{C}))^n \times \Delta_n &\rightarrow H(n, \mathbb{C}) \\ [U], (\lambda_1 > \dots > \lambda_n) &\mapsto U \text{diag}(\lambda_1, \dots, \lambda_n) U^{-1} \end{aligned}$$

d'image ouverte dense, on peut déterminer la loi jointe des valeurs propres sous la mesure du GUE :

$$d\mathbb{P}[x_1 \geq x_2 \geq \dots \geq x_n] = \frac{1}{Z_n} \Delta(x_1, \dots, x_n)^2 e^{-\frac{1}{2} \sum_{i=1}^n (x_i)^2} dx_1 \dots dx_n$$

où $\Delta(x) = \prod_{i < j} (x_j - x_i)$ est le Vandermonde de la matrice $((x_i)^{j-1})_{1 \leq i, j \leq n}$, et Z_n est une constante de normalisation.

Les valeurs propres d'une matrice M du GUE peuvent être regroupées en une mesure de probabilité (aléatoire) appelée **mesure empirique**, et définie par

$$m_M = \frac{1}{n} \sum_{i=1}^n \delta_{\lambda_i}.$$

Si l'on note $\tilde{m}(ds) = m(\sqrt{n} ds)$ la mesure empirique renormalisée, alors le résultat suivant est l'analogie pour les matrices aléatoires du théorème 3.3 :

2. Le calcul du noyau associé à la mesure de Plancherel poissonisée de paramètre θ sera néanmoins repris et généralisé dans la section 7.1.

3. Compte tenu de la décomposition de Cartan $GL(n, \mathbb{C}) \simeq U(n, \mathbb{C}) \times H(n, \mathbb{C})$, la donnée d'une mesure $U(n, \mathbb{C})$ -invariante sur $H(n, \mathbb{C})$ est équivalente à la donnée d'une mesure bi- $U(n, \mathbb{C})$ -invariante sur $GL(n, \mathbb{C})$, donc la loi du GUE fournit également une mesure gaussienne « symétrique » naturelle sur le groupe linéaire complexe.

Proposition 5.2 (Loi de Wigner, [Wig58]). *Au sens de la convergence en loi, $\tilde{m} \rightarrow \mu_\Omega$ en probabilité, où*

$$\mu_\Omega = \mathbb{1}_{[-2,2]}(s) \frac{\sqrt{4-s^2}}{2\pi} ds$$

est la loi de Wigner du demi-cercle — c'est aussi la mesure de transition de la forme limite Ω pour les diagrammes sous la mesure de Plancherel.

Une preuve possible de la loi de Wigner repose sur le développement topologique des espérances $\mathbb{E}[\text{tr } M^k]$ donné par la **formule de Wick**; cette formule rentre dans le cadre des techniques de diagrammes de Feynman en théorie quantique des champs. Ces interprétations combinatoires des intégrales gaussiennes peuvent toutes être comprises à partir du cas trivial de la dimension $n = 1$. Ainsi, pour calculer le k -ième moment

$$I_k = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} x^k e^{-\frac{x^2}{2}} dx$$

d'une gaussienne, il est utile de perturber le champ $x^2/2$ par un terme linéaire tx :

$$\frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} x^k e^{-\frac{x^2}{2}+tx} dx = \frac{\partial^k}{\partial t^k} \left(\frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-\frac{x^2}{2}+tx} dx \right) = \frac{\partial^k}{\partial t^k} \left(e^{\frac{t^2}{2}} \right).$$

Par conséquent, $I_k = \left. \frac{d^k}{dt^k} \right|_{t=0} \left(e^{\frac{t^2}{2}} \right)$, et lors de ce calcul, chaque dérivation de $e^{\frac{t^2}{2}}$ donne un terme t qui doit être compensé par une autre dérivation, car les termes en t^l restant à la fin du calcul sont annulés en évaluant en $t = 0$. On en déduit :

$$I_k = \text{nombre de façons d'apparier } k \text{ points} = \begin{cases} (2p-1)!! & \text{si } k = 2p \text{ est pair,} \\ 0 & \text{sinon.} \end{cases}$$

Exemple. Détaillons le calcul de I_4 et l'interprétation combinatoire en termes d'appariements :

$$\begin{aligned} I_4 &= \left. \frac{d^4}{dt^4} \right|_{t=0} \left(e^{\frac{t^2}{2}} \right) && \bullet \bullet \bullet \bullet \\ &= \left. \frac{d^3}{dt^3} \right|_{t=0} \left(t e^{\frac{t^2}{2}} \right) && \curvearrowright \bullet \bullet \bullet \\ &= \left. \frac{d^2}{dt^2} \right|_{t=0} \left(e^{\frac{t^2}{2}} + t^2 e^{\frac{t^2}{2}} \right) && \curvearrowleft \bullet \bullet \bullet + \curvearrowright \curvearrowright \bullet \bullet \bullet \\ &= \left. \frac{d}{dt} \right|_{t=0} \left(3t e^{\frac{t^2}{2}} \right) && \curvearrowleft \bullet \curvearrowright \bullet + \curvearrowright \curvearrowleft \bullet \bullet + \curvearrowleft \curvearrowright \bullet \bullet \\ &= 3 && \curvearrowleft \bullet \curvearrowleft \bullet + \curvearrowright \curvearrowright \bullet \bullet + \curvearrowleft \curvearrowright \bullet \bullet \end{aligned}$$

Les mêmes techniques s'appliquent au calcul de l'espérance $\mathbb{E}[\text{tr } M^k]$, qui pour des raisons de parité est nulle si k est impair. Ainsi :

$$\mathbb{E}[\text{tr } M^{2p}] = \sum_{i_1, \dots, i_{2p}} \mathbb{E}[M_{i_1 i_2} M_{i_2 i_3} \cdots M_{i_{2p} i_1}] = \sum_{i_1, \dots, i_{2p}} \left. \frac{\partial^{2p}}{\partial M_{i_1 i_2} \cdots \partial M_{i_{2p} i_1}} \right|_{M=0} \left(e^{\frac{\text{tr } M^2}{2}} \right).$$

