ANALYSE HARMONIQUE SUR DES GROUPES COMPACTS

P.-L. MÉLIOT

Si X est une variable aléatoire réelle, de nombreuses informations peuvent être extraites de sa *transformée de Fourier* $\phi_X(\xi) = \mathbb{E}[\mathrm{e}^{\mathrm{i}\xi X}]$, et des *moments* $\mathbb{E}[X^k]$ de X, qui sont les coefficients du développement en série de ϕ_X :

$$\phi_X(\xi) = \sum_{k=0}^{\infty} \mathbb{E}[X^k]\,\frac{(\mathrm{i}\xi)^k}{k!}.$$

En effet :

(1) Si $\phi_{X_n}(\xi) \to \phi_X(\xi)$, alors la suite de variables aléatoires $(X_n)_{n\in\mathbb{N}}$ converge en loi vers la variable aléatoire X :

$$\forall s \in \mathbb{R},\ \ \mathbb{P}[X_n \leq s] \to_{n\to\infty} \mathbb{P}[X \leq s].$$

(2) On peut mesurer cette convergence à l'aide de la distance de Kolmogorov :

$$\begin{aligned}d_{\mathrm{Kol}}(X_n, X) &= \sup_{s\in\mathbb{R}} |\mathbb{P}[X_n \leq s] - \mathbb{P}[X \leq s]| \\ &\leq_{(\text{Feller})} \frac{1}{\pi}\int_{-T}^{T} \frac{1}{\xi}\,|\phi_{X_n}(\xi) - \phi_X(\xi)|\,d\xi + \frac{24m}{\pi T}, \quad \text{où } m = \sup_{s\in\mathbb{R}} \frac{d\mathbb{P}_X(s)}{ds}.\end{aligned}$$

Exemple (Théorème central limite). Si $(A_n)_{n\in\mathbb{N}}$ est une suite de variables i.i.d. avec $\mathbb{E}[A_n] = 0$ et $\mathbb{E}[(A_n)^2] = \sigma^2$, alors

$$\frac{A_1 + A_2 + \cdots + A_n}{\sigma\sqrt{n}} = X_n \to_{\mathrm{loi}} X = \mathcal{N}(0,1).$$

De plus, si $\mathbb{E}[|A_n|^3] = \rho < \infty$, alors

$$d_{\mathrm{Kol}}(X_n, X) \leq \frac{3\rho}{\sigma^3\sqrt{n}} \quad \text{(estimées de Berry-Esseen).}$$

Objectif. Développer des outils semblables à la transformée de Fourier et aux moments pour des *matrices* aléatoires à valeurs dans des groupes compacts, *e.g.* $\mathrm{U}(N)$, $\mathrm{SU}(N)$, $\mathrm{SO}(N)$, $\mathrm{USp}(N)$, *etc.*

On va présenter cette théorie, avec en vue les deux résultats suivants :

(1) Diaconis-Shahshahani (1994). Si $M \sim \mathrm{Haar}(\mathrm{U}(N))$, alors les variables aléatoires $\operatorname{tr} M$, $\frac{1}{\sqrt{2}}\operatorname{tr} M^2$, $\frac{1}{\sqrt{3}}\operatorname{tr} M^3, \ldots$ convergent des gaussiennes complexes standards indépendantes.

(2) Méliot (2013). Si $(M_t)_{t\in\mathbb{R}_+}$ est un mouvement brownien sur $\mathrm{U}(N)$, alors

$$d_{\mathrm{TV}}(M_t, \mathrm{Haar}) = \begin{cases} 1 - o(1) & \text{si } t = 2(1-\varepsilon)\log N, \\ o(1) & \text{si } t = 2(1+\varepsilon)\log N. \end{cases}.$$

