

université
angers



CENTRE

HENRI LEBESGUE



UNIVERSITÉ DE NANTES CENTRE DE MATHÉMATIQUES

MÉMOIRE DE RAPPORT DE STAGE
UNIVERSITÉ DE NANTES
MASTER II MFA
PARCOURS ANALYSE ET PROBABILITÉS

AG ABOUBACRINE ASSADECK MOHAMED ALFAKI
SOUS LA DIRECTION DE MONSIEUR FABIEN PANLOUP

29 juin 2021

RÉPUBLIQUE FRANÇAISE
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

Université d'Angers
Faculté des sciences
Département de mathématiques du LAREMA



Mémoire de fin de cycle

En vue de l'obtention du grade de Master

Option :

Mathématiques fondamentales et applications parcours analyse et probabilités

Thème

Étude des systèmes de particules à champ moyen et applications

Présenté par :

M^r AG ABOUBACRINE ASSADECK Mohamed Alfaki

Encadré par :

M^r PANLOUP Fabien

Soutenance :

En juillet 2021 au département de mathématiques Jean Leray



UNIVERSITÉ DE NANTES



CENTRE

HENRI LEBESGUE

CENTRE DE MATHÉMATIQUES



REMERCIEMENTS

à toute personne qui, de près ou de loin, a contribué à ma réussite personnelle, en particulier mes parents, mes professeurs d'Angers, Nantes et Rennes, mes amis... Sans oublier surtout le centre Henri Lebesgue (CHL) qui a financé en grande partie si ce n'est en totalité ces années de master qui furent fructueuses. Je tiens à remercier aussi tous les personnels administratifs de ces quatre prestigieux pôles de connaissances dont on ne peut guère se passer. Je ne saurais exprimer ma gratitude par des mots, je ne peux que vous dire merci à tous autant que vous êtes de m'avoir épaulé durant toutes ces années d'études vraiment du tréfonds du coeur. Je tiens aussi à remercier le CROUS et toutes les instances liées à la vie étudiante pour leur omniprésence pour tout ce qui concerne le quotidien des étudiants. Je remercie infiniment le gouvernement français et l'OFPRA¹ sans lesquels rien de tout ça n'aurait été possible. Mes sincères remerciements à la ville d'Angers et toutes ses infrastructures grâce auxquelles durant ces quatre dernières années en tant que réfugié, j'ai pu apprendre à m'intégrer dans cette société et comprendre ses moeurs que je ne connaissais que par ouï-dire. Je crois avoir fait le tour, sinon je m'en excuse d'avance : j'ai une mémoire volatile. Et pour finir, j'aime croire que :

« La réussite est le fruit de l'effort constant et du travail bien fait. »

« Le travail nous éloigne de trois grands maux à savoir le vice, l'ennui et le besoin. »

Lucrece² :

« Si tu penses que les atomes, principes des choses, peuvent trouver le repos et dans ce repos engendrer toujours de nouveaux mouvements, tu te trompes et t'égaras loin de la vérité. Puisqu'ils errent dans le vide, il faut qu'ils soient tous emportés, soit par leur pesanteur propre, soit par le choc d'un autre corps. Car s'il leur arrive dans leur agitation de se rencontrer avec choc, aussitôt ils rebondissent en sens opposés : ce qui n'a rien d'étonnant puisqu'ils sont des corps très durs, pesants, denses, et que rien derrière eux ne les arrête. Et pour mieux comprendre comment s'agitent sans fin tous les éléments de la matière, souviens-toi qu'il n'y a dans l'univers entier aucun fond ni aucun lieu où puissent s'arrêter les atomes, puisque l'espace sans limite ni mesure est infini en tous sens, ainsi que je l'ai montré abondamment avec la plus sûre doctrine. Puisqu'il en est ainsi, il ne peut y avoir aucun repos pour les atomes à travers le vide immense ; au contraire agités d'un mouvement continu et divers, ils se heurtent, puis rebondissent, les uns à de grandes distances, les autres faiblement, et s'éloignent peu. »

1. Office Français de Protection des Réfugiés et Apatrides

2. Le philosophe et poète latin Lucrece (60 av. J.-C.) donne une remarquable description du mouvement des particules selon les principes d'Épicure dans son œuvre De la nature des choses.

RÉSUMÉ

Ce mémoire est une synthèse de mon stage de trois mois passé au LAREMA³ sous la direction de monsieur Fabien Panloup⁴. Je remercie également monsieur Nicolas Raymond⁵ pour ses conseils avisés et son aide précieuse durant toute cette année.

Le stage a porté sur l'étude en temps long d'équations différentielles stochastiques dont le terme de drift dépend de la loi du processus à l'instant considéré.

Ce type de non-linéarité apparaît notamment dans les équations dites de McKean Vlasov mais aussi dans d'autres problèmes tels que les dynamiques de type Fleming-Viot issues de l'étude des dynamiques de population conditionnées à rester en vie (ou plus généralement à ne pas être absorbées).

Ce stage a été plutôt orienté sur le premier type de non-linéarité dit de champ moyen qui est largement étudié dans la littérature mais qui a suscité un regain d'intérêt dans les dernières années avec son utilisation en optimisation pour la science des données. On peut citer sur ce point de récents articles⁶ pour l'utilisation des systèmes de particules pour l'optimisation des paramètres d'un réseau de neurones.

Dans un premier temps, l'objectif a été l'étude d'un article récent⁷ proposant de nouvelles inégalités fonctionnelles⁸ pour ce type de système.

Dans un second temps, plusieurs pistes ont été explorées :

- D'une part, l'étude de telles inégalités sur des systèmes dynamiques de type auto-attractif⁹.*
- D'autre part, avec plus d'intérêt, l'étude de récents articles d'applications¹⁰ orientés "optimisation", afin d'envisager l'intérêt pratique de l'étude précise du comportement en temps long de tels systèmes et de proposer de nouvelles bornes pour des schémas d'approximation issus de telles dynamiques.*

D'après H.Balzac, le hasard est le plus grand romancier du monde et pour être fécond, il n'y a qu'à l'étudier. Pour le coup, je suis d'accord avec lui; ces trois mois de stage ont été une vraie découverte de choses que j'ignorais telles que des dimensions analytiques des probabilités et leurs interprétations, comment mener à bien un travail de recherche...

Pour finir, ce large spectre correspondant aux thématiques de recherche sur lesquelles je souhaite travailler a conforté mon intérêt de continuer en thèse dans ce fascinant domaine.

BONNE LECTURE!

3. Laboratoire Angevin de REcherche en MATHématiques

4. Enseignant chercheur en probabilités et statistiques au LAREMA et responsable de l'équipe analyse-proba-stat avec monsieur Rodolphe GARBIT.

5. Directeur du Département de Mathématiques - Représentant du LAREMA au conseil scientifique du CHL, membre de l'équipe analyse-pro-stat et spécialisé en Théorie spectrale et analyse semiclassique.

6. Voir les références bibliographiques [1], [2], [3] et [4] à la fin du document.

7. Voir référence bibliographique [5].

8. Poincaré et log-Sobolev.

9. Systèmes "compagnons" des équations de McKean-Vlasov avec moyennisation en temps et non en espace.

10. Mean-Field Langevin Dynamics. Voir les quatre premières références bibliographiques.

Table des matières

REMERCIEMENTS	i
RÉSUMÉ	ii
I Problème de McKean-Vlasov?	3
1 Mise en situation du problème	4
1.1 Introduction générale	4
1.2 Système à champ moyen	5
1.3 Lien avec thermodynamique	9
2 \mathbb{V}, \mathbb{H} et inégalités fonctionnelles	13
2.1 Variance et entropie	13
2.2 Poincaré et log-Sobolev	14
2.3 PGD	16
3 Processus de Markov	17
3.1 Généralités	17
3.2 E.U de mesures invariantes	21
3.3 Convergence exponentielle	28
4 CV faible et Wasserstein	30
4.1 Caractérisations de la CV faible	30
4.2 Tension et pré-compacité	33
4.3 Wasserstein et couplage	35
II Sur les inégalités de Poincaré et log-Sobolev	36
5 Inégalités de Poincaré et log-Sobolev pour μ^N	37
5.1 Poincaré	37
5.2 Log-Sobolev	40
6 Convergence exponentielle pour le flot solution de l'équation de McKean-Vlasov	41
III Systèmes auto-attractifs et schémas d'Euler	48
7 Systèmes auto-attractifs	49
7.1 Introduction générale	49
7.2 Résultats préliminaires	52
7.3 Calcul différentiel sur $\mathcal{M}_1(\mathbb{R}^d, V)$	56

7.4	E.U du semi-flot	58
7.5	Énergie libre et points fixes	60
7.6	Sur les mesures d'occupation	61
8	Schémas d'Euler et moyennisation	62
8.1	Introduction générale	62
8.2	Stabilité et martingales	64
8.3	Discrétisation	70
8.4	Python	72
IV	Perspectives de thèse	88
V	Annexes	90
9	A1. $N \rightarrow +\infty$?	91
9.1	Vers un formalisme du problème	91
9.2	Stratégie de l'argument de couplage et chaos	95
9.3	Fonctions continues et polynômes sur $\mathbf{P}(E)$	98
9.4	$\mathcal{C}_{\text{sym}}(E^N)$ lorsque $N \rightarrow \infty$	101
9.5	Calcul différentiel sur $\mathbf{P}(E)$	104
10	A2. Résultats techniques	108
10.1	Un théorème de type Tietze-Urysohn	108
10.2	Le théorème de Stone-Weierstrass	108
10.3	Krein-Milman et Choquet	108
10.4	Le compactifié d'Alexandroff	110
10.5	Espace produit	111
11	A3. Probabilités	113
11.1	$\mathbb{E}[\cdot \mathcal{G}]$ et vecteurs gaussiens	116
11.2	Processus stochastiques	118
12	A4. Analyse	119
12.1	Éléments de théorie de la mesure	119
12.2	Éléments de théorie spectrale	123
	Bibliographie	128

Première partie

Problème de McKean-Vlasov?

Chapitre 1

Mise en situation du problème

Pour plus de détails, consulter en bibliographie principalement les références citées et [97].

1.1 Introduction générale

Dans ce mémoire, nous considérons les équations différentielles stochastiques (EDS) de McKean-Vlasov, qui sont des EDS dont les coefficients de dérive et de diffusion dépendent non seulement de l'état du processus inconnu, mais également de sa loi de probabilité. Ces EDS, également appelées EDS à champ moyen, ont d'abord été étudiées en physique statistique et représentent en quelque sorte le comportement moyen d'un nombre infini de particules.

Récemment, ce type d'équations a suscité un regain d'intérêt dans le contexte de la théorie des jeux à champ moyen. Cette théorie a été inventée par P.L. Lions et J.M. Lasry en 2006, pour résoudre le problème de l'existence d'un équilibre de Nash approximatif pour les jeux différentiels, avec un grand nombre de joueurs. Ces équations ont trouvé des applications dans divers domaines tels que la théorie des jeux, la finance mathématique, les réseaux de communication et la gestion des ressources pétrolières.

Plus exactement, nous allons considérer des équations non linéaires de McKean Vlasov avec un potentiel interne $V : \mathbb{R}^d \rightarrow \mathbb{R} \in \mathcal{C}^2$ et un potentiel d'interaction entre deux particules $W : \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R} \in \mathcal{C}^2$ (qu'on supposera la plupart du temps symétrique, c'est-à-dire $W(x, y) = W(y, x)$) tels que :

$$\partial_t \nu_t = \Delta \nu_t + \nabla \cdot (\nu_t \nabla V) + \nabla \cdot (\nu_t \nabla (W * \nu_t)) = \nabla \cdot (\nabla \nu_t + \nu_t \nabla V + \nu_t \nabla (W * \nu_t))$$

ou de façon équivalente¹ :

$$\partial_t \nu_t = \mathbf{div}(\nu_t \nabla (\log(\nu_t) + V + W * \nu_t))$$

avec $(\nu_t)_{t \geq 0}$ un flot de mesures de probabilité sur \mathbb{R}^d tel que ν_0 est donnée, ∇ le gradient, $\nabla \cdot$ la divergence, et :

$$W * \nu(x) := \int_{\mathbb{R}^d} W(x, y) d\nu(y)$$

De plus, l'équation différentielle stochastique correspondante dite diffusion d'auto-interaction est :

$$dX_t = \sqrt{2} dB_t - \nabla V(X_t) dt - \nabla (W * \nu_t)(X_t) dt$$

avec $\mathbb{P}_{X_t} = \nu_t$. Ainsi, on voit bien que l'étude d'une telle équation nécessite des propriétés particulières sur les deux potentiels V et W données plus bas et sous lesquelles toutes les affirmations qui suivent sont valides.

1. Si le flot est pour chaque instant une distribution régulière, c'est-à-dire une densité de probabilité.

Notre objectif est triple :

1. Dans un premier temps, on s'intéresse à la convergence exponentielle vers l'équilibre pour le flot solution de l'équation de McKean-Vlasov via un récent article sur les inégalités fonctionnelles pour ce type de dynamique non linéaire en référence [5] dans la bibliographie.
2. Dans un second temps, on s'intéresse aux équations différentielles stochastiques de type auto-attractif pour lesquelles on peut montrer sous des conditions adéquates que ses mesures invariantes coïncident avec celles de McKean-Vlasov via des articles en référence bibliographiques [93] et [94] . On essaie aussi d'adapter les résultats de l'article récent en [5] à ce type de dynamique.
3. Et pour finir, on établit des résultats de convergence de schémas d'Euler vers l'équilibre via les articles cités en références bibliographiques.

Du fait de la non linéarité de la dynamique de McKean-Vlasov, on est conduit à considérer un système de $N \geq 2$ particules de type Langevin dont le lien avec le processus de McKean-Vlasov tient dans la notion de propagation du chaos, c'est-à-dire si on fait tendre le nombre de particules N vers l'infini, on peut espérer qu'une particule prise au hasard se comporte comme une solution de l'équation de McKean-Vlasov. Le contenu de la suite de ce chapitre est essentiellement issu de la référence [5] en bibliographie.

1.2 Système à champ moyen

On considère un système noté (\star) dit de champ moyen de $N \geq 2$ particules dans \mathbb{R}^d défini par : $\forall i \in \{1, \dots, N\}$

$$dX_i^N(t) = \sqrt{2}dB_i(t) - \nabla V(X_i^N(t))dt - \frac{1}{N-1} \sum_{j \neq i} \nabla_x W(X_i^N(t), X_j^N(t))dt$$

avec B_1, \dots, B_N des mouvements browniens indépendants à valeurs dans \mathbb{R}^d . Le générateur infinitésimal de ce système est donné par :

$$\mathcal{L}^N = \sum_{i=1}^N \mathcal{L}_i^N$$

\mathcal{L}_i^N défini par : $\forall f \in \mathcal{C}^\infty((\mathbb{R}^d)^N)$

$$\mathcal{L}_i^N f(x_1, \dots, x_N) = \Delta_i f(x_1, \dots, x_N) - \nabla_i V(x_i) \cdot \nabla_i f(x_1, \dots, x_N) - \frac{1}{N-1} \sum_{j \neq i} \nabla_x W(x_i, x_j) \cdot \nabla_i f(x_1, \dots, x_N)$$

avec Δ_i et ∇_i respectivement le laplacien et le gradient par rapport à x_i , et $x \cdot y := \langle x, y \rangle$ désigne le produit scalaire euclidien standard sur \mathbb{R}^d .

En tant que "diffusion gradient", sous des conditions adéquates sur les potentiels V et W décrites plus bas, on peut montrer que le système ci-dessus admet une unique mesure de probabilité invariante μ^N absolument continue par rapport à la mesure de Lebesgue sur $(\mathbb{R}^d)^N$ donnée par :

$$\mu^N(dx_1, \dots, dx_N) = \frac{1}{Z_N} e^{-H_N(x_1, \dots, x_N)} dx_1 \cdots dx_N$$

avec H_N un hamiltonien défini par :

$$H_N(x_1, \dots, x_N) = \sum_{i=1}^N V(x_i) + \frac{1}{N-1} \sum_{1 \leq i < j \leq N} W(x_i, x_j)$$

et Z_N est une constante de normalisation, c'est-à-dire :

$$Z_N = \int_{\mathbb{R}^{dN}} e^{-H_N} dx_1 \cdots dx_N$$

On parle aussi de fonction de partition en mécanique statistique.

Remarques :

Sans interactions ($W(x, y) = c \in \mathbb{R}$ constante pour tout $x, y \in \mathbb{R}^d$), $\mu^N = \nu^{\otimes N}$ (produit tensoriel d'une mesure par elle-même N fois) et donc les particules sont indépendantes avec :

$$\nu(dx) := \frac{1}{C} e^{-V(x) - \frac{c}{2}} dx, \quad C = \int_{\mathbb{R}^d} e^{-V(x) - \frac{c}{2}} dx$$

Évidemment, pour l'instant, on a fait que des affirmations sans preuve à l'appui. Avant de prouver ces affirmations, on va établir quelques notations utiles pour la clarté de ce qui va suivre dans ce mémoire et choisir les "bons" potentiels de confinement et d'interaction à associer au système de particules à champ moyen.

Pour V un potentiel de confinement comme ci-dessus, $U \in \mathcal{C}^2((\mathbb{R}^d)^N)$ un potentiel d'interaction entre N particules et $H(x_1, \dots, x_N) = U(x_1, \dots, x_N) + \sum V(x_i)$ l'hamiltonien associé, on définit la mesure :

$$\mu(dx) = \frac{1}{Z} e^{-H(x)} dx, \quad Z := \int_{\mathbb{R}^{dN}} e^{-H(x)} dx$$

Et on suppose toujours $Z < +\infty$. On note par $\mu_i := \mu_i(dx_i | x^{\hat{i}})$ la distribution conditionnelle de x_i sachant $x^{\hat{i}} := (x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_N)$ sous μ , c'est-à-dire la loi de la i -ème particule sachant les $N - 1$ autres. Elle est donnée par :

$$\mu_i(dx_i) = \frac{1}{Z_i} e^{-U(x) - V(x_i)} dx_i, \quad Z_i := Z_i(x^{\hat{i}}) = \int_{\mathbb{R}^d} e^{-U(x) - V(x_i)} dx_i$$

Et on suppose toujours $Z_i < +\infty$.

On fait les hypothèses suivantes sur les potentiels qui assurent à la fois l'existence et l'unicité mentionnée plus haut ainsi que des propriétés de contraction/convergence à l'équilibre :

(H1) La hessienne de V est bornée inférieurement et il existe c_1 et c_2 deux constantes positives telles que :

$$\forall x \in \mathbb{R}^d, \quad \langle x, \nabla V(x) \rangle \geq c_1 \|x\|^2 - c_2$$

Cette hypothèse est une condition dite de Lyapunov.

(H2) La hessienne de W est bornée et :

$$\forall \lambda > 0, \quad \int e^{-[V(x) + V(y) + \lambda W(x, y)]} dx dy < +\infty$$

(H3) La constante de Lipschitz $c_{Lip, m}$ associée à la distribution conditionnelle d'une particule est telle que :

$$c_{Lip, m} := \frac{1}{4} \int_0^{+\infty} e^{\frac{1}{4} \int_0^s b_0(u) du} ds < +\infty$$

avec $b_0(r)$ le taux de drift d'une particule du système à la distance $r > 0$ et qui est défini par :

$$b_0(r) := \sup_{x, y, z \in \mathbb{R}^d, \|x - y\| = r} \left\langle \frac{y - x}{\|x - y\|}, \nabla V(x) - \nabla V(y) + \nabla_x W(x, z) - \nabla_x W(y, z) \right\rangle$$

Remarques : Cette dernière hypothèse rappelle les travaux d'Eberle (sans potentiel d'interaction) de convergence vers l'équilibre en distance L^1 -Wasserstein. Ici, le potentiel d'interaction est perçu comme une perturbation. On notera que **(H3)** est une hypothèse assez fine qui est destinée à obtenir de bons taux de convergence pour des estimées semi-quantitatives qu'on peut obtenir sous des hypothèses moins techniques, par exemple, de forte convexité...

On note $\nabla_{x_i, x_j}^2 H$, la quantité définie par :

$$\nabla_{x_i, x_j}^2 H := (\partial_{x_{ik} x_{jl}}^2 H)_{1 \leq k, l \leq d}, \quad x_i = (x_{i1}, \dots, x_{id}) \in \mathbb{R}^d$$

On définit aussi

$$\lambda_{1,m} := \inf_{N \geq 2} \left\{ \inf_{1 \leq i \leq N} \lambda_1(\mu_i) \right\}$$

avec $\lambda_1(\mu_i)$ la constante de Poincaré associée à μ_i définie plus bas dans le second chapitre de cette partie. On introduit aussi des coefficients (c_{ij}^Z) dits de dépendance de Zegarliniski de μ_j sur x_i tels que c_{ij}^Z est la meilleure constante positive vérifiant :

$$\left\| \nabla_i \left(\sqrt{\mathbb{E}_{\mu_j}[f^2]} \right) \right\| \leq \sqrt{\mathbb{E}_{\mu_j}[\|\nabla_i f\|^2]} + c_{ij}^Z \cdot \sqrt{\mathbb{E}_{\mu_j}[\|\nabla_j f\|^2]}$$

pour toute fonction lisse strictement positive $f : x := (x_1, \dots, x_N) \in (\mathbb{R}^d)^N \rightarrow f(x)$. En particulier, $c_{ii}^Z = 0$ et la matrice $c^Z := (c_{ij}^Z)_{1 \leq i, j \leq N}$ est appelée matrice d'interdépendance de Zegarliniski. Ces coefficients vont servir à établir une inégalité du type log-Sobolev pour μ^N la mesure invariante de (\star) .

On rappelle aussi la définition de l'entropie relative qui est utilisée pour la convergence exponentielle en temps long vers la mesure invariante du processus de McKean-Vlasov via notre système à champ moyen (\star) . On voit plus bas la nécessité des entropies.

Définition 1.1 (Entropie relative)

Soit $\mu \in \mathcal{P}(\mathbb{R}^d)$. On définit $\mathbf{H}[\cdot|\mu] : \mathcal{P}(\mathbb{R}^d) \rightarrow [0, +\infty]$ telle que $\mathbf{H}[v|\mu] = \mathbb{E}_v[\log(\frac{dv}{d\mu})] = \mathbf{Ent}_\mu[\frac{dv}{d\mu}]$ si $v \ll \mu$ et $+\infty$ sinon. Et on rappelle que dans le premier cas d'absolue continuité, $\frac{dv}{d\mu}$ est la densité de Radon-Nikodym de v par rapport à μ .

On rappelle aussi que : (distance L^p -Wasserstein)

$$\mathcal{W}_p(\mu, \nu) := \inf_{X \sim \mu, Y \sim \nu} (\mathbb{E}[\|X - Y\|^p])^{\frac{1}{p}}$$

De plus, l'ensemble des mesures de probabilités sur un espace polonais qui admettent un p -moment muni de cette métrique est aussi polonais². Dans ce mémoire, cet espace est noté $\mathcal{M}_1^p(\mathbb{R}^d)$. On définit aussi l'information de Fisher-Donsker-Varadhan de ν par rapport à μ par :

$$\mathbf{I}[v|\mu] = \int \left\| \nabla \left(\sqrt{\frac{dv}{d\mu}} \right) \right\|^2 d\mu$$

si $\nu \ll \mu$ et $\sqrt{\frac{dv}{d\mu}} \in \mathbf{H}_\mu^1$, et $\mathbf{I}[v|\mu] = +\infty$ sinon. \mathbf{H}_μ^1 est le domaine de la forme de Dirichlet

$$\mathcal{E}_\mu : g \mapsto \int \|\nabla g\|^2 d\mu.$$

Ici, la fonctionnelle qui joue le rôle de relative entropie pour notre système de particules en interaction associé à l'équation de McKean-Vlasov non linéaire est l'énergie libre $\mathbf{E}_f : \nu \in \mathcal{M}_1(\mathbb{R}^d) \rightarrow \mathbf{E}_f[\nu]$ telle que

$$\mathbf{E}_f[\nu] = \mathbf{H}[\nu|\alpha] + \frac{1}{2} \int \int W(x, y) \nu(dx) \nu(dy) := \mathbf{Ent}_\alpha \left[\frac{d\nu}{d\alpha} \right] + \frac{1}{2} \mathbb{E}_{\nu \otimes \nu}[W]$$

2. Voir référence [8] en bibliographie et annexes pour plus de détails.

si $\mathbf{H}[v|\alpha] < +\infty$ et $+\infty$ sinon, avec $\alpha(dx) := \frac{e^{-V(x)}}{C} dx$, et $C := \int_{\mathbb{R}^d} e^{-V(x)} dx < +\infty$ est le facteur de normalisation associé. Ou plus précisément, l'entropie de champ moyen correspondante

$$\mathbf{H}_W[v] := \mathbf{E}_f[v] - \inf_{\eta \in \mathcal{M}_1(\mathbb{R}^d)} \mathbf{E}_f[\eta].$$

On peut montrer que cette entropie est décroissante le long du flot solution de l'équation de McKean-Vlasov et admet un unique argument minimum qui est la mesure invariante associée à ce flot³.

Pour v telle que $v(dx) = f(x)dx$, $\int \|x\|^2 v(dx) < +\infty$ et $\nabla f \in L^1_{loc}(\mathbb{R}^d)$ au sens des distributions, on définit :

$$\mathbf{I}_W[v] := \frac{1}{4} \int \left\| \frac{\nabla f(x)}{f(x)} + \nabla V(x) + (\nabla_x W * v)(x) \right\|^2 v(dx)$$

et $\mathbf{I}_W[v] = +\infty$ sinon.

Pour mieux comprendre ce qui précède et ce qui va suivre, on peut faire le rapprochement avec la théorie de l'hypocoercivité⁴, par exemple, penser à des résultats du type Fokker-Planck⁵. En effet, l'équation cinétique de Fokker-Planck est un exemple de base de la théorie de l'hypocoercivité de Villani qui affirme qu'à défaut de coercivité, la décroissance exponentielle en temps de fonctionnelles d'entropie est un outil important et puissant pour l'étude de la convergence à long terme vers l'équilibre et des vitesses de convergence vers l'équilibre.

Dans la section suivante, on donne une interprétation physique de la notion fondamentale de limites de champ moyen via la thermodynamique et la mécanique statistique.

3. Voir deuxième partie du mémoire et [5] pour plus de détails.

4. Cette théorie a été développée précisément pour mesurer à la fois le taux de convergence dans l'espace des phases, c'est-à-dire en position et en vitesse, et pour définir une notion de distance dans laquelle on mesure ce taux.

5. Pour des détails, consulter le mémoire de Villani.

1.3 Lien avec thermodynamique

La thermodynamique est une discipline transversale de la physique, qui traite des transformations de l'énergie sous toutes ses formes. Elle est fondée sur deux principes fondamentaux :

1. le premier principe énonce de façon très générale la conservation de l'énergie : l'énergie peut être stockée par un système sous forme d'énergie interne ou d'énergie cinétique, et peut être échangée avec l'extérieur sous la forme de travail ou de chaleur ;
2. le second principe de la thermodynamique traite de l'évolution des systèmes, en introduisant la notion essentielle d'entropie.

La conjonction des deux principes permet de définir de façon très rigoureuse des conditions d'équilibre d'un système, c'est-à-dire l'état vers lequel il évoluera en fonction des conditions extérieures qui lui sont imposées. La thermodynamique de l'équilibre est une discipline essentielle pour l'ingénieur, et a des applications dans tous les domaines industriels : toute installation industrielle produit ou consomme de l'énergie, et est le siège de phénomènes physico-chimiques qui évoluent vers un état d'équilibre qui peut être prédit par la thermodynamique. La thermodynamique peut être abordée selon deux approches différentes et complémentaires :

1. phénoménologique qui s'appuie sur des considérations macroscopiques pour établir un nombre réduit de principes et de lois, issus d'observations expérimentales ;
2. statistique qui s'appuie quant à elle sur des considérations moléculaires et sur le calcul des probabilités appliqué à un grand nombre de particules. Elle s'attache à analyser la structure de la matière et à établir un lien entre ses propriétés et les principes de la thermodynamique phénoménologique.

Comme on le sait, un système physique à l'équilibre cherche à minimiser son énergie. En présence d'un grand nombre de particules en interaction, la mécanique statistique nous amène cependant à nuancer ce principe, en restreignant sa validité au cas où la température est nulle. À température non nulle, un mouvement désordonné des particules apparaît, qui entre en compétition avec le principe de minimisation de l'énergie. La quantité à minimiser devient alors l'énergie libre, c'est-à-dire l'énergie restant après soustraction de l'entropie, laquelle rend compte à l'échelle macroscopique de l'aléa introduit.

On va établir le lien des systèmes de particules à champ moyen avec la thermodynamique à partir d'un formalisme électrostatique. On sait que le champ électrique engendré par une particule ponctuelle placée en un point $y \in \mathbb{R}^d$ et de charge $e = -1$ dérive d'un potentiel V_y qui satisfait à l'équation de Poisson $\Delta V_y = \delta_y$. Il faut imposer une condition au bord pour assurer l'unicité du potentiel, et le cadre le plus réaliste consiste sans doute à travailler dans un domaine $\Omega \subset \mathbb{R}^d$ dont le bord est mis à la masse, de sorte que $V_y|_{\partial\Omega} = 0$. Par définition, on obtient alors $V_y(x) = G_\Omega(x, y)$, la fonction de Green de Ω .

En théorie du potentiel classique, on idéalise cette situation en travaillant avec la fonction de Green G_d de \mathbb{R}^d , c'est-à-dire l'unique solution invariante par translation de l'EDP $\Delta_x G_d = \delta_y$ qu'on obtient explicitement via la transformation de Fourier par :

$$G_d(x, y) = \mathbb{1}_{d=2} \cdot \frac{1}{2\pi} \log \|x - y\| - c_d \cdot \mathbb{1}_{d \geq 3} \cdot \|x - y\|^{2-d}$$

avec $\frac{1}{c_d} := (d-2) \text{Vol}(\mathbb{S}^{d-1})$. Ainsi pour $d \geq 3$, on a une mise à la masse à l'infini.

Deux particules identiques x et y de même charge $e = -1$ se repoussent mutuellement en cherchant à diminuer l'énergie d'interaction $\mathbf{E}_2(x, y) = -G_d(x, y)$.

On modélise l'interaction d'une configuration de N particules x_1, \dots, x_N en moyennant les contri-

butions de toutes les paires ⁶, c'est-à-dire :

$$\mathbf{E}_N(x_1, \dots, x_N) = \frac{1}{C_N^2} \sum_{1 \leq i < j \leq N} \mathbf{E}_2(x_i, x_j)$$

avec C_N^k le nombre de parties à k éléments dans un ensemble de cardinal N éléments. En cantonnant ces particules à un compact $K \subset \mathbb{R}^d$, qu'on pense comme un condensateur sur lequel les N particules se meuvent librement, celles-ci s'éloigneront de façon à minimiser l'énergie \mathbf{E}_N , et la position à l'équilibre sera donc donnée par une configuration $P \in K^N$ telle que $\inf_{K^N} \mathbf{E}_N = \mathbf{E}_N(P)$.

Ainsi définies, les configurations à l'équilibre sont loin d'être uniques en général. Fait remarquable, l'unicité est cependant restaurée asymptotiquement lorsque N tend vers l'infini. Plus précisément, une "distribution continue" de charges sur K est décrite par une mesure de probabilité ν , dont l'énergie est donnée par :

$$\mathbf{E}[\nu] = \int \int \mathbf{E}_2(x, y) d\nu(x) d\nu(y) = \mathbb{E}_{\nu \otimes \nu}[\mathbf{E}_2],$$

et on rappelle que l'ensemble $\mathcal{P}(K)$ des mesures de probabilité sur K est compact pour la topologie faible, et la fonctionnelle $\mathbf{E} : \mathcal{P}(K) \rightarrow \mathbb{R} \cup \{+\infty\}$ est semi-continue inférieurement par semi-continuité de \mathbf{E}_2 . Un point clé et non trivial de la théorie du potentiel est que \mathbf{E} est de plus strictement convexe.

Il résulte de ces propriétés la dichotomie suivante :

1. Si $\mathbf{E}[\nu] = +\infty$ pour toute mesure $\nu \in \mathcal{P}(K)$, on dit que K est polaire, c'est-à-dire négligeable ⁷ du point de vue de la théorie du potentiel sur \mathbb{R}^d .
2. Sinon, l'énergie à l'équilibre $\inf_{\mathcal{P}(K)} \mathbf{E}$ est atteinte en une unique mesure ν_{eq} , la mesure d'équilibre du compact K .

On peut montrer que l'énergie à l'équilibre "microscopique" tend vers l'énergie à l'équilibre "macroscopique", plus exactement :

$$\inf_{K^N} \mathbf{E}_N \xrightarrow{N \rightarrow +\infty} \inf_{\mathcal{P}(K)} \mathbf{E},$$

et l'unicité asymptotique évoquée plus haut signifie que toute suite de configurations $P_N \in K^N$ minimisant \mathbf{E}_N s'équi-répartit sur la mesure d'équilibre ν_{eq} lorsque N tend vers l'infini. En d'autres termes, on a convergence faible de $\delta_N(P_N)$ vers ν_{eq} avec :

$$\delta_N : K^N \rightarrow \mathcal{P}(K), \text{ l'application mesure empirique telle que pour tout } P \in K^N, \delta_N(P) = \frac{1}{N} \sum_{i=1}^N \delta_{x_i}.$$

S'il est naturel de considérer $(\mathcal{P}(K), \mathbf{E})$ comme la limite de (K^N, \mathbf{E}_N) , la signification précise de cette convergence n'est pas immédiate ⁸. En particulier, on notera que $\mathbf{E} \circ \delta_N \equiv +\infty$.

Introduisons maintenant de l'aléa dans le modèle précédent. On se donne une mesure de probabilité de référence ν_0 telle que $\mathbf{E}[\nu_0]$ est finie, et on considère une configuration aléatoire $P_N \in K^N$ formée de N particules indépendantes et identiquement distribuées de loi ν_0 , de sorte que $P_N \sim \nu_0^{\otimes N}$. Lorsque ces particules interagissent, cet aléa entre en compétition avec le principe de minimisation de l'énergie, et à l'équilibre, la loi de P_N est modifiée pour se concentrer autour des minima de \mathbf{E}_N , avec autant plus de vigueur que la température T diminue. Comme nous l'enseignent la mécanique statistique, cette nouvelle loi est donnée par la mesure de Gibbs :

$$\gamma_{\beta, N} := \frac{1}{Z_{\beta, N}} e^{-\beta N \mathbf{E}_N} \nu_0^{\otimes N}$$

6. On parle de champ moyen.

7. Un tel ensemble est de mesure de Lebesgue nulle, et même de dimension de Hausdorff au plus $d - 2$.

8. Voir dans annexes, A.1 pour plus de détails.

avec $\beta = \frac{1}{T}$ la température inverse et :

$$Z_{\beta,N} := \int_{K^N} e^{-\beta N E_N} \nu_0^{\otimes N}$$

le facteur de normalisation appelé fonction de partition. On peut montrer qu'on a équirépartition d'une configuration typique $P_N \in K^N$ de loi $\gamma_{\beta,N}$ sur une mesure d'équilibre ν_β en ce sens que la suite des mesures empiriques $\delta_N(P_N)$ converge en loi vers la mesure déterministe ν_β . La mesure ν_β est caractérisée comme l'unique argument minimum de la fonctionnelle d'énergie libre : $\mathbf{E} + \frac{1}{\beta} \mathbf{H}[\cdot | \nu_0]$.

De plus, ν_β est aussi une mesure de Gibbs, de la forme :

$$\nu_\beta = \frac{1}{Z_\beta} e^{-\beta U} \nu_0$$

avec U solution de l'équation "champ moyen de type Liouville" :

$$-\frac{1}{2} \Delta U = \frac{e^{-\beta U} \nu_0}{\int e^{-\beta U} \nu_0}$$

On a aussi :

$$\frac{1}{\beta} \log(Z_\beta) \xrightarrow{\beta \rightarrow +\infty} - \inf_{\mathcal{P}(K)} \left\{ \mathbf{E} + \frac{1}{\beta} \mathbf{H}[\cdot | \nu_0] \right\}$$

Cette convergence de la fonction de partition, quand la température tend vers 0, montre que la mesure de Gibbs ν_β se concentre exponentiellement vite vers les minima de U via un principe de grandes déviations, c'est donc l'ordre qui prédomine. Mais quand la température tend vers l'infini, la mesure de Gibbs converge vers ν_0 la loi de l'aléa initial, donc c'est le désordre qui l'emporte. Quant à la preuve de cette convergence quand la température tend vers 0, c'est une conséquence de la s.c.i, du théorème de Sanov et du principe variationnel de Gibbs.

Maintenant revenons à notre système à champ moyen (\star). En remarquant que :

$$\frac{1}{N} H_N(x_1, \dots, x_N) = \frac{1}{C_N^2} \sum_{1 \leq i < j \leq N} \Theta(x_i, x_j), \text{ avec } \Theta(x, y) = V(x) + V(y) + \frac{1}{2} W(x, y),$$

alors $\frac{1}{N} H_N$ est l'énergie microscopique d'ordre deux associée à notre système. Donc, sa limite macroscopique est l'énergie libre définie pour toute mesure de probabilité ν sur \mathbb{R}^d par :

$$\mathbb{E}_{\nu^{\otimes 2}}[\Theta] = 2 \cdot \mathbb{E}_\nu[V] + \frac{1}{2} \mathbb{E}_{\nu^{\otimes 2}}[W] = \mathbf{E}_f[\nu].$$

Ce qui justifie l'introduction dans notre problème de McKean-Vlasov des fonctionnelles \mathbf{E}_f et \mathbf{H}_W vu que comme on vient de le voir, elles régissent l'équilibre macroscopique qui correspond à une limite en champ moyen. De plus, l'énergie libre est donc s.c.i et sa borne inférieure est finie car c'est l'énergie macroscopique à l'équilibre, tout comme celle de $\frac{1}{N} H_N$ qui est l'énergie microscopique à l'équilibre convergeant vers la macroscopique. Et pour finir, les configurations microscopiques minimisant l'hamiltonien s'équirépartissent sur l'équilibre macroscopique lorsque N tend vers l'infini via l'application mesure empirique.

Définition 1.2 (Vocabulaire thermodynamique)

1. K est appelé espace des états internes.
2. $\varphi \in \mathcal{C}^0(K)$ est appelée observable interne.
3. $\mathcal{P}(K)$ espace des états macroscopiques qui correspondent aux distributions de particules identiques.
4. $\Phi \in \mathcal{C}^0(\mathcal{P})$ est appelée observable macroscopique. Et on remarque que toute observable interne définit une observable macroscopique par moyennisation.
5. $P \in K^N$ est appelé état microscopique. Et tout état microscopique définit un état macroscopique via δ_N l'application mesure empirique. Et notons que δ_N induit un homéomorphisme de K^N / \mathfrak{S}_N sur $\delta_N(K^N)$ avec \mathfrak{S}_N le groupe symétrique et :

$$\overline{\bigcup_N \delta_N(K^N)} = \mathcal{P}(K)$$

6. Pour chaque N , on modélise les interactions entre N points par une énergie d'interaction microscopique $\mathbf{E}_N : K^N \rightarrow \mathbb{R} \cup \{+\infty\}$ qui est, par définition, symétrique et semi-continue inférieurement. Les configurations à l'équilibre sont celles qui minimisent cette énergie.
7. Plus généralement, on dit que l'interaction est d'ordre r si \mathbf{E}_N est la moyenne des interactions entre r points, c'est-à-dire :

$$\mathbf{E}_N(x_1, \dots, x_N) = \frac{1}{C_N^r} \sum_{1 \leq i_1 < \dots < i_r \leq N} \mathbf{E}_r(x_{i_1}, \dots, x_{i_r})$$

Par symétrie, \mathbf{E}_N descend au quotient et on peut donc la voir comme une application $\widetilde{\mathbf{E}}_N : \mathcal{P}(K) \rightarrow \mathbb{R} \cup \{+\infty\}$ s.c.i telle que $\widetilde{\mathbf{E}}_N \equiv +\infty$ sur $\mathcal{P}(K) / \delta_N(K^N)$.

8. On dit que \mathbf{E}_N admet une limite macroscopique $\mathbf{E} : \mathcal{P}(K) \rightarrow \mathbb{R} \cup \{+\infty\}$ si $\widetilde{\mathbf{E}}_N$ converge faiblement vers \mathbf{E} .
9. si \mathbf{E}_N est d'ordre r , alors elle admet une limite macroscopique \mathbf{E} s.c.i définie par $\mathbf{E}[v] = \mathbb{E}_{v^{\otimes r}}[\mathbf{E}_r]$

Chapitre 2

\mathbb{V}, \mathbb{H} et inégalités fonctionnelles

Les inégalités de Sobolev logarithmiques doivent leur nom à un article célèbre de Gross paru en 1975. Ces inégalités fonctionnelles apparaissent en particulier comme une expression de la propriété d'hypercontractivité de semi-groupes markoviens, comme une traduction de la décroissance exponentielle de l'entropie le long de ces semi-groupes, comme une information de concentration gaussienne de la mesure, et enfin comme une façon de renforcer les inégalités de Poincaré, plus traditionnelles en analyse. Quant aux références, voir principalement en bibliographie [12], [14], \dots , [16], [18], [20] et [80], \dots , [92].

2.1 Variance et entropie

Définition 2.1 (Φ -entropie)

Soient $(\Omega, \mathcal{F}, \mathbb{P})$ un espace probabilisé, $\Phi : I \subset \mathbb{R}^d \rightarrow \mathbb{R}$ convexe et $X : \Omega \rightarrow I \subset \mathbb{R}^d$ un vecteur aléatoire tels que $X \in L^1$, $\Phi(X) \in L^1$ et $\mathbb{E}[X] \in I$. On appelle Φ -entropie de X , la quantité définie par :

$$\mathbb{E}^\Phi[X] := \mathbb{E}[\Phi(X)] - \Phi(\mathbb{E}[X]).$$

Par hypothèses, il est facile de voir que $\mathcal{D}^\Phi := \text{Dom}(\mathbb{E}^\Phi) \subset L^1$ est convexe et de plus, $\mathbb{E}^\Phi : \mathcal{D}^\Phi \rightarrow \mathbb{R}_+$ par inégalité de Jensen.

Remarques et exemples :

La variance et l'entropie sont des exemples de Φ -entropies pour respectivement $\Phi = \|\cdot\|^2$ et $\Phi : x \mapsto x \log(x)$. Il est facile de prouver¹ que la variance de X est exactement le carré de la distance en norme L^2 de X au sous espace des variables aléatoires presque sûrement constantes, c'est-à-dire :

$$\mathbb{V}[X] = \inf_{\lambda \in \mathbb{R}} \mathbb{E}[|X - \lambda|^2] = \inf_{\lambda \in \mathbb{R}} \mathbb{E}[\Phi(X) + \Phi(\lambda) - \Phi'(\lambda)X]$$

et cette borne inférieure est atteinte en $\lambda = \mathbb{E}[X]$. On parle de formulation variationnelle de Φ -entropies². De plus, on a :

$$\text{Ent}[(1 + \varepsilon X)^2] = 2\varepsilon^2 \mathbb{V}[X] + \mathbf{O}(\varepsilon^3)$$

et cette approximation montre avec les notations ci-dessous qu'une inégalité log-Sobolev de constante optimale ρ implique une inégalité de Poincaré de constante optimale $\lambda = \rho$.

Si Φ est strictement convexe, alors la Φ -entropie de X est nulle si et seulement si X est constante presque sûrement.

1. Par le théorème de la projection orthogonale sur un convexe fermé d'un espace de Hilbert.
2. Voir dualité de Monge-Kantorovitch et espaces de Wasserstein pour plus de détails.

2.2 Poincaré et log-Sobolev

Définition 2.2 (Inégalité de Poincaré)

Une mesure de probabilité μ sur \mathbb{R}^d vérifie une inégalité de Poincaré lorsqu'il existe une constante $\lambda > 0$ telle que pour toute fonction $f \in \mathcal{C}_b^1(\mathbb{R}^d)$:

$$\lambda \cdot \mathbb{V}_\mu[f] \leq \mathbb{E}_\mu \left[\|\nabla f\|^2 \right]$$

Et on appelle constante de Poincaré la meilleure constante $\lambda_1(\mu)$ pour laquelle on a une telle inégalité.

Poincaré et trou spectral : Cette constante optimale, dans certains cas bien particuliers, est aussi appelée constante d'écart spectral pour des raisons liées au spectre du générateur infinitésimal du semi-groupe markovien décrivant les lois d'un processus de diffusion de mesure invariante μ . C'est même équivalent, dans ces cas, au fait d'avoir une telle inégalité.³ Par exemple, dans le cas qui nous intéresse dans ce mémoire à savoir celui dans lequel μ est une mesure de Gibbs sur \mathbb{R}^d définie par :

$$\mu(dx) := \frac{1}{Z_H} e^{-H(x)} dx$$

avec $H \in \mathcal{C}^2(\mathbb{R}^d, \mathbb{R})$ telle que $Z_H := \int_{\mathbb{R}^d} e^{-H(x)} dx < +\infty$, μ apparaît comme la mesure invariante symétrique du processus de Kolmogorov $(X_t)_{t \geq 0}$ satisfaisant à l'équation différentielle stochastique :

$$dX_t = \sqrt{2} dB_t - \nabla H(X_t) dt$$

Cette équation décrit par exemple la vitesse X_t à l'instant $t \geq 0$ d'une particule qui subit à la fois une force d'agitation brownienne et une force de rappel associée au champ de potentiel H . Et à ce processus de diffusion, on associe un semi-groupe de Markov $(P_t)_{t \geq 0}$ tel que :

$$\forall (t, x) \in [0, +\infty[\times \mathbb{R}^d, f \in \mathcal{C}_b^0(\mathbb{R}^d, \mathbb{R}), P_t f(x) = \mathbb{E}_x[f(X_t)] := \mathbb{E}[f(X_t) | X_0 = x]$$

Ce semi-groupe de contraction de $L_{\mathbb{R}}^2(\mathbb{R}^d, \mu)$ a pour générateur infinitésimal⁴ l'opérateur différentiel \mathcal{L} défini par :

$$\mathcal{L} := \Delta - \nabla H \cdot \nabla$$

Et par définition, on sait que $\partial_t P_t = \mathcal{L} P_t = P_t \mathcal{L}$. Il est alors bien connu que l'inégalité de Poincaré de constante $\lambda_1(\mu)$ pour μ équivaut à l'existence d'un trou supérieur à $\lambda_1(\mu)$ dans le spectre de \mathcal{L} .

Définition 2.3 (Inégalité de Sobolev logarithmique)

Une mesure de probabilité μ sur \mathbb{R}^d vérifie une inégalité de Poincaré lorsqu'il existe une constante $\rho > 0$ telle que pour toute fonction $f \in \mathcal{C}_b^1(\mathbb{R}^d)$:

$$\rho \cdot \text{Ent}_\mu[f^2] \leq 2 \cdot \mathbb{E}_\mu \left[\|\nabla f\|^2 \right]$$

Et la meilleure constante $\rho_{LS}(\mu)$ pour laquelle une telle inégalité est vérifiée est appelée : constante de Sobolev logarithmique.

3. On parle souvent de propriété de trou spectral.

4. Via la formule d'Ito.

Sobolev et hypercontractivité : Quant à l'inégalité de Sobolev logarithmique pour la mesure de Gibbs μ , elle est équivalente⁵ à l'hypercontractivité⁶ du semi-groupe $(P_t)_{t \geq 0}$, c'est-à-dire :

$$\|P_t\|_{p \rightarrow q} \leq 1$$

avec $q \leq 1 + e^{4\rho_{LS}(\mu)t}$. Et ce résultat s'obtient en remarquant que la dérivée par rapport à p de $\|\cdot\|_p$ donne l'entropie⁷. De plus, on en déduit des inégalités de convergence dite exponentielle vers l'équilibre au sens de la variance pour l'inégalité de Poincaré et de l'entropie pour celle de Sobolev logarithmique, à savoir :

$$\begin{aligned} \mathbb{V}_\mu[P_t f] &\leq e^{-2t\lambda_1(\mu)} \mathbb{V}_\mu[f] \\ \text{Ent}_\mu[P_t f] &\leq e^{-4t\rho_{LS}} \text{Ent}_\mu[f] \end{aligned}$$

En effet, par exemple pour la variance nous avons :

Preuve

$$\partial_t \mathbb{V}_\mu[P_t f] = 2\mathbb{E}_\mu[P_t f \mathcal{L} P_t f] = -2\mathbb{E}_\mu \left[\|\nabla P_t f\|^2 \right] \leq -2\lambda_1(\mu) \mathbb{V}_\mu[P_t f]$$

par les théorèmes de la divergence, de Green et application de l'inégalité de Poincaré. Donc par le lemme de Gronwall, on en déduit que $\mathbb{V}_\mu[P_t f] \leq e^{-2t\lambda_1(\mu)} \mathbb{V}_\mu[P_0 f]$ et vu que $P_0 f = f$, on a bien le résultat attendu. \square

C'est ce genre de résultats qui nous intéresse et la deuxième partie de ce mémoire y est consacré pour notre système de champ moyen.

Inégalités fonctionnelles et courbure :

De plus, si le semi-groupe de diffusion associé à la mesure de Gibbs est ergodique et la hessienne de H est uniformément bornée au sens des applications linéaires, on a aussi les inégalités de Poincaré et de sobolev logarithmiques renversées (changement de sens des inégalités). Si $\exists \alpha > 0$ tel que $\forall x \in \mathbb{R}^d \nabla^2 H(x) \geq \alpha \mathbb{1}_{d \times d}$ au sens des formes quadratiques⁸, alors μ satisfait aux inégalités de Poincaré et de Sobolev logarithmique.⁹ Évidemment, ce n'est pas une équivalence vu que μ peut vérifier de telles inégalités sans convexité de H .

5. Théorème de GROSS.

6. Pour plus de détails, consulter le document intitulé diffusions hypercontractives de Dominique Bakry et Michel Émery.

7. Identités de GROSS.

8. On parle de coercivité de H , ellipticité de H , uniforme stricte convexité de H ou uniforme log-concavité de μ . C'est équivalent au fait que $\|\nabla P_t f\| \leq e^{-\alpha t} P_t \|\nabla f\|$.

9. Critère de courbure de Bakry et Emery.

2.3 PGD

Dans cette section, on rappelle quelques résultats sur la théorie des grandes déviations qui servent dans la suite de ce mémoire.

Théorème 2.1 (PGD et semi-groupes de diffusion)

On considère un semi-groupe de diffusion $(P_t)_{t \geq 0}$ sur \mathbb{R}^d de générateur infinitésimal $\mathcal{L} := \Delta - \nabla H \cdot \nabla$ tel que $\nabla^2 H$ est uniformément bornée en tant qu'endomorphisme de \mathbb{R}^d . Alors $\forall B \in \mathcal{B}(\mathbb{R}^d)$ et $x \in \mathbb{R}^d$

$$-\inf_{y \in \overset{\circ}{B}} J_x(y) \leq \liminf_{t \rightarrow 0^+} \{2t \log(P_t \mathbb{1}_B(x))\} \leq \limsup_{t \rightarrow 0^+} \{2t \log(P_t \mathbb{1}_B(x))\} \leq -\inf_{y \in \overline{B}} J_x(y)$$

avec $J_x := \frac{\|x - \cdot\|^2}{2}$ et ici la norme, comme dans ce qui précède est toujours euclidienne standard. On dit que la famille de distributions $(P_t^x)_{t > 0}$ vérifie un principe de grandes déviations de taux $\frac{1}{2t}$ et de fonction de taux J_x .

Le grand intérêt de ce théorème est le fait qu'il apporte une convergence exponentielle en loi pour des temps petits. Dans ce qui suit, on admet les résultats classiques de la théorie des grandes déviations tels que les théorèmes de Cramer, Sanov et Varadhan, le principe de contraction et leurs applications en physique statistique. Quant aux références, voir principalement en bibliographie [11], [35] et [65], [66], \dots , [71].

Le théorème précédent permet aussi d'obtenir des inégalités de concentration gaussienne¹⁰. Une question naturelle qu'on peut se poser est :

Est-ce que ce principe de grandes déviations en temps petits pour ce semi-groupe de Markov permettrait d'avoir des informations sur son comportement ergodique ?

Pour l'instant, on n'a pas encore de réponses (positives ou négatives) à cette question mais il serait intéressant d'y prêter attention, ce qui motive encore plus le fait qu'il figure dans ce rapport...

10. Pour plus de détails, voir par exemple l'article sur le sujet de Djalil Chafai.

Chapitre 3

Processus de Markov

Pour plus de détails sur ce chapitre, consulter principalement les références [96] et [98] en bibliographie.

3.1 Généralités

Dans cette section, Ω désigne l'espace des états considéré polonais et muni de sa tribu borélienne et d'une structure de variété différentiable, $\mu \in \mathcal{P}(\Omega)$ une mesure de probabilité sur cet espace d'états et $(P_t)_{t \geq 0}$ est un semi-groupe markovien qui est l'équivalent en temps continu des probabilités de transition pour les chaînes de Markov. Avant de pouvoir définir un tel semi-groupe, rappelons nous qu'il y a essentiellement deux types de processus de Markov dits respectivement inhomogènes et homogènes. Soit $(X_t)_{t \geq 0}$ un processus de Markov général à valeurs dans Ω . On définit une famille d'opérateurs $(P_{s,t})_{0 \leq s \leq t}$ associée à ce processus par :

$$P_{s,t}f(x) = \mathbb{E}_{X_s=x}[f(X_t)] = \mathbb{E}[f(X_t)|X_s = x],$$

et agissant à priori sur l'espace des fonctions mesurables bornées. Cette action s'étend à $L^2(\mu)$.

Définition 3.1 (Générateur de la famille)

On appelle générateur de $(P_{s,t})_{0 \leq s \leq t}$, l'application $t \mapsto \mathcal{L}_t$ telle que :

$$\mathcal{L}_t f = \lim_{s \rightarrow 0} \frac{P_{t,t+s}f - f}{s}, \text{ sur l'ensemble de toutes les fonctions telles que cette limite existe dans}$$

$L^2(\mu)$ et cet ensemble est appelé domaine du générateur et noté $\mathbf{Dom}(\mathcal{L}_t)$.

En toute rigueur, on ne parle de semi-groupe que dans le cas des processus de Markov homogènes, auquel cas $P_{s,t}$ ne dépend que de $t - s$ et sera simplement noté $P_t = P_{0,t}$, donc son générateur ne dépend pas du temps et sera noté \mathcal{L} . Notons que par dualité, l'action à droite de P_t sur l'espace des fonctions mesurables bornées implique une action à gauche sur les mesures de probabilité donnée pour toute mesure ν par :

$$\langle \nu P_t, f \rangle = \langle \nu, P_t f \rangle = \mathbb{E}_{X_0 \sim \nu}[f(X_t)], \text{ pour toute fonction test.}$$

On note \mathcal{A} l'algèbre des fonctions régulières dont toutes les dérivées appartiennent à $L^2(\mu)$. De plus, par l'égalité de Chapman-Kolmogorov $P_{s,t}P_{u,s} = P_{u,t}$, on a $\partial_t P_{s,t} = P_{s,t}\mathcal{L}_t$.

Définition 3.2 (Semi-groupe de Markov)

Un semi-groupe markovien $(P_t)_{t \geq 0}$ sur Ω est une famille de noyaux de probabilités telle que :

$$\forall (t, s) \in \mathbb{R}_+^2$$

$$P_0 := \mathbf{Id}, P_{t+s} := P_t \circ P_s$$

Et de plus il est entièrement décrit par son action sur les fonctions mesurables bornées, c'est-à-dire :

$$\forall f \in \mathcal{M}_b(\Omega), P_t f(x) := \int_{\Omega} f(y) P_t(x, dy)$$

On dit que $(P_t)_{t \geq 0}$ est fellerien si $\forall f \ t \mapsto P_t f$ est continue de $\mathbb{R}_+ \rightarrow L^2(\mu)$.

Étant donné ce semi-groupe, on lui associe un processus $(X_t)_{t \geq 0}$ de Markov et comme avec les chaînes de Markov, on a :

$$\mathbb{E}_x[f(X_t)] := \mathbb{E}[f(X_t) | X_0 = x] = P_t f(x)$$

On définit aussi les mesures stationnaires par analogie aux chaînes de Markov.

Définition 3.3 (Mesure invariante)

On dit que μ est invariante (ou stationnaire) pour $(P_t)_{t \geq 0}$ si $\forall f$

$$\mathbb{E}_{\mu}[P_t f] := \int_{\Omega} P_t f(x) \mu(dx) = \int_{\Omega} f(x) \mu(dx) := \mathbb{E}_{\mu}[f]$$

Dans tout ce qui suit, on suppose que le semi-groupe est toujours fellerien et dans ce cas, il définit un semi-groupe fortement continue d'opérateurs bornés de $L^2(\mu)$. Donc on peut lui associer un opérateur non borné $(D(\mathcal{L}), \mathcal{L})$ de $L^2(\mu)$ à domaine dense et dont la donnée équivaut à celle du semi-groupe car par définition : $\forall f \in D(\mathcal{L})$

$$\partial_t P_t f = \mathcal{L} P_t f = P_t \mathcal{L} f$$

On peut montrer aussi que μ est invariante pour le semi-groupe si et seulement si $\forall f \in D(\mathcal{L}) \ \mathbb{E}_{\mu}[\mathcal{L} f] = 0$ ou $\mathcal{L}' \mu = 0$ avec \mathcal{L}' l'adjoint de \mathcal{L} au sens des distributions. On va définir les formes dites de Dirichlet pour lesquelles on peut généraliser la notion d'inégalités de Poincaré et de Sobolev logarithmique dans un sens qu'on verra en fin de section.