Plaçons les entiers $1, 2, \dots, 2p$ aux sommets d'un $2p$ -gone régulier; la dérivée partielle d'indices i_1, \dots, i_{2p} évaluée en $M = 0$ est le nombre d'appariements des côtés du $2p$ -gone qui sont compatibles avec l'orientation canonique, et tels que deux sommets j_1 et j_2 sont identifiés si et seulement si $i_{j_1} = i_{j_2}$. C'est donc le nombre de **cartes**⁴ à p arêtes et $\text{card} \{i_1, i_2, \dots, i_{2p}\}$ sommets obtenues par recollement deux à deux des côtés d'un $2p$ -gone. Le nombre de sommets s'écrit $p + 1 - 2g$, où g est le genre de la surface obtenue. Par conséquent :

Proposition 5.3 (Formule de Wick, [t'H74]). *Si $k = 2p$ est un entier pair et si M est une matrice du GUE, alors*

$$\mathbb{E}[\text{tr } M^k] = \sum_{g \geq 0} n^{p+1-2g} |\text{Map}(k, g)|,$$

où $\text{Map}(k, g)$ est l'ensemble des cartes planaires de genre g obtenues par recollement des côtés d'un k -gone.

Exemple. À symétrie près, les cartes obtenues par recollement des côtés d'un hexagone sont toutes représentées sur la figure 5.1. L'espérance $\mathbb{E}[\text{tr } M^6]$ vaut donc $5n^4 + 10n^2$.

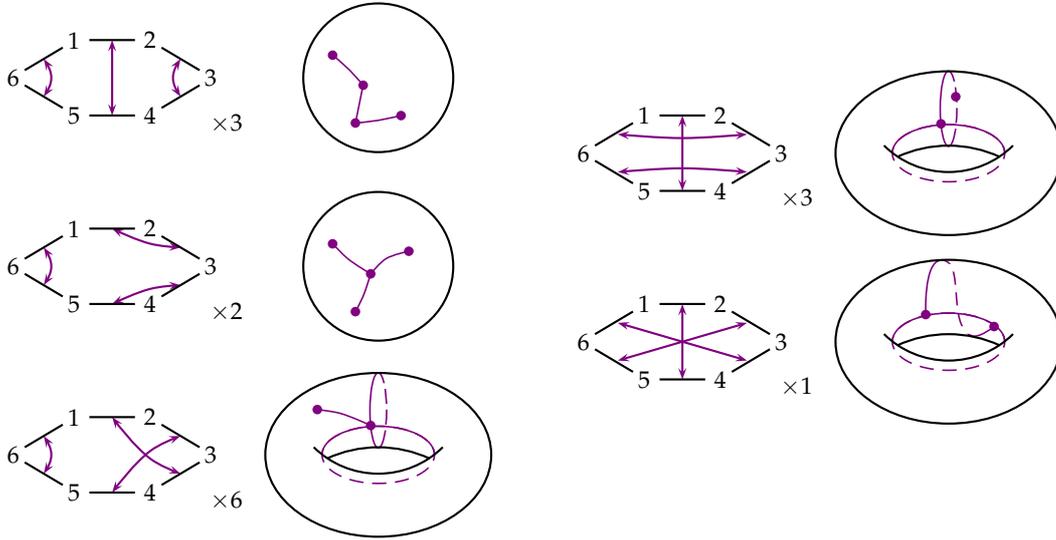


FIGURE 5.1 – Cartes obtenues par recollement des côtés d'un hexagone. Les symétries sont indiquées en dessous de chaque carte.

De façon générale, on obtient une carte planaire ($g = 0$) si et seulement si le graphe obtenu est un arbre, et il y a C_p façons différentes de faire ceci. Or, les cartes planaires correspondent au terme dominant dans le développement topologique de $\mathbb{E}[\text{tr } M^{2p}]$, donc :

$$\mathbb{E}[\text{tr } M^{2p}] = \mathbb{E}[\tilde{m}(x^{2p})] n^{p+1} = C_p n^{p+1} + O(n^{p-1}).$$

Comme la mesure μ_Ω a pour moments pairs les nombres de Catalan et a ses moments impairs nuls, on en déduit que $\mathbb{E}[\tilde{m}(x^k)]$ tend vers $\mu_\Omega(x^k)$ pour tout entier k . De plus, la formule de Wick se généralise au cas d'un produit de traces de puissances de M :

$$\mathbb{E} \left[\prod_{l=1}^s \text{tr } M^{k_l} \right] = \sum_S n^{\frac{k_1 + \dots + k_s}{2} + \chi(S) - s} |\text{Map}(k_1, \dots, k_s; S)|,$$

4. Une carte est la donnée d'une surface de Riemann et d'un graphe tracé dessus dont toutes les faces sont homéomorphes à des disques. On renvoie à [LZ04] pour la théorie des cartes et ses multiples incarnations; en particulier, le lien avec les modèles matriciels est évoqué dans le chapitre 3 de cet ouvrage.

la somme étant effectuée sur toutes les classes d'homéomorphismes de surfaces compactes S (éventuellement non connexes), et $\text{Map}(k_1, \dots, k_s; S)$ désignant l'ensemble des cartes tracées sur une surface de type S obtenues par recollement d'un k_1 -gone, d'un k_2 -gone, etc., et d'un k_s -gone.

Exemple. À symétrie près, les recollements possibles des côtés de deux carrés sont tous représentés sur la figure 5.2. L'espérance $\mathbb{E}[(\text{tr } M^4)^2]$ vaut donc $4n^6 + 24n^4 + 77n^2$.

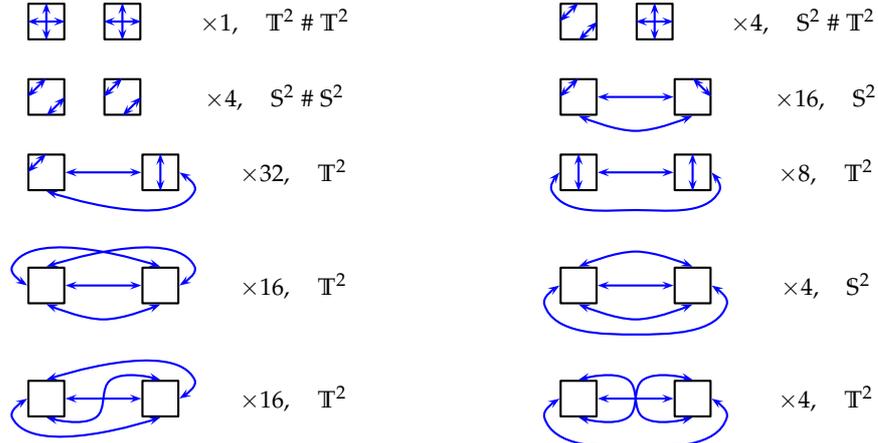


FIGURE 5.2 – Recollements possibles des côtés de deux carrés. Les symétries sont indiquées à côté de chaque carte, ainsi que le type de la surface obtenue.

