Aggregation versus Diffusion in Mathematical Biology

J. A. Carrillo

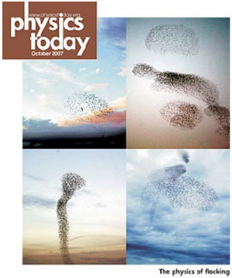
ICREA - Universitat Autònoma de Barcelona

UIMP, Santander, Spain, 2011

Outline

Outline

Swarming by Nature or by design?



Fish schools and Birds flocks.

Individual Based Models (Particle models)

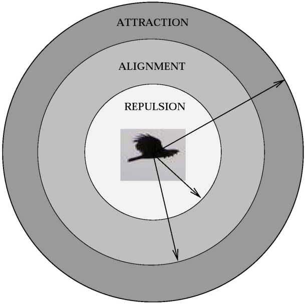
Swarming = Aggregation of agents of similar size and body type generally moving in a coordinated way.

Highly developed social organization: insects (locusts, ants, bees ...), fishes, birds, micro-organisms (myxo-bacteria, ...) and artificial robots for unmanned vehicle operation.

Interaction regions between individuals[a]

[a] Aoki, Helmerijk et al., Barbaro, Birnir et al.

- Repulsion Region: R_k.
- Attraction Region: A_k.
- Orientation Region: O_k.



Individual Based Models (Particle models)

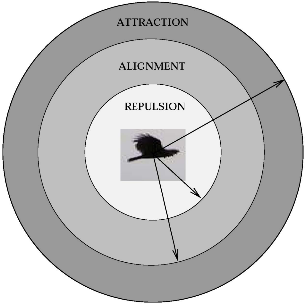
Swarming = Aggregation of agents of similar size and body type generally moving in a coordinated way.

Highly developed social organization: insects (locusts, ants, bees ...), fishes, birds, micro-organisms (myxo-bacteria, ...) and artificial robots for unmanned vehicle operation.

Interaction regions between individuals[a]

[a] Aoki, Helmerijk et al., Barbaro, Birnir et al.

- Repulsion Region: R_k.
- Attraction Region: A_k.
- Orientation Region: O_k.



Individual Based Models (Particle models)

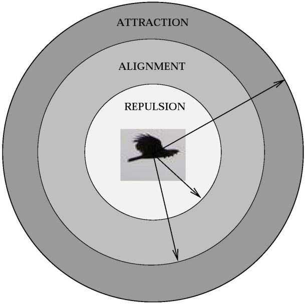
Swarming = Aggregation of agents of similar size and body type generally moving in a coordinated way.

Highly developed social organization: insects (locusts, ants, bees ...), fishes, birds, micro-organisms (myxo-bacteria, ...) and artificial robots for unmanned vehicle operation.

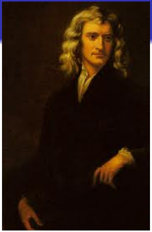
Interaction regions between individuals[a]

[a] Aoki, Helmerijk et al., Barbaro, Birnir et al.

- Repulsion Region: R_k.
- Attraction Region: A_k.
- Orientation Region: O_k.



2nd Order Model: Newton's like equations



D'Orsogna, Bertozzi et al. model (PRL 2006):

$$\begin{cases} \dfrac{dx_i}{dt} = v_i, \\[2ex] m\dfrac{dv_i}{dt} = (\alpha - \beta\,|v_i|^2)v_i - \displaystyle\sum_{j\neq i} \nabla U(|x_i - x_j|). \end{cases}$$

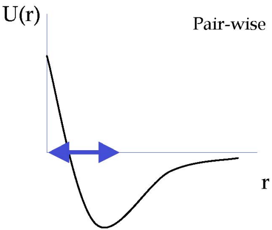
Model assumptions:

- Self-propulsion and friction terms determines an asymptotic speed of $\sqrt{\alpha/\beta}$.
- Attraction/Repulsion modeled by an effective pairwise potential $U(x)$.

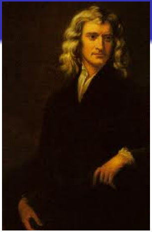
$$U(r) = -C_A e^{-r/\ell_A} + C_R e^{-r/\ell_R}.$$

One can also use Bessel functions in 2D and 3D to produce such a potential.

$C = C_R/C_A > 1$, $\ell = \ell_R/\ell_A < 1$ and $C\ell^2 < 1$:



2nd Order Model: Newton's like equations



D'Orsogna, Bertozzi et al. model (PRL 2006):

$$\begin{cases} \dfrac{dx_i}{dt} = v_i, \\ m\dfrac{dv_i}{dt} = (\alpha - \beta\,|v_i|^2)v_i - \displaystyle\sum_{j\neq i} \nabla U(|x_i - x_j|). \end{cases}$$

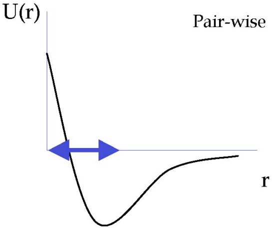
Model assumptions:

- Self-propulsion and friction terms determines an asymptotic speed of $\sqrt{\alpha/\beta}$.
- Attraction/Repulsion modeled by an effective pairwise potential $U(x)$.

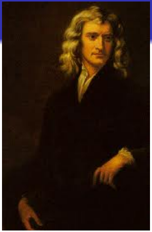
$$U(r) = -C_A e^{-r/\ell_A} + C_R e^{-r/\ell_R}.$$

One can also use Bessel functions in 2D and 3D to produce such a potential.

$C = C_R/C_A > 1$, $\ell = \ell_R/\ell_A < 1$ and $C\ell^2 < 1$:



2nd Order Model: Newton's like equations



D'Orsogna, Bertozzi et al. model (PRL 2006):

$$\begin{cases} \dfrac{dx_i}{dt} = v_i, \\ m\dfrac{dv_i}{dt} = (\alpha - \beta\,|v_i|^2)v_i - \displaystyle\sum_{j\neq i} \nabla U(|x_i - x_j|). \end{cases}$$

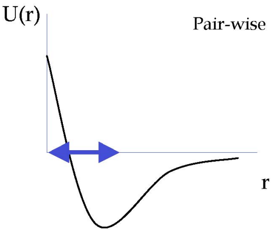
Model assumptions:

- Self-propulsion and friction terms determines an asymptotic speed of $\sqrt{\alpha/\beta}$.
- Attraction/Repulsion modeled by an effective pairwise potential $U(x)$.

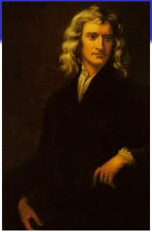
$$U(r) = -C_A e^{-r/\ell_A} + C_R e^{-r/\ell_R}.$$

One can also use Bessel functions in 2D and 3D to produce such a potential.

