

Théorème central limite pour les polynômes homogènes

FMC, É. Parcollet
Sous la supervision de M. Fathi

Mémoire de première année
2023-2024

Contents

1	Introduction	1
1.1	Le théorème central limite	1
1.2	Le sens direct	2
2	La méthode de Stein	3
3	Le théorème de caractérisation	3
3.1	Le cas des polynômes homogènes	3
3.2	Première borne par méthode de Stein	4
3.3	Gradients itérés du carré du champ	5
3.4	Méthode des chaos	6
3.5	Une condition suffisante	8
4	Conclusion	9
5	Annexe	10

1 Introduction

1.1 Le théorème central limite

Le théorème central limite (TCL) est un résultat fondamental de probabilités. Il établit la convergence en loi de la somme d'une suite de variables aléatoires vers la loi normale, permettant d'expliquer pourquoi les répartitions gaussiennes apparaissent naturellement lorsque on s'intéresse à des phénomènes naturels. Il s'énonce comme suit:

Théorème 1 (TCL). *Soit (X_i) une suite de variables variables aléatoires (v.a.) réelles indépendantes et identiquement distribuées (i.i.d.), d'espérance $\mu = \mathbb{E}(X_1)$ et d'écart-type $\sigma = \sqrt{\mathbb{V}(X_1)}$. Alors*

$$\frac{1}{\sigma\sqrt{n}} \left(\sum_{i=1}^n X_i - n\mu \right) \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N}(0, 1).$$

La généralité du résultat suscite de suite des envies de généralisation, dont certaines sont connues. Lorsque ces variables n'admettent pas de second moment un même résultat de convergence s'applique mais avec une différente limite. Sans second moment c'est l'expression $\frac{1}{\sqrt{n}}(\sum_{i=1}^n X_i - n\mu)$ qui converge, et lorsqu'elles n'admettent pas de premier moment c'est seulement $\frac{1}{\sqrt{n}} \sum_{i=1}^n X_i$ qui converge. La limite dépend du plus grand moment accepté par ces variables α , ainsi que d'un paramètre d'asymétrie β . Ces limites sont appelées les lois stables.

Théorème 2 ([1, Théorème C.2]). *Les lois stables non dégénérées (réduites à un atome) ont pour fonction caractéristique*

$$t \mapsto \exp \left(i\mu t + \frac{\sigma^\alpha}{\alpha} (-|t|^\alpha + it|t|^{\alpha-1}\beta\Phi_\alpha(t)) \right); \Phi_\alpha(t) = \begin{cases} \tan(\frac{\pi\alpha}{2}), & \text{si } \alpha \neq 1 \\ -\frac{2}{\pi} \ln |t|, & \text{si } \alpha = 1 \end{cases};$$

où $\alpha \in]0, 2]$, $\mu \in \mathbb{R}$ et $\beta \in [-1, 1]$. Le cas $\alpha = 2$ est le TCL classique.

Comme le font remarquer remarquer Chen, Goldstein et Shao dans [2, 1.1], nous pouvons aussi relaxer la condition d'indépendance sur les variables, ce qui donne lieu à de nombreuses déclinaisons et théorèmes sans l'universalité des résultats précédents.

Notre but est de démontrer une généralisation du TCL à des polynômes symétriques qui violent la linéarité des combinaisons précédentes. Nous présenterons le théorème du quatrième moment de Nualart et Peccati qui donne un critère optimal de convergence vers une loi gaussienne.

Ce critère s'applique de façon indépendante par degré: seuls les polynômes homogènes sont à considérer. De plus les monômes en un seul des X_i sont traités par l'application du TCL aux X_i^k . Nous utilisons donc une définition restreinte tout du long.

Définition 3 (Polynôme homogène). *Un polynôme homogène de degré k à N variables est un $P \in \mathbb{R}[X_1, \dots, X_N]$ de la forme*

$$P(X_1, \dots, X_N) = \sum_{i \in \llbracket 1, N \rrbracket^k} a_i \prod_{j=1}^k X_{i_j},$$

avec des réels (a_{i_1, \dots, i_k}) symétriques en les i_j et nuls sur la diagonale.

Le théorème s'énonce alors de manière très simple.

Théorème 4 (Convergence du kurtosis). *Soit (P_n) une suite de polynômes homogènes de degré k à (N_n) variables et (X_i) une suite de v.a. réelles i.i.d. selon des lois gaussiennes. On suppose que $Y_n = P_n(X_1, \dots, X_{N_n})$ est centrée réduite pour tout $n \in \mathbb{N}$. On a alors*

$$Y_n \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N}(0, 1) \iff \mathbb{E}_{\gamma_{N_n}}(Y_n^4) \xrightarrow[n \rightarrow +\infty]{} 3,$$

où γ_d est la mesure gaussienne sur \mathbb{R}^d .

Nous présenterons la preuve de Ledoux [3], qui donne aussi une vitesse de convergence explicite.

1.2 Le sens direct

Le sens direct est aisé.

