The Vlasov equation & mean field limit of a particle system

Simon Mayer

December 9, 2014

In this talk we present some results about how the Vlasov equation can be seen to arise as a mean field limit of a system of interacting particles. This is understood well for nice enough interaction potentials, and we re-express known results using notions from probability theory and optimal transportation. The situation gets more complicated as soon as one considers interactions which become singular at small distances such as the Coulomb force or the gravitational force, i.e. a force of the form $1/|x|^\alpha$ with $\alpha = d-1$ and d the dimension of the system. We explain key ideas in the proof of the mean field limit and the propagation of chaos in the case $\alpha < d-1$ but very close to $d-1$. This involves a control on the trajectories of particles that get very close to each other in position-velocity space.

Contents

1 Introduction

In this talk we will take a closer look at the Vlasov-Poisson equation, which is a type of mean field evolution PDE. It is given by

$$\begin{cases} \partial_t f(t,x,v) + v \cdot \nabla_x f(t,x,v) + E(t,x) \cdot \nabla_v f(t,x,v) = 0, \quad x, v \in \mathbb{R}^d, t \in \mathbb{R} \\ E(t,x) = \int_{\mathbb{R}^d} \rho(t,y) F(x-y) \, \mathrm{d}y \\ \rho(t,x) = \int_{\mathbb{R}^d} f(t,x,v) \, \mathrm{d}v. \end{cases} \tag{1}$$

The unknown $f = f(t,x,v)$ is the number density at time t of particles located at position x with velocity v and $F(x)$ is the interaction force. We shall be primarily interested in how the Vlasov equation (1) can be derived as the limit of an N-particle system. Denote by $X_i, V_i \in \mathbb{R}^d$ the position and velocity of the i-th particle. Assuming a two-body interaction, the system of ordinary differential equations governing the time evolution of the system is

$$\begin{cases} \frac{\mathrm{d}X_i}{\mathrm{d}t} = V_i, & i = 1, \ldots, N \\ \frac{\mathrm{d}V_i}{\mathrm{d}t} = \frac{1}{N} \sum_{j \neq i} F(X_i - X_j). \end{cases} \tag{2}$$

This is just Newton's second law, with a rescaled force term. This is referred to as mean field scaling and comes from the following fact. Let us generalize the system of ODEs a bit and start from an unscaled system

$$\frac{\mathrm{d}\hat{z}_i}{\mathrm{d}t} = \sum_{j \neq i} K(\hat{z}_i, \hat{z}_j)$$

for $\hat{z}_i \in \mathbb{R}^d$ and $K : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$ the interaction kernel. We want to choose a time variable $\hat{t}$ such that $\frac{\mathrm{d}\hat{z}_i}{\mathrm{d}t} = \mathcal{O}(1)$ for all i as $N \to \infty$. Assuming that each force term $K(\hat{z}_i, \hat{z}_j) = \mathcal{O}(1)$, we set $\hat{t} = \frac{t}{N}$. Then we have

$$\frac{\mathrm{d}\hat{z}_i}{\mathrm{d}\hat{t}} = \frac{1}{N} \sum_{j \neq i} K(\hat{z}_i, \hat{z}_j).$$

If we consider now this equation as our starting point and forget about the hats, we arrive at our desired mean field scaling.

The key idea for the mean field limit is the expected convergence of

$$\frac{1}{N} \sum_{j \neq i} K(z_i, z_j) \to \int_{\mathbb{R}^d} K(z_i, z') f(t, \mathrm{d}z')$$

as $N \to \infty$ when the z_i are distributed under the probability measure $f(t, \mathrm{d}z')$ (in some sense). Then we could replace the system of ODEs by

$$\frac{\mathrm{d}z}{\mathrm{d}t} = \int_{\mathbb{R}^d} K(z(t), z') f(t, \mathrm{d}z')$$

which is the equation of characteristics of the mean field PDE

$$\partial_t f + \mathrm{div}_z(f\mathcal{K}f) = 0,$$

where $\mathcal{K}f(t,z) := \int_{\mathbb{R}^d} K(z,z')f(t,\mathrm{d}z')$. This equation is to be understood in the sense of distributions, since f is a priori only a Borel probability measure:

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathbb{R}^d} \phi(z)f(t,\mathrm{d}z) = \int_{\mathbb{R}^d} \mathcal{K}f(t,z)\cdot\nabla\phi(z)f(t,\mathrm{d}z)$$

for each test function $\phi \in C_b^1(\mathbb{R}^d)$. To return from this general description to the Vlasov equation (1) given above, for $z = (x,v) \in \mathbb{R}^6$ set

$$K(z,z') = K(x,v,x',v') = (v - v', F(x - x')).$$

The question of the mean field limit is interesting mainly for two reasons. Firstly, one would gain a justification of the validity of the Vlasov equation and secondly, it would have practical applications for numerical simulations: Knowing that the discrete ODE model will converge to the continuous PDE we can simulate the ODE model with much fewer particles than the actual system possesses, dictated of course by the rate of convergence obtained.

In section 2, we present a general formalism to describe mean field limits that is valid for any mean field PDE with smooth (nice enough) interaction force. We derive a stability estimate for two different solutions of the mean field PDE and this enables us to derive the mean field limit which was formally described above.

In section 3 we look at more singular interactions that are related more closely to physically relevant interactions such as the Coulomb/gravitational force.

