Petit guide de champ moyen

Angelo Rosello

1 Introduction : champ moyen déterministe

1.1 Interaction à champ moyen : du discret au continu

On s'intéresse à un système composé de N particules, notées $X^{i,N}$, $i \in \{1, \ldots, N\}$, à valeurs dans $\mathbb{R}^d$, dont l'évolution au cours du temps est dictée par le système de N équations différentielles :

$$\begin{cases} \dfrac{dX_t^{i,N}}{dt} = \dfrac{1}{N} \displaystyle\sum_{j=1}^{N} a(X_t^{i,N}, X_t^{j,N}), \\ X_0^{i,N} = x_0^{i,N}, \end{cases} \qquad i \in \{1, \ldots, N\}. \tag{1.1}$$

Pour une donnée initiale $(x_0^{i,N})_{1 \leq i \leq N}$ fixée, ce système d'équation admet une unique solution, globale en temps, pour peu que $a : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$ soit localement Lipschitzien et sous-linéaire (c'est-à-dire $|a(x, x')| \lesssim 1 + |x| + |x'|$). Considérant un intervalle de temps borné $[0, T]$, on introduit l'espace des trajectoires

$$\mathscr{C} = C([0, T]; \mathbb{R}^d) \text{ muni de } \|x\|_\infty = \sup_{t \in [0,T]} |x_t|,$$

si bien que $X^{i,N} \in \mathscr{C}$ pour tout $i \in \{1, \ldots, N\}$. Dans nombre d'applications concrètes, le coefficient a est issu d'un potentiel d'interaction : $a(x, x') = -\nabla V(x - x')$. On pourra penser à un potentiel gravitationnel ou électrique. Dans cette première section introductive, on supposera que a est Lipschitzien.

Le système d'équation (1.1) est dit "à **champ moyen**" car l'interaction entre les particules est symétrique : chaque particule interagit en fait avec la moyenne de toutes les autres, au travers de la **mesure empirique**

$$\mu^N = \frac{1}{N} \sum_{i=1}^{N} \delta_{X^{i,N}} \in \mathcal{P}(\mathscr{C}). \tag{1.2}$$

Remarque 1.1. Tout au long de ces notes, on choisira de voir μ^N (et toutes les mesures dépendant du temps !) comme une mesure de probabilité sur l'espace des trajectoires $\mathscr{C}$.
Ce choix permettra une certaine praticité (et, de fait, une certaine élégance) dans l'énoncé des résultats. Bien-sûr, pour $\mu \in \mathcal{P}(\mathscr{C})$ et $t \in [0, T]$, en définissant naturellement $\mu_t \in \mathcal{P}(\mathbb{R}^d)$ comme la mesure image de μ par l'évaluation $x \in \mathscr{C} \mapsto x_t \in \mathbb{R}^d$, on constate facilement que l'on dispose du plongement continu

$$\begin{array}{ccc} \mathcal{P}(\mathscr{C}) & \longrightarrow & C([0, T]; \mathcal{P}(\mathbb{R}^d)) \\ \mu & \longmapsto & (\mu_t)_{t \in [0,T]} \end{array}$$

où $\mathcal{P}(\mathscr{C})$ et $\mathcal{P}(\mathbb{R}^d)$ sont munis de la topologie de la convergence étroite. En ce sens, un résultat de convergence énoncé dans $\mathcal{P}(\mathscr{C})$ est "plus fort" que dans $C([0, T]; \mathcal{P}(\mathbb{R}^d))$.

En termes de mesures empiriques, le système (1.1) peut se ré-écrire

$$\begin{cases} \dfrac{dX_t^{i,N}}{dt} = A[\mu_t^N](X_t^{i,N}), \\ X_0^{i,N} = x_0^{i,N}, \end{cases} \qquad i \in \{1, \ldots, N\}. \tag{1.3}$$

où pour tout $\mu \in \mathcal{P}(\mathbb{R}^d)$, l'operateur $A[\mu]$ est défini comme la "convolution"

$$A[\mu](x) = \int_{\mathbb{R}^d} a(x, y) d\mu(y). \tag{1.4}$$

Au système de particule (1.3), on associe naturellement l'EDP

$$\partial_t \mu + \nabla \cdot (A[\mu]\mu) = 0, \tag{1.5}$$

où l'opérateur $\nabla \cdot$ désigne la divergence.

Un simple calcul montre en effet que pour N fixé, la mesure empirique μ^N est une solution faible (au sens des EDP) de (1.5) : pour tout fonction test $\psi \in C_c^\infty(\mathbb{R}^d)$,

$$\begin{aligned}\frac{d}{dt}\langle \psi, \mu_t^N \rangle = \frac{d}{dt}\left(\frac{1}{N}\sum_{i=1}^N \psi(X_t^{i,N})\right) &= \frac{1}{N}\sum_{i=1}^N \nabla\psi(X_t^{i,N}) \cdot A[\mu_t^N](X_t^{i,N}) \\ &= \langle \nabla\psi \cdot A[\mu_t^N], \mu_t^N \rangle = -\Big\langle \psi, \nabla \cdot (A[\mu_t^N]\mu_t^N) \Big\rangle.\end{aligned}$$

L'équation (1.5) est une **équation de conservation non-linéaire** (du fait du terme d'interaction $A[\mu]$), portant sur une mesure. Notons que dans le cas particulier $a(x, x') = a(x)$ (absence d'interaction) on a $A[\mu] = a$ et on est ramené à une équation linéaire. Dans une équation de conservation, la masse totale est conservée au cours du temps. On s'attend ainsi à ce que l'évolution d'une solution μ consiste simplement en un "transport" de la distribution initiale μ_0. Ceci motive la définition suivante.

