Chapter 1

Introduction

Kinetic theory provides a mesoscopic description of many-particle systems, bridging the gap between microscopic particle dynamics and macroscopic continuum models. Rather than following every particle individually, one introduces a distribution function describing the statistical state of the system in phase space. This framework plays a central role in plasma physics, rarefied gas dynamics, collective behaviour, and many other applications. Particular emphasis will be placed on the Vlasov–Poisson and Vlasov–Maxwell systems, which serve as guiding examples throughout these notes.

The state of the system is encoded in the *distribution function* $f(t, x, v)$, where $x \in \mathbb{R}^n$ is the position variable, $v \in \mathbb{R}^n$ is the velocity variable, with $n = 1$, 2, or 3. More precisely,

$$f(t, x, v)\,dx\,dv \;=\; \text{number of particles in the phase-space volume } dx\,dv.$$

The distribution function satisfies the following transport equation:

$$\partial_t f + \nabla_x \cdot (v\,f) + \nabla_v \cdot (F(t, x, v)\,f) = 0.$$

When the force field satisfies $\nabla_v \cdot F = 0$ (as is the case for the Lorentz force), this simplifies to

$$\partial_t f + v \cdot \nabla_x f + F(t, x, v) \cdot \nabla_v f = 0.$$

The principal macroscopic observables are defined as velocity moments of f:

$$\begin{pmatrix} \rho(t,x) \\ \rho u(t,x) \\ e(t,x) \end{pmatrix} \;=\; \int_{\mathbb{R}^n} \begin{pmatrix} 1 \\ v \\ \dfrac{|v|^2}{2} \end{pmatrix} f(t, x, v)\,dv,$$

representing the particle density, momentum density, and kinetic energy density, respectively.

1.1 Electrostatic interactions

One of the most important examples of a kinetic model arises when particles interact through an electric field generated by their own charge distribution. The resulting coupling between transport and Poisson's equation leads to the Vlasov–Poisson system, which is a cornerstone of plasma physics and galactic dynamics.

Consider particles of charge q and mass m. The Coulomb force acting on a particle is

$$F(t, x) = q\,E(t, x) = -q\,\nabla_x \phi(t, x),$$

where the electrostatic potential ϕ is determined self-consistently by the particle distribution through Poisson's equation:

$$-\Delta_x \phi = \frac{q}{\epsilon_0}\,\rho_f.$$

Combining these with the transport equation yields the *Vlasov–Poisson system*:

$$\begin{cases} \partial_t f + v \cdot \nabla_x f - \dfrac{q}{m} \nabla_x \phi \cdot \nabla_v f = 0, \\ -\Delta_x \phi = \dfrac{q}{\epsilon_0} \rho_f, \end{cases}$$

where the charge density is

$$\rho_f(t,x) = \int_{\mathbb{R}^n} f(t,x,v)\, dv.$$

The Vlasov–Poisson system enjoys the following fundamental properties:

- **Positivity.** If $f^{\text{in}} \geq 0$, then $f(t) \geq 0$ for all $t \geq 0$.
- **Preservation of L^p norms** $(1 \leq p \leq \infty)$. Since the phase-space divergence vanishes,

  $$\nabla_{(x,v)} \cdot (v,\, E(t,x)) = 0,$$

  the flow is incompressible in phase space and therefore

  $$\|f(t)\|_{L^p} = \|f^{\text{in}}\|_{L^p}, \qquad \forall\, t \geq 0.$$

- **Conservation of energy.** The total energy

  $$\mathcal{E}(t) := \frac{m}{2} \int_{\mathbb{T}^d \times \mathbb{R}^d} |v|^2 f(t,x,v)\, dv\, dx + \frac{\epsilon_0}{2} \int_{\mathbb{T}^d} |\nabla_x \phi(t,x)|^2\, dx$$

  is conserved for all $t \geq 0$.
- **Velocity moments.** For $p \geq 1$, the p-th order velocity moment is

  $$m_p(t) = \int_{\mathbb{T}^d \times \mathbb{R}^d} |v|^p f(t,x,v)\, dx\, dv.$$

1.2 Electromagnetic interactions

The electrostatic approximation becomes insufficient when magnetic effects play a significant role. In that case, particles evolve under the full Lorentz force and the self-consistent fields satisfy Maxwell's equations. This leads to the Vlasov–Maxwell system, one of the fundamental models of collisionless plasma dynamics.

The Lorentz force acting on a particle of charge q and velocity v is

$$F = q(E + v \times B)\,.$$

Coupling this with Maxwell's equations yields the *Vlasov–Maxwell system*:

$$\begin{cases} \partial_t f + v \cdot \nabla_x f + \dfrac{q}{m}(E + v \times B) \cdot \nabla_v f = 0, \\ \nabla \cdot B = 0, \quad \partial_t B + \nabla \times E = 0, \\ \nabla \cdot E = \dfrac{q}{\epsilon_0} \rho_f, \quad \dfrac{1}{c^2} \partial_t E - \nabla \times B = -\mu_0\, q\, J_f, \end{cases}$$

where the charge and current densities are

$$\rho_f = \int_{\mathbb{R}^n} f\, dv, \qquad J_f = \int_{\mathbb{R}^n} v\, f\, dv.$$

1.3 The aggregation equation

Transport equations also arise in many contexts outside plasma physics. Aggregation models describe the collective motion of interacting agents, particles, or organisms. Depending on the interaction kernel, these models may exhibit clustering, pattern formation, or the emergence of coherent structures.

The *aggregation equation* arises in granular media models, swarming models for animal collective behaviour, equilibrium problems in self-assembly, and mean-field games in socioeconomics, among other applications:

$$\begin{cases} \partial_t \rho + \nabla_x \cdot (\mathcal{K}[\rho]\, \rho) = 0, \\ \mathcal{K}[\rho](t,x) = \displaystyle\int_{\mathbb{R}^n} K(x,x')\, \rho(t,x')\, dx', \end{cases}$$

where the kernel K encodes the pairwise interaction law. A common and physically motivated choice is

$$K(x,y) = -\nabla_x W(x-y),$$

where $W : \mathbb{R}^n \to \mathbb{R}$ is an interaction potential. Here $K(x,y)$ represents the force exerted on a particle at x by an infinitesimal mass located at y.

Chapter 2

Transport equations

Transport equations form the mathematical foundation of kinetic theory. Their solutions are propagated along characteristic curves determined by an underlying velocity field. Understanding these characteristic trajectories provides both an intuitive interpretation of the dynamics and a powerful analytical tool. In this chapter we introduce the method of characteristics and derive several fundamental properties that will be repeatedly used later.

2.1 The free transport equation

We begin with the simplest transport model, corresponding to particles moving freely without external forces. Although elementary, this equation already contains the essential geometric ideas of transport theory and serves as a useful prototype for more general kinetic equations.

Consider the Cauchy problem

$$\partial_t f + v \cdot \nabla_x f = 0,$$

with initial datum $f(0,x,v) = f^{\text{in}}(x,v)$. The characteristic curve of this equation passing through the point $y \in \mathbb{R}^n$ at time $t = 0$ is the straight line $\gamma(t) = y + tv$, and the solution is constant along each characteristic:

$$\frac{d}{dt} f(t, \gamma(t), v) = 0.$$

Theorem 2.1.1. *For $f^{\text{in}} \in \mathscr{C}^1(\mathbb{R}^n \times \mathbb{R}^n)$, the Cauchy problem*

$$\partial_t f + v \cdot \nabla_x f = 0, \quad f(0,x,v) = f^{\text{in}}(x,v)$$

has a unique solution $f \in \mathscr{C}^1(\mathbb{R}_+ \times \mathbb{R}^{2n})$ given by

$$f(t,x,v) = f^{\text{in}}(x - tv,\ v).$$

Proof. We first prove uniqueness, and then verify that the formula defines a solution.