Dém. Soit $Y \sim \mathcal{N}(0, 1)$, ses moments sont calculables par intégration par parties (IPP) répétée. En particulier,

$$\begin{aligned} \mathbb{E}_\gamma(Y^4) &= \int_{\mathbb{R}} Y^4 d\gamma \\ &= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} x^4 e^{-\frac{x^2}{2}} dx \\ &= \frac{3}{\sqrt{2\pi}} \int_{\mathbb{R}} x^2 e^{-\frac{x^2}{2}} dx \\ &= \frac{3}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-\frac{x^2}{2}} dx, \\ \mathbb{E}_\gamma(Y^4) &= 3. \end{aligned}$$

Supposons alors que $Y_n \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N}(0, 1)$, en particulier (Y_n) est uniformément bornée. Donc on a aussi convergence des moments et pour le kurtosis il vient bien $\mathbb{E}_{\gamma_{N_n}}(Y_n^4) \xrightarrow[n \rightarrow +\infty]{} 3$. \square

Le reste de ce mémoire a pour but d'expliciter le sens indirect. Une première section est dédiée à expliciter la méthode de Stein, qui va donner une borne sur une distance de Y_n à son éventuelle limite. Cette borne s'exprime à partir de la variance de ce qu'on appelle le *carré du champ*. Avec les polynômes homogènes c'est:

$$d(\mu, \gamma) \leq C \sigma_{\gamma_N}(\|P\|^2),$$

avec d qui caractérise la distance entre μ , la loi Y , et la loi gaussienne.

Après une brève présentation de la méthode de Stein, nous nous intéressons à établir cette borne pour des fonctions propres d'opérateurs markoviens, puis à majorer cette borne à partir de conditions spectrales autour de certaines fonctions propres appelées *chaos*, ce qui amènera à la majoration

$$\sigma_{\gamma_N}(\|P\|^2)^2 \leq C_k (\mathbb{E}_{\gamma_N}(Y^4) - 3),$$

avec k le degré de $P \in \mathbb{R}[X_1, \dots, X_N]$.

2 La méthode de Stein

Le lemme de Stein est une caractérisation de la loi normale souvent énoncée comme suit.

Lemme 5 (Stein, voir [2]). *Pour une v.a. réelle W , si pour toute f continue et C^1 par morceaux telle que $\mathbb{E}(|f'(Z)|) < \infty$ pour $Z \sim \mathcal{N}(0,1)$,*

$$\mathbb{E}(Wf(W)) = \mathbb{E}(f'(W)),$$

alors W suit une loi normale centrée réduite.

La méthode de Stein suit de cette caractérisation. Elle permet de donner certaines bornes sur des distances entre distributions de la forme $d(X, Y) = \sup_{h \in \mathcal{H}} |\int h dX - \int h dY| = \sup_{h \in \mathcal{H}} |\mathbb{E}(h(X) - h(Y))|$. La distance de variation totale en est une, avec $\mathcal{H} = \{\mathbb{1}_B | B \text{ borélien}\}$. Elle consiste en:

- Trouver un opérateur L et une classe de fonctions \mathcal{F} pour lesquels $(\forall f \in \mathcal{F}, \mathbb{E}(L(f)(X)) = 0) \iff X$ suit une loi connue;
- Étudier les solutions f_h de $L(f_h) = h - \mathbb{E}(h(X))$ pour $h \in \mathcal{H}$ et X suivant la loi ci-dessus;
- En prenant l'espérance conclure de propriétés des f_h des relations sur $\mathbb{E}(L(f_h)(Y))$ donc sur $\mathbb{E}(h(Y) - h(X))$ et donc sur $d(X, Y)$.

Selon comment L s'exprime il peut y avoir une condition analytique forte qui apparaît à partir de cette écriture tautologique.

3 Le théorème de caractérisation

Le théorème que nous allons exposer s'applique en plus grande généralité que seulement les polynômes homogènes. Ceux-ci sont en fait les fonctions propres de l'opérateur d'Ornstein-Uhlenbeck, qui vérifie des propriétés générales, ce qui permet de généraliser notre étude.

3.1 Le cas des polynômes homogènes

Définition 6. *L'opérateur d'Ornstein-Uhlenbeck est l'opérateur L défini sur $L^2(\mathbb{R}^N, \gamma_N) \cap C^\infty$ par*

$$L = \Delta - \langle x, \nabla \rangle.$$

Au vu de la généralité des notations Δ , ∇ et \langle, \rangle , nous écrivons aussi L indépendamment de la dimension N .

Proposition 7. *Soit $L = \Delta - \langle x, \nabla \rangle$, on a $Sp(L) = -\mathbb{N}$ et les espaces propres de L sont les $E_k = \text{Vect}(H_i, |i| = k)$, où $(H_i, i \in \mathbb{N}(\mathbb{N}))$ est la base des polynômes d'Hermite.*

Dém. Soit $i \in \mathbb{N}^N$, on a

$$\begin{aligned} L(H_i) &= \Delta H_i - \langle x, \nabla H_i \rangle \\ &= \Delta \left(\prod_{j=1}^N H_{i_j}(x_j) \right) - \left\langle x, \nabla \left(\prod_{j=1}^N H_{i_j}(x_j) \right) \right\rangle \\ &= \sum_{k=1}^N H''_{i_k}(x_k) \prod_{j=1, j \neq k}^N H_{i_j}(x_j) - \sum_{k=1}^N x_k H'_{i_k}(x_k) \prod_{j=1, j \neq k}^N H_{i_j}(x_j) \\ &= \sum_{k=1}^N (H''_{i_k}(x_k) - x_k H'_{i_k}(x_k)) \prod_{j=1, j \neq k}^N H_{i_j}(x_j) \\ &= - \sum_{k=1}^N i_k H_{i_k}(x_k) \prod_{j=1, j \neq k}^N H_{i_j}(x_j), \\ L(H_i) &= -|i|H_i. \end{aligned}$$

□

Ce qui suit étant valable dans le cas général, nous nous placerons dans le cadre d'un espace mesurable (E, \mathcal{F}) , avec μ la mesure invariante associée à l'opérateur L .