2 A general formalism for mean field limits

2.1 Regularity

For each N-tuple $Z_N = (z_1,\ldots,z_N) \in (\mathbb{R}^d)^N$ we define the empirical measure

$$\mu_{Z_N} := \frac{1}{N}\sum_{k=1}^{N}\delta_{z_k},$$

where δ_{z_k} is the Dirac measure at position z_k.

Given a function $K \in C^1(\mathbb{R}^d \times \mathbb{R}^d, \mathbb{R}^d)$, we define the integral operator $\mathcal{K}$ by acting on the probability measure $f(t,\mathrm{d}z)$ as

$$\mathcal{K}f(t,z) := \int_{\mathbb{R}^d} K(z,z')f(t,\mathrm{d}z').$$

Theorem 1 *Assume that the interaction kernel $K \in C^1(\mathbb{R}^d \times \mathbb{R}^d, \mathbb{R}^d)$ satisfies*

1. *K is skew-symmetric, i.e. $K(z,z') + K(z',z) = 0$ for all $z,z' \in \mathbb{R}^d$,*

2. K has bounded partial derivatives of order 1, i.e. there exists a constant $L \geq 0$ such that $\sup_{z'} |\nabla_z K(z, z')| \leq L$ and $\sup_z |\nabla_{z'} K(z, z')| \leq L$.

Then we have

1. for each $N \geq 1$ and each N-tuple $Z_N^{in} = (z_1^{in}, \dots, z_N^{in})$, the Cauchy problem for the N-particle ODE system

$$\begin{cases} \dot{z}_i(t) = \frac{1}{N} \sum_{j=1}^N K(z_i(t), z_j(t)) \quad i = 1, \dots, N, \\ z_i(0) = z_i^{in} \end{cases} \tag{3}$$

has a unique solution of class C^1 on $\mathbb{R}$ denoted by $t \mapsto Z_N(t) = (z_1(t), \dots, z_N(t)) =: T_t Z_N^{in}$,

2. the empirical measure $f(t, \mathrm{d}z) = \mu_{T_t Z_N^{in}}$ is a weak solution of the Cauchy problem for the mean field PDE

$$\begin{cases} \partial_t f + \mathrm{div}_z (f \mathcal{K} f) = 0, \\ f|_{t=0} = f^{in}. \end{cases} \tag{4}$$

Recall the definition of the push-forward measure. For two measurable spaces $(X, \mathcal{A})$ and $(Y, \mathcal{B})$, a measurable map $\Phi : (X, \mathcal{A}) \to (Y, \mathcal{B})$ and a measure m on $(X, \mathcal{A})$, the push-forward measure of m under Φ is the measure on $(Y, \mathcal{B})$ defined by

$$\Phi \# m(B) := m(\Phi^{-1}(B)) \quad \text{for all } B \in \mathcal{B}.$$

Theorem 2 *Under the assumptions on K from theorem 1, for each $\zeta^{in} \in \mathbb{R}^d$ and each Borel probability measure $\mu^{in} \in \mathcal{P}_1(\mathbb{R}^d)$, there exists a unique solution denoted $t \mapsto Z(t, \zeta^{in}, \mu^{in})$ of class C^1 of the problem*

$$\begin{cases} \partial_t Z(t, \zeta^{in}, \mu^{in}) = \mathcal{K}\mu(t)(Z(t, \zeta^{in}, \mu^{in})), \\ \mu(t) = Z(t, \cdot, \mu^{in}) \# \mu^{in}, \\ Z(0, \zeta^{in}, \mu^{in}) = \zeta^{in}. \end{cases}$$

We call the solution of theorem 2 the mean field characteristic flow. It is related to the flow T_t associated to the N-particle ODE system via the following

Proposition 3 *Under the same assumptions as in theorem 1, we have for each $Z_N^{in} = (z_1^{in}, \dots, z_N^{in})$ that the solution $T_t Z_N^{in} = (z_1(t), \dots, z_N(t))$ of the N-body problem and the mean field characteristic flow $Z(t, \zeta^{in}, \mu^{in})$ satisfy*

$$z_i(t) = Z(t, z_i^{in}, \mu_{Z_N^{in}}), \quad i = 1, \dots, N,$$

for all $t \in \mathbb{R}$.

2.2 Dobrushin's stability estimate

The Wasserstein (or Monge-Kantorovich) distance of order $r \geq 1$ between two (Borel) probability measures with r-th moments $\mu, \nu \in \mathcal{P}_r(\mathbb{R}^d)$ is defined by

$$W_r(\mu, \nu) = \inf_{\pi \in \Pi(\mu,\nu)} \left(\int_{\mathbb{R}^d \times \mathbb{R}^d} |x-y|^r \pi(\mathrm{d}x, \mathrm{d}y) \right)^{1/r}.$$

Here, $\Pi(\mu, \nu)$ is the space of couplings of μ and ν, i.e. it is the set of probability measures π on $\mathbb{R}^d \times \mathbb{R}^d$ which have first marginal μ and second marginal ν. In the special case $r = 1$ it is actually true that

$$W_1(\mu, \nu) = \sup_{\mathrm{Lip}(\phi) \leq 1} \left| \int_{\mathbb{R}^d} \phi(z)\mu(\mathrm{d}z) - \int_{\mathbb{R}^d} \phi(z)\nu(\mathrm{d}z) \right|,$$

where $\mathrm{Lip}(\phi)$ denotes the Lipschitz constant of ϕ and the supremum is taken over all Lipschitz functions on $\mathbb{R}^d$. This fact comes from a duality argument in optimization.

