Swarming Models with Repulsive-Attractive Effects

J. A. Carrillo

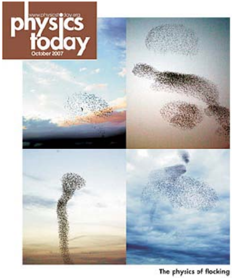
Imperial College London

Lecture 1, Ravello 2013

Outline

Outline

Swarming by Nature or by design?



Fish schools and Birds flocks.

Individual Based Models (Particle models)

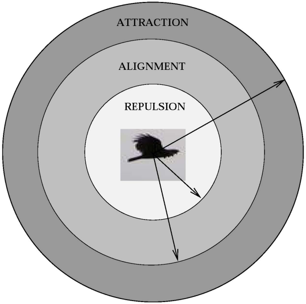
Swarming = Aggregation of agents of similar size and body type generally moving in a coordinated way.

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.



Metric versus Topological Interaction

Individual Based Models (Particle models)

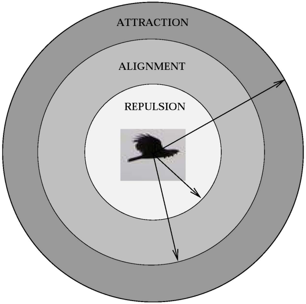
Swarming = Aggregation of agents of similar size and body type generally moving in a coordinated way.

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.



Metric versus Topological Interaction

Individual Based Models (Particle models)

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

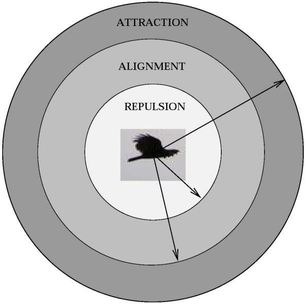
Highly developed social organization: insects (locusts, ants, bees ...), fishes, birds, micro-organisms (myxo-bacteria, ...) and artificial robots for unmanned vehicle operation.

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.

Metric versus Topological Interaction



Individual Based Models (Particle models)

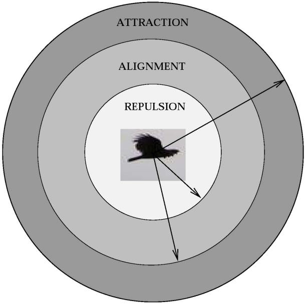
Swarming = Aggregation of agents of similar size and body type generally moving in a coordinated way.

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]

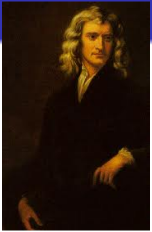
[a] Aoki, Helmerijk et al., Barbaro, Birnir et al.

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



Metric versus Topological Interaction

2nd Order Model: Newton's like equations



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

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

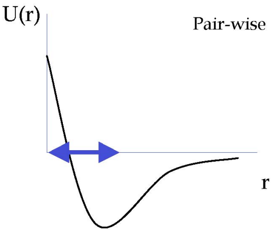
Model assumptions:

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

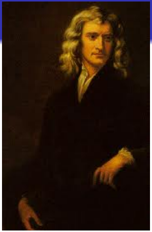
$$U(r) = -C_A e^{-r/\ell_A} + C_R e^{-r/\ell_R}.$$

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):

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

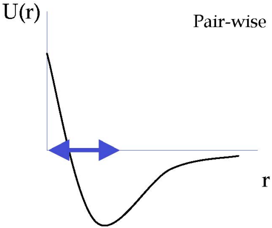
Model assumptions:

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

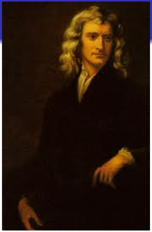
$$U(r) = -C_A e^{-r/\ell_A} + C_R e^{-r/\ell_R}.$$

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):

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

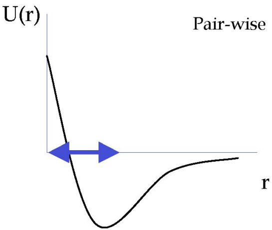
Model assumptions:

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

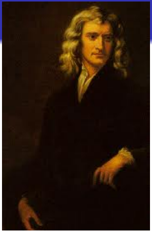
$$U(r) = -C_A e^{-r/\ell_A} + C_R e^{-r/\ell_R}.$$

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):

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

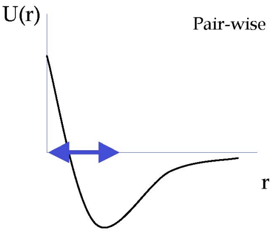
Model assumptions:

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

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

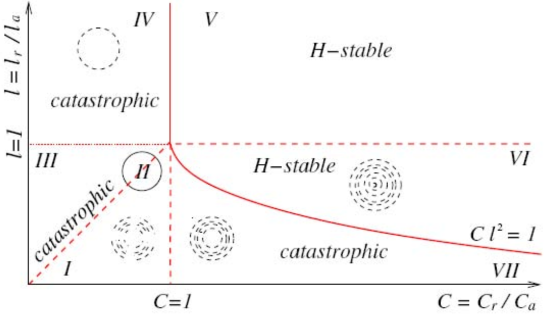
One can also use Bessel functions in 2D and 3D to produce such a potential.

