# The Principle of Least Action and its Geometric Aspects

**… What I understood …**

UCLA IPAM, February 2016

Bruno Lévy
**ALICE Géométrie & Lumière**
**CENTRE INRIA Nancy Grand-Est**

# OVERVIEW

**Part. 1.** The Least Action Principle

**Part. 2.** Optimal Transport – Elementary intro.

**Part. 3.** The Semi-Discrete Case

**Part. 4.** Understanding What's Going On…

**Part. 5.** Concluding Words

Inria

# OVERVIEW

**Yann Brenier**

The polar factorization theorem

(Brenier Transport)

"Each time the Laplace operator is used in a PDE, it can be replaced with the Monte-Ampère operator, and then interesting things occur"

# OVERVIEW


**Yann Brenier**

The polar factorization theorem
(Brenier Transport)

**Cédric Villani**

Optimal Transport Old & New
Topics on Optimal Transport

"Each time the Laplace operator is used in a PDE, it can be replaced with the Monte-Ampère operator, and then interesting things occur"

# OVERVIEW

*A computer programmer's adventure in MathLand*

June 2015
Institut Fourier

March 2015
Bonn

Febr 2015
← BIRS (Canada)

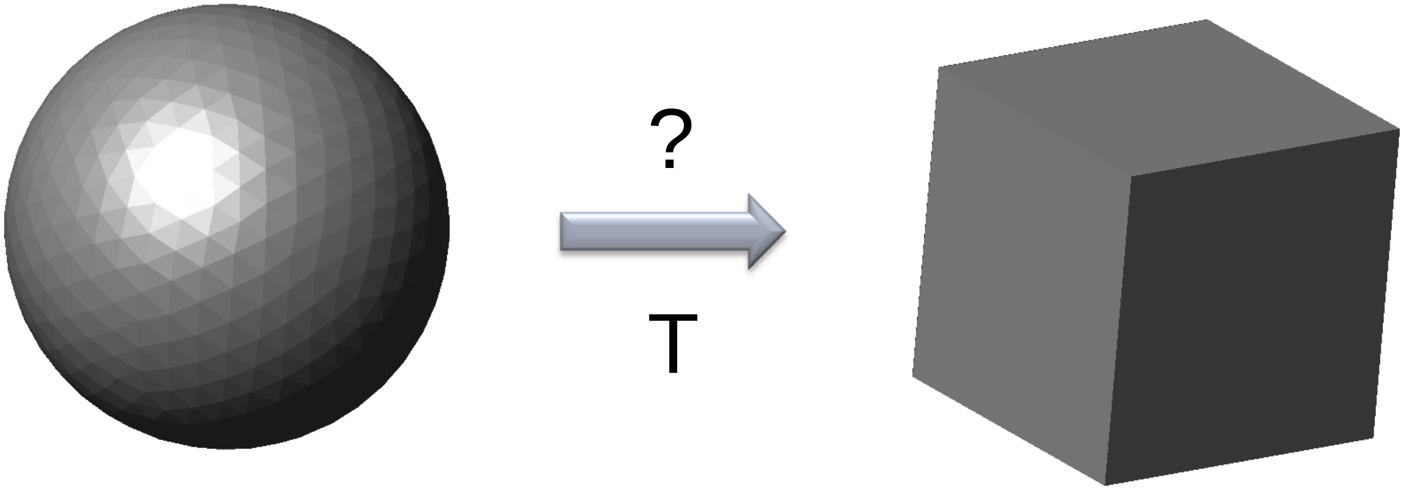
Febr 2015
LJLL

Discuss. with Quentin Mérigot (computational geometry – geometric measure theory)

# 1

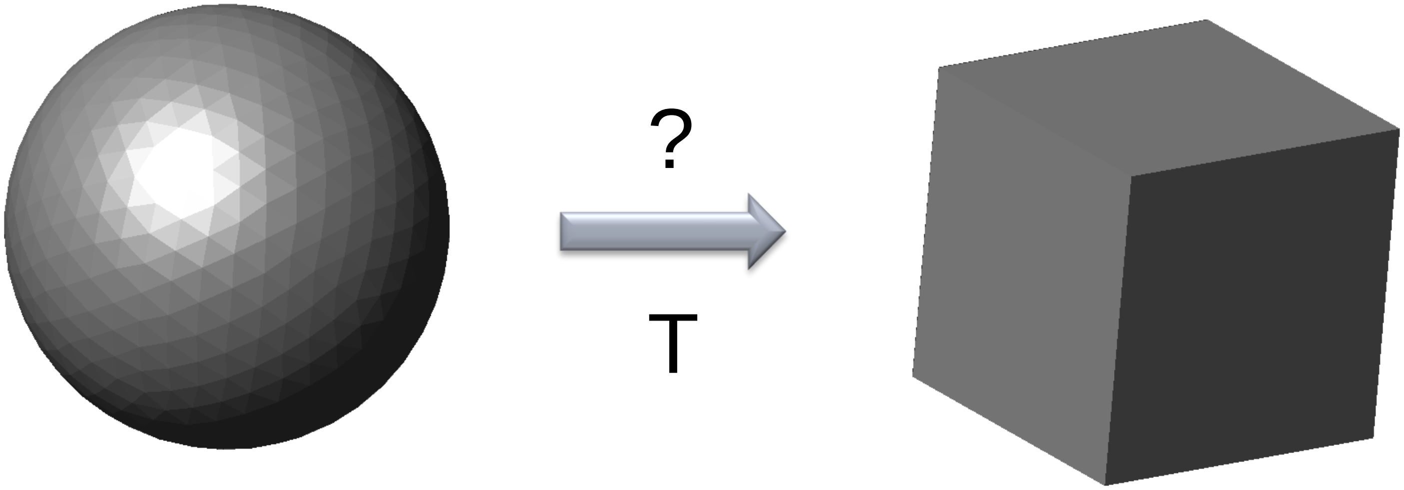
# The Least Action Principle

# **Part. 1** The Least Action Principle



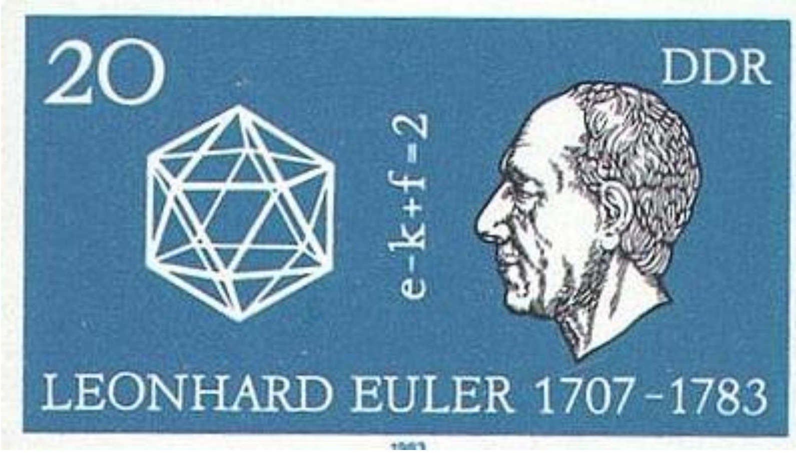

How to “morph” a shape into another one of same mass while minimizing the “effort” ?

# **Part. 1** The Least Action Principle




How to "morph" a shape into another one of same mass while minimizing the "effort" ?

The "effort" of the best T defines a **distance** between the shapes




Euler

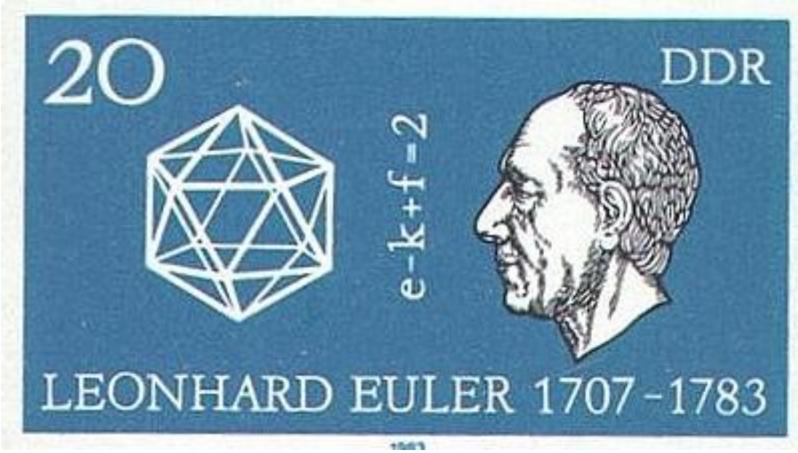
Hamilton,
Legendre,
Maupertuis

Lagrange

# The Least Action Principle

**Axiom 1**: There exists a function $L(x,\dot{x},t)$ that describes the state of a physical system

Short summary of the 1st chapter of Landau,Lifshitz Course of Theoretical Physics




Euler

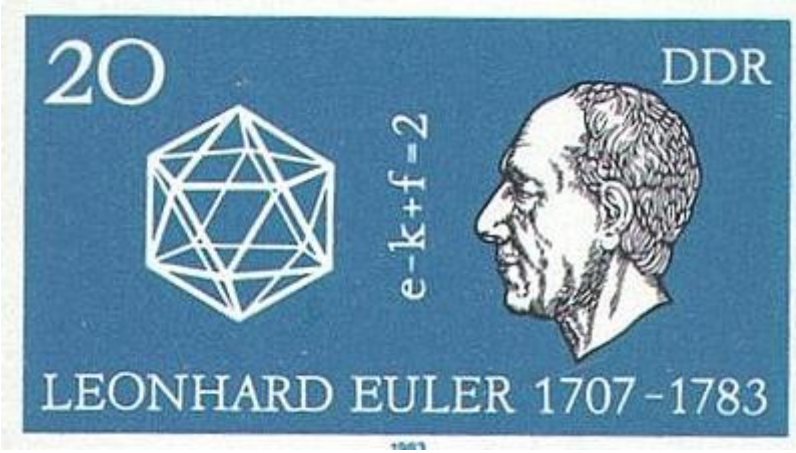
Hamilton,
Legendre,
Maupertuis

Lagrange

# The Least Action Principle

**Axiom 1**: There exists a function $L(x,\dot{x},t)$ that describes the state of a physical system

position




Euler

Hamilton,
Legendre,
Maupertuis

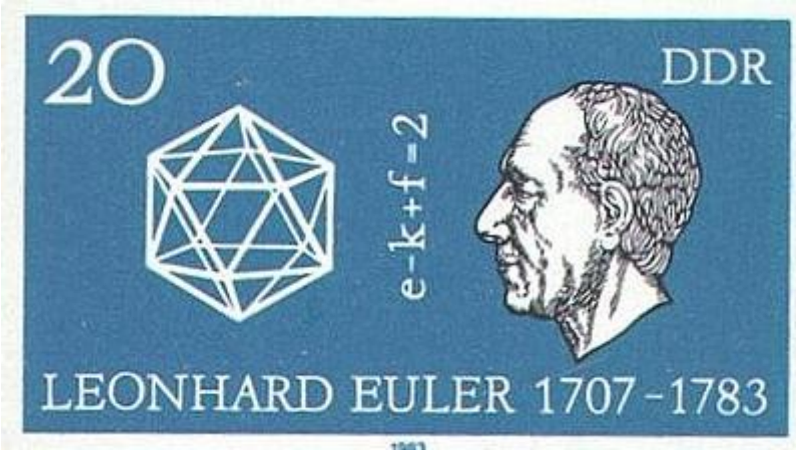
Lagrange

# The Least Action Principle

**Axiom 1**: There exists a function $L(x,\dot{x},t)$ that describes the state of a physical system

position

speed




Euler

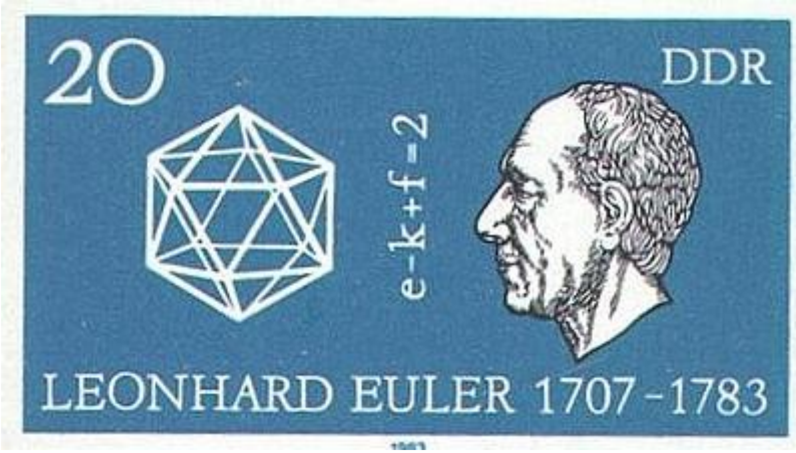
Hamilton,
Legendre,
Maupertuis

Lagrange

# The Least Action Principle

**Axiom 1**: There exists a function $L(x,\dot{x},t)$ that describes the state of a physical system

position ($x$), speed ($\dot{x}$), time ($t$)




Euler

Hamilton,
Legendre,
Maupertuis

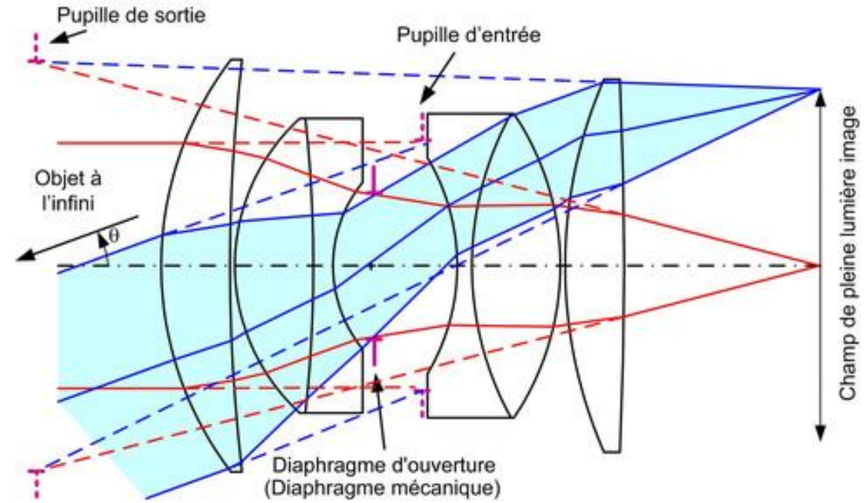
Lagrange

# The Least Action Principle

**Axiom 1**: There exists a function $L(x,\dot{x},t)$ that describes the state of a physical system

**Axiom 2**: The movement (time evolution) of the physical system minimizes the following integral

$$\int_{t_1}^{t_2} L(x,\dot{x},t)\,dt$$




# The Least Action Principle

**Axiom 1**: There exists a function $L(x,\dot{x},t)$ that describes the state of a physical system

**Axiom 2**: The movement (time evolution) of the physical system minimizes the following integral

$$\int_{t_1}^{t_2} L(x,\dot{x},t)\, dt$$

# The Least Action Principle

**Axiom 1**: There exists a function $L(x,\dot{x},t)$ that describes the state of a physical system

**Axiom 2**: The movement (time evolution) of the physical system minimizes the following integral

$$\int_{t_1}^{t_2} L(x,\dot{x},t)\, dt$$

**Theorem 1:** (Lagrange equation):

$$\frac{\partial L}{\partial x} = \frac{d}{dt}\frac{\partial L}{\partial \dot{x}}$$

# The Least Action Principle

**Axiom 1**: There exists L

**Axiom 2**: The movement minimizes

$$\int_{t_1}^{t_2} L(x,\dot{x},t)\,dt$$

**Axiom 3:**
Invariance w.r.t. change of Gallileo frame + hom. + isotrop. :

$$\begin{matrix} x' \\ t' \end{matrix} = \begin{matrix} x+vt \\ t \end{matrix}$$

**Theorem 1:** (Lagrange equation):

$$\frac{\partial L}{\partial x} = \frac{d}{dt}\,\frac{\partial L}{\partial \dot{x}}$$

# The Least Action Principle

**Axiom 1**: There exists L

**Axiom 2**: The movement minimizes

$$\int_{t_1}^{t_2} L(x,\dot{x},t)\, dt$$

**Theorem 1:** (Lagrange equation):

$$\frac{\partial L}{\partial x} = \frac{d}{dt}\frac{\partial L}{\partial \dot{x}}$$

**Axiom 3:**

Invariance w.r.t. change of Gallileo frame + hom. + isotrop. :

$$\begin{matrix} x' \\ t' \end{matrix} = \begin{matrix} x+vt \\ t \end{matrix}$$

**Theorem 2:**

$$\dot{x}\frac{\partial L}{\partial \dot{x}} - L = cte$$

# The Least Action Principle

**Axiom 1**: There exists L

**Axiom 2**: The movement minimizes

$$\int_{t_1}^{t_2} L(x,\dot{x},t)\, dt$$

**Theorem 1:** (Lagrange equation):

$$\frac{\partial L}{\partial x} = \frac{d}{dt} \frac{\partial L}{\partial \dot{x}}$$

**Axiom 3:**

Invariance w.r.t. change of Gallileo frame + hom. + isotrop. :

$$\begin{matrix} x' \\ t' \end{matrix} = \begin{matrix} x+vt \\ t \end{matrix}$$

**Theorem 2:**

$$\dot{x} \frac{\partial L}{\partial \dot{x}} - L = cte$$

Homogeneity of time → Preservation of **energy**

# The Least Action Principle

**Axiom 1**: There exists L
**Axiom 2**: The movement minimizes

$$\int_{t_1}^{t_2} L(x,\dot{x},t)\, dt$$

**Theorem 1:** (Lagrange equation):

$$\frac{\partial L}{\partial x} = \frac{d}{dt}\frac{\partial L}{\partial \dot{x}}$$

**Axiom 3:**
Invariance w.r.t. change of
Gallileo frame + hom. + isotrop. :

$$\begin{matrix} x' \\ t' \end{matrix} = \begin{matrix} x+vt \\ t \end{matrix}$$

**Theorem 2:**

$$\dot{x}\frac{\partial L}{\partial \dot{x}} - L = cte$$

Homogeneity of time → Preservation of **energy**

Homogeneity of space → Preservation of **momentum**

# The Least Action Principle

**Axiom 1**: There exists L

**Axiom 2**: The movement minimizes

$$\int_{t_1}^{t_2} L(x,\dot{x},t)\, dt$$

**Theorem 1:** (Lagrange equation):

$$\frac{\partial L}{\partial x} = \frac{d}{dt}\frac{\partial L}{\partial \dot{x}}$$

**Axiom 3:**
Invariance w.r.t. change of Gallileo frame + hom. + isotrop. :

$$\begin{matrix} x' \\ t' \end{matrix} = \begin{matrix} x+vt \\ t \end{matrix}$$

**Theorem 2:**

$$\dot{x}\frac{\partial L}{\partial \dot{x}} - L = cte$$

Homogeneity of time → Preservation of **energy**

Homogeneity of space → Preservation of **momentum**

Isotropy of space → Preservation of **angular momentum**

# The Least Action Principle

**Axiom 1**: There exists L

**Axiom 2**: The movement minimizes

$$\int_{t_1}^{t_2} L(x,\dot{x},t)\, dt$$

*Preserved quantities*
*"Integrals of Motion"*
*Noeter's theorem*

**Theorem 1:** (Lagrange equation):

$$\frac{\partial L}{\partial x} = \frac{d}{dt}\frac{\partial L}{\partial \dot{x}}$$

**Axiom 3:**
Invariance w.r.t. change of Gallileo frame + hom. + isotrop. :

$$\begin{pmatrix} x' \\ t' \end{pmatrix} = \begin{pmatrix} x+vt \\ t \end{pmatrix}$$

**Theorem 2:**

$$\dot{x}\frac{\partial L}{\partial \dot{x}} - L = cte$$

Homogeneity of time → Preservation of **energy**

Homogeneity of space → Preservation of **momentum**

Isotropy of space → Preservation of **angular momentum**

# The Least Action Principle

**Axiom 1**: There exists L

**Axiom 2**: The movement minimizes $\int L$

**Theorem 1:** (Lagrange equation):

$$\frac{\partial L}{\partial x} = \frac{d}{dt} \frac{\partial L}{\partial \dot{x}}$$

**Axiom 3:**
Invariance w.r.t. change of
Gallileo frame + hom. + isotrop. :

$$\begin{matrix} x' \\ t' \end{matrix} = \begin{matrix} x+vt \\ t \end{matrix}$$

**Theorem 2:**

$$\dot{x} \frac{\partial L}{\partial \dot{x}} - L = cte$$

Homogeneity of time →
Preservation of **energy**

Homogeneity of space →
Preservation of **momentum**

Isotropy of space →
Preservation of **angular momentum**

## Free particle:

**Theorem 3:** v = cte *(Newton's law I)*

# The Least Action Principle

**Axiom 1**: There exists L

**Axiom 2**: The movement minimizes $\int L$

**Theorem 1:** (Lagrange equation):

$$\frac{\partial L}{\partial x} = \frac{d}{dt}\frac{\partial L}{\partial \dot{x}}$$

**Axiom 3:**
Invariance w.r.t. change of Gallileo frame + hom. + isotrop. :

$$\begin{matrix} x' \\ t' \end{matrix} = \begin{matrix} x+vt \\ t \end{matrix}$$

**Theorem 2:**

$$\dot{x}\frac{\partial L}{\partial \dot{x}} - L = cte$$

Homogeneity of time →
Preservation of **energy**

Homogeneity of space →
Preservation of **momentum**

Isotropy of space →
Preservation of **angular momentum**

## <u>Free particle:</u>

**Theorem 3:** v = cte *(Newton's law I)*

Expression of the Lagrangian:

$$L = \tfrac{1}{2}\, m\, v^2$$

# The Least Action Principle

**Axiom 1**: There exists L

**Axiom 2**: The movement minimizes $\int L$

**Theorem 1:** (Lagrange equation):

$$\frac{\partial L}{\partial x} = \frac{d}{dt} \frac{\partial L}{\partial \dot{x}}$$

**Free particle:**

**Theorem 3:** $v = cte$ *(Newton's law I)*

Expression of the Lagrangian:

$$L = ½ \, m \, v^2$$

Expression of the Energy:

$$E = ½ \, m \, v^2$$

**Axiom 3:**
Invariance w.r.t. change of Gallileo frame + hom. + isotrop. :

$$\begin{matrix} x' \\ t' \end{matrix} = \begin{matrix} x+vt \\ t \end{matrix}$$

**Theorem 2:**

$$\dot{x} \frac{\partial L}{\partial \dot{x}} - L = cte$$

Homogeneity of time → Preservation of **energy**

Homogeneity of space → Preservation of **momentum**

Isotropy of space → Preservation of **angular momentum**

# The Least Action Principle

**Axiom 1**: There exists L

**Axiom 2**: The movement minimizes $\int L$

**Theorem 1:** (Lagrange equation):

$$\frac{\partial L}{\partial x} = \frac{d}{dt} \frac{\partial L}{\partial \dot{x}}$$

**Axiom 3:**
Invariance w.r.t. change of Gallileo frame + hom. + isotrop. :

$$\begin{matrix} x' \\ t' \end{matrix} = \begin{matrix} x+vt \\ t \end{matrix}$$

## Free particle:

**Theorem 3:** $v = cte$ ***(Newton's law I)***

Expression of the Lagrangian:

$$L = \frac{1}{2} m v^2$$

Expression of the Energy:

$$E = \frac{1}{2} m v^2$$

## Particle in a field:

Expression of the Lagrangian:

$$L = \frac{1}{2} m v^2 - U(x)$$

# The Least Action Principle

**Axiom 1**: There exists L

**Axiom 2**: The movement minimizes $\int L$

**Theorem 1:** (Lagrange equation):

$$\frac{\partial L}{\partial x} = \frac{d}{dt}\frac{\partial L}{\partial \dot{x}}$$

**Axiom 3:**
Invariance w.r.t. change of Gallileo frame + hom. + isotrop. :

$$\begin{matrix} x' \\ t' \end{matrix} = \begin{matrix} x+vt \\ t \end{matrix}$$

## Free particle:

**Theorem 3:** v = cte *(Newton's law I)*

Expression of the Lagrangian:

$$L = \frac{1}{2} m v^2$$

Expression of the Energy:

$$E = \frac{1}{2} m v^2$$

## Particle in a field:

Expression of the Lagrangian:

$$L = \frac{1}{2} m v^2 - U(x)$$

Expression of the Energy:

$$E = \frac{1}{2} m v^2 + U(x)$$

Inria

# The Least Action Principle

**Axiom 1**: There exists L

**Axiom 2**: The movement minimizes $\int L$

**Theorem 1:** (Lagrange equation):

$$\frac{\partial L}{\partial x} = \frac{d}{dt}\frac{\partial L}{\partial \dot{x}}$$

**Axiom 3:**
Invariance w.r.t. change of Gallileo frame + hom. + isotrop. :

$$\begin{matrix} x' \\ t' \end{matrix} = \begin{matrix} x+vt \\ t \end{matrix}$$

## Free particle:

**Theorem 3:** $v = cte$ ***(Newton's law I)***

Expression of the Lagrangian:

$$L = \frac{1}{2} m v^2$$

Expression of the Energy:

$$E = \frac{1}{2} m v^2$$

## Particle in a field:

Expression of the Lagrangian:

$$L = \frac{1}{2} m v^2 - U(x)$$

Expression of the Energy:

$$E = \frac{1}{2} m v^2 + U(x)$$

**Theorem 4:**

$m\ddot{x} = -\nabla U$ ***(Newton's law II)***

# The Least Action Principle

## *(relativistic setting)*

**Axiom 1**: There exists L

**Axiom 2**: The movement minimizes $\int L$

**Theorem 1:** (Lagrange equation):

$$\frac{\partial L}{\partial x} = \frac{d}{dt} \frac{\partial L}{\partial \dot{x}}$$

**Axiom 3:**
Invariance w.r.t. Lorentz change of frame

$$\begin{pmatrix} x' \\ t' \end{pmatrix} = \begin{pmatrix} (x-vt) \times \gamma \\ (t - vx/c^2) \times \gamma \end{pmatrix}$$

$$\gamma = 1 / \sqrt{(1 - v^2 / c^2)}$$

# The Least Action Principle
## *(relativistic setting – just for fun…)*

**Axiom 1**: There exists L

**Axiom 2**: The movement minimizes $\int L$

**Theorem 1:** (Lagrange equation):

$$\frac{\partial L}{\partial x} = \frac{d}{dt} \frac{\partial L}{\partial \dot{x}}$$

**Axiom 3:**
Invariance w.r.t. Lorentz change of frame

$$\begin{matrix} x' \\ t' \end{matrix} = \begin{matrix} (x-vt) \times \gamma \\ (t - vx/c^2) \times \gamma \end{matrix}$$

$$\gamma = 1 / \sqrt{(1 - v^2 / c^2)}$$

**Theorem 5:**

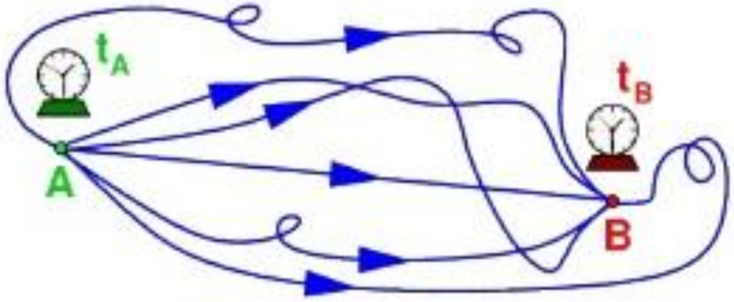
$$E = ½ \gamma m v^2 + mc^2$$

# The Least Action Principle

*(quantum physics setting – just for fun…)*

In quantum mechanics non just the extreme path contributes to the probability amplitude

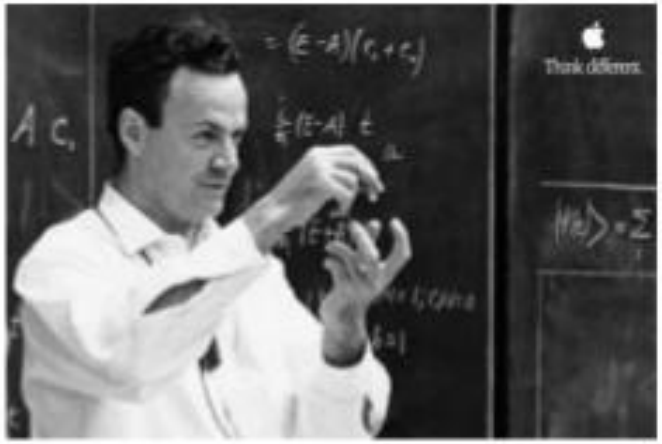
$$P(B,A)=|K(2,1)|^2$$




$$K(B,A)=\sum_{\text{over all possible paths}} \phi[x(t)]$$

where

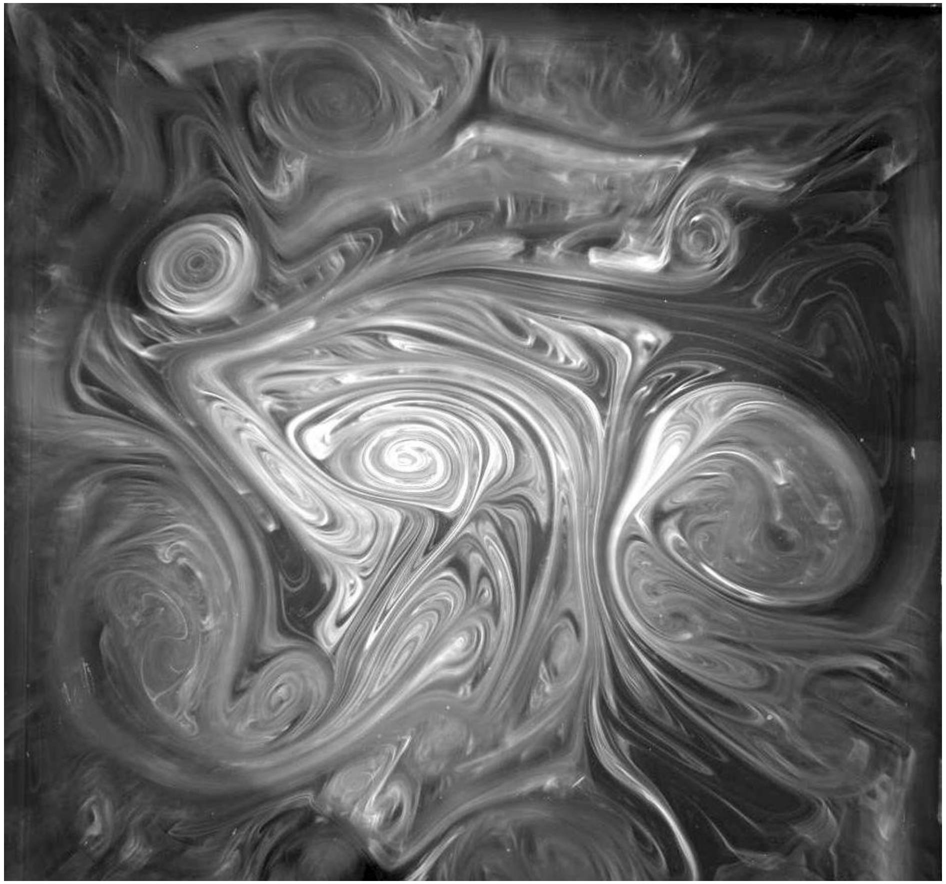
$$\phi[x(t)]=A\exp\left\{\frac{i}{\hbar}S[x(t)]\right\}$$




**Feynman's path integral formula**

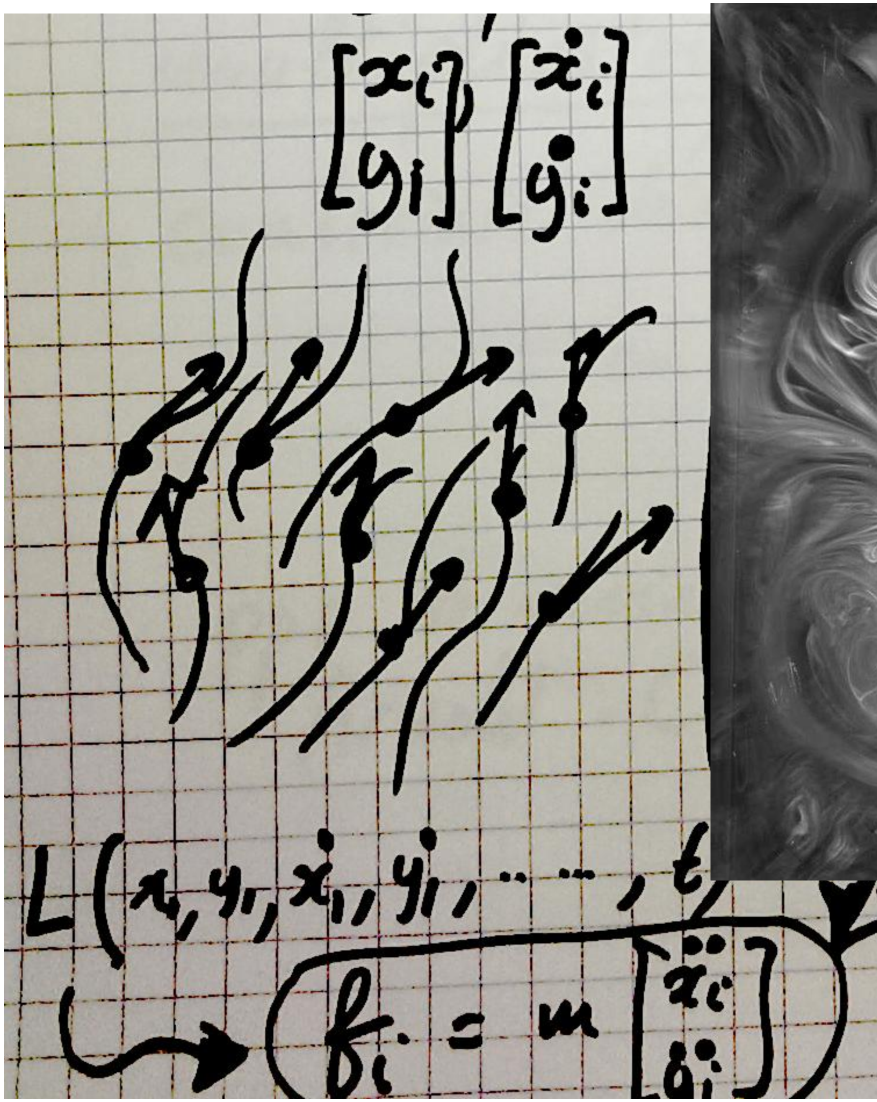
$$K(B,A)=\int_A^B \exp\left(\frac{i}{\hbar}S[B,A]Dx(t)\right)$$

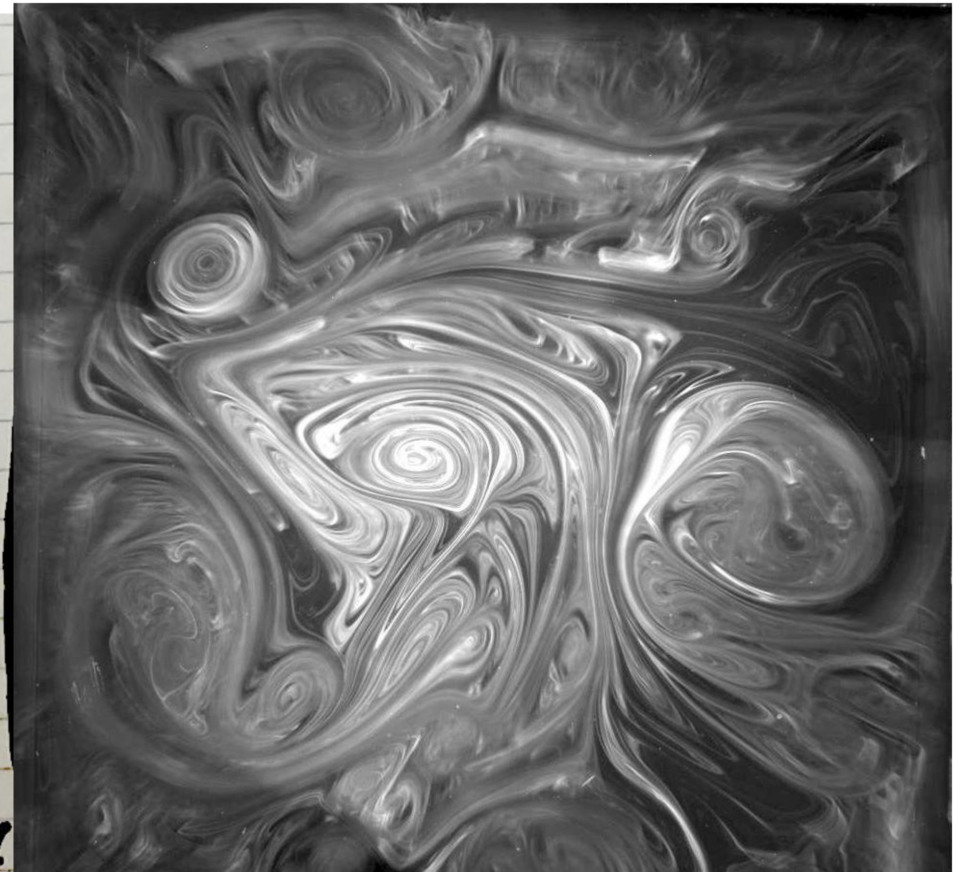
# Fluids



# Fluids

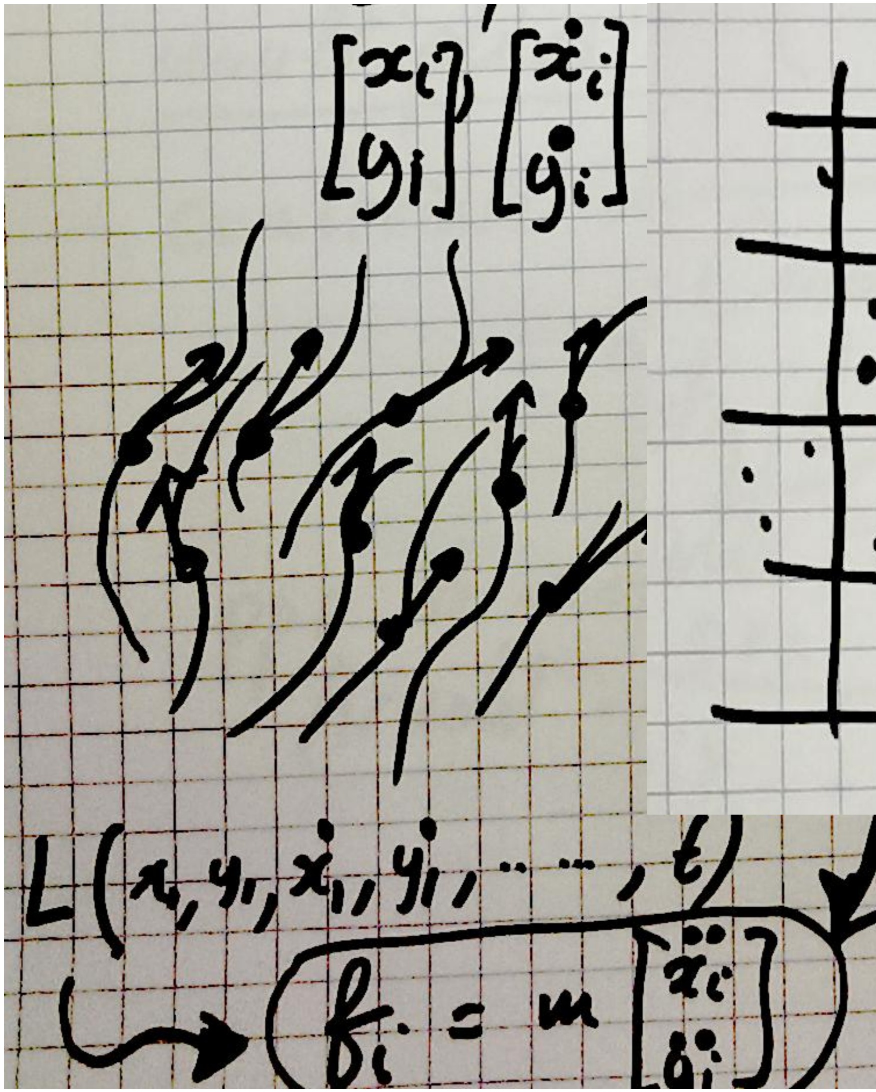
"Lagrange" point of view



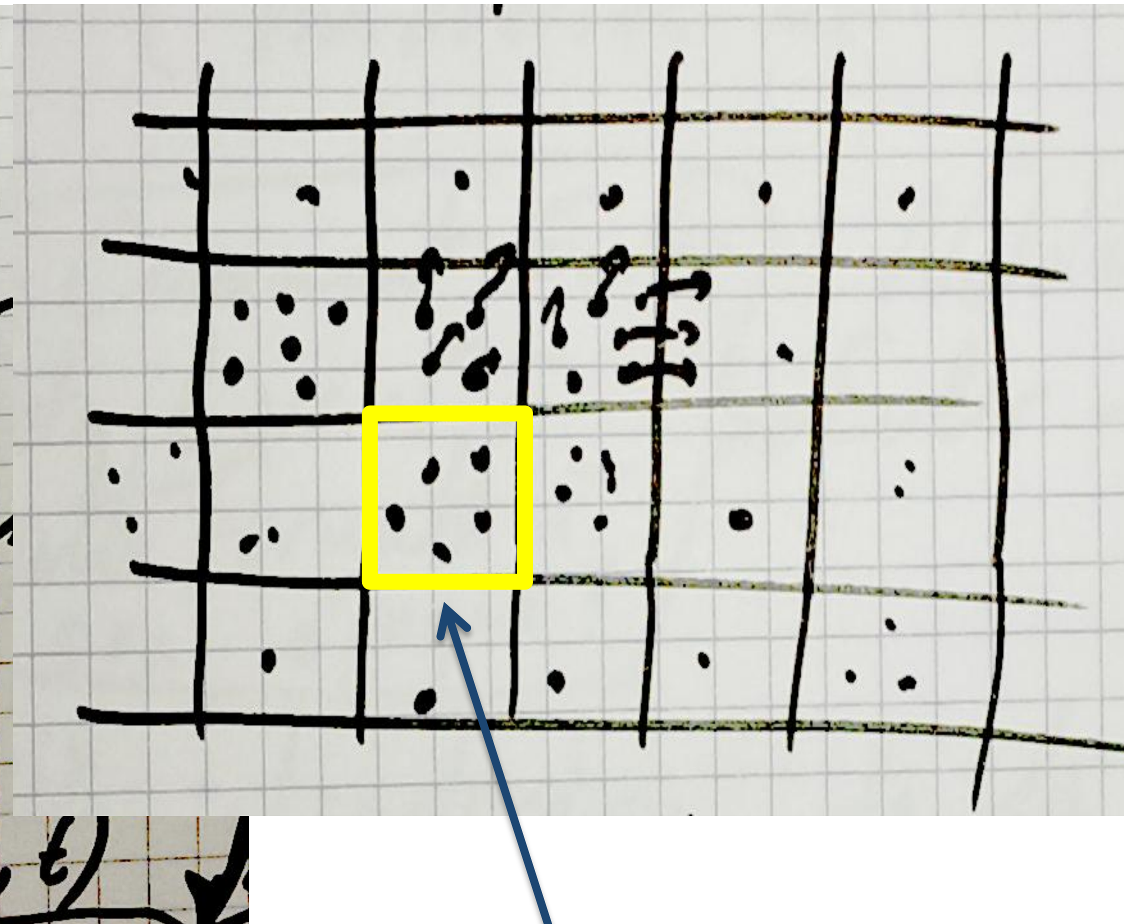


# Fluids

“Lagrange” point of view



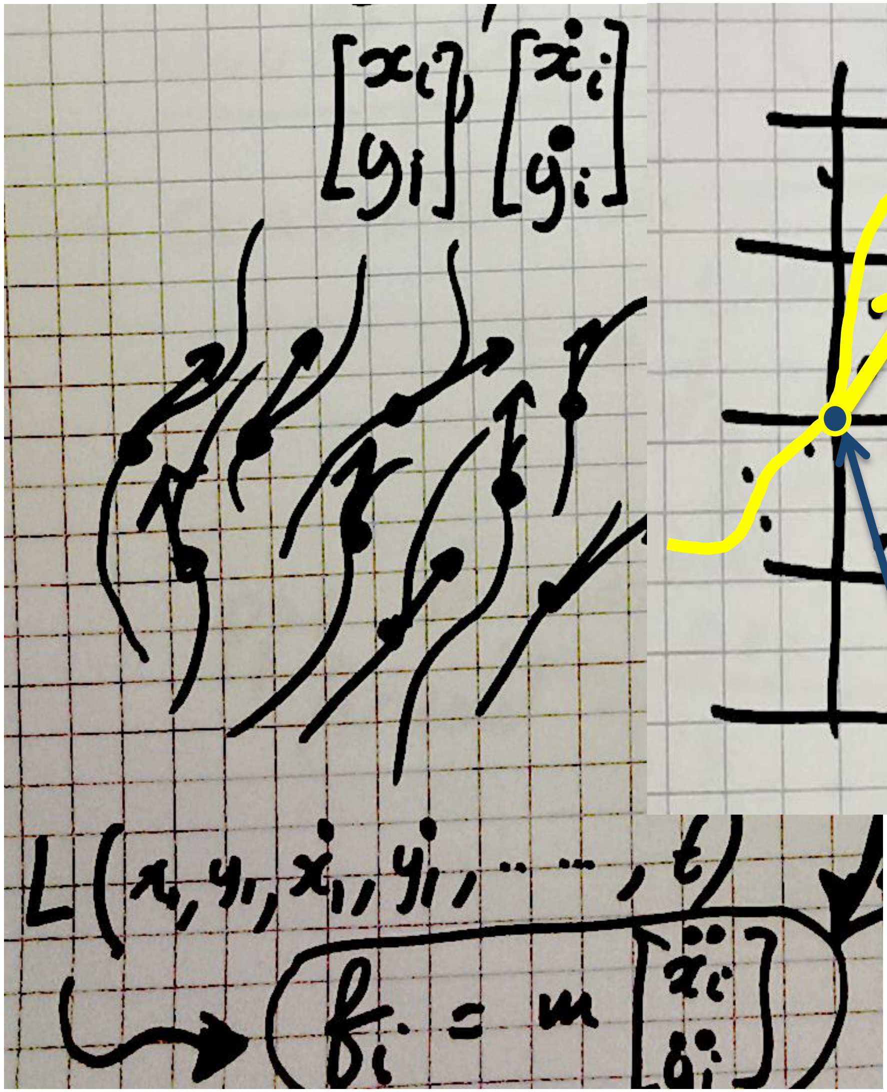
“Euler” point of view



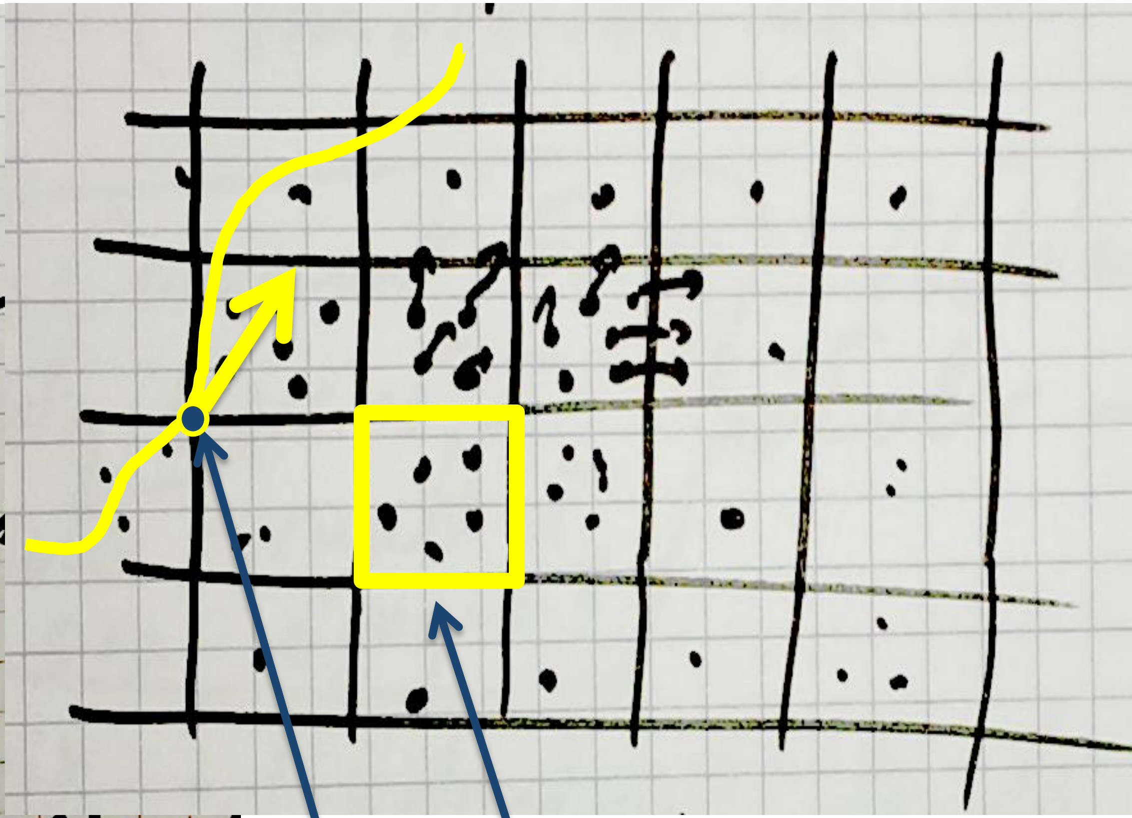
ρ “nb particles per square”