Uniqueness. Let f be a $\mathscr{C}^1$ solution and fix $y, v \in \mathbb{R}^n$. Along the characteristic $\gamma(t) = y + tv$,

$$\frac{d}{dt} f(t,\ y + tv,\ v) = \partial_t f + v \cdot \nabla_x f = 0.$$

Hence

$$f(t,\ y + tv,\ v) = f(0, y, v) = f^{\text{in}}(y, v).$$

Setting $x = y + tv$ gives $f(t,x,v) = f^{\text{in}}(x - tv, v)$, which is uniquely determined by f^{in}.

Existence. Define

$$f(t,x,v) = f^{\text{in}}(x - tv, v).$$

Clearly $f(0,x,v) = f^{\text{in}}(x,v)$, and a direct computation gives

$$\partial_t f + v \cdot \nabla_x f = -v \cdot (\nabla_x f^{\text{in}})(x - tv, v) + v \cdot (\nabla_x f^{\text{in}})(x - tv, v) = 0. \quad \square$$

2.2 Non-conservative equation

We now consider a more general transport equation in which the transported quantity is not necessarily conserved along the flow. Such equations naturally appear after changes of variables, linearization procedures, or when source terms are incorporated into the dynamics.

Setting $z = (x, v) \in \mathbb{R}^d$ and letting $V(t, z) \in \mathbb{R}^d$ be a time-dependent vector field, we consider the Cauchy problem

$$\begin{cases} \partial_t f + V(t, z) \cdot \nabla_z f = 0, \\ f(0, z) = f^{\text{in}}(z). \end{cases}$$

We impose the following assumptions on the velocity field:

(H1) $V \in \mathscr{C}([0, T] \times \mathbb{R}^d; \mathbb{R}^d)$.

(H2) $\nabla_z V \in \mathscr{C}([0, T] \times \mathbb{R}^d; \mathbb{R}^{d\times d})$.

(H3) *Linear growth:* there exists $\kappa > 0$ such that

$$\|V(t, z)\| \leq \kappa(1 + \|z\|), \qquad \forall (t, z) \in [0, T] \times \mathbb{R}^d.$$

Definition 2.2.1 (Characteristic curves)**.** *Given $(t, z) \in [0, T] \times \mathbb{R}^d$, the* characteristic curve *of the vector field V issued from z at time t is the solution $s \mapsto \gamma(s)$ of*

$$\begin{cases} \dot\gamma(s) = V(s, \gamma(s)), \\ \gamma(t) = z. \end{cases}$$

We denote this solution by $Z(s, t, z) := \gamma(s)$.

Theorem 2.2.2 (Cauchy–Lipschitz for characteristics)**.** *Under assumptions* (H1)–(H3)*, for each $(t, z) \in [0, T] \times \mathbb{R}^d$ there exists a unique solution*

$$s \mapsto Z(s, t, z) \in \mathscr{C}^1([0, T]; \mathbb{R}^d).$$

Moreover, $Z \in \mathscr{C}^1([0, T]^2 \times \mathbb{R}^d; \mathbb{R}^d)$.

Proof. The velocity field V satisfies assumptions (H1)–(H2), so by the Cauchy–Lipschitz theorem there exists a unique maximal $\mathscr{C}^1$ solution defined on some interval $I(t, z) \subset [0, T]$. For any $s \in I(t, z)$, integrating the ODE gives

$$Z(s, t, z) - z = \int_t^s V(\tau, Z(\tau, t, z))\, d\tau,$$

hence

$$|Z(s, t, z)| \leq |z| + \int_t^s |V(\tau, Z(\tau, t, z))|\, d\tau.$$

Applying assumption (H3) yields

$$|Z(s, t, z)| \leq |z| + \kappa \int_t^s (1 + |Z(\tau, t, z)|)\, d\tau.$$

Gronwall's lemma then gives the global bound

$$|Z(s, t, z)| \leq (|z| + \kappa T)\, e^{\kappa T},$$

which shows that the solution cannot blow up, so $\overline{I(t, z)} = [0, T]$. Continuous and differentiable dependence on the parameters (t, z) follows from another application of the Gronwall inequality to the variational equation. □

The flow map satisfies the following additional structural properties.

Theorem 2.2.3. *The flow $Z(s,t,z)$ satisfies:*

(i) ***Flow (semigroup) property:*** $Z(t_3,t_2,Z(t_2,t_1,z)) = Z(t_3,t_1,z)$ *for all* $t_1,t_2,t_3 \in [0,T]$.

(ii) ***Diffeomorphism:*** $Z(s,t,\cdot)$ *is a* $\mathscr{C}^1$*-diffeomorphism of* $\mathbb{R}^d$ *with inverse* $Z(s,t,\cdot)^{-1} = Z(t,s,\cdot)$.

(iii) ***Jacobian:*** *the determinant* $J(s,t,z) = \det(D_z Z(s,t,z))$ *satisfies*

$$\partial_s J = (\nabla_z \cdot V)\,(s, Z(s,t,z))\, J, \quad J(t,t,z) = 1,$$

which integrates to

$$J(s,t,z) = \exp\left(\int_t^s (\nabla_z \cdot V)(\tau, Z(\tau,t,z))\, d\tau\right).$$

Proof. (i). For fixed $t_1,t_2 \in [0,T]$ and $z \in \mathbb{R}^d$, both maps

$$t_3 \mapsto Z(t_3,t_2,Z(t_2,t_1,z)) \quad \text{and} \quad t_3 \mapsto Z(t_3,t_1,z)$$

are integral curves of the vector field V that pass through $Z(t_2,t_1,z)$ at time t_2, resp. through z at time t_1. Both pass through the same point at $t_3 = t_2$: indeed $Z(t_2,t_1,z)$ in the first case, and $Z(t_2,t_1,z)$ in the second. By uniqueness of the ODE, the two curves coincide.

(ii). Setting $t_3 = t_1$ in (i):

$$Z(t_1,t_2,Z(t_2,t_1,z)) = Z(t_1,t_1,z) = z.$$

Hence $z \mapsto Z(t_2,t_1,z)$ is invertible with inverse $Z(t_1,t_2,\cdot)$, which belongs to $\mathscr{C}^1(\mathbb{R}^d)$ by Theorem 2.2.2.

(iii). Define $J(s,t,z) = \det(D_z Z(s,t,z))$. Since differentiation and taking the determinant commute smoothly, we use the identity for the differential of the determinant: for $A \in \mathrm{GL}_d(\mathbb{R})$,

$$B \mapsto (D\det A)\cdot B = \det(A)\ \mathrm{tr}(A^{-1}B).$$

Differentiating J with respect to s and using $\partial_s(D_z Z) = D_z(\partial_s Z) = D_z(V(s,Z(s,t,z)))$, we obtain

$$\begin{aligned}
\partial_s J(s,t,z) &= J(s,t,z)\ \mathrm{tr}\big((D_z Z)^{-1}(s,t,z)\ \nabla_z V(s,Z(s,t,z))\, D_z Z(s,t,z)\big)\\
&= J(s,t,z)\ \mathrm{tr}(D_z V(s,Z(s,t,z)))\\
&= J(s,t,z)\,(\nabla_z \cdot V)(s,Z(s,t,z)),
\end{aligned}$$

where we used the cyclic property of the trace. The initial condition $J(t,t,z) = 1$ follows from $D_z Z(t,t,z) = I_d$. □

Remark 2.2.4. The Jacobian $J(s,t,z) = \det(D_z Z(s,t,z))$ measures the local volume distortion of the flow. When $\nabla_z \cdot V = 0$ we have $J \equiv 1$: the flow is *incompressible* and the non-conservative equation coincides with the conservative form

$$\begin{cases} \partial_t f + \nabla_z \cdot (V(t,z)\, f) = 0,\\ f(0,z) = f^{\mathrm{in}}(z),\end{cases}$$

which implies conservation of total mass:

$$\int_{\mathbb{R}^d} f(t)\, dz = \int_{\mathbb{R}^d} f^{\mathrm{in}}\, dz.$$

When $J \not\equiv 1$ the flow is compressible and mass is not preserved.