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



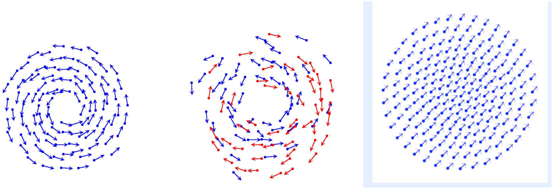
Model with an asymptotic velocity

Classification of possible patterns: Morse potential. D'Orsogna, Bertozzi et al. model (PRL 2006).



Model with an asymptotic speed

Typical patterns: milling, double milling or flocking:



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: $\gamma < 1/2$; Cucker-Smale.
General Proof for $0 < \gamma \leq 1/2$; C.-Fornasier-Rosado-Toscani.

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: $\gamma < 1/2$; Cucker-Smale.
General Proof for $0 < \gamma \leq 1/2$; C.-Fornasier-Rosado-Toscani.

Outline

Leadership, Geometrical Constraints, and Cone of Influence

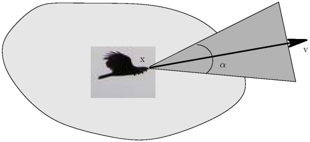
Cucker-Smale with local influence regions:

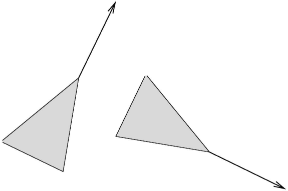
$$\begin{cases} \dfrac{dx_i}{dt} = v_i \ , \\ \dfrac{dv_i}{dt} = \displaystyle\sum_{j \in \Sigma_i(t)} a(|x_i - x_j|)(v_j - v_i) \ , \end{cases}$$

where $\Sigma_i(t) \subset \{1, \dots, N\}$ is the set of dependence, given by

$$\Sigma_i(t) := \left\{ 1 \leq \ell \leq N : \frac{(x_\ell - x_i) \cdot v_i}{|x_\ell - x_i||v_i|} \geq \alpha \right\} .$$

Cone of Vision:





Leadership, Geometrical Constraints, and Cone of Influence

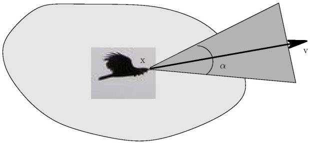
Cucker-Smale with local influence regions:

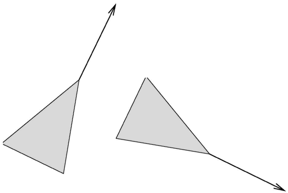
$$\begin{cases} \dfrac{dx_i}{dt} = v_i \ , \\ \dfrac{dv_i}{dt} = \displaystyle\sum_{j \in \Sigma_i(t)} a(|x_i - x_j|)(v_j - v_i) \ , \end{cases}$$

where $\Sigma_i(t) \subset \{1, \dots, N\}$ is the set of dependence, given by

$$\Sigma_i(t) := \left\{ 1 \le \ell \le N : \frac{(x_\ell - x_i) \cdot v_i}{|x_\ell - x_i||v_i|} \ge \alpha \right\}.$$

Cone of Vision:





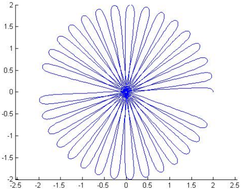
Roosting Forces

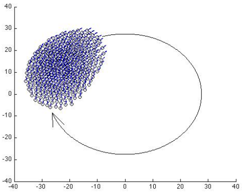
Adding a roosting area to the model:

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

with the roosting potential ϕ given by $\phi(x) := \dfrac{b}{4}\left(\dfrac{|x|}{R_{\text{Roost}}}\right)^4$.

Roosting effect: milling flocks $N = 400, R_{\text{roost}} = 20$.