Theorem 4 (Dobrushin [1]) *Under the assumptions of theorem 1, for $\mu_1^{in}, \mu_2^{in} \in \mathcal{P}_1(\mathbb{R}^d)$ and $t \in \mathbb{R}$, we let*

$$\mu_1(t) = Z(t, \cdot, \mu_1^{in})\#\mu_1^{in}, \quad \mu_2(t) = Z(t, \cdot, \mu_2^{in})\#\mu_2^{in}.$$

Then for all $t \in \mathbb{R}$, we have

$$W_1(\mu_1(t), \mu_2(t)) \leq \mathrm{e}^{2L|t|}\, W_1(\mu_1^{in}, \mu_2^{in}).$$

Theorem 5 (Mean field limit) *Assume that the interaction kernel $K \in C^1(\mathbb{R}^d \times \mathbb{R}^d, \mathbb{R}^d)$ satisfies the assumptions of theorem 1. Let f^{in} be a probability density on $\mathbb{R}^d$ such that*

$$\int_{\mathbb{R}^d} |z| f^{in}(z)\, \mathrm{d}z < \infty.$$

Then the Cauchy problem for the mean field PDE (4) with initial data f^{in} has a unique weak solution $f \in C(\mathbb{R}, L^1(\mathbb{R}^d))$. For each $N \geq 1$, let $\mathfrak{Z}(N) = (z_{1,N}^{in}, \ldots, z_{N,N}^{in}) \in (\mathbb{R}^d)^N$ be such that the empirical measure $\mu_{\mathfrak{Z}(N)}$ satisfies (where $\mathcal{L}^d$ denotes Lebesgue measure in d dimensions)

$$W_1(\mu_{\mathfrak{Z}(N)}, f^{in}\mathcal{L}^d) \to 0 \quad \textit{as } N \to \infty. \tag{5}$$

Let $t \mapsto T_t\mathfrak{Z}(N) = (z_{1,N}(t), \ldots, z_{N,N}(t)) \in (\mathbb{R}^d)^N$ be the solution of the N-particle ODE (3) with initial data $\mathfrak{Z}(N)$. Then

$$\mu_{T_t\mathfrak{Z}(N)} \to f(t, \cdot)\mathcal{L}^d \quad \textit{as } N \to \infty$$

in the weak topology of probability measures, with convergence rate

$$W_1(\mu_{T_t\mathfrak{Z}(N)}, f(t, \cdot)\mathcal{L}^d) \leq \mathrm{e}^{2L|t|}\, W_1(\mu_{\mathfrak{Z}(N)}, f^{in}\mathcal{L}^d) \to 0 \quad \textit{as } N \to \infty$$

for each $t \in \mathbb{R}$.

Take a probability density f^{in} on $\mathbb{R}^d$ that satisfies

$$\int_{\mathbb{R}^d} |z|^2 f^{\text{in}}(z)\,\mathrm{d}z < \infty.$$

Let $\Omega = (\mathbb{R}^d)^{\mathbb{N}}$, i.e. it is the set of sequences of points in $\mathbb{R}^d$ indexed by $\mathbb{N}$. Let

$$\mathcal{F} = \left\langle \prod_{n\geq 1} B_n \right\rangle$$

be the σ-algebra on Ω generated by cylinders, B_n are Borel sets in $\mathbb{R}^d$ where $B_n = \mathbb{R}^d$ for all but finitely many n. The measurable space $(\Omega, \mathcal{F})$ is then endowed with the probability measure $\mathbb{P} = (f^{\text{in}})^{\otimes\infty}$, which acts on the cylinder sets as

$$\mathbb{P}\left(\prod_{n\geq 1} B_n\right) = \prod_{n\geq 1} f^{\text{in}}(B_n).$$

The condition (5) is guaranteed by the following

Theorem 6 *For each $\mathbf{z}^{in} = (z_k^{in})_{k\geq 1} \in \Omega$, let $Z_N^{in} = (z_1^{in}, \dots, z_N^{in})$. Then*

$$W_1(\mu_{Z_N^{in}}, f^{in}\mathcal{L}^d) \to 0 \quad \text{as } N \to \infty$$

for $\mathbb{P}$-almost every $\mathbf{z}^{in} \in \Omega$.

Lemma 7 ([7]) *The Wasserstein distance of order one metrizes the topology of weak convergence. This means for a sequence $(\mu_n)_{n\geq 1}$ of probability measures in $\mathcal{P}_1(\mathbb{R}^d)$ and $\mu \in \mathcal{P}_1(\mathbb{R}^d)$ the two following statements are equivalent:*

1. *$W_1(\mu_n, \mu) \to 0$ as $n \to \infty$,*
2. *$\mu_n \to \mu$ in weak sense as $n \to \infty$ and the "tightness condition"*

   $$\sup_n \int_{|z|\geq R} |z|\mu_n(\mathrm{d}z) \to 0 \quad \text{as } R \to \infty$$

   holds.