Définition 1.1 (Solution de la forme transport)**.**
Soit $\mu \in \mathcal{P}(\mathscr{C})$ satisfaisant $\int_{\mathscr{C}} \|x\|_\infty d\mu(x) < \infty$. Le flot des **caractéristiques non-linéaires** $X^\mu = (X_t^\mu(x))_{x\in\mathbb{R}^d,\, t\in[0,T]}$ est défini de la façon suivante : pour tout $x \in \mathbb{R}^d$,

$$\begin{cases} \dfrac{dX_t^\mu(x)}{dt} = A[\mu_t](X_t^\mu), \quad t \in [0,T], \\ X_0^\mu(x) = x. \end{cases} \tag{1.6}$$

La mesure $\mu \in \mathcal{P}(\mathscr{C})$ est dite "**de la forme transport**" si elle satisfait le point fixe

$$\mu = (X^\mu)^*\mu_0,$$

où $(X^\mu)^*\mu_0$ désigne la mesure image de $\mu_0 \in \mathcal{P}(\mathbb{R}^d)$ par le flot $X^\mu : x \in \mathbb{R}^d \to X^\mu(x) \in \mathscr{C}$.
Autrement dit : pour tous $(t_1, \ldots, t_k) \in [0,T]^k$ et $\Psi : (\mathbb{R}^d)^k \to \mathbb{R}$ mesurable bornée,

$$\int_{x\in\mathscr{C}} \Psi(x_{t_1}, \ldots, x_{t_k}) d\mu(x) = \int_{x\in\mathbb{R}^d} \Psi\left(X_{t_1}^\mu(x), \ldots, X_{t_k}^\mu(x)\right) d\mu_0(x).$$

Remarque 1.2. La solution $X^\mu(x)$ de (1.6) est bien définie, unique et globale en temps sous l'hypothèse où a est Lipschitzien, grâce au moment de μ. En effet :

$$A[\mu_t](x) = \int_{R^d} a(x,y)\, d\mu_t(y) = \int_{y\in\mathscr{C}} a(x, y_t)\, d\mu(y)$$

et comme $|a(x,y)| \lesssim 1 + |x| + |y|$, on déduit facilement que $(t,x) \mapsto A[\mu_t](x)$ est continu par convergence dominée, avec une domination $C(1 + \|y\|_\infty)$. On voit ensuite que $A[\mu_t]$ est Lipschitzien (uniformément en t) puisque

$$|A[\mu_t](x) - A[\mu_t](x')| \leq \int |a(x,y) - a(x', y)|\, d\mu_t(y) \lesssim |x - x'|.$$

Remarque 1.3. Si le coefficient a est suffisamment régulier, il en va de même pour $A[\mu_t]$. Un résultat classique garantit alors que pour tout $t \in [0,T]$, le flot $x \mapsto X_t^\mu(x)$ définit un C^1-difféomorphisme. Il en résulte que si μ est de la forme transport et si la mesure initiale est à densité $\mu_0(dx) = f_0(x)dx$, alors pour tout $t \in [0,T]$, pour tout $\psi \in C_b(\mathbb{R}^d)$, par changement de variable,

$$\langle \psi, \mu_t \rangle = \int_{\mathbb{R}^d} \psi(X_t^\mu(x)) f_0(x) dx = \int_{\mathbb{R}^d} \psi(x) J_t^\mu(x)\, f_0\Big((X_t^\mu)^{-1}(x)\Big)\, dx = \int_{R^d} \psi(x) f_t(x) dx$$

de sorte que $\mu_t(dx) = f_t(x)dx$ avec $f_t(x) = J_t^\mu(x)\, f_0((X_t^\mu)^{-1}(x))$, où $J_t^\mu(x)$ désigne le déterminant Jacobien de $(X_t^\mu)^{-1}$ en x.

Remarque 1.4. Pour tout $N \in \mathbb{N}^*$ fixé, la mesure empirique $\mu^N \in \mathcal{P}(\mathscr{C})$ est de la forme transport ! L'équation (1.3) montre bien que $X^{i,N} = X^{\mu^N}(x_0^{i,N})$, de sorte que

$$\mu^N = \frac{1}{N}\sum_{i=1}^N \delta_{X^{i,N}} = \frac{1}{N}\sum_{i=1}^N \delta_{X^{\mu^N}(x_0^{i,N})} = (X^{\mu^N})^* \left(\frac{1}{N}\sum_{i=1}^N \delta_{x_0^{i,N}}\right) = (X^{\mu^N})^* \mu_0^N.$$

Proposition 1.1. Toute mesure $\mu \in \mathcal{P}(\mathscr{C})$ de la forme transport est solution (faible) de (1.5).