Adding Noise

Self-Propelling/Friction/Interaction with Noise Particle Model:

$$\begin{cases} \dot{x}_i = v_i, \\ dv_i = \left[(\alpha - \beta \, |v_i|^2) v_i - \nabla_{x_i} \sum_{j \neq i} U(|x_i - x_j|) \right] dt + \sqrt{2\sigma} \, d\Gamma_i(t) \ , \end{cases}$$

where $\Gamma_i(t)$ are N independent copies of standard Wiener processes with values in $\mathbb{R}^d$ and $\sigma > 0$ is the noise strength. The Cucker–Smale Particle Model with Noise:

$$\begin{cases} dx_i = v_i dt \ , \\ dv_i = \sum_{j=1}^{N} a(|x_j - x_i|)(v_j - v_i) \, dt + \sqrt{2\sigma \sum_{j=1}^{N} a(|x_j - x_i|)} \, d\Gamma_i(t) \ . \end{cases}$$

Adding Noise

Self-Propelling/Friction/Interaction with Noise Particle Model:

$$\begin{cases} \dot{x}_i = v_i, \\ dv_i = \left[(\alpha - \beta\,|v_i|^2) v_i - \nabla_{x_i} \sum_{j \neq i} U(|x_i - x_j|) \right] dt + \sqrt{2\sigma}\, d\Gamma_i(t) \,, \end{cases}$$

where $\Gamma_i(t)$ are N independent copies of standard Wiener processes with values in $\mathbb{R}^d$ and $\sigma > 0$ is the noise strength. The Cucker–Smale Particle Model with Noise:

$$\begin{cases} dx_i = v_i dt \,, \\ dv_i = \sum_{j=1}^{N} a(|x_j - x_i|)(v_j - v_i)\, dt + \sqrt{2\sigma \sum_{j=1}^{N} a(|x_j - x_i|)}\, d\Gamma_i(t) \,. \end{cases}$$

Outline

Vicsek's model

Assume N particles moving at unit speed: reorientation & diffusion:

$$\begin{cases} dX_t^i = V_t^i\, dt, \\ dV_t^i = \sqrt{2}\, P(V_t^i) \circ dB_t^i - P(V_t^i) \left(\frac{1}{N} \sum_{j=1}^{N} K(X_t^i - X_t^j)(V_t^i - V_t^j) \right) dt. \end{cases}$$

Here $P(v)$ is the projection operator on the tangent space at $v/|v|$ to the unit sphere in $\mathbb{R}^d$, i.e.,

$$P(v) = I - \frac{v \otimes v}{|v|^2}\,.$$

Noise in the Stratatonovich sense: imposed by the rigorous construction of the Brownian motion on a manifold. Rigorous derivation: Bolley-Cañizo-C.

Main issue: phase transition? Degond-Liu-Frouvelle.

Vicsek's model

Assume N particles moving at unit speed: reorientation & diffusion:

$$\begin{cases} dX_t^i = V_t^i \, dt, \\ dV_t^i = \sqrt{2}\, P(V_t^i) \circ dB_t^i - P(V_t^i) \left(\dfrac{1}{N} \displaystyle\sum_{j=1}^{N} K(X_t^i - X_t^j)(V_t^i - V_t^j) \right) dt. \end{cases}$$

Here $P(v)$ is the projection operator on the tangent space at $v/|v|$ to the unit sphere in $\mathbb{R}^d$, i.e.,

$$P(v) = I - \frac{v \otimes v}{|v|^2}.$$

Noise in the Stratatonovich sense: imposed by the rigorous construction of the Brownian motion on a manifold. Rigorous derivation: Bolley-Cañizo-C.

Main issue: phase transition? Degond-Liu-Frouvelle.

Vicsek's model

Assume N particles moving at unit speed: reorientation & diffusion:

$$\begin{cases} dX_t^i = V_t^i\, dt, \\ dV_t^i = \sqrt{2}\, P(V_t^i) \circ dB_t^i - P(V_t^i) \left(\dfrac{1}{N} \displaystyle\sum_{j=1}^{N} K(X_t^i - X_t^j)(V_t^i - V_t^j) \right) dt. \end{cases}$$

Here $P(v)$ is the projection operator on the tangent space at $v/|v|$ to the unit sphere in $\mathbb{R}^d$, i.e.,

$$P(v) = I - \frac{v \otimes v}{|v|^2}\,.$$

Noise in the Stratatonovich sense: imposed by the rigorous construction of the Brownian motion on a manifold. Rigorous derivation: Bolley-Cañizo-C.

Main issue: phase transition? Degond-Liu-Frouvelle.

