

Rapport de stage de M1

Différentes notions de chaos pour le problème de dérivation de l'équation de Boltzmann

Adèle DEBERGE, sous la supervision de Sergio SIMONELLA
à La Sapienza, Università di Roma

Mars 2023 - Juin 2023

Résumé

Lors de ce stage je me suis familiarisée avec différents concepts de théorie cinétique des gaz et autour de l'équation de Boltzmann. Je me suis en particulier concentrée sur la formulation de la notion de chaos pour le passage de la dynamique microscopique à un système thermodynamique. J'ai étudié différentes façons de formuler mathématiquement cette notion et les liens qu'il existait entre elles.

Table des matières

1	Introduction	2
1.1	Un peu de physique : le raisonnement de Boltzmann	2
1.2	Le problème de dérivation de l'équation	3
1.3	Déroulement du stage	3
2	Trois formulations équivalentes du chaos faible	3
2.1	Introduction au chaos faible	4
2.2	Deux types de distances sur $\mathcal{P}(\mathcal{E})$	4
2.2.1	Métriser la convergence faible	4
2.2.2	Les distances de Monge-Kantorovitch-Wassertsein	6
2.3	L'espace $\mathcal{P}(\mathcal{P}(E))$	6
2.3.1	Une famille déterminante pour la convergence faible sur $\mathcal{P}(\mathcal{P}(\mathcal{E}))$	7
2.3.2	Changement de perspective : de $\mathcal{P}_{sym}(E^N)$ vers $\mathcal{P}(\mathcal{P}(E))$	8
2.4	Preuve	9
2.4.1	Chaos 2 \Rightarrow Chaos 3	10
2.4.2	Chaos 3 \Rightarrow Chaos 1	10
3	Quelques notions plus fortes de chaos	10
3.1	Chaos par rapport à une fonctionnelle extensive	11
3.2	Information de Fisher et entropie	12
3.2.1	L'entropie H_d	12
3.2.2	L'information de Fisher I_d	14
3.3	L'inégalité HWI	15
3.3.1	Quelques notions sur le réarrangement convexe de probabilités	15
3.3.2	Preuve de l'inégalité	17
3.4	Preuve	18
3.4.1	(IC) \Rightarrow (IBC)	19
3.4.2	(IBC) \Rightarrow (KC)	19
3.4.3	(HC) \Rightarrow (KC)	19

1 Introduction

1.1 Un peu de physique : le raisonnement de Boltzmann

Comment décrit-on le comportement d'un gaz ? Un gaz est un système constitué d'un grand nombre de particules N de poids négligeable, qui interagissent entre elles et évoluent dans un domaine $\Omega \subset \mathbb{R}^d$ de l'espace. L'état d'une particule i "ponctuelle" est caractérisé par son vecteur position $x_i \in \Omega$ et son vecteur vitesse $v_i \in \mathbb{R}^d$. On appellera $z_i := (x_i, v_i)$ son état. L'état microscopique du gaz est alors un vecteur $Z^N = (z_i)_{i=1}^N$ et l'espace des états possibles pour le gaz est appelé espace des phases et est $(\Omega \times \mathbb{R}^d)^N$. On peut se dire que pour connaître le comportement du gaz, il suffit de connaître sa configuration initiale et de résoudre les équations de la mécanique classique. Cependant, plusieurs problèmes se posent. Par exemple, on se retrouve avec une quantité immense de variables et la donnée initiale est difficile d'accès. Le grand nombre de particules suggère qu'une approche statistique serait plus adaptée. Désormais, plutôt que de décrire l'état d'un gaz au court du temps par un vecteur Z_t , on va chercher à le décrire par une densité $f_t(x, v)$ représentant la proportion de particules du système présente au alentours de la position x et ayant une vitesse proche de v . Physiquement, on souhaiterait que $f_t \in L_{loc}^1(\Omega; L^1(\mathbb{R}^d))$, ce qui signifie qu'il y a un nombre fini de particules dans toute région bornée. Existe-t-il une équation d'évolution sur f_t ?

Dans le cas où les particules n'interagissent pas entre elles, en utilisant la conservation de la masse, les particules évoluant chacune dans un mouvement rectiligne uniforme, on obtient l'équation de transport libre à l'intérieur de Ω :

$$\partial_t f_t(x, v) + v \cdot \nabla_x f_t(x, v) = 0$$

En réalité, les particules interagissent entre elles, par exemple par des chocs si ce sont des sphères dures, ou par des interactions de type Coulomb. On dit que des particules collisionnent lorsqu'elles interagissent et leurs trajectoires s'en retrouvent fortement modifiées. Pour aboutir à une équation d'évolution sur f_t , Boltzmann a fait 5 hypothèses. La dernière est l'hypothèse du chaos de Boltzmann et se formule de la façon suivante :

Hypothèse du chaos de Boltzmann : Les vitesses de deux particules sur le point de collisionner sont décorréées, c'est-à-dire :

$$\underbrace{f_t^2(z, z^*)}_{\text{densité sur les couples de particules}} = f_t^{\otimes 2}(z, z^*)$$

densité sur les couples de particules

Avec ces hypothèses, on arrive à l'équation de Boltzmann :

$$\partial_t f_t + v \cdot \nabla f_t = Q(f, f) \text{ avec } Q(f, f) = \int_{v^*} dv^* \int_{|\sigma|=1} d\sigma B(v - v^*, \sigma) (f' f'^* - f f^*)$$

où σ représente en quelque sorte la déviation de chacune des particules, les grandeurs indicées étant prises en sortie de collision et celle non indicées en entrée de collision.

Avec l'équation de Boltzmann, bien qu'encore en cours d'étude mathématiquement, fournit le célèbre résultat de physique statistique :

à l'équilibre, la densité des états d'un gaz suit une distribution de Maxwell-Boltzmann : $f(v) = A \exp(-\frac{|v-\bar{v}|^2}{k_B T})$.

1.2 Le problème de dérivation de l'équation

Cette équation a plus de 100 ans mais la théorie autour d'elle est encore inaboutie. Il y a encore de nombreuses choses à vérifier mathématiquement sur ce modèle, notamment qu'il est bien adapté pour décrire le comportement d'un fluide (ou d'au moins certains gaz). A partir de mesures f_t^N , décrivant la dynamique microscopique d'un système à N particules, peut-on, en effectuant une limite d'échelle (en dézoomant en quelque sorte), retrouver une densité limite f_t solution de l'équation de Boltzmann? Ce problème est le problème de dérivation de l'équation. Il est assez complexe, rien que dans sa formulation. Pour le modèle déterministe dont je viens de parler, le théorème n'a été montré que dans le cas de gaz évoluant dans le vide (sur un temps infini) et sur des intervalles de temps très courts dans le cas de la dynamique des sphères dures : c'est le théorème de Lanford. Pour plus de détails à ce sujet, on pourra se référer aux premiers chapitre des ouvrages cités en bibliographie sur l'équation de Boltzmann.

Une autre approche existe, celle-ci étant stochastique, dans laquelle on supprime la donnée spatiale et on ne considère que la dynamique dans l'espace des vitesses. Il s'agit du modèle de Kac. Ce modèle, bien que simplifié, demeure intéressant, et permet d'aboutir à plus de résultats, notamment à un théorème de dérivation de l'équation pour tout temps positif. On trouve dans l'article de Carlen et al cité en référence une revue des résultats et problèmes ouverts sur le modèle de Kac.

Un des aspects du problème de dérivation est de définir une notion sur les suites $(f^N)_N$ et la densité finale f permettant d'aboutir à l'hypothèse du chaos de Boltzmann. Formellement, on souhaite que les marginales d'ordre 2 $f^{N,2}$ "ressemblent" à $f^{\otimes 2}$ lorsque N est grand. On dira alors que (f^N) est chaotique par rapport à f . Les résultats de dérivation sont pour l'instant obtenus pour une notion très faible de chaos. Pour obtenir de nouveaux résultats, il peut être intéressant de comprendre cette notion et de s'intéresser à d'autres façon de la formuler.

1.3 Déroulement du stage

Lors du premier mois du stage, je me suis familiarisée avec les concepts de théorie cinétique des gaz. Lors du deuxième mois, j'ai lu des articles sur le modèle de Kac. J'ai ensuite étudié, à la demande de mon maître de stage, trois façons équivalentes d'écrire l'hypothèse de chaos faible. Je me suis référée pour cela à la revue de L-P.Chaintron et A. Diesz, dans lequel ce travail est rédigé de façon plutôt accessible. Je présente ce résultat dans la première partie du rapport. Je me suis ensuite penchée sur un article de Mischler et Hauray, dans lequel sont présentées d'autres façons de définir le chaos et les liens qui existent entre elles. Cette démonstration figure dans la seconde partie du rapport. Mon travail a été principalement d'étudier ces preuves et de les expliquer à mon maître de stage. J'ai également recherché des contre-exemples permettant de montrer que ces notions ne sont pas équivalentes, mais n'ai pas la place de les exposer ici.

Lors de ce stage à la Sapienza, j'ai travaillé en autonomie. J'avais des entretiens toutes les semaines avec mon encadrant pour lui parler de mon avancée. J'avais également la possibilité d'assister à des séminaires de théorie cinétique, ce que j'ai fait à quelques reprises.

J'ai perçu une bourse Erasmus+ qui m'a aidée à financer les frais de logement.

2 Trois formulations équivalentes du chaos faible

Cette partie s'appuie et s'inspire largement de la partie trois de la revue de L-P Chaintron et A.Diesz. Au lieu de travailler avec des grandeurs à densité " f^N " comme dans l'introduction je travaillerai avec des mesures de probabilité " μ^N ".

2.1 Introduction au chaos faible

Le chaos faible (ou chaos de Kac) est la notion la plus faible de chaos. Elle permet d'obtenir le théorème de dérivation de Kac de façon complète. Commençons par introduire quelques notions et notations. [Les concepts et lemmes essentiels à la démonstration seront suivis d'un \otimes]

Dans cette partie, étant donné un espace E topologique et N un entier strictement positif, on notera $\mathcal{P}_{sym}(E^N)$ l'ensemble des probabilités boréliennes symétriques sur E^N (i.e. invariante par permutation des différentes coordonnées).

Pour μ^N dans $\mathcal{P}_{sym}(E^N)$ et k entier strictement positif, on notera $\mu^{N,k}$ la marginale en les k premières coordonnées de μ^N [$\forall (B_i) \in \mathcal{B}(E)^k, \mu^{N,k}(\prod_{i=1}^k B_i) = \mu^N(\prod_{i=1}^k B_i \times E^{N-k})$]. Pour un espace topologique \mathcal{E} , on note $\mathcal{C}_b(\mathcal{E})$ l'ensemble des fonctions continues bornées de \mathcal{E} dans \mathbb{R} et $\mathcal{C}_0(\mathcal{E})$ le sous-ensemble des fonctions tendant vers 0 en l'infini dans le sens où ϕ tend vers 0 vers l'infini lorsque $\forall \epsilon \in \mathbb{R}^{+*}, \exists K \subset \mathcal{E}$ compact, $\sup_K |\phi| \leq \epsilon$. On notera $\mathcal{C}_c(\mathcal{E})$ l'ensemble des fonctions continues à support compact de \mathcal{E} dans \mathbb{R} .

