Chapter 1

Introduction

Mean Field Particle Systems is about the study of particles which are represented by (stochastic) differential equations. This course in particular is concerned with the behaviour of the system as the size grows to infinity:

Definition 1.0.1 (Toy Mean Field Particle System)**.** Let $N \in \mathbb{N}$ then a Mean Field Particle System of first order is given by :

$$x_1(t), \ldots, x_n(t) \in C^1([0,T]; \mathbb{R}^d) \qquad x_i(0) = c_i.$$

Where each particle satisfies

$$dx_i = \frac{1}{N}\sum_{j=1}^{N} K(x_i, x_j)dt + \sigma dB_i(t).$$

Where B_i is a Brownian motion; For $\sigma = 0$ the system is called deterministic.

Example 1.0.2. Example choices for K are :

$$K(x_i, x_j) = \nabla(|x_i - x_j|^2).$$

or :

$$o\gamma = \frac{x_i - x_j}{|x_i - x_j|^d}.$$

Goal is to study what happens at $N \to \infty$, to do so we consider how the measure of a system converges

Definition 1.0.3 ((Empirical) Measure of a System)**.** Consider the point measure for every x_i : $\delta_{x_i(t)}$, then the measure of the System of order N is :

$$\mu_N(t) = \frac{1}{N}\sum_{i=1}^{N} \delta_{x_i(t)}.$$

Assumption 1.0.4. For initial data the empirical measure of a system converges $\mu_N(0) \to \mu(0)$ where μ is absolutely continuous with respect to the Lebesgue Measure

Corollary 1.0.5. *By Radon Nikodym*

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

It can be shown that μ solves a PDE, to do so we compute the derivative of μ using test functions

$$\forall \phi \in C_0^\infty(\mathbb{R}^d).$$

$$\begin{aligned}
\frac{d}{dt} < \mu_N(t), \phi >= \frac{d}{dt}\int_{\mathbb{R}^d} \phi(x) d\mu_N(t)(x) &= \frac{d}{dt}\int \frac{1}{N}\sum_{i=1}^N \phi(x) d\delta_{x_i(t)} \\
&= \frac{1}{N}\sum_{i=1}^N \frac{d}{dt}\phi(x_i(t)) \\
&\overset{\text{Chain.}}{=} \frac{1}{N}\sum_{i=1}^N \nabla\phi(x_i(t))\frac{d}{dt}x_i(t) \\
&= \frac{1}{N}\sum_{i=1}^N \nabla_x\phi(x_i(t)) * \underbrace{\frac{1}{N}\sum_{j=1}^N K(x_i(t), x_j(t))}_{\text{Def.}} \\
&= \frac{1}{N}\sum_{i=1}^N \nabla_x\phi(x_i(t)) * \frac{1}{N}\sum_{j=1}^N \int_{\mathbb{R}^d} K_(x_i(t), y) d\delta_{x_j(t)}(y) \\
&= \frac{1}{N}\sum_{i=1}^N \nabla_x\phi(x_i(t)) * \int_{\mathbb{R}^d} K_(x_i(t), y) d\mu_N(t) \\
&= \int_{\mathbb{R}^d} \nabla\phi(x) \int_{\mathbb{R}^d} K(x,y) d\mu_N(t,y) d\mu_N(t,x)
\end{aligned}$$

.

Where the last line can be rewritten by using Integration by Parts (Divergence Theorem) :

$$\int_{\mathbb{R}^d} \nabla\phi(x) \int_{\mathbb{R}^d} K(x,y) d\mu_N(t,y) d\mu_N(t,x) \overset{\text{Part.}}{=} - < \nabla * (\mu_N \int_{\mathbb{R}^d K(\cdot,y) d\mu_N(y)}), \phi >$$

.