Vicsek's model

Assume N particles moving at unit speed: reorientation & diffusion:

$$\begin{cases} dX_t^i = V_t^i \, dt, \\ dV_t^i = \sqrt{2}\, P(V_t^i) \circ dB_t^i - P(V_t^i) \left(\dfrac{1}{N} \displaystyle\sum_{j=1}^{N} K(X_t^i - X_t^j)(V_t^i - V_t^j) \right) dt. \end{cases}$$

Here $P(v)$ is the projection operator on the tangent space at $v/|v|$ to the unit sphere in $\mathbb{R}^d$, i.e.,

$$P(v) = I - \frac{v \otimes v}{|v|^2}\,.$$

Noise in the Stratatonovich sense: imposed by the rigorous construction of the Brownian motion on a manifold. Rigorous derivation: Bolley-Cañizo-C.

Main issue: phase transition? Degond-Liu-Frouvelle.

Vicsek's model

Assume N particles moving at unit speed: reorientation & diffusion:

$$\begin{cases} dX_t^i = V_t^i\, dt, \\ dV_t^i = \sqrt{2}\, P(V_t^i) \circ dB_t^i - P(V_t^i) \left(\frac{1}{N} \sum_{j=1}^{N} K(X_t^i - X_t^j)(V_t^i - V_t^j) \right) dt. \end{cases}$$

Here $P(v)$ is the projection operator on the tangent space at $v/|v|$ to the unit sphere in $\mathbb{R}^d$, i.e.,

$$P(v) = I - \frac{v \otimes v}{|v|^2}.$$

Noise in the Stratatonovich sense: imposed by the rigorous construction of the Brownian motion on a manifold. Rigorous derivation: Bolley-Cañizo-C.

Main issue: phase transition? Degond-Liu-Frouvelle.

Outline

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

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

Flock Solutions: stationary states x_i^s of the 1st order model are connected to particular solutions of the Bertozzi etal 2nd order model of the form

$$x_i(t) = x_i^s + tv_0$$

with v_0 fixed with $|v_0|^2 = \frac{\alpha}{\beta}$.

For which potentials do we evolve towards some nontrivial steady states/patterns?
Is there any implication of the stability from first to 2nd order models?

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 $\Rightarrow$

$$\begin{cases} \frac{\partial\rho}{\partial t} + \operatorname{div}(\rho v) = 0 \\ v = -\nabla U * \rho \end{cases}$$

Flock Solutions: stationary states x_i^s of the 1st order model are connected to particular solutions of the Bertozzi etal 2nd order model of the form

$$x_i(t) = x_i^s + tv_0$$

with v_0 fixed with $|v_0|^2 = \frac{\alpha}{\beta}$.

For which potentials do we evolve towards some nontrivial steady states/patterns? Is there any implication of the stability from first to 2nd order models?

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

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

Flock Solutions: stationary states x_i^s of the 1st order model are connected to particular solutions of the Bertozzi etal 2nd order model of the form

$$x_i(t) = x_i^s + tv_0$$

with v_0 fixed with $|v_0|^2 = \frac{\alpha}{\beta}$.

For which potentials do we evolve towards some nontrivial steady states/patterns? Is there any implication of the stability from first to 2nd order models?

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:

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

so finally, we obtain

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

Flock Solutions: stationary states x_i^s of the 1st order model are connected to particular solutions of the Bertozzi etal 2nd order model of the form

$$x_i(t) = x_i^s + t v_0$$

with v_0 fixed with $|v_0|^2 = \frac{\alpha}{\beta}$.

For which potentials do we evolve towards some nontrivial steady states/patterns?
Is there any implication of the stability from first to 2nd order models?

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

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

Flock Solutions: stationary states x_i^s of the 1st order model are connected to particular solutions of the Bertozzi etal 2nd order model of the form

$$x_i(t) = x_i^s + tv_0$$

with v_0 fixed with $|v_0|^2 = \frac{\alpha}{\beta}$.

For which potentials do we evolve towards some nontrivial steady states/patterns?
Is there any implication of the stability from first to 2nd order models?

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?
- What is the good analytical framework to deal with particles and continuum solutions at the same time?
- How to deal with the stability of patterns, which perturbations?

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?
- What is the good analytical framework to deal with particles and continuum solutions at the same time?
- How to deal with the stability of patterns, which perturbations?

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?
- What is the good analytical framework to deal with particles and continuum solutions at the same time?
- How to deal with the stability of patterns, which perturbations?

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?
- What is the good analytical framework to deal with particles and continuum solutions at the same time?
- How to deal with the stability of patterns, which perturbations?