Définition 2.1. (*Chaos faible*) Soit E un espace topologique (pensons-y comme à l'espace des vitesses d'une particule dans le modèle de Kac, ou à l'espace des phases dans le modèle déterministe). Pour tout n dans \mathbb{N}^* , soit μ^N une probabilité borélienne symétrique sur E^N et soit μ une probabilité borélienne sur E . On dit que (μ^N) est μ -chaotique (faiblement) lorsque $\forall k \in \mathbb{N}^*, \mu^{N,k} \rightharpoonup \mu^{\otimes k}$ (je préciserai plus tard de quelle convergence il s'agit).

Cette définition est en tant que telle un peu délicate à manipuler, notamment pour montrer qu'une suite est chaotique par rapport à une mesure de probabilité, puisqu'il faut vérifier la convergence pour toutes les marginales. Il s'avère en fait que cette définition est équivalente à la convergence seulement des marginales d'ordre 1 et 2. La preuve ne se fait pas de manière directe, et on va être obligé.e.s d'effectuer un détour et d'introduire de nouveaux concepts. Une autre difficulté avec cette notion est que les μ^N ne sont pas définies sur le même espace probabilisable. On va voir qu'on peut en fait changer de point de vue et plonger les μ^N dans un espace commun et obtenir une nouvelle caractérisation du chaos s'exprimant comme une convergence sur un espace métrique. Cet espace commun sera l'ensemble $\mathcal{P}(\mathcal{P}(E))$ des probabilités sur $\mathcal{P}(E)$. A toute probabilité symétrique μ^N sur E^N , on associera une probabilité sur $\mathcal{P}(E)$ notée $\hat{\mu}^N$. Je montrerai dans cette partie le résultat suivant. (Les notions manquantes seront précisées plus bas).

Théorème 2.1. (*Trois formulations équivalentes du chaos faible*) Soit E un espace métrique séparable et complet. Pour tout N dans \mathbb{N}^* , soit $\mu^N \in \mathcal{P}_{sym}(E^N)$ et soit $\mu \in \mathcal{P}(E)$. Il est équivalent de dire :

1. (μ^N) est μ chaotique.
2. $\mu^{N,2} \rightharpoonup \mu^{\otimes 2}$ (au sens de la convergence faible-* dans $\mathcal{P}(E^2)$)
3. $\hat{\mu}^N \rightharpoonup \delta_\mu$ (au sens de la convergence faible-* dans $\mathcal{P}(\mathcal{P}(E))$)

2.2 Deux types de distances sur $\mathcal{P}(\mathcal{E})$

2.2.1 Métriser la convergence faible

Commençons par introduire les notions dont on va avoir besoin pour l'énoncé et la démonstration. On se donne ici un espace métrique \mathcal{E} .

Définition 2.2. (*Convergence faible \otimes*) Pour $(\mu_n)_{n \in \mathbb{N}}$ une suite de probabilités sur un \mathcal{E} et μ une autre probabilité sur cette espace, on dit que (μ_n) converge faiblement vers μ et on note $\mu_n \rightharpoonup \mu$ lorsque $\forall \phi \in \mathcal{C}_b(\mathcal{E}), \langle \mu_n, \phi \rangle \xrightarrow{n \rightarrow \infty} \langle \mu, \phi \rangle$.

Remarque : Cette notion de convergence est celle héritée de la topologie faible \rightarrow^* sur $\mathcal{C}_b(\mathcal{E})^* \supset \mathcal{P}(\mathcal{E})$. D'un point de vue probabiliste, elle correspond à la convergence en loi (si l'on considère les μ_n comme loi de variables aléatoire).

Il peut être utile de se restreindre à certains ensembles de fonctions, par exemple des fonctions plus régulières, pour montrer des convergences faibles, ce qui motive, entre autre, la définition suivante.

Définition 2.3. (Ensemble déterminant pour la convergence en loi) Un ensemble (ou une famille) $\mathcal{F} \subset \mathcal{C}_b(\mathcal{E})$ est dit déterminant pour la convergence en loi lorsque $\forall (\mu_n), \mu \in \mathcal{P}(\mathcal{E})$,

$$[\mu_n \rightharpoonup \mu] \Leftrightarrow [\forall \phi \in \mathcal{F}, \langle \mu_n, \phi \rangle \xrightarrow{n \rightarrow \infty} \langle \mu, \phi \rangle]$$

Exemple 1.

- Par le théorème de Portemanteau, l'ensemble des fonctions uniformément continues bornées $\mathcal{UC}_b(\mathcal{E})$ est déterminant pour la convergence en loi.
- Supposons que (\mathcal{E}, d) soit séparable et complet. Alors pour k entier, $\mathcal{C}_b(\mathcal{E})^{\otimes k}$ l'ensemble des fonctions continues bornées sur \mathcal{E}^k s'écrivant comme produit de fonctions continues bornées d'un seul argument est déterminant pour la convergence faible sur $\mathcal{P}(\mathcal{E}^k)$. On peut le trouver dans la proposition 4.6 du Ethier Kurz.
- Un sous-ensemble dense pour $\| \cdot \|_\infty$ d'un ensemble déterminant pour la convergence faible l'est également.

Théorème 2.2. \otimes Supposons que (\mathcal{E}, d) soit un espace métrique séparable. Alors il existe un ensemble dénombrable déterminant pour la convergence faible sur \mathcal{E} . On peut de plus choisir des fonctions lipschitziennes et bornées en valeur absolue par 1 en les construisant astucieusement.

Comme corollaire presque immédiat, on obtient que lorsque (\mathcal{E}, d) est séparable, la convergence faible est métrisable / prendre par exemple étant donné (φ_k) déterminante pour la convergence $D_1(\mu, \nu) := \sum_{k=1}^{\infty} \frac{1}{2^k} |\langle \mu - \nu, \varphi_k \rangle|$.

Preuve. (Esquisse de preuve)

1. Dans le cas où \mathcal{E} est compact, utiliser le théorème de Stone-Weierstrass pour l'algèbre engendrée par des fonctions bien choisie pour trouver une famille dénombrable dense (pour la norme infini) de fonctions continues bornées (prendre l'algèbre engendrée par les $d(\cdot, x_i)$ où (x_i) est une suite dense).
2. On suppose désormais que \mathcal{E} est localement compact et séparable, et quitte à considérer $\min(d, 1)$ que d est bornée. Montrer que $\mathcal{C}_0(\mathcal{E})$ est séparable. On peut par exemple utiliser le compactifié d'Alessandrov \mathcal{E}^* de \mathcal{E} pour se ramener au cas précédent, et montrer que $\mathcal{C}(\mathcal{E}) \simeq \mathcal{C}_0(\mathcal{E}^*) + \mathbb{R}$. On conclut pour le cas \mathcal{E} localement compact.
3. On suppose désormais que \mathcal{E} est simplement séparable. On choisit une suite dense $(x_i)_{i \in \mathbb{N}}$.

Notons $\phi := \begin{cases} (\mathcal{E}, d) & \longrightarrow & ([0, 1]^{\mathbb{N}}, d_\infty) \\ x & \longmapsto & (\min(d(x, x_i), 1)_{i \in \mathbb{N}}, 1) \end{cases}$, où $d_\infty(u, v) := \sum_{k=1}^{\infty} \frac{1}{2^k} |u_k - v_k|$

est une distance métrisant le compact $[0, 1]^{\mathbb{N}}$ pour la topologie produit. Enfin, pour tout $(x, y) \in \mathcal{E}^2$, notons $\tilde{d}(x, y) := d_\infty(\phi(x), \phi(y))$. Alors \tilde{d} est une autre distance sur \mathcal{E} . (\mathcal{E}, \tilde{d}) est un espace métrique totalement borné (ie de toute suite on peut extraire une suite de Cauchy) par compacité de $([0, 1]^{\mathbb{N}}, d_\infty)$. Le complété $(\tilde{\mathcal{E}}, \tilde{d})$ est donc compact. On peut alors appliquer 1. Or, $(\mathcal{UC}_b(\mathcal{E}, \tilde{d}), \| \cdot \|_\infty) \simeq (\mathcal{C}_b(\tilde{\mathcal{E}}, \tilde{d}), \| \cdot \|_\infty)$, ce dernier étant séparable par compacité et on peut montrer que $\tilde{d} \leq d$ donc $\mathcal{UC}_b(\mathcal{E}, \tilde{d}) \supset \mathcal{UC}_b(\mathcal{E}, d)$. $\mathcal{UC}_b(\mathcal{E}, d)$ est donc séparable comme sous-ensemble d'un espace métrique séparable, et le théorème de Portemanteau permet de conclure.

Remarque : La métrique \tilde{d} est en fait (localement) équivalente à d . Le raisonnement suivant est exposé en grande partie page 45 de la revue de Chaintron-Diesz.

2.2.2 Les distances de Monge-Kantorovitch-Wassertsein

Dans la partie précédente, on a introduit une distance D_1 sur $\mathcal{P}(\mathcal{E})$ pour tout espace métrique séparable (\mathcal{E}, d) . Il existe une autre façon assez naturelle, héritée du transport optimal, d'introduire une notion de proximité sur $\mathcal{P}(\mathcal{E})$: les distances de Monge-Kantorovitch-Wasserstein (MKW). Je me contenterai ici d'introduire la notion et quelques théorèmes utiles pour la suite sans démontrer de résultat.

Définition 2.4. (*p-distance de MKW*) Soit $\mu \in \mathcal{P}(\mathcal{E})$. Soit $p \geq 1$. On dit que μ admet un moment d'ordre p lorsque $\exists x_0, \int_x d(x, x_0)^p d\mu(x) < +\infty$.

Pour deux mesures μ et ν admettant un moment d'ordre p , on définit la p -distance de MKW par $W_{d,p}(\mu, \nu) := \inf_{\pi \in \Pi(\mu, \nu)} \left(\int_{(x,y)} d(x,y)^p d\pi(x,y) \right)^{\frac{1}{p}}$ avec $\Pi(\mu, \nu)$ l'ensemble des mesures de probabilités boréliennes sur $\mathcal{E} \times \mathcal{E}$ de première marginale μ et de deuxième marginale ν .

$W_{d,p}$ définit alors une distance sur l'ensemble des mesures de probabilités boréliennes sur \mathcal{E} admettant un moment d'ordre p .

On peut formuler $W_{d,p}$ d'une façon plus probabiliste, ce qui sera un point de vue utile pour la suite : $W_{d,p} := \inf_{Loi(X) \sim \mu, Loi(Y) \sim \nu} \mathbb{E}[d(X, Y)^p]^{\frac{1}{p}}$ (où X et Y sont des variables aléatoires d'un même espace probabilisé et à valeurs dans \mathcal{E}). Dans cette partie on se servira uniquement de distances d'ordre 1, on omettra donc de préciser l'ordre. On peut ainsi créer une famille de distances sur $\mathcal{P}(\mathcal{E})$ en faisant varier d .

Définition 2.5. (*Variation totale*) Pour deux mesures de probabilités boréliennes μ et ν , on définit : $TV(\mu, \nu) := \inf_{Loi(X) \sim \mu, Loi(Y) \sim \nu} \mathbb{P}(X \neq Y)$ ($= W_{1-\delta_{x,y},1}$).

