MEAN FIELD PARTICLE SYSTEMS AND THEIR LIMITS TO NONLOCAL PD'S

Li Chen

October 18, 2023

Abstract

This lecture aims to give an introduction on the mean field derivation of a family
of non-local partial differential equations with and without diffusion

Chapter 1

Model description and Introduction

The following chapter will outline how the relevant particle models are defined, we differentiate between first and second order systems focusing here on first order systems while leaving the second order setting as exercises

1.1   1st Order Particle Systems

Definition 1.1.1 (1st Order Particle System). We consider a system of N particles and denote by $(x_1(t), x_2(t), \ldots, x_N(t)) \in \mathcal{C}^1([0, T]; \mathbb{R}^d)$, $i = 1, \ldots, N$ the trajectories of the particles.

Our first order system is then governed by the system of ordinary differential equations

$$\begin{cases} dx_i(t) & = \frac{1}{N} \sum_{j=1}^{N} K(x_i, x_j) dt + \sigma dW_i(t), \quad 1 \leq i \leq N \\ x_i(t)|_{t=0} & = x_i(0) \end{cases}.$$

where $K : \mathbb{R}^{2d} \to \mathbb{R}^d$ is a given function.

For $\sigma = 0$ we say the system is deterministic

We consider the following examples for K

Example. A common example for a well-behaved K is

$$K(x, y) = \nabla(|x - y|^2).$$

which is a locally Lipschitz continuous function.

Another typical interaction force which is not continuous is the potential

field given by Coulomb potential, namely

$$K(x, y) = \nabla \frac{1}{|x - y|^{d-2}} = \frac{x - y}{|x - y|^d}.$$

Definition 1.1.2 (Empirical Measure). For a set of particles
$(x_1(t), x_2(t), \ldots, x_N(t)) \in \mathcal{C}^1([0, T]; \mathbb{R}^d)$, $i = 1, \ldots, N$ we define the empirical measure by

$$\mu^N(t) \triangleq \frac{1}{N} \sum_{j=1}^{N} \delta_{x_i(t)}.$$

Our goal is the study of the limit of this system as $N \to \infty$. An appropriate quantity is to consider the empirical measure 1.1.2. If the initial empirical measure converges in some sense to a measure $\mu(0)$ i.e.

$$\mu^N(0) \to \mu(0).$$

would $\mu^N(t)$ also converge to some measure $\mu(t)$?

$$\mu^N(t) \xrightarrow{?} \mu(t).$$

Furthermore, can we find an equation which $\mu(t)$ satisfies and in which sense does it satisfy this equation?

Note. Consider the following case when the limit measure $\mu(t)$ is absolutely continuous with respect the Lebesgue measure, this means that

$$d\mu(0, x) = \rho_0(x)dx \quad \rho_0 \in L^1(\mathbb{R}^d).$$

would the limit function have the same property ?

1.2   Motivation For Partial Differential Equation

Let the following Proposition serve as a motivation on which partial differential equation $\mu(t)$ should satisfy and consider only the deterministic case for now.

Proposition 1.2.1. We say $\mu(t)$ solves the following partial differential equation (in the sense of distribution)

$$\partial_t \mu(t,x) + \nabla \cdot \left(\mu(t,x) \int_{\mathbb{R}^d} K(\cdot,y) d\mu(t,y) \right) = 0.$$

Proof. Take $\varphi \in \mathcal{C}_0^\infty(\mathbb{R}^d)$ and calculate

$$\frac{d}{dt} \langle \mu^N(t), \varphi \rangle \triangleq \frac{d}{dt} \int_{\mathbb{R}^d} \varphi(x) d\mu^N(t,x)$$

$$\stackrel{\text{Def.}}{=} \frac{d}{dt} \int_{\mathbb{R}^d} \frac{1}{N} \sum_{j=1}^N \varphi(x) d\delta_{x_i(t)}$$

$$\stackrel{\text{Lin.}}{=} \frac{1}{N} \sum_{j=1}^N \frac{d}{dt} \varphi(x_i(t))$$

$$= \frac{1}{N} \sum_{j=1}^N \nabla \varphi(x_i(t)) \cdot \frac{d}{dt} x_i(t)$$

$$= \frac{1}{N} \sum_{j=1}^N \nabla \varphi(x_i(t)) \cdot \frac{1}{N} \sum_{j=1}^N K(x_i, x_j)$$

$$= \frac{1}{N} \sum_{j=1}^N \nabla \varphi(x_i(t)) \cdot \frac{1}{N} \sum_{j=1}^N \int_{\mathbb{R}^d} K(x_i, y) d\delta_{x_i(t)}(y)$$

$$\stackrel{\text{Emp.}}{=} \frac{1}{N} \sum_{j=1}^N \nabla \varphi(x_i(t)) \cdot \int_{\mathbb{R}^d} K(x_i, y) d\mu^N(t,y)$$

$$= \frac{1}{N} \sum_{j=1}^N \int_{\mathbb{R}^d} \nabla \varphi(x) \cdot \int_{\mathbb{R}^d} K(x,y) d\mu^N(t,y) d\delta_{x_i(t)}(x)$$

$$= \int_{\mathbb{R}^d} \nabla \varphi(x) \cdot \int_{\mathbb{R}^d} K(x,y) d\mu^N(t,y) d\mu^N(t,x)$$

$$= -\left\langle \nabla \cdot \left(\mu^N(t,\cdot) \int_{\mathbb{R}^d} K(\cdot,y) d\mu^N(t,y) \right), \varphi \right\rangle.$$

i.e μ^N is a solution to

$$\partial_t \mu^N(t,x) + \nabla \cdot \left(\mu^N(t,x) \int_{\mathbb{R}^d} K(\cdot,y) d\mu^N(t,y) \right) = 0.$$