On peut ainsi calculer $\mathbb{E}[(\tilde{m}(x^k))^2]$, et pour $k = 2p$ pair, le terme dominant correspond aux cartes à deux composantes connexes planaires — c'est ce qui maximise la caractéristique d'Euler $\chi(S)$. Ainsi, $\mathbb{E}[(\tilde{m}(x^{2p}))^2] \simeq (C_p)^2$, donc :

$$\mathbb{E}[(\tilde{m}(x^{2p}) - \mu_\Omega(x^{2p}))^2] \rightarrow 0.$$

Ceci permet de conclure quant à la convergence de \tilde{m} vers la loi du demi-cercle μ_Ω . L'interprétation géométrique des espérances de polynômes traciaux $\mathbb{E}[\prod_{i=1}^s \text{tr } M^{k_i}]$ jouera de nouveau un rôle essentiel dans une preuve de l'équivalence de Baik-Deift-Johansson, voir la section 5.3.

Toujours dans le cadre du GUE, la déviation renormalisée $\sqrt{n}(\tilde{m} - \mu_\Omega)$ peut être décrite par un théorème central limite analogue à celui de la section 3.4, voir [Joh98]. D'autre part, les valeurs propres $\lambda_1, \dots, \lambda_n$ peuvent être envisagées comme un processus ponctuel sur \mathbb{R} de densité

$$dQ[x_1, \dots, x_n] = \frac{1}{n! Z_n} \Delta(x_1, \dots, x_n)^2 e^{-\frac{1}{2} \sum_{i=1}^n (x_i)^2} dx_1 \cdots dx_n,$$

c'est-à-dire que la k -ième fonction de corrélation du processus s'écrit :

$$\rho^k(x_1, \dots, x_k) = \frac{n!}{(n-k)!} \int_{\mathbb{R}^{n-k}} \frac{dQ(x_1, \dots, x_k, x_{k+1}, \dots, x_n)}{dx} dx_{k+1} \cdots dx_n$$

Dans ce contexte, l'identité

$$\mathbb{E} \left[\prod_{i=1}^n (1 + f(x_i)) \right] = \sum_{k=0}^n \frac{1}{k!} \int_{\mathbb{R}^k} \rho^k(x_1, \dots, x_k) f(x_1) \cdots f(x_k) dx_1 \cdots dx_k$$

permet de nombreux calculs de probabilités ; par exemple, en prenant $f = -\mathbb{1}_A$ avec $A \subset \mathbb{R}$, on obtient

$$\mathbb{P}[\forall i, x_i \notin A] = \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} \int_{A^k} \rho^k(x_1, \dots, x_k) dx_1 \cdots dx_k.$$

Envisageons plus généralement un processus ponctuel sur \mathbb{R} dont la densité s'écrit

$$d\mathbf{Q}(x) = Z^{-1} \Delta(x)^2 e^{-\frac{1}{2} \sum_{i=1}^n R(x_i)} dx,$$

où $R(x)$ est une fonction suffisamment grande à l'infini. Soit $(\pi_k)_{k \in \mathbb{N}}$ la famille des polygones orthogonaux associés à la mesure $\exp(-R(x)/2) dx$, c'est-à-dire que $\pi_k(x) = x^k + \cdots$ et

$$\int_{\mathbb{R}} \pi_i(x) \pi_j(x) e^{-\frac{R(x)}{2}} dx = c_i c_j \delta_{ij}.$$

Les fonctions $\phi_k(x) = (c_k)^{-1} \exp(-R(x)/4) \pi_k(x)$ forment une base orthonormale de $L^2(\mathbb{R})$, et d'autre part, $\Delta(x_1, \dots, x_k) = \det((\pi_i(x_j))_{1 \leq i, j \leq k})$ pour tout entier k . Notons $K_n(x, y)$ le noyau $\sum_{i=0}^{n-1} \phi_i(x) \phi_i(y)$. La structure particulière de ce noyau (dit reproduisant) implique le caractère déterminantal du processus ponctuel considéré :

Proposition 5.4 (Polynômes orthogonaux et processus ponctuels déterminantaux, [Meho4]).
 Dans le cadre précédent, le processus ponctuel $\{x_1, \dots, x_n\}$ a pour densité

$$d\mathbf{Q}(x) = \frac{1}{n!} \det((K_n(x_i, x_j))_{1 \leq i, j \leq n}) dx,$$

et les fonctions de corrélations s'écrivent $\rho^k(x_1, \dots, x_k) = \det((K_n(x_i, x_j))_{1 \leq i, j \leq k})$. De plus, le noyau $K_n(x, y)$ s'écrit à l'aide de la formule de Christoffel-Darboux :

$$K_n(x, y) = \frac{c_{n-1}}{c_n} \frac{\phi_n(x) \phi_{n-1}(y) - \phi_n(y) \phi_{n-1}(x)}{x - y}.$$

Exemple. Les valeurs propres d'une matrice du GUE renormalisées par un facteur \sqrt{n} correspondent au cas où $R(x) = nx^2$. Le polynôme π_k s'écrit $\pi_k(x) = H_k(\sqrt{n}x)/n^{k/2}$, où H_k est le k -ième polynôme de Hermite. L'asymptotique du processus ponctuel associé aux valeurs propres renormalisées se déduit donc dans ce cas des propriétés à l'infini des fonctions de Hermite (voir [Sze39], en particulier le chapitre 8 et les formules de Plancherel-Rotach). Précisons quelque peu ces résultats. Si B est une partie (borélienne) de \mathbb{R} , alors

$$\text{card} \{i \mid x_i \in B\} \sim n \mu_{\Omega}(B)$$

lorsque n tend vers l'infini. En particulier, si $B = B_{n,u} = [u - l/2n, u + l/2n]$ est un intervalle centré en u avec $u \in]-2, 2[$ point à l'intérieur du spectre limite, alors

$$\text{card} \{i \mid x_i \in B_{n,u}\} \sim \rho_{\Omega}(u) = \frac{\sqrt{4 - u^2}}{2\pi}.$$