Preuve. Soit μ de la forme transport Pour toute fonction test $\psi \in C_c^\infty(\mathbb{R}^d)$ on a ainsi

$$\langle \psi, \mu_t \rangle = \int_{R^d} \psi(X_t^\mu(x)) d\mu_0(x).$$

On a vu que $|A[\mu_t](x)| \lesssim 1+|x|$. Avec (1.6), un lemme de Grönwall classique garantit alors l'estimation $|X_t^\mu(x)| \lesssim 1+|x|$ (pour $t \in [0,T]$). Il en résulte

$$\left|\frac{d}{dt}\psi(X_t^\mu(x))\right| = \left|\nabla\psi(X_t^\mu(x)) \cdot A[\mu_t](X_t^\mu(x))\right| \lesssim 1+|x|.$$

Puisque $\int_{\mathbb{R}^d} |x| d\mu_0(x) \leq \int_{x\in\mathscr{C}} \|x\|_\infty d\mu(x) < \infty$, on peut dériver sous l'intégrale pour obtenir

$$\frac{d}{dt}\langle \psi, \mu_t \rangle = \int_{\mathbb{R}^d} \nabla\psi(X_t^\mu(x)) \cdot A[\mu_t](X_t^\mu(x)) d\mu_0(x) = \int_{R^d} \nabla\psi(x) \cdot A[\mu_t](x) d\mu_t(x) = \langle \nabla\psi \cdot A[\mu_t], \mu_t \rangle.$$

□

L'enjeu des **limites de champ moyen** est d'établir un lien entre la description discrète donnée par le système de particules (1.1), et la description continue donnée par l'EDP : dans la limite où le nombre N de particules tend vers l'infini, on s'attend à ce que la mesure empirique μ^N converge vers une solution μ de (1.5). Il convient évidemment de préciser la notion de convergence considérée. L'espace $\mathcal{P}(\mathscr{C})$ est naturellement muni de la convergence étroite, qui peut être maîtrisée de différentes manières (par exemple avec la distance de Prokhorov). Ceci-étant, la convergence étroite (i.e faible !) toute seule n'est pas suffisante pour espérer passer à la limite dans l'équation non linéaire (1.5). On a vu de plus que la notion de mesure de la forme transport conduit naturellement à supposer l'existence d'un moment (d'ordre 1) pour μ. Ceci nous invite à considérer les **espaces de Wasserstein**.

Définition 1.2 (Espace de Wasserstein)**.**
Soit $(E, \|\cdot\|)$ un espace vectoriel normé séparable et complet. Pour tout $p \geq 1$, on définit

$$\mathcal{P}_p(E) = \left\{\mu \in \mathcal{P}(E) \;\middle|\; \int_E \|x\|^p d\mu(x) < \infty\right\}.$$

Cet ensemble est muni de la distance de Wasserstein, notée W_p et définie par

$$W_p[\mu,\nu] = \inf_{\pi\in\Pi(\mu,\nu)} \left(\int_{E\times E} \|x-y\|^p d\pi(x,y)\right)^{1/p},$$

où $\Pi(\mu,\nu) = \left\{\pi \in \mathcal{P}(E\times E) \;\middle|\; \int_E \pi(\cdot, dy) = \mu \text{ et } \int_E \pi(dx, \cdot) = \nu\right\}.$

Une mesure $\pi \in \Pi(\mu,\nu)$ est appelée un **couplage** de μ et ν. La distance de Wasserstein répond naturellement à un problème de transport optimal. Plus de détails sur ces espaces peuvent être trouvés dans [ref :Villani]. On se contentera ici d'admettre que $\mathcal{P}_p(E)$ muni de W_p est séparable et complet, ainsi que le critère de convergence :

$$\mu^N \to \mu \text{ dans } \mathcal{P}_p(E) \iff \langle \psi, \mu^N, \rangle \to \langle \psi, \mu \rangle \text{ pour tout } \psi \in C(E) \text{ avec } \psi(x) \lesssim 1 + \|x\|^p.$$

Il s'agit donc d'une topologie sensiblement meilleure que la convergence étroite, qui garantit en particulier la convergence des moments jusqu'à l'ordre p.

Remarque 1.5. Attention : dans la suite, on utilisera la même notation $W_p[\mu,\nu]$ pour désigner la distance entre deux mesures dans $\mathcal{P}_p(\mathbb{R}^d)$ ou dans $\mathcal{P}_p(\mathscr{C})$ indifféremment !

1.2 Limite de champ moyen

Théorème 1. Supposons que $a : \mathbb{R}^d \to \mathbb{R}^d$ est (globalement) Lipschitzien.

On considère μ^N la mesure empirique associée au système de particule (1.1).

Si $\mu_0^N \xrightarrow[N\to\infty]{} \mu_0$ dans $\mathcal{P}_p(\mathbb{R}^d)$ pour un $p \geq 1$, alors $\mu^N \xrightarrow[N\to\infty]{} \mu$ dans $\mathcal{P}_p(\mathscr{C})$,
où μ est l'unique solution de (1.5) de la forme transport avec la donnée initiale μ_0.

Remarque 1.6. Le plongement naturel $\begin{array}{ccc} \mathcal{P}_p(\mathscr{C}) & \longrightarrow & C([0,T];\mathcal{P}_p(\mathbb{R}^d)) \\ \mu & \longmapsto & (\mu_t)_{t\in[0,T]} \end{array}$ est continu.