Théorème 2.3. (*Formule de dualité* \otimes)

- $TV(\mu, \nu) = \sup_{\|\varphi\|_\infty} < \mu - \nu, \varphi >$
- $W_d(\mu, \nu) = \sup_{\|\varphi\|_{Lip,d}} < \mu - \nu, \varphi >$ avec $\|\varphi\|_{Lip,d} = \sup_{x \neq y} \frac{|\varphi(x) - \varphi(y)|}{d(x,y)}$

Remarque : Les distances de Wasserstein définissent une topologie légèrement plus forte que la convergence en loi. $W_{d,p}$ métrise la convergence en loi sur l'ensemble des probabilités admettant un moment d'ordre p (en ajoutant de plus la convergence des moments d'ordres p). On peut trouver une démonstration de ce théorème dans Topics in optimal Transportation.

2.3 L'espace $\mathcal{P}(\mathcal{P}(E))$

Soit (E, d) un espace métrique séparable complet. Avec ce qui a été fait dans la partie précédente, une distance D_1 sur $\mathcal{P}(E)$ bornée et métrisant la convergence en loi. $(\mathcal{P}(E), D_1)$ est un espace métrique. En le munissant de sa tribu borélienne par rapport à D_1 , on en fait un espace probabilisable. On notera $\mathcal{P}(\mathcal{P}(E))$ l'ensemble des probabilités sur l'espace probabilisable $(\mathcal{P}(E), D_1, \mathcal{B}(\mathcal{P}(E)))$. D_1 étant bornée, on peut s'intéresser à W_{D_1} pour définir une distance sur $\mathcal{P}(\mathcal{P}(E))$.

Comme on l'a vu plus haut, on peut aussi regarder $\mathcal{P}(\mathcal{P}(E))$ avec la topologie induite par la topologie faible- \star sur $\mathcal{C}_b(\mathcal{P}(E), D_1)^*$.

Ces deux topologies sur $\mathcal{P}(\mathcal{P}(E))$ se parlent entre elles. En fait, la convergence W_{D_1} est plus forte que la convergence faible \star .

Lemme 2.1. \otimes Soient $(\mu_n)_{n \in \mathbb{N}}$ et μ des probabilités aléatoires sur E (ie $M_n = Loi(\mu_n) \in \mathcal{P}(\mathcal{P}(E))$), de même que $M = Loi(\mu)$.

1. $W_{D_1}(Loi(\mu_n), Loi(\mu)) \rightarrow 0 \Rightarrow \mu_n \xrightarrow{(\mathcal{L})} \mu$ (en tant que variables aléatoires, ce qui équivaut à $Loi(\mu_n) \rightarrow \delta_\mu$ dans $\mathcal{P}(\mathcal{P}(E))$).
2. Si les probabilités aléatoires partent d'un même espace probabilisé $\mu_n, \mu : \Omega \rightarrow \mathcal{P}(E)$, alors $(\forall \varphi \in \mathcal{UC}_b(E)), \mathbb{E}[< \mu_n - \mu, \varphi >] \rightarrow 0 \Rightarrow W_{D_1}(Loi(\mu_n), Loi(\mu)) \rightarrow 0$.

Preuve. 1. Avec le théorème de Portemanteau, on peut se contenter de montrer la convergence pour des fonctions dans $\mathcal{UC}_b(\mathcal{P}(E), D_1)$. Soit φ une telle fonction test. Soit $\epsilon > 0$. Choisissons δ module de continuité de φ adapté à ϵ . Soient $\mu'_n \sim M_n$ et $\mu' \sim M$.

$$\begin{aligned} |\mathbb{E}[\varphi(\mu_n)] - \mathbb{E}[\varphi(\mu)]| &\leq \mathbb{E}[|\varphi(\mu'_n) - \varphi(\mu')|] \\ &\leq \mathbb{E}[\mathbb{1}_{|\varphi(\mu'_n) - \varphi(\mu')| < \delta} |\varphi(\mu'_n) - \varphi(\mu')|] + \mathbb{E}[\mathbb{1}_{|\varphi(\mu'_n) - \varphi(\mu')| > \delta} |\varphi(\mu'_n) - \varphi(\mu')|] \end{aligned}$$

Avec le choix de δ , $\mathbb{E}[\mathbb{1}_{|\varphi(\mu'_n) - \varphi(\mu')| < \delta} |\varphi(\mu'_n) - \varphi(\mu')|] \leq \epsilon$.
 φ étant bornée, $\mathbb{E}[\mathbb{1}_{|\varphi(\mu'_n) - \varphi(\mu')| > \delta} |\varphi(\mu'_n) - \varphi(\mu')|] \leq 2 \|\varphi\|_\infty \mathbb{P}(D_1(\mu'_n, \mu') > \delta)$.

Par Markov, $\mathbb{P}(D_1(\mu'_n, \mu') > \delta) \leq \frac{\mathbb{E}[D_1(\mu'_n, \mu')]}{\delta}$.

En prenant l'inf sur μ'_n et μ' avec la formulation probabiliste des distances de Wasserstein :

$$|\mathbb{E}[\varphi(\mu_n)] - \mathbb{E}[\varphi(\mu)]| \leq \epsilon + 2 \|\varphi\|_\infty \frac{W_{D_1}(\mu_n, \mu)}{\delta}$$

Par hypothèse, $W_{D_1}(\mu_n, \mu) \xrightarrow[n \rightarrow \infty]{} 0$, ce qui conclut.

2. Supposons que $\varphi \in \mathcal{C}_b(E)$, $\mathbb{E}[|\langle \mu_n - \mu, \varphi \rangle|] \xrightarrow[n \rightarrow \infty]{} 0$. Avec le théorème de convergence dominée, en appliquant cela aux φ_k utilisés pour définir D_1 , $\sum_{k \in \mathbb{N}^*} \frac{1}{2^k} \mathbb{E}[|\langle \mu_n - \mu, \varphi_k \rangle|] \xrightarrow[n \rightarrow \infty]{} 0$. Or par Fubini-Tonelli, le terme de gauche est égal à $\mathbb{E}[D_1(\mu_n, \mu)]$, qui domine $W_{D_1}(\mu_n, \mu)$.

Dans le cas où la mesure aléatoire finale est déterministe (ie $\mathcal{P}(\mathcal{P}(E)) \ni \mu = \bar{\mu} \in \mathcal{P}(E)$ presque sûrement), les deux implications du lemme précédent sont en fait des équivalences.

Lemme 2.2. On conserve le contexte du 2. du lemme précédent. Supposons que μ soit déterministe.

Alors si $\mu_n \xrightarrow[n \rightarrow \infty]{(\mathcal{L})} \mu$, $\forall \varphi \in \mathcal{C}_b(E)$, $\mathbb{E}[|\langle \mu_n - \mu, \varphi \rangle|] \xrightarrow[n \rightarrow \infty]{} 0$.

Preuve. $\mu = \bar{\mu} \in \mathcal{P}(E)$ p.s.

Soit $\varphi \in \mathcal{C}_b(E)$. Notons $\phi := [\nu \mapsto |\langle \nu - \bar{\mu}, \varphi \rangle|]$. ϕ est continue pour la topologie faible- \star sur $\mathcal{P}(E)$ comme composée d'une application continue pour cette topologie de $\mathcal{P}(E)$ dans \mathbb{R} par une application continue de \mathbb{R} dans \mathbb{R} . ϕ est donc continue pour D_1 , qui métrise cette topologie. Elle est de plus bornée par $\|\varphi\|_\infty$. Supposons que $\mu_n \xrightarrow[n \rightarrow \infty]{(\mathcal{L})} \bar{\mu}$. Par définition de la convergence en loi de variables aléatoires, comme $\phi \in \mathcal{C}_b(\mathcal{P}(E), D_1)$,

$$\begin{aligned} \mathbb{E}[\phi(\mu_n)] &\xrightarrow[n \rightarrow \infty]{} \mathbb{E}[\phi(\bar{\mu})] = 0 \\ &\parallel \\ &\mathbb{E}[|\langle \mu_n - \mu, \varphi \rangle|] \end{aligned}$$

Introduisons finalement un dernier outil sur cet espace : les fonctions polynomiales.

2.3.1 Une famille déterminante pour la convergence faible sur $\mathcal{P}(\mathcal{P}(E))$

L'espace $\mathcal{C}_b(\mathcal{P}(E), D_1)$ est beaucoup plus complexe que $\mathcal{C}_b(E, d)$ et semble moins naturel. La convergence faible- \star sur $\mathcal{P}(\mathcal{P}(E))$ semble a priori beaucoup plus délicate à étudier que celle sur $\mathcal{P}(E)$. Cependant, on dispose d'une famille de fonctions relativement intuitive permettant de caractériser parfaitement cette convergence.

Définition 2.6. (Monômes) Soit $k \in \mathbb{N}^*$. Pour $\varphi_k \in \mathcal{C}_b(E^k)$, on définit le monôme d'ordre k

$$\text{par } R_{\varphi_k} := \left| \begin{array}{ccc} (\mathcal{P}(E), D_1) & \longrightarrow & (\mathbb{R}, |\cdot|) \\ \nu & \longmapsto & \langle \nu^{\otimes k}, \varphi_k \rangle \end{array} \right.$$

Lemme 2.3. 1. $\forall \varphi_k \in \mathcal{C}_b(E^k), R_{\varphi_k} \in \mathcal{C}_b(\mathcal{P}(E))$ \otimes

2. La famille des R_{φ_k} est déterminante pour la convergence faible- \star sur $\mathcal{P}(\mathcal{P}(E))$.

Preuve. 1. Facilement, $\|R_{\varphi_k}\|_\infty \leq \|\varphi_k\|_\infty$. Pour la continuité, puisqu'on travaille avec des topologies métrisables, il suffit de montrer la continuité séquentielle.

Soit $\mu_n \rightarrow \mu$. Soit $k \geq 1$. Travaillons d'abord dans le cas où $\varphi_k = \bigotimes_{1 \leq j \leq k} \varphi_k^j \in \mathcal{C}_b(E)^{\otimes k}$. Par le théorème de Fubini :

$$R_{\varphi_k} \mu_n = \prod_{1 \leq j \leq k} \underbrace{\langle \mu_n, \varphi_k^j \rangle}_{\xrightarrow[n \rightarrow \infty]{} \langle \mu, \varphi_k^j \rangle} \xrightarrow[n \rightarrow \infty]{} \prod_{1 \leq j \leq k} \langle \mu, \varphi_k^j \rangle = R_{\varphi_k} \mu$$

On a donc $\forall \varphi_k \in \mathcal{C}_b(E)^{\otimes k}, R_{\varphi_k} \mu_n \rightarrow R_{\varphi_k} \mu$, soit, de façon plus explicite, $\forall \varphi_k \in \mathcal{C}_b(E)^{\otimes k}, \langle \mu_n^{\otimes k}, \varphi_k \rangle \xrightarrow[n \rightarrow \infty]{} \langle \mu_n^{\otimes k}, \varphi_k \rangle$. Or, $\mathcal{C}_b(E)^{\otimes k}$ est déterminant pour la convergence faible- \star sur $\mathcal{P}(E)$, et le résultat s'étend alors à tout φ_k de $\mathcal{C}_b(E^k)$.

2. On admet ce résultat qui est démontré page 48 de la revue de Chaintron et Diesz et qui s'appuie sur des théorèmes de Markov processes : characterization and convergence.

Ainsi, pour montrer une convergence d'une suite de mesures aléatoires, il suffit de regarder ce qu'il se passe pour les monômes.

Maintenant qu'on a les outils nécessaires, revenons progressivement au problème.