Définition 3.4 (Formes de Dirichlet)

Étant donné un générateur infinitésimal \mathcal{L} d'un semi-groupe de Markov, on appelle carré du champ l'application bilinéaire symétrique positive ${}^a \Gamma$ définie par :

$$\Gamma(f, g) := \frac{1}{2} \left(\mathcal{L}(fg) - g\mathcal{L}(f) - f\mathcal{L}(g) \right)$$

On appelle forme de Dirichlet la forme bilinéaire symétrique positive \mathcal{E} définie par :

$$\mathcal{E}(f, g) := \int_{\Omega} \Gamma(f, g) \mu(dx)$$

De plus, dans le cas des mesures de Gibbs, un calcul direct montre que $\Gamma(f, f) = \|\nabla f\|^2$.

^a. vu que $\Gamma(f, f) = \lim_{t \rightarrow 0} \frac{P_t(f^2) - (P_t f)^2}{2t}$ par Cauchy-Schwarz

Définition 3.5 (Symétrie)

On dira que $(P_t)_{t \geq 0}$ est μ -symétrique si $\mathcal{L}^* = \mathcal{L}$ ou de façon équivalente : $\forall f, g$

$$\langle P_t f, g \rangle_\mu = \langle f, P_t g \rangle_\mu$$

Définition 3.6 (Résolvante)

On appelle résolvante du semi-groupe $(P_t)_{t \geq 0}$ l'application :

$$R_\lambda : \lambda \in]0, +\infty[\longmapsto \int_0^{+\infty} e^{-\lambda t} P_t dt$$

qui est la transformée de Laplace en temps de ce semi-groupe.

On peut montrer sans peine par une intégration par parties que $R_\lambda = (\lambda \mathbf{Id} - \mathcal{L})^{-1}$, ce qui justifie le nom de résolvante et de plus $\|R_\lambda\|_\infty \leq \frac{1}{\lambda}$. Il est facile de vérifier via l'inégalité d'Holder et l'invariance que $(P_t)_{t \geq 0}$ est un semi-groupe de contraction de $L^p(\mu)$ ¹. Donc, on peut étudier la nature du spectre du générateur infinitésimal via cette résolvante, par exemple, si la résolvante est compacte au sens des opérateurs sur un espace de Banach, alors le spectre du générateur est discret². Une application importante des processus de Markov³, qui sort un peu du cadre des probabilités, est la résolution de certaines équations aux dérivées partielles d'évolution de la forme $\partial_t u = \mathcal{L}u$ et $u(0, \cdot) = f$. En effet, si \mathcal{L} est le générateur d'un processus de Markov $(P_t)_{t \geq 0}$, on montre que :

$$u(t, x) = P_t f(x) = \mathbb{E}_x[f(X_t)]$$

pour toute condition initiale f admissible. Donc, la solution dépend linéairement de la condition initiale f et avec des propriétés de régularité sur le semi-groupe, cette solution dépend de façon continûment linéaire de f et l'EDP est régularisante. Un exemple classique de ce genre de problème d'évolution est donné par l'équation de la chaleur sur \mathbb{R}^d dont le générateur est $\frac{1}{2}\Delta$ et le semi-groupe est celui du mouvement brownien (processus de Wiener standard qui est un processus de Lévy à trajectoires continues et de dérive nulle) tel que $\forall f$:

$$P_t f(x) = \mathbb{E}_x[f(X_t)] = \mathbb{E}[f(x + B_t)] = \int_{\mathbb{R}^d} f(x + y) p_t(y) dy = \int_{\mathbb{R}^d} f(z) p_t(z - x) dz = f * p_t(x)$$

avec p_t la densité de probabilité de $B_t \sim \mathcal{N}(0, t \cdot \mathbf{Id})$.

Maintenant, on va voir comment on étend les inégalités fonctionnelles aux mesures invariantes des semi-groupes de Markov relativement à l'opérateur carré du champ.

1. Donc, par Hille-Yosida en annexes, son générateur est maximal dissipatif à domaine dense.

2. Que des valeurs propres de multiplicité finie qu'on démontre via les projecteurs de Riesz et la théorie de Fredholm.

3. On parle de la formule de Feynman-Kac dont un cas particulier est donné par l'équation backward de Kolmogorov. Une autre EDP intéressante est donnée par l'équation forward de Kolmogorov. Pour plus de détails, consulter en bibliographie la référence [118].

Définition 3.7 (Inégalités fonctionnelles)

On dit que μ satisfait une inégalité de :

1. Poincaré pour Γ s'il existe $\lambda > 0$ telle que $\forall f \in \mathcal{A}$:

$$\lambda \cdot \mathbb{V}_\mu[f] \leq \mathcal{E}(f, f)$$

2. Sobolev logarithmique pour Γ s'il existe $\rho > 0$ telle que $\forall f \in \mathcal{A}$:

$$\rho \cdot \mathbf{Ent}_\mu[f^2] \leq 2\mathcal{E}(f, f)$$

Les inégalités de Poincaré et de Sobolev logarithmiques pour μ sont équivalentes à des décroissances exponentielles du semi-groupe $(P_t)_{t \geq 0}$ respectivement en variance et en entropie, c'est-à-dire :

— **Poincaré :**

$$\forall f \in L^2(\mu) \quad t \geq 0, \quad \|P_t f - \langle \mu, f \rangle\|_{L^2(\mu)} \leq e^{-\lambda_1(\mu)t} \|f - \langle \mu, f \rangle\|_{L^2(\mu)} .$$

— **Log-Sobolev :**

$$\forall f \in L^1(\mu) \log L^1(\mu) \quad t \geq 0, \quad \mathbf{Ent}_\mu[P_t f] \leq e^{-\rho_{LS}(\mu)t} \mathbf{Ent}_\mu[f] .$$

Ici, la notation $L^1(\mu) \log L^1(\mu)$ désigne le domaine de définition de l'entropie sous μ .

Courbure et inégalités fonctionnelles :

Les critères de courbure de Bakry-Emery (pour les diffusions Markoviennes) et de Lott-Sturm-Villani (pour les espaces métriques mesurés) offrent un cadre systématique pour aborder les inégalités fonctionnelles...

Définition 3.8 (Inégalités de transport de Talagrand et généralités)

On dit que μ satisfait une :

- T_1 -inégalité de transport s'il existe $\alpha > 0$ telle que $\mathcal{W}_1(\cdot, \mu) \leq \sqrt{\alpha \mathbf{H}[\cdot|\mu]}$.
- T_2 -inégalité de transport s'il existe $\alpha > 0$ telle que $\mathcal{W}_2(\cdot, \mu) \leq \sqrt{\alpha \mathbf{H}[\cdot|\mu]}$.

De plus, comme avec les inégalités de Poincaré et de Sobolev logarithmiques, la seconde implique la première. La classe des probabilités vérifiant T_1 est identique à celle ayant un moment exponentiel d'ordre 2 fini^a. L'inégalité T_2 est nettement plus structurée que l'inégalité T_1 vu qu'elle implique une inégalité de trou spectral^b.

a. D'après Djellout, Guillin et Wu.

b. D'après Otto et Villani.

Otto-Villani : Une mesure de probabilité vérifiant l'inégalité de Poincaré a en particulier un support connexe, ce qui impose d'emblée une restriction géométrique à la classe des mesures de probabilités vérifiant T_2 .

Pour ce qui est du cas particulier des mesures de Gibbs⁴, consulter le chapitre 4 section 4.3 sur le couplage et les métriques de Wasserstein...

4. Qui sont les mesures pour lesquelles on établit de telles inégalités dans ce mémoire.

3.2 E.U de mesures invariantes

Dans cette section, notre objectif est d'établir des conditions pour que le processus de diffusion d'Ito homogène à valeurs dans \mathbb{R}^d :

$$dX_t = b(X_t)dt + \sigma(X_t)dB_t$$

soit bien défini (existences locale et globale) et admette des mesures invariantes (existence et unicité) avec $b : \mathbb{R}^d \rightarrow \mathbb{R}^d$, $\sigma : \mathbb{R}^d \rightarrow \mathbb{R}^{d \times m}$ et $(B_t)_{t \geq 0}$ respectivement les termes de dérive, de diffusion ou volatilité et brownien m -dimensionnel. Déjà la théorie des EDS nous dit que :

1. Si b et σ sont localement lipschitziens, alors on sait que $\forall x \in \mathbb{R}^d$, il y a existence locale sur un intervalle aléatoire $[0, \tau]$ avec $\tau > 0$ presque sûrement.
2. Si b et σ sont à croissance sous linéaire, il y a existence globale ($\tau = +\infty$ presque sûrement).

On note dans ce qui suit $L := b \cdot \nabla + \frac{1}{2} \text{Tr}(\sigma \sigma^* \times \nabla^2)$ le générateur⁵ de l'EDS de diffusion homogène. On rappelle aussi que d'après la formule d'Ito⁶, $\forall \varphi \in \mathcal{C}^{1,2}(\mathbb{R}_+ \times \mathbb{R}^d)$:

$$d\varphi(t, X_t) = \frac{\partial \varphi}{\partial t}(t, X_t)dt + \nabla \varphi(t, X_t) \cdot dX_t + \frac{1}{2} \text{Tr}(\sigma \sigma^* \times \nabla^2 \varphi)(t, X_t)dt$$

On en déduit que :

$$d\varphi(t, X_t) = \frac{\partial \varphi}{\partial t}(t, X_t)dt + L\varphi(t, X_t)dt + \nabla \varphi(t, X_t) \cdot \sigma(X_t)dB_t$$

Ainsi :

$$d\varphi(t, X_t) = (\partial_t + L)(\varphi)(t, X_t)dt + dM_t^\varphi = L_t \varphi(t, X_t)dt + dM_t^\varphi$$

avec $L_t := \partial_t + L$ opérateur différentiel non autonome et $M_t^\varphi := \int_0^t \nabla \varphi(s, X_s) \cdot \sigma(X_s)dB_s$ un terme martingale dont la nature dépend de φ et σ . On parle souvent de problème de martingales. Remarquons que L_t et L coïncident sur l'espaces des fonctions qui ne dépendent pas du temps.

3.2.1 Non explosion et fonction de Lyapunov

Définition 3.9 (Fonction de Lyapunov)

Soit $U : \mathbb{R}^d \rightarrow \mathbb{R}$ une fonction continue. On dit qu'elle est de Lyapunov si $U \geq 1$ et $\lim_{\|x\| \rightarrow +\infty} U(x) = +\infty$. En particulier, les ensembles de niveau $K_a := \{x \in \mathbb{R}^d, U(x) \leq a\}$ sont compacts.

Théorème 3.1 (Existence globale)

Si b et σ sont localement lipschitziens et il existe une constante c et une fonction de Lyapunov U telles que $LU \leq c \cdot U$, alors il y a existence globale des solutions pour tout point de départ $x \in \mathbb{R}^d$ et de plus en notant $(P_t)_{t \geq 0}$ le semi-groupe markovien associé à la diffusion, on a $P_t U(x) := \mathbb{E}[U(X_t^x)] \leq e^{ct} U(x)$.

5. Opérateur différentiel du second ordre donné par la formule d'Ito.
6. Condition de semi-martingale continue sur la diffusion pour appliquer cette formule.

Preuve

Étant donné $x \in \mathbb{R}^d$, soit (X_t^x) la solution issue de x de l'EDS et définie sur $[0, \tau]$. Soit :

$$\tau_n := \inf\{t > 0, U(X_t^x) > n\}$$

le temps de sortie du compact $K_n := U^{-1}([0, n])$. Si $n > U(x)$, alors l'équation différentielle stochastique de coefficients $b_n := b \wedge n$ et $\sigma_n := \sigma \wedge n$ globalement lipschitziens admet une unique solution issue de x qu'on note (Y_t^x) . On en déduit que $\tau > \tau_n$, $X_t^x = Y_t^x$ sur $t \leq \tau_n$ et en appliquant la formule d'Ito sur $[0, t \wedge \tau_n]$ à $M_t := U(X_t^x)e^{-ct}$, on obtient :

$$M_{t \wedge \tau_n} = M_0 + \int_0^{t \wedge \tau_n} e^{-cs} \nabla U(X_s^x) \cdot \sigma(X_s^x) dB_s + \int_0^{t \wedge \tau_n} e^{-cs} (LU - c \cdot U)(X_s^x) ds$$

et le terme :

$$N_{t \wedge \tau_n} := M_0 + \int_0^{t \wedge \tau_n} e^{-cs} \nabla U(X_s^x) \cdot \sigma(X_s^x) dB_s = U(x) + \int_0^{t \wedge \tau_n} e^{-cs} \nabla U(X_s^x) \cdot \sigma(X_s^x) dB_s$$

est une martingale locale continue et positive car $M_t \geq 0$ et $LU - c \cdot U \leq 0$, elle définit donc une vraie sur-martingale et en particulier, on a :

$$\mathbb{E}_x[M_{t \wedge \tau_n}] \leq \mathbb{E}_x[N_{t \wedge \tau_n}] \leq \mathbb{E}_x[M_0] = U(x)$$

On en déduit que :

$$U(x) \geq \mathbb{E}_x[U(X_{t \wedge \tau_n}^x) e^{-ct \wedge \tau_n}] \geq \mathbb{E}_x[U(X_{\tau_n}^x) \mathbb{1}_{\{\tau_n \leq t\}}] = n \mathbb{P}_x(\tau_n \leq t)$$

Ainsi $\forall t > 0$, on a $\mathbb{P}_x(\tau_n \leq t) \xrightarrow{n \rightarrow +\infty} 0$ ($\implies \tau_n \xrightarrow{\mathbb{P}_x - p.s} +\infty$). Il en découle que $\tau \geq \limsup_n \tau_n = +\infty$ \mathbb{P}_x -presque sûrement, ce qui prouve l'existence globale.

Quant à l'inégalité du théorème, à t fixé, par le lemme de Fatou, on a :

$$U(x) \geq \liminf_n \mathbb{E}_x[U(X_{t \wedge \tau_n}^x) e^{-ct \wedge \tau_n}] \geq \mathbb{E}_x[\liminf_n U(X_{t \wedge \tau_n}^x) e^{-ct \wedge \tau_n}] = \mathbb{E}_x[U(X_t^x) e^{-ct}]$$

Ce qui achève la preuve du théorème. \square

3.2.2 Existence de mesures invariantes

Pour l'existence de mesures invariantes, comme la plupart du temps, il est difficile voire impossible de trouver des distributions de probabilité qui appartiennent au noyau du dual au sens des distributions de l'opérateur différentiel L , on va utiliser la notion de tension⁷ qui assure une pré-compacité⁸.

Théorème 3.2 (Condition suffisante de tension)

Soient E un espace polonais et $\mathcal{S} \subset \mathcal{P}(E)$. Si $\exists U : E \rightarrow \mathbb{R}$ telle que $U \geq 0$, U tend vers l'infini à l'infini et :

$$\sup_{\mu \in \mathcal{S}} \langle \mu, U \rangle < +\infty,$$

alors \mathcal{S} est tendue.

Preuve

En effet, pour $\mu \in \mathcal{S}$ et $\alpha > 0$, par l'inégalité de Markov, on a :

$$\mu(U > \alpha) \leq \frac{\langle \mu, U \rangle}{\alpha} \leq \frac{C}{\alpha} \text{ avec } C := \sup_{\mu \in \mathcal{S}} \langle \mu, U \rangle$$

7. Voir chapitre suivant sur la convergence faible et le théorème de Prokhorov pour plus de clarté.

8. Les deux notions sont même équivalentes sur un espace polonais ce qui est notre cas d'étude.

On en déduit que pour α assez grand :

$$\mu(U > \alpha) \leq \varepsilon$$

Soit K un compact tel que $U > \alpha$ sur K^c , alors :

$$\mu(K^c) \leq \mu(U > \alpha) \leq \varepsilon$$

On en conclut que :

$$\sup_{\mu \in \mathcal{S}} \mu(K^c) \leq \varepsilon$$

Ce qu'il fallait prouver. \square

Si de plus, E est localement compact, alors la réciproque de ce théorème est vraie.

Preuve

En effet, dans ce cas, il existe une suite croissante (K_n) de compacts telle que $\sup_{\mu \in \mathcal{S}} \mu(K_n^c) \leq 2^{-2n}$, et on peut, quitte à les grossir, supposer que $E = \bigcup K_n$ et dans ce cas, la fonction U suivante convient :

$$U := \sum_n 2^n (\mathbb{1}_{K_n} - \mathbb{1}_{K_{n-1}})$$

Ce qui termine la preuve. \square

Théorème 3.3 (Existence de mesures invariantes)

Si le semi-groupe (P_t) est fellerien et il existe une fonction de Lyapunov U et $x \in E$ tels que :

$$\sup_t P_t U(x) < +\infty,$$

alors il existe une mesure de probabilité invariante.

Preuve

En effet, il suffit de considérer le flot de mesures de probabilité (μ_t) défini $\forall \varphi$ admissibles par :

$$\langle \mu_t, \varphi \rangle := \frac{1}{t} \int_0^t P_s \varphi(x) ds$$

On a de façon directe par hypothèses :

$$\sup_t \langle \mu_t, U \rangle \leq \sup_t P_t U(x) < +\infty$$

ce qui prouve que la famille (μ_t) est tendue, donc par le théorème de Prokhorov, il existe une sous suite (μ_{t_n}) (avec $t_n \rightarrow +\infty$) qui converge faiblement vers μ_∞ une mesure de probabilité. En d'autres termes, $\forall \varphi$ continue bornée $\langle \mu_{t_n}, \varphi \rangle \rightarrow \langle \mu_\infty, \varphi \rangle$. Or pour $r > 0$, vu que le semi-groupe est fellerien, $\forall \varphi$ continue bornée, $P_r \varphi$ est aussi continue bornée, et donc $\langle \mu_{t_n}, P_r \varphi \rangle \rightarrow \langle \mu_\infty, P_r \varphi \rangle$. De plus :

$$\langle \mu_{t_n}, P_r \varphi \rangle := \frac{1}{t_n} \int_0^{t_n} P_s P_r \varphi(x) ds = \frac{1}{t_n} \int_r^{r+t_n} P_s \varphi(x) ds = \frac{1}{t_n} \left((t_n + r) \langle \mu_{t_n+r}, \varphi \rangle - r \langle \mu_r, \varphi \rangle \right) \rightarrow \langle \mu_\infty, \varphi \rangle$$

Ce qui prouve que $\langle \mu_\infty, P_r \varphi \rangle = \langle \mu_\infty, \varphi \rangle$. On a ainsi montré que μ_∞ est une mesure invariante pour le semi-groupe, ce qui achève la preuve de l'existence d'une telle mesure. \square

On va maintenant donner des conditions suffisantes sur la fonction de Lyapunov d'existence de mesures de probabilité invariantes. Remarquons que les deux hypothèses du théorème ci-dessous sont entraînées par l'unique inégalité $LU \leq -\alpha \cdot U + \beta$ pour deux constantes $\alpha > 0$ et $\beta \in [0, +\infty[$ vu que dans ce cas, $LU \leq \beta \leq \beta \cdot U$ et de plus :

$$\partial_t P_t U = LP_t U = P_t LU \leq -\alpha \cdot P_t U + \beta$$

Et en intégrant cette inégalité différentielle⁹, on obtient :

$$P_t U \leq e^{-\alpha t} \left(U + \beta \int_0^t e^{\alpha s} ds \right) = e^{-\alpha t} U + \frac{\beta}{\alpha} (1 - e^{-\alpha t}) \leq e^{-\alpha t} U + \frac{\beta}{\alpha}$$

Donc, à $t > 0$ fixé, si on considère le compact $K := \{U \leq a\}$ pour a assez grand de telle sorte que $\kappa = e^{-\alpha t} + \frac{\beta}{\alpha a} < 1$, alors sur K^c , $U > a$ et :

$$P_t U \leq e^{-\alpha t} U + \frac{\beta}{\alpha} \leq \kappa \cdot U$$

Ce qui montre le résultat attendu.

Théorème 3.4 (Existence via Lyapunov)

Soient $c > 0$, $b < +\infty$, $\kappa \in]0, 1[$, $t_0 > 0$ et U une fonction de Lyapunov telles que :

1. $LU \leq c \cdot U$;
2. $P_{t_0} U \leq \kappa \cdot U + b \cdot \mathbb{1}_K$;

alors il y a une probabilité invariante.

Preuve

En effet, on a :

$$P_{2t_0} U \leq \kappa \cdot P_{t_0} U + b \cdot P_{t_0} \mathbb{1}_K \leq \kappa^2 \cdot U + \kappa b \cdot \mathbb{1}_K + b$$

Et donc par récurrence immédiate :

$$P_{nt_0} U \leq \kappa^n \cdot U + b \left(\kappa^{n-1} \cdot \mathbb{1}_K + \sum_{j=0}^{n-2} \kappa^j \right) = \kappa^n \cdot U + b \left(\kappa^{n-1} \cdot \mathbb{1}_K + \frac{1 - \kappa^{n-1}}{1 - \kappa} \right) \leq \kappa^n \cdot U + \frac{b}{1 - \kappa}$$

Vu que $LU \leq c \cdot U \implies \forall s, P_s U \leq e^{cs} U$, on a pour $t \in [nt_0, (n+1)t_0]$:

$$P_t U = P_{nt_0} P_{t-nt_0} U \leq e^{c(t-nt_0)} P_{nt_0} U \leq e^{ct_0} P_{nt_0} U \leq e^{ct_0} \left(\kappa^n \cdot U + \frac{b}{1 - \kappa} \right) \longrightarrow \frac{b}{1 - \kappa} e^{ct_0}$$

C'est donc une suite convergente, donc bornée, ainsi on conclut par le théorème précédemment prouvé. \square

On peut sous les mêmes hypothèses établir une propriété de récurrence donnée par le théorème ci-dessous.

Théorème 3.5 (Atteinte en temps fini d'un compact)

Sous les hypothèses du théorème précédent, si :

$$\tau_K := \inf\{t > 0, X_t^x \in K\}$$

désigne le premier temps de passage (retour) en K , alors il existe $\delta > 0$ telle que $\forall x, \mathbb{E}_x[e^{\delta \tau_K}] < +\infty$. En particulier, pour tout point de départ x , on atteint le compact K en un temps fini presque sûrement.

Preuve

Quitte à grossir K , on peut supposer que $K = \{U \leq a\}$. On va montrer par récurrence que $\tau := \tau_K$ satisfait $\forall x \notin K$:

$$\mathbb{P}_x(\tau > nt_0) \leq \frac{\kappa^n}{a} U(x)$$

9. Ou en utilisant directement le lemme de Gronwall.

En effet, par la propriété de Markov, vu que sur $\tau > t_0$, $X_{t_0}^x \notin K$, on a :

$$\mathbb{P}_x(\tau > (n+1)t_0) = \mathbb{P}_x(\tau > t_0, \{\tau > nt_0\} \circ \Theta_{t_0}) = \mathbb{E}_x[\mathbb{1}_{\{\tau > t_0\}} \mathbb{P}_{X_{t_0}^x}(\tau > nt_0)] \leq \mathbb{E}_x[\mathbb{1}_{\{\tau > t_0\}} \frac{\kappa^n}{a} U(X_{t_0}^x)]$$

On en déduit que :

$$\mathbb{P}_x(\tau > (n+1)t_0) \leq \frac{\kappa^n}{a} \mathbb{E}_x[U(X_{t_0}^x)] = \frac{\kappa^n}{a} P_{t_0} U(x) \leq \frac{\kappa^{n+1}}{a} U(x)$$

On montre aisément l'existence d'un moment exponentiel pour la variable aléatoire $\tau = \tau_K$ vu que de l'encadrement $t \in [nt_0, (n+1)t_0]$ et ce qui précède, on a pour une constante $\gamma > 0$:

$$\mathbb{P}_x(\tau > t) \leq e^{-\gamma t}$$

Et par Fubini-Tonelli, on sait que pour tout $\delta > 0$:

$$\mathbb{E}_x[e^{\delta\tau}] = \int_0^{+\infty} \mathbb{P}_x(e^{\delta\tau} > t) dt = \int_0^{+\infty} \mathbb{P}_x\left(\tau > \frac{1}{\delta} \log(t)\right) dt = \int_0^1 \mathbb{P}_x\left(\tau > \frac{1}{\delta} \log(t)\right) dt + \int_1^{+\infty} \mathbb{P}_x\left(\tau > \frac{1}{\delta} \log(t)\right) dt$$

Or vu que sur $[0, 1]$, le logarithme est négatif et qu'un temps d'arrêt est positif, on a :

$$\int_0^1 \mathbb{P}_x\left(\tau > \frac{1}{\delta} \log(t)\right) dt = 1$$

D'autre part, on a :

$$\int_1^{+\infty} \mathbb{P}_x\left(\tau > \frac{1}{\delta} \log(t)\right) dt \leq \int_1^{+\infty} t^{-\frac{\gamma}{\delta}} dt$$

et cette dernière intégrale est classique (impropre de Riemann), de plus nous savons qu'elle converge si et seulement si $\frac{\gamma}{\delta} > 1$. En conclusion pour avoir un δ -moment exponentiel pour τ , il suffit de prendre $\delta < \gamma$. Quant à la finitude presque sûrement de τ , elle découle du fait que toute fonction intégrable est finie presque partout, donc pour $\delta < \gamma$, $e^{\delta\tau} < +\infty$ \mathbb{P}_x -p.s, ce qui montre que τ aussi. \square

3.2.3 Unicité de mesures invariantes

Pour l'unicité, on va faire deux hypothèses supplémentaires en plus de celles de l'existence à savoir :

1. Une hypothèse de régularité sur le semi-groupe. On rappelle que $(P_t)_{t \geq 0}$ est fellerien s'il envoie l'espace des fonctions continues bornées dans lui-même (on parle de sous-espace stable)¹⁰ et on sait par définition que $P_t L^\infty \subset L^\infty$ avec L^∞ l'espace des fonctions mesurables et essentiellement bornées¹¹. On dit que $(P_t)_{t \geq 0}$ est fortement fellerien si pour tout $t > 0$, $P_t L^\infty \subset \mathcal{C}_b$. Il est dit régularisant si pour tout $t > 0$, il admet une densité lisse et si $P_t L^\infty \subset \mathcal{C}^\infty$. On établit cette propriété en utilisant le critère de Hormander.
2. Une hypothèse d'irréductibilité qui stipule qu'il existe $t_0 > 0$ tel que $\forall O$ ouvert, on a $P_{t_0} \mathbb{1}_O > 0$. On peut montrer par les équations de Chapman-Kolmogorov que cette propriété est vérifiée pour tout $t \geq t_0$. Et cette seconde hypothèse peut être établie en utilisant le théorème de support de Stroock-Varadhan.

On rappelle le théorème ergodique de Birkhoff qui est très important pour l'étude de mesures invariantes. Avec les notations du théorème de Birkhoff, on dit que la probabilité \mathbb{P} ou le semi-groupe $(\Phi_t)_{t \geq 0}$ sont dits ergodiques si la tribu \mathcal{I} des invariants est presque sûrement triviale, c'est-à-dire que pour tout $A \in \mathcal{I}$, $\mathbb{P}(A) = 0$ ou 1 , ou encore les variables \mathcal{I} -mesurables sont \mathbb{P} -presque

10. On peut le démontrer en utilisant la continuité de la solution d'une EDS par rapport aux paramètres.

11. Il est facile de montrer que le semi-groupe vu comme une famille d'opérateurs sur cet espace est un semi-groupe de contraction fortement continu, ce qui permet d'avoir des résultats très importants du type Hille-Yosida.

sûrement constantes et dans ce cas pour X une telle variable de plus intégrable, on a presque sûrement :

$$\mathbb{E}[X|\mathcal{F}] = \int X d\mathbb{P} := \mathbb{E}[X]$$

Étant donnée une mesure invariante π pour le semi-groupe $(P_t)_{t \geq 0}$, on note \mathbb{P}_π la loi du processus de loi initiale π , c'est-à-dire on se place sur l'espace canonique avec les applications coordonnées $X_t(\omega) = \omega(t)$ et le semi-groupe du shift défini par :

$$\Theta_t(\omega)(s) = \omega(t + s)$$

Dans ce cas, on a :

$$\mathbb{P}_\pi(A) = \int \mathbb{P}_x(A) \pi(dx) = \mathbb{P}_\pi(X \in A).$$

Et le shift préserve la mesure \mathbb{P}_π car sous \mathbb{P}_π , X_0 a pour loi π , donc X_t a pour loi $\pi P_t = \pi$ et par la propriété de Markov :

$$\mathbb{E}_\pi[f \circ \Theta_t(\omega)] = \mathbb{E}_\pi[\mathbb{E}_{X_t}[f(\omega)]] = \int \mathbb{P}_\pi(X_t \in dx) \mathbb{E}_x[f] = \mathbb{E}_\pi[f]$$

Théorème 3.6 (Ergodique de Birkhoff)

Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace probabilisé. Soit $(\Phi_t)_{t \geq 0}$ un semi-groupe de transformations de Ω qui préserve la mesure de probabilité \mathbb{P} , c'est-à-dire $\Phi_t : \Omega \rightarrow \Omega$, $\Phi_{t+s} = \Phi_t \circ \Phi_s$ et $\forall A \in \mathcal{F}$, $\mathbb{P}(\Phi_t^{-1}(A)) = \mathbb{P}(A)$. Alors, $\forall X \in L^1(\mathbb{P})$:

$$\frac{1}{t} \int_0^t X \circ \Phi_s(\omega) ds \xrightarrow{t \rightarrow +\infty} \mathbb{E}[X|\mathcal{I}] \mathbb{P} - ps \text{ tout } \omega \in \Omega$$

Et la variable limite désigne l'espérance conditionnelle de X par rapport à \mathbb{P} et la tribu des invariants $\mathcal{I} := \{A \in \mathcal{F}, \Phi \cdot A = A\}$

Théorème 3.7

Si le semi-groupe $(P_t)_{t \geq 0}$ admet deux mesure invariantes π_1 et π_2 distinctes, alors $\forall \alpha \in (0, 1)$, si $\pi = \alpha\pi_1 + (1 - \alpha)\pi_2$, la probabilité \mathbb{P}_π n'est pas ergodique et il existe une fonction h positive, bornée, non constante et invariante pour le semi-groupe.

Preuve

Pour ce faire, raisonnons par l'absurde et supposons que \mathbb{P}_π est ergodique. Ainsi pour tout $A \in \mathcal{I}$, $\mathbb{P}_\pi(A) = 0$ ou 1 et on voit que : $\mathbb{P}_\pi(A) = 0 \implies \mathbb{P}_{\pi_1}(A) = \mathbb{P}_{\pi_2}(A) = 0$ et $\mathbb{P}_\pi(A) = 1 \implies \mathbb{P}_{\pi_1}(A) = \mathbb{P}_{\pi_2}(A) = 1$. Pour f mesurable bornée, par le théorème ergodique de Birkhoff appliqué à la fonction $F(\omega) = f(X_0(\omega))$, il existe A tel que $\mathbb{P}_\pi(A) = 1$ et $\forall \omega \in A$:

$$\frac{1}{t} \int_0^t f(\omega(s)) ds \xrightarrow{t \rightarrow +\infty} m := \int f d\pi = \langle \pi, f \rangle$$

Vu que $\mathbb{P}_{\pi_1}(A) = \mathbb{P}_{\pi_2}(A) = 1$, par convergence dominée de Lebesgue, pour $i = 1, 2$:

$$m = \lim_{t \rightarrow +\infty} \frac{1}{t} \int_0^t \mathbb{E}_{\pi_i}[f(\omega(s))] ds = \mathbb{E}_{\pi_i}[f] = \langle \pi_i, f \rangle$$

On en déduit que $\pi_1 = \pi_2$, ce qui est absurde vu que par hypothèses elles sont distinctes. Ainsi \mathbb{P}_π n'est pas ergodique. Il reste juste à prouver l'existence de la fonction h et pour ça, il suffit de remarquer que la non

ergodicité implique qu'il existe $A \in \mathcal{F}$ tel que $\mathbb{P}_\pi(A) \in]0, 1[$. En posant $h(x) = \mathbb{P}_x(A) \in (0, 1)$, par la propriété de Markov, on a :

$$h(x) = \mathbb{P}_x(\Theta_t^{-1}(A)) = \mathbb{E}_x[\mathbb{P}_{X_t}(A)] = \mathbb{E}_x[h(X_t)] = P_t h(x)$$

Il en découle que $(h(X_t))_{t \geq 0}$ est une martingale positive qui converge presque sûrement et :

$$h(X_t) = \mathbb{P}_{X_t}(A) = \mathbb{E}[\mathbb{1}_A \circ \Theta_t | \mathcal{F}_t] = \mathbb{E}[\mathbb{1}_A | \mathcal{F}_t] \xrightarrow{t \rightarrow +\infty} \mathbb{1}_A \text{ p.s}$$

Donc, si la fonction h était constante, cette constante appartiendrait à la paire $\{0, 1\}$ et cela entraînerait que $\mathbb{P}_\pi(A) \in \{0, 1\}$, ce qui serait contradictoire. Ce qu'il fallait démontrer. \square

Théorème 3.8 (Unicité de la mesure invariante)

En admettant les hypothèses du théorème 3.1 d'existence globale des solutions d'une EDS, si le semi-groupe est irréductible et fortement fellerien, alors, il existe une unique mesure de probabilité invariante.

Quant à la preuve de ce théorème, elle utilise essentiellement les notions de récurrence et d'ergodicité. Nous allons voir maintenant que l'hypothèse d'irréductibilité entraîne la récurrence des ouverts.

Théorème 3.9 (Récurrence des ouverts)

Sous les hypothèses du théorème 3.8 d'unicité de la mesure invariante, tout ouvert non vide est récurrent pour la chaîne de Markov $(X_{n t_0})_{n \in \mathbb{N}}$. En conséquence, toute fonction f mesurable bornée ou positive, qui est invariante pour la chaîne $(P_{t_0} f = f)$, est constante presque sûrement.

Preuve

Soit O un ouvert non vide. Vu que $\mathbb{1}_O$ est bornée, la fonction $x \mapsto P_{t_0} \mathbb{1}_O(x)$ est continue, et on a par irréductibilité et compacité :

$$\delta := \inf_{x \in K} P_{t_0} \mathbb{1}_O(x) = \min_{x \in K} P_{t_0} \mathbb{1}_O(x) > 0$$

En considérant la chaîne de Markov définie par $Y_n := X_{n t_0}$, de matrice de transition $Q := P_{t_0}$, et le schéma de succès-échec défini par $\tau_0 = 0$ et :

$$\tau_p = \inf\{n > \tau_{p-1} + 1, Y_n \in K\}, \xi_p = \mathbb{1}_{\{Y_{\tau_p+1} \in O\}},$$

on a par le théorème 3.5 d'atteinte en temps fini d'un compact, $\forall p, \tau_p$ est fini presque sûrement et :

$$\mathbb{P}_x(\xi_1 = 0) = \mathbb{P}_x(Y_{\tau_K+1} \notin O) = \mathbb{E}_x[\mathbb{P}_{Y_{\tau_K}}(Y_1 \notin O)] \leq 1 - c.$$

On en déduit par récurrence et la propriété de Markov que :

$$\mathbb{P}_x\left(\bigcap_{j=1}^p \{\xi_j = 0\}\right) = \mathbb{E}_x\left[\mathbb{1}_{\{Y_{\tau_K+1} \notin O\}} \mathbb{P}_{Y_{\tau_K+1}}\left(\bigcap_{j=1}^{p-1} \{\xi_j = 0\}\right)\right] \leq (1 - c)^p.$$

Ce qui entraîne que :

$$\mathbb{P}_x\left(\bigcap_{\text{tous les } p} \{\xi_p = 0\}\right) = 0, \text{ et donc } O \text{ est récurrent.}$$

Soit maintenant f mesurable, bornée ou positive et invariante pour la chaîne. Vu que le semi-groupe est fortement fellerien, $P_{t_0} f = f$ est continue. Si f était non constante, il existerait $\alpha < \beta$ deux constantes telles que les ensembles $O_1 := \{f < \alpha\}$ et $O_2 := \{f > \beta\}$ soient des ouverts non vides, donc récurrents. Vu que $M_n := f(Y_n)$ est une martingale bornée ou positive, donc convergeant presque sûrement vers $M_\infty \in L^1$, par récurrence des ouverts O_1 et O_2 , on aurait presque sûrement $\beta \leq M_\infty \leq \alpha$, ce qui serait contradictoire car $\alpha < \beta$. Ce qui montre que f est bel et bien constante. \square

3.3 Convergence exponentielle

Pour une chaîne de Markov sur un espace d'états fini, l'irréductibilité de sa matrice de transition assure la convergence exponentielle vers l'unique mesure invariante via le théorème de Perron-Frobenius.

Théorème 3.10 (Perron-Frobenius)

Soit P une matrice stochastique sur un espace d'états fini telle qu'il existe k , pour tout $n \geq k$, la matrice P^n est à coefficients strictement positifs, c'est-à-dire pour tous (i, j) , $(P^n)_{i,j} > 0$. Alors P admet une unique mesure de probabilité invariante π ($\pi P = \pi$) et de plus, il existe $\gamma > 0$ et $C > 0$ des constantes telles que :

$$\sup_{x,y} |P^n_{x,y} - \pi(y)| \leq C e^{-\gamma n}$$

L'analogie en théorie des semi-groupes existe, si on considère des semi-groupes compacts vus comme opérateurs sur un espace de Banach \mathcal{H}_U dépendant d'une fonction de Lyapunov U et défini par :

$$\mathcal{H}_U := \left\{ f : E \rightarrow \mathbb{R} \text{ boréliennes telles que } \|f\|_U < +\infty \right\}, \text{ avec } \|f\|_U := \sup_{x \in E} \frac{|f(x)|}{U(x)}$$

Et on remarque que $\|f\|_U \leq \|f\|_\infty$, ce qui donne une injection continue de L^∞ dans cet espace.

Théorème 3.11 (Convergence exponentielle)

On suppose que :

1. le semi-groupe est irréductible et régularisant (ici, $E = \mathbb{R}^d$);
2. il existe $c > 0$ et une fonction de Lyapunov U telles que $LU \leq c \cdot U$;
3. il existe une suite de compacts (K_n) , des constantes $b_n \in (0, +\infty)$ et $\kappa_n \in]0, 1[$ telles que $\lim \kappa_n = 0$ et pour un $t_0 > 0$:

$$P_{t_0} U \leq \kappa_n \cdot U + b_n \cdot \mathbb{1}_{K_n}$$

Alors, $(P_t)_{t \geq 0}$ définit un semi-groupe d'opérateurs sur \mathcal{H}_U qui sont compacts pour tout $t \geq t_0$. De plus, le processus admet une unique mesure invariante π et il existe des constantes $\gamma > 0$ et $C > 0$ telles que :

$$\|P_t - \pi\|_U \leq C e^{-\gamma t}$$

C'est-à-dire que pour tout $f \in \mathcal{H}_U$:

$$\|P_t f - \langle \pi, f \rangle\|_U \leq C e^{-\gamma t} \|f\|_U$$

Preuve

Vu que par le point (2), $P_t U \leq e^{ct} U$, P_t est un opérateur linéaire sur \mathcal{H}_U tel que $\|P_t\|_U \leq e^{ct}$ parce qu'on a successivement $\forall f \in \mathcal{H}_U$:

$$|P_t f| \leq P_t |f| \leq \|f\|_U P_t U \leq e^{ct} \|f\|_U U \iff \|P_t f\|_U \leq e^{ct} \|f\|_U.$$

On va montrer que P_{2t_0} est compact. Pour ce faire, observons par (3) que pour tout $f \in \mathcal{H}_U$:

$$\|_{K_n^c} P_{t_0} f| \leq \|_{K_n^c} P_{t_0} |f| \leq \|_{K_n^c} \|f\|_U P_{t_0} U \leq \|_{K_n^c} \|f\|_U \cdot \kappa_n \cdot U$$

Ainsi :

$$\| \|_{K_n^c} P_{t_0} \|_U \leq \kappa_n \rightarrow 0$$

On va prouver que $\Lambda = \mathbb{1}_{K_n} P_{t_0} \mathbb{1}_{K_n}$ est compact et pour ça, il nous suffit de montrer que de toute famille $(f_i)_{i \in I}$ de \mathcal{H}_U telle que pour tout $i \in I$, $\|f_i\|_U \leq 1$, on peut extraire une sous-suite (f_p) telle que (Λf_p) converge dans \mathcal{H}_U . Remarquons que comme le semi-groupe est fortement fellerien par (1), et $\mathbb{1}_{K_n} f_i$ est bornée, la fonction $P_{t_0} \mathbb{1}_{K_n} f_i$ est continue, et $(\Lambda f_i)_{i \in I}$ est une famille de fonctions de $\mathcal{C}(K_n)$ l'espace des fonctions continues sur le compact K_n . De plus, elle est équibornée vu que par (3), $\forall x \in K_n$:

$$|\Lambda f_i(x)| \leq \|f_i\|_U P_{t_0} U(x) \leq \kappa_n \cdot U(x) + b_n \leq \kappa_n \cdot \sup_{y \in K_n} U(y) + b_n$$

Cette famille est également équicontinue vu que le semi-groupe est régularisant par (1), et donc par inégalité des accroissements finis et compacité, $\forall (x, y) \in K_n \times K_n$:

$$|\Lambda f_i(x) - \Lambda f_i(y)| \leq \int_{K_n} |p_{t_0}(x, z) - p_{t_0}(y, z)| \cdot |f_i(z)| dz \leq \|x - y\| \sup_{u, v \in K_n} \|\nabla_u p_{t_0}(u, v)\| \cdot \|f\|_U \int_{K_n} U(z) dz \leq C_n \|x - y\|$$

ce qui prouve une équilipschitzienité qui implique l'équicontinuité. Ainsi, par le théorème d'Arzela-Ascoli, la famille contient une sous suite (f_p) telle que (Λf_p) converge dans $\mathcal{C}(K_n)$ vers f , ce qui entraîne comme $U \geq 1$, que (Λf_p) converge dans \mathcal{H}_U vers $f \mathbb{1}_{K_n}$. Et pour conclure, il suffit d'écrire :

$$P_{2t_0} = \mathbb{1}_{K_n^c} P_{t_0} \circ P_{t_0} + \mathbb{1}_{K_n} P_{t_0} \circ \mathbb{1}_{K_n^c} P_{t_0} + \Lambda,$$

et remarquer que les deux premiers opérateurs convergent vers 0 dans \mathcal{H}_U , et le troisième est compact, donc P_{2t_0} aussi. Maintenant, on va montrer que P_{t_0} n'admet que 1 comme valeur propre de module 1 et c'est une valeur propre simple. Raisonnons par l'absurde et supposons que 1 n'est pas simple, alors il existe une fonction non constante $f \in \mathcal{H}_U$ telle $P_{t_0} f = f$. Quitte à diviser f par sa norme, on suppose que $\|f\|_U \leq 1$, alors $g = |f|$ est sous harmonique vu que $g = |P_{t_0} f| \leq P_{t_0} |f| = P_{t_0} g$. On en déduit que :

$$g \leq P_{nt_0} g \leq P_{nt_0} U \leq \kappa^n \cdot U + \frac{b}{1 - \kappa} \implies g \leq \frac{b}{1 - \kappa}$$

Donc, f est P_{t_0} -invariante et bornée, ainsi par le théorème 3.9 sur la récurrence des ouverts, f est constante. Maintenant, supposons que f soit un vecteur propre associé à la valeur propre λ de module 1. Si on note $\chi = E_\lambda(P_{t_0})$ l'espace propre associé à la valeur propre λ qui est de dimension finie par compacité de P_{t_0} , du fait que $\forall t$, $P_{t_0} P_t = P_t P_{t_0}$, χ est stable par P_t , et donc sur ce sous espace vectoriel de dimension finie, $P_t = e^{tL}$. On en déduit que si $\lambda = e^{it_0 \alpha}$ sur χ on a $P_t f = e^{it_0 \alpha} f$, et donc pour $t = 2\phi \alpha k > t_0$, $P_t f = f$. Ceci est contradictoire avec l'hypothèse d'irréductibilité appliquée ici en t au lieu de t_0 et le théorème 3.9. On va maintenant prouver la convergence exponentielle. Comme P_{t_0} est un opérateur compact, son spectre $\sigma(P_{t_0})$ est au plus dénombrable, avec 0 comme seul point d'accumulation possible et en outre, toute valeur non nulle du spectre est une valeur propre de multiplicité finie. On en déduit qu'on peut décomposer \mathcal{H}_U en somme directe $\mathcal{H}_U = H_1 \oplus H_2$ avec H_1 sous espace propre associé à la valeur propre 1 (une droite vectorielle décrite par les fonctions constantes) et H_2 sur lequel le rayon spectral de P_{t_0} est strictement inférieur à 1 ($\rho(P_{t_0}) < 1$ sur H_2). Comme $P_t = P_{t-t_0} P_{t_0}$ est compact pour tout $t \geq t_0$, il est clair (voir par exemple Davies theorem 3.19) que $\sigma(P_t) = \{0\} \cup e^{t\sigma(L)}$. On en déduit qu'il existe des constantes $\gamma > 0$ et $C > 0$ telles que :

$$\|P_t\|_{H_2} \leq C e^{-\gamma t}$$

Ce qui est exactement la convergence attendue vu que $f \mapsto \langle \pi, f \rangle$ est la projection sur H_1 , c'est-à-dire qu'on a $f - \langle \pi, f \rangle \in H_2$. \square

Chapitre 4

CV faible et Wasserstein

Pour plus de détails concernant ce chapitre, consulter principalement les références [20], [45], [95].

4.1 Caractérisations de la CV faible

On considère dans ce chapitre (E, d) un espace métrique, $\mathcal{B}(E)$ sa tribu borélienne, $\mathcal{C}_b(E)$ l'espace de Banach des fonctions scalaires continues bornées sur E et $\mathcal{P}(E)$ l'espace des mesures de probabilité sur $(E, \mathcal{B}(E))$. On rappelle la définition de la convergence faible¹.

Définition 4.1 (Convergence faible)

Soient $\mu \in \mathcal{P}(E)$ et $(\mu_n) \subset \mathcal{P}(E)$. On dit que (μ_n) converge faiblement vers μ si et seulement si $\forall \varphi \in \mathcal{C}_b(E)$:

$$\langle \mu_n, \varphi \rangle := \int_E f(x) d\mu_n(x) \longrightarrow \langle \mu, \varphi \rangle$$

Pour une suite de variables aléatoires (X_n) à valeurs dans (E, d) , on dit que cette suite converge faiblement vers la variable aléatoire X_∞ si et seulement si la suite des mesures images (lois) (\mathbb{P}_{X_n}) converge faiblement vers \mathbb{P}_{X_∞} .

Remarques et exemples :

Un exemple trivial de convergence faible est donné par n'importe quelle suite convergente $(x_n) \in E^{\mathbb{N}}$. En effet, dans ce cas, la suite de masses de Dirac (δ_{x_n}) converge faiblement par continuité séquentielle vers la masse de Dirac δ_x avec $x := \lim_n x_n$. Pour des détails sur les relations entre les convergences en loi, en probabilité, presque sûre et en norme L^p , voir partie annexes². Il est important de noter que la convergence faible n'implique pas la convergence forte³ qui est aussi une convergence en variations totales. Si l'espace métrique (E, d) est fini ($D := \mathbf{Card}(E) < +\infty$), alors $\mathcal{P}(E)$ s'identifie à une partie convexe et compacte de \mathbb{R}^D . Donc, dans ce cas, les modes de convergence coïncident. Dans le cas général, on utilise le théorème du Portemanteau qui donne plusieurs caractérisations de la convergence en loi.

Soient (X_n) une suite de variables aléatoires à valeurs dans (E, d) et X une variable aléatoire à valeurs dans (E, d) aussi. Le théorème ci-dessous dit du Portemanteau permet de montrer que

1. Convergence en loi ou au sens des distributions.
2. A3 chapitre 11.
3. Même définition que la convergence faible à une différence à savoir qu'on remplace l'espace des fonctions continues bornées par celui de celles qui sont mesurables bornées.

dans le cas $E = \mathbb{R}$ si $X_n \xrightarrow{\mathcal{L}} X$, alors la suite des fonctions de répartition (F_{X_n}) converge simplement sur $\{x \text{ tels que } F_X \text{ soit continue en } x\}$ vers F_X la fonction de répartition de X . Cette propriété est en fait une équivalence, et sert souvent, dans le cas des variables aléatoires réelles, de définition de la convergence en loi. En effet, d'un point de vue pédagogique, elle permet d'utiliser efficacement cette notion sans pour autant avoir eu à construire préalablement la théorie de la mesure. Quant à la preuve de cette propriété, en posant $F_n := F_{X_n}$ et $F := F_X$, par le point (5) du théorème du Portemanteau, on a :

Preuve

Par définition d'une fonction de répartition, la propriété $\lim_n F_n(x) = F(x)$ s'écrit sous la forme :

$$\lim_n \mathbb{P}(X_n \in A_x) = \mathbb{P}(X \in A_x),$$

avec $A_x =]-\infty, x]$ ($\implies \partial A_x = \{x\}$). On en déduit que :

$$\mathbb{P}(X \in \partial A_x) = \mathbb{P}(X = x) = F(x) - F(x_-),$$

qui est nul si et seulement si F est continue à gauche en x , c'est-à-dire si et seulement si F est continue en x (en effet, une fonction de répartition est partout continue à droite). \square

Théorème 4.1 (Du Portemanteau)

Les cinq assertions suivantes sont équivalentes :

1. X_n converge en loi vers X ;
2. pour toute fonction φ bornée et uniformément continue sur E , $\lim_n \mathbb{E}[\varphi(X_n)] = \mathbb{E}[\varphi(X)]$;
3. pour tout fermé F de E , $\limsup_n \mathbb{P}(X_n \in F) \leq \mathbb{P}(X \in F)$;
4. pour tout ouvert O de E , $\liminf_n \mathbb{P}(X_n \in O) \geq \mathbb{P}(X \in O)$;
5. pour tout borélien A de E tel que $\mathbb{P}(X \in \partial A) = 0$, $\lim_n \mathbb{P}(X_n \in A) = \mathbb{P}(X \in A)$.

Preuve

(1) \implies (2) : En effet, si $\lim_n \mathbb{E}[\varphi(X_n)] = \mathbb{E}[\varphi(X)]$ est vrai pour toute fonction φ bornée et continue sur E , alors cela est vrai en particulier pour toute fonction φ bornée et uniformément continue sur E .

(2) \implies (3) : Soit F un fermé de E . Pour tout $k \geq 1$, on note $F_k := \{x \in E \mid d(x, F) \leq \frac{1}{k}\}$. De plus, pour tout $x \in E$ et $k > 0$, on pose :

$$\varphi_k(x) := f(k d(x, F)) = \begin{cases} 1 - k d(x, F) & \text{si } x \in F_k \\ 0 & \text{si } x \notin F_k \end{cases}$$

où f est une fonction plateau définie $\forall y \in \mathbb{R}$ par :

$$f(y) := \begin{cases} 1 - y & \text{si } y \in [0, 1] \\ 0 & \text{si } y \geq 1 \\ 1 & \text{si } y \leq 0 \end{cases}$$

Ainsi, $\mathbb{1}_F \leq \varphi_k \leq \mathbb{1}_{F_k}$. Puisque $\mathbb{1}_F \leq \varphi_k$, on a $\limsup_n \mathbb{P}(X_n \in F) \leq \limsup_n \mathbb{E}[\varphi_k(X_n)]$. Comme les fonctions f et $x \mapsto d(x, F)$ sont 1-lipschitziennes, φ_k est uniformément continue, et d'après l'hypothèse (2) :

$$\limsup_n \mathbb{E}[\varphi_k(X_n)] = \lim_n \mathbb{E}[\varphi_k(X_n)] = \mathbb{E}[\varphi_k(X)]$$

De $\varphi_k \leq \mathbb{1}_{F_k}$, on déduit que :

$$\mathbb{E}[\varphi_k(X)] \leq \mathbb{P}(X \in F_k)$$

Par conséquent, pour tout $k \geq 1$:

$$\limsup_n \mathbb{P}(X_n \in F) \leq \mathbb{P}(X \in F_k)$$

Enfin, comme $(F_k)_k$ est décroissante et comme $\bigcap_{k \geq 1} F_k = \overline{F} = F$, on a :

$$\inf_k \mathbb{P}(X \in F_k) = \lim_k \mathbb{P}(X \in F_k) = \mathbb{P}(X \in F)$$

D'où le résultat.

(3) \iff (4) : Supposons le point (3) vrai, et considérons un ouvert \mathcal{O} de E . Alors \mathcal{O}^c est un fermé et on a, en vertu, entre autres, du point (3) :

$$\begin{aligned} \liminf_n \mathbb{P}(X_n \in \mathcal{O}) &= \liminf_n (1 - \mathbb{P}(X_n \in \mathcal{O}^c)) \\ &= 1 - \limsup_n \mathbb{P}(X_n \in \mathcal{O}^c) \\ &\geq 1 - \mathbb{P}(X \in \mathcal{O}^c) \\ &= \mathbb{P}(X \in \mathcal{O}). \end{aligned}$$

La démonstration de (4) \implies (3) est identique.

(3), (4) \implies (5) : Soit A un borélien de E tel que $\mathbb{P}(X \in \partial A) = 0$. Comme $\overset{\circ}{A} \subset A \subset \overline{A} = \overset{\circ}{A} \cup \partial A$, on en déduit que :

$$\mathbb{P}(X \in \overline{A}) = \mathbb{P}(X \in A) = \mathbb{P}(X \in \overset{\circ}{A})$$

Or, d'après le point (3) :

$$\limsup_n \mathbb{P}(X_n \in A) \leq \limsup_n \mathbb{P}(X_n \in \overline{A}) \leq \mathbb{P}(X \in \overline{A})$$

et d'après le point (4) :

$$\liminf_n \mathbb{P}(X_n \in A) \geq \liminf_n \mathbb{P}(X_n \in \overset{\circ}{A}) \geq \mathbb{P}(X \in \overset{\circ}{A})$$

Finalement :

$$\lim_n \mathbb{P}(X_n \in A) = \mathbb{P}(X \in A)$$

Ce qu'il fallait démontrer.

(5) \implies (1) : Commençons par traiter le cas où ψ est une fonction continue et bornée, telle que $0 < \psi < 1$. Alors, comme $0 < \psi < 1$, on a :

$$\mathbb{E}[\psi(X)] = \int_0^1 \mathbb{P}[\psi(X) > x] dx$$

et de même pour les X_n .

De plus, comme ψ est continue, on a aussi :

$$\partial\psi^{-1}(]x, +\infty[) \subset \psi^{-1}(\{x\})$$

Donc :

$$\mathbb{P}(X \in \partial\psi^{-1}(]x, +\infty[)) = 0$$

sauf pour un ensemble au plus dénombrable D_0 de valeurs de x et de même, pour tout $n \geq 1$,

$$\mathbb{P}(X_n \in \partial\psi^{-1}(]x, +\infty[)) = 0$$

sauf pour un ensemble au plus dénombrable D_n de valeurs de x . L'ensemble $\cup_{n \geq 0} D_n$ est au plus dénombrable donc Lebesgue-négligeable. Par conséquent, en vertu du point (5), pour presque tout :

$$x \in [0, 1], \mathbb{P}[\psi(X_n) > x] \xrightarrow{n \rightarrow \infty} \mathbb{P}[\psi(X) > x]$$

On conclut par convergence dominée :

$$\mathbb{E}[\psi(X_n)] = \int_0^1 \mathbb{P}[\psi(X_n) > x] dx \xrightarrow{n \rightarrow \infty} \int_0^1 \mathbb{P}[\psi(X) > x] dx = \mathbb{E}[\psi(X)]$$

Enfin, dans le cas général, pour une fonction continue bornée φ , telle que $a < \varphi < b$, on se ramène au cas précédent en posant :

$$\psi := \frac{\varphi - a}{b - a}$$

de sorte que $0 < \psi < 1$. Ce qui achève la preuve du théorème du Portemanteau. \square

On a aussi le théorème de l'application continue très important dont la preuve est très facile.

Théorème 4.2 (de l'application continue)

Soient E_1 et E_2 deux espaces métriques, $\Psi : E_1 \rightarrow E_2$ continue et (X_n) une suite de variables aléatoires à valeurs dans E_1 . Si $X_n \xrightarrow{\mathcal{L}} X$, alors $\Psi(X_n) \xrightarrow{\mathcal{L}} \Psi(X)$

Preuve

Il suffit juste de remarquer que $\forall \varphi \in \mathcal{C}_b(E_2)$, $\varphi \circ \Psi \in \mathcal{C}_b(E_1)$. Et vu que par définition de la convergence faible, $\forall \phi \in \mathcal{C}_b(E_1)$:

$$\lim_n \mathbb{E}[\phi(X_n)] = \mathbb{E}[\phi(X)],$$

en posant $\phi := \varphi \circ \Psi$, on obtient $\forall \varphi \in \mathcal{C}_b(E_2)$:

$$\lim_n \mathbb{E}[\varphi(\Psi(X_n))] = \mathbb{E}[\varphi(\Psi(X))]$$

ce qui est la définition de la convergence faible pour $(\Psi(X_n))$. \square

4.2 Tension et pré-compacité

On rappelle aussi la définition d'une famille tendue de $\mathcal{P}(E)$ qui intervient dans la caractérisation des parties pré-compactes de $\mathcal{P}(E)$ pour la convergence faible. On note $\mathcal{C}_0(E)$ le sous-espace de $\mathcal{C}_b(E)$ des fonctions nulles à l'infini (c'est-à-dire les fonctions f telles que pour tout $\varepsilon > 0$, il existe un compact $K \subset E$ tel que $\forall x \in E \setminus K \quad |f(x)| < \varepsilon$). Intuitivement, une famille de mesures est tendue si elle ne « s'échappe pas vers l'infini ». La tension est souvent un critère nécessaire pour démontrer la convergence faible d'une suite de mesures de probabilité, plus particulièrement quand l'espace des mesures est de dimension infinie.

Définition 4.2 (Famille de mesures tendue)

Soit $\mathcal{S} \subset \mathcal{P}(E)$ une famille de mesures de probabilité. On dit que \mathcal{S} est tendue si $\forall \varepsilon > 0 \exists K \subset E$ compact tels que :

$$\sup_{\mu \in \mathcal{S}} \mu(K^c) \leq \varepsilon \iff \forall \mu \in \mathcal{S} \quad \mu(K) > 1 - \varepsilon$$

Dans le cas $\mathcal{S} = \{\mu_0\}$, on parle de mesure tendue ou intérieurement régulière.