1. Transformée de Fourier non commutative sur les groupes compacts

Lorsqu'on étudie des variables aléatoires réelles, on dispose de la transformée de Fourier

$$\begin{aligned} \mathrm{L}^2(\mathbb{R}, dx) &\to \mathrm{L}^2(\mathbb{R}, dx) \\ f &\mapsto \widehat{f}(\xi) = \int_{\mathbb{R}} f(x)\,\mathrm{e}^{\mathrm{i}x\xi}\,dx \end{aligned}$$

et de la formule d'inversion de Parseval

$$f(x) = \frac{1}{2\pi}\int_{\mathbb{R}} \widehat{f}(\xi)\,\mathrm{e}^{-\mathrm{i}x\xi}\,d\xi,$$

et on les utilise le plus souvent avec f densité d'une variable aléatoire. Si f est maintenant une fonction sur un groupe compact G, l'analogue de cette théorie est fournie par le théorème de Peter-Weyl. Soit G un groupe compact. Une *représentation* (linéaire, complexe) de G est donnée par une paire (V, ρ), où V est un espace linéaire complexe de dimension finie, et $\rho : G \to \mathrm{GL}(V)$ est un morphisme de groupes continu. On peut alors faire agir G sur V par :

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

La représentation (ρ, V) est dite *irréductible* s'il n'existe pas de sous-espace de V non-trivial et stable par G. D'autre part, on peut toujours trouver un produit scalaire sur V qui est invariant par G, et donc tel que le morphisme ρ prenne ses valeurs dans le groupe unitaire $\mathrm{U}(V)$ pour cette structure. Dans ce qui suit, $*$ désigne l'adjonction sur $\mathrm{End}(V)$ liée à cette structure.

Théorème 1 (Peter-Weyl)**.** *Soit dg la mesure de Haar sur G, et $\widehat{G}$ l'ensemble des classes d'isomorphisme de représentations irréductibles de G. On note une classe $\lambda = (V^\lambda, \rho^\lambda)$, et $d^\lambda = \dim_{\mathbb{C}} V^\lambda$. L'algèbre $\mathrm{End}(V^\lambda)$ est munie du produit scalaire $\langle M \mid N\rangle = d^\lambda \operatorname{tr}(M^* N)$.*

(1) L'application

$$\begin{aligned} \mathrm{L}^2(G, dg) &\to \bigoplus_{\lambda \in \widehat{G}}^{\perp} \mathrm{End}(V^\lambda) \\ f &\mapsto \sum_{\lambda \in \widehat{G}} \widehat{f}(\lambda), \quad \text{avec } \widehat{f}(\lambda) = \int_G f(g)\,\rho^\lambda(g)\,dg \end{aligned}$$

est un isomorphisme d'algèbres ; une isométrie d'espaces de Hilbert ; et un morphisme de (G,G)-bimodules. Ici, $\mathrm{L}^2(G)$ est muni du produit de convolution et de sa structure usuelle d'espace hilbertien

$$\langle f_1 \mid f_2\rangle = \int_G \overline{f_1(g)}\, f_2(g)\,dg.$$

(2) On a la formule d'inversion :

$$f(g) = \sum_{\lambda \in \widehat{G}} d^\lambda \operatorname{tr}\Big(\widehat{f}(\lambda)\,\rho^\lambda(g^{-1})\Big).$$

Conséquence. Pour les fonctions sur G, le rôle de la transformée de Fourier est joué par la famille de matrices $(\widehat{f}(\lambda))_{\lambda \in \widehat{G}}$.

Un cas particulier est celui des fonctions f invariantes sur les classes de conjugaison de G (penser : la densité d'une mesure qui ne dépend que des valeurs propres de la matrice aléatoire). Notons $\mathrm{ch}^\lambda(\cdot) = \mathrm{tr}\,\rho^\lambda(\cdot)$ le *caractère* de la représentation irréductible $\lambda = (V^\lambda, \rho^\lambda)$.

Proposition 2. *Une fonction f appartient à $\mathrm{L}^2(G, dg)^G$ si et seulement si c'est une combinaison linéaire des caractères irréductibles. Alors,*

$$f(g) = \sum_{\lambda \in \widehat{G}} \left\langle \mathrm{ch}^\lambda \,\middle|\, f \right\rangle \mathrm{ch}^\lambda(g).$$

Dans ce cadre, si f est la densité d'un modèle de matrices aléatoires X à valeurs dans G et invariant par conjugaison, alors les moments de X vont être remplacés par les coefficients

$$\left\langle f \,\middle|\, \mathrm{ch}^\lambda \right\rangle = \mathbb{E}\left[\mathrm{ch}^\lambda(X)\right].$$

Cas particulier. Si $f = 1$ et si X suit la mesure de Haar, alors $\mathbb{E}[\mathrm{ch}^\lambda(X)] = 0$ à moins que λ soit la représentation triviale de dimension 1.