2.3.2 Changement de perspective : de $\mathcal{P}_{sym}(E^N)$ vers $\mathcal{P}(\mathcal{P}(E))$

Dans la première formulation du chaos faible, on étudie $(\mu^N)_{N \in \mathbb{N}^*}$ où $\mu^N \in \mathcal{P}_{sym}(E^N)$. Un inconvénient de ce problème est que les mesures ne sont pas sur le même espace. Tel quel, on ne peut pas formuler le chaos comme convergence d'une suite vers un objet limite. Cependant, on peut s'y ramener en plongeant tous nos objets dans $\mathcal{P}(\mathcal{P}(E))$.

Soit $N \geq 1$ un entier (on pensera ici à N comme au nombre de particules du système).

Définition 2.7. (Mesure empirique \otimes) Pour tout $Z^N = (z_i)_{i=1}^N \in E^N$, notons $\omega_{Z^N} := \frac{1}{N} \sum_{1 \leq j \leq N} \delta_{z_j} \in \mathcal{P}(E)$ la mesure empirique associée à Z^N .

En utilisant une caractérisation séquentielle, on montre facilement la proposition suivante.

Proposition 2.1. ω est continue de $E^N \rightarrow (\mathcal{P}(E), D_1)$, donc mesurable de $(E^N, \mathcal{B}(E^N)) \rightarrow (\mathcal{P}(E), \mathcal{B}(\mathcal{P}(E)))$.

Ainsi, pour toute mesure $\mu^N \in \mathcal{P}(E^N)$, on peut considérer la mesure image de μ^N par ω , que l'on notera $\hat{\mu}^N (\in \mathcal{P}(\mathcal{P}(E)))$.

D'un point de vue plus probabiliste, supposons que X^N soit une variable aléatoire de loi μ^N . Alors $Loi(\omega_{X^N}) = \hat{\mu}^N$.

Définition 2.8. (Moment d'ordre k d'une mesure aléatoire \otimes) Soit μ une probabilité aléatoire sur \mathcal{E} de loi $M \in \mathcal{P}(\mathcal{P}(E))$. Soit $k \geq 1$ un entier. La mesure moment d'ordre k de M (ou de μ) par dualité de la façon suivante : $\forall \varphi_k \in \mathcal{C}_b(E^k), \langle \hat{\mu}, \varphi_k \rangle = \langle M^k, \varphi_k \rangle := \int_{\nu \in \mathcal{P}(E)} \langle \nu^{\otimes k}, \varphi_k \rangle dM(\nu)$. Avec le théorème de Riesz-Markov-Kakutani, il est bien possible de définir cette mesure par dualité dès lors que E est localement compact. Cette mesure est une mesure de probabilité sur E^k .

Les moments encodent bien le comportement $N \rightarrow \infty$ des marginales d'ordre fixe d'une suite $(\hat{\mu}^N)$ dès lors que les μ^N sont symétriques.

Proposition 2.2. (Taux d'approximation des marginales \otimes) Pour tout N dans \mathbb{N}^* , soit μ^N dans $\mathcal{P}_{sym}(E^N)$. Alors $TV(\mu^{N,k}, \hat{\mu}^{N,k}) \leq \frac{2k(k-1)}{N}$ pour tout k entier plus petit que N .

Preuve. La preuve consiste en du dénombrement et l'utilisation de la formule de dualité sur TV en utilisant la symétrie de μ^N . Soit $\varphi_k \in \mathcal{C}_b(E^k)$. Pour tout $Z^N \in E^N$, notons $Z_1^k := (z_i)_{i=1}^k \in E^k$.

$$\begin{aligned} \langle \mu^{N,k}, \varphi_k \rangle &= \int_{\mathcal{E}^N} \varphi_k(Z_1^k) d\mu^N(Z) \\ &= \frac{1}{N!} \sum_{\sigma \in \mathfrak{S}_N} \int_{\mathcal{E}^N} \varphi_k((Z \circ \sigma)_1^k) d\mu^N(Z) \quad (\text{par symétrie de } \mu^N) \end{aligned}$$

$$\text{et } \langle \hat{\mu}^{N,k}, \varphi_k \rangle = \text{int}_{\mathcal{E}^N} \langle \omega_{Z^N}^{\otimes k}, \varphi_k \rangle d\mu^N(Z)$$

donc $|\langle \mu^{N,k}, \varphi_k \rangle - \langle \hat{\mu}^{N,k}, \varphi_k \rangle| \leq \sup_{Z \in \mathcal{E}^N} \left| \frac{1}{N!} \sum_{\sigma \in \mathfrak{S}_N} \varphi_k((Z \circ \sigma)_1^k) - \langle \omega_{Z^N}^{\otimes k}, \varphi_k \rangle \right|$
D'autre part, $S_N^k := \frac{1}{N!} \sum_{\sigma \in \mathfrak{S}_N} \varphi_k((Z \circ \sigma)_1^k) = \frac{1}{A_N^k} \sum_{\sigma \in \mathcal{A}_N^k} \varphi_k((Z \circ \sigma)_1^k)$ avec A_N^k l'ensemble des k arrangements de $\{1, \dots, N\}$ et $A_N^k := \#\mathcal{A}_N^k = \frac{N!}{(N-k)!}$.

Etudions l'autre terme. Avec une récurrence finie sur k , on obtient :

$$\langle \omega_{Z^N}^{\otimes k}, \varphi_k \rangle = \frac{1}{N^k} \sum_{i \in \{1, \dots, N\}^k} \varphi_k((z_{i_j})_{j=1}^k)$$

$\begin{array}{ccc} i \text{ arrangement} & & \text{reste } R_{k,N} \\ \text{occurrence : } A_N^k & \swarrow \quad \searrow & \text{occurrence : } N^k - A_N^k \end{array}$

On obtient $\langle \omega_{Z^N}^{\otimes k}, \varphi_k \rangle = \frac{1}{N^k} \sum_{k \in \mathcal{A}_N^k} \varphi_k((Z \circ \sigma)_1^k) + R_{k,N}$ avec $|R_{k,N}| \leq (1 - \frac{A_N^k}{N^k}) \|\varphi_k\|_\infty$ (et ce pour tout Z^N).

$$\begin{aligned} \text{On a alors } (*) &:= \left| \frac{1}{N!} \sum_{\sigma \in \mathfrak{S}_N} \varphi_k((Z \circ \sigma)_1^k) - \langle \omega_{Z^N}^{\otimes k}, \varphi_k \rangle \right| \leq \left(\frac{1}{A_N^k} - \frac{1}{N^k} \right) \sum_{\sigma \in \mathcal{A}_N^k} |\varphi_k((Z \circ \sigma)_1^k)| + |R_{k,N}| \\ &\leq 2 \left(1 - \frac{A_N^k}{N^k} \right) \|\varphi_k\|_\infty \end{aligned}$$

or $A_N^k \geq (N-k+1)^k$ donc $\frac{A_N^k}{N^k} \geq (1 - \frac{k-1}{N})^k \geq 1 - \frac{k(k-1)}{N}$ (par convexité de $x \mapsto (1+x)^k$ sur $[0, 1]$) donc $(*) \leq 2 \frac{k(k-1)}{N} \|\varphi_k\|_\infty$.

Avec la formulation duale de TV, on obtient le résultat.

Remarque : Ce résultat est valable pour (μ^N) très générale, pas forcément chaotique par rapport à une certaine mesure. On a alors le corollaire suivant.

Nous sommes désormais suffisamment armé.e.s pour démontrer le théorème qui nous intéresse.

2.4 Preuve

Je vais commencer par donner un plan de la preuve.

Plan de la preuve :

- Chaos 1 \Rightarrow Chaos 2 est directe.
- Pour Chaos 2 \Rightarrow Chaos 3, on utilise le lemme 2.1 : Si les probabilités aléatoires partent d'un même espace probabilisé $\mu_n, \mu : \Omega \rightarrow \mathcal{P}(E)$, alors $(\forall \varphi \in \mathcal{UC}_b(E)), \mathbb{E}[|\langle \mu_n - \mu, \varphi \rangle|] \rightarrow 0 \Rightarrow W_{D_1}(Loi(\mu_n), Loi(\mu)) \rightarrow 0$.
- Pour Chaos 3 \Rightarrow Chaos 1, on utilise le lemme 2.3, qui énonce que la convergence faible-* sur $\mathcal{P}(E)$ implique la convergence faible des moments d'ordre k et la proposition 2.2 qui compare $\mu^{N,k}$ et $\hat{\mu}^{N,k}$.

2.4.1 Chaos 2 \Rightarrow Chaos 3

Soit (μ^N) et μ vérifiant le chaos 2.

Pour tout $N \geq 1$, soit $X^N = (X_i)_{i=1}^N$ une variable aléatoire sur E^N de loi μ^N . En développant et par formule de transfert :

$$\begin{aligned} \mathbb{E}[\langle \omega_{X^N} - \delta_{\mu}, \varphi \rangle^2] &= \frac{1}{N^2} \sum_{1 \leq i, j \leq N} \mathbb{E}[\varphi(X_i)\varphi(X_j)] - \underbrace{\frac{2}{N} \sum_{i=1}^N \mathbb{E}[\varphi(X_i)]}_{\substack{\downarrow \text{(par le chaos 2 et par formule de transfert)} \\ 2 \langle \mu, \varphi \rangle}} \langle \mu, \varphi \rangle + \langle \mu, \varphi \rangle^2 \end{aligned}$$

Or par symétrie de μ^N ,

$$\begin{aligned} \frac{1}{N^2} \sum_{i, j} \mathbb{E}[\varphi(X_i)\varphi(X_j)] &= \frac{1}{N^2} (N\mathbb{E}[\varphi(X_1)^2] + (N^2 - N)\mathbb{E}[\varphi(X_1)\varphi(X_2)]) \\ &= \frac{1}{N} \mathbb{E}[\varphi(X_1)^2] + \frac{N-1}{N} \mathbb{E}[\varphi(X_1)\varphi(X_2)] \end{aligned}$$

$\frac{1}{N} \mathbb{E}[\varphi(X_1)^2] \xrightarrow[N \rightarrow \infty]{} 0$ et par hypothèse (chaos 2) $\frac{N-1}{N} \mathbb{E}[\varphi(X_1)\varphi(X_2)] \xrightarrow[N \rightarrow \infty]{} \langle \mu, \varphi \rangle^2$ donc $\mathbb{E}[\langle \omega_{X^N} - \delta_{\mu}, \varphi \rangle^2] \xrightarrow[N \rightarrow \infty]{} 0$. Le lemme 2.1 et Cauchy-Schwarz concluent ! Remarque : Avec le lemme

2.4.2 Chaos 3 \Rightarrow Chaos 1

Soient (μ^N) et μ vérifiant le chaos 3 et choisissons (X^N) comme dans la partie précédente.

$$\begin{aligned} \mathbb{E}[R_{\varphi_k}[\omega_{X^N}]] &\xrightarrow[N \rightarrow \infty]{} R_{\varphi_k}[\mu] \\ \omega_{X^N} \xrightarrow{X} \delta_{\mu} \text{ donc Avec le lemme 2.3 : } \forall k \in \mathbb{N}^*, \forall \varphi_k \in \mathcal{C}_b(E^k), &\mathbb{E}[\langle \omega_{X^N}^{\otimes k}, \varphi_k \rangle] \xrightarrow[N \rightarrow \infty]{} \langle \mu^{\otimes k}, \varphi_k \rangle \\ &\parallel \\ &\langle \hat{\mu}^{N,k}, \varphi_k \rangle = \langle \mu^N, \varphi_k \rangle + o_N(1) \end{aligned}$$

puisque par la proposition 2.2 et la formule de dualité, $|\langle \hat{\mu}^{N,k} - \mu^{N,k}, \varphi_k \rangle| \leq TV(\hat{\mu}^{N,k}, \mu^{N,k}) \parallel \varphi_k \parallel_{\infty} \xrightarrow[N \rightarrow \infty]{} 0$. On obtient finalement que $\forall k \in \mathbb{N}^*, \forall \varphi_k \in \mathcal{C}_b(E^k), \langle \mu^N, \varphi_k \rangle \xrightarrow[N \rightarrow \infty]{} \langle \mu^{\otimes k}, \varphi_k \rangle$, c'est-à-dire que le chaos 1 est vérifié.