We now use the flow to construct the solution of the non-conservative equation.

Theorem 2.2.5. *Under assumptions* (H1)–(H2)*, the unique* $\mathscr{C}^1$ *solution of the non-conservative transport equation is*

$$f(t,z) = f^{\mathrm{in}}(Z(0,t,z))\,,$$

and satisfies $\|f(t)\|_{L^\infty} = \|f^{\mathrm{in}}\|_{L^\infty}$.

Proof. We follow the same strategy as in the free transport case.

Uniqueness. Let $f(t,z)$ be a $\mathscr{C}^1$ solution. For any $(t,z) \in \mathbb{R}\times\mathbb{R}^d$,

$$\begin{aligned}\frac{d}{dt}f(t,Z(t,0,z)) &= \left(\partial_t f + \partial_s Z(s,0,z)\big|_{s=t}\cdot\nabla_z f\right)(t,Z(t,0,z))\\ &= \left(\partial_t f + V\cdot\nabla_z f\right)(t,Z(t,0,z)) = 0,\end{aligned}$$

since f is a solution. Hence $f(t,Z(t,0,z)) = f(0,z) = f^{\mathrm{in}}(z)$. Setting $y = Z(t,0,z)$ and inverting via $z = Z(0,t,y)$ yields

$$f(t,y) = f^{\mathrm{in}}(Z(0,t,y)),$$

which proves uniqueness.

Existence. The formula $f(t,z) = f^{\mathrm{in}}(Z(0,t,z))$ defines a $\mathscr{C}^1$ function. We verify that it solves the PDE by differentiating and using the identity

$$\partial_t Z(s,t,z) + \left(V(t,z)\cdot\nabla_z\right)Z(s,t,z) = 0.$$

To establish this identity, we differentiate the flow property provided in Theorem 2.2.3(i), as

$$Z(t_3,t_2,Z(t_2,t_1,z)) = Z(t_3,t_1,z)$$

with respect to t_2:

$$\left(\partial_t + V(t_2,Z(t_2,t_1,z))\cdot\nabla_z\right)Z(t_3,t_2,Z(t_2,t_1,z)) = 0.$$

Setting $t_2 = t_1 = t$ and $t_3 = s$ yields the claimed identity. □

2.3 Conservative equation

Conservative formulations play a central role in applications because they directly express the conservation of mass or probability. They are particularly well adapted to weak formulations and numerical discretizations based on conservation principles.

Let $V(t,z)\in\mathbb{R}^d$ be a time-dependent vector field and consider the Cauchy problem

$$\begin{cases}\partial_t\mu + \nabla_z\cdot(V(t,z)\,\mu) = 0,\\[2mm] \mu(0,z) = \mu^{\mathrm{in}}(z).\end{cases}$$

When μ^{in} is only a positive Radon measure or an $L^1(\mathbb{R}^d)$ function, we resort to a notion of weak solution. In the context of kinetic theory, it is natural to work with measure-valued solutions, since the distribution functions of particle systems are nonnegative by definition.

Definition 2.3.1 (Weak solution)**.** *A* weak solution *of the Cauchy problem for the conservative transport equation is an element* $\mu \in \mathscr{C}([0,T];\, w\text{-}\mathcal{M}(\mathbb{R}^d))$ *that satisfies the initial condition* $\mu(0) = \mu^{\mathrm{in}}$ *and, for every test function* $\varphi \in \mathscr{C}_c^1((0,T)\times\mathbb{R}^d)$,

$$\int_0^T\int_{\mathbb{R}^d}\left(\partial_t\varphi + V(t,z)\cdot\nabla_z\varphi\right)(t,z)\,\mu(t,dz)\,dt = 0.$$

Definition 2.3.2 (Push-forward of measures)**.** *Let μ be a positive Radon measure on $\mathbb{R}^d$ and let $\Phi : \mathbb{R}^d \to \mathbb{R}^d$ be a measurable map. The* push-forward *of μ by Φ is the measure $\Phi_{\#}\mu$ defined by*

$$(\Phi_{\#}\mu)(B) = \mu(\Phi^{-1}(B))$$

for every Borel set $B \subset \mathbb{R}^d$. Equivalently, for any bounded measurable function ψ,

$$\int_{\mathbb{R}^d} \psi(y)\,(\Phi_{\#}\mu)(dy) = \int_{\mathbb{R}^d} \psi(\Phi(x))\,\mu(dx).$$

Theorem 2.3.3. *Let $V \equiv V(t,z)$ satisfy* (H1)–(H2)*, and let $\mu^{\mathrm{in}} \in \mathcal{M}_+(\mathbb{R}^d)$ be a positive Radon measure on $\mathbb{R}^d$. Then the unique weak solution in the sense of Definition 2.3.1 is*

$$\mu(t) = Z(t,0,\cdot)_{\#}\,\mu^{\mathrm{in}},$$

that is, for every test function $\varphi \in \mathscr{C}(\mathbb{R}^d)$,

$$\int_{\mathbb{R}^d} \varphi(z)\,\mu(t,dz) = \int_{\mathbb{R}^d} \varphi(Z(t,0,y))\,\mu^{\mathrm{in}}(dy).$$

Proof. Existence. Set $\mu(t) = Z(t,0,\cdot)_{\#}\mu^{\mathrm{in}}$. Since $Z(t,0,\cdot)$ is a measurable map, $\mu(t) \in \mathcal{M}_+(\mathbb{R}^d)$. If μ^{in} has finite total mass, then

$$\int_{\mathbb{R}^d} \mu(t,dz) = \int_{\mathbb{R}^d} \mu^{\mathrm{in}}(dz),$$

by the definition of push-forward. Now let $\varphi \in \mathscr{C}^1_c((0,T) \times \mathbb{R}^d)$ and define

$$\mathcal{I}(t) = \int_{\mathbb{R}^d} \varphi(t, Z(t,0,y))\,\mu^{\mathrm{in}}(dy).$$

Differentiating under the integral sign (justified by the compact support of φ):

$$\begin{aligned}
\frac{\mathrm{d}\mathcal{I}}{\mathrm{d}t}(t) &= \int_{\mathbb{R}^d} (\partial_t\varphi + V\cdot\nabla_z\varphi)\,(t, Z(t,0,y))\,\mu^{\mathrm{in}}(dy)\\
&= \int_{\mathbb{R}^d} (\partial_t\varphi(t,z) + V(t,z)\cdot\nabla_z\varphi(t,z))\,\mu(t,dz).
\end{aligned}$$

Since φ is compactly supported in time, integrating over $[0,T]$ gives

$$0 = \mathcal{I}(T) - \mathcal{I}(0) = \int_0^T \int_{\mathbb{R}^d} (\partial_t\varphi + V\cdot\nabla_z\varphi)\,(t,z)\,\mu(t,dz)\,dt,$$

so μ is a weak solution.