2. Caractères des groupes unitaires et fonctions de Schur

Supposons que $G = \mathrm{U}(N)$ soit le *groupe unitaire* d'ordre N. Une fonction $M \mapsto f(M)$ sur G est invariante sur les classes de conjugaison de G si et seulement si elle ne dépend que des valeurs propres $x_1 = \mathrm{e}^{\mathrm{i}\theta_1}, \ldots, x_N = \mathrm{e}^{\mathrm{i}\theta_N}$ de M. Alors, son espérance sous la mesure de Haar se réécrit sous la forme

$$\begin{aligned}\mathbb{E}_{\mathrm{Haar}}[f(M)] &= \int_G f(g)\,dg \\ &=_{(\mathrm{Weyl})} \frac{1}{(2\pi)^N\,N!} \int_{[0,2\pi]^N} f(\mathrm{e}^{\mathrm{i}\theta_1}, \ldots, \mathrm{e}^{\mathrm{i}\theta_N}) \prod_{1 \leq i < j \leq N} \left| \mathrm{e}^{\mathrm{i}\theta_i} - \mathrm{e}^{\mathrm{i}\theta_j} \right|^2 d\theta_1 \cdots d\theta_N.\end{aligned}$$

La théorie de Cartan-Weyl (systèmes de racines, réseaux des poids) permet d'identifier les représentations irréductibles de $\mathrm{U}(N)$. Appelons *poids dominant* de $\mathrm{U}(N)$ une suite décroissante d'entiers $\lambda = (\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_N)$.

Théorème 3. *L'ensemble des classes d'isomorphisme de représentations irréductibles de* $\mathrm{U}(N)$ *est en bijection avec l'ensemble des poids dominants. Si λ est une tel poids, alors*

$$d^\lambda = \prod_{1 \leq i < j \leq N} \frac{\lambda_i - \lambda_j + j - i}{j - i},$$

et le caractère irréductible associé est donné par la fonction de Schur de type λ :

$$\mathrm{ch}^\lambda(M) = s_\lambda(z_1, \ldots, z_N) = \frac{\det(z_i^{\lambda_j + N - j})_{1 \leq i,j \leq N}}{\det(z_i^{N-j})_{1 \leq i,j \leq N}}.$$

En particulier, si λ et ρ sont deux poids dominants, alors

$$\begin{aligned}\delta_{\lambda,\rho} &= \left\langle \mathrm{ch}^\lambda \,\middle|\, \mathrm{ch}^\rho \right\rangle_{\mathrm{U}(N)} \\ &= \int_{[0,2\pi]^N} \overline{s_\lambda(\mathrm{e}^{\mathrm{i}\theta_1}, \ldots, \mathrm{e}^{\mathrm{i}\theta_N})}\, s_\rho(\mathrm{e}^{\mathrm{i}\theta_1}, \ldots, \mathrm{e}^{\mathrm{i}\theta_N})\, \frac{|\Delta(\mathrm{e}^{\mathrm{i}\theta_1}, \ldots, \mathrm{e}^{\mathrm{i}\theta_N})|^2\,d\theta}{(2\pi)^N\,N!}.\end{aligned}$$

Conséquence. Les fonctions de Schur et leurs espérances jouent le rôle des moments pour un modèle de matrice aléatoire invariant par conjugaison sur $\mathrm{U}(N)$.