If we can now take the limit $N \to \infty$ we obtain that μ should satisfy the proposed PDE $\qquad \square$

Corollary. If $\sigma > 0$ i.e our system is stochastic then we expect the limit partial differential equation to share a similar structure

$$\partial_t \mu(t,x) + \nabla \cdot \left(\mu(t,x) \int_{\mathbb{R}^d} K(\cdot, y) d\mu(t,y) \right) = \Delta \mu(t,x).$$

We define the stochastic case in detail later

1.3   2nd Order Particle Systems

We define a second order particle system as follows

Definition 1.3.1. Given the N particles

$$((x_1(t), v_1(t)), \ldots, (x_N(t), v_N(t))) \in \mathcal{C}^1([0,T]; \mathbb{R}^{2d}).$$

with initial values $x_i(0)$ for $i = 1, \ldots, N$

Then our second order system is then governed by

$$(\text{MPS}) \begin{cases} \frac{d}{dt} x_i(t) & = v_i(t) \\ \frac{d}{dt} v_i(t) & = \frac{1}{N} \sum_{j=1}^N F(x_i(t), v_i(t); x_j(t), v_j(t)) \quad 1 \le i \le N \end{cases}.$$

In this setting $(x_i(t), v_i(t))$ mean the position and velocity of the i-th particle respectively. An example for F would be

$$F(x, v; y, u) = \frac{x-y}{|x-y|^d}.$$

The empirical measure from Definition 1.1.2 can be rewritten to include the velocity as well

$$\mu^N \triangleq \frac{1}{N} \sum_{j=1}^N \delta_{x_i(t), v_i(t)}.$$

Exercise. Calculate for $\forall \varphi \in \mathcal{C}_0^\infty(\mathbb{R}^{2d})$ the following in the second order case

$$\frac{d}{dt} \langle \mu^N(t), \varphi \rangle.$$

1.4 Lecture Structure

In Chapter 1, we are going to discuss the deterministic case for "Good" interaction forces (2-3 weeks) while giving a brief review of the well-posedness theory of ordinary differential equation. And prove the mean field limit in the framework of 1-Wasserstein distance.

The stochastic case will be studied in Chapter 2. Where we first review the mandatory concepts of probability theory, the definition of the Itô integral, and the well-posedness of stochastic differential equations. Then the propagation of chaos result of the interacting SDE system is studied, where the well-posedness of Mckean-Vlasov equation plays an important role. If time allows, we will study non-smooth interaction forces in chapter 3.

The first result is the convergence in probability, which implies the weak convergence of propagation of chaos. The second topic is to introduce the relative entropy method to get the convengence in L^1 space.

Chapter 2

MEAN-FIELD LIMIT IN THE DETERMINISTIC SETTING

In this chapter we focus on the deterministic version of the mean-field limit. Namely, we start from a system of deterministic interacting particle system with mean-field structure and prove that the corresponding empirical measure converges weakly to the measure valued solution of the corresponding partial differential equation. We are going to work only with the first order system, recall

> **Definition** (1st Order Particle System). We consider a system of N particles and denote by $(x_1(t), x_2(t), \ldots, x_N(t)) \in \mathcal{C}^1([0, T]; \mathbb{R}^d)$, $i = 1, \ldots, N$ the trajectories of the particles.
>
> Our first order system is then governed by the system of ordinary differential equations
>
> $$\begin{cases} dx_i(t) & = \frac{1}{N} \sum_{j=1}^{N} K(x_i, x_j) dt \quad 1 \leq i \leq N \\ x_i(t)|_{t=0} & = x_i(0) \in \mathbb{R}^d \end{cases}.$$
>
> where $K : \mathbb{R}^{2d} \to \mathbb{R}^d$ is a given function.

In the case of higher dimensional vectors we sometimes use the following notation

$$X_N(t) = \big(x_1(t), x_2(t), \ldots, x_N(t)\big)^T \in \mathbb{R}^{dN}.$$

2.1   Review Of ODE Theory

Definition 2.1.1 (Intial Value Problem). For $\forall T > 0$ and $f \ [0, T] \times \mathbb{R}^d \to \mathbb{R}^d$ we consider the initial value problem given by

$$(\text{IVP}) \begin{cases} \frac{d}{dt}x(t) &= f(t, x) \quad t \in [0, T] \\ x|_{t=0} &= x_0 \in \mathbb{R}^d \end{cases}.$$

Assumption A. $f \in \mathcal{C}([0, T] \times \mathbb{R}^d; \mathbb{R}^d)$ and f is Lipschitz continuous in x, which means there $\exists L > 0$ such that $\forall (t, x), (t, y) \in [0, T] \times \mathbb{R}^d$

$$|f(t, x) - f(t, y)| \le L|x - y|.$$

Theorem 2.1.1 (Existence and uniqueness of IVP). If Assumption A holds then the (IVP) has a unique solution $x \in \mathcal{C}^1([0, T]; \mathbb{R}^d)$

Proof. We use Picard iteration to prove the existence, we can define the equivalent way of solving the (IVP) by considering the integral equation

$$x(t) - x_0 = \int_0^t f(s, x(s))ds \quad \forall t \in [0, T].$$

Then our Picard iteration is given by the following

$$x_1(t) = x_0 + \int_0^t f(s, x_0)ds$$

$$x_2(t) = x_0 + \int_0^t f(s, x_1(s))ds$$

$$\vdots$$

$$x_m(t) = x_0 + \int_0^t f(s, x_{m-1}(s))ds.$$

By Assumption A and properties of integration we have $x_m(t) \in \mathcal{C}^1([0, T]; \mathbb{R}^d)$.