Schedule:

- Lecture 2: Second order Models - Kinetic Equations for Swarming: measure solutions - mean field limit with/without noise.
- Lectures 3-4: First order Models - Aggregation Equations: derivation and mean-field limit, stability/instability of steady states for repulsive/attractive potentials. Qualitative properties of Steady States.
- Lecture 5-6: Second order Models - Kinetic Equations for Swarming: Flock solutions: Stability. Mill Solutions: Instability. Hydrodynamic models: mills and double mills. Asymptotic Speed Models as friction limits.

Schedule:

- Lecture 2: Second order Models - Kinetic Equations for Swarming: measure solutions - mean field limit with/without noise.
- Lectures 3-4: First order Models - Aggregation Equations: derivation and mean-field limit, stability/instability of steady states for repulsive/attractive potentials. Qualitative properties of Steady States.
- Lecture 5-6: Second order Models - Kinetic Equations for Swarming: Flock solutions: Stability. Mill Solutions: Instability. Hydrodynamic models: mills and double mills. Asymptotic Speed Models as friction limits.

Schedule:

- Lecture 2: Second order Models - Kinetic Equations for Swarming: measure solutions - mean field limit with/without noise.
- Lectures 3-4: First order Models - Aggregation Equations: derivation and mean-field limit, stability/instability of steady states for repulsive/attractive potentials. Qualitative properties of Steady States.
- Lecture 5-6: Second order Models - Kinetic Equations for Swarming: Flock solutions: Stability. Mill Solutions: Instability. Hydrodynamic models: mills and double mills. Asymptotic Speed Models as friction limits.

Outline

Definition of the distance[1]

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.,

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

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

Definition of the distance[1]

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.,

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

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

Definition of the distance[1]

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.,

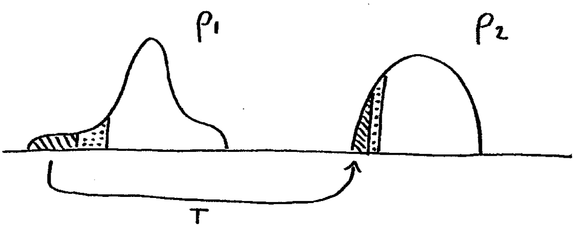
$$\int_{\mathbb{R}^d} \varphi(x) \, d\mu = \int_{\Omega} (\varphi \circ X) \, dP = \mathbb{E}\left[\varphi(X)\right].$$

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

Two piles of sand!

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

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



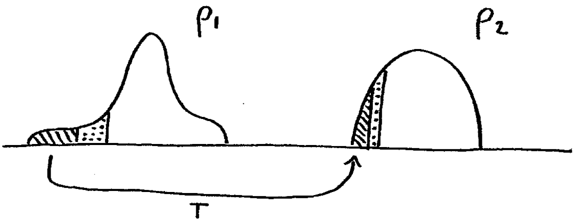
$W_2^2(\rho_1, \rho_2)$ = Among all possible ways to transport the mass from ρ_1 to ρ_2, find the one that minimizes the total energy

$$W_2^2(\rho_1, \rho_2) = \int_{\mathbb{R}^d} |x - T(x)|^2 \, d\rho_1(x)$$

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

$$W_2^2(\rho_1, \rho_2) = \int_{\mathbb{R}^d} |x - T(x)|^2 \, d\rho_1(x)$$

Definition of the distance

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

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

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.

$$W_\infty(\mu,\nu) = \inf_{\pi \in \Pi(\mu,\nu)} \sup_{(x,y) \in \mathrm{supp}(\pi)} |x-y| \, ,$$

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 $p = 1, 2, \infty$:

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

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.

$$W_\infty(\mu,\nu) = \inf_{\pi \in \Pi(\mu,\nu)} \sup_{(x,y) \in \mathrm{supp}(\pi)} |x-y| \,,$$

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 $p = 1, 2, \infty$:

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

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.

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

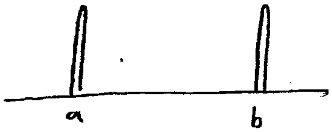
Monge's optimal mass transport problem:

Find

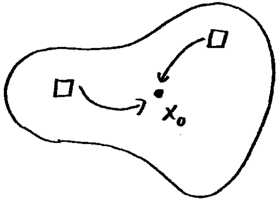
$$I := \inf_T \left\{ \int_{\mathbb{R}^d} |x - T(x)|^p \, d\mu(x); \; \nu = T\#\mu \right\}^{1/p}.$$

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
$$W_2(\mu, \nu) \leq \liminf_{n\to\infty} W_2(\mu_n, \nu_n).$$
3. **Completeness:** The space $\mathcal{P}_2(\mathbb{R}^d)$ endowed with the distance W_2 is a complete metric space.

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

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

$1 \leq p \leq \infty$.

[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

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

$1 \le p \le \infty$.

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