3 Quelques notions plus fortes de chaos

Le chaos est une façon de formuler le chaos pour le problème de dérivation. Même si cette notion permet d'aboutir à des théorèmes de dérivation, il peut être intéressant d'étudier ce qu'il se passe lorsqu'on choisit une formulation différente, plus forte.

Dans leur article, Mischler et Hauray proposent des formes différentes de chaos, faisant intervenir des grandeurs que l'on pourrait qualifier de thermodynamiques ou statistiques. Ils montrent qu'elles impliquent bien le chaos faible (sous certaines hypothèses), et établissent une hiérarchie entre elles. Dans cette partie je redémontre les résultats de la troisième partie du papier en essayant de rendre cela le plus pédagogique possible. Je vais commencer par énoncer le résultat et préciserai plus loin les définitions. Dans cette partie, $E = \mathbb{R}^d$ avec $d \geq 1$ entier.

L'entropie H_d et l'information de Fisher I_d sont des fonctionnelles partant de $\bigcup_{N \in \mathbb{N}^*} \mathcal{P}(E^N)$ et arrivant dans $\mathbb{R} \cup \{+\infty\}$. On peut définir des notions de chaos par rapport à celles-ci, que l'on verra plus bas. Dans cette partie, je montrerai le théorème suivant :

Théorème 3.1. (Comparaison des formes de chaos) Pour tout $N \in \mathbb{N}^*$, soit $\mu^N \in \mathcal{P}_{sym}(E^N)$. Soit $\mu \in \mathcal{P}(E)$. On suppose de plus que $(\langle \mu^{N,1}, |z|^p \rangle)$ la suite des moments d'ordre p des premières marginales est bornée pour un certain $p > 2$. Chacune des propositions implique la suivante.

- IC- (μ^N) est I_d -chaotique par rapport à μ .
- IBC- (μ^N) est faiblement chaotique par rapport à μ et $(I_d(\mu^N))$ est bornée.
- HC- (μ^N) est H_d -chaotique par rapport à μ .
- KC- (μ^N) est faiblement chaotique par rapport à μ .

Remarque : Dans leur papier, Mischler et Hauray montrent ce résultat pour E plus général, à savoir déformation bilipshitz et avec préservation du volume d'un ouvert ou de l'adhérence d'un ouvert convexe. Je me contenterais ici d'effectuer la démonstration dans le cas où E est l'espace entier, qui englobe l'essentiel du raisonnement effectué pour démontrer le théorème.

Plan de la démonstration :

- Pour $(IC) \Rightarrow (IBC)$ et $(HC) \Rightarrow (KC)$, on va se servir d'un même lemme, le lemme 3.1, qui établit que le chaos par rapport à un certain type de fonctionnelles est plus fort que le chaos faible et utiliser le fait que I_d et H_d sont toutes deux des fonctionnelles de ce type, l'hypothèse sur les moments d'ordre p assurant la tension par Markov.
- Pour $(IBC) \Rightarrow (HC)$, on utilise l'inégalité HWI, qui permet de relier la variation d'entropie entre deux probabilités à I_d et à des distances de Wasserstein ainsi qu'une formulation quantitative du chaos faible, démontrée dans la deuxième partie du papier de Mischler et Hauray.

3.1 Chaos par rapport à une fonctionnelle extensive

Dans cette partie je définis la notion de fonctionnelle extensive (appellation personnelle), une notion de chaos par rapport à une fonctionnelle de ce type et montre une hiérarchie entre ce chaos-là et le chaos faible.

Pour tout $p > 0$ et $N \geq 1$ entier, on note $\Gamma_p^N := \left| \begin{array}{ccc} \mathcal{E}^N & \longrightarrow & \mathbb{R}^*_+ \\ Z = (z_i)_{1 \leq i \leq N} & \longmapsto & g_p^N \exp(-\sum_i |z_i|^p) \end{array} \right.$,

où g_p est choisi pour faire de Γ_p^1 une densité de probabilité sur E . On note $\mathcal{P}_p(E^N)$ l'ensemble des probabilités boréliennes sur E^N admettant des moments d'ordres p par rapport à chacune des variables, c'est-à-dire $\mathcal{P}_p(E^N) = \{\mu^N \in \mathcal{P}(E^N), \langle \mu^N, \Gamma_p^N \rangle < +\infty\}$.

Définition 3.1. (Fonctionnelle extensive) Pour tout $N \in \mathbb{N}^*$, soit $F^N : \mathcal{P}(E^N) \rightarrow \mathbb{R} \cup \{+\infty\}$ une fonctionnelle telle que sur $\mathcal{P}_p(E^N)$, F^N soit convexe propre inférieurement semi-continue pour la topologie faible- $*$ sur $\mathcal{P}_p(E^N)$ (on a de plus une hypothèse sur les moments) et définissons F par $\forall \mu^N \in \mathcal{P}_p(E^N)$, $F(\mu^N) := \frac{1}{N} F^N(\mu^N)$.

Lorsque ces fonctionnelles sont telles que :

1. $\forall N, \forall \mu^N \in \mathcal{P}_{sym}(E^N), \forall \mu \in \mathcal{P}(E), \mu^{N,1} = \mu$ et $F^1(\mu) < +\infty \Rightarrow F^N(\mu^N) \geq N F^1(\mu)$ avec égalité ssi $\mu^N = \mu^{\otimes N}$
2. $\forall \mu^{N+M} \in \mathcal{P}_{sym}(\mathcal{E}^{N+M}), F^{N+M}(\mu^{N+M}) < +\infty, F^{N+M}(\mu^{N+M}) \geq F^N(\mu^{N+M,N}) + F^M(\mu^{N+M,M})$

je dirais que F est une fonctionnelle extensive.

On se donne dans cette partie une fonctionnelle extensive F .

Définition 3.2. Soit $p > 0$. Pour tout $N \in \mathbb{N}^*$, soit $\mu^N \in \mathcal{P}_{sym,p}(E^N)$ et soit $\mu \in \mathcal{P}_p(E)$. On dit que (μ^N) est F -chaotique par rapport à μ lorsque $\mu^{N,1} \rightarrow \mu$ et $F(\mu^N) \rightarrow F(\mu) < +\infty$.

Proposition 3.1. Soit $(\mu^N) \in \prod_{N \in \mathbb{N}^*} \mathcal{P}_{sym}(E^N)$. Supposons qu'il existe $k \geq 1$ entier et $C \in \mathbb{R}$:

1. $\mu^{N,k} \rightharpoonup \nu^k$ où $\nu^k \in \mathcal{P}_{sym}(\mathcal{E}^k)$ et $F(\nu^k) < \infty$
2. $\forall N \geq k, \forall 1 \leq r \leq k-1, F^r(\mu^{N,r}) \geq C$

Alors $F(\nu^k) \leq \liminf F(\mu^N)$.

Preuve. Pour tout $N \geq k$, notons $q_N := \lfloor \frac{N}{k} \rfloor$ et $r_N := N - q_N k$. Par convention, pour ne pas distinguer le cas $r_N = 0$ dans les calculs, on considérera que $F^0(\mu^{N,0}) = 0$ pour tout N . Soit $N \geq k$.

$$\begin{aligned}
& F^{q_N k}(\mu^{N,q_N k}) + F^{r_N}(\mu^{N,r_N}) \leq F^N(\mu^N) && \text{(par sur-additivité de } F) \\
\text{or } q_N F^k(\mu^{N,k}) & \leq F^{q_N k}(\mu^{N,q_N k}) && \text{(par sur-additivité et récurrence descendante)} \\
& \text{et } C \leq F^{r_N}(\mu^{N,r_N}) && \text{(par hypothèse)} \\
\text{donc } \frac{q_N}{N} k F(\mu^{N,k}) + \frac{C}{N} & \leq F(\mu^N) && \text{(en divisant par } N)
\end{aligned}$$

F étant inférieurement semi-continue et comme $\frac{q_N}{N} k \xrightarrow{N \rightarrow \infty} 1$, on obtient le résultat en passant à la \liminf .

Lemme 3.1. (*F-chaos et chaos faible*) Soit (μ^N) F -chaotique pour une mesure de probabilité μ . Supposons de plus que $(\mu^{N,k})$ soit tendue pour un certain $k \geq 2$. Alors (μ^N) est faiblement chaotique par rapport à μ .

Preuve. Comme notre suite $(\mu^{N,k})$ est tendue, par le théorème de Prohorov, il suffit de montrer que la seules valeurs d'adhérence de cette suite pour la topologie faible-* est $\mu^{\otimes k}$.

Soit $(\mu^{N',k})$ une extraite de $(\mu^{N,k})$ convergeant vers $\nu^k \in \mathcal{P}_{sym}(\mathcal{E}^k)$. Par la proposition précédente (en l'adaptant légèrement), $F(\nu^k) \leq \liminf F(\mu^{N'}) = F(\mu)$ par F -chaos. Or $(\mu^{N',k})^1 = \mu^{N',1} \rightharpoonup \mu$ par F -chaos donc $\nu^{k,1} = \mu$ et F étant extensive, $F(\nu^k) \geq F(\mu)$. Finalement, $\nu^{k,1} = \mu$ et $F(\nu^k) = F(\mu) < \infty$ donc par extensivité de F $\nu^k = \mu^{\otimes k}$. On obtient donc que $\mu^{N,k} \rightharpoonup \mu^{\otimes k}$. Comme $k \geq 2$, $\mu^{N,2} \rightharpoonup \mu^{\otimes 2}$ et on conclut avec la deuxième formulation du chaos faible (Chaos 2).

Remarque : On pourrait définir cette forme de chaos pour E espace polonais général.

3.2 Information de Fisher et entropie

Dans l'article, les deux fonctionnelles étudiées sont l'entropie et l'information de Fisher. Dans cette partie, je vais les définir et énoncer certaines propriétés qu'elles vérifient. Pour plus de précision on pourra se référer au passage p28-34 de On Kac's chaos and related problems.

3.2.1 L'entropie H_d

L'entropie est une notion que l'on retrouve dans de nombreux domaines, notamment en statistiques, théorie de l'information, physique statistique et thermodynamique. D'une certaine façon, cette grandeur permet d'estimer la complexité d'un signal ou d'une répartition par rapport à une donnée de référence.

Notons $h(s) := s \log(s) - s + 1$ et $h^*(t) := e^t - 1$ pour tout $s \geq 0$ et t réel. On notera λ^m la mesure de Lebesgue sur \mathbb{R}^m pour tout $m \in \mathbb{N}^*$. h est une fonction convexe, continue et positive, ne s'annulant qu'en 1.