Par contre, si $u = \pm 2$ est un point du bord, alors il faut choisir une renormalisation de l'intervalle B en $n^{2/3}$, car :

$$\text{card} \{i \mid x_i \in [2 - \varepsilon_n, 2]\} \sim \frac{n}{2\pi} \int_{2 - \varepsilon_n}^2 \sqrt{4 - x^2} dx \sim \frac{2}{3\pi} n(\varepsilon_n)^{3/2}.$$

Pour comprendre l'asymptotique « locale » du processus ponctuel des valeurs propres d'une matrice du GUE, on introduit donc deux noyaux renormalisés :

$$\begin{aligned}\tilde{K}_n^{\text{bulk}}(x, y) &= \frac{1}{n \rho_\Omega(u)} K_n \left(u + \frac{x}{n \rho_\Omega(u)}, u + \frac{y}{n \rho_\Omega(u)} \right); \\ \tilde{K}_n^{\text{edge}}(x, y) &= \frac{1}{n^{2/3}} K_n \left(2 + \frac{x}{n^{2/3}}, 2 + \frac{y}{n^{2/3}} \right).\end{aligned}$$

Proposition 5.5 (Asymptotique du processus ponctuel des valeurs propres, [TW94]). *Lorsque n tend vers l'infini, le noyau renormalisé à l'intérieur du spectre converge vers le noyau sinus :*

$$\forall u, \forall x, y, \tilde{K}_n^{\text{bulk}}(x, y) \rightarrow \frac{\sin \pi(x - y)}{\pi(x - y)}.$$

De même, le noyau renormalisé au bord du spectre converge vers le noyau d'Airy :

$$\forall x, y, \tilde{K}_n^{\text{edge}}(x, y) \rightarrow \frac{\text{Ai}(x) \text{Ai}'(y) - \text{Ai}'(x) \text{Ai}(y)}{x - y},$$

où $\text{Ai}(x) = \frac{1}{\pi} \int_0^\infty \cos(tx + t^3/3) dt$ est la fonction d'Airy, c'est-à-dire l'unique solution de $y'' - xy = 0$ qui ait pour asymptotique

$$\text{Ai}(x) \sim_{x \rightarrow \infty} \frac{e^{-2x^{3/2}/3}}{2\sqrt{\pi} x^{1/4}}.$$

En combinant ces résultats et la formule donnée pour la probabilité $\mathbb{P}[\forall i, x_i \notin A]$, C. Tracy et H. Widom sont parvenus à préciser l'asymptotique de la plus grande valeur propre λ_{\max} d'une matrice du GUE. Ainsi,

$$\begin{aligned}\mathbb{P}[2Y_1 \leq a] &= \mathbb{P} \left[n^{2/3} \left(\frac{\lambda_{\max}}{\sqrt{n}} - 2 \right) \leq a \right] = \sum_{k=0}^N \frac{(-1)^k}{k!} \int_{]a, \infty[^k} \det \left(\tilde{K}_n^{\text{edge}}(x_i, x_j) \right) dx^k \\ &\rightarrow \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} \int_{]a, \infty[^k} \det \left(K^{\text{Airy}}(x_i, x_j) \right) dx^k = \det \left(\text{id} - \mathcal{K}^{\text{Airy}} \Big|_{L^2(]a, \infty[)} \right).\end{aligned}$$

La loi obtenue, dite **loi de Tracy-Widom**, a pour fonction de répartition

$$F(x) = \exp \left(- \int_x^\infty (s - x) q^2(s) ds \right),$$

où q est l'unique solution de l'équation différentielle de Painlevé II

$$q''(x) = xq(x) + 2q^3(x)$$

qui ait pour asymptotique $q(x) \sim \text{Ai}(x)$. En particulier, la distribution de la plus grande valeur propre du GUE est liée à la théorie des équations de Painlevé et à la hiérarchie de Toda, voir les articles [FW01, FW02]. L'objectif de ce chapitre est de montrer que les premières lignes (ou colonnes) d'une partition sous la mesure de Plancherel ont la même loi asymptotique.

5.2 Équivalence de Baik-Deift-Johansson : l'approche déterminantale

Dans la section 4.2, nous avons vu que si $M_{z, \mathcal{B}(t, \zeta)}$ est la mesure sur les partitions obtenue en modulant le système cohérent des z -mesures $(M_{z, n})_{n \in \mathbb{N}}$ par la loi binomiale négative $\mathcal{B}(t, \zeta)$, alors les processus ponctuels $\mathcal{F}_*(M_{z, \mathcal{B}(t, \zeta)})$ et $\mathcal{D}_*(M_{z, \mathcal{B}(t, \zeta)})$ sont déterminantaux. De plus, on peut donner une expression explicite des noyaux correspondants en termes de fonctions hypergéométriques (voir le théorème 3.3 de [BOO]). Lorsque z tend vers l'infini et $\zeta = \theta/t$ tend vers 0, la loi $M_{z, n}$ tend vers la mesure de Plancherel M_n , et la loi binomiale négative $\mathcal{B}(t, \zeta)$ tend vers la loi de Poisson $\mathcal{P}(\theta)$. Par suite, la **mesure de Plancherel poissonisée**

$$M_{\mathcal{P}(\theta)} = \mathcal{P}(\theta)(|\lambda|) M_{|\lambda|}(\lambda) = \theta^{|\lambda|} e^{-\theta} \left(\frac{\dim \lambda}{|\lambda|!} \right)^2$$

correspond à un processus ponctuel déterminantal $\mathcal{D}_*(M_{\mathcal{P}(\theta)})$ dont le noyau s'obtient par passage à la limite du cas des z -mesures, cf. [BOO, théorèmes 1 et 2].