# Fluids

"Lagrange" point of view



"Euler" point of view



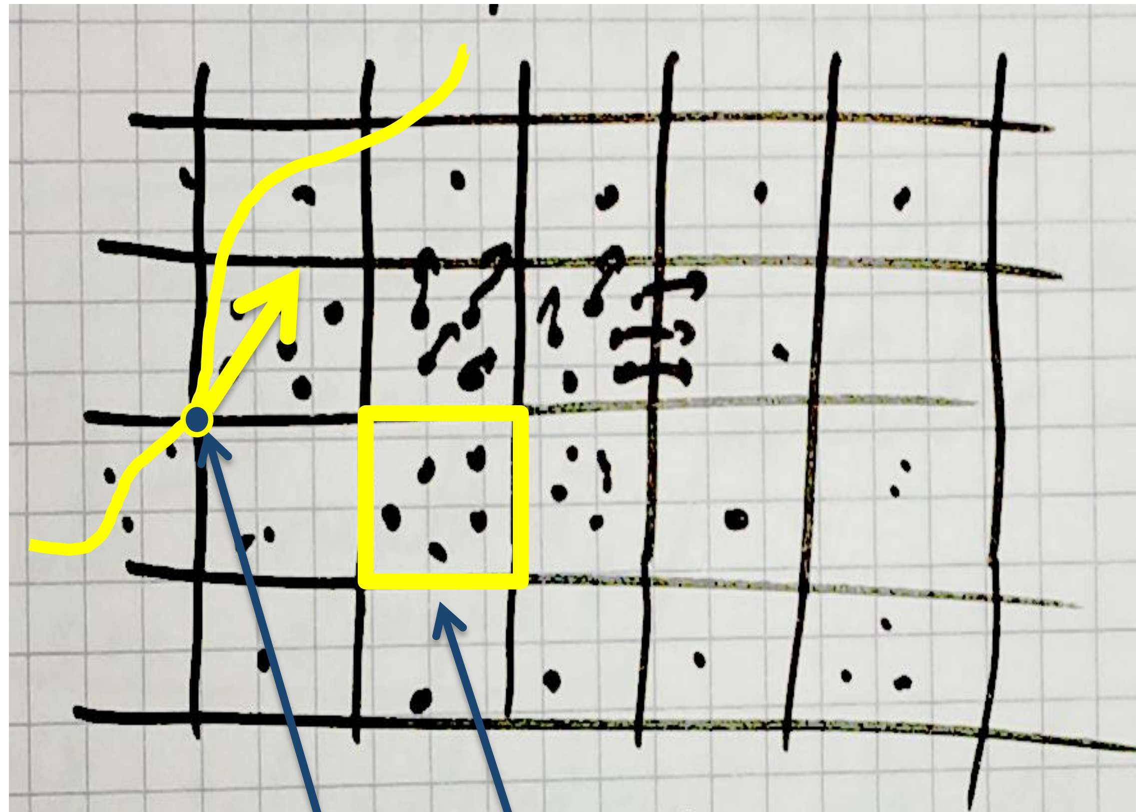
ρ(x,y,t) "nb particles per square"

v(x,y,t) speed of the particle under "grid point" (x,y) at time t

# Fluids

"Euler" point of view

Q1: how to compute the acceleration of the particles from v(x,y,t) ?



ρ(x,y,t) "nb particles per square"

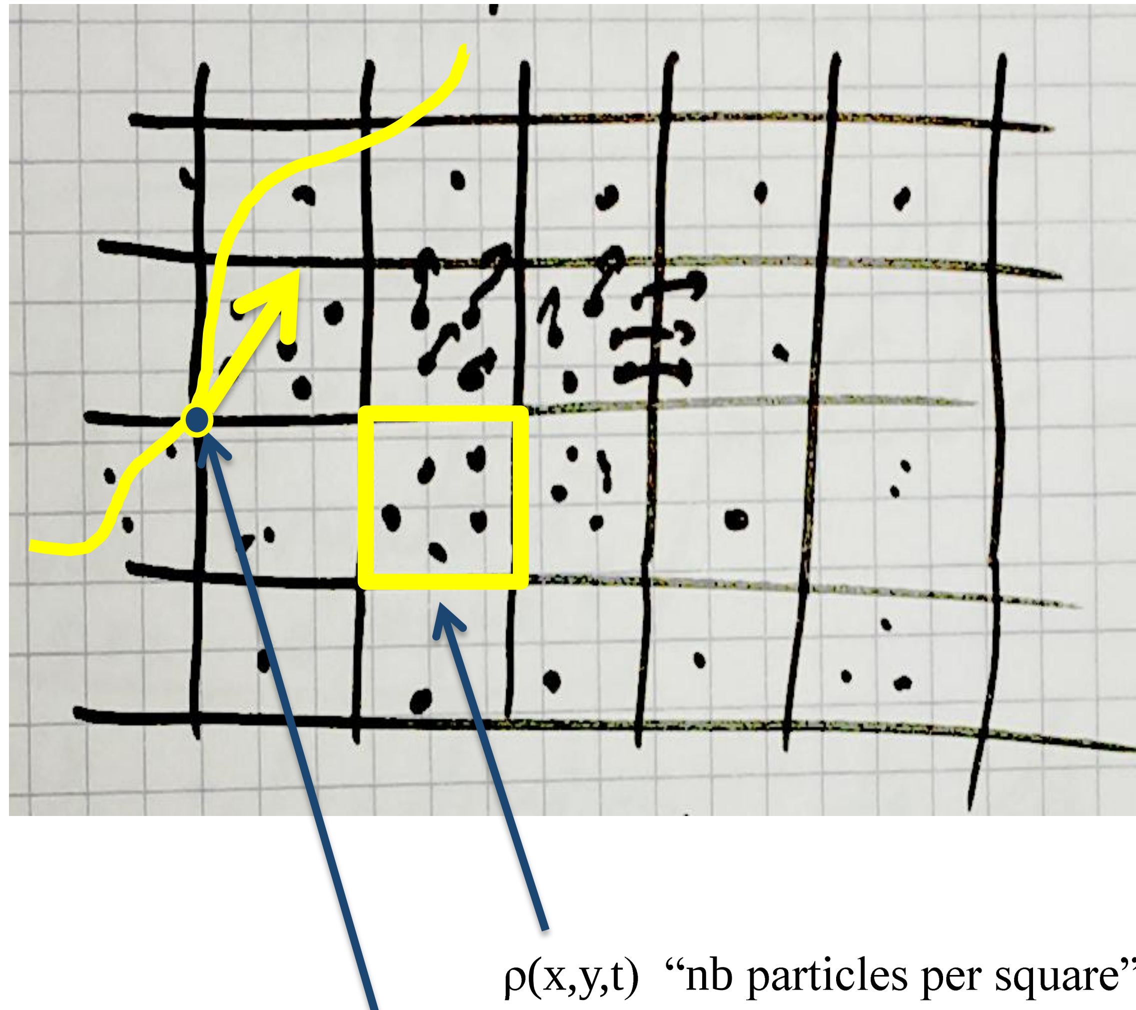
v(x,y,t) speed of the particle under "grid point" (x,y) at time t

# Fluids

"Euler" point of view

Q1: how to compute the acceleration of the particles from v(x,y,t) ?

$$\frac{dv}{dt} = \frac{\partial v}{\partial t} + \frac{\partial v}{\partial x}\frac{dx}{dt} + \frac{\partial v}{\partial y}\frac{dy}{dt}$$



ρ(x,y,t) "nb particles per square"

v(x,y,t) speed of the particle under "grid point" (x,y) at time t

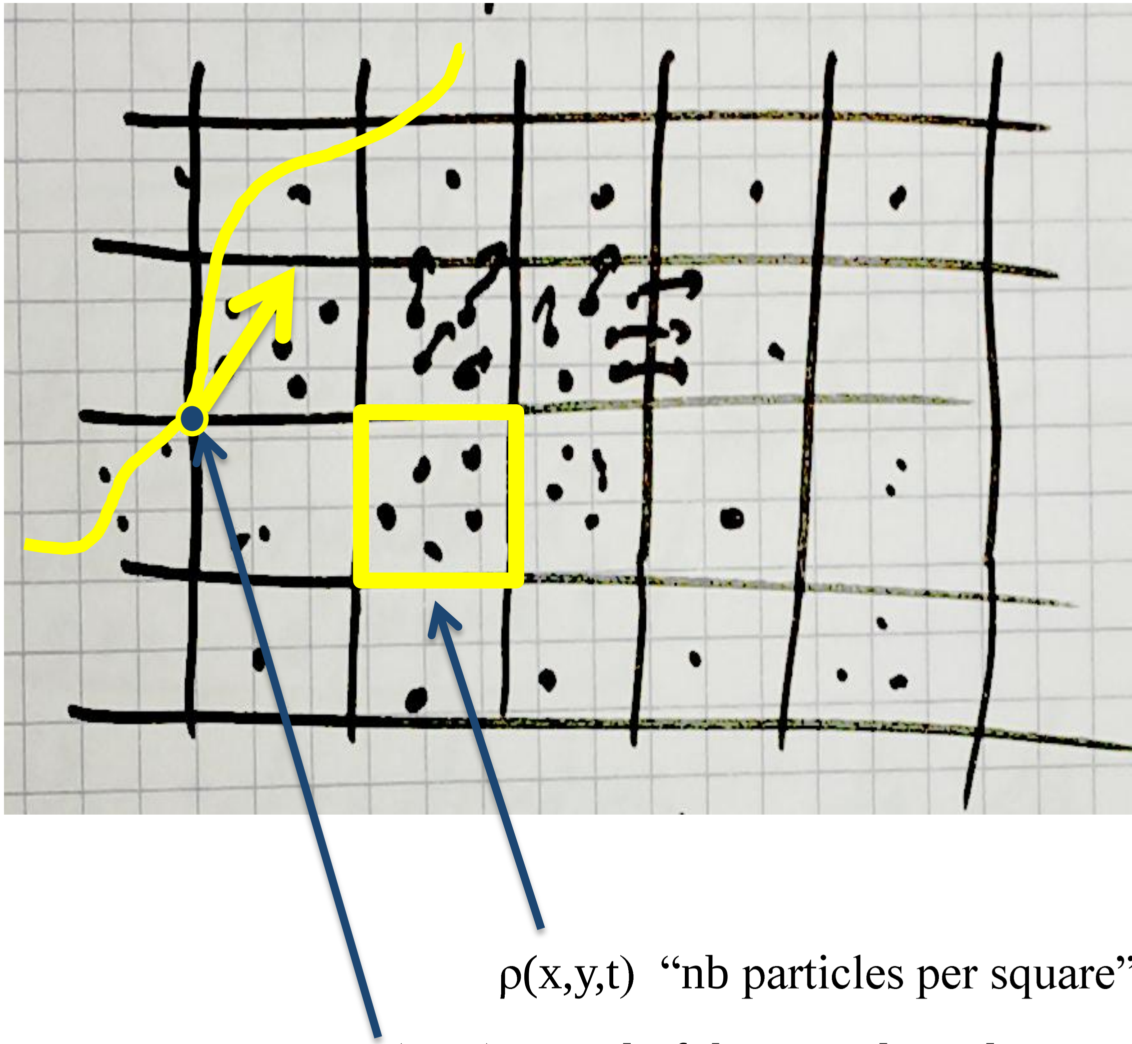
# Fluids

"Euler" point of view

Q1: how to compute the acceleration of the particles from v(x,y,t) ?

$$\frac{dv}{dt} = \frac{\partial v}{\partial t} + \frac{\partial v}{\partial x}\frac{dx}{dt} + \frac{\partial v}{\partial y}\frac{dy}{dt}$$

$$= \frac{\partial v}{\partial t} + \frac{\partial v}{\partial x} v_x + \frac{\partial v}{\partial y} v_y$$



ρ(x,y,t) "nb particles per square"

v(x,y,t) speed of the particle under "grid point" (x,y) at time t

# Fluids

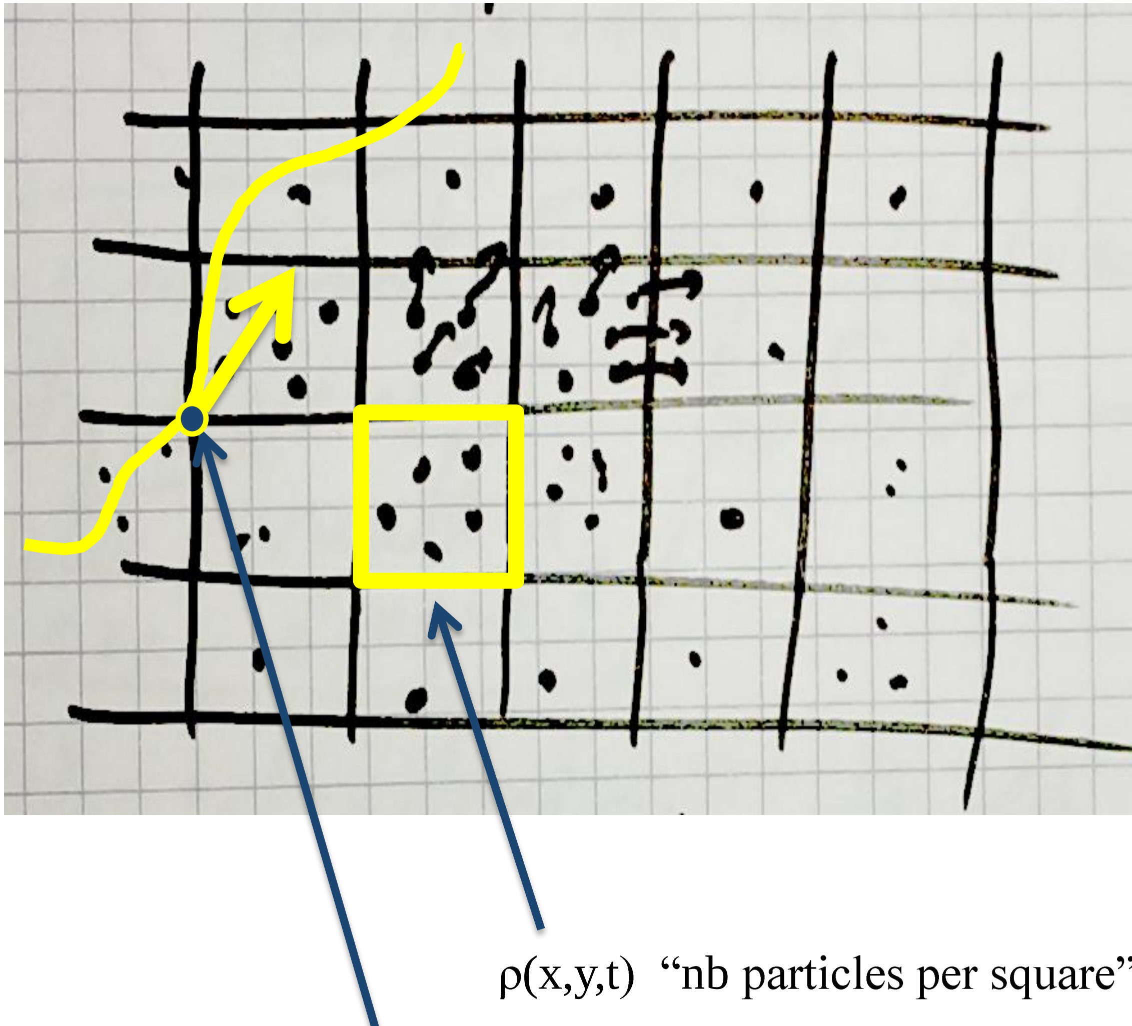
"Euler" point of view

Q1: how to compute the acceleration of the particles from v(x,y,t) ?

$$\frac{dv}{dt} = \frac{\partial v}{\partial t} + \frac{\partial v}{\partial x}\frac{dx}{dt} + \frac{\partial v}{\partial y}\frac{dy}{dt}$$

$$= \frac{\partial v}{\partial t} + \frac{\partial v}{\partial x} v_x + \frac{\partial v}{\partial y} v_y$$

$$= \frac{\partial v}{\partial t} + v.\nabla v$$



ρ(x,y,t) "nb particles per square"

v(x,y,t) speed of the particle under "grid point" (x,y) at time t

# Fluids

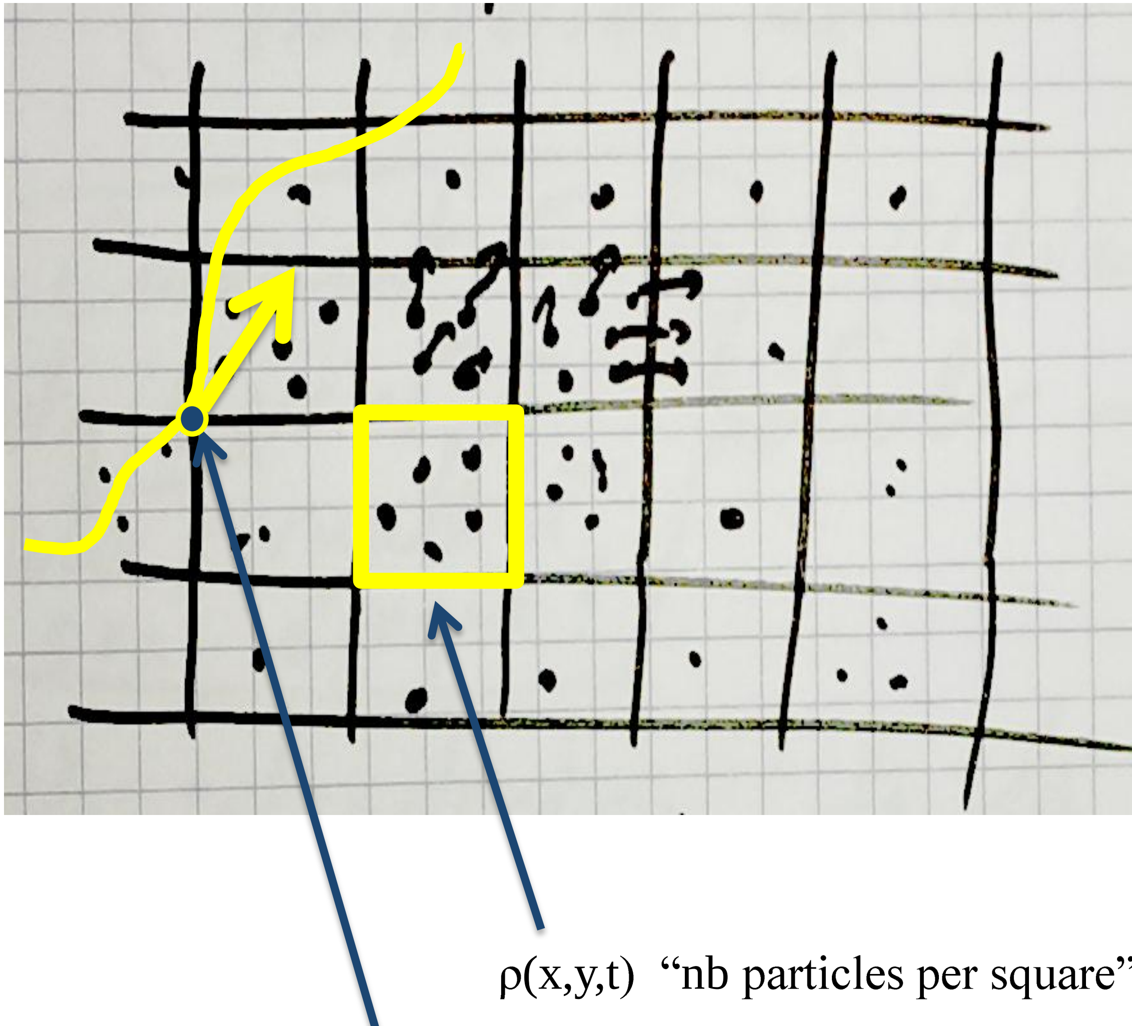
"Euler" point of view

Q1: how to compute the acceleration of the particles from v(x,y,t) ?

$$\frac{dv}{dt} = \frac{\partial v}{\partial t} + \frac{\partial v}{\partial x}\frac{dx}{dt} + \frac{\partial v}{\partial y}\frac{dy}{dt}$$

$$= \frac{\partial v}{\partial t} + \frac{\partial v}{\partial x} v_x + \frac{\partial v}{\partial y} v_y$$

$$= \frac{\partial v}{\partial t} + v.\nabla v$$



ρ(x,y,t) "nb particles per square"

v(x,y,t) speed of the particle under "grid point" (x,y) at time t

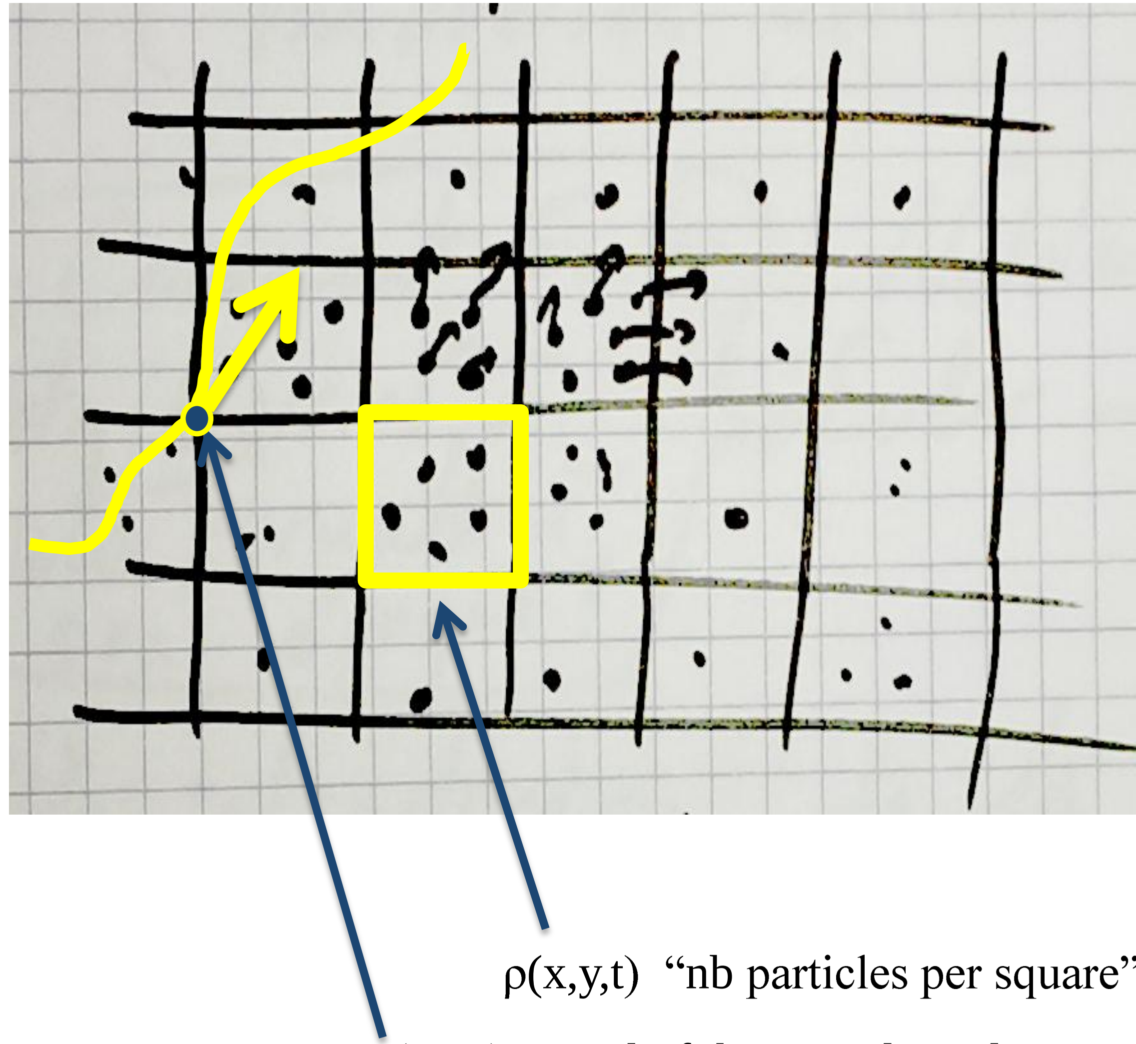
# Fluids

"Euler" point of view

Q1: how to compute the acceleration of the particles from v(x,y,t) ?

$$\frac{dv}{dt} = \frac{\partial v}{\partial t} + v.\nabla v$$

Q2: incompressible fluids ?



$\rho(x,y,t)$ "nb particles per square"

v(x,y,t) speed of the particle under "grid point" (x,y) at time t

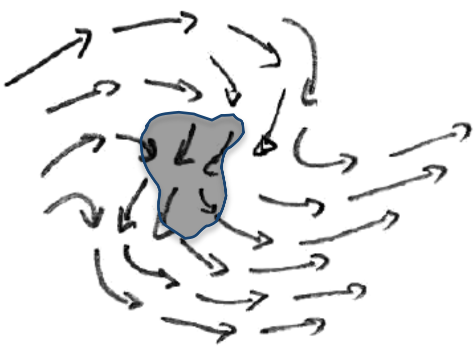
# Fluids

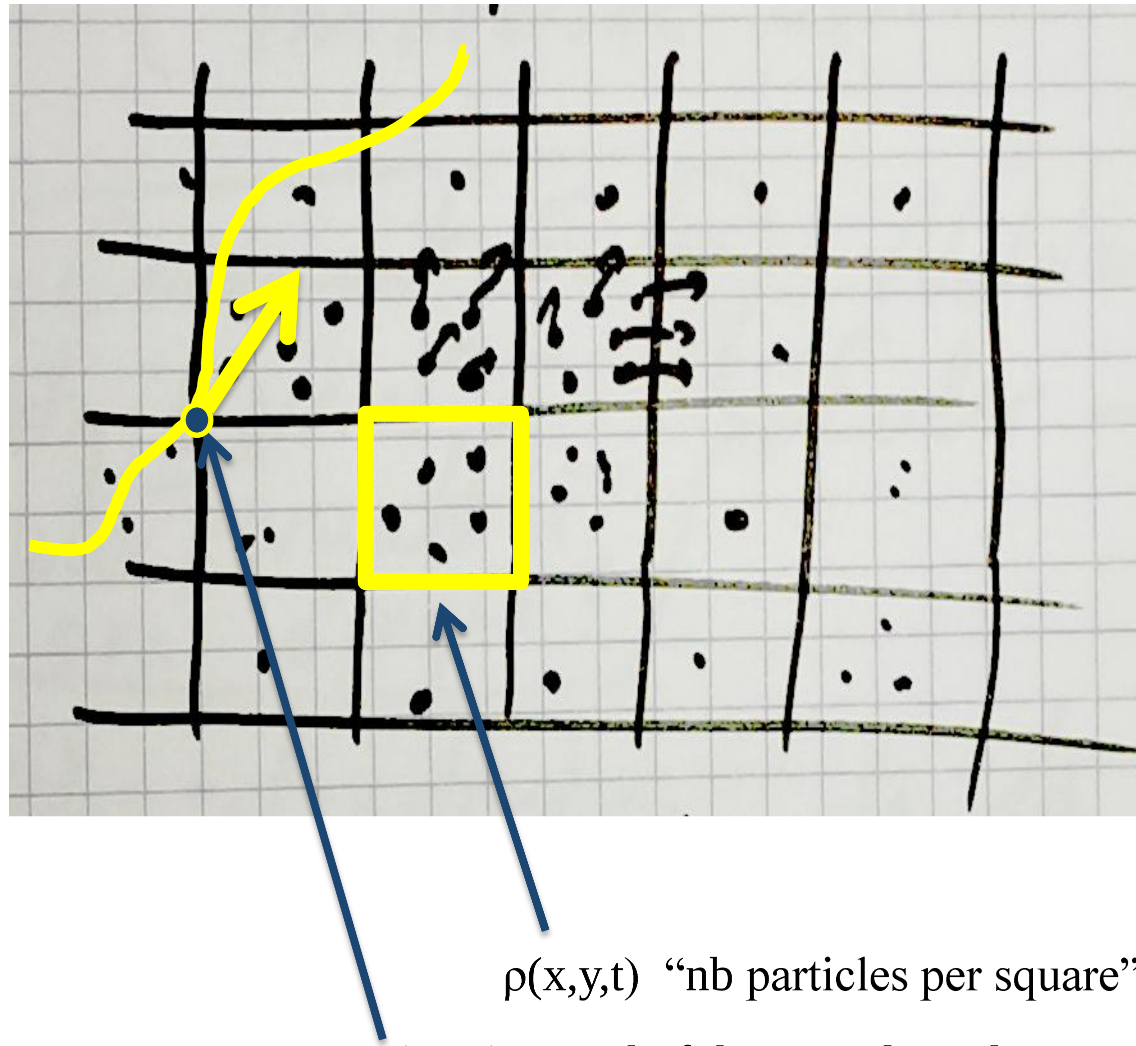
"Euler" point of view

Q1: how to compute the acceleration of the particles from v(x,y,t) ?

$$\frac{dv}{dt} = \frac{\partial v}{\partial t} + v.\nabla v$$

Q2: incompressible fluids ?





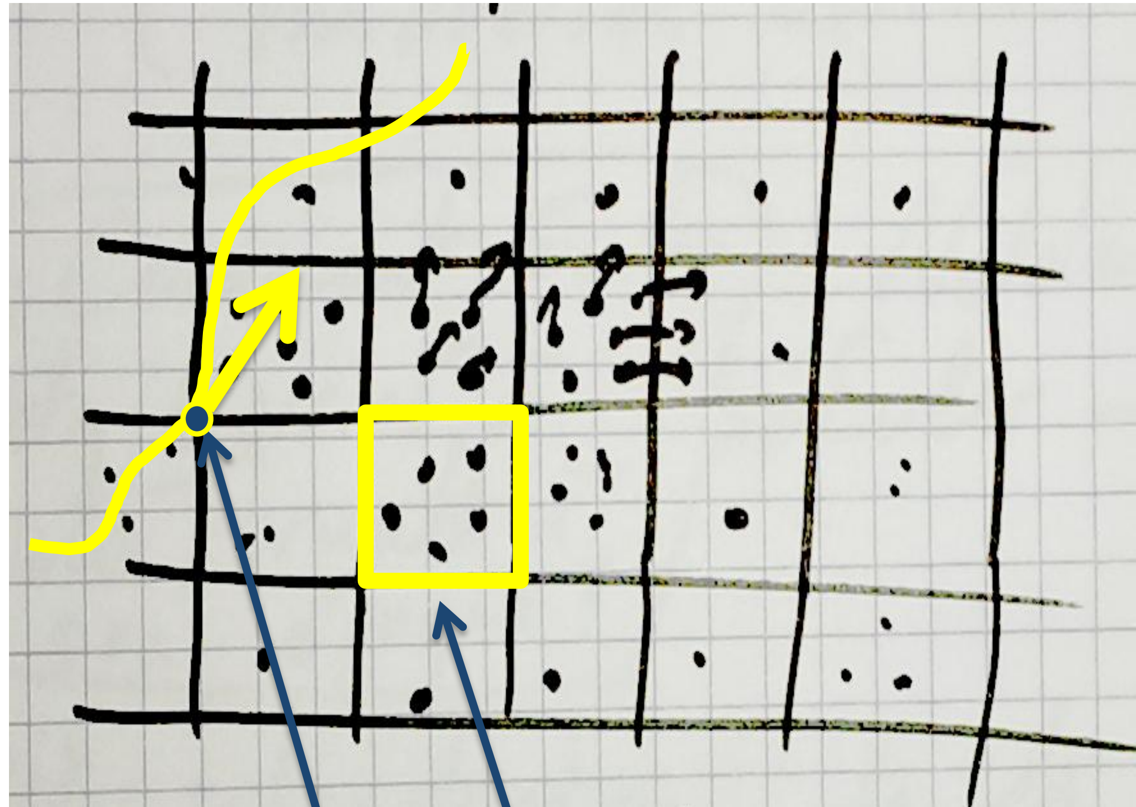
# Fluids

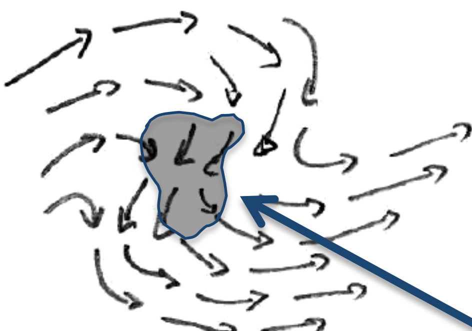
"Euler" point of view

Q1: how to compute the acceleration of the particles from v(x,y,t) ?

$$\frac{dv}{dt} = \frac{\partial v}{\partial t} + v.\nabla v$$

Q2: incompressible fluids ?





what goes in = what goes out

ρ(x,y,t) "nb particles per square"

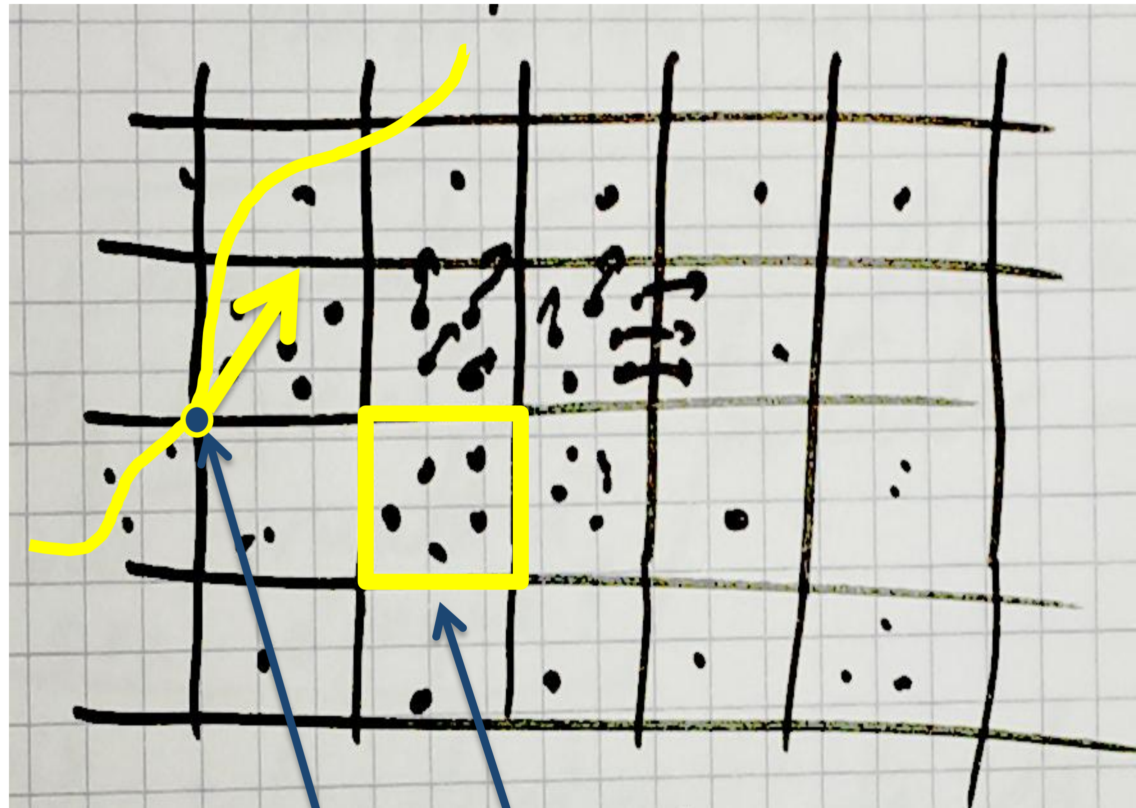
v(x,y,t) speed of the particle under "grid point" (x,y) at time t

# Fluids

"Euler" point of view

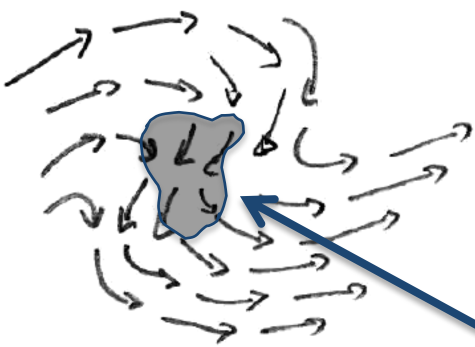
Q1: how to compute the acceleration of the particles from v(x,y,t) ?

$$\frac{dv}{dt} = \frac{\partial v}{\partial t} + v.\nabla v$$



Q2: incompressible fluids ?

$$div(v) = 0$$



what goes in =
what goes out

ρ(x,y,t) "nb particles per square"

v(x,y,t) speed of the particle under "grid point" (x,y) at time t

# Fluids

"Euler" point of view

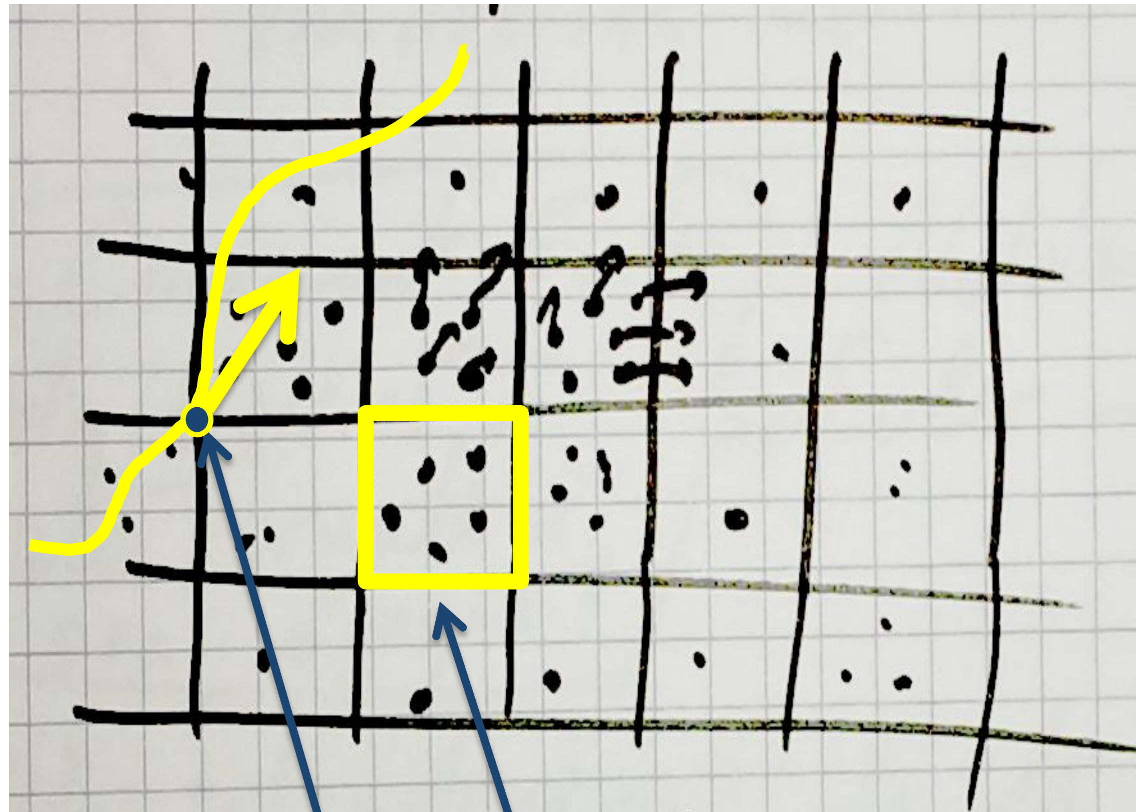
Q1: how to compute the acceleration of the particles from v(x,y,t) ?

$$\frac{dv}{dt} = \frac{\partial v}{\partial t} + v.\nabla v$$

Q2: incompressible fluids ?

$$div(v) = 0$$

Q3: mass preservation ?