Comme nous l'avons expliqué, nous allons borner la distance à une loi gaussienne grâce à la variance de l'opérateur carré du champs défini comme suit.

Définition 8. Soit L un opérateur. Le carré du champ de L est l'opérateur bilinéaire Γ défini par

$$\Gamma(f, g) = \frac{1}{2}(L(fg) - fL(g) - L(f)g).$$

Lorsque $f = g$, on notera $\Gamma(f)$ pour $\Gamma(f, f)$.

L'une des propriétés que l'on souhaite pour notre opérateur est la suivante:

Définition 9. On dit que L est un opérateur de diffusion si pour toute fonction lisse $\varphi : \mathbb{R} \rightarrow \mathbb{R}$, on a

$$L(\varphi(f)) = \varphi'(f)L(f) + \varphi''(f)\Gamma(f).$$

Le carré du champ de l'opérateur d'Ornstein–Uhlenbeck est

$$\begin{aligned} \Gamma(f, g) &= \frac{1}{2}(\Delta fg - \langle x, \nabla fg \rangle - f\Delta g + f\langle x, \nabla g \rangle - g\Delta f + g\langle x, \nabla f \rangle) \\ &= \frac{1}{2}(\Delta fg - f\Delta g - g\Delta f) \\ &= \langle \nabla f, \nabla g \rangle. \end{aligned}$$

Ainsi $\Gamma(f) = \|\nabla f\|^2$. De plus, on a

$$\begin{aligned} L(\varphi(f)) &= \Delta\varphi(f) - \langle x, \nabla\varphi(f) \rangle \\ &= \varphi''(f)\|\nabla f\|^2 + \varphi'(f)\Delta f - \varphi'(f)\langle x, \nabla f \rangle \\ &= \varphi''(f)\Gamma(f) + \varphi'(f)L(f). \end{aligned}$$

Donc l'opérateur d'Ornstein–Uhlenbeck est un opérateur de diffusion.

3.2 Première borne par méthode de Stein

Nous utilisons alors la méthode de Stein présentée précédemment pour caractériser la distance à une loi gaussienne.

Théorème 10 (Application de la méthode de Stein). Soit F une fonction propre de $-L$, un opérateur de diffusion, associée à $\lambda > 0$. Quitte à la normaliser, on suppose que $\mathbb{E}_\mu(F^2) = 1$. Notons μ_F la mesure image de μ par F . Soit $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ une fonction lisse bornée et notons $\psi : \mathbb{R} \rightarrow \mathbb{R}$ l'unique solution bornée de l'équation de Stein:

$$\varphi - \mathbb{E}_\gamma(\varphi) = \psi' - x\psi.$$

Alors on a

$$|\mathbb{E}_{\mu_F}(\varphi) - \mathbb{E}_\gamma(\varphi)| \leq \frac{C_\varphi}{\lambda} \sigma_\mu(\Gamma(F)),$$

où $C_\varphi = \|\psi'\|_\infty$ ne dépend que de φ . On a donc une estimation de la distance de variation totale entre μ_F et γ qui est donnée par $\sigma_\mu(\Gamma(F))$.

Dém. On a $\mathbb{E}_{\mu_F}(\varphi) - \mathbb{E}_\gamma(\varphi) = \mathbb{E}_\mu(\varphi(F)) - \mathbb{E}_\gamma(\varphi) = \mathbb{E}_\mu(\psi'(F) - F\psi(F))$. Or, $\psi'(F) - F\psi(F) = \psi'(F) + \frac{1}{\lambda}L(F)\psi(F)$ donc avec la propriété de diffusion et en utilisant Cauchy-Schwarz,

$$\begin{aligned} \mathbb{E}_{\mu_F}(\varphi) - \mathbb{E}_\gamma(\varphi) &= \mathbb{E}_\mu\left(\psi'(F) + \frac{1}{\lambda}L(F)\psi(F)\right) \\ &= \mathbb{E}_\mu(\psi'(F)) - \frac{1}{\lambda}\mathbb{E}_\mu(\psi'(F)\Gamma(F)) \\ &= \mathbb{E}_\mu\left(\psi'(F)\left(1 - \frac{1}{\lambda}\Gamma(F)\right)\right) \\ &\leq \mathbb{E}_\mu(\psi'(F)^2)^{\frac{1}{2}} \mathbb{E}_\mu\left(\left(1 - \frac{1}{\lambda}\Gamma(F)\right)^2\right)^{\frac{1}{2}} \end{aligned}$$

Or, $\mathbb{E}_\mu(\psi'(F)^2)^{\frac{1}{2}} \leq \|\psi'\|_\infty$ et comme F est normalisée, $\mathbb{E}_\mu(\Gamma(F)) = -\mathbb{E}_\mu(FL(F)) = \lambda$, donc $\mathbb{E}_\mu\left(\left(1 - \frac{1}{\lambda}\Gamma(F)\right)^2\right)^{\frac{1}{2}} = \sigma_\mu(\Gamma(F))$, d'où la majoration. □

Une expression de ψ permet alors de donner une majoration simple en fonction de $\|\varphi\|_\infty$.