Une autre approche possible pour le calcul du noyau de ce processus est l'utilisation du formalisme des mesures de Schur; en effet, $M_{\mathcal{P}(\theta)}$ est la mesure de Schur de paramètres $t = t' = (\sqrt{\theta}, 0, 0, \dots)$, voir page 55. Si r est un paramètre réel, on rappelle que la **fonction de Bessel** de première espèce J_r est la série de fonctions

$$J_r(z) = \sum_{n=0}^{\infty} \frac{(-1)^n}{n! \Gamma(r+n+1)} \left(\frac{z}{2} \right)^{r+2n},$$

de sorte que $J_m(z)$ est le coefficient de x^m dans le développement en série de Laurent de $e^{zx/2} e^{-z/2x}$, cf. [Wat44, chapitre 2]. Or, pour les paramètres t et t' précédemment décrits, la fonction $J(x) = \exp(T(x))$ décrite dans la section 4.2 est exactement $e^{zx/2} e^{-z/2x}$ avec $z = 2\sqrt{\theta}$. Par conséquent,

$$\begin{aligned} K_{\mathcal{D}_*(M_{\mathcal{P}(\theta)})}(x, y) &= \sum_{n=0}^{\infty} J_{x+n+1/2}(2\sqrt{\theta}) J_{-y-n-1/2}(-2\sqrt{\theta}) \\ &= \sum_{n=0}^{\infty} J_{x+n+1/2}(2\sqrt{\theta}) J_{y+n+1/2}(2\sqrt{\theta}) \\ &= \frac{\sqrt{\theta} \left(J_{x-1/2}(2\sqrt{\theta}) J_{y+1/2}(2\sqrt{\theta}) - J_{x+1/2}(2\sqrt{\theta}) J_{y-1/2}(2\sqrt{\theta}) \right)}{x-y}, \end{aligned}$$

la seconde égalité découlant de l'invariance de la fonction J par la transformation $x \mapsto -x^{-1}$, et la troisième égalité étant une conséquence des relations de Lommel pour un produit $J_\mu(z) J_\nu(z)$ de fonctions de Bessel, voir [Wat44, §5.4]. En utilisant les formules de Debye pour l'asymptotique des fonctions de Bessel lorsque l'argument complexe z tend vers l'infini (cf. [Wat44, chapitre 8]), ainsi que des techniques de « dépoissonisation » reliant $M_{\mathcal{P}(\theta)}$ et M_n lorsque $\theta \sim n \rightarrow \infty$, on en déduit le résultat suivant :

Proposition 5.6 (Asymptotique du processus ponctuel associé à la mesure de Plancherel, [BOO]). *Soit $(x^n)_{n \in \mathbb{N}} = (x_1^n < \dots < x_k^n)_{n \in \mathbb{N}}$ une suite de vecteurs d'entiers telles que les limites finies ou infinies*

$$a_i = \lim_{n \rightarrow \infty} \frac{x_i^n}{\sqrt{n}} \quad ; \quad d_{ij} = \lim_{n \rightarrow \infty} (a_i - a_j)$$

existent pour tous indices i et j . On note ρ_n la fonction de corrélation discrète du processus ponctuel $\mathcal{D}_*(M_n)$, où M_n est la mesure de Plancherel d'ordre n .

1. Si x^n se scinde en deux parties y^n et z^n dont la distance tend vers l'infini, alors $\lim_{n \rightarrow \infty} \rho_n(x^n) = (\lim_{n \rightarrow \infty} \rho_n(y^n)) \times (\lim_{n \rightarrow \infty} \rho_n(z^n))$. Autrement dit, les comportements locaux en des endroits distincts sont asymptotiquement indépendants.
2. Si x^n reste non scindée et si $a = a_1 = \dots = a_k$ est un point de $] -2, 2[$, alors $\lim_{n \rightarrow \infty} \rho_n(x^n) = \det(S^a(d_{ij}))$, où S^a est le noyau sinus discret :

$$S^a(x) = \frac{\sin(\arccos(a/2)x)}{\pi x}.$$

Au bord de la forme limite (c'est-à-dire avec $a = \pm 2$), le noyau sinus discret convenablement renormalisé dégénère en le noyau d'Airy, ce qui compte tenu de la proposition 5.5 fournit une preuve de l'équivalence de Baik-Deift-Johansson, cf. [BOOoo, §4]. Cette preuve est assurément la plus naturelle, puisqu'elle repose simplement sur l'étude asymptotique des fonctions de corrélation des processus $X = (X_1, X_2, \dots)$ et $Y = (Y_1, Y_2, \dots)$. En contrepartie, elle utilise des propriétés fines des fonctions de Bessel et des fonctions d'Airy, et ces arguments semblent difficiles à généraliser ; nous reviendrons sur ce point dans la section 7.1.

5.3 Équivalence de Baik-Deift-Johansson : l'approche géométrique

Une seconde preuve de nature plus conceptuelle est due à A. Okounkov ([Okooo]), et elle repose sur des interprétations combinatoires (géométriques) des polynômes traciaux de matrices aléatoires et des polynômes traciaux en les éléments de Jucys-Murphy dans l'algèbre du groupe symétrique. Nous concluons ce chapitre et la partie d'introduction de ce mémoire en exposant les arguments principaux de cette approche, qui laisse entrevoir des connections inattendues entre la théorie asymptotique des représentations et des problèmes géométriques tels que la théorie des invariants de Gromov-Witten et le décompte de structures combinatoires sur des surfaces (cartes et revêtements). Cette approche a motivé la recherche d'une interprétation géométrique des observables de diagrammes présentées dans le chapitre 2 ; on renvoie en particulier à [Šo6a, Šo6b] pour de plus amples détails sur ce point.