2.3 BBGKY hierarchy, propagation of chaos

We define the space of symmetric probability measures on the N-particle phase space by

$$\mathcal{P}_{\text{sym}}((\mathbb{R}^d)^N) = \{P \in \mathcal{P}((\mathbb{R}^d)^N) : S_\sigma \# P = P \text{ for all } \sigma \in \mathfrak{S}_N\},$$

where for a permutation σ the transformation S_σ on $(\mathbb{R}^d)^N$ is defined by

$$S_\sigma(z_1, \dots, z_N) = (z_{\sigma(1)}, \dots, z_{\sigma(N)}).$$

For each $N \in \mathbb{N}$, each $P_N \in \mathcal{P}_{\mathrm{sym}}((\mathbb{R}^d)^N)$ and each $k = 1, \dots, N$, we define the k-particle marginal $P_{N:k} \in \mathcal{P}_{\mathrm{sym}}((\mathbb{R}^d)^k)$ by the formula

$$\int_{(\mathbb{R}^d)^k)} \phi(z_1, \dots, z_k) P_{N:k}(\mathrm{d}z_1 \dots \mathrm{d}z_k) = \int_{(\mathbb{R}^d)^N} \phi(z_1, \dots, z_k) P_N(\mathrm{d}z_1 \dots \mathrm{d}z_N)$$

for test functions $\phi \in C_b((\mathbb{R}^d)^k)$.

Now consider an N-particle symmetric probability measure $F_N^{\mathrm{in}} \in \mathcal{P}_{\mathrm{sym}}((\mathbb{R}^d)^N)$ and define its time evolution via the push-forward unter the flow associated to the N-particle ODE (3),

$$F_N(t) := T_t \# F_N^{\mathrm{in}}, \quad t \in \mathbb{R}.$$

Then $F_N(t)$ is the unique weak solution in $C(\mathbb{R}, \text{w-}\mathcal{P}((\mathbb{R}^d)^N))$ of the Cauchy problem for the N-particle Liouville equation

$$\begin{cases} \partial_t F_N + \frac{1}{N} \sum_{i,j=1}^N \mathrm{div}_{z_i}(F_N K(z_i, z_j)) = 0, \quad z_1, \dots, z_N \in \mathbb{R}^d, t \in \mathbb{R}, \\ F_N|_{t=0} = F_N^{\mathrm{in}}. \end{cases} \tag{6}$$

Now it is important to notice that the symmetry condition on F_N^{in} is being propagated by the flow of the N-particle Liouville equation.

Theorem 8 *Assume $K \in C^1(\mathbb{R}^d \times \mathbb{R}^d, \mathbb{R}^d)$ satisfies the assumptions of theorem 1 and let $F_N^{in} \in \mathcal{P}_{1,sym}((\mathbb{R}^d)^N)$ with time evolution $F_N(t) = T_t \# F_N^{in}$ for all $t \in \mathbb{R}$. The sequence of marginal distributions $F_{N:j}$ of F_N for $j = 1, \dots, N$ is a weak solution of the string of equations*

$$\begin{cases} \partial_t F_{N:1} + \frac{N-1}{N} \mathrm{div}_{z_1}[K(z_1, z_2) F_{N:2}]_{:1} = 0, \\ \partial_t F_{N:j} + \frac{N-j}{N} \sum_{l=1}^j \mathrm{div}_{z_l}[K(z_l, z_{j+1}) F_{N:j+1}]_{:j} \\ \quad + \frac{1}{N} \sum_{k,l=1}^j \mathrm{div}_{z_l}(K(z_l, z_k) F_{N:j}) = 0, \qquad j = 2, \dots N-1, \\ \partial_t F_{N:N} + \frac{1}{N} \sum_{k,l=1}^N \mathrm{div}_{z_l}(K(z_l, z_k) F_{N:N}) = 0 \end{cases}$$

with initial conditions $F_{N:j}|_{t=0} = F_{N:j}^{in}$, $j = 1, \dots, N$.

The string of equations in theorem 8 is called the BBGKY hierarchy. It is equivalent to the N-particle Liouville equation (6) and the advantage of deriving it lies in the following argument. Formally, we want to pass to the limit $N \to \infty$ in (6) but this poses the question as to what the limiting object will be, as it is a symmetric function of infinitely many variables. If we take the BBKGY hierarchy and pass for a fixed j to the limit in $F_{N:j}$, the limit function F_j (assuming it exists) will be a priori a symmetric function of j variables. For the middle term in the equations, we have for $N \to \infty$

$$\frac{N-j}{N} \int_{\mathbb{R}^d} K(z_l, z_{j+1}) F_{N:j+1}(\mathrm{d}z_{j+1}) \to \int_{\mathbb{R}^d} K(z_l, z_{j+1}) F_{j+1}(\mathrm{d}z_{j+1})$$

while for the last term

$$\frac{1}{N} K(z_l, z_k) F_{N:j} \to 0.$$

Formally, we have that the limit function satisfies

$$\partial_t F_j + \sum_{l=1}^{j} \mathrm{div}_{z_l} \int_{\mathbb{R}^d} K(z_l, z_{j+1}) F_{j+1}(\mathrm{d}z_{j+1}) = 0, \quad j \geq 1. \tag{7}$$

This string of equations is called the (infinite) Vlasov hierarchy and it looks already quite similar to the mean field PDE, which was our motivation in deriving it. The precise statement is given in the following

Proposition 9 *Under the assumptions of theorem 1 on K, let f^{in} be a smooth (at least C^1) probability density on $\mathbb{R}^d$ with finite first moment, i.e.*

$$\int_{\mathbb{R}^d} |z| f^{in}(z)\,\mathrm{d}z < \infty.$$

Assume that the Cauchy problem for the mean field equation (4) with initial data f^{in} has a classical (at least C^1) solution f. Set $f_j(t,\cdot) = f(t,\cdot)^{\otimes j}$ which means

$$f_j(t, z_1, \ldots, z_j) = \prod_{k=1}^{j} f(t, z_k)$$

for each $t \in \mathbb{R}$ and $z_1, \ldots, z_j \in \mathbb{R}^d$. Then the sequence $(f_j)_{j\geq 1}$ is a solution of the infinite mean field hierarchy

$$\partial_t f_j(z_1, \ldots, z_j) + \sum_{l=1}^{j} \mathrm{div}_{z_l} \int_{\mathbb{R}^d} K(z_l, z_{j+1}) f_{j+1}(z_1, \ldots, z_{j+1})\,\mathrm{d}z_{j+1} = 0, \quad j \geq 1.$$

