Mathématiques - Informatique

Ecole Jeunes Chercheurs GDR IM
Géométrie Numérique

Partie I: Transport Optimal (Lundi) *B. Lévy*
Partie II: Résolution de systèmes linéaires (Jeudi) *N. Ray*

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

OVERVIEW

Part. 1. Goals and Motivations

Part. 2. Introduction to Optimal Transport

Part. 3. Semi-Discrete Optimal Transport

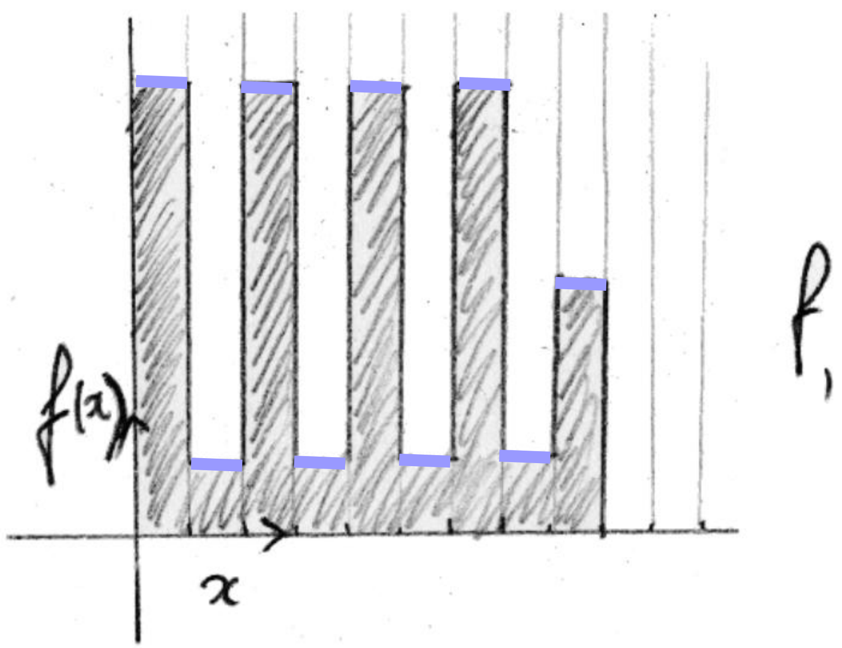
Part. 4. Applications in Computational Physics

1

Goals and Motivations

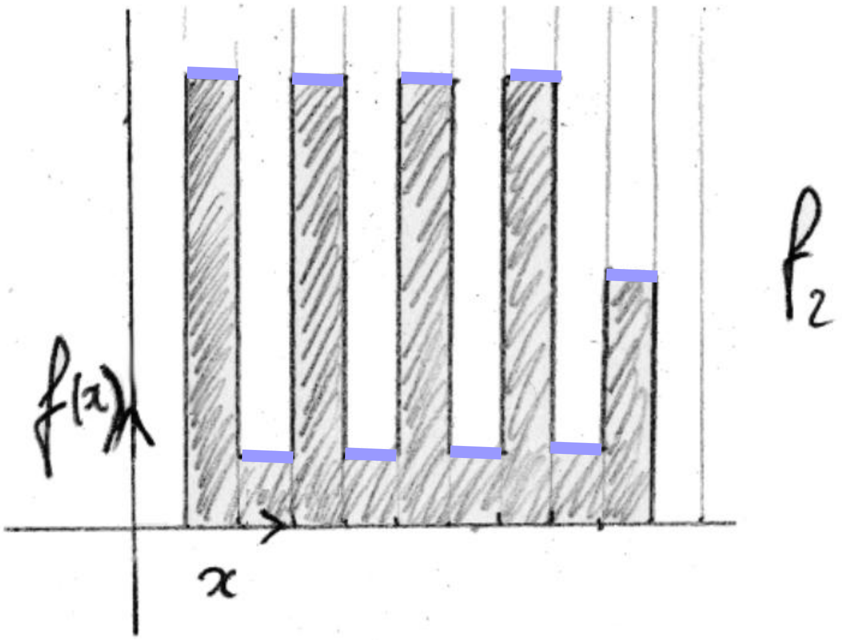
Part. 1 Optimal Transport

Measuring distances between functions



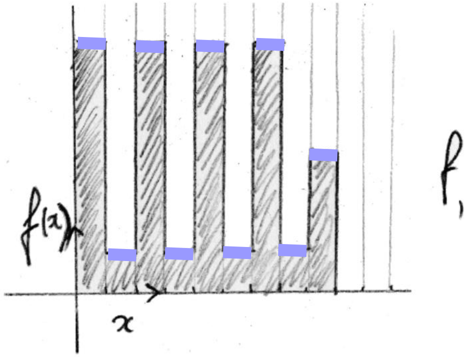
Part. 1 Optimal Transport

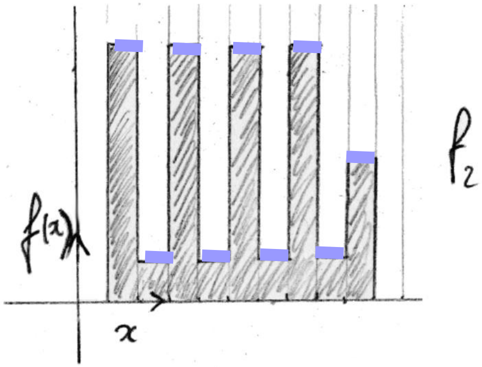
Measuring distances between functions



Part. 1 Optimal Transport

Measuring distances between functions

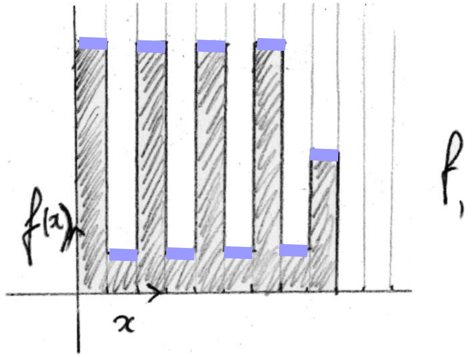


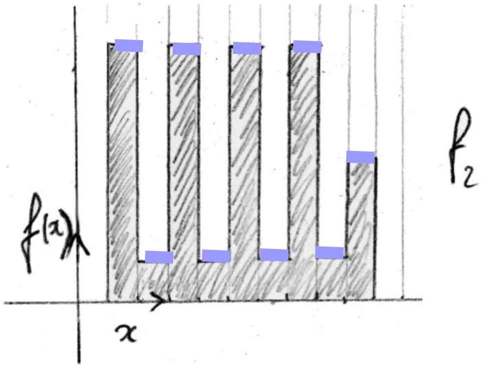


$$d_{L_2}(f_1, f_2) = \int (f_1(x) - f_2(x))^2 \, dx$$

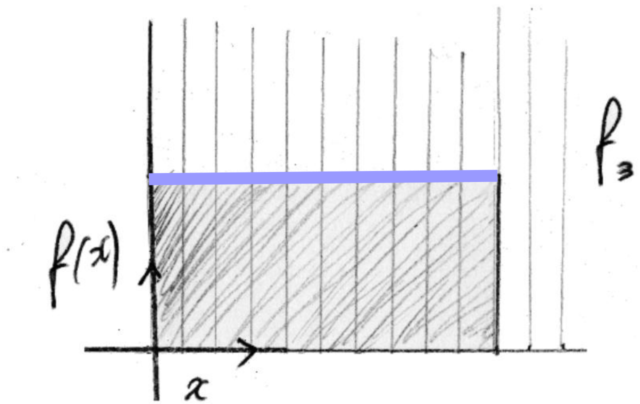
Part. 1 Optimal Transport

Measuring distances between function

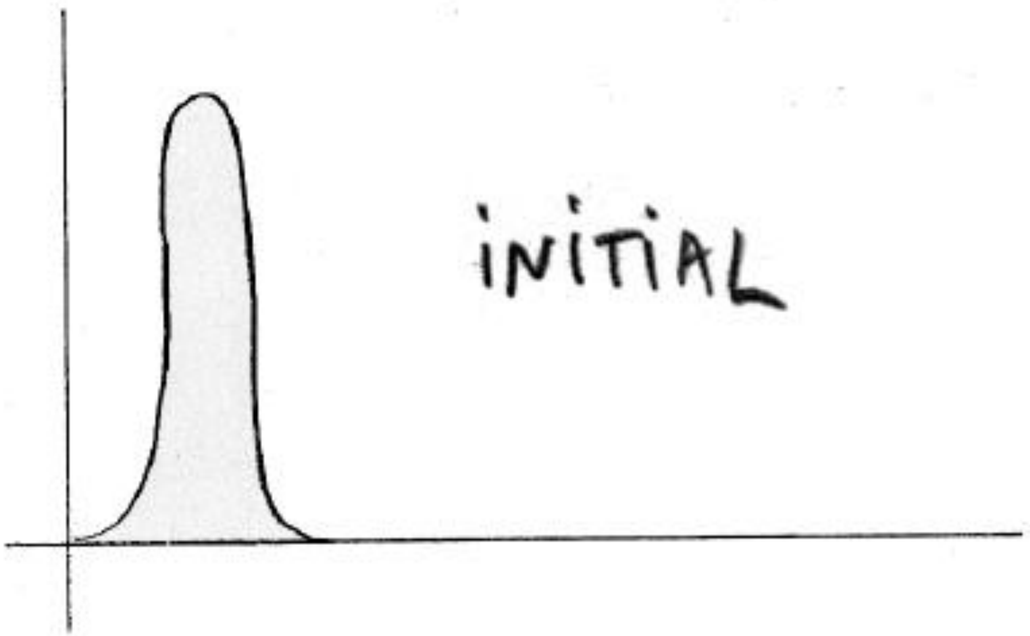




$$d_{L_2}(f_1, f_2) = \int (f_1(x) - f_2(x))^2 \, dx$$



Part. 1 Optimal Transport
Interpolating functions

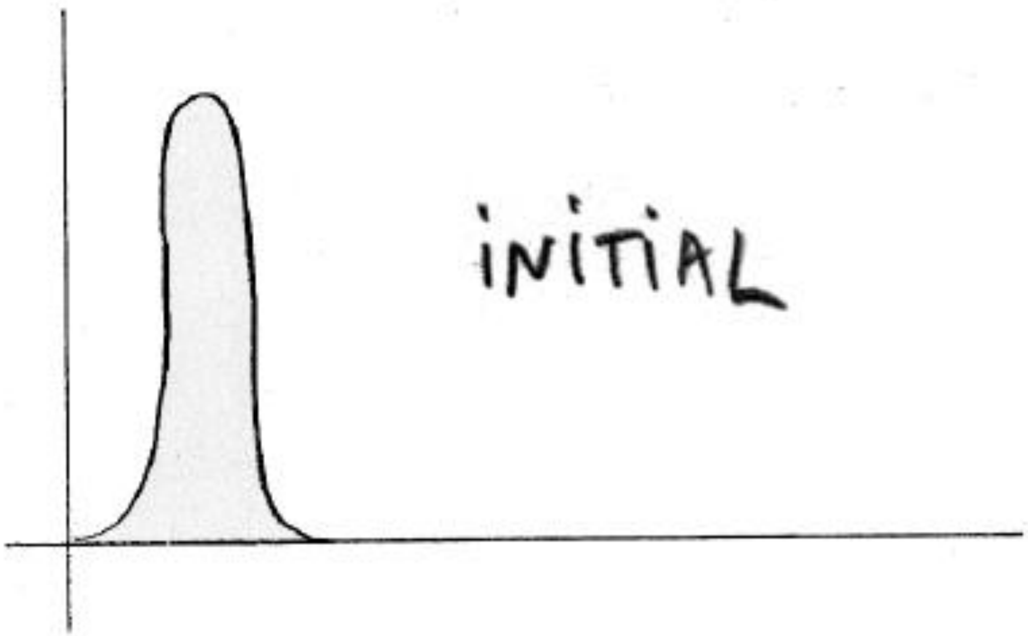


Part. 1 Optimal Transport

Interpolating functions

Part. 1 Optimal Transport

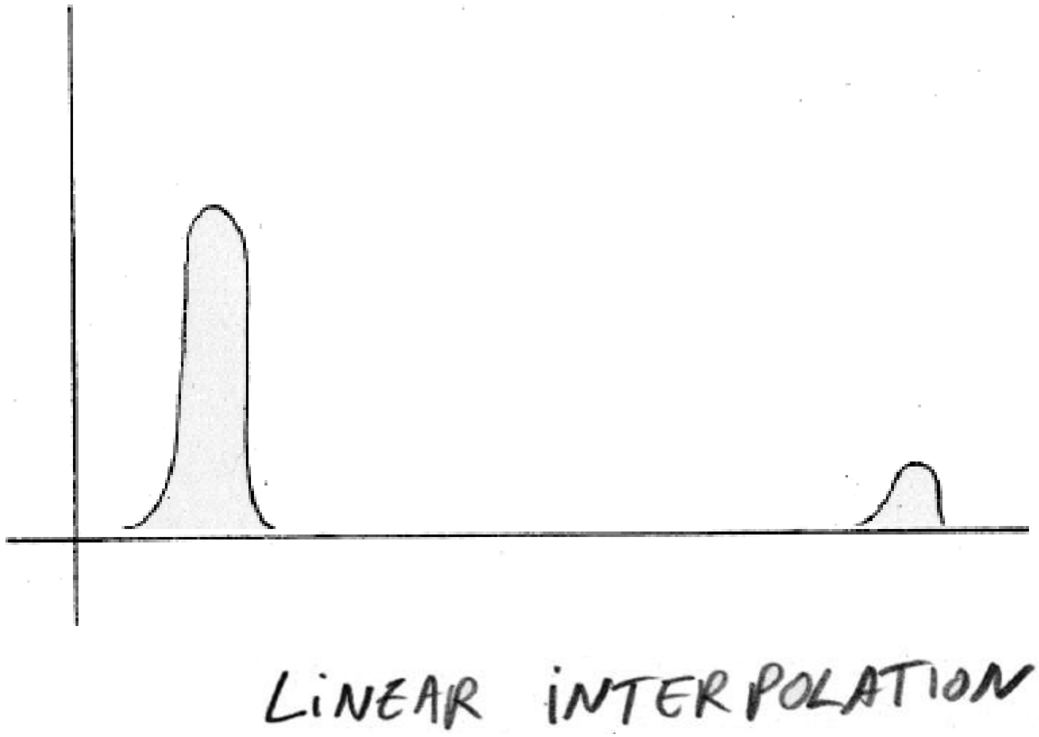
Interpolating functions



LiNEAR iNTERPOLATiON

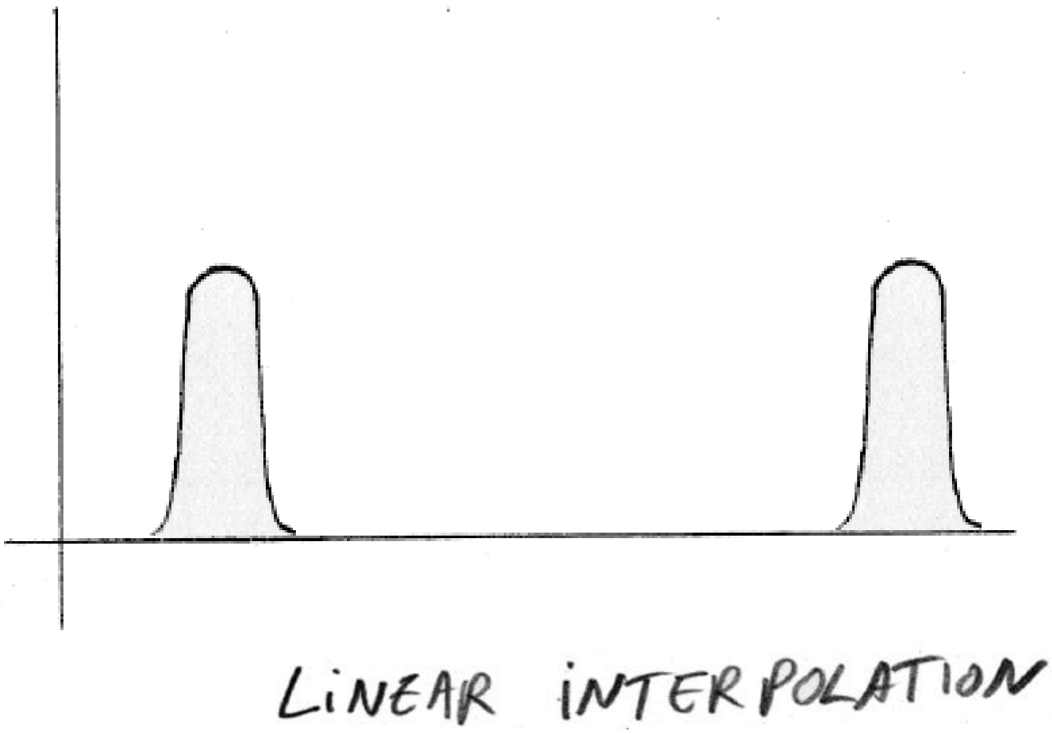
Part. 1 Optimal Transport

Interpolating functions



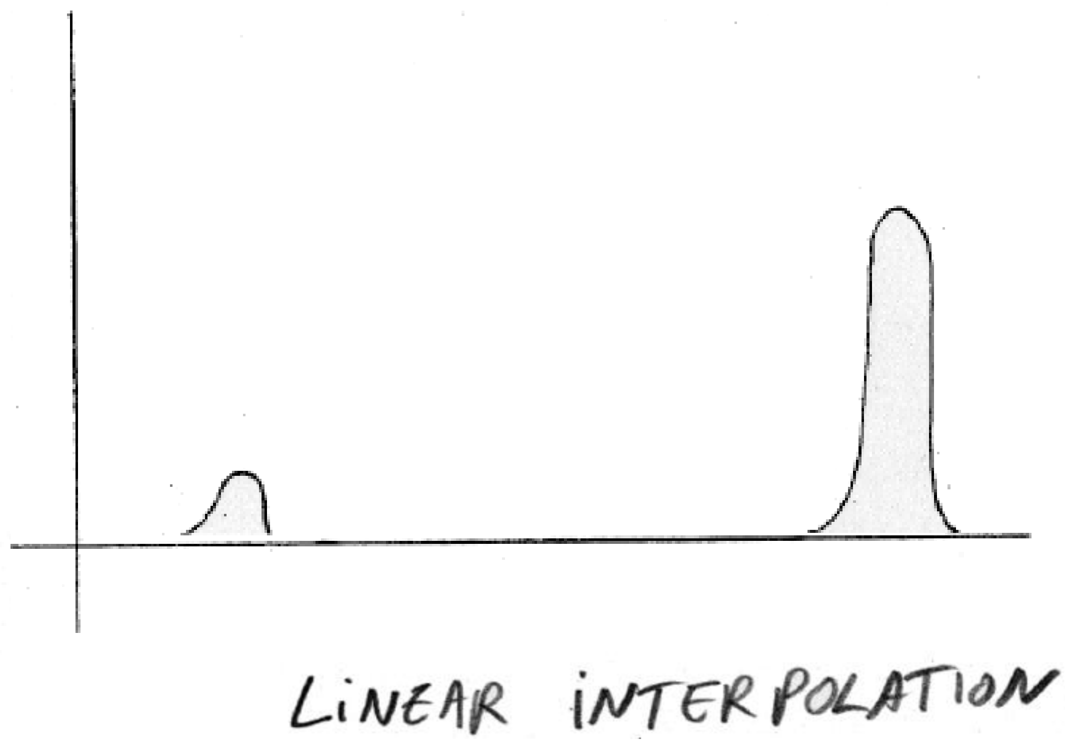
Part. 1 Optimal Transport

Interpolating functions



Part. 1 Optimal Transport

Interpolating functions

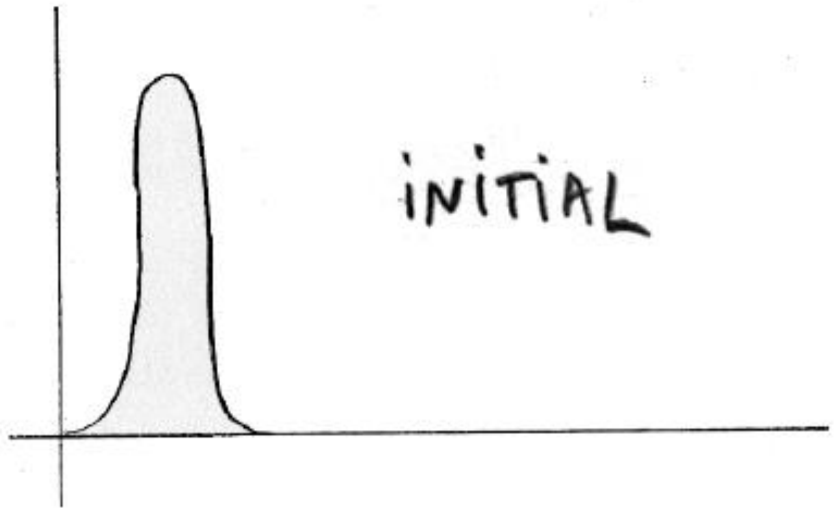


Interpolating functions

LINEAR INTERPOLATION

Part. 1 Optimal Transport

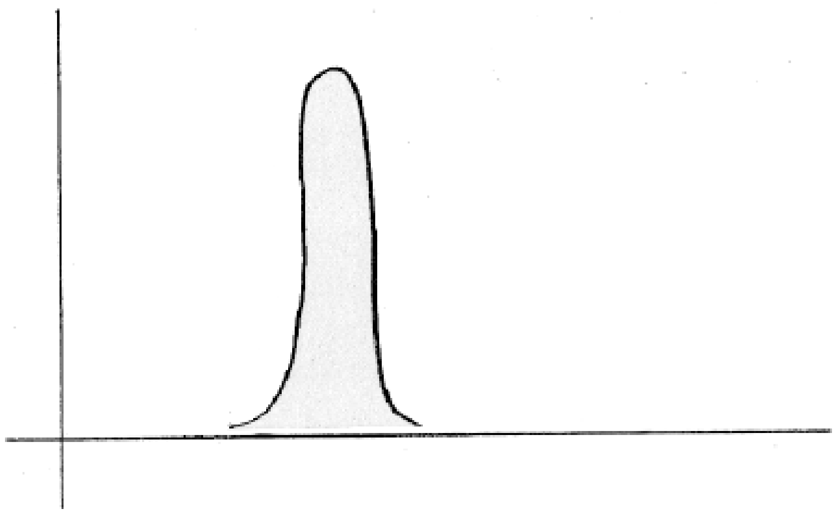
Interpolating functions



DISPLACEMENT INTERPOLATION

Part. 1 Optimal Transport

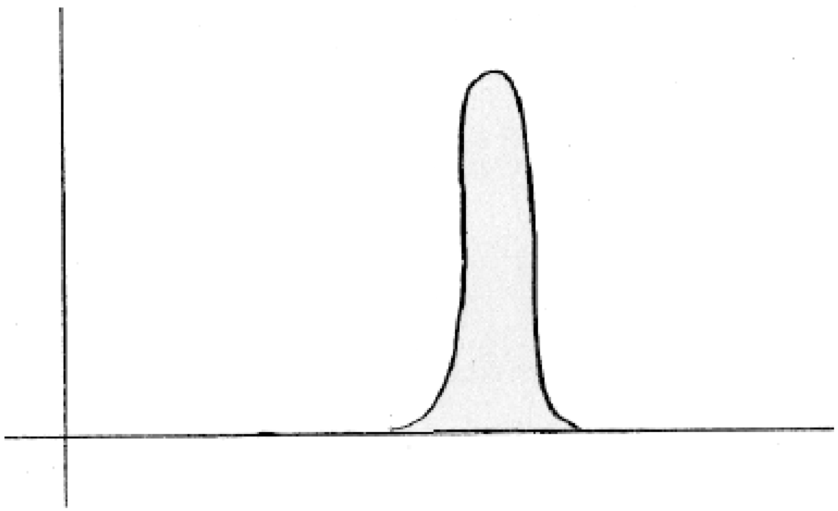
Interpolating functions



DISPLACEMENT INTERPOLATION

Part. 1 Optimal Transport

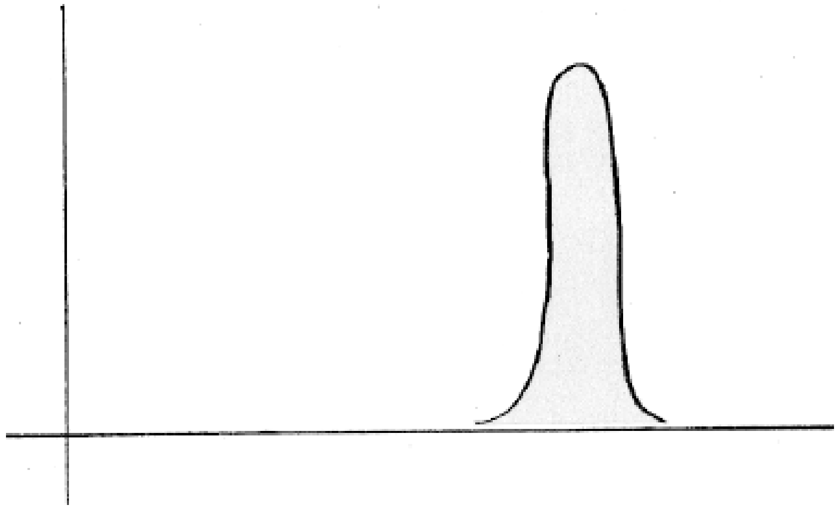
Interpolating functions



DISPLACEMENT INTERPOLATION

Part. 1 Optimal Transport

Interpolating functions



DISPLACEMENT INTERPOLATION

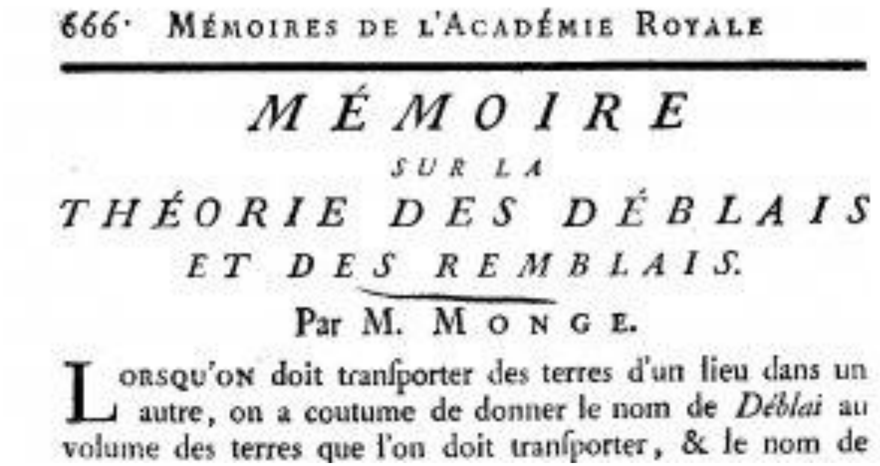
Part. 1 Optimal Transport

Interpolating functions

DISPLACEMENT INTERPOLATION

Part. 1 Optimal Transport

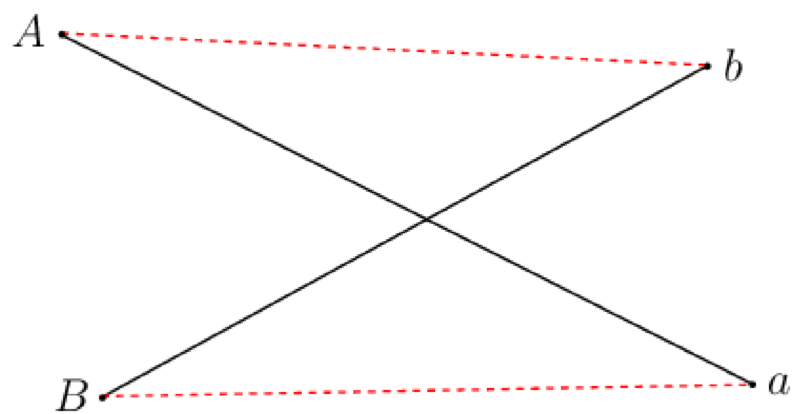
Gaspard Monge - 1784


666· MÉMOIRES DE L'ACADÉMIE ROYALE

MÉMOIRE SUR LA THÉORIE DES DÉBLAIS ET DES REMBLAIS.

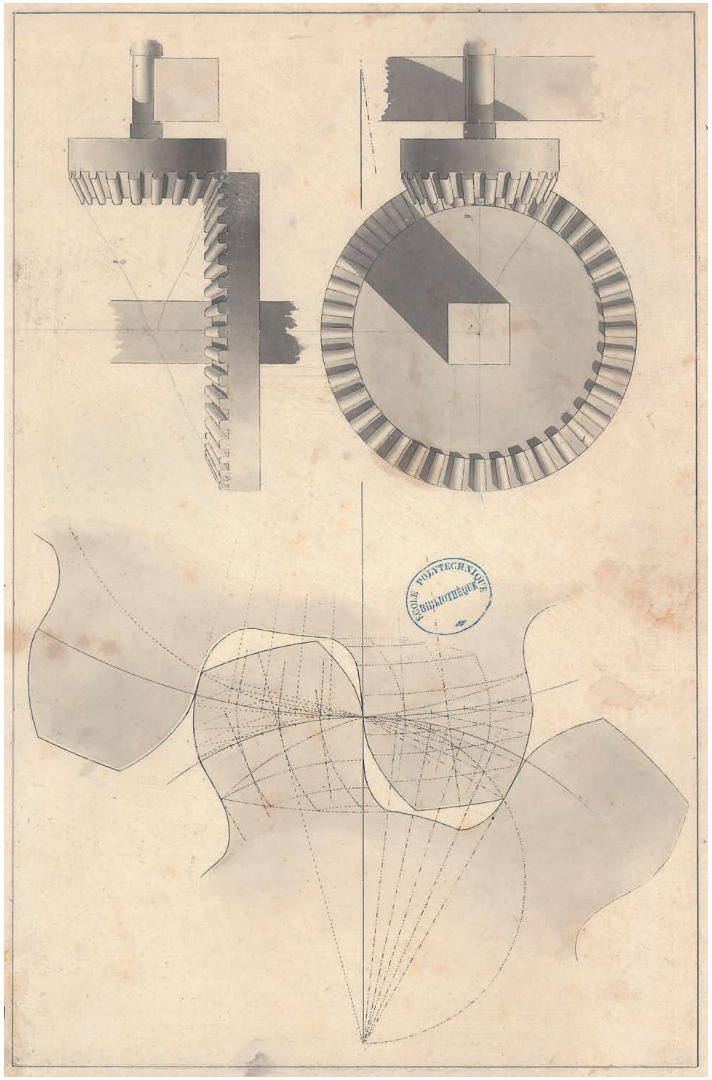
Par M. MONGE.