Compte tenu du théorème de Lévy reliant la convergence en loi d'une suite de variables aléatoires à celle des fonctions caractéristiques (voir l'exemple 5.5 dans [Bil69, chapitre 1]), le théorème 5.1 est équivalent à l'assertion suivante : si $\widehat{X}(\xi) = \sum_{j=1}^{\infty} \exp(X_j \xi)$ et $\widehat{Y}(\xi) = \sum_{j=1}^{\infty} \exp(Y_j \xi)$ sont les transformées de Laplace des mesures $\sum_{j=1}^{\infty} \delta_{X_j}$ et $\sum_{j=1}^{\infty} \delta_{Y_j}$, alors pour tout entier $s \geq 1$ et tous paramètres ξ_1, \dots, ξ_s strictement positifs,

$$\lim_{n \rightarrow \infty} \mathbb{E} \left[\widehat{X}(\xi_1) \cdots \widehat{X}(\xi_s) \right] = \lim_{n \rightarrow \infty} \mathbb{E} \left[\widehat{Y}(\xi_1) \cdots \widehat{Y}(\xi_s) \right],$$

l'existence de la limite pour le terme de droite étant garantie par le théorème de Tracy-Widom. Fixons des paramètres positifs ξ_1, \dots, ξ_s , et des entiers positifs pairs $k_i \sim \xi_i n^{2/3}$; on note $k = \sum_{i=1}^s k_i$. Le polynôme tracial

$$M(k_1, \dots, k_s) = (2\sqrt{n})^{-k} \mathbb{E} \left[\prod_{i=1}^s \text{tr } M^{k_i} \right]$$

est principalement déterminé par les puissances des plus grandes valeurs propres en valeur absolue (donc, celles au bord du spectre). Comme

$$\left(\frac{\lambda_j}{2\sqrt{n}}\right)^{k_i} \sim \left(1 + \frac{Y_j}{n^{2/3}}\right)^{\xi_i n^{2/3}} \rightarrow \exp(Y_j \xi_i),$$

$M(k_1, \dots, k_s)$ a par conséquent la même asymptotique que $\mathbb{E}[\widehat{Y}(\xi_1) \cdots \widehat{Y}(\xi_s)]$. De plus, la formule de Wick fournit un développement topologique de $M(k_1, \dots, k_s)$:

$$M(k_1, \dots, k_s) = \frac{1}{2^k} \sum_S n^{\chi(S)-s} |\text{Map}(k_1, \dots, k_s; S)|,$$

la somme étant effectuée sur les classes d'homéomorphisme de surfaces compactes éventuellement non connexes. Ainsi, du côté des matrices aléatoires, le problème est ramené à celui de l'asymptotique des nombres de cartes $|\text{Map}(k_1, \dots, k_s; S)|$.

Du côté des partitions aléatoires, les analogues des fonctions $M(k_1, \dots, k_s)$ sont les polynômes traciaux $C(k_1, \dots, k_s) = 2^{-k} \sqrt{n}^{s-k} \mathbb{E}[\prod_{i=1}^s (L_i)^{k_i}]$, où les L_i sont les éléments de Jucys-Murphy

$$L_1 = (1, 2) + (1, 3) + \cdots + (1, n) \quad ; \quad L_2 = (2, 3) + \cdots + (2, n) \quad ; \quad \cdots$$

et où $\mathbb{E}[\cdot]$ désigne l'espérance de l'espace de probabilité non commutatif $\mathbb{C}\mathfrak{S}_n$, c'est-à-dire $(1/n!) \text{tr}(\cdot)$. Notons que les L_i ici définis diffèrent légèrement des J_i du paragraphe 1.5 ; ils ont néanmoins les mêmes propriétés, et en particulier, leurs actions sur un module de Specht V^μ ont pour valeurs propres les contenus des cases de μ . Ainsi,

$$\begin{aligned} \mathbb{E}[L_1^k] &= \frac{1}{n!} \sum_{\mu \in \mathcal{Y}_n} (\dim \mu) \varsigma^\mu(L_1^k) = \frac{1}{n!} \sum_{\mu \in \mathcal{Y}_n} \dim \mu \sum_{(\mu_i, i) \text{ coin de } \mu} [\dim(\mu \setminus (\mu_i, i))] (\mu_i - i)^k \\ &= \sum_{\mu \in \mathcal{Y}_n} M_n(\mu) \sum_{(\mu_i, i) \text{ coin de } \mu} \delta_i^*(\mu) (\mu_i - i)^k, \end{aligned}$$

où les $\delta_i^*(\mu) = [\dim(\mu \setminus (\mu_i, i))] / [\dim \mu]$ sont les probabilités de cotransition du processus de Plancherel, c'est-à-dire les probabilités de transition de l'unique processus markovien sur \mathcal{Y} qui est décroissant, compatible avec les règles de branchement et qui conserve les mesures de Plancherel. L'existence d'une forme limite Ω pour les diagrammes sous la mesure de Plancherel implique la convergence en probabilité

$$\forall i, \delta_i^*(\mu) \sqrt{n} \rightarrow 1,$$

voir [Okooo, §3.2.2] — l'asymptotique est également valable pour les probabilités de transition du processus de Plancherel. Fixons des paramètres positifs ξ_1, \dots, ξ_s , et des entiers positifs $k'_i \sim \xi_i n^{1/3}$; on note comme précédemment $k' = \sum_{i=1}^s k'_i$. D'après ce qui précède,

$$\frac{1}{2^{k'_1} (\sqrt{n})^{k'_1-1}} \mathbb{E}[(L_1)^{k'_1}] = \sum_{\mu \in \mathcal{Y}_n} M_n(\mu) \sum_{(\mu_i, i) \text{ coin de } \mu} (\delta_i^*(\mu) \sqrt{n}) \left(\frac{\mu_i - i}{2\sqrt{n}}\right)^{k'_1}$$

et le rapport $|(\mu_i - i) / 2\sqrt{n}|$ est maximal pour $i = 1, 2, \dots$ et $i = \ell(\mu), \ell(\mu) - 1, \dots$. Dans le premier cas,

$$\left(\frac{\mu_i - i}{2\sqrt{n}}\right)^{k'_1} \sim \left(\frac{\mu_i}{2\sqrt{n}}\right)^{k'_1}$$

car $k'_1 \propto n^{1/3}$ et $\mu_i \propto n^{1/2}$; dans le second cas,

$$\left(\frac{\mu_{\ell(\mu)-i} - (\ell(\mu) - i)}{2\sqrt{n}} \right)^{k'_1} \sim \left(-\frac{\mu'_i}{2\sqrt{n}} \right)^{k'_1}.$$

Dans chaque cas, $\delta_i^*(\mu) \sqrt{n}$ tend vers 1. En utilisant l'invariance par conjugaison des diagrammes de la mesure de Plancherel, on conclut que :

$$\frac{1}{2^{k'_1} (\sqrt{n})^{k'_1-1}} \mathbb{E} \left[(L_1)^{k'_1} \right] \sim \sum_{\mu \in \mathcal{O}_n} M_n(\mu) \left\{ \sum_{i=1}^{\infty} \left(\frac{\mu_i}{2\sqrt{n}} \right)^{k'_1} + \left(-\frac{\mu'_i}{2\sqrt{n}} \right)^{k'_1} \right\}.$$