$C = C_R/C_A > 1$, $\ell = \ell_R/\ell_A < 1$ and $C\ell^2 < 1$:



2nd Order Model: Newton's like equations



D'Orsogna, Bertozzi et al. model (PRL 2006):

$$\begin{cases} \dfrac{dx_i}{dt} = v_i, \\ m\dfrac{dv_i}{dt} = (\alpha - \beta\,|v_i|^2)v_i - \displaystyle\sum_{j\neq i} \nabla U(|x_i - x_j|). \end{cases}$$

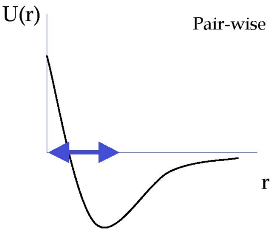
Model assumptions:

- Self-propulsion and friction terms determines an asymptotic speed of $\sqrt{\alpha/\beta}$.
- Attraction/Repulsion modeled by an effective pairwise potential $U(x)$.

$$U(r) = -C_A e^{-r/\ell_A} + C_R e^{-r/\ell_R}.$$

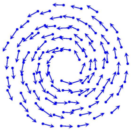
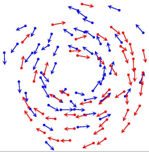
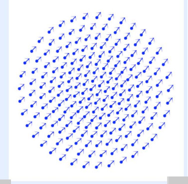
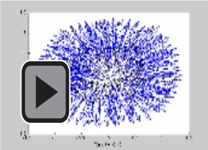
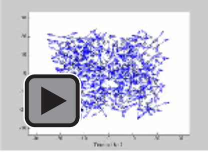
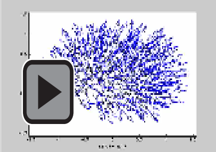
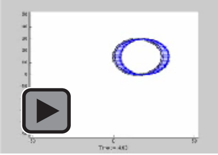
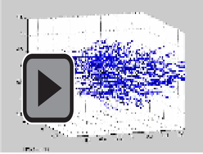
One can also use Bessel functions in 2D and 3D to produce such a potential.

$C = C_R/C_A > 1$, $\ell = \ell_R/\ell_A < 1$ and $C\ell^2 < 1$:



Model with an asymptotic speed

Typical patterns: milling, double milling or flocking:

1st Order Friction Model:

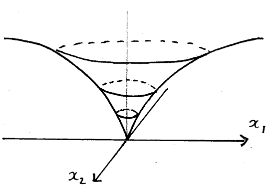
Edelshtein-Keshet, Mogilner (JMB 2000): Assume the variations of the velocity and speed are much smaller than spatial variations, then from Newton's equation:

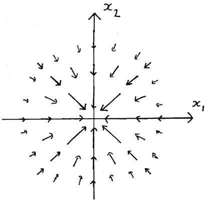
$$m\frac{d^2x_i}{d^2t} + \alpha\frac{dx_i}{dt} + \sum_{j\neq i}\nabla U(|x_i - x_j|) = 0$$

so finally, we obtain

$$\frac{dx_i}{dt} = -\sum_{j\neq i}\nabla U(|x_i - x_j|) \qquad \text{in the continuum setting} \Rrightarrow \qquad \begin{cases} \frac{\partial\rho}{\partial t} + \operatorname{div}(\rho v) = 0 \\ v = -\nabla U * \rho \end{cases}$$

Purely Aggregative Case: $U : \mathbb{R}^d \longrightarrow \mathbb{R}$





1st Order Friction Model:

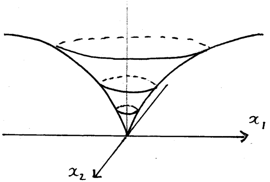
Edelshtein-Keshet, Mogilner (JMB 2000): Assume the variations of the velocity and speed are much smaller than spatial variations, then from Newton's equation:

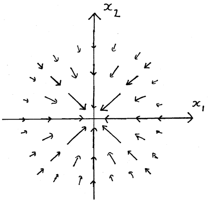
$$m\frac{d^2x_i}{d^2t} + \alpha\frac{dx_i}{dt} + \sum_{j\neq i}\nabla U(|x_i - x_j|) = 0$$

so finally, we obtain

$$\frac{dx_i}{dt} = -\sum_{j\neq i}\nabla U(|x_i - x_j|) \qquad \text{in the continuum setting} \Rrightarrow \qquad \begin{cases} \frac{\partial\rho}{\partial t} + \operatorname{div}(\rho v) = 0 \\ v = -\nabla U * \rho \end{cases}$$

Purely Aggregative Case: $U : \mathbb{R}^d \longrightarrow \mathbb{R}$





1st Order Friction Model:

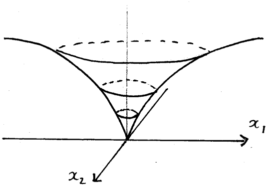
Edelshtein-Keshet, Mogilner (JMB 2000): Assume the variations of the velocity and speed are much smaller than spatial variations, then from Newton's equation:

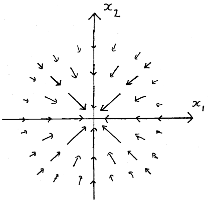
$$m\frac{d^2x_i}{d^2t} + \alpha\frac{dx_i}{dt} + \sum_{j\neq i} \nabla U(|x_i - x_j|) = 0$$

so finally, we obtain

$$\frac{dx_i}{dt} = -\sum_{j\neq i} \nabla U(|x_i - x_j|) \qquad \text{in the continuum setting} \Rightarrow \qquad \begin{cases} \frac{\partial\rho}{\partial t} + \operatorname{div}(\rho v) = 0 \\ v = -\nabla U * \rho \end{cases}$$

Purely Aggregative Case: $U : \mathbb{R}^d \longrightarrow \mathbb{R}$





1st Order Friction Model:

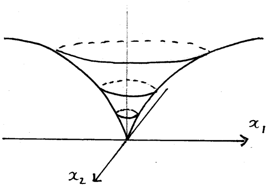
Edelshtein-Keshet, Mogilner (JMB 2000): Assume the variations of the velocity and speed are much smaller than spatial variations, then from Newton's equation:

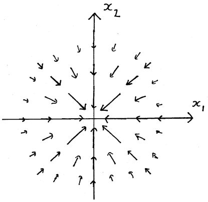
$$m\frac{d^2x_i}{d^2t} + \alpha\frac{dx_i}{dt} + \sum_{j\neq i} \nabla U(|x_i - x_j|) = 0$$

so finally, we obtain

$$\frac{dx_i}{dt} = -\sum_{j\neq i} \nabla U(|x_i - x_j|)$$ in the continuum setting $\Rightarrow$ $$\begin{cases} \frac{\partial \rho}{\partial t} + \mathrm{div}\,(\rho v) = 0 \\ v = -\nabla U * \rho \end{cases}$$

Purely Aggregative Case: $U : \mathbb{R}^d \longrightarrow \mathbb{R}$





Velocity consensus model

Cucker-Smale Model (IEEE Automatic Control 2007):

$$\begin{cases} \dfrac{dx_i}{dt} = v_i, \\ \dfrac{dv_i}{dt} = \displaystyle\sum_{j=1}^{N} a_{ij}\,(v_j - v_i)\,, \end{cases}$$

with the communication rate, $\gamma \geq 0$:

$$a_{ij} = a(|x_i - x_j|) = \frac{1}{(1 + |x_i - x_j|^2)^\gamma}\,.$$

Asymptotic flocking: $\gamma \leq 1/2$; Cucker-Smale.