Remarque. Un cas particulier est celui où le poids λ ne contient que des entiers positifs ; on parle alors de *partition* de longueur plus petite que N, et la taille de la partition est $|\lambda| = \lambda_1 + \lambda_2 + \cdots + \lambda_N$. Dans ce cas, $s_\lambda(x_1, \ldots, x_N)$ est un polynôme homogène de degré $|\lambda|$ et symétrique en les variables $x_1, \ldots, x_N$:

$$\forall \sigma \in \mathfrak{S}(N),\ \ s_\lambda(x_{\sigma(1)}, \ldots, x_{\sigma(N)}) = s_\lambda(x_1, \ldots, x_N).$$

N'importe quel polynôme symétrique homogène de degré n en les variables $x_1, \ldots, x_N$ se décompose en une combinaison linéaire de fonctions de Schur s_λ avec $|\lambda| = n$.

3. Normalité asymptotique des traces des puissances

Soit M une matrice aléatoire suivant la mesure de Haar sur $\mathrm{U}(N)$. On s'intéresse aux variables aléatoires complexes $X_k = \mathrm{tr}\, M^k$. En termes des valeurs propres,

$$\mathrm{tr}\, M^k = \sum_{i=1}^{N} (\mathrm{e}^{\mathrm{i}\theta_i})^k = p_k(\mathrm{e}^{\mathrm{i}\theta_1}, \ldots, \mathrm{e}^{\mathrm{i}\theta_N}),$$

où $p_k(x_1, \ldots, x_N) = \sum_{i=1}^{N} (x_i)^k$ est la k-ième somme de puissances. On note aussi

$$p_\mu(X) = p_{\mu_1}(X) p_{\mu_2}(X) \cdots p_{\mu_\ell}(X), \quad \mu \text{ partition.}$$

Théorème 4 (Diaconis-Shahshahani). *Soit $M \sim Haar(\mathrm{U}(N))$. Lorsque N tend vers l'infini, le vecteur aléatoire*

$$\left(\mathrm{tr}\, M, \frac{1}{\sqrt{2}}\, \mathrm{tr}\, M^2, \frac{1}{\sqrt{3}}\, \mathrm{tr}\, M^3, \ldots \right)$$

converge en loi vers un vecteur de gaussiennes complexes indépendantes standards $\mathcal{N}_{\mathbb{C}}(0,1) = \mathcal{N}_{\mathbb{R}}(0, \frac{1}{2}) + \mathrm{i}\mathcal{N}_{\mathbb{R}}(0, \frac{1}{2})$.

On rappelle que les moments d'une gaussienne complexe sont $\mathbb{E}\left[\overline{X^m} X^{m'}\right] = \delta_{m,m'}\, m!$. Par conséquent, le théorème est équivalent à la convergence des moments :

$$\lim_{N \to \infty} \mathbb{E}\left[\left(\prod_{i=1}^{s} \overline{(X_i)^{m_i}} \right) \left(\prod_{i=1}^{s} (X_i)^{m_i'} \right) \right] = \prod_{i=1}^{s} \delta_{m_i, m_i'}\, i^{m_i}\, m_i!.$$

Si μ est une partition de taille n avec m_1 parts de taille 1, m_2 parts de taille 2, *etc.*, on note $z_\mu = \prod_{i=1}^{s} i^{m_i}\, m_i!$. C'est aussi la taille du centralisateur d'une permutation de type cyclique μ dans $\mathfrak{S}(n)$. On note aussi $X_\mu = \prod_{i=1}^{s} (\mathrm{tr}\, M^i)^{m_i}$. On va montrer que pour N assez grand, on a en fait

$$\mathbb{E}_{\mathrm{U}(N)}\left[\overline{X_\mu} X_{\mu'} \right] = \delta_{\mu,\mu'}\, z_\mu.$$

On peut supposer $|\mu| = |\mu'| = n$, car sinon l'espérance vaut 0 par invariance de la mesure par action du cercle. Comme $p_\mu \in \mathrm{Sym}$, on peut la décomposer sur la base des fonctions de Schur : $p_\mu = \sum_{|\lambda| = |\mu|} c_{\lambda\mu}\, s_\lambda$. Alors,

$$\mathbb{E}_{\mathrm{U}(N)}\left[\overline{X_\mu} X_{\mu'} \right] = \left\langle p_\mu \,\middle|\, p_{\mu'} \right\rangle_{\mathrm{U}(N)} = \sum_{|\lambda| = n} \overline{c_{\lambda\mu}}\, c_{\lambda\mu'}$$

car les fonctions de Schur sont orthonormales par rapport à $\langle \cdot \mid \cdot \rangle_{\mathrm{U}(N)}$, si $N \geq n$. Pour calculer $c_{\lambda\mu}$, on utilise :