Proposition 11. *L'unique solution bornée de l'équation de Stein est*

$$\psi(x) = e^{\frac{x^2}{2}} \int_x^{+\infty} e^{-\frac{t^2}{2}} (\mathbb{E}_\gamma(\varphi) - \varphi(t)) dt.$$

De plus, pour C_φ on a la majoration $\|\psi'\|_\infty \leq 4\|\varphi\|_\infty$.

Dém. Les solutions de l'équation homogène étant les $x \mapsto \lambda e^{\frac{x^2}{2}}$, on déduit grâce à la méthode de la variation de la constante les solutions de l'équation de Stein:

$$\left\{ x \mapsto e^{\frac{x^2}{2}} \left(\int_0^x e^{-\frac{t^2}{2}} (\varphi(t) - \mathbb{E}_\gamma(\varphi)) dt + \lambda \right), \lambda \in \mathbb{R} \right\} = \left\{ x \mapsto e^{\frac{x^2}{2}} \left(\int_x^{+\infty} e^{-\frac{t^2}{2}} (\mathbb{E}_\gamma(\varphi) - \varphi(t)) dt + \lambda' \right), \lambda' \in \mathbb{R} \right\}.$$

Or, en notant que $\int_x^{+\infty} e^{-\frac{t^2}{2}} (\mathbb{E}_\gamma(\varphi) - \varphi(t)) dt = -\int_{-\infty}^x e^{-\frac{t^2}{2}} (\mathbb{E}_\gamma(\varphi) - \varphi(t)) dt$, on a pour $x \geq 0$

$$\begin{aligned} |x\psi(x)| &\leq \int_x^{+\infty} \left| x e^{\frac{x^2-t^2}{2}} (\mathbb{E}_\gamma(\varphi) - \varphi(t)) \right| dt \\ &\leq 2\|\varphi\|_\infty t \int_x^{+\infty} e^{\frac{(x-t)(x+t)}{2}} dt \\ &= 2\|\varphi\|_\infty \int_0^{+\infty} t e^{\frac{-t(2x+t)}{2}} dt \\ &\leq 2\|\varphi\|_\infty \int_0^{+\infty} t e^{\frac{-t^2}{2}} dt \\ &= 2\|\varphi\|_\infty. \end{aligned}$$

Et pour $x \leq 0$

$$\begin{aligned} |x\psi(x)| &\leq \int_{-\infty}^x \left| x e^{\frac{x^2-t^2}{2}} (\mathbb{E}_\gamma(\varphi) - \varphi(t)) \right| dt \\ &\leq 2\|\varphi\|_\infty \int_{-\infty}^x -t e^{\frac{(x-t)(x+t)}{2}} dt \\ &= 2\|\varphi\|_\infty \int_{-\infty}^0 -t e^{\frac{-t(2x+t)}{2}} dt \\ &\leq 2\|\varphi\|_\infty \int_{-\infty}^0 -t e^{\frac{-t^2}{2}} dt \\ &= 2\|\varphi\|_\infty. \end{aligned}$$

Donc ψ est la seule solution bornée de l'équation de Stein et on a $|\psi'(x)| \leq |x\psi(x)| + |\varphi(x) - \mathbb{E}_\gamma(\varphi)| \leq 4\|\varphi\|_\infty$. \square

On obtient alors la première inégalité annoncée

3.3 Gradients itérés du carré du champ

La seconde partie de la démonstration utilise des outils plus algébriques, qui font apparaitre certaines conditions sur le spectre de notre opérateur L .

Définition 12. *On peut itérer le carré du champ. On pose $\Gamma_0(f, g) = fg$ et pour tout $m \in \mathbb{N}$, Γ_{m+1} est l'opérateur bilinéaire sur (E, \mathcal{F}) défini par*

$$\Gamma_{m+1}(f, g) = \frac{1}{2}(L(\Gamma_m(f, g)) - \Gamma_m(f, L(g)) - \Gamma_m(L(f), g)).$$

Proposition 13. *Soit F une fonction propre de $-L$ associée à λ . Pour tout $m \in \mathbb{N}^*$,*

$$\Gamma_m(F) = \left(\frac{1}{2}L + \lambda Id \right)^{m-1} \Gamma(F).$$

De plus, pour tout $(m, n) \in (\mathbb{N}^*)^2$,

$$\mathbb{E}_\mu(\Gamma_n(F)\Gamma_m(F)) = \mathbb{E}_\mu(\Gamma_{n-1}\Gamma_{m+1}(F)).$$

En particulier, pour $n = 1$,

$$\mathbb{E}_\mu(\Gamma(F)\Gamma_m(F)) = \mathbb{E}_\mu(F^2\Gamma_{m+1}(F)).$$

Dém. La première égalité découle de

$$\Gamma_{m+1}(F) = \frac{1}{2}L(\Gamma_m(F)) + \lambda\Gamma_m(F) = \left(\frac{1}{2}L + \lambda Id\right)\Gamma_m.$$