Uniqueness. Let μ be any weak solution and define

$$\nu(t) := Z(0,t,\cdot)_{\#}\,\mu(t).$$

Fix $\psi \in \mathscr{C}^1_c(\mathbb{R}^d)$. For any $\chi \in \mathscr{C}^\infty_c(0,T)$, we compute

$$-\int_0^T \chi'(t)\left(\int_{\mathbb{R}^d} \psi(x)\,\nu(t,dx)\right)dt = -\int_0^T \chi'(t)\left(\int_{\mathbb{R}^d} \psi(Z(0,t,y))\,\mu(t,dy)\right)dt.$$

Using the key identity

$$(\partial_t + V(t,z)\cdot\nabla_z)\,\psi(Z(0,t,z)) = 0$$

obtained from Theorem 2.2.5), and the fact that μ is a weak solution, one obtains

$$\frac{d}{dt}\int_{\mathbb{R}^d}\psi(x)\,\nu(t,dx)=0\quad\text{in }\mathcal{D}'(0,T).$$

Hence $\nu(t)=\nu(0)=\mu^{\text{in}}$, which gives $\mu(t)=Z(t,0,\cdot)_{\#}\mu^{\text{in}}$.

To justify the differentiation and the integration by parts, we need the support of $(t,y)\mapsto\psi(Z(0,t,y))$ to be compact. This follows from the linear growth bound (H3): for all $s,t\in[0,T]$ and $y\in\mathbb{R}^d$,

$$|Z(s,t,y)|\le(|y|+\kappa T)\,e^{\kappa T},$$

which ensures that the support remains bounded. □

We finally characterise smooth solutions when the initial datum is a function.

Theorem 2.3.4. *Let V satisfy* (H1)–(H3) *and let $f^{\text{in}}\in\mathscr{C}^1(\mathbb{R}^d)$. Then the conservative transport equation*

$$\partial_t f+\nabla_z\cdot(V(t,z)\,f)=0,\quad f\big|_{t=0}=f^{\text{in}}$$

has a unique $\mathscr{C}^1$ solution given by

$$f(t,z)=f^{\text{in}}(Z(0,t,z))\,J(0,t,z),$$

where $J(s,t,z)=\det(D_zZ(s,t,z))$ is the Jacobian of the flow.

Proof. Uniqueness. Let g be a solution with zero initial datum and set $u(t)=g(t,Z(t,0,y))$. Differentiating and using the transport equation in conservative form gives

$$u'(t)=-u(t)\,(\operatorname{div}_z V)(t,Z(t,0,y)),\quad u(0)=0.$$

This is a linear ODE with a damping/amplifying coefficient and zero initial datum, so $u(t)=0$ for all t, hence $g\equiv 0$.

Existence. Set $f(t,z)=f^{\text{in}}(Z(0,t,z))\,J(0,t,z)$. Differentiating and using the evolution equation for J from Theorem 2.2.3 (*iii*) shows that f satisfies the PDE. □

Chapter 3

From Particle Systems to Mean Field PDEs

3.1 Introduction and motivation

Many important models in physics and applied mathematics describe the collective behavior of a large number of interacting particles. When the number of particles N tends to infinity, it is often fruitful to pass to a *mean-field limit*, in which the discrete system is approximated by a continuous evolution governed by a partial differential equation (PDE) for the empirical density. This passage from a microscopic, finite-dimensional description to a macroscopic, infinite-dimensional one is both mathematically rich and physically significant: it underlies kinetic theory, plasma physics, stellar dynamics, and the modeling of collective phenomena such as flocking or chemotaxis.

The fundamental idea is the following. When N is large and the particles are approximately "indistinguishable" (in the sense that the initial data is exchangeable, or more precisely *chaotic*), each individual particle feels only the *average* influence of all others. This average, or *mean field*, replaces the N-body interaction by a single effective external field, reducing the problem to a nonlinear PDE in one-particle phase space.

The goal of this chapter is to present a rigorous and general formalism for such limits in the context of classical mechanics. We focus on systems where the interaction between particles is given by a kernel K that satisfies Newton's third law (skew-symmetry). This framework encompasses important examples such as:

- **Point vortex dynamics**: in two-dimensional incompressible fluid mechanics, point vortices interact via the Biot–Savart kernel, and the mean-field limit yields the two-dimensional Euler equation in vorticity form.

- **Vlasov–Poisson system**: in plasma physics, charged particles interact via the Coulomb potential, and the mean-field limit yields the Vlasov–Poisson system. The same model (with an attractive sign) governs gravitational dynamics in stellar systems.

- **Cucker–Smale flocking**: models for the collective alignment of self-propelled agents.

We will derive the mean-field PDE, introduce the associated characteristic flow, establish well-posedness via a Picard iteration argument, prove stability via Dobrushin's estimates, and finally justify the convergence of the N-particle system to the mean-field limit.

The chapter is organized as follows. Section 3.2 sets up the general formalism and introduces the empirical measure. Section 3.3 constructs the mean-field characteristic flow and proves its well-posedness. Section 3.4 presents Dobrushin's stability estimate, which is the cornerstone of the convergence proof.

3.2 A general formalism for mean field limits in classical mechanics

3.2.1 The N-particle system

Consider a system of N particles whose states at time t are described by phase-space coordinates

$$\hat{z}_1(t), \ldots, \hat{z}_N(t) \in \mathbb{R}^d.$$

Here d depends on the model: for point vortices in two dimensions, $d = 2$ and $\hat{z}_i$ is the position of the i-th vortex; for the Vlasov–Poisson system, $d = 6$ and $\hat{z}_i = (x_i, v_i)$ encodes both the position and the velocity of particle i. The interaction between particles i and j is encoded in a kernel $K : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$.

The N-particle dynamics are governed by the system of ODEs

$$\frac{d\hat{z}_i}{dt}(t) = \frac{1}{N} \sum_{\substack{j=1 \\ j \neq i}}^{N} K(\hat{z}_i(t), \hat{z}_j(t)), \qquad i = 1, \ldots, N. \tag{3.1}$$

The factor $\frac{1}{N}$ normalizes the interaction so that the total force on each particle remains of order $O(1)$ as $N \to \infty$; this is sometimes called the *mean-field scaling*.

3.2.2 Newton's third law and simplification

A crucial structural assumption is that the interaction respects Newton's third law:

$$K(z, z') = -K(z', z), \qquad \forall z, z' \in \mathbb{R}^d. \tag{3.2}$$

This is a skew-symmetry condition. Setting $z = z'$ in (3.2) immediately gives $K(z, z) = 0$, meaning that particles do not exert a force on themselves. Thanks to this vanishing of the diagonal, the self-interaction term ($j = i$) contributes nothing, and the system (3.1) simplifies to

$$\dot{z}_i(t) = \frac{1}{N} \sum_{j=1}^{N} K(z_i(t), z_j(t)), \qquad i = 1, \ldots, N. \tag{3.3}$$

This simplification is crucial: it allows us to include the $j = i$ term in the sum without changing the dynamics, which will later make the empirical measure formula cleaner.

3.2.3 The mean-field limit

In the mean-field regime ($N \to \infty$), if the particles are distributed according to a probability measure $f(t, dz)$, the empirical average converges formally to an integral:

$$\frac{1}{N} \sum_{j=1}^{N} K(z_i(t), z_j(t)) \;\longrightarrow\; \int_{\mathbb{R}^d} K(z_i(t), z')\, f(t, dz').$$

The heuristic behind this limit is the *law of large numbers*: when the z_j are approximately independent and identically distributed according to $f(t, dz)$, the empirical average concentrates around its expectation.