D'après le théorème de représentation de Riesz, le dual de $\mathcal{C}_0(E)$ est l'espace $\mathcal{R}_b(E)$ des mesures de Radon bornées. Il contient l'ensemble $\mathcal{R}_b^+(E)$ des mesures positives bornées et le sous-ensemble $\mathcal{P}(E)$ des mesures de probabilité.

Remarques et exemples

Si E est un espace compact, alors toute famille de mesures (éventuellement complexes) sur E est tendue. Si E est un espace polonais, alors toute mesure de probabilité sur E est tendue. De plus, par le théorème de Prokhorov, une famille de mesures de probabilité est tendue si et seulement si elle est relativement compacte pour la topologie de la convergence faible des mesures. Considérons la droite réelle \mathbb{R} munie de la topologie borélienne. Soit δ_x la mesure de Dirac ayant une unique masse au point x . La famille $M_1 := \{\delta_n | n \in \mathbb{N}\}$ n'est pas tendue, puisque les sous-ensembles compacts de \mathbb{R} sont précisément les ensembles fermés bornés, et ces ensembles ont une masse nulle pour les mesures δ_n pour n suffisamment grand.

Cependant, la famille $M_2 := \{\delta_{1/n} | n \in \mathbb{N}\}$ est tendue, en effet, l'intervalle $[0, 1]$ est considéré comme K_η pour tout $\eta > 0$. En général, une famille de mesures de Dirac sur \mathbb{R}^d est tendue si et seulement si la famille de leur support est bornée.

Il existe une autre notion de tension dite exponentielle. La tension exponentielle est une généralisation de la tension des mesures qui a des applications pour le principe de grandes déviations. Une famille de lois de probabilité $(\mu_\delta, \delta > 0)$ sur un espace topologique séparé X est exponentiellement tendue si, pour tout $\eta > 0$, il existe un sous-ensemble compact K_η de X tel que :

$$\limsup_{\delta \downarrow 0} \delta \log \mu_\delta(X \setminus K_\eta) < -\eta.$$

On va maintenant donner l'énoncé du théorème de Prokhorov qui relie le concept de tension à la compacité relative dans l'espace des mesures de probabilité, ou plus généralement des mesures finies. Ce théorème porte le nom du mathématicien Iouri Prokhorov.

Théorème 4.3 (De Prokhorov)

Soit $\mathcal{S} \subset \mathcal{P}(E)$. Si \mathcal{S} est tendue, alors elle est relativement compacte dans $\mathcal{P}(E)$ pour la topologie de la convergence faible, c'est-à-dire séquentiellement, pour toute suite $(\mu_n) \subset \mathcal{S} \exists \sigma : \mathbb{N} \rightarrow \mathbb{N}$ une extractrice telle que :

$$\mu_{\sigma(n)} \xrightarrow{\text{faiblement}} \mu \in \mathcal{P}(E)$$

ou de façon équivalente, \mathcal{S} est d'adhérence compacte dans $\mathcal{P}(E)$ muni de sa topologie de convergence faible. Réciproquement, si \mathcal{S} est pré-compacte et si E est polonais, alors elle est tendue.

On énonce aussi un théorème de représentation dit de Skorokhod qui nous est utile pour certaines preuves. Le théorème de représentation de Skorokhod montre qu'une suite de variables aléatoires convergeant en loi peut toujours, en un certain sens, être représentée par une suite de variables aléatoires convergeant presque sûrement. Ce théorème porte le nom du mathématicien ukrainien A.V. Skorokhod.

Théorème 4.4 (De Skorokhod)

On suppose que E est séparable. Soit $(\mu_n) \subset \mathcal{P}(E)$ qui converge faiblement vers $\mu \in \mathcal{P}(E)$. Alors il existe un espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$ et des variables aléatoires définies sur cet espace (X_n) et X à valeurs dans E telles que $X_n \sim \mu_n$, $X \sim \mu$ et $X_n \xrightarrow{\mathbb{P}-ps} X$.

Comme exemple d'application des théorèmes 4.2 de l'application continue et 4.4 de représentation, on a le théorème important suivant qui donne la convergence pour le crochet de dualité sur une classe de fonctions non bornées.

Théorème 4.5

On suppose que (μ_n) converge faiblement vers μ dans $\mathcal{P}(E)$ avec E séparable. Si $\varphi : E \rightarrow \mathbb{R}$ est continue et uniformément intégrable en ce sens que :

$$\lim_{r \rightarrow +\infty} \sup_n \int_{\{|\varphi| \geq r\}} |\varphi| d\mu_n = 0$$

alors :

$$\langle \mu_n, \varphi \rangle \xrightarrow{n \rightarrow +\infty} \langle \mu, \varphi \rangle$$

Preuve

Par le théorème de l'application continue, on sait que $(\mu_n \circ \varphi^{-1})$ converge faiblement vers $\mu \circ \varphi^{-1}$ dans $\mathcal{P}(\mathbb{R})$. Et par le théorème de représentation, il existe $(\Omega, \mathcal{F}, \mathbb{P})$ un espace probabilisé, (X_n) et X des variables aléatoires définies sur cet espace à valeurs dans \mathbb{R} telles que $X_n \sim \mu_n \circ \varphi^{-1}$, $X \sim \mu \circ \varphi^{-1}$ et (X_n) converge \mathbb{P} -presque sûrement vers X . Par la formule de transfert et changement de variables abstrait, notre hypothèse d'uniforme intégrabilité devient :

$$\lim_{r \rightarrow +\infty} \sup_n \mathbb{E}[|X_n| \mathbb{1}_{\{|X_n| \geq r\}}] = 0$$

Ce qui est par définition, l'uniforme intégrabilité de la suite (X_n) et vu qu'on a de plus une convergence presque sûre de cette suite vers X , on a par convergence dominée de Lebesgue :

$$\lim_{n \rightarrow +\infty} \langle \mu_n, \varphi \rangle = \lim_{n \rightarrow +\infty} \mathbb{E}[X_n] = \mathbb{E}[X] = \langle \mu, \varphi \rangle$$

Ce qui prouve qu'on a bien le résultat attendu. \square

4.3 Wasserstein et couplage

Une façon d'obtenir des bornes quantitatives pour le comportement en temps long d'un processus de diffusion est d'utiliser un argument de couplage pour les distances de Wasserstein ou la norme en variation totale dont on rappelle l'expression :

$$\|\mu - \nu\|_{TV} := \inf \left\{ \mathbb{P}(X = Y), \quad X \sim \mu, \quad Y \sim \nu \right\} .$$

En toute généralité, la distance en variation totale est majorée par la racine carrée de l'entropie relative⁴.

Définition 4.3 (T_2 -inégalité de transport de Talagrand)

On dit que $\mu \in \mathcal{M}_1^2(\mathbb{R}^d)$ vérifie une inégalité de transport T_2 de constante α si, et seulement si $\mathcal{W}_2(\mu, \cdot) \leq \sqrt{\alpha \mathbf{H}[\cdot | \mu]}$.

Quelques résultats : Soit μ une mesure de probabilité sur \mathbb{R}^d qui vérifie une inégalité de Sobolev logarithmique de constante ρ . Alors elle vérifie une inégalité de transport T_2 de constante $\frac{1}{\rho}$. Et la réciproque est vraie si μ est une mesure de Gibbs d'hamiltonien H uniformément elliptique (convexe, coercif...). Chercher des estimations en distance de Wasserstein est donc essentiellement payant lorsque l'inégalité de Sobolev logarithmique est en défaut. Une autre propriété que nous chercherons à établir dans différents contextes est le phénomène de concentration gaussienne⁵. De plus, une T_1 inégalité de Talagrand pour une mesure de Gibbs est équivalente à l'inégalité de Poincaré...

4. C'est l'inégalité de Csiszár-Kullback-Pinsker.

5. Consulter [20] pour plus de détails.

Deuxième partie

Sur les inégalités de Poincaré et log-Sobolev

Chapitre 5

Inégalités de Poincaré et log-Sobolev pour μ^N

Dans ce chapitre de cette seconde partie, on établit des inégalités fonctionnelles pour la mesure invariante μ^N de notre système à champ moyen (\star). Pour plus de détails, consulter principalement la référence [5] en bibliographie.

5.1 Poincaré

Le théorème suivant est issu de [5].

Théorème 5.1 (Inégalité de Poincaré pour μ^N)

1. Sous les hypothèses faites sur les potentiels, on a $\lambda_{1,m} c_{\text{Lip},m} \geq 1$.
2. On suppose que $\exists h > -\lambda_{1,m}$ telle que $\forall (x_1, \dots, x_N) \in (\mathbb{R}^d)^N$

$$\frac{1}{N-1} (\mathbb{1}_{i \neq j} \nabla_{x,y}^2 W(x_i, x_j))_{1 \leq i, j \leq N} \geq h \mathbf{Id}_{dN}$$

au sens des matrices symétriques définies positives.

Alors, μ^N satisfait une inégalité de Poincaré dont une constante est $h + \lambda_{1,m} > 0$ ou encore de façon équivalente l'écart spectral $\lambda_1(\mu^N)$ de \mathcal{L}^N sur $L^2(\mu^N)$ (défini comme étant la borne inférieure des éléments $\lambda > 0$ du spectre de ce générateur infinitésimal) vérifie :

$$\lambda_1(\mu^N) \geq h + \lambda_{1,m} \geq \frac{1}{c_{\text{Lip},m}} + h$$

Résultats préliminaires : Avant de pouvoir démontrer ce théorème, on va prouver des résultats préliminaires. On rappelle que la distribution conditionnelle $\mu_i(dx_i) := \mu_i(dx_i | x_j, j \neq i)$ de x_i sachant $\hat{x}^i := (x_j)_{j \neq i}$ de notre mesure de champ moyen μ^N est donnée par :

$$\mu_i(dx_i) = \frac{1}{Z_i} \exp \left\{ -V(x_i) - \frac{1}{N-1} \sum_{j \neq i} W(x_i, x_j) \right\} dx_i$$

avec Z_i le facteur de normalisation associé. On pose :

$$H_i(x_i) := V(x_i) + \frac{1}{N-1} \sum_{j \neq i} W(x_i, x_j)$$

l'hamiltonien associé à cette mesure de Gibbs. Le générateur $\mathcal{L}_i^N = \Delta_i - \nabla_i H_i \cdot \nabla_i$, avec \hat{x}^i fixé, est μ_i -symétrique. Par définition de $b_0(r)$, on a $\forall x, y \in (\mathbb{R}^d)^N : (x^{\hat{i}, y_i} \in (\mathbb{R}^d)^N \text{ et } (x^{\hat{i}, y_i})_j = x_j \text{ si } j \neq i \text{ et } y_i \text{ sinon})$

$$\left\langle \frac{x_i - y_i}{\|x_i - y_i\|}, \nabla_i H(x^{\hat{i}}, y_i) - \nabla_i H(x_i) \right\rangle = \frac{1}{N-1} \sum_{j \neq i} \left\langle \frac{x_i - y_i}{\|x_i - y_i\|}, \gamma_j \right\rangle \leq b_0(\|x_i - y_i\|)$$

avec $\gamma_j := -(\nabla V(x_i) - \nabla V(y_i) + \nabla_x W(x_i, x_j) - \nabla_x W(y_i, x_j))$. On en déduit la proposition importante :

Proposition 1 (Contrôle de trous spectraux) *Sous l'hypothèse (H3), l'opérateur de Poisson $(\mathcal{L}_i)^{-1}$ sur l'espace de Banach $\mathcal{C}_{Lip,0}(\mathbb{R}^d) := \mathcal{C}_{Lip}(\mathbb{R}^d) \cap \{f, \mu_i(f) = 0\}$ équipé de la norme $\|\cdot\|_{Lip}$ est borné et sa norme vérifie :*

$$\|(-\mathcal{L}_i)^{-1}\|_{Lip} \leq c_{Lip,m}$$

En particulier, $\lambda_1(\mu_i)$ l'écart spectral de \mathcal{L}_i dans $L^2(\mu_i)$ vérifie l'inégalité $\lambda_1(\mu_i) \geq \frac{1}{c_{Lip,m}}$.

Quant à la preuve de cette proposition, voir bibliographie (référence [6]). Cette proposition nous donne des contrôles des trous spectraux des générateurs associés aux mesures de Gibbs conditionnelles...

Proposition 2 (Ledoux et minoration de l'écart spectral) *On rappelle que $Z < +\infty$ et $\forall i$ et $x^{\hat{i}}$, $Z_i < +\infty$ les constantes de normalisation associées aux mesures de Gibbs μ et ses distributions conditionnelles μ_i d'hamiltoniens respectifs :*

$$H(x_1, \dots, x_N) = U(x_1, \dots, x_N) + \sum_{i=1}^N V(x_i) \text{ et } H_i \text{ définis dans la première partie du mémoire.}$$

Si :

1. les μ_i satisfont une inégalité de Poincaré uniforme, c'est-à-dire :

$$\lambda_{1,m} := \inf_{1 \leq i \leq N, x^{\hat{i}} \in (\mathbb{R}^d)^{N-1}} \lambda_1(\mu_i) > 0$$

2. $\exists h \in \mathbb{R}$ telle que :(au sens des matrices symétriques définies positives)

$$(\|_{i \neq j} \nabla_{x_i, x_j}^2 U)_{1 \leq i, j \leq N} \geq h \mathbf{Id}_{dN}$$

alors, $\lambda_1(\mu) \geq h + \lambda_{1,m}$.

Cette proposition est essentiellement due à Ledoux. En effet, dans le cas $d = 1$, si $\mathbf{Hess}(U) \geq \underline{\lambda} \mathbf{Id}_N$ et $\forall i$ et $x^{\hat{i}}$ $\partial_{ii}^2 U \leq \bar{\lambda}$, on a $\forall (v_1, \dots, v_N) \in \mathbb{R}^N$

$$\sum_{i \neq j} v_i \partial_{ij}^2 U v_j = \langle \mathbf{Hess}(U) v, v \rangle - \sum_i \partial_{ii}^2 U v_i^2 \geq (\underline{\lambda} - \bar{\lambda}) \|v\|^2$$

c'est-à-dire que la deuxième condition de la proposition est vérifiée pour $h = \underline{\lambda} - \bar{\lambda}$ et $\lambda_1(\mu) \geq \lambda_{1,m} + \underline{\lambda} - \bar{\lambda}$. Ce qui est le résultat original de Ledoux.

Preuve

On suppose sans perte de généralité que $\lambda_{1,m} + h > 0$. Soit $\mathcal{L} = \Delta - \nabla H \cdot \nabla$ le générateur symétrique associé à la mesure de Gibbs μ . Par description duale de l'inégalité de Poincaré, la conclusion est équivalente à :

$$\int (\mathcal{L}f)^2 d\mu \geq (\lambda_{1,m} + h) \int \|\nabla f\|^2 d\mu$$

En utilisant la formule de Bakry-Emery, on a $\int \Gamma_2(f) d\mu = \int (\mathcal{L}f)^2 d\mu$ et $\Gamma_2(f) = \|\nabla^2 f\|_{HS}^2 + \langle \nabla^2 H \nabla f, \nabla f \rangle$ avec pour toute matrice A , $\|A\|_{HS} := \sqrt{\text{Tr}(AA^*)} = \sqrt{\sum |a_{ij}|^2}$ est la norme de Hilbert-Schmidt de la matrice A (ou encore norme de Frobenius ou norme hermitienne standard de A). On en déduit du fait :

$$\int (\mathcal{L}f)^2 d\mu = \int \left(\|\nabla^2 f\|_{HS}^2 + \langle \nabla^2 H \nabla f, \nabla f \rangle \right) d\mu$$

qu'on a :

$$\int (\mathcal{L}f)^2 d\mu = \int \left(\|\nabla^2 f\|_{HS}^2 + \langle \text{Hess}(U) \nabla f, \nabla f \rangle + \sum_{1 \leq i \leq N} \langle \text{Hess}(V)(x_i) \nabla_{x_i} f, \nabla_{x_i} f \rangle \right) d\mu$$

Il en découle que :

$$\int (\mathcal{L}f)^2 d\mu \geq \sum_{1 \leq i \leq N} \int \int_{\mathbb{R}^d} \left(\|\nabla_{x_i}^2 f\|_{HS}^2 + \langle \text{Hess}(V)(x_i) + \nabla_{x_i, x_i}^2 U \nabla_{x_i} f, \nabla_{x_i} f \rangle \right) d\mu_i d\mu + \int \sum_{i \neq j} \langle \nabla_{x_i, x_j}^2 U \nabla_{x_i} f, \nabla_{x_j} f \rangle d\mu$$

Par caractérisation de l'inégalité de Poincaré pour les mesures conditionnelles, on a $\forall i$ et $x^{\hat{i}}$:

$$\int_{\mathbb{R}^d} \left(\|\nabla_{x_i}^2 f\|_{HS}^2 + \langle \text{Hess}(V)(x_i) + \nabla_{x_i, x_i}^2 U \nabla_{x_i} f, \nabla_{x_i} f \rangle \right) d\mu_i \geq \lambda_{1,m} \int_{\mathbb{R}^d} \|\nabla_{x_i} f\|^2 d\mu_i$$

On déduit de la seconde condition de la proposition que :

$$\int \sum_{i \neq j} \langle \nabla_{x_i, x_j}^2 U \nabla_{x_i} f, \nabla_{x_j} f \rangle d\mu \geq h \int \|\nabla f\|^2 d\mu$$

Et cette inégalité, combinée avec les précédentes, nous donne l'inégalité voulue. \square

Preuve du théorème sur l'inégalité de Poincaré :

Preuve

En appliquant la proposition précédente à $\mu = \mu^N$, le potentiel U est donné par :

$$U(x) = \frac{1}{N-1} \sum_{1 \leq i < j \leq N} W(x_i, x_j) = \frac{1}{2} \sum_{i=1}^N U_i(x), \quad U_i(x) = \frac{1}{N-1} \sum_{j \neq i} W(x_i, x_j)$$

De plus, $\forall i \neq j$:

$$\nabla_{x_i, x_j}^2 U = \frac{1}{N-1} \nabla_{x, y}^2 W(x_i, x_j)$$

L'hypothèse du théorème implique la deuxième condition d'existence de h de la proposition juste ci-dessus. En appliquant la proposition (1) sur l'opérateur de Poisson à $\mu_i(dx_i) = \frac{1}{Z_i} e^{-V(x_i) - U_i(x)} dx_i$ d'hamiltonien $H_i(x_i) = V(x_i) + U_i(x)$, on obtient $\lambda_1(\mu_i) \geq \frac{1}{c_{Lip, m}}$ (on a ainsi obtenu la première condition d'uniforme inégalité de Poincaré de la proposition ci-dessus). Donc, on peut appliquer la proposition ci-dessus pour en déduire qu'on a bel et bien $\lambda_1(\mu^N) \geq h + \lambda_{1,m} \geq \frac{1}{c_{Lip, m}} + h$. \square

5.2 Log-Sobolev

Le théorème suivant est issu de [5].

Théorème 5.2 (Inégalité log-Sobolev pour μ^N)

On suppose que :

1. $\exists \rho_{LS,m} > 0$ telle que : (log-sobolev uniforme) $\forall f \in \mathcal{C}_b^1(\mathbb{R}^d)$

$$\rho_{LS,m} \cdot \mathbf{Ent}_{\mu_i}[f^2] \leq 2 \cdot \mathbb{E}_{\mu_i}[\|\nabla f\|^2]$$

$\forall i$ et $x^{\hat{i}}$.

2. on a une condition de translation de Zegarliniski à savoir :

$$\gamma_0 := c_{Lip,m} \sup_{x,y \in \mathbb{R}^d, \|z\|=1} \|\nabla_{x,y}^2 W(x,y)z\| < 1$$

Dans ce cas, μ^N satisfait $\forall f \in \mathcal{C}_b^1((\mathbb{R}^d)^N)$:

$$\rho_{LS,m}(1 - \gamma_0)^2 \cdot \mathbf{Ent}_{\mu^N}[f^2] \leq 2 \cdot \mathbb{E}_{\mu^N}[\|\nabla f\|^2]$$

c'est-à-dire $\rho_{LS}(\mu^N) \geq \rho_{LS,m}(1 - \gamma_0)^2$.

Ce théorème est une conséquence directe du théorème suivant dit de Zegarliniski pour des mesures de Gibbs dont la preuve détaillée est à trouver en référence [7] dans la bibliographie.

Théorème 5.3 (De Zegarliniski)

On suppose que :

1. les distributions conditionnelles (μ_i) associées à μ satisfont une inégalité log-Sobolev uniforme, c'est-à-dire :

$$\rho_{LS,m} := \inf_{1 \leq i \leq N, x^{\hat{i}} \in (\mathbb{R}^d)^{N-1}} \rho_{LS}(\mu_i) > 0$$

2. la condition de Zegarliniski suivante est vérifiée :

$$\gamma := \sup_{1 \leq i \leq N} \max \left\{ \sum_{j=1}^N c_{ji}^Z, \sum_{j=1}^N c_{ij}^Z \right\} < 1$$

Alors, la mesure de Gibbs μ d'hamiltonien H satisfait une inégalité log-Sobolev à savoir :

$$\rho_{LS,m}(1 - \gamma)^2 \cdot \mathbf{Ent}_{\mu}[f^2] \leq 2 \cdot \mathbb{E}_{\mu}[\|f\|^2]$$

pour toute fonction $f \in \mathcal{C}_b^1((\mathbb{R}^d)^N)$, c'est-à-dire :

$$\rho_{LS}(\mu) \geq \rho_{LS,m}(1 - \gamma)^2$$

Chapitre 6

Convergence exponentielle pour le flot solution de l'équation de McKean-Vlasov

Dans ce chapitre, on établit via les résultats précédents une convergence exponentielle en temps long du flot solution de l'équation de McKean-Vlasov non linéaire vers son équilibre en métrique L^2 -Wasserstein. Donc, on admet les hypothèses émises dans le chapitre précédent. Pour plus de détails, consulter principalement la référence [5] en bibliographie.

Théorème 6.1 (De convergence vers l'équilibre)

Sous les hypothèses du théorème de Zegarlinski :

1. $\exists! \nu_\infty \in \mathcal{M}_1(\mathbb{R}^d)$ telle que :

$$\nu_\infty = \mathbf{argmin} \left\{ \mathbf{H}_W[\nu], \nu \in \mathcal{M}_1(\mathbb{R}^d) \right\} ,$$

avec \mathbf{H}_W l'entropie de champ moyen définie dans le chapitre 1 section 1.2.

2. $\rho_{LS} := \limsup_{N \rightarrow +\infty} \rho_{LS}(\mu^N) \geq \rho_{LS,m}(1 - \gamma_0)^2$ vérifie :

$$\rho_{LS} \mathbf{H}_W[\nu] \leq 2 \cdot \mathbf{I}_W[\nu], \quad \rho_{LS} \mathcal{W}_2^2(\nu, \nu_\infty) \leq 2 \cdot \mathbf{H}_W[\nu] .$$

pour tout $\nu \in \mathcal{M}_1(\mathbb{R}^d)$. On dit qu'on a une inégalité log-Sobolev non linéaire pour la première et une inégalité de transport de Talagrand pour la seconde.

3. Pour le flot $(\nu_t)_{t \geq 0}$ solution de l'équation de McKean-Vlasov non linéaire de distribution initiale ν_0 admettant un moment d'ordre 2, on a $\forall t \geq 0$:

$$\mathbf{H}_W[\nu_t] \leq \mathbf{H}_W[\nu_0] \cdot e^{-\frac{t}{2} \rho_{LS}}, \quad \mathcal{W}_2^2(\nu_t, \nu_\infty) \leq \frac{2}{\rho_{LS}} \mathbf{H}_W[\nu_0] \cdot e^{-\frac{t}{2} \rho_{LS}} .$$

Résultats préliminaires : Pour prouver ce théorème, on a besoin d'établir quelques résultats ¹.

Proposition 3 (Convergence en entropie relative) *Pour toute mesure de probabilité ν sur \mathbb{R}^d telle que $\mathbf{H}[\nu|\alpha] < +\infty$, on a :*

$$\frac{1}{N} \mathbf{H}[\nu^{\otimes N} | \mu^N] \xrightarrow{N \rightarrow +\infty} \mathbf{H}_W[\nu] .$$

1. Pour plus de détails, consulter la référence [11] en bibliographie.

Preuve

On rappelle que $\alpha(dx) := \frac{e^{-V(x)}}{C} dx$. Sous l'hypothèse **(H1)**, on sait que $\exists \lambda_0 > 0$ telle que :

$$\int_{\mathbb{R}^d} e^{\lambda_0 \|x\|^2} \alpha(dx) < +\infty$$

En posant :

$$\tilde{Z}_N := \int_{(\mathbb{R}^d)^N} e^{-\frac{1}{2(N-1)} \sum_{i \neq j} W(x_i, x_j)} \alpha^{\otimes N}(dx_1, \dots, dx_N) \quad ,$$

on obtient : (par Fubini-Tonelli)

$$\mu^N(dx_1, \dots, dx_N) = \frac{1}{\tilde{Z}_N} e^{-\frac{1}{2(N-1)} \sum_{i \neq j} W(x_i, x_j)} \alpha^{\otimes N}(dx_1, \dots, dx_N)$$

Soit $\nu \in \mathcal{M}_1(\mathbb{R}^d)$ telle que $\mathbf{H}[\nu|\alpha] < +\infty$. Vu que $\mathbf{H}[\nu^{\otimes 2}|\alpha^{\otimes 2}] = 2\mathbf{H}[\nu|\alpha]$, $x \mapsto e^{\lambda_0 \|x\|^2} \in L^1(\alpha)$ et $|W(x, y)| \leq \beta(1 + \|x\|^2 + \|y\|^2)$ par bornitude de sa hessienne $\nabla^2 W$ (hypothèse **(H2)**), d'après la formule variationnelle de Donsker-Varadhan de l'entropie, on a $W \in L^1(\nu^{\otimes 2})$. On a successivement : (par un calcul direct et application du théorème de Fubini-Tonelli)

$$\frac{1}{N} \mathbf{H}[\nu^{\otimes N}|\mu^N] = \frac{1}{N} \mathbf{Ent}_{\mu^N} \left[\frac{d\nu^{\otimes N}}{d\mu^N} \right] = \frac{1}{N} \int_{(\mathbb{R}^d)^N} \log \left(\frac{d\nu^{\otimes N}}{d\mu^N} \right) d\nu^{\otimes N}$$

On en déduit que :

$$\frac{1}{N} \mathbf{H}[\nu^{\otimes N}|\mu^N] = \frac{1}{N} \int \sum_{i=1}^N \log \left(\frac{d\nu}{d\alpha}(x_i) \right) d\nu^{\otimes N} + \frac{1}{2N(N-1)} \int \sum_{i \neq j} W(x_i, x_j) d\nu^{\otimes N} + \frac{1}{N} \log(\tilde{Z}_N)$$

Or $\lim_{N \rightarrow +\infty} \frac{1}{N} \log(\tilde{Z}_N) = - \inf_{\eta \in \mathcal{M}_1(\mathbb{R}^d)} \mathbf{E}_f[\eta]$ (voir référence [11]), $\frac{1}{N} \int \sum_{i=1}^N \log \left(\frac{d\nu}{d\alpha}(x_i) \right) d\nu^{\otimes N} = \mathbf{H}[\nu|\alpha]$ et enfin, on a aussi :

$$\frac{1}{2N(N-1)} \int \sum_{i \neq j} W(x_i, x_j) d\nu^{\otimes N} = \frac{1}{2} \int \int W(x, y) \nu(dx) \nu(dy)$$

Ainsi :

$$\frac{1}{N} \mathbf{H}[\nu^{\otimes N}|\mu^N] \xrightarrow{N \rightarrow +\infty} \mathbf{H}[\nu|\alpha] + \frac{1}{2} \int \int W(x, y) \nu(dx) \nu(dy) - \inf_{\eta \in \mathcal{M}_1(\mathbb{R}^d)} \mathbf{E}_f[\eta] = \mathbf{H}_W[\nu]$$

Ce qu'il fallait prouver. \square

Proposition 4 (Entropie relative et produit tensoriel) Soient $\prod_{i=1}^N \alpha_i$ et Q respectivement une mesure de probabilité produit et une mesure de probabilité définies sur $E_1 \times \dots \times E_N$ un produit d'espaces polonais. En notant Q_i la distribution marginale de x_i sous Q , on a :

$$\mathbf{H}[Q|\prod_{i=1}^N \alpha_i] \geq \sum_{i=1}^N \mathbf{H}[Q_i|\alpha_i] \quad .$$

Preuve

Soit $Q^i(\cdot|x_{[1, i-1]})$ la distribution conditionnelle de x_i sachant $x_{[1, i-1]} := (x_1, \dots, x_{i-1})$. (ne sachant rien si $i = 1$)

On a :

$$\mathbf{H}[Q|\prod_{i=1}^N \alpha_i] = \mathbb{E}_Q \left[\log \left(\frac{dQ}{d\prod_{i=1}^N \alpha_i} \right) \right] = \mathbb{E}_Q \left[\sum_{i=1}^N \log \left(\frac{Q^i(dx_i|x_{[1, i-1]})}{\alpha_i(dx_i)} \right) \right] = \mathbb{E}_Q \left[\sum_{i=1}^N \mathbf{H}[Q^i(\cdot|x_{[1, i-1]})|\alpha_i] \right]$$

Vu que $\mathbb{E}_Q[Q^i(\cdot|x_{[1,i-1]})] = Q_i(\cdot)$, on obtient par convexité de l'entropie relative (inégalité de Jensen) :

$$\mathbb{E}_Q \left[\mathbf{H}[Q^i(\cdot|x_{[1,i-1]})|\alpha_i] \right] \geq \mathbf{H}[Q_i|\alpha_i]$$

Ce qui montre qu'on a bien le resultat de la proposition. \square

Proposition 5 (Entropie relative et mesure de Boltzmann) Soit μ une mesure de probabilité sur un espace polonais E et $U : E \rightarrow (-\infty, +\infty]$ un potentiel mesurable tel que :

$$\int e^{-pU} d\mu < +\infty$$

pour un certain $p > 1$. En considérant la mesure de probabilité de Boltzmann $\mu_U := \frac{e^{-U}}{C} d\mu$, si pour une mesure ν , $\mathbf{H}[\nu|\mu_U] < +\infty$, on a successivement :

1. $\mathbf{H}[\nu|\mu] < +\infty$ et $U \in L^1(\nu)$.
2. $\mathbf{H}[\nu|\mu_U] = \mathbf{H}[\nu|\mu] + \int U d\nu + \log \int e^{-U} d\mu$.

Preuve

Pour f une fonction mesurable sur E , on définit :

$$\Lambda_\mu(f) := \log(\mathbb{E}_\mu[e^f]) = \log \int e^f d\mu \in (-\infty, +\infty]$$

la transformation log-Laplace sous μ qui est convexe en f par inégalité de Holder. On a :

$$\Lambda_{\mu_U}(f) = \log \int e^f d\mu_U = \Lambda_\mu(f - U) - \Lambda_\mu(-U) \leq \frac{1}{p} \Lambda_\mu(-pU) + \frac{1}{q} \Lambda_\mu(qf) - \Lambda_\mu(-U)$$

par inégalité de Holder en considérant l'exposant conjugué $q := \frac{p}{p-1}$ de p . Par la formule variationnelle de Donsker-Varadhan, on en déduit que :

$$\mathbf{H}[\nu|\mu_U] = \sup_{f \in \mathcal{M}_b(E)} \left\{ \int f d\nu - \Lambda_{\mu_U}(f) \right\} \geq \sup_{f \in \mathcal{M}_b(E)} \left\{ \int f d\nu - \frac{1}{q} \Lambda_\mu(qf) \right\} + \Lambda_\mu(-U) - \frac{1}{p} \Lambda_\mu(-pU)$$

Or :

$$\sup_{f \in \mathcal{M}_b(E)} \left\{ \int f d\nu - \frac{1}{q} \Lambda_\mu(qf) \right\} + \Lambda_\mu(-U) - \frac{1}{p} \Lambda_\mu(-pU) = \frac{1}{q} \mathbf{H}[\nu|\mu] + \Lambda_\mu(-U) - \frac{1}{p} \Lambda_\mu(-pU)$$

Donc si $\mathbf{H}[\nu|\mu_U] < +\infty$, $\mathbf{H}[\nu|\mu] < +\infty$ ou de façon équivalente, $\log\left(\frac{d\nu}{d\mu}\right) \in L^1(\nu)$ et :

$$\log\left(\frac{d\nu}{d\mu_U}\right) = \log\left(\frac{d\nu}{d\mu}\right) + U + \Lambda_\mu(-U) \in L^1(\nu)$$

Ce qui prouve la proposition. \square

Proposition 6 (Propagation du chaos) Soit $(\nu_t)_{t \geq 0}$ le flot solution de l'équation de McKean-Vlasov de distribution initiale ν_0 donnée telle que $\int \|x\|^2 d\nu_0(x) < +\infty$. Soient μ_t^N la loi de la solution de notre système à champ moyen (\star) $X^N(t) = (X_1^N(t), \dots, X_N^N(t))$ telle que $\mu_0^N = \nu_0^{\otimes N}$, et $\mu_t^{N,I}$ la loi de $(X_i^N(t))_{i \in I}$ pour un ensemble d'indices $I \subset \mathbb{N}^*$. Alors, pour tous t et I tels que $\text{Card}(I) < +\infty$, on a :

$$\mu_t^{N,I} \xrightarrow{N \rightarrow +\infty} \nu_t^{\otimes \text{Card}(I)}$$

en distance L^2 -Wasserstein \mathcal{W}_2 .

Pour la démonstration de cette proposition, voir références [9] et [10] en bibliographie. Quant à l'interprétation de cette proposition sur la propagation du chaos, asymptotiquement², k particules arbitrairement choisies parmi les N sont indépendantes et identiquement distribuées de lois données par le flot $(\nu_t)_{t \geq 0}$ des distributions du processus de McKean-Vlasov.

Proposition 7 (Unicité de l'argument minimum de H_W) Si :

$$c_{Lip,m} \|\nabla_{x,y}^2 W\|_\infty < 1,$$

alors l'argument minimum ν_∞ de l'énergie libre E_f est unique.

Preuve

On va d'abord montrer qu'un tel argument minimum existe. En effet, d'après la référence [11], sous l'hypothèse **(H2)**, si $\mathbf{H}[v|\alpha] < +\infty$, $\mathbb{E}_{v \otimes 2}[W] < +\infty$ et $E_f : \mathcal{M}_1(\mathbb{R}^d) \rightarrow \mathbb{R}$ est inf-compacte. Ce qui prouve qu'il existe un argument minimum ν_∞ pour E_f . Il reste juste à montrer l'unicité. Si ν une mesure de probabilité est un argument minimum de E_f , $\mathbf{H}[v|\alpha] < +\infty$, et donc $\int \|x\|^2 d\nu(x) < +\infty$ par l'hypothèse **(H1)**. En regardant les dérivées au sens de Gateaux, on voit que ν est un point fixe de Φ définie par :

$$\Phi(\nu) := \frac{1}{Z'} e^{-V-W*\nu} dx$$

avec Z' un facteur de normalisation. Ici, $W * \nu(x) := \mathbb{E}_\nu[W(x, \cdot)]$ est bien définie parce que par bornitude de $\nabla^2 W$, $|W(x, y)| \leq \beta(1 + \|x\|^2 + \|y\|^2)$. On en déduit que :

$$\Phi : \mathcal{M}_1^2(\mathbb{R}^d) \rightarrow \mathcal{M}_1^2(\mathbb{R}^d)$$

En effet, vu que l'hamiltonien $H_\nu = V + W * \nu$ (pour $\nu \in \mathcal{M}_1^2(\mathbb{R}^d)$) satisfait la condition suivante de dissipativité $\forall x, y \in \mathbb{R}^d$:

$$-\left\langle \frac{x-y}{\|x-y\|}, \nabla H_\nu(x) - \nabla H_\nu(y) \right\rangle \leq b_0(\|x-y\|) \quad ,$$

le générateur $\mathcal{L}_\nu = \Delta - \nabla H_\nu \cdot \nabla$ associé satisfait à une estimation d'écart spectral lipschitzien d'après le théorème sur l'inégalité de Poincaré pour les mesures de Gibbs précédemment prouvé. Ce qui implique un écart spectral pour la mesure $\nu' = \Phi(\nu)$, et en particulier pour $\delta > 0$, $\int e^{\delta\|x\|} d\nu'(x) < +\infty$. (voir référence [12] en bibliographie.) On en déduit que si $\nu \in \mathcal{M}_1^2(\mathbb{R}^d)$, alors $\Phi(\nu)$ aussi. Quant à la preuve de l'unicité, on va prouver que Φ est une contraction de $(\mathcal{M}_1^2(\mathbb{R}^d), \mathcal{W}_1)$. Soient ν_0 et ν_1 deux mesures de probabilité admettant des 2-moments. Posons :

$$\nu_t = (1-t)\nu_0 + t\nu_1, \quad \mu_t = \Phi(\nu_t)$$

Pour une fonction 1-Lipschitzienne f , on a :

$$\frac{d}{dt} \mu_t(f) = \frac{d}{dt} \mathbb{E}_{\mu_t}[f] = \mathbf{Cov}_{\mu_t}[f, -\partial_t(W * \nu_t)] = \mathbf{Cov}_{\mu_t}[f, -W * (\nu_1 - \nu_0)]$$

De plus :

$$\|\nabla_x(W * (\nu_1 - \nu_0))\| = \|(\nabla_x W) * (\nu_1 - \nu_0)\| \leq \|\nabla_{x,y}^2 W\|_\infty \mathcal{W}_1(\nu_0, \nu_1)$$

On a aussi via l'opérateur de Poisson associé à \mathcal{L}_{ν_t} :

$$\mathbf{Cov}_{\mu_t}[f, -W * (\nu_1 - \nu_0)] = \left\langle (-\mathcal{L}_{\nu_t})^{-1} f, \mathcal{L}_{\nu_t} W * (\nu_1 - \nu_0) \right\rangle_{\mu_t} = \int \left\langle \nabla(-\mathcal{L}_{\nu_t})^{-1} f, \nabla W * (\nu_1 - \nu_0) \right\rangle d\mu_t$$

Il en découle que par Cauchy-Schwarz et ce qui précède :

$$\mathbf{Cov}_{\mu_t}[f, -W * (\nu_1 - \nu_0)] \leq c_{Lip,m} \|\nabla_{x,y}^2 W\|_\infty \mathcal{W}_1(\nu_0, \nu_1)$$

2. Quand le nombre de particules tend vers l'infini

Et vu que : (théorème fondamental de l'analyse)

$$\mu_1(f) - \mu_0(f) = \int_0^1 \frac{d}{dt} \mu_t(f) dt \leq c_{Lip,m} \|\nabla_{x,y}^2 W\|_\infty \mathcal{W}_1(v_0, v_1),$$

par la relation de dualité de Kantorovitch-Rubinstein, on a bien :

$$\mathcal{W}_1(\Phi(v_0), \Phi(v_1)) \leq c_{Lip,m} \|\nabla_{x,y}^2 W\|_\infty \mathcal{W}_1(v_0, v_1)$$

et $c_{Lip,m} \|\nabla_{x,y}^2 W\|_\infty < 1$. Donc, il existe un unique argument minimum car sinon, on aurait $c_{Lip,m} \|\nabla_{x,y}^2 W\|_\infty \geq 1$, ce qui est absurde par hypothèses. Ici, on n'a pas utilisé de théorèmes du point fixe du type Banach ou Picard ($(\mathcal{M}_1^2(\mathbb{R}^d), \mathcal{W}_1)$ n'est pas complet!), c'est juste la contraction qu'on utilise en raisonnant par l'absurde vu qu'on a prouvé l'existence avant. \square

Proposition 8 (Convergence en information de Fisher) Si $\mathbf{I}[v|\alpha] < +\infty$, on a :

$$\frac{1}{N} \mathbf{I}[v^{\otimes N} | \mu^N] \xrightarrow{N \rightarrow +\infty} \mathbf{I}_W[v] \quad .$$

Preuve

Pour toute mesure de probabilité ν sur \mathbb{R}^d telle que $\mathbf{I}[v|\alpha] < +\infty$, par la condition de Lyapunov (H1) sur le potentiel V (voir référence [13] en bibliographie), on a :

$$c_1 \int \|x\|^2 d\nu \leq c_2 + \mathbf{I}[v|\alpha] < +\infty$$

Comme les dérivées du second ordre de W sont bornées par la condition sur sa hessienne, $\nabla_x W$ a une croissance linéaire. Donc $\nabla_x W \in L^2(\nu^{\otimes 2})$. Par la loi des grands nombres pour les suites indépendantes et identiquement distribuées, on a successivement :

$$\frac{1}{N} \mathbf{I}[v^{\otimes N} | \mu^N] = \frac{1}{4N} \int \left\| \nabla \log \left(\frac{d\nu^{\otimes N}}{d\mu^N} \right) \right\|^2 d\nu^{\otimes N} = \frac{1}{4N} \int \sum_{i=1}^N \left\| \nabla_{x_i} \log \left(\frac{d\nu^{\otimes N}}{d\alpha^{\otimes N}} \right) + \frac{1}{N-1} \sum_{j \neq i} \nabla_x W(x_i, x_j) \right\|^2 d\nu^{\otimes N}$$

Or :

$$\frac{1}{4N} \int \sum_{i=1}^N \left\| \nabla_{x_i} \log \left(\frac{d\nu^{\otimes N}}{d\alpha^{\otimes N}} \right) + \frac{1}{N-1} \sum_{j \neq i} \nabla_x W(x_i, x_j) \right\|^2 d\nu^{\otimes N}$$

est égal à :

$$\frac{1}{4} \int \left\| \nabla \log \left(\frac{d\nu}{d\alpha} \right) (x_1) + \frac{1}{N-1} \sum_{j=2}^N \nabla_x W(x_1, x_j) \right\|^2 d\nu^{\otimes N}$$

Et quand N tend vers l'infini, on a :

$$\frac{1}{4} \int \left\| \nabla \log \left(\frac{d\nu}{d\alpha} \right) (x_1) + \frac{1}{N-1} \sum_{j=2}^N \nabla_x W(x_1, x_j) \right\|^2 d\nu^{\otimes N} \xrightarrow{N \rightarrow +\infty} \frac{1}{4} \int \left\| \nabla \log \left(\frac{d\nu}{d\alpha} \right) (x_1) + \int \nabla_x W(x_1, y) d\nu(y) \right\|^2 d\nu(x_1)$$

De plus :

$$\frac{1}{4} \int \left\| \nabla \log \left(\frac{d\nu}{d\alpha} \right) (x_1) + \int \nabla_x W(x_1, y) d\nu(y) \right\|^2 d\nu(x_1) = \mathbf{I}_W[v]$$

Ce qui prouve la proposition. \square

Preuve du théorème de convergence exponentielle : Maintenant, on peut démontrer le théorème sur la convergence exponentielle.

Preuve

Le premier point du théorème concernant l'argument minimum v_∞ de \mathbf{H}_W est juste une conséquence directe de l'une des propositions (celle sur l'argument minimum de l'énergie libre \mathbf{E}_f) ci-dessus. Il reste juste à prouver les autres points. On suppose sans perte de généralité que $\mathbf{I}[v|\alpha] < +\infty$ (sinon $\mathbf{I}_W[v] = +\infty$). Vu que la hessienne $\nabla^2 V$ de V est bornée par le bas, et V satisfait une condition de Lyapunov, par Cattiaux-Guillin-Wu (voir référence [14] en bibliographie), α satisfait une inégalité de Sobolev logarithmique. Donc $\mathbf{H}[v|\alpha] < +\infty$. Par l'inégalité de Sobolev logarithmique satisfaite par μ^N , on a :

$$\rho_{LS}(\mu^N)\mathbf{H}[v^{\otimes N}|\mu^N] \leq 2\mathbf{I}[v^{\otimes N}|\mu^N]$$

parce qu'une telle inégalité pour μ^N est équivalente au fait que $\forall \eta \in \mathcal{M}_1((\mathbb{R}^d)^N)$:

$$\rho_{LS}(\mu^N)\mathbf{H}[\eta|\mu^N] \leq 2\mathbf{I}[\eta|\mu^N]$$

Et on rappelle que $\rho_{LS}(\mu^N) \geq \rho_{LS,m}(1 - \gamma_0)^2$. En multipliant cette inégalité par $\frac{1}{N}$ et en faisant tendre N vers l'infini, on obtient par les propositions précédentes que :

$$\rho_{LS}\mathbf{H}_W[v] \leq 2\mathbf{I}_W[v]$$

Par Otto-Villani (voir référence [15]) ou Bobkov-Gentil-Ledoux (voir référence [16]), l'inégalité de Sobolev logarithmique implique une T_2 -inégalité de Talagrand, c'est-à-dire $\forall Q \in \mathcal{M}_1((\mathbb{R}^d)^N)$:

$$\rho_{LS}(\mu^N)\mathcal{W}_2^2(Q, \mu^N) \leq 2\mathbf{H}[Q|\mu^N]$$

Pour $Q = v^{\otimes N}$ avec $\mathbf{H}[v|\alpha] < +\infty$, on a :

$$\rho_{LS}(\mu^N)\frac{1}{N}\mathcal{W}_2^2(v^{\otimes N}, \mu^N) \leq \frac{2}{N}\mathbf{H}[v^{\otimes N}|\mu^N]$$

Notons que :

$$\mathcal{W}_2^2(v^{\otimes N}, \mu^N) \geq \sum_{i=1}^N \mathcal{W}_2^2(v, \mu^{N,i}) = N\mathcal{W}_2^2(v, \mu^{N,1})$$

avec $\mu^{N,i}$ la distribution marginale de x_i sous μ^N et toutes ces distributions marginales sont identiques par symétrie de μ^N . En outre, par unicité de v_∞ et principe de grandes déviations (théorème de Sanov) pour :

$$\frac{1}{N} \sum_{i=1}^N \delta_{x_i}$$

sous μ^N (voir référence [11]), $\forall f \in \mathcal{C}_b(\mathbb{R}^d)$:

$$\mu^{N,1}(f) = \int \frac{1}{N} \sum_{i=1}^N f(x_i) d\mu^N \xrightarrow{N \rightarrow +\infty} v_\infty(f) = \int f dv_\infty$$

Ou de façon équivalente, $\mu^{N,1}$ converge faiblement (en loi ou au sens des distributions) vers v_∞ . Par les propositions précédentes et semi-continuité inférieure de \mathcal{W}_2 , on a :

$$\rho_{LS}\mathcal{W}_2^2(v, v_\infty) \leq \rho_{LS} \liminf_{N \rightarrow +\infty} \mathcal{W}_2^2(v, \mu^{N,1}) \leq 2\mathbf{H}_W[v]$$

Ce qui prouve la T_2 -inégalité de Talagrand pour l'équation de McKean-Vlasov. Ainsi, on a prouvé le second point du théorème. Il reste juste la preuve du troisième point du théorème sur le flot solution de l'équation de McKean-Vlasov. On suppose que $\mathbf{H}_W[v_0] < +\infty$ sans perte de généralité. Fixons $t > 0$. On sait d'après les propositions que :

$$\lim_{N \rightarrow +\infty} \frac{1}{N} \mathbf{H}[v_0^{\otimes N}|\mu^N] = \mathbf{H}_W[v_0]$$

De plus, par équivalence entre l'inégalité de Sobolev logarithmique pour μ^N et la convergence exponentielle en entropie pour μ_t^N la loi de $X^N(t) = (X_i^N(t))_{i=1, \dots, N}$, on a :

$$\frac{1}{N} \mathbf{H}[\mu_t^N | \mu^N] \leq \frac{1}{N} \mathbf{H}[\mu_0^N | \mu^N] \cdot e^{-\rho_{LS}(\mu^N) \frac{t}{2}} = \frac{1}{N} \mathbf{H}[\nu_0^{\otimes N} | \mu^N] \cdot e^{-\rho_{LS}(\mu^N) \frac{t}{2}} < +\infty$$

Par l'une des propositions, on a aussi $\mathbf{H}[\mu_t^N | \alpha^{\otimes N}] < +\infty$. Vu que μ_t^N admet un moment d'ordre 2 (facile par la théorie des EDS), et W est à croissance quadratique, on a :

$$W(x_i, x_j) \in L^1(\mu_t^N)$$

D'après les propositions encore :

$$\frac{1}{N} \mathbf{H}[\mu_t^N | \alpha^{\otimes N}] \geq \mathbf{H}[\mu_t^{N,1} | \alpha]$$

Et par propagation du chaos et semi-continuité inférieure de l'entropie relative $\nu \mapsto \mathbf{H}[\nu | \alpha]$:

$$\liminf_{N \rightarrow +\infty} \mathbf{H}[\mu_t^{N,1} | \alpha] \geq \mathbf{H}[\nu_t | \alpha]$$

On en déduit que :

$$\liminf_{N \rightarrow +\infty} \frac{1}{N} \mathbf{H}[\mu_t^N | \mu^N] = \liminf_{N \rightarrow +\infty} \left(\frac{1}{N} \mathbf{H}[\mu_t^N | \alpha^{\otimes N}] + \int \frac{1}{N(N-1)} \sum_{1 \leq i < j \leq N} W(x_i, x_j) d\mu_t^N + \frac{1}{N} \log(\tilde{Z}_N) \right)$$

Or :

$$\liminf_{N \rightarrow +\infty} \left(\frac{1}{N} \mathbf{H}[\mu_t^N | \alpha^{\otimes N}] + \int \frac{1}{N(N-1)} \sum_{1 \leq i < j \leq N} W(x_i, x_j) d\mu_t^N + \frac{1}{N} \log(\tilde{Z}_N) \right)$$

est minoré par :

$$\mathbf{H}[\nu_t | \alpha] + \liminf_{N \rightarrow +\infty} \left(\frac{1}{2} \int W(x_1, x_2) d\nu_t \right) - \inf_{\eta \in \mathcal{M}_1(\mathbb{R}^d)} \mathbf{E}_f[\eta] = \mathbf{H}[\nu_t | \alpha] + \frac{1}{2} \int \int W(x_1, x_2) d\nu_t(x_1) d\nu_t(x_2) - \inf_{\eta \in \mathcal{M}_1(\mathbb{R}^d)} \mathbf{E}_f[\eta]$$

Et :

$$\mathbf{H}[\nu_t | \alpha] + \frac{1}{2} \int \int W(x_1, x_2) d\nu_t(x_1) d\nu_t(x_2) - \inf_{\eta \in \mathcal{M}_1(\mathbb{R}^d)} \mathbf{E}_f[\eta] = \mathbf{H}_W[\nu_t]$$

Donc :

$$\liminf_{N \rightarrow +\infty} \frac{1}{N} \mathbf{H}[\mu_t^N | \mu^N] \geq \mathbf{H}_W[\nu_t]$$

Ce qui prouve modulo ce qui précède le dernier point du théorème de convergence exponentielle. \square

Troisième partie

Systemes auto-attractifs et schémas d'Euler

Chapitre 7

Systèmes auto-attractifs

Pour plus de détails sur ce chapitre, consulter principalement les références [93] et [94] en bibliographie.

7.1 Introduction générale

On considère l'équation différentielle stochastique dite diffusion d'auto-interaction :

$$dX_t = \sigma dB_t - (\nabla V(X_t) + \nabla W * \mu_t(X_t)) dt$$

avec $\mu_t := \frac{1}{t} \int_0^t \delta_{X_s} ds$ mesure d'occupation, V et W sont respectivement des potentiels de confinement et d'interaction. Dans ce qui suit, on suppose que $W(x, y) = W_0(x - y)$ avec $W_0 : \mathbb{R}^d \rightarrow \mathbb{R}$ une fonction paire au moins deux fois continûment différentiable. Dans ce cas, on a :

$$\nabla W * \mu_t(X_t) = \frac{1}{t} \int_0^t \nabla W_0(X_t - X_s) ds$$

On fait les hypothèses suivantes sur les potentiels qui, comme dans McKean-Vlasov, vont nous donner des propriétés adéquates :

(P1) $V \in \mathcal{C}^2(\mathbb{R}^d)$, $W \in \mathcal{C}^2(\mathbb{R}^d \times \mathbb{R}^d)$, $V \geq 1$ et $W \geq 0$. On dit qu'on a des propriétés de régularité et de positivité.

(P2) $\exists C > 0$ telle que $\forall (x, y) \in \mathbb{R}^d \times \mathbb{R}^d$: (propriété de croissance)

$$\|\nabla V(x) - \nabla V(y)\| \leq C \min(\|x - y\|, 1)(V(x) + V(y)).$$

(P3) $\exists \kappa \geq 1$ telle que $\forall (x, y) \in \mathbb{R}^d \times \mathbb{R}^d$: (propriétés de domination)

$$W(x, y) \leq \kappa(V(x) + V(y)), \|\nabla_{x,x}^2 W(x, y)\| + \|\nabla_x W(x, y)\| \leq \kappa(V(x) + W(x, y))$$

et :

$$\lim_{\|x\| \rightarrow +\infty} \sup_{y \in \mathbb{R}^d} \frac{\|\nabla V(x)\|^2 + 2\langle \nabla V(x), \nabla_x W(x, y) \rangle}{V(x) + W(x, y)} = +\infty.$$

(P4) $\exists \alpha > 0$, $a > 0$, $\delta > 1$ et $M \in \mathbb{R}$ telles que $\forall (x, y, z) \in \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^d$: (propriétés de courbure)

$$\langle x, \nabla V(x) + \nabla_x W(x, y) \rangle \geq a\|x\|^{2\delta} - \alpha \text{ et } \langle (\nabla^2 V(x) + \nabla_{x,y}^2 W(x, y))z, z \rangle \geq M\|z\|^2.$$

Remarques :

On voit que ce processus comme celui de McKean-Vlasov n'est pas markovien vu qu'il dépend de son "passé" mais tout comme pour McKean-Vlasov, (X_t, μ_t) est markovien. Si on fige la mesure, c'est-à-dire en considérant $\eta \in \mathcal{M}_1(\mathbb{R}^d)$ fixée, l'EDS :

$$dY_t = \sigma dB_t - (\nabla V(Y_t) + \nabla W * \eta(Y_t))dt$$

est une diffusion "gradient" d'hamiltonien $H_\eta(x) := V(x) + W * \eta(x)$, de générateur :

$$\mathcal{L}_\eta := \frac{\sigma^2}{2} \Delta - \nabla H_\eta \cdot \nabla$$

et sous des conditions adéquates sur la régularité de l'hamiltonien, il existe une unique mesure de Gibbs invariante donnée par :

$$\mu_\eta(dx) := \frac{e^{-\frac{2}{\sigma^2} H_\eta(x)}}{Z_\eta} dx$$

avec Z_η le facteur de normalisation et on peut montrer que cette mesure est l'unique argument minimum de la fonctionnelle dite d'énergie libre définie $\forall \nu \in \mathcal{M}_1(\mathbb{R}^d)$ par :

$$\mathbf{E}_\eta[\nu] := \mathbb{E}_\nu[H_\eta] + \frac{\sigma^2}{2} \mathbf{H}[\nu|dx] \quad .$$

De plus :

$$\mu_\eta \xrightarrow{\sigma \rightarrow 0} \delta_{\text{argmin} H_\eta} \quad .$$

Par le théorème ergodique de Birkhoff, on a :

$$\frac{1}{t} \int_0^t \delta_{Y_s} ds \xrightarrow{t \rightarrow +\infty} \mu_\eta \quad .$$

Pour la clarté de ce qui suit, on introduit les notations suivantes :

$$\forall r > 0 \text{ et } \mu \in \mathcal{M}_1(\mathbb{R}^d) \quad \mu_t^r := \frac{r}{r+t} \mu + \frac{1}{r+t} \int_0^t \delta_{X_s} ds.$$

$$\Pi : \eta \in \mathcal{M}_1(\mathbb{R}^d) \longmapsto \mu_\eta.$$

$$\forall \varphi, \|\varphi\|_V := \sup_{x \in \mathbb{R}^d} \frac{|\varphi(x)|}{V(x)}. \text{ Et vu que } V \geq 1, \|\cdot\|_V \leq \|\cdot\|_\infty.$$

$$\mathcal{C}^0(\mathbb{R}^d, V) := \{\varphi \in \mathcal{C}^0(\mathbb{R}^d), \|\varphi\|_V < +\infty\} \text{ et } \mathcal{C}^k(\mathbb{R}^d, V) := \mathcal{C}^0(\mathbb{R}^d, V) \cap \mathcal{C}^k(\mathbb{R}^d).$$

En notant respectivement $\mathcal{M}(\mathbb{R}^d)$ et $\mathcal{M}_1(\mathbb{R}^d)$ les espaces des mesures signées bornées et de probabilité sur \mathbb{R}^d , on définit :

$$\mathcal{M}(\mathbb{R}^d, V) := \left\{ \mu \in \mathcal{M}(\mathbb{R}^d), \int_{\mathbb{R}^d} V(x) |\mu|(dx) < \infty \right\}$$

avec $|\mu| := \mu^+ + \mu^-$ la variation de μ et (μ^+, μ^-) la décomposition de Hahn-Jordan de μ . Remarquons que $\gamma(dx) := \frac{e^{-\frac{2}{\sigma^2} V(x)}}{Z_\sigma} dx$ appartient à cet espace. On définit aussi $\forall \mu \in \mathcal{M}(\mathbb{R}^d, V)$:

$$\|\mu\|_V := \sup_{\varphi \in \mathcal{C}^0(\mathbb{R}^d, V), \|\varphi\|_V \leq 1} |\langle \mu, \varphi \rangle|$$

la norme duale associée à $\|\cdot\|_V$ et on rappelle que $\langle \mu, \varphi \rangle := \int \varphi d\mu$ est le crochet de dualité. Et $(\mathcal{M}(\mathbb{R}^d, V), \|\cdot\|_V)$ est un espace de Banach, ainsi $\mathcal{M}_1(\mathbb{R}^d, V)$ est un espace métrique complet pour la distance induite par cette norme. On définit la topologie faible associée par :

$$\mu_n \xrightarrow{V\text{-faiblement}} \mu \iff \forall \varphi \in \mathcal{C}^0(\mathbb{R}^d, V), \quad \langle \mu_n, \varphi \rangle \xrightarrow{n \rightarrow +\infty} \langle \mu, \varphi \rangle,$$

et vu que $\mathcal{C}_b^0(\mathbb{R}^d) \subset \mathcal{C}^0(\mathbb{R}^d, V)$, cette convergence implique la convergence faible classique. On rappelle que $\mathcal{M}_1(\mathbb{R}^d, V)$ muni de cette topologie peut être associé à une métrique donnée par :

$$d_V(\mu, \nu) := \sum_{j=1}^{+\infty} 2^{-j} |\mathbb{E}_\mu[\varphi_j] - \mathbb{E}_\nu[\varphi_j]|,$$

avec $(\varphi_j)_{j \geq 1}$ une suite de fonctions dense dans :

$$\left\{ \varphi \in \mathcal{C}^0(\mathbb{R}^d, V), \|\varphi\|_V \leq 1 \right\}, \text{ et une telle suite existe par séparabilité.}$$

Ainsi la topologie faible est celle engendrée par cette métrique.