Velocity consensus model

Cucker-Smale Model (IEEE Automatic Control 2007):

$$\begin{cases} \dfrac{dx_i}{dt} = v_i, \\ \dfrac{dv_i}{dt} = \displaystyle\sum_{j=1}^{N} a_{ij} \left(v_j - v_i \right), \end{cases}$$

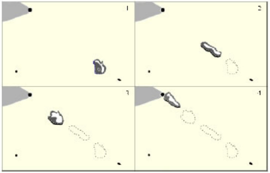
with the communication rate, $\gamma \geq 0$:

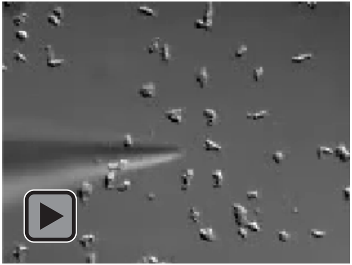
$$a_{ij} = a(|x_i - x_j|) = \frac{1}{(1 + |x_i - x_j|^2)^{\gamma}}.$$

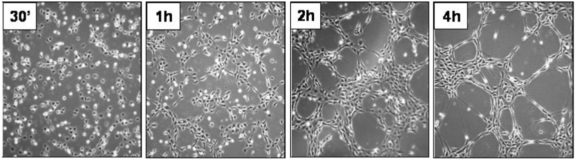
Asymptotic flocking: $\gamma \leq 1/2$; Cucker-Smale.

Outline

A reason for cell's motility: Chemotaxis







Cell movement and aggregation by chemical interaction.

KS System

Keller-Segel System:

Cells positions are assumed to fluctuate, in the sense of a Brownian motion, around the dominated trend to follow the trail of the largest concentration of chemoattractant:

$$x' = \nabla c(t,x) + \Gamma(t).$$

where $\Gamma(t)$ is a Wiener process with fixed variance. The chemoattractant diffuses spatially and is produced by the cells themselves.

$$\begin{cases} \dfrac{\partial \rho}{\partial t}(t,x) = \Delta \rho(t,x) - \nabla \cdot (\rho(t,x) \nabla c(t,x)) & x \in \mathbb{R}^2,\ t > 0, \\ \dfrac{\partial c}{\partial t}(t,x) - \Delta c(t,x) = \rho(t,x) - \alpha c(t,x) & x \in \mathbb{R}^2,\ t > 0, \\ \rho(t=0,x) = \rho_0 \geq 0 & x \in \mathbb{R}^2. \end{cases}$$

Patlak (1953), Keller-Segel (1971), Nanjundiah (1973).

KS System

Keller-Segel System:

Cells positions are assumed to fluctuate, in the sense of a Brownian motion, around the dominated trend to follow the trail of the largest concentration of chemoattractant:

$$x' = \nabla c(t,x) + \Gamma(t).$$

where $\Gamma(t)$ is a Wiener process with fixed variance. The chemoattractant diffuses spatially and is produced by the cells themselves.

$$\begin{cases} \dfrac{\partial \rho}{\partial t}(t,x) = \Delta \rho(t,x) - \nabla \cdot (\rho(t,x) \nabla c(t,x)) & x \in \mathbb{R}^2\,,\ t > 0\,, \\ \dfrac{\partial c}{\partial t}(t,x) - \Delta c(t,x) = \rho(t,x) - \alpha c(t,x) & x \in \mathbb{R}^2\,,\ t > 0\,, \\ \rho(t=0,x) = \rho_0 \geq 0 & x \in \mathbb{R}^2\,. \end{cases}$$

Patlak (1953), Keller-Segel (1971), Nanjundiah (1973).

KS System

Keller-Segel System:

Cells positions are assumed to fluctuate, in the sense of a Brownian motion, around the dominated trend to follow the trail of the largest concentration of chemoattractant:

$$x' = \nabla c(t,x) + \Gamma(t).$$

where $\Gamma(t)$ is a Wiener process with fixed variance. The chemoattractant diffuses spatially and is produced by the cells themselves.

$$\begin{cases} \dfrac{\partial \rho}{\partial t}(t,x) = \Delta \rho(t,x) - \nabla \cdot (\rho(t,x) \nabla c(t,x)) & x \in \mathbb{R}^2\,,\ t > 0\,, \\ \dfrac{\partial c}{\partial t}(t,x) - \Delta c(t,x) = \rho(t,x) - \alpha c(t,x) & x \in \mathbb{R}^2\,,\ t > 0\,, \\ \rho(t=0,x) = \rho_0 \geq 0 & x \in \mathbb{R}^2\,. \end{cases}$$

Patlak (1953), Keller-Segel (1971), Nanjundiah (1973).

PKS System

$$\begin{cases} \dfrac{\partial \rho}{\partial t}(t,x) = \Delta \rho(t,x) - \nabla \cdot (\rho(t,x) \nabla c(t,x)) & x \in \mathbb{R}^2 ,\ t > 0 , \\ c(t,x) = -\dfrac{1}{2\pi} \displaystyle\int_{\mathbb{R}^2} \log|x-y|\, \rho(t,y)\, dy & x \in \mathbb{R}^2 ,\ t > 0 , \\ \rho(x, t=0) = \rho_0 \geq 0 & x \in \mathbb{R}^2 . \end{cases}$$

Huge Literature: Horstmann reviews (2003& 2004), Perthame review (2004).

Conservations: mass and center of mass:

$$\frac{d}{dt} \int_{\mathbb{R}^2} \rho(t,x)\, dx = \frac{d}{dt} \int_{\mathbb{R}^2} x\, \rho(t,x)\, dx = 0 .$$

Free energy:

$$\mathcal{F}_{\mathrm{PKS}}[\rho] := \int_{\mathbb{R}^2} \rho(x) \log \rho(x)\, dx + \frac{1}{4\pi} \iint_{\mathbb{R}^2 \times \mathbb{R}^2} \rho(x)\, \rho(y)\, \log|x-y|\, dx\, dy$$

PKS System

$$\begin{cases} \dfrac{\partial \rho}{\partial t}(t,x) = \Delta \rho(t,x) - \nabla \cdot (\rho(t,x) \nabla c(t,x)) & x \in \mathbb{R}^2\,,\ t > 0\,, \\ c(t,x) = -\dfrac{1}{2\pi} \displaystyle\int_{\mathbb{R}^2} \log|x-y|\, \rho(t,y)\, dy & x \in \mathbb{R}^2\,,\ t > 0\,, \\ \rho(x, t=0) = \rho_0 \geq 0 & x \in \mathbb{R}^2\,. \end{cases}$$

Huge Literature: Horstmann reviews (2003& 2004), Perthame review (2004).

Conservations: mass and center of mass:

$$\frac{d}{dt} \int_{\mathbb{R}^2} \rho(t,x)\, dx = \frac{d}{dt} \int_{\mathbb{R}^2} x\, \rho(t,x)\, dx = 0\,.$$

Free energy:

$$\mathcal{F}_{\mathrm{PKS}}[\rho] := \int_{\mathbb{R}^2} \rho(x) \log \rho(x)\, dx + \frac{1}{4\pi} \iint_{\mathbb{R}^2 \times \mathbb{R}^2} \rho(x)\, \rho(y)\, \log|x-y|\, dx\, dy$$

PKS System

$$\begin{cases} \dfrac{\partial \rho}{\partial t}(t,x) = \Delta \rho(t,x) - \nabla \cdot \left(\rho(t,x) \nabla c(t,x)\right) & x \in \mathbb{R}^2\,,\ t > 0\,, \\ c(t,x) = -\dfrac{1}{2\pi} \displaystyle\int_{\mathbb{R}^2} \log|x-y|\, \rho(t,y)\, dy & x \in \mathbb{R}^2\,,\ t > 0\,, \\ \rho(x, t=0) = \rho_0 \geq 0 & x \in \mathbb{R}^2\,. \end{cases}$$

Huge Literature: Horstmann reviews (2003& 2004), Perthame review (2004).