Due to completeness of $\mathcal{C}^1([0, T]; \mathbb{R}^d)$ we only need to show that $(x_m(t))_{m \in \mathbb{N}}$ is a Cauchy sequence to get the existence. We first prove by induction that for $m \ge 2$ it holds for some constant M that

$$|x_m(t) - x_{m-1}(t)| \le \frac{ML^{m-1}|t|^m}{m!}.$$

IA For $m = 1$ it holds

$$|x_2(t) - x_1(t)| \overset{\text{Tri.}}{\leq} \int_0^t |f(s, x_1(s)) - f(s, x_0)| ds$$

$$\leq L \int_0^t |x_1(s_0) - x_0| ds_0$$

$$\leq L \int_0^t \int_0^{s_0} |f(s_1, x_0)| ds_1 ds_0$$

$$\leq ML \int_0^t (s_0 - 0) ds_0$$

$$= \frac{MLt^2}{2}.$$

where we chose $M \geq \max_{s \in [0,T]} |f(s, x_0)|$

IV Suppose for $m \in \mathbb{N}$ it holds

$$|x_m(t) - x_{m-1}(t)| \leq \frac{ML^{m-1}|t|^m}{m!}.$$

IS $m \to m + 1$

$$|x_{m+1}(t) - x_m(t)| = \left| \int_0^t f(s, x_m(s)) - f(s, x_{m-1}(s)) ds \right|$$

$$\overset{\text{Tri.}}{\leq} \int_0^t |f(s, x_m(s)) - f(s, x_{m-1})(s)| ds$$

$$\leq L \int_0^t |x_m(s) - x_{m-1}(s)| ds$$

$$\overset{\text{IV}}{\leq} L \int_0^t \frac{ML^{m-1}|s|^m}{m!} ds$$

$$= \frac{ML^m |t|^{m+1}}{(m+1)!}.$$

Now take arbitrary $p, m \in \mathbb{N}$ then by triangle inequality we obtain for $\forall t \in [0, T]$ that

$$|x_{m+p} - x_m(t)| \leq \sum_{k=m+1}^{m+p} |x_k(t) - x_{k-1}(t)|$$

$$\leq \sum_{k=m+1}^{m+p} M \frac{L^{k-1} T^k}{k!}$$

$$= \frac{M}{L} \sum_{k=m+1}^{m+p} \frac{(LT)^k}{k!}$$

$$.$$

Continuing on the next page

$$\frac{M}{L} \sum_{k=m+1}^{m+p} \frac{(LT)^k}{k!} \leq \frac{M}{L} \frac{(LT)^{m+1}}{(m+1)!} \sum_{k=0}^{p-1} \frac{(LT)^k}{k!}$$

$$\leq \frac{M}{L} \frac{(LT)^{m+1}}{(m+1)!} \sum_{k=0}^{\infty} \frac{(LT)^k}{k!}$$

$$= \frac{M}{L} \frac{(LT)^{m+1}}{(m+1)!} e^{LT} \xrightarrow{m\to\infty} 0 \text{ uniformly in } t \in [0,T].$$

Therefore $x_m(t)$ has a limit $x(t) \in \mathcal{C}^1([0,T];\mathbb{R}^d)$ with

$$\max_{t\in[0,T]} |x_m(t) - x(t)| \xrightarrow{m\to\infty} 0.$$

Then by taking $m \to \infty$ in

$$x_m(t) = x_0 + \int_{t_0}^{t} f(s, x_{m-1}(s))ds.$$

we get

$$x(t) = x_0 + \int_{t_0}^{t} f(s, x(s))ds.$$

which means $x(t) \in \mathcal{C}^1([0,T];\mathbb{R}^d)$ is a solution of the equivalent integral equation.

To prove the uniqueness suppose we have two solutions $x(t), \tilde{x}(t) \in \mathcal{C}^1([0,T];\mathbb{R}^d)$ then they satisfy

$$x(t) = x_0 + \int_{t_0}^{t} f(s, x(s))ds$$

$$\tilde{x}(t) = x_0 + \int_{t_0}^{t} f(s, \tilde{x}(s))ds.$$

By taking the difference of these two solutions and using the Lipschitz continuity of f in x we obtain

$$|x(t) - \tilde{x}(t)| \leq \int_{0}^{t} |f(s, x(s)) - f(s, x(\tilde{s}))|ds + |x_0 - x_0|$$

$$\leq L \int_{)}^{t} |x(s) - \tilde{x}(s)|ds$$

$$\leq L \int_{0}^{t} e^{-\alpha s} |x(s) - \tilde{x}(s)| e^{\alpha s} ds.$$

For any $\alpha > 0$. By considering the quantity $P(t) = e^{-\alpha t}|x(t) - \tilde{x}(t)|$, we

obtain

$$|x(t) - \tilde{x}(t)| \leq L \int_0^t \max_{0 \leq s \leq t} \{e^{-\alpha s}|x(s) - \tilde{x}(s)|\} e^{\alpha s} ds$$

$$\leq L \max_{0 \leq s \leq t} \{e^{-\alpha s}|x(s) - \tilde{x}(s)|\} \int_0^t e^{\alpha s} ds.$$

We obtain

$$P(t) = e^{-\alpha t}|x(t) - \tilde{x}(t)| \leq \max_{t \in [0,T]} P(t) \leq \frac{L}{\alpha} \max_{t \in [0,T]} P(t) \quad \forall t \in [0, T].$$