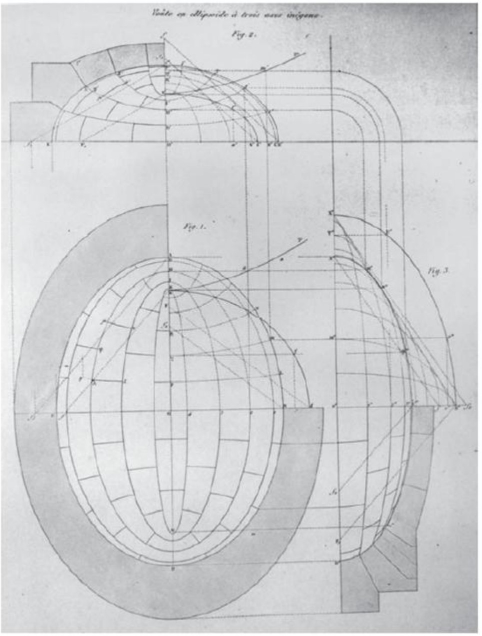
Lorsqu'on doit tranſporter des terres d'un lieu dans un autre, on a coutume de donner le nom de *Déblai* au volume des terres que l'on doit tranſporter, & le nom de



Part. 1 Optimal Transport

Gaspard Monge – geometry and light





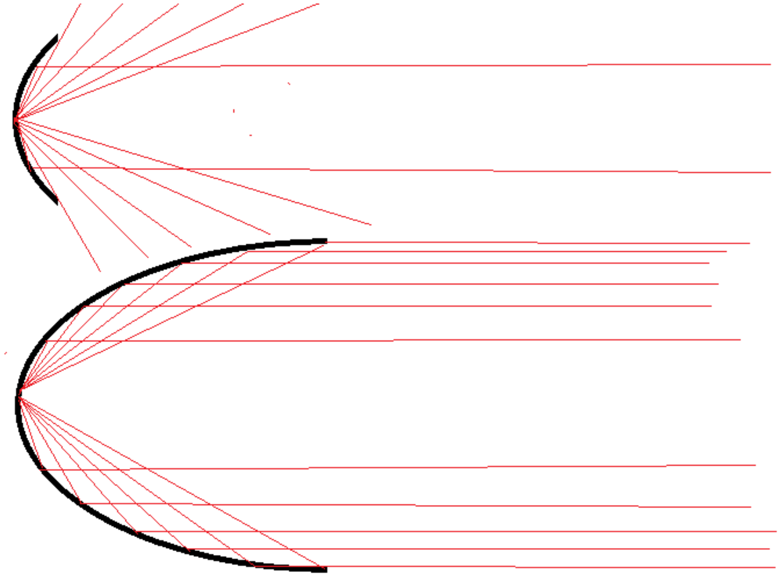
Part. 1 Optimal Transport

Monge-Brenier-Villani, the french connection

Cédric Villani
Optimal Transport Old & New
Topics on Optimal Transport

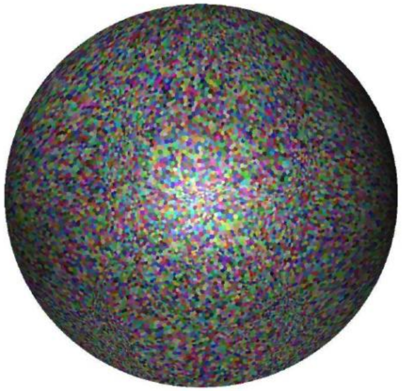
Yann Brenier
The polar factorization theorem
(Brenier Transport)

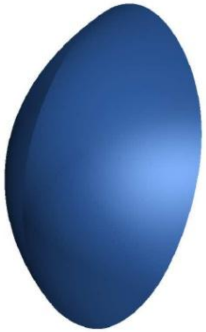
Part. 1 Optimal Transport



Optimal transport
geometry and light

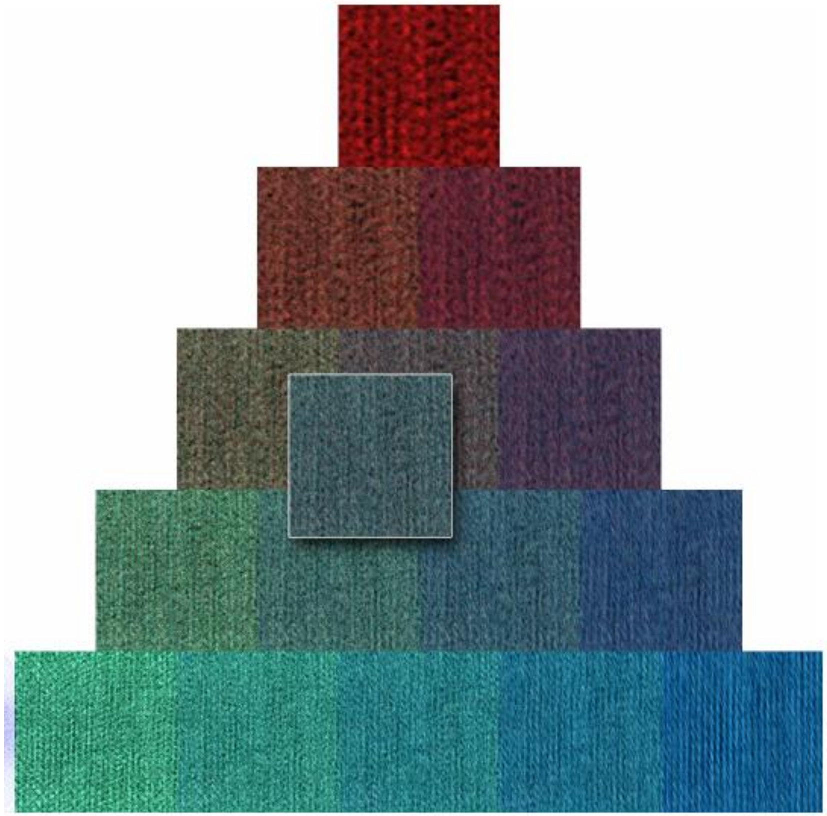
[Caffarelli, Kochengin, and Oliker 1999]





[Castro, Merigot, Thibert 2014]

Part. 1 Optimal Transport – Image Processing



Barycenters / mixing textures

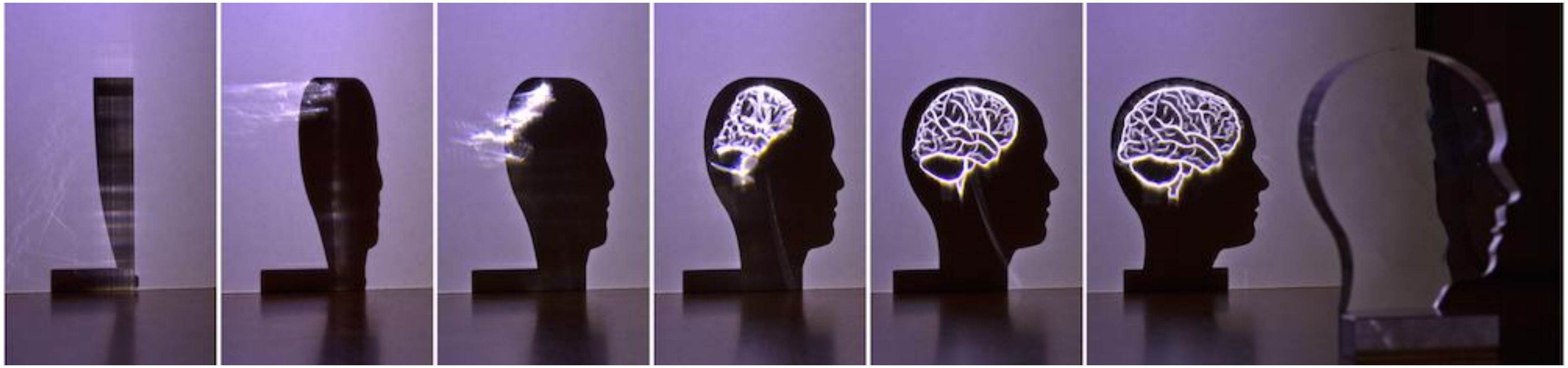
[Nicolas Bonneel, Julien Rabin, Gabriel Peyré, Hanspeter Pfister]

Video-style transfer,
A.I., "data sciences"

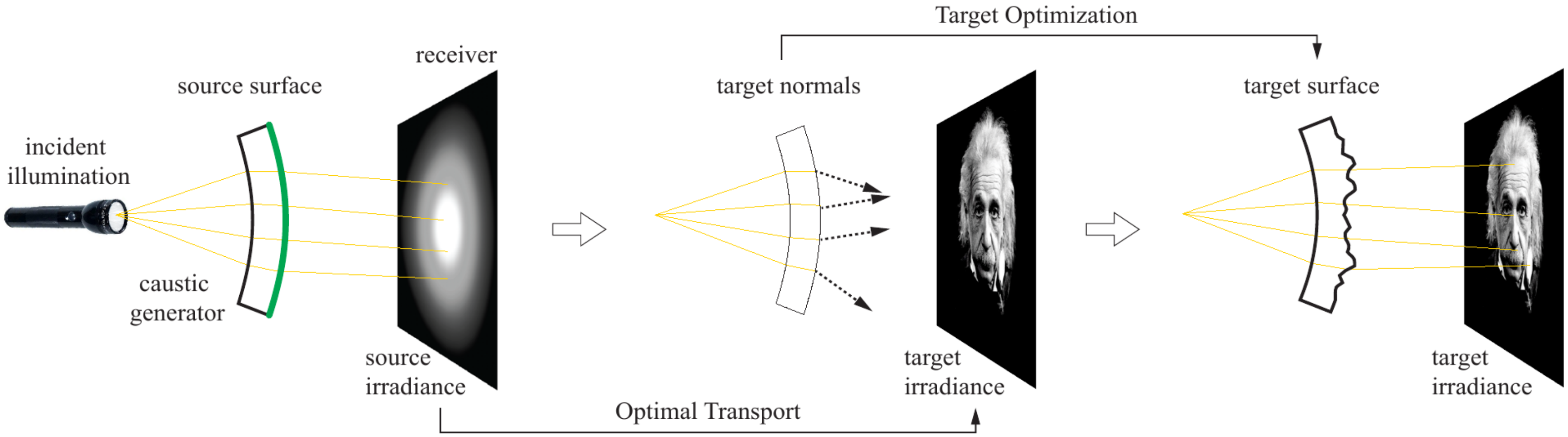
[Nicolas Bonneel, Kalyan Sunkavalli, Sylvain Paris, Hanspeter Pfister]
[Marco Cuturi, Gabriel Peyré]

Part. 1 Optimal Transport

Optimal transport - geometry and light



[Chwartzburg, Testuz, Tagliasacchi, Pauly, SIGGRAPH 2014]

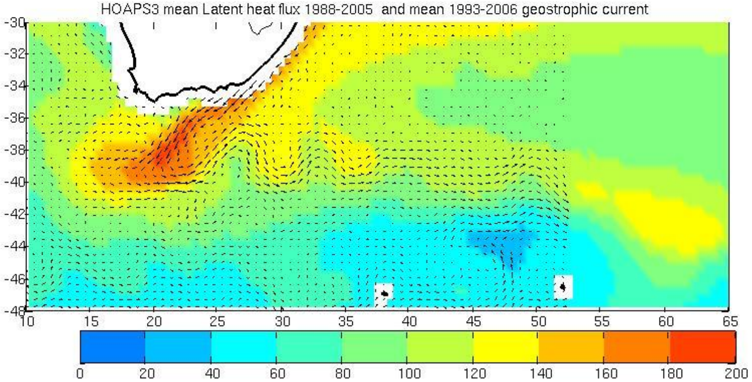


Part. 1. Motivations

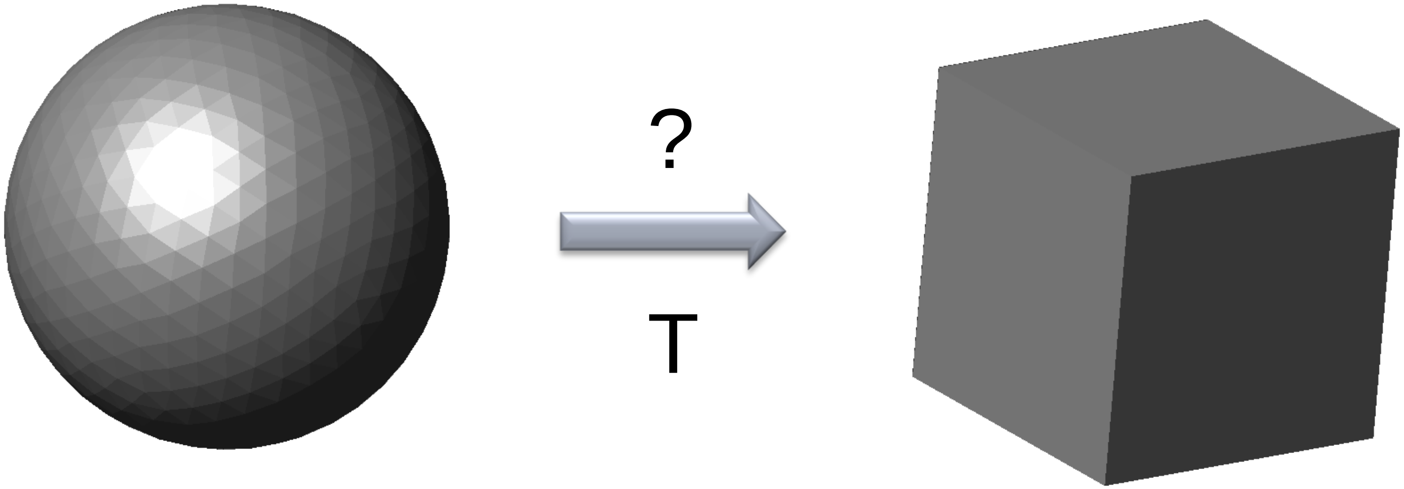
Discretization of functionals involving the Monge-Ampère operator,
Benamou, Carlier, Mérigot, Oudet
arXiv:1408.4536

The variational formulation of the Fokker-Planck equation
Jordan, Kinderlehrer and Otto
SIAM J. on Mathematical Analysis

Geostrophic current

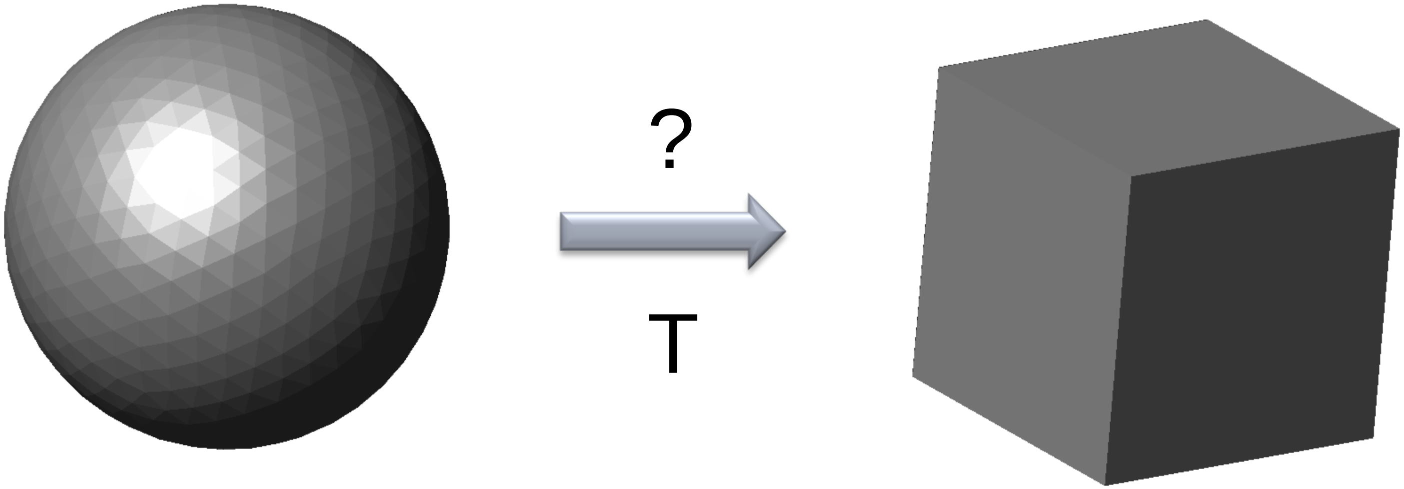


Part. 1 Optimal Transport



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

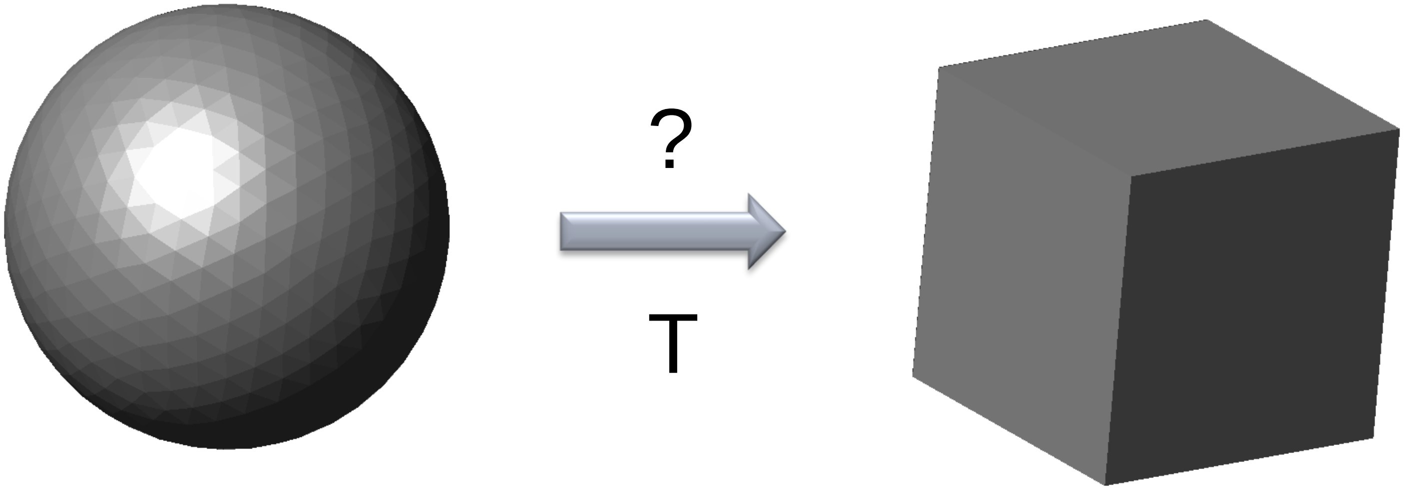
Part. 1 Optimal Transport



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

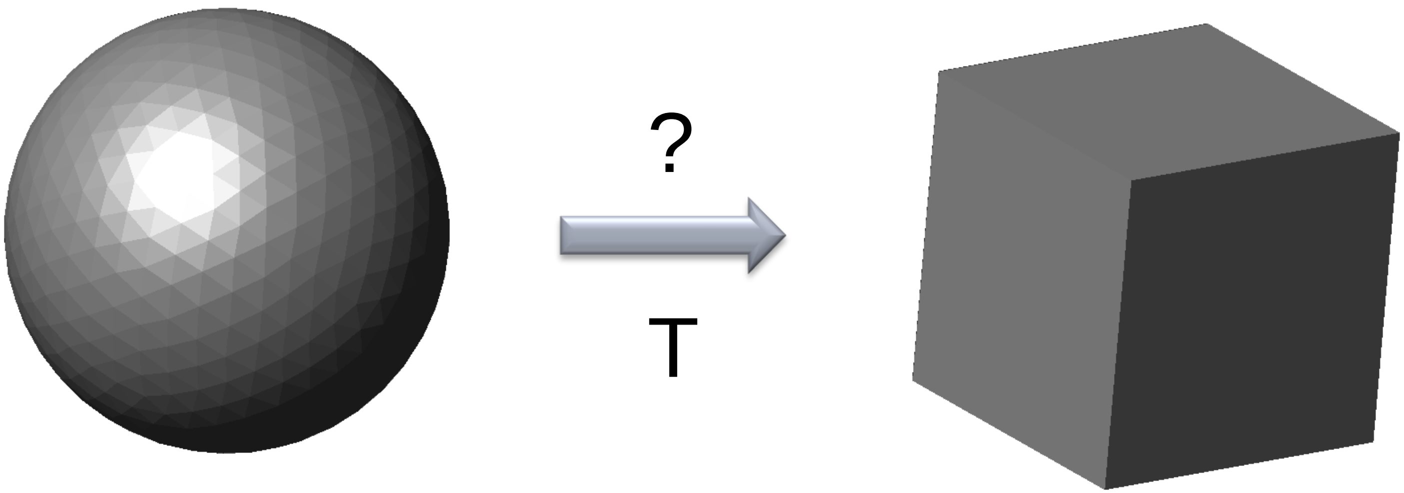
Part. 1 Optimal Transport



How to "morph" a shape into another one
while preserving mass and minimizing the effort ?

Part. 1 Optimal Transport

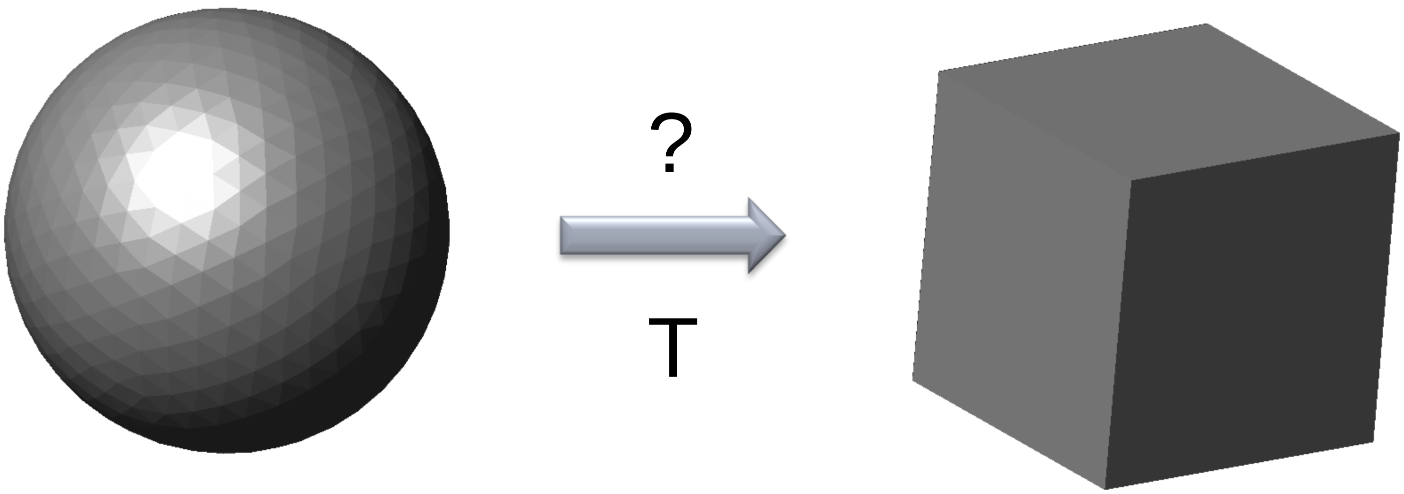
Part. 1 Optimal Transport



How to "morph" a shape into another one
while preserving mass and **minimizing the effort** ?

"minimum action principle"

Part. 1 Optimal Transport



How to “morph” a shape into another one
while **preserving mass** and **minimizing the effort** ?

“conservation law”

“minimum action principle”

Part. 1 Optimal Transport

OT=

“minimum action principle subject to conservation law”

Yann Brenier:

“Each time the Laplace operator is used in a PDE,
it can be replaced with the Monge-Ampère operator”

Part. 1 Optimal Transport

OT=

“minimum action principle subject to conservation law”

Yann Brenier:

“Each time the Laplace operator is used in a PDE,
it can be replaced with the Monge-Ampère operator”

New ways of simulating physics with a computer

Part. 1 Optimal Transport

OT=

"minimum action principle subject to conservation law"

Yann Brenier:

"Each time the <u>Laplace operator</u> is used in a PDE,
it can be replaced with the Monge-Ampère operator"

<u>Fast Fourier Transform</u>

New ways of simulating physics with a computer

Part. 1 Optimal Transport

OT=

"minimum action principle subject to conservation law"

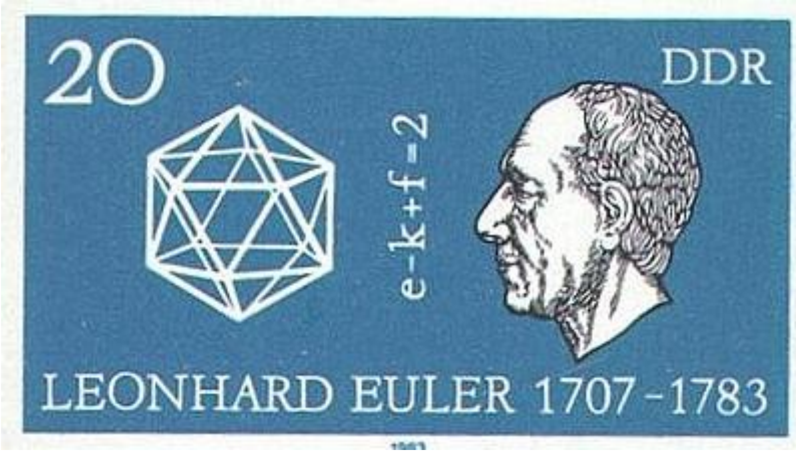
Yann Brenier:

"Each time the Laplace operator is used in a PDE,
it can be replaced with the Monge-Ampère operator"

Fast Fourier Transform

Fast OT algo. ???

New ways of simulating physics with a computer



Euler

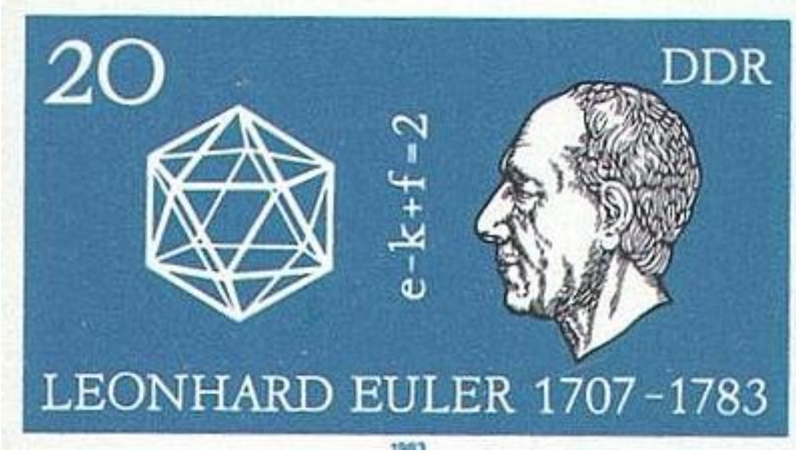
Hamilton,
Legendre,
Maupertuis

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

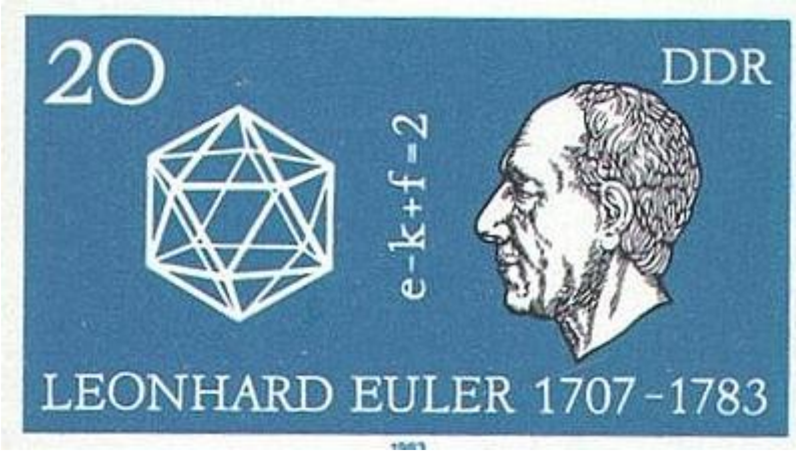
Hamilton,
Legendre,
Maupertuis

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

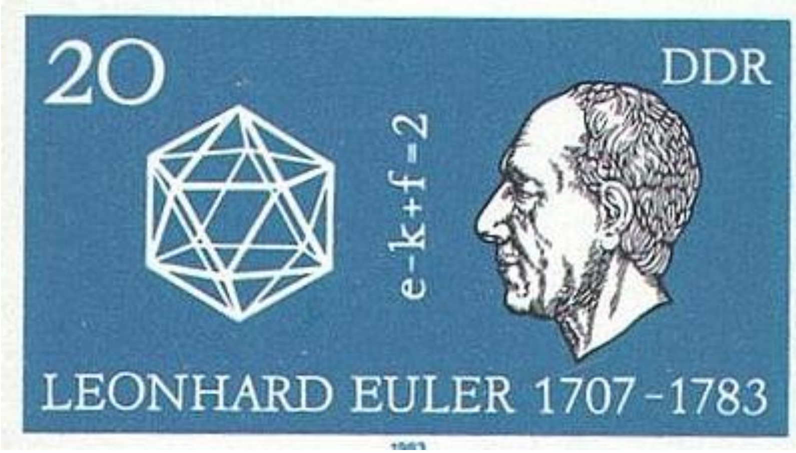
Lagrange

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

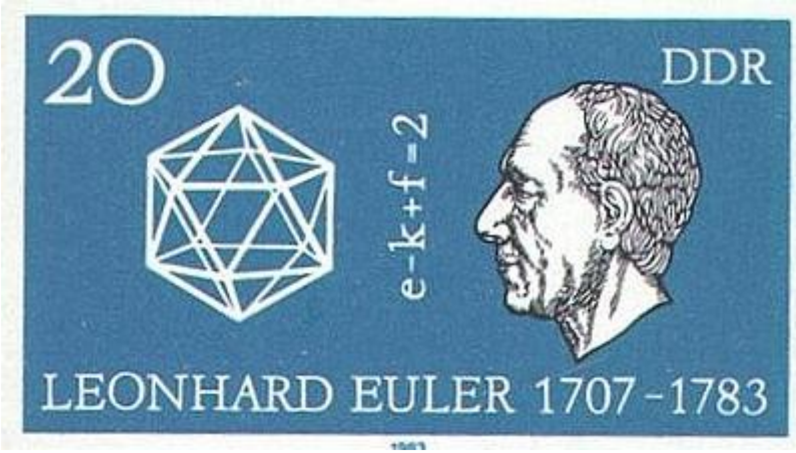
Hamilton,
Legendre,
Maupertuis

Lagrange

The Least Action Principle

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

- x: position
- $\dot{x}$: speed
- t: time



Euler

Hamilton,
Legendre,
Maupertuis

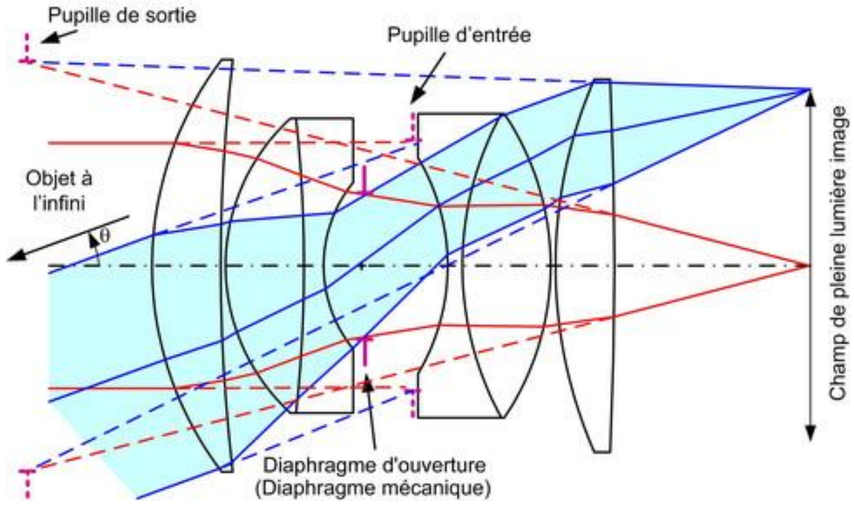
Lagrange

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

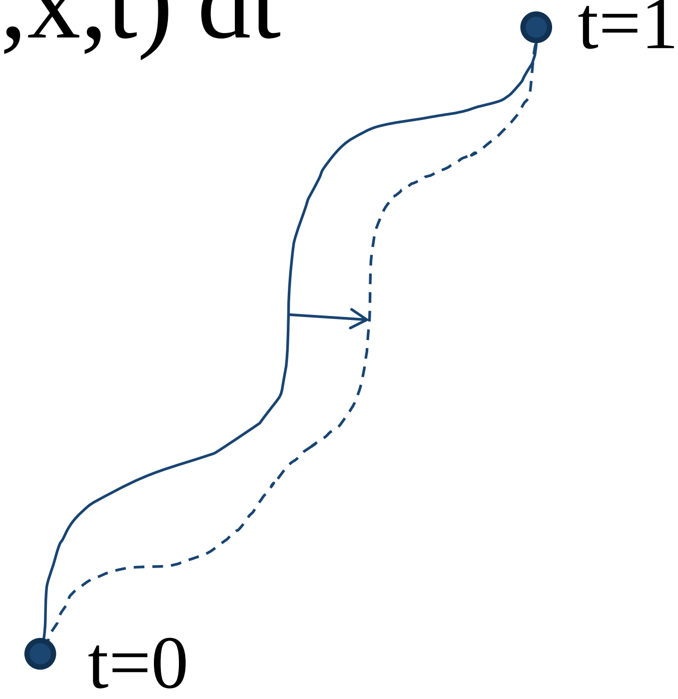
$$\int_{t_1}^{t_2} L(x,\dot{x},t)\,dt$$

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

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

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):

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

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

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

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):

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

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

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

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):

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

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"
Noether'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**

<u>Free particle:</u>

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):

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

Free particle:

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

Expression of the Lagrangian:

$$L = ½\, m\, v^2$$

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. :

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

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)***

Expression of the Lagrangian:

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

Expression of the Energy:

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

The Least Action Principle

Axiom 1: There exists L

Axiom 2: The movement minimizes $\int L$

Theorem 1: (Lagrange equation):

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

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:

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

Expression of the Energy:

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

Particle in a field:

Expression of the Lagrangian:

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

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:

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

Expression of the Energy:

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

Particle in a field:

Expression of the Lagrangian:

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

Expression of the Energy:

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

The Least Action Principle

Axiom 1: There exists L

Axiom 2: The movement minimizes $\int L$

Theorem 1: (Lagrange equation):

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

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:

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

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

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

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 – 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

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

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

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:

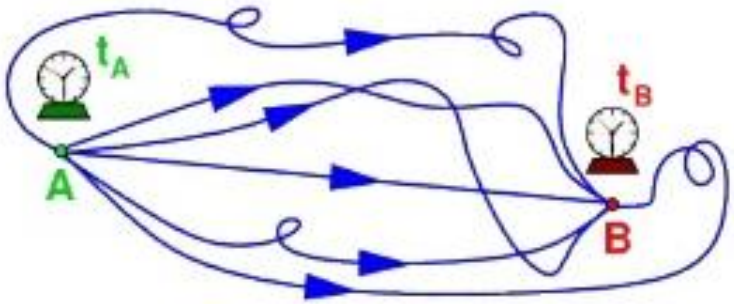
$$E = \tfrac{1}{2} \gamma m v^2 + mc^2$$

The Least Action Principle

(quantum physics setting – just for fun…)