Now we see how the method of hierarchies enables us to prove the mean field limit. Choosing factorized initial data for the N-particle Liouville equation, i.e. for a probability density f^{in} on $\mathbb{R}^d$ with finite first order moment, we look at the N-particle Liouville equation (6) with this initial data. Assume we can prove $F_{N:j} \to F_j$ (in some sense) for each $j \geq 1$, where F_j is a solution of the infinite hierarchy (7) and that we can prove the uniqueness of this solution.

Now let f be a solution of the mean field PDE (4) with initial data f^{in}. From proposition 9 we know that $f_j := f^{\otimes j}$ is a solution of the infinite mean field hierarchy (7) with initial data $(f^{\mathrm{in}})^{\otimes j}$ and (assuming uniqueness) we thus have

$$F_{N:j} \to F_j = f^{\otimes j} \quad \text{as } N \to \infty, j \geq 1.$$

For $j = 1$ in particular, we have

$$F_{N:1} \to f,$$

which says that the first marginal of the solution of the N-particle Liouville equation with factorized initial data converges to the solution of the mean field PDE in the large N limit.

Now we characterize the convergence of the marginals. Let p be a probability measure on $\mathbb{R}^d$. A sequence P_N of symmetric N-particle probability measures on $(\mathbb{R}^d)^N$ is said to be p-chaotic if

$$P_{N:j} \to p^{\otimes j} \quad \text{weakly in } \mathcal{P}((\mathbb{R}^d)^j)$$

as $N \to \infty$ for all $j \geq 1$ fixed. This convergence can be characterized using the empirical measure.

Theorem 10 *Let $p \in \mathcal{P}(\mathbb{R}^d)$ and $P_N \in \mathcal{P}_{sym}((\mathbb{R}^d)^N)$. Then the two following statements are equivalent:*

1. *P_N is p-chaotic;*
2. *for each $\phi \in C_b(\mathbb{R}^d)$ and each $\epsilon > 0$, we have*

$$P_N(\{Z_N \in (\mathbb{R}^d)^N : |\langle \mu_{Z_N} - p, \phi\rangle| \geq \epsilon\}) \to 0$$

as $N \to \infty$, where μ_{Z_N} is the empirical measure at the coordinates Z_N.

Theorem 11 *Under assumptions on K of theorem 1, take a probability density f^{in} on $\mathbb{R}^d$ such that*

$$\int_{\mathbb{R}^d} |z|^{d+5} f^{in}(z)\,\mathrm{d}z < \infty.$$

Let $F_N^{in} = (f^{in}\mathcal{L}^d)^{\otimes N}$ be the initial data to the N-particle Liouville equation (6) and denote its solution by $F_N(t) = T_t\#F_N^{in}$. Then we have, for each $j \geq 1$ that F_N is $f(t,\cdot)\mathcal{L}^d$-chaotic, i.e.

$$F_{N:j}(t) \to (f(t,\cdot)\mathcal{L}^d)^{\otimes j} \quad \textit{weakly in } \mathcal{P}((\mathbb{R}^d)^j)$$

as $N \to \infty$, where the probability density $f(t,\cdot)$ is the solution of the mean field PDE (4).

Theorem 12 (Horowitz-Karandikar [5]) *For all $p \in \mathcal{P}(\mathbb{R}^d)$ such that $a := \langle p, |z|^{d+5}\rangle < \infty$ we have that*

$$\int_{(\mathbb{R}^d)^N} W_2(\mu_{Z_N}, p)^2 p^{\otimes N}(\mathrm{d}Z_N) \leq \frac{C(a,d)^2}{N^{2/(d+4)}},$$

where the constant $C(a,d)$ depends on the constant a and the dimension d.

Theorem 13 *Under the assumptions of theorem 1 on K, let $F_N^{in} \in \mathcal{P}_{1,sym}((\mathbb{R}^d)^N)$ and $F_N(t)$ its time evolution via the N-particle Liouville equation (6). Then*

$$\int_{(\mathbb{R}^d)^N} \mu_{T_t Z_N^{in}}^{\otimes m} F_N^{in}(\mathrm{d}Z_N^{in}) = \frac{N!}{(N-m)!N^m} F_{N:m}(t) + R_{N,m}(t),$$

where $R_{N,m}(t)$ is a positive Radon measure on $(\mathbb{R}^d)^m$ with total mass

$$\langle R_{N,m}(t), 1\rangle = 1 - \frac{N!}{(N-m)!N^m} \leq \frac{m(m-1)}{2N}.$$

Furthermore, assuming the initial data is factorized, $F_N^{in} = (f^{in}\mathcal{L}^d)^{\otimes N}$, with

$$a = \int_{\mathbb{R}^d} |z|^{d+5} f^{in}(z)\,\mathrm{d}z < \infty,$$

we have

$$W_1(F_{N:1}(t), f(t)\mathcal{L}^d) \le \frac{C(a,d)\,\mathrm{e}^{2L|t|}}{N^{1/(d+4)}}.$$