Plus précisément, pour tous $\mu,\nu \in \mathcal{P}_p(\mathscr{C})$, on a l'inégalité : $\boxed{\sup_{t\in[0,T]} W_p[\mu_t,\nu_t] \leq W_p[\mu,\nu]}$

Preuve. Soit $\pi \in \Pi(\mu,\nu)$ un couplage de μ et ν. Pour tout $t \in [0,T]$, la mesure image de π par $(x,y) \in \mathscr{C}^2 \mapsto (x_t,y_t) \in (\mathbb{R}^d)^2$, que l'on note π_t, est ainsi un couplage de μ_t et ν_t. Il en résulte :

$$W_p[\mu_t,\nu_t]^p \leq \int_{\mathbb{R}^d\times\mathbb{R}^d} |x-y|^p d\pi_t(x,y) = \int_{\mathscr{C}\times\mathscr{C}} |x_t-y_t|^p d\pi(x,y) \leq \int_{\mathscr{C}\times\mathscr{C}} \|x-y\|_\infty^p d\pi(x,y).$$

Cette dernière inégalité est valable pour tout couplage $\pi \in \Pi(\mu,\nu)$. En passant à l'infimum sur π, on obtient $W_p[\mu_t,\nu_t]^p \leq W_p[\mu,\nu]^p$. C'est valable pour tout $t \in [0,T]$, d'où le résultat annoncé.

$\square$

Commençons par établir quelques résultats préliminaires.

Lemme 1.1.
Soit $p \geq 1$. Pour toutes $\mu,\nu \in \mathcal{P}_p(\mathbb{R}^d)$, et $x,x' \in \mathbb{R}^d$, $\quad \Big|A[\mu](x) - A[\nu](x')\Big| \lesssim |x-x'| + W_p[\mu,\nu]$.

Preuve du Lemme 1.1. On remarque que, pour tout couplage $\pi \in \Pi(\mu,\nu)$,

$$A[\mu](x) - A[\nu](x') = \int_{\mathbb{R}^d} a(x,y)d\mu(y) - \int_{\mathbb{R}^d} a(x',y')d\nu(y') = \int_{\mathbb{R}^d\times\mathbb{R}^d} (a(x,y)-a(x',y'))d\pi(y,y'),$$

d'où l'on tire

$$\Big|A[\mu](x) - A[\nu](x')\Big| \leq \int_{\mathbb{R}^d\times\mathbb{R}^d} |a(x,y)-a(x',y')| d\pi(y,y') \lesssim \int_{\mathbb{R}^d\times\mathbb{R}^d} (|x-x'|+|y-y'|)d\pi(y,y')$$

et donc

$$\Big|A[\mu](x) - A[\nu](x')\Big| \lesssim |x-x'| + \int_{\mathbb{R}^d\times\mathbb{R}^d} |y-y'|d\pi(y,y') \leq |x-x'| + \left(\int_{\mathbb{R}^d\times\mathbb{R}^d} |y-y'|^p d\pi(y,y')\right)^{1/p}.$$

En passant à l'infimum sur tous les couplages $\pi \in \Pi(\mu,\nu)$, on obtient l'estimation annoncée.

$\square$

La preuve du Théorème 1. reposera essentiellement sur le lemme suivant :

Lemme 1.2 (Lemme de contraction)**.**
Pour toutes $\mu,\nu \in \mathcal{P}_p(\mathscr{C})$ de la forme transport, $W_p[\mu,\nu] \lesssim W_p[\mu_0,\nu_0]$.
(La constante mise en jeu dans $\lesssim$ dépend de a et de T uniquement.)

Preuve du Lemme 1.2.
Considérons un couplage $\pi_0 \in \Pi(\mu_0, \nu_0)$.
Pour tout $t \in [0,T]$, puisque $\mu_t = (X_t^\mu)^*\mu_0$ et $\nu_t = (X_t^\nu)^*\nu_0$, la mesure image de π_0 par $(x,y) \mapsto (X_t^\mu(x), X_t^\nu(x))$, que l'on note π_t, est un couplage de μ_t, ν_t, de sorte que

$$W_p[\mu_t,\nu_t]^p \leq \int_{\mathbb{R}^d\times\mathbb{R}^d} |x-y|^p d\pi_t(x,y) = \int_{\mathbb{R}^d\times\mathbb{R}^d} \Big|X_t^\mu(x) - X_t^\nu(y)\Big|^p d\pi_0(x,y). \tag{1.7}$$

Pour tous $x, y \in \mathbb{R}^d$, l'estimation du Lemme 1.1 donne

$$\left|\frac{d}{dt}\Big(X_t^\mu(x) - X_t^\nu(y)\Big)\right| = \Big|A[\mu_t](X_t^\mu(x)) - A[\nu_t](X_t^\nu(y))\Big| \lesssim \Big|X_t^\mu(x) - X_t^\nu(y)\Big| + W_p[\mu_t,\nu_t],$$

c'est à dire, sous forme intégrée,

$$\Big|X_t^\mu(x) - X_t^\nu(y)\Big| \lesssim |x-y| + \int_0^t \Big|X_s^\mu(x) - X_s^\nu(y)\Big| ds + \int_0^t W_p[\mu_s,\nu_s] ds,$$

et donc

$$\Big|X_t^\mu(x) - X_t^\nu(y)\Big|^p \lesssim |x-y|^p + \int_0^t \Big|X_s^\mu(x) - X_s^\nu(y)\Big|^p ds + \int_0^t W_p[\mu_s,\nu_s]^p ds. \tag{1.8}$$