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

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



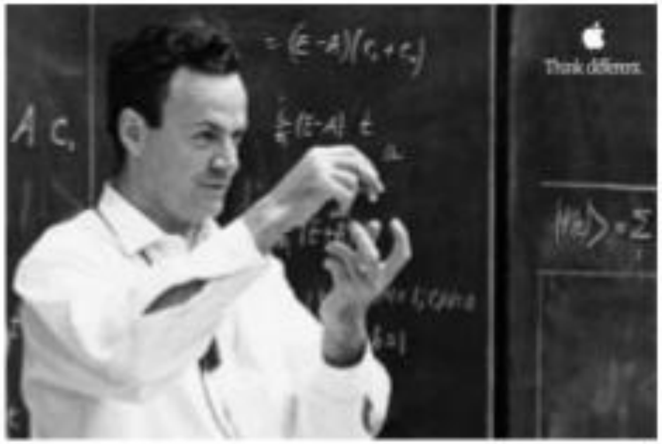
$$K(B,A)=\sum_{over\,all\,possible\,paths}\phi[x(t)]$$

where

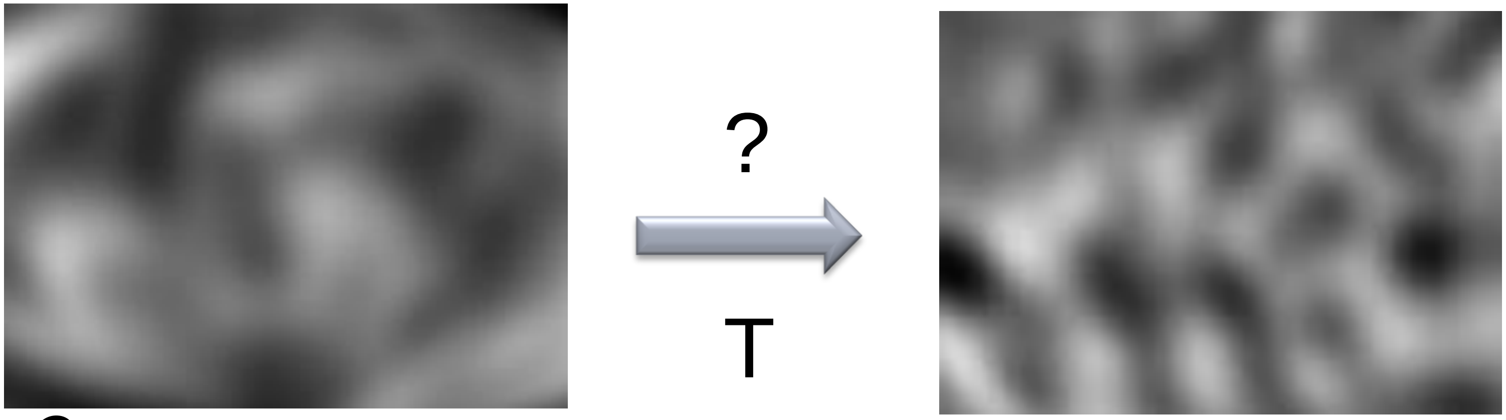
$$\phi[x(t)]=A\exp\left\{\frac{i}{\hbar}S[x(t)]\right\}$$

Feynman's path integral formula

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

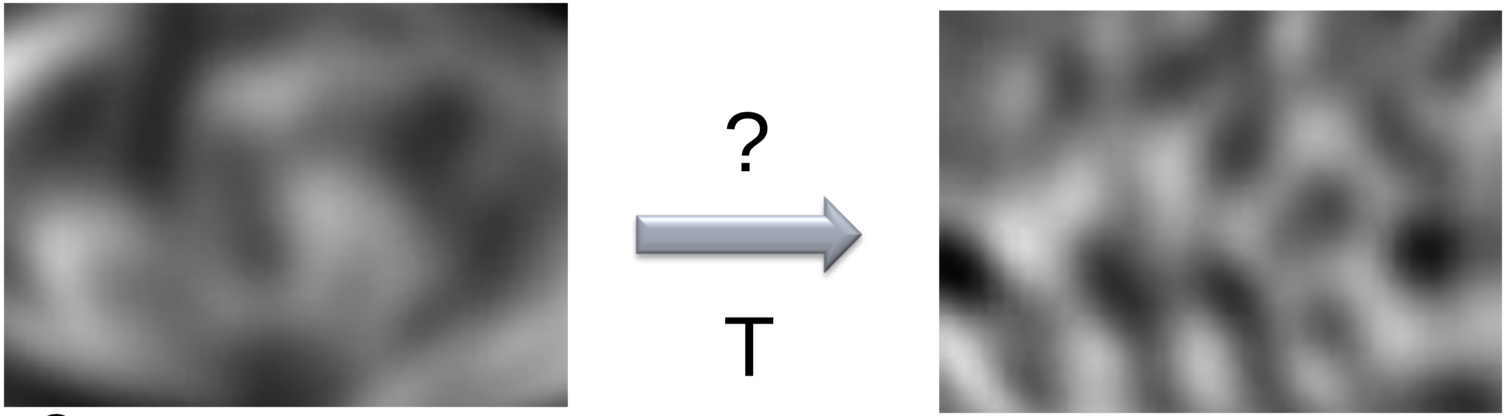


Fluids – Benamou Brenier



ρ_1 ρ_2

Fluids – Benamou Brenier



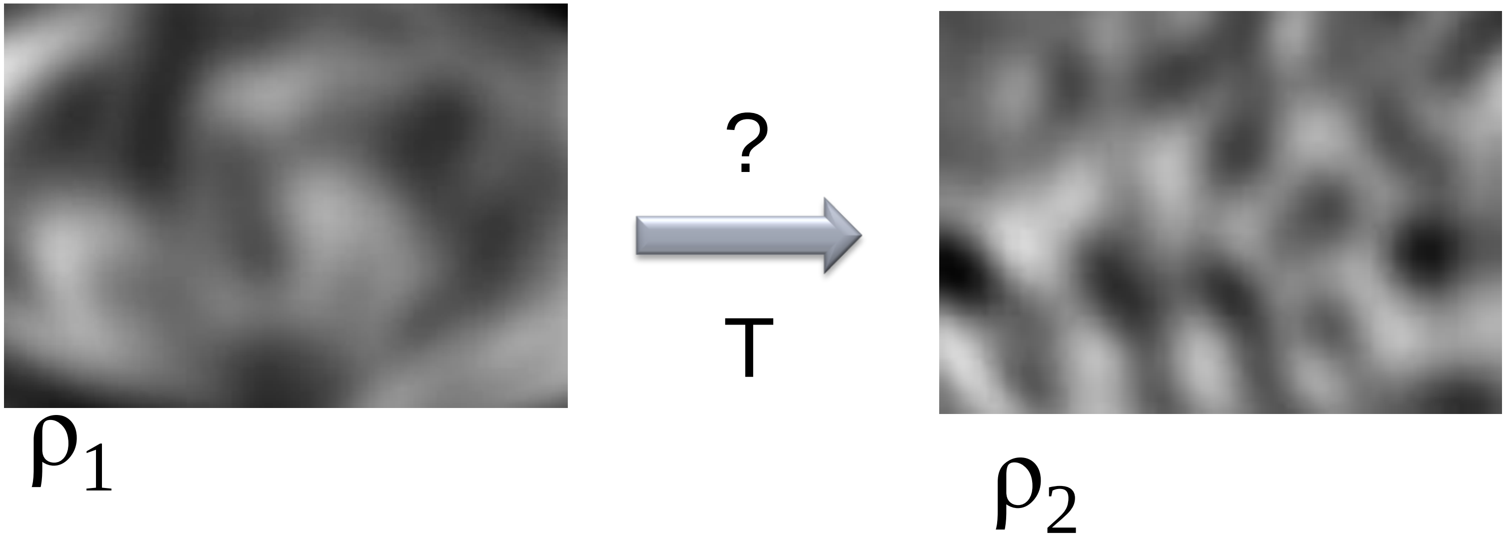
ρ_1 ρ_2

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

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

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

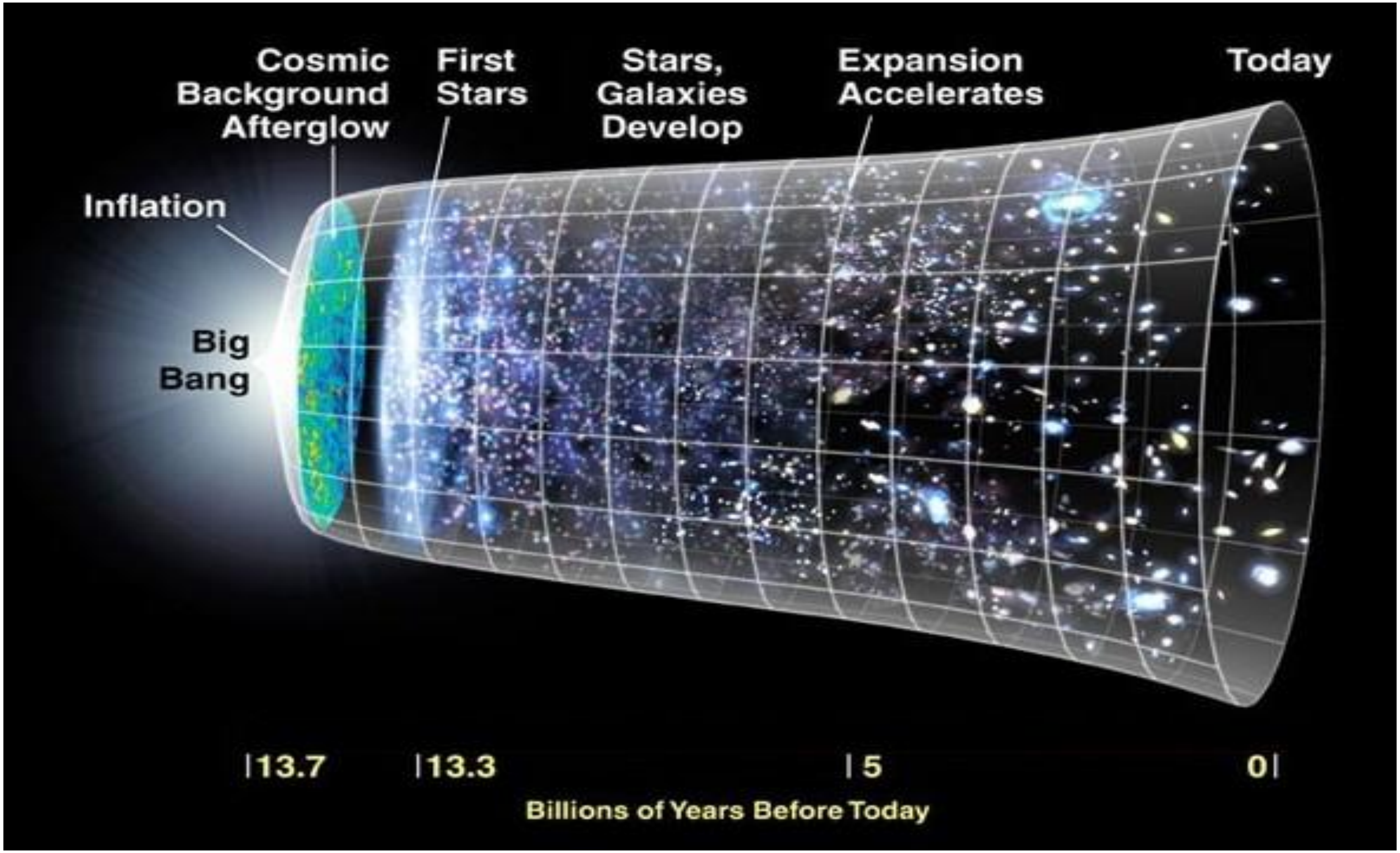
⟷

Minimize C(T) =

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

s.t. T is measure-preserving

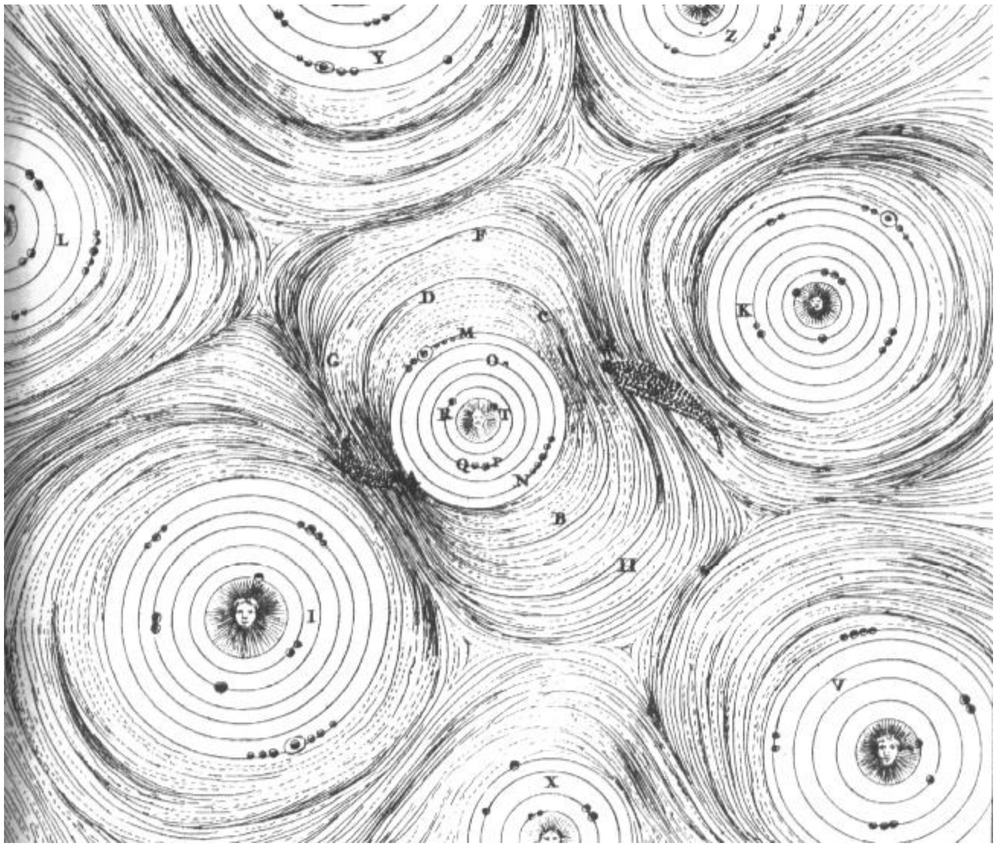
Simuler tout l'univers dans un ordinateur…

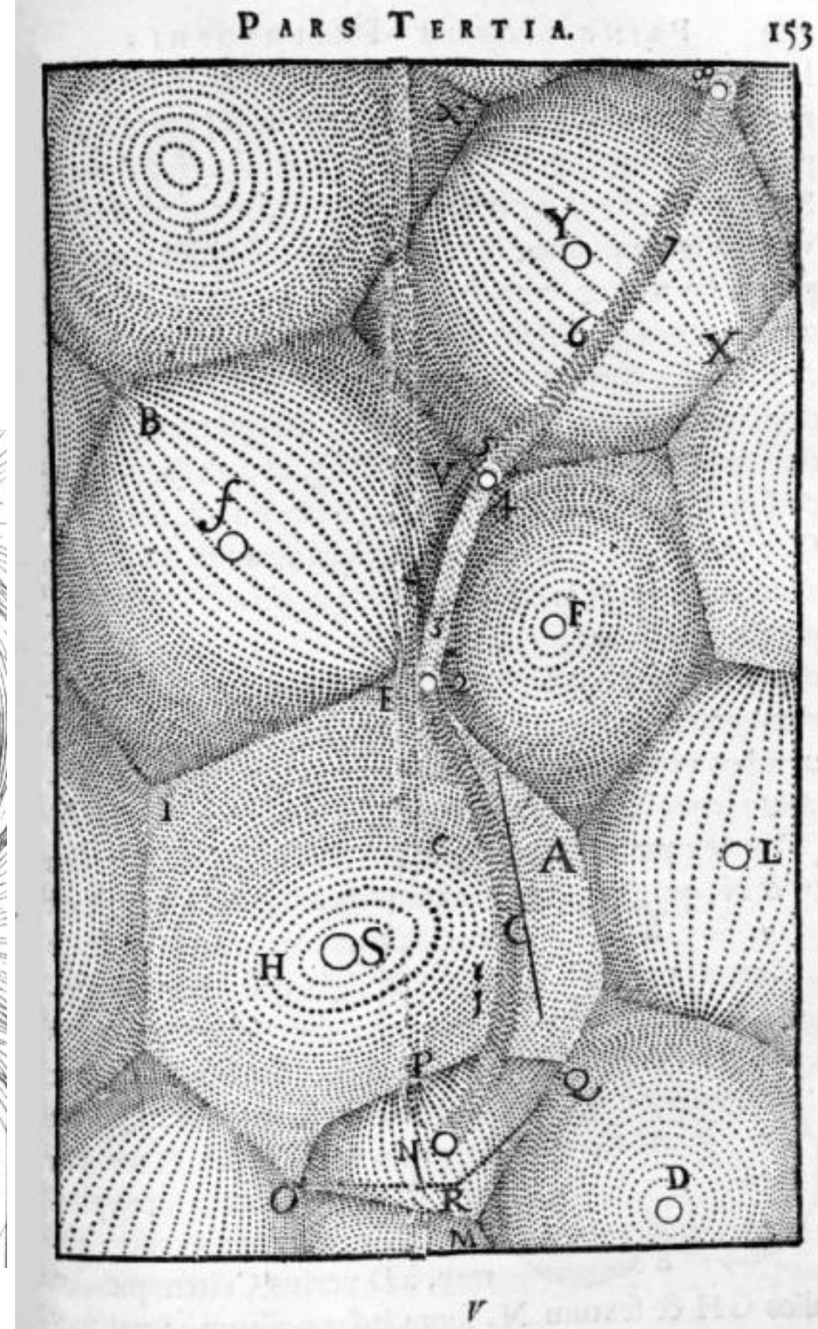


Part. 1 Optimal Transport

Des tourbillons dans l'éther ?

René Descartes - 1663



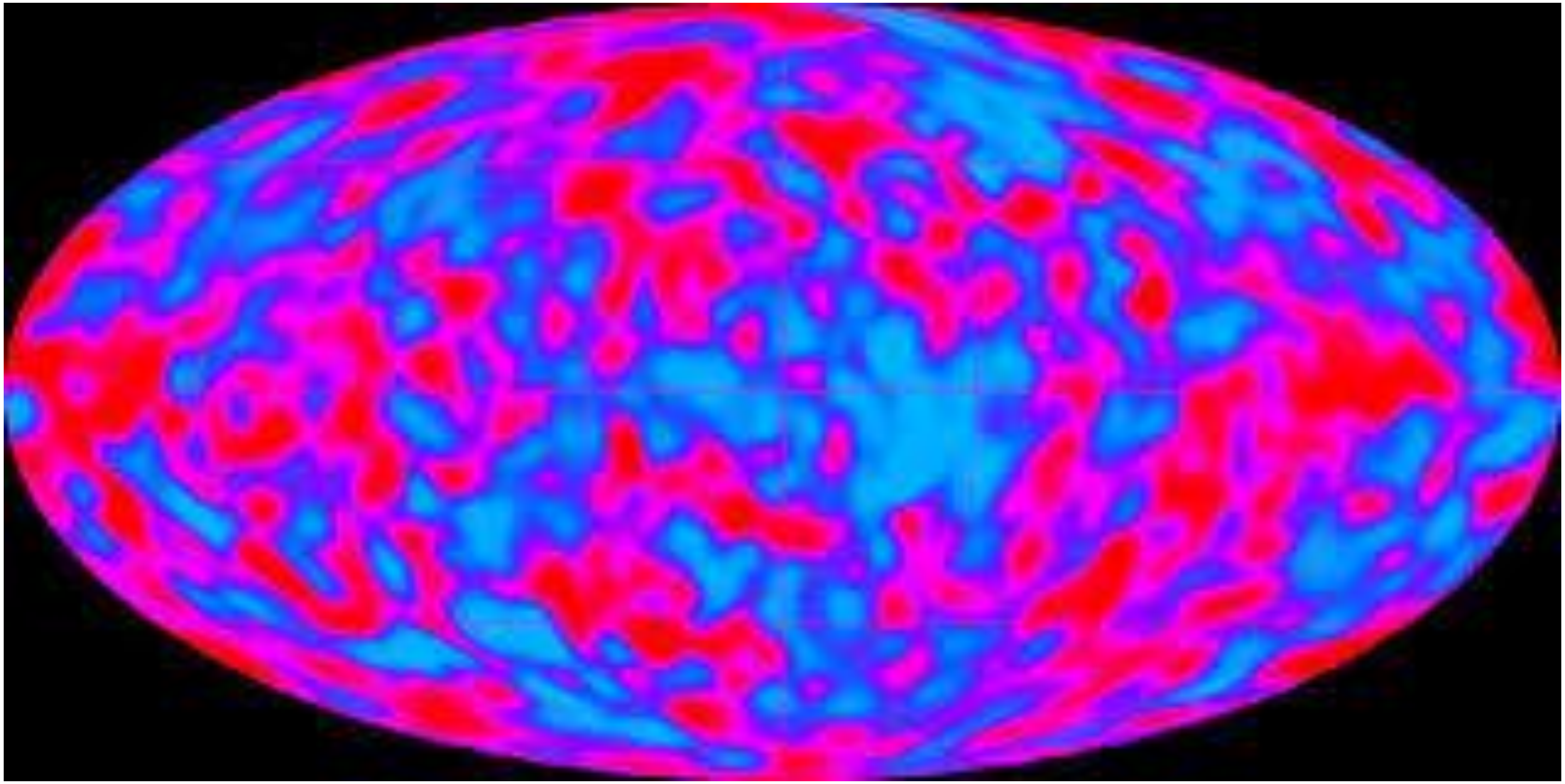


Vers l'infini et au delà …

COBE 1992

Reconstruction de l univers primordial

Les données – Fond de rayonnement cosmologique

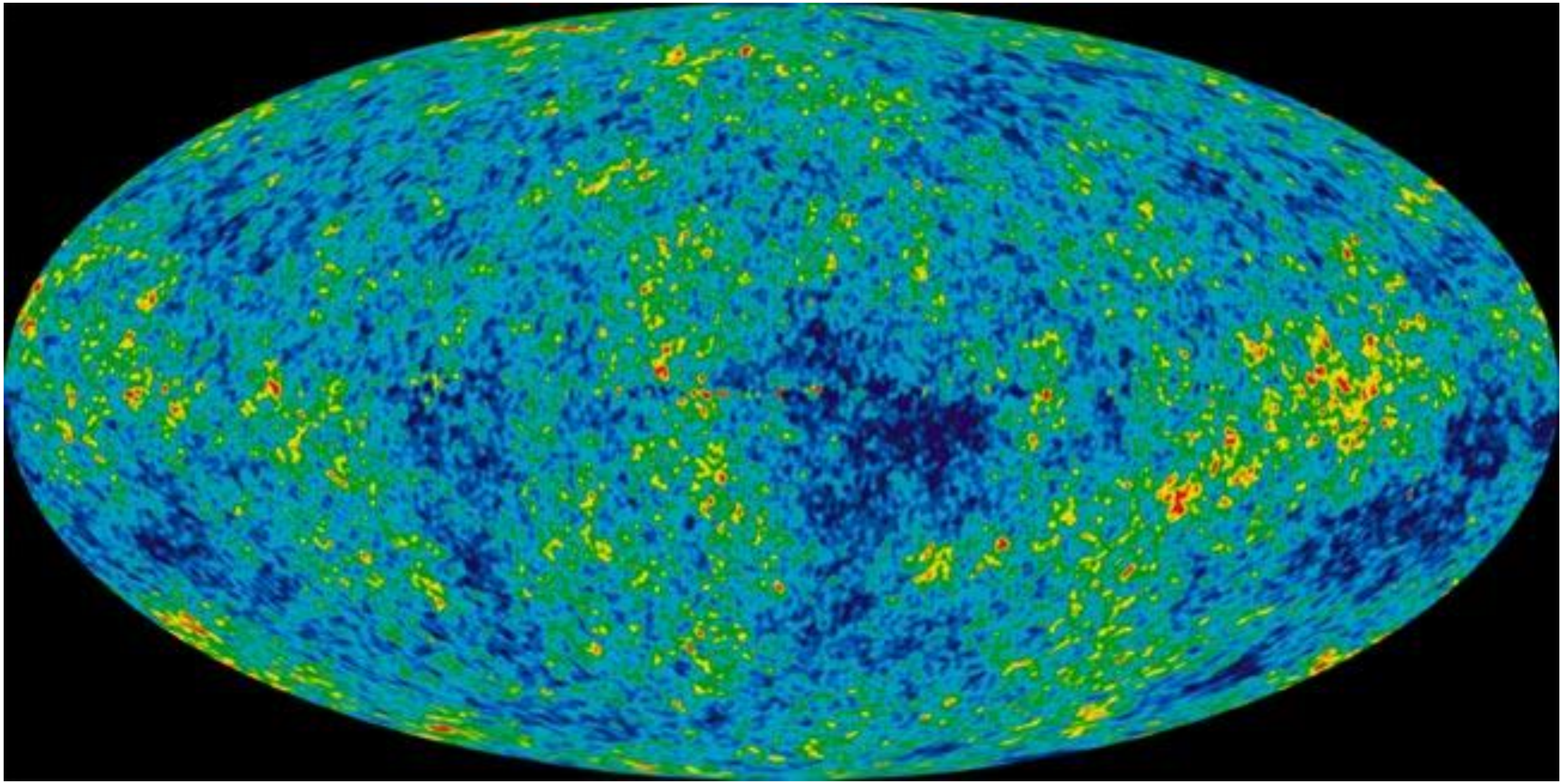


Vers l'infini et au delà …

Reconstruction de l univers primordial

Les données – Fond de rayonnement cosmologique

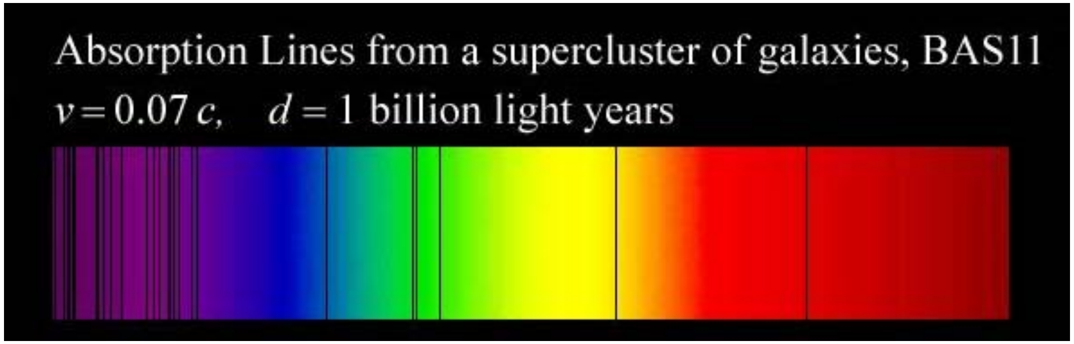
WMAP 2003
2006
2008
2010



Vers l'infini et au delà …

Reconstruction de l univers primordial

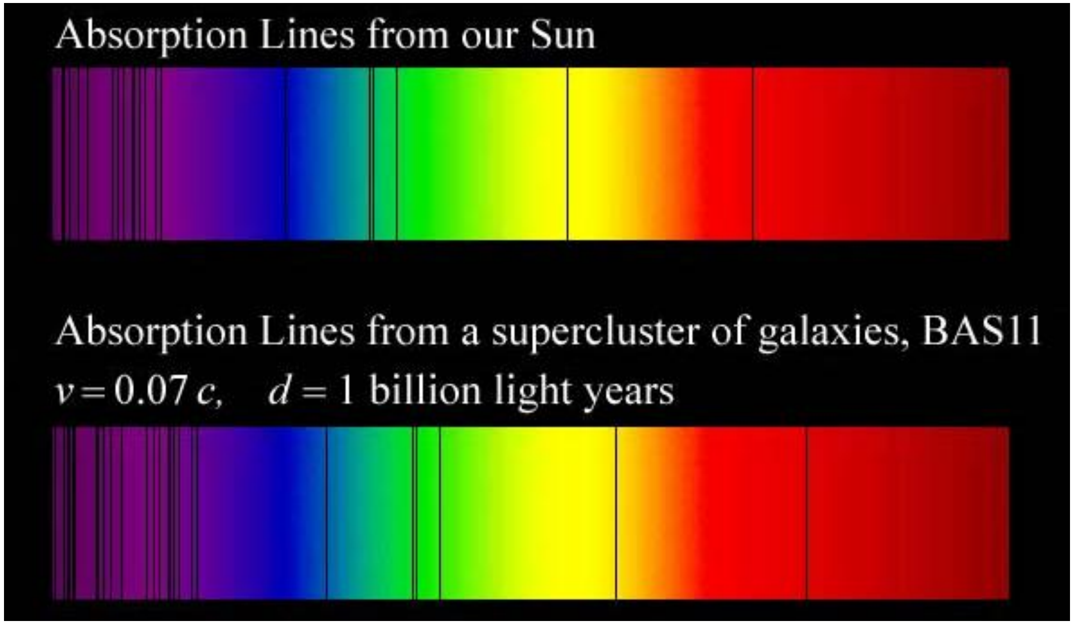
Les données – campagnes d’acquisition red-shift



Vers l'infini et au delà …

Reconstruction de l univers primordial

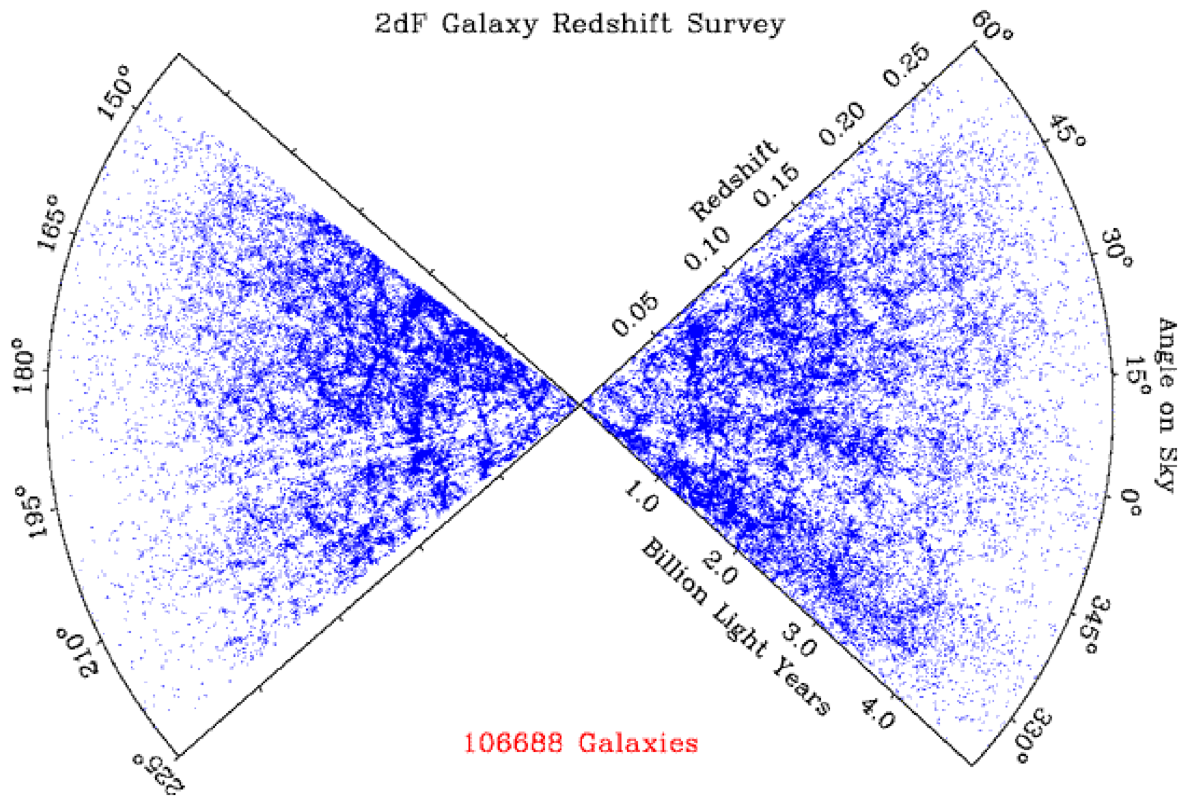
Les données – campagnes d’acquisition red-shift