En utilisant ce résultat sur n , on a donc

$$\mathbb{E}_\mu(\Gamma_n(F)\Gamma_m(F)) = \frac{1}{2}\mathbb{E}_\mu(L(\Gamma_{n-1}(F))\Gamma_m(F)) + \lambda\mathbb{E}_\mu(\Gamma_{n-1}\Gamma_m(F)),$$

et de l'autre coté, en l'utilisant sur $m+1$,

$$\mathbb{E}_\mu(\Gamma_{n-1}(F)\Gamma_{m+1}(F)) = \frac{1}{2}\mathbb{E}_\mu(L(\Gamma_{n-1}(F))\Gamma_m(F)) + \lambda\mathbb{E}_\mu(\Gamma_{n-1}\Gamma_m(F)).$$

□

On peut donner des interprétations géométriques aux gradients itérés du carré du champs. En particulier, Γ_2 est relié à la courbure de L . Cela correspond à la courbure de la variété différentielle pour laquelle L est le générateur de la marche aléatoire canonique.

Proposition 14. *Supposons que L soit de courbure $\rho > 0$, c'est à dire que pour tout f , $\Gamma_2(f) \geq \rho\Gamma(f)$. Si F est une fonction propre de $-L$ associée à ρ alors $\Gamma(F)$ est constante.*

En particulier, si L est un opérateur de diffusion, d'après le théorème 10, la distribution de F est alors gaussienne.

Dém. Le proposition précédent donne $\Gamma_2(F) = \frac{1}{2}L(\Gamma(F)) + \rho\Gamma(F)$, donc $L(\Gamma(F))$ est positive. Ainsi,

$$0 \leq \mathbb{E}_\mu(\Gamma(F)L(\Gamma(F))) = -\mathbb{E}_\mu(\Gamma(\Gamma(F))) \leq 0.$$

Donc $\Gamma(\Gamma(F)) = 0$ et $\Gamma(F)$ est constante. □

Pour l'opérateur d'Ornstein-Uhlenbeck, on a $\Gamma_2(f) = \frac{1}{2}\Delta\langle\nabla f, \nabla f\rangle - \langle x, \nabla\langle\nabla f, \nabla f\rangle\rangle - \langle\nabla f, \nabla(\Delta f - \langle x, \nabla f\rangle)\rangle$. Or les calculs donnent:

$$\begin{aligned}\Delta\langle\nabla f, \nabla f\rangle &= 2\sum_{i,j}\left(\frac{\partial^2 f}{\partial x_i\partial x_j}\right)^2 + \frac{\partial f}{\partial x_i}\frac{\partial^3 f}{\partial x_i\partial x_j^2}; \\ \langle x, \nabla\langle\nabla f, \nabla f\rangle\rangle &= \sum_{i,j}x_j\frac{\partial f}{\partial x_i}\frac{\partial^2 f}{\partial x_i\partial x_j}; \\ \langle\nabla f, \nabla\Delta f\rangle &= \sum_{i,j}\frac{\partial f}{\partial x_i}\frac{\partial^3 f}{\partial x_i\partial x_j^2}; \\ \langle\nabla f, \nabla\langle x, \nabla f\rangle\rangle &= \sum_i\left(\frac{\partial f}{\partial x_i}\right)^2 + \sum_{i,j}x_j\frac{\partial f}{\partial x_i}\frac{\partial^2 f}{\partial x_i\partial x_j}.\end{aligned}$$

Ainsi $\Gamma_2(f) = \sum\left(\frac{\partial^2 f}{\partial x_i\partial x_j}\right)^2 + \Gamma(f)$, donc on en déduit que L est de courbure 1. Or, les fonctions propres associées à 1 sont les polynômes homogènes de degré 1, donc on retrouve le fait qu'une somme de loi gaussiennes est aussi une loi gaussiennes.

3.4 Méthode des chaos

La seconde inégalité repose sur l'étude des *chaos* de l'opérateur.

Définition 15 (Chaos). *Considérons un opérateur $-L$ dont le spectre est $0 = \lambda_0 < \lambda_1 < \lambda_2 < \dots$. Pour tout $k \in \mathbb{N}$, on pose*

$$Q_k = \prod_{i=0}^{k-1}(X - \lambda_i) = \sum_{i=0}^k a_i X^i.$$

On définit alors la forme bilinéaire $Q_k(\Gamma)$ comme

$$Q_k(\Gamma) = \sum_{i=0}^k a_i \Gamma_i.$$

On dit alors qu'une fonction propre F de L associée à $-\lambda_k$ est un chaos d'ordre k si on a $Q_{k+1}(\Gamma)(F) = 0$.

On définit aussi deux polynômes auxiliaires, qui interviendront dans le contrôle de $\mathbb{E}(\Gamma(F)^2)$.

Définition 16. Pour tout $k \in \mathbb{N}^*$, on pose

$$R_k = \frac{1}{X^2}(Q_k(X) - Q'_k(0)X) = \sum_{i=2}^k a_i X^{i-2},$$

$$T_k = R_k(X + \lambda_{k-1}) - R_k(\lambda_{k-1}).$$

En utilisant l'érgodicité de L , on obtiens une caractérisation des chaos.