This leads to the *mean-field characteristic equation*

$$\frac{dz}{dt}(t) = \int_{\mathbb{R}^d} K(z(t), z')\, f(t, dz') =: \mathcal{K}[f](t, z(t)). \tag{3.4}$$

Each particle now moves under the influence of the *collective* field $\mathcal{K}[f]$ generated by the whole distribution f. The distribution f in turn evolves by the transport (or continuity) equation

$$\partial_t f + \nabla_z \cdot (\mathcal{K}[f]\, f) = 0, \tag{3.5}$$

where the nonlinear velocity field is

$$\mathcal{K}[f](t,z) := \int_{\mathbb{R}^d} K(z,z')\, f(t,dz'). \tag{3.6}$$

Equation (3.5) is *nonlinear* in f: the velocity field $\mathcal{K}[f]$ depends on f itself. This self-consistent structure is the hallmark of mean-field equations.

The PDE (3.5) is understood in the weak (distributional) sense: a locally finite measure f is a weak solution if, for every test function $\phi \in \mathscr{C}_c^1(\mathbb{R}^d)$ and a.e. t,

$$\frac{d}{dt}\int_{\mathbb{R}^d} \phi(z)\, f(t,dz) = \int_{\mathbb{R}^d} \mathcal{K}[f](t,z)\cdot \nabla\phi(z)\, f(t,dz). \tag{3.7}$$

This formulation requires only one derivative on ϕ, and makes sense even for measure-valued solutions f, which is important since the empirical measure μ_{Z_N} of the particle system is such a measure.

3.2.4 Assumptions on the kernel

We now impose precise regularity assumptions on the kernel that will be used throughout:

(HK1) **Skew-symmetry**: $K(z,z') = -K(z',z)$ for all $z,z' \in \mathbb{R}^d$.

(HK2) **Lipschitz regularity**: $K \in C^1(\mathbb{R}^d \times \mathbb{R}^d; \mathbb{R}^d)$ with bounded first derivatives, i.e. there exists a constant $L \geq 0$ such that

$$\sup_{z'} |\nabla_z K(z,z')| \leq L \qquad \text{and} \qquad \sup_{z} |\nabla_{z'} K(z,z')| \leq L.$$

These conditions have several immediate consequences:

1. From (HK1), K is Lipschitz in each variable uniformly in the other, with Lipschitz constant L:
$$|K(z,z') - K(w,z')| \leq L|z-w|, \quad |K(z,z') - K(z,w')| \leq L|z'-w'|.$$

2. From (HK1) and $K(z,z)=0$, together with (HK2), K grows at most linearly:
$$|K(z,z')| \leq L(|z|+|z'|).$$

3. Consequently, the associated integral operator $\mathcal{K}[\cdot]$ extends naturally to the space of probability measures with finite first moment
$$\mathcal{P}_1(\mathbb{R}^d) := \left\{ \mu \in \mathcal{P}(\mathbb{R}^d) \;\middle|\; \int_{\mathbb{R}^d} |z|\, \mu(dz) < \infty \right\},$$
and the resulting velocity field $\mathcal{K}[\mu]$ grows at most linearly in z for any $\mu \in \mathcal{P}_1(\mathbb{R}^d)$.

3.2.5 The empirical measure

A central object connecting the microscopic and macroscopic descriptions is the empirical measure.

Definition 3.2.1 (Empirical measure)**.** *To an N-tuple $Z_N = (z_1, \ldots, z_N) \in (\mathbb{R}^d)^N$, one associates the* empirical measure

$$\mu_{Z_N} := \frac{1}{N} \sum_{j=1}^{N} \delta_{z_j} \;\in\; \mathcal{P}(\mathbb{R}^d).$$

Intuitively, μ_{Z_N} places a point mass of weight $\frac{1}{N}$ at each particle's location. For any continuous function ϕ, we have

$$\int_{\mathbb{R}^d} \phi(z)\, \mu_{Z_N}(dz) = \frac{1}{N} \sum_{j=1}^{N} \phi(z_j),$$

which is precisely the empirical average. The key observation, made precise in Theorem 3.2.2 below, is that $\mu_{Z_N(t)}$ satisfies the mean-field PDE (3.5) exactly, as a measure-valued solution, whenever $Z_N(t)$ satisfies the ODE system (3.3).

With the assumptions above, one arrives at the following existence and uniqueness result for the N-body ODE system and its associated empirical measure.

Theorem 3.2.2. *Assume (HK1)–(HK2). The Cauchy problem for the N-particle ODE system*

$$\begin{cases} \dfrac{dz_i}{dt}(t) \;=\; \dfrac{1}{N} \displaystyle\sum_{j=1}^{N} K(z_i(t), z_j(t)), \qquad i = 1, \ldots, N, \\ z_i(0) = z_i^{in}, \end{cases}$$

has a unique global $\mathscr{C}^1$ solution on $\mathbb{R}$. Moreover, the empirical measure $\mu_{Z_N(t)}$ is a weak solution of the mean-field PDE

$$\begin{cases} \partial_t \mu + \nabla_z \cdot (\mathcal{K}[\mu]\, \mu) = 0, \\ \mu|_{t=0} = \mu^{\mathrm{in}} := \mu_{Z_N^{in}}, \end{cases}$$

in the sense of (3.7)*.*

Proof. The right-hand side of the ODE is a $\mathscr{C}^1$ function of $(z_1, \ldots, z_N)$. Moreover, assumption (HK2) gives the linear growth bound

$$|K(z, z')| \le L(|z| + |z'|),$$

so that the vector field satisfies the hypotheses of the Cauchy–Lipschitz theorem with linear growth, as developed in Chapter 2. Specifically, the right-hand side satisfies

$$\left| \frac{1}{N} \sum_{j=1}^{N} K(z_i, z_j) \right| \le \frac{L}{N} \sum_{j=1}^{N} (|z_i| + |z_j|) \le L \left(|z_i| + \frac{1}{N} \sum_{j=1}^{N} |z_j| \right),$$

and one checks that the growth in $(z_1, \ldots, z_N)$ is indeed at most linear. This implies both local existence and uniqueness, and global existence (no finite-time blow-up).

For the second part, let $\varphi \in \mathscr{C}_c^1(\mathbb{R}^d)$ be a test function. We compute the time derivative of $\int \varphi \, d\mu_{Z_N(t)}$ directly:

$$\begin{aligned}
\frac{d}{dt}\int_{\mathbb{R}^d} \varphi(z)\, \mu_{Z_N(t)}(dz) &= \frac{1}{N}\sum_{i=1}^N \frac{d}{dt}\varphi(z_i(t)) \\
&= \frac{1}{N}\sum_{i=1}^N \nabla\varphi(z_i(t)) \cdot \dot{z}_i(t) \\
&= \frac{1}{N}\sum_{i=1}^N \nabla\varphi(z_i(t)) \cdot \frac{1}{N}\sum_{j=1}^N K(z_i(t), z_j(t)) \\
&= \int_{\mathbb{R}^d} \nabla\varphi(z) \cdot \mathcal{K}[\mu_{Z_N(t)}](z)\, \mu_{Z_N(t)}(dz),
\end{aligned}$$

where in the last step we recognized the double empirical sum as a double integral against $\mu_{Z_N(t)} \otimes \mu_{Z_N(t)}$, and used the definition (3.6) of $\mathcal{K}$. This is exactly the weak formulation (3.7), so $\mu_{Z_N(t)}$ is a weak solution of the mean-field PDE.