Vers l'infini et au delà …

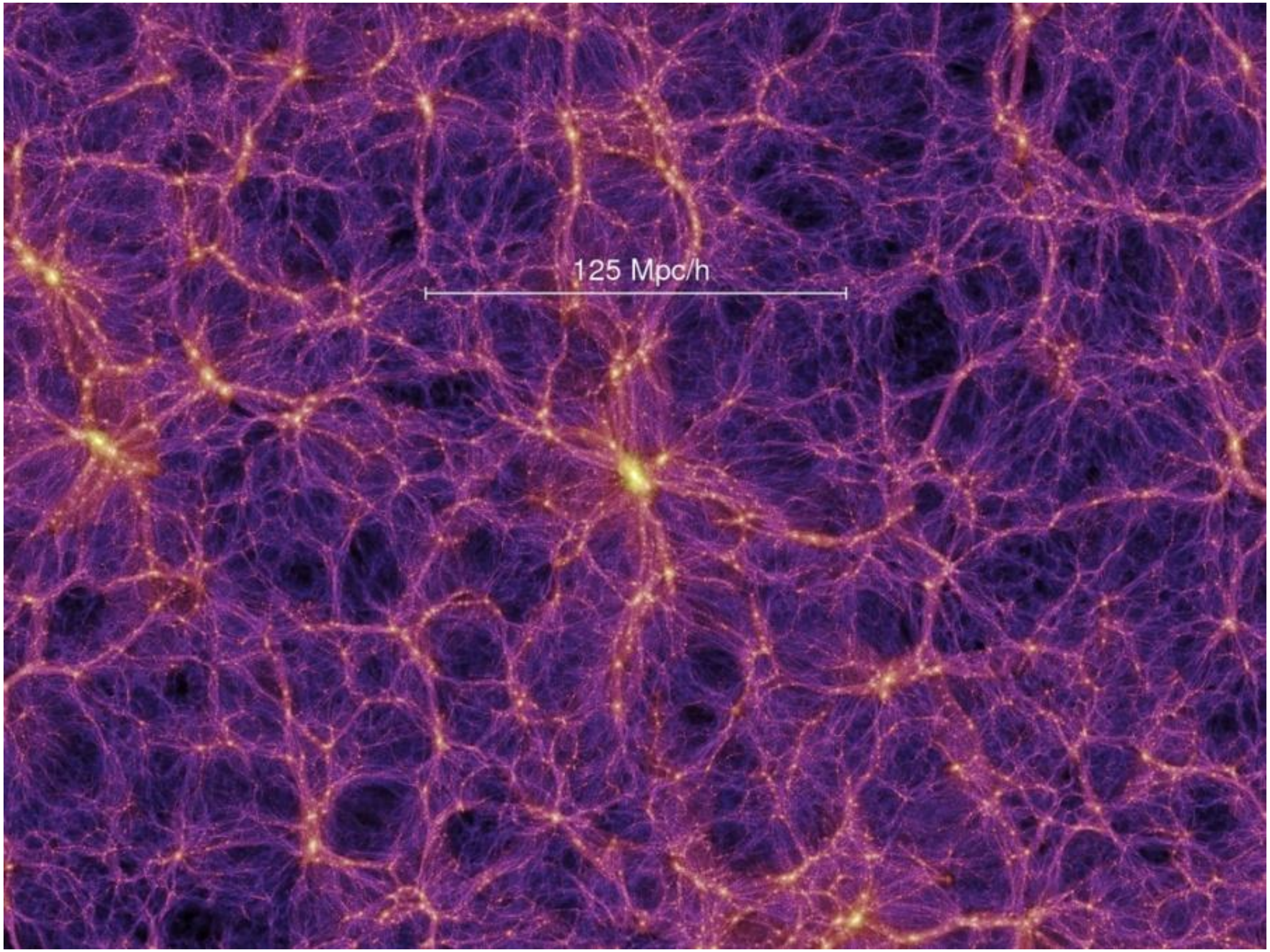
Reconstruction de l univers primordial

Les données – campagnes d’acquisition red-shift



Vers l'infini et au delà …

pc/h : parsec (= 3.2 années lumières)

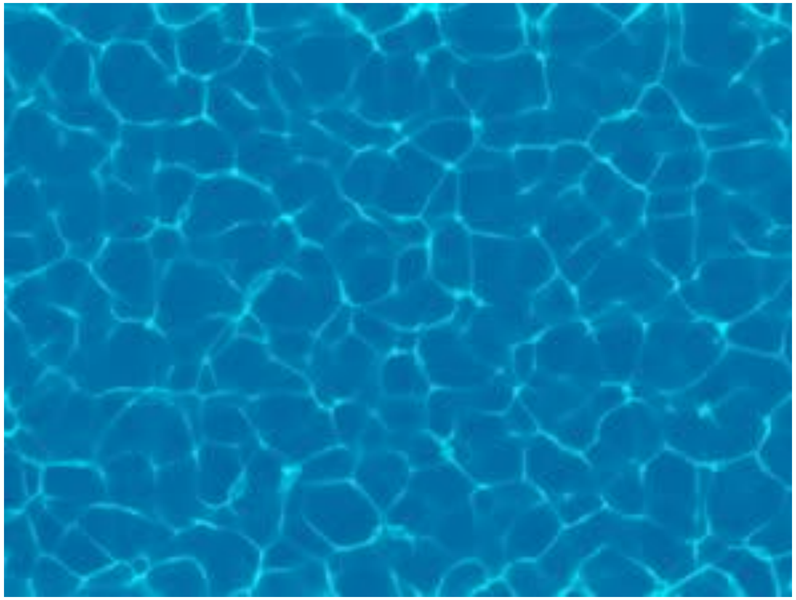


The millenium simulation project, Max Planck Institute fur Astrophysik

Vers l'infini et au delà …

Reconstruction de l état primordial de l univers

La piscine universelle



Vers l'infini et au delà …

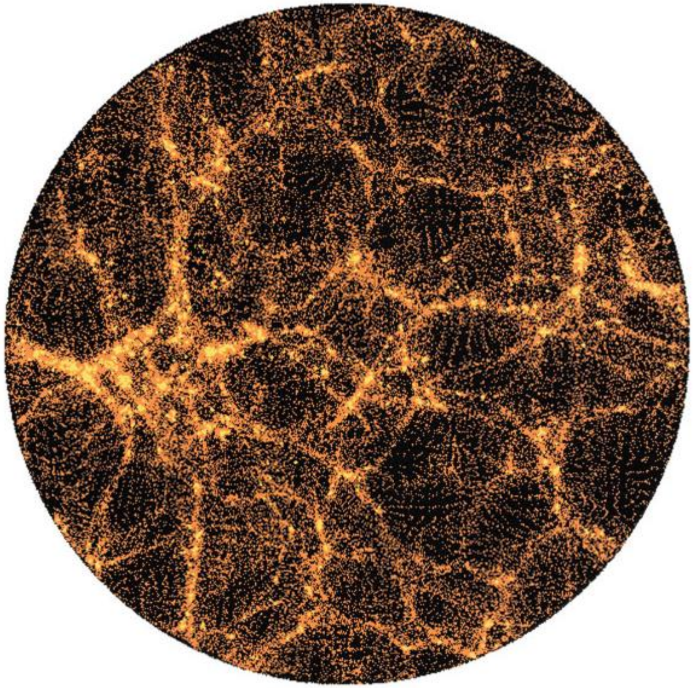
Inverser les équations de Newton / Einstein pour remonter le temps de 14 milliards d années

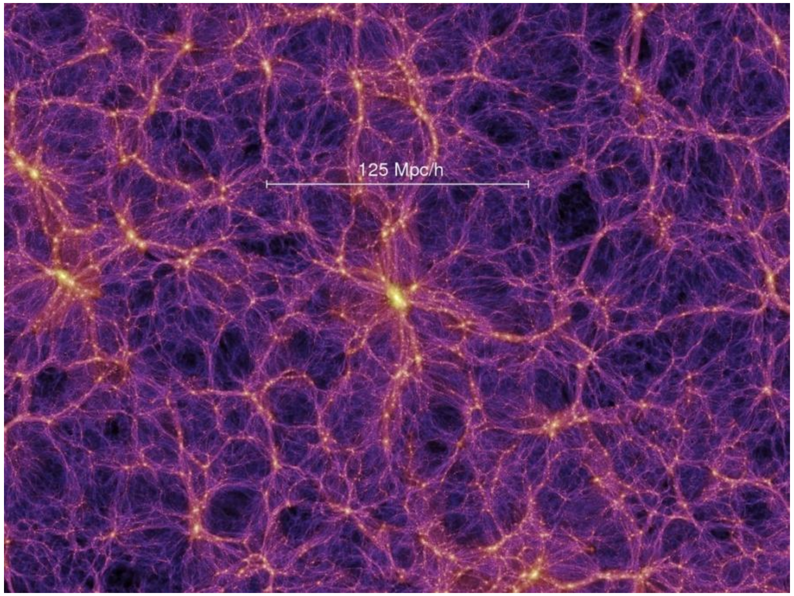
RECONSTRUCTION OF THE EARLY UNIVERSE, ZELDOVICH APPROXIMATION AND MONGE-AMPÈRE GRAVITATION

YANN BRENIER

U. Frisch, S. Matarrese, R. Mohayaee, A. Sobolevski, *A reconstruction of the initial conditions of the Universe by optimal mass transportation*, Nature 417 (2002) 260-262.

pc/h : parsec (= 3.2 light years)





The millenium simulation project,
Max Planck Institute fur Astrophysik

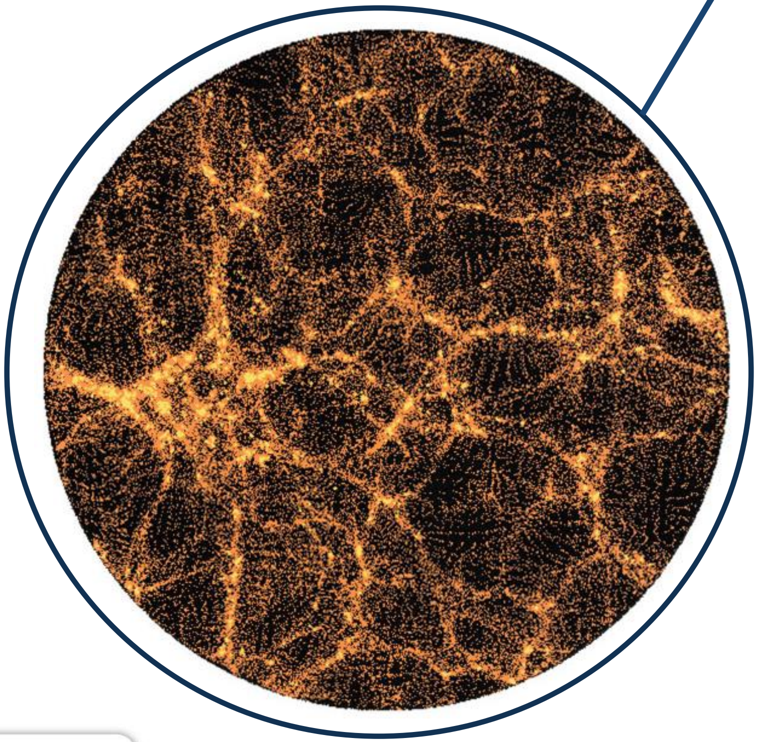
Vers l'infini et au delà ...

En 2002: 5 heures de calcul
sur un super-ordinateur / **5000** points

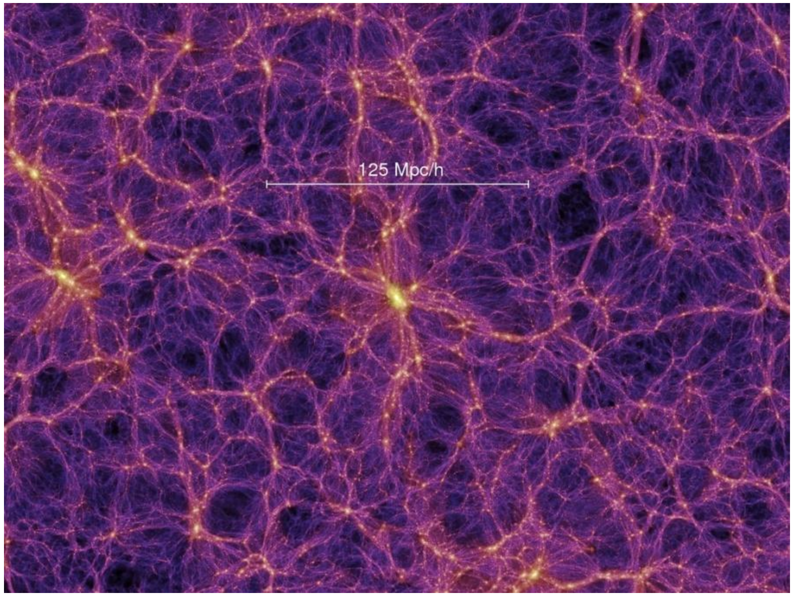
RECONSTRUCTION OF THE EARLY UNIVERSE, ZELDOVICH APPROXIMATION AND MONGE-AMPÈRE GRAVITATION

YANN BRENIER

U. Frisch, S. Matarrese, R. Mohayaee, A. Sobolevski, *A reconstruction of the initial conditions of the Universe by optimal mass transportation*, Nature 417 (2002) 260-262.



pc/h : parsec (= 3.2 années lumières)



The millenium simulation project,
Max Planck Institute fur Astrophysik

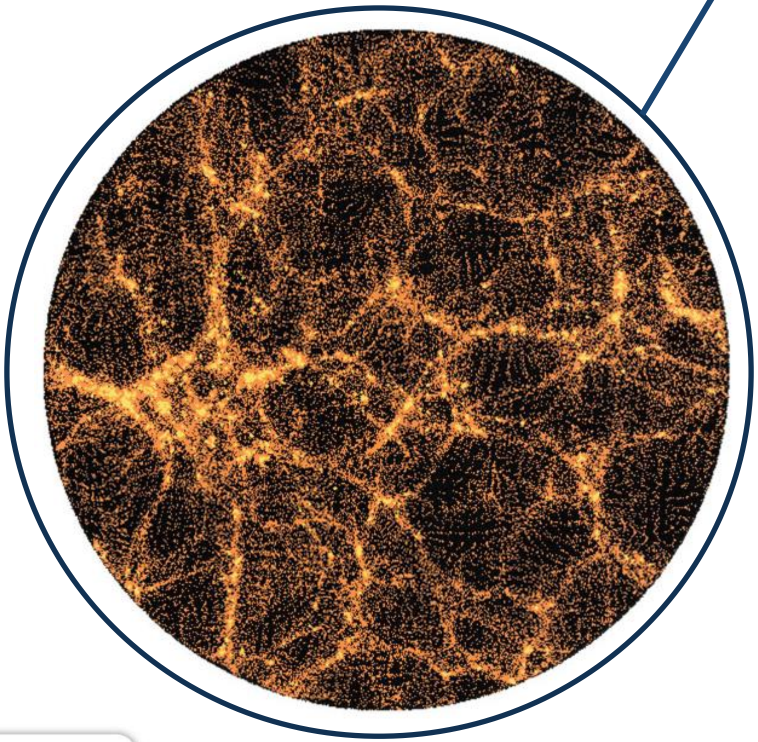
Vers l'infini et au delà …

En 2002: 5 heures de calcul
sur un super-ordinateur / **5000** points
Il serait déraisonnable de faire le calcul avec plus de 100 000 points

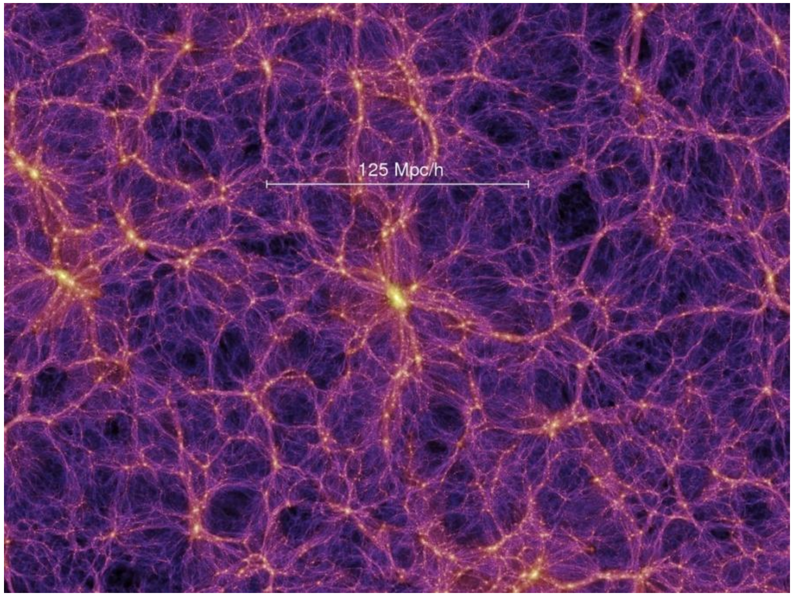
RECONSTRUCTION OF THE EARLY UNIVERSE, ZELDOVICH APPROXIMATION AND MONGE-AMPÈRE GRAVITATION

YANN BRENIER

U. Frisch, S. Matarrese, R. Mohayaee, A. Sobolevski, *A reconstruction of the initial conditions of the Universe by optimal mass transportation,* Nature 417 (2002) 260-262.



pc/h : parsec (= 3.2 années lumières)



The millenium simulation project,
Max Planck Institute fur Astrophysik

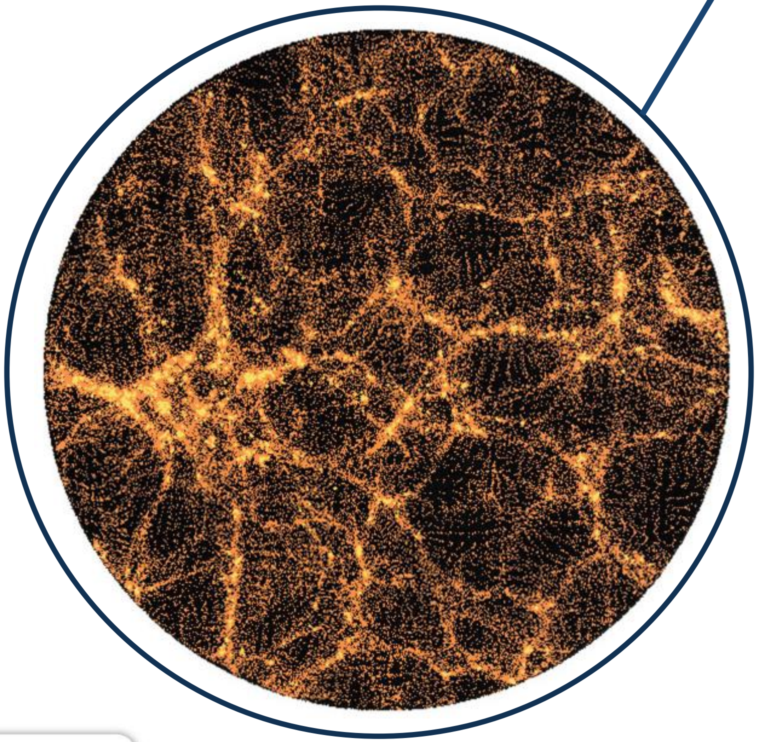
Vers l'infini et au delà …

En 2002: 5 heures de calcul
sur un super-ordinateur / **5000** points
Il serait déraisonnable de faire le calcul avec plus de 100 000 points
Peut-on faire le calcul avec **1 000 000** points ?

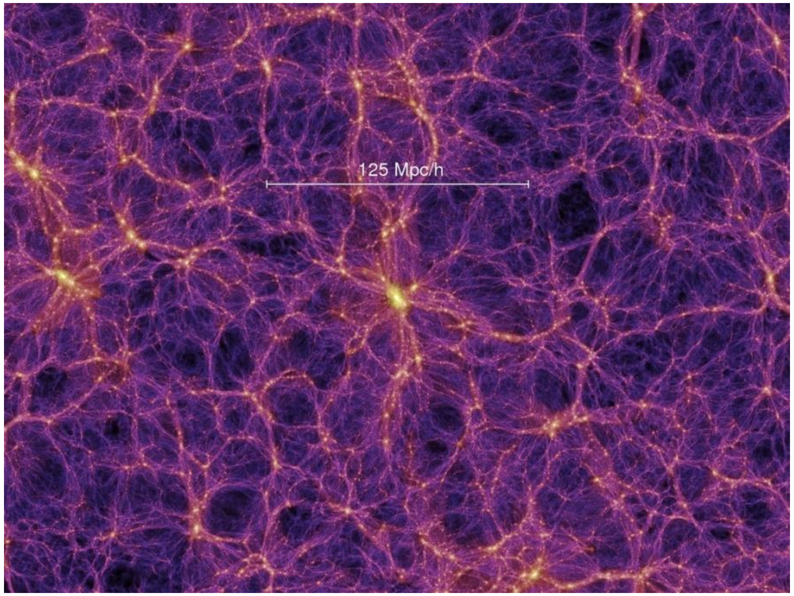
RECONSTRUCTION OF THE EARLY UNIVERSE, ZELDOVICH APPROXIMATION AND MONGE-AMPÈRE GRAVITATION

YANN BRENIER

U. Frisch, S. Matarrese, R. Mohayaee, A. Sobolevski, *A reconstruction of the initial conditions of the Universe by optimal mass transportation,* Nature 417 (2002) 260-262.



pc/h : parsec (= 3.2 années lumières)



The millenium simulation project,
Max Planck Institute fur Astrophysik

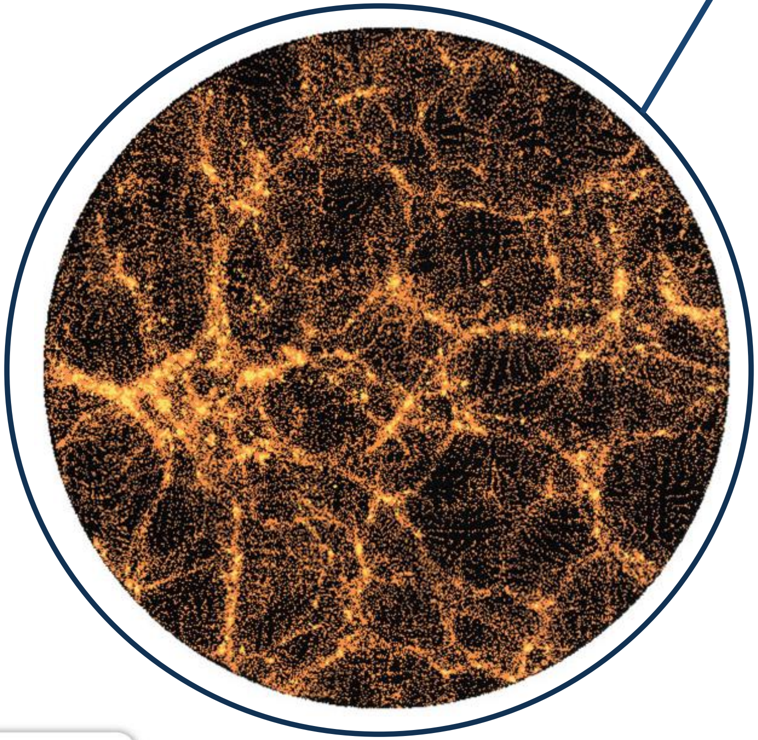
Vers l'infini et au delà …

En 2002: 5 heures de calcul
sur un super-ordinateur / **<u>5000</u>** points
Il serait déraisonnable de faire le calcul avec plus de 100 000 points
Peut-on faire le calcul avec
<u>1 000 000</u> points ?
Oui si on attend (4500 ans !!)

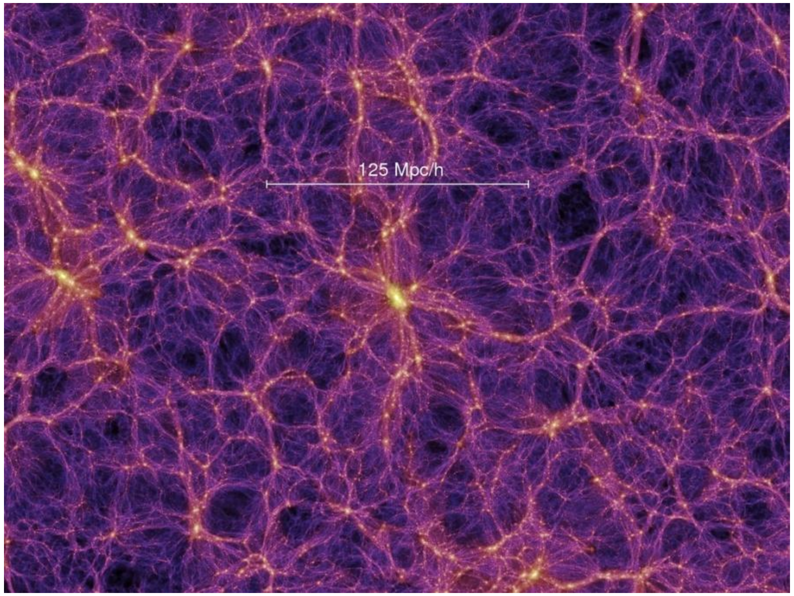
RECONSTRUCTION OF THE EARLY UNIVERSE, ZELDOVICH APPROXIMATION AND MONGE-AMPÈRE GRAVITATION

YANN BRENIER

U. Frisch, S. Matarrese, R. Mohayaee, A. Sobolevski, *A reconstruction of the initial conditions of the Universe by optimal mass transportation,* Nature 417 (2002) 260-262.



pc/h : parsec (= 3.2 années lumières)



The millenium simulation project,
Max Planck Institute fur Astrophysik

Transport and Physics

- Euler fluids – geodesics [Galouet, Mirebeau, Mérigot]
- Euler fluids – Cauchy problem
- Gradient flows – Jordan Kinderlehrer Otto

Transport and Physics

- Euler fluids – geodesics [Galouet, Mirebeau, Mérigot]
- Euler fluids – Cauchy problem
- Gradient flows – Jordan Kinderlehrer Otto

Nice connections with least action principle (Louiville)
Theory characterizes irregular objects (sums of Dirac masses, probability measures supported by meshes)

Transport and Physics

- Euler fluids – geodesics [Galouet, Mirebeau, Mérigot]
- Euler fluids – Cauchy problem
- Gradient flows – Jordan Kinderlehrer Otto

Nice connections with least action principle (Louiville)
Theory characterizes irregular objects (sums of Dirac masses, probability measures supported by meshes)

Iterative schemes, need to solve multiple OT problems

Transport and Physics

- Euler fluids – geodesics [Galouet, Mirebeau, Mérigot]
- Euler fluids – Cauchy problem
- Gradient flows – Jordan Kinderlehrer Otto

Nice connections with least action principle (Louiville)
Theory characterizes irregular objects (sums of Dirac masses, probability measures supported by meshes)

Iterative schemes, need to solve multiple OT problems

Need for an efficient solver
(there is no “FFT for OT”, what can we do instead ?)

2

Optimal Transport

an elementary introduction

Part. 2 Optimal Transport – Monge's problem

(X;μ)

(Y;v)

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

Part. 2 Optimal Transport – Monge's problem

(X;μ)

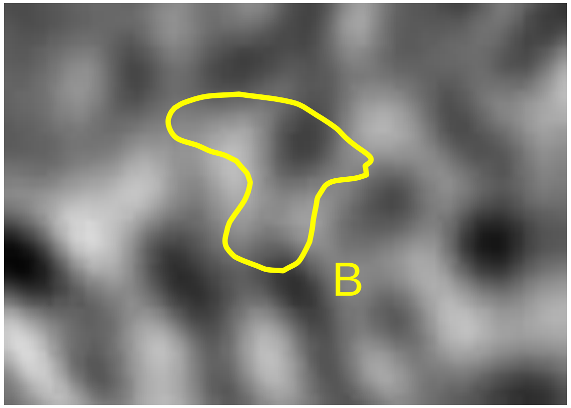
(Y;ν)

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)

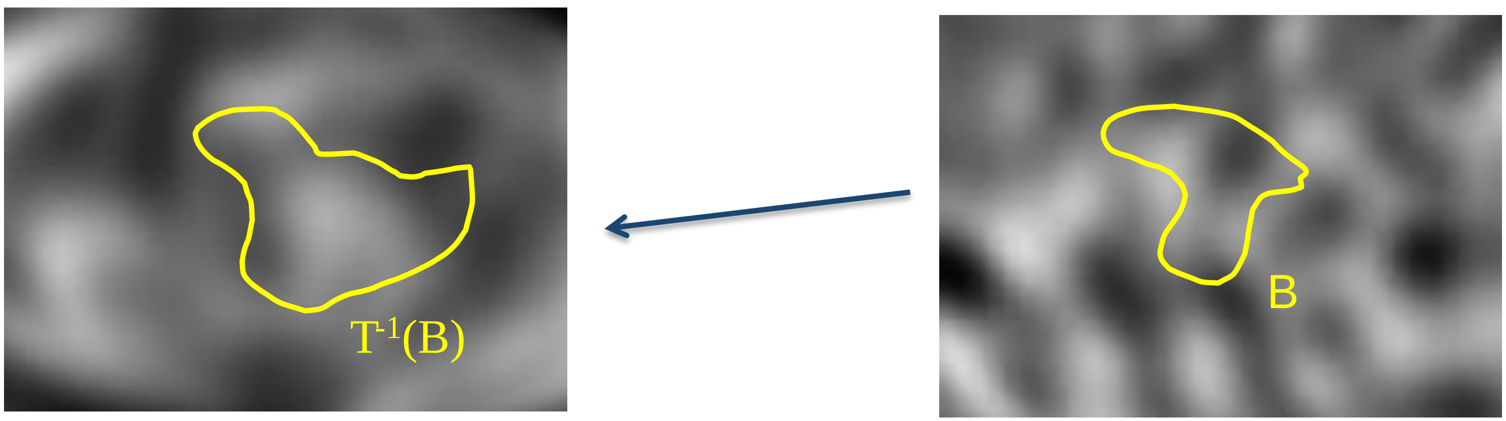


(Y;ν)

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



(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

(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

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

Part. 2 Optimal Transport – Monge's problem

(X;μ)

(Y;ν)

Monge's problem:

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:

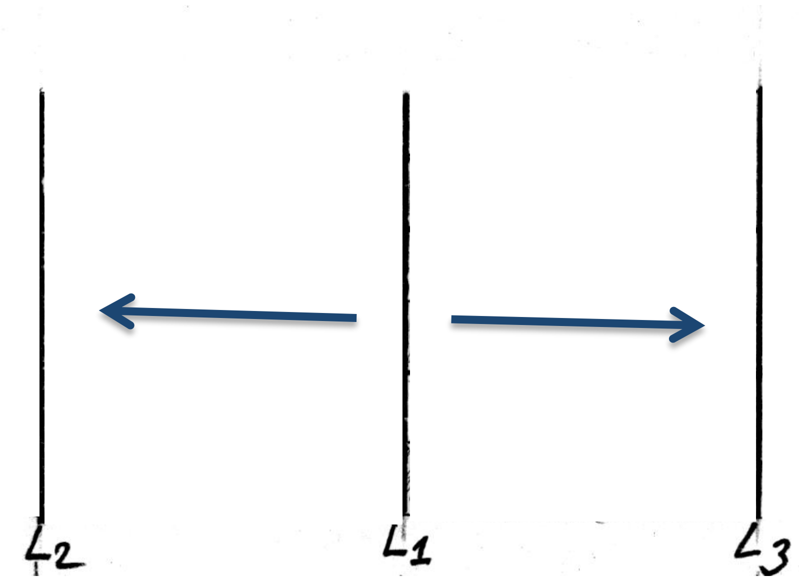
Find a transport map T that minimizes $C(T) = \int_X \| x - T(x) \|^2 \, d\mu(x)$

- 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 – Monge's problem

Monge's problem:

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

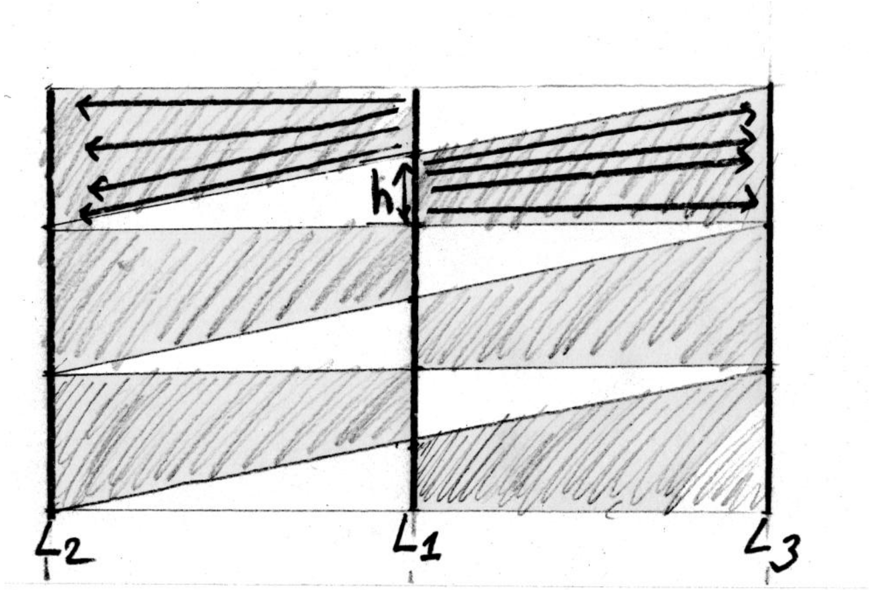


Transport from a measure concentrated on L_1 onto another one concentrated on L_2 and L_3

Part. 2 Optimal Transport – Monge's problem

Monge's problem:

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

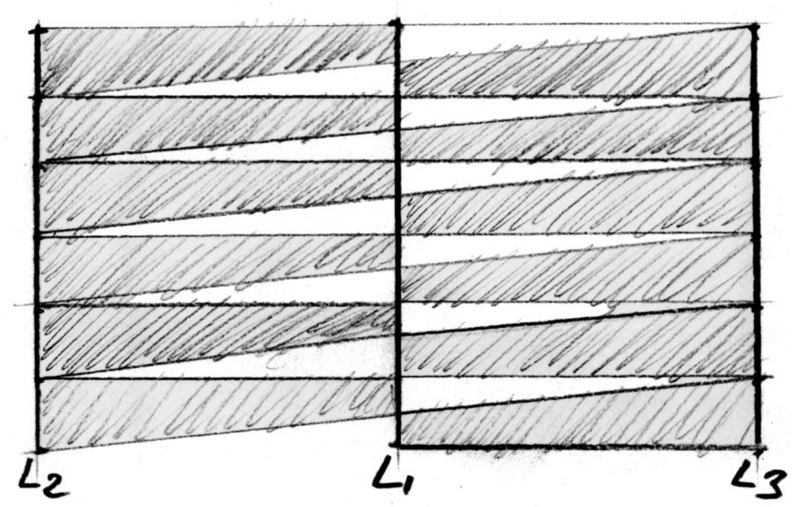


Transport from a measure concentrated on L_1 onto another one concentrated on L_2 and L_3

Part. 2 Optimal Transport – Monge's problem

<u>Monge's problem:</u>

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

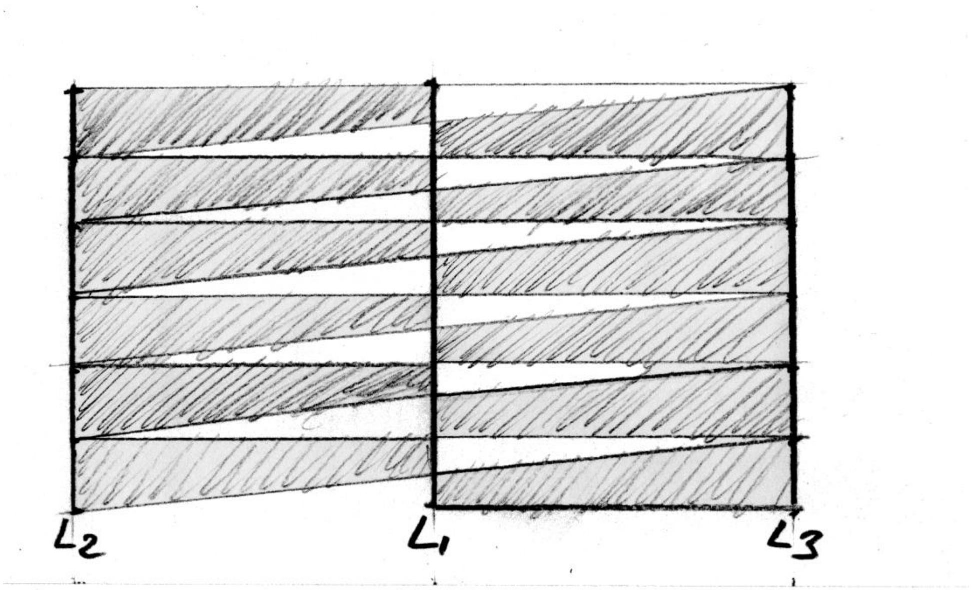


Transport from a measure concentrated on L_1 onto another one concentrated on L_2 and L_3

Part. 2 Optimal Transport – Monge's problem

Monge's problem:

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



Transport from a measure concentrated on L_1 onto another one concentrated on L_2 and L_3

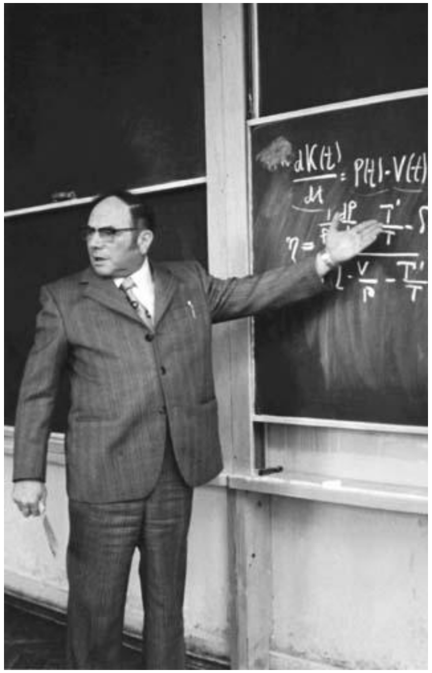
The infimum is never realized by a map, need for a relaxation

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 $X \times Y$

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

<u>Monge's problem:</u>

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

<u>Kantorovich's problem:</u>

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

"$\gamma(x,y)$" :
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

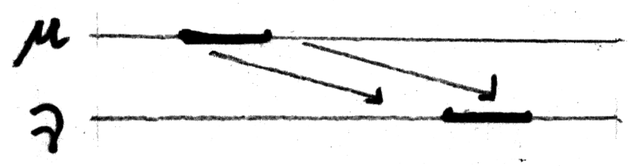
such that $\int_{x \text{ in } X} d\gamma(x,y) = d\nu(y)$

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

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

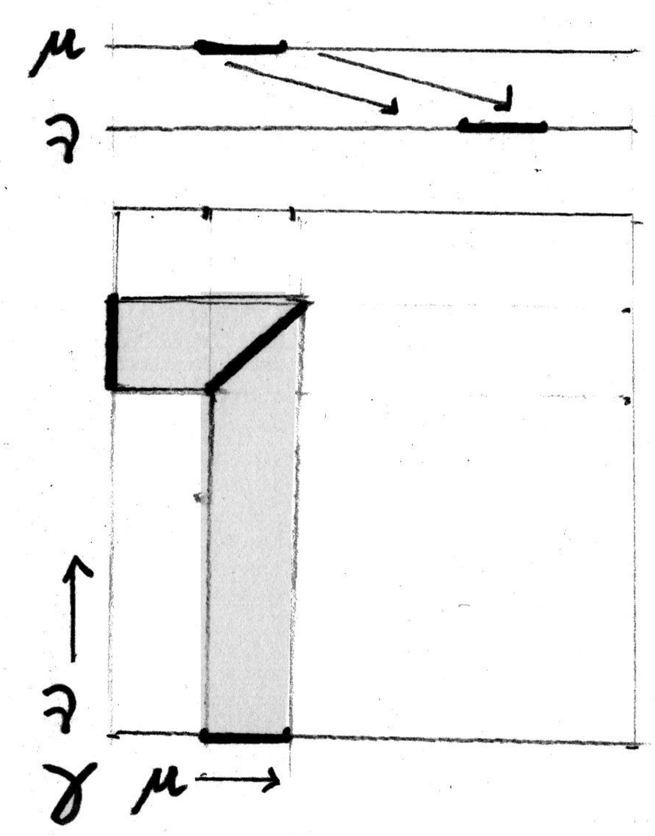
that minimizes $\iint_{X \times Y} \| x - y \|^2 \, d\gamma(x,y)$

Part. 2 Optimal Transport – Kantorovich



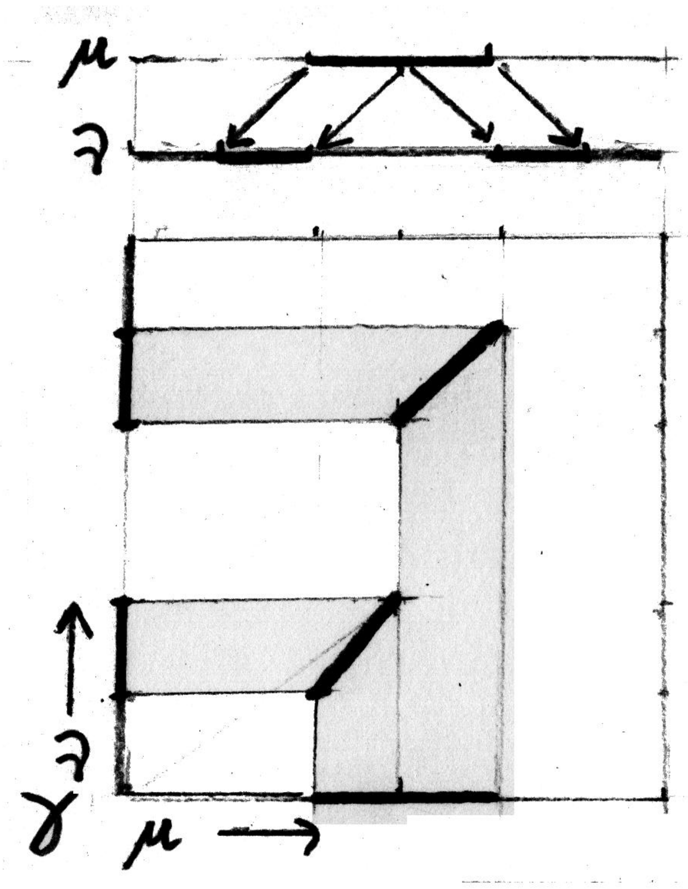
Transport plan – example 1/4 : translation of a segment

Part. 2 Optimal Transport – Kantorovich



Transport plan – example 1/4 : translation of a segment

Part. 2 Optimal Transport – Kantorovich



Transport plan – example 2/4 : spitting a segment

Part. 2 Optimal Transport – Kantorovich

Observation 1. *If* $(Id \times T)\sharp\mu \in \pi(\mu, \nu)$, *then* T *pushes* μ *to* ν.

Part. 2 Optimal Transport – Kantorovich

Observation 1. *If* $(Id \times T)\sharp\mu \in \pi(\mu, \nu)$, *then* T *pushes* μ *to* ν.

Proof. $(Id{\times}T)\sharp\mu$ belongs to $\pi(\mu,\nu)$, therefore $(P_2)\sharp(Id{\times}T)\sharp\mu = \nu$, or $((P_2) \circ (Id \times T))\,\sharp\mu = \nu$, thus $T\sharp\mu = \nu$ $\square$

Part. 2 Optimal Transport – Kantorovich

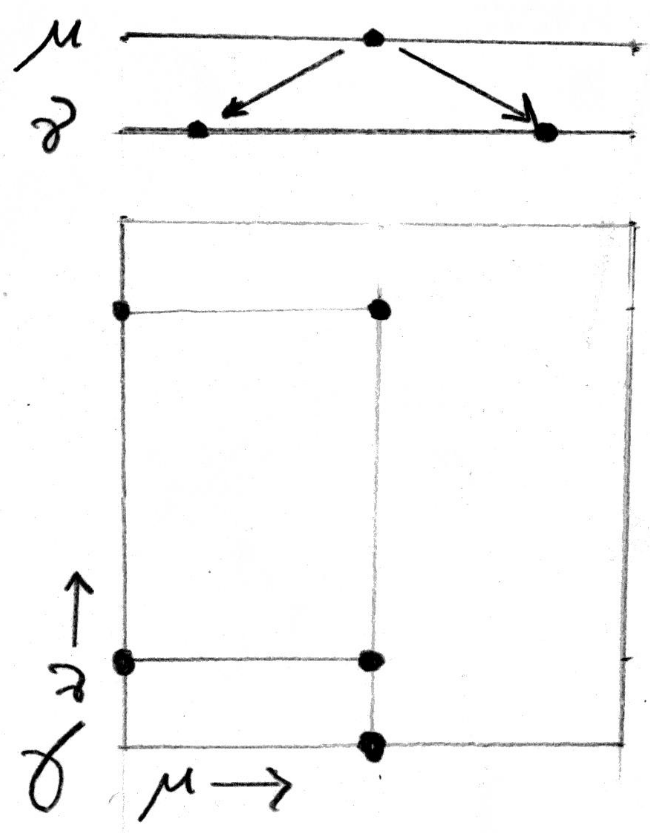
Observation 1. *If* $(Id \times T)\sharp\mu \in \pi(\mu,\nu)$, *then* T *pushes* μ *to* ν.

Proof. $(Id \times T)\sharp\mu$ belongs to $\pi(\mu,\nu)$, therefore $(P_2)\sharp(Id \times T)\sharp\mu = \nu$, or $((P_2)\circ(Id \times T))\,\sharp\mu = \nu$, thus $T\sharp\mu = \nu$ $\square$

With this observation, for transport plans of the form $\gamma = (Id \times T)\sharp\mu$, (K) becomes

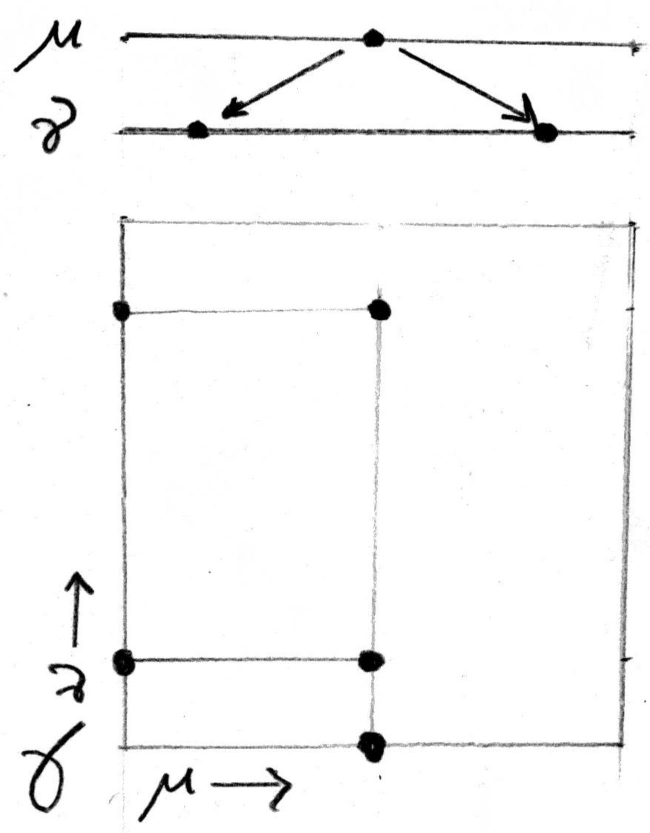
$$\min\left\{\int_{\Omega\times\Omega} c(x,y)d\left((Id \times T)\sharp\mu\right)\right\} \quad = \quad \min\left\{\int_{\Omega} c(x,T(x))d\mu\right)$$

Part. 2 Optimal Transport – Kantorovich



Transport plan – example 3/4 : splitting a Dirac into two Diracs

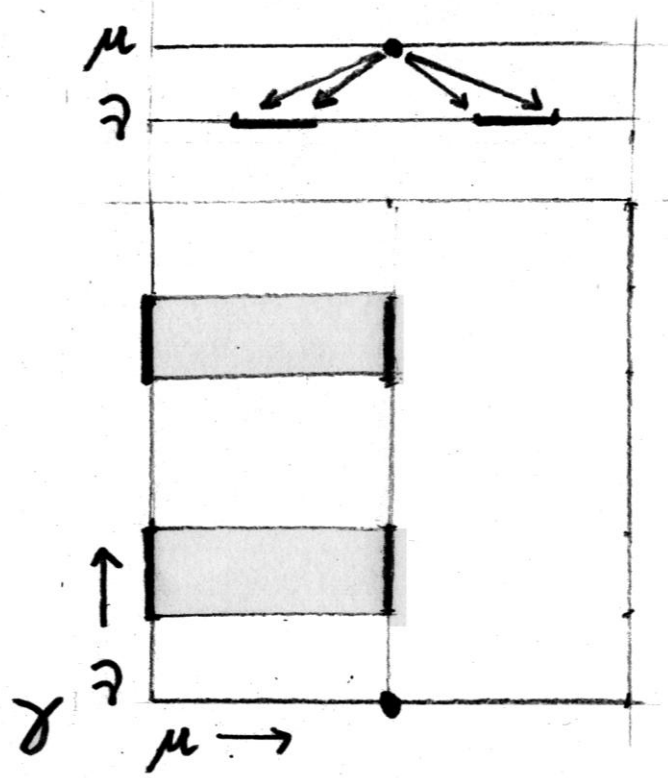
Part. 2 Optimal Transport – Kantorovich



Transport plan – example 3/4 : splitting a Dirac into two Diracs

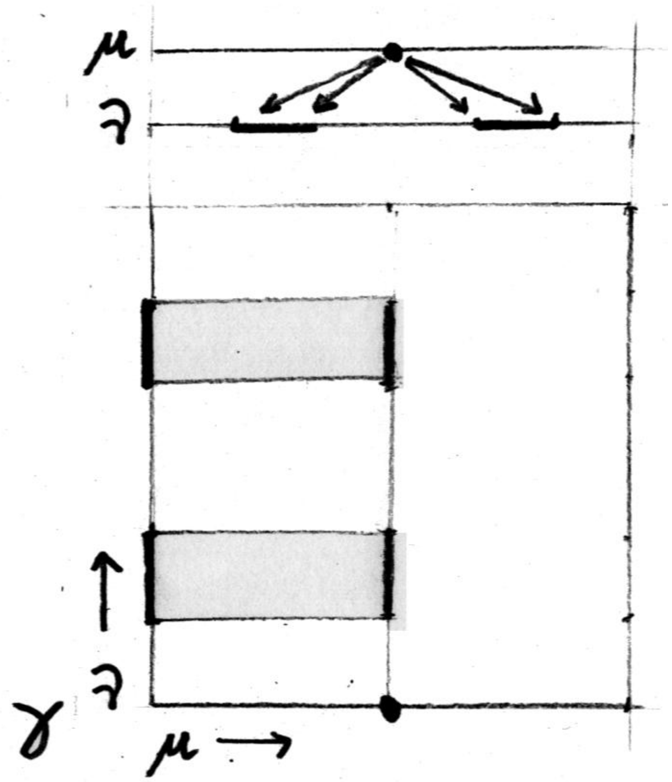
(No transport map)

Part. 2 Optimal Transport – Kantorovich



Transport plan – example 4/4 : splitting a Dirac into two segments

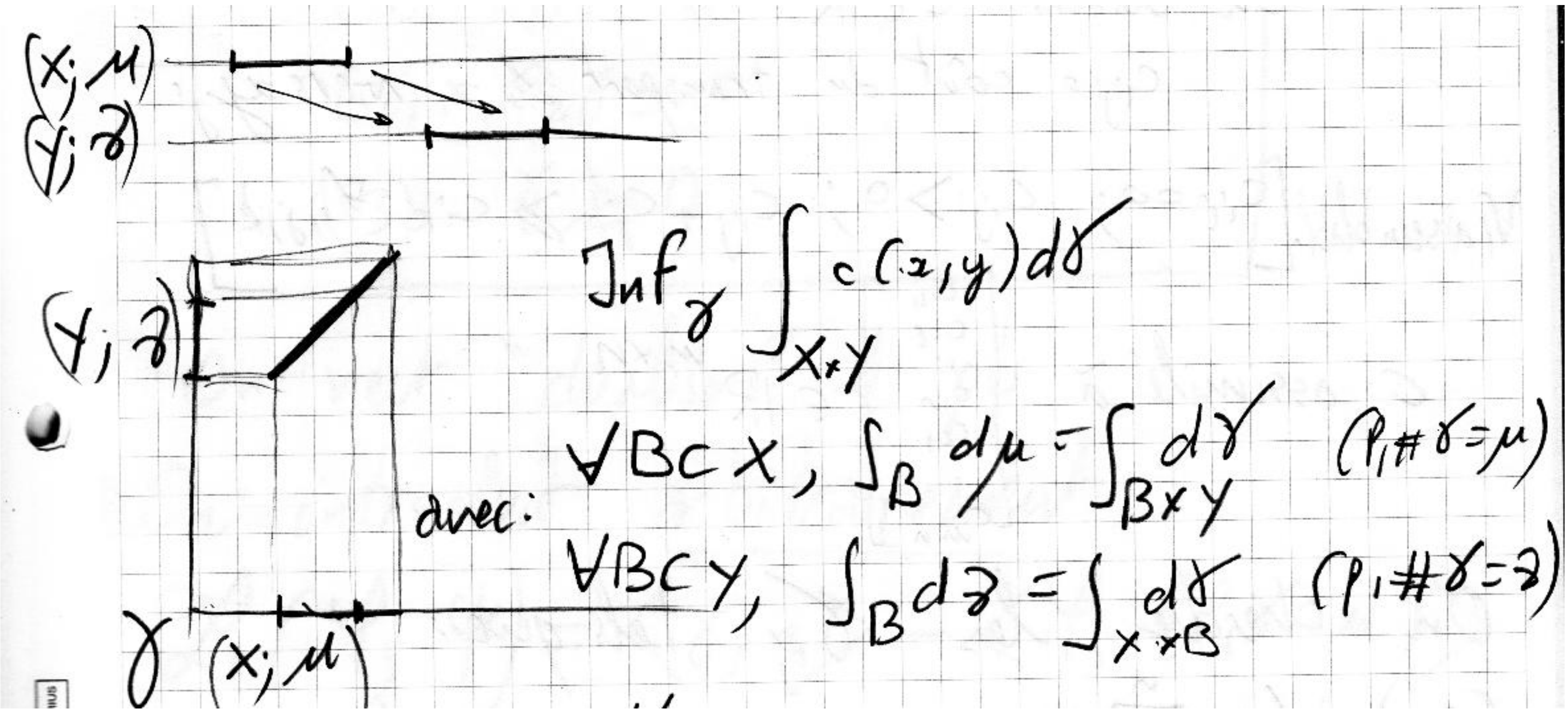
Part. 2 Optimal Transport – Kantorovich



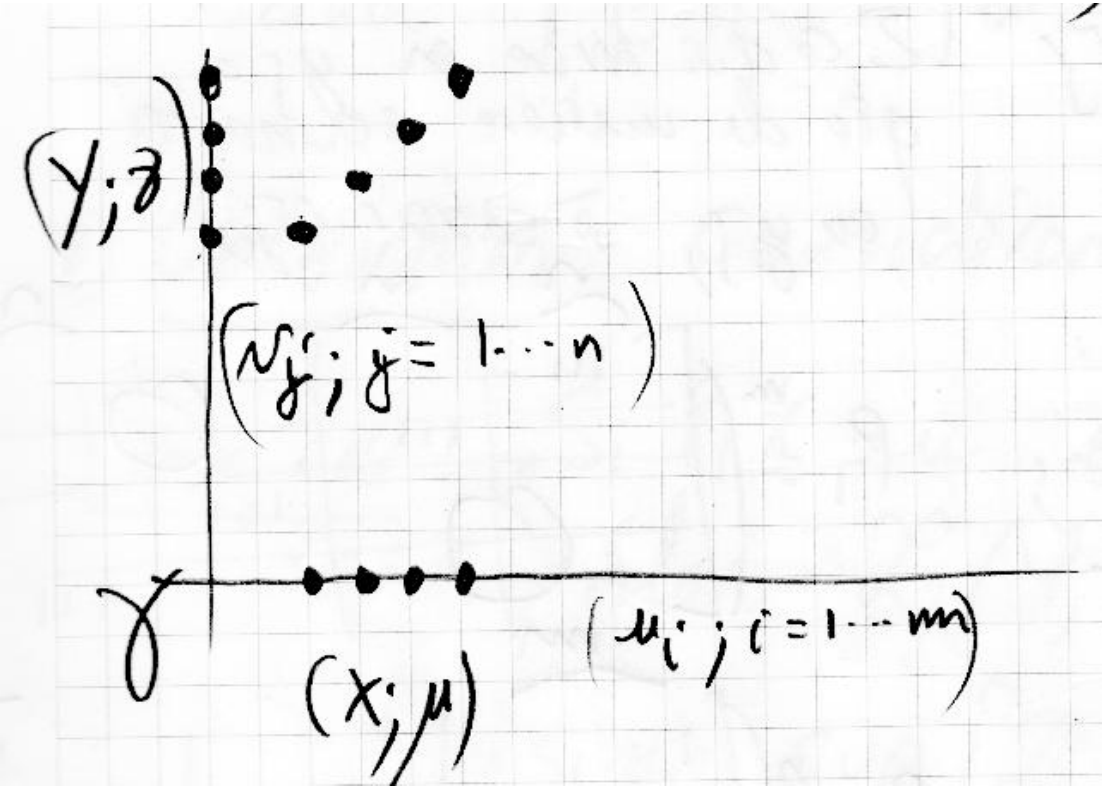
Transport plan – example 4/4 : splitting a Dirac into two segments

(No transport map)

Part. 2 Optimal Transport – Duality

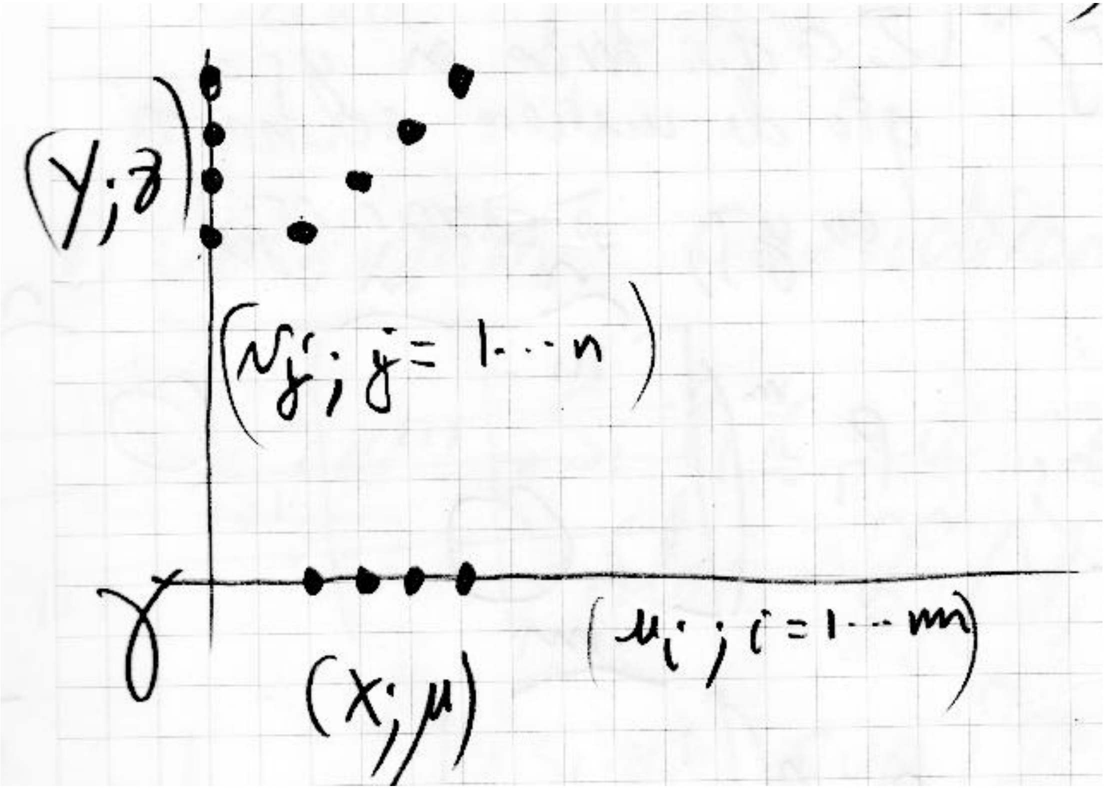


Part. 2 Optimal Transport – Duality



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

Part. 2 Optimal Transport – Duality

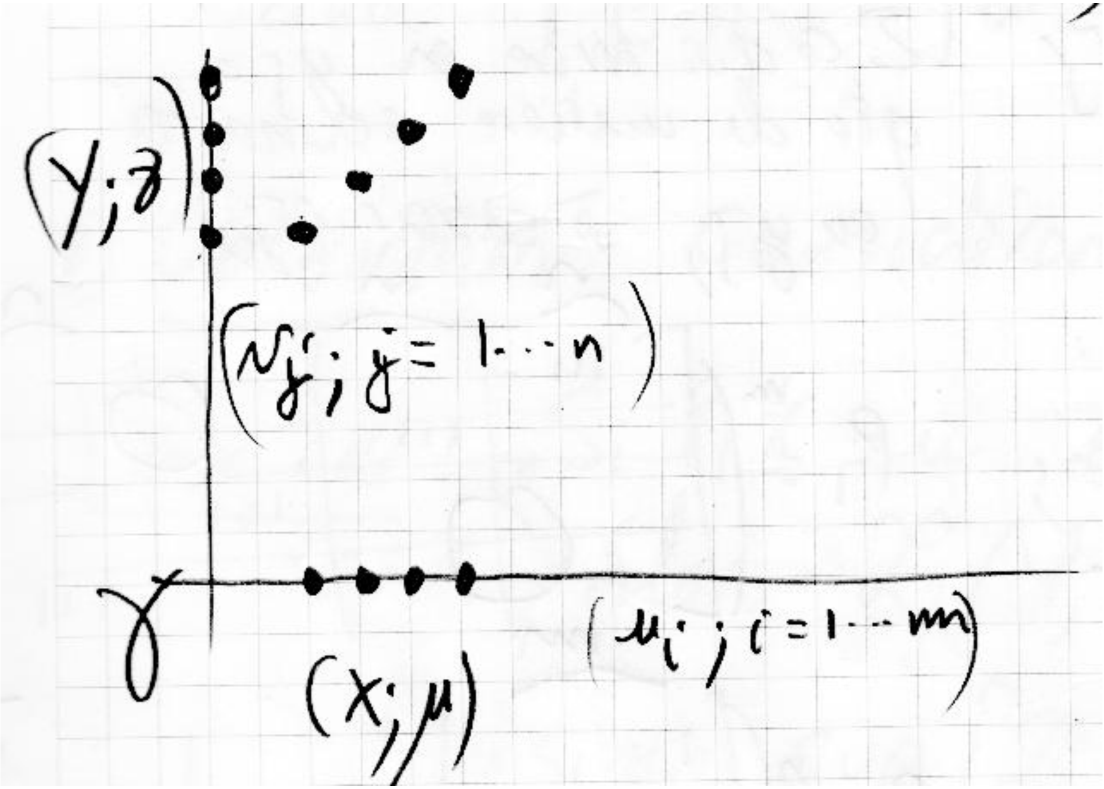


(DMK):

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

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

Part. 2 Optimal Transport – Duality



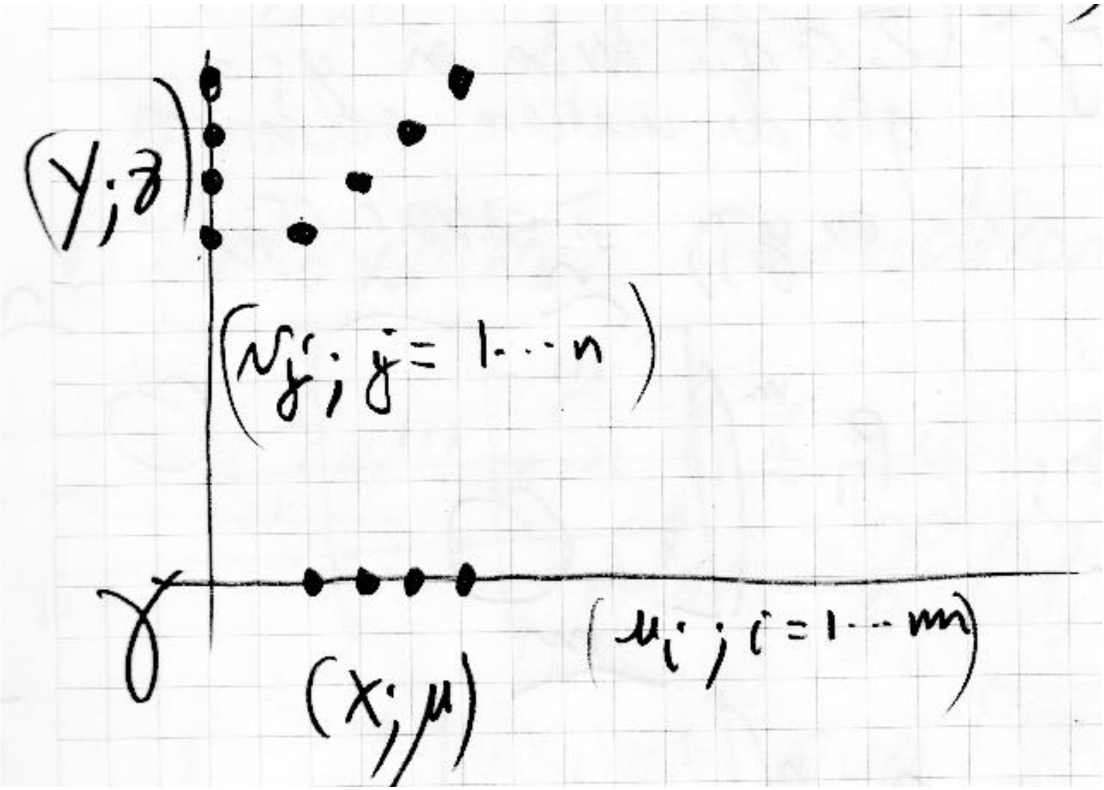
(DMK):

Min $\langle c, \gamma \rangle$

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

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

Part. 2 Optimal Transport – Duality



(DMK):