This means μ satisfies :

$$\partial_t \mu_N + \nabla * (\mu_N * \int K(\cdot, y) d\mu_N(y)) = 0 \quad \xrightarrow{N\to\infty} \partial_t \mu + \nabla * (\mu * \int K(\cdot, y) d\mu(y)) = 0.$$

In practical applications (Theoretical Physics , Biology) systems that are considered are often of second order

Definition 1.0.6 (Toy Second Order System)**.** Given $N \in \mathbb{N}$ a Second Order System is given by

$$(x_i(t), v_i(t)), \ldots, (x_N(t), v_N(t)) \in \mathbb{R}^{2d}.$$

Such that :

$$\begin{aligned}
&\frac{d}{dt} x_i(t) = v_i(t) \\
&\frac{d}{dt} v_i(t) = \frac{1}{N} \sum_{j=1}^{N} F(\underbrace{x_i(t), v_i(t)}_{\text{Position and Velocity of itself}} ; x_j(t), v_j(t)) + \sigma \frac{dB_t}{dt}
\end{aligned}$$

.

Example 1.0.7 (Gravitational Force)**.** An example of F could be :

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

Definition 1.0.8 (Second Order Measure)**.** The Measure of a second order System is given by :

$$\mu_N(x, v) = \frac{1}{N} \sum_{i=1}^{N} \delta_{(x_i(t), v_i(t))}.$$

Exercise 1.0.9. *Show what PDE μ solves for $\sigma = 0$, Hint : Calculate $\frac{d}{dt} < \mu_N, \phi >$ for some test function $\phi \in C_0^\infty(\mathbb{R}^{2d})$*

Chapter 2

Deterministic Mean Field Particle Systems

Definition 2.0.1 (Deterministic Mean Field Particle System)**.** For $N \in \mathbb{N}$ a deterministic mean field particle system is given by N particles :

$$x_1(t), \ldots, x_n(t) \in C^1([0,T];\mathbb{R}^d) \qquad x_i(0) = c_i.$$

With initial points :

$$x_i(0) = x_{i,0} \in \mathbb{R}^d.$$

And the relation :

$$\frac{d}{dt}x_i = \frac{1}{N}\sum_{j=1}^{N} K(x_i, x_j).$$

The system is then given by :

$$X_N = \begin{pmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_N(t) \end{pmatrix} \in \mathbb{R}^{dN}.$$

The goal is to solve the above $d * N$ dimensional system under assumptions on K.

2.1 ODE Theory

Definition 2.1.1 (Initial Value Problem (standard))**.** For $\forall T > 0$ the standard ode system is given by :

$$\begin{aligned} x' &= f(t,x) \\ x|_{t=0} &= x_0 \in \mathbb{R}^n. \end{aligned}$$

with $t \in [0,T],\ x(t) \in \mathbb{R}^n$

Assumption 2.1.2 (A)**.** Condition for global existence : f is continuous in $(t,x) \in [0,T] \times \mathbb{R}^n$
Condition for uniqueness : f is Lipschitz continuous in x

Theorem 2.1.3. *Whenever assumption 2.1.2 holds the standard IVP has a unique solution $x \in C^1([0,T];\mathbb{R}^n)$*

Proof. Rewriting the IVP using integration :

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

We construct the following sequence by :

$$\begin{aligned} x_1(t) &= x_0 + \int_0^t f(s,x_0)ds \\ x_2(t) &= x_0 + \int_0^t f(s,x(s))ds \qquad\qquad \vdots \\ x_m(t) &= x_0 + \int_0^t f(s,x_{m-1}(s))ds. \end{aligned}$$

Step 1 of the proof consists of proving the above sequence is converging , step 2 is then showing the limit is a solution to the IVP

Under Assumption 2.1.2 we know that f is continuous such that $(x_n)_{n\in N} \subset C^1([0,T];\mathbb{R}^n)$, As C^1 is complete any sequence that is cauchy must also converge against a limit in the space.