$\forall \beta > 1$, on définit :

$$\mathcal{M}_{1,\beta}(\mathbb{R}^d, V) := \left\{ \mu \in \mathcal{M}_1(\mathbb{R}^d, V), \int V(x) \mu(dx) \leq \beta \right\}$$

C'est une partie faiblement compacte de $\mathcal{M}_1(\mathbb{R}^d, V)$.

Pour $\eta \in \mathcal{M}_1(\mathbb{R}^d, V)$, on a $\mathcal{C}^0(\mathbb{R}^d, V) \subset L^2(\Pi(\eta))$ et $\forall \varphi$ et ψ , on définit :

$$\langle \varphi, \psi \rangle_{\Pi(\eta)} := \int_{\mathbb{R}^d} \varphi(x) \overline{\psi}(x) \Pi(\eta)(dx)$$

On introduit la notion de pseudo-trajectoire asymptotique relativement à un semi-flot qu'on peut définir dans un cadre plus général à savoir celui d'un espace métrique...

Définition 7.1 (Pseudo-trajectoire asymptotique)

Une fonction continue $\xi : \mathbb{R}_+ \rightarrow \mathcal{M}_1(\mathbb{R}^d, V)$ est appelée pseudo-trajectoire asymptotique relativement à un semi-flot Φ si $\forall T > 0$:

$$\lim_{t \rightarrow +\infty} \sup_{0 \leq s \leq T} d_V(\xi_{t+s}, \Phi_s(\xi_t)) = 0.$$

Dans ce qui suit, on note $\Phi : \mathbb{R}_+ \times \mathcal{M}_1(\mathbb{R}^d, V) \rightarrow \mathcal{M}_1(\mathbb{R}^d, V)$ le semi-flot engendré par :

$$\Phi_t(\eta) = e^{-t} \eta + \int_0^t e^{s-t} \Pi \circ \Phi_s(\eta) ds, \text{ avec } \Phi_0(\eta) = \eta,$$

dont l'existence est prouvée plus bas. On verra que $(\mu_{h(t)}^r)_{t \geq 0}$ définit une pseudo-trajectoire asymptotique pour ce semi-flot avec $h(t) := r(e^t - 1)$.

Définition 7.2 (ω -limite)

Soit $\xi : \mathbb{R}_+ \rightarrow \mathcal{M}_1(\mathbb{R}^d, V)$ une fonction continue. L'ensemble des ω -limites de ξ noté $\omega(\xi_t, t \geq 0)$ est celui des limites faibles des suites (ξ_{t_n}) avec (t_n) tendant vers l'infini en croissant, c'est-à-dire :

$$\omega(\xi_t, t \geq 0) := \bigcap_{t \geq 0} \overline{\xi([t, +\infty))}.$$

Et ici, $\overline{\xi([t, +\infty))}$ est la fermeture faible de $\xi([t, +\infty))$. En particulier, c'est un fermé pour la topologie de la convergence faible.

7.2 Résultats préliminaires

On note $(X_t^\eta)_{t \geq 0}$ l'unique solution forte issue de $x \in \mathbb{R}^d$ de la diffusion "gradient" ci-dessus avec toujours les mêmes conditions sur les potentiels de confinement et d'interaction.

Proposition 9 $\forall \mu \in \mathcal{M}(\mathbb{R}^d, V)$, on a $W * \mu \in \mathcal{C}^2(\mathbb{R}^d, V)$ et $\exists \kappa \geq 1$ telle que $\forall x \in \mathbb{R}^d$:

$$|W * \mu(x)| \leq 2\kappa \|\mu\|_V V(x)$$

Il existe $D > 0$ telle que $\forall \mu \in \mathcal{M}(\mathbb{R}^d, V)$:

$$|\Delta H_\mu| \leq D \cdot H_\mu$$

Preuve

En effet, en se rappelant que $H_\mu := V + W * \mu$, c'est une conséquence directe des propriétés **(P2)** ($\implies \Delta V \leq aV$) et **(P3)** de croissance et de domination sur les potentiels. \square

On en déduit le corollaire suivant :

Corollaire 1 Soit $\beta > 1$. $\forall \mu \in \mathcal{M}_{1,\beta}(\mathbb{R}^d, V)$, on a l'inégalité de borne supérieure :

$$\int_{\mathbb{R}^d} e^{-H_\mu(x)} dx \geq e^{-2\kappa\beta} \int_{\mathbb{R}^d} e^{-2(\kappa+1)V(x)} dx \geq \int_{\mathbb{R}^d} e^{-2(2\kappa\beta+1)V(x)} dx$$

De plus H_μ est une bonne fonction de Lyapunov telle que :

$$V \leq H_\mu \leq 3\kappa\beta V$$

Proposition 10 (Bonne fonction de Lyapunov) $\forall \eta \in \mathcal{M}_{1,\beta}(\mathbb{R}^d, V)$, $(X_t^\eta)_{t \geq 0}$ n'explose jamais presque sûrement.

Preuve

En effet, d'après la formule d'Ito et ce qui précède, on a $\mathcal{L}_\eta H_\eta \leq DH_\eta$. Il en découle que :

$$\mathbb{E}[H_\eta(X_t^\eta)] \leq H_\eta(x) e^{Dt},$$

ce qui prouve le résultat attendu. \square

Le théorème ergodique classique est vrai pour $(X_t^\eta)_{t \geq 0}$, c'est-à-dire presque sûrement $\forall f \in \mathcal{C}^0(\mathbb{R}^d, V)$:

$$\frac{1}{t} \int_0^t f(X_s^\eta) ds \xrightarrow{t \rightarrow +\infty} \mu_\eta(f) := \langle \mu_\eta, f \rangle = \int_{\mathbb{R}^d} f(x) \mu_\eta(dx).$$

On note aussi respectivement $(P_t^\eta)_{t \geq 0}$ le semi-groupe fellerien, Q_η le noyau fondamental¹ et K_η la projection orthogonale associés à la diffusion "gradient" de potentiel H_η . On rappelle que $\forall f \in \mathcal{C}^\infty(\mathbb{R}^d, V)$:

$$Q_\eta f = \int_0^{+\infty} \left(P_t^\eta f - \langle \Pi(\eta), f \rangle \right) dt, \quad K_\eta f = f - \langle \Pi(\eta), f \rangle.$$

Donc par invariance de la mesure $\Pi(\eta)$ pour $(P_t^\eta)_{t \geq 0}$, on en déduit que :

$$Q_\eta = \int_0^{+\infty} K_\eta \circ P_t^\eta dt.$$

De plus, via l'équation de Poisson, on a :

$$\mathcal{L}_\eta \circ Q_\eta = Q_\eta \circ \mathcal{L}_\eta = -K_\eta.$$

Par hypothèses et définition, $(P_t^\eta)_{t \geq 0}$ est pour tout $\eta \in \mathcal{M}_1(\mathbb{R}^d, V)$ un semi-groupe d'opérateurs uniformément bornés en temps de $L^2(\Pi(\eta)) \rightarrow L^\infty$. Par ultracontractivité², le semi-groupe est V -uniformément ergodique en ce sens que $\forall f \in \mathcal{C}^0(\mathbb{R}^d, V)$:

$$\|K_\eta \circ P_t^\eta f\|_V \leq c(\eta) e^{-\lambda(\eta)t} \|f\|_V, \quad c(\eta) > 0 \text{ et } \lambda(\eta) > 0.$$

Il en découle que :

$$\|Q_\eta f\|_V \leq c(\eta) \|f\|_V \int_0^{+\infty} e^{-\lambda(\eta)t} dt = \frac{c(\eta)}{\lambda(\eta)} \|f\|_V.$$

Ainsi, on est en droit d'appliquer le théorème de prolongement des applications uniformément continues sur une partie dense d'un espace métrique à valeurs dans un espace métrique complet. Par application de la formule d'Ito :

$$Q_\eta f(X_t^\eta) = Q_\eta f(x) + \int_0^t \mathcal{L}_\eta \circ Q_\eta f(X_s^\eta) ds + \sigma \int_0^t \nabla Q_\eta f(X_s^\eta) \cdot dB_s = Q_\eta f(x) - \int_0^t K_\eta f(X_s^\eta) ds + M_t$$

avec M_t le terme martingale associé. Remarquons que ce terme martingale est négligeable par rapport au temps.

Proposition 11 (Critère d'ultracontractivité) Soit $(P_t)_{t \geq 0}$ le semi-groupe de Markov sur \mathbb{R}^d de générateur $\mathcal{L} = \frac{\sigma^2}{2} \Delta - \nabla H \cdot \nabla$ tel qu'il existe $\hbar > 0$ pour lequel, $\nabla^2 H \geq -\hbar \cdot \mathbf{Id}_d$ au sens des formes quadratiques. Supposons qu'il existe une fonction continue et croissante $\chi : \mathbb{R}_+ \rightarrow \mathbb{R}_+^*$ telle que :

1. $\lim_{r \rightarrow +\infty} \frac{\chi(r)}{r} = +\infty$;
2. $\forall m > 0$, l'application $r \mapsto r \chi(m \log(r))$ est convexe sur $[1, +\infty)$;
3. il existe $b > 0$ pour laquelle $\mathcal{L} \|x\|^2 \leq b - \chi(\|x\|^2)$.

Dans ce cas, le semi-groupe admet une unique mesure invariante. Si de plus, $\forall m > 0$:

$$\int_2^{+\infty} \frac{1}{r \chi(m \log(r))} dr < +\infty, \quad \text{le semi-groupe est ultracontractif.}$$

Dans le cas particulier $\chi(r) = \chi \cdot r^\delta$ avec $\chi > 0$ et $\delta > 1$, $\exists c := c(b, \chi) > 0$ telle que $\forall t \in (0, 1]$:

$$\sup_{f \in \mathcal{C}^\infty(\mathbb{R}^d), f \neq 0} \frac{\|P_t f\|_\infty}{\|f\|_2} \leq e^{ct^{-\frac{\delta}{\delta-1}}}.$$

1. Solution de l'équation de Poisson associée.
2. Qu'on démontre plus bas.

On en déduit de cette proposition que la famille de semi-groupes $(P^\eta)_\eta$ est uniformément ultracontractive.

Preuve

En effet, il suffit de prendre $H = H_\eta$. En appliquant la propriété **(P4)** de courbure, il existe $a > 0$ et $b > 0$ telles que $\forall \eta \in \mathcal{M}_1(\mathbb{R}^d, V)$:

$$\mathcal{L}_\eta \|x\|^2 = d\sigma^2 - 2\nabla H_\eta \cdot x \leq b - a\|x\|^{2\delta}.$$

Ainsi, on peut prendre $\chi(r) = a \cdot r^\delta$ et la constante c ne dépend pas de η . \square

Il en découle que :

$$\sup_{f \in \mathcal{C}^\infty(\mathbb{R}^d, V), f \neq 0} \frac{\|P_t f\|_\infty}{\|f\|_{L^2(\Pi(\eta))}} \leq e^{ct^{-\frac{\delta}{\delta-1}}}.$$

Corollaire 2 (Log-Sobolev uniforme) *La famille de mesures invariantes $(\Pi(\eta))_\eta$ admet une constante de type log-Sobolev universelle en η , c'est-à-dire $\exists \rho > 0$ telle que $\forall \eta$ et $f \in \mathcal{C}^\infty(\mathbb{R}^d, V)$:*

$$\|P_t^\eta \circ K_\eta f\|_{L^2(\Pi(\eta))} \leq e^{-\frac{t}{\rho}} \|K_\eta f\|_{L^2(\Pi(\eta))}.$$

Preuve

ultracontractivité \implies hypercontractivité \iff existence d'une inégalité log-Sobolev. Quant à l'uniformité, elle découle de celle de l'ultracontractivité. \square

Théorème 7.1 (E.U de la solution forte)

Soit :

$$dX_t = \sigma dB_t - \nabla H_{\mu_t^r}(X_t) dt.$$

$\forall x \in \mathbb{R}^d, \mu \in \mathcal{M}_1(\mathbb{R}^d, V)$ et $r > 0$, il existe une unique solution forte $((X_t, \mu_t^r))_{t \geq 0}$.

Preuve

Soit (τ_n) la suite croissante de temps d'arrêt définie par :

$$\tau_0 = 0, \quad \tau_n = \inf \left\{ t \geq \tau_{n-1}, \quad H_{\mu_t^r}(X_t) + \int_0^t \|\nabla H_{\mu_s^r}(X_s)\|^2 ds > n \right\}.$$

En utilisant la fonction de Lyapunov $(x, \mu) \mapsto H_\mu(x)$, le processus $(t, x) \mapsto H_{\mu_t^r}(x)$ est \mathcal{C}^2 en espace et c'est une \mathcal{C}^1 -semi-martingale en temps. En appliquant la formule d'Ito généralisée (ou d'Ito-Ventzell) à $(t, x) \mapsto H_{\mu_{t \wedge \tau_n}^r}(x)$, on obtient :

$$H_{\mu_{t \wedge \tau_n}^r}(X_{t \wedge \tau_n}) = H_\mu(x) + \int_0^{t \wedge \tau_n} \nabla H_{\mu_s^r}(X_s) \cdot dB_s - \int_0^{t \wedge \tau_n} \|\nabla H_{\mu_s^r}(X_s)\|^2 ds + \frac{1}{2} \int_0^{t \wedge \tau_n} \Delta H_{\mu_s^r}(X_s) ds + A$$

avec :

$$A := \int_0^{t \wedge \tau_n} \left(W(X_s, X_s) - W * \mu_s^r(X_s) \right) \frac{ds}{s+r}.$$

Notons que le second terme de cette somme est une vraie martingale. En posant $k := a + \frac{2k}{r} + D$, on en déduit que :

$$\mathbb{E}[H_{\mu_{t \wedge \tau_n}^r}(X_{t \wedge \tau_n})] \leq H_\mu(x) + k \log(t+1) \int_0^t \mathbb{E}[H_{\mu_{s \wedge \tau_n}^r}(X_{s \wedge \tau_n})] ds.$$

Et par le lemme de Gronwall :

$$\mathbb{E}[V(X_{t \wedge \tau_n})] \leq \mathbb{E}[H_{\mu_{t \wedge \tau_n}}(X_{t \wedge \tau_n})] \leq H_{\mu}(x)e^{kt \log(t+1)}.$$

Vu que V est de Lyapunov, $(X_t)_{t \geq 0}$ n'explose pas en temps fini et il existe une unique solution globale. \square

Proposition 12 $\forall \varepsilon > 0 \exists \hbar > 0$ telles que $\forall x \in \mathbb{R}^d, \eta \in \mathcal{M}_{1,\beta}(\mathbb{R}^d, V)$ et $f \in \mathcal{C}^0(\mathbb{R}^d, V)$:

$$|Q_{\eta}f(x)| \leq (\varepsilon V(x) + \hbar)\|f\|_V.$$

De plus, $\forall \varepsilon > 0 \exists \hbar_1 > 0$ telles que $\forall \eta \in \mathcal{M}_{1,\beta}(\mathbb{R}^d, V)$ et $f \in \mathcal{C}^{\infty}(\mathbb{R}^d, V)$, $Q_{\eta}f \in \mathcal{C}^1(\mathbb{R}^d)$ et :

$$\|\nabla Q_{\eta}f\| \leq (\varepsilon V + \hbar_1)\|f\|_V.$$

Preuve

Soit $t_0 \in (0, 1]$. On a :

$$|Q_{\eta}f| \leq \int_0^{+\infty} |P_t^{\eta} \circ K_{\eta}f| dt = \int_0^{t_0} |P_t^{\eta} \circ K_{\eta}f| dt + \int_{t_0}^{+\infty} |P_t^{\eta} \circ K_{\eta}f| dt.$$

En utilisant la propriété de semi-groupe, l'uniforme ultracontractivité et l'uniforme écart spectral, on a aussi :

$$\int_{t_0}^{+\infty} |P_t^{\eta} \circ K_{\eta}f| dt \leq e^{ct_0^{-\frac{\delta}{\delta-1}}} \|K_{\eta}f\|_{L^2(\Pi(\eta))} \int_0^{+\infty} e^{-\frac{t}{\rho}} dt \leq \rho e^{ct_0^{-\frac{\delta}{\delta-1}}} \|f\|_{L^2(\Pi(\eta))}$$

Comme K_{η} est une projection orthogonale sur 1^{\perp} de $L^2(\Pi(\eta))$. Vu que :

$$\|f\|_{L^2(\Pi(\eta))} \leq \|V\|_{L^2(\Pi(\eta))} \|f\|_V,$$

on en déduit que :

$$\int_{t_0}^{+\infty} |P_t^{\eta} \circ K_{\eta}f| dt \leq \rho e^{ct_0^{-\frac{\delta}{\delta-1}}} \|V\|_{L^2(\Pi(\eta))} \|f\|_V.$$

D'autre part :

$$|P_t^{\eta}f| \leq \|f\|_V P_t^{\eta}V \leq \|f\|_V \mathbb{E}[H_{\eta}(X_t^{\eta})] \leq \|f\|_V e^{Dt} H_{\eta}(x).$$

Or :

$$\int_0^{t_0} \mathbb{E}[H_{\eta}(X_t^{\eta})] dt = \mathbf{O}(t_0) \implies \text{le résultat attendu.}$$

Quant à la seconde affirmation, on introduit l'opérateur :

$$\Gamma_2^{\eta}(f) = \|\nabla^2 f\|^2 + \langle \nabla^2 H_{\eta} \cdot \nabla f, \nabla f \rangle, \text{ associé au carré du champ.}$$

Par la propriété (P4) de courbure et le critère de Bakry-Emery, on a :

$$\Gamma_2^{\eta}(f) \geq M \|\nabla f\|^2 \implies \|\nabla P_t^{\eta} \circ K_{\eta}f\|^2 \leq \frac{M}{e^{2Mt} - 1} |P_t^{\eta}(K_{\eta}f)|^2, \quad \forall t > 0 \text{ et } f.$$

On en déduit que :

$$\int_{t_0}^{+\infty} \|\nabla P_t^{\eta} \circ K_{\eta}f\| dt \leq 2\rho \sqrt{\frac{M}{e^{2Mt_0} - 1}} e^{ct_0^{-\frac{\delta}{\delta-1}}} \|V\|_{L^4(\Pi(\eta))} \|f\|_V.$$

Et on conclut comme précédemment. \square

7.3 Calcul différentiel sur $\mathcal{M}_1(\mathbb{R}^d, V)$

L'espace $\mathcal{M}_1(\mathbb{R}^d, V)$ a une structure de variété différentiable de dimension infinie. Cette structure va nous servir pour l'étude du semi flot introduit dans la première section de ce chapitre. Pour plus de détails, voir annexes.

Pour $\mu \in \mathcal{M}_1(\mathbb{R}^d, V)$, on note $\forall k \geq 1$, $\mathcal{C}^k(\mu)$ l'ensemble des chemins $\psi : (-\varepsilon, \varepsilon) \rightarrow \mathcal{M}_1(\mathbb{R}^d, V)$ tels que $\psi(0) = \mu$ et $\psi \in \mathcal{C}^k$.

Définition 7.3 (Fonctions régulières)

On dit que $\phi : \mathcal{M}_1(\mathbb{R}^d, V) \rightarrow \mathbb{R}$ est de classe \mathcal{C}^k si $\forall \mu$ et ψ , la fonction numérique $\phi \circ \psi : (-\varepsilon, \varepsilon) \rightarrow \mathbb{R}$ est de classe \mathcal{C}^k . Par définition, on a :

$$\mathbf{T}_\mu \mathcal{M}_1(\mathbb{R}^d, V) = \left\{ \frac{d}{dt} \psi(t) \Big|_{t=0}, \quad \psi \in \mathcal{C}^1(\mu) \right\} = \mathcal{M}_0(\mathbb{R}^d, V) := \mathcal{M}(\mathbb{R}^d, V) \cap \left\{ \eta, \quad \eta(\mathbb{R}^d) = 0 \right\}.$$

$\forall \mu \in \mathcal{M}_1(\mathbb{R}^d, V)$ et $\eta \in \mathcal{M}_0(\mathbb{R}^d, V)$, on a :

$$\mathcal{D}_\mu \phi(\eta) := \frac{d}{dt} \phi(\mu + t\eta) \Big|_{t=0}$$

qui est une définition de la différentielle par dérivées directionnelles.

Par exemple, $\forall x \in \mathbb{R}^d$, $\mu \mapsto W * \mu(x) \in \mathcal{C}^\infty$ et l'application mesure invariante Π aussi par application du théorème de Lebesgue.

On considère l'espace de Banach des opérateurs bornés sur $\mathcal{C}^\infty(\mathbb{R}^d, V) \subset L^2(\gamma)$ qu'on peut munir de la norme du graphe de \mathcal{L}_μ et dans ce cas, \mathcal{L}_μ appartient au sous espace fermé des opérateurs \mathcal{L} de cet espace tels que $\mathcal{L}1 = 0$.

Proposition 13 Les applications $\mu \mapsto \mathcal{L}_\mu$ et $\mu \mapsto K_\mu$ sont \mathcal{C}^∞ . De plus, $\forall f \in \mathcal{C}^\infty(\mathbb{R}^d, V)$, l'application $\mu \mapsto Q_\mu f$ est aussi \mathcal{C}^∞ pour la topologie forte et :

1. $\mathcal{D}(\mathcal{L}_\mu f) \cdot \eta = -\langle \nabla W * \eta, \nabla f \rangle$ avec ici, $\langle \cdot, \cdot \rangle$ le produit scalaire euclidien canonique;
2. $\mathcal{D}(K_\mu f) \cdot \eta = -\langle \mathcal{D}_\mu \Pi(\eta), f \rangle$ avec ici, $\langle \cdot, \cdot \rangle$ le crochet de dualité;
3. $\mathcal{D}(Q_\mu f) \cdot \eta = \langle \mathcal{D}_\mu \Pi(\eta), Q_\mu f \rangle + Q_\mu \langle \nabla W * \eta, \nabla Q_\mu f \rangle$ avec ici, $\langle \cdot, \cdot \rangle$ indifféremment le crochet de dualité et le produit scalaire euclidien canonique.

Preuve

Les applications $\mu \mapsto \mathcal{L}_\mu$ et $\mu \mapsto K_\mu$ sont \mathcal{C}^∞ car $\forall x \in \mathbb{R}^d$, $\mu \mapsto W * \mu(x) \in \mathcal{C}^\infty$ et l'application mesure invariante Π aussi par application du théorème de Lebesgue. Quant à $\mu \mapsto Q_\mu f$, on va utiliser la résolvante du semi-groupe et des approximations de Q_μ , c'est-à-dire :

$$\forall \lambda > 0, \quad R_\lambda^\mu = \int_0^{+\infty} e^{-\lambda t} P_t^\mu dt = (\lambda \mathbf{Id} - \mathcal{L}_\mu)^{-1}, \quad Q_\mu^\lambda = \int_0^{+\infty} e^{-\lambda t} P_t^\mu \circ K_\mu dt = K_\mu \circ (\lambda \mathbf{Id} - \mathcal{L}_\mu)^{-1}.$$

Il en découle par composition que $\mu \mapsto Q_\mu^\lambda f$ est \mathcal{C}^∞ . Par l'uniforme écart spectral, on montre qu'il existe $c_1 > 0$ et $c_2 > 0$ telles que :

$$\|Q_\mu f - Q_\mu^\lambda f\|_V \leq \int_0^{+\infty} (1 - e^{-\lambda t}) \|P_t^\mu \circ K_\mu f\|_V dt \leq \lambda c_1 \|f\|_V \int_0^{+\infty} t e^{-c_2 t} dt = \frac{\lambda c_1}{c_2^2} \|f\|_V.$$

Donc la convergence (quand $\lambda \rightarrow 0$) en norme subordonnée à la norme $\|\cdot\|_V$ de $(Q_\mu^\lambda)_\lambda$ vers Q_μ est uniforme en μ . On en déduit que $\mu \mapsto Q_\mu f \in \mathcal{C}^0$. De plus :

$$\mathcal{D}(Q_\mu^\lambda) \cdot \eta = \left(\mathcal{D}(K_\mu) \cdot \eta \right) (\lambda \mathbf{Id} - \mathcal{L}_\mu)^{-1} + K_\mu (\lambda \mathbf{Id} - \mathcal{L}_\mu)^{-1} \left(\mathcal{D}(\mathcal{L}_\mu) \cdot \eta \right) (\lambda \mathbf{Id} - \mathcal{L}_\mu)^{-1},$$

et en remplaçant $\mathcal{D}(\mathcal{L}_\mu)$ et $\mathcal{D}(K_\mu)$ par leurs expressions respectives, on va prouver que chacun des termes de cette égalité converge uniformément. En effet, on a :

Pour le premier :

$$\lim_{\lambda \rightarrow 0} \langle \mathcal{D}_\mu \Pi(\eta), (\lambda \mathbf{Id} - \mathcal{L}_\mu)^{-1} f \rangle = \langle \mathcal{D}_\mu \Pi(\eta), Q_\mu f \rangle, \quad \text{uniformément,}$$

comme $\langle \mathcal{D}_\mu \Pi(\eta), (\lambda \mathbf{Id} - \mathcal{L}_\mu)^{-1} f \rangle = \langle \mathcal{D}_\mu \Pi(\eta), Q_\mu^\lambda f \rangle$.

Pour le second :

$$K_\mu (\lambda \mathbf{Id} - \mathcal{L}_\mu)^{-1} \left(\mathcal{D}(\mathcal{L}_\mu) \cdot \eta \right) (\lambda \mathbf{Id} - \mathcal{L}_\mu)^{-1} = Q_\mu^\lambda \langle W * \eta, \nabla Q_\mu^\lambda f \rangle,$$

et $\nabla Q_\mu^\lambda f \xrightarrow{\lambda \rightarrow 0} \nabla Q_\mu f$ uniformément (en μ) vu que par définition de Q_μ^λ , on a :

$$\left\| \nabla Q_\mu^\lambda f - \nabla Q_\mu f \right\| \leq \int_0^{+\infty} \left\| \nabla P_t^\mu \circ K_\mu f \right\| (1 - e^{-\lambda t}) dt.$$

On conclut par la proposition précédente que cette famille de différentielles converge uniformément en μ , et donc que $\mu \mapsto Q_\mu f \in \mathcal{C}^1$. En remarquant que $\mathcal{D}(Q_\mu f) \in \mathcal{C}^1$ par composition, on montre par induction le caractère \mathcal{C}^∞ . \square

Corollaire 3 $\forall f \in \mathcal{C}^\infty(\mathbb{R}^d, V)$:

$$|(\mathcal{D}(Q_\mu) \cdot \eta)(f)(x)| \leq (\varepsilon V^2(x) + k(\varepsilon)) \|f\|_V \|\eta\|_V.$$

Ce qui est une propriété de continuité de la forme linéaire $\eta \in \mathbf{T}_\mu \mathcal{M}_1(\mathbb{R}^d, V) \mapsto (\mathcal{D}(Q_\mu) \cdot \eta)(f)(x)$.

7.4 E.U du semi-flot

La proposition suivante est utile pour trouver un bon cylindre de sécurité.

Proposition 14 (Lipschitz) $\forall \beta > 1$, la restriction $\Pi|_{\mathcal{M}_{1,\beta}(\mathbb{R}^d, V)}$ est lipschitzienne et bornée.

Preuve

En effet, d'après ce qui précède, on a :

$$\|\Pi(\mu)\|_V \leq \left(\int_{\mathbb{R}^d} e^{-4\kappa\beta V(x)} \gamma(dx) \right)^{-1} \int_{\mathbb{R}^d} V(x) \gamma(dx) =: c_\beta < +\infty.$$

De plus, on sait que $\Pi \in \mathcal{C}^\infty$ sur $\mathcal{M}_1(\mathbb{R}^d, V)$ muni de sa topologie forte. Sa différentielle en μ est l'application linéaire continue $\mathcal{D}_\mu \Pi : \mathcal{M}_0(\mathbb{R}^d, V) \rightarrow \mathcal{M}_0(\mathbb{R}^d, V)$ définie $\forall \eta$ par :

$$\mathcal{D}_\mu \Pi(\eta)(dx) = -2 \left(W * \eta(x) - \int_{\mathbb{R}^d} W * \eta(y) \Pi(\mu)(dy) \right) \Pi(\mu)(dx).$$

On en déduit que :

$$\|\mathcal{D}_\mu \Pi(\eta)\|_V \leq 4\kappa(c_\beta + 1) \|\eta\|_V \|V\|_{L^2(\Pi(\mu))}^2,$$

et le calcul fait pour la bornitude de la restriction de Π permet un contrôle uniforme en μ de la norme $L^2(\Pi(\mu))$ de V , ce qui montre que cette différentielle est aussi bornée disons par $c'_\beta > 0$ (\implies la restriction est lipschitzienne). On a donc bel et bien le résultat de la proposition. \square

Théorème 7.2 (Existence locale)

$\forall \mu \in \mathcal{M}_1(\mathbb{R}^d, V)$, l'équation :

$$\Psi_t(\mu) = e^{-t} \mu + \int_0^t e^{s-t} \Pi \circ \Psi_s(\mu) ds, \text{ avec } \Psi_0(\mu) = \mu$$

admet une solution locale qui définit un \mathcal{C}^∞ semi-flot Φ pour la topologie forte.

Preuve

Soit $\beta > 2\|\mu\|_V$ de sorte que $\mu \in \mathcal{M}_{1,\beta}(\mathbb{R}^d, V)$. On introduit le schéma d'approximation de Picard :

$$\mu_t^{(0)} = \mu, \quad \forall n \geq 1 \quad \mu_t^{(n)} = e^{-t} \mu + \int_0^t e^{s-t} \Pi(\mu_s^{(n-1)}) ds.$$

Soit $\varepsilon > 0$ assez petit tel que $\|\mu\|_V + (1 - e^{-\varepsilon})c_\beta \leq \beta$ et $\varepsilon c'_\beta < 1$. On en déduit que $\forall n, \mu_t^{(n)} \in \mathcal{M}_{1,\beta}(\mathbb{R}^d, V)$, ce qui montre que $[0, \varepsilon) \times \mathcal{M}_{1,\beta}(\mathbb{R}^d, V)$ est un bon cylindre de sécurité. On a :

$$\forall t < \varepsilon, \quad \|\mu_t^{(n+1)} - \mu_t^{(n)}\|_V \leq (1 - e^{-\varepsilon}) c'_\beta \sup_{t < \varepsilon} \|\mu_t^{(n)} - \mu_t^{(n-1)}\|_V.$$

Vu que la série de terme général $\sup_{t < \varepsilon} \|\mu_t^{(n)} - \mu_t^{(n-1)}\|_V$ converge, la suite de fonctions $(\mu^{(n)})$ est de Cauchy pour la topologie de la convergence uniforme. Comme $\mathcal{M}_1(\mathbb{R}^d, V)$ est complet, on a bien construit une solution sur $[0, \varepsilon)$. Vu que $\Pi \in \mathcal{C}^\infty$ pour la topologie forte, chaque approximation $\mu \mapsto \mu_t^{(n)}$ l'est aussi par induction, et il suffit de passer à la limite uniformément en μ sur $\mathcal{M}_{1,\beta}(\mathbb{R}^d, V)$ pour conclure que $\Phi \in \mathcal{C}^\infty$. \square

Définition 7.4 (Invariance)

On dit que $A \subset \mathcal{M}_1(\mathbb{R}^d, V)$ est :

- positivement invariant pour Φ si $\forall t \geq 0, \Phi_t(A) \subset A$;
- négativement invariant pour Φ si $\forall t \geq 0, A \subset \Phi_t(A)$;
- invariant pour Φ si $\forall t \geq 0, \Phi_t(A) = A$.

Proposition 15 (Conditions sur l'interaction de non explosion du semi-flot) *Si W est symétrique ($W(x, y) = W(y, x)$) ou uniformément borné en la seconde variable ($W(x, y) \leq \kappa V(x)$), alors le semi-flot n'explose pas.*

Preuve

Dans le second cas, il suffit d'appliquer la proposition 14. En effet, on montre que Π est globalement lipschitzienne de constante C , ce qui signifie que $\Phi_t(\mu)$ appartient à $\mathcal{M}_{1,C}(\mathbb{R}^d, V)$, il ne peut donc pas exploser. Quant au premier cas de figure, remarquons d'abord que l'énergie libre n'est pas une fonction de Lyapunov pour le semi-flot vu qu'en général, $\Phi_t(\mu)$ n'est pas absolument continue par rapport à la mesure de Lebesgue, et donc $\mathcal{F}(\Phi_t(\mu)) = +\infty$. Soit la fonction de Lyapunov $\mathcal{J}(\mu) = \mathcal{F}(\Pi(\mu))$ qu'on peut considérer comme l'énergie libre restreinte au sous-espace des mesures de probabilités absolument continues par rapport à la mesure de Lebesgue de référence. Cette fonctionnelle de Lyapunov est \mathcal{C}^∞ pour la topologie forte et par symétrie :

$$\forall \eta \in \mathcal{M}_0(\mathbb{R}^d, V), \quad \mathcal{D}_\mu \mathcal{F}(\eta) = \int_{\mathbb{R}^d} \left(\log \left(\frac{d\mu}{d\gamma}(x) \right) + 2W * \mu(x) \right) \eta(dx).$$

Il en découle que par calcul de la différentielle d'une composée :

$$\forall \eta \in \mathcal{M}_0(\mathbb{R}^d, V), \quad \mathcal{D}_\mu \mathcal{J}(\eta) = -4 \int_{\mathbb{R}^d} \left(W * \Pi(\mu) - W * \mu \right) \left(W * \eta - \int_{\mathbb{R}^d} W * \eta d\Pi(\mu) \right) d\Pi(\mu),$$

et pour $\eta = \Pi(\mu) - \mu$, on en déduit que :

$$\frac{1}{4} \frac{d}{dt} \mathcal{J}(\Phi_t(\mu)) = -\mathbb{V}_{\Pi(\mu)} [W * \eta] \leq 0.$$

Ainsi, $\forall c > 0, \{ \mu, \mathcal{J}(\mu) \leq c \}$ est positivement invariant. Ce qui prouve la faible compacité de ces ensembles, donc la non explosion du semi-flot. \square

On a ainsi défini un semi-flot lisse pour la topologie forte. Mais, pour étudier le comportement asymptotique de $(\mu_t')_{t \geq 0}$, il est techniquement plus facile de travailler avec la topologie faible. Par conséquent, on considérera également le semi-flot avec la topologie faible.

Proposition 16 (Semi-flot faible) Φ définit un semi-flot continu par rapport à la topologie faible.

Preuve

En effet, vu que $\mu \mapsto W * \mu(x)$ est faiblement continue d'après la propriété (P3) de domination, Π est faiblement continue aussi. Maintenant en reprenant le schéma d'approximation de Picard, on obtient que :

$$\forall n \geq 1 \quad t \geq 0, \quad \mu \mapsto \mu_t^{(n)} \quad \text{est faiblement continue,}$$

et on conclut en passant à la limite. \square

7.5 Énergie libre et points fixes

On définit l'énergie libre, comme dans McKean-Vlasov, par :

$$\mathcal{F} : \mathcal{M}_1(\mathbb{R}^d, V) \longrightarrow [0, +\infty], \quad \mu \longmapsto \begin{cases} \mathbf{H}[\mu|\gamma] + \mathbb{E}_{\mu^{\otimes 2}}[W], & \text{si } \mathbf{H}[\mu|\gamma] < +\infty, \\ +\infty, & \text{sinon.} \end{cases}$$

Cette fonctionnelle est donc obtenue comme somme d'une entropie et d'une énergie macroscopique induite par le potentiel d'interaction W . Par semi-continuité et convexité, cette énergie libre admet un unique argument minimum. Tout comme dans la partie précédente sur McKean-Vlasov, on a une relation liant les points fixes de l'application mesure invariante Π et les arguments minimum de l'énergie libre \mathcal{F} via le théorème suivant.

Théorème 7.3 (Énergie libre et points fixes)

Par symétrie du potentiel d'interaction W , les points fixes de Π sont les minima de \mathcal{F} .

Preuve

Le calcul fait précédemment de la différentielle de l'énergie libre donne :

$$\mathcal{D}_\mu \mathcal{F} \equiv 0 \iff \Pi(\mu) = \mu, \quad \text{ce qui prouve que les points critiques de l'énergie libre sont exactement}$$

les points fixes de Π . Il reste à prouver que ce sont bien des minima et pour ce faire, on va utiliser le critère de la hessienne vu que l'énergie libre est lisse. On a :

$$\mathcal{D}_\mu^{(2)}(\eta_1, \eta_2) = \int_{\mathbb{R}^d} \eta_1(x)\eta_2(x)\mu(x)^{-1}\gamma(x)dx + 2 \int_{\mathbb{R}^d \times \mathbb{R}^d} W(x, y)\eta_1(dx)\eta_2(dy) \geq 0,$$

ce qui permet de conclure par convexité. \square

Quant à l'unicité, on a le corollaire suivant.

Corollaire 4 (Unicité du minima) *De plus, si $\forall y \in \mathbb{R}^d, x \longmapsto V(x) + W(x, y)$ est strictement convexe, alors l'énergie libre admet un unique argument minimum μ_∞ tel que $\Phi_t(\mu) \xrightarrow{t \rightarrow +\infty} \mu_\infty$.*

Théorème 7.4 (Nature de l'ensemble des points fixes)

Si l'interaction est symétrique ou bornée en la seconde variable, alors l'ensemble des points fixes $\{\mu \in \mathcal{M}_1(\mathbb{R}^d, V), \quad \Pi(\mu) = \mu\}$ est une partie non vide faiblement compacte de $\mathcal{M}_1(\mathbb{R}^d, V)$.

Preuve

En effet, si W est borné en la variable y , c'est-à-dire $W(x, y) \leq \kappa V(x)$, alors pour :

$$\beta := \frac{\mathbb{E}_\gamma[V]}{\mathbb{E}_\gamma[e^{-2\kappa V}]}, \quad \text{d'après ce qui précède, l'application mesure invariante est faiblement continue de}$$

$\mathcal{M}_{1,\beta}(\mathbb{R}^d, V)$ partie compacte et convexe vers elle-même. Donc par le théorème du point fixe de Leray-Schauder, l'ensemble $\{\mu \in \mathcal{M}_{1,\beta}(\mathbb{R}^d, V), \quad \Pi(\mu) = \mu\}$ est non vide.

Supposons maintenant que W est symétrique et considérons l'énergie libre $\mathcal{I} = \mathcal{F} \circ \Pi$. Soit :

$$m := \inf_{\mu \in \mathcal{M}_1(\mathbb{R}^d, V)} \mathcal{I}(\mu). \quad \text{Il existe une suite de mesures de probabilité absolument continues par rapport à la}$$

mesure de Lebesgue de référence qu'on note (μ_n) telle que $0 \leq \mathcal{I}(\mu_n) - m \leq \frac{1}{n}$. Ainsi, $\forall c > 0$, l'ensemble $\{\mu, \mathcal{I}(\mu) \leq c\}$ est faiblement compact, on peut donc extraire une sous suite $(\mu_{\sigma(n)})$ convergeant faiblement vers μ_∞ . Par continuité faible de $\mu \mapsto W * \mu$ et Π, \mathcal{I} aussi, et on en déduit que $\mathcal{I}(\mu_\infty) = m$ puis on conclut par le théorème 7.3 sur les points fixes et minima. \square

7.6 Sur les mesures d'occupation

Dans cette section, on s'intéresse au comportement de $(\mu_t^r)_{t \geq 0}$ qui régit la dynamique à l'équilibre de la diffusion auto-attractive.

Théorème 7.5 (Tension presque sûrement)

$$\sup_{t \geq 0} \int V(x) \mu_t^r(dx) < +\infty, \quad \mathbb{P} - p.s.$$

Preuve

En considérant

$$\phi(t) := \int_0^t V(X_s) ds, \quad \text{il suffit de prouver que } \phi(t) = \mathcal{O}(t) \quad \mathbb{P} - p.s.$$

Pour ce faire, on réutilise la fonctionnelle de Lyapunov définie par $H_\mu := V + W * \mu$ et la formule d'Ito-Ventzell appliquée à $H_{\mu_t^r}(X_t)$. D'après la propriété (P3) de domination sur les potentiels, on a :

$$\forall \varepsilon > 0 \quad \exists \eta > 0, \quad |x| \geq \eta \implies V(x) + W(x, y) \leq \varepsilon \left(\|\nabla V(x)\|^2 + 2 \langle \nabla V(x), \nabla_x W(x, y) \rangle \right).$$

On en déduit que pour tout $\varepsilon > 0$, il existe k_ε telle que $H_{\mu_t^r}(x) \leq k_\varepsilon + \varepsilon \|\nabla H_{\mu_t^r}(x)\|^2$. D'une part, si :

$$\int_0^{+\infty} \|\nabla H_{\mu_s^r}(X_s)\|^2 ds < +\infty, \quad \text{presque sûrement, la loi forte des grands nombres pour les martingales}$$

assure que :

$$\int_0^{+\infty} \langle \nabla H_{\mu_s^r}(X_s), dB_s \rangle \quad \text{converge presque sûrement vers } M_\infty, \quad \text{et la preuve est similaire à celle ci-dessous.}$$

En effet, le théorème ergodique implique que ce cas ne se produit pas. D'autre part, si :

$$\int_0^{+\infty} \|\nabla H_{\mu_s^r}(X_s)\|^2 ds = +\infty, \quad \text{presque sûrement, et alors presque sûrement, il existe } T(\omega) \text{ telle que}$$

$$\forall t \geq T(\omega), \quad \int_0^t \langle \nabla H_{\mu_s^r}(X_s), dB_s \rangle \leq \frac{1}{2} \int_0^t \|\nabla H_{\mu_s^r}(X_s)\|^2 ds, \quad \text{et donc pour tout temps aléatoire assez large}$$

$$\int_0^t \|\nabla H_{\mu_s^r}(X_s)\|^2 ds \leq 2H_\mu(x) + \int_0^t \Delta H_{\mu_s^r}(X_s) ds + \frac{2}{r} \int_0^t W(X_s, X_s) ds.$$

Par la propriété de domination, on sait que :

$$W(X_s, X_s) \leq 2\kappa V(X_s) \leq 2\kappa \left(k_\varepsilon + \varepsilon \|\nabla H_{\mu_s^r}(X_s)\|^2 \right), \quad \text{et il en découle aussi que}$$

$$\Delta H_{\mu_s^r} \leq DH_{\mu_s^r} \leq D \left(k_\varepsilon + \varepsilon \|\nabla H_{\mu_s^r}(X_s)\|^2 \right). \quad \text{On en déduit l'inégalité suivante}$$

$$\left(1 - \varepsilon \left(D + \frac{4\kappa}{r} \right) \right) \int_0^t \|\nabla H_{\mu_s^r}(X_s)\|^2 ds \leq 2H_\mu(x) + k_\varepsilon t \left(D + \frac{4\kappa}{r} \right).$$

Pour $\varepsilon = \frac{\left(D + \frac{4\kappa}{r} \right)^{-1}}{2}$, on a le résultat attendu à savoir :

$$\exists C > 0, \quad \phi(t) \leq \int_0^t H_{\mu_s^r}(X_s) ds \leq C(t+1).$$

Ce qui achève la preuve de la tension presque sûre. \square

Chapitre 8

Schémas d'Euler et moyennisation

Pour plus de détails sur ce chapitre, consulter principalement les références [21], [93], [94], [98], [99] et [100] en bibliographie.

8.1 Introduction générale

Soit l'EDS de diffusion standard :

$$dX_t = b(t, X_t)dt + \sigma(t, X_t)dB_t$$

avec respectivement $b : \mathbb{R}_+ \times \mathbb{R}^d \rightarrow \mathbb{R}^d$ et $\sigma : \mathbb{R}_+ \times \mathbb{R}^d \rightarrow \mathbb{R}^{d \times d}$ les coefficients de dérive et de diffusion. On rappelle que $(B_t)_{t \geq 0}$ est un mouvement brownien d -dimensionnel standard. On rappelle aussi ce qu'on entend par solution de l'EDS :

Définition 8.1 (Solution forte)

Une solution forte de l'EDS ci-dessus issue de $x \in \mathbb{R}^d$ sur un intervalle de temps de longueur $T > 0$ est un processus vectoriel $(X_t)_{0 \leq t \leq T}$ adapté à la filtration brownienne tel que

1. presque sûrement :

$$\int_0^T \|b(t, X_t)\| + \|\sigma(t, X_t)\|^2 dt < +\infty$$

2. $\forall t \in [0, T]$, presque sûrement :

$$X_t = x + \int_0^t b(s, X_s)ds + \int_0^t \sigma(s, X_s)dB_s$$

On a la condition suivante suffisante d'existence et d'unicité d'une telle solution :

Proposition 17 (Existence et unicité) Si $\exists K > 0$ telle que $\forall t \in [0, T]$ et $(x, y) \in \mathbb{R}^d \times \mathbb{R}^d$:

1. $\|b(t, x) - b(t, y)\| + \|\sigma(t, x) - \sigma(t, y)\| \leq K\|x - y\|$;
2. $\|b(t, x)\| + \|\sigma(t, x)\| \leq K(1 + \|x\|)$;

alors l'EDS ci-dessus admet une unique solution forte $(X_t)_{t \in [0, T]}$ telle que :

$$\mathbb{E} \left[\sup_{t \in [0, T]} \|X_t\|^2 \right] < +\infty$$

Soit le problème de Cauchy :

$$y'(t) = F(t, y(t)), \quad y(0) = y_0$$

On sait via le théorème de Cauchy-Lipschitz que sous des conditions adéquates sur F , il y a existence et unicité d'une solution maximale. Quand une telle solution n'est pas explicite, on peut construire un schéma d'approximation sur $[0, T]$ dit d'Euler comme suit :

1. On se donne une subdivision (ici régulière par exemple) de $[0, T]$ dont les points $(t_k)_{k=0,1,\dots,n}$ sont tels que $\forall k, t_k := \frac{kT}{n}$.
2. On approxime la solution aux points de cette subdivision, c'est-à-dire $\hat{y}(0) = y_0$ et $\forall k \geq 1$:

$$\hat{y}(t_k) = \hat{y}(t_{k-1}) + \frac{T}{n} F(t_{k-1}, \hat{y}(t_{k-1}))$$

3. On approxime la solution sur $[0, T]$ par le processus linéaire par morceaux passant par les points $(t_k, \hat{y}(t_k))_{k=0,1,\dots,n}$.

Sous des hypothèses très faibles sur la régularité de F , on a :

$$\|y - \hat{y}\|_\infty \xrightarrow{n \rightarrow +\infty} 0$$

Quant aux EDS, c'est essentiellement la même idée que pour les EDO. En effet :

1. On se donne une subdivision (ici régulière par exemple) de $[0, T]$ dont les points $(t_k)_{k=0,1,\dots,n}$ sont tels que $\forall k, t_k := \frac{kT}{n}$.
2. On approxime la solution aux points de cette subdivision, c'est-à-dire $\hat{X}_0^n = x$ et $\forall k \geq 1$:

$$\hat{X}_{t_k}^n = \hat{X}_{t_{k-1}}^n + \frac{T}{n} b(t_{k-1}, \hat{X}_{t_{k-1}}^n) + \sigma(t_{k-1}, \hat{X}_{t_{k-1}}^n)(B_{t_k} - B_{t_{k-1}})$$

3. On approxime la solution sur $[0, T]$ par le processus linéaire par morceaux passant par les points $(t_k, \hat{X}_{t_k}^n)_{k=0,1,\dots,n}$, c'est-à-dire $\forall t \in [t_k, t_{k+1}]$:

$$\hat{X}_t^n = \hat{X}_{t_k}^n + (t - t_k)b(t_k, \hat{X}_{t_k}^n) + \sigma(t_k, \hat{X}_{t_k}^n)(B_t - B_{t_k})$$

Et la mise en oeuvre pratique du schéma d'Euler est très simple. En effet, il suffit de générer n vecteurs gaussiens indépendants $(G_k)_{k=1,\dots,n}$ tels que $G_k \sim \mathcal{N}(0, \mathbf{Id}_d)$. Par définition du processus de Wiener standard, on a :

$$B_{t_k} - B_{t_{k-1}} \sim \sqrt{t_k - t_{k-1}} G_k$$

On a le critère suivant de convergence forte du schéma d'Euler :

Proposition 18 (Convergence forte) Si $\exists K > 0$ et $\alpha > 0$ telles que $\forall (t, s) \in [0, T]^2$ et $x \in \mathbb{R}^d$:

$$\|b(t, x) - b(s, x)\| + \|\sigma(t, x) - \sigma(s, x)\| \leq K|t - s|^\alpha(1 + \|x\|)$$

alors pour $\beta := \min(\alpha, \frac{1}{2})$, on a $\forall p \geq 1 \exists C > 0 \forall n \in \mathbb{N}^*$:

$$\mathbb{E} \left[\sup_{t \in [0, T]} \|\hat{X}_t^n - X_t\|^{2p} \right] \leq \frac{C}{n^{2\beta p}}$$

Ainsi, $\forall \gamma < \beta$, presque sûrement :

$$n^\gamma \sup_{t \in [0, T]} \|\hat{X}_t^n - X_t\| \xrightarrow{n \rightarrow +\infty} 0$$

8.2 Stabilité et martingales

Dans cette section, on rappelle quelques résultats de stabilité de processus de diffusions homogènes et la méthode de martingales permettant d'étudier le comportement asymptotique de tels processus. Dans ce qui suit, on note $(P_t)_{t \geq 0}$ le semi-groupe de Markov associé à ce processus et \mathcal{L} son générateur infinitésimal. On a vu dans la première partie de ce mémoire au chapitre processus de Markov qu'une condition suffisante d'existence d'une solution pour l'EDS associée est que ses coefficients soient localement lipschitziens et qu'il existe une fonction de Lyapunov U telle que :

$$\mathcal{L}U \leq cU, \quad c > 0.$$

Définition 8.2 (Condition de stabilité)

Une diffusion de générateur \mathcal{L} satisfait une condition de stabilité s'il existe $U \in \mathcal{C}^2(\mathbb{R}^d, [1, +\infty[)$ de Lyapunov et $\phi \in \mathcal{C}^0([1, +\infty[, \mathbb{R}_+)$ telles que :

1. $\forall u \geq 1, \phi(u) \leq u$ et $\lim_{u \rightarrow +\infty} \phi(u) = +\infty$;
2. $\exists \alpha > 0$ et $\beta \geq 0$ telles que $\mathcal{L}U \leq -\alpha \cdot \phi \circ U + \beta$.

Si $\forall u, \phi(u) = u$, on dit que la condition de stabilité est forte et faible sinon.

Par abus, on dira par la suite que \mathcal{L} vérifie une condition de stabilité de type (U, ϕ, α) . Lorsque la connaissance du α n'est pas importante, on parlera de condition de stabilité de type (U, ϕ) . Le fait que $\alpha > 0$ est primordial.

Remarques :

Une condition de stabilité pour \mathcal{L} implique l'existence d'une mesure invariante. En général, $\phi(u) = u^a$ avec $a \in]0, 1]$ ou $\phi(u) = \log(u)$. Dans le premier cas, on pourra consulter la référence [99] en bibliographie.

On énonce le théorème suivant qui permet d'avoir une bonne connaissance du comportement du processus de diffusion de générateur \mathcal{L} sous l'unique condition de stabilité.

Théorème 8.1

On suppose que le générateur \mathcal{L} vérifie une condition de stabilité de type (U, ϕ) . Alors :

1. $\forall \Theta \in \mathcal{C}^1(\mathbb{R}_+,]0, +\infty[)$ décroissante et telle que $\int_0^{+\infty} \Theta(u) du < +\infty$, on a :

$$\int_0^{+\infty} \Theta(s) P_s \phi \circ U ds \stackrel{\text{Fubini}}{=} \mathbb{E} \left[\int_0^{+\infty} \Theta(s) \cdot \phi \circ U(X_s) ds \right] < +\infty,$$

et :

$$\lim_{t \rightarrow +\infty} \Theta(t) U(X_t) = 0 \quad \mathbb{P} - p.s.$$

2. si la condition de stabilité est forte, on a :

$$\sup_{t \geq 0} P_t U = \sup_{t \geq 0} \mathbb{E} [U(X_t)] < +\infty,$$

ce qui assure la tension du flot des distributions de $(X_t)_{t \geq 0}$, donc l'existence d'une sous-suite convergente par le théorème de Prokhorov.

Quant à la preuve de ce théorème, c'est une conséquence de la convergence de surmartingales positives, des lemmes de Kronecker, Fatou et de la formule d'Ito. On voit bien que la condition de

stabilité seule ne suffit pas pour avoir des résultats de tension des lois; il faut ajouter des hypothèses de domination du coefficient de diffusion et du gradient de la fonction de Lyapunov dans la direction de σ , c'est-à-dire qu'on va supposer une domination de $\mathbf{Tr}(\sigma^*(\nabla U)^{\otimes 2}\sigma)$.

Proposition 19 (Transfert de stabilité) *Supposons que \mathcal{L} vérifie une condition de stabilité de type (U, ϕ, α) . S'il existe $\xi \in [0, 1]$ et $C_{U,\sigma} > 0$ telles que :*

$$\mathbf{Tr}(\sigma^*(\nabla U)^{\otimes 2}\sigma) \leq C_{U,\sigma} U^{1-\xi} \phi \circ U,$$

alors \mathcal{L} vérifie une condition de stabilité de type $(\Psi \circ U, \phi_\Psi)$, avec :

$$\forall u \geq 1, \quad \Psi(u) = \begin{cases} u^p \text{ avec } p-1 \in]0, \frac{2\alpha}{C_{U,\sigma}}[, & \text{si } \xi = 0 \\ e^{\lambda u^\xi} \text{ avec } \lambda < \frac{2\alpha}{C_{U,\sigma}}, & \text{si } \xi \in]0, 1] \end{cases}, \quad \phi_\Psi(u) = u \left(\frac{\phi \circ \Psi^{-1}}{\Psi} \right)(u).$$

Et Ψ^{-1} est l'application réciproque de Ψ .

Cette proposition est très importante. En effet, le fait que la condition de stabilité se "transfère" à des fonctions plus grandes (polynomiales, exponentielles...) sous une condition de domination sur le coefficient de diffusion est fondamental. D'après la proposition précédente, on a :

$$\frac{e^{\lambda U(X_t)}}{t \log^{1+\varepsilon}(t)} \xrightarrow{t \rightarrow +\infty} 0 \quad \mathbb{P} - p.s.$$

Et ce type de convergence permet d'avoir une meilleure connaissance du comportement du processus mais aussi de ses mesures d'occupation. La combinaison des conditions de stabilité et de domination permet d'établir la tension des mesures d'occupation du processus.

En fait, la stabilité seule suffit à prouver la tension des mesures d'occupations et donc l'existence d'une mesure invariante. Quant à la condition supplémentaire de domination sur le coefficient de diffusion, elle permet de montrer que toute valeur d'adhérence de la famille des mesures d'occupation est une probabilité¹ sous laquelle on peut intégrer des fonctions à croissance polynomiale² voire exponentielle. Plus précisément, on a le résultat suivant :

Corollaire 5 (Sur les valeurs d'adhérence)

$$\begin{cases} \forall p < \frac{1}{2} + \frac{\alpha}{C_{U,\sigma}}, & \overline{\lim}_{t \rightarrow +\infty} \frac{1}{t} \int_0^t (U^{p-1} \phi \circ U)(X_s) ds \leq C \quad \mathbb{P} - p.s. \quad \text{si } \xi = 0, \\ \forall \lambda < \frac{\alpha}{\xi C_{U,\sigma}}, & \overline{\lim}_{t \rightarrow +\infty} \frac{1}{t} \int_0^t (e^{\lambda U^\xi} \frac{\phi \circ U}{U})(X_s) ds \leq C \quad \mathbb{P} - p.s. \quad \text{si } \xi \in]0, 1]. \end{cases}$$

Observons aussi que le paramètre ξ sert de jauge pour le coefficient de diffusion en ce sens que plus il diminue, plus σ peut avoir une croissance importante.

On va maintenant rappeler le théorème d'Echeverria-Weiss qui est très important pour caractériser les mesures invariantes. Soit E un espace métrique séparable et

$$\mathcal{C}_0(E) := \left\{ f \in \mathcal{C}^0(E), \quad \forall \varepsilon > 0 \quad \exists K \subset E \text{ compact tel que } \forall x \in K^c, \quad |f(x)| \leq \varepsilon \right\}$$

1. Il est bien connu que les valeurs d'adhérence des mesures d'occupation d'un tel processus sont des probabilités stationnaires.

2. On dit aussi à croissance lente.

l'espace des fonctions continues sur E nulles à l'infini. On rappelle que le dual topologique de cet espace³ est l'ensemble des mesures de Radon et qu'un opérateur $\mathfrak{A} : \mathcal{D}(\mathfrak{A}) \subset \mathcal{C}_0(E) \rightarrow \mathcal{C}_0(E)$ vérifie le principe du maximum positif si :

$$\forall \varphi \in \mathcal{D}(\mathfrak{A}), \quad \varphi(x_0) = \sup_{x \in E} \varphi(x) \implies \mathfrak{A}\varphi \leq 0.$$

On va énoncer deux versions du théorème qui nous intéresse à savoir lorsque E est compact et le cas général de locale compacité.