By choosing $\alpha = 2L$ we have

$$\max_{t \in [0,T]} e^{-2Lt}|x(t) - \tilde{x}(t)| = 0.$$

i.e

$$x(t) = \tilde{x}(t) \quad \forall t \in [0, T].$$

This concludes the uniqueness proof $\qquad \square$

Remark. An alternative proof for uniqueness uses Gronwall's inequality which we give in the following. Furthermore similar to the uniqueness proof, one can obtain that the solution $x(t; t_0, x_0)$ is continuously dependent on initial data

Lemma 2.1.1 (Gronwall's inequality). Let $\alpha, \beta, \varphi \in \mathcal{C}([a, b]; \mathbb{R}^d)$ and $\beta(t) \geq 0$ for $\forall t \in [a, b]$ such that

$$0 \leq \varphi(t) \leq \alpha(t) + \int_a^t b(s)\varphi(s)ds \quad \forall t \in [a, b].$$

then

$$\varphi(t) \leq \alpha(t) + \int_a^t \beta(s) e^{\int_s^t \beta(\tau)d\tau} \alpha(s)ds \quad \forall t \in [a, b].$$

Specially if $\alpha(t) \equiv M$ then we have

$$\varphi(t) \leq M e^{\int_a^b \beta(\tau)d\tau} \quad \forall t \in [a, b].$$

Proof. Define

$$\psi(t) = \int_a^t \beta(\tau)\varphi(\tau)d\tau \quad \forall t \in [a, b].$$

because of the continuity of β and φ we get that ψ is differentiable on $[a, b]$ and

$$\psi'(t) = \beta(t)\varphi(t).$$

Since $\beta(t) \geq$ we have

$$\psi'(t) = \beta(t)\varphi(t) \leq \beta(t)(\alpha(t) + \psi(t)) \quad \forall t \in [a, b].$$

Then by multiplying both sides with $e^{-\int_a^t \beta(\tau)d\tau}$ we obtain

$$\frac{d}{dt}(e^{-\int_a^t \beta(\tau)d\tau}\psi(t)) = e^{-\int_a^t \beta(\tau)d\tau}(\psi'(t) - \beta(t)\psi(t))$$

$$\leq \beta(t)\alpha(t)e^{-\int_a^t \beta(\tau)d\tau}.$$

Integrate the above inequality from a to t to get

$$e^{-\int_a^t \beta(\tau)d\tau}\psi(t) - e^{-\int_a^t \beta(\tau)d\tau}\psi(a) \leq \int_a^t \beta(s)\alpha(s)e^{-\int_a^s \beta(\tau)d\tau}ds.$$

Which implies

$$\psi(t) \leq \int_a^t \beta(s)\alpha(s)e^{\int_s^t \beta(\tau)d\tau}ds.$$

and

$$\varphi(t) \leq \alpha(t) + \psi(t) \leq \alpha(t) + \int_a^t \beta(s)\alpha(s)e^{\int_s^t \beta(\tau)d\tau}ds.$$

The case with $\alpha(t) \equiv M$ is handled by using the main theorem of Differential and Integral calculus

$$\varphi(t) \leq M\left(1 + \int_a^t \beta(s)e^{\int_s^t \beta(\tau)d\tau}ds\right)$$

$$= M(1 - e^{\int_s^t \beta(\tau)d\tau}|_a^t)$$

$$= Me^{\int_a^t \beta(\tau)d\tau}.$$

$\square$

2.2 Mean-field particle system, well-posedness and problem setting

Let us again give the model and problem setting

> **Definition 2.2.1** (1st Order Particle System)**.** We consider a system of N particles and denote by $(x_1(t), x_2(t), \ldots, x_N(t)) \in \mathcal{C}^1([0,T];\mathbb{R}^d)$, $i = 1, \ldots, N$ the trajectories of the particles.
>
> Our first order system is then governed by the system of ordinary differential equations
>
> $$\text{(MPS)} \begin{cases} dx_i(t) & = \frac{1}{N}\sum_{j=1}^{N} K(x_i, x_j)dt \quad 1 \le i \le N \\ x_i(t)|_{t=0} & = x_i(0) \in \mathbb{R}^d \end{cases}.$$
>
> where $K : \mathbb{R}^{2d} \to \mathbb{R}^d$ is anti-symmetric i.e
>
> $$K(x,y) = -K(y,x) \quad K(x,x) = 0.$$

> **Assumption B.** $K \in \mathcal{C}^1(\mathbb{R}^d \times \mathbb{R}^d; \mathbb{R}^d)$ and there exists some $L > 0$ such that $\forall x, y \in \mathbb{R}^d$ it holds
>
> $$\sup_y |\nabla_x K(x,y)| + \sup_x |\nabla_y K(x,y)| \le L.$$

> **Lemma 2.2.1.** When Assumption B holds for K then for $\forall T > 0$ the (MPS) has a unique solution
>
> $$X_N(t) = (x_1(t), x_2(t), \ldots, x_N(t)) \in \mathcal{C}^1([0,T];\mathbb{R}^{dN}).$$
>
> and for any fixed $t \in [0,T]$ the map
>
> $$X_N(t, \cdot) : \mathbb{R}^{dN} \to \mathbb{R}^{dN} \; : \; x \mapsto X_N(t,x)$$
>
> is a bijection

In the introduction we saw that the empirical measure satisfies a partial differential equation