Théorème 5 (Schur-Weyl)**.** *Le commutant de l'action de* $\mathrm{U}(N)$ *sur* $(\mathbb{C}^N)^{\otimes n}$ *est l'algèbre* $\mathbb{C}\mathfrak{S}(n)$ *engendrée par le groupe symétrique (elle agit par permutation des tenseurs). On a la décomposition comme* $(\mathrm{U}(N), \mathfrak{S}(n))$*-bimodule*

$$(\mathbb{C}^N)^{\otimes n} = \sum_{|\lambda|=n} {}_{\mathrm{U}(N)\hookrightarrow}(V^\lambda) \otimes_{\mathbb{C}} (S^\lambda)_{\hookleftarrow \mathfrak{S}(n)},$$

où V^λ *est la représentation irréductible de* $\mathrm{U}(N)$ *de caractère* s_λ*, et* $(S^\lambda)_{|\lambda|=n}$ *est une collection complète de représentations irréductibles de* $\mathfrak{S}(n)$*.*

Considérons alors l'action de la paire $(\mathrm{diag}(x_1, \ldots, x_N), \sigma_\mu)$, où les x_i sont des nombres complexes de module 1, et où $\sigma_\mu \in \mathfrak{S}(n)$ est de type cyclique μ. La trace de l'action de cet opérateur sur $(\mathbb{C}^N)^{\otimes n}$ est $p_\mu(x_1, \ldots, x_N)$, donc si $N \geq n$, alors

$$p_\mu(x_1, \ldots, x_N) = \sum_{|\lambda|=|\mu|} s_\lambda(x_1, \ldots, x_N) \, \mathrm{ch}^{S^\lambda}(\sigma_\mu)$$

et $c_{\lambda\mu} = \mathrm{ch}^{S^\lambda}(\sigma_\mu)$. On doit donc calculer

$$\sum_{|\lambda|=n} \overline{\mathrm{ch}^{S^\lambda}(\sigma_\mu)} \, \mathrm{ch}^{S^\lambda}(\sigma_{\mu'}) = \sum_{|\lambda|=n} \mathrm{ch}^{S^\lambda}(\sigma_\mu^{-1}) \, \mathrm{ch}^{S^\lambda}(\sigma_{\mu'}).$$

Or, on a aussi par le théorème de Peter-Weyl

$$\mathbb{C}\mathfrak{S}(n) = \mathrm{L}^2(\mathfrak{S}(n)) = \sum_{|\lambda|=n} {}_{\mathfrak{S}(N)\hookrightarrow}(S^\lambda) \otimes_{\mathbb{C}} (S^{\lambda *})_{\hookleftarrow \mathfrak{S}(n)},$$

donc la quantité ci-dessus est la bitrace de l'action par multiplication de $(\sigma_\mu^{-1}, \sigma_{\mu'})$ sur $\mathbb{C}\mathfrak{S}(n)$. C'est donc le nombre de permutations τ telles que $\sigma_\mu^{-1}\tau\sigma_{\mu'} = \tau$, c'est-à-dire z_μ si $\mu = \mu'$, et 0 sinon.

4. Mouvement brownien sur $\mathrm{U}(N)$

Le mouvement brownien sur $\mathrm{U}(N)$ est l'unique processus markovien $(M_t)_{t \in \mathbb{R}_+}$ issu de $M_0 = I_N$, et de générateur infinitésimal $\frac{1}{2}\Delta$, où Δ est l'opérateur de Laplace-Beltrami associé au produit scalaire $\langle X \mid Y \rangle = -N \,\mathrm{tr}(XY)$ sur l'algèbre de Lie $\mathfrak{u}(N)$. Les densités p_t des matrices aléatoires M_t par rapport à la mesure de Haar sont donc les solutions de l'équation aux dérivées partielles

$$\frac{dp_t}{dt} = \frac{1}{2} \, \Delta p_t,$$

et elles sont invariantes par conjugaison. L'opérateur Δ agit diagonalement sur les fonctions de Schur s_λ, donc,

$$p_t(M) = \sum_\lambda d^\lambda \, \mathrm{e}^{-\frac{c_\lambda t}{2}} \, s_\lambda(x_1, \ldots, x_N),$$

où $-c_\lambda$ est la valeur propre de Δ agissant sur s_λ, et est donnée par la formule

$$c_\lambda = \frac{1}{N} \left(\sum_{i=1}^N (\lambda_i)^2 + (N + 1 - 2i)\lambda_i \right).$$