Définition 3.3. (*Entropie H_d^N*) Pour tout $\mu^N \in \mathcal{P}(E^N)$, notons

$$H_d^N(\mu^N) := \begin{cases} \int f^N \log(f^N) & \text{si } d\mu^N = f^N d\lambda^{Nd} \text{ et } f^N \log(f^N) \in L^1(E^N) \\ +\infty & \text{sinon} \end{cases}$$

Définition 3.4. (Entropie sur $\mathcal{P}_p(\mathcal{E}^N)$) Pour tout $p > 0$ et $N \geq 1$ entier, on définit la fonctionnelle convexe $H_{d,p}^N := \begin{cases} \mathcal{P}_p(E^N) & \rightarrow \mathbb{R} \cup \{+\infty\} \\ \mu^N & \mapsto \sup_{\varphi^N \in \mathcal{C}_b(E^N)} (\langle \mu^N, \varphi^N \rangle - \int h^*(\varphi^N) \Gamma_p^N) + \langle \mu^N, \Gamma_p^N \rangle \end{cases}$.

On admet le lemme suivant.

Lemme 3.2. • Pour tout $N \geq 1$, les $H_{d,p}^N$ coïncident sur leurs domaines communs et coïncident avec H_d^N . On peut donc définir H_d^N de façon univoque.

• L'écriture intégrale de H_d^N s'étend en fait à $\mathcal{P}_p(E^N) \cap L^1(E^N)$:

$$H_d^N(\mu^N) = \underbrace{\int h\left(\frac{f^N}{\Gamma_p^N}\right) \Gamma_p^N + \langle \mu^N, \log(\Gamma_p^N) \rangle}_{0 \leq \text{comme } h \text{ est positive}} = \int f^N \log(f^N) \text{ (qui est donc bien défini dans } \mathbb{R} \cup \{+\infty\})$$

- $H_d^N : \mathcal{P}_p(E^N) \rightarrow \mathbb{R} \cup \{+\infty\}$ est inférieurement continue dans le sens suivant : si $\mu_n^N \rightarrow \mu^N$ et qu'il existe $m > p$ tel que $(\langle \mu_n^N, \sum |z_i|^m \rangle)_n$ est bornée, alors $\liminf_n H_d^N(\mu_n^N) \geq H_d^N(\mu^N)$.
- $\forall \mu^N \in \mathcal{P}_p(E^N)$, $H_d^N(\mu^N) \geq N \log(g_p) - \langle \mu^N, \sum_i |z_i|^p \rangle$ (ce qui se voit facilement avec la formulation intégrale dans le cas où μ^N admet une densité et reste vrai sinon, le terme de gauche étant infini et terme de droite étant fini)

Remarque : Dans la suite, on ne se servira pas trop de cette définition par dualité, mais on utilisera l'expression sous forme intégrale du deuxième point.

Définition 3.5. (Entropie renormalisée) On définit alors la fonction H_d renormalisée "suivant le nombre de particules N ". Pour tout $N \in \mathbb{N}^*$, on note $H_d(\mu^N) := \frac{H_d^N(\mu^N)}{N}$.

Montrons désormais que H_d est une fonctionnelle extensive. Il nous reste à faire les points 1 et 2, l'aspect inférieurement semi-continu étant ici admis.

Proposition 3.2. Soit $N \geq 1$ entier. Soit $\mu \in \mathcal{P}(E)$ admettant f comme densité par rapport à la mesure de Lebesgue et ayant une entropie finie. Alors $H_d(\mu^{\otimes N}) = H_d(\mu)$.

Preuve. Si μ admet une entropie finie, il en va de même de μ^N (en effectuant le calcul précédent un calcul similaire, qui permet dans un premier temps de montrer l'intégrabilité de $f^N \log(f^N)$).

$$\begin{aligned} H_d(\mu) &= \int_E f \log(f) = \int_{E^N} f^{\otimes N}(Z^N) \log(f)(z_i) \\ &= \frac{1}{N} \sum_i \int_{E^N} f^{\otimes N}(Z^N) \log(f)(z_i) = \frac{1}{N} \int_{E^N} f^{\otimes N}(Z^N) \sum_i \log(f)(z_i) \\ &= \frac{1}{N} \int_{E^N} f^{\otimes N} \log(f^{\otimes N}) = H_d(\mu^{\otimes N}) \end{aligned}$$

Proposition 3.3. (Positivité de l'entropie relative) Soient f^N et g^N deux densités de probabilités sur E^N . Notons respectivement μ^N et ν^N les mesures de probabilité qui leur sont associées. Notons $H_d^N(\mu^N | \nu^N) := \int f^N \log(\frac{f^N}{g^N})$. Cette quantité est toujours bien définie et est positive. De plus, $H_d^N(\mu^N | \nu^N) = 0$ ssi $\mu^N = \nu^N$.

Preuve. Plus haut, j'ai défini la fonction $h(s) = s \log(s) - s + 1$. Cette fonction est continue et positive, ne s'annulant que en 1.

$$h\left(\frac{f^N}{g^N}\right) g^N = f^N \log\left(\frac{f^N}{g^N}\right) - \underbrace{f^N + g^N}_{\text{intégrable et d'intégrale nulle}}$$

$H_d^N(\mu^N | \nu^N)$ est donc bien définie et vaut $\int h(\frac{f^N}{g^N})g^N$. Cette quantité est bien positive. Supposons qu'elle soit nulle. Alors $h(\frac{f^N}{g^N})g^N$ est nulle presque partout pour la mesure de Lebesgue. En particulier, le sous-ensemble de $\{g^N > 0\}$ sur lequel $h(\frac{f^N}{g^N}) > 0$ est de mesure de Lebesgue nulle, donc sur cet ensemble, $\frac{f^N}{g^N} = 1$ presque partout, donc $f^N = g^N$ sur $\{g^N > 0\}$. On a alors $\mu^N(\{g^N > 0\}) = 1$ et donc $f^N = 0$ presque partout en dehors de $\{g^N > 0\}$, donc $f^N = g^N$ presque partout et $\mu^N = \nu^N$. L'implication réciproque s'obtient facilement.

Corollaire. H_d est extensive.

Preuve. 1. Soit $\mu^{N,1}$ telle que $\mu^{N,1} = f\lambda$ admet une entropie finie. Le cas où $H_d(\mu^N)$ est infini est trivial. Supposons désormais que $H_d(\mu^N)$ est fini, et notons alors f^N la densité de μ^N par rapport à la mesure de Lebesgue. Par symétrie de f^N et caractérisation des marginales

$$\begin{aligned} H_d(f^{\otimes N}) &= \frac{1}{N} \sum_i \int f^N(Z) \log(f(z_i)) \\ &= \frac{1}{N} \int f^N(Z) \log(f^{\otimes N}(Z)) \\ &= \frac{1}{N} \int f^N(Z) (-\log(\frac{f^N(Z)}{f^{\otimes N}(Z)}) + \log(f^N)) \\ (\text{la somme étant bien définie}) &= \frac{1}{N} \int f^N(Z) (-\log(\frac{f^N(Z)}{f^{\otimes N}(Z)}) + \log(f^N)) \\ &= -\frac{1}{N} H_d^N(\mu^N | ((\mu^{N,1})^{\otimes N})) + H_d(\mu^N) \\ &(\text{par le lemme précédent}) \leq H_d(\mu^N) \end{aligned}$$

Les deux entropies sont égales ssi $H_d^N(\mu^N | ((\mu^{N,1})^{\otimes N})) = 0$, donc avec le lemme précédent ssi $\mu^N = f^{\otimes N} d\lambda^{Nd}$.

2. Le raisonnement est similaire à celui effectué au point précédent, mais cette fois on s'intéresse à $H_d^{N+M}(\mu^{N+M} | (\mu^{N+M,N} \otimes \mu^{N+M,M}))$.

3.2.2 L'information de Fisher I_d

Je vais désormais définir l'information de Fisher I_d . J'énoncerai les résultats sans démonstration.

Définition 3.6. (Information de Fisher) Soit $N \geq 1$ entier. Pour tout $\mu^N \in \mathcal{P}(E^N)$, notons $I_d^N(\mu^N) := \begin{cases} \int f^N | \nabla \log(f^N) |^2 & \text{si } d\mu^N = f^N d\lambda^{Nd} \text{ et } f^N \in W^{1,1}(E^N) \\ = +\infty & \text{sinon} \end{cases}$. On notera alors $I_d(\mu^N) = \frac{1}{N} I^N(\mu^N)$.

On admet la proposition suivante (on peut se référer aux lemmes 3.6 et 3.7 de l'article de Mishler et Hauray).

Proposition 3.4.

- I_d^N est propre, convexe et inférieurement semi-continue sur $\mathcal{P}(E^N)$.
- I_d est positive.
- I_d est une fonctionnelle extensive.

Remarque : I_d vérifie des propriétés plus fortes qu'une fonctionnelle extensive.

Définition 3.7. (Information de Fisher relative) Soient f^N et g^N deux densités de probabilité sur E^N , de mesures respectives μ^N et ν^N . Notons $I_d^N(\mu^N | \nu^N) := \int | \nabla \log(\frac{f^N}{g^N}) |^2 g^N$ si $\mu^N \ll \nu^N$ et $+\infty$ sinon.

3.3 L'inégalité HWI

Ce résultat a été énoncé et démontré dans le papier d'Otto et Villani. Le théorème permet de relier dans certains cas l'entropie relative, l'information de Fisher relative et W_2^m la distance de Wasserstein d'ordre 2 pour la distance euclidienne de deux mesures de probabilité sur \mathbb{R}^m avec $m \geq 1$.

Théorème 3.2. (Inégalité relative HWI sur \mathbb{R}^m) Soit $m \geq 1$. Soient μ_0 et μ_1 dans $\mathcal{P}(\mathbb{R}^m)$ absolument continues par rapport à la mesure de Lebesgue. Soit $\psi \in \mathcal{C}^2(\mathbb{R}^m)$ telle qu'il existe $K \in \mathbb{R}$ tel que $\nabla^2 \psi \geq K \text{Id}_m$ (dans le sens $\nabla^2 \psi - K \text{Id}_m$ est une matrice symétrique positive), que $e^{-\psi}$ soit une densité de probabilité admettant un moment d'ordre 2. Supposons que les deux mesures admettent un moment d'ordre deux.

Alors $H^m(\mu_0 | e^{-\psi}) - H^m(\mu_1 | e^{-\psi}) \leq \sqrt{I^m(\mu_0 | e^{-\psi})} W_2^m(\mu_0, \mu_1) - \frac{K}{2} W_2^m(\mu_0, \mu_1)^2$ où les grandeurs définies correspondent à celles définies plus haut pour " $d=1$ ".

De cette inégalité relative on peut déduire une inégalité avec les grandeurs absolues et renormalisées, qui nous servira à montrer la deuxième implication du théorème de la section.

Théorème 3.3. (Inégalité HWI absolue renormalisée sur $(\mathbb{R}^d)^N$) Soient μ_0^N et μ_1^N des probabilités dans $\mathcal{P}((\mathbb{R}^d)^N)$, absolument continues par rapport à la mesure de Lebesgue et admettant toutes deux un moment d'ordre 2. Alors

$$H_d(\mu_0^N) - H_d(\mu_1^N) \leq \frac{1}{\sqrt{N}} \sqrt{I_d(\mu_0^N)} W_2^{Nd}(\mu_0^N, \mu_1^N)$$

Montrons qu'on peut obtenir le théorème 3.3 avec le théorème 3.2.