Definition 2.2.2 (PDE Problem)**.** Let $\mu^N(t)$ be the empirical measure

$$\mu^N(t) \triangleq \frac{1}{N} \sum_{j=1}^{N} \delta_{x_i(t)}.$$

Then from the introduction we know that for $\forall \varphi \in \mathcal{C}_0^\infty(\mathbb{R}^d)$ the empirical measure satisfies

$$\frac{d}{dt}\langle \mu^N(t), \varphi \rangle = \langle \mu^N(t), \nabla \varphi \cdot \mathcal{K}\mu^N(t) \rangle.$$

where

$$\mathcal{K}\mu^N(\cdot) = \int_{\mathbb{R}^d} K(\cdot, y) d\mu^N(y).$$

Idea. If $\mu^N \to \mu$ in some sense, then the limiting measure μ should also satisfy

$$\begin{cases} \partial_t \mu + \nabla \cdot (\mu \mathcal{K}\mu) = 0 \\ \mu^N(0) \to \mu_0 \end{cases}.$$

in the sense weak sense i.e. the "sense of distributions" which we define in the following section

2.3   A short introduction for Distributions

Definition 2.3.1. Let $\Omega \subset \mathbb{R}^d$ be an open subset then the space of test functions $\mathcal{D}(\Omega)$ consists of all the functions in $\mathcal{C}_0^\infty(\Omega)$ supplemented by the following convergence

We say $\varphi_m \to \varphi \in \mathcal{C}_0^\infty(\Omega)$ iff

1. There exists a compact set $\exists K \subset \Omega$ such that $\operatorname{supp}\varphi_m \subset K$ for $\forall m$

2. For all multi indices α it holds

$$\sup_K |\partial^\alpha \varphi_m - \partial^\alpha \varphi| \xrightarrow{m \to \infty} 0.$$

Remark. $\mathcal{D}(\Omega)$ is a linear space

Definition 2.3.2 (Multi-Index)**.** A multi-index $\alpha \in \mathbb{N}_0^n$ of length $|\alpha| = \sum_i \alpha_i$ for example $\alpha = (0, 2, 1) \in \mathbb{N}_0^3$ can be used to denote partial derivatives of higher order as such :

$$\partial^\alpha = \prod_i (\frac{\partial}{\partial x_i})^{\alpha_i}.$$

Definition 2.3.3 (Distribution)**.** The space of Distributions is denoted by $\mathcal{D}'(\Omega)$ and is the dual space of $\mathcal{D}(\Omega)$ i.e. it is the linear space of all continuous linear functions on $\mathcal{D}(\Omega)$

We say a functional $T : \mathcal{D}(\Omega) \to \mathbb{C}$ is continuous linear iff

1. $\langle T, \alpha\varphi_1 + \beta\varphi_2 \rangle = \alpha\langle T, \varphi_1 \rangle + \beta\langle T, \varphi_2 \rangle$

2. If $\varphi_m \to \varphi$ in $\mathcal{D}(\Omega)$ then $\langle T, \varphi_m \rangle \to \langle T, \varphi \rangle$

We can define several operations on the space of distributions but since most of them are not used in this Lecture we only define the multiplication with a smooth function

Definition 2.3.4. For a smooth function $f \in \mathcal{C}^\infty$ and a distribution $T \in \mathcal{D}'$ the product is defined as follows

$$\langle Tf, \varphi \rangle = \langle T, f\varphi \rangle \quad \forall \varphi \in \mathcal{D}.$$

Remark. Multiplication between two Distributions $T, F \in \mathcal{D}'$ is not well defined, instead the convolution of two Distributions is defined

Example. For functions $f \in L^1_{\text{loc}}(\Omega)$ we can define the associated distribution $T_f \in \mathcal{D}'(\Omega)$ is defined by

$$\langle T_f, \varphi \rangle = \int_\Omega f(x)\varphi(x)dx \quad \forall \varphi \in \mathcal{D}(\Omega).$$

and say $L^1_{\text{loc}}(\Omega) \subset \mathcal{D}'(\Omega)$

Similarly $L^p_{\text{loc}} \subset \mathcal{D}'(\Omega)$, using Hölder's inequality one obtains $L^p_{\text{loc}}(\Omega) \subset L^q_{\text{loc}}(\Omega)$ for $1 < q < p < \infty$

Remark. The support of a distribution is also well-defined

Theorem 2.3.1. L^1_{loc} functions are uniquely deterimed by distributions. More precisely for two functions $f, g \in L^1_{\text{loc}}(\Omega)$ if

$$\int_\Omega f\varphi dx = \int_\Omega g\varphi dx \quad \forall \varphi \in \mathcal{D}(\Omega).$$

then $f = g$ a.e. in Ω

Proof. This proof is left as an exercise $\qquad\square$

Example. The set of probability density functions on $\mathbb{R}$ is a subset of $\mathcal{D}'(\mathbb{R})$. For any probability density function $P(x)$ the associated distribution $T_P \in \mathcal{D}'(\mathbb{R})$ is defined by

$$\langle T_P, \varphi \rangle = \int_\mathbb{R} \varphi(x)P(x)dx \quad \forall \varphi \in \mathcal{D}(\mathbb{R}).$$

Example. The set of measures $\mathcal{M}(\Omega)$ is a subset of $\mathcal{D}'(\Omega)$. For any $\mu \in \mathcal{M}(\Omega)$ the associated distribution T_μ is defined by

$$\langle T_\mu, \varphi \rangle = \int_\Omega \varphi(x)d\mu \quad \forall \varphi \in \mathcal{D}(\Omega).$$