Conservations: mass and center of mass:

$$\frac{d}{dt} \int_{\mathbb{R}^2} \rho(t,x)\, dx = \frac{d}{dt} \int_{\mathbb{R}^2} x\, \rho(t,x)\, dx = 0\,.$$

Free energy:

$$\mathcal{F}_{\mathrm{PKS}}[\rho] := \int_{\mathbb{R}^2} \rho(x) \log \rho(x)\, dx + \frac{1}{4\pi} \iint_{\mathbb{R}^2 \times \mathbb{R}^2} \rho(x)\, \rho(y)\, \log|x-y|\, dx\, dy$$

Nonlinear diffusion PKS system

Volume effects can be taken into account by considering nonlinear diffusion (Calvez& C., JMPA 2006) as:

$$\begin{cases} \dfrac{\partial \rho}{\partial t}(t,x) = \operatorname{div}\left[\nabla \rho^m(t,x) - \rho(t,x)\nabla c(t,x)\right] & t>0\,,\ x\in\mathbb{R}^d\,, \\ -\Delta c(t,x) = \rho(t,x)\,, & t>0\,,\ x\in\mathbb{R}^d\,, \end{cases}$$

Free Energy:

The corresponding free energy is

$$\mathcal{F}_m[\rho](t) := \int_{\mathbb{R}^d} \frac{\rho^m}{m-1}\,dx - \frac{c_d}{2}\iint_{\mathbb{R}^d\times\mathbb{R}^d} \frac{1}{|x-y|^{d-2}}\,\rho(t,x)\,\rho(t,y)\,dx\,dy$$

with $c_d^{-1} := (d-2)2\,\pi^{d/2}/\Gamma(d/2)$.

Diffusion to compensate exactly drift by scaling (Blanchet, C. & Laurençot, CVPDE 2008) is

$$m_d := 2(d-1)/d.$$

Nonlinear diffusion PKS system

Volume effects can be taken into account by considering nonlinear diffusion (Calvez& C., JMPA 2006) as:

$$\begin{cases} \dfrac{\partial \rho}{\partial t}(t,x) & = \operatorname{div}\left[\nabla \rho^m(t,x) - \rho(t,x)\nabla c(t,x)\right] \qquad t>0\,,\ x\in\mathbb{R}^d\,, \\ -\Delta c(t,x) & = \rho(t,x)\,, \qquad t>0\,,\ x\in\mathbb{R}^d\,, \end{cases}$$

Free Energy:

The corresponding free energy is

$$\mathcal{F}_m[\rho](t) := \int_{\mathbb{R}^d} \frac{\rho^m}{m-1}\,dx - \frac{c_d}{2}\iint_{\mathbb{R}^d\times\mathbb{R}^d} \frac{1}{|x-y|^{d-2}}\,\rho(t,x)\,\rho(t,y)\,dx\,dy$$

with $c_d^{-1} := (d-2)2\,\pi^{d/2}/\Gamma(d/2)$.

Diffusion to compensate exactly drift by scaling (Blanchet, C. & Laurençot, CVPDE 2008) is

$$m_d := 2(d-1)/d.$$

Nonlinear diffusion PKS system

Volume effects can be taken into account by considering nonlinear diffusion (Calvez& C., JMPA 2006) as:

$$\begin{cases} \dfrac{\partial \rho}{\partial t}(t,x) = \operatorname{div}\left[\nabla \rho^m(t,x) - \rho(t,x)\nabla c(t,x)\right] & t > 0\,,\ x \in \mathbb{R}^d\,, \\ -\Delta c(t,x) = \rho(t,x)\,, & t > 0\,,\ x \in \mathbb{R}^d\,, \end{cases}$$

Free Energy:

The corresponding free energy is

$$\mathcal{F}_m[\rho](t) := \int_{\mathbb{R}^d} \frac{\rho^m}{m-1}\,dx - \frac{c_d}{2}\iint_{\mathbb{R}^d\times\mathbb{R}^d} \frac{1}{|x-y|^{d-2}}\,\rho(t,x)\,\rho(t,y)\,dx\,dy$$

with $c_d^{-1} := (d-2)2\,\pi^{d/2}/\Gamma(d/2)$.

Diffusion to compensate exactly drift by scaling (Blanchet, C. & Laurençot, CVPDE 2008) is

$$m_d := 2(d-1)/d.$$

Outline

1. Motivations
   - Applied Mathematics: Collective Behavior Models
   - Applied Mathematics: Modelling Chemotaxis
   - Pure Mathematics: Gradient Flows

2. Outline of the course

3. Transversal Tool: Wasserstein Distances
   - Definition
   - Properties

General Entropy Functional[1]

$$\mathcal{F}[\rho] = \Pi[\rho] + \mathcal{V}[\rho] + \mathcal{U}[\rho]$$

with

$$\Pi[\rho] = \int_{\mathbb{R}^d} \pi(\rho(x))\, dx \quad \text{internal energy}$$

$$\mathcal{V}[\rho] = \int_{\mathbb{R}^d} V(x)\rho(x)\, dx \quad \text{confinement energy}$$

$$\mathcal{U}[\rho] = \frac{1}{2}\int_{\mathbb{R}^{2d}} U(x-y)\rho(x)\rho(y)\, dx\, dy \quad \text{interaction energy}$$

Let us write the formal gradient flow equation as before:

$$\frac{\partial \rho}{\partial t} = \operatorname{div}\left(\rho \nabla \frac{\delta \mathcal{F}}{\delta \rho}\right), \qquad (x \in \mathbb{R}^d, t > 0).$$

and the dissipation of entropy is defined as

$$\frac{d}{dt}\mathcal{F}[\rho] = -D[\rho] \quad \text{with} \quad D[\rho] = \int_{\mathbb{R}^d} |\xi|^2 \rho(x)\, dx,$$

with

$$\xi = \nabla\left[\pi'[\rho] + V + U * \rho\right] = \nabla \frac{\delta \mathcal{F}}{\delta \rho}.$$

[1] J.A. Carrillo, R.J. McCann & C. Villani, RMI 2003 & ARMA 2006.

General Entropy Functional[1]

$$\mathcal{F}[\rho] = \Pi[\rho] + \mathcal{V}[\rho] + \mathcal{U}[\rho]$$

with

$$\Pi[\rho] = \int_{\mathbb{R}^d} \pi(\rho(x))\, dx \quad \text{internal energy}$$

$$\mathcal{V}[\rho] = \int_{\mathbb{R}^d} V(x)\rho(x)\, dx \quad \text{confinement energy}$$

$$\mathcal{U}[\rho] = \frac{1}{2}\int_{\mathbb{R}^{2d}} U(x-y)\rho(x)\rho(y)\, dx\, dy \quad \text{interaction energy}$$

Let us write the formal gradient flow equation as before:

$$\frac{\partial \rho}{\partial t} = \operatorname{div}\left(\rho \nabla \frac{\delta \mathcal{F}}{\delta \rho}\right), \qquad (x \in \mathbb{R}^d, t > 0).$$

and the dissipation of entropy is defined as

$$\frac{d}{dt}\mathcal{F}[\rho] = -D[\rho] \quad \text{with} \quad D[\rho] = \int_{\mathbb{R}^d} |\xi|^2 \rho(x)\, dx,$$

with

$$\xi = \nabla\left[\pi'[\rho] + V + U * \rho\right] = \nabla \frac{\delta \mathcal{F}}{\delta \rho}.$$

[1] J.A. Carrillo, R.J. McCann & C. Villani, RMI 2003 & ARMA 2006.