ρ(x,y,t) "nb particles per square"

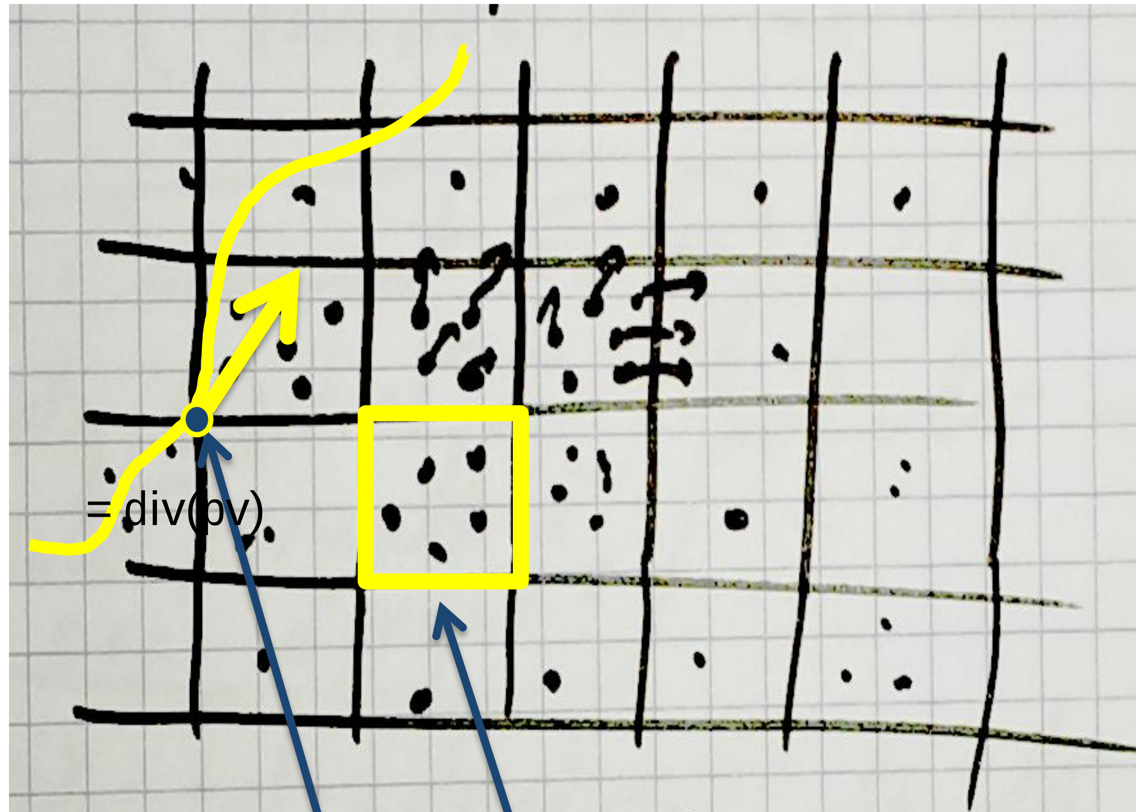
v(x,y,t) speed of the particle under "grid point" (x,y) at time t

# Fluids

"Euler" point of view

Q1: how to compute the acceleration of the particles from v(x,y,t) ?

$$\frac{dv}{dt} = \frac{\partial v}{\partial t} + v.\nabla v$$




ρ(x,y,t) "nb particles per square"

v(x,y,t) speed of the particle under "grid point" (x,y) at time t

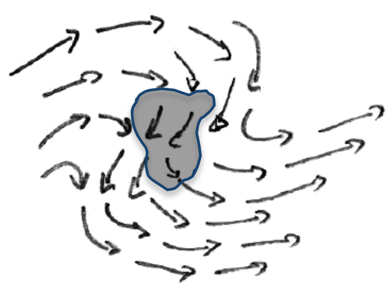
Q2: incompressible fluids ?

$$div(v) = 0$$

Q3: mass preservation ?

$$\frac{d\rho}{dt} = - div(\rho v)$$

(Continuity equation)



# Fluids

Start with Lagrange coordinates:

"particle trajectories": X(t,x)

($\rho$ = cte)

Minimize Action: $$1/2 \int_{t_1}^{t_2} \int_{\Omega} \left\| \frac{\partial X(t,x)}{\partial t} \right\|^2 dxdt$$

s.t. X satisfies mass preservation (X is measure-preserving, more on this later…)

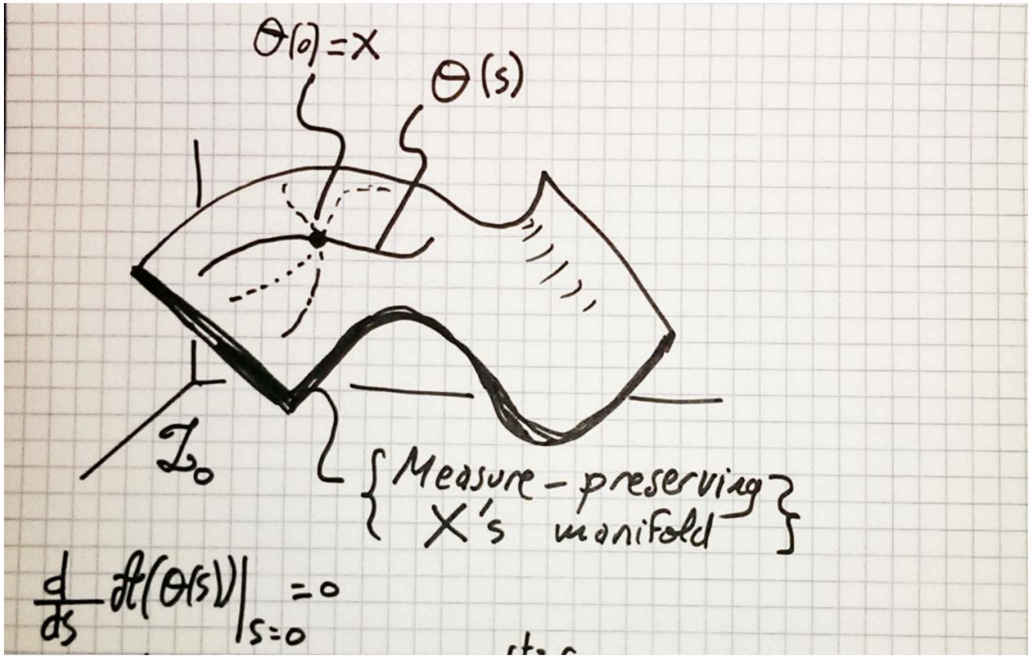
# Fluids

Start with Lagrange coordinates:

"particle trajectories": $X(t,x)$

$(\rho = cte)$

Minimize Action: $$\frac{1}{2}\int_{t_1}^{t_2}\int_{\Omega}\left\|\frac{\partial X(t,x)}{\partial t}\right\|^2 dxdt$$




s.t. X satisfies mass preservation
(X is measure-preserving, more on this later…)

# Fluids

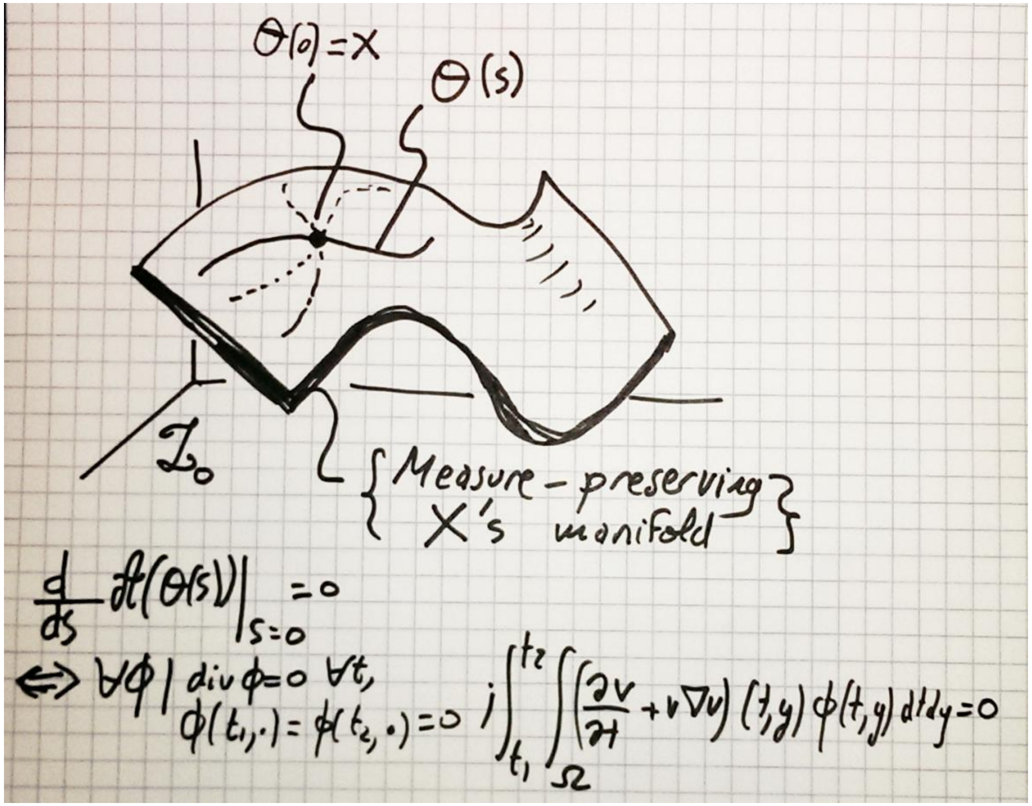
Start with Lagrange coordinates:

"particle trajectories": $X(t,x)$

($\rho$ = cte)

Minimize Action: $$\frac{1}{2}\int_{t_1}^{t_2}\int_{\Omega}\left\|\frac{\partial X(t,x)}{\partial t}\right\|^2 dxdt$$

s.t. X satisfies mass preservation (X is measure-preserving, more on this later…)



# Fluids

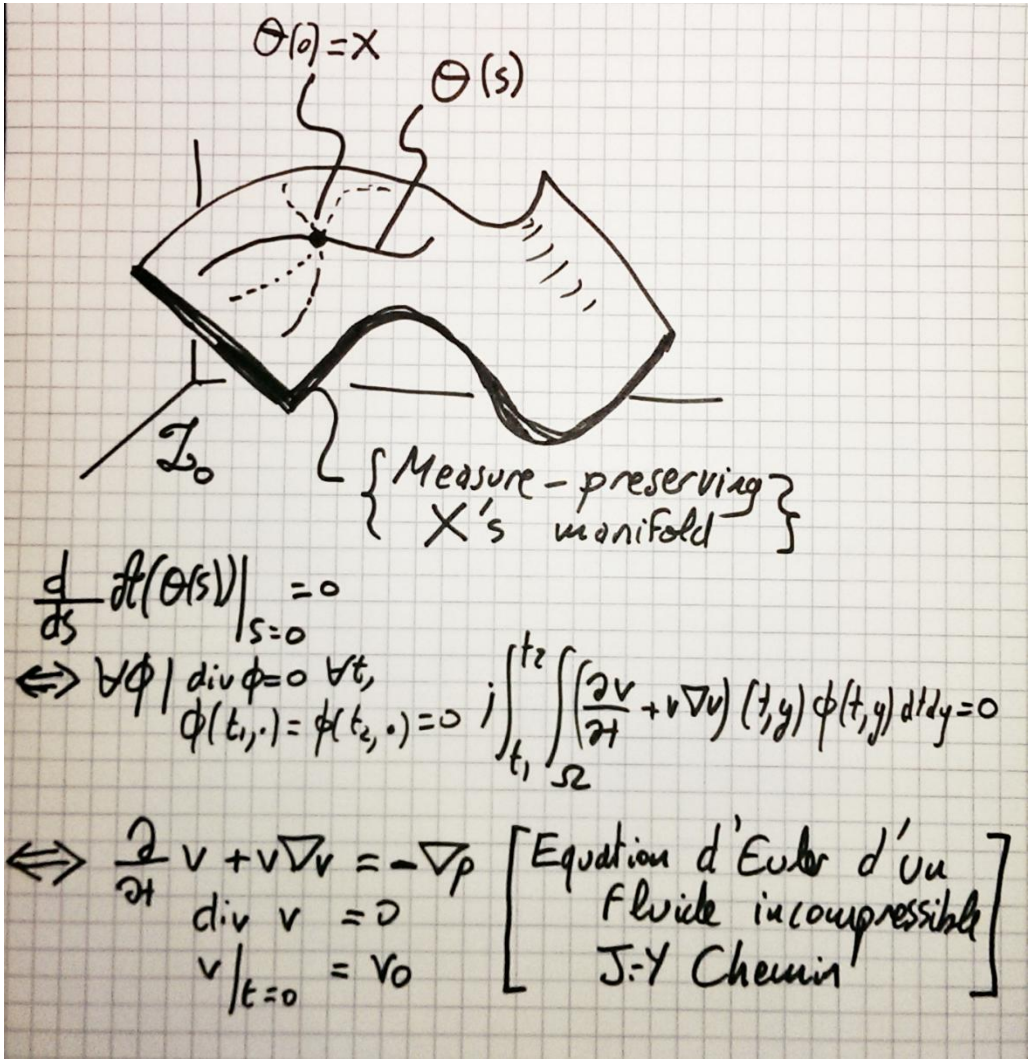
Start with Lagrange coordinates:

"particle trajectories": $X(t,x)$

($\rho$ = cte)

Minimize Action: $$\frac{1}{2}\int_{t_1}^{t_2}\int_{\Omega}\left\|\frac{\partial X(t,x)}{\partial t}\right\|^2 dx\,dt$$

s.t. X satisfies mass preservation (X is measure-preserving, more on this later…)



# Fluids

Start with Lagrange coordinates:

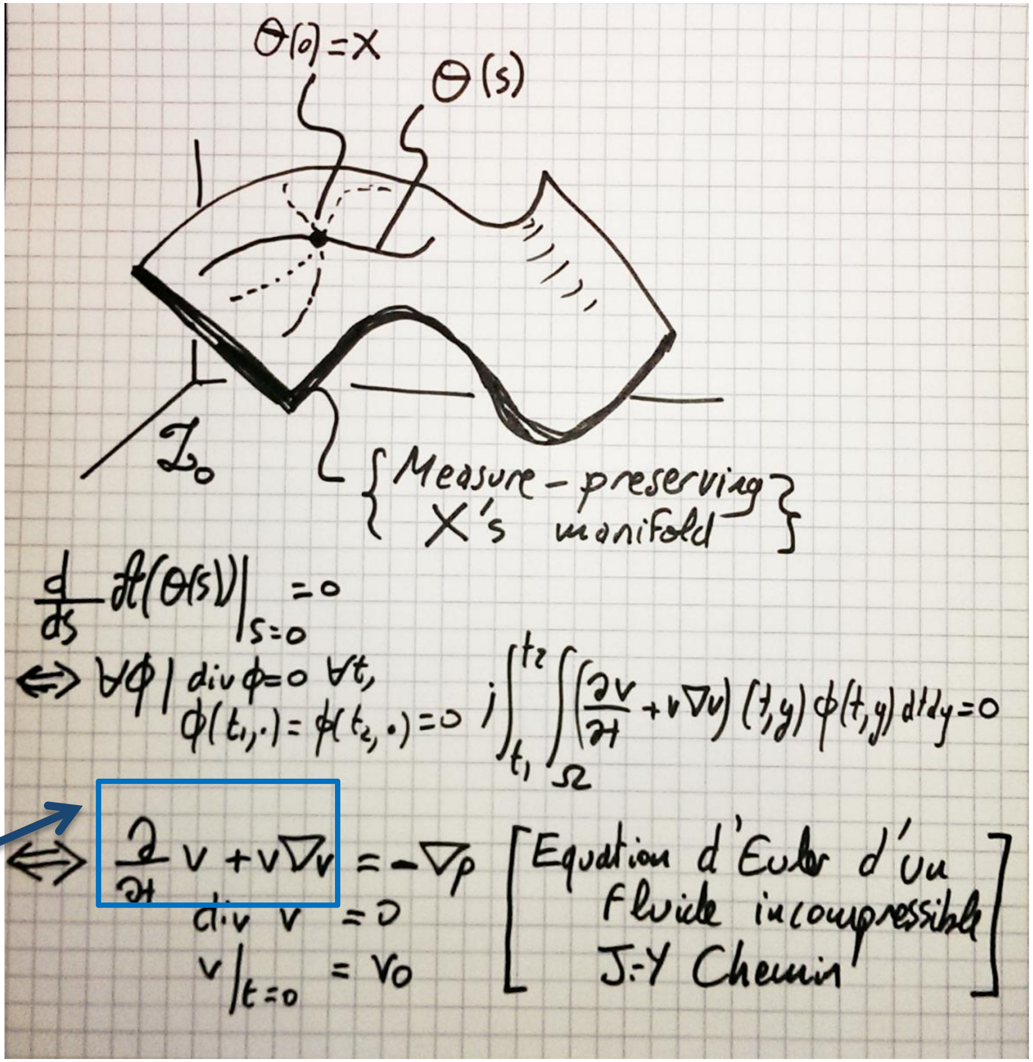
"particle trajectories": X(t,x)

($\rho$ = cte)

Minimize Action: $$1/2 \int_{t_1}^{t_2} \int_{\Omega} \left\| \frac{\partial X(t,x)}{\partial t} \right\|^2 dxdt$$

s.t. X satisfies mass preservation
(X is measure-preserving, more on this later…)

Acceleration of the particle "under the grid"



# Fluids

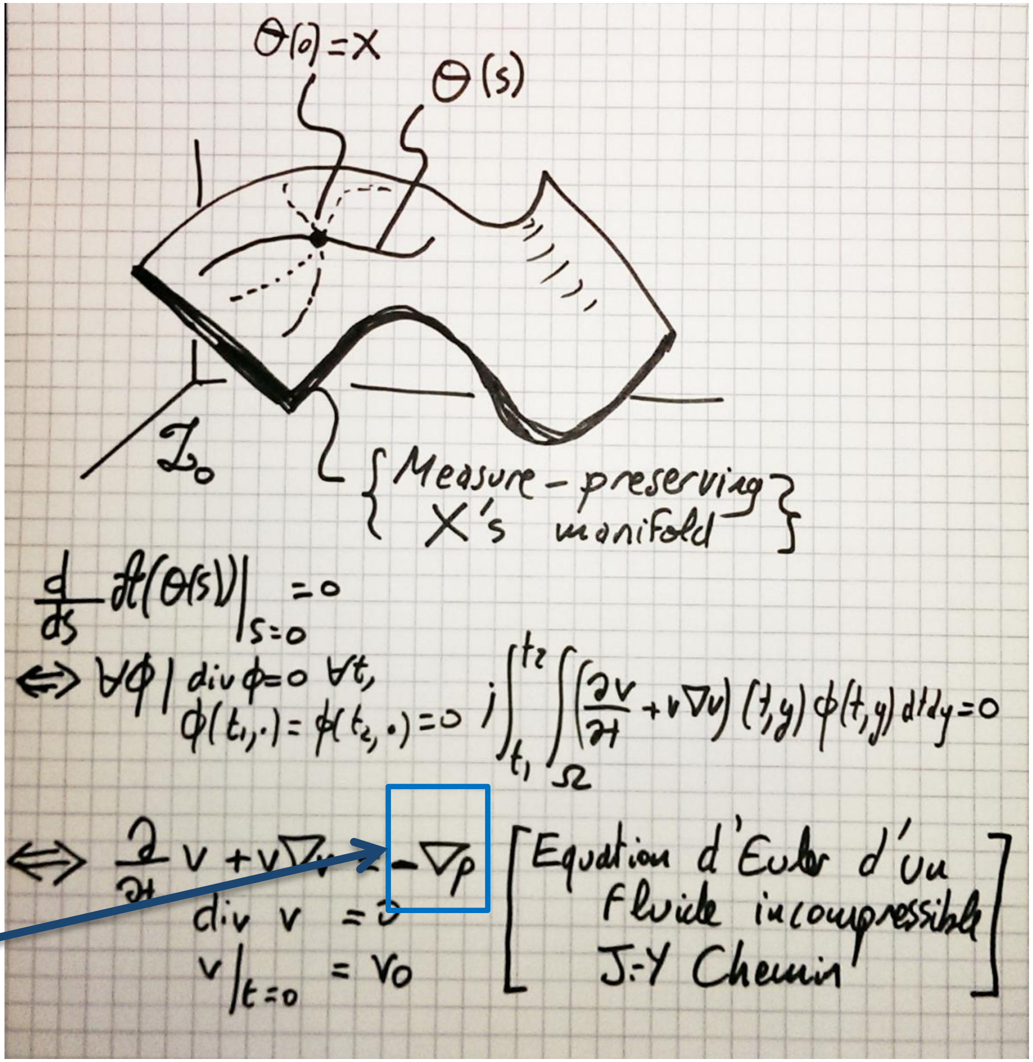
Start with Lagrange coordinates:
"particle trajectories": $X(t,x)$

($\rho$ = cte)

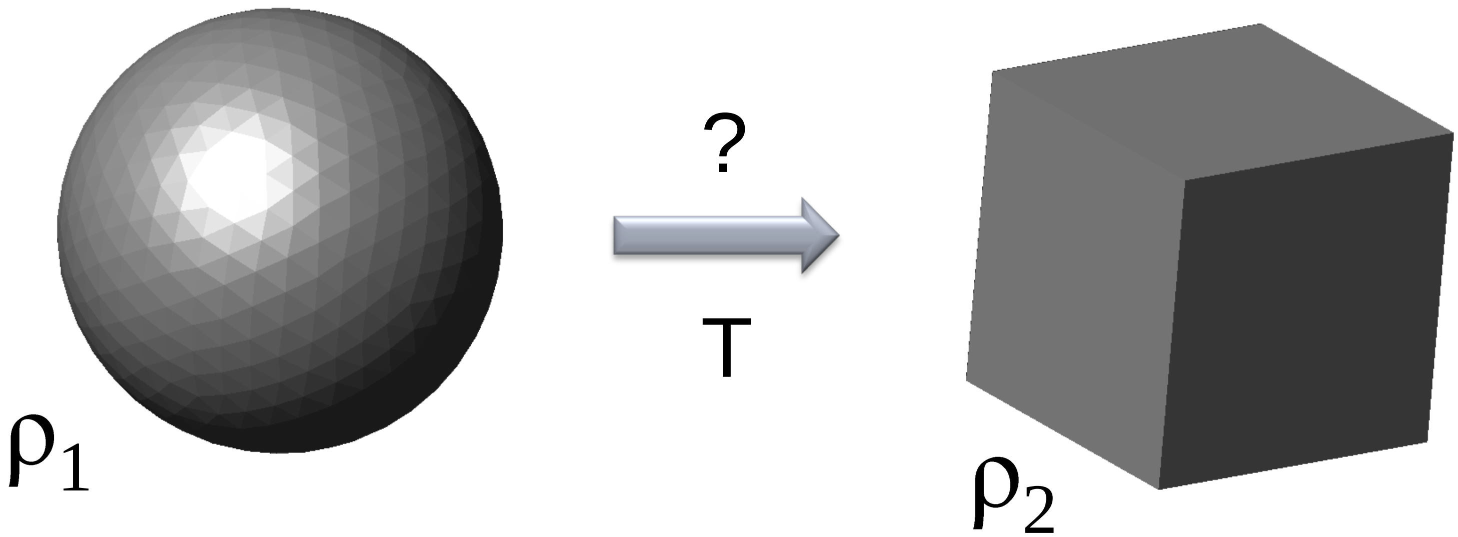
Minimize Action: $$1/2 \int_{t_1}^{t_2} \int_{\Omega} \left\| \frac{\partial X(t,x)}{\partial t} \right\|^2 dxdt$$

s.t. X satisfies mass preservation (X is measure-preserving, more on this later…)

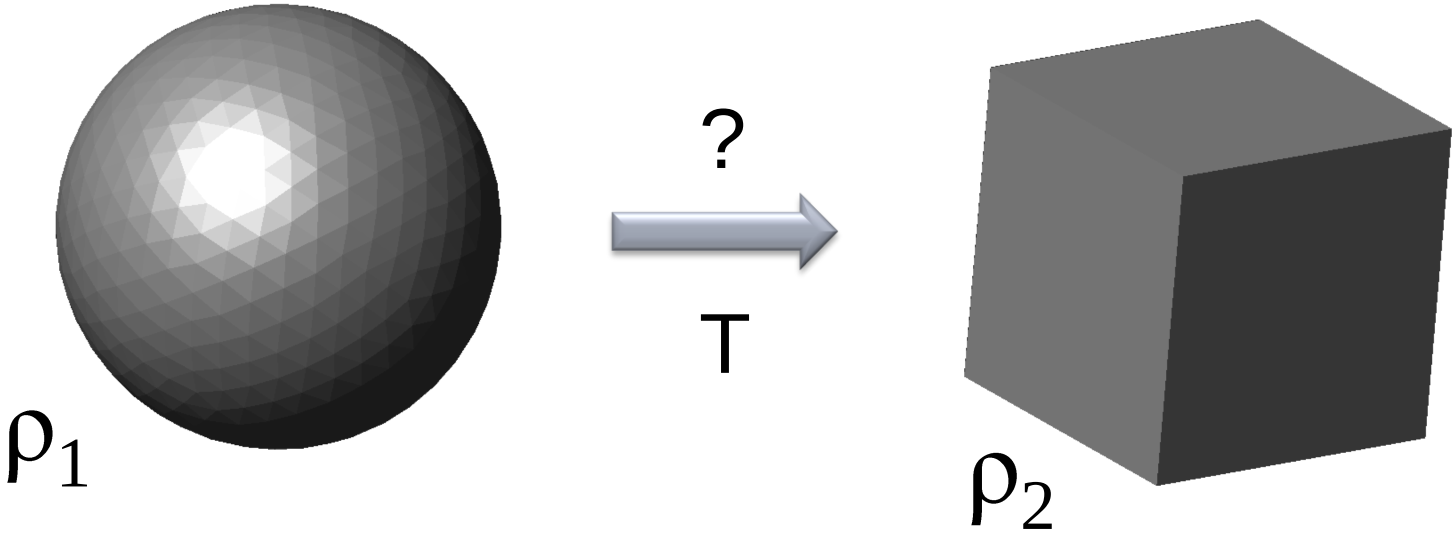
The "Lagrange multiplier" for the constraint = pressure



# Fluids – Benamou Brenier



# Fluids – Benamou Brenier

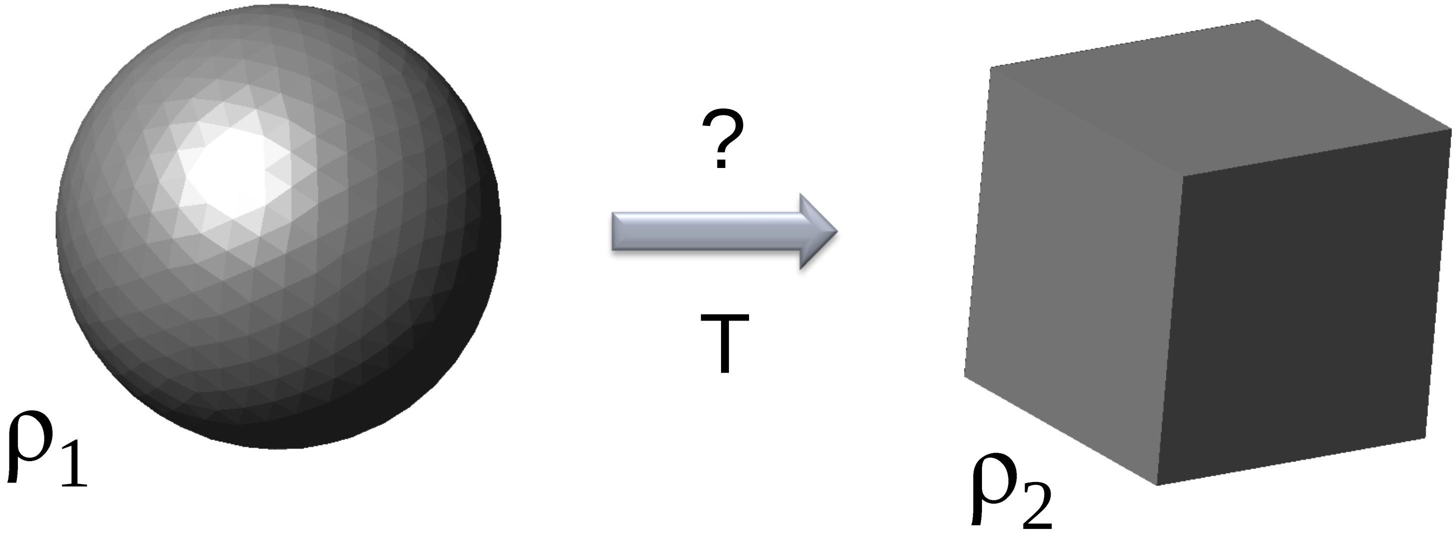



Minimize

$$A(\rho,v) = (t_2-t_1) \int_{t_1}^{t_2} \int_{\Omega} \rho(x,t)\, \|v(t,x)\|^2 dx dt$$

s.t. $\rho(t_1,.) = \rho_1$ ; $\rho(t_2,.) = \rho_2$ ; $\frac{d\rho}{dt} = -\operatorname{div}(\rho v)$

# Fluids – Benamou Brenier




Minimize

$$A(\rho,v) = (t_2 - t_1) \int_{t_1}^{t_2} \int_{\Omega} \rho(x,t)\, \|v(t,x)\|^2 dxdt$$

s.t. $\rho(t_1,.) = \rho_1$ ; $\rho(t_2,.) = \rho_2$ ; $\frac{d\rho}{dt} = -\operatorname{div}(\rho v)$

⟷

Minimize C(T) =

$$\int_{\Omega} \rho_1(x) \| x - T(x) \|2\, dx$$

s.t. T is measure-preserving

# 2

## Optimal Transport

### an elementary introduction

# **Part. 2** Optimal Transport – Monge's problem

(X;μ)

(Y;ν)

Two measures μ, ν such that $\int_X d\mu(x) = \int_Y d\nu(x)$

# **Part. 2** Optimal Transport – Monge's problem

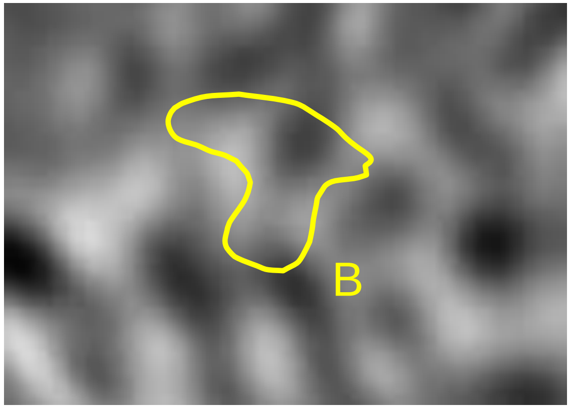
(X;μ)

(Y;ν)

A map T is a ***transport map*** between μ and ν if
$\mu(T^{-1}(B)) = \nu(B)$ for any Borel subset B of Y

# **Part. 2** Optimal Transport – Monge's problem

(X;μ)

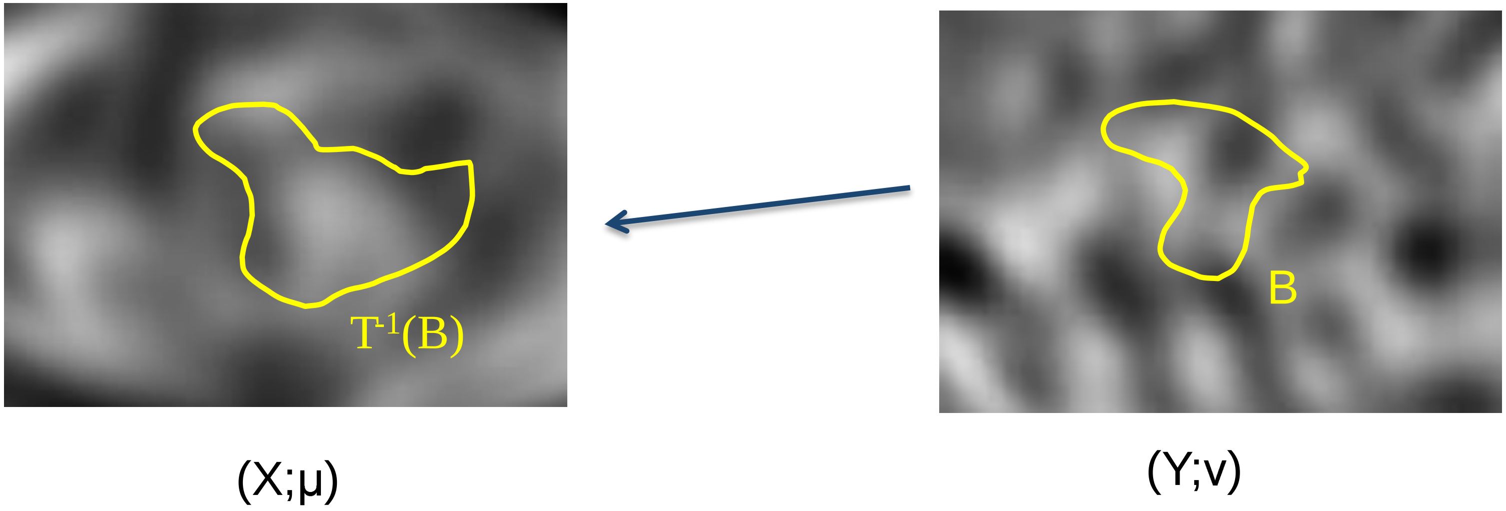



(Y;ν)

A map T is a ***transport map*** between μ and ν if

$\mu(T^{-1}(B)) = \nu(B)$ for any Borel subset B

# Part. 2 Optimal Transport – Monge's problem




A map T is a ***transport map*** between μ and ν if

$$\mu(T^{-1}(B)) = \nu(B) \text{ for any Borel subset } B$$

# Part. 2 Optimal Transport – Monge's problem

$(X;\mu)$

$(Y;\nu)$

A map T is a ***transport map*** between $\mu$ and $\nu$ if

$\mu(T^{-1}(B)) = \nu(B)$ for any Borel subset B

(or $\nu = T\#\mu$ the ***pushforward*** of $\mu$)

# **Part. 2** Optimal Transport – Monge's problem

(X;μ)

(Y;ν)

Monge's problem (1787):

Find a transport map T that minimizes $C(T) = \int_X \| x - T(x) \|^2 \, d\mu(x)$

# **Part. 2** Optimal Transport – Monge's problem

***Monge's problem:***

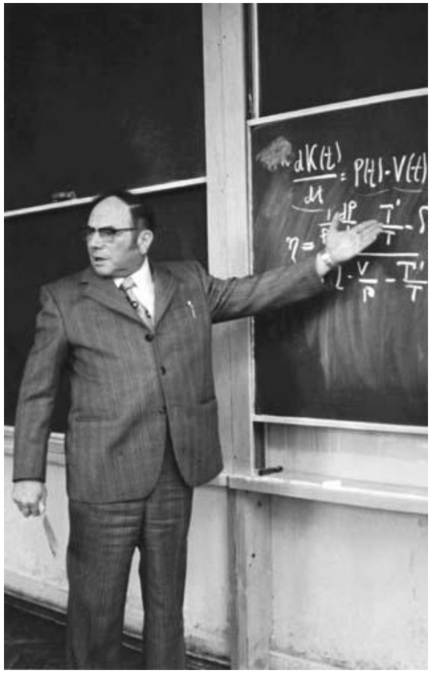
Find a transport map T that minimizes $C(T) = \int_X \| x - T(x) \|^2 \, d\mu(x)$

- Difficult to study
- If μ has an atom (isolated Dirac),
  it can only be mapped to another Dirac
  (T needs to be a map)

# **Part. 2** Optimal Transport – Kantorovich

***Monge's problem:***

Find a transport map T that minimizes $C(T) = \int_X \| x - T(x) \|^2 \, d\mu(x)$




***Kantorovich's problem* (1942):**

Find a measure γ defined on X x Y

such that $\int_{x \text{ in } X} d\gamma(x,y) = d\nu(y)$

and $\int_{y \text{ in } Y} d\gamma(x,y) = d\mu(x)$

that minimizes $\iint_{X \times Y} \| x - y \|^2 \, d\gamma(x,y)$

# Part. 2 Optimal Transport – Kantorovich

***Monge's problem:***

Find a transport map T that minimizes $C(T) = \int_X || x - T(x) ||^2 d\mu(x)$

***Kantorovich's problem:***

Find a measure γ defined on X x Y

such that $\int_{x \text{ in } X} d\gamma(x,y) = d\nu(y)$

and $\int_{y \text{ in } Y} d\gamma(x,y) = d\mu(x)$

that minimizes $\iint_{X \times Y} || x - y ||^2 d\gamma(x,y)$

"γ(x,y)" :
How much sand goes from x to y

# **Part. 2** Optimal Transport – Kantorovich

***Monge's problem:***

Find a transport map T that minimizes $C(T) = \int_X || x - T(x) ||^2 \, d\mu(x)$

***Kantorovich's problem:***

Find a measure γ defined on X x Y

such that $\int_{x \text{ in } X} d\gamma(x,y) = d\nu(y)$

and $\int_{y \text{ in } Y} d\gamma(x,y) = d\mu(x)$

Everything that is transported **from x** sums to "μ(x)"

that minimizes $\iint_{X \times Y} || x - y ||^2 \, d\gamma(x,y)$

# **Part. 2** Optimal Transport – Kantorovich

***Monge's problem:***

Find a transport map T that minimizes $C(T) = \int_X \|x - T(x)\|^2 \, d\mu(x)$

***Kantorovich's problem:***

Find a measure γ defined on X x Y

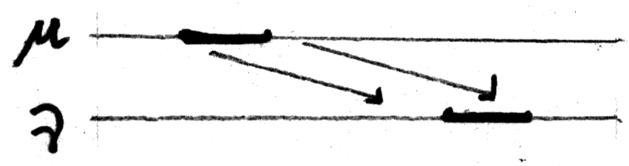
such that $\int_{x \text{ in } X} d\gamma(x,y) = d\nu(y)$

and $\int_{y \text{ in } Y} d\gamma(x,y) = d\mu(x)$

Everything that is transported **to y** sums to "v(y)"

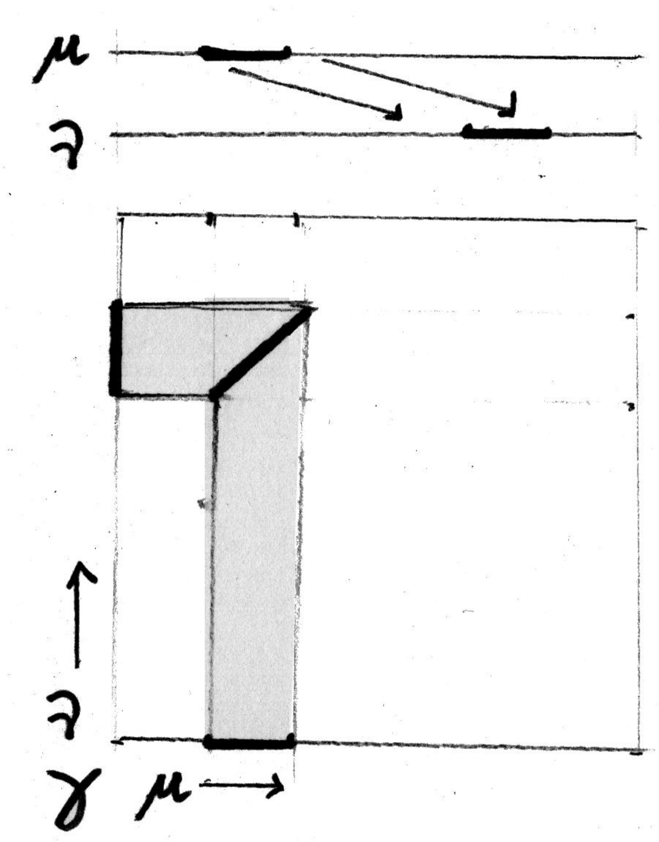
that minimizes $\iint_{X \times Y} \|x - y\|^2 \, d\gamma(x,y)$

# **Part. 2** Optimal Transport – Kantorovich



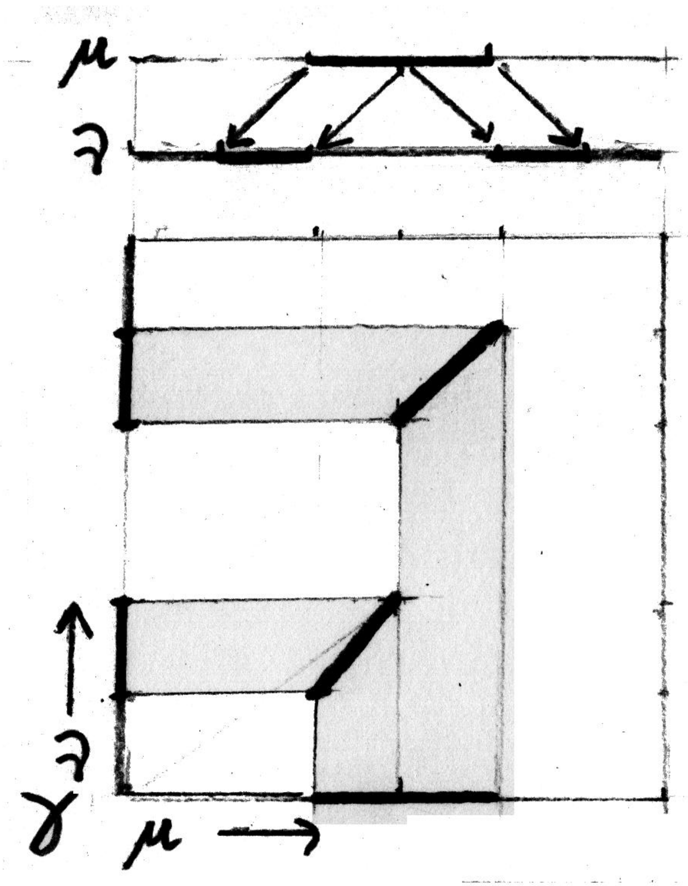

Transport plan – example 1/2 : translation of a segment

# **Part. 2** Optimal Transport – Kantorovich



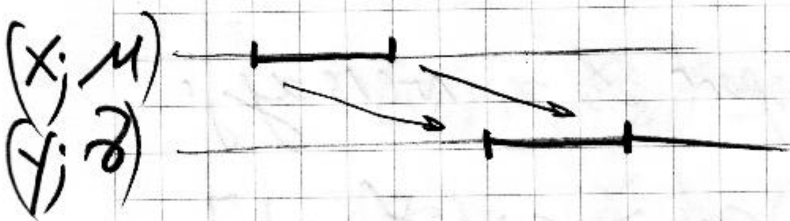

Transport plan – example 1/2 : translation of a segment

# **Part. 2** Optimal Transport – Kantorovich




Transport plan – example 2/2 : spitting a segment

# **Part. 2** Optimal Transport – Duality





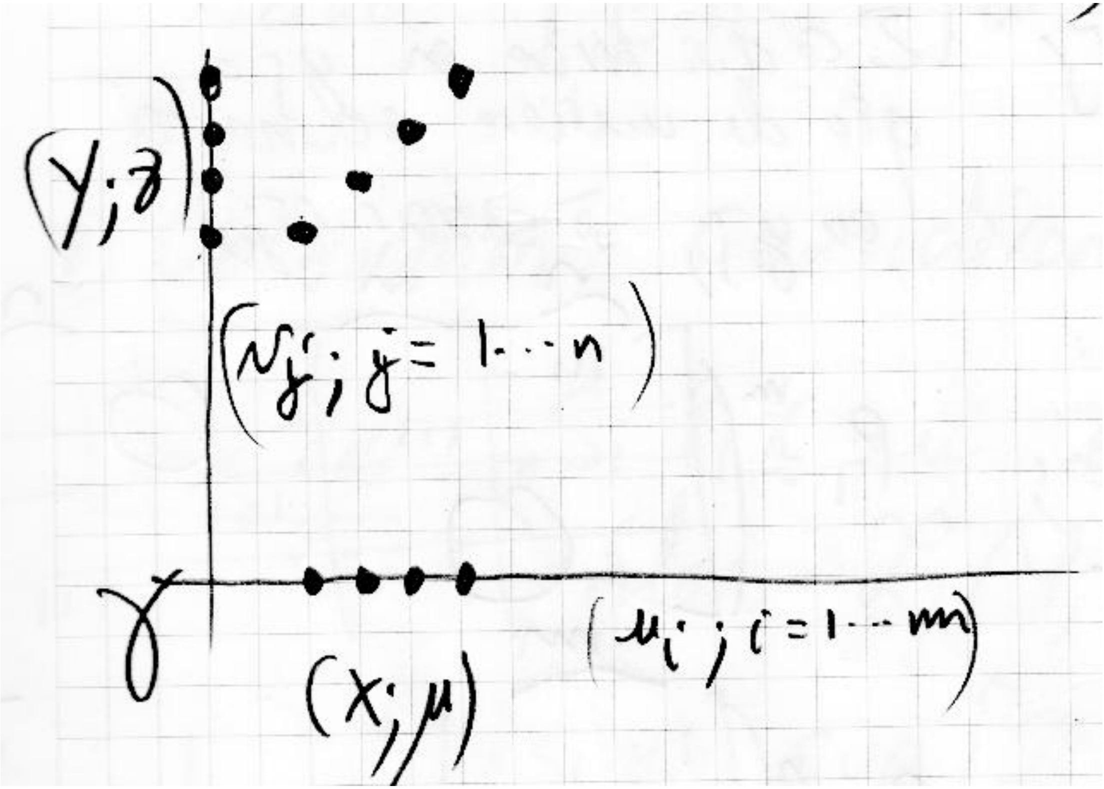
$$\text{Inf}_{\gamma} \int_{X\times Y} c(x,y)\, d\gamma$$

avec:

$\forall B \subset X, \int_B d\mu = \int_{B\times Y} d\gamma \quad (P_1 \# \gamma = \mu)$

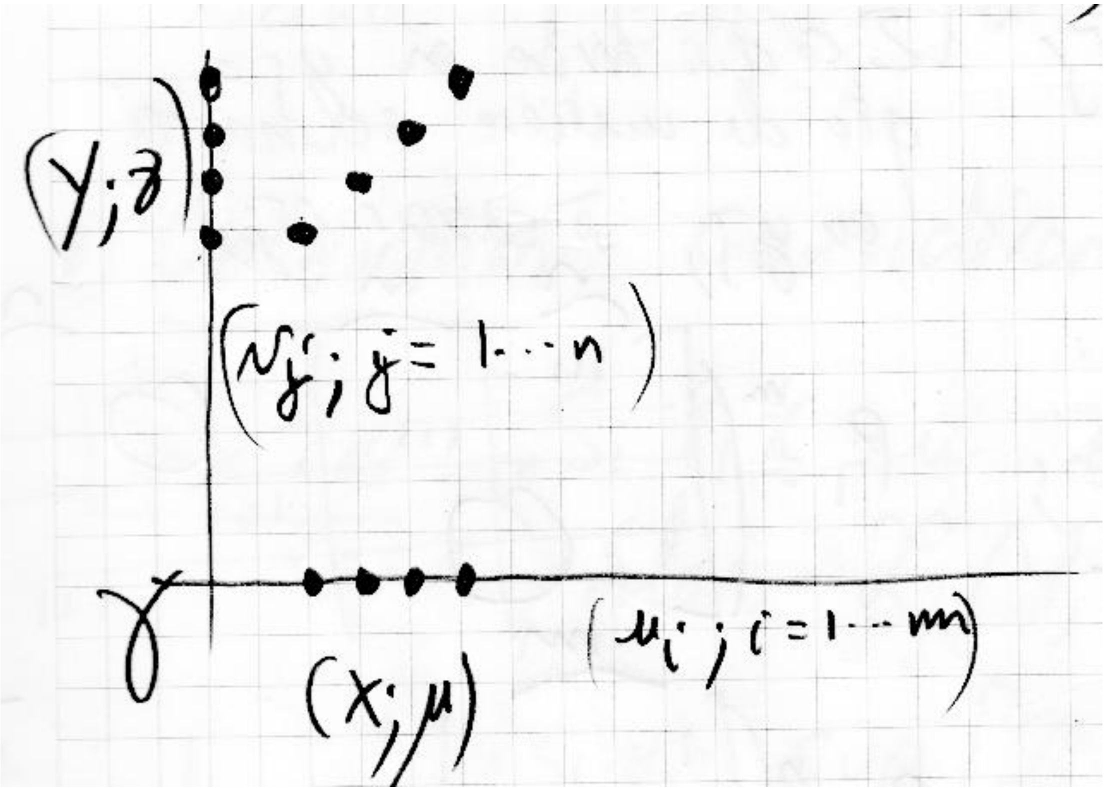
$\forall B \subset Y, \int_B d\nu = \int_{X\times B} d\gamma \quad (P_1 \# \gamma = \nu)$

# **Part. 2** Optimal Transport – Duality




Duality is easier to understand with a discrete version
Then we'll go back to the continuous setting.