Example. An important example of a distribution which is not defined in the above way is the Delta distribution $\delta_y(x)$ (concentrated on $y \in \mathbb{R}^d$)

$$\langle \delta_y, \varphi \rangle = \int_{\mathbb{R}^d} \varphi(x) d\delta_y(x) = \varphi(y) \quad \forall \varphi \in \mathcal{D}(\Omega).$$

where

$$\delta_y(E) = \begin{cases} 1, & y \in E \\ 0, & y \notin E \end{cases}.$$

The empirical measure μ^N is actually given by using the Delta distribution

$$\mu^N(t) \triangleq \frac{1}{N} \sum_{j=1}^{N} \delta_{x_i(t)} \quad \langle \mu^N, \varphi \rangle = \frac{1}{N} \sum_{j=1}^{N} \varphi(x_i(t)).$$

We define the convergence for a sequence of distributions as follows

Definition 2.3.5. For a sequence of distributions $(T_m)_{m\in\mathbb{N}} \subset \mathcal{D}'(\Omega)$ we was it converges against a limit $T \in \mathcal{D}'(\Omega)$ iff

$$\langle T_m, \varphi \rangle \to \langle T, \varphi \rangle \quad \forall \varphi \in \mathcal{D}(\Omega).$$

Based on this convergence we give some examples in the approximation of $\delta_0(x)$

Example (Heat Kernel)**.** The heat kernel for $x \in \mathbb{R}$ and $t > 0$ is given by

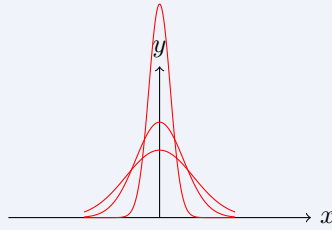$$f_t(x) = \frac{1}{(4\pi t)^{\frac{1}{2}}} e^{-\frac{|x|^2}{4t}}.$$

Figure 2.1: Heat Kernel for different t

Lemma 2.3.1. The sequence of distributions associated to the heat kernel converge to the Delta distribution

Proof. We consider the limit $t \to 0^+$ and obtain $\forall \varphi \in \mathcal{C}_0^\infty(\Omega)$

$$\lim_{t \to 0^+} \int_{\mathbb{R}} f_t(x)\varphi(x) = \lim_{t \to 0^+} \int_{\mathbb{R}} \frac{1}{(4\pi t)^{\frac{1}{2}}} e^{-\frac{|x|^2}{4t}} \varphi(x)$$

$$= \lim_{t \to 0^+} \int_{-\infty}^{\infty} \frac{1}{\sqrt{\pi}} e^{-y^2} \varphi(2\sqrt{t}y) dy$$

$$= \varphi(0) = \langle \delta_0, \varphi \rangle.$$

where we used $x = 2\sqrt{t}y$ □

Example. For the rectangular functions

$$Q_n(x) = \begin{cases} \frac{n}{2}, & |x| \leq \frac{1}{n} \\ 0, & |x| > \frac{1}{n} \end{cases}.$$
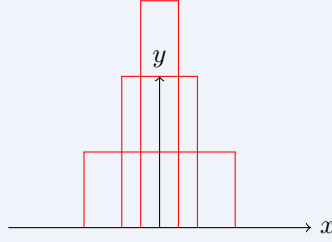
Then

$$Q_n \xrightarrow{n \to \infty} \delta_0(x).$$

Figure 2.2: Rectangular functions for different n

Example. The Dirichlet kernel

$$D_n(x) = \frac{\sin(n + \frac{1}{2})x}{\sin \frac{x}{2}} = 1 + 2\sum_{k=1}^{n} \cos(kx).$$
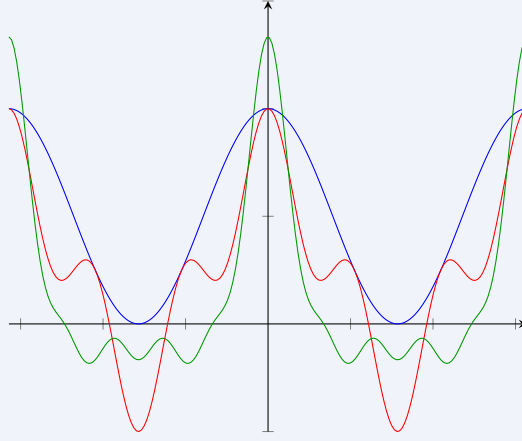
Then

$$D_n \xrightarrow{n \to \infty} 2\pi\delta_0(x).$$

Figure 2.3: Dirichlet kernel for different n

2.4 Weak Derivative Of Distributions

Definition 2.4.1. For all distributions $\forall T \in \mathcal{D}'(\Omega)$ we define the derivative $\partial_i T$ by

$$\langle \partial_i T, \varphi \rangle := -\langle T, \partial_i \varphi \rangle \quad \forall \varphi \in \mathcal{D}(\Omega) \langle \partial_i^\alpha T, \varphi \rangle \quad := (-1)^{|\alpha|}\langle T, \partial_i^\alpha \varphi \rangle \quad \forall \varphi \in \mathcal{D}(\Omega).$$

For multi index α

Exercise. Prove the function $-\langle T, \partial_i \varphi \rangle$ is a continuous and linear function

Hint: Consider the case where $T := T_f$ for $f \in L^1_{\text{loc}}$

We give a couple examples

Example. For $\forall \varphi \in \mathcal{D}(\Omega)$ the weak derivative of the Dirac Delta distribution is given by

$$\langle \delta_0', \varphi \rangle = -\langle \delta_0, \varphi' \rangle = -\varphi(0)$$
$$\langle \delta_0^{(k)}, \varphi \rangle = (-1)^k \varphi^{(k)}(0).$$