Théorème 8.2 (Echeverria-Weiss dans le cas compact)

Soient E un espace métrique compact séparable, $\mathfrak{A} : \mathcal{D}(\mathfrak{A}) \subset \mathcal{C}(E) \rightarrow \mathcal{C}(E)$ un opérateur linéaire de domaine une sous-algèbre unitaire dense vérifiant le principe du maximum positif et tel que $\mathfrak{A}\mathbb{1} = 0$. Alors, si $\mu \in \mathcal{P}(E)$ est telle que **image**(\mathfrak{A}) $\perp \mathbb{1}$ dans $L^2(\mu)$, il existe une solution stationnaire au problème de martingale associé à (\mathfrak{A}, μ) , c'est-à-dire un processus $(X_t)_{t \geq 0}$ tel que :

$$\forall \varphi \in \mathcal{D}(\mathfrak{A}) \quad t \geq 0, \quad M_t^\varphi := \varphi(X_t) - \int_0^t \mathfrak{A}\varphi(X_s) ds \text{ définit une martingale relativement à la filtration canonique de } (X_t)_{t \geq 0}, \text{ et de loi initiale } \mu.$$

Le passage du cas compact au cas localement compact⁴ ne pose pas de problème mis à part qu'il faut vérifier que la solution n'explose pas⁵.

Théorème 8.3 (Echeverria-Weiss dans le cas général)

Soient E un espace métrique localement compact séparable, $\mathfrak{A} : \mathcal{D}(\mathfrak{A}) \subset \mathcal{C}_0(E) \rightarrow \mathcal{C}_0(E)$ un opérateur linéaire de domaine une sous-algèbre dense vérifiant le principe du maximum positif. S'il existe une suite de fonctions strictement positives $(U_n)_{n \geq 0}$ telle que :

1. $(U_n)_{n \geq 0}$ croît vers une fonction $U > 0$ telle que pour toute suite $(K_n)_{n \geq 1}$ de compacts tendant vers E , on a $\lim_{n \rightarrow +\infty} \inf_{x \notin K_n} U(x) = +\infty$;
2. $\forall n \geq 0, \exists K_n \subset E$ compact tel que $\forall x \in K_n^c, U_n(x) = C_n$, et $U_n - C_n \in \mathcal{D}(\mathfrak{A})$;
3. $\exists c > 0, \forall n \geq 0, \mathfrak{A}(U_n - C_n) \leq cU_n$;

et s'il existe $\mu \in \mathcal{P}(E)$ telle que $U \in L^1(\mu)$ et **image**(\mathfrak{A}) $\perp \mathbb{1}$ dans $L^2(\mu)$, alors il existe une solution stationnaire au problème de martingale pour (\mathfrak{A}, μ) .

Proposition 20 (Caractérisation des mesures invariantes) Soit $(P_t)_{t \geq 0}$ un semi-groupe de diffusion sur \mathbb{R}^d de générateur \mathcal{L} . S'il existe une fonction de Lyapunov $U \in \mathcal{C}^2(\mathbb{R}^d)$ telle que $\mathcal{L}U \leq cU$ avec $c > 0$ et une mesure de probabilité μ telle que $U \in L^1(\mu)$ et $\forall \varphi \in \mathcal{C}_c^2(\mathbb{R}^d), \mathbb{E}_\mu[\mathcal{L}\varphi] = 0$, alors μ est une mesure invariante pour $(P_t)_{t \geq 0}$, c'est-à-dire $\forall t \geq 0, \mu P_t = \mu$.

3. $\mathcal{C}_c^0(E) \subset \mathcal{C}_0(E) \subset \mathcal{C}_b^0(E) \subset L^\infty(E)$.

4. Qui est notre cadre car on considère des équations sur \mathbb{R}^d .

5. $\mathbb{P}\left(\bigcap_{t \geq 0} \{X_t \in \mathbb{R}^d\}\right) = 1$.

Preuve

Il s'agit de vérifier les hypothèses du théorème d'Echeverria-Weiss dans le cas localement compact. Ici, $E = \mathbb{R}^d$, $\mathfrak{A} = \mathcal{L}$ et $\mathcal{D}(\mathfrak{A}) = \mathcal{C}_c^2(\mathbb{R}^d)$. Cet opérateur de diffusion est linéaire et vérifie le principe du maximum positif en tant qu'opérateur différentiel elliptique du second ordre. De plus, son domaine est dense dans $\mathcal{C}_0(\mathbb{R}^d)$ par densité de l'espace des fonctions tests via le théorème d'approximation de l'unité. Soit $(\Psi_n)_{n \geq 1}$ la suite de fonctions telle que :

$$\forall n \geq 1 \quad x \geq 1, \Psi_n(x) := \begin{cases} x & \text{si } x \in [1, n], \\ x - (x - n)^3 + \frac{1}{2}(x - n)^4 & \text{si } x \in [n, n + 1], \\ n + \frac{1}{2} & \text{si } x \geq n + 1. \end{cases}$$

Il est aisé de voir que $\Psi_n \in \mathcal{C}^2$, concave ($\Psi_n'' \leq 0$) et $\text{image}(\Psi_n') \subset [0, 1]$. Posons $U_n = \Psi_n \circ U$ de sorte que la suite $(U_n)_{n \geq 1}$ est croissante et tend vers U . $\forall n \geq 1$, par la formule d'Ito, on a :

$$\mathcal{L}U_n = \Psi_n'(U)\mathcal{L}U + \Psi_n''(U)(\nabla U)^{\otimes 2} \leq c\Psi_n'(U)U.$$

Vu que $\forall x \geq 1, x\Psi_n'(x) \leq \Psi_n(x)$, on en déduit que $\mathcal{L}U_n \leq cU_n$. De plus, la fonction $U_n - n - 1$ est de classe \mathcal{C}^2 , à support compact et $\mathcal{L}(U_n - n - 1) = \mathcal{L}U_n \leq cU_n$. Ce qui montre qu'on a bel et bien les hypothèses du théorème d'Echeverria-Weiss. \square

Le grand intérêt du théorème d'Echeverria-Weiss est qu'il fournit une caractérisation des mesures invariantes plus aisée à manipuler en pratique que celle via le semi-groupe. En effet, la caractérisation de telles mesures par le semi-groupe est pénible, par exemple, quand on étudie les mesures empiriques⁶ du schéma d'Euler à pas décroissant.

L'étude par les méthodes de martingales du comportement ergodique d'un processus de diffusion donne de nombreuses informations sur le comportement asymptotique de ce processus ainsi que sur ses mesures d'occupation. Ces techniques permettent de retrouver un théorème ergodique en faisant abstraction du théorème ergodique ponctuel. De plus, on obtient un résultat de convergence même en présence de plus d'une mesure invariante⁷.

Théorème 8.4 (Partielle ergodicité)

Supposons que \mathcal{L} vérifie une condition de stabilité de type (U, ϕ, α) et une condition de domination de son coefficient de diffusion de type $(\xi, C_{U, \sigma})$. Alors l'ensemble \mathcal{I} des mesures invariantes de $(P_t)_{t \geq 0}$ est non vide, et en notant \mathbf{d}_e la distance de la convergence étroite, on a :

$$\lim_{t \rightarrow +\infty} \mathbf{d}_e\left(\frac{1}{t} \int_0^t \delta_{X_s} ds, \mathcal{I}\right) = 0 \quad \mathbb{P}. - p.s.$$

De plus, si $\mathcal{I} = \{\mu\}$, alors pour toute fonction mesurable φ telle que $|\varphi| = o\left(\frac{\phi \circ U}{U}(\sqrt{\Psi} \circ U)\right)$:

$$\frac{1}{t} \int_0^t \varphi(X_s) ds \xrightarrow{t \rightarrow +\infty} \langle \mu, \varphi \rangle \quad \mathbb{P}. - p.s.$$

Preuve

D'après ce qui précède, en notant (ν_t) la famille des mesures d'occupation, on a :

$$\limsup_{t \rightarrow +\infty} \left\langle \nu_t, \frac{\phi \circ U}{U}(\sqrt{\Psi} \circ U) \right\rangle < +\infty \quad \mathbb{P}. - p.s.$$

6. Ergodicité du schéma.

7. Situation qui peut avoir lieu si le coefficient de diffusion est dégénéré.

Et donc, elle est tendue et par le théorème de Prokhorov, on a relative compacité de cette famille pour la topologie de la convergence étroite (faible). Soit ν_∞ une valeur d'adhérence de cette famille et $(\nu_{a(t)})_{t \geq 0}$ une sous suite convergeant étroitement vers cette valeur d'adhérence, alors :

$$\forall \varphi \in \mathcal{C}_c^2(\mathbb{R}^d), \quad \frac{1}{a(t)} \int_0^{a(t)} \mathcal{L} \varphi(X'_s) ds = \frac{\varphi(X'_{a(t)}) - \varphi(X'_0)}{a(t)} - \frac{1}{a(t)} \int_0^{a(t)} \nabla \varphi(X'_s) \cdot \sigma(X'_s) dB_s,$$

et comme φ est bornée, $\lim_{t \rightarrow +\infty} \frac{\varphi(X'_{a(t)}) - \varphi(X'_0)}{a(t)} = 0$. Et par propriété de martingale⁸, on en déduit que :

$$\int_{\mathbb{R}^d} \mathcal{L} \varphi(x) \nu_\infty(dx) = \lim_{t \rightarrow +\infty} \frac{1}{a(t)} \int_0^{a(t)} \nabla \varphi(X'_s) \cdot \sigma(X'_s) dB_s = 0.$$

Et par le théorème d'Echeverria-Weiss, ν_∞ est une mesure invariante. L'unicité de la mesure invariante implique la convergence étroite de ν_t vers μ . \square

Théorème 8.5 (Nature des valeurs d'adhérence des mesure d'occupation)

Si les mesures d'occupation (ν_t) sont tendues, alors tout candidat limite est une mesure invariante pour $(P_t)_{t \geq 0}$.

Preuve

Soit $t_0 > 0$ fixé et $(\mathcal{G}_k)_{k \geq 0}$ la filtration définie $\forall k \geq 0$ par $\mathcal{G}_k := \sigma(X'_s, s \leq kt_0)$. Soit $\varphi \in \mathcal{C}_c^0(\mathbb{R}^d, \mathbb{R})$. Montrons dans un premier temps que :

$$\frac{1}{t} \int_0^t \varphi(X'_s) ds - \frac{1}{t} \int_0^t P_{t_0} \varphi(X'_s) ds \xrightarrow{t \rightarrow +\infty} 0 \quad \mathbb{P} - p.s.$$

Pour ce faire, on considère la variable aléatoire :

$$Z_k := \int_{(k-1)t_0}^{kt_0} \varphi(X'_s) ds, \quad \mathcal{G}_k - \text{mesurable, de sorte que} \quad \frac{1}{nt_0} \int_0^{nt_0} \varphi(X'_s) ds = \frac{1}{nt_0} \sum_{k=2}^n Z_k.$$

Par convergence de la martingale :

$$\left(\sum_{k=2}^n \frac{1}{kt_0} (Z_k - \mathbb{E}[Z_k | \mathcal{G}_{k-1}]) \right)_{n \geq 0}$$

et le lemme de Kronecker, on montre que :

$$\lim_{n \rightarrow +\infty} \frac{1}{nt_0} \sum_{k=2}^n (Z_k - \mathbb{E}[Z_k | \mathcal{G}_{k-1}]) = 0 \quad \mathbb{P} - p.s.$$

De même, en réitérant cet argument, on a :

$$\lim_{n \rightarrow +\infty} \frac{1}{nt_0} \sum_{k=2}^n (Z_k - \mathbb{E}[Z_k | \mathcal{G}_{k-2}]) = 0 \quad \mathbb{P} - p.s.$$

Or :

$$\forall k \geq 2, \quad \mathbb{E}[Z_k | \mathcal{G}_{k-2}] = \int_{(k-1)t_0}^{kt_0} \mathbb{E}[\varphi(X'_s) | \mathcal{G}_{k-2}] ds, \quad \text{et} \quad \forall s \in [(k-1)t_0, kt_0], \quad \mathcal{G}_{k-2} \subset \mathcal{F}_{s-t_0}.$$

Il en découle que :

$$\mathbb{E}[\varphi(X'_s) | \mathcal{G}_{k-2}] = \mathbb{E} \left[\mathbb{E}[\varphi(X'_s) | \mathcal{F}_{s-t_0}] \middle| \mathcal{G}_{k-2} \right] = \mathbb{E}[P_{t_0} \varphi(X'_{s-t_0}) | \mathcal{G}_{k-2}].$$

On en déduit que :

$$\mathbb{E}[Z_k | \mathcal{G}_{k-2}] = \mathbb{E} \left[\int_{(k-1)t_0}^{kt_0} P_{t_0} \varphi(X'_{s-t_0}) ds \middle| \mathcal{G}_{k-2} \right] = \mathbb{E} \left[\int_{(k-2)t_0}^{(k-1)t_0} P_{t_0} \varphi(X'_s) ds \middle| \mathcal{G}_{k-1} \right],$$

8. Convergence vers zéro du crochet de la martingale et application du lemme de Kronecker.

et on montre comme précédemment que :

$$\frac{1}{nt_0} \sum_{k=2}^n \int_{(k-1)t_0}^{kt_0} P_{t_0} \varphi(X_s) ds - \mathbb{E}[Z_k | \mathcal{G}_{k-2}] \xrightarrow{n \rightarrow +\infty} 0 \quad \mathbb{P}. - p.s.$$

Par synthèse de ce qui précède, on en tire que :

$$\frac{1}{nt_0} \int_{t_0}^{nt_0} \varphi(X_s) ds - \frac{1}{nt_0} \int_{t_0}^{nt_0} P_{t_0} \varphi(X_s) ds \xrightarrow{t \rightarrow +\infty} 0 \quad \mathbb{P}. - p.s.$$

et l'on déduit aisément le résultat attendu (φ étant continue à support compact, elle est bornée). On conclut en prenant une suite extraite $(\nu_{a(t)})_{t \geq 0}$ qui converge étroitement vers une mesure ν_∞ vérifiant alors :

$$\forall t_0 \geq 0, \quad \int_{\mathbb{R}^d} \varphi(x) \nu_\infty(dx) = \int_{\mathbb{R}^d} P_{t_0} \varphi(x) \nu_\infty(dx),$$

ce qui prouve que cette valeur d'adhérence est invariante pour le semi-groupe. \square

8.3 Discrétisation

Soit $(\gamma_n)_{n \geq 1}$ une suite de réels positifs décroissante et tendant vers 0 appelée SUITE DE PAS telle que :

$$\sum_{n \geq 1} \gamma_n = +\infty, \quad \text{on parle de terme général d'une série divergente et on note } (\Gamma_n)_{n \geq 0}$$

la suite des sommes partielles avec $\Gamma_0 = 0$. $(\Gamma_n)_{n \geq 0}$ permet de discrétiser le temps, et cette discrétisation est de plus en plus fine comme la suite de pas tend vers 0. Il n'est en général pas nécessaire de supposer que la suite de pas est décroissante pour l'étude du schéma $(X_n)_{n \geq 0}$ construit comme suit :

$$X_{n+1} = X_n + \gamma_{n+1} b(X_n) + \sqrt{\gamma_{n+1}} \sigma(X_n) U_{n+1}, \quad X_0 = x \in \mathbb{R}^d,$$

et la suite $(U_n)_{n \geq 1}$ est une suite de vecteurs aléatoires indépendants, centrés et de matrice de covariance \mathbf{Id} . La suite $(\sqrt{\gamma_n} U_n)_{n \geq 1}$ sert à discrétiser le brownien $(B_t)_{t \geq 0}$ entre les instants Γ_n et Γ_{n+1} , donc on pourrait supposer que $U_1 \sim \mathcal{N}(0, \mathbf{Id})$. On parle souvent de bruit blanc.

On rappelle la formule de Taylor :

$$\varphi(x+h) = \varphi(x) + \langle \nabla \varphi(x), h \rangle + \frac{1}{2} \langle \nabla^2 \varphi(\xi_{x,h}) h, h \rangle, \quad \xi_{x,h} \in \{tx + (1-t)h, \quad t \in [0, 1]\},$$

ou encore :

$$\varphi(x+h) = \sum_{j=0}^n \frac{1}{j!} \mathcal{D}_x^{(j)} \varphi(h^{\otimes j}) + \mathbf{o}(\|h\|^n).$$

Introduisons maintenant les mesures empiriques pondérées du schéma $(X_n)_{n \geq 0}$. On se donne une suite $(\eta_n)_{n \geq 1}$ de réels positifs telle que

$$H_n := \sum_{j=1}^n \eta_j \xrightarrow{n \rightarrow +\infty} +\infty, \quad \text{et on définit la suite des mesures empiriques du schéma associée par}$$

$$\nu_n^\eta := \frac{1}{H_n} \sum_{j=1}^n \eta_j \delta_{X_j}, \quad \text{et on remarque que la suite } (\eta_n)_{n \geq 1} \text{ n'est pas supposée tendre vers 0.}$$

Et un choix naturel est $\eta_n = 1$ ou $\eta_n = \gamma_n$. $(\eta_n)_{n \geq 1}$ est appelée suite de poids associée au schéma d'Euler. Le fait d'étudier la convergence de $(\nu_n^\eta)_{n \geq 1}$ pour toute une famille de suites de poids est motivé par l'étude des vitesses de convergence qui permet d'obtenir une suite de poids optimale.

Revenons maintenant à notre diffusion auto-attractrice. On définit $(\bar{X}_t)_{t \geq 0}$ le schéma d'Euler associé au processus auto-attractif par :

$$\bar{X}_0 = x \in \mathbb{R}^d, \forall n \geq 0, \quad \bar{X}_{\Gamma_{n+1}} = \bar{X}_{\Gamma_n} - \gamma_{n+1} \left(\nabla V(\bar{X}_{\Gamma_n}) + \frac{1}{\Gamma_n + r} \int_0^{\Gamma_n} \nabla W_0(\bar{X}_{\Gamma_n} - \bar{X}_{\underline{s}}) ds \right) + \sigma \sqrt{\gamma_{n+1}} U_{n+1},$$

avec $\underline{s} := \max\{\Gamma_n, \Gamma_n \leq s\}$. Finalement, sur l'intervalle $[\Gamma_n, \Gamma_{n+1})$, on définit \bar{X}_s de différentes façons mais principalement par interpolation. Et on définit les mesures empiriques pondérées de la façon suivante :

$$\bar{\nu}_n := \frac{1}{H_n} \sum_{j=1}^n \eta_j \delta_{\bar{X}_{\Gamma_{j-1}}}.$$

Quant à la discrétisation du processus de McKean-Vlasov, de façon analogue, on définit son schéma par :

$$Y_{n+1} = Y_n - \gamma_{n+1} \left(\nabla V(Y_n) + \frac{1}{\Gamma_n} \sum_{j=0}^{n-1} \gamma_{j+1} \nabla_x W(Y_n, Y_j) \right) + \sqrt{2\gamma_{n+1}} U_{n+1},$$

et on définit aussi de façon similaire la suite des mesures empiriques pondérées qu'on note $(\bar{\mu}_n)_n$. Pour $\eta_n = \gamma_n$, on a naturellement :

$$\bar{\mu}_{n+1} = \bar{\mu}_n + \frac{\gamma_{n+1}}{\Gamma_n} (\delta_{Y_{n+1}} - \bar{\mu}_n), \quad \bar{\nu}_{n+1} = \bar{\nu}_n + \frac{\gamma_{n+1}}{\Gamma_n} (\delta_{\bar{X}_{n+1}} - \bar{\nu}_n), \quad \text{même pour toute suite de poids,}$$

ce qui définit des algorithmes stochastiques dont les suites de pas sont données par $(\frac{\gamma_{n+1}}{\Gamma_n})_{n \geq 1}$, et donc l'implémentation numérique aisée à réaliser. Le coût de calcul (complexité) est de l'ordre de n^2 avec n l'horizon de temps. Ces mesures empiriques pondérées représentent en quelque sorte des versions discrétisées des mesures d'occupation, donc leurs valeurs d'adhérence sont les mesures invariantes de nos processus respectifs⁹.

Une autre façon d'approcher le régime stationnaire dans le cas spécifique à McKean-Vlasov est de considérer la suite des schémas d'Euler donnée par :

$$\forall N \geq 2, \quad Y_{n+1}^N = Y_n^N - \gamma_{n+1} \nabla H_N(Y_n^N) + \sqrt{2\gamma_{n+1}} U_{n+1}^N,$$

avec U_{n+1}^N représentant l'adjonction de N copies de U_{n+1} qui sont indépendantes. Cette suite est une discrétisation du système à champ moyen associé au processus de McKean-Vlasov, donc elle permet une approximation de la mesure invariante μ^N de ce système.

9. Non des schémas parce que les approximations ne sont pas biaisées.

8.4 Python

Dans ce qui suit, tous les codes sont supposés précédés de :

```

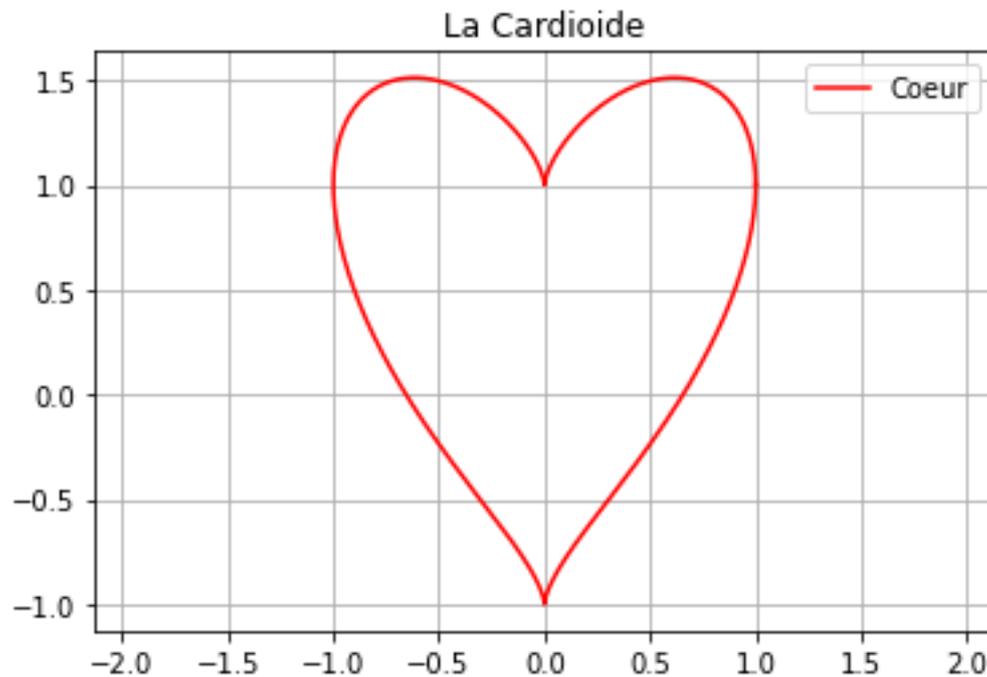
1 import numpy as np
2 import matplotlib as mpl
3 from matplotlib import pyplot as plt
4 from mpl_toolkits.mplot3d import Axes3D
5 import scipy.linalg as lg
6 import math as m
7 import random

```

Listing 8.1 – Préambule python nécessaire.

8.4.1 Quelques rappels

On donne ci-dessous un exemple de représentation en 2D sous python.



```

1 def Card():
2     theta=np.linspace(0,2*np.pi,100000)
3     X=np.cos(theta)
4     Y=np.sin(theta)+(X**2)**(1/3)
5     plt.plot(X,Y,'red',label='Coeur')
6     plt.legend()
7     plt.title("La Cardioïde")
8     plt.grid()
9     plt.axis('equal')

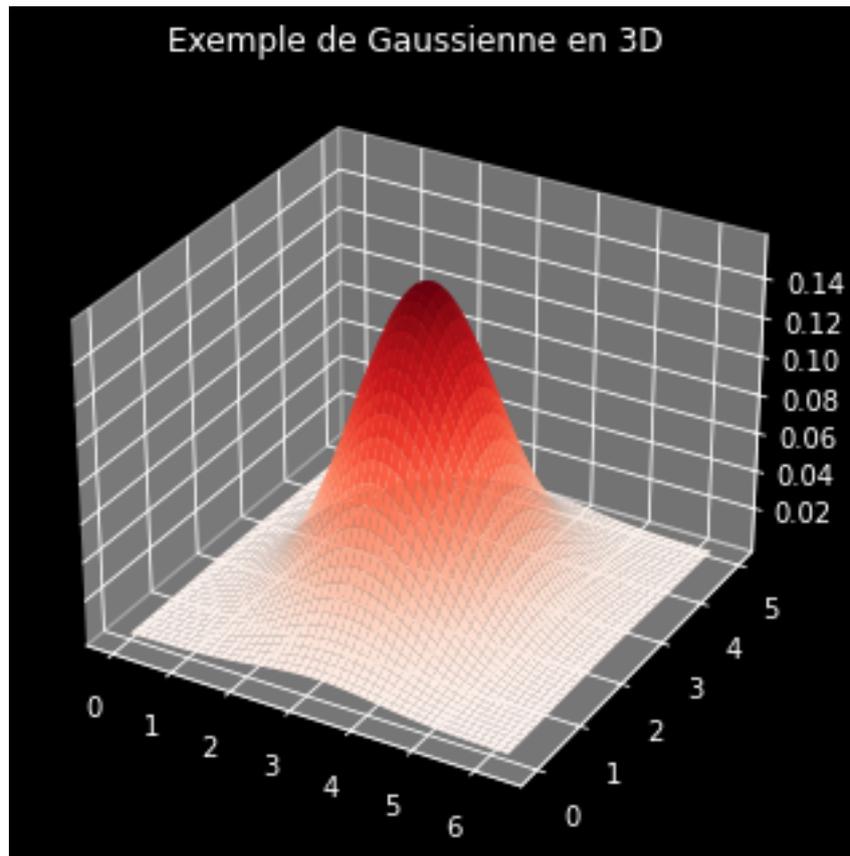
```

Listing 8.2 – Code python donnant la cardioïde.

La cardioïde est une courbe algébrique plane, trajectoire d'un point fixé à un cercle qui roule sans glisser sur un second cercle de même diamètre. Il s'agit donc d'une courbe cycloïdale dont la directrice est un cercle (ou épicycloïde). Son équation est :

$$x^2 + (y - x^{\frac{2}{3}})^2 = 1.$$

Maintenant, on donne trois exemples de représentation en 3D sous python.

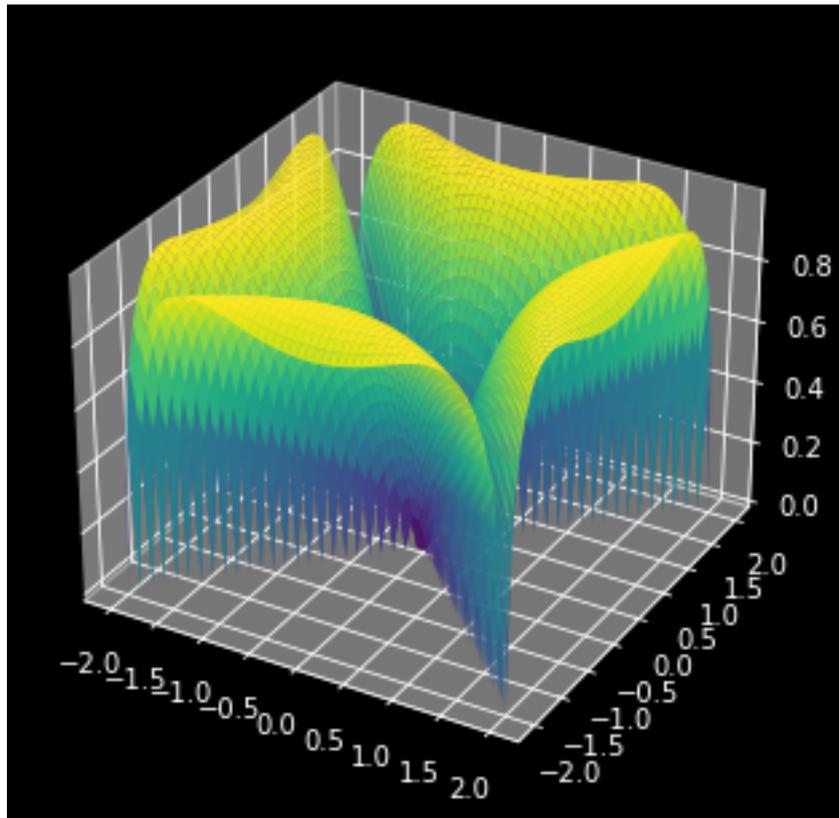


```

1 def gamma(x, y, mu_x, mu_y, sigma):
2     return (np.exp(-(np.power(x - mu_x, 2) + np.power(y - mu_y, 2))
3             / (2*sigma**2))
4             / (2 * np.pi * sigma**2))
5
6
7 def Gauss3D():
8     with plt.style.context(('dark_background')):
9         x = np.linspace(0, 2*np.pi)
10        y = np.linspace(0, 5)
11        v_x, v_y= np.meshgrid(x, y)
12        v_z = gamma(v_x, v_y, np.pi, 2.5, 1)
13        fig = plt.figure()
14        ax = Axes3D(fig)
15        ax.plot_surface(v_x, v_y, v_z, cmap='Reds')
16        plt.title('Exemple de Gaussienne en 3D')

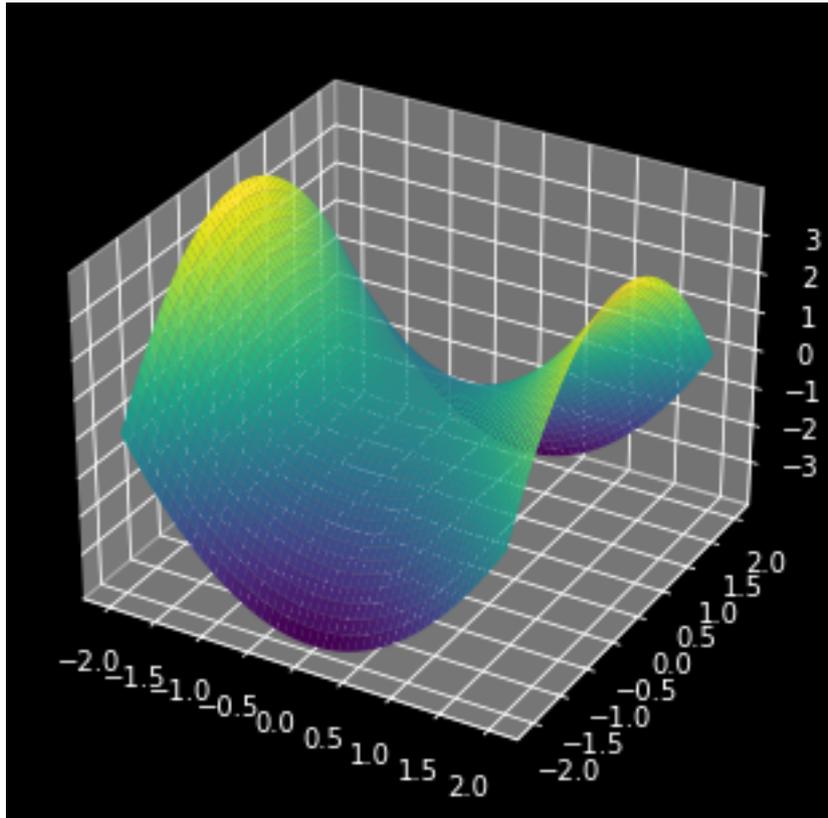
```

Listing 8.3 – Code python donnant une gaussienne en 3D.



```
1 def autre(x, y):
2     return np.sin(np.sqrt(np.abs(np.power(x,2)-np.power(y,2))))
3 def fig3D():
4     x = np.linspace(-2, 2, 50)
5     X, Y = np.meshgrid(x, x)
6     Z = autre(X, Y)
7     with plt.style.context(('dark_background')):
8         fig = plt.figure()
9         ax = Axes3D(fig)
10        ax.plot_surface(X, Y, Z, cmap='viridis')
```

Listing 8.4 – Code pour représentation 3D second exemple.



```

1 def col(x, y):
2     return np.power(x, 2) - np.power(y, 2)
3
4 def fig3D():
5     x = np.linspace(-2, 2, 50)
6     X, Y = np.meshgrid(x, x)
7     Z = col(X, Y)
8     with plt.style.context(('dark_background')):
9         fig = plt.figure()
10        ax = Axes3D(fig)
11        ax.plot_surface(X, Y, Z, cmap='viridis');

```

Listing 8.5 – Code python pour représentation 3D troisième exemple.

Pour générer un vecteur gaussien numériquement sur python, on peut utiliser le code suivant :

```

1 def G(d, esp, sigma):
2     x=np.ones(d)
3     for i in range(d):
4         x[i]=random.gauss(esp[i],m.sqrt(sigma[i,i]))
5     return x

```

Listing 8.6 – Fonction python générant un vecteur gaussien

Maintenant, on va illustrer quelques propriétés du schéma d'Euler pour les EDO dans le cas particulier de l'oscillateur harmonique. On sait que quand on étudie une méthode numérique, on s'intéresse à sa convergence¹⁰, sa stabilité et sa consistance. On rappelle que l'oscillateur harmonique est régi par l'EDO :

$$\frac{d}{dt}Y(t) = \begin{pmatrix} 0 & 1 \\ -\omega^2 & 0 \end{pmatrix} Y(t), \quad \text{avec } \omega \text{ la pulsation.}$$

10. Étude de l'erreur et vitesse de convergence.

Et cette équation différentielle est très facile à résoudre via l'exponentielle de sa matrice qu'on peut calculer, par exemple, par la décomposition de Dunford-Jordan ¹¹.

```

1 def pas_euler(systeme,h,tn,Yn):
2     deriv = systeme(Yn,tn)
3     return Yn+h*deriv
4
5
6 def euler(systeme,Yi,T,h):
7     Y = Yi
8     t = 0.0
9     liste_y = [Y]
10    liste_t = [t]
11    while t<T:
12        Y = pas_euler(systeme,h,t,Y)
13        t += h
14        liste_y.append(Y)
15        liste_t.append(t)
16    return (np.array(liste_t),np.array(liste_y))
17
18 def euler_bis(systeme,Yi,T,h):
19    Y = Yi
20    t = 0.0
21    tableau_y = np.array([Y])
22    tableau_t = np.array([t])
23    while t<T:
24        Y = pas_euler(systeme,h,t,Y)
25        t += h
26        tableau_y = np.append(tableau_y,[Y],axis=0)
27        tableau_t = np.append(tableau_t,t)
28    return (tableau_t,tableau_y)

```

Listing 8.7 – Codes pour le schéma d'Euler explicite.

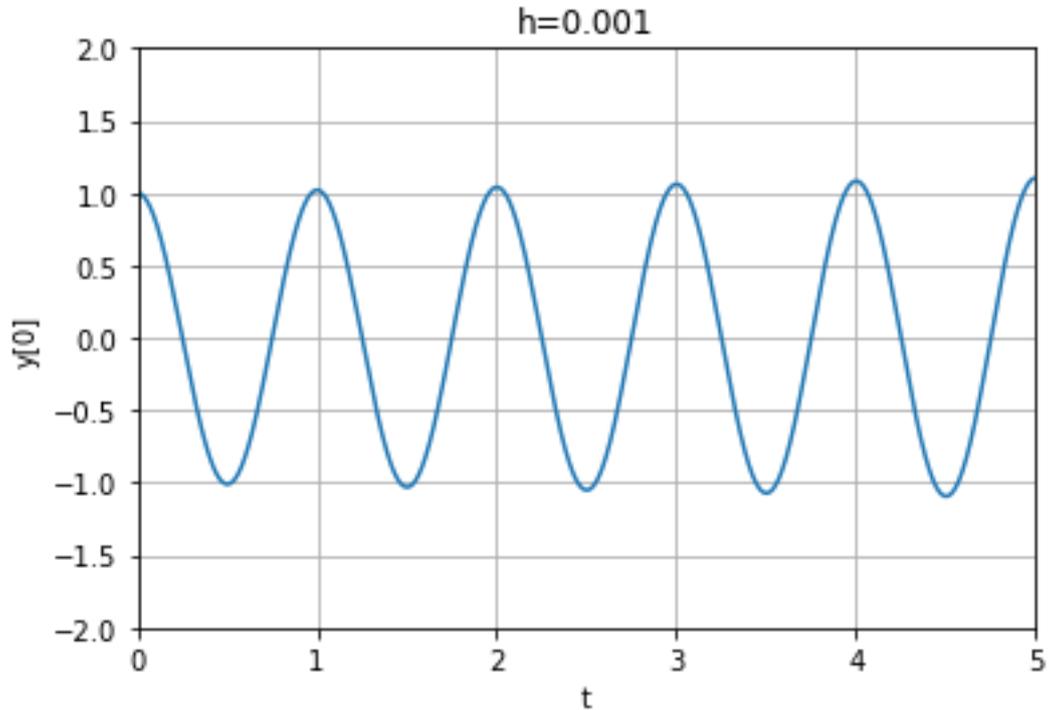
```

1 w2 = (2*np.pi)**2
2 def oscillateur(Y,t):
3     return np.array([Y[1],-w2*Y[0]])

```

Listing 8.8 – Pulsation pour exemples et définition de l'oscillateur harmonique associé.

¹¹. Réduction d'endomorphismes.



```

1 T = 5.0
2 h = 1.0e-3
3 Yi = [1.0,0]
4 (t,tab_y) = euler(oscillateur ,Yi ,T,h)
5 x = tab_y[:,0]
6 plt.figure(figsize=(6,4))
7 plt.plot(t,x)
8 plt.xlabel('t')
9 plt.ylabel('y[0]')
10 plt.axis([0,T,-2,2])
11 plt.title('h=0.001')
12 plt.grid()

```

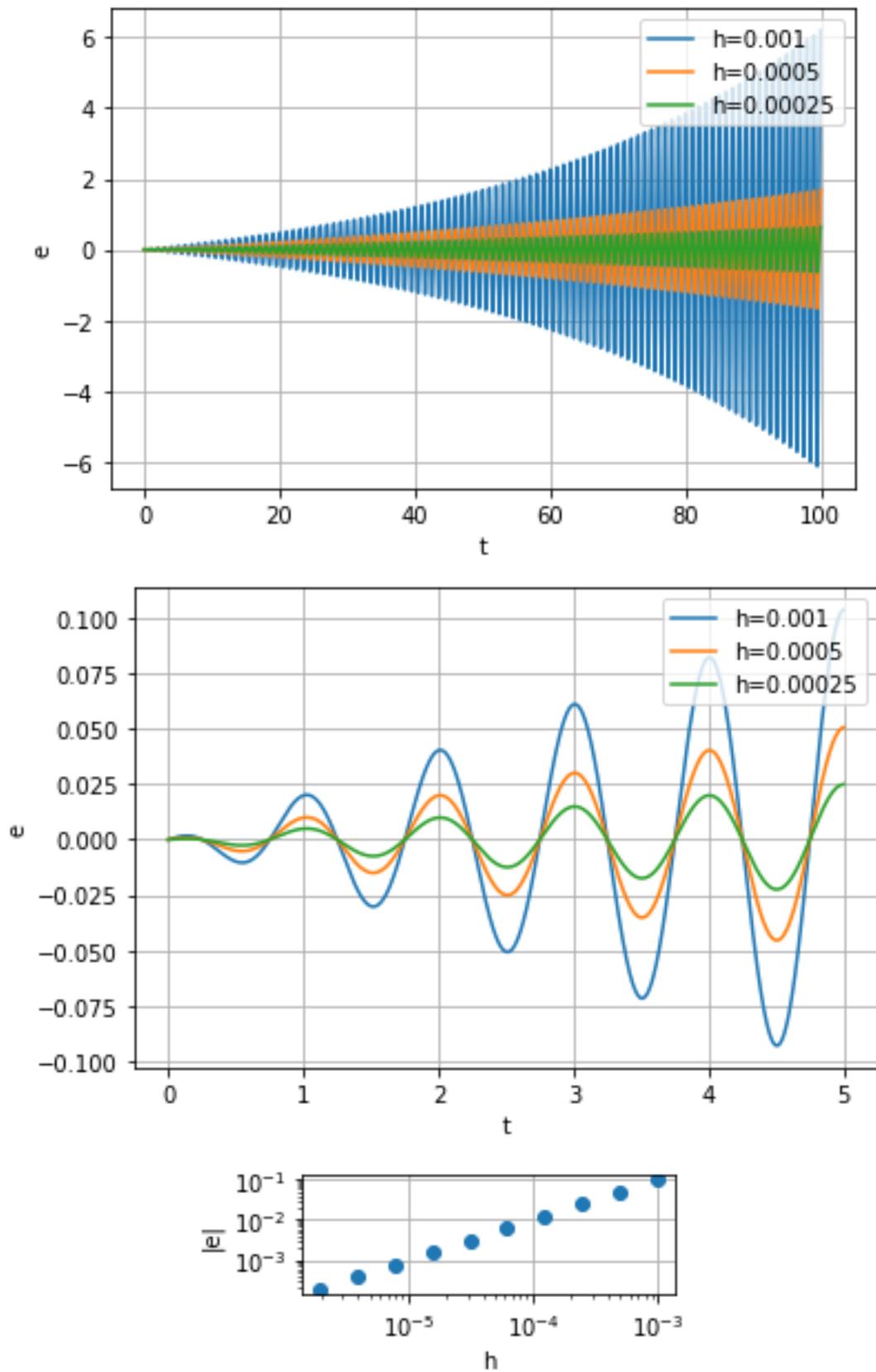
Listing 8.9 – Exemple de représentation de l'oscillateur harmonique.

```

1 def erreur_euler(systeme ,solution ,Yi ,T,h):
2     Y = Yi
3     t = 0.0
4     while t<T:
5         Y = pas_euler(systeme ,h,t,Y)
6         t += h
7     return abs(Y[0]-solution(t))

```

Listing 8.10 – Code python donnant l'erreur.



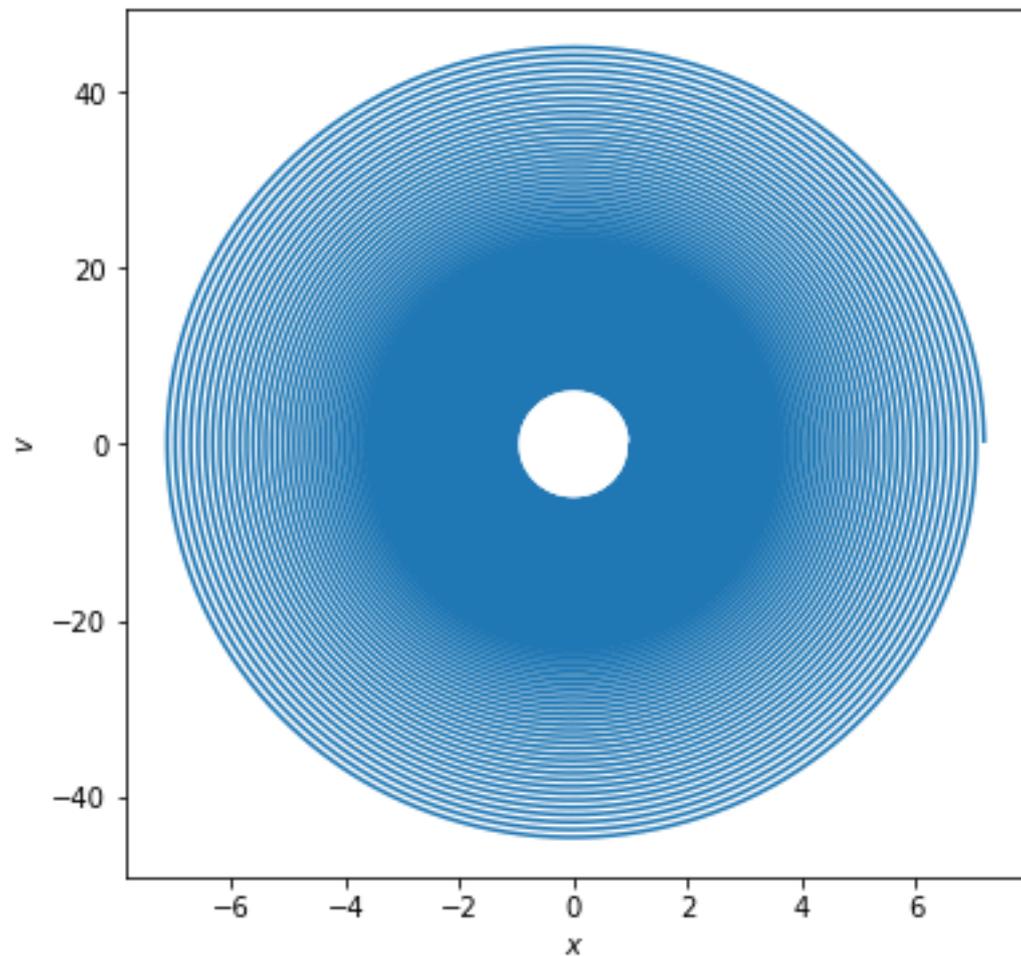
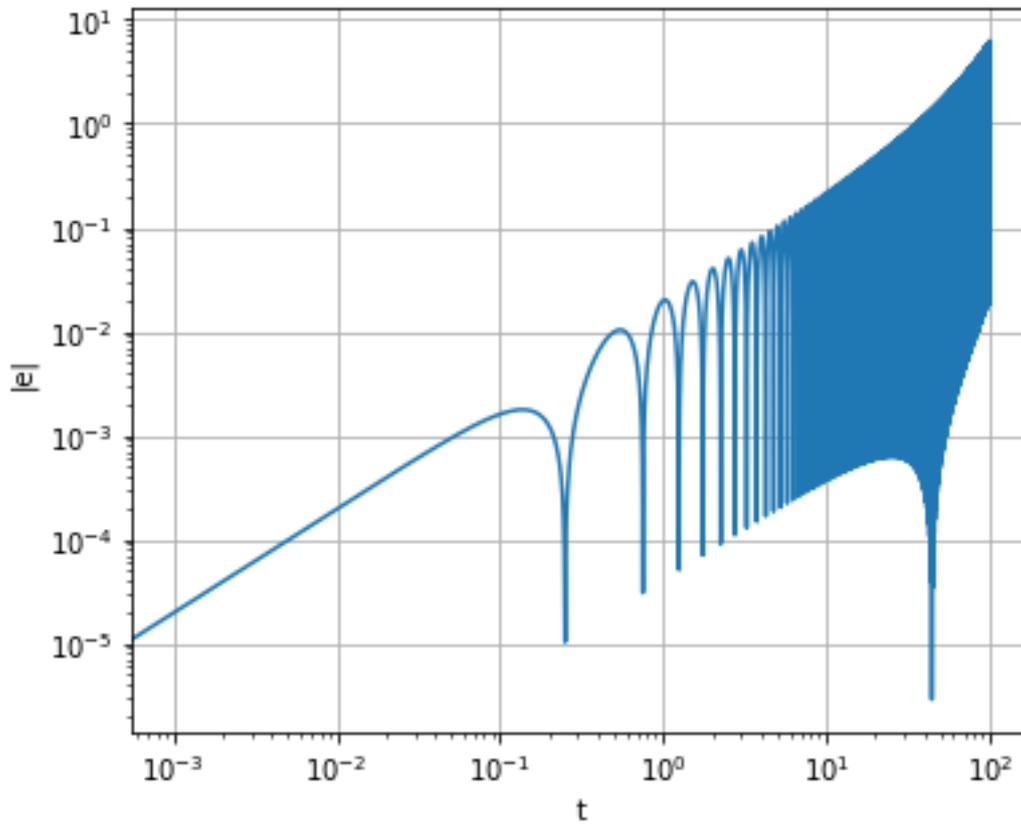
Dans la page suivante, figure le code correspondant à ces graphiques.

```

1  erreur = x-np.cos(2*np.pi*t)
2  plt.figure(figsize=(6,4))
3  plt.plot(t,erreur,label='h=0.001')
4  plt.xlabel('t')
5  plt.ylabel('e')
6  plt.legend()
7  plt.grid()
8
9  h = 1.0e-3
10 h = h/2
11 (t2,tab_y) = euler(oscillateur,Yi,T,h)
12 yt = tab_y.transpose()
13 x = yt[0]
14 erreur2 = x-np.cos(2*np.pi*t2)
15 plt.plot(t2,erreur2,label='h=0.0005')
16 h = h/2
17 (t3,tab_y) = euler(oscillateur,Yi,T,h)
18 x = tab_y[:,0]
19 erreur3 = x.copy()
20 for i in range(x.size):
21     erreur3[i] = x[i]-np.cos(2*np.pi*t3[i])
22 plt.plot(t3,erreur3,label='h=0.00025')
23 plt.legend(loc='upper right')
24
25 liste_h = []
26 liste_e = []
27 T = 5.0
28 h = 0.001
29 def solution(t):
30     return np.cos(2*np.pi*t)
31 for i in range(10):
32     liste_h.append(h)
33     liste_e.append(erreur_euler(oscillateur,solution,Yi,T,h))
34     h = h/2
35 plt.figure(figsize=(3,1))
36 plt.plot(liste_h,list_e,'o')
37 plt.xlabel('h')
38 plt.ylabel('|e|')
39 plt.xscale('log')
40 plt.yscale('log')
41 plt.grid()

```

Listing 8.11 – Code donnant l'évolution de l'erreur par rapport au pas.



Le code correspondant à ces graphiques est dans la page qui suit.

```
1 T = 100.0
2 h = 1.0e-3
3 (t,tab_y) = euler(oscillateur,Yi,T,h)
4 x = tab_y[:,0]
5 erreurbis= np.absolute(x-np.cos(2*np.pi*t))
6 plt.figure(figsize=(6,5))
7 plt.plot(t,erreurbis)
8 plt.xlabel('t')
9 plt.ylabel('|e|')
10 plt.yscale('log')
11 plt.xscale('log')
12 plt.grid()
13
14 v = tab_y[:,1]
15 plt.figure(figsize=(6,6))
16 plt.plot(x,v)
17 plt.xlabel('$x$')
18 plt.ylabel('$v$')
```

Listing 8.12 – Code relatif à l’instabilité de l’oscillateur harmonique.

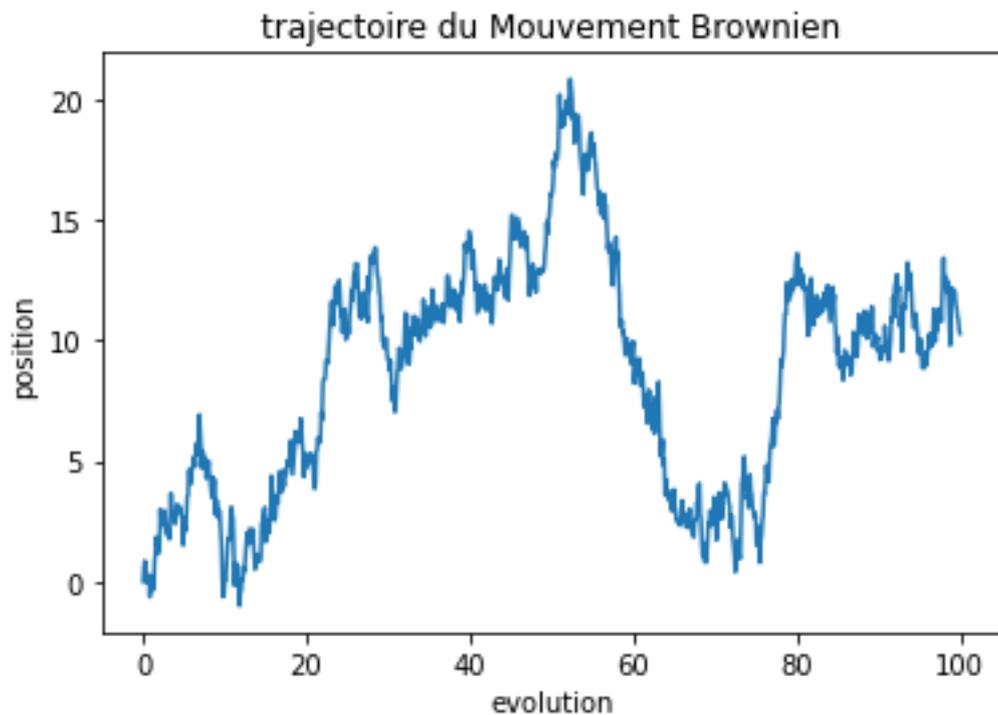
8.4.2 Schéma numérique et exemples

```

1 def Points(x,b,sigma,T,n):
2     t0=[k*T/n for k in range(n+1)]
3     Y0=x
4     X=[x]
5     for i in range(1,n+1):
6         Y0=Y0+(T/n)*b(t0[i-1],Y0)+np.dot(sigma(t0[i-1],Y0),G(np.size(x)
7             ,np.zeros(np.size(x)),np.sqrt(T/n)*np.eye(np.size(x))))
8     X=X+[Y0]
9     plt.plot(t0,X)
10    return X

```

Listing 8.13 – Exemple de code pour le schéma d'Euler.

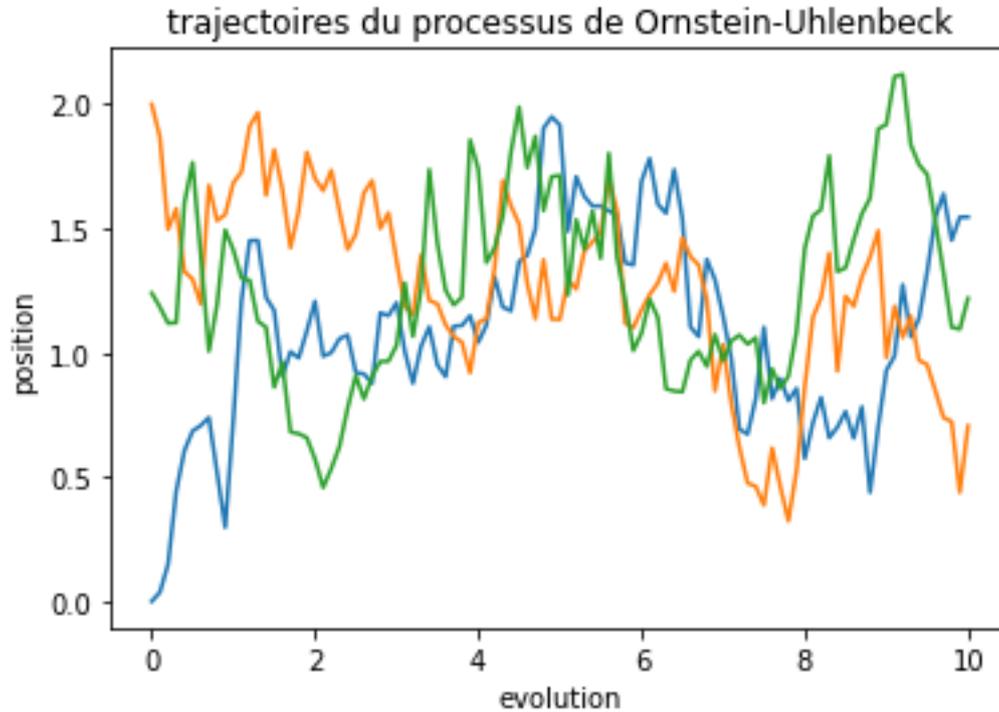


```

1 def drift1(t,x):
2     return 0
3
4 def diffusion1(t,x):
5     return 1
6
7 def fig2D():
8     Points(0,drift1,diffusion1,100,1000)
9     plt.xlabel('evolution')
10    plt.ylabel('position')
11    plt.title('trajectoire du Mouvement Brownien')

```

Listing 8.14 – Code simulant une trajectoire du mouvement brownien standard.



```

1 def drift2(t,x):
2     return 1.2-x
3
4 def diffusion2(t,x):
5     return 0.3
6
7 def fig2D():
8     Points(0,drift2,diffusion2,10,100)
9     Points(2,drift2,diffusion2,10,100)
10    Points(random.gauss(1.2,(0.3**2)/2),drift2,diffusion2,10,100)
11    plt.title('trajectoires du processus de Ornstein-Uhlenbeck')
12    plt.xlabel('evolution')
13    plt.ylabel('position')

```

Listing 8.15 – Code simulant trois trajectoires du processus d’Ornstein-Uhlenbeck.

8.4.3 McKean-Vlasov

```

1 def H(N,V,W,x):
2     H=0
3     for i in range(N):
4         for j in range(N):
5             if i!=j:
6                 H=H+W(x[i],x[j])
7     H0=1/(2*(N-1))*H
8     return sum(V(x[i]) for i in range(N))+H0

```

Listing 8.16 – Code définissant l’hamiltonien associé au système à champ moyen

On donne ci-dessous des exemples de potentiels V et W , le premier exemple donne un système de particules sans confinement et le second un système de particules de type Ornstein-Uhlenbeck.

```

1 def confinement1(x):
2     return 0
3
4 def interaction1(x,y):
5     return x*y
6
7 def drift3(t,x):
8     N=1000
9     y=np.array(x)
10    return 1/(N-1)*(y-np.dot(y,np.ones(N))*np.ones(N))
11 def diffusion3(t,x):
12    return np.sqrt(2)
13
14 def confinement2(x):
15    return np.dot(x,x)
16 def interaction2(x,y):
17    return confinement2(x)+confinement2(y)
18
19 def drift4(t,x):
20    return -4*np.array(x)
21 def diffusion4(t,x):
22    return np.sqrt(2)

```

Listing 8.17 – Exemples de potentiels.

Et remarquons que le second exemple vérifie toutes les hypothèses émises sur les potentiels.