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

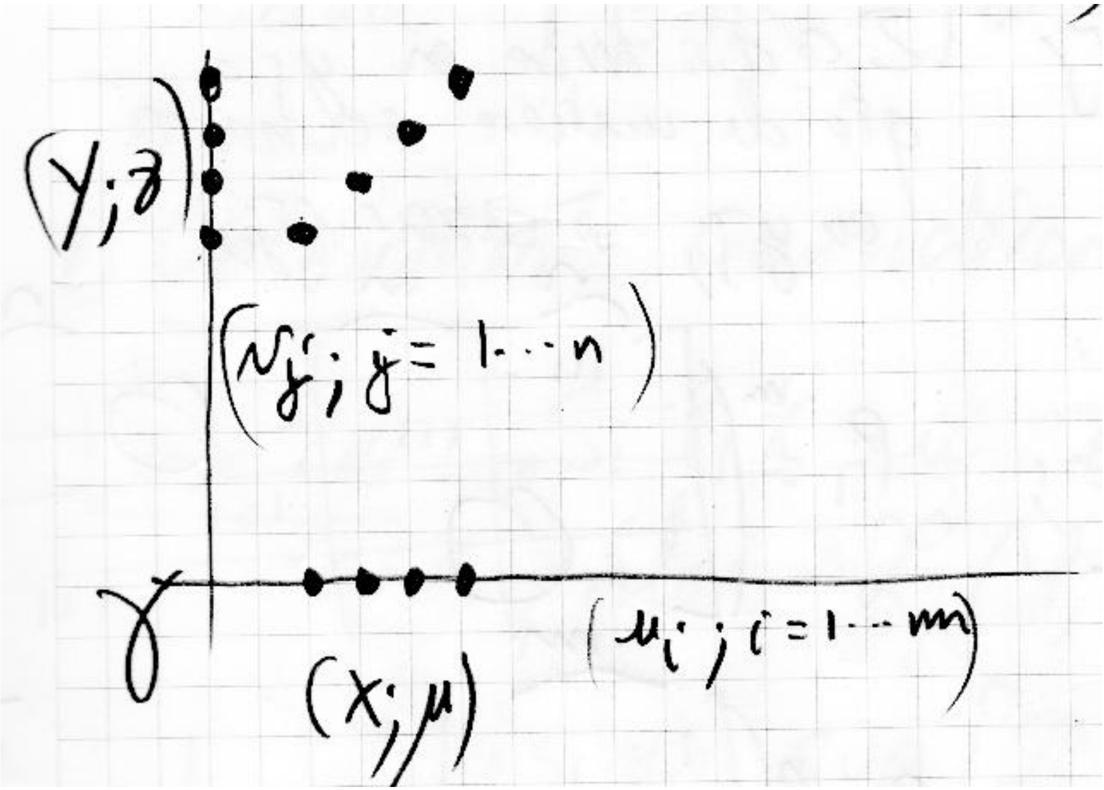
$$s.t. \begin{cases} P_1 \gamma = u \\ P_2 \gamma = v \\ \gamma \geq 0 \end{cases}$$

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

Part. 2 Optimal Transport – Duality

(DMK):

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



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

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

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

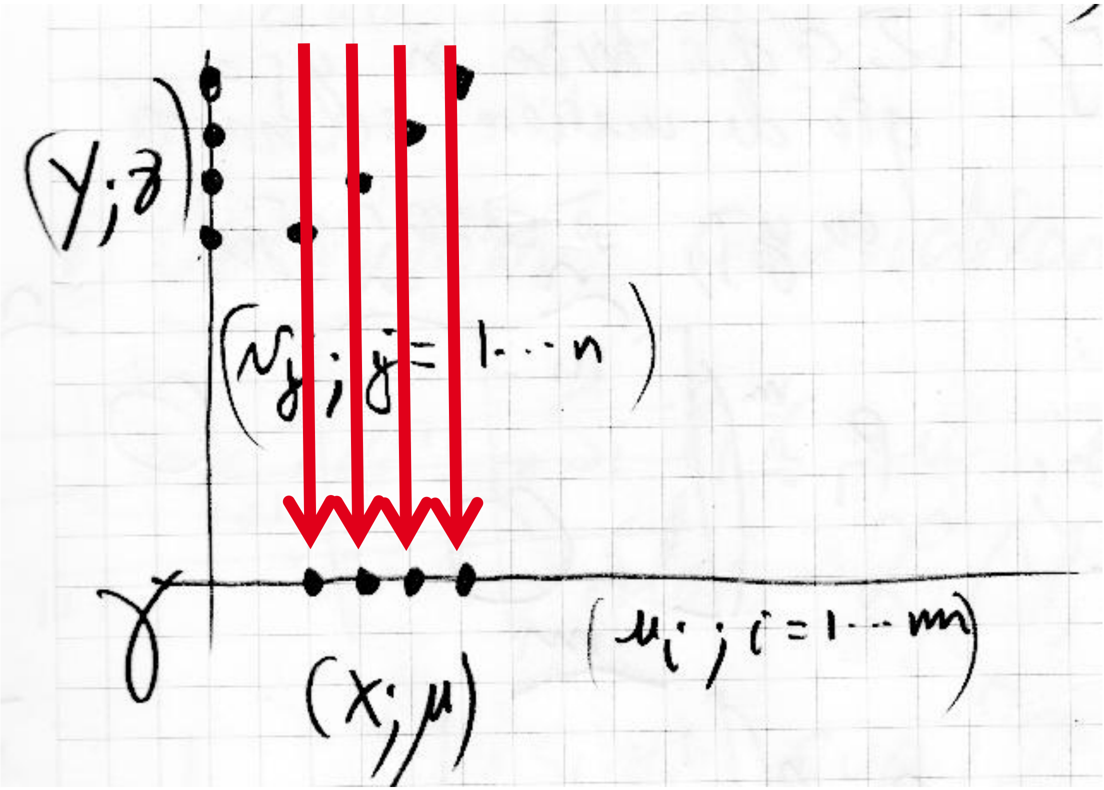
Part. 2 Optimal Transport – Duality

(DMK):

Min $\langle c, \gamma \rangle$

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

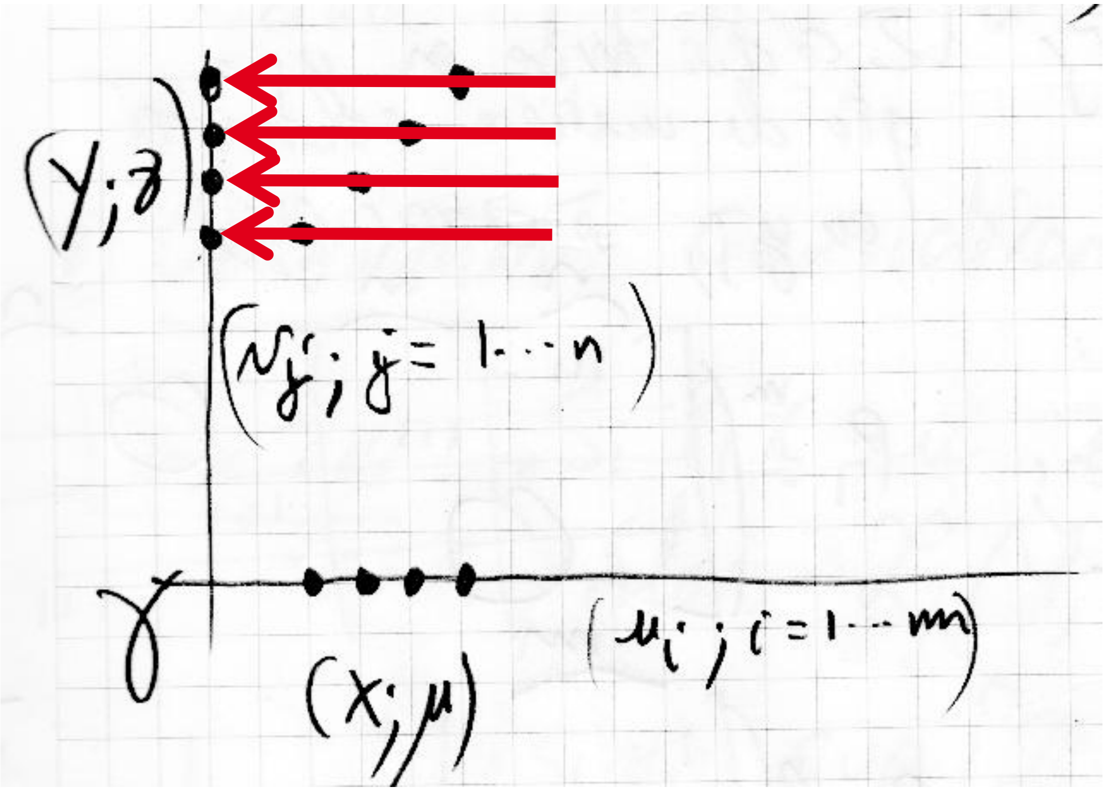
mn X m → P_1



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

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

Part. 2 Optimal Transport – Duality



(DMK):

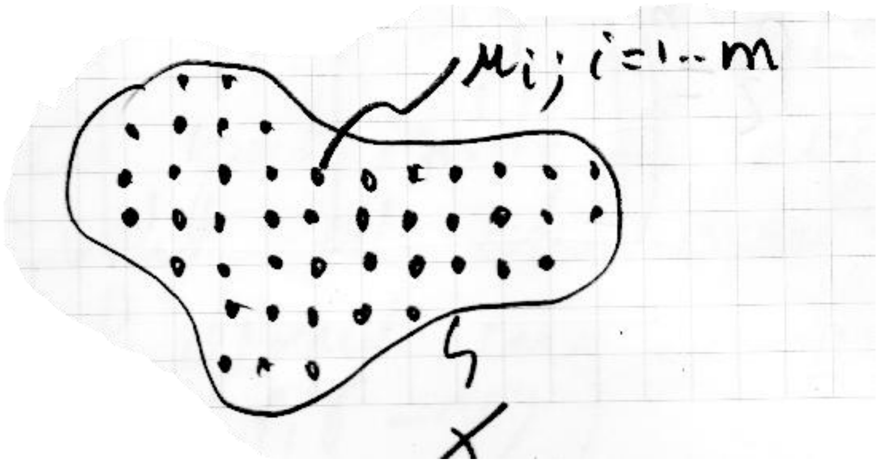
Min $\langle c, \gamma \rangle$

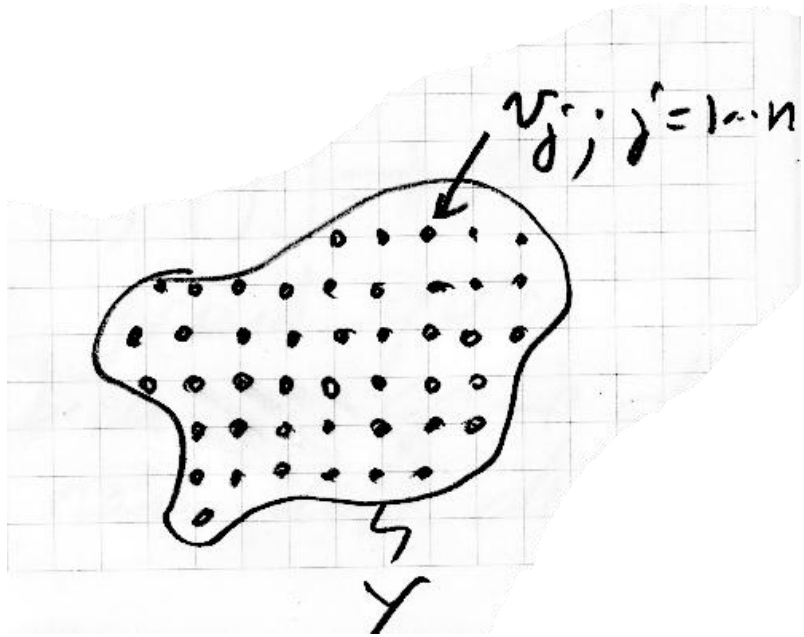
s.t.
- mn X m → $P_1 \gamma = u$
- **mn X n** → $\mathbf{P_2} \gamma = v$
- $\gamma \geq 0$

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

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

Part. 2 Optimal Transport – Duality





(DMK):

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

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

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

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

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

$= +\infty$ otherwise

Consider now: $\inf_{\gamma \geq 0} \left[\sup_{\substack{\varphi \in IR^m \\ \psi \in IR^n}} [\, \mathcal{L}(\varphi,\psi) \,] \right]$

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

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

$= +\infty$ otherwise

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

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

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

$= +\infty$ otherwise

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

Part. 2 Optimal Transport – Duality

(DMK):

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

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

Part. 2 Optimal Transport – Duality

(DMK):

Min $\langle c, \gamma \rangle$

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

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

Exchange Inf Sup

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

Part. 2 Optimal Transport – Duality

(DMK):

Min $\langle c, \gamma \rangle$

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

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

Exchange Inf Sup

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

Expand/Reorder/Collect

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

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):

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

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

Interpret

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

Part. 2 Optimal Transport – Duality

(DMK):

Min $<c, \gamma>$

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

$$\operatorname*{Sup}_{\substack{\varphi \in IR^m \\ \psi \in IR^n}} \Big[\operatorname*{Inf}_{\gamma \geq 0} \big[< \gamma, c - P_1^t \varphi - P_2^t \psi > + <\varphi, u> + <\psi, v> \big] \Big]$$

Interpret

$$\operatorname*{Sup}_{\substack{\varphi \in IR^m \\ \psi \in IR^n \\ P_1^t \varphi + P_2^t \psi \leq c}} \Big[<\varphi, u> + <\psi, v> \Big]$$

(DDMK)

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

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

Dual formulation of Kantorovich's problem (Continuous):

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 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, $\varphi(x) + \psi(y) \leq ½\| x - y \|^2$

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 $X \times Y$
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 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 $\int_X \varphi d\mu + \int_Y \psi d\nu$

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

Part. 2 Optimal Transport – Kantorovich dual

Kantorovich's problem:

Find a measure γ 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 φ in $L^1(\mu)$ and ψ in $L^1(\nu)$
Such that for all x,y, $\varphi(x) + \psi(y) \leq ½\| x - y \|^2$

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

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, $\varphi(x) + \psi(y) \leq ½|| x - y ||^2$

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

Part. 2 Optimal Transport – c-conjugate functions

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

If we got two functions φ and ψ that satisfy the constraint

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

Part. 2 Optimal Transport – c-conjugate functions

Dual formulation of Kantorovich's problem:

Find two functions φ in $L^1(\mu)$ and ψ in $L^1(\nu)$
Such that for all x,y, $\varphi(x) + \psi(y) \leq ½\| x - y \|^2$

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

If we got two functions φ and ψ that satisfy the constraint

Then it is possible to obtain a better solution by replacing ψ with φ^c defined by:
For all y, $\varphi^c(y) = \inf_{x \in X} ½\| x - y \|^2 - \varphi(y)$

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

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 ψ

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

ψ is called a **"Kantorovich potential"**

Part. 2 Optimal Transport – c-subdifferential

Dual formulation of Kantorovich's problem:

Find a c-concave function ψ

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

ψ 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 ψ

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

ψ 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* ψ.

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* ψ.

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* ψ.

Proof: see OTON, chap. 10.

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

Part. 2 Optimal Transport – c-subdifferential

Theorem 1.

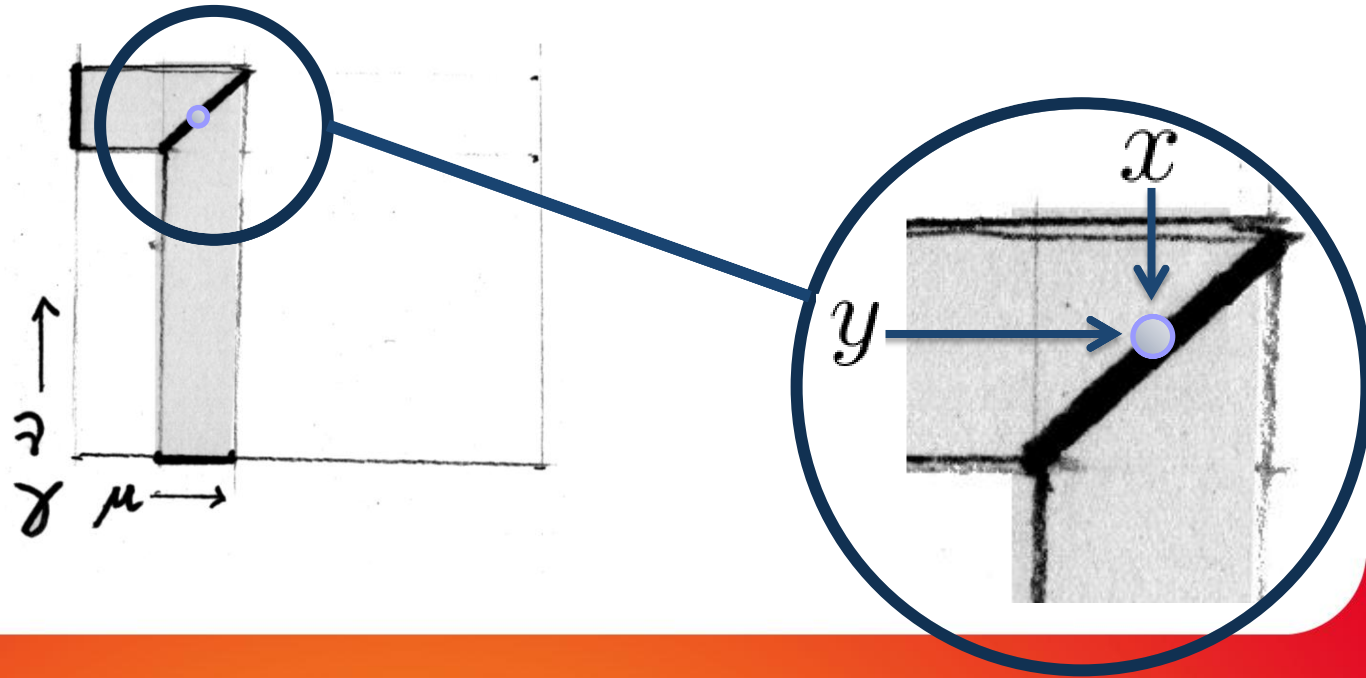
$$\forall (x,y) \in \partial_c \psi, \nabla \psi(x) - \nabla_x c(x,y) = 0$$

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

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.

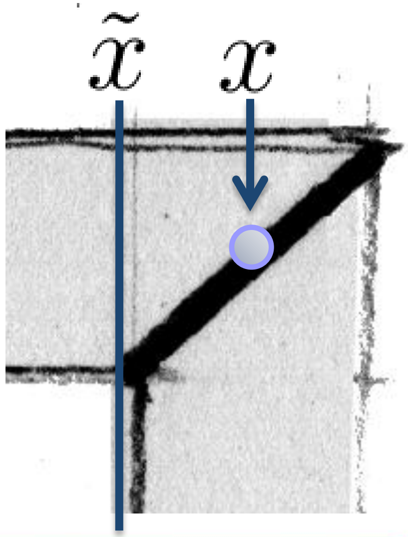
$$\forall (x,y) \in \partial_c \psi, \nabla \psi(x) - \nabla_x c(x,y) = 0$$

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

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.

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

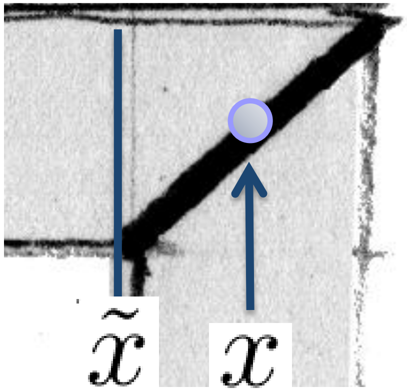
where $\partial_c\psi = \{(x,y)|\phi(x) + \psi(y) = c(x,y)\}$ *denotes the so-called* c-subdifferential *of* ψ.

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 ϵ and starting at x. One can compute the gradient along an arbitrary direction w by taking the limit when ϵ 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.

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

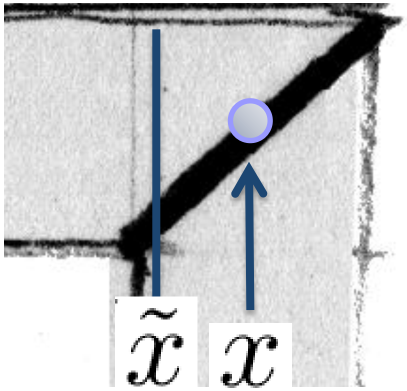
where $\partial_c \psi = \{(x,y) | \phi(x) + \psi(y) = c(x,y)\}$ *denotes the so-called* c-subdifferential *of* ψ.

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 ϵ and starting at x. One can compute the gradient along an arbitrary direction w by taking the limit when ϵ 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.

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

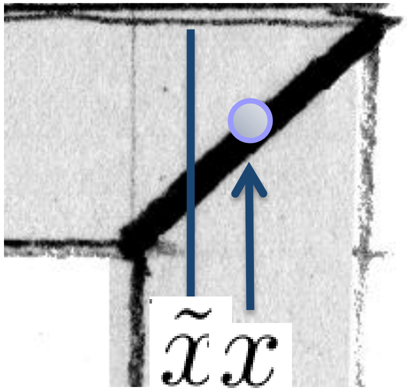
where $\partial_c \psi = \{(x,y) | \phi(x) + \psi(y) = c(x,y)\}$ *denotes the so-called* c-subdifferential *of* ψ.

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 ϵ and starting at x. One can compute the gradient along an arbitrary direction w by taking the limit when ϵ 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.

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

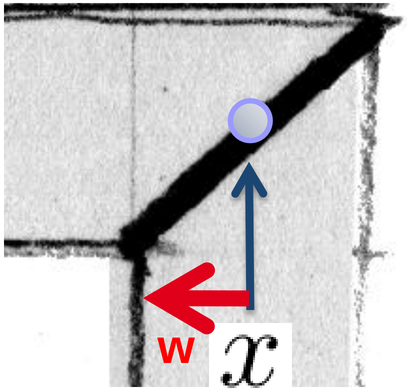
where $\partial_c \psi = \{(x,y) | \phi(x) + \psi(y) = c(x,y)\}$ *denotes the so-called* c-subdifferential *of* ψ.

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 ϵ and starting at x. One can compute the gradient along an arbitrary direction w by taking the limit when ϵ 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.

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

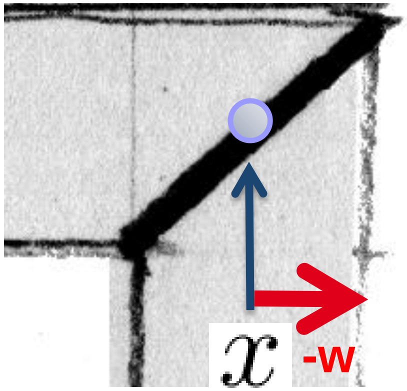
where $\partial_c \psi = \{(x,y) | \phi(x) + \psi(y) = c(x,y)\}$ *denotes the so-called* c-subdifferential *of* ψ.

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 ϵ and starting at x. One can compute the gradient along an arbitrary direction w by taking the limit when ϵ 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 $\nabla \psi(x) \cdot w \leq \nabla_x c(x,y) \cdot w$
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 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 $T(x) = x - \nabla\psi(x) = \nabla(\|x\|^2/2 - \psi(x))$.

Part. 2 Optimal Transport – c-subdifferential

<u>Dual formulation of Kantorovich's problem:</u>

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

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

$\overline{\psi}$ 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 x²- ψ(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:

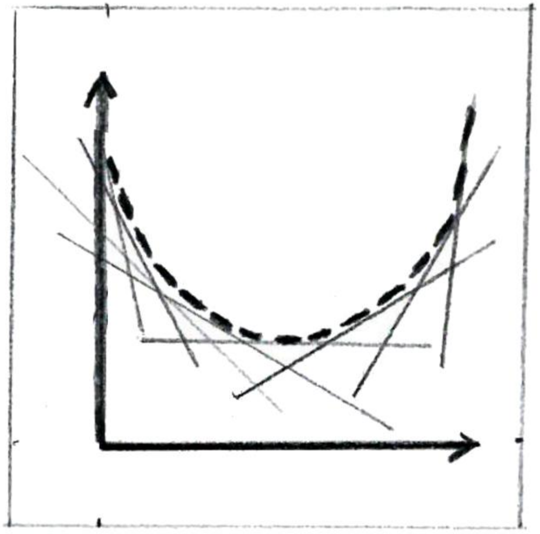
Find a c-concave function ψ

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 $\bar{\psi}(x) := (½\, x^2 - \psi(x))$

$\bar{\psi}$ 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 ψ

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"

Part. 2 Optimal Transport – no collision

Dual formulation of Kantorovich's problem:

Find a c-concave function ψ

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

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"

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 ψ

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:

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

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

Part. 2 Optimal Transport – Monge-Ampere

Dual formulation of Kantorovich's problem:

Find a c-concave function ψ

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

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

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

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

for all borel set A, $\int_A d\mu = \int_{T(A)} (|JT|)\, dv$ (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 ψ

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

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

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

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

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

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

Part. 2 Optimal Transport – Monge-Ampere

Dual formulation of Kantorovich's problem:

Find a c-concave function ψ

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

What about our initial problem ?

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

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

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

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

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

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

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

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

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

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

Part. 2 Optimal Transport – Isoperimetric inequality

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

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

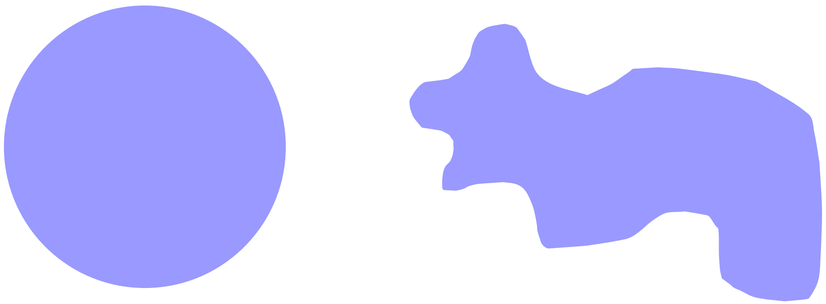
Explanation in **[Dario Cordero Erauquin]** course notes

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 $Vol(\Omega) = Vol(B_2^3)$
and f = the indicatrix function of Ω

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



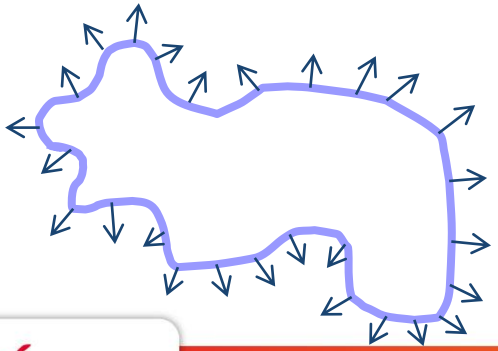
Part. 2 Optimal Transport – Isoperimetric inequality

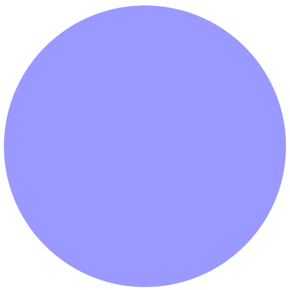
L_1 Sobolev inegality: Given $f: \mathbb{R}^n \rightarrow \mathbb{R}$ sufficiently regular

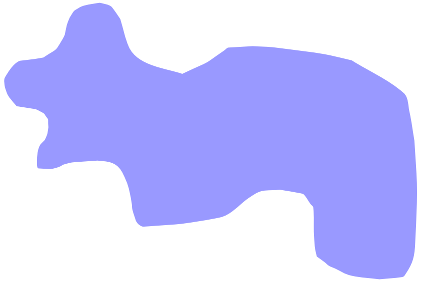
Consider a compact set Ω such that $\mathrm{Vol}(\Omega) = \mathrm{Vol}(B_2^3)$
and f = the indicatrix function of Ω

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

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







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

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

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

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

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

Part. 2 Optimal Transport – Isoperimetric inequality

L_1 Sobolev inegality: a proof with OT [Gromov]

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

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 $\int f^{\,n/(n-1)} = 1$

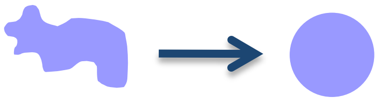
Part. 2 Optimal Transport – Isoperimetric inequality

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 $\mathbf{1}B_2^n/\mathrm{Vol}(B_2^n)dx$



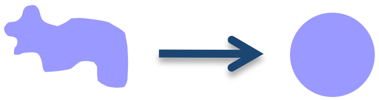
Part. 2 Optimal Transport – Isoperimetric inequality

L_1 Sobolev inegality: a proof with OT [Gromov]

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

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

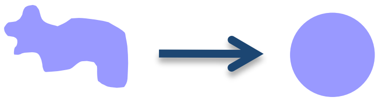
Part. 2 Optimal Transport – Isoperimetric inequality

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

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

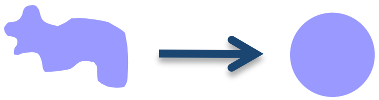
Part. 2 Optimal Transport – Isoperimetric inequality

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

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

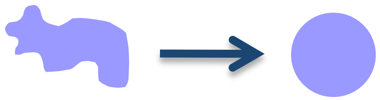
Part. 2 Optimal Transport – Isoperimetric inequality

L_1 Sobolev inegality: a proof with OT [Gromov]

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

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

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

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

Part. 2 Optimal Transport – Isoperimetric inequality

L_1 Sobolev inegality: a proof with OT [Gromov]

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

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

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

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

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 $\int f^{\, n/(n-1)} = 1$

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

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

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

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 $\int f^{\ n/(n-1)} = 1$

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

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

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

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

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

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

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

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 $\int f^{\,n/(n-1)} = 1$

Monge-Ampère equation:

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

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

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

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

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 $\int f^{\,n/(n-1)} = 1$

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

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

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

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

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

3

Semi-Discrete Optimal Transport

Part. 3 Optimal Transport – how to program ?

Continuous

(X;μ)

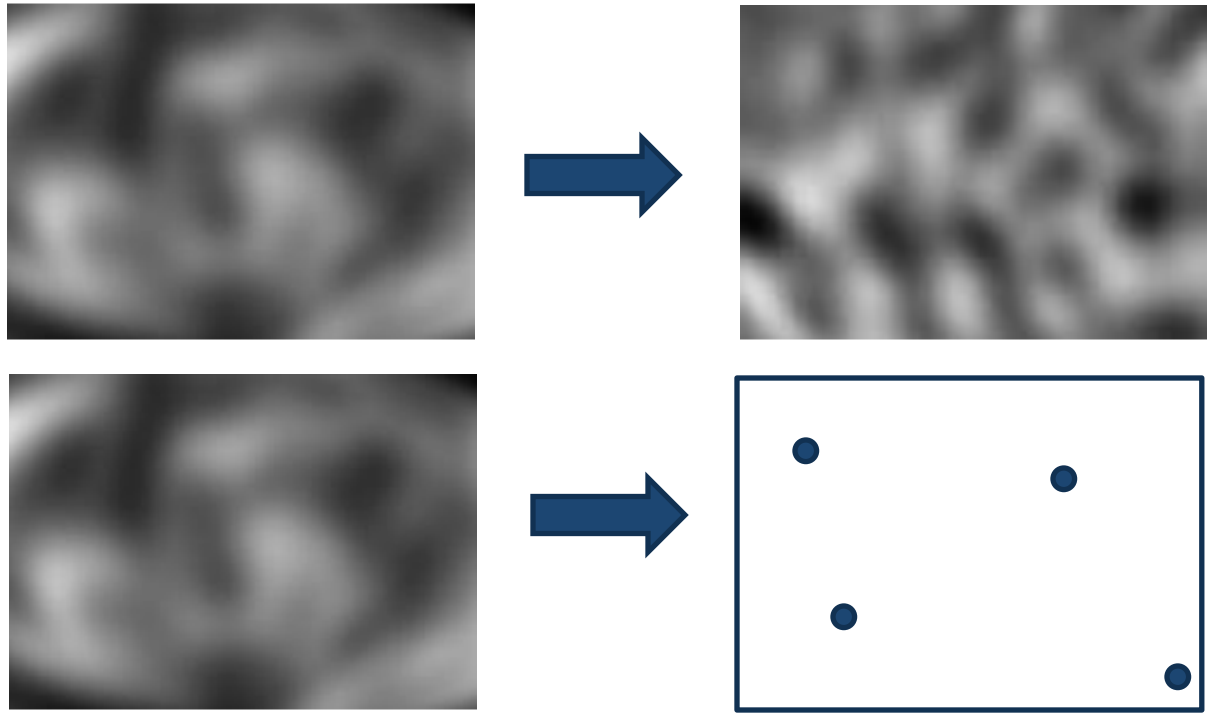
(Y;ν)

Part. 3 Optimal Transport – how to program ?