On souhaite mesurer la distance en variation totale

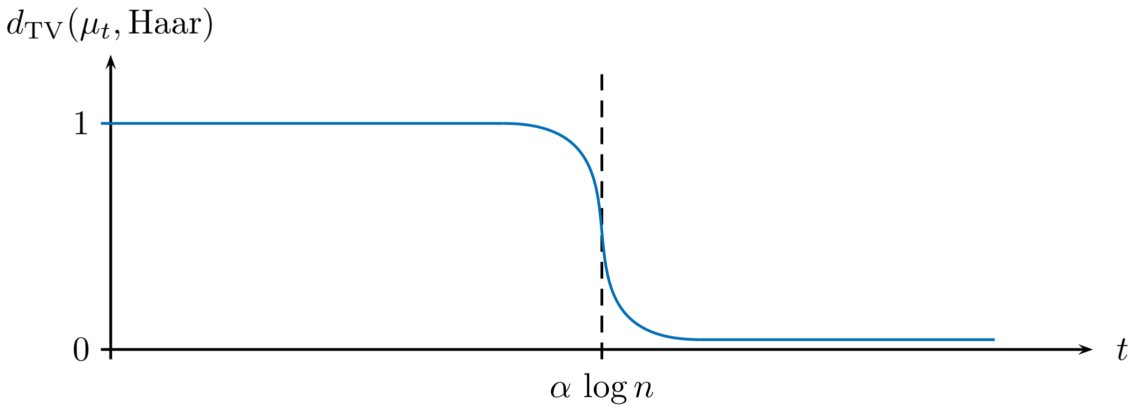
$$d_{\mathrm{TV}}(M_t, \mathrm{Haar}) = \frac{1}{2}\|p_t - 1\|_{\mathrm{L}^1(\mathrm{SU}(N),\mathrm{Haar})} \leq \frac{1}{2}\|p_t - 1\|_{\mathrm{L}^2(\mathrm{SU}(N),\mathrm{Haar})}$$

$$\begin{aligned}\left(\|p_t - 1\|_{\mathrm{L}^2(\mathrm{SU}(N),\mathrm{Haar})}\right)^2 &\leq \sum_{\lambda} |\,\langle s_\lambda \mid p_t - 1\rangle\,|^2 = {\sum_{\lambda}}' |\,\langle s_\lambda \mid p_t\rangle\,|^2 \\ &= {\sum_{\lambda}}' (d^\lambda)^2\, \mathrm{e}^{-c_\lambda t}\end{aligned}$$

où le $'$ indique que l'on retire le poids trivial $\lambda = (0, \ldots, 0)$ de la somme. Notons que d^λ croît avec les parts de λ, mais cette croissance est compensée par la décroissance de $\mathrm{e}^{-c_\lambda t}$.

Théorème 6 (Méliot)**.** *Pour $t = 2 \log N$, $(d^\lambda)^2\, \mathrm{e}^{-c_\lambda t}$ est uniformément borné (indépendamment de N et de λ). Pour $t = 2(1+\varepsilon)\log N$, la somme $\sum'_\lambda (d^\lambda)^2\, \mathrm{e}^{-c_\lambda t}$ est un $o(1)$.*

On peut montrer que réciproquement, pour $t = 2(1-\varepsilon)\log N$, la distance en variation totale reste proche de 1 ; on a donc un phénomène de coupure.



Le résultat se généralise à n'importe quel mouvement brownien tracé sur un groupe de Lie compact G, ou sur un espace symétrique G/K (sphères, espaces projectifs, grassmanniennes, *etc.*).