# **Part. 2** Optimal Transport – Duality

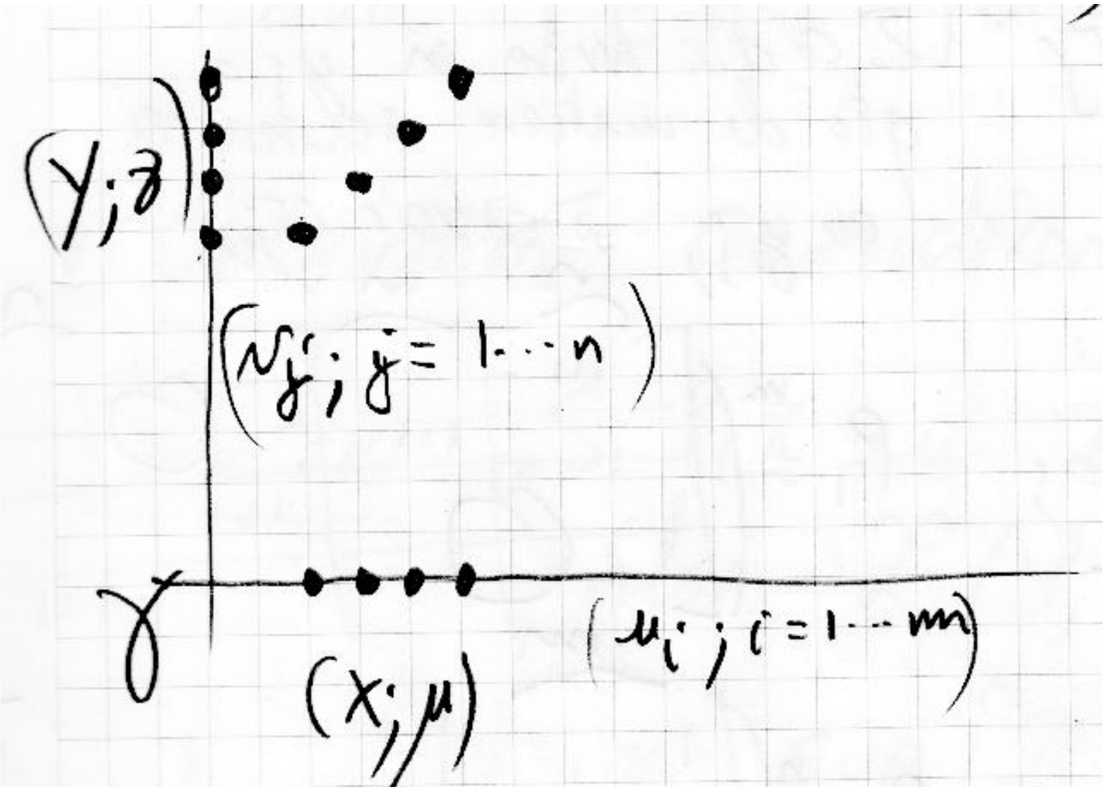



**(DMK):**

$$\text{Min} \quad \langle c, \gamma \rangle$$

$$s.t. \quad \begin{cases} P_1 \gamma = u \\ P_2 \gamma = v \\ \gamma \geq 0 \end{cases}$$

# **Part. 2** Optimal Transport – Duality



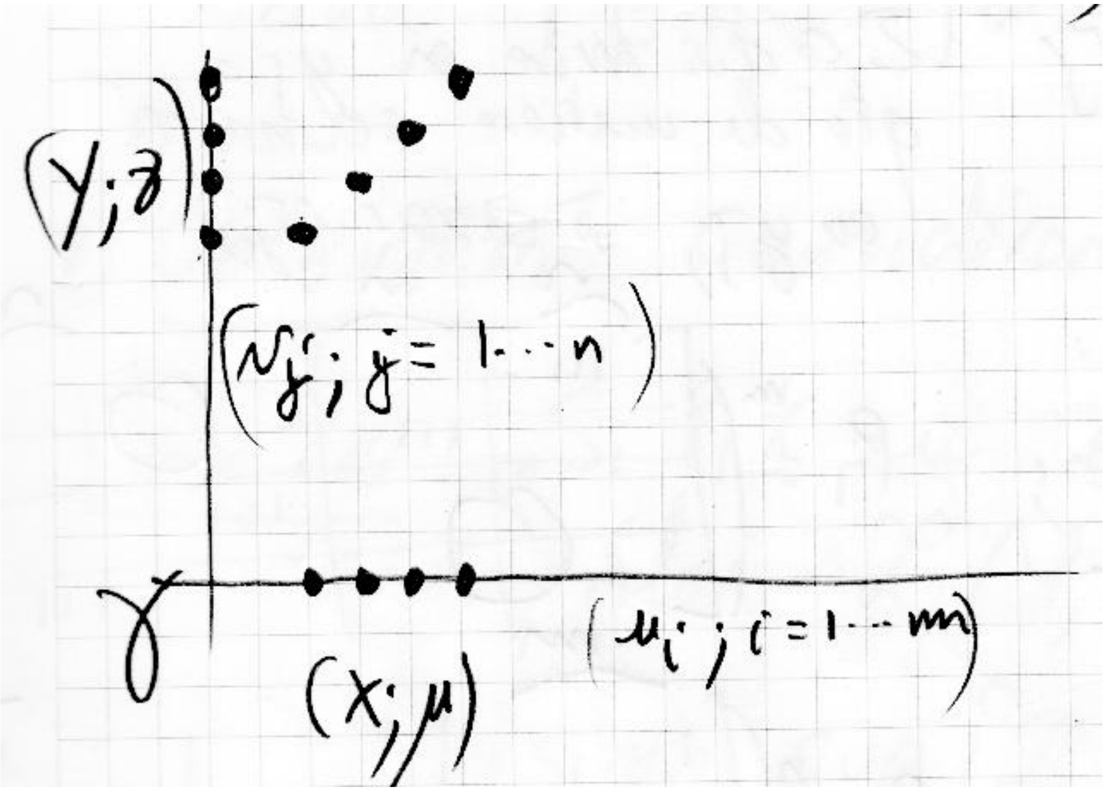

**(DMK):**

Min $<c, \gamma>$

$$s.t. \begin{cases} P_1 \gamma = u \\ P_2 \gamma = v \\ \gamma \geq 0 \end{cases}$$

$$\underset{\in \mathrm{IR}^{mn}}{\gamma} = \begin{bmatrix} \gamma_{11} \\ \gamma_{12} \\ \dots \\ \gamma_{1n} \\ \gamma_{22} \\ \dots \\ \gamma_{2n} \\ \dots \\ \dots \\ \gamma_{mn} \end{bmatrix}$$

# Part. 2 Optimal Transport – Duality




**(DMK):**

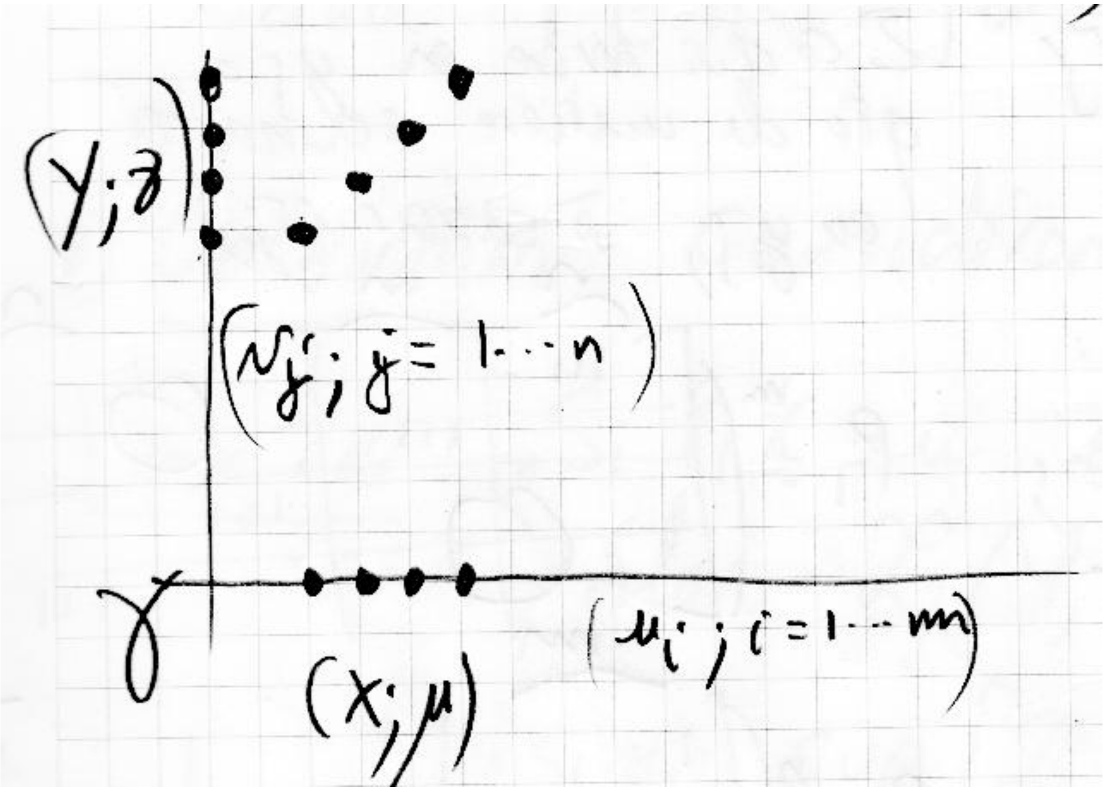
Min $<c, \gamma>$

$$s.t. \begin{cases} P_1 \gamma = u \\ P_2 \gamma = v \\ \gamma \geq 0 \end{cases}$$

$$c = \begin{bmatrix} c_{11} \\ c_{12} \\ \dots \\ c_{1n} \\ c_{22} \\ \dots \\ c_{2n} \\ \dots \\ \dots \\ c_{mn} \end{bmatrix} \qquad \gamma = \begin{bmatrix} \gamma_{11} \\ \gamma_{12} \\ \dots \\ \gamma_{1n} \\ \gamma_{22} \\ \dots \\ \gamma_{2n} \\ \dots \\ \dots \\ \gamma_{mn} \end{bmatrix}$$

$c \in IR^{mn}$ $\quad$ $\gamma \in IR^{mn}$

# **Part. 2** Optimal Transport – Duality



**(DMK):**

Min $\langle c, \gamma \rangle$

$$s.t. \begin{cases} P_1 \gamma = u \\ P_2 \gamma = v \\ \gamma \geq 0 \end{cases}$$

$$c = \begin{bmatrix} c_{11} \\ c_{12} \\ \dots \\ c_{1n} \\ c_{22} \\ \dots \\ c_{2n} \\ \dots \\ \dots \\ c_{mn} \end{bmatrix} \qquad \gamma = \begin{bmatrix} \gamma_{11} \\ \gamma_{12} \\ \dots \\ \gamma_{1n} \\ \gamma_{22} \\ \dots \\ \gamma_{2n} \\ \dots \\ \dots \\ \gamma_{mn} \end{bmatrix}$$

$c \in \mathbb{R}^{mn}$ $\qquad$ $\gamma \in \mathbb{R}^{mn}$

$c_{ij} = \| x_i - y_j \|^2$

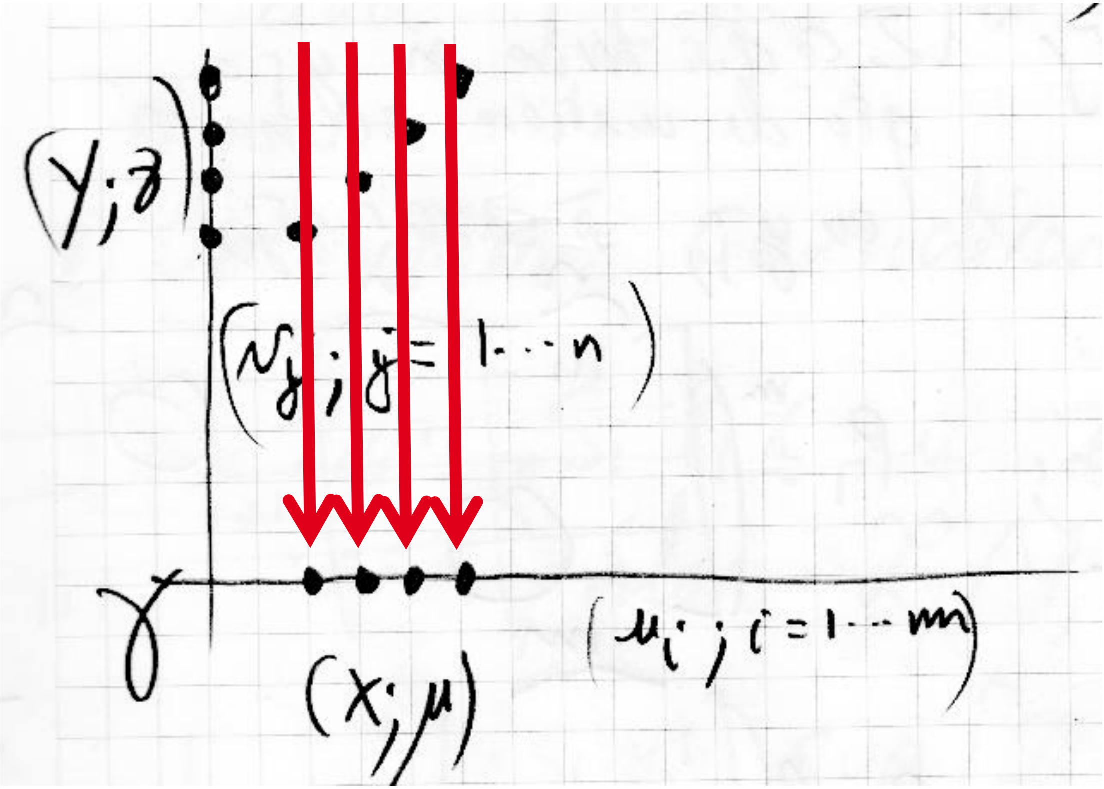
# Part. 2 Optimal Transport – Duality

(DMK):

Min $\langle c, \gamma \rangle$

s.t. $\begin{cases} P_1 \gamma = u \\ P_2 \gamma = v \\ \gamma \geq 0 \end{cases}$

mn X m → $P_1$




$c = \begin{bmatrix} c_{11} \\ c_{12} \\ \dots \\ c_{1n} \\ c_{22} \\ \dots \\ c_{2n} \\ \dots \\ \dots \\ c_{mn} \end{bmatrix}$ $\in \mathbb{R}^{mn}$

$\gamma = \begin{bmatrix} \gamma_{11} \\ \gamma_{12} \\ \dots \\ \gamma_{1n} \\ \gamma_{22} \\ \dots \\ \gamma_{2n} \\ \dots \\ \dots \\ \gamma_{mn} \end{bmatrix}$ $\in \mathbb{R}^{mn}$

$c_{ij} = || x_i - y_j ||^2$

# Part. 2 Optimal Transport – Duality

**(DMK):**

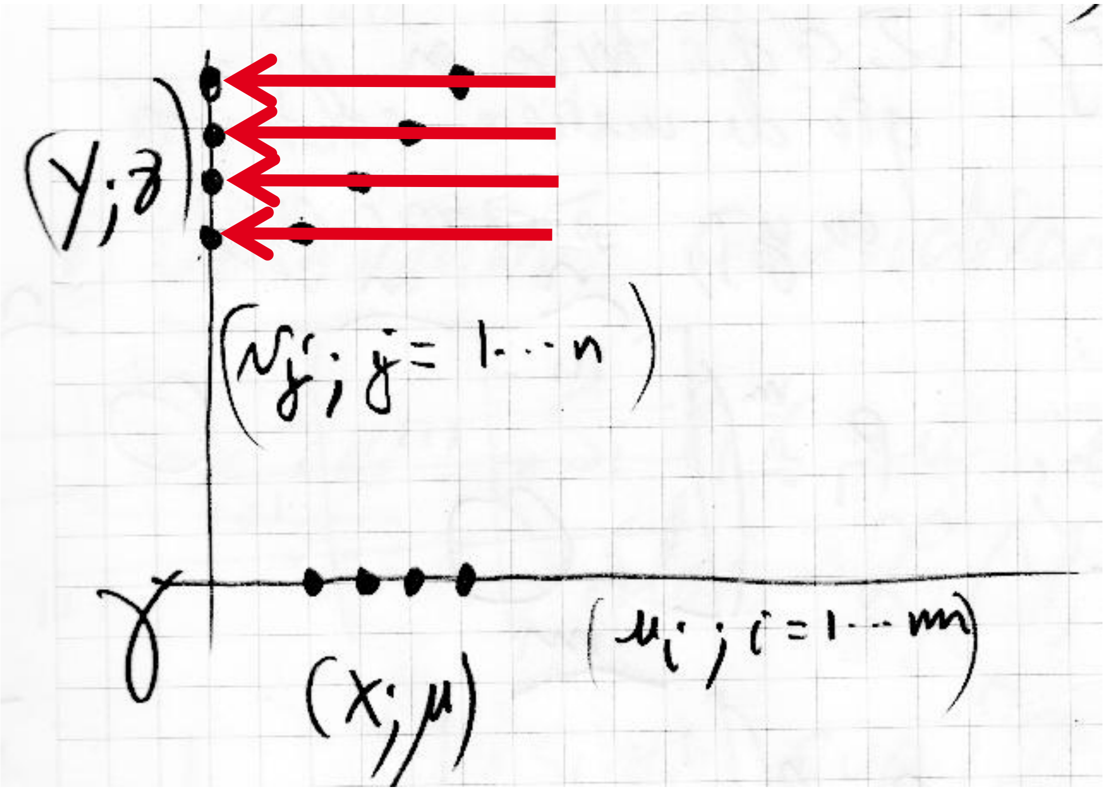
Min $<c, \gamma>$

s.t.

mn X m → $P_1 \gamma = u$

**mn X n** → $\mathbf{P_2} \gamma = v$

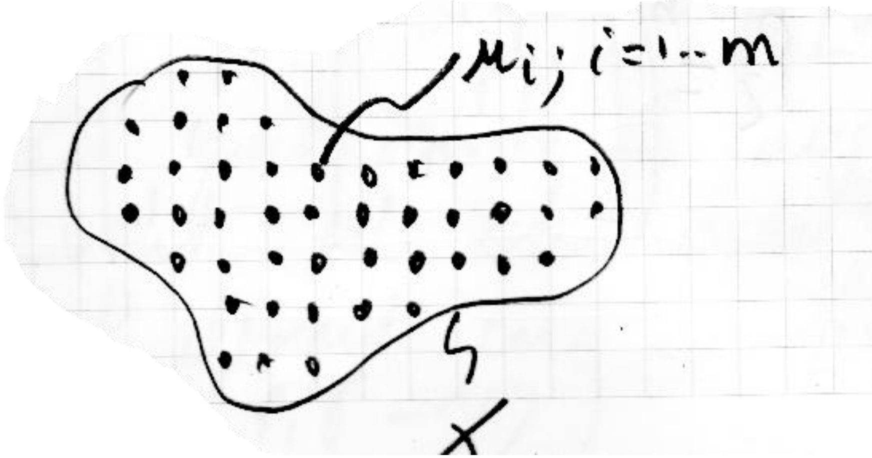
$\gamma \geq 0$

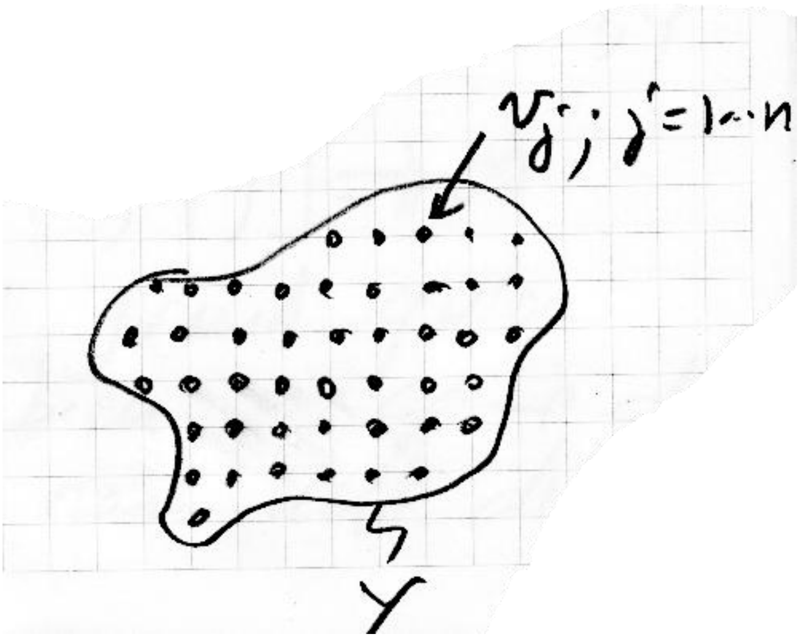



$$c = \begin{bmatrix} c_{11} \\ c_{12} \\ \dots \\ c_{1n} \\ c_{22} \\ \dots \\ c_{2n} \\ \dots \\ \dots \\ c_{mn} \end{bmatrix} \in \mathbb{R}^{mn} \qquad \gamma = \begin{bmatrix} \gamma_{11} \\ \gamma_{12} \\ \dots \\ \gamma_{1n} \\ \gamma_{22} \\ \dots \\ \gamma_{2n} \\ \dots \\ \dots \\ \gamma_{mn} \end{bmatrix} \in \mathbb{R}^{mn}$$

$c_{ij} = || x_i - y_j ||^2$

# **Part. 2** Optimal Transport – Duality




**(DMK):**

$$\text{Min} \; \langle c, \gamma \rangle$$

s.t.
- (mn X m) $P_1 \gamma = u$
- (mn X n) $P_2 \gamma = v$
- $\gamma \geq 0$

$$c = \begin{bmatrix} c_{11} \\ c_{12} \\ \dots \\ c_{1n} \\ c_{22} \\ \dots \\ c_{2n} \\ \dots \\ \dots \\ c_{mn} \end{bmatrix} \in \mathbb{R}^{mn} \qquad \gamma = \begin{bmatrix} \gamma_{11} \\ \gamma_{12} \\ \dots \\ \gamma_{1n} \\ \gamma_{22} \\ \dots \\ \gamma_{2n} \\ \dots \\ \dots \\ \gamma_{mn} \end{bmatrix} \in \mathbb{R}^{mn}$$

# **Part. 2** Optimal Transport – Duality

**(DMK):**

Min $<c, \gamma>$

$$
s.t. \begin{cases} P_1 \gamma = u \\ P_2 \gamma = v \\ \gamma \geq 0 \end{cases}
$$

*$< u, v >$ denotes the dot product between u and v*

Consider $L(\varphi,\psi) = <c, \gamma> - <\varphi, P_1 \gamma - u> - <\psi, P_2 \gamma - v>$

# **Part. 2** Optimal Transport – Duality

**(DMK):**

$$\text{Min } \langle c, \gamma \rangle \quad s.t. \begin{cases} P_1 \gamma = u \\ P_2 \gamma = v \\ \gamma \geq 0 \end{cases}$$

Consider $L(\varphi,\psi) = \langle c, \gamma \rangle - \langle \varphi, P_1 \gamma - u \rangle - \langle \psi, P_2 \gamma - v \rangle$

Remark: $\sup_{\substack{\varphi \in \mathbb{R}^m \\ \psi \in \mathbb{R}^n}} [\ L(\varphi,\psi)\ ] = \langle c, \gamma \rangle$ if $P_1 \gamma = u$ and $P_2 \gamma = v$

# Part. 2 Optimal Transport – Duality

**(DMK):**

$$\text{Min } \langle c, \gamma \rangle \quad s.t. \begin{cases} P_1 \gamma = u \\ P_2 \gamma = v \\ \gamma \geq 0 \end{cases}$$

Consider $L(\varphi,\psi) = \langle c, \gamma \rangle - \langle \varphi, P_1 \gamma - u \rangle - \langle \psi, P_2 \gamma - v \rangle$

Remark: $\sup_{\substack{\varphi \in \mathbb{R}^m \\ \psi \in \mathbb{R}^n}} [\, L(\varphi,\psi) \,] = \langle c, \gamma \rangle$ if $P_1 \gamma = u$ and $P_2 \gamma = v$

$= +\infty$ otherwise

# Part. 2 Optimal Transport – Duality

(DMK):

$$\text{Min } \langle c, \gamma \rangle \quad s.t. \begin{cases} P_1 \gamma = u \\ P_2 \gamma = v \\ \gamma \geq 0 \end{cases}$$

Consider $L(\varphi,\psi) = \langle c, \gamma\rangle - \langle \varphi, P_1 \gamma - u\rangle - \langle \psi, P_2 \gamma - v\rangle$

Remark: $\underset{\substack{\varphi \in \mathbb{R}^m \\ \psi \in \mathbb{R}^n}}{\text{Sup}}[\ L(\varphi,\psi)\ ] = \langle c, \gamma\rangle$ if $P_1 \gamma = u$ and $P_2 \gamma = v$

$= +\infty$ otherwise

Consider now: $\underset{\gamma \geq 0}{\text{Inf}} \left[ \underset{\substack{\varphi \in \mathbb{R}^m \\ \psi \in \mathbb{R}^n}}{\text{Sup}}[\ L(\varphi,\psi)\ ] \right]$

# Part. 2 Optimal Transport – Duality

**(DMK):**

Min $\langle c, \gamma \rangle$

$$s.t. \begin{cases} P_1 \gamma = u \\ P_2 \gamma = v \\ \gamma \geq 0 \end{cases}$$

Consider $L(\varphi,\psi) = \langle c, \gamma \rangle - \langle \varphi, P_1 \gamma - u \rangle - \langle \psi, P_2 \gamma - v \rangle$

Remark: $\sup_{\substack{\varphi \in \mathbb{R}^m \\ \psi \in \mathbb{R}^n}} [\ L(\varphi,\psi)\ ] = \langle c, \gamma \rangle$ if $P_1 \gamma = u$ and $P_2 \gamma = v$

$= +\infty$ otherwise

Consider now: $\inf_{\gamma \geq 0} \left[ \sup_{\substack{\varphi \in \mathbb{R}^m \\ \psi \in \mathbb{R}^n}} [\ L(\varphi,\psi)\ ] \right] = \inf_{\substack{\gamma \geq 0 \\ P_1 \gamma = u \\ P_2 \gamma = v}} [\ \langle c, \gamma \rangle\ ]$

# Part. 2 Optimal Transport – Duality

**(DMK):**

$$\text{Min } \langle c, \gamma \rangle \quad s.t. \begin{cases} P_1 \gamma = u \\ P_2 \gamma = v \\ \gamma \geq 0 \end{cases}$$

Consider $L(\varphi,\psi) = \langle c, \gamma \rangle - \langle \varphi, P_1 \gamma - u \rangle - \langle \psi, P_2 \gamma - v \rangle$

Remark: $\sup_{\varphi \in \mathbb{R}^m, \psi \in \mathbb{R}^n} [\, L(\varphi,\psi)\,] = \langle c, \gamma \rangle$ if $P_1 \gamma = u$ and $P_2 \gamma = v$

$= +\infty$ otherwise

Consider now: $\inf_{\gamma \geq 0} \left[ \sup_{\varphi \in \mathbb{R}^m, \psi \in \mathbb{R}^n} [\, L(\varphi,\psi)\,] \right] = \inf_{\gamma \geq 0,\ P_1 \gamma = u,\ P_2 \gamma = v} [\, \langle c, \gamma \rangle \,]$ **(DMK)**

# **Part. 2** Optimal Transport – Duality

**(DMK):**

$$\text{Min} \ \langle c, \gamma \rangle \quad s.t. \begin{cases} P_1 \gamma = u \\ P_2 \gamma = v \\ \gamma \geq 0 \end{cases}$$

$$\underset{\gamma \geq 0}{\text{Inf}} \left[ \underset{\substack{\varphi \in \mathbb{R}^m \\ \psi \in \mathbb{R}^n}}{\text{Sup}} [ \langle c, \gamma \rangle - \langle \varphi, P_1 \gamma - u \rangle - \langle \psi, P_2 \gamma - v \rangle ] \right]$$

# Part. 2 Optimal Transport – Duality

**(DMK):**

Min $\langle c, \gamma \rangle$

$$s.t. \begin{cases} P_1 \gamma = u \\ P_2 \gamma = v \\ \gamma \geq 0 \end{cases}$$

$$\inf_{\gamma \geq 0} \Big[ \sup_{\substack{\varphi \in IR^m \\ \psi \in IR^n}} [ \langle c, \gamma \rangle - \langle \varphi, P_1 \gamma - u \rangle - \langle \psi, P_2 \gamma - v \rangle ] \Big]$$

Exchange Inf Sup

$$\sup_{\substack{\varphi \in IR^m \\ \psi \in IR^n}} \Big[ \inf_{\gamma \geq 0} [ \langle c, \gamma \rangle - \langle \varphi, P_1 \gamma - u \rangle - \langle \psi, P_2 \gamma - v \rangle ] \Big]$$

# Part. 2 Optimal Transport – Duality

**(DMK):**

$$\text{Min} \; \langle c, \gamma \rangle \quad s.t. \begin{cases} P_1 \gamma = u \\ P_2 \gamma = v \\ \gamma \geq 0 \end{cases}$$

$$\inf_{\gamma \geq 0} \Big[ \sup_{\varphi \in \mathbb{R}^m, \, \psi \in \mathbb{R}^n} \big[ \langle c, \gamma \rangle - \langle \varphi, P_1 \gamma - u \rangle - \langle \psi, P_2 \gamma - v \rangle \big] \Big]$$

Exchange Inf Sup

$$\sup_{\varphi \in \mathbb{R}^m, \, \psi \in \mathbb{R}^n} \Big[ \inf_{\gamma \geq 0} \big[ \langle c, \gamma \rangle - \langle \varphi, P_1 \gamma - u \rangle - \langle \psi, P_2 \gamma - v \rangle \big] \Big]$$

Expand/Reorder/Collect

$$\sup_{\varphi \in \mathbb{R}^m, \, \psi \in \mathbb{R}^n} \Big[ \inf_{\gamma \geq 0} \big[ \langle \gamma, c - P_1^t \varphi - P_2^t \psi \rangle + \langle \varphi, u \rangle + \langle \psi, v \rangle \big] \Big]$$

# **Part. 2** Optimal Transport – Duality

**(DMK):**

$$\text{Min} \; \langle c, \gamma \rangle \quad s.t. \; \begin{cases} P_1 \gamma = u \\ P_2 \gamma = v \\ \gamma \geq 0 \end{cases}$$

$$\inf_{\gamma \geq 0} \Big[ \sup_{\varphi \in \mathbb{R}^m, \; \psi \in \mathbb{R}^n} [ \langle c, \gamma \rangle - \langle \varphi, P_1 \gamma - u \rangle - \langle \psi, P_2 \gamma - v \rangle ] \Big]$$

Exchange Inf Sup

$$\sup_{\varphi \in \mathbb{R}^m, \; \psi \in \mathbb{R}^n} \Big[ \inf_{\gamma \geq 0} [ \langle c, \gamma \rangle - \langle \varphi, P_1 \gamma - u \rangle - \langle \psi, P_2 \gamma - v \rangle ] \Big]$$

Expand/Reorder/Collect

$$\sup_{\varphi \in \mathbb{R}^m, \; \psi \in \mathbb{R}^n} \Big[ \inf_{\gamma \geq 0} [ \langle \gamma, c - P_1^t \varphi - P_2^t \psi \rangle + \langle \varphi, u \rangle + \langle \psi, v \rangle ] \Big]$$

Interpret

# Part. 2 Optimal Transport – Duality

**(DMK):**

$$\text{Min } \langle c, \gamma \rangle \quad s.t. \begin{cases} P_1 \gamma = u \\ P_2 \gamma = v \\ \gamma \geq 0 \end{cases}$$

$$\underset{\substack{\varphi \in \mathbb{R}^m \\ \psi \in \mathbb{R}^n}}{\text{Sup}}\Big[ \underset{\gamma \geq 0}{\text{Inf}}\big[ \langle \gamma, c - P_1^t \varphi - P_2^t \psi \rangle + \langle \varphi, u \rangle + \langle \psi, v \rangle \big] \Big]$$

Interpret

$$\underset{\substack{\varphi \in \mathbb{R}^m \\ \psi \in \mathbb{R}^n \\ P_1^t \varphi + P_2^t \psi \leq c}}{\text{Sup}}\Big[ \langle \varphi, u \rangle + \langle \psi, v \rangle \Big] \qquad \textbf{(DDMK)}$$

# Part. 2 Optimal Transport – Duality

**(DMK):**

$$\text{Min} \ \langle c, \gamma \rangle \quad s.t. \begin{cases} P_1 \gamma = u \\ P_2 \gamma = v \\ \gamma \geq 0 \end{cases}$$

$$\sup_{\substack{\varphi \in \mathbb{R}^m \\ \psi \in \mathbb{R}^n}} \Big[ \inf_{\gamma \geq 0} \big[ \langle \gamma, c - P_1^t \varphi - P_2^t \psi \rangle + \langle \varphi, u \rangle + \langle \psi, v \rangle \big] \Big]$$

Interpret

$$\sup_{\substack{\varphi \in \mathbb{R}^m \\ \psi \in \mathbb{R}^n \\ P_1^t \varphi + P_2^t \psi \leq c}} \Big[ \langle \varphi, u \rangle + \langle \psi, v \rangle \Big] \qquad \textbf{(DDMK)}$$

$\varphi_i + \psi_j \leq c_{ij} \quad \forall (i,j)$

# Part. 2 Optimal Transport – Kantorovich dual

***Kantorovich's problem:***

Find a measure $\gamma$ defined on $X \times Y$

such that $\int_{x \text{ in } X} d\gamma(x,y) = d\mu(x)$

and $\int_{y \text{ in } Y} d\gamma(x,y) = d\nu(x)$

that minimizes $\iint_{X \times Y} \| x - y \|^2 \, d\gamma(x,y)$

***Dual formulation of Kantorovich's problem (Continuous):***

Find two functions $\varphi$ in $L^1(\mu)$ and $\psi$ in $L^1(\nu)$
Such that for all x,y, $\varphi(x) + \psi(y) \leq \frac{1}{2}\| x - y \|^2$

that maximize $\int_X \varphi d\mu + \int_Y \psi d\nu$

# **Part. 2** Optimal Transport – Kantorovich dual

***Kantorovich's problem:***

Find a measure γ defined on X x Y

such that $\int_{x \text{ in } X} d\gamma(x,y) = d\mu(x)$

and $\int_{y \text{ in } Y} d\gamma(x,y) = d\nu(x)$

that minimizes $\iint_{X \times Y} \| x - y \|^2 \, d\gamma(x,y)$

Your point of view:
Try to minimize transport cost

***Dual formulation of Kantorovich's problem:***

Find two functions φ in $L^1(\mu)$ and ψ in $L^1(\nu)$
Such that for all x,y, $\varphi(x) + \psi(y) \leq \frac{1}{2}\| x - y \|^2$

that maximize $\int_X \varphi d\mu + \int_Y \psi d\nu$

# **Part. 2** Optimal Transport – Kantorovich dual

***Kantorovich's problem:***

Find a measure $\gamma$ defined on $X \times Y$

such that $\int_{x \text{ in } X} d\gamma(x,y) = d\mu(x)$

and $\int_{y \text{ in } Y} d\gamma(x,y) = d\nu(x)$

that minimizes $\iint_{X \times Y} \| x - y \|^2 \, d\gamma(x,y)$

Your point of view:
Try to minimize transport cost

***Dual formulation of Kantorovich's problem:***

Find two functions $\varphi$ in $L^1(\mu)$ and $\psi$ in $L^1(\nu)$
Such that for all x,y, $\varphi(x) + \psi(y) \leq \frac{1}{2}\| x - y \|^2$

that maximize $\int_X \varphi d\mu + \int_Y \psi d\nu$

Point of view of a "transport company":
Try to maximize transport price

# Part. 2 Optimal Transport – Kantorovich dual

***Kantorovich's problem:***

Find a measure $\gamma$ defined on X x Y

such that $\int_{x \in X} d\gamma(x,y) = d\mu(x)$

and $\int_{y \in Y} d\gamma(x,y) = d\nu(x)$

that minimizes $\iint_{X \times Y} \| x - y \|^2 \, d\gamma(x,y)$

Your point of view:
Try to minimize transport cost

***Dual formulation of Kantorovich's problem:***

Find two functions $\varphi$ in $L^1(\mu)$ and $\psi$ in $L^1(\nu)$
Such that for all x,y, $\varphi(x) + \psi(y) \leq \frac{1}{2}\| x - y \|^2$

that maximize $\int_X \varphi(x) d\mu + \int_Y \psi(y) d\nu$

What they charge for loading at x

# **Part. 2** Optimal Transport – Kantorovich dual

***Kantorovich's problem:***

Find a measure $\gamma$ defined on $X \times Y$
such that $\int_{x \text{ in } X} d\gamma(x,y) = d\mu(x)$
and $\int_{y \text{ in } Y} d\gamma(x,y) = d\nu(x)$

that minimizes $\iint_{X \times Y} \| x - y \|^2 \, d\gamma(x,y)$

Your point of view:
Try to minimize transport cost

***Dual formulation of Kantorovich's problem:***

Find two functions $\varphi$ in $L^1(\mu)$ and $\psi$ in $L^1(\nu)$
Such that for all x,y, $\varphi(x) + \psi(y) \leq \frac{1}{2}\| x - y \|^2$

that maximize $\int_X \varphi(x) d\mu + \int_Y \psi(y) d\nu$

What they charge for loading at x

What they charge for unloading at y

# **Part. 2** Optimal Transport – Kantorovich dual

***Kantorovich's problem:***

Find a measure γ defined on X x Y
such that $\int_{x \text{ in } X} d\gamma(x,y) = d\mu(x)$
and $\int_{y \text{ in } Y} d\gamma(x,y) = d\nu(x)$

that minimizes $\iint_{X \times Y} \| x - y \|^2 \, d\gamma(x,y)$

Your point of view:
Try to minimize transport cost

***Dual formulation of Kantorovich's problem:***

Find two functions φ in $L^1(\mu)$ and ψ in $L^1(\nu)$
Such that for all x,y, $\varphi(x) + \psi(y) \leq \frac{1}{2}\| x - y \|^2$

that maximize $\int_X \varphi(x) d\mu + \int_Y \psi(y) d\nu$

Price (loading + unloading) cannot be greater than transport cost (else you do the job yourself)

What they charge for loading at x

What they charge for unloading at y

# **Part. 2** Optimal Transport – c-conjugate functions

***Dual formulation of Kantorovich's problem:***

Find two functions $\varphi$ in $L^1(\mu)$ and $\psi$ in $L^1(\nu)$
Such that for all x,y, $\varphi(x) + \psi(y) \leq \frac{1}{2}\| x - y \|^2$

that maximize $\int_X \varphi(x) d\mu + \int_Y \psi(y) d\nu$

# **Part. 2** Optimal Transport – c-conjugate functions

***Dual formulation of Kantorovich's problem:***

Find two functions $\varphi$ in $L^1(\mu)$ and $\psi$ in $L^1(\nu)$
Such that for all x,y, $\varphi(x) + \psi(y) \leq \frac{1}{2}\| x - y \|^2$

that maximize $\int_X \varphi(x) d\mu + \int_Y \psi(y) d\nu$

If we got two functions $\varphi$ and $\psi$ that satisfy the constraint

Then it is possible to obtain a better solution by replacing $\psi$ with $\varphi^c$ defined by:
For all y, $\varphi^c(y) = \inf_{x \in X} \frac{1}{2}\| x - y \|^2 - \varphi(y)$

# Part. 2 Optimal Transport – c-conjugate functions

***Dual formulation of Kantorovich's problem:***

Find two functions $\varphi$ in $L^1(\mu)$ and $\psi$ in $L^1(\nu)$
Such that for all x,y, $\varphi(x) + \psi(y) \leq \frac{1}{2}\| x - y \|^2$

that maximize $\int_X \varphi(x)d\mu + \int_Y \psi(y)d\nu$

If we got two functions $\varphi$ and $\psi$ that satisfy the constraint

Then it is possible to obtain a better solution by replacing $\psi$ with $\varphi^c$ defined by:
For all y, $\varphi^c(y) = \inf_{x \text{ in } X} \frac{1}{2}\| x - y \|^2 - \varphi(y)$

- $\varphi^c$ is called the **c-conjugate** function of $\varphi$
- If there is a function $\varphi$ such that $\psi = \varphi^c$ then $\psi$ is said to be **c-concave**
- If $\psi$ is c-concave, then $\psi^{cc} = \psi$

# **Part. 2** Optimal Transport – c-conjugate functions

***Dual formulation of Kantorovich's problem:***

Find a c-concave function $\psi$

that maximizes $\int_X \psi(x) d\mu + \int_Y \psi^c(y) d\nu$

# **Part. 2** Optimal Transport – c-conjugate functions

***Dual formulation of Kantorovich's problem:***

Find a c-concave function $\psi$

that maximizes $\int_X \psi(x) d\mu + \int_Y \psi^c(y) d\nu$

$\psi$ is called a **"Kantorovich potential"**

# **Part. 2** Optimal Transport – c-subdifferential

***Dual formulation of Kantorovich's problem:***

Find a c-concave function $\psi$

that maximizes $\int_X \psi(x) d\mu + \int_Y \psi^c(y) d\nu$

$\psi$ is called a **"Kantorovich potential"**

What about our initial problem ?

# **Part. 2** Optimal Transport – c-subdifferential

***Dual formulation of Kantorovich's problem:***

Find a c-concave function $\psi$

that maximizes $\int_X \psi(x) d\mu + \int_Y \psi^c(y) d\nu$

$\psi$ is called a **"Kantorovich potential"**

What about our initial problem ? (i.e., this is T() that we want to find …)

# **Part. 2** Optimal Transport – c-subdifferential

**Theorem 1.**

$$\forall (x,y) \in \partial_c \psi, \nabla \psi(x) - \nabla_x c(x,y) = 0$$

*where* $\partial_c \psi = \{(x,y) | \phi(x) + \psi(y) = c(x,y)\}$ *denotes the so-called* c-subdifferential *of* $\psi$.

# **Part. 2** Optimal Transport – c-subdifferential

**Theorem 1.**

$$\forall (x,y) \in \partial_c \psi, \nabla \psi(x) - \nabla_x c(x,y) = 0$$

*where* $\partial_c \psi = \{(x,y) | \phi(x) + \psi(y) = c(x,y)\}$ *denotes the so-called* c-subdifferential *of* $\psi$.

**Proof**: see OTON, chap. 10.

# **Part. 2** Optimal Transport – c-subdifferential

**Theorem 1.**

$$\forall (x,y) \in \partial_c \psi, \nabla \psi(x) - \nabla_x c(x,y) = 0$$

*where* $\partial_c \psi = \{(x,y) | \phi(x) + \psi(y) = c(x,y)\}$ *denotes the so-called* c-subdifferential *of* $\psi$*.*

**Proof**: see OTON, chap. 10.

Heuristic argument (at the beginning of the same chapter):

# Part. 2 Optimal Transport – c-subdifferential

**Theorem 1.**

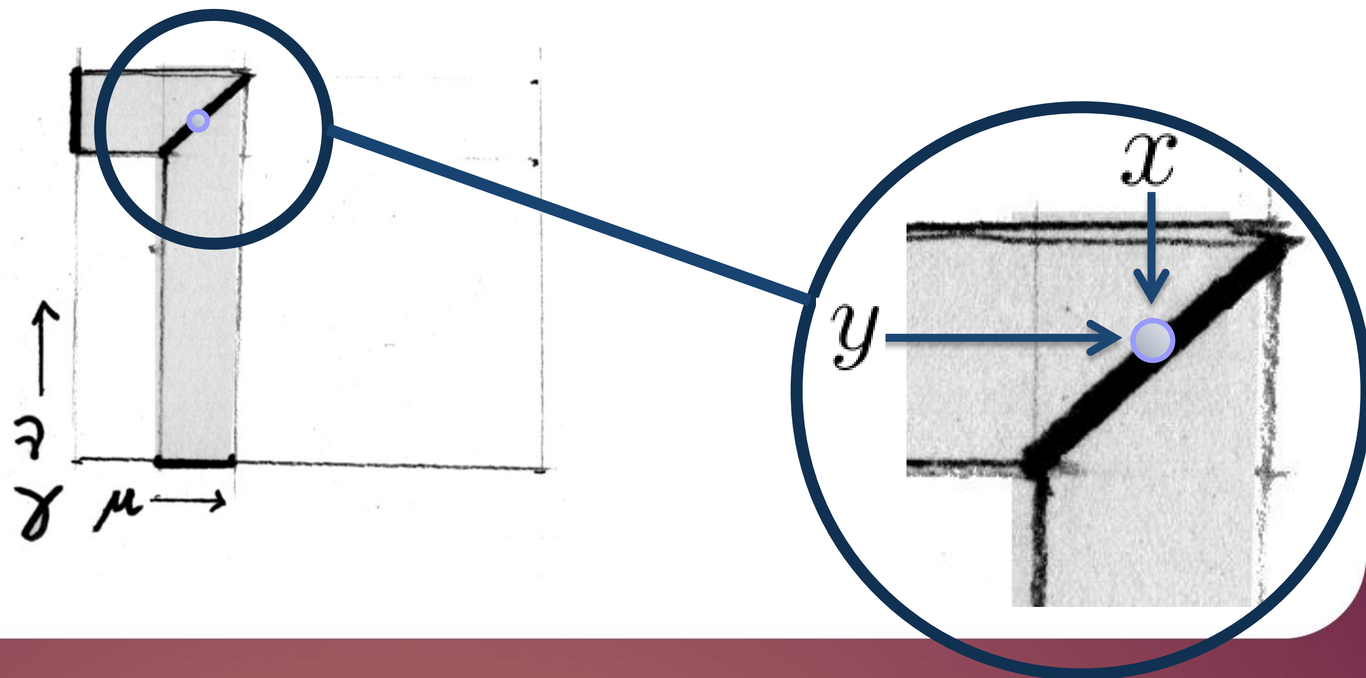
$$\forall (x,y) \in \partial_c \psi, \nabla \psi(x) - \nabla_x c(x,y) = 0$$

*where* $\partial_c \psi = \{(x,y) | \phi(x) + \psi(y) = c(x,y)\}$ *denotes the so-called* c-subdifferential *of* $\psi$.