Preuve

En effet, pour $V(x) = \|x\|^2$ et $W(x,y) = V(x) + V(y)$, il est évident que **(H1)**, **(H2)** et **(H3)** sont vérifiées vu que :

$\nabla V(x) = 2x$, $\nabla^2 V(x) = 2\mathbf{Id}$, $\nabla^2 W(x,y) = 2\mathbf{Id}$, $\forall \lambda > 0$, on a par Fubini-Tonelli et passage en polaire

$$\int_{\mathbb{R}^d \times \mathbb{R}^d} e^{-[V(x)+V(y)+\lambda W(x,y)]} dx dy = \left(\frac{\pi}{\lambda+1}\right)^d < +\infty, \text{ et de plus, on a } b_0(r) = -4r, \quad c_{Lip,m} = \frac{1}{4} < +\infty,$$

et :

$$H_N(x) = 2 \sum_{j=1}^N V(x_j) = 2\|x\|^2, \quad \nabla H_N(x) = 4x, \quad \mathcal{L}^N = \Delta - \nabla H_N \cdot \nabla = \Delta - 4x \cdot \nabla, \text{ ce qui montre qu'on a bien}$$

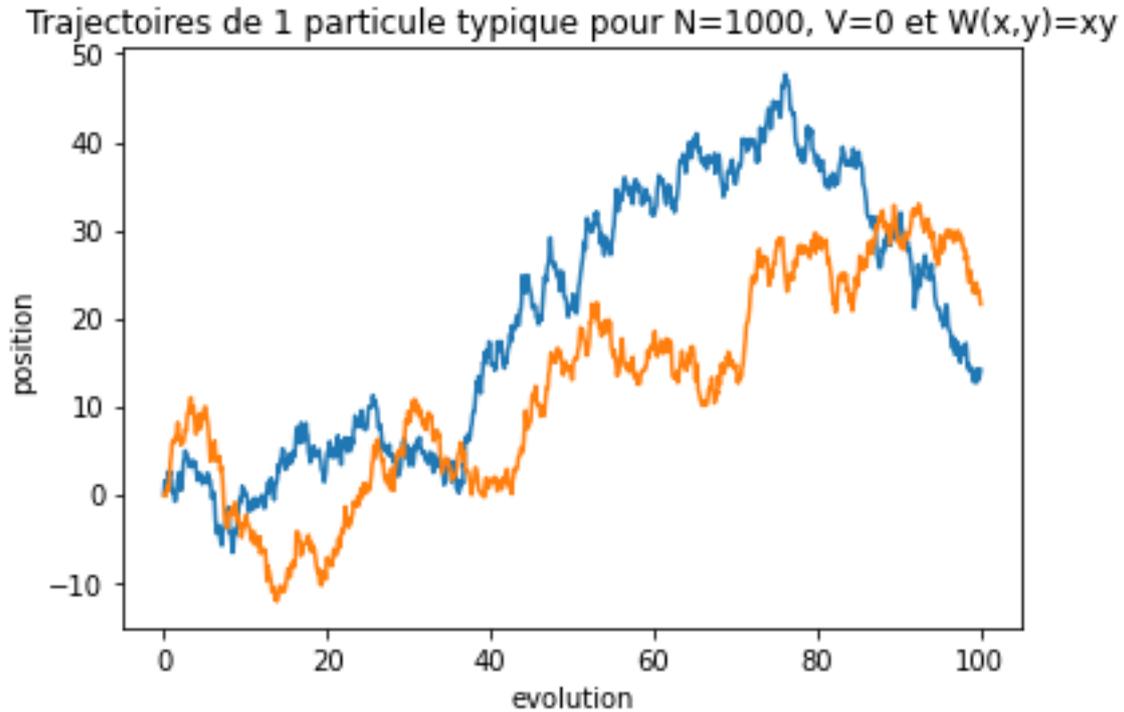
un système de particules régi par un processus d'Ornstein-Uhlenbeck. \square

```

1 def Points1(x,b,sigma,T,n):
2     t0=[k*T/n for k in range(n+1)]
3     Y0=x
4     X=[x]
5     for i in range(1,n+1):
6         Y0=Y0+(T/n)*b(t0[i-1],Y0)+np.dot(sigma(t0[i-1],Y0),G(np.size(x),
7             np.zeros(np.size(x)),np.sqrt(T/n)*np.eye(np.size(x))))
8         X=X+[Y0]
9     plt.plot(t0,[X[i][1] for i in range(len(X))])
10    return X

```

Listing 8.18 – Code pour approximation de McKean-Vlasov.

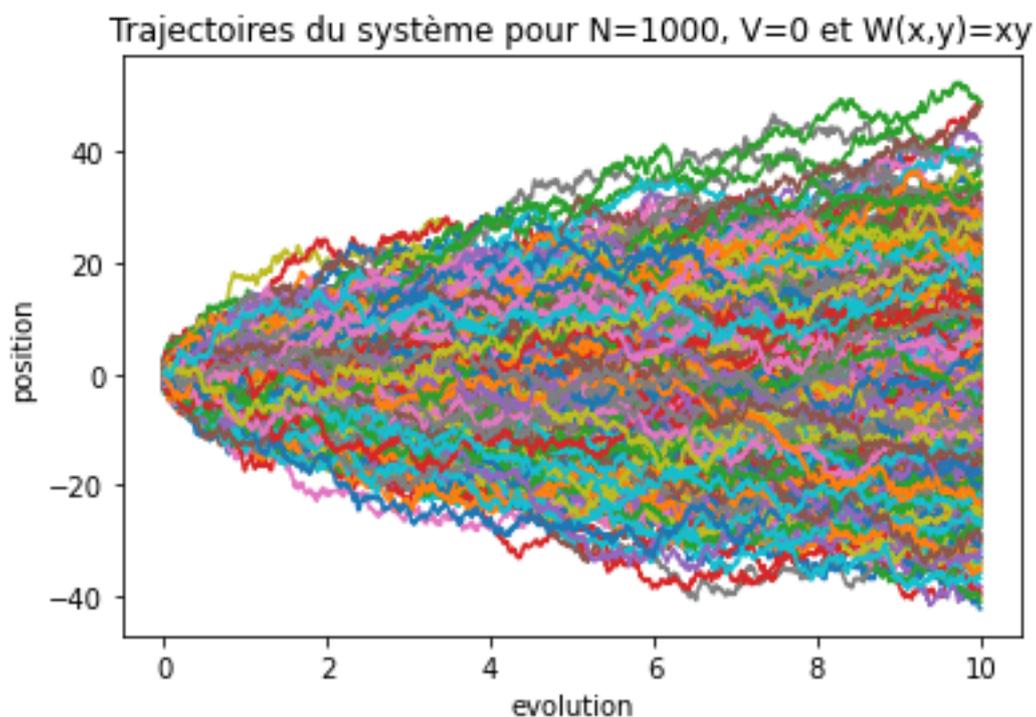


```

1 def fig():
2     Points1(1/1000*np.ones(1000),drift3,diffusion3,100,1000)
3     Points1(G(1000,np.zeros(1000),np.eye(1000)),drift3,diffusion3,100,1
4         000)
5     plt.title('Trajectoires de 1 particule typique pour N=1000, V=0 et
6         W(x,y)=xy')
7     plt.xlabel('evolution')
8     plt.ylabel('position')

```

Listing 8.19 – Code simulant deux trajectoires d'une particule typique du système.



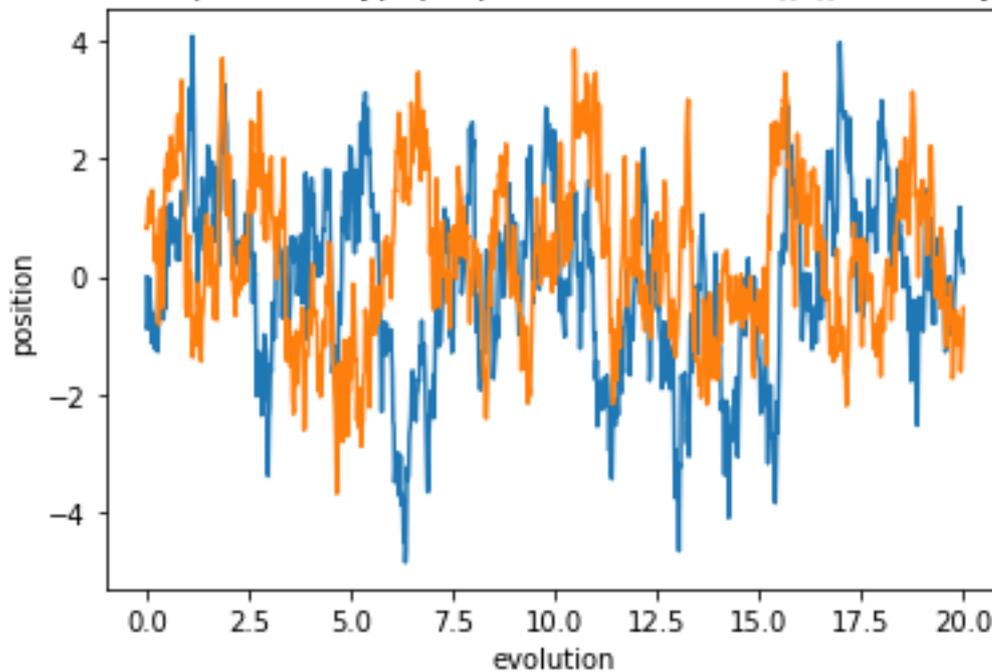
```

1 def fig():
2     Points(G(1000,np.zeros(1000),np.eye(1000)),drift3,diffusion3,10,100
3         0)
4     plt.title('Trajectoires du systeme pour N=1000, V=0 et W(x,y)=xy')
5     plt.xlabel('evolution')
6     plt.ylabel('position')

```

Listing 8.20 – Code simulant des trajectoires des particules du système.

Trajectoires de 1 particule typique pour $N=1000$, $V = ||.||^2$ et $W(x,y)=V(x)+V(y)$

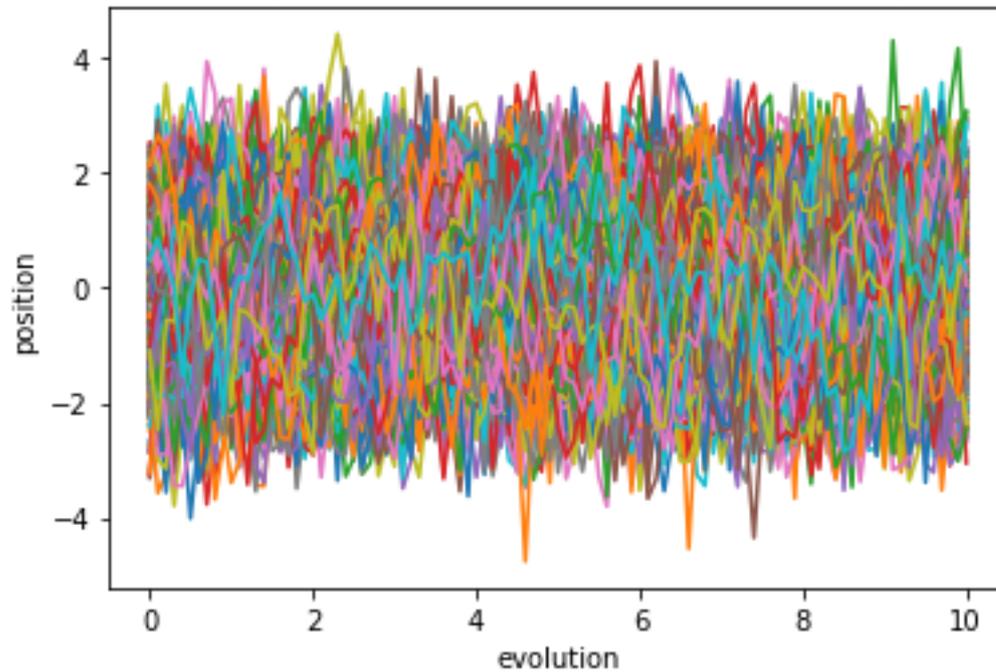


```

1 def fig():
2     Points1(1/1000*np.ones(1000),drift4,diffusion4,20,1000)
3     Points1(G(1000,np.zeros(1000),np.eye(1000)),drift4,diffusion4,20,10
4         00)
5     plt.title('Trajectoires de 1 particule typique pour N=1000, $V
6         =||.||^2$ et W(x,y)=V(x)+V(y)')
7     plt.xlabel('evolution')
8     plt.ylabel('position')

```

Listing 8.21 – Second exemple de simulation des trajectoires d'une particule typique.

Trajectoires du système pour $N=1000$, $V=||.||^2$ et $W(x,y)=V(x)+V(y)$ 

```

1 def fig():
2     Points(G(1000,np.zeros(1000),np.eye(1000)),drift4,diffusion4,10,100
3     )
4     plt.title('Trajectoires du systeme pour N=1000, V=$||.||^2$ et W(x,
5     y)=V(x)+V(y)')
6     plt.xlabel('evolution')
7     plt.ylabel('position')

```

Listing 8.22 – Second exemple de simulation des trajectoires des particules du système

Quatrième partie

Perspectives de thèse

On a vu durant ce stage, à travers les articles d'Arnaud Guillin, Aline Kurtzmann, Michel Benaim, Olivier Raimond et leurs co-auteurs que les équations de McKean-Vlasov et auto-attractives ont la même dynamique limite et vérifient des inégalités de transport de type Talagrand, sous des hypothèses adéquates sur les potentiels de confinement et d'interaction. Plusieurs idées sont envisageables pour mon début de thèse et elles sont essentiellement inspirées par l'étude de ces articles.

— **Comparaison et amélioration des hypothèses sur les potentiels**

Q1 : Les hypothèses d'Arnaud Guillin et d'Aline Kurtzmann sont-elles optimales?

Q2 : Les hypothèses d'Arnaud Guillin sont-elles équivalentes à celles d'Aline Kurtzmann?

— **Nature des potentiels**

Q1 : Peut-on avoir les mêmes résultats, pour l'approche champ moyen, avec un potentiel W faisant interagir plus de deux particules?

Q2 : Qu'en est-il si le coefficient de diffusion est non constant et celui de drift ne dérive pas d'un potentiel?

— **Comparaison des approches champ moyen et Méthode de l'EDO**

Q1 : Des résultats de la méthode de l'EDO peuvent-ils, via l'approche champ moyen, s'adapter à l'équation de McKean-Vlasov et vice-versa?

Q2 : Peut-on améliorer ces résultats?

— **Ouverture sur les schémas de discrétisation**

Q1 : Les inégalités de Sobolev logarithmiques établies suffisent-elles pour la bonne approximation du régime stationnaire via les mesures empiriques pondérées du schéma d'Euler à pas décroissant?

Q2 : Pour quelles classes de poids sont-elles suffisantes?

Et bien d'autres questions...

Pour répondre partiellement à une de ces questions, je pense qu'on peut espérer que des résultats de la méthode de l'EDO s'adaptent à l'EDS de McKean-Vlasov standard. Mais quant à la réciproque, je n'en suis pas convaincu...

Pour clore cette partie sur mes perspectives de thèse, je souhaite partager les citations suivantes qui m'ont motivé en particulier dans mes études...

S'il n'y avait pas d'hiver, le printemps ne serait pas si agréable : si nous ne goûtions pas à l'adversité, la réussite ne serait pas tant appréciée.



Anne Bradstreet

www.citation-celebre.com

La vie, c'est comme une bicyclette, il faut avancer pour ne pas perdre l'équilibre.



Albert Einstein

www.citation-celebre.com

Choisissez un travail que vous aimez et vous n'aurez pas à travailler un seul jour de votre vie.



Confucius

www.citation-celebre.com

La réussite est liée à la patience mais elle dépend également de beaucoup de bonne volonté.

Gilbert Brévert

www.citation-celebre.com

C'est dans l'effort que l'on trouve la satisfaction et non dans la réussite. Un plein effort est une pleine victoire.



Gandhi

www.citation-celebre.com

Si l'esprit d'un homme s'égare, faites-lui étudier les mathématiques car dans les démonstrations, pour peu qu'il s'écarte, il sera obligé de recommencer.



Francis Bacon (philosophe)

www.citation-celebre.com

Cinquième partie

Annexes

Chapitre 9

A1. $N \rightarrow +\infty$?

Pour plus de détails concernant ce chapitre, voir référence [17] en bibliographie.

9.1 Vers un formalisme du problème

On considère un espace d'états ou de configurations admissibles E . On prendra $E \subset \mathbb{R}^d$. On considère un système de particules (ou individus/cellules/agrégats, ...) dont chaque particule est totalement décrite par sa variable d'état $y \in E$ (typiquement ici y est la position de la particule, et on note $y = x \in E$, ou sa vitesse, et on note $y = v \in E$, mais on peut imaginer bien d'autres situations, en particulier un couple $y = (x, v)$ position-vitesse, une masse $y = m \geq 0, \dots$). Un système de N particules est donc décrit par une variable d'état

$$Y = (y_1, \dots, y_N) \in E^N,$$

que l'on notera parfois $Y = Y^N = (y_j^N)_{1 \leq j \leq N}$ lorsque l'on voudra insister sur la dépendance en N . On suppose également le plus souvent que les particules sont indistinguables (échangeables dans la terminologie probabiliste) ce qui signifie qu'on ne fait pas de différence entre l'état Y et l'état $Y_\sigma = (y_{\sigma(1)}, \dots, y_{\sigma(N)})$ pour toute permutation $\sigma \in \mathfrak{S}_N = \{\text{bijections de } \{1, 2, \dots, N\}\}$ donnée. Dans ce cas le système est en fait décrit par la classe de Y obtenue par permutation des indices, l'espace des états du système est donc E^N / \mathfrak{S}_N .

Le problème à N particules. La configuration du système à N particules est obtenue comme solution d'un problème d'optimisation (minimisation d'une entropie ou problème de minmax, c'est un problème stationnaire ou à l'équilibre) ou d'une équation d'évolution (problème hors équilibre) mettant en jeu la variable Y (ou une fonction de densité de la variable Y). Plus précisément, soit le système est déterministe et est décrit par un point dans l'espace des phases

$$\begin{aligned} \bar{Y} &= \text{solution optimale d'un problème stationnaire} \in E^N \text{ (ou } E^N / \mathfrak{S}_N), \\ Y(t) &= \text{solution d'un problème d'évolution} \in E^N \text{ (ou } E^N / \mathfrak{S}_N). \end{aligned}$$

Soit le système est stochastique et les variables ci-dessus sont aléatoires. Dans ce cas, on peut préférer une description par densité de probabilité qui correspond à la loi de la variables aléatoire précédente, le système est alors décrit par

$$\begin{aligned} \bar{F}^N &= \text{solution optimale d'un problème stationnaire} \in \mathbf{P}(E^N) \text{ (ou } \mathbf{P}_{\text{sym}}(E^N)), \\ F^N(t) &= \text{solution d'un problème d'évolution} \in \mathbf{P}(E^N) \text{ (ou } \mathbf{P}_{\text{sym}}(E^N)), \end{aligned}$$

où $\mathbf{P}_{\text{sym}}(E^N)$ est l'espace des probabilités symétriques de E^N , c'est à dire invariantes par permutations des indices des variables. Le fait de considérer $\mathbf{P}_{\text{sym}}(E^N)$ et non pas $\mathbf{P}(E^N)$ provient de ce

que l'on suppose que les particules sont indistinguables. Il est important de signaler que dans le cas d'un problème d'évolution, la source d'aléas peut provenir :

- de la dynamique d'évolution elle-même;
- de la donnée initiale non connue avec certitude, la dynamique d'évolution étant elle déterministe.

Question 1 : y a-t-il une limite lorsque $N \rightarrow \infty$?

Une première question est de savoir si on peut identifier une limite lorsque $N \rightarrow \infty$. Pour que cette question ait un sens il faut d'une part que le problème soit observé "*à la bonne échelle*" afin que le système (ou plutôt certaines quantités fabriquées à partir de la description complète du système) "reste d'ordre 1" dans la limite $N \rightarrow \infty$. Il faut d'autre part que l'on identifie "*une limite*" (c'est à dire un objet limite, éventuellement défini comme solution d'une équation/d'un problème limite).

Il est possible d'imaginer plusieurs types de limite, et essentiellement trois suivant la précision de l'observation souhaitée.

- *échelle microscopique*. On peut souhaiter rester au niveau des particules/individus/ cellules et l'espace de configuration limite est alors E^∞ , c'est-à-dire, soit l'espace des suites $E^{\mathbb{N}}$ (ce qui peut être le cas pour des modèles de coalescence ou de fragmentation d'un système de particules décrit par la suite de masses $(m_i)_{i \geq 1}$, $m_1 \geq m_2 \geq \dots \geq m_i \geq \dots \geq 0$) soit un espace structuré spatialement du type $E^{\mathbb{Z}^d}$ (ce serait le cas pour un système de particules en interaction et contraint à rester sur un réseau).

- *échelle mésoscopique/statistique*. Souvent lorsqu'on s'intéresse à un système composé d'un nombre très grand de particules ce n'est pas tant l'état de chaque particule prise individuellement qui nous importe mais bien l'état du système globalement (car chaque particule prise individuellement n'a qu'une action très petite sur le système et son action propre ne peut tout simplement pas être observée). L'échelle mésoscopique correspond à l'échelle où on s'intéresse au comportement statistique ou typique des particules, et on répond à la question : quelle chance ai-je de trouver une particule dans un état donné $y \in E$ ou plutôt dans une portion de l'espace des états $A \subset E$?

- *échelle macroscopique*. Allons plus loin. Ce qui nous importe vraiment c'est comment agir sur un système de particules ou quelle est l'action qu'exerce celui-ci sur son environnement. Dans les deux cas les quantités pertinentes sont les mêmes et ce sont les "*observables du système*" (les quantités que l'on peut observer grâce à des appareils de mesure) qui sont obtenues comme moyennes sur la densité de particules : la densité moyenne, la vitesse moyenne, la température, ...

Cette distinction entre ces trois échelles est particulièrement pertinente pour les gaz d'atomes pour lesquels on peut avoir une description microscopique (la dynamique est donnée par les équations de Newton de la mécanique classique), une description mésoscopique (la dynamique est donnée par les équations cinétiques de Boltzmann ou Vlasov, ...), une description macroscopique (la dynamique est donnée par les équations de la mécanique des fluides de Euler, Navier-Stokes, ...).

Limite de champ moyen. Nous nous intéressons ici uniquement à la deuxième situation ci-dessus. En résumé, on souhaite établir une description statistique du système à partir d'une description individuelle des particules et dans la limite où le nombre de particules tend vers l'infini. En termes probabilistes, cela correspond à une limite de type loi des grands nombres (LGN). Deux objets permettent de décrire un tel système limite. Le plus simple est une densité de probabilité $f \in \mathbf{P}(E)$ qui représente la répartition statistique des particules dans le système. De manière un peu plus élaborée, on peut s'intéresser à un processus stochastique décrivant une particule typique, et dont la loi correspond à la probabilité f ci-dessus. Pour être plus précis dans la terminologie, nous

allons nous intéresser à des “*limites de champ moyen*” qui correspondent au cas où l’action de chaque particule du système de N particules devient de plus en plus petite lorsque $N \rightarrow \infty$, mais que l’action moyenne des particules est d’ordre 1.

Comment effectuer une limite $N \rightarrow \infty$?

Un problème majeur réside dans le fait que l’espace E^N ou $\mathbf{P}(E^N)$ dans lequel vit la solution du problème de N particules change avec N , et est donc a priori différent de l’espace dans lequel va vivre sa limite.

Rappelons que lorsque le système est déterministe, il est décrit par

— sa variable d’état $Y = (y_1, \dots, y_N) \in E^N$,

ou de manière équivalente lorsque les particules sont indiscernables par

— la mesure empirique associée $\mu_Y^N \in \mathbf{P}(E)$,

avec par définition

$$\mu_Y^N(dy) := \frac{1}{N} \sum_{i=1}^N \delta_{y_i}(dy) \in \mathbf{P}(E). \quad (9.1.1)$$

Dans le cas stochastique le système est décrit pas la variable aléatoire $Y \in E^N$ ou la mesure empirique aléatoire associée $\mu_Y^N \in \mathbf{P}(E)$. En termes de lois, le système est donc décrit par

— la loi $F^N \in \mathbf{P}(E^N)$ de la va Y , en fait $F^N \in \mathbf{P}_{sym}(E^N)$,

— ou la loi $\hat{F}^N \in \mathbf{P}(\mathbf{P}(E))$ de la va μ_Y^N .

Le système limite est quant à lui décrit par une densité typique $\bar{f} \in \mathbf{P}(E)$ ou par un état (éventuellement aléatoire) typique $\bar{Y} \in E$. Afin d’établir la convergence du système de N particules, qui est décrit soit par une suite d’états $Y^N = (y_1^N, \dots, y_N^N)$ soit par une suite de densités $F^N \in \mathbf{P}_{sym}(E^N)$ ou donc $\hat{F}^N \in \mathbf{P}(\mathbf{P}(E))$, vers un comportement typique, on peut procéder de différentes manières.

— (1) On travaille dans $\mathbf{P}(E)$, et on prouve que

$$\mu_{Y^N}^N \rightharpoonup \bar{f} \quad \text{faiblement dans } \mathbf{P}(E).$$

— (2) On travaille toujours dans $\mathbf{P}(E)$, et on prouve qu’il existe $\bar{\pi}_1 \in \mathbf{P}(E)$ tel que

$$F_1^N := \int_{E^{N-1}} F^N dy_2 \dots dy_N \rightharpoonup \bar{\pi}_1 \quad \text{faiblement dans } \mathbf{P}(E).$$

Dans (1) et (2) les quantités $\mu_{Y^N}^N$ et F_1^N représentent la densité de probabilité d’une particule du système de N particules de se trouver dans un état donné. Il s’avère que la convergence (1) n’est envisageable que pour un système déterministe et que nous ne sommes pas toujours capable d’identifier la limite avec la seule information (2). On peut alors, comme le fait Kac (1956), s’intéresser à une densité de probabilité des couples de particules du système de N .

— (2′) On travaille dans $\mathbf{P}(E^2)$, et on prouve qu’il existe $\bar{\pi}_2 \in \mathbf{P}_{sym}(E^2)$ tel que

$$F_2^N := \int_{E^{N-2}} F^N dy_3 \dots dy_N \rightharpoonup \bar{\pi}_2 \quad \text{faiblement dans } \mathbf{P}(E^2).$$

Là encore l’information est souvent trop partielle pour identifier la limite, et nous sommes alors amenés à mettre en œuvre l’une des stratégies suivantes.

— (2′′) On travaille dans $\mathbf{P}(E^k)$, pour tout $k \in \mathbb{N}^*$ fixé, et on prouve qu’il existe $\bar{\pi}_k \in \mathbf{P}_{sym}(E^k)$ tel que

$$F_k^N := \int_{E^{N-k}} F^N dy_{k+1} \dots dy_N \rightharpoonup \bar{\pi}_k \quad \text{faiblement dans } \mathbf{P}(E^k).$$

— (3) On travaille dans E^N . Dans le cas déterministe, on prouve que

$$d_1(Y^N, \bar{Y}^N) := \frac{1}{N} \sum_{i=1}^N |y_i^N - \bar{y}_i^N| \rightarrow 0, \quad (9.1.2)$$

avec $\bar{y}_i = \bar{y} \in E$ fixé. Dans le cas stochastique, on prouve encore (9.1.2) au sens p.s. ou en espérance, avec $\bar{Y}^N = (\bar{y}_1, \dots, \bar{y}_N)$ un vecteur dont les coordonnées sont des va indépendantes et de loi identique \bar{f} fixée. En particulier, on prouve

$$W_1(F^N, \bar{f}^{\otimes N}) \leq \mathbf{E}(d_1(Y^N, \bar{Y}^N)) \rightarrow 0,$$

où W_1 désigne la distance de Monge-Kantorovich-Wasserstein associée à la distance d_1 . On peut également démontrer la variante

$$\mathbf{E}\left(W_1(\mu_{Y^N}^N, \mu_{\bar{Y}^N}^N)\right) := \mathbf{E}\left(\inf_{\sigma \in \mathcal{S}_N} \frac{1}{N} \sum_{i=1}^N |y_i^N - \bar{y}_{\sigma(i)}^N|\right) \rightarrow 0,$$

qui peut sembler plus faible, mais ne l'est en fait pas pour des particules indistinguables, i.e. lorsque $F^N \in \mathbf{P}_{sym}(E^N)$.

— (4) On travaille dans $\mathbf{P}(\mathbf{P}(E))$, et on prouve qu'il existe $\pi \in \mathbf{P}(\mathbf{P}(E))$ telle que

$$\hat{F}^N \rightarrow \hat{\pi} \quad \text{faiblement dans } \mathbf{P}(\mathbf{P}(E)),$$

ou de manière un peu plus précis on travaille dans le cadre des va à valeurs dans $\mathbf{P}(E)$, le système est décrit par la suite de va $\mu_{Y^N}^N$ et on montre qu'il existe une va \bar{Y} à valeurs dans $\mathbf{P}(E)$ telle que $\mu_{Y^N}^N \rightarrow \bar{Y}$ presque sûrement.

On ne discutera pas maintenant précisément le sens de ces convergences et des relations entre elles. Faisons toutefois quelques commentaires. A part dans le cas (3), l'espace dans lequel on travaille est fixe, on peut donc faire dans ces cas des arguments de compacité et ne pas connaître a priori la limite cherchée \bar{f} , $\bar{\pi}_k$ ou π . Les convergences (2') (avec $\bar{\pi}_2 = \bar{f}^{\otimes 2}$), (2'') (avec $\bar{\pi}_k = \bar{f}^{\otimes k}$), (3) et (4) (avec $\hat{\pi} = \delta_{\bar{f}}$) sont équivalentes et impliquent les convergences (1) et (2), ces deux dernières convergences étant équivalentes. La convergence (1) (et (2) avec $k = 1$) répond bien à la question 1 : la distribution "moyenne" des N particules converge vers une distribution "typique".

Question 2 : a-t-on le chaos ou la propagation du chaos ?

On dit qu'une suite $Y = Y^N \in E^N$ de va est chaotique, ou plus précisément \bar{f} -chaotique avec $\bar{f} \in \mathbf{P}(E)$, si sa loi F^N satisfait

$$\forall k \in \mathbb{N}^* \quad F_k^N \rightarrow \bar{f}^{\otimes k} \quad \text{faiblement dans } \mathbf{P}(E^k). \quad (9.1.3)$$

Cela est exactement la convergence (2'') avec $\bar{\pi}_k = \bar{f}^{\otimes k}$ pour tout $k \geq 1$ et c'est une formulation précise de l'idée vague selon laquelle Y^N est "asymptotiquement proche d'une va de coordonnées indépendantes lorsque $N \rightarrow \infty$ " ou dit autrement est "presque une va de coordonnées indépendantes dans l'asymptotique $N \rightarrow \infty$ ". De la même façon, on dit que $F^N \in \mathbf{P}_{sym}(E^N)$ est \bar{f} -chaotique si (9.1.3) a lieu.

Une deuxième question est alors de savoir si dans le cas d'un problème stationnaire F^N (ou Y^N) est \bar{f} -chaotique (**chaos**), ou si dans le cas d'un problème d'évolution $F^N(0)$ (ou $Y^N(0)$) est $\bar{f}(0)$ -chaotique entraîne que $F^N(t)$ (ou $Y^N(t)$) est $\bar{f}(t)$ -chaotique pour tout $t \geq 0$ (**propagation du chaos**).

Plusieurs remarques s'imposent :

- La question de la chaotité est plus exigeante et subtile que celle d'une simple limite de champ moyen puisque (9.1.3) implique en particulier (1) et (2).

- Il est important de retenir que dès que le système à N particules est en interaction, la dynamique d'une particule donnée est **toujours corrélée** / n'est **jamais indépendante** à celle des autres particules, ce phénomène d'indépendance (qu'est le chaos) n'est donc valable qu'à la limite.

- Soulignons également que pour certains problèmes (par exemple dans le cas du problème de Boltzmann) on doit répondre à la deuxième question (limite chaotique) pour répondre à la première question (équation sur la loi d'une particule typique) : on ne sait démontrer (2) avec $k = 1$ qu'en passant par la preuve préalable de (9.1.3).

9.2 Stratégie de l'argument de couplage et chaos

Soit $X = (X^{N,1}, \dots, X^{N,N}) \in E^N$ la solution du système de N particules

$$\begin{aligned} dX^i &= \frac{1}{N} \sum_{j \neq i} a(X^i - X^j) dt + dB_t^i \\ &= (a \star \mu_X^N)(X^i) dt + dB_t^i, \end{aligned}$$

où (B^i) est une collection de mouvements Brownien indépendants et X_0^N est un vecteur aléatoire dont les coordonnées sont indépendantes et de même loi f_0 (indépendant de N).

- La limite de champ moyen est

$$\partial_t f = \frac{1}{2} \Delta f + \operatorname{div}((a \star f) f),$$

et on peut montrer encore que

$$\forall f_0, g_0 \quad W_1(f_t, g_t) \leq C_t W_1(f_0, g_0) \quad \text{ou} \quad \leq C e^{-\lambda t} W_1(f_0, g_0).$$

Le problème ici est que μ_X^N n'est pas solution de l'EDP non linéaire et l'argument valable pour Vlasov ne tient plus.

- Mais on va faire mieux. On introduit la dynamique de particules fictives (pour les mêmes mouvements Brownien)

$$dY^i = (a \star f)(Y^i) dt + dB_t^i, \quad Y_0^i = X_0^i,$$

et la formule d'Itô implique que $\mathcal{L}(Y^i) = f$ car $\mathcal{L}(Y^i)$ est solution de la même EDP (pour laquelle il y a unicité des solutions). On démontre (sous forme intégrale) que

$$\frac{d}{dt} \frac{1}{N} \sum_{k=1}^N |X_t^k - Y_t^k| \leq 2 \|a\|_{Lip} \frac{1}{N} \sum_{k=1}^N |X_t^k - Y_t^k| + \mathcal{Y}(t)$$

avec $\mathcal{Y}(t)$ qui ne dépend que de $Y(t)$. On en déduit

$$\frac{d}{dt} \mathbf{E} |X_t^1 - Y_t^1| \leq 2 \|a\|_{Lip} \mathbf{E} |X_t^1 - Y_t^1| + \mathbf{E}(\mathcal{Y}(t)).$$

Or on peut montrer que $\mathbf{E}(\mathcal{Y}(t)) = \mathcal{O}(1/\sqrt{N})$: cela provient juste du fait que les Y^i forme une suite de va i.i.d. et que $\mathbf{E}(\mathcal{Y}(t))$ est une mesure de l'erreur entre la loi d'une moyenne empirique des Y^i et sa limite déterministe.

On en déduit d'une part que pour tout $\varphi \in \operatorname{Lip}(E)$

$$|\langle f_1^N, \varphi \rangle - \langle f, \varphi \rangle| = |\mathbf{E}\varphi(X_t^1) - \mathbf{E}\varphi(Y_t^1)| = \mathcal{O}(1/\sqrt{N}),$$

ce qui signifie bien que $f_1^N = \mathcal{L}(X_1) \rightarrow f$ lorsque $N \rightarrow \infty$.

Et d'autre part, pour tout $\varphi \in \operatorname{Lip}(E^j)$

$$\begin{aligned} |\langle f_j^N, \varphi \rangle - \langle f^{\otimes j}, \varphi \rangle| &= |\mathbf{E}\varphi(X_t^1, \dots, X_t^j) - \mathbf{E}\varphi(Y_t^1, \dots, Y_t^j)| \\ &\leq \|\varphi\|_{Lip} \sum_{k=1}^j \mathbf{E} |X_t^k - Y_t^k| \leq j \mathcal{O}(1/\sqrt{N}), \end{aligned}$$

ce qui signifie que f^N est f -chaotique.

On remarquera que la même preuve implique le même résultat (propagation du chaos) pour le modèle de Vlasov (avec donnée initiale aléatoire mais de coordonnées indépendantes).

Problème 1. Peut-on établir le même résultat en restant au niveau des lois des systèmes de particules (X^1, \dots, X^N) et (Y^1, \dots, Y^N) ? Il faut bien sûr faire attention à quand prendre les marginales : le plus tard possible!

- On considère le système de particules

$$dX_t^i = dB_t^i + \frac{1}{N} \sum_{j=1}^N F(X_t^i - X_t^j) dt$$

où les B^i sont des mouvements Browniens indépendants et $F \in (Lip \cap L^\infty)(\mathbb{R}^d)$. En posant

$$G(x) := \left(\frac{1}{N} \sum_{j=1}^N F(x^1 - x^j), \dots, \frac{1}{N} \sum_{j=1}^N F(x^N - x^j) \right)$$

on voit que la loi m^N de X_t résout

$$\partial_t m^N = \frac{1}{2} \Delta m^N - \nabla \cdot (G m^N).$$

- On considère la solution ρ_t de l'équation parabolique non linéaire

$$\partial_t \rho = \frac{1}{2} \Delta \rho - \nabla \cdot [(F * \rho) \rho] \quad \mathbb{R}^d. \quad (9.2.1) \quad \boxed{\text{Mc}}$$

• On va mettre en place une stratégie de couplage. On considère Y_t^1, \dots, Y_t^N des va indépendantes et solutions de l'équation différentiel stochastique

$$dY_t^i = dB_t^i + (F * \rho_t)(Y_t^i) dt, \quad Y \sim \rho_0^{\otimes N}.$$

Sa loi η^i est solution de

$$\partial_t \eta^i = \frac{1}{2} \Delta \eta^i - \nabla \cdot [(F * \rho) \eta^i] \quad \mathbb{R}^d,$$

et comme ρ est l'unique solution de Mc (9.2.1) on a $\eta^i = \rho$ pour tout $i = 1, \dots, N$.

Proposition 9.2.1 (Couplage 1). On suppose F Lipschitzienne et bornée. Alors

$$\sup_{[0, T]} \mathbb{E} |X_t^1 - Y_t^1| \leq \frac{C_T}{\sqrt{N}}.$$

Preuve de la proposition (9.2.1). Par définition, on a en notant $b(x, y) = F(x - y)$

$$\begin{aligned} \frac{d}{dt} (X_t^i - Y_t^i) &= \frac{1}{N} \sum_{j=1}^N b(X_t^i, X_t^j) - \int_{\mathbb{R}^d} b(Y_t^i, y) \rho_t(dy) \\ &= \frac{1}{N} \sum_{j=1}^N [b(X_t^i, X_t^j) - b(Y_t^i, X_t^j)] \\ &\quad + \frac{1}{N} \sum_{j=1}^N [b(Y_t^i, X_t^j) - b(Y_t^i, X_t^j)] \\ &\quad + \frac{1}{N} \sum_{j=1}^N [b(Y_t^i, Y_t^j) - \int_{\mathbb{R}^d} b(Y_t^i, y) \rho_t(dy)], \end{aligned}$$

d'où la majoration

$$\begin{aligned} \frac{d}{dt} |X_t^i - Y_t^i| &\leq \|b\|_{Lip} |X_t^i - Y_t^i| \\ &+ \frac{1}{N} \sum_{j=1}^N \|b\|_{Lip} |X_t^j - Y_t^j| \\ &+ \left| \frac{1}{N} \sum_{j=1}^N [b(Y_t^i, Y_t^j) - \int_{\mathbb{R}^d} b(Y_t^i, y) \rho_t(dy)] \right|. \end{aligned}$$

En sommant ces N inégalités il vient

$$\begin{aligned} \frac{d}{dt} \frac{1}{N} \sum_{j=1}^N |X_t^j - Y_t^j| &\leq 2 \|b\|_{Lip} \frac{1}{N} \sum_{k=1}^N |X_t^k - Y_t^k| \\ &+ \frac{1}{N} \sum_{i=1}^N \left| \frac{1}{N} \sum_{j=1}^N [b(Y_t^i, Y_t^j) - \int_{\mathbb{R}^d} b(Y_t^i, y) \rho_t(dy)] \right|. \end{aligned}$$

En prenant l'espérance et en utilisant la symétrie des lois de X_t et Y_t par rapport à l'indice des particules ($\mathcal{L}(X_t^i) = \mathcal{L}(X_t^j)$ et $\mathcal{L}(Y_t^i) = \mathcal{L}(Y_t^j) = \rho_t$ pour tout i, j), il vient

$$\begin{aligned} \frac{d}{dt} \mathbf{E} |X_t^1 - Y_t^1| &\leq 2 \|b\|_{Lip} \mathbf{E} |X_t^1 - Y_t^1| + A(t) \\ \text{avec } A(t) &:= \mathbf{E} \left| \frac{1}{N} \sum_{j=1}^N [b(Y_t^1, Y_t^j) - \int_{\mathbb{R}^d} b(Y_t^1, y) \rho_t(dy)] \right|. \end{aligned}$$

Nous allons maintenant estimer le terme $A(t)$. Par inégalité de Cauchy-Schwarz et en développant entièrement on obtient

$$\begin{aligned} A(t)^2 &\leq \mathbf{E} \left| \frac{1}{N} \sum_{j=1}^N [b(Y_t^1, Y_t^j) - \int_{\mathbb{R}^d} b(Y_t^1, y) \rho_t(dy)] \right|^2 \\ &\leq \frac{1}{N^2} \sum_{j,k=1}^N \mathbf{E} \left[b(Y_t^1, Y_t^j) - \int_{\mathbb{R}^d} b(Y_t^1, y) \rho_t(dy) \right] \left[b(Y_t^1, Y_t^k) - \int_{\mathbb{R}^d} b(Y_t^1, y) \rho_t(dy) \right] \\ &\leq \frac{1}{N^2} \sum_{(j,k) \in B} \left[\mathbf{E} b(Y_t^1, Y_t^j) - \mathbf{E} \int_{\mathbb{R}^d} b(Y_t^1, y) \rho_t(dy) \right] \left[\mathbf{E} b(Y_t^1, Y_t^k) - \mathbf{E} \int_{\mathbb{R}^d} b(Y_t^1, y) \rho_t(dy) \right] \\ &+ \frac{1}{N^2} \sum_{(j,k) \in B^c} \mathbf{E} \left[b(Y_t^1, Y_t^j) - \int_{\mathbb{R}^d} b(Y_t^1, y) \rho_t(dy) \right] \left[b(Y_t^1, Y_t^k) - \int_{\mathbb{R}^d} b(Y_t^1, y) \rho_t(dy) \right] \end{aligned}$$

où $B := \{(j, k); j \neq k, j \neq 1, k \neq 1\}$ de sorte que Y_t^j, Y_t^k et Y_t^1 sont indépendants si $(j, k) \in B$, les quatre termes intervenant dans la somme sur l'ensemble B sont égaux (ils valent tous $\langle \rho^{\otimes 2}, b \rangle$) et cette somme est donc nulle, et où le cardinal de B^c vaut $3N - 2$. Ainsi

$$A(t) \leq \sqrt{\frac{1}{N^2} (3N - 2) (2 \|b\|_{L^\infty})^2} \leq \frac{\sqrt{12} \|b\|_{L^\infty}}{\sqrt{N}}.$$

De plus, le lemme de Gronwall appliqué à l'inégalité différentielle

$$u' \leq \alpha u + \beta, \quad u(0) = 0$$

implique $u(t) \leq \frac{\beta}{\alpha} e^{\alpha t}$. On obtient ainsi le résultat avec $C_T = \frac{2 \|b\|_{L^\infty}}{\|b\|_{Lip}} e^{2 \|b\|_{Lip} T}$. \square

Proposition 9.2.2 (Propagation du chaos). *On suppose F Lipschitzienne et bornée. Pour toute fonction $\varphi \in C_b(\mathbb{R}^d) \cap Lip(\mathbb{R}^d)$ on a*

$$\sup_{[0, T]} \mathbf{E} \left| \int_{\mathbb{R}^d} \varphi(y) \hat{\mu}_{X_t}^N(dy) - \int_{\mathbb{R}^d} \varphi(y) \rho_t(dy) \right| \leq \frac{C_T}{\sqrt{N}}.$$

Preuve de la Proposition (9.2.2). On sépare en deux morceaux

$$\begin{aligned} & \mathbf{E} \left| \int_{\mathbb{R}^d} \varphi(y) \hat{\mu}_{X_t}^N(dy) - \int_{\mathbb{R}^d} \varphi(y) \rho_t(dy) \right| \leq \\ & \leq \mathbf{E} \left| \int_{\mathbb{R}^d} \varphi \hat{\mu}_{X_t}^N - \int_{\mathbb{R}^d} \varphi \hat{\mu}_{Y_t}^N \right| + \mathbf{E} \left| \int_{\mathbb{R}^d} \varphi \hat{\mu}_{Y_t}^N - \int_{\mathbb{R}^d} \varphi \rho_t \right|. \end{aligned}$$

Le premier terme s'écrit

$$\begin{aligned} & \mathbf{E} \left| \int_{\mathbb{R}^d} \varphi \hat{\mu}_{X_t}^N - \int_{\mathbb{R}^d} \varphi \hat{\mu}_{Y_t}^N \right| = \mathbf{E} \left| \frac{1}{N} \sum_{i=1}^N \varphi(X_t^i) - \frac{1}{N} \sum_{i=1}^N \varphi(Y_t^i) \right| \\ & \leq \frac{1}{N} \sum_{i=1}^N \mathbf{E} \left| \varphi(X_t^i) - \varphi(Y_t^i) \right| = \mathbf{E} \left| \varphi(X_t^1) - \varphi(Y_t^1) \right| \\ & \leq \|b\|_{Lip} \mathbf{E} |X_t^1 - Y_t^1| \leq \frac{C_T}{\sqrt{N}}, \end{aligned}$$

grâce à la proposition (9.2.1). On se rappelle maintenant que $Y_t \sim \rho_t^{\otimes N}$ et en reprenant le premier point de la preuve d'un théorème de l'annexe, on a

$$\left(\mathbf{E} \left| \int_{\mathbb{R}^d} \varphi \hat{\mu}_{Y_t}^N - \int_{\mathbb{R}^d} \varphi \rho_t \right| \right)^2 \leq \mathbf{E} \left(\int_{\mathbb{R}^d} \varphi d(\hat{\mu}_{Y_t}^N - \rho_t) \right)^2 \leq \frac{\|\varphi\|_{L^2}^2}{N},$$

ce qui donne également une contribution en $1/\sqrt{N}$ pour le second terme. \square

9.3 Fonctions continues et polynômes sur $\mathbf{P}(E)$

Dans cette section, E désignera un espace polonais, un espace polonais localement compact ou un espace compact.

Définition 9.3.3 *Soit E un espace polonais. On dit qu'une fonction $U : \mathbf{P}(E) \rightarrow \mathbb{R}$ est continue si $\rho_n \rightarrow \rho$ faiblement implique $U(\rho_n) \rightarrow U(\rho)$. On note $U \in C(\mathbf{P}(E))$. On note $C_b(\mathbf{P}(E))$ l'espace des fonctions continues et bornées. On note $UC(\mathbf{P}(E))$, l'espace des fonctions uniformément continues et bornées sur $\mathbf{P}(E)$: étant fixée une distance D métrisant la topologie faible, il existe un module de continuité ω (c'est-à-dire une fonction $\omega : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ continue et telle que $\omega(0) = 0$) tel que*

$$\forall \rho, \eta \in \mathbf{P}(E) \quad |U(\eta) - U(\rho)| \leq \omega(D(\rho, \eta)).$$

Lorsque E est compact alors $\mathbf{P}(E)$ est compact et $UC(\mathbf{P}(E)) = C_b(\mathbf{P}(E)) = C(\mathbf{P}(E))$.

• Le prototype de la fonction continue est le monôme de degré 1 : étant donnée une fonction $\varphi \in C_b(E)$ on pose

$$\forall \rho \in \mathbf{P}(E) \quad R_\varphi(\rho) = \int_E \varphi d\rho.$$

On définit bien ainsi $R_\varphi \in C_b(\mathbf{P}(E))$. A noter également que si $\varphi \in Lip(E)$ alors $R_\varphi \in Lip(\mathbf{P}(E)) \subset UC(\mathbf{P}(E))$, où $Lip(\mathbf{P}(E))$ correspond au cas où le module de continuité peut être pris de la forme $\omega(s) = Ls$, $L \geq 0$, dans la définition (9.3.3).

• Plus généralement, pour $k \in \mathbb{N}^*$ et $\varphi \in C_b(E^k)$ donnés, on définit $R_\varphi : \mathbf{P}(E) \rightarrow \mathbb{R}$ par

$$R_\varphi(m) := \int_{E^k} \varphi(x_1, \dots, x_k) m(dx_1) \dots m(dx_k) = \langle m^{\otimes k}, \varphi \rangle.$$

On dit que R_φ est un polynôme de degré (au plus) k , et φ correspond aux coefficients. On définit la multiplication par un scalaire de manière évidente. On définit l'addition de deux polynômes R_1 et R_2 , respectivement associés aux fonctions $\varphi_1 \in C(E^{k_1})$ et $\varphi_2 \in C(E^{k_2})$, avec disons $k_1 \leq k_2$, comme le polynôme de degré au plus k_2 associé à la fonction

$$\varphi(x_1, \dots, x_{k_2}) = \varphi_1(x_1, \dots, x_{k_1}) + \varphi_2(x_1, \dots, x_{k_2}).$$

On a bien ainsi $R_\varphi(m) = R_1(m) + R_2(m)$ pour tout $m \in \mathbf{P}(E)$. On définit le produit de R_1 et R_2 comme le polynôme de degré au plus $k_1 + k_2$ associé à la fonction

$$\psi(x_1, \dots, x_{k_1}, x_{k_1+1}, \dots, x_{k_1+k_2}) = \varphi_1(x_1, \dots, x_{k_1}) \varphi_2(x_{k_1+1}, \dots, x_{k_1+k_2}).$$

On a bien ainsi $R_\psi(m) = R_1(m) R_2(m)$ pour tout $m \in \mathbf{P}(E)$. On note $\mathbb{P}(\mathbf{P}(E))$ l'ensemble des polynômes qui a donc une structure d'algèbre unitaire.

Lemme 9.3.4 *Soit E un espace polonais localement compact. Alors $\mathbb{P}(\mathbf{P}(E)) \subset C_b(\mathbf{P}(E))$*

Preuve du Lemme (9.3.4). On fixe $\varphi \in C_b(E^k)$. D'une part clairement R_φ est bornée. D'autre part, étant donnée une suite (ρ_n) de $\mathbf{P}(E)$ et $\rho \in \mathbf{P}(E)$ telles que $\rho_n \rightarrow \rho$ faiblement dans $\mathbf{P}(E)$, on a en particulier $\rho_n^{\otimes k} \rightarrow \rho^{\otimes k}$ pour la dualité de $C_c(E)^{\otimes k}$, donc pour la dualité de $C_c(E^k)$ (ici on utilise le théorème de Stone-Weierstrass) et donc enfin pour la dualité de $C_b(E^k)$ (ici on utilise le lemme d'équivalence des convergences). On en déduit $R_\varphi(\rho_n) \rightarrow R_\varphi(\rho)$. \square

Théorème 9.3.5 *Lorsque E est compact, l'algèbre des polynômes $\mathbb{P}(\mathbf{P}(E))$ est dense dans $C(\mathbf{P}(E))$. Cela signifie que pour toute fonction $U \in C(\mathbf{P}(E))$ il existe une suite de polynômes $R_j = R_{\varphi_j}$, avec $\varphi \in C(E^j)$, (et $j \rightarrow \infty$ si U n'est pas un polynôme!) telle que $R_j \rightarrow U$ dans $C(\mathbf{P}(E))$:*

$$\sup_{m \in \mathbf{P}(E)} |U(m) - R_j(m)| \rightarrow 0 \text{ lorsque } j \rightarrow \infty.$$

Preuve du théorème (9.3.5). Comme E est compact, l'espace $\mathbf{P}(E)$ est compact.

Il est clair que $\mathbb{P}(\mathbf{P}(E))$ sépare les points de $\mathbf{P}(E)$: pour $\rho_1 \neq \rho_2 \in \mathbf{P}(E)$ il suffit de prendre une fonction $\varphi \in C(E)$ telle que $\langle \rho_2, \varphi \rangle \neq \langle \rho_1, \varphi \rangle$. On applique le théorème de Stone-Weierstrass (mais cette fois sur le compact $\mathbf{P}(E)$) et on obtient que l'algèbre unitaire $\mathbb{P}(\mathbf{P}(E))$ est dense dans $C(\mathbf{P}(E))$. \square

• On note $\mathbb{M}(\mathbf{P}(E))$ l'ensemble des monômes : ce sont les polynômes associés aux fonctions tensorisées $\varphi = \varphi_1 \otimes \dots \otimes \varphi_k$, $\varphi_i \in C(E)$. Lorsque E est compact, on montre comme ci-dessus que l'algèbre engendrée par les monômes est dense dans $C(\mathbf{P}(E))$. Cela explique pourquoi on peut se restreindre à considérer des fonctions testes tensorisées lorsque l'on souhaite montrer des théorèmes de convergence.

• Considérons $\phi : \mathbb{R}^2 \rightarrow \mathbb{R}$ de classe C^1 , $\phi(\cdot, 0) = 0$ et $S_t : \mathbf{P}(\mathbb{R}) \rightarrow \mathbf{P}(\mathbb{R})$ le semi-groupe de la chaleur sur \mathbb{R} . Comme $S_t : \mathbf{P}(\mathbb{R}) \rightarrow (L^1 \cap L^\infty)(\mathbb{R})$, $\rho \mapsto \rho_t := \gamma_t * \rho$, pour tout $t > 0$, on peut définir

$$\Phi_t(\rho) := \int_{\mathbb{R}^2} \phi(x, \rho_t(y)) dx dy,$$

et $\Phi_t \in C_b(\mathbf{P}(\mathbb{R}))$. Il est à noter que si $\phi(x, s) = \psi(x)$, $\psi \in C_c(\mathbb{R})$ alors Φ_t est le polynôme associée à $\psi * \gamma_t$. Dans le cas général, Φ_t n'est pas un polynôme.

• Considérons $S_t : \mathbf{P}(\mathbb{R}^d) \rightarrow \mathbf{P}(\mathbb{R}^d)$, $\rho \mapsto \rho_t$ un semi-groupe non linéaire (par exemple celui obtenu en résolvant l'équation de Boltzmann ou une équation de la chaleur non linéaire). Pour tout $\phi \in C_c(\mathbb{R}^d)$, la fonction $\Phi_t(\rho) := R_\phi(\rho_t)$ appartient à $C_b(\mathbf{P}(E))$ mais n'est pas un polynôme.

• La construction des polynômes ci-dessus nous a permis de définir une application

$$C_b(E^k) \rightarrow C_b(\mathbf{P}(E)), \quad \varphi \mapsto R_\varphi. \quad (9.3.2)$$

• **L'application** π_C^N . Inversement, étant donnée une fonction $U \in C_b(\mathbf{P}(E))$, on définit pour tout N une fonction symétrique sur E^N en posant

$$\forall X \in E^N \quad u^N(X) := U(\hat{\mu}_X^N). \quad (9.3.3)$$

Par exemple, pour un polynôme $R = R_\varphi$ de degré 1, avec donc $\varphi \in C_b(E)$, on calcule

$$X = (x_1, \dots, x_N) \mapsto R_\varphi(\mu_X^N) = \frac{1}{N} \sum_{i=1}^N \varphi(x_i).$$

• Pour tout $k, N \in \mathbb{N}^*$ en combinant les points précédents, on définit une application

$$C_b(E^k) \rightarrow C_{\text{sym}}(E^N), \quad \varphi \mapsto \bar{\varphi}^{(N)}(X) = R_\varphi(\mu_X^N). \quad (9.3.4)$$

Plus précisément, pour tout $\varphi \in C_b(E^k)$ et tout $X \in E^N$, on a

$$\begin{aligned} R_\varphi(\mu_X^N) &= \int_{E^k} \varphi(y_1, \dots, y_k) \left(\frac{1}{N} \sum_{i_1=1}^N \delta_{x_{i_1}} \right) (dy_1) \dots \left(\frac{1}{N} \sum_{i_k=1}^N \delta_{x_{i_k}} \right) (dy_k) \\ &= \frac{1}{N^k} \sum_{i_1, \dots, i_k=1}^N \varphi(x_{i_1}, \dots, x_{i_k}) = \bar{\varphi}^{(N)}(X). \end{aligned} \quad (9.3.5)$$

D'autre part, pour une fonction $\psi \in C(E^N)$ notons $\tilde{\psi}$ la "fonction symétrisée de ψ ", c'est-à-dire celle obtenue en posant

$$\tilde{\psi}(X) := \frac{1}{\#\mathfrak{S}_N} \sum_{\sigma \in \mathfrak{S}_N} \psi(x_{\sigma(1)}, \dots, x_{\sigma(N)}).$$

Le résultat suivant établit que pour $\psi = \varphi \otimes \mathbf{1}^{N-k}$, $\varphi \in C_b(E^k)$, on a $\tilde{\psi} \sim \bar{\varphi}$ asymptotiquement lorsque $N \rightarrow \infty$.