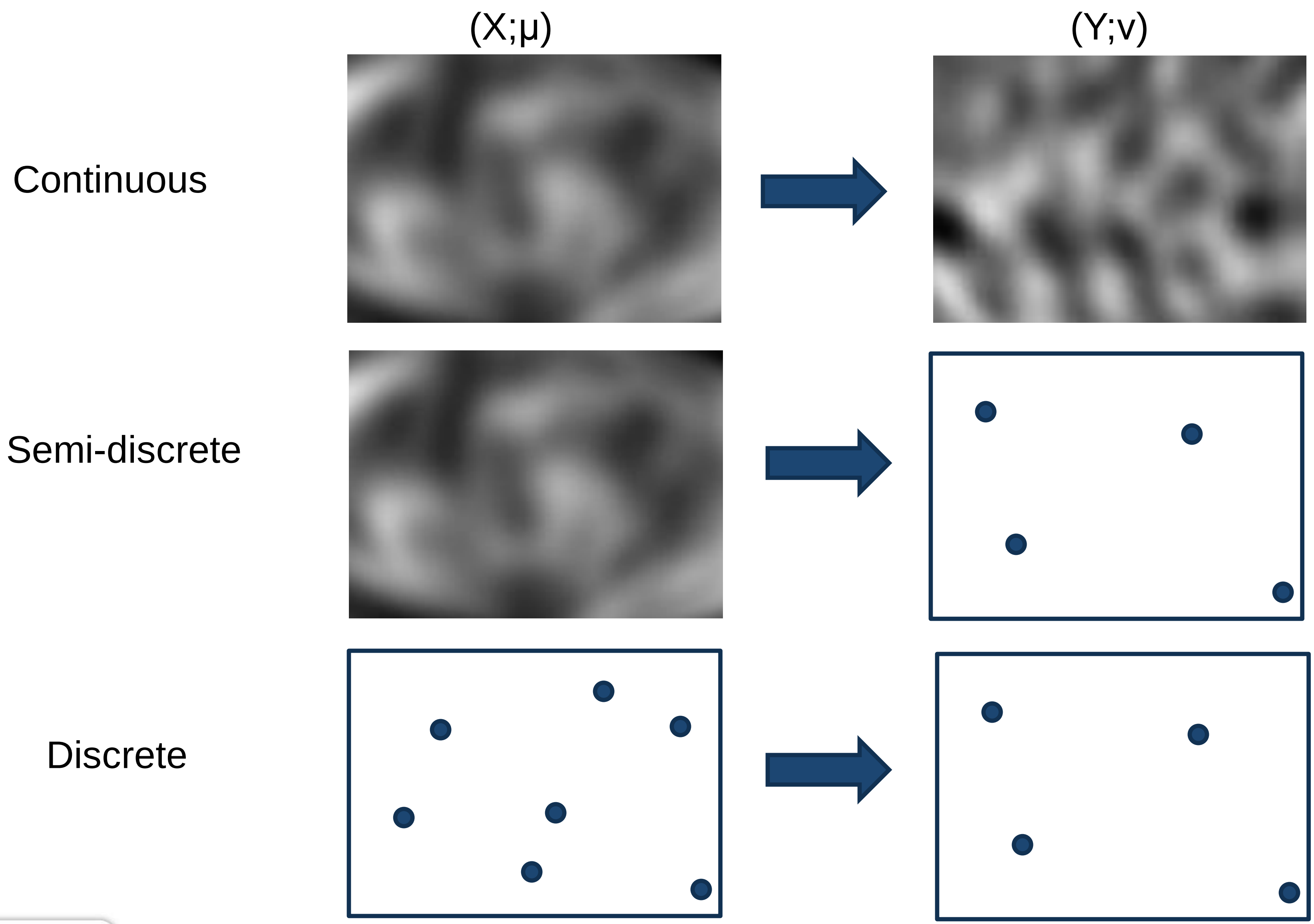
(X;μ) (Y;ν)

Continuous

Semi-discrete



Part. 3 Optimal Transport – how to program ?



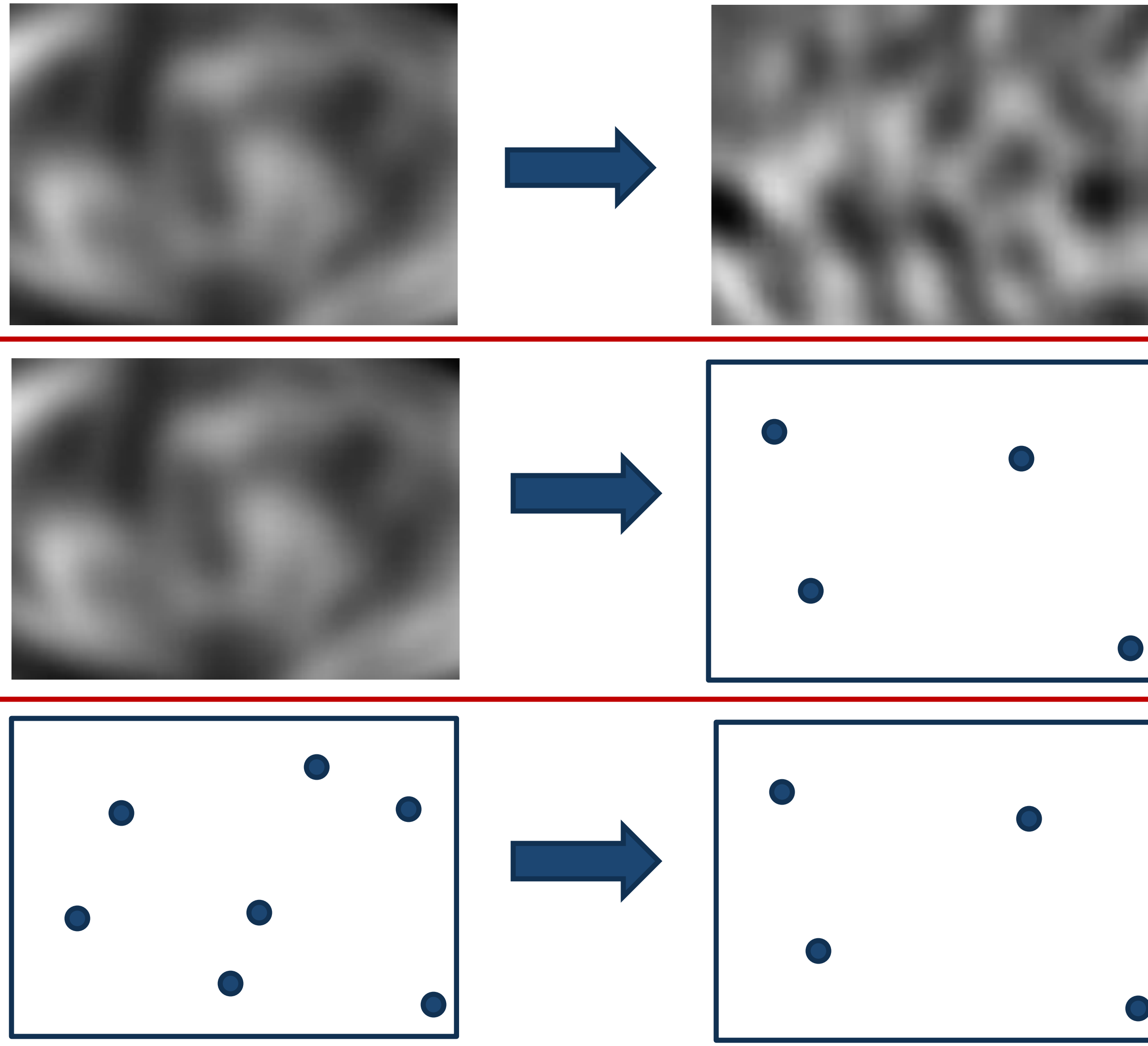
Part. 3 Optimal Transport – how to program ?

(X;μ) (Y;ν)

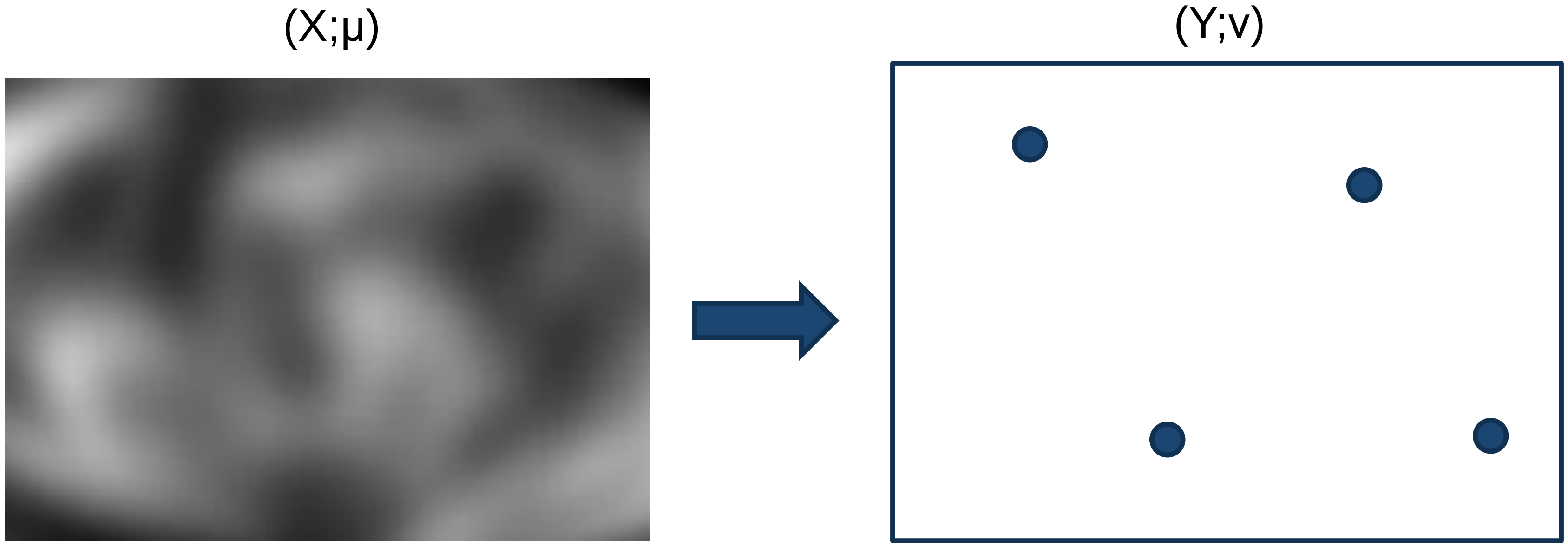
Continuous

Semi-discrete

Discrete

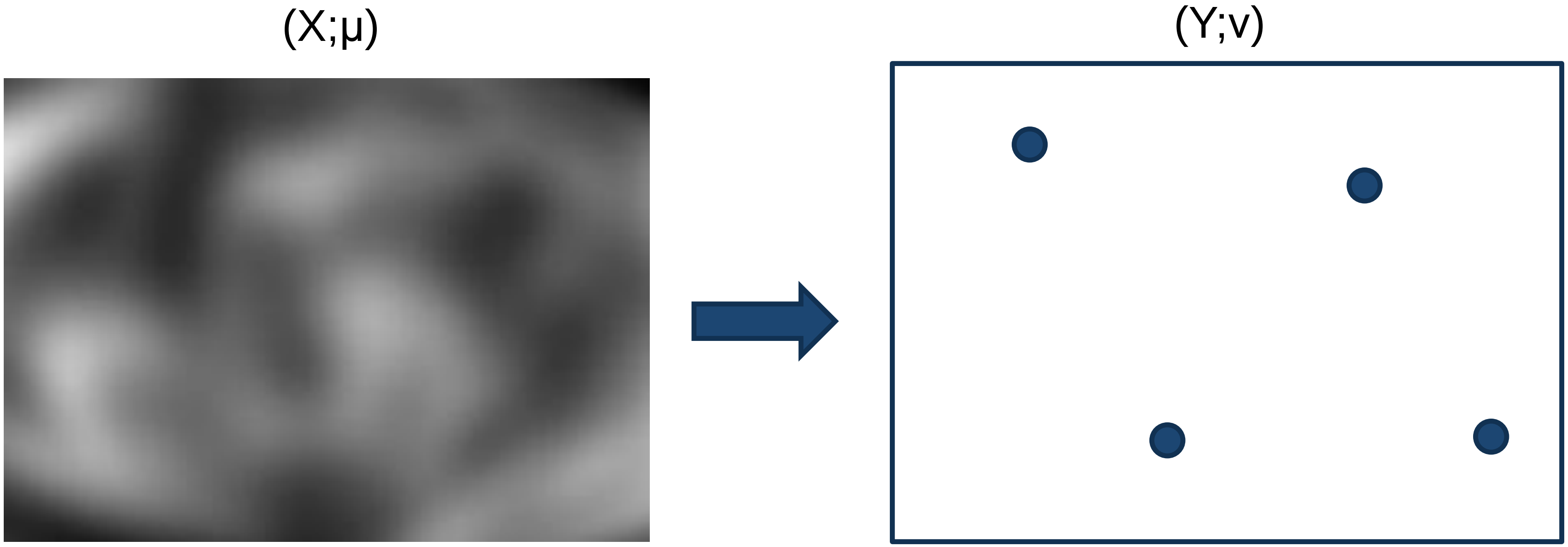


Part. 3 Optimal Transport – semi-discrete



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

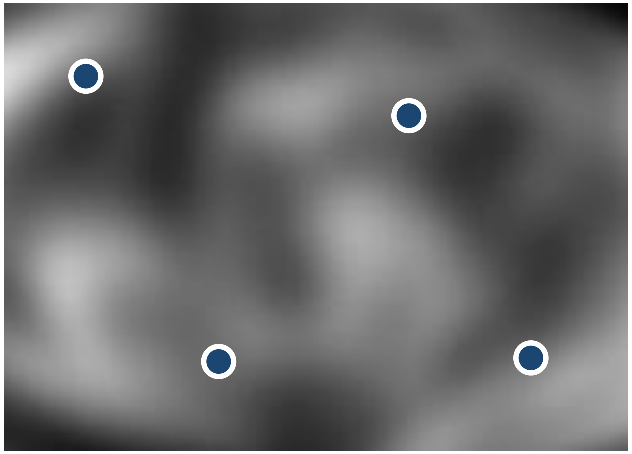
Part. 3 Optimal Transport – semi-discrete



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

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

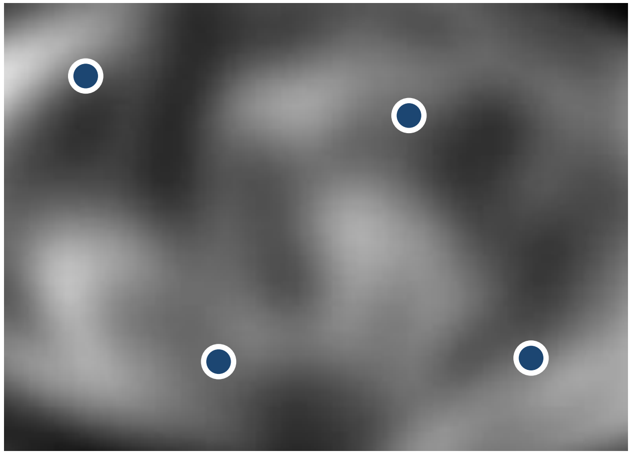
Part. 3 Optimal Transport – semi-discrete



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

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

Part. 3 Optimal Transport – semi-discrete

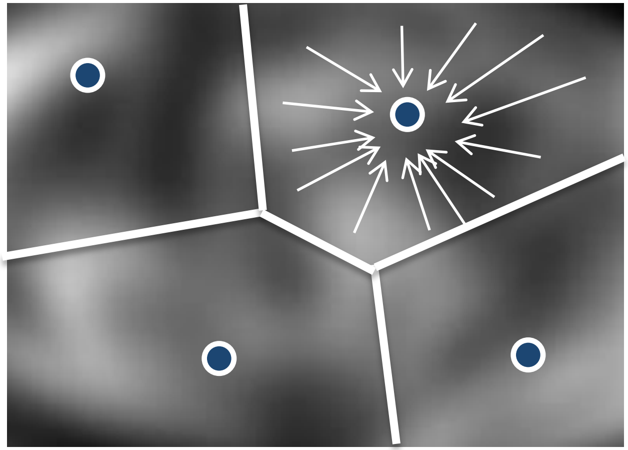


(DMK) $\sup_{\psi \in \Psi^c} \int_X \psi^c(x) d\mu + \int_Y \psi(y) d\nu$

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

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

Part. 3 Optimal Transport – semi-discrete



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

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

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

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

Part. 3 Optimal Transport – semi-discrete

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

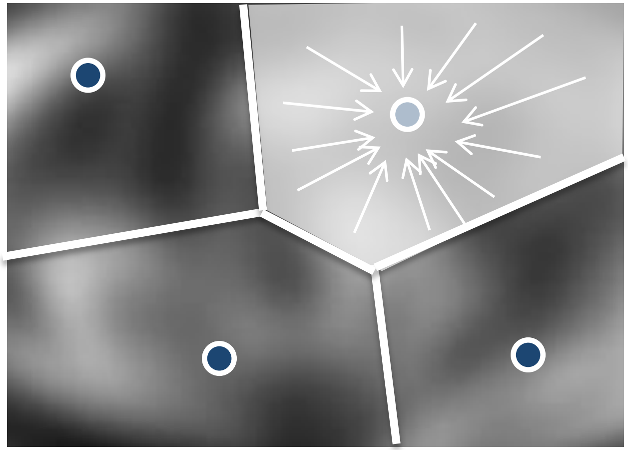
Where: $Lag_{\psi}(y_j) = \left\{ x \mid \| x - y_j \|^2 - \psi(y_j) < \| x - y_{j'} \|^2 - \psi(y_{j'}) \right\}$ for all $j' \neq j$

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 $j' \neq j$

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



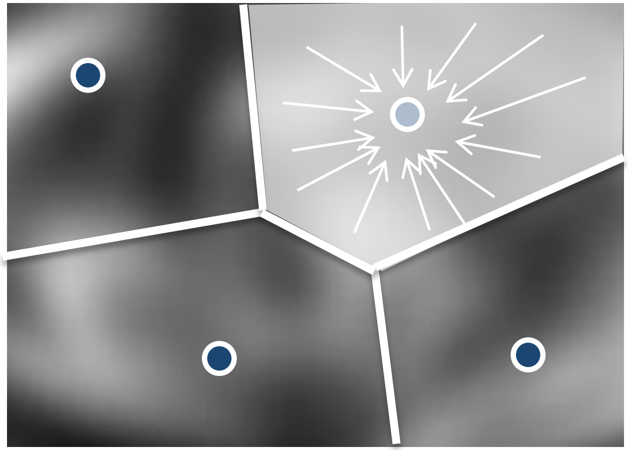
Part. 3 Optimal Transport – semi-discrete

(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



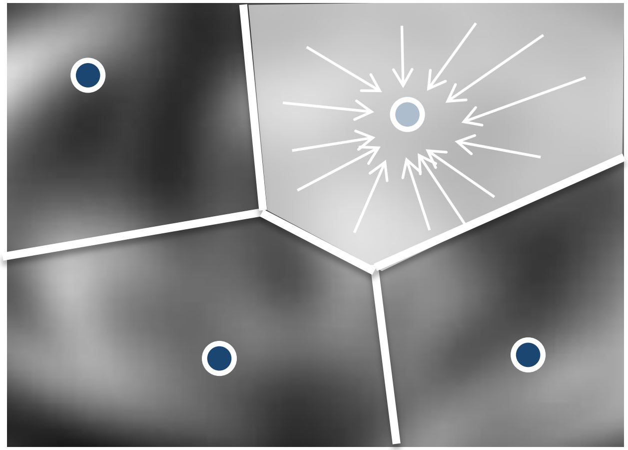
Part. 3 Optimal Transport – semi-discrete

(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 $j' \neq j$

Weight of y_j in the power diagram

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



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

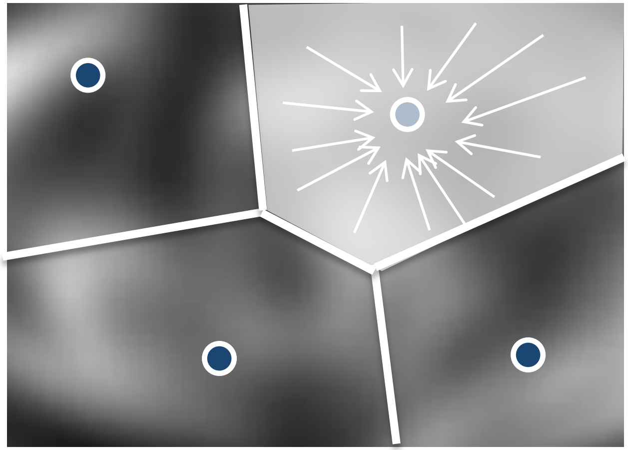
Part. 3 Optimal Transport – semi-discrete

(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 $j' \neq j$

Weight of y_j in the power diagram

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



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

For all weight vector, ψ is c-concave

Part. 3 Power Diagrams

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

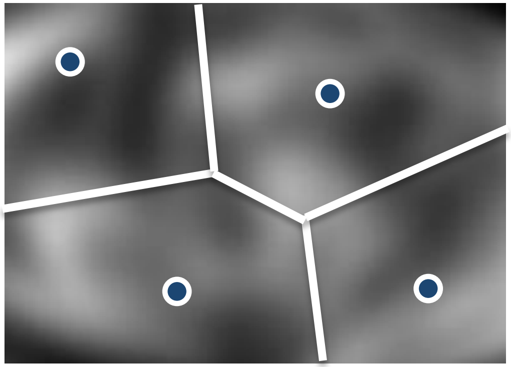
Part. 3 Power Diagrams

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

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

Part. 3 Power Diagrams

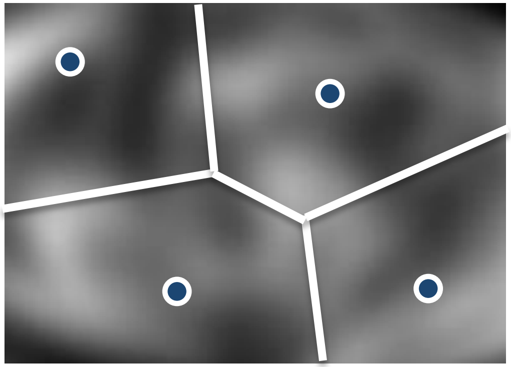
Part. 3 Optimal Transport



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 $v = \sum v_j\, \delta(y_j)$

Part. 3 Optimal Transport



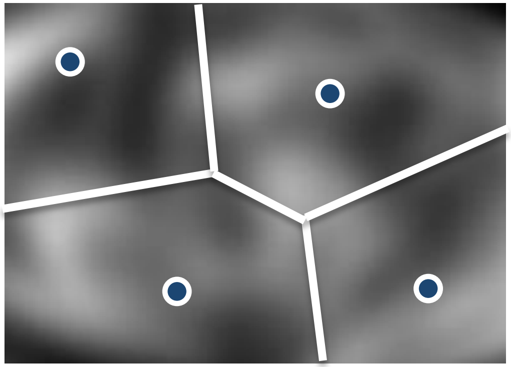
Theorem: *(direct consequence of MK duality*
alternative proof in [Aurenhammer, Hoffmann, Aronov 98]):

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

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

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

Part. 3 Optimal Transport – the AHA paper

Idea of the proof

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

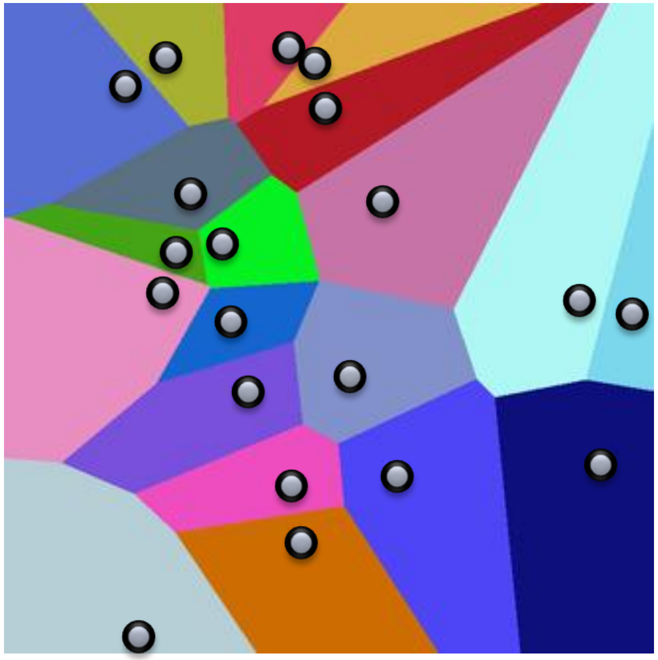
The (unknown) weights $W = [\psi(y_1)\ \psi(y_2)\ \ldots\ \psi(y_m)]$

Part. 3 Optimal Transport – the AHA paper

Idea of the proof

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

T : an arbitrary but fixed assignment.



Part. 3 Optimal Transport – the AHA paper

Idea of the proof

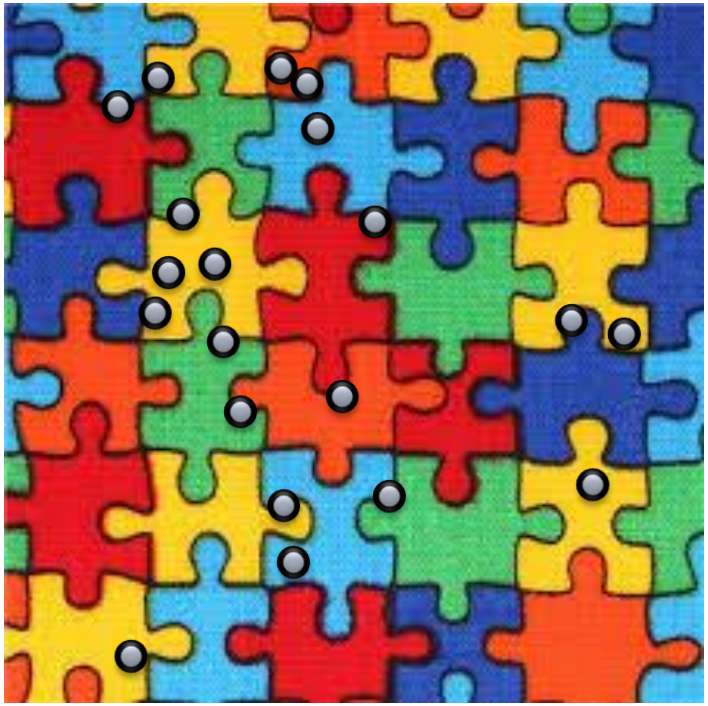
Consider the function $f_T(W) = \int \left(\| x - T(x) \|^2 - \psi\left(T(X)\right) \right) d\mu(x)$

T : an arbitrary but fixed assignment.

Part. 3 Optimal Transport – the AHA paper

Idea of the proof

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

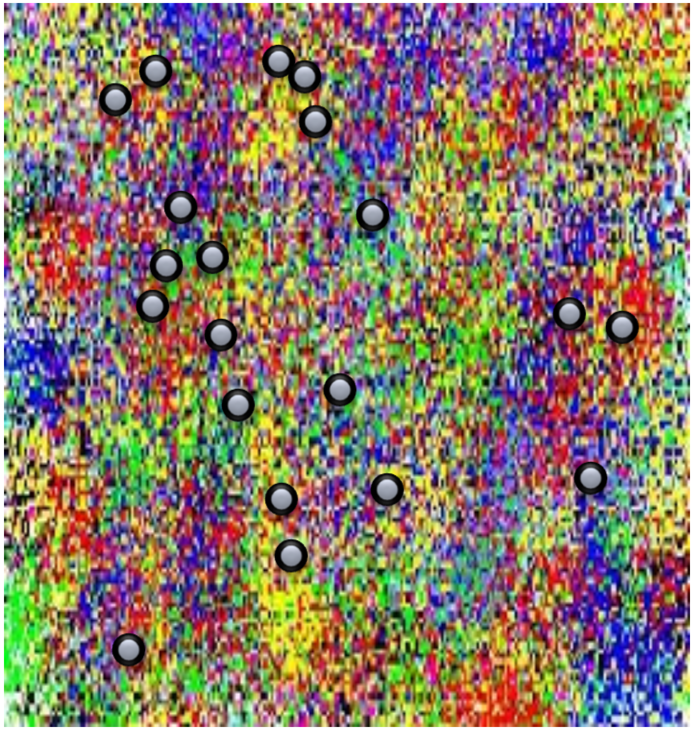


T : an arbitrary but fixed assignment.

Part. 3 Optimal Transport – the AHA paper

Idea of the proof

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

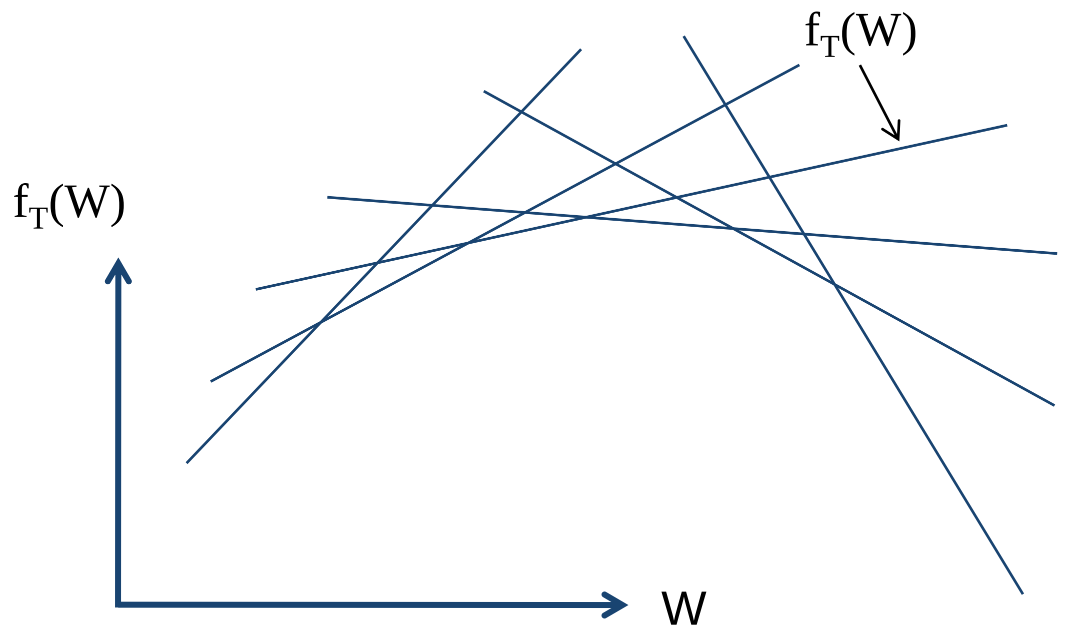


T : an arbitrary but fixed assignment.