Corollaire 17. F est un chaos d'ordre k de L ssi $Q_k(\Gamma)(F)$ est constante.

Dém. En prenant F normalisée dans $L^2(\mu)$ on a en utilisant l'IPP associée à L répétée on a la formule dans $L^2(\mu)$:

$$\langle Q_k(\Gamma)(F) \rangle = \langle F, Q_k(-L)F \rangle.$$

Plus prosaïquement: $\int Q_k(\Gamma)(F)d\mu = \int FQ_k(-L)Fd\mu = Q_k(\lambda_k)$, puisque F est normalisée et fonction propre de $-L$.

Si $Q_k(\Gamma)(F)$ est constante on a qu'il s'agit donc de $Q_k(\lambda_k)$. D'autre part $LQ_k(\Gamma)(F) = 2Q_{k+1}(\Gamma)(F)$. L annule les constantes: $Q_{k+1}(\Gamma)(F) = 0$.

Réciproquement si $Q_{k+1}(\Gamma)(F) = 0$, l'espace associé à λ_0 ne contient que des constantes, c'est-à-dire $-L$ est ergodique, donc c'est le cas de $Q_k(\Gamma)(F)$. \square

Dans le cas de l'opérateur d'Ornstein-Uhlenbeck, le calcul donne simplement $Q_k(\Gamma)(f) = \|\nabla^k f\|^2$ donc si F est un polynôme homogène de degré k , c'est bien un chaos d'ordre k . Ainsi, on peut bien appliquer les résultats qui suivent à n'importe quel polynôme homogène, sans se soucier de la distinction entre fonction propre et chaos.

Le fait que F soit un chaos de L permet alors de donner une expression de $\mathbb{E}_\mu(\Gamma(F)^2)$ en fonction de $\mathbb{E}_\mu(F^2\Gamma(F))$ et donc de $\mathbb{E}_\mu(F^4)$.

Proposition 18. Avec les mêmes notations que précédemment, on pose pour tout $k \in \mathbb{N}$,

$$\pi_k = \prod_{i=1}^k \lambda_i.$$

Si F est un chaos d'ordre $k \leq 1$ de L associé à $-\lambda_k$, alors

$$\pi_{k-1}\mathbb{E}_\mu(\Gamma(F)^2) = \pi_k\mathbb{E}_\mu(F^2\Gamma(F)) + (-1)^k\mathbb{E}_\mu\left(\Gamma(F)T_{k+1}\left(\frac{L}{2}\right)(\Gamma(F))\right).$$

Dém. D'après la définition d'un chaos,

$$\begin{aligned} 0 &= \mathbb{E}_\mu(F^2Q_{k+1}(\Gamma)(F)) \\ &= \sum_{i=1}^{k+1} a_i\mathbb{E}_\mu(F^2\Gamma_i(F)) \\ &= a_1\mathbb{E}_\mu(F^2\Gamma(F)) + \sum_{i=2}^{k+1} a_i\mathbb{E}_\mu(\Gamma(F)\Gamma_{i-1}(F)). \end{aligned}$$

Or, $a_1 = (-1)^k\pi_k$. Aussi,

$$\begin{aligned} \sum_{i=2}^{k+1} a_i\mathbb{E}_\mu(\Gamma(F)\Gamma_{i-1}(F)) &= \sum_{i=2}^{k+1} a_i\mathbb{E}_\mu\left(\Gamma(F)\left(\frac{1}{2}L + \lambda_k I_d\right)^{i-2}\Gamma(F)\right) \\ &= \sum_{i=0}^{k-1} a_{i+2}\mathbb{E}_\mu\left(\Gamma(F)\left(\sum_{j=0}^i \lambda_k^{i-j}\left(\frac{L}{2}\right)^j\right)\Gamma(F)\right) \\ &= \sum_{i=0}^{k-1} \sum_{j=0}^i a_{i+2}\lambda_k^{i-j}\mathbb{E}_\mu\left(\Gamma(F)\left(\frac{L}{2}\right)^j\Gamma(F)\right). \end{aligned}$$

Or, on a bien

$$T_{k+1} = \sum_{i=0}^{k-1} a_{i+2}(X + \lambda_k)^i - R_{k+1}(\lambda_k) = \sum_{i=0}^{k-1} \sum_{j=0}^i a_{i+2}\lambda_k^{i-j}X^j - (-1)^{k-1}\pi_{k-1},$$

d'où l'égalité. \square

Si l'on est dans un cas favorable, on a alors majoration, ce qui permet d'établir l'inégalité.

Corollaire 19. *Si on a également*

$$\forall n \in \mathbb{N}, (-1)^k T_{k+1} \left(-\frac{\lambda_n}{2} \right) \leq 0,$$

alors $\mathbb{E}_\mu (\Gamma(F)^2) \leq \lambda_k \mathbb{E}_\mu (F^2 \Gamma(F))$. Donc si F est normalisée et si L est un opérateur de diffusion, on a

$$\sigma_\mu(\Gamma(F))^2 \leq \lambda_k^2 \left(\frac{1}{3} \mathbb{E}_\mu (F^4) - 1 \right).$$