Lemma 2.4.1. The weak derivative of the 1-D Heaviside function

$$H(x) = \begin{cases} 1, & x \geq 0, \\ 0, & x < 0 \end{cases}.$$

is the Dirac Delta distribution

Proof. For $\forall \varphi \in \mathcal{D}(\Omega)$ it holds

$$\langle H', \varphi \rangle \stackrel{\text{Def.}}{=} -\langle H, \varphi' \rangle$$
$$= -\int_{-\infty}^{\infty} H(x)\varphi'(x)dx$$
$$= -\int_{0}^{\infty} \varphi'(x)dx$$
$$= \varphi(0)$$
$$= \langle \delta_0, \varphi \rangle.$$

Therefore

$$H' = \delta_0.$$

$\square$

We can now go on to properly formulate the mean field partial differential equation in a weak sense

2.5 Weak Formulation Of The Mean Field Partial Differential Equation

Using the notation of the empirical measure we can rewrite our earlier definition of the (MPS) as follows

$$\begin{cases} \frac{d}{dt}x_i(t) &= \langle K(x_i, \cdot), \mu^N(t, \cdot) \rangle = \int_{\mathbb{R}^d} K(x_i, y)d\mu^N(t, y) \\ x_i(0) &= x_{i,0} \in \mathbb{R}^d, t \in [0, T] \end{cases}.$$

As has been discussed before, the empirical measure satisfies for $\forall \varphi \in \mathcal{C}_0^\infty(\mathbb{R}^D)$

$$\frac{d}{dt}\langle \mu^N, \varphi \rangle = \langle \mu^N \mathcal{K}\mu^N, \nabla\varphi \rangle = \langle -\operatorname{div}(\mu^N \mathcal{K}\mu^N), \varphi \rangle.$$

where

$$\mathcal{K}\mu^N(x) = \int_{\mathbb{R}^d} K(x,y)d\mu^N(y),$$

which means that the empirical measure μ^N satisfies the following equation in the sense of distribution

$$(\text{MPDE}) \quad \partial_t \mu^N + \text{div}(\mu^N \mathcal{K}\mu^N) = 0.$$

Exercise. Show $\mu^N \mathcal{K}\mu^N$ is a distribution for smooth $K(x,y)$

Next we concentrate on the following PDE

Definition 2.5.1 (Mean Field Equation (MFE)). Define the mean field equation as

$$(\text{MFE}) \begin{cases} \partial_t + \text{div}(\mu \mathcal{K}\mu) &= 0 \\ \mu|_{t=0} &= \mu_0 \end{cases}.$$

where

$$\mathcal{K}\mu^N(x) = \int_{\mathbb{R}^d} K(x,y)d\mu^N(y),$$

We give the definition of the weak solution of (MFE)

Definition 2.5.2 (Weak Solution of MFE). For all $t \in [0,T]$, $\mu(t) \in \mathcal{M}(\mathbb{R}^d)$ is called a weak solution of (MFE), where $\mathcal{M}(\mathbb{R}^d)$ denotes the space of measures on $\mathbb{R}^d$

For $\forall \varphi \in \mathcal{C}_0^\infty(\mathbb{R}^d)$ it holds

$$\langle \mu(t), \varphi \rangle - \langle \mu_0, \varphi \rangle = \int_0^t \langle \mu(s)\mathcal{K}\mu(s), \nabla\varphi \rangle.$$

Remark. If $\mu_0 = \mu^N(0)$ i.e. the initial data is given by an empirical measure, then $\mu^N(t, \cdot)$ is a weak solution of the (MFE)

We define the following initial value problem, the so called characteristics equation

Definition 2.5.3 (Push Forwad Measure)**.** For a measurable function X and a measure $\mu_0 \in \mathcal{M}(\mathbb{R}^d)$ denote the push forward measure for any Borel set $B \subset \mathbb{R}^d$ by

$$X \# \mu_0 \coloneqq \mu_0(X^{-1}(B)).$$

Definition 2.5.4 (Characteristics equation)**.**

$$
\begin{cases}
\frac{d}{dt}x(t, x_0, \mu_0) & = \int_{\mathbb{R}^d} K(x(t, x_0, \mu_0), y) d\mu(y, t) \\
x(0, x_0, \mu_0) & = x_0 \quad \forall x_0 \in \mathbb{R}^d \\
\mu(\cdot, t) & = x(t, \cdot, \mu_0) \# \mu_0
\end{cases}.
$$

The solution flow $x(t, \cdot, \mu_0)$ gives for any time $t > 0$ a map

$$x(t, \cdot, \mu_0) \; : \; \mathbb{R}^d \to \mathbb{R}^d.$$

Remark. It can be easily checked that the push forward measure $\mu(t)$ obtained in the Characteristics equation is a weak solution of the (MFE)

Remark. The solution space of the Characteristics equation is given by

$$\mathcal{P}_1(\mathbb{R}^d) = \{\mu \in \mathcal{P}(\mathbb{R}^d) \; : \; \int_{\mathbb{R}^d} |x| d\mu(x) < \infty\}.$$

where $\mathcal{P}(\mathbb{R}^d)$ is the space of all probability measures

Assumption C (Regularity)**.** We say an interaction force K is regular if $K \in \mathcal{C}^1(\mathbb{R}^d \times \mathbb{R}^d; \mathbb{R}^d)$ and there exists an $L > 0$ such that

$$\sup_y |\nabla_x K(x, y)| + \sup_x |\nabla_y K(x, y)| \leq L.$$

Actually this assumption has already been used in order to show the well-posedness of the particle system

Theorem 2.5.1 (Existence and Uniqueness of Characteristics Equation)**.** Let Assumption C hold for K and $\mu_0 \in \mathcal{P}_1(\mathbb{R}^d)$ then the Characteristics equation has a unique solution $x(t, x_0, \mu_0) \in \mathcal{C}^1(\mathbb{R}; \mathbb{R}^d)$ and $x(t, \cdot, \mu_0) \# \mu_0 \in \mathcal{P}_1$ for $\forall t > 0$

Proof. The proof is based on Picard iteration.