En réitérant ce raisonnement avec les éléments L_2, \dots, L_s , on voit plus généralement que le polynôme tracial en les éléments de Jucys-Murphy $C(k'_1, \dots, k'_s)$ a la même asymptotique que

$$\sum_{\mu \in \mathcal{O}_n} M_n(\lambda) \left\{ \sum_{i_1, \dots, i_s=1}^{\infty} \frac{(\mu_{i_1})^{k'_1} \cdots (\mu_{i_s})^{k'_s}}{(2\sqrt{n})^{k'}} + \dots \right\}$$

où les points de suspension dans la somme correspondent à $2^s - 1$ autres termes mettant en jeu les $-\mu'_i$. De plus, comme dans le cas des matrices aléatoires, $C(k'_1, \dots, k'_s)$ a la même asymptotique que $\mathbb{E}[\widehat{X}(\xi_1) \cdots \widehat{X}(\xi_s)]$, car

$$\left(\frac{\mu_j}{2\sqrt{n}} \right)^{k'_i} \sim \left(1 + \frac{X_j}{n^{1/3}} \right)^{\xi_i n^{1/3}} \rightarrow \exp(X_j \xi_i).$$

La preuve « combinatoire » de l'équivalence de Baik-Deift-Johansson est donc ramenée à l'équivalence asymptotique $M(k_1, \dots, k_s) \sim C(k'_1, \dots, k'_s)$.

Ceci étant, on peut donner un développement topologique de $C(k'_1, \dots, k'_s)$ analogue à celui des fonctions M , mais en termes de revêtements ramifiés de la sphère. Pour commencer, on remplace chaque $L_i = \sum_{j>i} (i, j)$ par

$$\widetilde{L}_i = L_i - \sum_{j=i+1}^s (i, j) = (i, s+1) + (i, s+2) + \cdots + (i, n);$$

ceci ne change pas l'asymptotique des polynômes traciaux $C(k'_1, \dots, k'_s)$ — nous noterons $\widetilde{C}(k'_1, \dots, k'_s)$ les polynômes modifiés. Maintenant, comme $\mathbb{E}[\sigma] = \mathbb{1}_{\sigma=\text{id}}$ pour toute permutation σ , $\widetilde{C}(k'_1, \dots, k'_s)$ est le nombre de solutions dans \mathfrak{S}_n de l'équation

$$(1, \tau_1)(1, \tau_2) \cdots (1, \tau_{k'_1})(2, \tau_{k'_1+1}) \cdots (2, \tau_{k'_1+k'_2}) \cdots (s, \tau_{k'}) = \text{id}$$

avec chaque τ_i dans $\llbracket s+1, n \rrbracket$. Le groupe \mathfrak{S}_{n-s} agit sur cet intervalle, et donc sur l'ensemble des solutions de l'équation par conjugaison; de plus, l'orbite d'une solution τ est de cardinal $(n-s)(n-s-1) \cdots (n-s-d(\tau)-1) \sim n^{d(\tau)}$, où $d(\tau) = \text{card} \{\tau_1, \dots, \tau_{k'}\}$. Ainsi,

$$\widetilde{C}(k'_1, \dots, k'_s) \sim \sum_{\{\tau\}/\mathfrak{S}_{n-s}} n^{d(\tau)}.$$

À toute solution $\tau = \{\tau'_1, \dots, \tau'_k\}$, on peut associer un revêtement ramifié de la sphère $\mathbb{C}\mathbb{P}^1$ en procédant comme suit⁵ :

5. On renvoie à [LZ04, §1.2] pour des détails sur le lien entre équations dans \mathfrak{S}_n et revêtements ramifiés de $\mathbb{C}\mathbb{P}^1$ donné par la monodromie dans une fibre au-dessus d'un point base. Ce point sera également évoqué dans le chapitre 13, avec quelques rappels sur la notion de revêtement ramifié.

1. Sur la droite projective complexe \mathbb{CP}^1 , on place un point base 0, et on répartit k' points autour — on les notera $1, 2, \dots, k'$. Le point 0 sera non ramifié, et les points $1, \dots, k'$ seront simplement ramifiés.
2. Le revêtement a $s + d(\tau)$ feuillettes, dont s feuillettes spéciaux marqués au-dessus du point base 0.
3. Si γ_i est la boucle issue du point base et contournant le point i , alors la monodromie de cette boucle au dessus de 0 est la transposition du feuillet spécial i avec un feuillet non marqué τ_i , voir la figure 5.3.

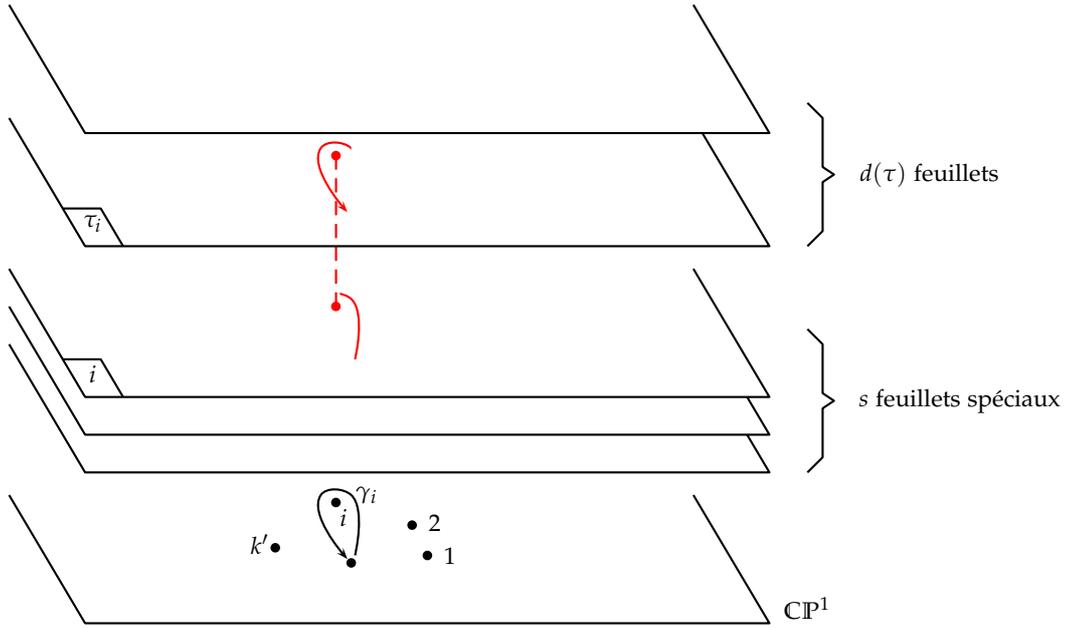


FIGURE 5.3 – Revêtement ramifié de Jucys-Murphy associé à une équation τ .