Alternatively, one observes that $\mu_{Z_N(t)}$ is the pushforward of $\mu_{Z_N^{in}}$ under the flow map

$$Z(t, \cdot, \mu_{Z_N^{in}}) : \mathbb{R}^d \to \mathbb{R}^d,$$

and the weak solution property follows from the chain of equalities:

$$\int_{\mathbb{R}^d} \varphi(z)\mu_{Z_N}(t, dz) = \int_{\mathbb{R}^d} \varphi(Z(t, y))\mu_{Z_N}^{in}(dy),$$

which holds since each particle follows the characteristic flow. □

Remark 3.2.3. Theorem 3.2.2 says that the empirical measure provides an *exact* solution of the mean-field PDE, with no approximation error. The approximation comes later: we compare $\mu_{Z_N(t)}$ with a smooth solution $f(t)$ of (3.5) when the initial data $\mu_{Z_N^{in}}$ is close to f^{in} in the Monge–Kantorovich distance.

3.3 The mean field characteristic flow

The goal of this section is to construct a global solution to the mean-field PDE (3.5) via the method of characteristics. The idea is to solve the characteristic ODE (3.4) for all initial conditions simultaneously, yielding a family of trajectories that together encode the full solution. The main difficulty is the self-referential nature of the problem: the velocity field $\mathcal{K}[\mu(t)]$ depends on the solution $\mu(t)$, which is itself defined as the pushforward of μ^{in} by the very flow we are trying to construct.

We address this by a fixed-point (Picard iteration) argument in a suitable function space, which is a natural variant of the classical proof of the Cauchy–Lipschitz theorem.

Theorem 3.3.1. *Assume (HK1)–(HK2). For each initial position $\zeta^{in} \in \mathbb{R}^d$ and each initial measure $\mu^{in} \in \mathcal{P}_1(\mathbb{R}^d)$, there exists a unique $\mathscr{C}^1$ solution*

$$t \mapsto Z(t, \zeta^{in}, \mu^{in}) \in \mathbb{R}^d$$

of the coupled mean-field characteristic system

$$\begin{cases}
\partial_t Z(t, \zeta^{in}, \mu^{in}) = \mathcal{K}[\mu(t)]\big(Z(t, \zeta^{in}, \mu^{in})\big), \\
\mu(t) \ = \ Z(t, \cdot, \mu^{in})_{\#}\mu^{in}, \\
Z(0, \zeta^{in}, \mu^{in}) = \zeta^{in}.
\end{cases} \tag{3.8}$$

Here $Z(t,\cdot,\mu^{in})_{\#}\mu^{in}$ denotes the pushforward of μ^{in} by the map $\zeta \mapsto Z(t,\zeta,\mu^{in})$.

Remark 3.3.2. The second equation in (3.8) expresses the self-consistency condition: $\mu(t)$ is the distribution of particles at time t, obtained by transporting μ^{in} along the flow Z. This turns (3.8) into a nonlinear ODE for Z as a function of the initial label ζ.

Proof. Let $\zeta^{in} \in \mathbb{R}^d$ and $\mu^{in} \in \mathcal{P}_1(\mathbb{R}^d)$. Denote the first moment of the initial measure by

$$C_1 := \int_{\mathbb{R}^d} |z| \, \mu^{in}(dz) < \infty.$$

The strategy is to rewrite the coupled system (3.8) as a fixed-point equation for the function $\zeta \mapsto Z(t,\zeta)$ in a suitable Banach space, then apply the Banach fixed-point theorem via a Picard iteration.

Step 1: Function space. Define the Banach space

$$X := \left\{ v \in \mathscr{C}(\mathbb{R}^d;\mathbb{R}^d) \text{ such that } \|v\|_X := \sup_{z\in\mathbb{R}^d} \frac{|v(z)|}{1+|z|} < \infty \right\}.$$

Elements of X are continuous functions that grow at most linearly. The norm $\|\cdot\|_X$ measures the linear growth rate. We look for the flow map $Z(t,\cdot) \in X$ for each fixed t.

Step 2: Lipschitz estimate for the interaction operator. By assumption (HK2) on the interaction kernel K, for each v, $w \in X$ and each $z \in \mathbb{R}^d$:

$$\begin{aligned}
&\left| \int_{\mathbb{R}^d} K(v(z),v(z'))\,\mu^{in}(dz') - \int_{\mathbb{R}^d} K(w(z),w(z'))\,\mu^{in}(dz') \right| \\
&\quad \le L \int_{\mathbb{R}^d} \big(|v(z)-w(z)| + |v(z')-w(z')|\big)\,\mu^{in}(dz') \\
&\quad \le L\|v-w\|_X(1+|z|) + L\|v-w\|_X \int_{\mathbb{R}^d}(1+|z'|)\mu^{in}(dz') \\
&\quad = L\|v-w\|_X(1+|z|+1+C_1) \\
&\quad \le L\|v-w\|_X(2+C_1)(1+|z|),
\end{aligned}$$

where in the third line we used that $|v(z)-w(z)| \le \|v-w\|_X(1+|z|)$ and $\int(1+|z'|)\mu^{in}(dz') = 1 + C_1$. Dividing by $(1+|z|)$, this gives the operator Lipschitz bound

$$\left\| \int_{\mathbb{R}^d} K(v(\cdot),v(z'))\mu^{in}(dz') - \int_{\mathbb{R}^d} K(w(\cdot),w(z'))\mu^{in}(dz') \right\|_X \le L(2+C_1)\|v-w\|_X. \tag{3.9}$$

Step 3: Picard iteration. Define a sequence $(Z_n)_{n\ge 0}$ of approximate flows by the iteration:

$$\begin{cases} Z_{n+1}(t,\zeta) = \zeta + \displaystyle\int_0^t \int_{\mathbb{R}^d} K(Z_n(s,\zeta),Z_n(s,\zeta'))\,\mu^{in}(d\zeta')\,ds, \\ Z_0(t,\zeta) \ = \ \zeta. \end{cases} \tag{3.10}$$

The zeroth iterate $Z_0(t,\zeta) = \zeta$ is constant in time: each particle stays at its initial position. Each subsequent iterate Z_{n+1} is obtained by integrating in time the interaction force computed along the previous iterate Z_n.

Using (3.9), one checks by induction that for each $n \in \mathbb{N}$:

$$\|Z_{n+1}(t,\cdot) - Z_n(t,\cdot)\|_X \le \frac{((2+C_1)L|t|)^n}{n!}\|Z_1(t,\cdot) - Z_0(t,\cdot)\|_X.$$

Indeed, for the base case $n = 0$: the difference $Z_1(t,\zeta) - Z_0(t,\zeta) = Z_1(t,\zeta) - \zeta$ satisfies

$$\begin{aligned}|Z_1(t,\zeta) - \zeta| &= \left|\int_0^t \int_{\mathbb{R}^d} K(\zeta,\zeta')\,\mu^{in}(d\zeta')\,ds\right| \\ &\le \int_0^{|t|} \int_{\mathbb{R}^d} L(|\zeta| + |\zeta'|)\,\mu^{in}(d\zeta')\,ds \\ &= L(|\zeta| + C_1)|t| \\ &\le L(1 + C_1)(1 + |\zeta|)|t|,\end{aligned}$$

so that $\|Z_1(t,\cdot) - Z_0(t,\cdot)\|_X \le L(1 + C_1)|t|$, which is finite. The induction step then follows by applying (3.9) at each level.