$$\begin{aligned} |x_2 - x_1| = |\int_0^t f(s,x_1(s))ds - \int_0^t f(s,x_0(s))ds| &= |\int_0^t f(s,x_{m-1}(s)) - f(s,x_{n-1}(s))ds| \\ &\leq \int_0^t |f(s,x_1(s)) - f(s,x_0(s))|ds \\ &\overset{\text{Lip.}}{\leq} L\int_0^t |x_1(s) - x_0(s)|ds \\ &= L\int_0^t |\int_0^{s_0} f(s,x_0)ds|ds_0 \\ &\leq L * \int_0^t \int_0^{s_0} |f(s,x_0)|dsds_0 \\ &\leq L \underbrace{M}_{=\max_{s\in[0,T]}|f(s,x_0)|} \frac{t^2}{2}. \end{aligned}$$

By repeatedly using the Lipschitz continuity of f the following induction assumption is motivated :

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

(IS) : $m \longrightarrow m+1$

$$
\begin{aligned}
|x_{m+1}(t) - x_m(t)| &\overset{\text{Lip.}}{\leq} L \int_0^t |x_m(s) - x_{m-1}(s)| ds \\
&\overset{\text{IA.}}{\leq} L \int_0^t \frac{ML^{m-1}s^m}{m!} ds = ML^m \frac{t^{m+1}}{(m+1)!}.
\end{aligned}
$$

For any $n, m \in \mathbb{N}$ and assuming without loss of generality that $n > m$ such that $n = m + p$ for $p \in N$:

$$
\begin{aligned}
|x_n(t) - x_m(t)| = |x_{m+p}(t) - x_m(t)| \leq \sum_{k=m+1}^{m+p} |x_k(t) - x_{k-1}(t)| &\overset{\text{Ind.}}{\leq} M \sum_{k=m+1}^{m+p} \frac{L^{k-1}T^k}{k!} \\
= \frac{M}{L} \sum_{k=m+1}^{m+p} \frac{(LT)^k}{k!} &= \frac{M}{L} \frac{(LT)^{m+1}}{(m+1)!} \sum_{k=0}^{p-1} \frac{(LT)^k}{k!} \\
&\leq \frac{M}{L} \frac{(LT)^{m+1}}{(m+1)!} e^{LT} \xrightarrow{m\to\infty} 0 \text{ uniformly in } t \in [0, T].
\end{aligned}
$$

This shows that $(x_m)_{m\in\mathbb{N}}$ is Cauchy and has a limit $x \in C([0, T]; \mathbb{R}^n)$ with :

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

It remains to show that $x(t)$ is a solution to the IVP i.e :

$$
x(t) = \lim_{m\to\infty} x_0 + \int_0^t f(s, x_{m-1}(s)) ds \leftrightarrow x_0 + \int_0^t f(s, x(s)) ds.
$$

Which can be shown by :

$$
\begin{aligned}
|\lim_{m\to\infty} \int_0^t f(s, x_{m-1}(s)) - f(s, x(s)) ds| &\leq \lim_{m\to\infty} \int_0^t |f(s, x_{m-1}(s)) - f(s, x(s))| ds \\
&\leq \lim_{m\to\infty} L \int_0^t |x_{m-1}(s) - x(s)| ds \\
&\leq \lim_{m\to\infty} Lt * \max_{s\in[0,t]} |x_{m-1}(s) x(s)| \\
&\leq \lim_{m\to\infty} Lt * \max_{s\in[0,T]} |x_{m-1}(s) x(s)| \\
&= 0.
\end{aligned}
$$

It remains to show that the solution is unique, for that assume $x, \hat{x} \in C([0, T]; \mathbb{R}^n)$ are both solutions to the IVP. Meaning that :

$$
\begin{aligned}
x(t) &= x_0 + \int_0^t f(s, x(s)) ds \\
\hat{x}(t) &= x_0 + \int_0^t f(s, \hat{x}(s)) ds.
\end{aligned}
$$

Then :

$$
\begin{aligned}
|x-\hat{x}| &\leq \int_0^t |f(s,x(s))-f(s,\hat{x}(s))|ds \leq L * \int_0^t |x(s)-\hat{x}(s)|ds \\
&= L\int_0^t \underbrace{e^{-\alpha s}|x(s)-\hat{x}(s)|}_{=\rho(s)} e^{\alpha s} ds \\
&\leq L \max_{t\in[0,T]} \rho(t) * \frac{1}{\alpha}(e^{\alpha t}-1) \\
&\leq L \max_{t\in[0,T]} \rho(t) * \frac{1}{\alpha} * e^{\alpha t}.
\end{aligned}
$$