En intégrant ceci contre $d\pi_0(x,y)$ et en utilisant l'inégalité (1.7), il vient

$$\int_{\mathbb{R}^d\times\mathbb{R}^d} \Big|X_t^\mu(x) - X_t^\nu(y)\Big|^p d\pi_0(x,y) \lesssim \int_{\mathbb{R}^d\times\mathbb{R}^d} |x-y|^p d\pi_0(x,y) + \int_0^t \int_{\mathbb{R}^d\times\mathbb{R}^d} \Big|X_s^\mu(x) - X_s^\nu(y)\Big|^p d\pi_0(x,y) ds.$$

Par suite, un simple lemme de Grönwall donne

$$\forall t \in [0,T], \quad \int_{\mathbb{R}^d\times\mathbb{R}^d} \Big|X_t^\mu(x) - X_t^\nu(y)\Big|^p d\pi_0(x,y) \lesssim \int_{\mathbb{R}^d\times\mathbb{R}^d} |x-y|^p d\pi_0(x,y)$$

et donc, en utilisant l'inégalité (1.7) de nouveau,

$$\forall t \in [0,T], \quad W_p[\mu_t,\nu_t]^p \lesssim \int_{\mathbb{R}^d\times\mathbb{R}^d} |x-y|^p d\pi_0(x,y).$$

En passant à l'infimum sur $\pi_0 \in \Pi(\mu_0,\nu_0)$ on obtient ainsi le contrôle :

$$\forall t \in [0,T], \quad W_p[\mu_t,\nu_t]^p \lesssim W_p[\mu_0,\nu_0]^p. \tag{1.9}$$

Ce n'est pas encore tout à fait le résultat annoncé ! On peut maintenant revenir à (1.8) et déduire, pour tous $x, y \in \mathbb{R}^d$,

$$\Big|X_t^\mu(x) - X_t^\nu(y)\Big|^p \lesssim |x-y|^p + W_p[\mu_0,\nu_0]^p + \int_0^t \Big|X_s^\mu(x) - X_s^\nu(y)\Big|^p ds.$$

Un nouveau lemme de Grönwall donne ainsi le contrôle :

$$\forall t \in [0,T], \quad \Big|X_t^\mu(x) - X_t^\nu(y)\Big|^p \lesssim |x-y|^p + W_p[\mu_0,\nu_0]^p. \tag{1.10}$$

On peut maintenant conclure : puisque $\mu = (X^\mu)^*\mu_0$ et $\nu = (X^\nu)^*\nu_0$, la mesure image de π_0 par $(x,y) \mapsto (X^\mu(x), X^\nu(y))$, que l'on note π, est un couplage de μ et ν. Il en résulte que

$$W_p[\mu,\nu]^p \leq \int_{\mathscr{C}\times\mathscr{C}} \|x-y\|_\infty^p d\pi(x,y) = \int_{\mathbb{R}^d\times\mathbb{R}^d} \|X^\mu(x) - X^\nu(y)\|_\infty^p d\pi_0(x,y),$$

c'est à dire

$$W_p[\mu,\nu]^p \leq \int_{\mathbb{R}^d\times\mathbb{R}^d} \sup_{t\in[0,T]} \Big|X_t^\mu(x) - X_t^\nu(y)\Big|^p d\pi_0(x,y). \tag{1.11}$$

En utilisant (1.10), on obtient

$$W_p[\mu,\nu]^p \lesssim \int_{\mathbb{R}^d\times\mathbb{R}^d} |x-y|^p d\pi_0(x,y) + W_p[\mu_0,\nu_0]^p,$$

puis en passant à l'infimum sur $\pi_0 \in \Pi(\mu_0,\nu_0)$, $W_p[\mu,\nu]^p \lesssim W_p[\mu_0,\nu_0]^p$, ce qui conclut la preuve.

□

On peut à présent démontrer le Théorème 1.

Preuve du Théorème 1. On a vu que les mesures empiriques μ^N sont de la forme transport. L'inégalité du Lemme de contraction 1.2 donne ainsi, pour tous $N, M \in \mathbb{N}$, $W_p[\mu^N, \mu^M] \lesssim W_p[\mu_0^N, \mu_0^M]$. Puisque la suite $(\mu_0^N)_{N\geq 1}$ converge dans $\mathcal{P}_p(\mathbb{R}^d)$, elle est de Cauchy : on en déduit que $(\mu^N)_{N\geq 1}$ est elle-même de Cauchy dans $\mathcal{P}_p(\mathscr{C})$, donc converge vers un certain $\mu \in \mathcal{P}_p(\mathscr{C})$.

Montrons que μ est de la forme transport. On peut introduire les caractéristiques non-linéaires X^μ (solution de (1.6)), et donc la mesure $\nu = (X^\mu)^*\mu_0$. Il s'agit ainsi de justifier que $\mu = \nu$.
Pour se faire, on se contente de légèrement adapter la preuve du Lemme de contraction 1.2.