Consider the notion of statistical solution. For the Cauchy problem of the ODE

$$\begin{cases} \dot{x}(t) = v(x(t)), \\ x(0) = x_0 \end{cases}$$

we have existence of a global solution flow X if the vector field v is Lipschitz by the standard theorems such that $t \mapsto X(t, x_0)$ is a solution to the ODE with initial condition $X(0, x_0) = x_0$. Now suppose instead of the initial condition x_0 we are given a probability distribution μ_0 set on the space of initial data $\mathbb{R}^d$. Then define $\mu(t) = X(t,\cdot)\#\mu_0$ for $t \in \mathbb{R}$. By the method of characteristics, we have that $\mu(t)$ solves the Cauchy problem

$$\begin{cases} \partial_t \mu(t) + \operatorname{div}(\mu(t)v) = 0, \\ \mu(0) = \mu_0. \end{cases}$$

In our case, the initial data is set on the space of probability measures with finite first order moments on $\mathbb{R}^d$, $\mathcal{P}_1(\mathbb{R}^d)$. So the notion of statistical solution would be to take initial data ν_0 from the probability measures on $\mathcal{P}_1(\mathbb{R}^d)$. We define the one-parameter group $G_t f^{\mathrm{in}} = f(t,\cdot)$. Then the time evolution of ν_0 is analogously defined as $\nu(t) = G_t\#\nu_0$. To determine the nature of $\nu(t)$, we write

$$\int_{\mathcal{P}_1(\mathbb{R}^d)} p^{\otimes j}\nu(t,\mathrm{d}p) = \int_{\mathcal{P}_1(\mathbb{R}^d)} (G_t p)^{\otimes j}\nu_0(\mathrm{d}p),$$

where we would like to test for appropriate continuous functions $\mathcal{F}$ on some functional space. Certainly, polynomials should be contained in such a class. Thus we consider monomials of degree k on $\mathcal{P}(\mathbb{R}^d)$:

$$M_{k,\phi}(p) := \int_{(\mathbb{R}^d)} \phi(z_1,\dots,z_k)p(\mathrm{d}z_1)\dots p(\mathrm{d}z_k) = \langle p^{\otimes k}, \phi\rangle.$$

The equality above becomes, if we specialize to $\mathcal{F} = M_{j,\phi}$

$$\int_{\mathcal{P}_1(\mathbb{R}^d)} p^{\otimes j}\nu(t,\mathrm{d}p) = \int_{\mathcal{P}_1(\mathbb{R}^d)} (G_t p)^{\otimes j}\nu_0(\mathrm{d}p).$$

Defining

$$F_j(t) := \int_{\mathcal{P}_1(\mathbb{R}^d)} (G_t p)^{\otimes j}\nu_0(\mathrm{d}p), \quad j \ge 1,$$

we have that F_j is a solution of the infinite mean field hierarchy (7): Since $(G_t p)^{\otimes j}$ is a solution of (7) and the hierarchy is a sequence of linear equations it follows that F_j is also a solution as it is an average under ν_0 of solutions.

Let $\Omega = (\mathbb{R}^d)^{\mathbb{N}}$ equipped with product topology and associated Borel algebra $\mathcal{B}(\Omega)$. Denote by $p^{\otimes\infty}$ the probability measure defined by

$$p^{\otimes\infty}\left(\prod_{k\geq 1} E_k\right) = \prod_{k\geq 1} p(E_k)$$

for each sequence E_k of Borel subsets of $\mathbb{R}^d$ such that $E_k = \mathbb{R}^d$ for all but finitely many k. Note $(p^{\otimes\infty})_{:j} = p^{\otimes j}$, $j \geq 1$. Define

$$\mathbf{F}(t) := \int_{\mathcal{P}_1(\mathbb{R}^d)} p^{\otimes\infty}\nu(t, \mathrm{d}p) = \int_{\mathcal{P}_1(\mathbb{R}^d)} (G_t p)^{\otimes\infty}\nu_0(\mathrm{d}p),$$

then for each $j \geq 1$ we have $\mathbf{F}(t)_{:j} = F_j(t)$ and $\mathbf{F}(t)_{:j}$ is a solution of the infinite mean field hierarchy. The natural extension regarding symmetric probability measures on Ω is defined by the condition

$$U_\sigma \# \mu = \mu, \quad U_\sigma(z_1, z_2, \dots) = (z_{\sigma(1)}, \dots, z_{\sigma(N)}, z_{N+1}, \dots).$$

We define a map $t \mapsto \mathbf{P}(t)$ to be a measure-valued solution of the mean field hierarchy if and only if

$$\partial_t \langle \mathbf{P}(t), \psi_j \rangle = \langle \mathbf{P}(t), \sum_{i=1}^{j} K(z_i, z_{j+1} \cdot \nabla_{z_i} \psi_j \rangle$$

for $t \in I \subset \mathbb{R}$, where I is an interval and $\psi_j \in C_0^1((\mathbb{R}^d)^j)$, $j \geq 1$. With this definition, $t \mapsto \mathbf{F}(t) \in \mathcal{P}_{\mathrm{sym}}(\Omega)$ defined before is a measure-valued solution of the infinite mean field hierarchy with initial condition

$$\mathbf{F}(0) = \int_{\mathcal{P}_1(\mathbb{R}^d)} p^{\otimes j}\nu_0(\mathrm{d}p).$$

Two questions come to mind:

1. Are these measure-valued solutions uniquely determined by the initial data?
2. Are all measure-valued solutions of the mean field hierarchy defined by statistical solutions of the mean field PDE?

The answers are given by the two following theorems.