Preuve. Soit $N \geq 1$. Pour tout $\lambda > 0$, notons $\gamma_\lambda^N := (v_i)_{i=1}^d \in (\mathbb{R}^d)^N \mapsto \frac{1}{(2\pi\lambda)^{\frac{Nd}{2}}} e^{-\frac{\sum |v_i|^2}{2\lambda}}$. Pour tout λ , γ_λ^N est une densité de probabilité sur $(\mathbb{R}^d)^N$ vérifiant des hypothèses requise par " $e^{-\psi}$ " dans le théorème 3.2 pour " $m = Nd$ " et " $K = 0$ ". En appliquant le théorème 3.2 et en divisant par N , on obtient

$$H_d(\mu_0^N | \gamma_\lambda^N) - H_d(\mu_1^N | \gamma_\lambda^N) \leq \frac{1}{\sqrt{N}} \sqrt{I_d(\mu_0^N | \gamma_\lambda^N)} W_2^{Nd}(\mu_0^N, \mu_1^N)$$

or pour $i=0$ ou 1 ,

$$\begin{aligned} H_d(\mu_i^N | \gamma_\lambda^N) &= H_d(\mu_i^N) - \int \log((\gamma_\lambda^N) d\mu_i^N) = H_d(\mu_i^N) + \frac{\langle \mu_i^N, |V|^2 \rangle}{2\lambda N} + \frac{d}{2} \log(2\pi\lambda) \\ I_d(\mu_0^N | \gamma_\lambda^N) &= I_d(\mu_0^N) - \frac{2d}{\lambda} + \frac{\langle \mu_0^N, |V|^2 \rangle}{\lambda^2 N} \end{aligned}$$

donc

$$\begin{aligned} H_d(\mu_0^N | \gamma_\lambda^N) - H_d(\mu_1^N | \gamma_\lambda^N) &= H_d(\mu_0^N) - H_d(\mu_1^N) + o_{\lambda \rightarrow \infty}(1) \\ I_d(\mu_0^N | \gamma_\lambda^N) &= I_d(\mu_0^N) + o_{\lambda \rightarrow \infty}(1) \end{aligned}$$

En faisant tendre λ vers $+\infty$ dans l'inégalité, on obtient l'inégalité désirée.

Dans ce qui suit on se donne ψ comme dans l'énoncé de l'inégalité.

3.3.1 Quelques notions sur le réarrangement convexe de probabilités

Dans cette partie l'espace étant fixé, on omettra les indices pour plus de lisibilité. La démonstration repose sur deux idées. La première est d'écrire μ_1 comme une mesure image de μ_0 par une fonction Φ de \mathbb{R}^m dans \mathbb{R}^m et d'ainsi ré-écrire les grandeurs qui nous intéressent en fonction de Φ . La deuxième est de s'intéresser au comportement de l'entropie lors du "transport de μ_0 vers μ_1 ".

J'énonce ici les théorèmes nécessaires à cette démarche.

Théorème 3.4. (Réarrangement convexe, Mc Cann-Brenier) Soient μ et $\nu \ll \lambda$ dans $\mathcal{P}(\mathbb{R}^m)$ admettant un moment d'ordre 2 toutes deux.

Alors il existe une fonction ϕ convexe, propre, inférieurement semi-continue telle que $\mu = \nabla\phi\#\nu$, le terme de droite étant la mesure image de ν par $\nabla\phi$. De plus, $W_2(\mu, \nu) = \sqrt{\int |\nabla\phi|^2 d\nu}$ où $\nabla\phi := \nabla\phi - id$. $\nabla\phi$ est de plus uniquement défini ν presque partout.

Ce théorème est énoncé pour la première fois par Brenier dans des cas particuliers, puis étendu par Mc-Cann. On utilise ici une version légèrement simplifiée, énoncée dans les notes de cours données en référence.

Définition 3.8. (Différentielle au sens d'Aleksandrov) Soit ϕ une fonction convexe de $\Omega \subset \mathbb{R}^m$ ouvert convexe dans \mathbb{R} . Soit $x \in \Omega$. On dit que ϕ est différentiable au sens d'Aleksandrov en x lorsque ϕ est différentiable en x et qu'il existe $\nabla_A^2\phi \in M_m(\mathbb{R})$ telle que

$$\phi(h+x) = \phi(x) + \langle \nabla\phi(x), h \rangle + \frac{1}{2} \langle \nabla_A^2\phi(x)h, h \rangle + o_{h \rightarrow 0}(|h|^2)$$

On trouve ce théorème dans Measure theory and fine properties of functions et dans A convexity principle for interacting gases.

Théorème 3.5. (d'Aleksandrov) Soit ϕ une fonction convexe de $\Omega \subset \mathbb{R}^d$ ouvert convexe dans \mathbb{R} . Soit $x \in \Omega$. Alors ϕ admet une différentielle au sens d'Aleksandrov Lebesgue-presque partout sur Ω . $\nabla_A^2\phi$ est de plus symétrique et positive.

Pour ϕ une fonction convexe (finie) de partant d'un domaine $\Omega \subset \mathbb{R}^m$, on peut également définir (ϕ étant alors continue) pour tout $1 \leq i, j \leq d$ les distributions $\partial_i\partial_j\phi : \mathcal{C}_c^2(\Omega) \rightarrow \mathbb{R}$. Il s'avère que ces distributions sont telles que $\forall i, j, \partial_i\partial_j\phi = \partial_j\partial_i\phi$ et que pour tout $i, \partial_i\partial_i\phi$ est une distribution positive. On définit alors $\Delta\phi := \sum_i \partial_i\partial_i$. Ces distributions peuvent en fait s'écrire comme mesures de Radon signées, c'est-à-dire, il existe des mesures boréliennes sur Ω localement finies $D^{i,j}$ telles que appliquer $\partial_i\partial_j\phi$ à une fonction test revienne à l'intégrer contre la mesure $D^{i,j}$ (voir p239 de Measure theory and fine properties of functions). Par le théorème de Radon-Nikodym, on peut écrire $D^{i,i} = F^{i,i}\lambda^d + D_s^{i,i}$ où $F^{i,i}$ est $L^1_{loc}(\Omega)$ et positive et $D^{i,i}$ est étrangère à la mesure de Lebesgue et est une mesure positive.

La proposition suivante est utilisée dans le papier d'Otto et Villani sans être véritablement énoncée.

Proposition 3.5. (Comparaison de $\Delta\phi$ et $\Delta^A\phi$) Soit $\phi : \Omega \rightarrow \mathbb{R}$ convexe. On définit $\Delta^A\phi := \sum_i \partial_{i,i}^A\phi$. On admet que $\Delta^A\phi$ permet de définir une distribution sur Ω . Alors $\Delta^A\phi \leq \Delta\phi$ au sens des distributions.

Preuve. (ou plutôt tentative de justification) Je n'ai pas trouvé de référence expliquant clairement ce résultat, mais, si l'on suppose que les $F^{i,i}$ correspondent presque partout au sens de la mesure de Lebesgue aux $\partial_{i,i}^A\phi$ (ce qui au vue du livre d'Evans et Gariépy, ne semble pas abérant), on obtient que

$$D^{i,i} = \underbrace{\partial_{i,i}^A\phi \lambda^d + D_s^{i,i}}_{\text{mesure positive}}$$

Théorème 3.6. (Changement de variable, Mc Cann) Soient μ et ν des mesures de probabilité boréliennes absolument continues par rapport à λ , de densités respectives f et g et ϕ convexe telle que $\mu = \nabla\phi\#\nu$.

Notons $X \in \mathcal{B}(\mathbb{R}^m)$ le domaine où $\nabla_A^2\phi$ est définie et inversible. Alors $\nu(X) = 1$. De plus, si $A : \mathbb{R}^+ \rightarrow \mathbb{R}$ est mesurable et telle que $A(0)=0$ alors

$$\int A(f)d\lambda = \int_X A\left(\frac{g}{\det(\nabla_A^2\phi)}\right) \det(\nabla_A^2\phi)d\lambda \in \mathbb{R} \cup \{\pm\infty\}$$

Remarque : On se servira du théorème suivant pour réécrire l'entropie en l'appliquant à " $A(x) = x \log(x)$ ".

3.3.2 Preuve de l'inégalité

Le raisonnement de fond de la démonstration n'est pas difficile à comprendre, mais la justification des différentes étapes est assez technique et requière l'utilisation de nombreux résultats de réarrangement convexe et d'analyse convexe. On commence par effectuer la démonstration dans le cas où μ_0 et μ_1 admettent des densités C^∞ à supports compacts. Par un argument d'approximation détaillé dans Topics in Optimal Transportation p 299, cela sera suffisant pour avoir l'inégalité générale.

Soient $d\mu_0 = f_0 d\lambda$ et $d\mu_1 = f_1 d\lambda$ de telles probabilités sur \mathbb{R}^m . Notons $\nabla\phi = id + \nabla\varphi$ le réarrangement convexe de μ_0 vers μ_1 . $\mu_1 = \nabla\phi\#\mu_0$.

Pour tout t dans $[0,1]$, notons $\mu_t := (id + t\nabla\varphi)\#\mu_0$, qui est alors une mesure de probabilité sur \mathbb{R}^d . μ_0 et μ_1 étant à supports bornés, $\nabla\varphi$ est borné sur le support de μ_0 et donc les μ_t admettent un moment d'ordre 2. Pour tout $0 \leq t \leq 1$, $id + t\nabla\varphi = (1-t)id + t\nabla\phi$ est combinaison linéaire positive de deux gradients de fonctions convexes donc est gradient d'une fonction convexe. Par la proposition 1.3 A convexity principle for interacting gases, les μ_t admettent toutes une densité f_t par rapport à la mesure de Lebesgue. Par l'unicité μ_0 presque partout du réarrangement convexe et par définition des μ_t , $\forall t \in [0,1], W_2(\mu_0, \mu_t) = \sqrt{\int |t\nabla\varphi|^2 d\mu_0} = tW_2(\mu_0, \mu_1)$.

Intéressons-nous désormais à l'entropie. Je vais étudier $t \mapsto H(\mu_t | e^{-\psi})$ et montrer une inégalité sur l'accroissement de cette fonction entre 0 et 1, qui donnera ensuite l'inégalité recherchée. Soit $t \in [0,1]$.

Par le changement de variable de Mc-Cann, $id + t\nabla\varphi$ étant gradient d'une fonction convexe sur le support de μ_0 , $\int \log(f_t)f_t = \int [\frac{f_0}{\det(I_d + t\nabla^2_A\varphi)} \log(\frac{f_0}{\det(I_d + t\nabla^2_A\varphi)})] \det(I_d + t\nabla^2_A\varphi) = \int f_0 \log(f_0) - f_0 \log(\det(I_d + t\nabla^2_A\varphi))$. Par continuité de f_0 , son support étant de plus compact, $f_0 \log(f_0)$ est intégrable. On peut donc sortir la somme de l'intégrale. De la même façon, $\int \psi(id + t\nabla\varphi)f_0 = \int \psi f_t$ étant un réel bien défini (avec les hypothèses sur ψ et f_0) on peut écrire $H(\mu_t | e^{-\psi}) = \int \log(\frac{f_t}{e^{-\psi}})f_t = \int f_t \log(f_t) + \int \psi(id + t\nabla\varphi)f_0$. On obtient :

$$H(\mu_t | e^{-\psi}) = \int f_0 \log(f_0) - \int f_0 \log(\det(I_d + t\nabla^2_A\varphi)) + \int \psi(id + t\nabla\varphi)f_0$$

Notons $\Psi := t \mapsto \int \psi(id + t\nabla\varphi)f_0$ et $D := t \mapsto \int f_0 \log(\det(I_d + t\nabla^2_A\varphi))$.