Lemme 9.3.6 Soit E un espace métrique quelconque et soit $\varphi \in C_b(E^k)$. Alors

$$\sup_{X \in E^N / \mathfrak{S}_N} |\varphi \otimes \mathbf{1}^{\otimes(N-k)}(X) - R_\varphi(\hat{\mu}_X^N)| \leq \frac{2k^2}{N} \|\varphi\|_\infty \quad (9.3.6)$$

et donc

$$\limsup_{N \rightarrow \infty} \|\varphi \otimes \mathbf{1}^{\otimes(N-k)}\|_{L^\infty(E^N)} \leq \sup_{\rho \in \mathbf{P}(E)} |R_\varphi(\rho)|. \quad (9.3.7)$$

Preuve du lemme (9.3.6). D'une part, par définition

$$\varphi \otimes \mathbf{1}^{\otimes(N-k)}(X) = \frac{1}{\#\mathfrak{S}_N} \sum_{\sigma \in \mathfrak{S}_N} \varphi(x_{\sigma(1)}, \dots, x_{\sigma(k)}).$$

D'autre part, on définit pour tout $k \leq N$: $A_{N,k} := \{(i_1, \dots, i_k); i_\ell \neq i_{\ell'} \forall \ell \neq \ell'\}$, son complémentaire $B_{N,k} := \{(i_1, \dots, i_k); \exists \ell \neq \ell' \text{ t.q. } i_\ell = i_{\ell'}\}$ et pour tout $(i_1, \dots, i_k) \in A_{N,k}$ l'ensemble $C_{N,k}(i_1, \dots, i_k) := \{\sigma \in \mathfrak{S}_N; \sigma(1) = i_1, \dots, \sigma(k) = i_k\}$. Estimons le cardinal de $B_{N,k}$ et des $C_{N,k}$. Pour tout $(i_1, \dots, i_k) \in A_{N,k}$ on

a $\#C_{N,k}(i_1, \dots, i_k) = (N-k)!$ puisqu'il convient de choisir $\sigma(k+1)$ parmi $N-k$ termes, \dots , $\sigma(N-2)$ parmi 2 termes restants. Pour la même raison $\#A_{N,k} = N \dots (N-k+1)$. En particulier $\#A_{N,k} \#C_{N,k} = N! = \#\mathfrak{S}_N$ et

$$\begin{aligned} \frac{\#B_{N,k}}{N^k} &= 1 - \left(1 - \frac{1}{N}\right) \dots \left(1 - \frac{k-1}{N}\right) \\ &= 1 - \exp\left(\sum_{\ell=0}^{k-1} \ln\left(1 - \frac{\ell}{N}\right)\right) \\ &\leq 1 - \exp\left(-2 \sum_{\ell=0}^{k-1} \frac{\ell}{N}\right) \\ &\leq \frac{2}{N} + \dots + \frac{2(k-1)}{N} \leq \frac{k^2}{N}, \end{aligned}$$

car $\ln(1-x) \geq -2x \forall x \in [0, 1/2]$ et $e^x \geq 1+x \forall x \in \mathbb{R}$. On a alors

$$\begin{aligned} R_\varphi(\mu_X^N) &= \frac{1}{N^k} \sum_{(i_1, \dots, i_k) \in A_{N,k}} \varphi(x_{i_1}, \dots, x_{i_k}) + \frac{1}{N^k} \sum_{(i_1, \dots, i_k) \in B_{N,k}} \varphi(x_{i_1}, \dots, x_{i_k}) \\ &= \frac{1}{\#A_{N,k}} \sum_{(i_1, \dots, i_k) \in A_{N,k}} \frac{1}{\#C_{N,k}} \sum_{\sigma \in C_{N,k}(i_1, \dots, i_k)} \varphi(x_{i_1}, \dots, x_{i_k}) \\ &\quad + \left(\frac{1}{N^k} - \frac{1}{\#A_{N,k}}\right) \sum_{(i_1, \dots, i_k) \in A_{N,k}} \varphi(x_{i_1}, \dots, x_{i_k}) + \frac{1}{N^k} \sum_{(i_1, \dots, i_k) \in B_{N,k}} \varphi(x_{i_1}, \dots, x_{i_k}) \\ &= \frac{1}{\#\mathfrak{S}_N} \sum_{\sigma \in \mathfrak{S}_N} \varphi(x_{\sigma(1)}, \dots, x_{\sigma(k)}) \\ &\quad + \mathcal{O}\left(\left|\frac{1}{N^k} - \frac{1}{\#A_{N,k}}\right| \#A_{N,k} + \frac{1}{N^k} \#B_{N,k}\right) \|\varphi\|_\infty, \end{aligned}$$

ce qui prouve bien (9.3.6). On en déduit

$$\|\varphi \otimes \overline{1^{\otimes(N-k)}}\|_{L^\infty(E^N)} \leq \sup_{X \in E^N} R_\varphi(\mu_X^N) + \frac{2k^2}{N} \|\varphi\|_{L^\infty(E^k)}$$

et (9.3.7) à la limite $N \rightarrow \infty$. □

9.4 $\mathcal{C}_{\text{sym}}(E^N)$ lorsque $N \rightarrow \infty$

Supposons dans un premier temps E compact. Etant donnée une fonction $U \in C(\mathbf{P}(E))$ on définit pour tout N une fonction symétrique sur E^N en posant

$$\forall X \in E^N \quad u_N(X) := U(\mu_X^N). \quad (9.4.8)$$

On définit ainsi une application $\pi_C^N : C(\mathbf{P}(E)) \rightarrow C(E^N)$. Cette application est d'une certaine manière l'application "duale" de l'application π_E^N introduite à la définition de la mesure empirique.

Le résultat suivant concerne le problème inverse.

Proposition 9.4.7 *Pour toute fonction $u_N \in C(E^N)$ symétrique il existe une fonction $U_N \in C(\mathcal{P}_N(E))$ telle que (9.4.8). De plus, il existe une fonction $\bar{U}_N \in C(\mathbf{P}(E))$ telle que $\bar{U}_{|\mathcal{P}_N(E)} \equiv U_N$ et donc telle que (9.4.8). Plus précisément, fixons une distance D sur $\mathbf{P}(E)$ et notons d_D la distance sur E^N / \mathfrak{S}_N définie par $d_D(X, Y) = D(\mu_X^N, \mu_Y^N)$ pour tout $X, Y \in E^N$. Si u_N satisfait*

$$\forall X, Y \in E \quad |u_N(Y) - u_N(X)| \leq \omega(d_D(X, Y)),$$

pour un module de continuité ω , alors U_N et \bar{U}_N satisfont uniformément en N

$$\begin{aligned} \forall f, g \in \mathcal{P}_N(E) \quad & |U_N(g) - U_N(f)| \leq \omega(D(f, g)), \\ \forall f, g \in \mathbf{P}(E) \quad & |\bar{U}_N(g) - \bar{U}_N(f)| \leq \omega(D(f, g)). \end{aligned}$$

Preuve de la proposition (9.4.7). D'une part, il suffit de définir $U_N(v) := u_N(X)$ pour tout $v \in \mathcal{P}_N(E)$ en choisissant $X \in E^N$ quelconque tel que $\mu_X^N = v$. D'autre part, on définit \bar{U}_N grâce au Lemme de Tietze-Urysohn en prenant $K = \mathbf{P}(E)$, $A = \mathcal{P}_N(E)$. \square

Plus intéressant pour les applications, on a le résultat suivant.

Théorème 9.4.8 *Soit (u_N) une suite de fonctions telle que $u_N : E^N \rightarrow \mathbb{R}$ symétrique. On suppose qu'il existe C et un module de continuité (sous-linéaire) ω tels que*

$$\forall N, \forall X, Y \in E^N \quad |u_N(X)| \leq C, \quad |u_N(Y) - u_N(X)| \leq \omega(d_D(X, Y)).$$

Alors il existe une sous-suite, notée $(u_{N'})$, et une fonction $U \in C(\mathbf{P}(E))$ telle que

$$\sup_{X \in E^{N'}} |u_{N'}(X) - U(\mu_X^{N'})| \xrightarrow{N' \rightarrow \infty} 0. \quad (9.4.9)$$

On notera parfois $u_{N'} \rightarrow U$ fort dans $C_b(\mathcal{P}_{N'}(E))_{\forall N'}$.

Preuve 1 du Théorème (9.4.8). On définit une suite (U_N) de $C(\mathcal{P}_N(E))$ en posant

$$U_N(\mu_X^N) := u_N(X) \quad \forall \mu_X^N \in \mathcal{P}_N(E).$$

Comme $D(\mu_Y^N, \mu_X^N) = d_D(Y, X)$, U_N est continue avec module de continuité ω sur l'espace $(\mathcal{P}_N(E), D)$. On prolonge U_N à $(\mathbf{P}(E), W_2)$ en une fonction \bar{U}_N grâce à la proposition (9.4.7). Celle-ci a pour module de continuité ω et est bornée par $C + \omega(\text{diam}(\mathbf{P}(E)))$. Grâce au théorème d'Ascoli il existe $U \in C(\mathbf{P}(E))$ et une sous-suite qui converge vers U au sens de la norme sup :

$$\sup_{m \in \mathbf{P}(E)} |U(m) - \bar{U}_N(m)| \rightarrow 0 \quad \text{lorsque } N \rightarrow \infty.$$

On conclut en se restreignant à $m = \mu_X^N$.

Preuve 2 du Théorème. On considère d'abord des entiers du type $N = 2^n$, $n \in \mathbb{N}^*$, et on remarque qu'on a ainsi une injection (isométrique) $\mathbf{P}_{2^k}(E) \subset \mathbf{P}_{2^n}(E)$ si $k \leq n$. Avec les notations de la Proposition, on définit $U_{k,n} : \mathbf{P}_{2^k}(E) \rightarrow \mathbb{R}$ par $U_{k,n} := U_{N|_{\mathbf{P}_{2^k}(E)}}$ ou encore

$$\forall X \in E^{2^k} \quad U_{k,n}(\mu_X^{2^k}) = U_{2^n}(\mu_X^{2^k}) = u_{2^n}(\underbrace{X, \dots, X}_{2^{n-k} \text{ fois}}) =: u_{2^n}(\tilde{X}^n).$$

Par hypothèse, on a pour tout $\mu_X^{2^k}, \mu_Y^{2^k} \in \mathbf{P}_{2^k}(E)$

$$\begin{aligned} |U_{k,n}(\mu_X^{2^k}) - U_{k,n}(\mu_Y^{2^k})| &\leq \omega(w_p(\tilde{X}^n, \tilde{Y}^n)) = \omega(W_p(\mu_{\tilde{X}^n}^{2^n}, \mu_{\tilde{Y}^n}^{2^n})) \\ &= \omega(W_p(\mu_X^{2^k}, \mu_Y^{2^k})). \end{aligned}$$

En appliquant le théorème d'Ascoli à la suite de fonctions $(U_{k,n})_{n \geq 1}$ on déduit qu'il existe une sous-suite $(U_{k,n_\ell}^k)_{\ell \geq 1}$ avec $\{n_\ell^{k+1}, \ell \geq 1\} \subset \{n_\ell^k, \ell \geq 1\}$ et une fonction $U_k \in C(\mathbf{P}_{2^k}(E))$ telles que

$$\forall \mu_X^{2^k} \in \mathbf{P}_{2^k}(E) \quad U_{k,n_\ell^k}(\mu_X^{2^k}) \rightarrow U_k(\mu_X^{2^k}) \quad \text{lorsque } \ell \rightarrow \infty,$$

et pour tout $\mu_X^{2^k}, \mu_Y^{2^k} \in \mathbf{P}_{2^k}(E)$

$$|U_k(\mu_X^{2^k}) - U_k(\mu_Y^{2^k})| \leq \omega(W_p(\mu_X^{2^k}, \mu_Y^{2^k})). \quad (9.4.10)$$

De $\mathbf{P}_{2^k}(E) \subset \mathbf{P}_{2^{k+1}}(E)$ on tire que $U_{k+1}(\mu_X^{2^k}) = U_k(\mu_X^{2^k})$ pour tout $\mu_X^{2^k} \in \mathbf{P}_{2^k}(E)$. On définit alors $U \in C(\mathbf{P}(E))$ en posant $U(\rho) = U_k(\rho)$ si $\rho \in \mathbf{P}_{2^k}(E)$ et en l'étendant par continuité à $\mathbf{P}(E)$ grâce à la proposition précédente. La fonction U ainsi construite admet comme module de continuité ω pour la distance W_p . Enfin, par un procédé d'extraction diagonale, on peut construire une seule sous-suite U_{N_ℓ} , $N_\ell = 2^{n_\ell}$, telle que

$$\forall \rho \in \bigcup_{k \geq 1} \mathbf{P}_{2^k}(E) \quad |U_{N_\ell}(\rho) - U(\rho)| \rightarrow 0 \quad \text{lorsque } \ell \rightarrow \infty,$$

ce qui est bien le résultat attendu lorsque l'on traduit cela en terme de (u_N) . \square

Remarque 9.4.9 1. Pour $u \in C^1(E^N)$, on a

$$|u(Y) - u(X)| \leq \sum_{i=1}^N \|\nabla_i u\|_{L^\infty} |y_i - x_i|,$$

d'où on tire

$$\begin{aligned} (i) \quad \|\nabla_i u\|_{L^\infty} \leq \frac{C}{N} &\implies |u(Y) - u(X)| \leq \max \|\nabla_i u\|_{L^\infty} \sum_{i=1}^N |y_i - x_i| \leq C w_1(Y, X), \\ (ii) \quad \sum_i \|\nabla_i u\|_{L^\infty}^2 \leq \frac{C}{N} &\implies |u(Y) - u(X)| \leq \left(\sum_{i=1}^N \|\nabla_i u\|_{L^\infty}^2 \right)^{1/2} \left(\sum_{i=1}^N |y_i - x_i|^2 \right)^{1/2} \\ &\leq C w_2(Y, X), \\ (iii) \quad \sum_i \|\nabla_i u\|_{L^\infty} \leq C &\implies |u(Y) - u(X)| \leq \sum_{i=1}^N \|\nabla_i u\|_{L^\infty} \max_{1 \leq i \leq N} |y_i - x_i| \\ &\leq C w_\infty(Y, X), \end{aligned}$$

et bien sûr $(i) \Rightarrow (ii) \Rightarrow (iii)$.

2. On a pour tout $X \in E^N / \mathfrak{S}_N$

$$u^k(X) := \max_{1 \leq i \leq k} x_i \longrightarrow U(m_X), \quad \text{avec } U(m) := \sup(\text{supp } m).$$

3. Si (u^N) est une suite de fonctions continues symétriques justes bornées ($\exists C > 0 \forall N |u^N| \leq C$), alors les limites sup et les limites inf suivantes existent

$$U^*(m) = \limsup_{N \rightarrow \infty, m_X^N \rightarrow m} u^N(X), \quad U_*(m) = \liminf_{N \rightarrow \infty, m_X^N \rightarrow m} u^N(X).$$

Est-il clair que si $U^* = U_* =: U$ alors $U \in C(\mathbf{P}(E))$ et a posteriori (u_N) admet un module de continuité (au moins sur la sous-suite qui converge)?

On généralise au cas non compact la variante du résultat d'Ascoli déjà établie dans le cas d'un espace compact.

Théorème 9.4.10 (variante d'Ascoli). Soit $E = \mathbb{R}^d$ (ou E un espace polonais localement compact). Soit (u_N) une suite de $C_b(E^N)$ telle que

$$\forall X, Y \in E^N \quad |u_N(X)| \leq C, \quad |u_N(X) - u_N(Y)| \leq \omega(W_p(\mu_X^N, \mu_Y^N)),$$

pour une constante C , un module de continuité ω et un indice $p \in (0, \infty)$. Alors il existe une sous-suite (N') , une fonction $U \in C_b(\mathbf{P}_p(E) - W_p)$ telles que pour tout $k > p$, $a > 0$

$$\sup_{X \in E^{N'}, M_k(\mu_X^{N'}) \leq a} |u_{N'}(X) - U(\mu_X^{N'})| \rightarrow 0 \quad \text{lorsque } N' \rightarrow \infty.$$

Preuve du Théorème (9.4.10). On reprend la deuxième preuve du Théorème précédent. On considère des entiers du type $N = 2^n$, $n \in \mathbb{N}^*$, et on remarque qu'on a ainsi une injection (isométrique) $\mathcal{P}_{2^j}(E) \subset \mathcal{P}_{2^n}(E)$ si $j \leq n$. On définit $U_{j,n} : \mathcal{P}_{2^j}(E) \rightarrow \mathbb{R}$, par

$$\forall \mu_X^{2^j} \in \mathcal{P}_{2^j}(E) \quad U_{j,n}(\mu_X^{2^j}) = u_{2^n}(\underbrace{X, \dots, X}_{2^{n-j} \text{ fois}}) =: u_{2^n}(\tilde{X}^n),$$

et on note simplement $U_N = U_{n,n} \in C_b(\mathcal{P}_N(E))$. Par hypothèse, pour tout $\mu_X^{2^j}, \mu_Y^{2^j} \in \mathcal{P}_{2^j}(E)$, on a

$$\begin{aligned} |U_{j,n}(\mu_X^{2^j}) - U_{j,n}(\mu_Y^{2^j})| &\leq \omega(w_p(\tilde{X}^n, \tilde{Y}^n)) = \omega(W_p(\mu_{\tilde{X}^n}^{2^n}, \mu_{\tilde{Y}^n}^{2^n})) \\ &= \omega(W_p(\mu_X^{2^j}, \mu_Y^{2^j})). \end{aligned}$$

Pour tout j fixé, l'ensemble $\mathcal{P}_{2^j}(E) \cap \mathcal{B}\mathbf{P}_{k,a}$ étant compact (par exemple pour la distance W_p), on peut appliquer le théorème d'Ascoli à la suite de fonctions $(U_{j,n})_{n \geq 1}$ de $C(\mathcal{P}_{2^j}(E) \cap \mathcal{B}\mathbf{P}_{k,a})$, et on déduit qu'il existe une sous-suite $(U_{j,n_\ell})_{\ell \geq 1}$ avec $\{n_\ell^{j+1}, \ell \geq 1\} \subset \{n_\ell^j, \ell \geq 1\}$ et une fonction $U_j \in C(\mathcal{P}_{2^j}(E) \cap \mathcal{B}\mathbf{P}_{k,a})$ telles que

$$\forall \mu_X^{2^j} \in \mathcal{P}_{2^j}(E) \cap \mathcal{B}\mathbf{P}_{k,a} \quad U_{j,n_\ell}(\mu_X^{2^j}) \rightarrow U_j(\mu_X^{2^j}) \text{ lorsque } \ell \rightarrow \infty, \quad (9.4.11)$$

et pour tout $\mu_X^{2^j}, \mu_Y^{2^j} \in \mathcal{P}_{2^j}(E) \cap \mathcal{B}\mathbf{P}_{k,a}$

$$|U_j(\mu_X^{2^j})| \leq C, \quad |U_j(\mu_X^{2^j}) - U_j(\mu_Y^{2^j})| \leq \omega(W_p(\mu_X^{2^j}, \mu_Y^{2^j})). \quad (9.4.12)$$

De $\mathcal{P}_{2^j}(E) \subset \mathcal{P}_{2^{j+1}}(E)$ on tire que $U_{j+1}(\mu_X^{2^j}) = U_j(\mu_X^{2^j})$ pour tout $\mu_X^{2^j} \in \mathcal{P}_{2^j}(E)$. On définit alors $U \in C_b(\mathbf{P}_p(E)_{-W_p})$ en posant $U(\rho) = U_j(\rho)$ si $\rho \in \mathcal{P}_{2^j}(E)$ et en l'étendant par continuité à $\mathbf{P}_k(E)$, ou même $\mathbf{P}_p(E)$, grâce à ce qui précède et au Théorème de densité de Krein-Milman. La fonction U ainsi construite admet comme module de continuité ω pour la distance W_p . Enfin, par un procédé d'extraction diagonale, on peut construire une seule sous-suite, notée $U_{N'}$, $N' = 2^{n'}$, telle que

$$\forall \rho \in K_a := \bigcup_{j \geq 1} \mathcal{P}_{2^j}(E) \cap \mathcal{B}\mathbf{P}_{k,a} \quad |U_{N'}(\rho) - U(\rho)| \rightarrow 0 \text{ lorsque } N' \rightarrow \infty.$$

Cette dernière convergence est en fait uniforme sur $K_a^{n'}$ avec $K_a^j := \mathcal{P}_{2^j} \cap \mathcal{B}\mathbf{P}_{k,a}$. En effet, on écrit

$$\begin{aligned} \sup_{\rho \in K_a^{N'}} |U_{N'}(\rho) - U(\rho)| &\leq \sup_{\eta \in K_a^{N'}} \mathbf{inf}_{\eta \in K_a^j} \{ |U_{N'}(\rho) - U_{N'}(\eta)| + |U_{N'}(\eta) - U(\eta)| + |U(\rho) - U(\eta)| \} \\ &\leq \sup_{\eta \in K_a^j} |U_{N'}(\eta) - U(\eta)| + 2 \sup_{\rho \in \mathcal{B}\mathbf{P}_{k,a}} \omega(W_p(\rho, K_a^j)), \end{aligned}$$

et le premier terme tend vers 0 lorsque N tend vers l'infini pour tout j fixé (c'est une des conséquences du théorème d'Ascoli appliqué dans $C(\mathcal{P}_{2^j}(E) \cap \mathcal{B}\mathbf{P}_{k,a})$) alors que le second terme tend vers 0 lorsque j tend vers l'infini (c'est la conclusion du Théorème de Krein-Milman). \square

9.5 Calcul différentiel sur $\mathbf{P}(E)$

9.5.1 Quatre définitions possibles

Il existe un grand nombre de définitions de la notion de différentiabilité et de classe de régularité possible. Nous en donnons quatre définitions possibles, chacune pouvant conduire à des

variantes. Le point essentiel à comprendre et qui explique la difficulté de donner une seule définition est que d'une part il existe de nombreuses distances topologiquement équivalentes mais non métriquement équivalentes, la notion de différentiabilité dépendra donc très fortement du choix de la métrique considérée, et que d'autre part, pour une distance donnée il n'existe pas en général de norme à laquelle elle serait associée. Cela fait que $\mathbf{P}(E)$ possède naturellement une structure de variété métrique et seulement parfois une structure de variété métrique plongée dans un espace vectoriel normé.

Définition 9.5.1 Soient $\mathcal{O} \subset \mathbf{P}(E)$ muni d'une distance D et $U : \mathcal{O} \rightarrow \mathbb{R}$.

1. On dit que U est différentiable en $f \in \mathcal{O}$ si pour tout chemin $\gamma = (\gamma_t)_{t \in [0,1]}$ sur \mathcal{O} optimal et de vitesse constante au sens de D et issu de f , on a

$$\langle DU(f), \dot{\gamma}(0) \rangle := \lim_{t \rightarrow 0} \frac{1}{t} \{u(\gamma_t) - u(f)\} \text{ existe.}$$

Ici, le sens du terme de gauche est juste la limite définie à droite de l'égalité. En particulier, on ne donne pas de sens précis au "vecteur tangent" $\dot{\gamma}(0)$ ni à la dualité $\langle \cdot, \cdot \rangle$. Le sens à donner au fait que γ soit un chemin optimal et de vitesse constante issu de f est que

$$\forall 0 \leq s < t \leq 1 \quad D(\gamma_s, \gamma_t) = (t-s)D(f, \gamma_1), \quad \gamma_0 = f.$$

2. On dit que u est continûment et uniformément différentiable sur \mathcal{O} , on note $U \in UC^1(\mathcal{O})$, si U est différentiable en tout point $f \in \mathcal{O}$, pour tout $f, g \in \mathcal{O}$ il existe un unique chemin optimal $\gamma = \gamma_{f,g}$ joignant f à g (au sens où $\gamma(0) = f$ et $\gamma(1) = g$), l'application $\mathcal{O} \times \mathcal{O} \rightarrow \mathbb{R}$, $(f, g) \mapsto \langle DU(f), \dot{\gamma}_{f,g}(0) \rangle$ est continue et

$$|U(g) - U(f) - \langle DU(f), \dot{\gamma}_{f,g}(0) \rangle| \leq \Omega(D(f, g)),$$

avec $\Omega(s)/s := \omega(s) \rightarrow 0$ lorsque $s \rightarrow 0$.

La différentiabilité au sens de cette définition est de loin la plus classique. Elle n'est rien d'autre que celle introduite par F. Otto et largement étudiée dans son ouvrage dans le cas $D = W_p$ avec $p \in (1, \infty)$ ou $p = 2$ et avec $\mathcal{O} = \mathbf{P}_p(E)$ ou $\mathcal{O} = L_p^1(E)$.

Définition 9.5.2 Soient un ensemble convexe $E \subset \mathbb{R}^d$, un ensemble convexe $\mathcal{O} \subset \mathbf{P}(E)$ muni de la distance $D := W_p$, $1 \leq p \leq \infty$, et $U : \mathbf{P}(E) \rightarrow \mathbb{R}$.

1. On dit que U est "différentiable" en $f \in \mathcal{O}$ si U est "directionnellement différentiable" en $f \in \mathcal{O}$, au sens où il existe un espace de probabilité $(\Omega, \mathcal{A}, \mathbb{P})$ et une variable aléatoire $\mathcal{X}_0 \in L^p(\Omega, E)$ telle que sa loi $\mathcal{L}(\mathcal{X}_0)$ satisfait $\mathcal{L}(\mathcal{X}_0) = f$, tels que la fonction $\mathcal{U} : L^p(\Omega, E) \rightarrow \mathbb{R}$, $\mathcal{Y} \mapsto \mathcal{U}(\mathcal{Y}) := U(\mathcal{L}(\mathcal{Y}))$ admette des dérivées directionnelles en \mathcal{X}_0 : pour tout $\mathcal{Y} \in L^p(\Omega, E)$ on a

$$\mathcal{U}'(\mathcal{X}_0; \mathcal{Y} - \mathcal{X}_0) := \lim_{t \rightarrow 0} \frac{1}{t} \{\mathcal{U}((1-t)\mathcal{X}_0 + t\mathcal{Y}) - \mathcal{U}(\mathcal{X}_0)\} \text{ existe,}$$

et tels que enfin il existe $\Xi_{\mathcal{X}_0} \in L^{p'}$ de sorte que $\mathcal{U}'(\mathcal{X}_0; \mathcal{Y} - \mathcal{X}_0) = \mathbb{E}(\langle \Xi_{\mathcal{X}_0}, \mathcal{Y} - \mathcal{X}_0 \rangle)$ pour tout $\mathcal{Y} \in L^p$, où ici $\langle \cdot, \cdot \rangle$ désigne le produit scalaire euclidien de E et \mathbb{E} désigne l'espérance associée à \mathbb{P} .

2. On dit que U est continûment et uniformément différentiable sur \mathcal{O} s'il existe existe $\xi : \mathbf{P}_p(E) \times E \rightarrow E$ continue telle que U est différentiable en tout point $f \in \mathcal{O}$, et pour tout $\mathcal{X} \in L^p(\Omega)$, $\mathcal{L}(\mathcal{X}) = f$, l'application \mathcal{U} définie ci-dessus est différentiable en \mathcal{X} , $\Xi_{\mathcal{X}} = \xi(f, \mathcal{X})$ et

$$|U(g) - U(f) - \mathbb{E}(\langle \xi(f, \mathcal{X}), \mathcal{Y} - \mathcal{X} \rangle)| \leq \Omega(D(f, g)).$$

Cette définition est due à P.-L. Lions. Il est assez instructif de la réécrire uniquement en termes de mesures de probabilité. La notion de différentiabilité devient : U est différentiable en $f \in \mathcal{O}$, si pour tout plan de transport $\pi \in \mathbf{P}(E \times E)$ tel que le chemin (π_t) défini par

$$\langle \pi_t, \varphi \rangle := \int_{E \times E} \varphi((1-t)x + ty) \pi(dx, dy)$$

vérifie $\pi_t \in \mathcal{O}$ pour tout $t \in (0, 1]$, $\pi_0 = f$ on a

$$\frac{1}{t} \lim_{t \rightarrow 0} \frac{1}{t} \{u(\pi_t) - u(f)\} \text{ existe.}$$

De plus, si U est continûment différentiable, la notion de "dérivée directionnelle" peut se réécrire en termes de mesures de probabilités de la façon suivante : pour tout $f, g \in \mathbf{P}_p(E)$, pour tout plan de transport $\pi \in \Pi(f, g)$, et toutes variables aléatoires $\mathcal{X}, \mathcal{Y} \in L^p(\Omega; E)$ avec $\mathcal{L}(\mathcal{X}) = f$, $\mathcal{L}(\mathcal{Y}) = g$, on a

$$\mathcal{U}'(\mathcal{X}; \mathcal{Y} - \mathcal{X}) = \mathbb{E}(\langle \Xi_{\mathcal{X}}, \mathcal{Y} - \mathcal{X} \rangle) = \int_{E^2} \langle \xi(f, x), y - x \rangle \pi(dx, dy).$$

Définition 9.5.3 Soient un ensemble convexe $E \subset \mathbb{R}^d$, un ensemble convexe $\mathcal{O} \subset \mathbf{P}(E)$ muni de la distance W_p , $1 \leq p \leq \infty$, et $U : \mathbf{P}(E) \rightarrow \mathbb{R}$.

2. On dit que U est continûment et uniformément différentiable sur \mathcal{O} , on note $U \in UC^1(\mathcal{O})$, si l'application $u^N : E^N \rightarrow \mathbb{R}$ définie par $u^N(X) := U(\mu_X^N)$ est de classe C^1 et il existe une constante C telle que

$$|u^N(X+Y) + u^N(X-Y) - 2u^N(X)| \leq \Omega\left(D(\mu_X^N, \mu_{X+Y}^N) + D(\mu_X^N, \mu_{X-Y}^N)\right).$$

Définition 9.5.4 Soient $E \subset \mathbb{R}^d$, un ensemble convexe $\mathcal{O} \subset \mathbf{P}(E)$, \mathcal{G} un espace vectoriel normé tel que $\mathcal{O} - \mathcal{O} \subset \mathcal{G}$, et $U : \mathcal{O} \rightarrow \mathbb{R}$.

2. On dit que U est continûment et uniformément différentiable sur \mathcal{O} , on note $U \in UC^1(\mathcal{O})$, s'il existe $DU : \mathcal{O} \rightarrow \mathcal{G}'$ telle que

$$|U(g) - U(f) - \langle DU(f), g - f \rangle| \leq \Omega(\|g - f\|_{\mathcal{G}}).$$

On peut commencer par un lemme facile qui permet de relier ces différentes notions.

Lemme 9.5.5 0. Si U est différentiable en un des sens ci-dessus pour une distance/norme alors U est différentiable dans le même sens pour toute distance/norme plus forte (au sens d'une inégalité Lipschitz).

1. Si $\varphi \in \text{Lip}(E^k)$ alors le polynôme R_φ satisfait $R_\varphi \in UC^1_{(4)}(\mathcal{O})$ avec $\mathcal{O} := \mathbf{P}_1(E)$, $\mathcal{G} := (\text{Lip}_0(E))^'$.
2. Si $U \in UC^1_{(4)}(\mathcal{O})$ alors $U \in UC^1_{(2)}(\mathcal{O})$ et $U \in UC^1_{(3)}(\mathcal{O})$.
3. Si $U \in UC^1_{(2)}(\mathcal{O})$ alors $U \in UC^1_{(1)}(\mathcal{O})$.

9.5.2 Deuxième définition (relèvement dans un espace de variables aléatoires)

Nous nous intéressons pour le moment au cas $E = \mathbb{R}^d$ et $p = 2$. Commençons par deux remarques.

- Pour une variable aléatoire \mathcal{X} définie sur Ω , dire que $\mathcal{X} \in L^2(\Omega)$ est équivalent à dire que $\mathcal{L}(\mathcal{X}) \in \mathbf{P}_2(E)$.

- Pour un monôme¹ $U = R_\varphi$ défini par

$$U(f) := \int_{E^\ell} \varphi(X) f^{\otimes \ell}(dX), \quad \varphi := \varphi_1 \otimes \dots \otimes \varphi_\ell, \quad \varphi_i \in C_b(E^\ell; \mathbb{R}),$$

1. quelle serait l'écriture pour un polynôme?

la fonction \mathcal{U} est simplement définie par

$$\mathcal{U}(\mathcal{X}) := \prod_{i=1}^{\ell} \mathbb{E}(\varphi_i(\mathcal{X})) \quad \forall \mathcal{X} \in L^2(\Omega; E).$$

Si $\varphi_i \in W^{1,\infty}(E)$, alors R_φ est différentiable sur $\mathbf{P}_2(E)$. En effet, il suffit de le voir lorsque $\ell = 1$. Dans ce cas, pour tout $f, g \in \mathbf{P}_2(E)$, $\pi \in \Pi(f, g)$, $\mathcal{L}(\mathcal{X}, \mathcal{Y}) = \pi$, on a

$$\begin{aligned} \lim_{t \rightarrow 0} \frac{1}{t} \{ \mathcal{U}((1-t)\mathcal{X} + t\mathcal{Y}) - \mathcal{U}(\mathcal{X}) \} &= \\ &= \lim_{t \rightarrow 0} \mathbb{E} \left(\frac{1}{t} \{ \varphi((1-t)\mathcal{X} + t\mathcal{Y}) - \varphi(\mathcal{X}) \} \right) \\ &= \lim_{t \rightarrow 0} \int_{E^2} \left(\frac{1}{t} \{ \varphi((1-t)x + ty) - \varphi(x) \} \right) \pi(dx, dy) \\ &= \int_{E^2} \langle \nabla \varphi(x), y - x \rangle \pi(dx, dy), \end{aligned}$$

de sorte que $\xi(f, x) := \nabla \varphi(x)$ et $\Xi_{\mathcal{X}} := \xi(f, \mathcal{X}) \in L^2(\Omega; E)$ (et même ici $\Xi_{\mathcal{X}} \in L^\infty(\Omega; E)$). Dans le cas $\ell \geq 1$, on en déduit

$$\xi(f, x) := \sum_{i=1}^{\ell} \prod_{j \neq i} \langle f, \varphi_j \rangle \nabla \varphi_i(x).$$

On énonce les théorèmes suivants liés à la loi gradient.

Théorème 9.5.6 Soit $U : \mathbf{P}_2(E) \rightarrow \mathbb{R}$ différentiable en un point $f \in \mathbf{P}_2(E)$. Alors pour tout $\mathcal{X} \in L^2(\Omega)$ tel que $\mathcal{L}(\mathcal{X}) = f$, la fonction $\mathcal{U} : L^2 \rightarrow \mathbb{R}$ est Fréchet différentiable en X et $\mathcal{U}'(\mathcal{X}) \in L^2(\Omega)$ ne dépend pas de \mathcal{X} .

Théorème 9.5.7 Soit $U : \mathbf{P}_2(E) \rightarrow \mathbb{R}$ de classe C^1 , $f \in \mathbf{P}_2(E)$. Alors, il existe $\xi \in L^2(E, df; E)$ tel que pour tout $\mathcal{X} \in L^2(\Omega)$, $\mathcal{L}(\mathcal{X}) = f$ on a $D\mathcal{U}[\mathcal{X}] = \xi(\mathcal{X})$ \mathbb{P} -p.s. On peut donc noter $\xi = \xi(f, \cdot)$.

Chapitre 10

A2. Résultats techniques

Pour plus de détails, voir principalement référence [17] en bibliographie.

10.1 Un théorème de type Tietze-Urysohn

Lemme 10.1.8 (variante de Tietze-Urysohn). Soit (K, d_K) un espace compact, $A \subset K$ fermé, $u : A \rightarrow \mathbb{R}$ une fonction continue de module de continuité ω (que l'on peut supposer sous-linéaire car K est compact). Alors il existe $\bar{u} : K \rightarrow \mathbb{R}$ continue, de module de continuité ω et telle que $\bar{u}|_A = u$. Ici on ne prétend pas $\|\bar{u}\|_\infty = \|u\|_\infty$ mais seulement $\|\bar{u}\|_\infty \leq \|u\|_\infty + \omega(\text{diam}(K))$.

Preuve du Lemme 10.1.8. On définit

$$\bar{u}(x) := \inf_{a \in A} [u(a) + \omega(d_K(x, a))].$$

Alors pour tout $x, y \in K$ et si l'infimum est atteint en $a \in A$ dans la définition de $\bar{u}(x)$, on a

$$\bar{u}(y) \leq u(a) + \omega(d_K(y, a)) \leq u(a) + \omega(d_K(x, a)) + \omega(d_K(x, y)) \leq \bar{u}(x) + \omega(d_K(x, y)).$$

En inversant les rôles de x et y , on trouve $|\bar{u}(y) - \bar{u}(x)| \leq \omega(d_K(y, x))$. Enfin, lorsque $x \in A$, on a $u(x) \leq u(a) + \omega(d_K(x, a))$ pour tout $a \in A$, et donc $\bar{u}(x) = u(x)$. \square

10.2 Le théorème de Stone-Weierstrass

Théorème 10.1 (Le théorème de Stone-Weierstrass)

Soit E un espace métrique compact et $\mathcal{A} \subset C(E; \mathbb{R})$. On suppose que

- (i) \mathcal{A} est une algèbre (stable par addition, multiplication et multiplication par un scalaire);
- (ii) \mathcal{A} contient les constantes;
- (iii) \mathcal{A} sépare les points de E (pour tout $x, y \in E$, $x \neq y$, il existe $\varphi \in \mathcal{A}$ telle que $\varphi(x) \neq \varphi(y)$).

Alors \mathcal{A} est dense (au sens de la convergence uniforme) dans $C(E, \mathbb{R})$.

10.3 Krein-Milman et Choquet

Définition 10.3.9 Soit X un espace vectoriel et $K \subset X$ un convexe compact.

(i) On dit que $x \in K$ est un point extrémal de K si pour tout $y, z \in K$, $t \in (0, 1)$ on a $(1-t)y + tz = x$ implique $y = z = x$. On note $\mathcal{E}(K)$ l'ensemble des points extrémaux de K .

(ii) Plus généralement, on dit que $M \subset K$ est un ensemble extrémal de K si M est convexe et compact, et pour tout $y, z \in K$, $t \in (0, 1)$ on a $(1 - t)y + tz \in M$ implique $y, z \in M$.

(iii) Ainsi un singleton $M = \{x\}$ est un ensemble extrémal si, et seulement si, x est un point extrémal.

Lemme 10.3.10 (de Krein-Milman). Soit X un evn et $K \subset X$ un convexe compact.

a) - Pour tout $f \in X'$, les ensembles

$$K_f^+ := \{x, f(x) = \max_{y \in K} f(y)\}, \quad K_f^- := \{x, f(x) = \min_{y \in K} f(y)\}.$$

sont des ensembles extrémaux.

b) - Tout ensemble extrémal $M \subset K$ contient un point extrémal.

Idée fondamentale. L'idée fondamentale est de couper un ensemble convexe compact non vide K par des hyperplans $H_\alpha := [f = \alpha]$ pour tout $\alpha \in \mathbb{R}$, et tout $f \in X' \setminus \{0\}$ fixé. Alors les deux hyperplans extrêmes pour lesquels l'intersection n'est pas vide sont des sous-ensembles extrémaux K_f^\pm . En itérant, on trouve ainsi des singletons qui sont des points extrémaux.

Preuve du Lemme de Krein-Milman. L'assertion a) est immédiate et il suffit de montrer l'assertion b) pour l'ensemble extrémal K . On note \mathcal{M} la collection des ensembles compacts extrémaux de K . On a $K \in \mathcal{M}$. Soit (M_i) une famille totalement ordonnée (pour l'inclusion) d'éléments de \mathcal{M} . Alors $\bar{M} := \bigcap M_i$ est un compact, convexe, non vide. C'est également un ensemble extrémal qui est un minorant de la famille (M_i) . \mathcal{M} est donc un ensemble inductif, et par le lemme de Zorn, \mathcal{M} admet au moins un élément minimal, notons le M_1 . Montrons que $M_1 = \{x_1\}$. En effet, s'il existe $x_2 \in M_1$, $x_2 \neq x_1$, il existe $f \in X'$ qui sépare $\{x_1\}$ et $\{x_2\}$ et donc on peut supposer $\langle f, x_1 \rangle < \langle f, x_2 \rangle$. On définit

$$M_0 := \{x \in M_1; \langle f, x \rangle = m_1 := \inf_{y \in M_1} \langle f, y \rangle\}.$$

On a $M_0 \subset M_1$, $M_0 \neq M_1$ (puisque $x_2 \notin M_0$), M_0 est non vide, convexe, compact, et M_0 est un ensemble extrémal. En effet, si $x, y \in K$, $t \in (0, 1)$ et $(1 - t)x + ty \in M_0 \subset M_1$ alors $x, y \in M_1$ (puisque M_1 est extrémal) et alors, par définition de M_0 , on doit avoir $\langle f, x \rangle = \langle f, y \rangle = m_1$, ce qui implique $x, y \in M_0$. Cela contredit la "minimalité" de M_1 . En conclusion M_1 est un singleton. \square

Théorème 10.2 (Krein-Milman cadre abstrait)

Soit K un convexe compact (non vide) d'un evn X . Alors K est l'adhérence de l'enveloppe convexe de l'ensemble (non vide!) de ses points extrémaux. : $K = \overline{\text{conv}(\mathcal{E}(K))}$.

Preuve du théorème de Krein-Milman. L'inclusion $\overline{\text{conv}(A)} \subset K$ est claire. Supposons par l'absurde qu'il existe $b \in K \setminus \overline{\text{conv}(A)}$. Alors, par la forme géométrique du théorème de Hahn-Banach, il existe $f \in X'$, $\ell \in \mathbb{R}$ telle que

$$\forall a \in \overline{\text{conv}(A)} \quad \langle f, a \rangle \leq \ell < \langle f, b \rangle.$$

On définit $L := \{x \in K; f(x) = m := \sup_{y \in K} f(y)\}$. De $m \geq \langle f, b \rangle > \ell$ on déduit que $L \cap A = \emptyset$. De plus, il est clair que L est un sous-ensemble compact convexe (non vide) de K , et est un ensemble extrémal de K . Par le lemme de Krein-Milman il existe $c \in L$ élément extrémal de L , donc également élément extrémal de K . Cela implique $c \in A \cap L$, ce qui est absurde. \square

Corollaire 10.3.11 (Krein-Milman pour les mesures). Soit E un espace métrique compact. Les masses de Dirac sont les éléments extrémaux de $\mathbf{P}(E)$ et donc toute mesure de probabilité est limite faible d'une suite de combinaisons convexes de masses de Dirac.

Corollaire 10.3.12 (Krein-Milman pour $\mathcal{M}\mathcal{B}_N$). L'ensemble des matrices de permutations $\mathcal{M}\mathcal{P}_N$ forme l'ensemble des points extrémaux de l'ensemble des matrices bistochastiques \mathcal{B}_N , et donc toute matrice bistochastique est limite faible d'une suite de combinaisons convexes de matrices de permutations.

On peut préciser le théorème de Krein-Milman grâce au résultat suivant.

Théorème 10.3 (Choquet)

Soit K un convexe compact d'un evn X tel que l'ensemble des points extrémaux $\mathcal{E}(K)$ est fermé (donc compact). Pour tout $x \in K$ il existe une mesure $\hat{\pi}_x \in \mathbf{P}(\mathcal{E}(K))$ telle que

$$x = \int_{\mathcal{E}(K)} y \hat{\pi}_x(dy).$$

Preuve du théorème de Choquet. On considère une suite (x_n) telle que $x_n \in \text{conv}(\mathcal{E}(K))$ et $x_n \rightarrow x$. La condition $x_n \in \text{conv}(\mathcal{E}(K))$ s'écrit également

$$\exists \pi_n \in \mathbf{P}(\mathcal{E}(K)) \quad x_n = \int_{\mathcal{E}(K)} y \pi_n(dy)$$

Comme $\mathbf{P}(\mathcal{E}(K))$ est compact, on peut extraire une sous-suite $(\pi_{n'})$ telle que $\pi_{n'} \rightarrow \pi$ faiblement dans $\mathbf{P}(\mathcal{E}(K))$ pour un certain $\pi \in \mathbf{P}(\mathcal{E}(K))$. On conclut en passant à la limite dans la formule de représentation ci-dessus, et on obtient donc $\pi_x := \pi$. \square

Le théorème de Hewitt et Savage est alors un "avatar" du théorème de Choquet. En effet, pour tout $\pi \in \mathbf{P}_{\text{sym}}(E^{\mathbb{N}})$ il existe d'après le théorème de Choquet une mesure de probabilité $\hat{\pi} \in \mathbf{P}(\mathcal{E}(\mathbf{P}_{\text{sym}}(E^{\mathbb{N}})))$ sur les points extrémaux de $\mathbf{P}_{\text{sym}}(E^{\mathbb{N}})$ telle que

$$\pi = \int_{\mathcal{E}(\mathbf{P}_{\text{sym}}(E^{\mathbb{N}}))} e \hat{\pi}(de) = \int_{\mathbf{P}(E)} \rho^{\otimes \infty} \hat{\pi}(d\rho),$$

la première égalité provenant du $\mathcal{E}(\mathbf{P}_{\text{sym}}(E^{\mathbb{N}}))$ est un fermé de $\mathbf{P}_{\text{sym}}(E^{\mathbb{N}})$ et la deuxième égalité reposant sur le fait que $\mathbf{P}(E) \approx \mathcal{E}(\mathbf{P}_{\text{sym}}(E^{\mathbb{N}}))$ via l'application $\rho \mapsto \rho^{\otimes \mathbb{N}}$.

10.4 Le compactifié d'Alexandroff

La procédé de compactification d'Alexandroff permet partant d'un espace non compact E de le rendre compact en adjoignant un "point à l'infini" ∞ à l'espace E .

Théorème 10.4 (Compactifié d'Alexandroff)

Soit (E, d) un espace métrique localement compact, non compact, séparable. Soit ∞ un point (à l'infini) n'appartenant pas à E . Alors, on peut munir l'ensemble $\hat{E} = E \cup \{\infty\}$ d'une distance \hat{d} telle que (\hat{E}, \hat{d}) est un espace métrique compact et $\hat{d}|_E$ définit la même topologie que d sur E .

Preuve du théorème d'Alexandroff. Etape 1. On définit la topologie $\hat{\mathcal{T}}$ de \hat{E} par

$$\hat{\mathcal{T}} = \{\mathcal{O}, \mathcal{O} \subset E \text{ ouvert}\} \cup \{\hat{E} \setminus K, K \subset E \text{ compact}\}.$$

On a clairement que $\hat{\mathcal{T}}|_E$ est la topologie \mathcal{T} de E , et que $(\hat{E}, \hat{\mathcal{T}})$ est un espace (séparé et) compact et séparable.

- \hat{E} est séparé :** Soit x_1, x_2 deux points distincts de \hat{E} . Si $x_1, x_2 \in E$ alors $\exists \mathcal{O}_1, \mathcal{O}_2$ deux ouverts disjoints tels que $x_1 \in \mathcal{O}_1$ et $x_2 \in \mathcal{O}_2$ car E est séparé. Si $x_1 \in E$ et $x_2 = \infty$ il existe un ouvert $\mathcal{O}_1 \subset E$ tel que $\bar{\mathcal{O}}_1$ est compact. Ainsi $\mathcal{O}_2 = \hat{E} \setminus \bar{\mathcal{O}}_1$ est un ouvert de \hat{E} qui contient x_2 et $\mathcal{O}_1 \cap \mathcal{O}_2 = \emptyset$.

2. **\hat{E} est compact** : Soit $\{\mathcal{O}_n\}_{n \in \mathbb{N}}$ un recouvrement de \hat{E} par des ouverts. Au moins un de ces ensembles, disons \mathcal{O}_N , $N \in \mathbb{N}$, contient ∞ . Par définition de $\hat{\mathcal{T}}$, on a $\mathcal{O}_N = \hat{E} \setminus K$ pour un ensemble compact $K \subset E$. Comme K est compact, il existe un ensemble fini $J \subset \mathbb{N}$ tel que $\{\mathcal{O}_j\}_{j \in J}$ recouvre K . Ainsi $\mathcal{O}_N \cup \{\mathcal{O}_j\}_{j \in J}$ est un recouvrement fini de \hat{E} .

Etape 2. Soit (z_p) une suite dénombrable et dense de E et soit (\mathcal{O}_j) une base dénombrable d'ouverts relativement compacts de E de la forme

$$\{\mathcal{O}_j, j \in \mathbb{N}\} = \{B(z_p, 2^{-q}), p \in \mathbb{N}^*, q \in \mathbb{N}^*\},$$

plus précisément, on ne considérera que les couples (p, q) tels que de plus $B(z_p, 2^{1-q})$ est un compact de E . Pour tout $j \in \mathbb{N}$, $\mathcal{O}_j = B(z_{p_j}, 2^{-q_j})$, choisissons une fonction $\phi_j \in C_c(E)$ telle que

$$0 \leq \phi_j \leq 1, \quad \phi_j \equiv 1 \text{ sur } \mathcal{O}_j, \quad \text{supp } \phi_j \subset B(z_{p_j}, 2^{1-q_j})$$

et posons $\phi_j(\infty) := 0$. On définit la distance

$$\forall x, y \in \hat{E} \quad \hat{d}_E(x, y) := \sum_{j=0}^{\infty} 2^{-j} |\phi_j(x) - \phi_j(y)|.$$

On vérifie alors sans trop de difficulté que la topologie associée à \hat{d} est $\hat{\mathcal{T}}$ et que la topologie associée à $\hat{d}|_{E \times E}$ est celle de E . \square

10.5 Espace produit

On se donne un espace polonais E muni de sa distance $d = d_E$, de sa topologie induite $\mathcal{O} = \mathcal{O}_E$ et de sa tribu borélienne $\mathcal{B} = \mathcal{B}_E$.

Espace produit. On définit $E^\infty = E^{\mathbb{N}}$ l'espace produit (des suites de E) : $X \in E^{\mathbb{N}}$ si $X = (x_n)_{n \in \mathbb{N}}$ avec $x_n \in E$.

Distance produit. On définit la fonction $D : E^{\mathbb{N}} \times E^{\mathbb{N}} \rightarrow \mathbb{R}_+$ par

$$D(X, Y) = \sum_{n=0}^{\infty} \frac{1}{2^n} \min\{d(x_n, y_n), 1\}, \quad X = (x_n)_{n \in \mathbb{N}}, \quad Y = (y_n)_{n \in \mathbb{N}}.$$

Alors D est une distance.

Topologie produit. On définit la topologie produit \mathcal{O}_1 sur E^∞ comme la famille des ensembles de la forme C_O lorsque $O \in \mathcal{O}_{E^k}$ et $k \geq 1$. On définit la topologie produit \mathcal{O}_2 sur E^∞ comme la topologie induite par la distance D . On montre alors $\mathcal{O}_1 = \mathcal{O}_2$, topologie que l'on note \mathcal{O}_{E^∞} . Si E est compact alors E^∞ est compact et si E est polonais alors E^∞ est polonais.

Tribu produit. On définit la tribu produit \mathcal{T}_1 sur E^∞ comme étant la tribu engendrée par les ensembles $A = A_0 \times A_1 \times \dots \subset E^\infty$ avec $A_n \in \mathcal{B}$ et il existe J sous-ensemble fini de \mathbb{N} tel que $A_n = E \forall n \notin J$. On définit la tribu produit \mathcal{T}_2 sur E^∞ comme la tribu engendrée par la distance produit D / la topologie produit \mathcal{B}_{E^∞} . On montre alors $\mathcal{T}_1 = \mathcal{T}_2$, tribu que l'on note \mathcal{B}_{E^∞} .

Probabilités sur l'espace produit. On note $\Pi_k : E^{\mathbb{N}} \rightarrow E^k$, $k < N \leq \infty$ l'application qui à $X = (x_n)_{n \in \mathbb{N}} \in E^{\mathbb{N}}$ associe $\Pi_k X = (x_n)_{n \leq k} \in E^k$. On dit qu'une suite $(\pi_N) \in \mathbf{P}(E^{\mathbb{N}})$, $N \in \mathbb{N}^*$, est compatible si $\Pi_k \pi_N = \pi_k$ si $N \geq k$ où par définition

$$\Pi_k \pi_N \in \mathbf{P}(E^k) \quad (\Pi_k \pi_N)(A_1 \times \dots \times A_k) = \pi_N(A_1 \times \dots \times A_k \times E \times \dots).$$

Il est clair que pour toute mesure de probabilité $\pi \in \mathbf{P}(E^\infty)$ on définit une famille de mesures de probabilité compatibles en posant $\pi_n := \Pi_n \pi$. Inversement, le théorème de Kolmogorov affirme qu'étant donnée une famille (π_n) de mesures de probabilité compatibles il existe une (unique) mesure de probabilité $\pi \in \mathbf{P}(E^\infty)$ telle que $\pi_n := \Pi_n \pi$ pour tout $n \geq 1$.

Convergence faible sur l'espace des probabilités sur l'espace produit.

Lemme 10.5.13 *Pour une suite (α^j) de $\mathbf{P}(E^\infty)$ et $\alpha \in \mathbf{P}(E^\infty)$ il y a équivalence entre*

- (1) $\alpha^j \rightarrow \alpha$ au sens de la convergence faible de $\mathbf{P}(E^\infty)$;
- (2) $\Pi_k(\alpha^j) \rightarrow \Pi_k(\alpha)$ au sens de la convergence faible de $\mathbf{P}(E^k)$.

Preuve du Lemme 10.5.17. Comme il n'est pas forcément pratique de définir (1) à l'aide de la définition usuelle

$$\langle \alpha^j, \Phi \rangle \rightarrow \langle \alpha, \Phi \rangle \quad \forall \Phi \in C_b(E^\infty),$$

nous utilisons le critère (ii) du Théorème de convergence faible. On a évidemment (c'est la définition de la topologie \mathcal{O}_{E^∞}) l'équivalence suivante

- (1') $\liminf \alpha^j(C_O) \geq \alpha(C_O)$ pour tout $C_O \in \mathcal{O}_{E^\infty}$;
- (2') $\liminf \alpha_k^j(O) \geq \alpha_k(O)$ pour tout $O \in \mathcal{O}_{E^k}$ et tout $k \geq 1$.

Et on conclut en utilisant l'équivalence (i) \Leftrightarrow (ii) dans le Théorème de convergence faible. \square

(i) *Etant donné un espace polonais E , on définit $E^\infty = E^{\mathbb{N}^*}$ l'espace produit (des suites de E) qui est un espace polonais lorsqu'il est muni de la distance canonique. On définit sa tribu borélienne \mathcal{B}_{E^∞} qui est également la tribu engendrée par les cylindres, i.e. les ensembles de la forme*

$$C_A = A \times E \times \dots \times E \times \dots, \quad .$$

avec $A \in \mathcal{B}_{E^k}$ ou même $A = A_1 \times \dots \times A_k \in \mathcal{B}_E^{\otimes k}$.

C'est exactement le théorème de Kolmogorov qui affirme qu'une mesure sur un espace produit infini est bien défini, et de manière unique, par l'ensemble de ses marginales ou formulé d'une autre manière, par une famille de mesures de $E^{\mathbb{N}}$ compatibles : il existe $\tilde{\pi} \in \mathbf{P}(E^{\mathbb{N}})$ telle que

$$\tilde{\pi}\left(\prod_{j=1}^{\infty} \tilde{A}_j\right) = \pi_k\left(\prod_{i=1}^k A_i\right).$$

pour tout $\prod_{j=1}^{\infty} \tilde{A}_j \in \mathcal{B}(E)^{\otimes \mathbb{N}}$ tel que $\tilde{A}_j = E$ pour tout $j \neq j_i$, $i = 1, \dots, k$ et $\tilde{A}_{j_i} = A_i$ pour tout $i = 1, \dots, k$.

Chapitre 11

A3. Probabilités

Dans ce chapitre, on définit des objets probabilistes qui nous sont fort utiles dans ce mémoire à savoir :

1. L'espérance conditionnelle et les vecteurs gaussiens.
2. Les processus stochastiques : les sur, sous, semi et locales martingales et les théorèmes classiques les concernant. Puis les processus de Wiener, Levy et Markov.
3. L'intégrale stochastique et la formule d'Ito.
4. Les équations différentielles stochastiques et edps classiques.

Dans tout ce qui suit $(\Omega, \mathcal{T}, \mathcal{F}, \mathbb{P})$ est un espace probabilisé filtré et pour toute variable aléatoire Y , on note $\sigma(Y)$ la tribu engendrée par Y , c'est-à-dire la plus petite tribu rendant mesurable Y .

On rappelle aussi que :

$L^p(\Omega, \mathbb{P}) := \{X \text{ variables aléatoires telles que } \mathbb{E}[|X|^p] < \infty\}$.

$L^\infty(\Omega, \mathbb{P}) \subset \dots \subset L^1(\Omega, \mathbb{P})$ et toutes ces injections sont continues.

$\forall \varepsilon > 0$ et X variable aléatoire presque sûrement positive ou nulle $\mathbb{P}(X \geq \varepsilon) \leq \frac{\mathbb{E}[X]}{\varepsilon}$.¹

$X_n \xrightarrow{\mathbb{P}} X \iff \forall \varepsilon > 0 \mathbb{P}(|X_n - X| \geq \varepsilon) \xrightarrow{n \rightarrow \infty} 0$.

$X_n \xrightarrow{\mathcal{L}} X \iff \forall f \in \mathcal{C}_b^0(\mathbb{R}), \mathbb{E}[f(X_n)] \xrightarrow{n \rightarrow \infty} \mathbb{E}[f(X)]$.

Pour toute variable aléatoire X , l'application $\varphi_X : t \in \mathbb{R} \mapsto \mathbb{E}[e^{itX}] \in \mathcal{C}_b^0(\mathbb{R})$.²

$\forall X \in L^p(\Omega, \mathbb{P})$ et $Y \in L^q(\Omega, \mathbb{P})$ $\mathbb{E}[|XY|] \leq \mathbb{E}[|X|^p]^{\frac{1}{p}} \mathbb{E}[|Y|^q]^{\frac{1}{q}}$ avec $\frac{1}{p} + \frac{1}{q} = 1$.³

$\limsup_n A_n := \bigcap_{n \geq 0} \bigcup_{k \geq n} A_k$ et $\liminf_n A_n := \bigcup_{n \geq 0} \bigcap_{k \geq n} A_k$.

On dit que X est sans atomes si \mathbb{P}_X ne charge pas les points.

On appelle fonction de répartition de X , l'application càdlàg⁴

$$F_X : x \in \mathbb{R} \mapsto \mathbb{P}(X \leq x) := \mathbb{P}_X([-\infty, x])$$

On dit que X est à densité par rapport à la mesure de Lebesgue si F_X est absolument continue.

Soit μ une mesure borélienne finie sur $(\mathbb{R}^d, \mathcal{B}_{\mathbb{R}^d})$, alors on définit sa transformée de Fourier ou fonction caractéristique par :

$$\hat{\mu} : x \in \mathbb{R}^d \mapsto \int_{\mathbb{R}^d} e^{i\langle x, y \rangle} \mu(dy)$$

1. C'est l'inégalité dite de Markov qui implique celle de Bienaymé-Tchebychev qui stipule que pour toute variable aléatoire X de carré intégrable (même intégrable suffit!) et pour tout $\varepsilon > 0$ $\mathbb{P}(|X - \mathbb{E}[X]| \geq \varepsilon) \leq \frac{\mathbb{V}[X]}{\varepsilon^2}$

2. C'est la fonction caractéristique de X qui est en réalité à un facteur 2π près la transformée de Fourier inverse de sa mesure de probabilité \mathbb{P}_X .