General Entropy Functional[1]

$$\mathcal{F}[\rho] = \Pi[\rho] + \mathcal{V}[\rho] + \mathcal{U}[\rho]$$

with

$$\Pi[\rho] = \int_{\mathbb{R}^d} \pi(\rho(x))\,dx \quad \text{internal energy}$$

$$\mathcal{V}[\rho] = \int_{\mathbb{R}^d} V(x)\rho(x)\,dx \quad \text{confinement energy}$$

$$\mathcal{U}[\rho] = \frac{1}{2}\int_{\mathbb{R}^{2d}} U(x-y)\rho(x)\rho(y)\,dx\,dy \quad \text{interaction energy}$$

Let us write the formal gradient flow equation as before:

$$\frac{\partial \rho}{\partial t} = \operatorname{div}\left(\rho \nabla \frac{\delta \mathcal{F}}{\delta \rho}\right), \qquad (x \in \mathbb{R}^d, t > 0).$$

and the dissipation of entropy is defined as

$$\frac{d}{dt}\mathcal{F}[\rho] = -D[\rho] \quad \text{with} \quad D[\rho] = \int_{\mathbb{R}^d} |\xi|^2 \rho(x)\,dx,$$

with

$$\xi = \nabla\left[\pi'[\rho] + V + U * \rho\right] = \nabla \frac{\delta \mathcal{F}}{\delta \rho}.$$

[1] J.A. Carrillo, R.J. McCann & C. Villani, RMI 2003 & ARMA 2006.

General Entropy Functional

Formal Particular Examples:

- Aggregation Equations: $\pi(\rho) = 0$, $V(x) = 0$, U with possible singularity at the origin.-

$$\frac{\partial \rho}{\partial t}(t,x) = \operatorname{div}\left(\rho(t,x)\left[\nabla U * \rho\right](t,x)\right) \ .$$

- Patlak-Keller-Segel: $\pi(\rho) = \rho \log \rho$ or $\rho^m/(m-1)$, $V(x) = 0$, $U = -G$ with G the fundamental solution of $-\Delta$.-

$$\frac{\partial \rho}{\partial t} = \operatorname{div}\left(\rho\left[-G(x) * \rho + \rho \nabla \pi(\rho)\right]\right), \qquad (x \in \mathbb{R}^d, t > 0)$$

- Given U smooth (C^1) and convex (λ-convex), then the theory developed by Ambrosio-Gigli-Savaré (Birkhausser 2005) about gradient flows in the space of probability measures applies.
- In the first order models and in the PKS model, the interaction potential is not smooth at the origin or even singular. The convexity will help in some cases.

General Entropy Functional

Formal Particular Examples:

- Aggregation Equations: $\pi(\rho) = 0$, $V(x) = 0$, U with possible singularity at the origin.-

$$\frac{\partial \rho}{\partial t}(t,x) = \operatorname{div}\left(\rho(t,x)\left[\nabla U * \rho\right](t,x)\right) \; .$$

- Patlak-Keller-Segel: $\pi(\rho) = \rho \log \rho$ or $\rho^m/(m-1)$, $V(x) = 0$, $U = -G$ with G the fundamental solution of $-\Delta$.-

$$\frac{\partial \rho}{\partial t} = \operatorname{div}\left(\rho\left[-G(x) * \rho + \rho \nabla \pi(\rho)\right]\right), \qquad (x \in \mathbb{R}^d, t > 0)$$

- Given U smooth (C^1) and convex (λ-convex), then the theory developed by Ambrosio-Gigli-Savaré (Birkhausser 2005) about gradient flows in the space of probability measures applies.
- In the first order models and in the PKS model, the interaction potential is not smooth at the origin or even singular. The convexity will help in some cases.

General Entropy Functional

Formal Particular Examples:

- Aggregation Equations: $\pi(\rho) = 0$, $V(x) = 0$, U with possible singularity at the origin.-

$$\frac{\partial \rho}{\partial t}(t,x) = \operatorname{div}\left(\rho(t,x)\left[\nabla U * \rho\right](t,x)\right) .$$

- Patlak-Keller-Segel: $\pi(\rho) = \rho \log \rho$ or $\rho^m/(m-1)$, $V(x) = 0$, $U = -G$ with G the fundamental solution of $-\Delta$.-

$$\frac{\partial \rho}{\partial t} = \operatorname{div}\left(\rho\left[-G(x) * \rho + \rho \nabla \pi(\rho)\right]\right), \qquad (x \in \mathbb{R}^d, t > 0)$$

- Given U smooth (C^1) and convex (λ-convex), then the theory developed by Ambrosio-Gigli-Savaré (Birkhausser 2005) about gradient flows in the space of probability measures applies.
- In the first order models and in the PKS model, the interaction potential is not smooth at the origin or even singular. The convexity will help in some cases.

General Entropy Functional

Formal Particular Examples:

- Aggregation Equations: $\pi(\rho) = 0$, $V(x) = 0$, U with possible singularity at the origin.-

$$\frac{\partial \rho}{\partial t}(t,x) = \operatorname{div}\left(\rho(t,x)\left[\nabla U * \rho\right](t,x)\right) \ .$$

- Patlak-Keller-Segel: $\pi(\rho) = \rho \log \rho$ or $\rho^m/(m-1)$, $V(x) = 0$, $U = -G$ with G the fundamental solution of $-\Delta$.-

$$\frac{\partial \rho}{\partial t} = \operatorname{div}\left(\rho\left[-G(x) * \rho + \rho \nabla \pi(\rho)\right]\right), \qquad (x \in \mathbb{R}^d, t > 0)$$

- Given U smooth (C^1) and convex (λ-convex), then the theory developed by Ambrosio-Gigli-Savaré (Birkhausser 2005) about gradient flows in the space of probability measures applies.
- In the first order models and in the PKS model, the interaction potential is not smooth at the origin or even singular. The convexity will help in some cases.

Mathematical Questions:

- What are the continuum models associated to these systems as the number of individuals gets larger and larger? Mean-field limits.
- What is the good analytical framework to deal with the possible concentration of mass in finite/infinite time in space or in velocity? Stability of patterns?
- What is the good analytical framework to deal with particles and continuum solutions at the same time? We want to allow solutions with evolving absolutely continuous and atomic parts of the measures at the same time.
- How to deal with the competition between diffusion and aggregation? Are there important functional inequalities involved?

Mathematical Questions:

- What are the continuum models associated to these systems as the number of individuals gets larger and larger? Mean-field limits.
- What is the good analytical framework to deal with the possible concentration of mass in finite/infinite time in space or in velocity? Stability of patterns?
- What is the good analytical framework to deal with particles and continuum solutions at the same time? We want to allow solutions with evolving absolutely continuous and atomic parts of the measures at the same time.
- How to deal with the competition between diffusion and aggregation? Are there important functional inequalities involved?