Dém. Décomposons $\Gamma(F)$ dans une base orthogonale de fonctions propres de $-L$ en $\Gamma(F) = \sum_{n \in \mathbb{N}} G_n$, on a alors

$$T_{k+1} \left(\frac{L}{2} \right) (\Gamma(F)) = \sum_{n \in \mathbb{N}} T_{k+1} \left(-\frac{\lambda_n}{2} \right) G_n,$$

d'où

$$\mathbb{E}_\mu \left(\Gamma(F) T_{k+1} \left(\frac{L}{2} \right) (\Gamma(F)) \right) = \sum_{n \in \mathbb{N}} T_{k+1} \left(-\frac{\lambda_n}{2} \right) \mathbb{E}_\mu (G_n^2).$$

Avec la condition, on en déduit donc que $\mathbb{E}_\mu (\Gamma(F)^2) \leq \lambda_k \mathbb{E}_\mu (F^2 \Gamma(F))$. Aussi puisque la fonction F est normalisée, $\mathbb{E}_\mu(\Gamma(F)) = \lambda_k$ et $\lambda_k \mathbb{E}_\mu(F^4) = -\mathbb{E}_\mu(F^3 L(F)) = 3\mathbb{E}_\mu(F^2 \Gamma(F))$, d'où l'inégalité. \square

Théorème 20. *Lorsque le spectre de $-L$ est \mathbb{N} , la condition du corollaire précédent est vérifiée.*

Dém. Il s'agit de montrer que $\forall n \in \mathbb{N}$

$$\left(\frac{n}{2} - k \right)^{-1} \left[\prod_{i=0}^{k-1} \left(\frac{n}{2} - i \right) - k! \right] \geq (k-1)!.$$

Lorsque $n = 2k$, le membre de droite vaut $k! \sum_{i=1}^k \frac{1}{i}$, donc l'inégalité est vérifiée. Sinon, supposons d'abord que $n > 2k$, on a alors

$$\begin{aligned} \prod_{i=0}^{k-1} \left(\frac{n}{2} - i \right) &= \left(\frac{n}{2} - k + 1 \right) \prod_{i=0}^{k-2} \left(\frac{n}{2} - i \right) \\ &\geq \left(\frac{n}{2} - k + 1 \right) \prod_{i=0}^{k-2} (k - i) \\ &= \left(\frac{n}{2} - k + 1 \right) k!, \end{aligned}$$

d'où le résultat. Maintenant si $n < 2k$, posons $p = \lceil \frac{n}{2} \rceil$ on a $1 \leq p \leq k$ et il vient alors

$$\begin{aligned} \prod_{i=0}^{k-1} \left(\frac{n}{2} - i \right) &= \frac{n}{2} \prod_{i=1}^{p-1} \left(\frac{n}{2} - i \right) \prod_{i=p}^{k-1} \left(\frac{n}{2} - i \right) \\ &\leq \frac{n}{2} \prod_{i=1}^{p-1} (p - i) \prod_{i=p}^{k-1} (i - p + 1) \\ &= \frac{n}{2} (p-1)! (k-p)! \\ &\leq \frac{n}{2} (k-1)!, \end{aligned}$$

le corollaire est donc bien démontré. \square

3.5 Une condition suffisante

La question qu'on s'est posé maintenant est quelle signification donner à la condition $\forall n \in \mathbb{N} (-1)^k T_{k+1} \left(-\frac{\lambda_n}{2} \right) \leq 0$? L'idée est de se dire que bien que les positions des λ_n jouent un rôle dans notre étude, celles des $\frac{\lambda_n}{2}$ ne représentent rien par rapport à nos polynômes. Aussi, on remarque que dans le cas de \mathbb{N} , on a même $\forall x \in \mathbb{R}_-, (-1)^k T_{k+1}(x) \leq 0$. Nous renforçons donc cette condition suffisante en gardant en tête que l'on reste proche du cas général

Proposition 21. *On a*

$$\forall x \in \mathbb{R}_-, (-1)^k T_{k+1}(x) \leq 0 \iff \forall y \in]0, \lambda_k[, \prod_{i=1}^{k-1} (\lambda_i - y) \leq \pi_{k-1}.$$

Dém. On a $R_{k+1} = \frac{1}{X} \left(\prod_{i=1}^k (X - \lambda_i) - (-1)^k \pi_k \right)$ donc

$$\begin{aligned} (-1)^k T_{k+1} &= \frac{1}{\lambda_k + X} \left(\prod_{i=1}^k (\lambda_i - \lambda_k - X) - \pi_k \right) + \pi_{k-1} \\ &= \frac{1}{\lambda_k + X} \left(\prod_{i=1}^k (\lambda_i - \lambda_k - X) - \pi_k + \pi_k + X \pi_{k-1} \right) \\ &= \frac{X}{\lambda_k + X} \left(\pi_{k-1} - \prod_{i=1}^{k-1} (\lambda_i - \lambda_k - X) \right). \end{aligned}$$

Ainsi, la première condition est équivalente à $\pi_{k-1} \leq \prod_{i=1}^{k-1} (\lambda_i - \lambda_k - x)$ si $x \in]-\lambda_k, 0[$ et $\pi_{k-1} \geq \prod_{i=1}^{k-1} (\lambda_i - \lambda_k - x)$ si $x \in]-\infty, -\lambda_k[$ et avec un changement de variable c'est équivalent à $\pi_{k-1} \leq \prod_{i=1}^{k-1} (\lambda_i - y)$ si $y \in]0, \lambda_k[$ et $\pi_{k-1} \geq \prod_{i=1}^{k-1} (\lambda_i - y)$ si $y \in \mathbb{R}_-$. Or, cette seconde condition est toujours vérifiée puisque comme $\prod_{i=1}^{k-1} (\lambda_i - y)$ est scindé à racines simples sur \mathbb{R} , par le théorème de Rolle, ses extremums locaux sont dans $] \lambda_1, \lambda_{k-1} [$, donc ce polynôme est décroissant sur \mathbb{R}_- . □

Cette condition donne donc des informations sur la distance maximale entre les k premiers λ_i par rapport au gap initial entre λ_0 et λ_1 . En effet, deux λ_i consécutifs doivent toujours être assez proches.