Theorem 14 (Hewitt-Savage [4]) *For each $\mathbf{P} \in \mathbf{P}_{sym}(\Omega)$, there exists a unique probability measure π on $\mathcal{P}(\mathbb{R}^d)$ such that*

$$\mathbf{P} = \int_{\mathcal{P}(\mathbb{R}^d)} p^{\otimes\infty}\pi(\mathrm{d}p).$$

Theorem 15 (Spohn [6]) *Under the assumptions of theorem 1 on the interaction kernel K, for each probability measure ν^{in} on $\mathcal{P}_1(\mathbb{R}^d)$, the only measure-valued solution of the Vlasov hierarchy with initial data*

$$\mathbf{F}^{in} = \int_{\mathcal{P}_1(\mathbb{R}^d)} p^{\otimes\infty} \nu^{in}(\mathrm{d}p)$$

is

$$\mathbf{F}(t) = \int_{\mathcal{P}_1(\mathbb{R}^d)} (G_t p)^{\otimes\infty} \nu^{in}(\mathrm{d}p), \quad t \in \mathbb{R}.$$

For $P \in \mathcal{P}_{1,\mathrm{sym}}((\mathbb{R}^d)^M)$ and $Q \in \mathcal{P}_{1,\mathrm{sym}}((\mathbb{R}^d)^N)$ we define a variant of the Wasserstein distance by

$$\mathcal{W}_1(P,Q) = \inf_{\rho\in\Pi(P,Q)} \int_{\mathbb{R}^{dM}\times\mathbb{R}^{dN}} W_1(\mu_{X_M}, \mu_{Y_N}) \rho(\mathrm{d}X_M, \mathrm{d}Y_N).$$

In the same setting as Spohn's theorem, we can find a stability estimate for solutions of the N-particle Liouville equation with statistical initial data:

Theorem 16 (Golse-Mouhot-Ricci [2]) *For $M, N \geq 1$ let $P_M^{in} \in \mathcal{P}_{1,sym}((\mathbb{R}^d)^M)$ and $Q_N^{in} \in \mathcal{P}_{1,sym}((\mathbb{R}^d)^N)$ and assume that the interaction kernel satisfies the assumptions of theorem 1. Denote by $P_M(t)$ and $Q_N(t)$ the solutions to the M- and N-particle Liouville equations with initial data P_M^{in} and Q_N^{in} respectively. Then*

1. *For each $t \in \mathbb{R}$, one has*

$$\mathrm{D}_{MK,1}(P_M(t), Q_N(t)) \leq \mathrm{e}^{2L|t|} \mathrm{D}_{MK,1}(P_M^{in}, Q_N^{in}).$$

2. *For each $t \in \mathbb{R}$, $m, M, N \in \mathbb{N}$ such that $M, N \geq m$ and for each bounded Lipschitz continuous function ϕ_m on $(\mathbb{R}^d)^m$, one has*

$$\begin{aligned} &|\langle P_{M:m}(t) - Q_{N:m}(t), \phi_m\rangle| \\ &\quad\leq m\,\mathrm{e}^{2L|t|}\,\mathrm{Lip}(\phi_m)\mathrm{D}_{MK,1}(P_M^{in}, Q_N^{in}) + m(m-1)\|\phi_m\|_\infty \left(\frac{1}{M} + \frac{1}{N}\right). \end{aligned}$$

3 The case of singular interaction kernels: Coulomb and gravitational force

3.1 Main example: Coulomb/gravitational force

A well-known example for the interaction force is the Coulomb or the graviational force which is given by

$$F(x) = C\frac{x}{|x|^d}, \quad d \neq 2.$$

In two dimensions, the force is logarithmic in the distance, $F(x) = C\log(|x|)$. This comes from the fact that the Coulomb force is the negative gradient of the Coulomb potential which in turn is defined as the Green's function of the Laplacian:

$$V(x) = C\frac{1}{|x|^{d-2}}, \quad -\Delta V = \delta.$$

For positive C, this corresponds to the Coulomb force of charges with equal sign and for negative C it corresponds to gravitational interaction. The particles considered in this system could thus be ions or electrons in a plasma or stars (or clusters of stars, galaxies, ...). As the number of particles in such systems usually is very large (10^9 to 10^{25}), solving (2) numerically is already challenging. Notice that $F(x)$ is singular at the origin.

The singularity at the origin poses a problem if one attempts to obtain the same results as in section 2 as the Gronwall estimates are no longer as easy to establish. Instead we have to work with the structure of the equation itself and look at the trajectories of particles that get close to each other in position-velocity space.

3.2 Mean field limit

We want to quantify the strength of the singularity in the force. Therefore we define two different conditions for "weakly" and "strongly" singular forces. F is weakly singular if it satisfies the condition (S^α) for $0 < \alpha < 1$:

$$(S^\alpha) \quad \exists C > 0 \quad \forall x \in \mathbb{R}^d\backslash\{0\} \quad |F(x)| \leq \frac{C}{|x|^\alpha}, \, |\nabla F(x)| \leq \frac{C}{|x|^{\alpha+1}}.$$

Strongly singular forces will be cut off at the origin. We say that F satisfies (S^α_m) if

1. F satisfies (S^α) for $1 \leq \alpha < d-1$,
2. For $|x| \geq N^{-m}$ we have $F_N(x) = F(x)$,
3. For $|x| \leq N^{-m}$ we have $|F_N(x)| \leq N^{m\alpha}$.