Mathematical Questions:

- What are the continuum models associated to these systems as the number of individuals gets larger and larger? Mean-field limits.
- What is the good analytical framework to deal with the possible concentration of mass in finite/infinite time in space or in velocity? Stability of patterns?
- What is the good analytical framework to deal with particles and continuum solutions at the same time? We want to allow solutions with evolving absolutely continuous and atomic parts of the measures at the same time.
- How to deal with the competition between diffusion and aggregation? Are there important functional inequalities involved?

Mathematical Questions:

- What are the continuum models associated to these systems as the number of individuals gets larger and larger? Mean-field limits.
- What is the good analytical framework to deal with the possible concentration of mass in finite/infinite time in space or in velocity? Stability of patterns?
- What is the good analytical framework to deal with particles and continuum solutions at the same time? We want to allow solutions with evolving absolutely continuous and atomic parts of the measures at the same time.
- How to deal with the competition between diffusion and aggregation? Are there important functional inequalities involved?

Mathematical Questions:

- What are the continuum models associated to these systems as the number of individuals gets larger and larger? Mean-field limits.
- What is the good analytical framework to deal with the possible concentration of mass in finite/infinite time in space or in velocity? Stability of patterns?
- What is the good analytical framework to deal with particles and continuum solutions at the same time? We want to allow solutions with evolving absolutely continuous and atomic parts of the measures at the same time.
- How to deal with the competition between diffusion and aggregation? Are there important functional inequalities involved?

Mathematical Questions:

- What are the continuum models associated to these systems as the number of individuals gets larger and larger? Mean-field limits.
- What is the good analytical framework to deal with the possible concentration of mass in finite/infinite time in space or in velocity? Stability of patterns?
- What is the good analytical framework to deal with particles and continuum solutions at the same time? We want to allow solutions with evolving absolutely continuous and atomic parts of the measures at the same time.
- How to deal with the competition between diffusion and aggregation? Are there important functional inequalities involved?

Mathematical Questions:

- What are the continuum models associated to these systems as the number of individuals gets larger and larger? Mean-field limits.
- What is the good analytical framework to deal with the possible concentration of mass in finite/infinite time in space or in velocity? Stability of patterns?
- What is the good analytical framework to deal with particles and continuum solutions at the same time? We want to allow solutions with evolving absolutely continuous and atomic parts of the measures at the same time.
- How to deal with the competition between diffusion and aggregation? Are there important functional inequalities involved?

Schedule:

- Lecture 2: First order Models - Aggregation Equations: characterization of finite versus infinite time blow-up, global existence of measure solutions, stability/instability of steady states for attractive-repulsive potentials.
- Lecture 3: First order Models - Aggregation versus Diffusion: Critical Mass Classical Patlak-Keller-Segel model, relation to Log-HLS and fast diffusion equations, interplay with functional inequalities.
- Lecture 4: Second order Models - Kinetic Equations for Swarming: measure solutions - mean field limit with/without noise, asymptotic flocking for Cucker-Smale, discussion on patterns.

Schedule:

- Lecture 2: First order Models - Aggregation Equations: characterization of finite versus infinite time blow-up, global existence of measure solutions, stability/instability of steady states for attractive-repulsive potentials.
- Lecture 3: First order Models - Aggregation versus Diffusion: Critical Mass Classical Patlak-Keller-Segel model, relation to Log-HLS and fast diffusion equations, interplay with functional inequalities.
- Lecture 4: Second order Models - Kinetic Equations for Swarming: measure solutions - mean field limit with/without noise, asymptotic flocking for Cucker-Smale, discussion on patterns.

Schedule:

- Lecture 2: First order Models - Aggregation Equations: characterization of finite versus infinite time blow-up, global existence of measure solutions, stability/instability of steady states for attractive-repulsive potentials.
- Lecture 3: First order Models - Aggregation versus Diffusion: Critical Mass Classical Patlak-Keller-Segel model, relation to Log-HLS and fast diffusion equations, interplay with functional inequalities.
- Lecture 4: Second order Models - Kinetic Equations for Swarming: measure solutions - mean field limit with/without noise, asymptotic flocking for Cucker-Smale, discussion on patterns.

Potential $U(x) = |x|$ in 1D

The equation reads

$$\frac{\partial \rho}{\partial t}(t,x) = \operatorname{div}\left(\rho(t,x)\left[\operatorname{sign}(x) * \rho\right](t,x)\right) .$$

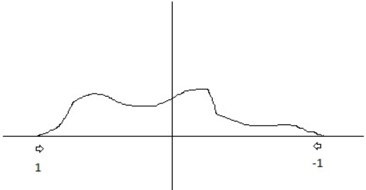
and thus the velocity field is given by

$$v(t,x) = -\int_{-\infty}^{x} \rho(t,x)\,dx + \int_{x}^{\infty} \rho(t,x)\,dx = 1 - 2\int_{-\infty}^{x} \rho(t,x)\,dx := 1 - 2F(t,x) .$$

Thus, the equation satisfied by F (Bodnar & Velazquez, JDE 2006) is a Burger's like equation of the form:

$$\frac{\partial F}{\partial t}(t,x) - \frac{\partial}{\partial x}\left(F^2\right)(t,x) = -F_x ,$$

which can be solved by characteristics: a jump formation in F in finite time equivalent to a formation of a Dirac Delta in finite time for ρ.



Potential $U(x) = |x|$ in 1D

The equation reads

$$\frac{\partial \rho}{\partial t}(t,x) = \operatorname{div}\left(\rho(t,x)\left[\operatorname{sign}(x) * \rho\right](t,x)\right) .$$

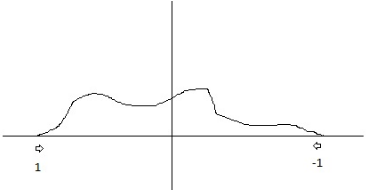
and thus the velocity field is given by

$$v(t,x) = -\int_{-\infty}^{x} \rho(t,x)\,dx + \int_{x}^{\infty} \rho(t,x)\,dx = 1 - 2\int_{-\infty}^{x} \rho(t,x)\,dx := 1 - 2F(t,x) .$$

Thus, the equation satisfied by F (Bodnar & Velazquez, JDE 2006) is a Burger's like equation of the form:

$$\frac{\partial F}{\partial t}(t,x) - \frac{\partial}{\partial x}\left(F^2\right)(t,x) = -F_x ,$$

which can be solved by characteristics: a jump formation in F in finite time equivalent to a formation of a Dirac Delta in finite time for ρ.



Potential $U(x) = |x|$ in 1D

The equation reads

$$\frac{\partial \rho}{\partial t}(t,x) = \mathrm{div}\left(\rho(t,x)\left[\mathrm{sign}(x) * \rho\right](t,x)\right) \ .$$

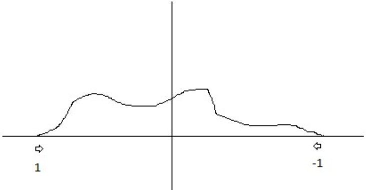
and thus the velocity field is given by

$$v(t,x) = -\int_{-\infty}^{x} \rho(t,x)\,dx + \int_{x}^{\infty} \rho(t,x)\,dx = 1 - 2\int_{-\infty}^{x} \rho(t,x)\,dx := 1 - 2F(t,x)\,.$$

Thus, the equation satisfied by F (Bodnar & Velazquez, JDE 2006) is a Burger's like equation of the form:

$$\frac{\partial F}{\partial t}(t,x) - \frac{\partial}{\partial x}\left(F^2\right)(t,x) = -F_x\,,$$

which can be solved by characteristics: a jump formation in F in finite time equivalent to a formation of a Dirac Delta in finite time for ρ.