**Proof**: see OTON, chap. 10.

Heuristic argument (at the beginning of the same chapter):

Consider a point $(x,y)$ on the c-subdifferential $\partial_c \psi$, that satisfies $\phi(y) + \psi(x) = c(x,y)$ (1).



# Part. 2 Optimal Transport – c-subdifferential

**Theorem 1.**

$$\forall (x,y) \in \partial_c \psi, \nabla \psi(x) - \nabla_x c(x,y) = 0$$

*where* $\partial_c \psi = \{(x,y) | \phi(x) + \psi(y) = c(x,y)\}$ *denotes the so-called* c-subdifferential *of* $\psi$.

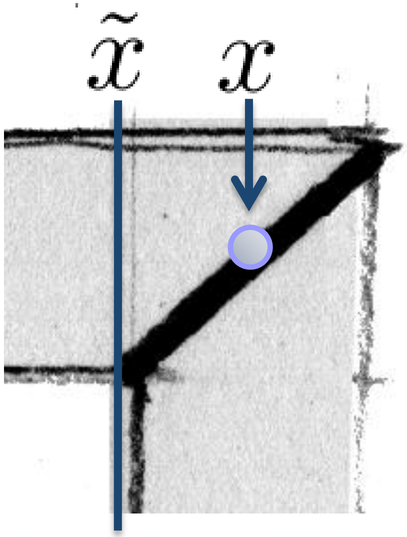
**Proof**: see OTON, chap. 10.

Heuristic argument (at the beginning of the same chapter):

Consider a point $(x,y)$ on the c-subdifferential $\partial_c \psi$, that satisfies $\phi(y) + \psi(x) = c(x,y)$ (1).

By definition, $\phi(y) = \psi^c(y) = \inf_x c(x,y) - \psi(x)$, thus $\forall \tilde{x}, \phi(y) \leq c(\tilde{x},y) - \psi(\tilde{x})$, or $\phi(y) + \psi(\tilde{x}) \leq c(\tilde{x},y)$ (2).

By substituting (1) into (2), one gets $\psi(\tilde{x}) - \psi(x) \leq c(\tilde{x},y) - c(x,y)$ for all $\tilde{x}$.



# Part. 2 Optimal Transport – c-subdifferential

**Theorem 1.**

$$\forall (x,y) \in \partial_c \psi, \nabla \psi(x) - \nabla_x c(x,y) = 0$$

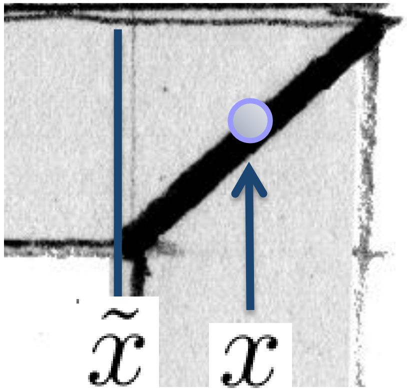
*where* $\partial_c \psi = \{(x,y) | \phi(x) + \psi(y) = c(x,y)\}$ *denotes the so-called* c-subdifferential *of* $\psi$.

**Proof**: see OTON, chap. 10.

Heuristic argument (at the beginning of the same chapter):

Consider a point $(x,y)$ on the c-subdifferential $\partial_c \psi$, that satisfies $\phi(y) + \psi(x) = c(x,y)$ (1).
By definition, $\phi(y) = \psi^c(y) = \inf_x c(x,y) - \psi(x)$, thus $\forall \tilde{x}, \phi(y) \leq c(\tilde{x},y) - \psi(\tilde{x})$, or $\phi(y) + \psi(\tilde{x}) \leq c(\tilde{x},y)$ (2).
By substituting (1) into (2), one gets $\psi(\tilde{x}) - \psi(x) \leq c(\tilde{x},y) - c(x,y)$ for all $\tilde{x}$.

Imagine now that $\tilde{x}$ follows a trajectory parameterized by $\epsilon$ and starting at $x$. One can compute the gradient along an arbitrary direction $w$ by taking the limit when $\epsilon$ tends to zero in the relation $\frac{\psi(\tilde{x}) - \psi(x)}{\epsilon} \leq \frac{c(\tilde{x},y) - c(x,y)}{\epsilon}$.



# Part. 2 Optimal Transport – c-subdifferential

**Theorem 1.**

$$\forall (x,y) \in \partial_c \psi, \nabla \psi(x) - \nabla_x c(x,y) = 0$$

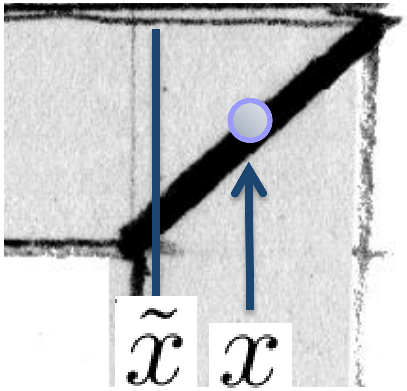
*where* $\partial_c \psi = \{(x,y) | \phi(x) + \psi(y) = c(x,y)\}$ *denotes the so-called* c-subdifferential *of* $\psi$.

**Proof**: see OTON, chap. 10.

Heuristic argument (at the beginning of the same chapter):

Consider a point $(x,y)$ on the c-subdifferential $\partial_c \psi$, that satisfies $\phi(y) + \psi(x) = c(x,y)$ (1).
By definition, $\phi(y) = \psi^c(y) = \inf_x c(x,y) - \psi(x)$, thus $\forall \tilde{x}, \phi(y) \le c(\tilde{x},y) - \psi(\tilde{x})$, or $\phi(y) + \psi(\tilde{x}) \le c(\tilde{x},y)$ (2).
By substituting (1) into (2), one gets $\psi(\tilde{x}) - \psi(x) \le c(\tilde{x},y) - c(x,y)$ for all $\tilde{x}$.

Imagine now that $\tilde{x}$ follows a trajectory parameterized by $\epsilon$ and starting at $x$. One can compute the gradient along an arbitrary direction $w$ by taking the limit when $\epsilon$ tends to zero in the relation $\frac{\psi(\tilde{x}) - \psi(x)}{\epsilon} \le \frac{c(\tilde{x},y) - c(x,y)}{\epsilon}$.



# Part. 2 Optimal Transport – c-subdifferential

**Theorem 1.**

$$\forall (x,y) \in \partial_c \psi, \nabla \psi(x) - \nabla_x c(x,y) = 0$$

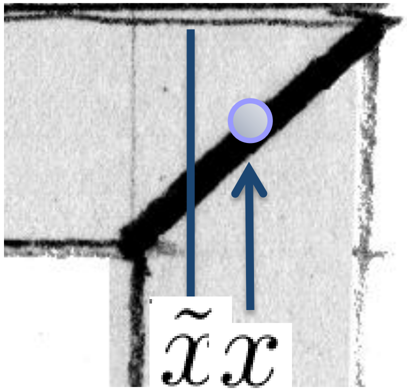
*where* $\partial_c \psi = \{(x,y) | \phi(x) + \psi(y) = c(x,y)\}$ *denotes the so-called* c-subdifferential *of* $\psi$.

**Proof**: see OTON, chap. 10.

Heuristic argument (at the beginning of the same chapter):

Consider a point $(x,y)$ on the c-subdifferential $\partial_c \psi$, that satisfies $\phi(y) + \psi(x) = c(x,y)$ (1).
By definition, $\phi(y) = \psi^c(y) = \inf_x c(x,y) - \psi(x)$, thus $\forall \tilde{x}, \phi(y) \le c(\tilde{x},y) - \psi(\tilde{x})$, or $\phi(y) + \psi(\tilde{x}) \le c(\tilde{x},y)$ (2).
By substituting (1) into (2), one gets $\psi(\tilde{x}) - \psi(x) \le c(\tilde{x},y) - c(x,y)$ for all $\tilde{x}$.

Imagine now that $\tilde{x}$ follows a trajectory parameterized by $\epsilon$ and starting at $x$. One can compute the gradient along an arbitrary direction $w$ by taking the limit when $\epsilon$ tends to zero in the relation $\frac{\psi(\tilde{x}) - \psi(x)}{\epsilon} \le \frac{c(\tilde{x},y) - c(x,y)}{\epsilon}$.



# **Part. 2** Optimal Transport – c-subdifferential

**Theorem 1.**

$$\forall (x,y) \in \partial_c \psi, \nabla \psi(x) - \nabla_x c(x,y) = 0$$

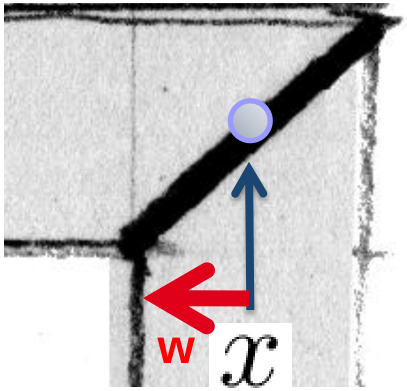
*where* $\partial_c \psi = \{(x,y) | \phi(x) + \psi(y) = c(x,y)\}$ *denotes the so-called* c-subdifferential *of* $\psi$.

**Proof**: see OTON, chap. 10.

Heuristic argument (at the beginning of the same chapter):

Consider a point $(x,y)$ on the c-subdifferential $\partial_c \psi$, that satisfies $\phi(y) + \psi(x) = c(x,y)$ (1).
By definition, $\phi(y) = \psi^c(y) = \inf_x c(x,y) - \psi(x)$, thus $\forall \tilde{x}, \phi(y) \leq c(\tilde{x},y) - \psi(\tilde{x})$, or $\phi(y) + \psi(\tilde{x}) \leq c(\tilde{x},y)$ (2).
By substituting (1) into (2), one gets $\psi(\tilde{x}) - \psi(x) \leq c(\tilde{x},y) - c(x,y)$ for all $\tilde{x}$.

Imagine now that $\tilde{x}$ follows a trajectory parameterized by $\epsilon$ and starting at $x$. One can compute the gradient along an arbitrary direction $w$ by taking the limit when $\epsilon$ tends to zero in the relation $\frac{\psi(\tilde{x}) - \psi(x)}{\epsilon} \leq \frac{c(\tilde{x},y) - c(x,y)}{\epsilon}$.
Thus we have $\nabla \psi(x) \cdot w \leq \nabla_x c(x,y) \cdot w$



# Part. 2 Optimal Transport – c-subdifferential

**Theorem 1.**

$$\forall (x,y) \in \partial_c \psi, \nabla \psi(x) - \nabla_x c(x,y) = 0$$

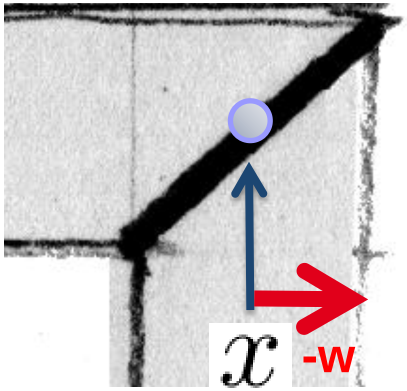
*where* $\partial_c \psi = \{(x,y) | \phi(x) + \psi(y) = c(x,y)\}$ *denotes the so-called* c-subdifferential *of* $\psi$.

**Proof**: see OTON, chap. 10.

Heuristic argument (at the beginning of the same chapter):

Consider a point $(x,y)$ on the c-subdifferential $\partial_c \psi$, that satisfies $\phi(y) + \psi(x) = c(x,y)$ (1).
By definition, $\phi(y) = \psi^c(y) = \inf_x c(x,y) - \psi(x)$, thus $\forall \tilde{x}, \phi(y) \le c(\tilde{x},y) - \psi(\tilde{x})$, or $\phi(y) + \psi(\tilde{x}) \le c(\tilde{x},y)$ (2).
By substituting (1) into (2), one gets $\psi(\tilde{x}) - \psi(x) \le c(\tilde{x},y) - c(x,y)$ for all $\tilde{x}$.

Imagine now that $\tilde{x}$ follows a trajectory parameterized by $\epsilon$ and starting at $x$. One can compute the gradient along an arbitrary direction $w$ by taking the limit when $\epsilon$ tends to zero in the relation $\frac{\psi(\tilde{x}) - \psi(x)}{\epsilon} \le \frac{c(\tilde{x},y) - c(x,y)}{\epsilon}$.
Thus we have $\nabla \psi(x) \cdot w \le \nabla_x c(x,y) \cdot w$
The same derivation can be done with $-w$ instead of $w$, and one gets:
$\forall w, \nabla \psi(x) \cdot w = \nabla_x c(x,y) \cdot w$, thus $\forall (x,y) \in \partial_c \psi, \nabla \psi(x) - \nabla_x c(x,y) = 0$.



# **Part. 2** Optimal Transport – c-subdifferential

***Dual formulation of Kantorovich's problem:***

Find a c-concave function $\psi$

that maximizes $\int_X \psi(x)d\mu + \int_Y \psi^c(y)d\nu$

In the $L_2$ case, i.e. $c(x,y) = 1/2\|x-y\|^2$, we have $\forall(x,y) \in \partial_c\psi, \nabla\psi(x) + y - x = 0$, thus, whenever the optimal transport map $T$ exists, we have $T(x) = x - \nabla\psi(x) = \nabla(\|x\|^2/2 - \psi(x))$.

# **Part. 2** Optimal Transport – c-subdifferential

***Dual formulation of Kantorovich's problem:***

Find a c-concave function $\psi$

that maximizes $\int_X \psi(x)d\mu + \int_Y \psi^c(y)d\nu$

In the $L_2$ case, i.e. $c(x,y) = 1/2\|x-y\|^2$, we have $\forall (x,y) \in \partial_c \psi, \nabla\psi(x) + y - x = 0$, thus, whenever the optimal transport map $T$ exists, we have $T(x) = x - \nabla\psi(x) = \nabla(\|x\|^2/2 - \psi(x))$.

grad $\overline{\psi}(x)$ with $\overline{\psi}(x) := (\frac{1}{2} x^2 - \psi(x))$

# **Part. 2** Optimal Transport – convexity

***Dual formulation of Kantorovich's problem:***

Find a c-concave function $\psi$

that maximizes $\int_X \psi(x)d\mu + \int_Y \psi^c(y)d\nu$

In the $L_2$ case, i.e. $c(x,y) = 1/2\|x-y\|^2$, we have $\forall(x,y) \in \partial_c\psi, \nabla\psi(x) + y - x = 0$, thus, whenever the optimal transport map $T$ exists, we have $T(x) = x - \nabla\psi(x) = \nabla(\|x\|^2/2 - \psi(x))$.

grad $\overline{\psi}(x)$ with $\overline{\psi}(x) := (½\, x^2 - \psi(x))$

$\overline{\psi}$ is convex

# Part. 2 Optimal Transport – convexity

***Dual formulation of Kantorovich's problem:***

Find a c-concave function ψ

that maximizes $\int_X \psi(x)d\mu + \int_Y \psi^c(y)d\nu$

In the $L_2$ case, i.e. $c(x,y) = 1/2\|x-y\|^2$, we have $\forall (x,y) \in \partial_c\psi, \nabla\psi(x) + y - x = 0$, thus, whenever the optimal transport map $T$ exists, we have $T(x) = x - \nabla\psi(x) = \nabla(\|x\|^2/2 - \psi(x))$.

grad $\overline{\psi}(x)$ with $\overline{\psi}(x) := (½\, x^2 - \psi(x))$
$\overline{\psi}$ is convex

*Proof.*

$$
\begin{aligned}
\psi(x) &= \inf_y \frac{|x-y|^2}{2} - \phi(y) \\
&= \inf_y \frac{\|x\|^2}{2} - x\cdot y + \frac{\|y\|^2}{2} - \phi(y) \\
-\bar{\psi}(x) &= \phi(x) - \frac{\|x\|^2}{2} = \inf_y -x\cdot y + \left(\frac{\|y\|^2}{2} - \phi(y)\right) \\
\bar{\psi}(x) &= \sup_y x\cdot y - \left(\frac{\|y\|^2}{2} - \phi(y)\right)
\end{aligned}
$$

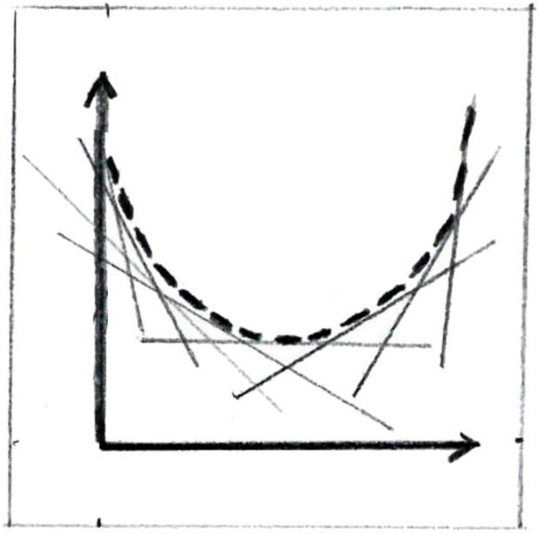
# **Part. 2** Optimal Transport – convexity

***Dual formulation of Kantorovich's problem:***

Find a c-concave function ψ

that maximizes $\int_X \psi(x)d\mu + \int_Y \psi^c(y)d\nu$

In the $L_2$ case, i.e. $c(x,y) = 1/2\|x-y\|^2$, we have $\forall(x,y) \in \partial_c\psi, \nabla\psi(x) + y - x = 0$, thus, whenever the optimal transport map $T$ exists, we have $T(x) = x - \nabla\psi(x) = \nabla(\|x\|^2/2 - \psi(x))$.



grad $\bar{\psi}(x)$ with $\bar{\psi}(x)$ := (½ $x^2$- ψ(x))
$\bar{\psi}$ is convex

$$\begin{aligned}
\psi(x) &= \inf_y \frac{|x-y|^2}{2} - \phi(y) \\
&= \inf_y \frac{\|x\|^2}{2} - x\cdot y + \frac{\|y\|^2}{2} - \phi(y) \\
-\bar{\psi}(x) &= \phi(x) - \frac{\|x\|^2}{2} = \inf_y -x\cdot y + \left(\frac{\|y\|^2}{2} - \phi(y)\right) \\
\bar{\psi}(x) &= \sup_y x\cdot y - \left(\frac{\|y\|^2}{2} - \phi(y)\right)
\end{aligned}$$

# Part. 2 Optimal Transport – no collision

***Dual formulation of Kantorovich's problem:***

Find a c-concave function $\psi$

that maximizes $\int_X \psi(x) d\mu + \int_Y \psi^c(y) d\nu$

If T(.) exists, then

$T(x) = x - \text{grad}\, \psi(x) = \text{grad}\, (\frac{1}{2} x^2 - \psi(x))$

$\underbrace{\frac{1}{2} x^2 - \psi(x)}_{\text{grad}\, \overline{\psi}(x)}$

$\overline{\psi}$ is convex

↓

***Two transported particles cannot "collide"***

# Part. 2 Optimal Transport – no collision

***Dual formulation of Kantorovich's problem:***

Find a c-concave function $\psi$

that maximizes $\int_X \psi(x) d\mu + \int_Y \psi^c(y) d\nu$

If T(.) exists, then

$$T(x) = x - \text{grad}\, \psi(x) = \text{grad}\, \underbrace{(\tfrac{1}{2} x^2 - \psi(x))}_{\text{grad}\, \overline{\psi}(x)}$$

$\overline{\psi}$ is convex

↓

***Two transported particles cannot "collide"***

*Proof.* By contradiction, suppose that you have $t \in (0,1)$ and $x_1 \neq x_2$ such that:

$$(1-t)x_1 + tT(x_1) = (1-t)x_2 + tT(x_2)$$

# Part. 2 Optimal Transport – no collision

***Dual formulation of Kantorovich's problem:***

Find a c-concave function $\psi$

that maximizes $\int_X \psi(x)d\mu + \int_Y \psi^c(y)d\nu$

If T(.) exists, then

$T(x) = x - \text{grad}\,\psi(x) = \text{grad}\,(\underbrace{\tfrac{1}{2}x^2 - \psi(x)}_{\text{grad}\,\bar{\psi}(x)})$

$\bar{\psi}$ is convex

↓

***Two transported particles cannot "collide"***

*Proof.* By contradiction, suppose that you have $t \in (0,1)$ and $x_1 \neq x_2$ such that:

$$(1-t)x_1 + tT(x_1) = (1-t)x_2 + tT(x_2)$$

$$\begin{aligned}
(1-t)x_1 + t\nabla\bar{\psi}(x_1) &= (1-t)x_2 + t\nabla\bar{\psi}(x_2) \\
(1-t)(x_1 - x_2) + t(\nabla\bar{\psi}(x_1) - \nabla\bar{\psi}(x_2)) &= 0 \\
\forall v, (1-t)v \cdot (x_1 - x_2) + tv \cdot (\nabla\bar{\psi}(x_1) - \nabla\bar{\psi}(x_2)) &= 0 \\
\text{take } v = (x_1 - x_2) & \\
(1-t)\|x_1 - x_2\|^2 + t(x_1 - x_2) \cdot (\nabla\bar{\psi}(x_1) - \nabla\bar{\psi}(x_2)) &= 0
\end{aligned}$$

# **Part. 2** Optimal Transport – Monge-Ampere

***Dual formulation of Kantorovich's problem:***

Find a c-concave function $\psi$

that maximizes $\int_X \psi(x) d\mu + \int_Y \psi^c(y) d\nu$

What about our initial problem ? If T(.) exists, then one can show that:

$T(x) = x - \text{grad}\, \psi(x) = \text{grad}\, (\frac{1}{2} x^2 - \psi(x))$

$\text{grad}\, \overline{\psi}(x)$ with $\overline{\psi}(x) := (\frac{1}{2} x^2 - \psi(x))$

for all borel set A, $\int_A d\mu = \int_{T(A)} (|JT|)\, d\nu$ (change of variable)

Jacobian of T (1st order derivatives)

# **Part. 2** Optimal Transport – Monge-Ampere

***Dual formulation of Kantorovich's problem:***

Find a c-concave function $\psi$

that maximizes $\int_X \psi(x) d\mu + \int_Y \psi^c(y) d\nu$

What about our initial problem ? If T(.) exists, then one can show that:

$T(x) = x - \text{grad}\, \psi(x) = \text{grad}\, (\frac{1}{2} x^2 - \psi(x))$

$\text{grad}\, \overline{\psi}(x)$ with $\overline{\psi}(x) := (\frac{1}{2} x^2 - \psi(x))$

for all borel set A, $\int_A d\mu = \int_{T(A)} (|JT|)\, d\nu = \int_{T(A)} (H\, \overline{\psi}\,)\, d\nu$

Det. of the Hessian of $\overline{\psi}$ ($2^{nd}$ order derivatives)

# **Part. 2** Optimal Transport – Monge-Ampere

***Dual formulation of Kantorovich's problem:***

Find a c-concave function $\psi$

that maximizes $\int_X \psi(x) d\mu + \int_Y \psi^c(y) d\nu$

What about our initial problem ?

$T(x) = x - \text{grad}\, \psi(x) = \underbrace{\text{grad}\, (\frac{1}{2} x^2 - \psi(x))}$

$\text{grad}\, \Phi(x)$ with $\Phi(x) := (\frac{1}{2} x^2 - \psi(x))$

for all borel set A, $\int_A d\mu = \int_{T(A)} (|JT|)\, d\nu = \int_{T(A)} (H\, \bar{\psi})\, d\nu$

When $\mu$ and $\nu$ have a density u and v, $(H\, \bar{\psi}(x)).\ v(\text{grad}\, \bar{\psi}(x)) = u(x)$ *Monge-Ampère equation*

# **Part. 2** Optimal Transport – summary

Find a transport map T that minimizes $C(T) = \int_X || x - T(x) ||^2 \, d\mu(x)$

# **Part. 2** Optimal Transport – summary

Find a transport map T that minimizes $C(T) = \int_X \| x - T(x) \|^2 \, d\mu(x)$

*After several rewrites and under some conditions….*
*(Kantorovich formulation, dual, c-convex functions)*

# **Part. 2** Optimal Transport – summary

Find a transport map T that minimizes $C(T) = \int_X \| x - T(x) \|^2 d\mu(x)$

*After several rewrites and under some conditions….*
*(Kantorovich formulation, dual, c-convex functions)*

Solve $(H\,\overline{\psi}(x)).\, v\big(\mathrm{grad}\,\overline{\psi}(x)\big) = u(x)$ *Monge-Ampère equation*
(When μ and ν have a density u and v resp.)

# Part. 2 Optimal Transport – summary

Find a transport map T that minimizes $C(T) = \int_X \| x - T(x) \|^2 \, d\mu(x)$

*After several rewrites and under some conditions….*
*(Kantorovich formulation, dual, c-convex functions)*

Solve $(H\, \bar{\psi}(x)).\, v\big(\mathrm{grad}\, \bar{\psi}(x)\big) = u(x)$ *Monge-Ampère equation*
(When $\mu$ and $\nu$ have a density u and v resp.)

**Brenier, Mc Cann, Trudinger**: ***The optimal transport map is then given by:***
$T(x) = \mathrm{grad}\, \bar{\psi}(x)$

# **Part. 2** Optimal Transport – Isoperimetric inequality


**For a given volume,**
**ball is the shape that minimizes border area**

# **Part. 2** Optimal Transport – Isoperimetric inequality

**$L_1$ Sobolev inegality:** Given $f: \mathbb{R}^n \rightarrow \mathbb{R}$ sufficiently regular

$$\int |\, \mathrm{grad}\ f\,| \geq n\ \mathrm{Vol}(B_2^n)^{1/n} \left(\int f^{\,n/(n-1)}\right)^{(n-1)/n}$$

Explanation in **[Dario Cordero Erauquin]** course notes

# Part. 2 Optimal Transport – Isoperimetric inequality

**$L_1$ Sobolev inegality:** Given f: $\mathbb{R}^n \rightarrow \mathbb{R}$ sufficiently regular

$$\int |\,\mathrm{grad}\, f\,| \;\geq\; n\, \mathrm{Vol}(B_2^n)^{1/n} \left( \int f^{\,n/(n-1)} \right)^{(n-1)/n}$$

Explanation in **[Dario Cordero Erauquin]** course notes

# **Part. 2** Optimal Transport – Isoperimetric inequality

**$L_1$ Sobolev inegality:** Given f: $\mathbb{R}^n \rightarrow \mathbb{R}$ sufficiently regular

$$\int |\text{grad } f| \geq n \, \text{Vol}(B_2^n)^{1/n} \left( \int f^{n/(n-1)} \right)^{(n-1)/n}$$

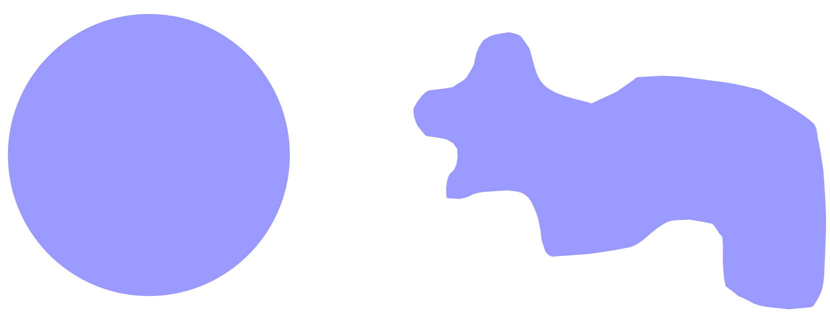
Explanation in **[Dario Cordero Erauquin]** course notes

# Part. 2 Optimal Transport – Isoperimetric inequality

**$L_1$ Sobolev inegality:** Given f: $\mathbb{R}^n \rightarrow \mathbb{R}$ sufficiently regular

Consider a compact set Ω such that $\mathrm{Vol}(\Omega) = \mathrm{Vol}(B_2^3)$
and f = the indicatrix function of Ω

$$\int | \mathrm{grad}\ f | \geq n\ \mathrm{Vol}(B_2^n)^{1/n} \left( \int f^{\,n/(n-1)} \right)^{(n-1)/n}$$



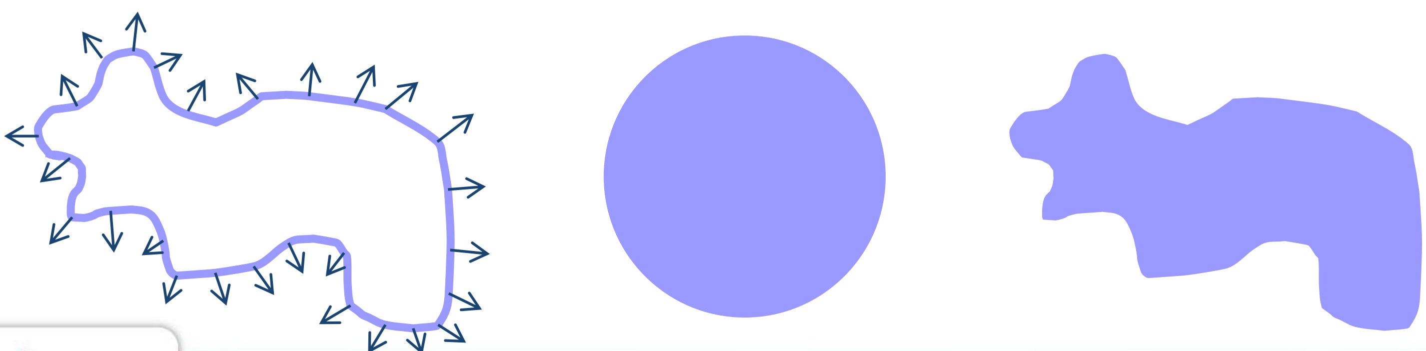
# Part. 2 Optimal Transport – Isoperimetric inequality

**$L_1$ Sobolev inegality:** Given f: $\mathbb{R}^n \rightarrow \mathbb{R}$ sufficiently regular

Consider a compact set Ω such that $Vol(\Omega) = Vol(B_2^3)$
and f = the indicatrix function of Ω

$$\int |\text{grad } f| \geq n \, Vol(B_2^n)^{1/n} \left( \int f^{n/(n-1)} \right)^{(n-1)/n}$$

$$Vol(\partial\Omega) \geq n \, Vol(B_2^3)^{1/3} \, Vol(B_2^3)^{2/3}$$



# **Part. 2** Optimal Transport – Isoperimetric inequality

**$L_1$ Sobolev inegality:** Given f: $\mathbb{R}^n \rightarrow \mathbb{R}$ sufficiently regular

Consider a compact set Ω such that $Vol(\Omega) = Vol(B_2^3)$
and f = the indicatrix function of Ω

$$\int |\text{grad } f| \geq n \, Vol(B_2^n)^{1/n} \left( \int f^{n/(n-1)} \right)^{(n-1)/n}$$

$$\downarrow \qquad \downarrow \qquad \downarrow$$

$$Vol(\partial\Omega) \geq n \, Vol(B_2^3)^{1/3} \, Vol(B_2^3)^{2/3}$$

$$Vol(\partial\Omega) \geq 4\pi = Vol(\partial B_2^3)$$

# **Part. 2** Optimal Transport – Isoperimetric inequality

**$L_1$ Sobolev inegality:** a proof with OT [Gromov]

$$\int | \text{grad } f | \geq n \, \text{Vol}(B_2^n)^{1/n} \left( \int f^{\,n/(n-1)} \right)^{(n-1)/n}$$

# **Part. 2** Optimal Transport – Isoperimetric inequality

**$L_1$ Sobolev inegality:** a proof with OT [Gromov]

$$\int |\text{grad } f| \geq n \, \text{Vol}(B_2^n)^{1/n} \left( \int f^{n/(n-1)} \right)^{(n-1)/n}$$

We suppose w.l.o.g. that $\int f^{n/(n-1)} = 1$

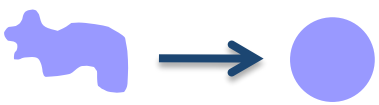
# Part. 2 Optimal Transport – Isoperimetric inequality

**$L_1$ Sobolev inegality:** a proof with OT [Gromov]

$$\int |\,\mathrm{grad}\, f\,| \;\geq n\, \mathrm{Vol}(B_2^n)^{1/n} \left(\int f^{\,n/(n-1)}\right)^{(n-1)/n}$$

We suppose w.l.o.g. that $\int f^{\,n/(n-1)} = 1$

There exists an optimal transport $T = \mathrm{grad}\overline{\Psi}$ between $f^{\,n/(n-1)}(x)dx$ and $\mathbf{1}B_2^n/\mathrm{Vol}(B_2^n)dx$



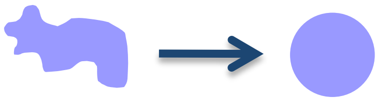
# Part. 2 Optimal Transport – Isoperimetric inequality

**$L_1$ Sobolev inegality:** a proof with OT [Gromov]

$$\int |\text{grad } f| \geq n \,\text{Vol}(B_2^n)^{1/n} \left(\int f^{n/(n-1)}\right)^{(n-1)/n}$$

We suppose w.l.o.g. that $\int f^{n/(n-1)} = 1$

There exists an optimal transport $T = \text{grad}\overline{\Psi}$ between $f^{n/(n-1)}(x)dx$ and $\mathbf{1}B_2^n/\text{Vol}(B_2^n)dx$



Monge-Ampère equation: $\text{Vol}(B_2^n)\, f^{n/(n-1)}(x) = \det \text{Hess}\, \overline{\Psi}$

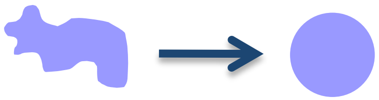
# **Part. 2** Optimal Transport – Isoperimetric inequality

**$L_1$ Sobolev inegality:** a proof with OT [Gromov]

$$\int |\,\mathrm{grad}\ f\,| \geq n\ \mathrm{Vol}(B_2^n)^{1/n} \left(\int f^{\ n/(n-1)}\right)^{(n-1)/n}$$

We suppose w.l.o.g. that $\int f^{\ n/(n-1)} = 1$

There exists an optimal transport $T = \mathrm{grad}\overline{\Psi}$ between $f^{\ n/(n-1)}(x)dx$ and $\mathbf{1}B_2^n/\mathrm{Vol}(B_2^n)dx$



Monge-Ampère equation: $\mathrm{Vol}(B_2^n)\ f^{n/(n-1)}(x) = \det \mathrm{Hess}\ \overline{\Psi}$

Arithmetico-geometric ineq: $\det(H)^{1/n} \leq \mathrm{trace}(H)/n$ if $H$ positive

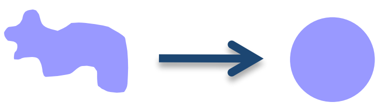
# Part. 2 Optimal Transport – Isoperimetric inequality

**$L_1$ Sobolev inegality:** a proof with OT [Gromov]

$$\int |\mathrm{grad}\, f| \geq n\, \mathrm{Vol}(B_2^n)^{1/n} \left(\int f^{\,n/(n-1)}\right)^{(n-1)/n}$$

We suppose w.l.o.g. that $\int f^{\,n/(n-1)} = 1$

There exists an optimal transport $T = \mathrm{grad}\overline{\Psi}$ between $f^{\,n/(n-1)}(x)dx$ and $\mathbf{1}B_2^n/\mathrm{Vol}(B_2^n)dx$



Monge-Ampère equation: $\mathrm{Vol}(B_2^n)\, f^{n/(n-1)}(x) = \det \mathrm{Hess}\, \overline{\Psi}$

Arithmetico-geometric ineq: $\det(H)^{1/n} \leq \mathrm{trace}(H)/n$ if $H$ positive

$\det(\mathrm{Hess}\, \overline{\Psi})^{1/n} \leq \mathrm{trace}(\mathrm{Hess}\, \overline{\Psi})/n$

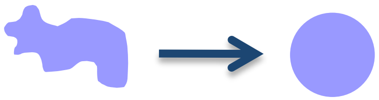
# Part. 2 Optimal Transport – Isoperimetric inequality

**$L_1$ Sobolev inegality:** a proof with OT [Gromov]

$$\int |\operatorname{grad} f| \geq n \, Vol(B_2^n)^{1/n} \left( \int f^{n/(n-1)} \right)^{(n-1)/n}$$

We suppose w.l.o.g. that $\int f^{n/(n-1)} = 1$

There exists an optimal transport $T = \operatorname{grad} \overline{\Psi}$ between $f^{n/(n-1)}(x)dx$ and $\mathbf{1}B_2^n / Vol(B_2^n) dx$



Monge-Ampère equation: $Vol(B_2^n) \, f^{n/(n-1)}(x) = \det \operatorname{Hess} \overline{\Psi}$

Arithmetico-geometric ineq: $\det(H)^{1/n} \leq \operatorname{trace}(H)/n$ if $H$ positive

$\det(\operatorname{Hess} \overline{\Psi})^{1/n} \leq \operatorname{trace}(\operatorname{Hess} \overline{\Psi})/n$

$\det(\operatorname{Hess} \overline{\Psi})^{1/n} \leq \Delta \overline{\Psi} / n$

# Part. 2 Optimal Transport – Isoperimetric inequality

**$L_1$ Sobolev inegality:** a proof with OT [Gromov]

$$\int |\text{grad } f| \geq n \, \text{Vol}(B_2^n)^{1/n} \left( \int f^{\,n/(n-1)} \right)^{(n-1)/n}$$

We suppose w.l.o.g. that $\int f^{\,n/(n-1)} = 1$

$\det(\text{Hess } \overline{\Psi})^{1/n} \leq (\Delta \overline{\Psi})/n$

Monge-Ampère equation:
$\text{Vol}(B_2^n) \, f^{n/(n-1)}(x) = \det \text{Hess } \overline{\Psi}$

# **Part. 2** Optimal Transport – Isoperimetric inequality

**$L_1$ Sobolev inegality:** a proof with OT [Gromov]

$$\int | \text{grad } f | \geq n \, \text{Vol}(B_2^n)^{1/n} \left( \int f^{\,n/(n-1)} \right)^{(n-1)/n}$$

We suppose w.l.o.g. that $\int f^{\,n/(n-1)} = 1$

Monge-Ampère equation:

$\text{Vol}(B_2^n)\, f^{n/(n-1)}(x) = \det \text{Hess}\, \overline{\Psi}$

$\det(\text{Hess}\, \overline{\Psi})^{1/n} \leq (\Delta \overline{\Psi})/n$

$$\text{Vol}(B_2^n) = \text{Vol}(B_2^n) \int f^{\,n/(n-1)} = \int f \;\, \text{Vol}(B_2^n)\, f^{\,1/(n-1)}$$

# Part. 2 Optimal Transport – Isoperimetric inequality

**$L_1$ Sobolev inegality:** a proof with OT [Gromov]

$$\int |\text{grad } f| \geq n \, \text{Vol}(B_2^n)^{1/n} \left( \int f^{\, n/(n-1)} \right)^{(n-1)/n}$$

We suppose w.l.o.g. that $\int f^{\, n/(n-1)} = 1$

Monge-Ampère equation:
$\text{Vol}(B_2^n) \, f^{n/(n-1)}(x) = \det \text{Hess } \overline{\Psi}$

$\det(\text{Hess } \overline{\Psi})^{1/n} \leq (\Delta \overline{\Psi})/n$

$$\text{Vol}(B_2^n) = \text{Vol}(B_2^n) \int f^{\, n/(n-1)} = \int f \, \text{Vol}(B_2^n) \, f^{\, 1/(n-1)} \leq 1/n \int f \, \Delta \overline{\Psi}$$

# Part. 2 Optimal Transport – Isoperimetric inequality

**$L_1$ Sobolev inegality:** a proof with OT [Gromov]

$$\int |\operatorname{grad} f| \geq n \, \mathrm{Vol}(B_2^n)^{1/n} \left( \int f^{\, n/(n-1)} \right)^{(n-1)/n}$$

We suppose w.l.o.g. that $\int f^{\, n/(n-1)} = 1$

Monge-Ampère equation:
$\mathrm{Vol}(B_2^n) \, f^{n/(n-1)}(x) = \det \operatorname{Hess} \overline{\Psi}$

$$\det(\operatorname{Hess} \overline{\Psi})^{1/n} \leq (\Delta \overline{\Psi})/n$$

$$\mathrm{Vol}(B_2^n) = \mathrm{Vol}(B_2^n) \int f^{\, n/(n-1)} = \int f \;\; \mathrm{Vol}(B_2^n) \, f^{\, 1/(n-1)} \leq 1/n \int f \, \Delta \overline{\Psi}$$

$$\int f \, \Delta \overline{\Psi} = - \int \operatorname{grad} f \, . \, \operatorname{grad} \overline{\Psi}$$

# Part. 2 Optimal Transport – Isoperimetric inequality

**$L_1$ Sobolev inegality:** a proof with OT [Gromov]

$$\int |\text{grad } f| \geq n \, \text{Vol}(B_2^n)^{1/n} \left( \int f^{\,n/(n-1)} \right)^{(n-1)/n}$$

We suppose w.l.o.g. that $\int f^{\,n/(n-1)} = 1$

$\det(\text{Hess } \overline{\Psi})^{1/n} \leq (\Delta \overline{\Psi})/n$

Monge-Ampère equation:
$\text{Vol}(B_2^n) \, f^{n/(n-1)}(x) = \det \text{Hess } \overline{\Psi}$

$$\text{Vol}(B_2^n) = \text{Vol}(B_2^n) \int f^{\,n/(n-1)} = \int f \; \text{Vol}(B_2^n) \, f^{\,1/(n-1)} \leq 1/n \int f \, \Delta \overline{\Psi}$$

$$\int f \, \Delta \overline{\Psi} = - \int \text{grad } f \, . \, \text{grad } \overline{\Psi} \quad \leq \int |\text{grad } f| \qquad (T = \text{grad } \overline{\Psi} \in B_2^n)$$

# Part. 2 Optimal Transport – Isoperimetric inequality

**$L_1$ Sobolev inegality:** a proof with OT [Gromov]

$$\int |\text{grad } f| \geq n \, Vol(B_2^n)^{1/n} \left(\int f^{\,n/(n-1)}\right)^{(n-1)/n}$$

We suppose w.l.o.g. that $\int f^{\,n/(n-1)} = 1$

Monge-Ampère equation:
$Vol(B_2^n) \, f^{n/(n-1)}(x) = \det \text{Hess } \overline{\Psi}$

$\det (\text{Hess } \overline{\Psi})^{1/n} \leq (\Delta \overline{\Psi})/n$

$$Vol(B_2^n) = Vol(B_2^n) \int f^{\,n/(n-1)} = \int f \; Vol(B_2^n) \, f^{\,1/(n-1)} \leq 1/n \int f \, \Delta \overline{\Psi}$$

$$\int f \, \Delta \overline{\Psi} = - \int \text{grad } f \, . \, \text{grad } \overline{\Psi} \leq \int |\text{grad } f| \quad (T = \text{grad } \overline{\Psi} \in B_2^n)$$

$$\int |\text{grad } f| \geq n \, Vol(B_2^n)^{1/n} \; \blacksquare$$

# 3

# Semi-Discrete Optimal Transport

# **Part. 3** Optimal Transport – how to program ?

Continuous

(X;μ)

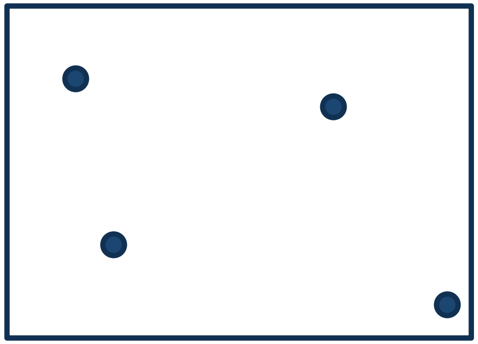
(Y;v)