By rearranging with the initial term :

$$
\rho(t) = e^{-\alpha t}|x(t)-\hat{x}(t)| \leq \frac{L}{\alpha} \max_{t\in[0,T]} \rho(t)
$$

$$
\max_{t\in[0,T]} \rho(t) \leq \frac{L}{\alpha} \max t \in [0,T]\rho(t).
$$

by choosing $\alpha = 2L$:

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

And the solutions must be equal for $\forall t \in [0,T]$. □

The reason this proof deviates from the standard Picard-Lindelöf theorem, is that for our systems we require Global existence, doing so by requiring f to be globally Lipschitz continuous.

Theorem 2.1.4. *The solution $x(t,t_0,x_0) \in C$ is continuously dependent on (t_0,x_0)*

Theorem 2.1.5 (Gronwalls inequality)**.** *For $\alpha,\beta,\phi \in C([a,b];\mathbb{R})$ $\beta \geq 0$ and*

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

then :

$$
\phi(t) \leq \alpha(t) + \int_a^t \beta(s)\exp\left(\int_s^t \beta(\tau)d\tau\right)\alpha(s)ds.
$$

Proof. Denote $\psi(t) = \int_a^t \beta(s)\phi(s)ds$ then

$$
\begin{aligned}
\psi'(t) &= \beta(t)\phi(t) \leq \beta(t)\alpha(t) + \beta(t)\psi(t) \\
&= \beta(t) * (\alpha(t)+\psi(t))
\end{aligned}
$$

.

Recall $\dot{x} + a(t)x + b(t) = 0$

$$
\begin{aligned}
(\dot{\psi}(t) - \beta(t)\psi(t))e^{-\int_a^t \beta(s)ds} &\leq \beta(t)\alpha(t) * e^{-\int_a^t \beta(s)ds} \\
(e^{-\int_a^t \beta(s)ds}\psi(t))' &\leq \beta(t)\alpha(t) * e^{-\int_a^t \beta(s)ds}.
\end{aligned}
$$

Integrating gives :

$$(e^{-\int_a^t \beta(s)ds}\psi(t)) \overset{\psi(a)=0}{\leq} \int_a^t \beta(s)\alpha(s)e^{-\int_a^s \beta(r)dr}ds.$$

□

Definition 2.1.6 (Regularity)**.** A function $K : \mathbb{R}^{2d} \to \mathbb{R}^d$ is called regular if :

1. $K \in C^1(\mathbb{R}^{2d}; \mathbb{R}^d)$ (gives local lipschitz)
2. And $\exists L > 0$ s.t. :

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

We further assume K has the following properties :

$$K(x, y) = -K(y, x) \qquad \text{(antisymmetric)}$$
$$K(x, x) = 0.$$

Theorem 2.1.7. *When the assumption of regularity on K holds, the MPS has a solution for all $T > 0$*

$$\begin{cases} \frac{d}{dt}x_i & = \frac{1}{N}\sum_{j=1}^N K(x_i, x_j), 1 \leq i \leq N \\ x_i(0) & = x_{i,0} \in \mathbb{R}^d \end{cases}.$$

has a unique solution by Picard-Iteration :

$$X_N(t) = \big(x_1(t), x_2(t), \ldots, x_N(t)\big) \in C^1([0, T]; \mathbb{R}^{dN}).$$

Definition 2.1.8 (Empirical Measure of a System)**.** Consider the point measure for every $x_i : \delta_{x_i(t)}$, then the measure of the System of order N is given by

$$\mu_N(t) = \frac{1}{N}\sum_{i=1}^N \delta_{x_i(t)}.$$

As shown in the introduction μ_N is a (weak-) solution to the following PDE

$$\partial_t \mu_N + \nabla * (\mu_N * \int K(\cdot, y)d\mu_N(y)) = 0.$$

Idea. *Now for $N \to \infty$ if we have $\mu_N \xrightarrow{\text{in some sense}} \mu$ then μ is a (weak) solution to*

$$\partial_t \mu + \nabla * (\mu * \int K(\cdot, y)d\mu(y)) = 0.$$

with

$$\mu_0 \leftarrow \mu_N(0).$$

2.2 Weak Solutions and Distributions

Distributions are a more general class of functions and can be seen as the dual space of the space of test functions