Outline

Definition of the distance[2]

Transporting measures:

Given $T : \mathbb{R}^d \longrightarrow \mathbb{R}^d$ mesurable, we say that $\nu = T\#\mu$, if $\nu[K] := \mu[T^{-1}(K)]$ for all mesurable sets $K \subset \mathbb{R}^d$, equivalently

$$\int_{\mathbb{R}^d} \varphi \, d\nu = \int_{\mathbb{R}^d} (\varphi \circ T) \, d\mu$$

for all $\varphi \in C_o(\mathbb{R}^d)$.

Random variables:

Say that X is a random variable with law given by μ, is to say $X : (\Omega, \mathcal{A}, P) \longrightarrow (\mathbb{R}^d, \mathcal{B}_d)$ is a mesurable map such that $X\#P = \mu$, i.e.,

$$\int_{\mathbb{R}^d} \varphi(x) \, d\mu = \int_{\Omega} (\varphi \circ X) \, dP = \mathbb{E}\left[\varphi(X)\right].$$

[2] C. Villani, AMS Graduate Texts (2003).

Definition of the distance[2]

Transporting measures:

Given $T : \mathbb{R}^d \longrightarrow \mathbb{R}^d$ mesurable, we say that $\nu = T\#\mu$, if $\nu[K] := \mu[T^{-1}(K)]$ for all mesurable sets $K \subset \mathbb{R}^d$, equivalently

$$\int_{\mathbb{R}^d} \varphi \, d\nu = \int_{\mathbb{R}^d} (\varphi \circ T) \, d\mu$$

for all $\varphi \in C_o(\mathbb{R}^d)$.

Random variables:

Say that X is a random variable with law given by μ, is to say $X : (\Omega, \mathcal{A}, P) \longrightarrow (\mathbb{R}^d, \mathcal{B}_d)$ is a mesurable map such that $X\#P = \mu$, i.e.,

$$\int_{\mathbb{R}^d} \varphi(x) \, d\mu = \int_{\Omega} (\varphi \circ X) \, dP = \mathbb{E}\left[\varphi(X)\right].$$

[2] C. Villani, AMS Graduate Texts (2003).

Definition of the distance[2]

Transporting measures:

Given $T : \mathbb{R}^d \longrightarrow \mathbb{R}^d$ mesurable, we say that $\nu = T\#\mu$, if $\nu[K] := \mu[T^{-1}(K)]$ for all mesurable sets $K \subset \mathbb{R}^d$, equivalently

$$\int_{\mathbb{R}^d} \varphi \, d\nu = \int_{\mathbb{R}^d} (\varphi \circ T) \, d\mu$$

for all $\varphi \in C_o(\mathbb{R}^d)$.

Random variables:

Say that X is a random variable with law given by μ, is to say $X : (\Omega, \mathcal{A}, P) \longrightarrow (\mathbb{R}^d, \mathcal{B}_d)$ is a mesurable map such that $X\#P = \mu$, i.e.,

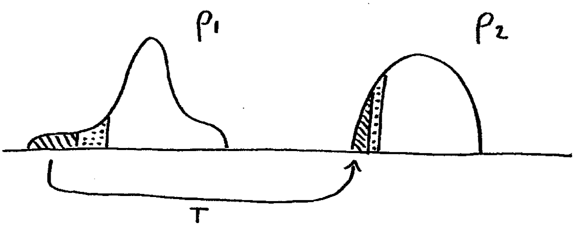
$$\int_{\mathbb{R}^d} \varphi(x) \, d\mu = \int_{\Omega} (\varphi \circ X) \, dP = \mathbb{E}\left[\varphi(X)\right].$$

[2] C. Villani, AMS Graduate Texts (2003).

Two piles of sand!

Energy needed to transport m kg of sand from $x = a$ to $x = b$:

$$\boxed{\text{energy} = m\,|a-b|^2}$$



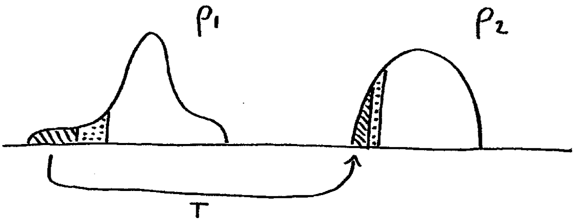
$W_2^2(\rho_1, \rho_2)$ = Among all possible ways to transport the mass from ρ_1 to ρ_2, find the one that minimizes the total energy

$$W_2^2(\rho_1, \rho_2) = \int |x - T(x)|^2 \, d\rho_1(x)$$

Two piles of sand!

Energy needed to transport m kg of sand from $x = a$ to $x = b$:

$$\boxed{\text{energy} = m\,|a-b|^2}$$



$W_2^2(\rho_1, \rho_2)$ = Among all possible ways to transport the mass from ρ_1 to ρ_2, find the one that minimizes the total energy

$$W_2^2(\rho_1, \rho_2) = \int |x - T(x)|^2 \, d\rho_1(x)$$

Definition of the distance

Kantorovich-Rubinstein-Wasserstein Distance $p = 1, 2$:

$$W_p^p(\mu,\nu) = \inf_\pi \left\{ \iint_{\mathbb{R}^d\times\mathbb{R}^d} |x-y|^p \, d\pi(x,y) \right\} = \inf_{(X,Y)} \left\{ \mathbb{E}\left[|X-Y|^p\right] \right\}$$

where the transference plan π runs over the set of joint probability measures on $\mathbb{R}^d \times \mathbb{R}^d$ with marginals μ and $\nu \in \mathcal{P}_p(\mathbb{R}^d)$ and (X,Y) are all possible couples of random variables with μ and ν as respective laws.

Monge's optimal mass transport problem:

Find

$$I := \inf_T \left\{ \int_{\mathbb{R}^d} |x - T(x)|^p \, d\mu(x);\ \nu = T\#\mu \right\}^{1/p}.$$

Take $\gamma_T = (1_{\mathbb{R}^d} \times T)\#\mu$ as transference plan π.

Definition of the distance

Kantorovich-Rubinstein-Wasserstein Distance $p = 1, 2$:

$$W_p^p(\mu, \nu) = \inf_\pi \left\{ \iint_{\mathbb{R}^d \times \mathbb{R}^d} |x - y|^p \, d\pi(x, y) \right\} = \inf_{(X,Y)} \left\{ \mathbb{E}\left[|X - Y|^p \right] \right\}$$

where the transference plan π runs over the set of joint probability measures on $\mathbb{R}^d \times \mathbb{R}^d$ with marginals μ and $\nu \in \mathcal{P}_p(\mathbb{R}^d)$ and (X, Y) are all possible couples of random variables with μ and ν as respective laws.