3. C'est l'inégalité de Holder sur les espaces de Lebesgue.

4. continue à droite et limite à gauche.

On dit que X est discrète si elle prend ses valeurs dans un ensemble au plus dénombrable. Pour X discrète, on appelle transformée en \mathcal{Z} de X ou fonction génératrice des probabilités, la série quand elle a un sens :

$$G_X : z \mapsto \mathbb{E}[z^X] := \sum_{x \in X(\Omega)} \mathbb{P}(X = x) z^x$$

Remarquons que formellement $\varphi_X(t) = G_X(e^{it})$.

on appelle vecteur aléatoire toute application mesurable de $\Omega \rightarrow \mathbb{K}^d$ dont les composantes sont des variables aléatoires.

Quant aux références, voir principalement en bibliographie [45].

Proposition 21 (Lemmes de Borel-Cantelli) Soit $(A_n)_{n \in \mathbb{N}}$ une suite d'évènements. Dans ce cas nous avons que :

1. si $\sum_{n \in \mathbb{N}} \mathbb{P}(A_n) < \infty$, alors $\mathbb{P}(\limsup_n A_n) = 0$
2. si ces évènements sont indépendants et $\sum_{n \in \mathbb{N}} \mathbb{P}(A_n) = \infty$, alors $\mathbb{P}(\limsup_n A_n) = 1$

Proposition 22 (Lois faible et forte des grands nombres) Soit $(X_n)_{n \in \mathbb{N}}$ une suite de variables aléatoires indépendantes identiquement distribuées qui admettent une espérance μ , alors :

$$\frac{1}{n} \sum_{k=0}^{n-1} X_k \xrightarrow{\mathbb{P}} \mu$$

De plus si ces variables aléatoires sont intégrables ($\mathbb{E}[|X_0|] < \infty$), alors cette moyenne empirique converge presque sûrement (\mathbb{P} -ps) vers son espérance ($\mathbb{E}[X_0]$).

Proposition 23 (TCL) Soit $(X_n)_{n \in \mathbb{N}}$ une suite de variables aléatoires indépendantes identiquement distribuées de carré intégrable, alors :

$$\frac{S_n - n\mathbb{E}[X_0]}{\sqrt{\mathbb{V}[X_0]n}} \xrightarrow{\mathcal{L}} \mathcal{N}(0, 1), S_n := \sum_{k=0}^{n-1} X_k$$

Proposition 24 (Continuité de Levy) Soient X, Y deux variables aléatoires. Alors :

$$\mathbb{P}_X = \mathbb{P}_Y \iff \varphi_X = \varphi_Y$$

Et de plus :

$$X_n \xrightarrow{\mathcal{L}} X \iff \forall x \in \mathbb{R}, \varphi_{X_n}(x) \xrightarrow{n \rightarrow \infty} \varphi_X(x)$$

Proposition 25 X est sans atomes $\iff F_X$ est continue.

Proposition 26 (Inversion de Fourier) Soit μ une mesure borélienne finie telle que $\hat{\mu} \in L^1(\mathbb{R}^d, dx)$. Alors cette mesure est à densité par rapport à la mesure de Lebesgue et de plus on a la formule dite d'inversion de Fourier :

$$\frac{\mu(dy)}{dy} = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} \hat{\mu}(x) e^{-i\langle y, x \rangle} \lambda_{\mathbb{R}^d}(dx)$$

Proposition 27 (Les lois de probabilités classiques) On a les lois usuelles suivantes :

1. $X \sim \delta_x \iff X \stackrel{\mathbb{P}-ps}{=} x$. De plus si X est discrète telle que $X(\Omega) := \{x_1, \dots, x_n, \dots\}$, alors :

$$\mathbb{P}_X := \sum_j \mathbb{P}(X = x_j) \delta_{x_j}$$

2.

$$X \sim b(p) \iff \mathbb{P}_X = p\delta_1 + (1-p)\delta_0$$

3.

$$X \sim b(n, p) \iff \mathbb{P}_X = \sum_{j=0}^n C_n^j p^j (1-p)^{n-j} \delta_j$$

4.

$$X \sim \mathcal{G}(p) \iff \mathbb{P}_X = \sum_{j=1}^{+\infty} p(1-p)^{j-1} \delta_j$$

5.

$$X \sim \mathcal{P}(\lambda) \iff \mathbb{P}_X = \sum_{j=0}^{+\infty} e^{-\lambda} \frac{\lambda^j}{j!} \delta_j$$

On suppose connues les autres lois (à densité) telles que la gaussienne, l'exponentielle, la loi Gamma, la loi de Cauchy, la loi du χ ...

11.1 $\mathbb{E}[\cdot|\mathcal{G}]$ et vecteurs gaussiens

11.1.1 Espérance conditionnelle

Définition 11.1 (Espérance conditionnelle)

Soient $X \in L^1(\Omega, \mathbb{P})$ et $\mathcal{G} \subset \mathcal{F}$ une sous-tribu. On appelle espérance conditionnelle de X sachant \mathcal{G} l'unique variable aléatoire $Y := \mathbb{E}[X|\mathcal{G}]$ \mathcal{G} -mesurable et intégrable telle que :

$$\forall A \in \mathcal{G}, \mathbb{E}[X \mathbb{1}_A] = \mathbb{E}[Y \mathbb{1}_A]$$

Si la variable aléatoire est de carré intégrable, alors l'espérance conditionnelle peut être vue comme un opérateur de projection orthogonale sur un sous-espace vectoriel fermé d'un espace de Hilbert.

Proposition 28 $L^2(\Omega, \mathcal{G}, \mathbb{P}) \subset L^2(\Omega, \mathcal{F}, \mathbb{P})$ est un sous-espace vectoriel fermé de l'espace de Hilbert des variables aléatoires de carré intégrable. De plus dans ce cas, l'espérance conditionnelle n'est autre que la projection orthogonale sur ce sous espace vectoriel.

Propriétés 1 (de l'espérance conditionnelle) On a les propriétés suivantes :

1. $\mathbb{E}[X + \lambda Y|\mathcal{G}] = \mathbb{E}[X|\mathcal{G}] + \lambda \mathbb{E}[Y|\mathcal{G}]$.
2. $\forall \mathcal{H} \subset \mathcal{G}$ une sous tribu $\mathbb{E}[\mathbb{E}[X|\mathcal{G}]|\mathcal{H}] = \mathbb{E}[X|\mathcal{H}]$.
3. $X \leq Y \implies \mathbb{E}[X|\mathcal{G}] \leq \mathbb{E}[Y|\mathcal{G}]$.
4. Si $(X_n)_{n \in \mathbb{N}}$ est une suite croissante de variables aléatoires positives qui converge presque sûrement vers X , alors

$$\mathbb{E}[X_n|\mathcal{G}] \xrightarrow{n \rightarrow \infty} \mathbb{E}[X|\mathcal{G}]$$

Plus généralement, le théorème de convergence dominée de Lebesgue et le lemme de Fatou s'appliquent naturellement aux espérances conditionnelles.

5. Si X est indépendante de \mathcal{G} , alors $\mathbb{E}[X|\mathcal{G}] = \mathbb{E}[X]$.
6. Si Y est \mathcal{G} -mesurable, alors $\mathbb{E}[XY|\mathcal{G}] = Y \mathbb{E}[X|\mathcal{G}]$.
7. Si φ est une fonction convexe telle que $\varphi(X)$ est intégrable, alors $\varphi(\mathbb{E}[X|\mathcal{G}]) \leq \mathbb{E}[\varphi(X)|\mathcal{G}]$.

11.1.2 Vecteurs gaussiens

Définition 11.2 (Vecteur gaussien)

Soit X un vecteur aléatoire à valeurs dans \mathbb{R}^d . On dit qu'il est gaussien si, et seulement si :

$$\forall u \in \mathbb{R}^d, \langle X, u \rangle \text{ est une variable aléatoire gaussienne.}$$

D'après la définition, les marginales d'un tel vecteur définissent des variables aléatoires gaussiennes. De plus, il est entièrement caractérisé par son vecteur d'espérance m_X et sa matrice de covariance Γ_X . En effet, par transfert, on a :

$$\varphi_X(u) := \mathbb{E} \left[e^{i\langle u, X \rangle} \right] = \varphi_{\langle u, X \rangle}(1) = e^{i\langle u, m_X \rangle} e^{-\frac{1}{2} \langle \Gamma_X u, u \rangle}.$$

On rappelle que la matrice de covariance est symétrique positive. On rappelle aussi une condition équivalente à l'indépendance spécifique aux vecteurs gaussiens.

Proposition 29 (Indépendance) Deux vecteurs gaussiens sont indépendants si, et seulement si les covariances croisées sont nulles. En particulier, les marginales d'un vecteur gaussien sont indépendantes si, et seulement si sa matrice de covariance est diagonale.

Proposition 30 (Densité gaussienne) Soit X un vecteur gaussien. si $\det(\Gamma_X) \neq 0$, alors Γ_X est une matrice de produit scalaire, X est à densité par rapport à la mesure de Lebesgue par inversion de Fourier et cette densité est donnée par :

$$f_X(x) := \frac{1}{\sqrt{(2\pi)^d \det(\Gamma_X)}} e^{-\frac{1}{2} \langle \Gamma_X^{-1}(x - m_X), x - m_X \rangle}.$$

Dans le cas dégénéré, il n'a pas de densité et de plus, la loi de $X - m_X$ est presque sûrement portée par un espace vectoriel engendré par les vecteurs propres associés aux valeurs propres non nulles de Γ_X .

Propriétés 2 Soit X un vecteur gaussien. Il existe $Z \sim \mathcal{N}(0, \mathbf{Id})$ gaussien tel que :

$$X = \sqrt{\Gamma_X} Z + m_X.$$

Posons $X := (Y, X_1, \dots, X_{d-1})$ gaussien. Alors :

$$\mathbb{E}[Y|X_1, \dots, X_{d-1}] \text{ est une fonction affine de } X_1, \dots, X_{d-1},$$

plus exactement, c'est la projection orthogonale dans $L^2(\Omega, \mathbb{P})$ de Y sur :

$$F := \mathbf{Vect} \{1, X_1, \dots, X_{d-1}\}.$$

Via la fonction caractéristique, on a un théorème central limite vectoriel, c'est-à-dire pour toute suite $(X_n)_{n \geq 1}$ de vecteurs aléatoires indépendants, identiquement distribués, admettant des moments d'ordre 2, de vecteur d'espérance m et de matrice de covariance Γ , on a :

$$\sqrt{n}(\bar{X}_n - m) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \Gamma).$$

11.2 Processus stochastiques

Définition 11.3 (Processus stochastique)

On désigne par $(\Omega, \mathcal{F}, \mathbb{P})$ un espace de probabilité, T un ensemble arbitraire et S un espace métrique muni de la tribu borélienne notée $\mathcal{B}(S)$. T est souvent appelé ensemble des indices (souvent, on aura $T = \mathbb{R}, \mathbb{R}_+, \mathbb{R}^p, \mathbb{N}, \dots$). T peut faire référence au temps, à l'espace ou aux deux à la fois.

L'indice $t \in T$ désigne alors un instant (\mathbb{R}_+), une date (\mathbb{N}), un point, ou encore un point à un certain instant. S est appelé espace d'état (souvent, on aura $S = \mathbb{R}, \mathbb{C}, \mathbb{C}^{\mathbb{N}}$, un ensemble fini ou dénombrable). Si l'espace d'état S est de la forme \mathbb{R}^d , on parle de champ aléatoire.

Un processus stochastique (ou aléatoire) est une famille de variables aléatoires (c'est-à-dire, des applications mesurables) définies sur le même espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$ indexée par T et à valeurs dans S . Si T est un sous-ensemble d'un espace multidimensionnel, on préfère utiliser la dénomination de champ stochastique. Un processus stochastique est noté par $(X_t)_{t \in T}$. La valeur de la variable aléatoire X_t en un certain $\omega \in \Omega$ est désignée par $X_t(\omega)$.

La famille de toutes les distributions finies-dimensionnelles de X s'appelle la loi spatiale du processus. Si $S \subseteq \mathbb{R}$, on parle de loi temporelle.

Exemples et processus importants

Un premier exemple très important est donné par les martingales qui sont en quelque sorte l'équivalent aléatoire des fonctions constantes.

Définition 11.4 (Martingales)

Un processus $(X_t)_{t \in T}$ avec T totalement ordonné est une $(\mathcal{F}_t)_{t \in T}$ -martingale (respectivement surmartingale, sous-martingale) si, et seulement si :

1. $(\mathcal{F}_t)_{t \in T}$ est une filtration;
2. X_t est \mathcal{F}_t -mesurable et intégrable;
3. $\forall s \leq t, \mathbb{E}[X_t | \mathcal{F}_s] = X_s$ (respectivement \leq, \geq).

Ainsi, à partir d'une martingale, il est aisé de construire via les fonctions convexes et concaves des surmartingales et des sous-martingales.

On rappelle que toute famille bornée dans $L^{1+\epsilon}(\Omega)$ est uniformément intégrable et pour une martingale, la convergence, la fermeture et l'uniforme intégrabilité sont équivalentes. Un exemple de martingales est donné par le mouvement brownien standard ou son carré moins le temps. Via le théorème de représentation de Lévy, toute martingale à trajectoires continues dont la variation quadratique est le temps est un mouvement brownien et réciproquement. On suppose connus les théorèmes classiques sur les martingales tels que ceux de Doob-Meyer, les inégalités maximales de Doob, les théorèmes limite... Et pour plus de détails, consulter les dernières références en bibliographie à savoir [101], \dots , [104] et [106].

D'autres exemples de processus sont donnés par ceux dits gaussiens (Ornstein-Uhlenbeck, Wiener standard...), de diffusion, PAIS (Lévy), de Markov...

Chapitre 12

A4. Analyse

Dans ce chapitre, comme dans les précédents, nous rappelons des résultats et notations d'analyse qui nous sont utiles pour la bonne compréhension de ce mémoire. Quant aux références, voir principalement en bibliographie [45].

12.1 Éléments de théorie de la mesure

Théorème 12.1 (de représentation de Riesz)

1. Si \mathcal{H} est un espace de Hilbert, alors $\mathcal{H}^* = \mathcal{H}$.
2. Si μ est une mesure σ -finie, alors $\forall p \geq 1$ $L^p(\mu)^* = L^q(\mu)$ avec q l'exposant conjugué de p . De plus $\forall p > 1$, $L^p(\mu)$ est un espace de Banach réflexif.

Preuve

1. Soit L une forme linéaire continue non nulle car si elle est nulle, la preuve suivante est triviale. Alors nous savons que $\ker(L) := \{x \in \mathcal{H}, L(x) = 0\}$ est fermé comme image réciproque d'un fermé par une application continue. Donc cet hyperplan est de codimension 1 et d'après le théorème de la projection orthogonale sur un convexe fermé d'un hilbert, on a $\ker(L)^\perp = \mathbb{K}v$ et $\ker(L) \oplus \ker(L)^\perp = \mathcal{H}$. On en déduit que :

$$\forall x \in \mathcal{H}, x = \underset{\in \ker(L)}{y} + \lambda \cdot v$$

Et de plus, il est aisé de voir que $\lambda = \frac{\langle x, v \rangle}{\|v\|^2}$. Il en découle :

$$L(x) = L(\lambda \cdot v) = \lambda L(v) = \langle x, \frac{\overline{L(v)}}{\|v\|^2} \cdot v \rangle = \langle x, w \rangle$$

Et réciproquement, si $L : x \in \mathcal{H} \mapsto \langle x, w \rangle$ avec $w \in \mathcal{H}$, alors $L \in \mathcal{H}^*$ car elle est linéaire par bilinéarité (ou sesquilinearité) du produit scalaire (ou hermitien) et par l'inégalité de Cauchy-Schwarz, on a :

$$\forall x \in \mathcal{H}, |L(x)| \leq \|w\| \cdot \|x\|$$

Donc l'application Φ définie par :

$$\mathcal{H} \longrightarrow \mathcal{H}^*, w \longmapsto \langle \cdot, w \rangle$$

est une isométrie linéaire (anti-linéaire dans le cas hermitien), c'est-à-dire elle est linéaire (respectivement anti-linéaire) et $\|\Phi(w)\| = \|w\|$ vu que $\|\Phi(w)\| \geq \frac{\Phi(w)(w)}{\|w\|} = \|w\|$ et par Cauchy-Schwarz $\|w\|$ est un majorant de $\|\Phi(w)\|$.

2. Pour ce second point, il suffit de considérer l'application bilinéaire continue $\Psi : L^p(\mu) \times L^q(\mu) \rightarrow L^1(\mu)$ $(f, g) \mapsto fg$ vu que par l'inégalité de Holder, on a :

$$\int |fg| d\mu \leq \|f\|_{L^p(\mu)} \|g\|_{L^q(\mu)}$$

On définit ainsi :

$$\Phi : L^q(\mu) \rightarrow L^p(\mu)^*, g \mapsto \langle \cdot, g \rangle_{L^2(\mu)}$$

Et cette application est une isométrie comme la précédente telle que $\|\Phi(g)\| = \|g\|_{L^q(\mu)}$ et de plus $L^p(\mu)^{**} = L^p(\mu)$ pour tout $p > 1$. Pour une preuve plus détaillée consulter le théorème de Riesz-Fisher sur la complétude des espaces de Lebesgue. \square

Soit (E, \mathcal{A}, μ) un espace mesuré.

Théorème 12.2 (Théorème de convergence monotone)

- Soit $(f_n)_{n \geq 0}$ une suite **croissante** de fonctions mesurables **positives** sur E . On a :

$$\lim_{n \rightarrow +\infty} \int f_n d\mu = \int \lim_{n \rightarrow +\infty} f_n d\mu$$

- Soit $(f_n)_{n \geq 0}$ une suite de fonctions mesurables **positives**. On a :

$$\int \sum_{n=0}^{+\infty} f_n d\mu = \sum_{n=0}^{+\infty} \int f_n d\mu$$

Théorème 12.3 (Théorème de convergence dominée de Lebesgue)

Soit $(f_n)_{n \geq 0}$ une suite de fonctions mesurables de E dans \mathbb{C} , et f une fonction mesurable de E . On suppose que

- **limite** : pour μ -presque tout $x \in E$, $f_n(x) \xrightarrow{n \rightarrow +\infty} f(x)$
 - **domination** : il existe une fonction $\varphi : E \rightarrow \mathbb{R}_+$ mesurable telle que $\int \varphi d\mu < +\infty$ et pour tout $n \in \mathbb{N}$, μ -presque tout $x \in E$, $|f_n(x)| \leq \varphi(x)$
- On a alors : f est intégrable, de même que f_n pour tout $n \in \mathbb{N}$,

$$\int |f_n - f| d\mu \xrightarrow{n \rightarrow +\infty} 0 \text{ et } \int f_n d\mu \xrightarrow{n \rightarrow +\infty} \int f d\mu$$

Théorème 12.4 (Théorème de continuité sous l'intégrale)

Soit $f : (t, x) \mapsto f(t, x)$ une fonction de $I \times E$ dans \mathbb{C} (où I est un espace métrique). On suppose que :

- **mesurabilité** : pour tout $t \in I$, $x \mapsto f(t, x)$ est mesurable
- **continuité** : pour μ -presque tout $x \in E$, $t \mapsto f(t, x)$ est continue sur I
- **domination** il existe une fonction $\varphi : E \rightarrow \mathbb{R}_+$ mesurable telle que $\int \varphi d\mu < +\infty$ et pour tout $t \in I$, pour μ -presque tout $x \in E$, $|f(t, x)| \leq \varphi(x)$

Alors la fonction

$$F : t \mapsto F(t) = \int f(t, x) d\mu(x)$$

est bien définie pour tout $t \in I$, et est continue sur I .

Théorème 12.5 (Théorème de dérivation sous l'intégrale)

Soit $f : (t, x) \mapsto f(t, x)$ une fonction de $I \times E$ dans \mathbb{C} . On suppose que :

- **existence de F** pour tout $t \in I$, $x \mapsto f(t, x)$ est intégrable,
- **dérivabilité** pour μ -presque tout $x \in E$, $t \mapsto f(t, x)$ est dérivable sur I de dérivée notée $\frac{\partial f}{\partial t}$
- **domination de la dérivée** il existe une fonction $\varphi : E \rightarrow \mathbb{R}_+$ mesurable telle que $\int \varphi d\mu < \infty$ et pour tout $t \in I$, pour μ -presque tout $x \in E$, $\left| \frac{\partial f}{\partial t}(t, x) \right| \leq \varphi(x)$

Alors la fonction

$$F : t \mapsto F(t) = \int f(t, x) d\mu(x)$$

est dérivable sur I et, pour tout $t \in I$,

$$F'(t) = \int \frac{\partial f}{\partial t}(t, x) d\mu(x)$$

Soient (E, \mathcal{A}, μ) et (F, \mathcal{B}, ν) deux espaces mesurés σ -finis.

Théorème 12.6 (Théorème de Fubini - Tonelli)

Pour toute fonction mesurable **positive** f sur $E \times F$,

$$\int_{E \times F} f(x, y) d(\mu \otimes \nu)(x, y) = \int_E \left(\int_F f(x, y) d\nu(y) \right) d\mu(x) = \int_F \left(\int_E f(x, y) d\mu(x) \right) d\nu(y) \quad (*)$$

Théorème 12.7 (Théorème de Fubini-Lebesgue)

Pour toute fonction mesurable $f : E \times F \rightarrow \mathbb{C}$, telle que

$$\int_{E \times F} |f(x, y)| d(\mu \otimes \nu)(x, y) < \infty$$

la suite d'égalités (*) reste vraie.

Par le théorème de Fubini-Tonelli, la condition équivaut à

$$\int_E \left(\int_F |f(x, y)| d\nu(y) \right) d\mu(x) < \infty \quad \text{ou} \quad \int_F \left(\int_E |f(x, y)| d\mu(x) \right) d\nu(y) < \infty.$$

Théorème 12.8 (Lemme de Fatou)

Soit (E, \mathcal{F}, μ) un espace mesuré. Si $(f_n)_{n \geq 0}$ est une suite de fonctions mesurables positives, alors :

$$\int_E \liminf_n f_n d\mu \leq \liminf_n \int_E f_n d\mu \quad .$$

On suppose connus les résultats classiques sur les distributions, transformées de Fourier et les espaces de Sobolev¹. On rappelle ci-dessous les théorèmes importants de changement de variables jacobien et de Radon-Nikodym.

1. Consulter [117] pour plus de détails.

Théorème 12.9 (Théorème de changement de variables dans \mathbb{R}^d)

Soit U, D des ouverts de \mathbb{R}^d . Soit $f : D \rightarrow \mathbb{C}$ mesurable, et $\varphi : U \rightarrow D$ un \mathcal{C}^1 -difféomorphisme

1. si f est positive, alors

$$\int_D f(x) dx = \int_U f(\varphi(u)) |J_\varphi(u)| du$$

et

$$\int_U f(\varphi(u)) du = \int_D f(x) |J_{\varphi^{-1}}(x)| dx$$

2. si f est intégrable sur D , la première égalité précédente a un sens (autrement dit, $u \mapsto f(\varphi(u)) |J_\varphi(u)|$ est intégrable sur U) et est vraie. Si $f \circ \varphi$ est intégrable sur U , alors il en est de même de la deuxième.

Théorème 12.10 (De Radon-Nikodym)

Soient μ et ν deux mesures σ -finies sur le même espace mesurable, alors il y a équivalence entre ν est absolument continue par rapport à μ et le fait que :

$$\exists \varphi, \quad \nu(dx) = \varphi(x) \mu(dx), \quad \text{c'est-à-dire existence d'une densité.}$$

On note $\frac{d\nu}{d\mu}$ une telle densité dite de Radon-Nikodym.

12.2 Éléments de théorie spectrale

On va rappeler le théorème de Lax-Milgram.

Le théorème de Lax-Milgram – des noms de Peter Lax et Arthur Milgram, auxquels on adjoint parfois celui de Jacques-Louis Lions – est un théorème de mathématiques s’appliquant à certains problèmes aux dérivées partielles exprimés sous une formulation faible (appelée également formulation variationnelle). Il est notamment l’un des fondements de la méthode des éléments finis.

Théorème 12.11 (de Lax-Milgram)

Soient :

1. \mathcal{H} un espace de Hilbert réel ou complexe muni de son produit scalaire noté $\langle \cdot, \cdot \rangle$, de norme associée notée $\| \cdot \|$;
2. $a(\cdot, \cdot)$ une forme bilinéaire (ou une forme sesquilinéaire si \mathcal{H} est complexe) qui est continue sur $\mathcal{H} \times \mathcal{H}$: $\exists c > 0 \quad \forall (u, v) \in \mathcal{H}^2 \quad |a(u, v)| \leq c \|u\| \|v\|$, coercive sur \mathcal{H} (certains auteurs disent plutôt \mathcal{H} -elliptique) : $\exists \alpha > 0 \quad \forall u \in \mathcal{H} \quad a(u, u) \geq \alpha \|u\|^2$ (dans le cas complexe, on prend la partie réelle) ;
3. $L(\cdot)$ une forme linéaire continue sur \mathcal{H} .

Sous ces hypothèses, il existe un unique u de \mathcal{H} tel que l’équation $a(u, v) = L(v)$ soit vérifiée pour tout v de \mathcal{H} : $\exists! u \in \mathcal{H} \quad \forall v \in \mathcal{H} \quad a(u, v) = L(v)$.

Si de plus la forme bilinéaire a est symétrique (respectivement hermitienne), alors u est l’unique élément de \mathcal{H} qui minimise la fonctionnelle $J : \mathcal{H} \rightarrow \mathbb{R}$ définie par $J(v) = \frac{1}{2} a(v, v) - L(v)$ pour tout v de \mathcal{H} , c’est-à-dire : $\exists! u \in \mathcal{H} \quad J(u) = \min_{v \in \mathcal{H}} J(v)$.

Preuve

Par application du théorème de Riesz sur les formes linéaires continues, il existe un vecteur $f \in \mathcal{H}$ tel que

$$\forall v \in \mathcal{H} \quad L(v) = \langle f, v \rangle$$

Par application de ce même théorème aux formes bilinéaires continues, il existe un endomorphisme linéaire continu $A \in \mathcal{L}(\mathcal{H})$ tel que

$$\forall u, v \in \mathcal{H} \quad a(u, v) = \langle Au, v \rangle$$

On en déduit alors :

$$\exists! u \in \mathcal{H} \quad Au = f$$

Pour prouver ce théorème, il suffit donc de montrer que A est une bijection de \mathcal{H} sur \mathcal{H} . On montre dans un premier temps que l’opérateur est injectif, puis qu’il est surjectif.

Par la coercivité de a et en appliquant l’inégalité de Cauchy-Schwarz, on a pour tout $v \in \mathcal{H}$

$$\alpha \|v\|^2 \leq a(v, v) = \langle Av, v \rangle \leq \|Av\| \|v\|$$

D’où $\|Av\| \geq \alpha \|v\|$ pour tout v de \mathcal{H} , ce qui montre que A est injectif et d’image fermée. Notons \mathcal{I} cette image. Par le théorème du supplémentaire orthogonal d’un fermé on sait que $\mathcal{H} = \mathcal{I} \oplus \mathcal{I}^\perp$. Soit ensuite un élément w de \mathcal{I}^\perp , on a par définition $\langle Aw, w \rangle = 0$ et donc :

$$\alpha \|w\|^2 \leq a(w, w) = \langle Aw, w \rangle = 0$$

D’où $w = 0$. Ainsi, \mathcal{I}^\perp est réduit à $\{0\}$, ce qui montre que A est surjectif.

L’endomorphisme A est bijectif ; il existe donc un unique u de \mathcal{H} tel que $Au = f$ et il est donné par $u = A^{-1}f$.

□

On aura besoin de résultats sur la théorie des semi-groupes dont l'origine vient de l'étude du flot d'une équation différentielle ordinaire autonome en dimension finie ainsi que de l'exponentielle d'opérateurs.

Définition 12.1 (Semi-groupe fortement continu)

Soit E un espace de Banach. On considère une famille d'opérateurs $(P_t)_{t \geq 0}$ telle que $\forall t P_t \in \mathcal{L}(E)$. On dit que c'est un semi-groupe fortement continu si :

$$P_0 = Id_E, \forall (t, s) P_{t+s} = P_t \circ P_s \text{ et } \forall x \in E \lim_{t \rightarrow 0^+} P_t x = x$$

$$(P_t)_{t \geq 0} \text{ est de contraction} \iff \forall t \|P_t\| := \sup_{x \in E, x \neq 0} \frac{\|P_t x\|}{\|x\|} \leq 1.$$

On appelle générateur infinitésimal de $(P_t)_{t \geq 0}$, l'opérateur $(Dom(T), T)$ défini par :

1. $Dom(T) := \{x \in E \text{ tels que } \lim_{t \rightarrow 0} \frac{P_t x - x}{t} \text{ existe}\}.$
2. $\forall x \in Dom(T) T x := \lim_{t \rightarrow 0} \frac{P_t x - x}{t}$

Remarquons que pour $T \in \mathcal{L}(E)$, la famille d'opérateurs $(e^{tT})_{t \geq 0}$ (définie classiquement par sa série!) est un semi-groupe fortement continu de générateur infinitésimal T . C'est pourquoi, on note parfois de façon abusive : $P_t = e^{tT}$.

On a le résultat suivant important concernant les semi-groupes de contraction :

Proposition 31 Soient E un espace de Banach, $(P_t)_{t \geq 0}$ un semi-groupe de contraction sur E de générateur infinitésimal $(Dom(T), T)$. Alors :

1. $\forall x \in E$ le flot $t \mapsto P_t x \in \mathcal{C}^0(\mathbb{R}^+, E)$.
2. $\forall x \in Dom(T)$ et $t \geq 0$, on a $P_t x \in Dom(T)$, de plus le flot $t \mapsto P_t x \in \mathcal{C}^1(\mathbb{R}^+, X)$ et vérifie l'équation différentielle $\frac{dy(t)}{dt} = T y(t)$.
3. $(Dom(T), T)$ est fermé et de domaine dense.

En théorie des semi-groupes, un théorème très important est celui de Hille-Yosida qui est un outil puissant et fondamental reliant les propriétés de dissipation de l'énergie d'un opérateur non borné $T : D(T) \subset E \rightarrow E$ à l'existence et l'unicité et la régularité des solutions d'une équation différentielle

$$\begin{cases} x'(t) = T x(t) \\ x(0) = x_0 \end{cases}$$

Ce résultat permet notamment de donner l'existence, l'unicité et la régularité des solutions d'une équations aux dérivées partielles plus efficacement que le théorème de Cauchy-Lipschitz-Picard, plus adapté aux équations différentielles ordinaires.

Avant d'énoncer ce théorème, on va rappeler quelques définitions et propriétés.

Définition 12.2 (Opérateur m-dissipatif)

Un opérateur $(T, D(T))$ est dissipatif si $\forall x \in D(A), \forall \lambda > 0, \|x - \lambda T x\| \geq \|x\|$. Dans le cas où $E = H$ est hilbertien on montre que T est dissipatif si et seulement si $\forall x \in D(T), \Re(\langle T x, x \rangle_H) \leq 0$.

Si de plus $\forall \lambda > 0, Id - \lambda T$ est surjectif on dit que $(T, D(T))$ est maximal-dissipatif.

Remarques :

Si $(T, D(T))$ est un opérateur dissipatif alors $\forall \lambda > 0$ l'opérateur $(\text{Id} - \lambda T)$ est injectif car

$$(\text{Id} - \lambda T)x = 0 \Rightarrow 0 \leq \|x\| \leq \|(\text{Id} - \lambda T)x\| = 0 \Rightarrow x = 0$$

On peut montrer que $\forall \lambda > 0, \text{Id} - \lambda T$ est surjectif si et seulement si $\exists \lambda_0, \text{Id} - \lambda_0 T$ surjectif.

En pratique pour montrer qu'un opérateur est m -dissipatif on montre d'abord à la main qu'il est dissipatif et on résout ensuite un problème variationnel pour une valeur λ_0 bien choisie (par exemple avec le théorème de Lax-Milgram, voir exemple de l'équation de la chaleur traité plus bas).

Dans ce cas l'opérateur $(\text{Id} - \lambda T)$ est un isomorphisme (a priori non continu) de $L(T, X)$ et on note $J_\lambda = (\text{Id} - \lambda T)^{-1}$, qui est au facteur près la résolvante de T . De plus,

$$\|J_\lambda y\|_E \leq \|(\text{Id} - \lambda T)[J_\lambda y]\|_E \leq \|y\|_E, J_\lambda \in \mathcal{L}((E, \|\cdot\|_E), (D(T), \|\cdot\|_E))$$

Nous allons voir que cette propriété de continuité peut être améliorée (on va rendre moins fine la topologie sur $(D(T), \|\cdot\|_E)$ en munissant $D(T)$ de sa norme du graphe).

Propriétés 3 On a les propriétés importantes suivantes :

1. Si $(T, D(T))$ est m -dissipatif alors c'est un opérateur fermé.
2. pour $x \in D(T)$ on pose $\|x\|_{D(T)} = \|x\|_E + \|Tx\|_E$. Alors $\|\cdot\|_{D(T)}$ est une norme pour laquelle $D(T)$ est un espace de Banach et $T \in \mathcal{L}(D(T), \|\cdot\|_{D(T)}, (E, \|\cdot\|_E))$.
3. Si H est un espace hilbertien et $T : D(T) \subset H \rightarrow H$ est m -dissipatif alors il est à domaine dense.
4. Réciproquement si $T : D(T) \subset H \rightarrow H$ est de domaine dense, dissipatif, fermé et tel que son adjoint $(T^*, D(T^*))$ est dissipatif alors $(T, D(T))$ est m -dissipatif.
5. Toujours dans le cadre hilbertien :
 si $(T, D(T))$ est dissipatif autoadjoint, alors il est m -dissipatif,
 si $(T, D(T))$ est anti-adjoint, alors il est m -dissipatif,
 dans ce dernier résultat, la condition de dissipativité n'est pas nécessaire car $(T, D(T))$ anti-adjoint entraîne que $\Re(\langle Tx, x \rangle_H) = 0$ donc la dissipativité, voir l'exemple de l'équation des ondes plus bas.

Théorème 12.12 (de Hille-Yosida)

Soit E un espace de Banach et $T : D(T) \subset E \rightarrow E$ un opérateur non borné. On a l'équivalence :

$(T, D(T))$ est m -dissipatif à domaine dense.

$(T, D(T))$ est le générateur infinitésimal d'un semi-groupe de contraction.

Rien de tel que des exemples pour mieux appréhender la puissance et l'utilité de ce théorème.

Équation de la chaleur :

On se donne Ω un ouvert borné de classe \mathcal{C}^2 de \mathbb{R}^n et on cherche à résoudre l'équation de la chaleur

$$\begin{cases} \partial_t u(x, t) - \Delta u(x, t) = 0 \\ u(x, 0) = u_0(x) \end{cases} \quad \text{sur } (x, t) \in \Omega \times [0, +\infty] \text{ pour une condition initiale donnée.}$$

On peut réécrire cette équation aux dérivées partielles sous la forme d'une équation différentielle ordinaire $y'(t) = Ty(t)$ en posant $E = H = L^2(\Omega)$, $y(t) = u(\cdot, t) \in H$ et en définissant $(T, D(T))$ par :

$D(T) = H^2(\Omega) \cap H_0^1(\Omega) \subset L^2(\Omega)$ et $Tx = \Delta x$ pour tout $x \in D(T)$. Nous sommes dans le bon cadre pour utiliser la théorie des semi-groupes et le théorème de Hille-Yosida; reste à montrer que l'opérateur T est m-dissipatif.

Il est bien connu que le laplacien est un opérateur autoadjoint : (formule de Green)

$$\langle Tu, v \rangle_H = \int_{\Omega} (\Delta u)v = - \int_{\Omega} \nabla u \cdot \nabla v = \int_{\Omega} u(\Delta v) = \langle u, Tv \rangle_H$$

par double intégration par parties, et que $D(T)$ est dense dans $L^2(\Omega)$, il suffit donc de montrer qu'il est dissipatif ou de façon équivalente que $\Re \langle Tx, x \rangle_H \leq 0$. Or tout $x \in D(T) = H^2(\Omega) \cap H_0^1(\Omega)$ est de trace nulle, donc en intégrant par parties :

$$\Re \langle Tx, x \rangle_H = - \int_{\Omega} \|\nabla x\|_{\mathbb{R}^n}^2 \leq 0$$

Les propriétés ci-dessus et le théorème de Hille-Yosida permettent enfin de conclure quant à l'existence-unicité et la régularité des solutions. On remarque de plus que :

$$\frac{d}{dt} (\|y(t)\|_H^2) = 2 \langle y'(t), y(t) \rangle_H = 2 \langle Ty(t), y(t) \rangle_H \leq 0$$

On retrouve, bien sûr, le côté dissipatif et irréversible de l'équation de la chaleur.

Équation des ondes :

L'équation des ondes homogène se formule dans un domaine Ω suffisamment régulier (c'est-à-dire \mathcal{C}^2 en pratique) et sur un intervalle de temps $[0, T)$ (avec $T > 0$) selon

$$\begin{cases} u_{tt}(t, x) - \Delta u(t, x) = 0 & \forall (t, x) \in (0, T) \times \Omega \\ u(0, x) = f(x) & \forall x \in \Omega \\ u_t(0, x) = g(x) & \forall x \in \Omega \end{cases}$$

On se place dans la théorie des semi-groupes en mettant l'équation précédente au premier ordre en temps.

On pose alors $\mathcal{A} = \begin{pmatrix} 0 & I \\ \Delta & 0 \end{pmatrix}$, $\mathcal{Y} = \begin{pmatrix} u \\ v \end{pmatrix}$ (avec $v = u'$) et $\mathcal{Y}_0 = \begin{pmatrix} f \\ g \end{pmatrix}$. L'équation devient alors

$$\begin{cases} \mathcal{Y}'(t) = \mathcal{A}\mathcal{Y}(t) \\ \mathcal{Y}(0) = \mathcal{Y}_0 \end{cases}$$

Le domaine du Laplacien étant $D(\Delta) = H^2(\Omega) \cap H_0^1(\Omega)$, celui de \mathcal{A} est $D(\mathcal{A}) = H^2(\Omega) \cap H_0^1(\Omega) \times H_0^1(\Omega)$ sur $H = H_0^1(\Omega) \times L^2(\Omega)$. Les conditions initiales seront alors prises dans H . Le produit scalaire dans H est défini pour tout couple (u, v) dans H ($u = (u_1, u_2)$ et $v = (v_1, v_2)$) par

$$(u, v)_H = (\nabla u_1, \nabla v_1)_{L^2(\Omega)} + (u_2, v_2)_{L^2(\Omega)}.$$

Reste à vérifier que nous sommes bien dans les conditions d'application du théorème de Hille-Yosida : $D(\mathcal{A})$ est dense dans H .

\mathcal{A} est fermé.

\mathcal{A} est dissipatif. Ce point mérite une preuve.

On peut utiliser les propriétés ci-dessus pour montrer que \mathcal{A} est m-dissipatif en montrant que \mathcal{A} est anti-adjoint. On a alors pour tout couple (u, v) dans $D(\mathcal{A})$

$$\begin{aligned} (\mathcal{A}u, v)_H &= (\nabla u_2, \nabla v_1)_{L^2(\Omega)} + (\Delta u_1, v_2)_{L^2(\Omega)} \\ &= -(u_2, \Delta v_1)_{L^2(\Omega)} - (\nabla u_1, \nabla v_2)_{L^2(\Omega)} \\ &= -(u, \mathcal{A}v)_H. \end{aligned}$$

Ainsi, \mathcal{A} est anti-adjoint et à domaine dense donc m-dissipatif.

Listings

8.1	Préambule python nécessaire.	72
8.2	Code python donnant la cardioide.	72
8.3	Code python donnant une gaussienne en 3D.	73
8.4	Code pour représentation 3D second exemple.	74
8.5	Code python pour représentation 3D troisième exemple.	75
8.6	Fonction python générant un vecteur gaussien	75
8.7	Codes pour le schéma d'Euler explicite.	76
8.8	Pulsation pour exemples et définition de l'oscillateur harmonique associé.	76
8.9	Exemple de représentation de l'oscillateur harmonique.	77
8.10	Code python donnant l'erreur.	77
8.11	Code donnant l'évolution de l'erreur par rapport au pas.	79
8.12	Code relatif à l'instabilité de l'oscillateur harmonique.	81
8.13	Exemple de code pour le schéma d'Euler.	82
8.14	Code simulant une trajectoire du mouvement brownien standard.	82
8.15	Code simulant trois trajectoires du processus d'Ornstein-Uhlenbeck.	83
8.16	Code définissant l'hamiltonien associé au système à champ moyen	83
8.17	Exemples de potentiels.	84
8.18	Code pour approximation de McKean-Vlasov.	84
8.19	Code simulant deux trajectoires d'une particule typique du système.	85
8.20	Code simulant des trajectoires des particules du système.	86
8.21	Second exemple de simulation des trajectoires d'une particule typique.	86
8.22	Second exemple de simulation des trajectoires des particules du système	87

Bibliographie

- [1] Qiang Liu and Dilin Wang. Stein variational gradient descent : A general purpose bayesian inference algorithm, 2019.
- [2] Jianfeng Lu, Yulong Lu, and James Nolen. Scaling limit of the Stein variational gradient descent : the mean field regime. *SIAM J. Math. Anal.*, 51(2) :648–671, 2019.
- [3] Jean-François Jabir, David Siska, and Lukasz Szpruch. Mean-field neural odes via relaxed optimal control, 2020.
- [4] Kaitong Hu, Zhenjie Ren, David Siska, and Lukasz Szpruch. Mean-field langevin dynamics and energy landscape of neural networks, 2020.
- [5] Arnaud Guillin, Wei Liu, Liming Wu, and Chaoen Zhang. Uniforme Poincaré an Logarithmic Sobolev Inequalities for Mean-Field Particles Systems. working paper or preprint, September 2019.
- [6] L. Wu. Gradient estimates of Poisson equations on Riemannian manifolds and applications. *J. Funct. Anal.*, 257(12) :4015–4033, 2009.
- [7] B. Zegarlinski. Dobrushin uniqueness theorem and logarithmic Sobolev inequalities. *J. Funct. Anal.*, 105(1) :77–111, 1992.
- [8] Cédric Villani. Optimal transport, volume 338 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 2009. Old and new.
- [9] Alain-Sol Sznitman. Topics in propagation of chaos. In *Ecole d’Été de Probabilités de Saint-Flour XIX—1989*, volume 1464 of *Lecture Notes in Math.*, pages 165–251. Springer, Berlin, 1991.
- [10] P. Cattiaux, A. Guillin, and F. Malrieu. Probabilistic approach for granular media equations in the non-uniformly convex case. *Probab. Theory Related Fields*, 140(1-2) :19–40, 2008.
- [11] L. Wu and W. Liu. Large deviations for empirical measures of mean-field gibbs measures. *Stoch. Proc. Appl.*. To appear, preprint 2018.
- [12] S. Bobkov and M. Ledoux. Poincaré’s inequalities and Talagrand’s concentration phenomenon for the exponential distribution. *Probab. Theory Related Fields*, 107(3) :383–400, 1997.
- [13] A. Guillin, C. Léonard, L. Wu, and N. Yao. Transportation-information inequalities for markov processes. *Probab. Theo. Rel. Fields*, 144(3-4) :669–695., 2009.
- [14] Patrick Cattiaux, Arnaud Guillin, and Li-Ming Wu. A note on Talagrand’s transportation inequality and logarithmic Sobolev inequality. *Probab. Theory Related Fields*, 148(1-2) :285–304, 2010.
- [15] F. Otto and C. Villani. Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality. *J. Funct. Anal.*, 173(2) :361–400, 2000.
- [16] Sergey G. Bobkov, Ivan Gentil, and Michel Ledoux. Hypercontractivity of Hamilton-Jacobi equations. *J. Math. Pures Appl. (9)*, 80(7) :669–696, 2001.
- [17] Cours de l’Ecole doctorale EDDIMO : Une introduction aux limites de champ moyen pour des systèmes de particules -Stéphane MISCHLER.

- [18] *Sur les inégalités de Sobolev logarithmiques en théorie de l'information et pour des systèmes de spins conservatifs en mécanique statistique* par Djalil Chafai.
- [19] *Quelques propriétés de diffusions infini-dimensionnelles liées à la mécanique statistique* par Désiré SEU.
- [20] *Inégalités fonctionnelles et comportement en temps long de quelques processus de Markov* par Florent Malrieu.
- [21] *Algorithmes stochastiques pour l'apprentissage, l'optimisation et l'approximation du régime stationnaire* par Sofiane Saadane.
- [22] Chung, K.L., and Williams, R.J. (1990). *Introduction to Stochastic Integration, Second Edition. Probability and Its Applications.*
- [23] Lévy, Paul. (1965). *Processus Stochastiques et Mouvement Brownien, Deuxième édition.*
- [24] Durrett, R., (1984), *Brownian motion and Martingales in Analysis.*
- [25] Evans, Lawrence (2010), *Partial differential equations. Second edition.*
- [26] Revuz, Daniel, Yor, Marc (1999), *Continuous martingales and Brownian motion. Third edition.*
- [27] Stroock, Daniel W.; Varadhan, S. R. Srinivasa (1979), *Multidimensional diffusion processes.*
- [28] Karatzas, Ioannis, Shreve, Steven E. (1991), *Brownian motion and stochastic calculus. Second edition.*
- [29] Chaudru de Raynal, P.-E., and Frikha, N., *Well-posedness for some non-linear diffusion processes and related PDE on the Wasserstein space. Preprint.*
- [30] H.P. McKean Jr., *Propagation of chaos for a class of non-linear parabolic equations.*, in : *Stochastic Differential Equations (Lecture Series in Differential Equations, Session 7, Catholic Univ., 1967), Air Force Office Sci. Res., Arlington, VA, 1967, pp. 41–57.*
- [31] Carrillo, J. A., Gvalani, R. S., Pavliotis, G. A., Schlichting, A. (2020). *Long-time behaviour and phase transitions for the McKean-Vlasov equation on the torus.*
- [32] Chaudru de Raynal, P. E. (2020), *Strong well posedness of McKean-Vlasov stochastic differential equations with Hölder drift.*
- [33] Kolokoltsov, Vassili N., Troeva, Marianna (2019), *On mean field games with common noise and McKean-Vlasov SPDEs.*
- [34] Eberle, Andreas, Guillin, Arnaud, Zimmer, Raphael (2019), *Quantitative Harris-type theorems for diffusions and McKean-Vlasov processes.*
- [35] Dawson, Donald A., Gärtner, Jürgen (1987), *Large deviations from the McKean-Vlasov limit for weakly interacting diffusions.*
- [36] Méléard, Sylvie (1996), *Asymptotic behaviour of some interacting particle systems; McKean-Vlasov and Boltzmann models.*
- [37] Bossy, Mireille, Talay, Denis (1997), *A stochastic particle method for the McKean-Vlasov and the Burgers equation.*
- [38] Veretennikov, A. Yu. (2006), *On ergodic measures for McKean-Vlasov stochastic equations.*
- [39] Crisan, Dan, Xiong, Jie (2007), *An approximate McKean-Vlasov model for the stochastic filtering problem.*
- [40] S. Benachour, B. Roynette, D. Talay, P. Vallois, *Nonlinear self-stabilizing processes. I. Existence, invariant probability, propagation of chaos, Stochastic Process. Appl. 75 (2) (1998) 173–201.*
- [41] S. Benachour, B. Roynette, P. Vallois, *Nonlinear self-stabilizing processes. II. Convergence to invariant probability, Stochastic Process. Appl. 75 (2) (1998) 203–224.*

- [42] P. Cattiaux, A. Guillin, and F. Malrieu. Probabilistic approach for granular media equations in the non-uniformly convex case. *Probab. Theory Related Fields*, 140(1-2) :19–40, 2008.
- [43] José A. Carrillo, Robert J. McCann, and Cédric Villani. Kinetic equilibration rates for granular media and related equations : entropy dissipation and mass transportation estimates. *Rev. Mat. Iberoamericana*, 19(3) :971–1018, 2003.
- [44] S. Herrmann, P. Imkeller, D. Peithmann, Large deviations and a Kramers'type law for self-stabilizing diffusions, *Ann. Appl. Probab.* 18 (4) (2008) 1379–1423.
- [45] Wikipédia et moteurs de recherche classiques.
- [46] Benedetto, D., Caglioti, E., Pulvirenti, M. (1997), A kinetic equation for granular media.
- [47] Pulvirenti, M. (1999), Kinetic and hydrodynamical equations for one-dimensional granular media.
- [48] Carrillo, J. A., Cercignani, C., Gamba, I. M. (2000), Steady states of a Boltzmann equation for driven granular media.
- [49] Malrieu, Florent (2003), Convergence to equilibrium for granular media equations and their Euler schemes.
- [50] Mischler, Stéphane, Mouhot, Clément (2009), Stability, convergence to the steady state and elastic limit for the Boltzmann equation for diffusively excited granular media.
- [51] Tristani, Isabelle (2016), Boltzmann equation for granular media with thermal force in a weakly inhomogeneous setting.
- [52] Herrmann, Samuel. (2003). Système de processus auto-stabilisants. (System of autostabilizing processes). *Dissertationes Mathematicae*. 414.
- [53] Duong, Hong, and Tugaut, Julian. Coupled McKean-Vlasov diffusions : wellposedness, propagation of chaos and invariant measures, *Stochastics*.
- [54] Thomas Hudson, Patrick van Meurs, Mark A. Peletier (2020), Atomistic origins of continuum dislocation dynamics.
- [55] M. Shiino (1987), Dynamical behavior of stochastic systems of infinitely many coupled nonlinear oscillators exhibiting phase transitions of mean-field type : H theorem on asymptotic approach to equilibrium and critical slowing down of order-parameter fluctuations.
- [56] T. Lepoutre, A. Moussa (2017), Entropic structure and duality for multiple species crossdiffusion systems.
- [57] M. Di Francesco, S. Fagioli (2013), Measure solutions for non-local interaction pdes with two species.
- [58] M. Di Francesco, S. Fagioli (2016), A nonlocal swarm model for predators-prey interactions.
- [59] Benaïm, Michel, Bréhier, Charles-Edouard (2019), Convergence analysis of adaptive biasing potential methods for diffusion processes.
- [60] Billionis, I., Koutsourelakis, P. S. (2012), Free energy computations by minimization of Kullback-Leibler divergence : an efficient adaptive biasing potential method for sparse representations.
- [61] Virginie Ehlacher, Tony Lelièvre, Pierre Monmarché (2019), Adaptive force biasing algorithms : new convergence results and tensor approximations of the bias.
- [62] Fort, G., Jourdain, B., Lelièvre, T., Stoltz, G. (2018), Convergence and efficiency of adaptive importance sampling techniques with partial biasing.
- [63] Alrachid, Houssam, Lelièvre, Tony (2015), Long-time convergence of an adaptive biasing force method : variance reduction by Helmholtz projection.

- [64] *Bouttier, Clément, Gavra, Ioana (2019), Convergence rate of a simulated annealing algorithm with noisy observations.*
- [65] *A. Dembo, O. Zeitouni (2010), Large deviations techniques and applications.*
- [66] *M. I. Freidlin and A. D. Wentzell (1979), Random perturbations of dynamical systems.*
- [67] *Herrmann, S., Imkeller, P., Peithmann, D. (2008), Large deviations and a Kramers' type law for self-stabilizing diffusions.*
- [68] *Delarue, François, Lacker, Daniel, Ramanan, Kavita (2020), From the master equation to mean field game limit theory : Large deviations and concentration of measure.*
- [69] *Mogulskii, A. A., Prokopenko, E. I. (2019), The large deviation principle in the phase space for a multidimensional second compound renewal process.*
- [70] *Andresen, Andreas, Imkeller, Peter, Perkowski, Nicolas (2013), Large deviations for Hilbert-space-valued Wiener processes : a sequence space approach.*
- [71] *Herrmann, Samuel, Imkeller, Peter, Peithmann, Dierk (2006), Transition times and stochastic resonance for multidimensional diffusions with time periodic drift : a large deviations approach.*
- [72] *François Bolley, Ivan Gentil, Arnaud Guillin (2012), Convergence to equilibrium in Wasserstein distance for Fokker-Planck equations*
- [73] *François Bolley, Ivan Gentil, Arnaud Guillin (2013). Uniform convergence to equilibrium for granular media.*
- [74] *Ambrosio, L., Gigli, N., Savaré, G. (2008), Gradient Flows in Metric Spaces and in The Space of Probability Measures.*
- [75] *Daneri, S., Savaré, G. (2012), Lecture notes on gradient flows and optimal transport.*
- [76] *A. Arnold, P. Markowich, G. Toscani, A. Unterreiter (2001), On convex Sobolev inequalities and the rate of convergence to equilibrium for Fokker-Planck type equations.*
- [77] *L.A. Caffarelli (1992), The regularity of mappings with a convex potential.*
- [78] *K.-T. Sturm, M.-K. von Renesse (2005), Transport inequalities, gradient estimates, entropy and Ricci curvature.*
- [79] *Benamou, J. D., Brenier, Y. (2001), Mixed L^2 -Wasserstein optimal mapping between prescribed density functions.*
- [80] *Ané, C., Blachère, S., Chafaï, D., Fougères, P., Gentil, I., Malrieu, F., Roberto, C., Scheffer, G. (2000). Sur les inégalités de Sobolev logarithmiques.*
- [81] *Malrieu, F. (2001). Logarithmic Sobolev inequalities for some nonlinear PDE's.*
- [82] *C. Ané, M. Ledoux (2000), On logarithmic Sobolev inequalities for continuous time random walks on graphs.*
- [83] *D. Bakry (1992), L'hypercontractivité et son utilisation en théorie des semigroupes.*
- [84] *D. Bakry, T. Coulhon, M. Ledoux, L. Saloff-Coste (1995), Sobolev inequalities in disguise.*
- [85] *Bobkov, S. G., Götze, F. (1999), Exponential integrability and transportation cost related to logarithmic Sobolev inequalities.*
- [86] *S. Bobkov, I. Gentil, M. Ledoux (2000), Hypercontractivity of Hamilton-Jacobi equations.*
- [87] *Figalli, Alessio, Glaudo, Federico (2020), On the Sharp Stability of Critical Points of the Sobolev Inequality.*
- [88] *Bolley, François, Gentil, Ivan, Guillin, Arnaud (2018), Dimensional improvements of the logarithmic Sobolev, Talagrand and Brascamp-Lieb inequalities.*

- [89] Bakry, Dominique; Barthe, Franck; Cattiaux, Patrick; Guillin, Arnaud (2008), *A simple proof of the Poincaré inequality for a large class of probability measures.*
- [90] D. Bakry, P. Cattiaux, A. Guillin (2008), *Rate of convergence for ergodic continuous Markov processes : Lyapunov versus Poincaré.*
- [91] Bonnefont, Michel, Joulin, Aldéric, Ma, Yutao (2016), *A note on spectral gap and weighted Poincaré inequalities for some one-dimensional diffusions.*
- [92] Boissard, E., Cattiaux, P., Guillin, A., Miclo, L. (2018), *Ornstein-Uhlenbeck pinball and the Poincaré inequality in a punctured domain.*
- [93] Benaïm, M., Ledoux, M. and Raimond, O. (2002). *Self-interacting diffusions. Probab. Theory Related Fields* 122 1–41.
- [94] August 2010, *The ODE method for some self-interacting diffusions on \mathbb{R}^d by Aline Kurtzmann.*
- [95] *Mean field games and interacting particle systems by Daniel Lacker, spring 2018.*
- [96] *Fonctions de Lyapunov : application à l'étude de diffusions par Philippe Carmona.*
- [97] *Limite thermodynamique et théorie du potentiel par S. Boucksom.*
- [98] *Hypocoercivité : approches alternatives et applications aux algorithmes stochastiques par Pierre MonMarché.*
- [99] G. Pagès : *Sur quelques algorithmes récursifs pour les probabilités numériques. ESAIM Probab. Statist., 5 :141–170 (electronic), 2001.*
- [100] *Estimation récursive de la mesure invariante d'un processus de diffusion par Vincent Le-maire.*
- [101] *Transparents du cours sur les processus stochastiques de Mr Ying Hu.*
- [102] *Transparents du cours de calcul stochastique de Mr Breton.*
- [103] *Transparents du cours de grandes déviations de Mr Mathias Rousset.*
- [104] *Transparents du cours de processus de Lévy et à sauts de Mr Mihai Gradinaru.*
- [105] *Transparents du cours de théorie spectrale de Mr Nicolas Raymond.*
- [106] *Transparents du cours sur les chaînes de Markov de Mr Franke Brice.*
- [107] *Transparents du cours d'hypocoercivité de Mr Frédéric Hérau.*
- [108] *Transparents du cours d'analyse microlocale de Mr Zied Ammari.*
- [109] *Transparents du cours de systèmes hamiltoniens de Mr Benoit Grébert.*
- [110] *Transparents du cours de distributions et EDPs de Mr Nicolas Depauw.*
- [111] *Transparents du cours sur les espaces de Sobolev de Mr Roger Lewandowski.*
- [112] *Transparents du cours de théorie pseudospectrale de Mr Joseph Viola.*
- [113] *Transparents du cours sur l'équation de Schrödinger de Mr Gabriel Rivière.*
- [114] *Transparents du cours de géométrie différentielle de Mr Erwan Brugallé.*
- [115] *Projet de L3 sur les séries de Fourier et applications aux équations classiques de la physique sous la direction de Mr Laurent MEERSSEMAN.*
- [116] *PRA de M1 sur l'application de l'analyse hilbertienne et de Fourier à la théorie du signal : échantillonnage de Shannon et filtrage sous la direction de Mr Laurent EVAÏN.*
- [117] *Analyse de Fourier dans les espaces fonctionnels par Mohammed EL AMRANI.*
- [118] *Long-time dynamics of stochastic differential equations by Nils Berglund. Institut Denis Poisson – UMR 7013 Université d'Orléans, Université de Tours, CNRS. Lecture notes Summer School — From kinetic equations to statistical mechanics, Saint Jean de Monts, June–July 2021.*