Definition 2.2.1 (Multi-Index)**.** A multi-index $\gamma \in \mathbb{N}_0^n$ of length $|\gamma| = \sum_i \gamma_i$ for example $\gamma = (0, 2, 1) \in \mathbb{N}_0^3$ can be used to denote partial derivatives of higher order as such :

$$\partial^\gamma = \prod_i (\frac{\partial}{\partial x_i})^{\gamma_i}.$$

Only sensible cause partial derivatives commute as otherwise the index would be ambiguous.

Definition 2.2.2 (Test Functions)**.** For $\Omega \subset \mathbb{R}^d$ the space of test functions $\mathcal{D}(\Omega) \supset C_0^\infty(\Omega)$. We say a sequence of test functions $(\phi_m)_{m\in\mathbb{N}} \subset C_0^\infty(\Omega)$ converges against some limit $\phi \in C_0^\infty(\Omega)$ iff.

1. $\exists$ a compact set $K \subset \Omega$ s.t. supp $\phi_m \subset K$ for all $m \in \mathbb{N}$
2. $\forall$ multi-indexes $\alpha \in \mathbb{N}_0^n$:

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

Remark 2.2.3. $\mathcal{D}(\Omega)$ is a linear space

Definition 2.2.4 (Distribution)**.** The space of distributions $\mathcal{D}(\Omega)'$ is the dual space of $\mathcal{D}(\Omega)$ i.e. $\mathcal{D}(\Omega)'$ contains all the continuous linear functionals T

$$T : \mathcal{D}(\Omega) \to \mathbb{K}.$$

Remark 2.2.5. Continuity refers to the notion that for a sequence $(\phi_m)_{m\in\mathbb{N}} \subset \mathcal{D}(\Omega)$ with limit ϕ then :

$$\phi_m \to \phi \implies T(\phi_m) \to T(\phi).$$

linearity :

$$T(\alpha\phi_1 + \beta\phi_2) = \alpha T(\phi_1) + \beta T(\phi_2).$$

We sometimes write $\langle T, \phi \rangle$ instead of $T(\phi)$

Example 2.2.6. Every locally integrable function $f \in L^1_{\text{loc}}(\Omega) := \{f \mid \forall K \subset \Omega, \int_K f(x)dx < \infty\}$ defines a Distribution by :

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

i.e. $L^1_{\text{loc}}(\Omega) \subset \mathcal{D}'(\Omega)$

Example 2.2.7. Probability densities are distributions in the same sense as for L^1_{loc} functions

Example 2.2.8. (Probability -) Measures $\mu \in \mathcal{M}(\Omega)$ define a distributions , by :

$$\langle T_\mu, \phi \rangle = \int_{\mathbb{R}^d} \phi(x) d\mu(x) < \infty \quad \forall \phi \in \mathcal{D}(\Omega).$$

A prominent example is the δ distribution defined by :

$$\langle \delta, \phi \rangle = \int_{\mathbb{R}^d} \phi(x) d\delta = \phi(0).$$

Remember for a measurable set E

$$\delta_x(E) = \begin{cases} 1, & \text{if } x \in E \\ 0 & \text{if } x \notin E \end{cases}.$$

Coming back to our empirical measure 2.1.8, we can see the corresponding distribution is defined by :

$$\langle \mu_n, \phi \rangle = \frac{1}{N} \sum_{i=1}^{N} \phi(x_i).$$

Chapter 3

Stochastic Mean Field Particle Systems

5-6 weeks

3.1 Basics of probability

3.2 Bad K

3.3 Convergence

Chapter 4

New Results

4.1 Relative entropy Method

Goal to prove "strong" convergence in L^1