Monge's optimal mass transport problem:

Find

$$I := \inf_T \left\{ \int_{\mathbb{R}^d} |x - T(x)|^p \, d\mu(x); \, \nu = T\#\mu \right\}^{1/p}.$$

Take $\gamma_T = (1_{\mathbb{R}^d} \times T)\#\mu$ as transference plan π.

Definition of the distance

Kantorovich-Rubinstein-Wasserstein Distance $p = 1, 2$:

$$W_p^p(\mu,\nu) = \inf_\pi \left\{ \iint_{\mathbb{R}^d \times \mathbb{R}^d} |x-y|^p \, d\pi(x,y) \right\} = \inf_{(X,Y)} \left\{ \mathbb{E}\left[|X-Y|^p \right] \right\}$$

where the transference plan π runs over the set of joint probability measures on $\mathbb{R}^d \times \mathbb{R}^d$ with marginals μ and $\nu \in \mathcal{P}_p(\mathbb{R}^d)$ and (X,Y) are all possible couples of random variables with μ and ν as respective laws.

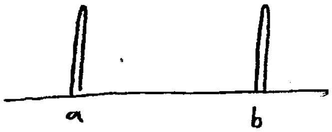
Monge's optimal mass transport problem:

Find

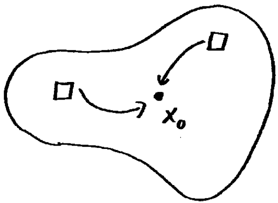
$$I := \inf_T \left\{ \int_{\mathbb{R}^d} |x - T(x)|^p \, d\mu(x); \; \nu = T\#\mu \right\}^{1/p}.$$

Take $\gamma_T = (1_{\mathbb{R}^d} \times T)\#\mu$ as transference plan π.

Three examples



$$W_2^2(\delta_a, \delta_b) = |a - b|^2$$



$$W_2^2(\rho, \delta_{X_0}) = \int |X_0 - y|^2 \, d\rho(y)$$
$$= \mathrm{Var}\,(\rho)$$

Outline

Euclidean Wasserstein Distance

Convergence Properties

1. **Convergence of measures:** $W_2(\mu_n, \mu) \to 0$ is equivalent to $\mu_n \rightharpoonup \mu$ weakly-* as measures and convergence of second moments.
2. **Weak lower semicontinuity:** Given $\mu_n \rightharpoonup \mu$ and $\nu_n \rightharpoonup \nu$ weakly-* as measures, then
$$W_2(\mu, \nu) \leq \liminf_{n\to\infty} W_2(\mu_n, \nu_n).$$
3. **Completeness:** The space $\mathcal{P}_2(\mathbb{R}^d)$ endowed with the distance W_2 is a complete metric space.

Euclidean Wasserstein Distance

Convergence Properties

1. **Convergence of measures:** $W_2(\mu_n,\mu)\to 0$ is equivalent to $\mu_n \rightharpoonup \mu$ weakly-* as measures and convergence of second moments.
2. **Weak lower semicontinuity:** Given $\mu_n \rightharpoonup \mu$ and $\nu_n \rightharpoonup \nu$ weakly-* as measures, then
$$W_2(\mu,\nu)\le \liminf_{n\to\infty} W_2(\mu_n,\nu_n).$$
3. **Completeness:** The space $\mathcal{P}_2(\mathbb{R}^d)$ endowed with the distance W_2 is a complete metric space.

Euclidean Wasserstein Distance

Convergence Properties

1. **Convergence of measures:** $W_2(\mu_n, \mu) \to 0$ is equivalent to $\mu_n \rightharpoonup \mu$ weakly-* as measures and convergence of second moments.
2. **Weak lower semicontinuity:** Given $\mu_n \rightharpoonup \mu$ and $\nu_n \rightharpoonup \nu$ weakly-* as measures, then
$$W_2(\mu, \nu) \le \liminf_{n \to \infty} W_2(\mu_n, \nu_n).$$
3. **Completeness:** The space $\mathcal{P}_2(\mathbb{R}^d)$ endowed with the distance W_2 is a complete metric space.

One dimensional Case

Distribution functions:

In one dimension, denoting by $F(x)$ the distribution function of μ,

$$F(x) = \int_{-\infty}^{x} d\mu,$$

we can define its pseudo inverse:

$$F^{-1}(\eta) = \inf\{x : F(x) > \eta\} \qquad \text{for } \eta \in (0,1),$$

we have $F^{-1} : ((0,1), \mathcal{B}_1), d\eta) \longrightarrow (\mathbb{R}, \mathcal{B}_1)$ is a random variable with law μ, i.e., $F^{-1}\#d\eta = \mu$

$$\int_{\mathbb{R}} \varphi(x)\, d\mu = \int_0^1 \varphi(F^{-1}(\eta))\, d\eta = \mathbb{E}\left[\varphi(X)\right].$$

One dimensional Case

Distribution functions:

In one dimension, denoting by $F(x)$ the distribution function of μ,

$$F(x) = \int_{-\infty}^{x} d\mu,$$

we can define its pseudo inverse:

$$F^{-1}(\eta) = \inf\{x : F(x) > \eta\} \qquad \text{for } \eta \in (0,1),$$

we have $F^{-1} : ((0,1), \mathcal{B}_1), d\eta) \longrightarrow (\mathbb{R}, \mathcal{B}_1)$ is a random variable with law μ, i.e., $F^{-1}\#d\eta = \mu$

$$\int_{\mathbb{R}} \varphi(x)\, d\mu = \int_0^1 \varphi(F^{-1}(\eta))\, d\eta = \mathbb{E}\left[\varphi(X)\right].$$

One dimensional Case

Wasserstein distance:

In one dimension, it can be checked[a] that given two measures μ and ν with distribution functions $F(x)$ and $G(y)$ then, $(F^{-1} \times G^{-1})\#d\eta$ has joint distribution function $H(x,y) = \min(F(x), G(y))$. Therefore, in one dimension, the optimal plan is given by $\pi_{opt}(x,y) = (F^{-1} \times G^{-1})\#d\eta$, and thus

$$W_p(\mu,\nu) = \left(\int_0^1 [F^{-1}(\eta) - G^{-1}(\eta)]^p \, d\eta\right)^{1/p} = \|F^{-1} - G^{-1}\|_{L^p(\mathbb{R})}$$

$1 \leq p < \infty$.

[a] W. Hoeffding (1940); M. Fréchet (1951); A. Pulvirenti, G. Toscani, Annali Mat. Pura Appl. (1996).

One dimensional Case

Wasserstein distance:

In one dimension, it can be checked[a] that given two measures μ and ν with distribution functions $F(x)$ and $G(y)$ then, $(F^{-1} \times G^{-1})\#d\eta$ has joint distribution function $H(x,y) = \min(F(x), G(y))$. Therefore, in one dimension, the optimal plan is given by $\pi_{opt}(x,y) = (F^{-1} \times G^{-1})\#d\eta$, and thus

$$W_p(\mu,\nu) = \left(\int_0^1 [F^{-1}(\eta) - G^{-1}(\eta)]^p \, d\eta \right)^{1/p} = \|F^{-1} - G^{-1}\|_{L^p(\mathbb{R})}$$

$1 \leq p < \infty$.

[a] W. Hoeffding (1940); M. Fréchet (1951); A. Pulvirenti, G. Toscani, Annali Mat. Pura Appl. (1996).