# **Part. 3** Optimal Transport – how to program ?

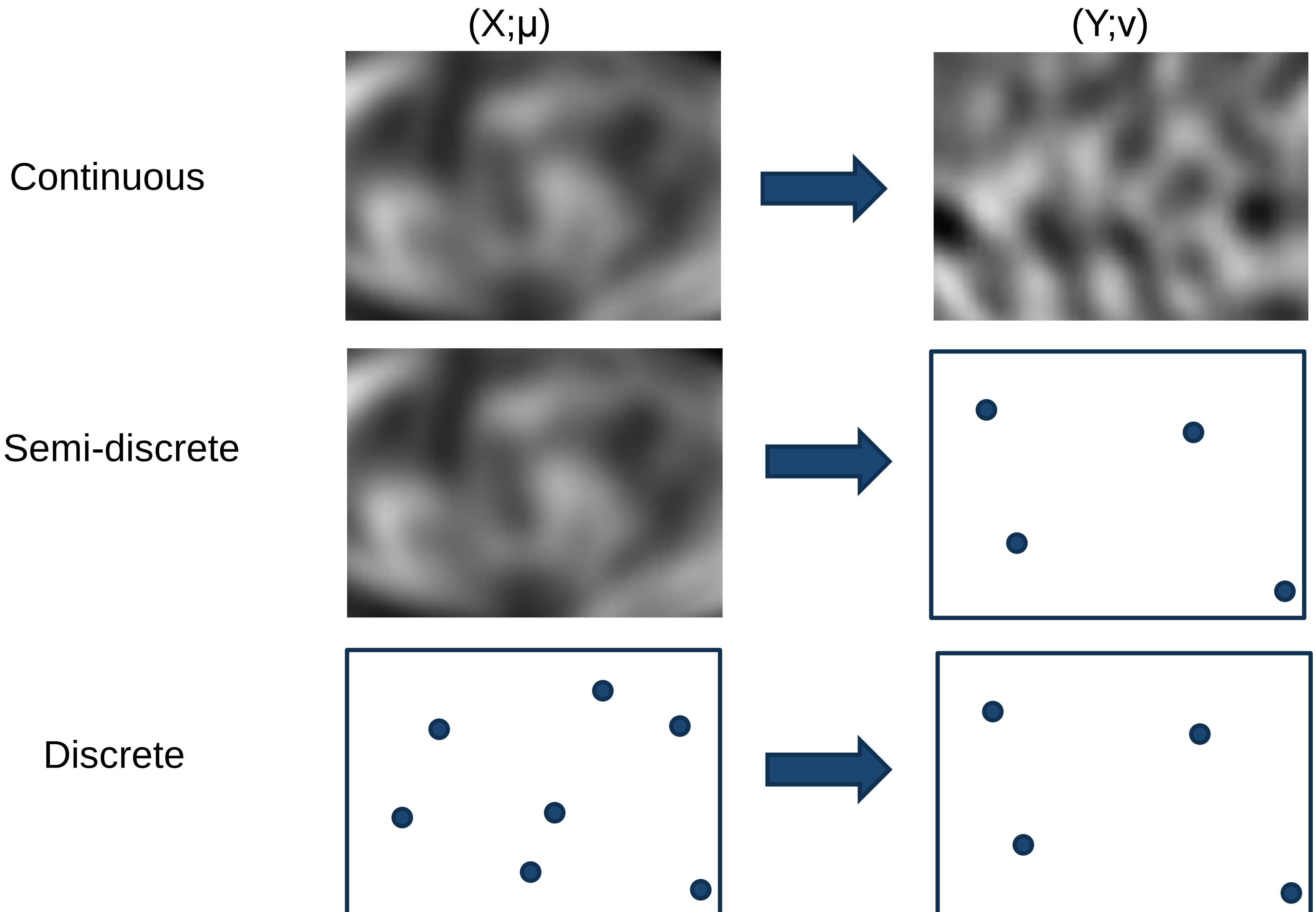
(X;μ) (Y;ν)

Continuous

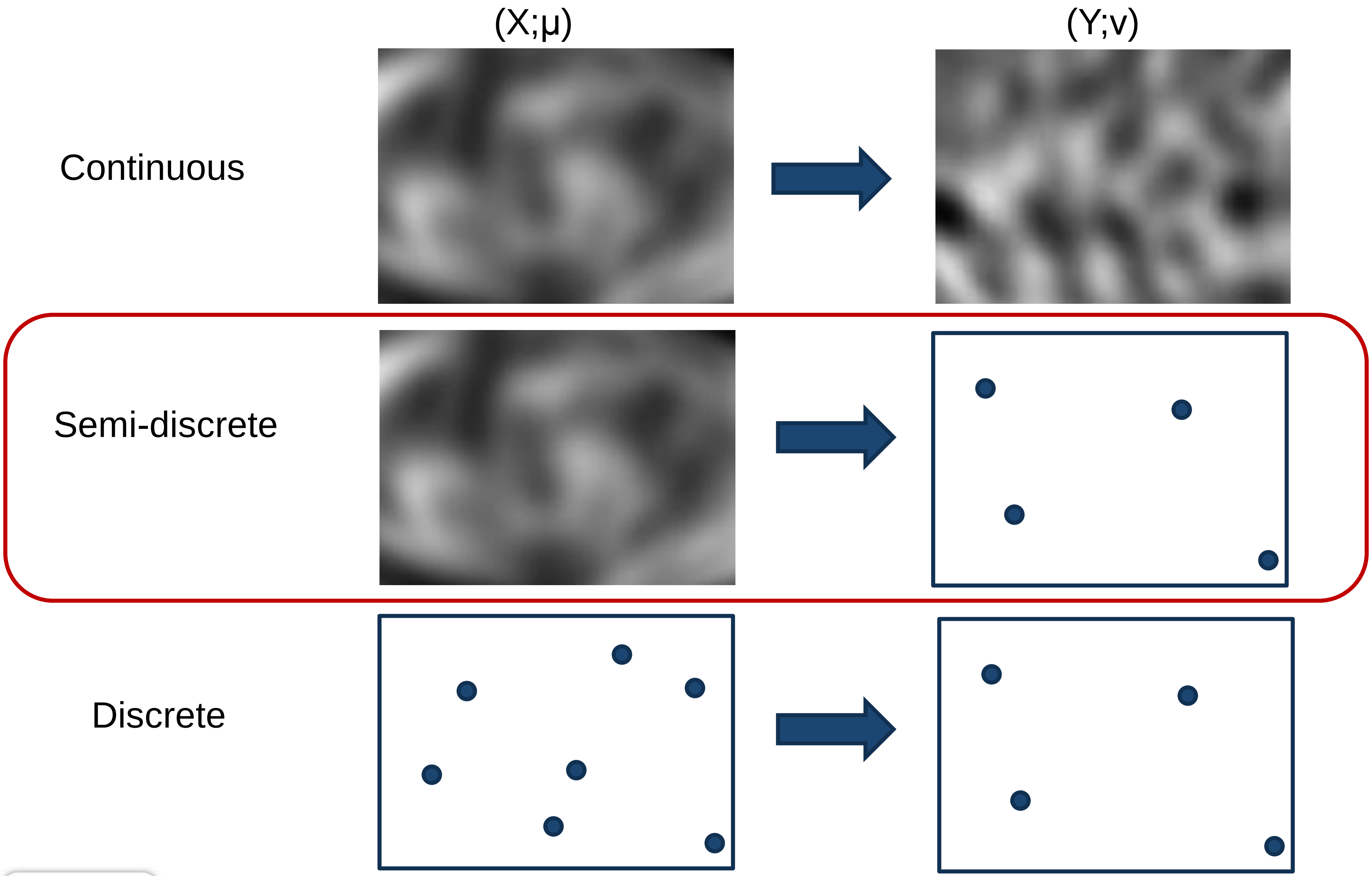
Semi-discrete



# Part. 3 Optimal Transport – how to program ?

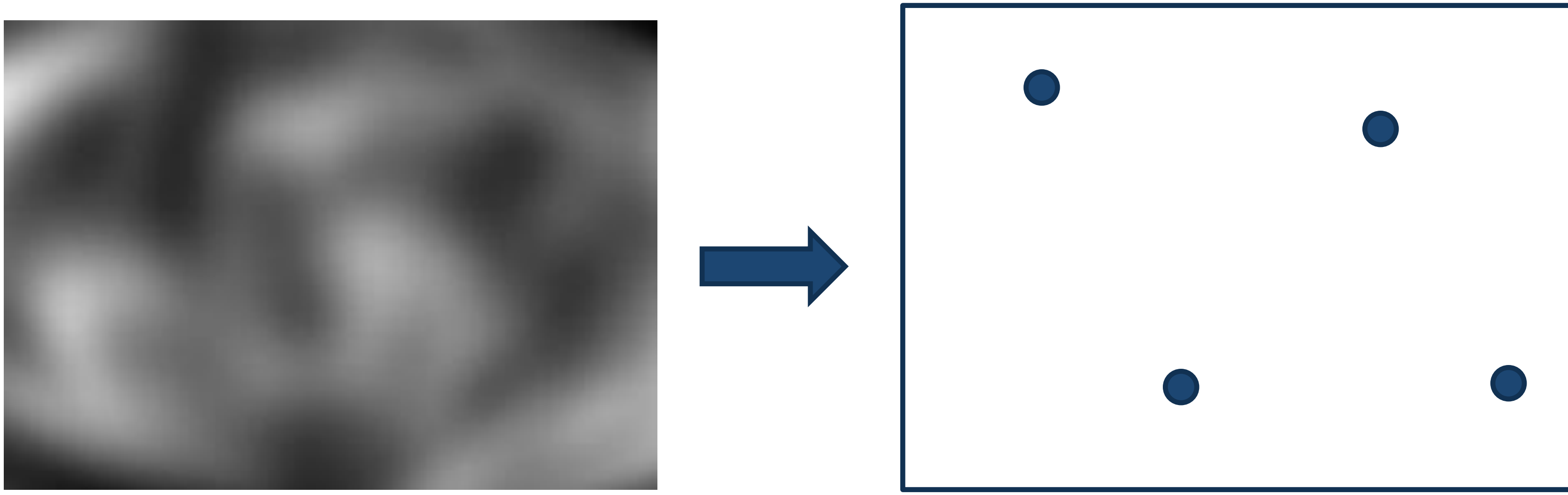


# **Part. 3** Optimal Transport – how to program ?



# **Part. 3** Optimal Transport – semi-discrete

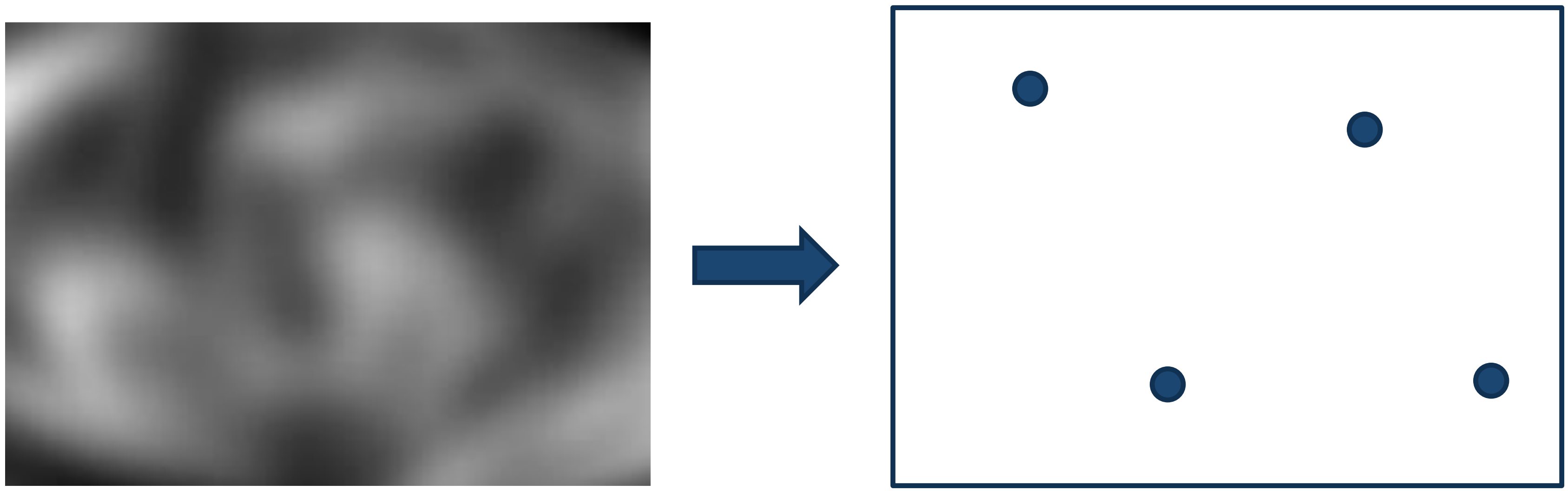
(X;μ) (Y;ν)



**(DMK)** $$\sup_{\psi \in \psi^c} \int_X \psi^c(x)\,d\mu + \int_Y \psi(y)\,d\nu$$

# **Part. 3** Optimal Transport – semi-discrete

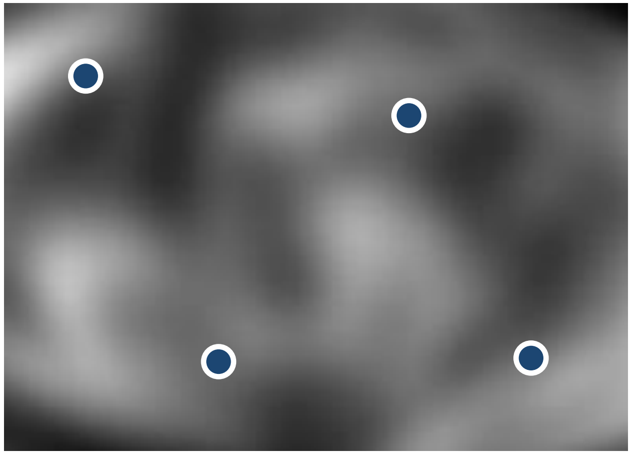
(X;μ) (Y;ν)



**(DMK)** $$\sup_{\psi \in \Psi^c} \int_X \psi^c(x)d\mu + \int_Y \psi(y)d\nu$$

$$\sum_j \psi(y_j)\, \nu_j$$

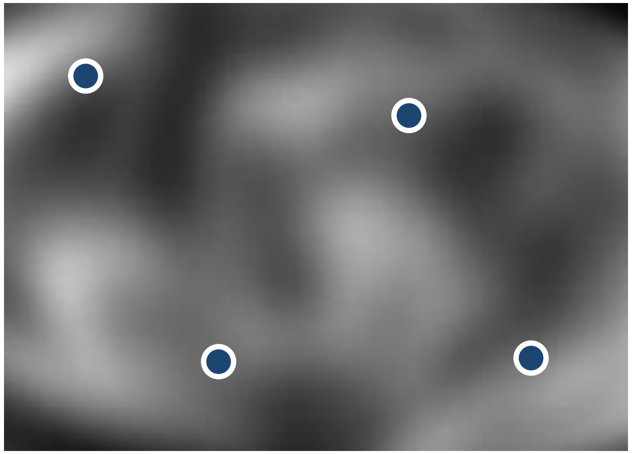
# Part. 3 Optimal Transport – semi-discrete



**(DMK)** $$\sup_{\psi \in \psi^c} \int_X \psi^c(x)d\mu + \int_Y \psi(y)d\nu$$

$$\sum_j \psi(y_j)\, \nu_j$$

# **Part. 3** Optimal Transport – semi-discrete

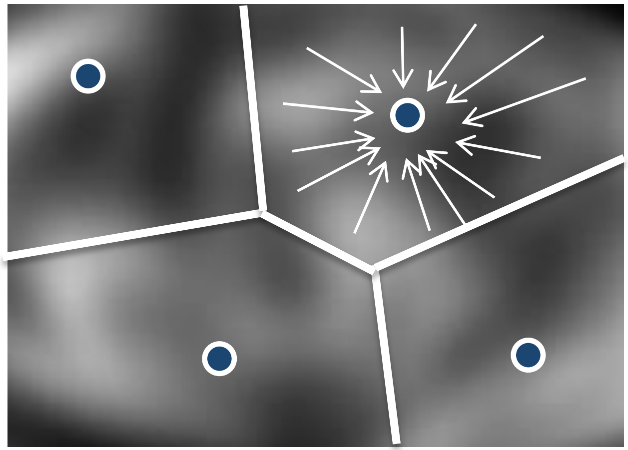


**(DMK)** $$\sup_{\psi \in \psi^c} \int_X \psi^c(x)d\mu + \int_Y \psi(y)dv$$

$$\int_X \inf_{y_j \in Y} [\, \| x - y_j \|^2 - \psi(y_j) \,]\, d\mu$$

$$\sum_j \psi(y_j)\, v_j$$

# Part. 3 Optimal Transport – semi-discrete



**(DMK)** $$\sup_{\psi \in \Psi^c} \int_X \psi^c(x)\,d\mu + \int_Y \psi(y)\,d\nu$$

$$\int_X \inf_{y_j \in Y} [\, \| x - y_j \|^2 - \psi(y_j) \,]\, d\mu$$

$$\sum_j \int_{Lag_\psi(y_j)} \| x - y_j \|^2 - \psi(y_j)\, d\mu$$

$$\sum_j \psi(y_j)\, \nu_j$$

# Part. 3 Optimal Transport – semi-discrete

**(DMK)** $\sup_{\psi \in \psi^c} G(\psi) = \sum_j \int_{Lag\ \psi(yj)} \| x - y_j \|^2 - \psi(y_j)\, d\mu + \sum_j \psi(y_j)\, v_j$

Where: $Lag\ \psi(yj) = \{ x \mid \| x - y_j \|^2 - \psi(y_j) < \| x - y_{j'} \|^2 - \psi(y_{j'}) \}$ for all $j' \neq j$

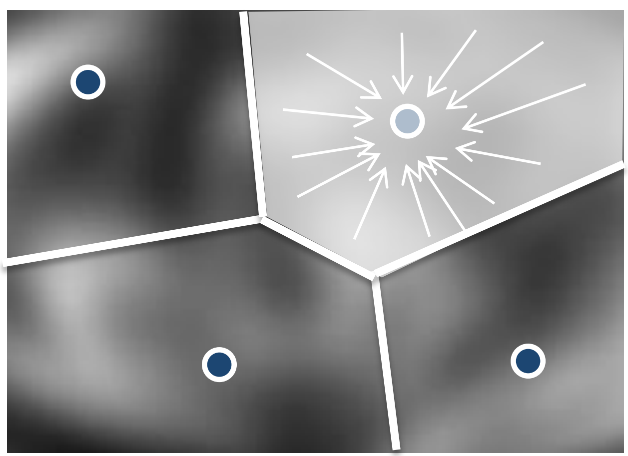
# **Part. 3** Optimal Transport – semi-discrete

**(DMK)**

$$\sup_{\psi \in \psi^c} G(\psi) = \sum_j \int_{\mathrm{Lag}\,\psi(yj)} \| x - y_j \|^2 - \psi(y_j)\, d\mu \; + \; \sum_j \psi(y_j)\, v_j$$

Where: $\mathrm{Lag}\,\psi(yj) = \{ x \mid \| x - y_j \|^2 - \psi(y_j) < \| x - y_{j'} \|^2 - \psi(y_{j'}) \}$ for all $j' \neq j$

**Laguerre diagram** of the $y_j$'s
(with the $L_2$ cost $\| x - y \|^2$ used here, Power diagram)



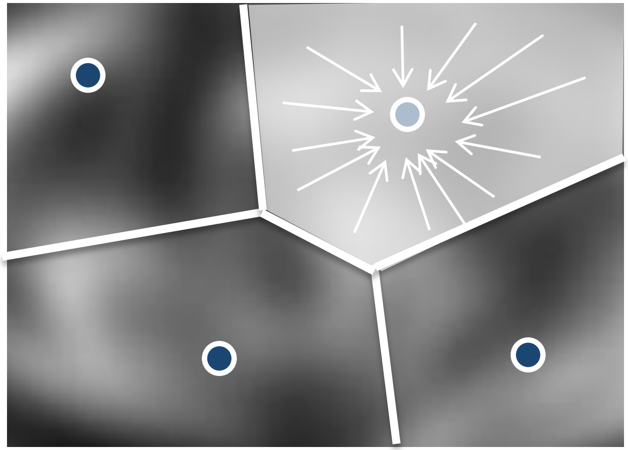
# **Part. 3** Optimal Transport – semi-discrete

**(DMK)** $\sup_{\psi \in \psi^c} G(\psi) = \sum_j \int_{Lag\ \psi(yj)} \| x - y_j \|^2 - \psi(y_j)\, d\mu \ + \ \sum_j \psi(y_j)\, v_j$

Where: $Lag\ \psi(yj) = \{ x \mid \| x - y_j \|^2 - \psi(y_j) < \| x - y_{j'} \|^2 - \psi(y_{j'}) \}$ for all $j' \neq j$

**Laguerre diagram** of the $y_j$'s
(with the $L_2$ cost $\| x - y \|^2$ used here, Power diagram)

**Weight** of $y_j$ in the power diagram



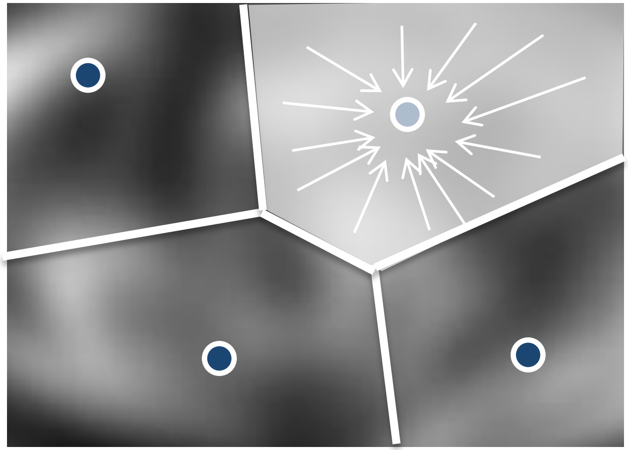
# **Part. 3** Optimal Transport – semi-discrete

**(DMK)** $$\sup_{\psi \in \Psi^c} G(\psi) = \sum_j \int_{Lag\ \psi(yj)} \| x - y_j \|^2 - \psi(y_j)\, d\mu \ + \ \sum_j \psi(y_j)\, v_j$$

Where: $Lag\ \psi(yj) = \{ x \mid \| x - y_j \|^2 - \psi(y_j) < \| x - y_{j'} \|^2 - \psi(y_{j'}) \}$ for all $j' \neq j$

**Laguerre diagram** of the $y_j$'s
(with the $L_2$ cost $\| x - y \|^2$ used here, Power diagram)

**Weight** of $y_j$ in the power diagram



$\psi$ is determined by the
**weight vector** $[\psi(y_1)\ \psi(y_2)\ \ldots\ \psi(y_m)]$

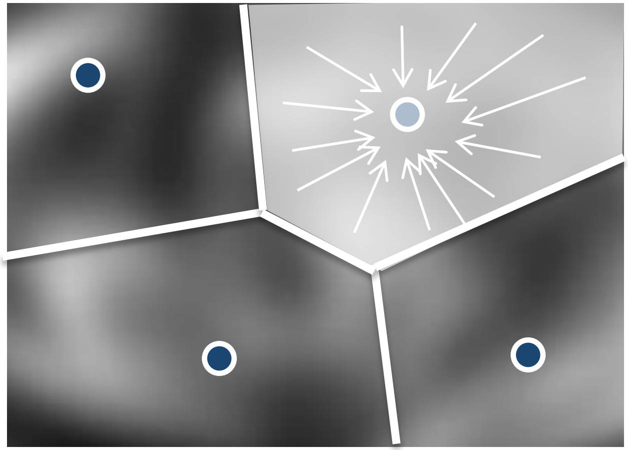
# Part. 3 Optimal Transport – semi-discrete

**(DMK)** $$\sup_{\psi \in \Psi^c} G(\psi) = \sum_j \int_{Lag\ \psi(yj)} \| x - y_j \|^2 - \psi(y_j)\, d\mu + \sum_j \psi(y_j)\, v_j$$

Where: $Lag\ \psi(yj) = \{ x \mid \| x - y_j \|^2 - \psi(y_j) < \| x - y_{j'} \|^2 - \psi(y_{j'}) \}$ for all $j' \neq j$

**Laguerre diagram** of the $y_j$'s
(with the $L_2$ cost $\| x - y \|^2$ used here, Power diagram)

**Weight** of $y_j$ in the power diagram



$\psi$ is determined by the
**weight vector** $[\psi(y_1)\ \psi(y_2)\ \dots\ \psi(y_m)]$

For all weight vector, $\psi$ is c-concave

# **Part. 3** Power Diagrams

**Voronoi diagram:** $Vor(x_i) = \{ x \mid d^2(x,x_i) \quad < d^2(x,x_j) \quad \}$

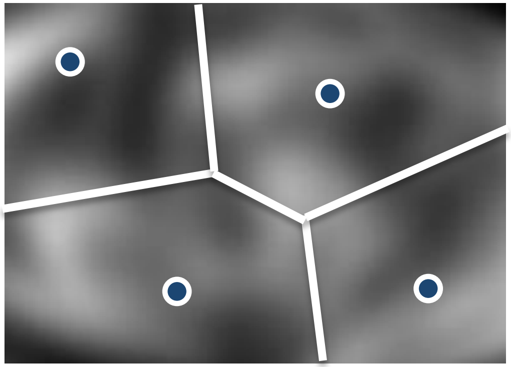
# **Part. 3** Power Diagrams

**Voronoi diagram:** $Vor(x_i) = \{ x \mid d^2(x,x_i) < d^2(x,x_j) \}$

**Power diagram:** $Pow(x_i) = \{ x \mid d^2(x,x_i) - \psi_i < d^2(x,x_j) - \psi_j \}$

# **Part. 3** Power Diagrams

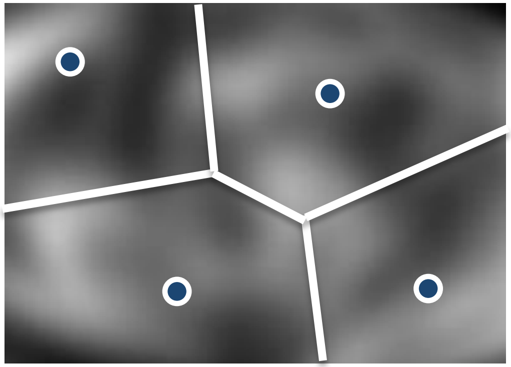
# **Part. 3** Optimal Transport



***Theorem:*** *(direct consequence of MK duality*
*alternative proof in* [Aurenhammer, Hoffmann, Aronov 98] ):

Given a measure μ with density, a set of points $(y_j)$, a set of positive coefficients $v_j$ such that $\sum v_j = \int d\mu(x)$, it is possible to find the weights $W = [\psi(y_1)\ \psi(y_2)\ \dots\ \psi(y_m)]$ such that the map $T_S^W$ is the unique optimal transport map between μ and $\nu = \sum v_j\, \delta(y_j)$

# **Part. 3** Optimal Transport



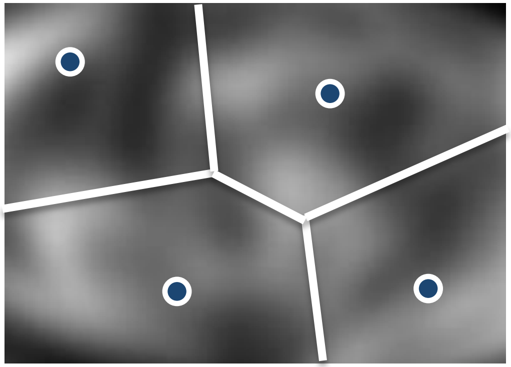
***Theorem:*** *(direct consequence of MK duality alternative proof in* [Aurenhammer, Hoffmann, Aronov 98] ):

Given a measure μ with density, a set of points $(y_j)$, a set of positive coefficients $v_j$ such that $\sum v_j = \int d\mu(x)$, it is possible to find the weights $W = [\psi(y_1)\ \psi(y_2)\ \ldots\ \psi(y_m)]$ such that the map $T_S^W$ is the unique optimal transport map between μ and $v = \sum v_j\, \delta(y_j)$

***Proof:*** $G(\psi) = \sum_j \int_{\mathrm{Lag}\ \psi(yj)} \| x - y_j \|^2 - \psi(y_j)\, d\mu \ + \ \sum_j \psi(y_j)\, v_j$

Is a concave function of the weight vector $[\psi(y_1)\ \psi(y_2)\ \ldots\ \psi(y_m)]$

# **Part. 3** Optimal Transport



***Theorem:*** *(direct consequence of MK duality*
*alternative proof in* [Aurenhammer, Hoffmann, Aronov 98] ):

Given a measure μ with density, a set of points $(y_j)$, a set of positive coefficients $v_j$ such that $\sum v_j = \int d\mu(x)$, it is possible to find the weights $W = [\psi(y_1)\ \psi(y_2)\ \ldots\ \psi(y_m)]$ such that the map $T_S^W$ is the unique optimal transport map between μ and $\nu = \sum v_j\, \delta(y_j)$

***Proof:*** $$G(\psi) = \sum_j \int_{Lag_{\psi}(y_j)} \| x - y_j \|^2 - \psi(y_j)\, d\mu + \sum_j \psi(y_j)\, v_j$$

Is a concave function of the weight vector $[\psi(y_1)\ \psi(y_2)\ \ldots\ \psi(y_m)]$

# **Part. 3** Optimal Transport – the AHA paper

***Idea of the proof***

Consider the function $$f_T(W) = \int \left( \| x - T(x) \|^2 - \psi\left(T(X)\right) \right) d\mu(x)$$

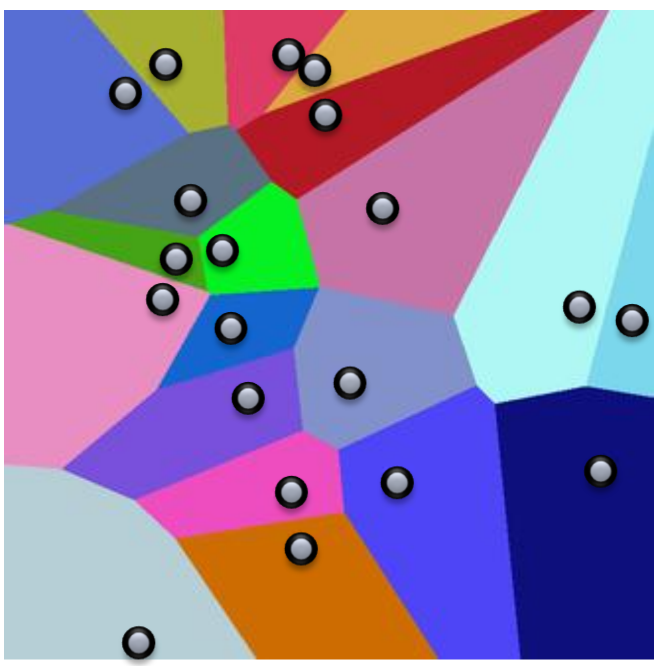
The (unknown) weights $W = [\psi(y_1)\ \psi(y_2)\ \dots\ \psi(y_m)]$

# **Part. 3** Optimal Transport – the AHA paper

***Idea of the proof***

Consider the function $$f_T(W) = \int \left( \| x - T(x) \|^2 - \psi\left(T(X)\right) \right) d\mu(x)$$

T : an arbitrary but fixed assignment.



# **Part. 3** Optimal Transport – the AHA paper

***Idea of the proof***

Consider the function $f_T(W) = \int \left( \| x - T(x) \|^2 - \psi\left(T(X)\right) \right) d\mu(x)$

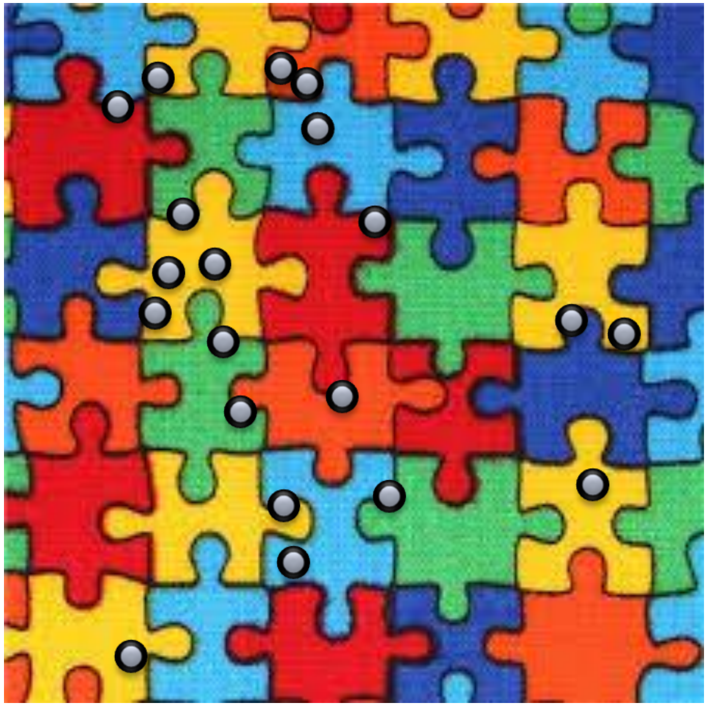
T : an arbitrary but fixed assignment.

# Part. 3 Optimal Transport – the AHA paper

***Idea of the proof***

Consider the function $f_T(W) = \int \left( \| x - T(x) \|^2 - \psi\left(T(X)\right) \right) d\mu(x)$

T : an arbitrary but fixed assignment.

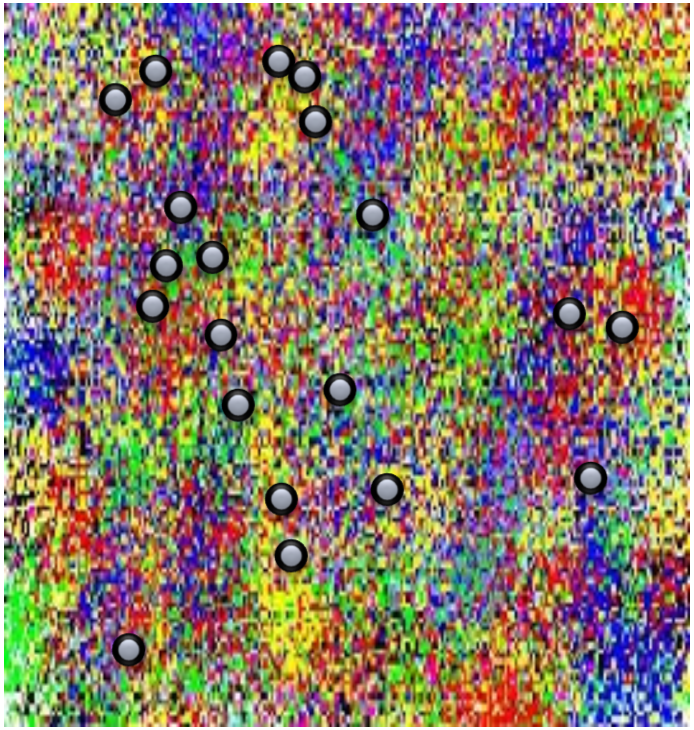


# **Part. 3** Optimal Transport – the AHA paper

***Idea of the proof***

Consider the function $f_T(W) = \int \left( \| x - T(x) \|^2 - \psi\left(T(X)\right) \right) d\mu(x)$

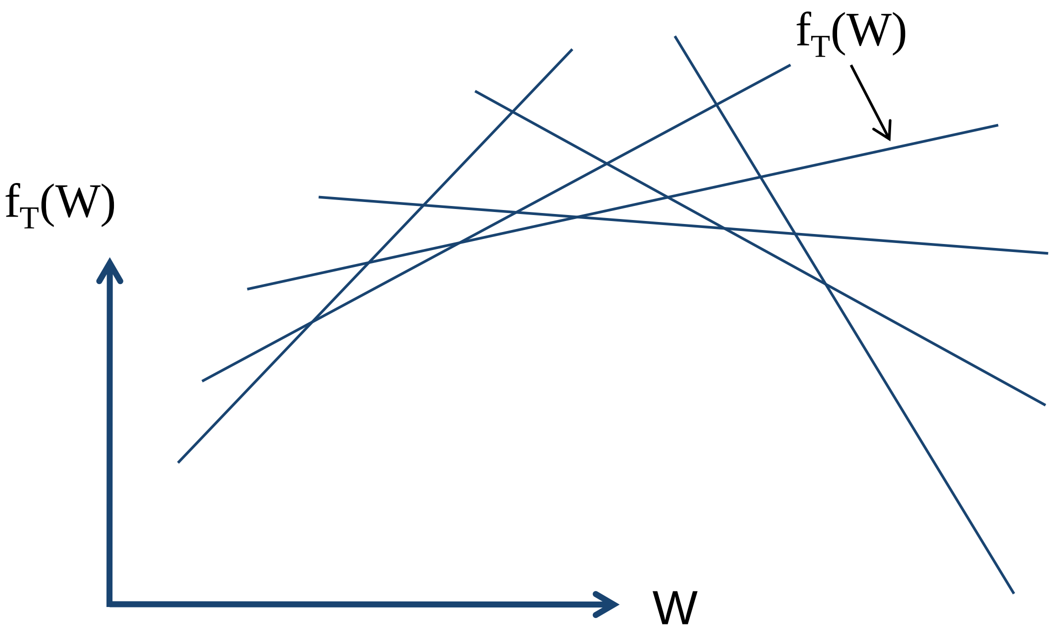
T : an arbitrary but fixed assignment.



# Part. 3 Optimal Transport – the AHA paper

***Idea of the proof***

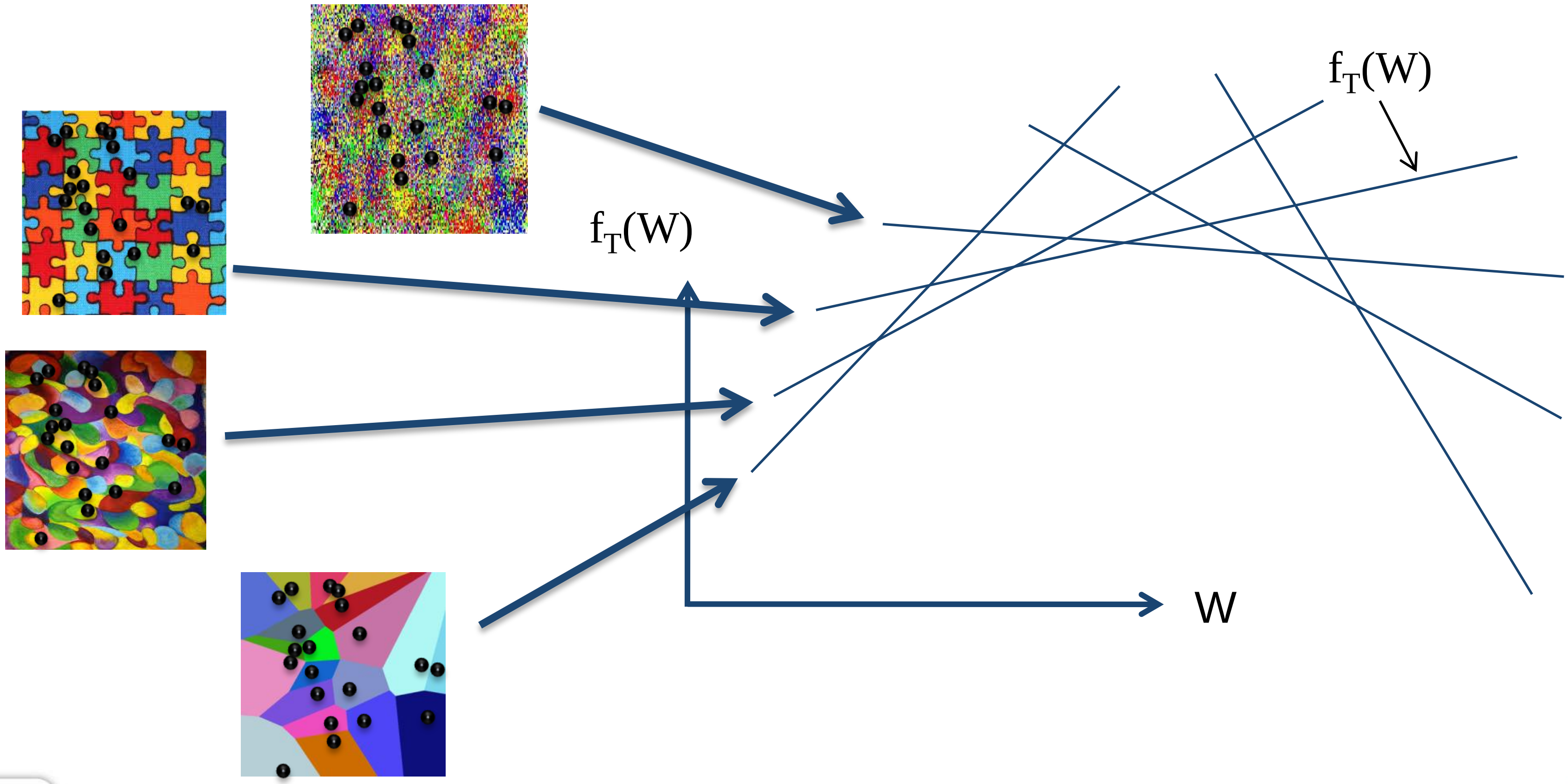
Consider the function $$f_T(W) = \int \left( \| x - T(x) \|^2 - \psi\left(T(X)\right) \right) d\mu(x)$$



# Part. 3 Optimal Transport – the AHA paper

***Idea of the proof***

Consider the function $f_T(W) = \int \left( \| x - T(x) \|^2 - \psi\left(T(X)\right)\right) d\mu(x)$

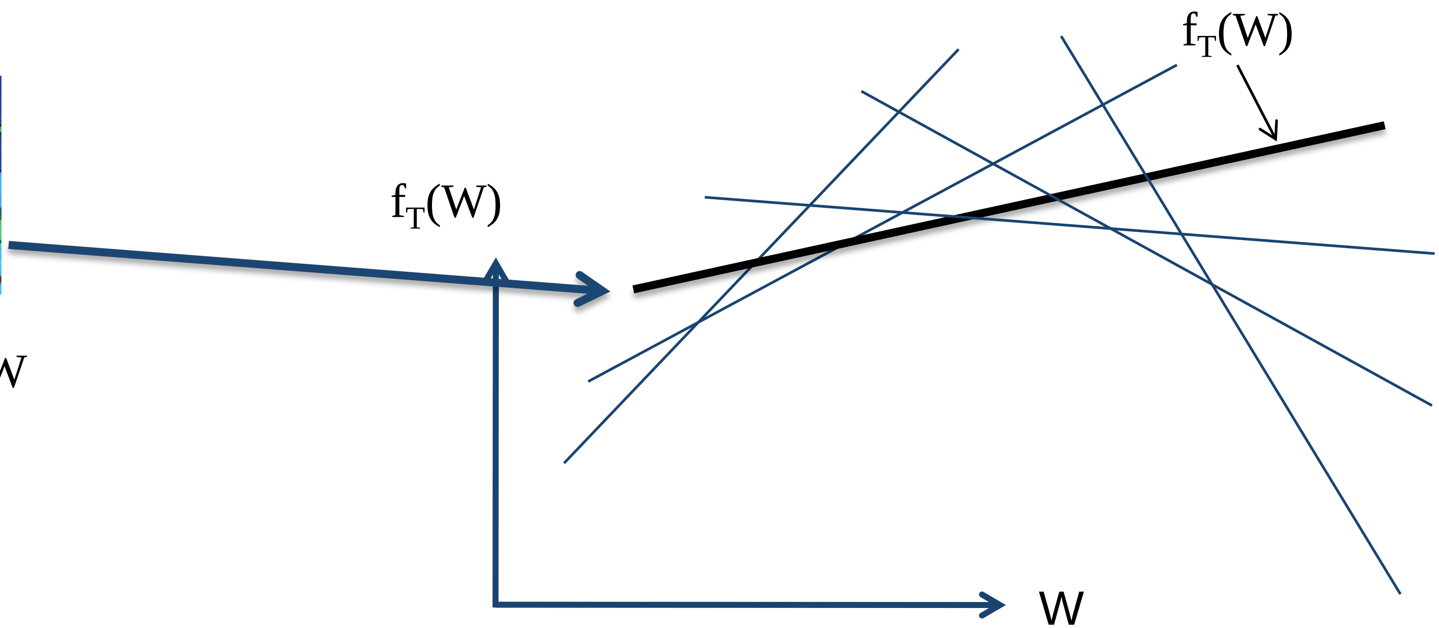


# **Part. 3** Optimal Transport – the AHA paper

***Idea of the proof***

Consider the function $f_T(W) = \int \left( \| x - T(x) \|^2 - \psi\left(T(X)\right) \right) d\mu(x)$

Fixed T




$f_T(W)$ is linear in W

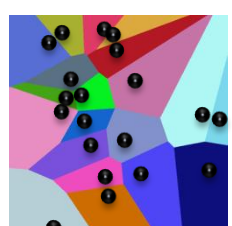
# **Part. 3** Optimal Transport – the AHA paper