Step 4: Convergence. Summing the geometric-like bound over n:

$$\sum_{n=0}^{\infty} \|Z_{n+1}(t,\cdot) - Z_n(t,\cdot)\|_X \le L(1 + C_1)|t| \sum_{n=0}^{\infty} \frac{((2 + C_1)L|t|)^n}{n!} = L(1 + C_1)|t|\, e^{(2+C_1)L|t|} < \infty,$$

so $(Z_n(t,\cdot))_{n\ge0}$ is a Cauchy sequence in X for each fixed t. Moreover, the convergence is uniform on every bounded time interval $[-\tau,\tau]$. The limit

$$Z(t,\cdot) := \lim_{n\to\infty} Z_n(t,\cdot) \;\in\; X$$

satisfies the integral equation

$$Z(t,\zeta) = \zeta + \int_0^t \int_{\mathbb{R}^d} K(Z(s,\zeta), Z(s,\zeta'))\,\mu^{in}(d\zeta')\,ds \tag{3.11}$$

for all $t \in \mathbb{R}$ and all $\zeta \in \mathbb{R}^d$, obtained by passing to the limit in (3.10).

Step 5: Uniqueness. If Z and $\widetilde{Z} \in \mathscr{C}(\mathbb{R}; X)$ both satisfy the integral equation (3.11), then their difference satisfies

$$Z(t,\zeta) - \widetilde{Z}(t,\zeta) = \int_0^t \int_{\mathbb{R}^d} \Big(K(Z(s,\zeta), Z(s,\zeta')) - K(\widetilde{Z}(s,\zeta), \widetilde{Z}(s,\zeta'))\Big)\,\mu^{in}(d\zeta')\,ds.$$

Applying (3.9):

$$\|Z(t,\cdot) - \widetilde{Z}(t,\cdot)\|_X \le L(2 + C_1)\left|\int_0^t \|Z(s,\cdot) - \widetilde{Z}(s,\cdot)\|_X\,ds\right|.$$

By Gronwall's inequality, this implies $\|Z(t,\cdot) - \widetilde{Z}(t,\cdot)\|_X = 0$ for all t, hence $Z = \widetilde{Z}$. Uniqueness is established.

Step 6: $\mathscr{C}^1$ regularity. Since $Z \in \mathscr{C}(\mathbb{R}; X)$ and $K \in \mathscr{C}^1(\mathbb{R}^d \times \mathbb{R}^d, \mathbb{R}^d)$ satisfies (HK2), and $\mu^{in} \in \mathcal{P}_1(\mathbb{R}^d)$, the integrand

$$s \mapsto \int_{\mathbb{R}^d} K(Z(s,\zeta), Z(s,\zeta'))\,\mu^{in}(d\zeta')$$

is continuous on $\mathbb{R}$ for each ζ. By the fundamental theorem of calculus, differentiating (3.11) with respect to t shows that $t \mapsto Z(t,\zeta)$ is of class $\mathscr{C}^1$ on $\mathbb{R}$ and satisfies the ODE

$$\begin{cases} \partial_t Z(t,\zeta) = \displaystyle\int_{\mathbb{R}^d} K(Z(t,\zeta), Z(t,\zeta'))\,\mu^{in}(d\zeta'), \\ Z(0,\zeta) = \zeta. \end{cases}$$

Step 7: Identification with the mean-field system. Substituting $z' = Z(t, \zeta')$ in the integral (i.e., performing the change of variables from the label ζ' to the current position z'):

$$\int_{\mathbb{R}^d} K(Z(t,\zeta), Z(t,\zeta'))\, \mu^{in}(d\zeta') = \int_{\mathbb{R}^d} K(Z(t,\zeta), z')\, \big(Z(t,\cdot)_{\#}\mu^{in}\big)(dz') = \mathcal{K}[\mu(t)](Z(t,\zeta)),$$

where $\mu(t) = Z(t,\cdot)_{\#}\mu^{in}$. This confirms that Z solves (3.8). □

Remark 3.3.3. The flow $Z(t, \zeta^{in}, \mu^{in})$ lives in the *single-particle* phase space $\mathbb{R}^d$, not in the N-particle phase space $(\mathbb{R}^d)^N$. For any solution $Z_N(t)$ of the N-particle system one has the consistency relation

$$z_i(t) = Z\big(t, z_i^{in}, \mu_{Z_N^{in}}\big), \qquad i = 1, \ldots, N.$$

This is because the N-particle empirical measure $\mu_{Z_N^{in}}$ serves as a valid initial measure for the mean-field flow, and each particle simply follows the corresponding characteristic.

3.4 Dobrushin's stability estimate

We now turn to the key quantitative stability result for the mean-field flow. This estimate, due to Dobrushin [?], will simultaneously imply well-posedness of the mean-field PDE and convergence of the particle system.

The right metric on the space of probability measures for this problem is the Monge–Kantorovich (or Wasserstein-1) distance, which we now define.

3.4.1 The Monge–Kantorovich distance

Given $\mu, \nu \in \mathcal{P}_1(\mathbb{R}^d)$, let $\Pi(\mu,\nu)$ denote the set of *transport plans* (or *couplings*) between μ and ν: these are Borel probability measures $\pi \in \mathcal{P}(\mathbb{R}^d \times \mathbb{R}^d)$ with marginals μ and ν, i.e.

$$\int_{\mathbb{R}^d\times\mathbb{R}^d} (\phi(x) + \psi(y))\, \pi(dx\, dy) = \int_{\mathbb{R}^d} \phi(x)\, \mu(dx) + \int_{\mathbb{R}^d} \psi(y)\, \nu(dy)$$

for all bounded continuous ϕ, ψ.

Definition 3.4.1. *For each μ, $\nu \in \mathcal{P}_1(\mathbb{R}^d)$, the* Monge–Kantorovich distance $\mathrm{dist}_1(\mu,\nu)$ *between μ and ν is defined by*

$$\mathrm{dist}_1(\mu,\nu) := \inf_{\pi\in\Pi(\mu,\nu)} \int_{\mathbb{R}^d\times\mathbb{R}^d} |x-y|\, \pi(dx\, dy). \tag{3.12}$$

By the Kantorovich–Rubinstein duality theorem (see Villani [?]), this equals

$$\mathrm{dist}_1(\mu,\nu) = \sup_{\substack{\phi\in\mathrm{Lip}(\mathbb{R}^d)\\ \mathrm{Lip}(\phi)\le 1}} \left| \int_{\mathbb{R}^d} \phi(z)\, \mu(dz) - \int_{\mathbb{R}^d} \phi(z)\, \nu(dz) \right|.$$

The intuition behind (3.12) is as follows: a coupling π specifies a joint distribution of pairs (x,y) with $x \sim \mu$ and $y \sim \nu$; the cost $\int |x-y|\, d\pi$ measures the expected distance between paired points. The infimum over all couplings finds the most efficient pairing (the optimal transport plan), and $\mathrm{dist}_1(\mu,\nu)$ is the minimal expected distance. The dual formulation gives a characterization in terms of integration against Lipschitz-1 test functions: two measures are close in dist_1 if and only if they give similar values to all 1-Lipschitz observables.

3.4.2 Dobrushin's key computation

The mean field characteristic flow contains all the relevant information about both the mean field PDE and the N-particle ODE system. Dobrushin's approach is based on proving stability of the mean field characteristic flow $Z(t, \zeta^{in}, \mu^{in})$ jointly in the initial position ζ^{in} and the initial distribution μ^{in}. The Monge–Kantorovich distance is precisely the right tool to measure this stability.