Considèrons un couplage $\pi_0^N \in \Pi(\mu_0^N, \mu_0)$.
Puisque $\mu^N = (X^{\mu^N})^*\mu_0^N$ et $\nu = (X^\mu)^*\mu_0$, la mesure image de π_0^N par $(x, y) \mapsto (X^{\mu^N}(x), X^\mu(y))$ est un couplage de μ^N et ν. On en déduit, comme en (1.11), que

$$W_p[\mu^N, \nu]^p \leq \int_{\mathbb{R}^d\times\mathbb{R}^d} \sup_{t\in[0,T]} \left|X_t^{\mu^N}(x) - X_t^\mu(y)\right|^p d\pi_0^N(x, y).$$

Or cette fois, on voit qu'on a, pour tous $x, y \in \mathbb{R}^d$,

$$\left|\frac{d}{dt}\Big(X_t^{\mu^N}(x) - X_t^\mu(y)\Big)\right| \lesssim \left|X_t^{\mu^N}(x) - X_t^\mu(y)\right| + W_p[\mu_t^N, \mu_t] \leq \left|X_t^{\mu^N}(x) - X_t^\mu(y)\right| + W_p[\mu^N, \mu].$$

Un lemme de Grönwall donnera ainsi

$$\forall t \in [0, T], \quad \left|X_t^{\mu^N}(x) - X_t^\mu(y)\right|^p \lesssim |x - y|^p + W_p[\mu^N, \mu]^p,$$

ce qui conduit à

$$W_p[\mu^N, \nu]^p \leq \int_{\mathbb{R}^d\times\mathbb{R}^d} |x - y|^p d\pi_0^N(x, y) + W_p[\mu^N, \mu]^p.$$

En passant à l'infimum sur $\pi_0^N \in \Pi(\mu_0^N, \mu_0)$, on obtient finalement l'estimation

$$W_p[\mu^N, \nu]^p \lesssim W_p[\mu_0^N, \mu_0]^p + W_p[\mu^N, \mu]^p.$$

Les deux termes du membre de droite tendent vers 0 quand N tend vers l'infini : on en déduit que $\mu^N \to \nu$ dans $\mathcal{P}_p(\mathscr{C})$ et donc que $\mu = \nu$ par unicité de la limite.

Enfin, l'unicité d'une mesure de la forme transport avec la donnée initiale μ_0 est à nouveau assurée par l'estimation du Lemme de contraction 1.2 : si ν est de la forme transport avec $\nu_0 = \mu_0$, alors

$$W_p[\mu, \nu] \lesssim W_p[\mu_0, \nu_0] = 0$$

et donc $\mu = \nu$. Ceci conclut la preuve du Théorème. □

1.3 Propagation du chaos

Un autre résultat central concernant les systèmes de particules à champ moyen, qui découle assez directement de la convergence que l'on vient d'établir, est la **propagation du chaos**. Considérons à nouveau le système d'équation

$$\begin{cases} \dfrac{dX_t^{i,N}}{dt} = \dfrac{1}{N}\displaystyle\sum_{j=1}^N a(X_t^{i,N}, X_t^{j,N}), & \qquad i \in \{1, \ldots, N\}, \\ X_0^{i,N} = x_0^i, & \end{cases}$$

avec, cette fois ci, des données initiales aléatoires : $(x_0^i)_{i\geq 1}$ est une suite de variables aléatoires indépendantes, identiquement distribuées selon une loi $\mu_0 \in \mathcal{P}(\mathbb{R}^d)$.

Cette première introduction d'aléa (dans la condition initiale) est relativement répandue dans l'étude des EDP. Il ne s'agit pas encore d'un "champ moyen stochastique" à proprement parler! Notons cependant que les $X^{i,N} \in \mathscr{C}$ et la mesure empirique $\mu^N \in \mathcal{P}(\mathscr{C})$ deviennent des variables aléatoires, mesurables par rapport à la tribu $\sigma(x_0^i, i \geq 1)$.

Lorsque le nombre N de particules devient grand, la loi des grands nombres suggère que la mesure empirique initiale (aléatoire) $\mu_0^N = \frac{1}{N}\sum_{i=1}^N \delta_{x_0^i}$ est "proche" de la mesure (deterministe) μ_0. On s'attend ainsi à ce que la convergence $\mu^N \to \mu$ du Théorème 1 s'applique dans ce cadre.
Plus précisément, on peut établir un résultat intéressant sur la convergence des particules $X^{i,N}$.

Théorème 2. Supposons que $a : \mathbb{R}^d \to \mathbb{R}^d$ est (globalement) Lipschitzien.

Soit $p \geq 1$ et $\mu_0 \in \mathcal{P}_p(\mathbb{R}^d)$.

Soit $(x_0^i)_{i\geq 1}$ une suite de variables aléatoires de $\mathbb{R}^d$, i.i.d de loi μ_0, définies sur $(\Omega, \mathcal{F}, P)$.

Soit $(X^{i,N})_{1\leq i\leq N}$ la solution de (1.1) avec $X^{i,N} = x_0^i$, et μ^N la mesure empirique associée.

Alors $\mu^N \xrightarrow[N\to\infty]{} \mu$ p.s dans $\mathcal{P}_p(\mathscr{C})$ (et dans $L^q(\Omega; \mathcal{P}_q(\mathscr{C}))$ pour tout $q < p$), où μ est l'unique solution de (1.5) de la forme transport avec la donnée initiale μ_0.