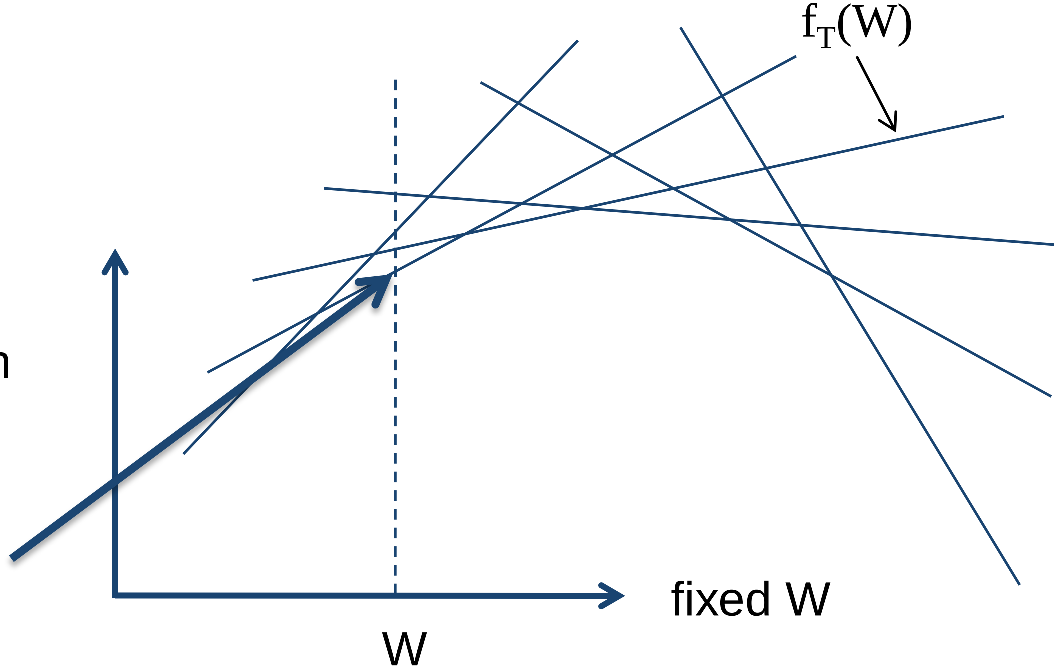
***Idea of the proof***

Consider the function $f_T(W) = \int \left( \| x - T(x) \|^2 - \psi\left(T(X)\right) \right) d\mu(x)$

$f_T(W)$ is linear in W

$f_{T_W}(W)$ : defined by power diagram





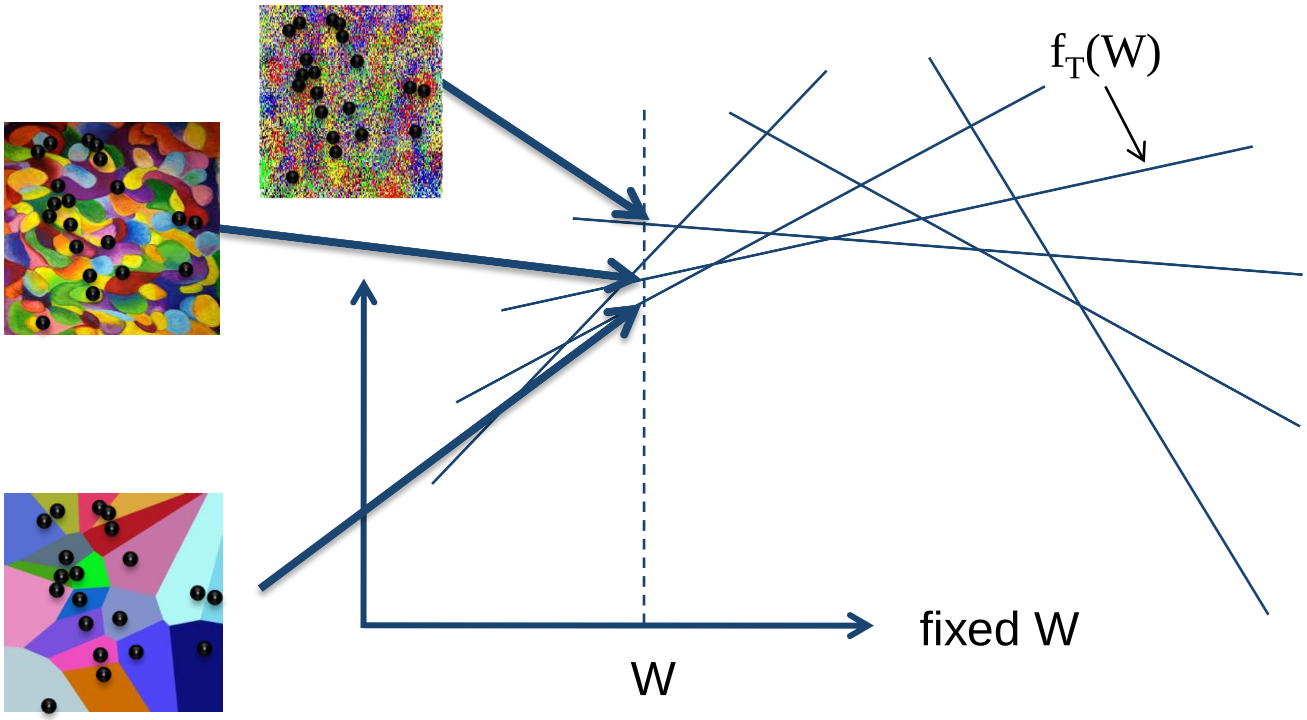
# Part. 3 Optimal Transport – the AHA paper

***Idea of the proof***

Consider the function $f_T(W) = \int \left( \| x - T(x) \|^2 - \psi(T(X)) \right) d\mu(x)$

$f_T(W)$ is linear in W

$f_{T_W}(W) = \min_T f_T(W)$



# **Part. 3** Optimal Transport – the AHA paper

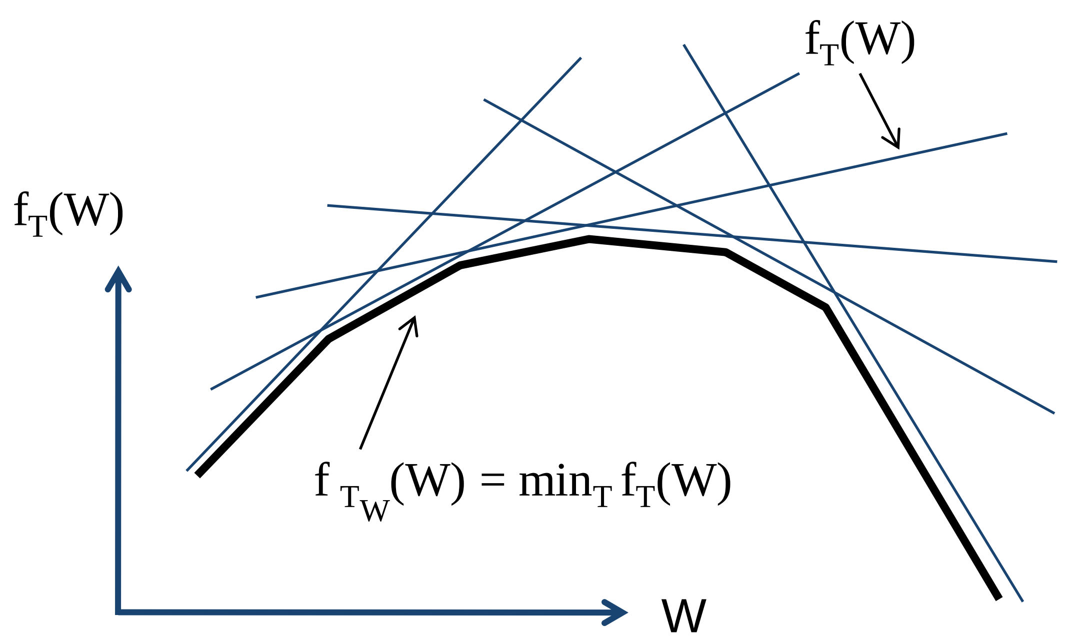
***Idea of the proof***

Consider the function $$f_T(W) = \int \left( \| x - T(x) \|^2 - \psi\left(T(X)\right) \right) d\mu(x)$$

$f_T(W)$ is linear in W

f: $W \rightarrow f_{T_W}(W)$ is **concave !!**

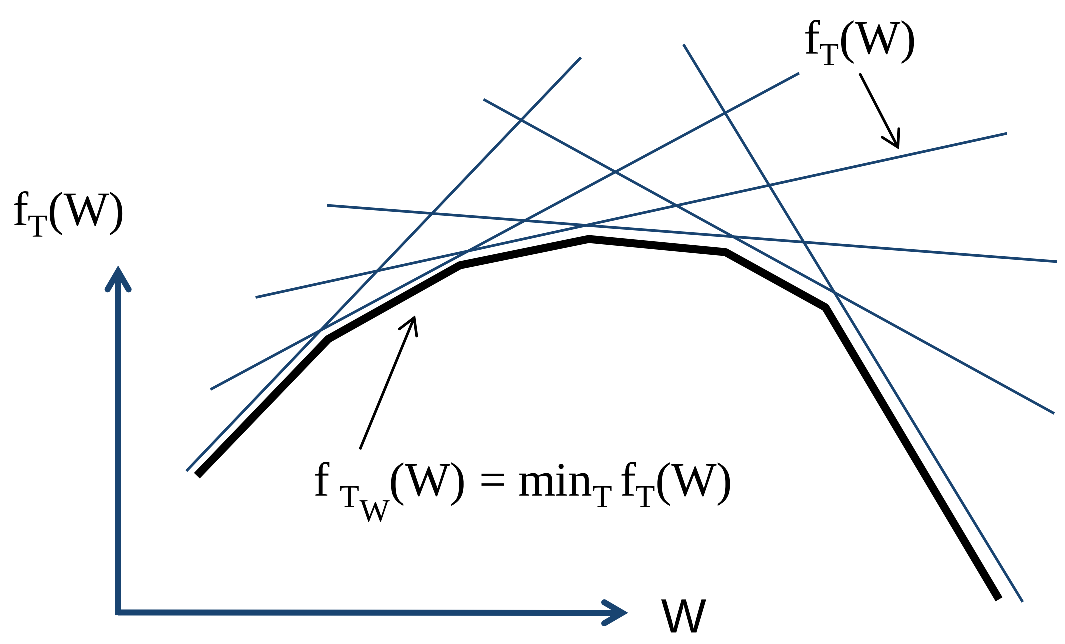
(because its graph is the lower enveloppe of linear functions)



# Part. 3 Optimal Transport – the AHA paper

***Idea of the proof***

Consider the function $f_T(W) = \int \left( \| x - T(x) \|^2 - \psi\left(T(X)\right) \right) d\mu(x)$



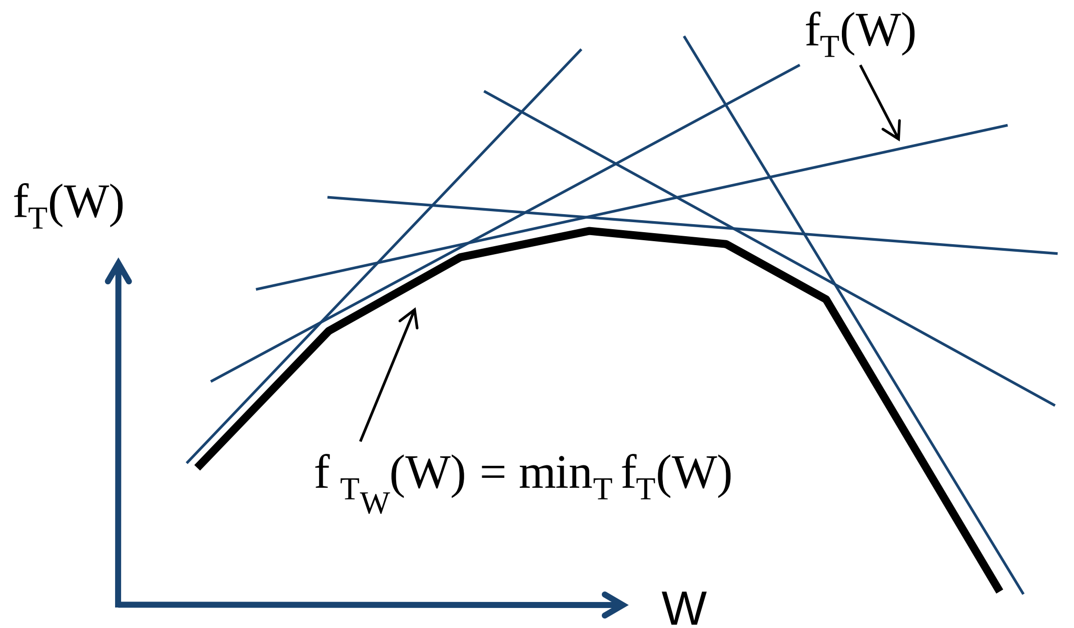

$f_T(W)$ is linear in W
$f: W \rightarrow f_{T_W}(W)$ is concave
(because its graph is the lower enveloppe of linear functions)

Consider $g(W) = f_{T_W}(W) + \sum v_j \psi_j$

# Part. 3 Optimal Transport – the AHA paper

***Idea of the proof***

Consider the function $f_T(W) = \int \left( \| x - T(x) \|^2 - \psi(T(X)) \right) d\mu(x)$




$f_T(W)$ is linear in W
f: $W \rightarrow f_{T_W}(W)$ is concave
(because its graph is the lower enveloppe of linear functions)

Consider $g(W) = f_{T_W}(W) + \sum v_j \psi_j$

$\partial g / \partial \psi_j = v_j - \int_{pow(yj)} \| x - y_j \|^2 d\mu(x)$ and g is concave.

# Part. 3 Optimal Transport – the algorithm

**Semi-discrete OT Summary:**

**(DMK)** $$\sup_{\psi \in \boldsymbol{\psi}^c} G(\psi) = \int_X \psi^c(x)d\mu + \int_Y \psi(y)d\nu$$

# Part. 3 Optimal Transport – the algorithm

**Semi-discrete OT Summary:**

**(DMK)** $\sup_{\psi \in \Psi^c} G(\psi) = \int_X \psi^c(x)d\mu + \int_Y \psi(y)d\nu$

$$G(\psi) = g(W) = \sum_j \int_{Lag_{\psi}(y_j)} \| x - y_j \|^2 - \psi(y_j)\, d\mu + \sum_j \psi(y_j)\, \nu_j$$ is concave

# Part. 3 Optimal Transport – the algorithm

**Semi-discrete OT Summary:**

**(DMK)** $\sup_{\psi \in \psi^c} G(\psi) = \int_X \psi^c(x) d\mu + \int_Y \psi(y) dv$

$G(\psi) = g(W) = \sum_j \int_{Lag\ \psi(yj)} \| x - y_j \|^2 - \psi(y_j)\, d\mu + \sum_j \psi(y_j)\, v_j$ is concave

$\partial G / \partial \psi_j = v_j - \int_{pow(yj)} \| x - y_j \|^2 d\mu(x)$ (= 0 at the maximum)

# Part. 3 Optimal Transport – the algorithm

**Semi-discrete OT Summary:**

**(DMK)** $\sup_{\psi \in \psi^c} G(\psi) = \int_X \psi^c(x)d\mu + \int_Y \psi(y)d\nu$

$G(\psi) = g(W) = \sum_j \int_{Lag_{\psi}(y_j)} \| x - y_j \|^2 - \psi(y_j)\, d\mu + \sum_j \psi(y_j)\, \nu_j$ is concave

$\partial G / \partial \psi_j = \nu_j - \int_{pow(y_j)} \| x - y_j \|^2 d\mu(x)$ (= 0 at the maximum)

Desired mass at $y_j$

Mass transported to $y_j$

# **Part. 3** Optimal Transport – the algorithm

**The [AHA] paper summary:**

- The optimal weights minimize a convex function
- The gradient of this convex function is easy to compute

Note: the weight w(s) correspond to the Kantorovich potential $\psi(x)$ (solves a “discrete Monge-Ampere” equation)

**The algorithm:**

**Input:** two tetrahedral meshes $M_1$ and $M_2$
**Output:** a morphing between $M_1$ and $M_2$

# **Part. 3** Optimal Transport – the algorithm

**The [AHA] paper summary:**

- The optimal weights minimize a convex function
- The gradient of this convex function is easy to compute

Note: the weight w(s) correspond to the Kantorovich potential $\psi(x)$
(solves a "discrete Monge-Ampere" equation)

**The algorithm:**

**Input:** two tetrahedral meshes $M_1$ and $M_2$
**Output:** a morphing between $M_1$ and $M_2$

**Step 1:** sample $M_2$ with N points ($s_1 \ldots s_N$)

# **Part. 3** Optimal Transport – the algorithm

**The [AHA] paper summary:**

- The optimal weights minimize a convex function
- The gradient of this convex function is easy to compute

Note: the weight w(s) correspond to the Kantorovich potential $\psi(x)$
(solves a "discrete Monge-Ampere" equation)

**The algorithm:**

**Input:** two tetrahedral meshes $M_1$ and $M_2$
**Output:** a morphing between $M_1$ and $M_2$

**Step 1:** sample $M_2$ with N points ($s_1 \ldots s_N$)
**Step 2:** initialize the weights ($w_1 \ldots w_N$) = (0 … 0)

# **Part. 3** Optimal Transport – the algorithm

**The [AHA] paper summary:**

- The optimal weights minimize a convex function
- The gradient of this convex function is easy to compute

Note: the weight w(s) correspond to the Kantorovich potential $\psi(x)$
(solves a "discrete Monge-Ampere" equation)

**The algorithm:**

**Input:** two tetrahedral meshes $M_1$ and $M_2$
**Output:** a morphing between $M_1$ and $M_2$

**Step 1:** sample $M_2$ with N points $(s_1 \dots s_N)$
**Step 2:** initialize the weights $(w_1 \dots w_N) = (0 \dots 0)$
**Step 3:** minimize $g(w_1 \dots w_N)$ with a quasi-Newton algorithm:

# **Part. 3** Optimal Transport – the algorithm

**The [AHA] paper summary:**

- The optimal weights minimize a convex function
- The gradient of this convex function is easy to compute

Note: the weight w(s) correspond to the Kantorovich potential $\psi(x)$
(solves a "discrete Monge-Ampere" equation)

**The algorithm:**

**Input:** two tetrahedral meshes $M_1$ and $M_2$
**Output:** a morphing between $M_1$ and $M_2$

**Step 1:** sample $M_2$ with N points $(s_1 \ldots s_N)$
**Step 2:** initialize the weights $(w_1 \ldots w_N) = (0 \ldots 0)$
**Step 3:** minimize $g(w_1 \ldots w_N)$ with a quasi-Newton algorithm:
For each iterate $(s_1 \ldots s_N)^{(k)}$:

# **Part. 3** Optimal Transport – the algorithm

**The [AHA] paper summary:**

- The optimal weights minimize a convex function
- The gradient of this convex function is easy to compute

Note: the weight w(s) correspond to the Kantorovich potential $\psi(x)$
(solves a "discrete Monge-Ampere" equation)

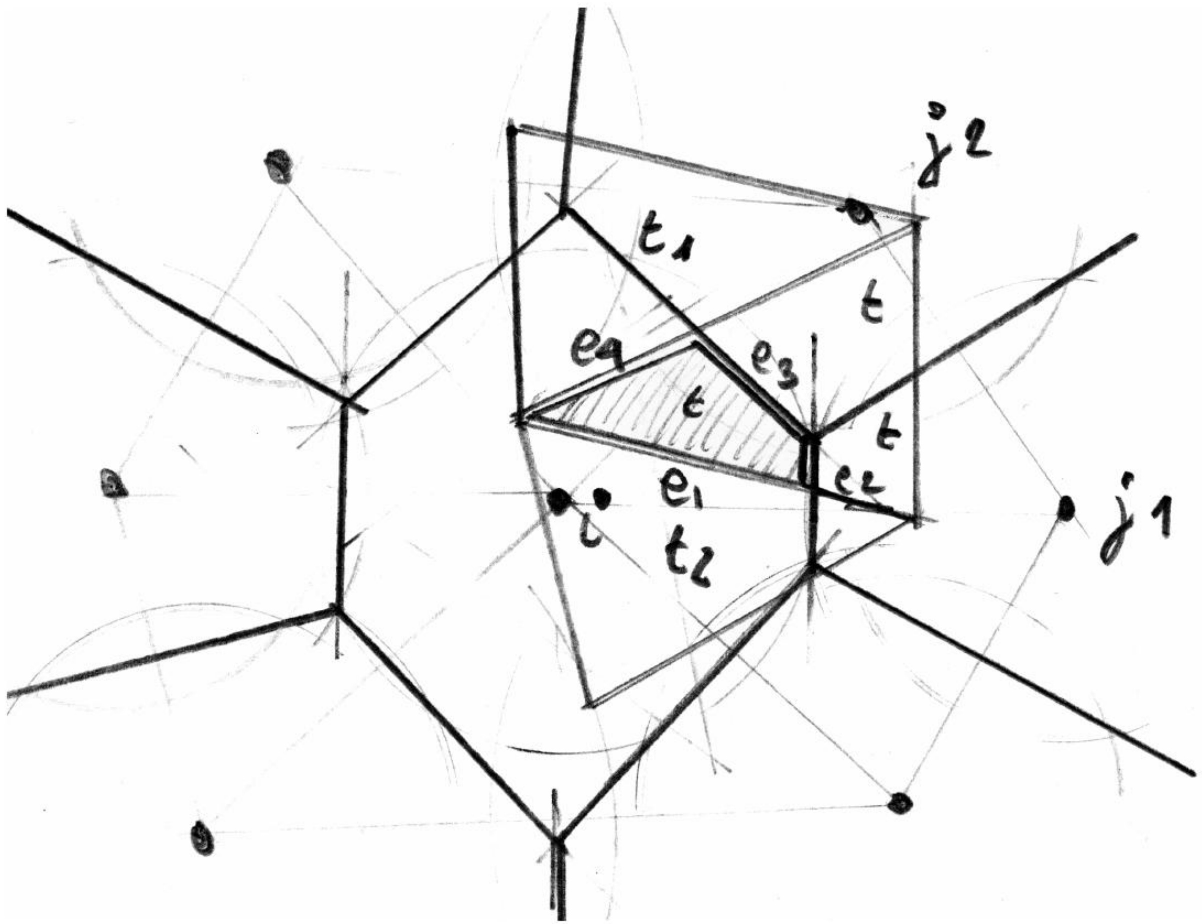
**The algorithm:**

**Input:** two tetrahedral meshes $M_1$ and $M_2$
**Output:** a morphing between $M_1$ and $M_2$

**Step 1:** sample $M_2$ with N points $(s_1 \ldots s_N)$
**Step 2:** initialize the weights $(w_1 \ldots w_N) = (0 \ldots 0)$
**Step 3:** minimize $g(w_1 \ldots w_N)$ with a quasi-Newton algorithm:
For each iterate $(s_1 \ldots s_N)^{(k)}$:
Compute $\mathrm{Pow}\big( (w_i, s_i) \big) \cap M_1$ [Nivoliers, **L** 2014, Curves and Surfaces]

# **Part. 3** Optimal Transport – the algorithm

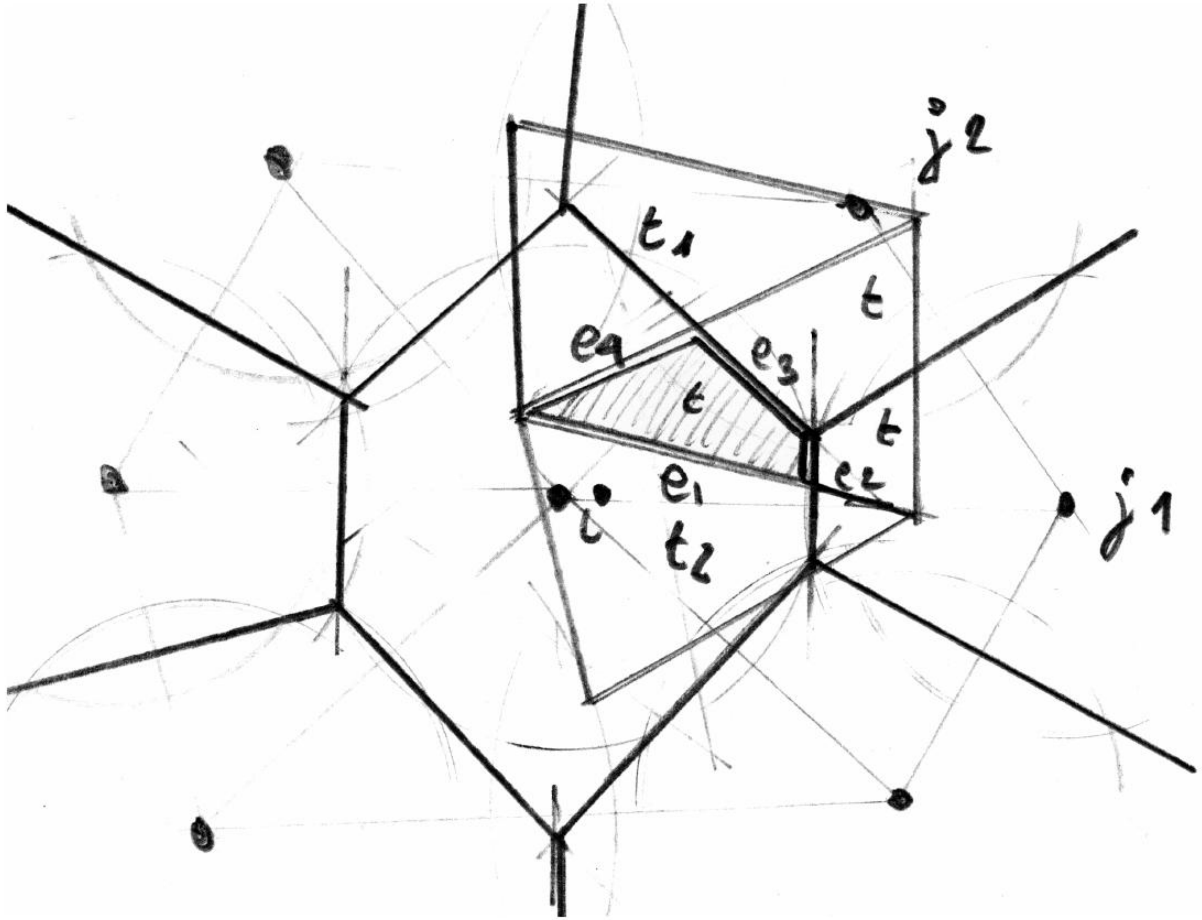
Compute Pow( ($w_i$, $s_i$) ) $\cap$ $M_1$ [Nivoliers, **L** 2014, Curves and Surfaces]



# **Part. 3** Optimal Transport – the algorithm

Compute Pow( $(w_i, s_i)$ ) $\cap$ $M_1$ [Nivoliers, **L** 2014, Curves and Surfaces]

Implementation in **GEOGRAM** (http://alice.loria.fr/software/geogram




Predicates + Predicate Construction Kit [**L** 2015] – available for download

# Part. 3 Optimal Transport – the algorithm

**The [AHA] paper summary:**

- The optimal weights minimize a convex function
- The gradient of this convex function is easy to compute

Note: the weight w(s) correspond to the Kantorovich potential $\psi(x)$
(solves a "discrete Monge-Ampere" equation)

**The algorithm:**

**Input:** two tetrahedral meshes $M_1$ and $M_2$
**Output:** a morphing between $M_1$ and $M_2$

**Step 1:** sample $M_2$ with N points $(s_1 \ldots s_N)$
**Step 2:** initialize the weights $(w_1 \ldots w_N) = (0 \ldots 0)$
**Step 3:** minimize $g(w_1 \ldots w_N)$ with a quasi-Newton algorithm:

For each iterate $(s_1 \ldots s_N)^{(k)}$:

Compute $\mathrm{Pow}\big( (w_i, s_i) \big) \cap M_1$ [**L** 2014, Curves and Surfaces]

Compute g and grad g

# **Part. 3** Optimal Transport – the algorithm

**The [AHA] paper summary:**

- The optimal weights minimize a convex function
- The gradient of this convex function is easy to compute

Note: the weight w(s) correspond to the Kantorovich potential $\psi(x)$
(solves a "discrete Monge-Ampere" equation)

**The algorithm:**

**Input:** two tetrahedral meshes $M_1$ and $M_2$
**Output:** a morphing between $M_1$ and $M_2$

**Step 1:** sample $M_2$ with N points $(s_1 \ldots s_N)$
**Step 2:** initialize the weights $(w_1 \ldots w_N) = (0 \ldots 0)$
**Step 3:** minimize $g(w_1 \ldots w_N)$ with a quasi-Newton algorithm:

For each iterate $(s_1 \ldots s_N)^{(k)}$:

Compute $\text{Pow}\big( (w_i, s_i) \big) \cap M_1$ [**L** 2014, Curves and Surfaces]

Compute g and grad g

\+ *Multilevel version [Merigot 2011] (2D),*
*[**L** 2014 arXiv, M2AN 2015] (3D & relation with Centroidal Voronoi Tesselation)*

# **Part. 3** Optimal Transport – the algorithm

**The [AHA] paper summary:**

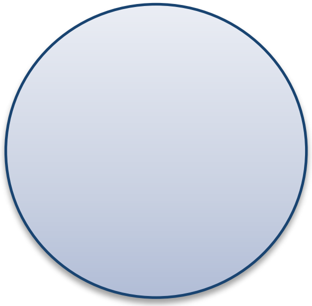
- The optimal weights minimize a convex function
- The gradient of this convex function is easy to compute

Note: the weight w(s) correspond to the Kantorovich potential $\psi(x)$
(solves a "discrete Monge-Ampere" equation)

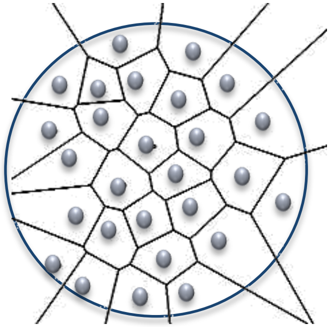
**The algorithm:**

**Summary:**

The algorithm computes the weights $w_i$ such that the power cells associated with the Diracs correspond to the preimages of the Diracs.



μ



ν

# **Part. 3** Optimal Transport – the algorithm

**<u>The [AHA] paper summary:</u>**

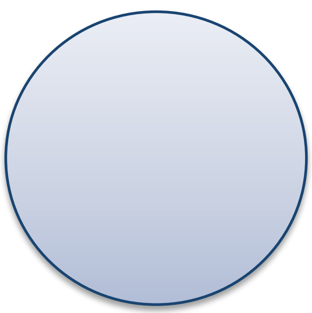
• The optimal weights minimize a convex function
• The gradient of this convex function is easy to compute

Note: the weight w(s) correspond to the Kantorovich potential $\psi(x)$
(solves a “discrete Monge-Ampere” equation)

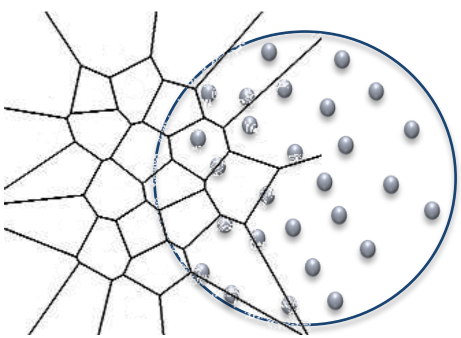
**<u>The algorithm:</u>**

**<u>Summary:</u>**

The algorithm computes the weights $w_i$ such that the power cells associated with the Diracs correspond to the preimages of the Diracs.



μ



ν

# **Part. 3** Optimal Transport – the algorithm

**The [AHA] paper summary:**

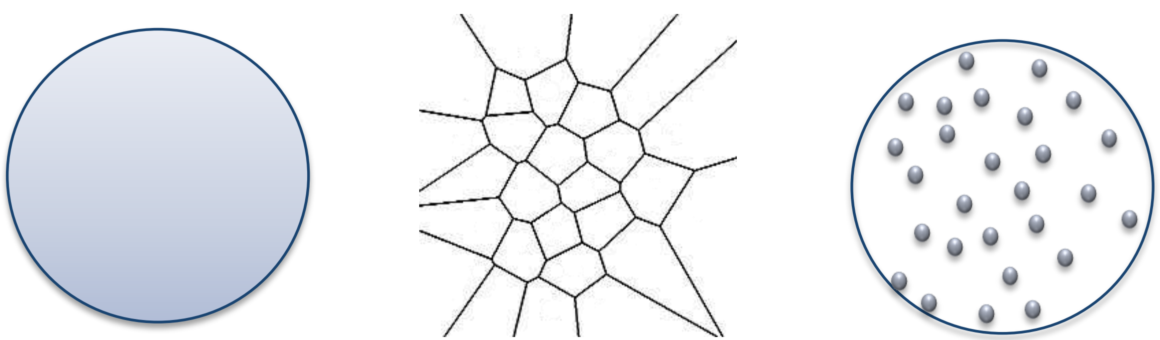
- The optimal weights minimize a convex function
- The gradient of this convex function is easy to compute

Note: the weight w(s) correspond to the Kantorovich potential $\psi(x)$
(solves a "discrete Monge-Ampere" equation)

**The algorithm:**

**Summary:**

The algorithm computes the weights $w_i$ such that the power cells associated with the Diracs correspond to the preimages of the Diracs.



μ ν

# **Part. 3** Optimal Transport – the algorithm

**The [AHA] paper summary:**

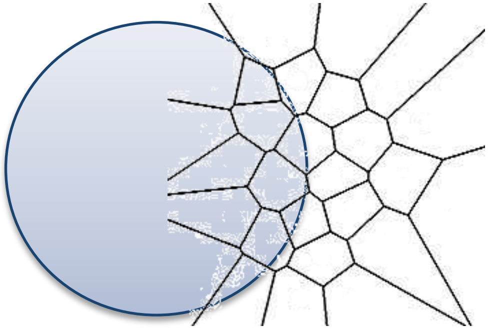
- The optimal weights minimize a convex function
- The gradient of this convex function is easy to compute

Note: the weight w(s) correspond to the Kantorovich potential $\psi(x)$
(solves a "discrete Monge-Ampere" equation)

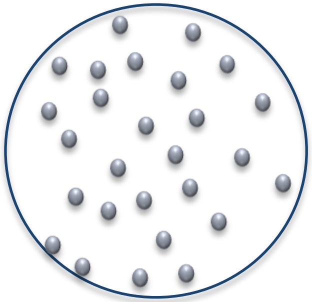
**The algorithm:**

**Summary:**

The algorithm computes the weights $w_i$ such that the power cells associated with the Diracs correspond to the preimages of the Diracs.



μ



ν

# Part. 3 Optimal Transport – the algorithm

**The [AHA] paper summary:**

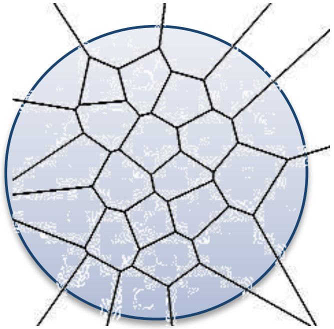
- The optimal weights minimize a convex function
- The gradient of this convex function is easy to compute

Note: the weight w(s) correspond to the Kantorovich potential $\psi(x)$
(solves a “discrete Monge-Ampere” equation)

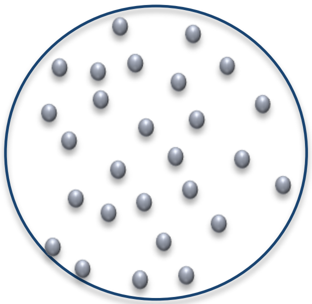
**The algorithm:**

**Summary:**

The algorithm computes the weights $w_i$ such that the power cells associated with the Diracs correspond to the preimages of the Diracs.

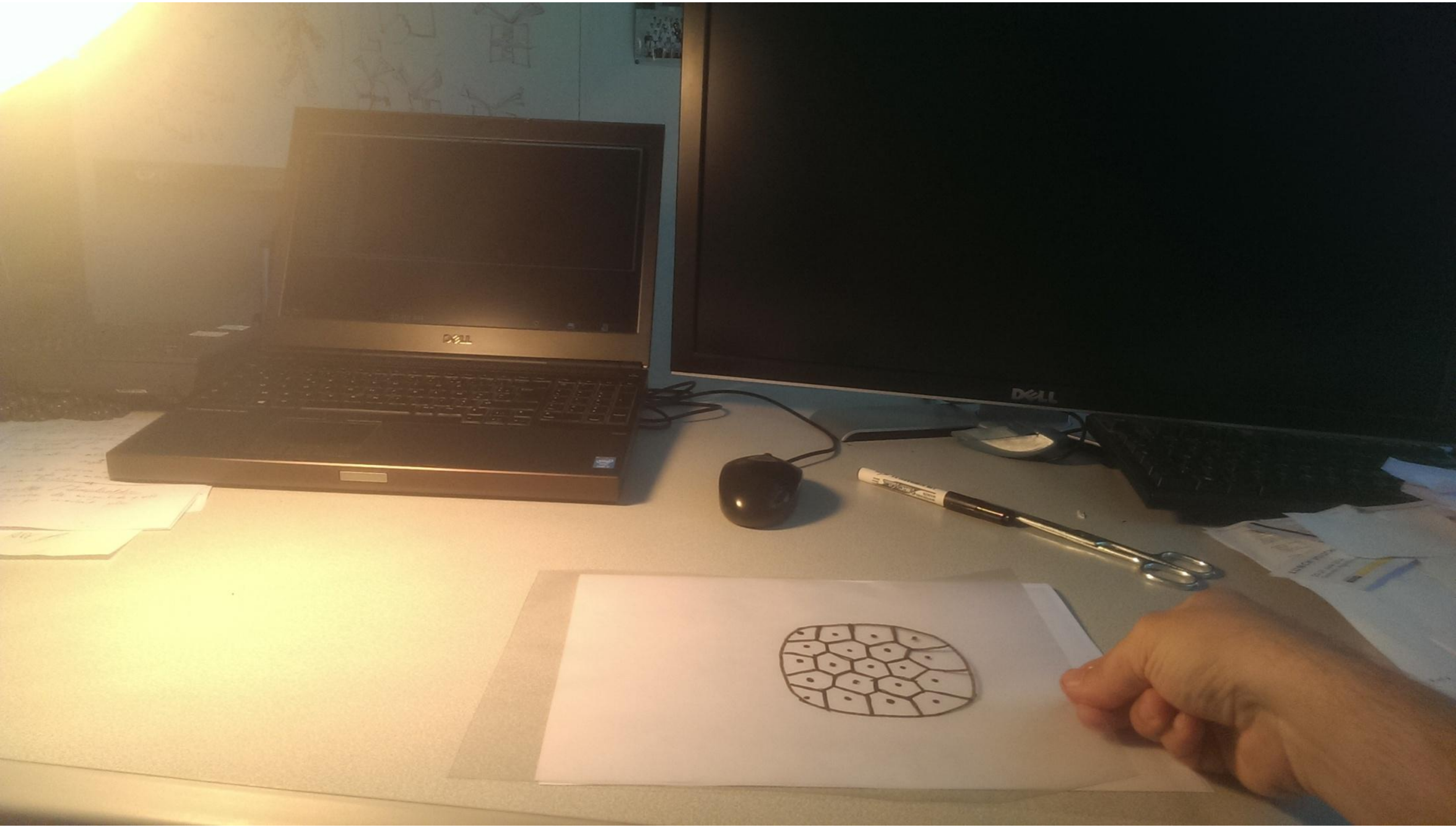


μ

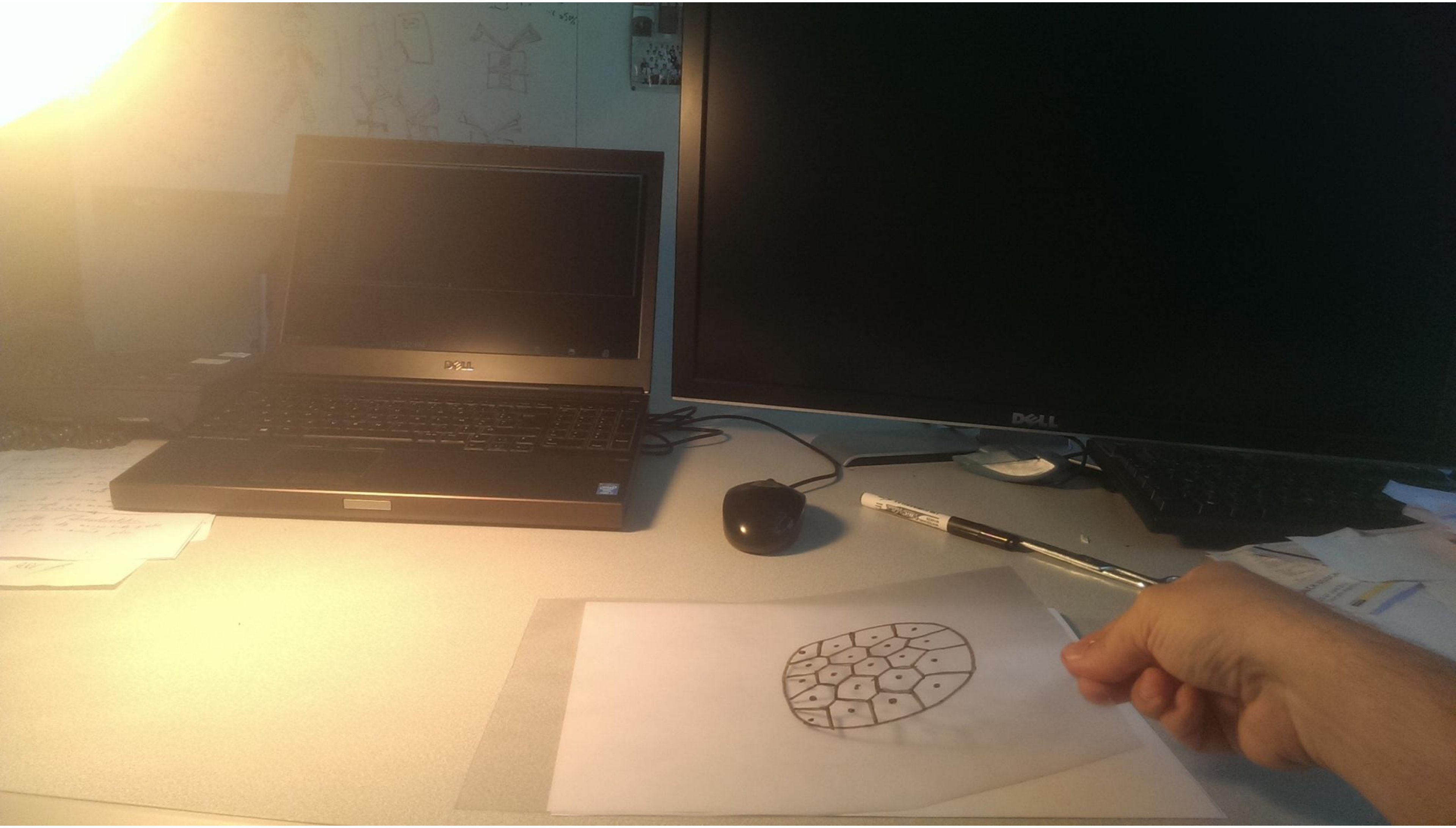


ν

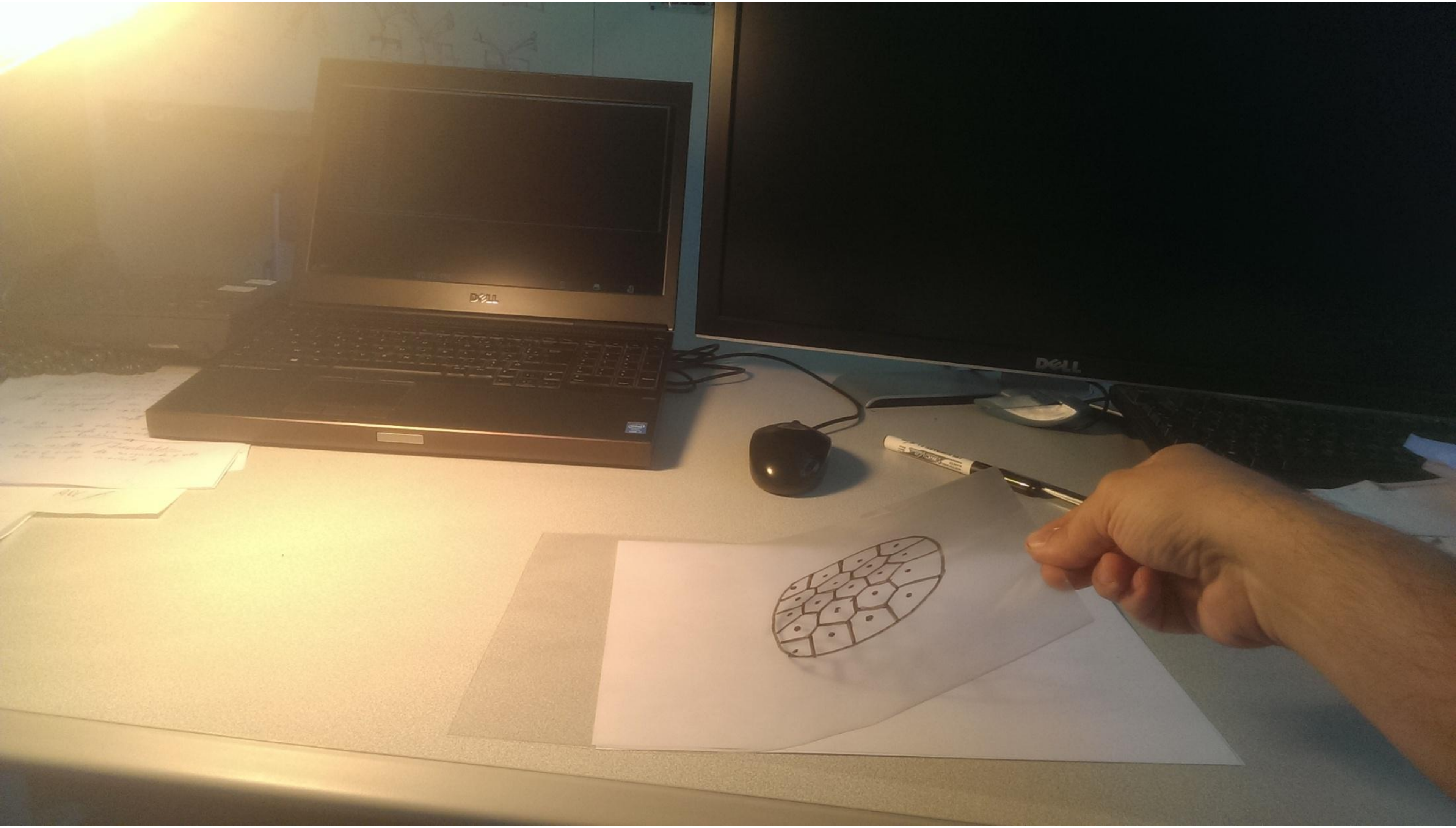
# **Part. 4** Power Diagrams & Transport