Let ζ_1^{in}, $\zeta_2^{in} \in \mathbb{R}^d$, and μ_1^{in}, $\mu_2^{in} \in \mathcal{P}_1(\mathbb{R}^d)$. Let $\pi^{in} \in \Pi(\mu_1^{in}, \mu_2^{in})$ be any coupling of the two initial measures. Denote $Z_j(t,\zeta) = Z(t,\zeta,\mu_j^{in})$ for $j = 1, 2$, and $\mu_j(t) = Z_j(t,\cdot)_{\#}\mu_j^{in}$.

Starting from Duhamel's formula (i.e., the integral formulation (3.11)):

$$\begin{aligned}\mathcal{I} &:= Z(t,\zeta_1,\mu_1^{in}) - Z(t,\zeta_2,\mu_2^{in}) \\ &= \zeta_1 - \zeta_2 + \int_0^t \left(\int_{\mathbb{R}^d} K(Z_1(s,\zeta_1), z')\mu_1(s,dz') - \int_{\mathbb{R}^d} K(Z_2(s,\zeta_2), z')\mu_2(s,dz') \right) ds.\end{aligned}$$

The key step is to rewrite both integrals over $\mu_j(s) = Z_j(s,\cdot)_{\#}\mu_j^{in}$ as integrals over μ_j^{in}, and then combine them using the coupling π^{in}:

$$\begin{aligned}\mathcal{I} &= \zeta_1 - \zeta_2 \\ &\quad + \int_0^t \int_{\mathbb{R}^d\times\mathbb{R}^d} \Big(K(Z_1(s,\zeta_1), Z_1(s,\zeta_1')) - K(Z_2(s,\zeta_2), Z_2(s,\zeta_2')) \Big)\, \pi^{in}(d\zeta_1' d\zeta_2')\, ds.\end{aligned} \tag{3.13}$$

This last equality is the key observation in Dobrushin's argument: by integrating against the coupling π^{in} simultaneously, the differences between the two flows appear in a form directly amenable to the Lipschitz bound (HK2). The coupling π^{in} allows us to compare trajectories starting from ζ_1' under μ_1^{in} with trajectories starting from ζ_2' under μ_2^{in}, for "matched" pairs (ζ_1', ζ_2').

We introduce the average distance between the two flows, integrated against the coupling:

$$D[\pi^{in}](s) := \int_{\mathbb{R}^d\times\mathbb{R}^d} |Z_1(s,\zeta_1') - Z_2(s,\zeta_2')|\, \pi^{in}(d\zeta_1' d\zeta_2').$$

Taking absolute values in (3.13) and applying (HK2):

$$|\mathcal{I}| \le |\zeta_1 - \zeta_2| + L\int_0^t |Z_1(s,\zeta_1) - Z_2(s,\zeta_2)|\, ds + L\int_0^t D[\pi^{in}](s)\, ds.$$

Integrating both sides against $\pi^{in}(d\zeta_1\, d\zeta_2)$:

$$\begin{aligned}D[\pi^{in}](t) &\le D[\pi^{in}](0) + L\int_0^t D[\pi^{in}](s)\, ds + L\int_0^t D[\pi^{in}](s)\, ds \\ &= D[\pi^{in}](0) + 2L\int_0^t D[\pi^{in}](s)\, ds.\end{aligned}$$

Note that the two L-terms come respectively from the difference in the "base point" ζ_j and from the difference in the "background measure" μ_j^{in}; both contribute with the same rate L, yielding the factor $2L$.

By Gronwall's inequality:

$$D[\pi^{in}](t) \le D[\pi^{in}](0)\, e^{2L|t|}. \tag{3.14}$$

3.4.3 Dobrushin's stability theorem

We can now state and prove the main stability result.

Theorem 3.4.2 (Dobrushin's estimate)**.** *Assume (HK1)–(HK2). Let μ_1^{in}, $\mu_2^{in} \in \mathcal{P}_1(\mathbb{R}^d)$, and let*

$$\mu_j(t) \;=\; Z(t,\cdot,\mu_j^{in})_{\#}\mu_j^{in}, \qquad j \;=\; 1,\,2,$$

be the corresponding mean-field solutions. Then for all $t \in \mathbb{R}$,

$$\mathrm{dist}_1(\mu_1(t), \mu_2(t)) \le e^{2L|t|}\ \mathrm{dist}_1(\mu_1^{in}, \mu_2^{in}). \tag{3.15}$$

Proof. We use the coupling approach. For any $\pi^{in} \in \Pi(\mu_1^{in}, \mu_2^{in})$, consider the pushforward coupling

$$\pi(t) := \Phi_{t\#}\pi^{in}, \quad \text{where } \Phi_t(\zeta_1, \zeta_2) := (Z_1(t,\zeta_1), Z_2(t,\zeta_2)).$$

Since π^{in} has marginals μ_j^{in} and $Z_j(t,\cdot)_{\#}\mu_j^{in} = \mu_j(t)$, it follows that $\pi(t) \in \Pi(\mu_1(t), \mu_2(t))$.

By definition of dist_1 and using $\pi(t)$ as a (not necessarily optimal) coupling:

$$\begin{aligned}\mathrm{dist}_1(\mu_1(t), \mu_2(t)) &\le \int_{\mathbb{R}^d\times\mathbb{R}^d} |\zeta_1 - \zeta_2|\,\pi(t)(d\zeta_1\,d\zeta_2)\\ &= \int_{\mathbb{R}^d\times\mathbb{R}^d} |Z_1(t,\zeta_1) - Z_2(t,\zeta_2)|\,\pi^{in}(d\zeta_1\,d\zeta_2) = D[\pi^{in}](t).\end{aligned}$$

Taking the infimum over all $\pi^{in} \in \Pi(\mu_1^{in}, \mu_2^{in})$, we get

$$\begin{aligned}\mathrm{dist}_1(\mu_1(t), \mu_2(t)) &\le \inf_{\pi^{in}\in\Pi(\mu_1^{in},\mu_2^{in})} D[\pi^{in}](t)\\ &\le e^{2L|t|} \inf_{\pi^{in}\in\Pi(\mu_1^{in},\mu_2^{in})} D[\pi^{in}](0)\\ &= e^{2L|t|}\ \mathrm{dist}_1(\mu_1^{in}, \mu_2^{in}),\end{aligned}$$

where we used (3.14) in the second inequality, and recognized the last infimum as $\mathrm{dist}_1(\mu_1^{in}, \mu_2^{in})$ by definition. This concludes the proof. □

Remark 3.4.3. The bound (3.15) has several important consequences:

1. **Uniqueness of the mean-field PDE**: if $\mu_1^{in} = \mu_2^{in}$, then $\mathrm{dist}_1(\mu_1(t), \mu_2(t)) = 0$ for all t, so $\mu_1(t) = \mu_2(t)$. This gives uniqueness of weak measure-valued solutions.
2. **Stability**: solutions depend continuously (in fact, Lipschitz continuously in time) on the initial data in the dist_1 metric.
3. **Mean-field limit**: applying (3.15) with $\mu_1^{in} = \mu^{in}$ (a smooth density) and $\mu_2^{in} = \mu_{Z_N^{in}}$ (the empirical measure of initial particle positions) gives the quantitative convergence estimate for the mean-field limit.

Using the fact that the empirical measure is an exact measure solution to the mean-field PDE (Theorem 3.2.2), and combining it with Dobrushin's estimate (Theorem 3.4.2), one can now prove the mean-field limit: the empirical measure $\mu_{Z_N(t)}$ converges to the solution $\mu(t)$ of the mean-field PDE as $N \to \infty$, at a rate controlled by $\mathrm{dist}_1(\mu_{Z_N^{in}}, \mu^{in})$. This convergence is the content of the next section.