De plus, pour tout $r \geq 1$ et $\Psi_1, \dots, \Psi_r \in C_b(\mathscr{C})$, $\mathbb{E}\Big[\prod_{i=1}^r \Psi_i(X^{i,N})\Big] \xrightarrow[N\to\infty]{} \prod_{i=1}^n \langle \Psi_i, \mu\rangle$.

Enfin, pour tout $i \geq 1$, en introduisant la solution X^i de $\begin{cases} \dfrac{dX_t^i}{dt} = A[\mu_t](X_t^i), \quad t \in [0,T], \\ X_0^i = x_0^i. \end{cases}$

alors $X^i \in \mathscr{C}$ est de loi μ et $X^{i,N} \xrightarrow[N\to\infty]{} X^i$ p.s dans $\mathscr{C}$ (et dans $L^q(\Omega; \mathscr{C})$ pour tout $q < p$).

Remarque 1.7.
• La mesure limite $\mu \in \mathcal{P}_p(\mathscr{C})$ est déterministe : l'aléa disparait par "moyennisation" quand $N \to \infty$.

• Le processus aléatoire $(X_t^i)_{t\geq 0}$ interagit avec μ_t, c'est à dire avec sa propre loi!
Un tel processus est appelé **processus de McKean-Vlasov**.

• Les $(X^i)_{i\geq 1}$ sont des variables aléatoires à valeurs dans $\mathscr{C}$ i.i.d de loi μ. Ainsi, la convergence $(X^{1,N}, \dots, X^{r,N}) \to (X^1, \dots, X^r)$ implique naturellement, pour les lois, celle de $\mathbb{E}\Big[\prod_{i=1}^r \Psi_i(X^{i,N})\Big]$.
On choisit malgré tout de présenter (dans le Théorème et dans sa preuve) ces deux convergences séparément car : d'une part, c'est ainsi que le résultat est le plus souvent énoncé; d'autre part, il est certains cas où l'on ne pourra pas obtenir la convergence presque-sûre, mais seulement celle des lois.

Donnons un éclairage sur le résultat de propagation du chaos. La donnée initiale $X_0^{i,N} = x_0^i$ peut être qualifiée de "chaotique" : à l'instant $t = 0$, chaque particule est distribuée indépendamment des autres. À N fixé, dès lors que $t > 0$, il est clair que deux particules $X_t^{i,N}$ et $X_t^{j,N}$ pour $i \neq j$ deviennent immediatement corrélées, du fait de leur interaction. Cependant, dans la limite où le nombre de particules est très grand, les corrélations fini-dimensionnelles disparaissent : si l'on observe un nombre fini ($r \geq 1$) de particules, tout se passe comme si elles évoluaient de façon totalement indépendante, suivant chacune la dynamique dictée par le processus limite X^i. Dans cette asymptotique, une particule donnée ne "voit" alors plus les autres, mais se contente d'interagir directement avec la mesure limite μ. Dans la limite $N \to \infty$, le "chaos" introduit initialement se "propage" aux temps positifs!

La propagation du chaos est une formalisation élégante du fait que chaque individu au sein d'un système de très grande envergure se sent impuissant à en altérer les rouages! Quand bien même ledit système est concrètement constitué d'une multitude d'individus en interaction, ces derniers ne se "voient" jamais vraiment : si l'on observe un nombre fini d'entre eux, chacun semble toujours évoluer indépendamment des autres, interagissant de la même façon avec un objet limite μ fixé, sur lequel il n'a sensiblement aucune d'emprise... Sur ces belles paroles, démontrons les résultats annoncés!

Preuve du Théorème 2. La loi des grands nombres donne, pour tout $\psi \in C(\mathbb{R}^d)$ avec $|\psi(x)| \lesssim 1+|x|^p$,

$$\langle \psi, \mu_0^N \rangle = \frac{1}{N} \sum_{i=1}^{N} \psi(x_0^i) \xrightarrow[N\to\infty]{} \mathbb{E}[\psi(x_0^1)] = \langle \psi, \mu_0 \rangle.$$

Ceci prouve que $\mu_0^N \to \mu_0$ presque sûrement dans $\mathcal{P}_p(\mathbb{R}^d)$. On applique le résultat du Théorème 1 (à $\omega \in \Omega$ fixé), qui garantit donc la convergence $\mu^N \to \mu$ dans $\mathcal{P}_p(\mathscr{C})$ presque sûrement, où μ est l'unique solution de (1.5) de la forme transport.
Pour tout $q < p$, on a bien sûr $W_q[\mu^N, \mu] \le W_p[\mu^N, \mu] \to 0$. De plus l'inégalité du Lemme 1.2

$$W_q[\mu^N, \mu]^q \lesssim W_q[\mu_0^N, \mu_0]^q \le \int_{R^d} |x|^q d\mu_0^N(x) + \int_{R^d} |x|^q d\mu_0(x)$$

conduit, en posant $\alpha = p/q > 1$, à

$$\mathbb{E}\Big[\Big(W_q[\mu^N, \mu]^q \Big)^\alpha \Big] \lesssim \mathbb{E}\Big[\int_{R^d} |x|^p d\mu_0^N(x) \Big] + \int_{R^d} |x|^p d\mu_0(x) = 2 \int_{R^d} |x|^p d\mu_0(x) < \infty,$$

ce qui montre l'uniforme intégrabilité de $(W_q[\mu^N, \mu]^q)_{N\ge 1}$. On conclut que $\mathbb{E}\Big[W_q[\mu^N, \mu]^q \Big] \to 0$, c'est à dire $\mu^N \to \mu$ dans $L^q(\Omega; \mathcal{P}_q(\mathscr{C}))$.