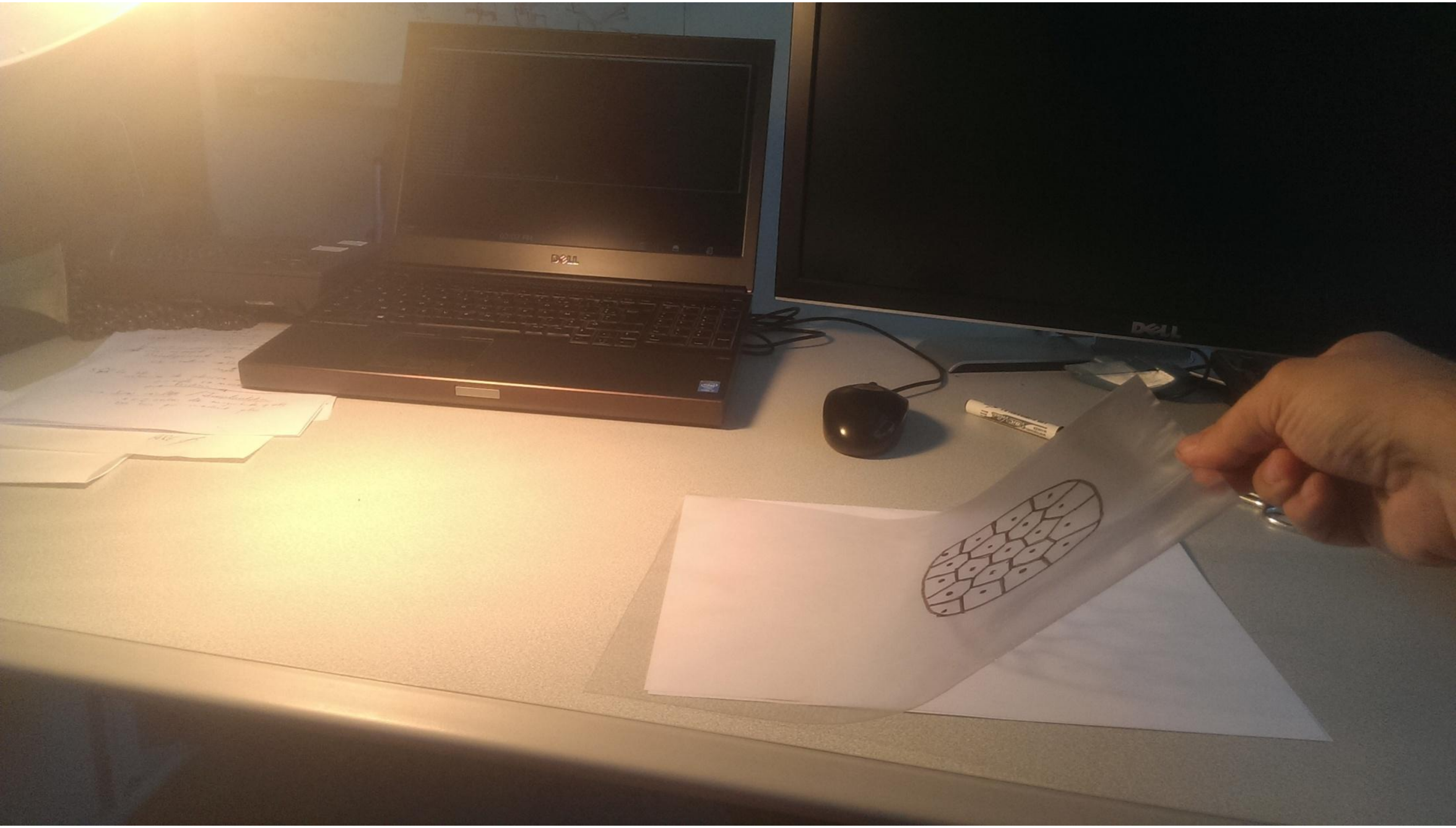
# **Part. 4** Power Diagrams & Transport



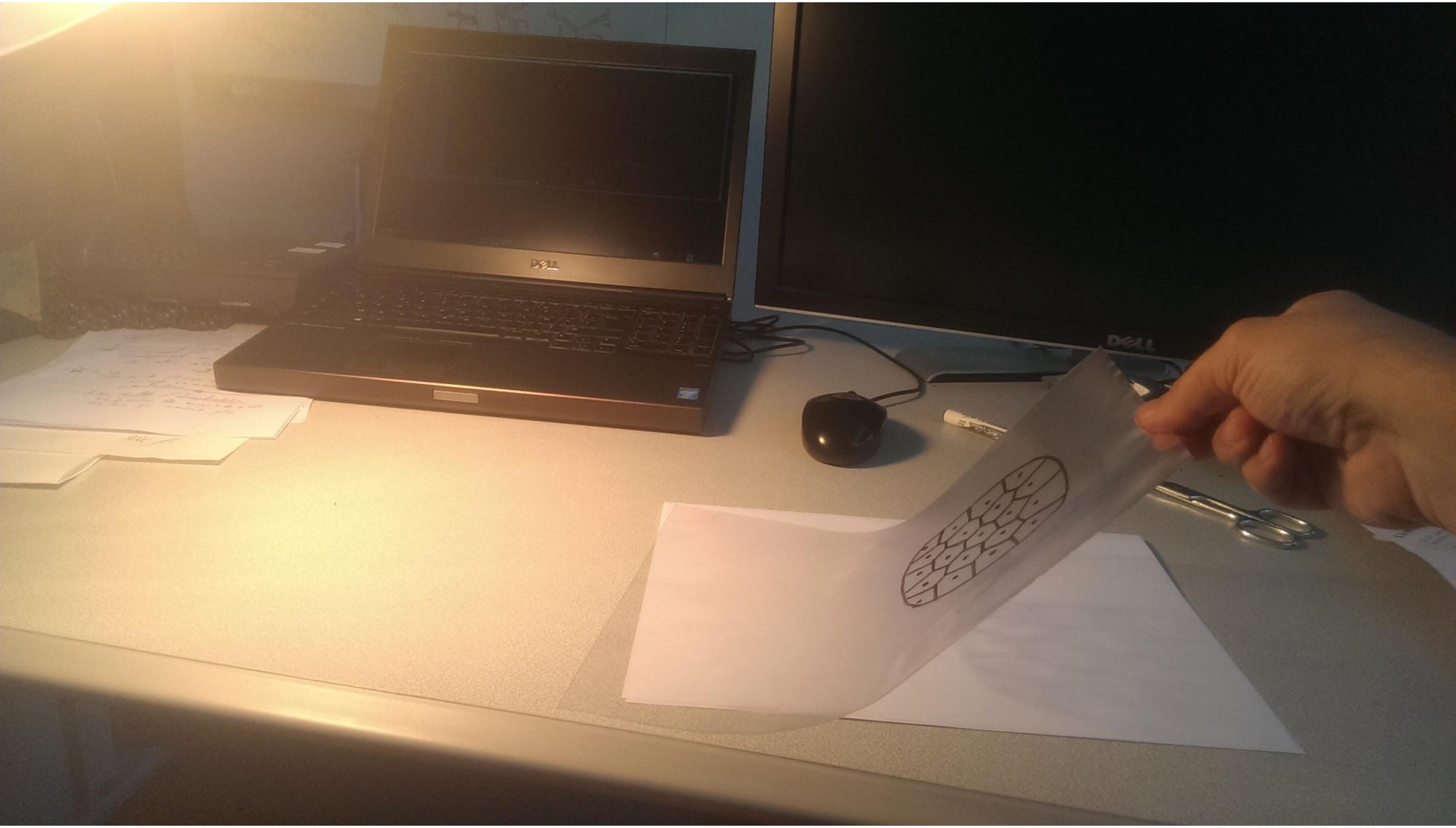
# **Part. 4** Power Diagrams & Transport



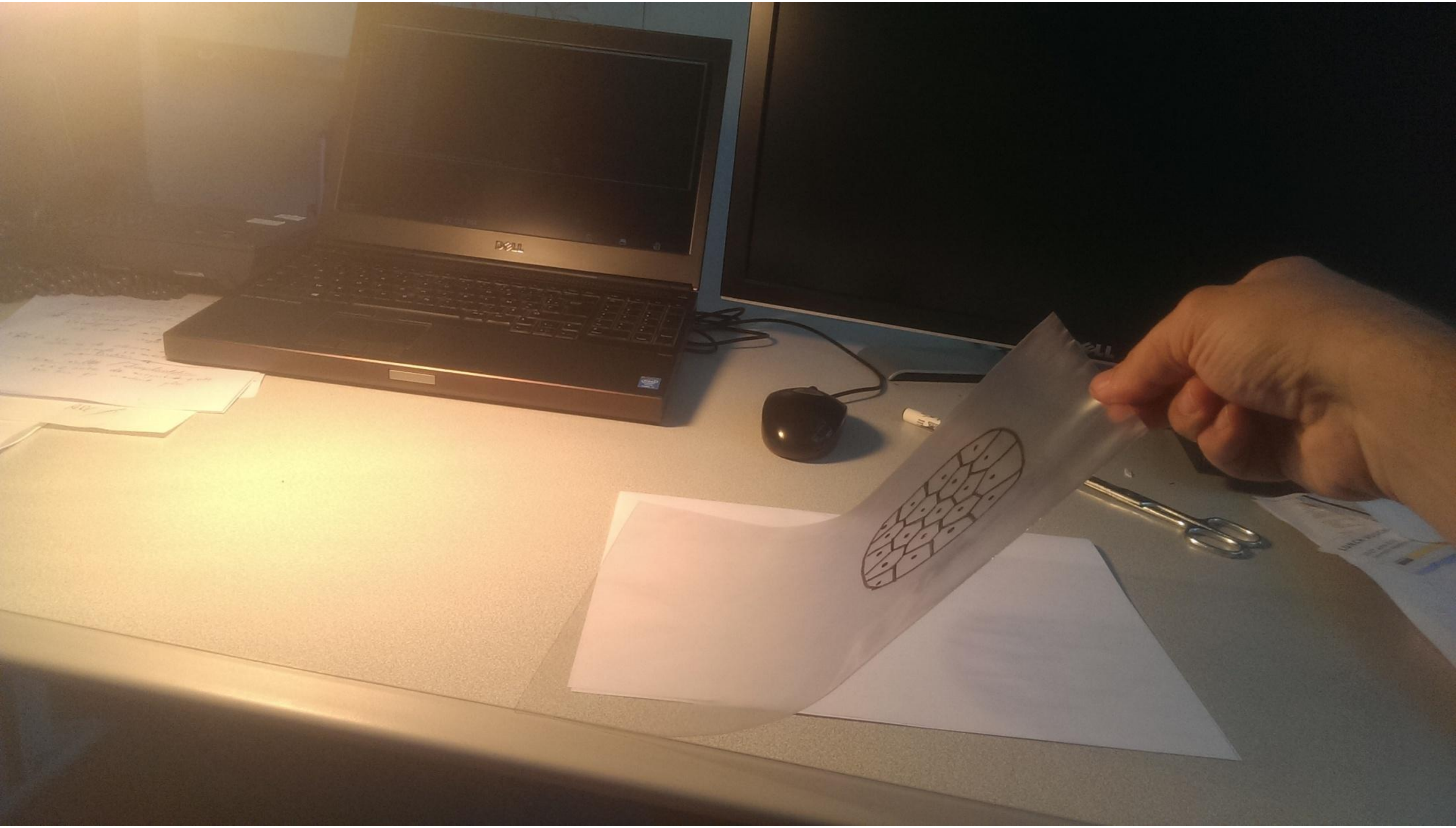
# **Part. 4** Power Diagrams & Transport



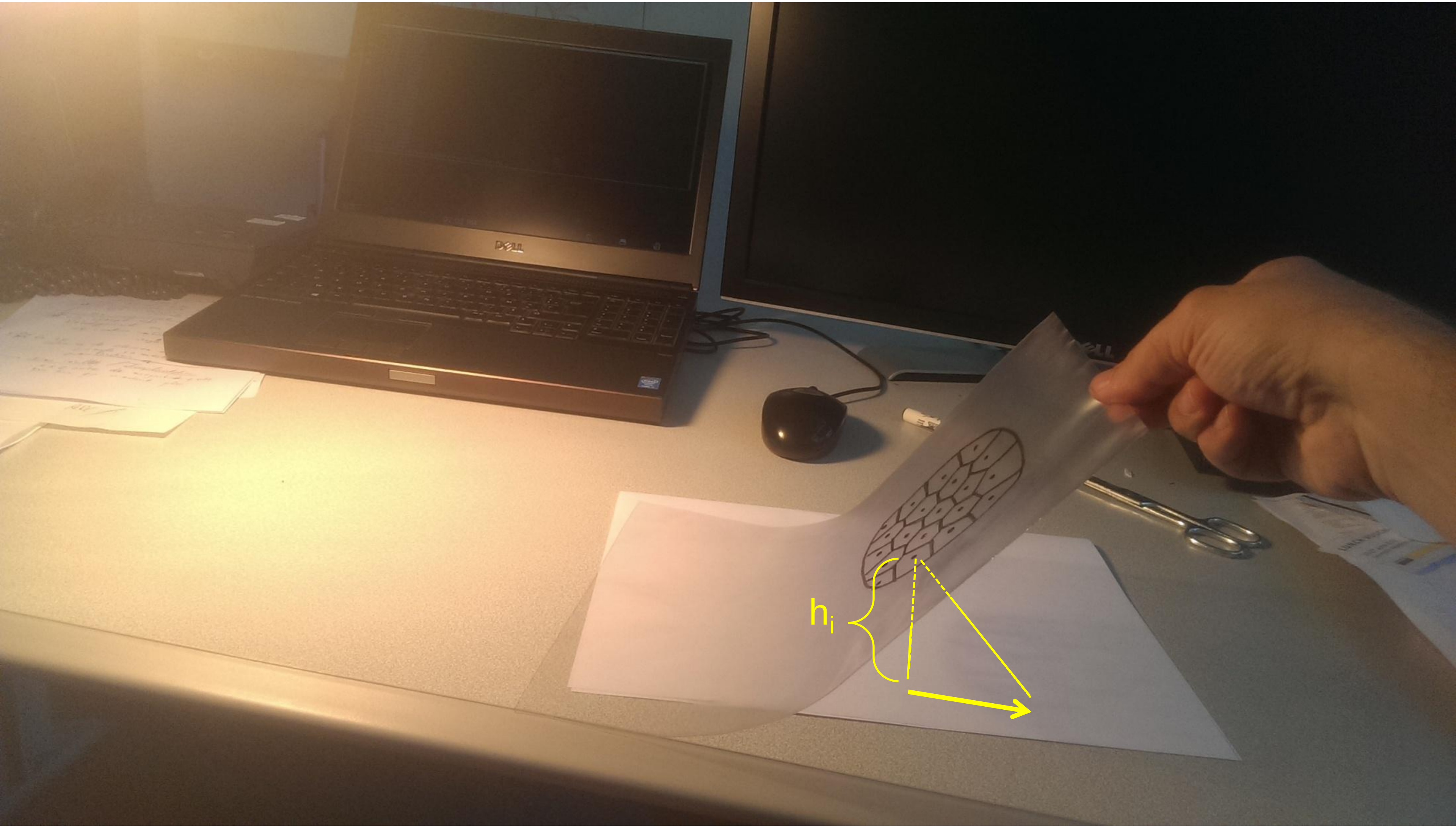
# Part. 4 Power Diagrams & Transport



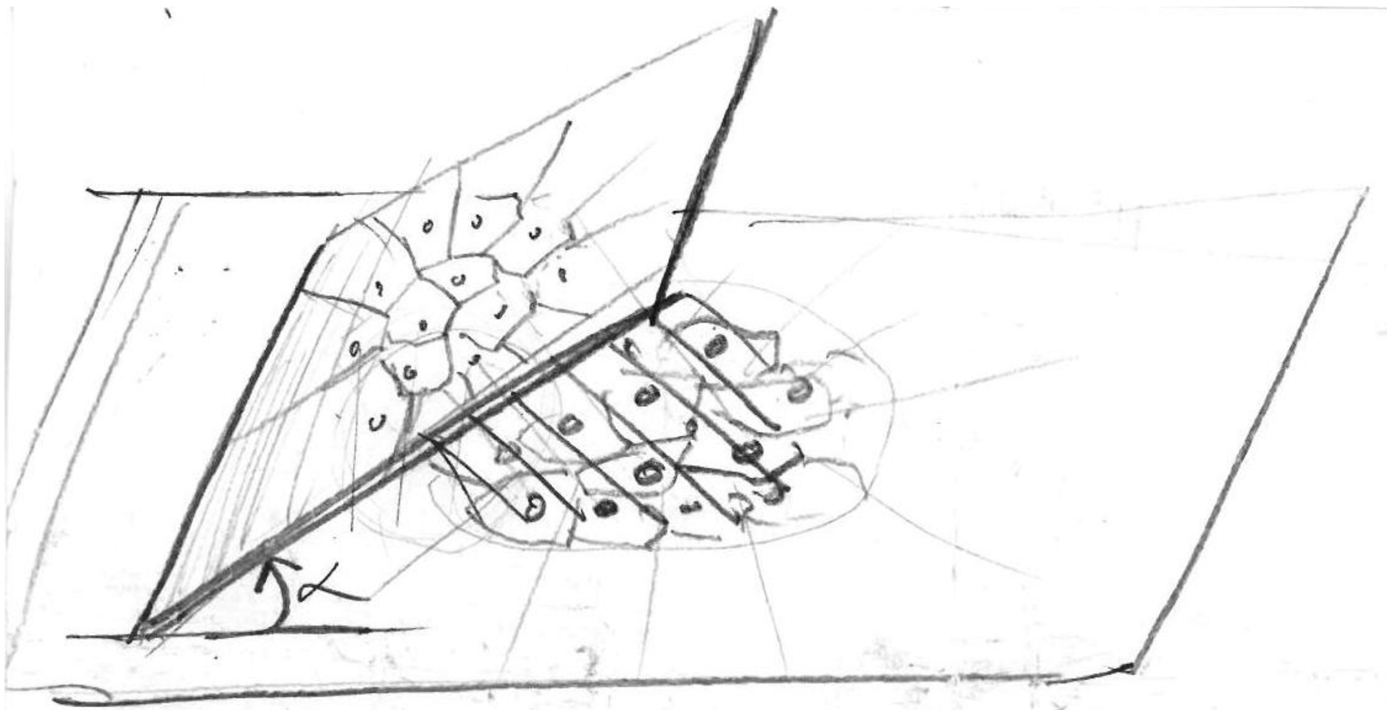
# **Part. 4** Power Diagrams & Transport



# **Part. 4** Power Diagrams & Transport



# **Part. 4** Power Diagrams & Transport

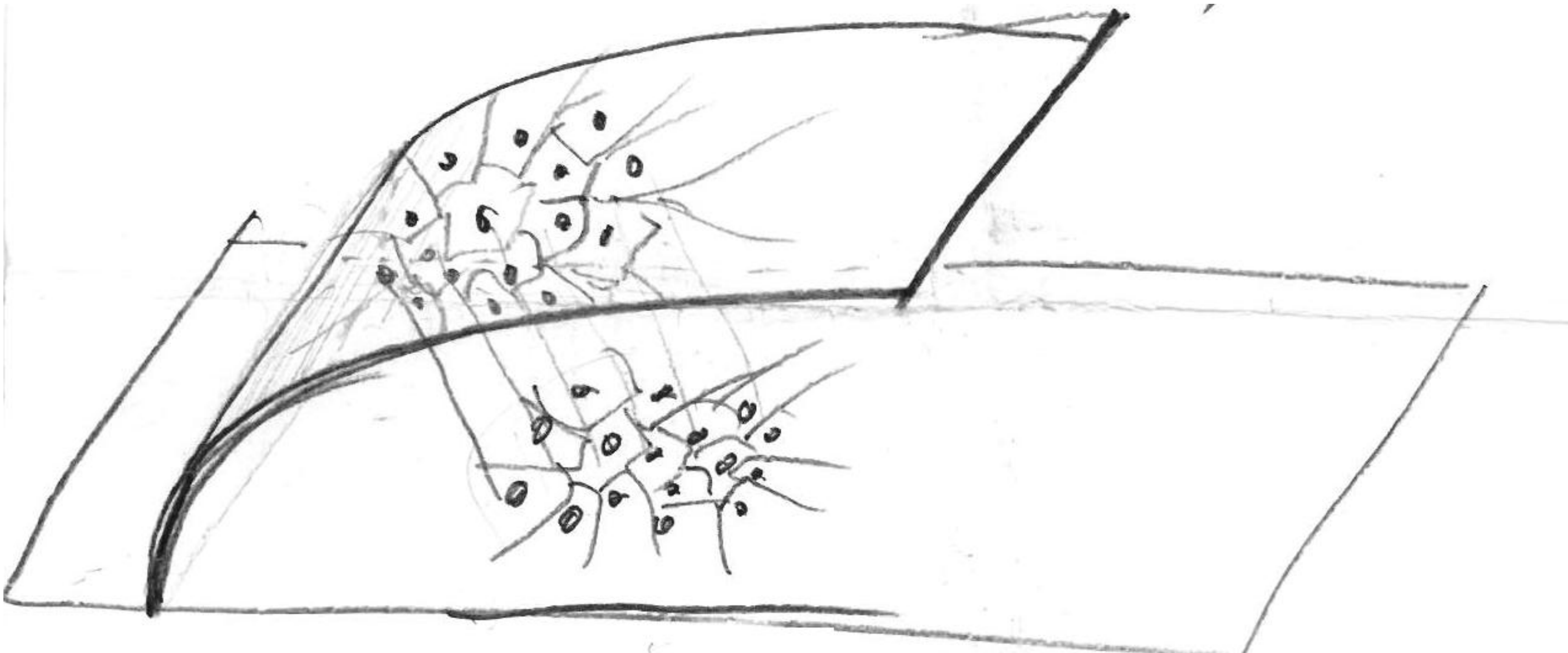



Translating a Voronoi diagram.

1st Try: linear lifting

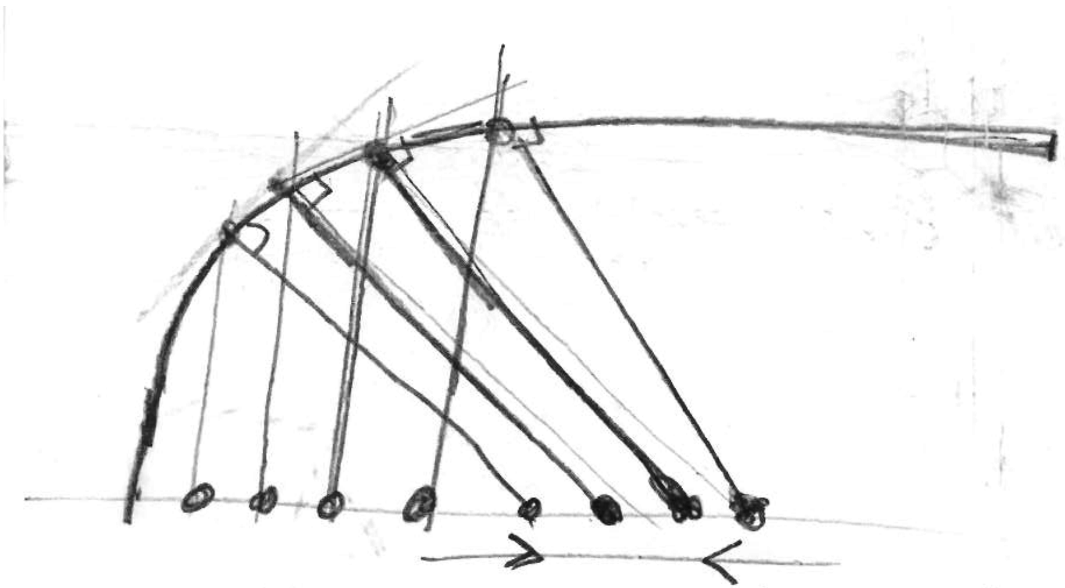
(FAIL: scales by 1/cos(α))

# **Part. 4** Power Diagrams & Transport



2nd Try : Curved lifting

# **Part. 4** Power Diagrams & Transport



"converging beams" can compensate the $1/\cos(\alpha)$ expansion by "re-concentrating" the points

# **Part. 4** Power Diagrams & Transport

(2)

$$d^2(p_i,q) \begin{vmatrix} + h_i^2 \\ - w_i \end{vmatrix} < d^2(p_j,q) \begin{matrix} + h_j^2 \\ - w_j \end{matrix} \quad \forall j$$

$$d^2(p_i, q-T) < d^2(p_j, q-T) \quad \forall j$$

$$(p_i - q + T)^2 < (p_j - q + T)^2 \quad \forall j$$

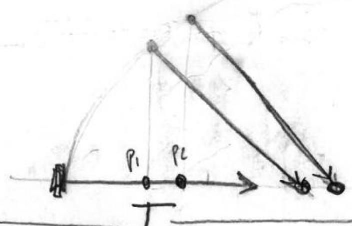
$$d^2(p_i,q) + 2T\cdot(p_i - q) + T^2 < d^2(p_j,q) + 2T(p_j - q) + T^2 \quad \forall j$$

$$d^2(p_i,q) + 2T\cdot p_i < d^2(p_j,q) + 2T\cdot p_j$$

$$\begin{matrix} h_i^2 \\ w_i \end{matrix} \Big| = -2T\cdot p_i + cte$$

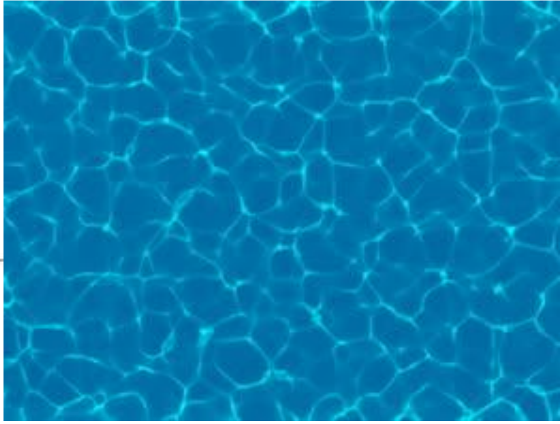
$$h_i^2 = (2T\cdot p_i + Cte)$$

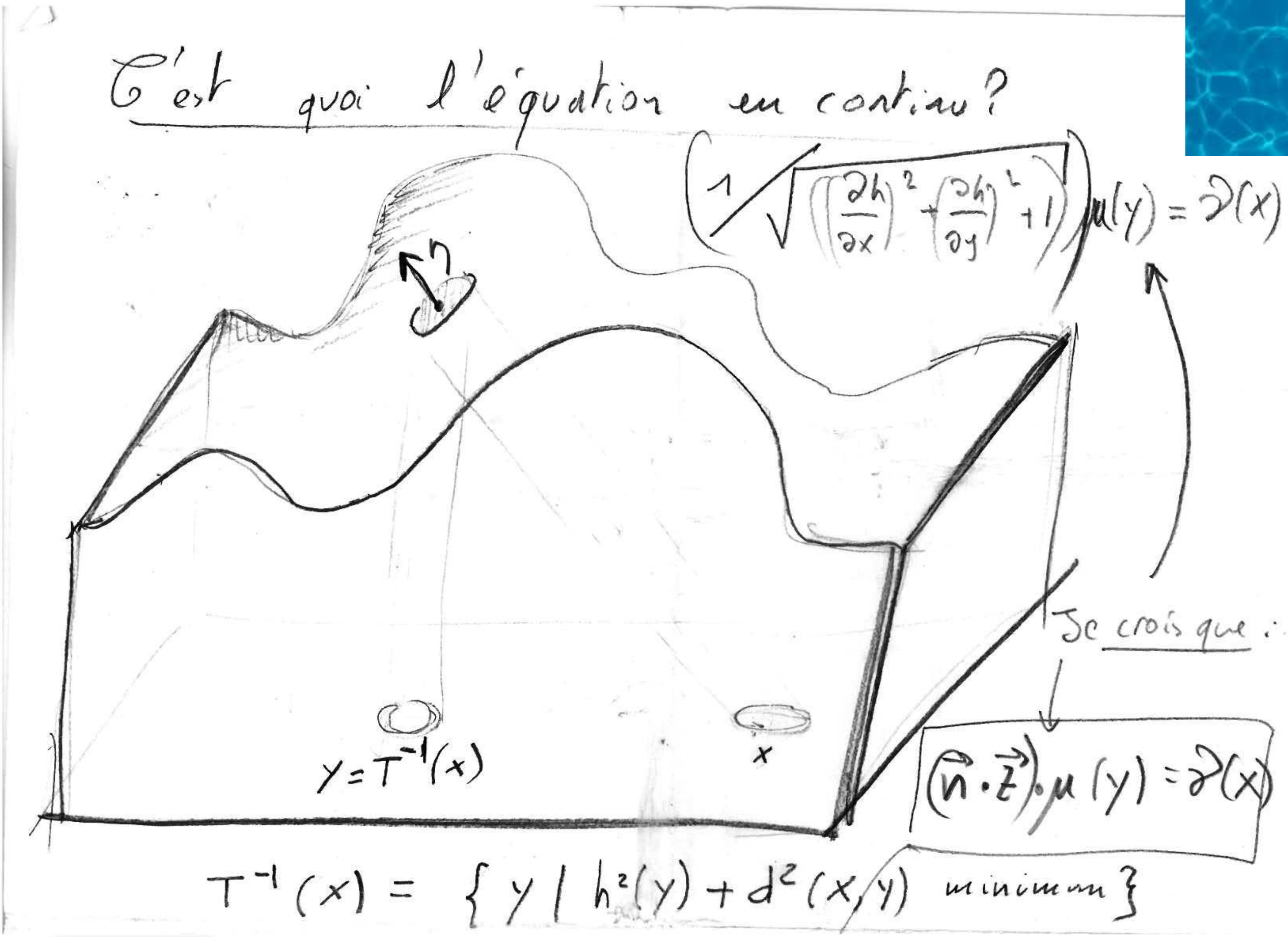
$$h_i = \sqrt{2T\cdot p_i - \min_i (T\cdot p_i)}$$



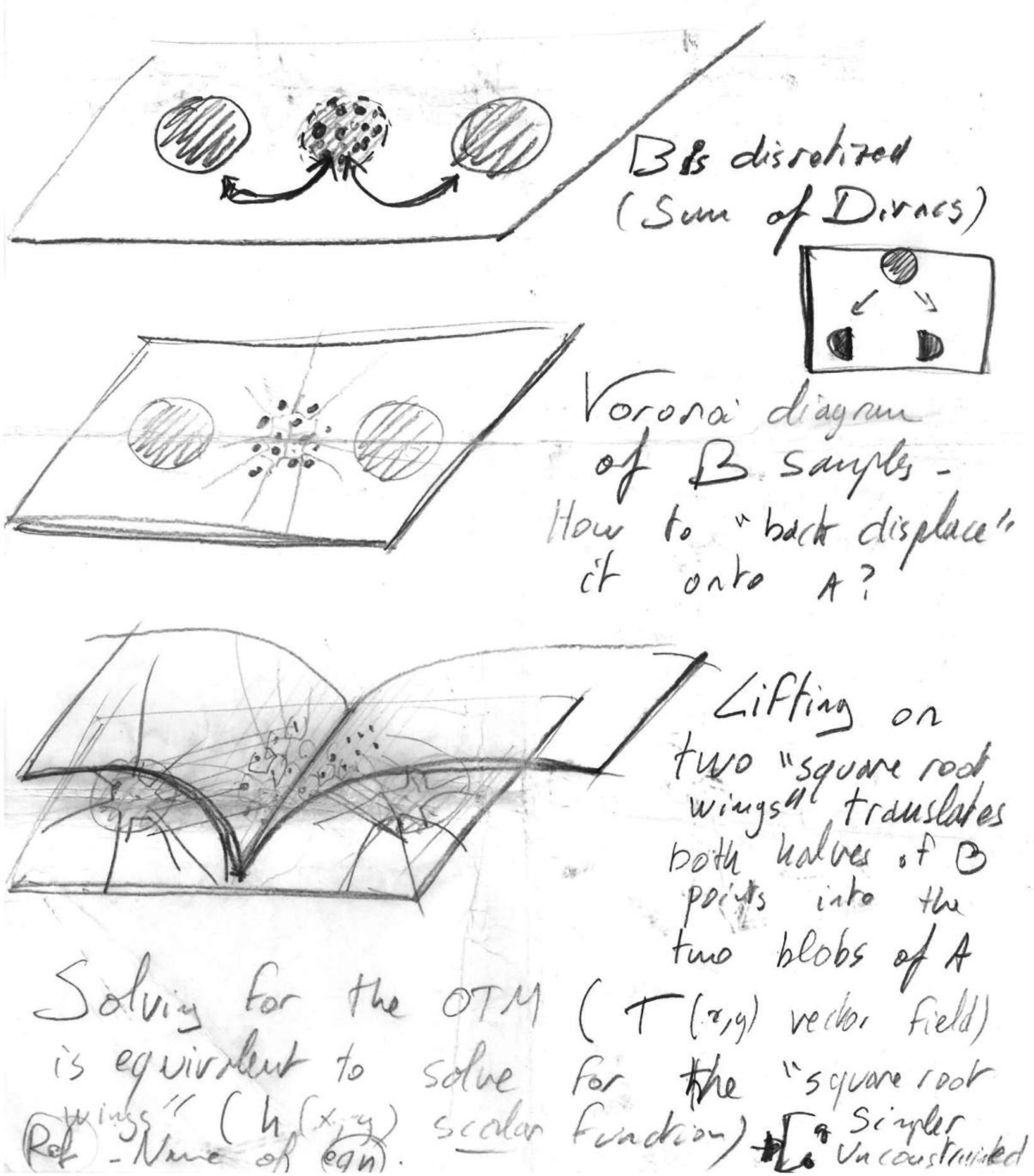

Translation d'un diagramme de Voronoi sectionnel - Relèvement en racine carrée -

# **Part. 4** Power Diagrams & Transport





# **Part. 4** Power Diagrams and Transport




Inria

# **Part. 4** Optimal Transport – 2D examples

## Numerical Experiment: *A disk becomes two disks*

# **Part. 4** Optimal Transport – 3D examples

## Numerical Experiment: *A sphere becomes a cube*

# **Part. 4** Optimal Transport – 3D examples

## Numerical Experiment: *A sphere becomes two spheres*

# **Part. 4** Optimal Transport – 3D examples

## Numerical Experiment: *Armadillo to sphere*

# **Part. 4** Optimal Transport – 3D examples

## Numerical Experiment: *Other examples*

# **Part. 4** Optimal Transport – 3D examples

## Numerical Experiment: *Varying density*

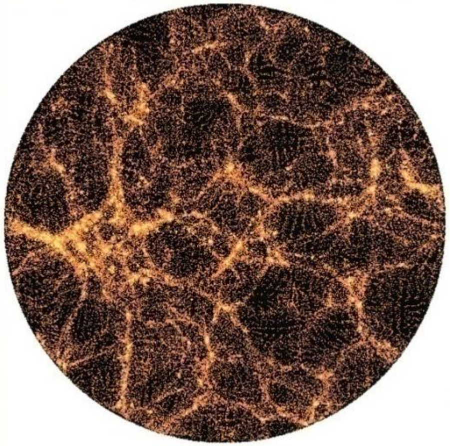
# **Part. 4** Optimal Transport – 3D examples

## Numerical Experiment: *Performances*

| nb masses | 1000 | 2000 | 5000 | 10000 | 30000 | 50000 | $10^5$ | $3 \times 10^5$ | $5 \times 10^5$ | $10^6$ |
|---|---|---|---|---|---|---|---|---|---|---|
| time (s) | 1.45 | 3.2 | 7.3 | 17.3 | 55 | 154 | 187 | 671 | 1262 | 2649 |

TABLE 4. Statistics for the Armadillo $\rightarrow$ sphere optimal transport with varying number of masses (see third row of Figure 12). Timings are given in seconds. The multi-level algorithm with BRIO pre-ordering and degree 2 regressions is used.

# **Part. 4** Optimal Transport – 3D examples
## Numerical Experiment: *Performances*



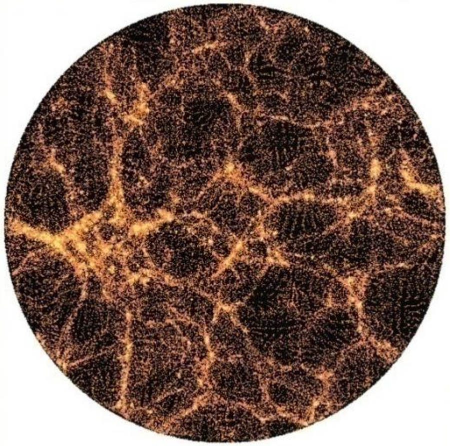
| nb masses | 1000 | 2000 | 5000 | 10000 | 30000 | 50000 | $10^5$ | $3 \times 10^5$ | $5 \times 10^5$ | $10^6$ |
|---|---|---|---|---|---|---|---|---|---|---|
| time (s) | 1.45 | 3.2 | 7.3 | 17.3 | 55 | 154 | 187 | 671 | 1262 | 2649 |

TABLE 4. Statistics for the Armadillo → sphere optimal transport with varying number of masses (see third row of Figure 12). Timings are given in seconds. The multi-level algorithm with BRIO pre-ordering and degree 2 regressions is used.

Note that a few years ago, several hours of supercomputer time were needed for computing OT with a few thousand Dirac masses, with a combinatorial algorithm in $O(n^3)$

# **Part. 4** Optimal Transport – 3D examples

## Numerical Experiment: *Performances*



| nb masses | 1000 | 2000 | 5000 | 10000 | 30000 | 50000 | $10^5$ | $3 \times 10^5$ | $5 \times 10^5$ | $10^6$ |
|---|---|---|---|---|---|---|---|---|---|---|
| time (s) | 1.45 | 3.2 | 7.3 | 17.3 | 55 | 154 | 187 | 671 | 1262 | 2649 |

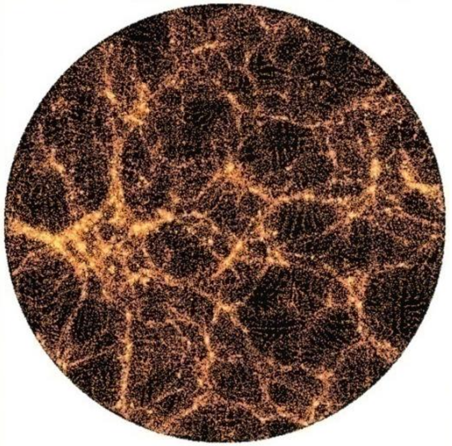
TABLE 4. Statistics for the Armadillo → sphere optimal transport with varying number of masses (see third row of Figure 12). Timings are given in seconds. The multi-level algorithm with BRIO pre-ordering and degree 2 regressions is used.

Note that a few years ago, several hours of supercomputer time were needed for computing OT with a few thousand Dirac masses, with a combinatorial algorithm in $O(n^3)$

With the semi-discrete algorithm, it takes **less than 10 seconds** on my laptop

# **Part. 4** Optimal Transport – 3D examples

## Numerical Experiment: *Performances*



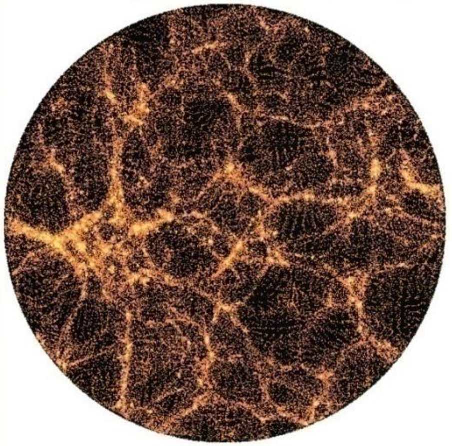
| nb masses | 1000 | 2000 | 5000 | 10000 | 30000 | 50000 | $10^5$ | $3 \times 10^5$ | $5 \times 10^5$ | $10^6$ |
|---|---|---|---|---|---|---|---|---|---|---|
| time (s) | 1.45 | 3.2 | 7.3 | 17.3 | 55 | 154 | 187 | 671 | 1262 | 2649 |

TABLE 4. Statistics for the Armadillo → sphere optimal transport with varying number of masses (see third row of Figure 12). Timings are given in seconds. The multi-level algorithm with BRIO pre-ordering and degree 2 regressions is used.

Note that a few years ago, several hours of supercomputer time were needed for computing OT with a few thousand Dirac masses, with a combinatorial algorithm in $O(n^3)$

In the semi-discrete setting, my 3D version of Merigot's multigrid algorithm computes OT for **1 million Dirac masses** in less than 1 hour on a laptop PC

# **Part. 4** Optimal Transport – 3D examples
## Numerical Experiment: *Performances*



| nb masses | 1000 | 2000 | 5000 | 10000 | 30000 | 50000 | $10^5$ | $3 \times 10^5$ | $5 \times 10^5$ | $10^6$ |
|---|---|---|---|---|---|---|---|---|---|---|
| time (s) | 1.45 | 3.2 | 7.3 | 17.3 | 55 | 154 | 187 | 671 | 1262 | 2649 |

TABLE 4. Statistics for the Armadillo → sphere optimal transport with varying number of masses (see third row of Figure 12). Timings are given in seconds. The multi-level algorithm with BRIO pre-ordering and degree 2 regressions is used.

Note that a few years ago, several hours of supercomputer time were needed for computing OT with a few thousand Dirac masses, with a combinatorial algorithm in $O(n^3)$

In the semi-discrete setting, my 3D version of Merigot's multigrid algorithm computes OT for **1 million Dirac masses** in less than 1 hour on a laptop PC

Even much faster convergence can probably be reached with a true Newton solver (and several acceleration tricks), still investigating…

# Other topics

- **Euler equation in more complicated setting**:
  [Merigot & Mirebeau]

- **Using semi-discrete OT to solve other PDEs**
  [Benamou, Carlier, Merigot , Oudet]

- **New fluid simulation methods**
  “power particles” [DeGoes et.al]

# Conclusions – Open questions

* **Connections with physics, Legendre transform and <u>entropy</u> ?**
  **[Cuturi & Peyré]** – regularized discrete optimal transport – why does it work ?
    Hint 1: Minimum action principle subject to conservation laws
    Hint 2: Entropy = dual of temperature ; Legendre = Fourier[(+,*) → (Max,+)]…

* **More continuous numerical algorithms ?**
  **[Benamou & Brenier]** fluid dynamics point of view **–** very elegant, but 4D problem !!
  FEM-type adaptive discretization of the subdifferential (graph of T) ?

* **Can we characterize OT in other semi-discrete settings ?**
  measures supported on unions of spheres
  piecewise linear densities

* **Connections with computational geometry ?**
  Singularity set **[Figalli]** = set of points where T is discontinuous
  Looks like a “mutual power diagram”, anisotropic Voronoi diagrams

# Conclusions - References

## Some references (that this presentation is based on)

A Multiscale Approach to Optimal Transport,
**Quentin Mérigot**, Computer Graphics Forum, 2011

Variational Principles for Minkowski Type Problems, Discrete Optimal Transport, and Discrete Monge-Ampere Equations
**Xianfeng Gu, Feng Luo, Jian Sun, S.-T. Yau**, ArXiv 2013

Minkowski-type theorems and least-squares clustering
**AHA! (Aurenhammer, Hoffmann, and Aronov),** SIAM J. on math. ana. 1998

Topics on Optimal Transportation, 2003
Optimal Transport Old and New, 2008
**Cédric Villani**

**Jean-David Benamou & Yann Brenier**
a fluid formulation of Optimal Transportation, 2000

**Laudau and Lifschitz** – Course of Theoretical Physics – Volumes I and III

# Conclusions - References

## Other references

Polar factorization and monotone rearrangement of vector-valued functions
**Yann Brenier**, Comm. On Pure and Applied Mathematics, June 1991

A computational fluid mechanics solution of the Monge-Kantorovich mass transfer problem, **J.-D. Benamou, Y. Brenier**, Numer. Math. 84 (2000), pp. 375-393

**Pogorelov, Alexandrov** – Gradient maps, Minkovsky problem (older than AHA paper, some overlap, in slightly different context, formalism used by Gu & Yau)

**Rockafeller** – Convex optimization – Theorem to switch inf(sup()) – sup(inf()) with convex functions (used to justify Kantorovich duality)

New textbook: **Filippo Santambrogio** – Optimal Transport for Applied Mathematician, Calculus of Variations, PDEs and Modeling – Jan 15, 2015

# Online resources

All the sourcecode/documentation available from:
alice.loria.fr/software/geogram

Computes semi-discrete OT in 3D
Scales up to millions Dirac masses on a laptop

**L**., A numerical algorithm for semi-discrete L2 OT in 3D,
ESAIM Math. Modeling and Analysis, accepted
(draft: http://arxiv.org/abs/1409.1279 <= to be fixed: bug in MA equation in this version, fixed in M2AN journal version)

Downloads: alice.loria.fr/software GEOGRAM & GRAPHITE
Video of course on OT: www.loria.fr/~levy

# Acknowledgements


Funding: European Research Council & ANR ("french NSF")

GOODSHAPE ERC-StG-205693
VORPALINE ERC-PoC-334829
ANR MORPHO, ANR BECASIM

New project: EXPLORAGRAM (Inria) with Q. Mérigot and J.-D. Benamou

*Quentin Merigot, Yann Brenier, Boris Thibert, Emmanuel Maitre, Jean-David Benamou, Filippo Santambrogio, Edouard Oudet, Hervé Pajot.*
ANR TOMMI, ANR GEOMETRYA