Theorem 17 (Hauray-Jabin [3]) *Assume the interaction force satisfies an (S^α) condition for $\alpha < 1$ and the initial data f^{in} is such that the Vlasov equation has a unique solution $f(t)$. Assume further that the initial conditions Z_N^{in} satisfy some compatibility conditions and denote $Z_N(t)$ the unique solution to the ODE system. Then for $N \geq \mathrm{e}^{cT}$*

$$W_1(\mu_{Z_N}(t), f(t)) \leq \mathrm{e}^{Ct}\left(W_1(\mu_{Z_N^{in}}, f^{in}) + 2N^{-\frac{\gamma}{2d}}\right),$$

where $\gamma \in (0,1)$ and $t \in [0,T]$; c and C denote positive constants.

The exact compatibility conditions on the initial data Z_N^{in} are

1. The local density does not exceed a certain bound,

$$\sup_{z\in\mathbb{R}^{2d}} N^\gamma \mu_{Z_N^{\mathrm{in}}}(B_{2d}(z, N^{-\gamma/(2d)})) \leq C$$

2. The coordinates are contained in a ball of finite radius $R > 0$,

$$\operatorname{supp} \mu_{Z_N^{\mathrm{in}}} \subset B_{2d}(0, R).$$

3. The initial interparticle distance is bounded below,

$$\inf_{i \neq j} |z_i^{\mathrm{in}} - z_j^{\mathrm{in}}| \geq N^{-\gamma(1+r)/(2d)}.$$

Here, we want to give some ideas for the proof of theorem 17:

- Use the Wasserstein distance of order ∞:

$$W_\infty(\mu, \nu) = \inf_{\pi \in \Pi(\mu,\nu)} \pi\text{-ess sup}\, |x - y|$$

  Since $W_1 \leq W_\infty$, a control on the infinite Wasserstein distance will imply control on the one Wasserstein distance

- Introduce a scale $\epsilon(N) = N^{-\gamma/(2d)}$ for $\gamma \in (0, 1)$. Notice that ϵ is larger than the average inter-particle distance of order $N^{-1/(2d)}$

- Use this scale to distinguish contributions from three different domains

- Define the minimal inter-particle distance

$$d_N(t) = \inf_{i \neq j} |z_i(t) - z_j(t)|.$$

Then we have to distinguish contributions from three different domains (compare figure 1). They are

1. Particles that are sufficiently far away from each other in position space
2. Particles that are ϵ-close in position space but have sufficiently different velocities
3. Particles that are ϵ-close in $\mathbb{R}^{2d}$

Carefully summing up contributions from the different domains gives two differential inequalities (for rescaled versions $\tilde{W}_\infty$ and $\tilde{d}_n$ of W_∞ and d_N that are of order one)

$$\begin{aligned} \frac{\tilde{W}_\infty(t) - \tilde{W}_\infty(t-\epsilon)}{\epsilon} &\leq C(\tilde{W}_\infty(t) + \mathrm{Rest}), \\ \frac{\tilde{d}_N(t) - \tilde{d}_N(t-\epsilon)}{\epsilon} &\geq -C(\tilde{d}_N(t) + \mathrm{Rest}). \end{aligned}$$

The terms in “Rest” come with small weights ϵ^β for positive β such that these inequalities provide uniform bounds until a critical time T_ϵ with $T_\epsilon \to \infty$ as $\epsilon \to 0$.

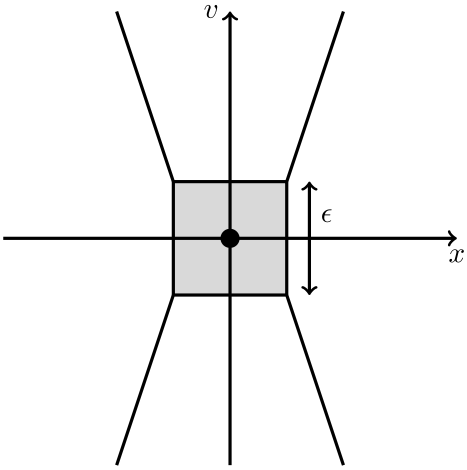

Figure 1: Imagine one particle sitting at the origin and the coordinates in this drawing give the distance to another particle. Then we distinguish contributions from three different domains.

3.3 Propagation of chaos

Similarly as before, we can prove propagation of chaos from the deterministic mean field limit. More delicate here: Need to show that the compatibility conditions on the initial data are satisfied with large probability in the limit when the initial data is chosen randomly with law $(f^{\text{in}})^{\otimes N}$. To conclude we use “large deviation bounds” from statistical theory.

References

[1] R. L. Dobrushin. “Vlasov equations”. In: *Func. Anal. Appl.* **13** (1979), pp. 115–223.

[2] F. Golse, C. Mouhot, and V. Ricci. *Empirical measures and Vlasov hierarchies.* 2013. arXiv: `1309.0222`.

[3] M. Hauray and P.-E. Jabin. *Particle approximation of Vlasov equations with singular forces: propagation of chaos.* 2014. arXiv: `1107.382`.

[4] E. Hewitt and L. J. Savage. “Symmetric measures on Cartesian products”. In: *Trans. Amer. Math. Soc.* (1955).

[5] J. Horowitz and R. Karandikar. “Mean rates of convergence of empirical measures in the Wasserstein metric”. In: *J. Comput. Appl. Math.* **55** (1994).

[6] H. Spohn. “On the Vlasov hierarchy”. In: *Math. Meth. in the Appl. Sci.* **3** (1981), pp. 445–455.

[7] C. Villani. *Topics in Optimal Transportation.* Providence, RI: American Mathematical Society, 2003.