Intéressons-nous maintenant aux particules $X^{1,N}, X^{2,N}, \ldots X^{r,N}$. La propagation du chaos se déduit de la convergence $\mu^N \to \mu$ à l'aide d'un argument de symétrie simple mais efficace !
On remarque, au vu du système (1.1) et des conditions initiales $(x_0^i)_{1\le i \le N}$ i.i.d, que pour tous $p_1, p_2, \ldots, p_r \in \{1, \ldots, N\}$ deux à deux distincts,

$$\Big(X^{p_1,N}, X^{p_2,N}, \ldots, X^{p_r,N} \Big) \overset{Loi}{=} \Big(X^{1,N}, X^{2,N}, \ldots, X^{r,N} \Big).$$

Écrivons $\Big| \mathbb{E}\Big[\prod_{i=1}^{r} \Psi_i(X^{i,N}) \Big] - \prod_{i=1}^{r} \langle \Psi_i, \mu \rangle \Big| \le A^N + B^N$, avec

$$A^N = \Big| \mathbb{E}\Big[\prod_{i=1}^{r} \Psi_i(X^{i,N}) \Big] - \mathbb{E}\Big[\prod_{i=1}^{r} \langle \Psi_i, \mu^N \rangle \Big] \Big|,$$
$$B^N = \Big| \mathbb{E}\Big[\prod_{i=1}^{r} \langle \Psi_i, \mu^N \rangle \Big] - \prod_{i=1}^{r} \langle \Psi_i, \mu \rangle \Big|.$$

Comme $\mu^N \to \mu$ p.s, on a directement $B^N \le \mathbb{E}\Big| \prod_{i=1}^{r} \langle \Psi_i, \mu^N \rangle - \prod_{i=1}^{r} \langle \Psi_i, \mu \rangle \Big| \to 0$ par convergence dominée. Par ailleurs, on note que

$$\mathbb{E}\Big[\prod_{i=1}^{r} \langle \Psi_i, \mu^N \rangle \Big] = \frac{1}{N^r} \sum_{1 \le p_1, \ldots, p_r \le N} \mathbb{E}\Big[\prod_{i=1}^{r} \Psi_i(X^{p_i,N}) \Big]$$

et, en utilisant la symétrie en loi,

$$\mathbb{E}\Big[\prod_{i=1}^{r} \Psi_i(X^{i,N}) \Big] = \frac{1}{N(N-1)\ldots(N-r+1)} \sum_{\substack{1 \le p_1, \ldots, p_r \le N \\ \text{2 à 2 distincts}}} \mathbb{E}\Big[\prod_{i=1}^{r} \Psi_i(X^{p_i,N}) \Big].$$

Ainsi, les Ψ_i étant bornés, on obtient facilement

$$A^N \lesssim \left(\frac{1}{N\ldots(N-r+1)} - \frac{1}{N^r} \right) N\ldots(N-r+1) + \frac{N^r - N\ldots(N-r+1)}{N^r}$$

soit

$$A^N \lesssim 2 \left(1 - \frac{N(N-1)\ldots(N-r+1)}{N^r} \right) \xrightarrow[N\to\infty]{} 0.$$

Enfin, étudions la convergence d'une particule $X^{i,N}$. En considérant les caractéristiques non-linéaires (1.6), on peut noter que $X^{i,N} = X^{\mu^N}(x_0^i)$ et $X^i = X^\mu(x_0^i)$. Ainsi, en adaptant à nouveau les arguments avancés dans la preuve du Lemme de contraction 1.2,

$$\left|\frac{d}{dt}\Big(X_t^{i,N} - X_t^i\Big)\right| \lesssim \left|X_t^{i,N} - X_t^i\right| + W_p[\mu_t^N, \mu_t] \leq \left|X_t^{i,N} - X_t^i\right| + W_p[\mu^N, \mu],$$

si bien qu'un lemme de Grönwall conduit à

$$\|X^{i,N} - X^i\|_\infty = \sup_{t\in[0,T]} \left|X_t^{i,N} - X_t^i\right| \lesssim W_p[\mu^N, \mu].$$

On en déduit que $\|X^{i,N} - X^i\|_\infty \to 0$ p.s. Pour $q < p$, le même argument (en remplaçant p par q) donne l'estimation

$$\|X^{i,N} - X^i\|_\infty \lesssim W_q[\mu^N, \mu],$$

et on déduit $\mathbb{E}\Big[\|X^{i,N} - X^i\|_\infty^q\Big] \lesssim \mathbb{E}\Big[W_q[\mu^N, \mu]^q\Big] \to 0$, c'est à dire $X^{i,N} \to X^i$ dans $L^q(\Omega; \mathscr{C})$, ce qui conclut la preuve du Théorème.

$\square$

Remarque 1.8. Le Théorème Central Limite suggère que la convergence $\mu_0^N \to \mu_0$ – et par suite, les convergences $\mu^N \to \mu$ et $X^{i,N} \to X^i$, grâce au Lemme de contraction – se fait à la vitesse $N^{-1/2}$.