Nous noterons $\text{Cov}_{\text{JM}}(k'_1, \dots, k'_s)$ l'ensemble des classes d'équivalence de **revêtements ramifiés de Jucys-Murphy** de paramètres k'_1, \dots, k'_s , et si S est une surface orientable compacte, nous noterons $\text{Cov}_{\text{JM}}(k'_1, \dots, k'_s; S)$ l'ensemble constitué de ceux qui sont homéomorphes à S . Par la formule de Riemann-Hurwitz (cf. [DS94, p. 39]), on a alors $\chi(S) = 2(s + d(\tau)) - k'$, donc

$$n^{d(\tau)} = n^{\frac{\chi(S)-s}{2}} \times n^{\frac{k'-s}{2}}.$$

D'autre part, deux équations donnent le même revêtement ramifié marqué si et seulement si l'une peut être obtenue à partir de l'autre par conjugaison par \mathfrak{S}_{n-s} . On en déduit le développement topologique (asymptotique)

$$C(k'_1, \dots, k'_s) \sim \frac{1}{2^k} \sum_S n^{\frac{\chi(S)-s}{2}} |\text{Cov}_{\text{JM}}(k'_1, \dots, k'_s; S)|,$$

la somme étant de nouveau effectuée sur les classes d'homéomorphisme de surfaces compactes. L'équivalence asymptotique $M(k_1, \dots, k_s) \sim C(k'_1, \dots, k'_s)$ découle dès lors du résultat suivant :

Proposition 5.7 (Équivalence asymptotique des nombres de cartes et des nombres de revêtements de Jucys-Murphy, [Ok000]). *Si $k_i \sim \zeta_i t$ avec t tendant vers l'infini, alors pour tout genre g , il existe un fonction $m_g(\zeta_1, \dots, \zeta_s)$ telle que*

$$\frac{|\text{Map}(k_1, \dots, k_s; \mathbb{T}_g)|}{2^k} \sim t^{3g-3+3s/2} m_g(\zeta_1, \dots, \zeta_s).$$

Le résultat s'étend sans difficulté au cas d'une surface S non connexe. Dans le même contexte,

$$|\text{Map}(k_1, \dots, k_s; \mathbb{T}_g)| \sim |\text{Cov}_{\text{JM}}(k_1, \dots, k_s; \mathbb{T}_g)|,$$

ce qui implique l'équivalence asymptotique $M(k_1, \dots, k_s) \sim C(k'_1, \dots, k'_s)$ avec les normalisations des k_i et des k'_i en $n^{2/3}$ et $n^{1/3}$.

Exemple. Si $g = 0$ et $s = 1$, alors $|\text{Map}(2k; g = 0)|$ est le nombre d'appariements sans croisement des côtés d'un $2k$ -gone, ou encore le nombre d'arbres généraux à k arêtes ; c'est donc le nombre de Catalan $C_k = \frac{1}{k+1} \binom{2k}{k}$. D'autre part, pour tout revêtement de Jucys-Murphy, appelons valence d'une feuille F non spéciale le nombre de points i de $\llbracket 1, k \rrbracket$ tels que la monodromie autour de i permute la feuille F . Alors, $\text{val}(F) \geq 2$, et

$$\sum_F (\text{val}(F) - 2) = k - 2d(\tau) = 2s - \chi(S).$$

Lorsque $g = 0$ et $s = 1$, $k = 2d(\tau)$, et la valence de tous les feuillet non spéciaux est donc 2. Autrement dit, $|\text{Cov}_{\text{JM}}(2k; g = 0)|$ est le nombre de solutions de l'équation

$$(1, i_1)(1, i_2) \cdots (1, i_{2k}) = \text{id}$$

telles que les i_j soient réunis par paires. Il n'est pas difficile de voir que ce nombre est encore le nombre de Catalan C_k ; dans ce cas particulier, l'équivalence asymptotique $|\text{Map}| \sim |\text{Cov}_{\text{JM}}|$ est donc une égalité. De plus, par la formule Stirling, la fonction $m_0(\zeta)$ existe bien, et elle vaut $\frac{1}{\sqrt{\pi}} \left(\frac{\zeta}{2}\right)^{-3/2}$.

La preuve de la proposition 5.7 donnée par Okounkov dans [Ok000] repose sur l'utilisation des **graphes de rubans** comme objets combinatoires intermédiaires entre les cartes et les revêtements de Jucys-Murphy. Les graphes de rubans (métrisés) fournissent un modèle combinatoire des espaces de module des courbes $\mathcal{M}_{s,g}$ (cf. [Kon92]), et on peut interpréter l'expression $M(k_1, \dots, k_s)$ comme une somme de Riemann d'une intégrale sur le modèle combinatoire de Kontsevich. D'autre part, il est bien connu que le décompte des revêtements ramifiés de la sphère est lié à la théorie de l'intersection sur les espaces de modules $\mathcal{M}_{s,g}$ — ce qu'il est convenu d'appeler la **théorie de Gromov-Witten** — et au calcul d'intégrales de classes de cohomologie sur le compactifié de cet espace (cf. [ELSV01, Zvo05] et le paragraphe 13.3). Par conséquent, les deux expressions $M(k_1, \dots, k_s)$ et $C(k'_1, \dots, k'_s)$ sont certainement des discrétisations d'expressions intégrales sur les espaces de modules $\mathcal{M}_{s,g}$, ce qui laisse entrevoir une connexion entre permutations aléatoires et espaces de modules analogue aux résultats de Kontsevich dans le contexte des matrices aléatoires, voir [Kon92, OP01, OP02]. Ceci apporte une motivation supplémentaire à l'étude asymptotique des représentations, et clôt notre première partie d'introduction.