Let $C_1 = \int_{\mathbb{R}^d} |x| d\mu_0(x)$ and define the following Banach space

$$X := \{v \in \mathcal{C}(\mathbb{R}^d) \ \|v\|_X < \infty\}.$$

Where

$$\|v\|_X := \sup_{x \in \mathbb{R}^d} \frac{|v(x)|}{1 + |x|}.$$

As preparations we need the following estimates for the nonlocal term, by using Assumption C for K we have for $\forall v, w \in X$

$$\left| \int_{\mathbb{R}^d} K(v(x), v(y)) d\mu_0(y) - \int_{\mathbb{R}^d} K(w(x), w(y)) d\mu_0(y) \right|$$

$$\leq L \int_{\mathbb{R}^d} |v(x) - w(x)| + |v(y) - w(y)| d\mu_0(y)$$

$$\leq L\|v - w\|_X (1 + |x|) + L\|v - w\|_X \int_{\mathbb{R}^d} (1 + |y|) d\mu_0(y)$$

$$\leq L(2 + C_1)\|v - w\|_X (1 + |x|).$$

Now define the Picard iteration for $\forall y \in \mathbb{R}^d$

$$x_0(t, y) = y$$

$$x_1(t, y) = y + \int_0^t \int_{\mathbb{R}^d} K(x_0(s, y), x_0(s, z)) d\mu_0(z) ds$$

$$\vdots$$

$$x_m(t, y) = y + \int_0^t \int_{\mathbb{R}^d} K(x_{m-1}(s, y), x_{m-1}(s, z)) d\mu_0(z) ds$$

.

Then we can bound the difference between x_1 and x_0 by

$$|x_1(t,y) - x_0(t,y)| = \left| \int_0^t \int_{\mathbb{R}^d} K(x_0(s,y), x_0(s,z)) d\mu_0(z) ds \right|$$

$$= \left| \int_0^t \int_{\mathbb{R}^d} K(y,z) d\mu_0(z) ds \right|$$

$$\leq \int_0^{|t|} \int_{\mathbb{R}^d} L(|y| + |z|) d\mu_0(z) ds$$

$$= \int_0^{|t|} L(|y| + C_1) ds$$

$$\leq L(1 + C_1)(1 + |y|)|t|.$$

Furthermore for $\forall m \geq 1$ we have

$$|x_m(t,y) - x_{m-1}(t,y)|$$

$$= \left| \int_0^t \int_{\mathbb{R}^d} \left(K(x_{m-1}(s,y), x_{m-1}(s,z)) - K(x_{m-2}(s,y), x_{m-2}(s,z)) \right) d\mu_0(z) ds \right|$$

$$\leq L(2 + C_1) \int_0^{|t|} \|x_{m-1}(s,\cdot) - x_{m-2}(s,\cdot)\|_X (1 + |y|) ds.$$

hence by dividing both sides by $1 + |y|$ we have

$$\|x_m(t,\cdot) - x_{m-1}(t,\cdot)\|_X \leq L(2 + C_1) \int_0^{|t|} \|x_{m-1}(s,\cdot) - x_{m-2}(s,\cdot)\|_X ds$$

$$\leq \frac{((2 + C_1)L|t|)^d}{(m-1)!}.$$

which implies for $\forall m > n \to \infty$

$$\|x_m(t,\cdot) - x_n(t,\cdot)\|_X \leq \sum_{i=n}^{m-1} \|x_{i+1}(t,\cdot) - x_i(t,\cdot)\|_X \to 0.$$

Therefore for $T > 0$

$$x_m(t,\cdot) \to x(t,\cdot) \text{ in } X \text{ uniformly in } [-T, T].$$

and $x \in \mathcal{C}(\mathbb{R}; \mathbb{R}^d)$ satisfies that, after taking the limit in Picard iteration $\forall y \in \mathbb{R}^d$

$$x(t,y) = y + \int_0^t \int_{\mathbb{R}^d} K(x(s,y), x(s,z)) d\mu_0(z) ds.$$

By the fundamental theorem of calculus and Assumption C we know that for $y \in \mathbb{R}^d$ and $x(t,y) \in \mathcal{C}^1(\mathbb{R}; \mathbb{R}^d)$

$$\frac{d}{dt} x(t,y) = \int_{\mathbb{R}^d} K(x(t,y), x(t,z)) d\mu_0(z) = \int_{\mathbb{R}^d} K(x(t,y), z') d\mu(z', t).$$

where $\mu(\cdot, t)$ is the push forward measure of μ_0 along $x(t, \cdot)$

For uniqueness consider two solutions $x, \tilde{x}$ then by taking the difference we have

$$x(t,y) - \tilde{x}(t,y) = \int_0^t \int_{\mathbb{R}^d} \left(K(x(s,y), x(s,z)) - K(\tilde{x}(s,y), \tilde{x}(s,z)) \right) d\mu_0(z) ds.$$

Using estimates similarly to before we obtain

$$\|x(t,\cdot) - \tilde{x}(t,\cdot)\|_X \leq L(2 + C_1) \int_0^{|t|} \|x(s,\cdot) - \tilde{x}(s,\cdot)\|_X ds.$$

By applying Gronwall's inequality we get

$$\|x(t,\cdot) - \tilde{x}(t,\cdot)\|_X = 0.$$

where clearly $\|x(0,\cdot) - \tilde{x}(0,\cdot)\|_X = 0$ $\qquad \square$