Etude de Ψ : Comme dit plus haut, $\nabla\varphi$ est bornée sur le support de μ_0 . Par application du théorème de convergence dominée, ψ étant C^2 , Ψ est C^2 sur $[0,1]$ et

$$\begin{aligned} \Psi'(t) &= \int \langle \nabla\psi(id + t\nabla\varphi), \nabla\varphi \rangle f_0 \\ \Psi''(t) &= \int \langle \nabla^2\psi(id + t\nabla\varphi)\nabla\varphi, \nabla\varphi \rangle f_0 \end{aligned}$$

Or par hypothèse, $\forall z \in \mathbb{R}^m, \langle \nabla^2\psi z, z \rangle \geq K |z|^2$, donc $\Psi''(t) \geq \int K |\nabla\varphi|^2 f_0 = KW_2(\mu_0, \mu_1)^2$ et donc

$$\Psi''(t) \geq KW_2(\mu_0, \mu_1)$$

$t \mapsto \Psi(t) - \frac{t^2}{2}KW_2(\mu_0, \mu_1)$ est donc C^2 et convexe car de dérivée seconde positive et est donc au-dessus de ses tangentes, donc $\Psi(0) + 1 \times (\Psi'(0) - 2 \times 0 \frac{K}{2}W_2(\mu_0, \mu_1)^2) \leq \Psi(1) - \frac{K}{2}W_2(\mu_0, \mu_1)^2$ soit

$$\Psi(0) - \Psi(1) \leq \int -\langle \nabla\psi, \nabla\varphi \rangle f_0 - \frac{K}{2}W_2(\mu_0, \mu_1)^2$$

Etude de D : D est concave et finie.

En effet, pour tout z dans l'ensemble de définition de $\nabla_A^2 \varphi$, notons $d_z(t) := \log(\det(Id_m + t\nabla_A^2 \varphi(z)))$. En choisissant une base orthornormée $e(z)$ dans laquelle la matrice de $\nabla_A^2 \varphi(z)$ est diagonale (cette endomorphisme étant symétrique), on obtient $d_z(t) = \sum_{i=1}^d \log(1 + t\partial_{e_i(z), e_i(z)}^A \varphi(z))$. Comme \log est concave, chaque d_z est concave sur $[0, 1]$, ce qui nous fournit la concavité de D par positivité de l'intégration et f_0 étant également positive. $H(\mu_{0,1} | e^{-\psi})$ étant finies et avec l'expression plus haut, le domaine où une fonction concave est finie étant convexe, le domaine où D est finie est un convexe de $[0, 1]$ contenant 0 et 1, c'est donc l'intervalle tout entier. Etudions le comportement de D au voisinage de 0. Soit z dans le domaine de définition de $\nabla_A^2 \varphi$. d_z est concave, donc

$$\forall 0 \leq h \leq h' \leq 1, d_z(1) - d_z(0) \leq \frac{d_z(h') - d_z(0)}{h'} \leq \frac{d_z(h) - d_z(0)}{h} \leq \lim_{u \rightarrow 0^+} \frac{d_z(u) - d_z(0)}{u}$$

or pour t suffisamment petit, $d'_z(t) = \sum_i \frac{\partial_{e_i(z), e_i(z)}^A \varphi(z)}{1 + t\partial_{e_i(z), e_i(z)}^A \varphi(z)}$

donc $d'_z(0) = \sum_i \partial_{e_i(z), e_i(z)}^A \varphi(z) = \Delta^A \varphi(z)$. $v \mapsto d_z(1) - d_z(0)$ étant μ_0 intégrable, on peut appliquer le théorème de convergence monotone pour obtenir que $D'(0)$ existe et :

$$D'(0) = \lim_{h \rightarrow 0^+} \frac{D(h) - D(0)}{h} = \int d'_z(0) d\mu_0 = \int \Delta^A \varphi f_0$$

or en terme de distribution, $\Delta^A \varphi \leq \Delta \varphi$. f_0 étant positive et C^∞ à support compact, donc $D'(0) \leq \int \Delta \varphi = - \int \langle \nabla \varphi, \nabla f_0 \rangle$.

Par concavité de D, on obtient :

$$D(1) - D(0) \leq - \int \langle \nabla \varphi, \nabla f_0 \rangle$$

Recombinons désormais les deux inégalités obtenues sur D et Ψ pour conclure.

$$\begin{aligned} H(\mu_0 | e^{-\psi}) - H(\mu_1 | e^{-\psi}) &= (\Psi(0) - \Psi(1)) + (D(1) - D(0)) \\ &\leq - \left(\int \langle \nabla \psi, \nabla \varphi \rangle f_0 + \int \langle \nabla, \nabla \varphi \rangle \right) - \frac{K}{2} W_2(\mu_0, \mu_1)^2 \\ &\leq \int \langle -\nabla(\log(\frac{f_0}{e^{-\psi}})), \nabla \varphi \rangle d\mu_0 - \frac{K}{2} W_2(\mu_0, \mu_1)^2 \\ (\text{par Cauchy-Schwarz}) &\leq \sqrt{\int |\nabla \log(\frac{f_0}{e^{-\psi}})|^2 d\mu_0} \sqrt{\int |\nabla \varphi|^2 d\mu_0} - \frac{K}{2} W_2(\mu_0, \mu_1)^2 \\ &= \sqrt{I(\mu_0 | e^{-\psi})} W_2(\mu_0, \mu_1) - \frac{K}{2} W_2(\mu_0, \mu_1)^2 \end{aligned}$$

3.4 Preuve

Pour tout $N \in \mathbb{N}^*$, soit $\mu^N \in \mathcal{P}_{sym}(E^N)$. Soit $\mu \in \mathcal{P}(E)$. On suppose de plus que $(\langle \mu^{N,1}, |z|^p \rangle)$ la suite des moments d'ordre p des premières marginales est bornée pour un certain $p > 2$. Pour tout N , notons $m^N := \langle \mu^N, |z|^p \rangle$ et choisissons $A > 0$ tel que $\forall N \in \mathbb{N}^*, 0 \leq m^N \leq A$.

Proposition 3.6. $(\mu^{N,2})_N$ est tendue.

Preuve. Pour tout $r > 0$ et tout N ,

$$\begin{aligned} \mu^{N,2}(\{|z_1|^p + |z_2|^p > r\}) &\leq \frac{1}{r} < \mu^{N,2}, \{|z_1|^p + |z_2|^p > r\} && \text{(par Markov)} \\ &\leq \frac{2m^N}{r} && \text{(par symétrie de } \mu^N \text{)} \\ &\leq \frac{2A}{r} && \text{(par hypothèse)} \end{aligned}$$

3.4.1 (IC) \Rightarrow (IBC)

Pour le premier point, si (μ^N) est Fisher-chaotique par rapport à μ , alors $(I_d(\mu^N))$ est convergente donc bornée. I_d étant positive, la proposition 3.6 et le lemme 3.1 donnent le chaos faible.

3.4.2 (IBC) \Rightarrow (KC)

Supposons que (μ^N) et μ vérifient (IBC). Soit $N \geq 1$ entier. En appliquant le théorème 3.3 pour " $\mu_0^N = \mu^N$ " et " $\mu_1^N = \mu^{\otimes N}$ " et vice versa et avec le caractère extensif de H_d et I_d ,

$$\begin{aligned} |H_d(\mu^N) - H_d(\mu)| &\leq (\sqrt{I_d(\mu^N)} + \sqrt{I_d(\mu)}) \frac{W_2^{Nd}(\mu^N, \mu^{\otimes N})}{\sqrt{N}} \\ &\leq M \frac{W_2^{Nd}(\mu^N, \mu^{\otimes N})}{\sqrt{N}} \end{aligned}$$

or, la suite des moments étant bornée (voir p 36 Mischler-Hauray) $\frac{1}{\sqrt{N}} W_2^{Nd}(\mu^N, \mu^{\otimes N}) \leq C(\frac{1}{\sqrt{N}} W_1^{Nd}(\mu^N, \mu^{\otimes N}))^{\frac{1}{2} - \frac{1}{p}}$ pour un certain $C > 0$ indépendant de N et par un résultat de la partie "Kac's chaos" de l'article de Mischler et Hauray, $\frac{1}{\sqrt{N}} W_1^{Nd}(\mu^N, \mu^{\otimes N}) \xrightarrow{N \rightarrow \infty} 0$ si (μ^N) est faiblement chaotique par rapport à μ , donc $|H_d(\mu^N) - H_d(\mu)| \rightarrow 0$.

3.4.3 (HC) \Rightarrow (KC)

Avec la proposition 3.6 et l'hypothèse sur les moments d'ordre p bornés, il reste à montrer que (μ^N) vérifie la deuxième hypothèse du lemme 3.1, soit que $(H_d^2(\mu^{N,2}))$ est minorée. Or par le lemme 3.2 pour tout N , $H_d^2(\mu^{N,2}) \geq 2 \log(g_p) - 2m^N \geq 2 \log(g_p) - 2A$ ce qui conclut.

4 Bibliographie

Bibliographie générale

- A. Diesz et L-P. Chaintron, Propagation of chaos : a review of models, methods and applications. I. Models and methods
- S. N. Ethier and T. G. Kurtz, Markov processes : characterization and convergence, Wiley series in probability and mathematical statistics, Wiley, New York, 1986.
- M. Hauray and S. Mischler, On Kac's chaos and related problems, J. Funct. Anal., 266 (2014), 6055–6157.
- F. Otto and C. Villani, Generalization of an Inequality by Talagrand and Links with the Logarithmic Sobolev Inequality, J. Funct. Anal., 173 (2000), 361–400
- C. Villani, Topics in Optimal Transportation, no. 58 in Graduate Studies in Mathematics, American Mathematical Society, 2003.

- McCann, R. J. A convexity principle for interacting gases. *Adv.Math.*128, 1 (1997), 153–179.
- Mc Cann, R. J. Existence and uniqueness of monotone measure-preserving maps. *Duke-Math.J.*80, 2 (1995), 309–323.
- Nathael Gozlan, Paul-Marie Samson, Pierre-André Zitt, Note de cours sur le transport optimal, 21 octobre 2022
- Evans, L.C., and Gariepy, R.F. *Measure theory and fine properties of functions. Studies in Advanced Mathematics*, CRC Press (1992).

Bibliographie complémentaire

Sur l'équation de Boltzmann

- C. Villani, A Review of Mathematical Topics in Collisional Kinetic Theory, in *Handbook of Mathematical Fluid Dynamics* (eds. S. Friedlander and D. Serre), vol. 1, Elsevier Science, 2002, 71–74, <https://linkinghub.elsevier.com/retrieve/pii/S1874579202800040>.
- C. Cercignani, R. Illner and M. Pulvirenti, *The Mathematical Theory of Dilute Gases*, no. 106 in *Applied Mathematical Sciences*, Springer-Verlag New York, 1994

Sur le modèle de Kac

- M. Kac, *Foundations of kinetic theory*, in *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability*, vol. 3, University of California Press Berkeley and Los Angeles, California, 1956, 171–197.
- E. Carlen, M. C. Carvalho, J. Le Roux, M. Loss and C. Villani, *Entropy and chaos in the Kac model*, *Kinet. Relat. Models*, 3 (2008), 85–122.