Corollaire 22. *Notons pour tout $i \in \mathbb{N}^*$, $\lambda_{i+1} - \lambda_i = \delta_i$. Soit $k \in \mathbb{N}^*$, si $(\delta_i)_{i \in [1, k]}$ est décroissante, la condition précédente est vérifiée pour T_{k+1} .*

Dém. Soit $y \in]0, \lambda_k[$, supposons que $\lambda_d < y < \lambda_{d+1}$, on écrit $y = \varepsilon + \sum_{j=1}^d \delta_j$, on a alors

$$\begin{aligned} \prod_{i=1}^{k-1} (\lambda_i - y) &= \prod_{i=1}^{k-1} \left(\sum_{j=1}^i \delta_j - \varepsilon - \sum_{j=0}^{d-1} \delta_j \right) \\ &= -\varepsilon \prod_{i=1}^{d-1} \left(-\varepsilon - \sum_{j=i}^d \delta_j \right) \prod_{i=d+1}^{k-1} \left(\sum_{j=d+1}^i \delta_j - \varepsilon \right) \\ &\leq \varepsilon \prod_{i=1}^{d-1} \left(\sum_{j=i}^{d+1} \delta_j \right) \prod_{i=d+1}^{k-1} \left(\sum_{j=d+1}^i \delta_j \right) \\ &\leq \varepsilon \delta_{d+1} \prod_{i=1}^{d-1} \lambda_{i+2} \prod_{i=d+2}^{k-1} \lambda_i \\ &\leq \pi_{k-1}. \end{aligned}$$

□

4 Conclusion

Initialement, le résultat auquel nous nous sommes intéressés avait été démontré par D. Nualart et G. Peccati dans [4] et utilisait surtout les intégrales stochastiques. Cependant, la preuve de M. Ledoux est non seulement plus élémentaire mais aussi valable dans un cas plus général.

En s'appuyant sur le lemme de Stein et l'étude des chaos, nous pouvons alors dégager des critères spectraux plus généraux. C'est ce que nous avons exploré dans notre dernière partie

En essayant d'approfondir notre étude à d'autres opérateurs, nous nous sommes toutefois heurtés à un manque d'exemples explicites d'opérateurs markoviens sur les réelles qui ne se factorisent pas en une action sur une chaîne de Markov discrète et action avec noyau.

5 Annexe

Définition 23. La mesure gaussienne sur $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ est la mesure de probabilité γ_d définie par densité par rapport à la mesure de Lebesgue:

$$d\gamma_d(x) = \frac{e^{-\|x\|^2/2}}{(2\pi)^{d/2}} d\lambda(x),$$

où $\|\cdot\|$ est la norme euclidienne sur \mathbb{R}^d . On note γ pour γ_1 .

Définition 24. Soit $n \in \mathbb{N}$, le n -ième polynôme d'Hermite H_n est défini par

$$H_n(x) = (-1)^n e^{\frac{x^2}{2}} D^n \left(e^{-\frac{x^2}{2}} \right) (x) = \sum_{k=0}^{\lfloor \frac{n}{2} \rfloor} (-1)^k \frac{1}{2^k} \binom{n}{2k} x^{n-2k}.$$

Les polynômes d'Hermite forment une base orthogonale de $L^2(\mathbb{R}, \gamma)$.

Ces polynômes vérifient également $H'_n = nH_{n-1}$, ce qui se déduit des deux équations suivantes:

$$H''_n - xH'_n + nH_n = 0 \text{ et } H_{n+1} - xH'_n + nH_{n-1} = 0.$$

En dimension supérieure, pour $N \in \mathbb{N}^*$, la famille définie par

$$(H_i)_{i \in \mathbb{N}^N} = \left(\prod_{j=1}^N H_{i_j}(X_j) \right)_{i \in \mathbb{N}^N} \subset \mathbb{R}[X_1, \dots, X_N]$$

est une base orthogonale de $L^2(\mathbb{R}^N, \gamma_N)$.

References

- [1] Vladimir M Zolotarev. *One-dimensional stable distributions*, volume 65. American Mathematical Soc., 1986.
- [2] L. H. Y. Chen, L. Goldstein, and Q. M. Shao. *Normal Approximation by Stein's Method*. Springer, 2011.
- [3] Michel Ledoux. Chaos of a Markov operator and the fourth moment condition. *The Annals of Probability*, 40(6):2439–2459, 2012.
- [4] D. Nualart and G. Peccati. Central limit theorems for sequences of multiple stochastic integrals. *The Annals of Probability*, 2005.