Part. 3 Optimal Transport – the AHA paper

Idea of the proof

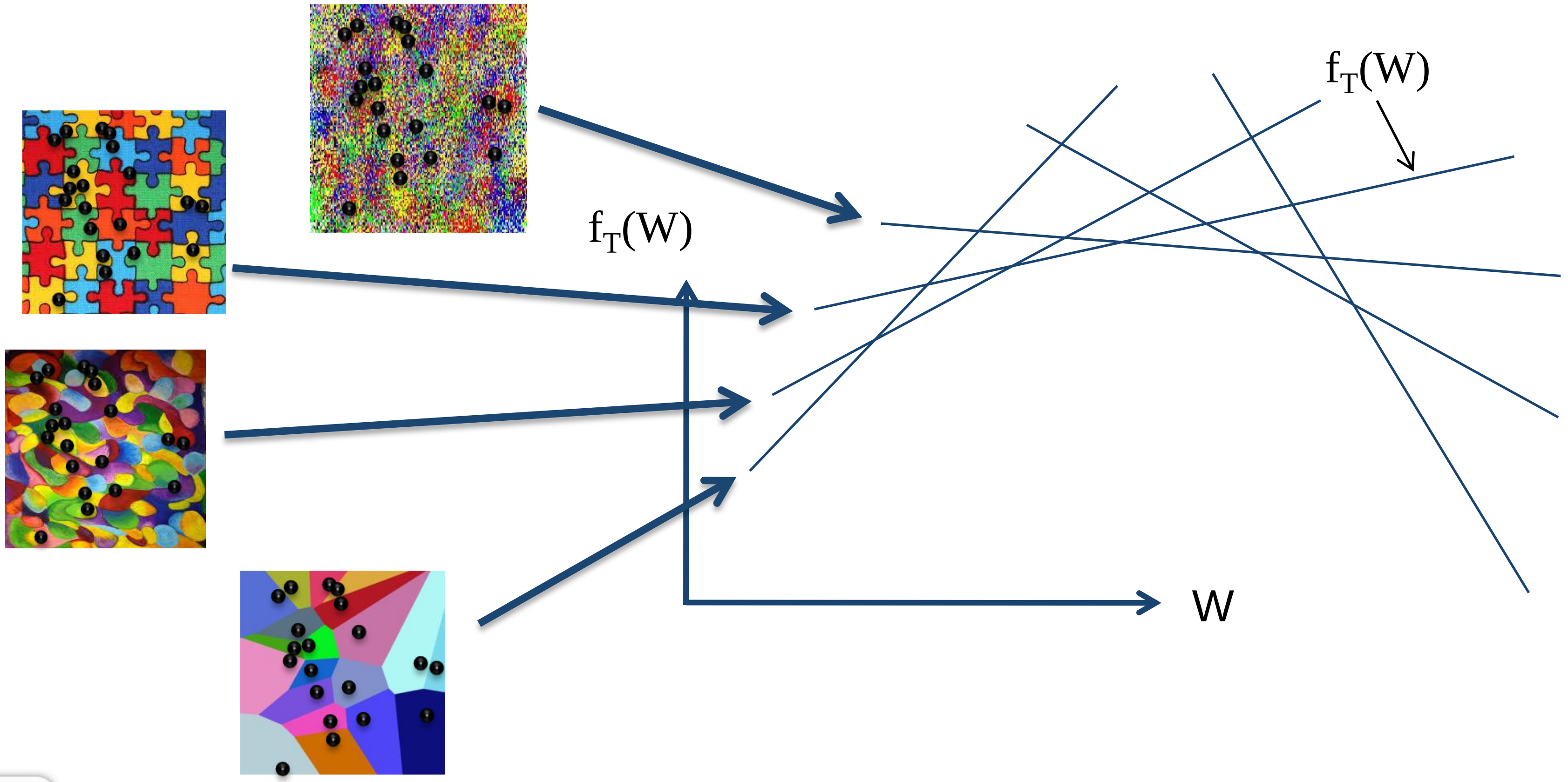
Consider the function $f_T(W) = \int \left(\| x - T(x) \|^2 - \psi\left(T(X)\right) \right) d\mu(x)$



Part. 3 Optimal Transport – the AHA paper

Idea of the proof

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

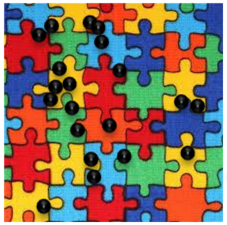


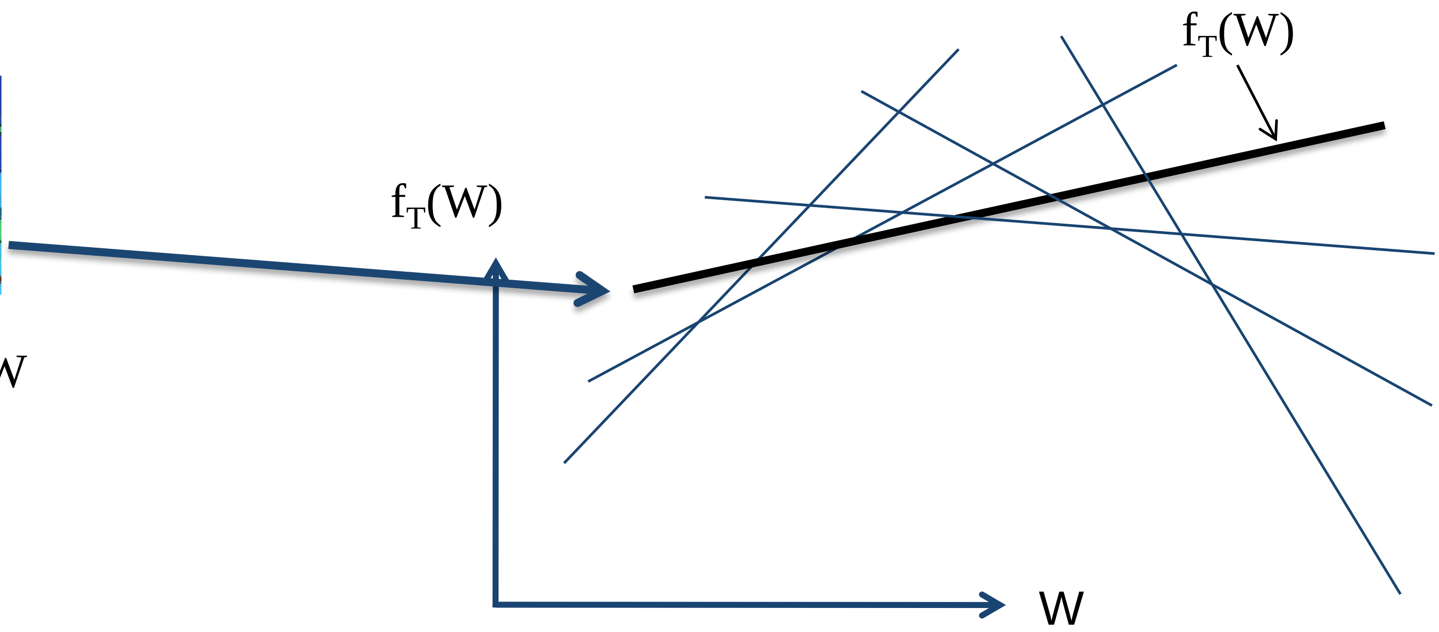
Part. 3 Optimal Transport – the AHA paper

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

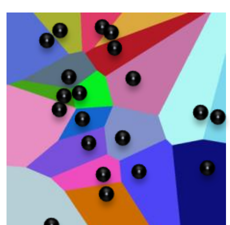
Part. 3 Optimal Transport – the AHA paper

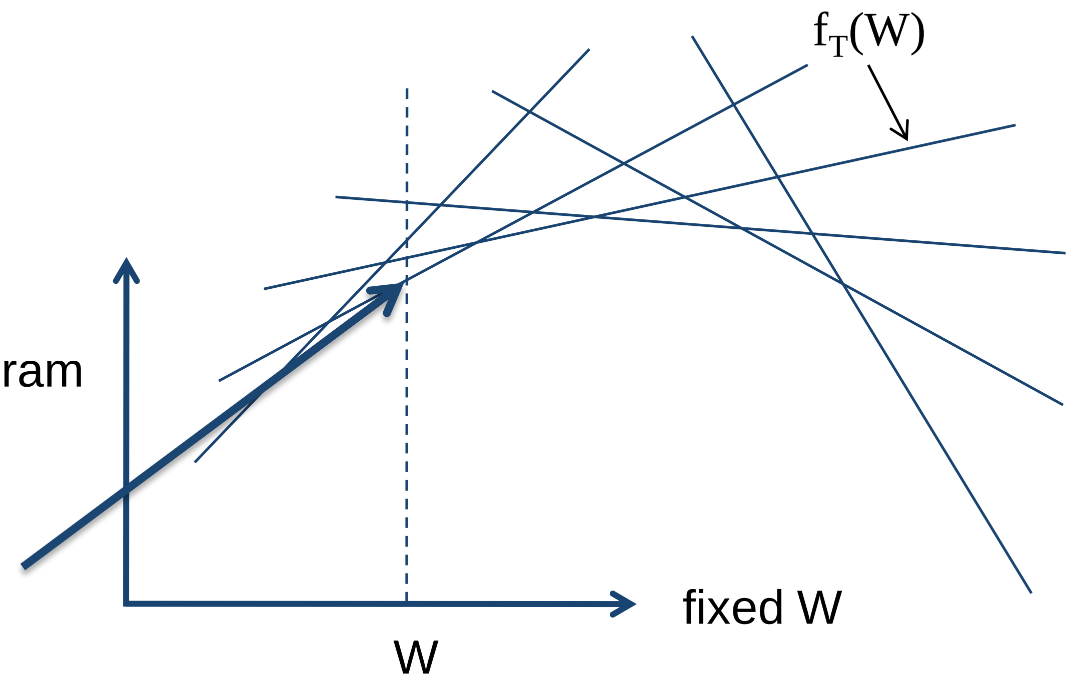
Idea of the proof

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

$f_T(W)$ is linear in W

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





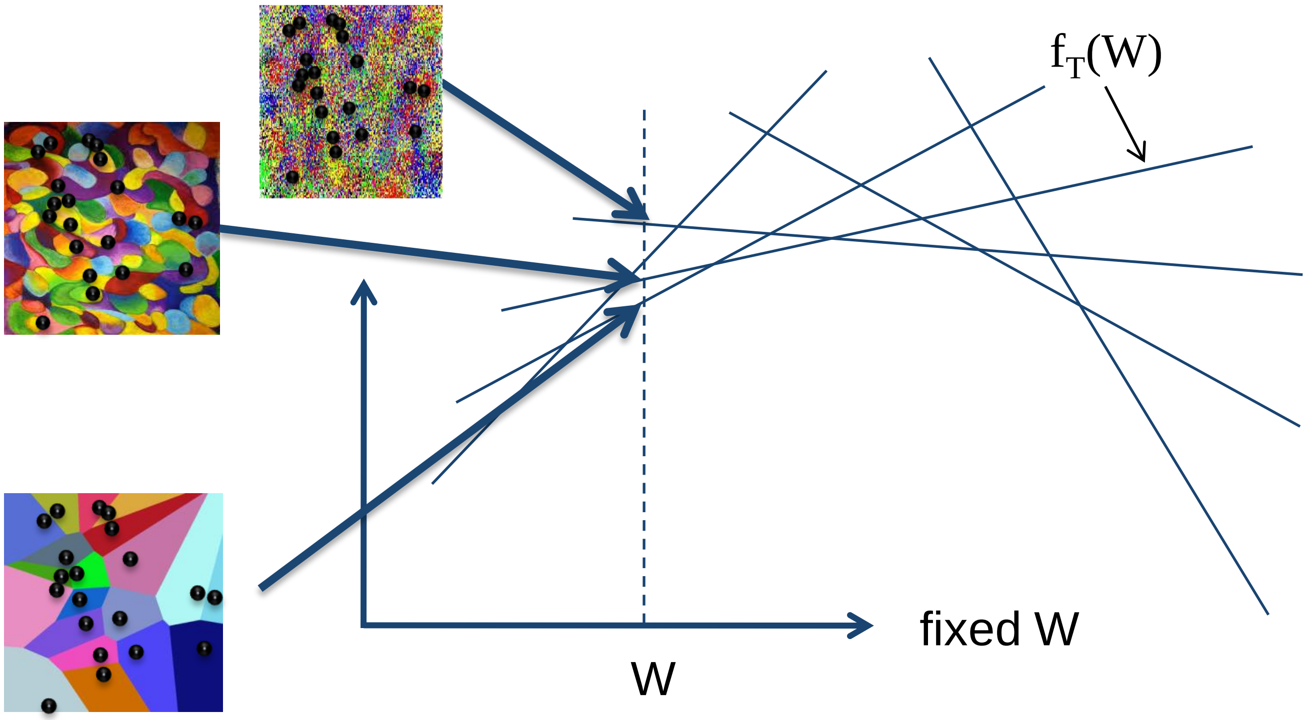
Part. 3 Optimal Transport – the AHA paper

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

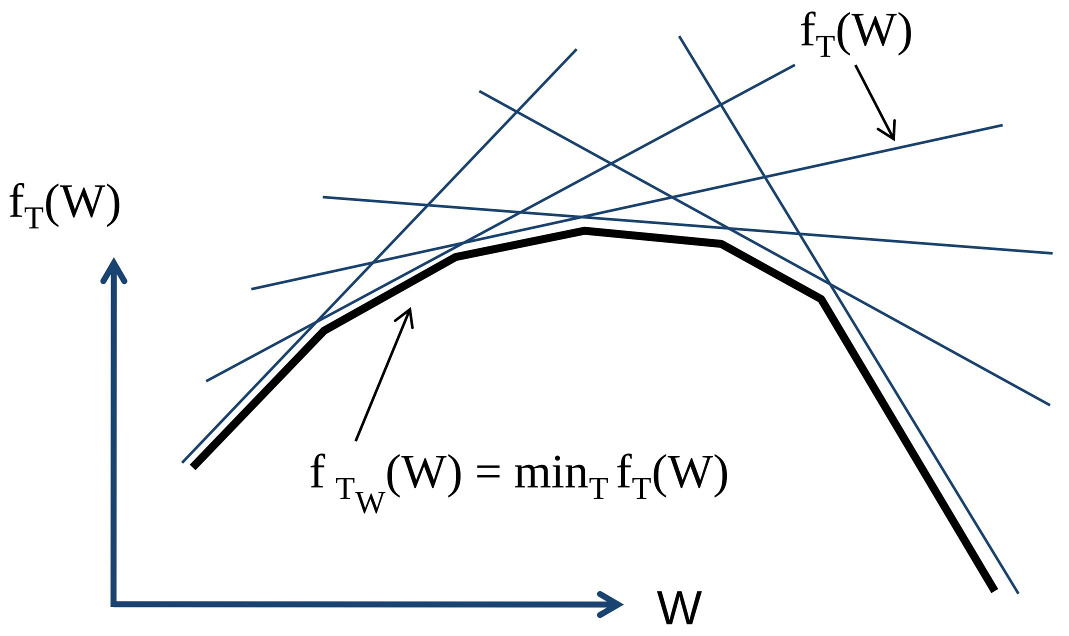
Idea of the proof

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

$f_T(W)$ is linear in W

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

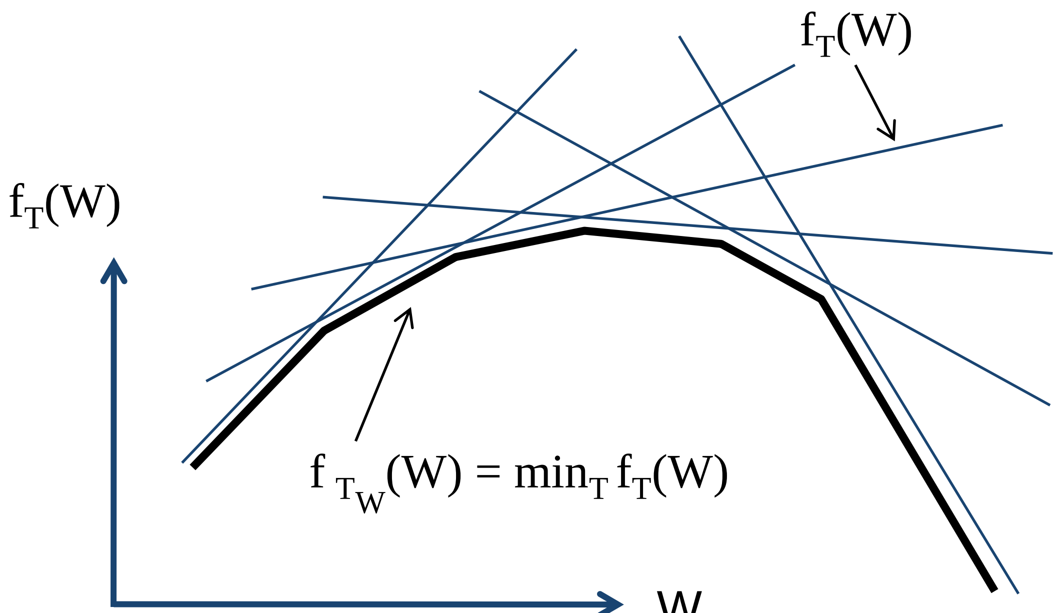
(because its graph is the lower enveloppe of linear functions)



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

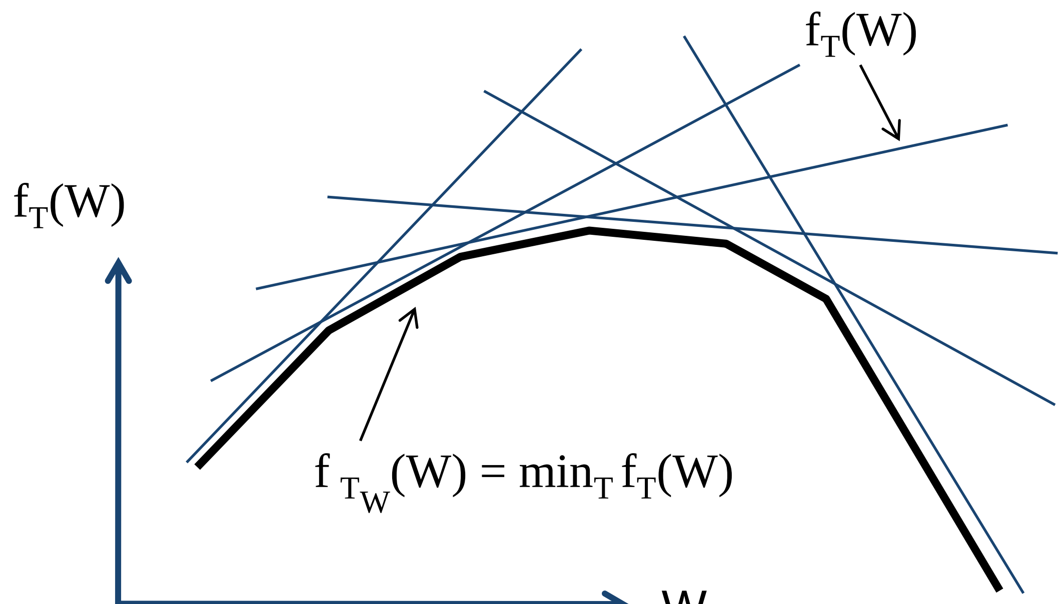
$f: W \rightarrow f_{T_W}(W)$ is concave (because its graph is the lower enveloppe of linear functions)

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

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

$$\partial g / \partial \psi_j = v_j - \int_{Lag^{\psi}(yj)} d\mu(x)$$ and g is concave.

Part. 3 Optimal Transport – the algorithm

Semi-discrete OT Summary:

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

Part. 3 Optimal Transport – the algorithm

Semi-discrete OT Summary:

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

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

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

$$\partial G / \partial \psi_j = v_j - \int_{Lag(yj)} d\mu(x) \quad (= 0 \text{ at the maximum})$$

Part. 3 Optimal Transport – the algorithm

Semi-discrete OT Summary:

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

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

$$\partial G / \partial \psi_j = v_j - \int_{Lag(yj)} d\mu(x) \quad (= 0 \text{ at the maximum})$$

Desired mass at y_j

Mass transported to y_j

Part. 3 Optimal Transport – the Hessian

$$\partial G / \partial \psi_j = V_j - \int_{Lag(yj)} d\mu(x)$$

Part. 3 Optimal Transport – the Hessian

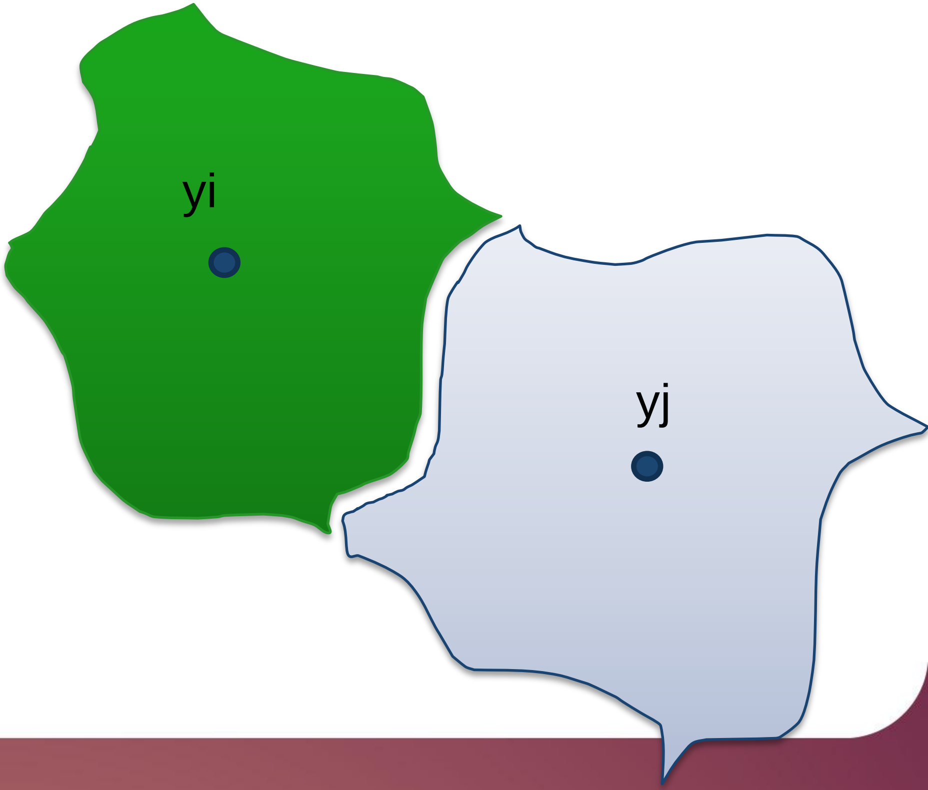
$$\partial G / \partial \Psi_j = v_j - \int_{Lag(yj)} d\mu(x)$$

$$\partial^2 G / \partial \Psi_i \Psi_j = - \partial / \partial \Psi_j \int_{Lag(yj)} d\mu(x)$$

Part. 3 Optimal Transport – the Hessian

$$\partial G / \partial \psi_j = v_j - \int_{Lag(yj)} d\mu(x)$$

$$\partial^2 G / \partial \psi_i \psi_j = - \partial / \partial \psi_j \int_{Lag(yj)} d\mu(x)$$

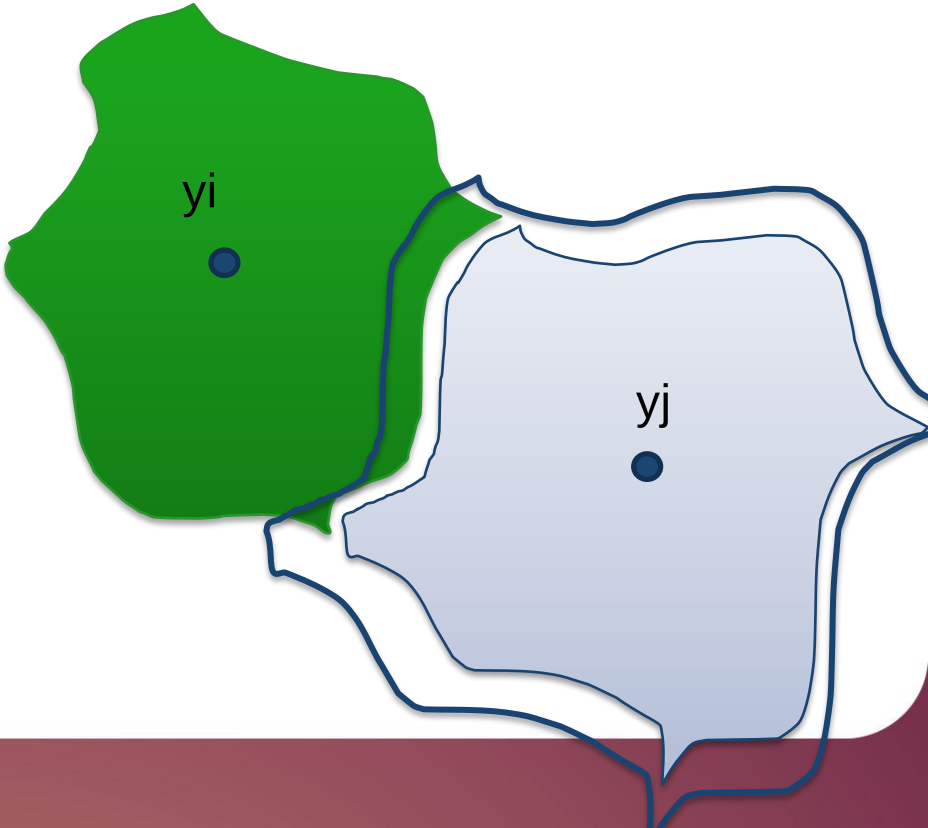


Part. 3 Optimal Transport – the Hessian

$$\partial G / \partial \psi_j = v_j - \int_{Lag(yj)} d\mu(x)$$

$$\partial^2 G / \partial \psi_i \psi_j = - \partial / \partial \psi_j \int_{Lag(yj)} d\mu(x)$$

$$\psi_j \leftarrow \psi_j + \delta\psi_j$$



Part. 3 Optimal Transport – the Hessian

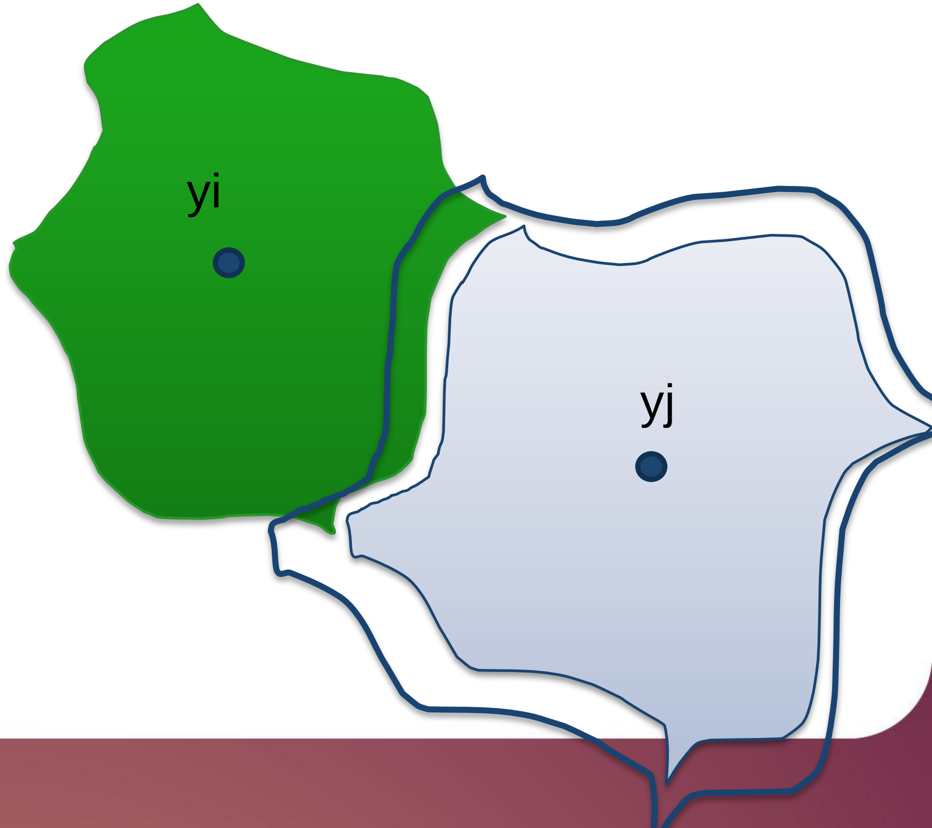
$$\partial G / \partial \psi_j = \nu_j - \int_{Lag(yj)} d\mu(x)$$

$$\partial^2 G / \partial \psi_i \psi_j = - \partial / \partial \psi_j \int_{Lag(yj)} d\mu(x)$$

$$\psi_j \leftarrow \psi_j + \delta\psi_j$$

Reynold's thm:

$$\partial / \partial \psi_j \int_{Lag(yj)} d\mu(x) = \int_{\partial Lag(yj)} v.n \, d\mu(x)$$

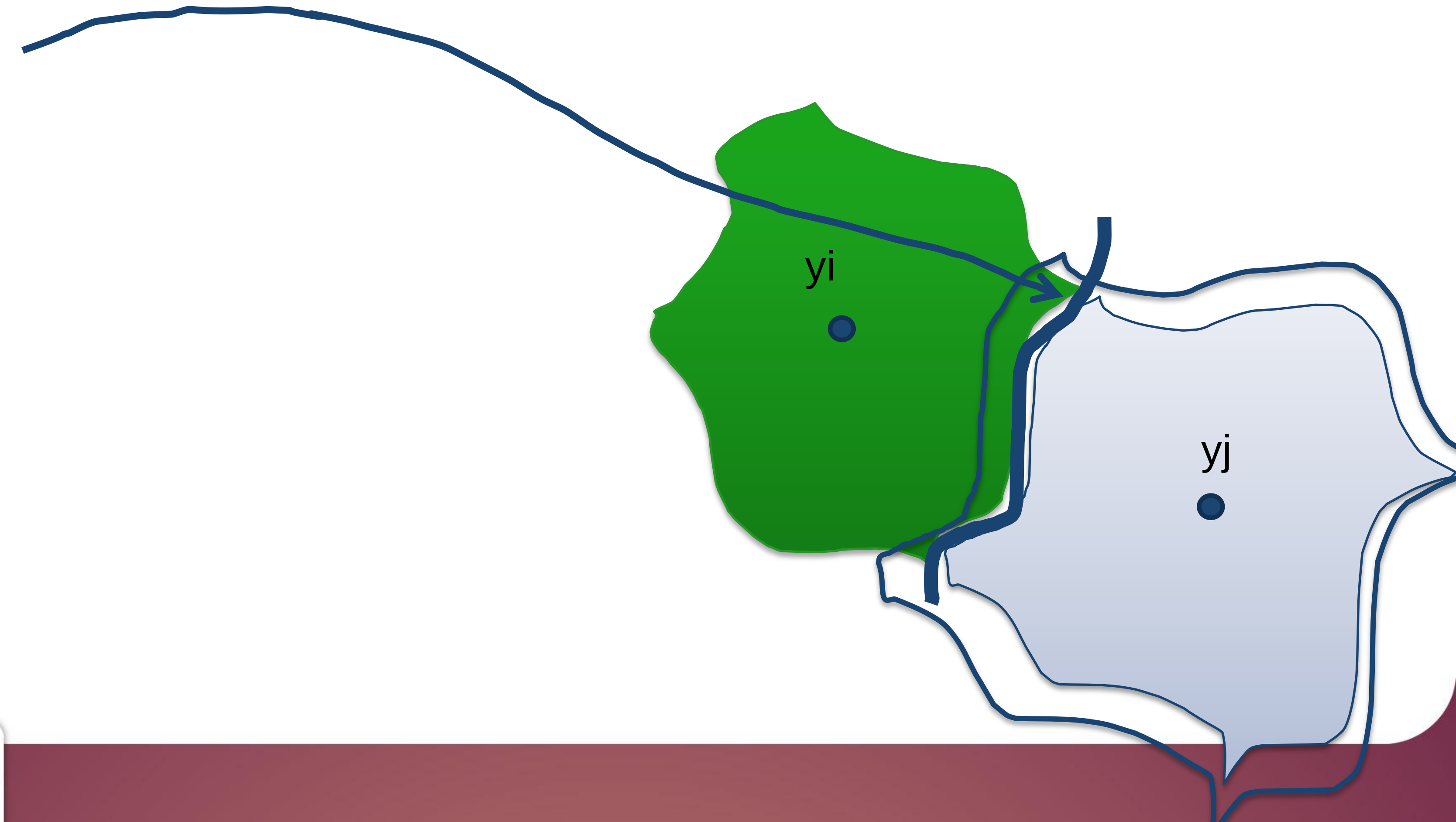


Part. 3 Optimal Transport – the Hessian

Reynold's thm:

$$\partial / \partial \psi_j \int_{Lag(yj)} d\mu(x) = \int_{\partial Lag(yj)} v.n \, d\mu(x)$$

$f_{ij}(x) = 0$



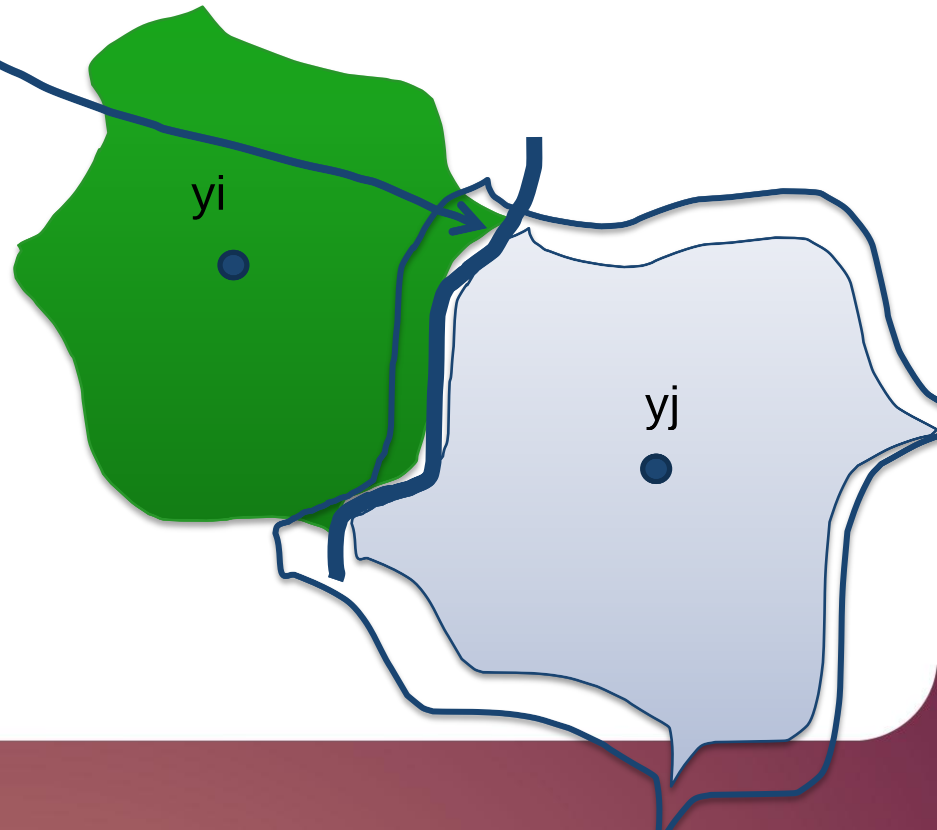
Part. 3 Optimal Transport – the Hessian

Reynold's thm:

$$\partial / \partial \psi_j \int_{Lag(yj)} d\mu(x) = \int_{\partial Lag(yj)} v.n \, d\mu(x)$$

$f_{ij}(x) = 0$

$c(x,y_i) - c(x,y_j) + \psi_j - \psi_i = 0$



Part. 3 the Hessian

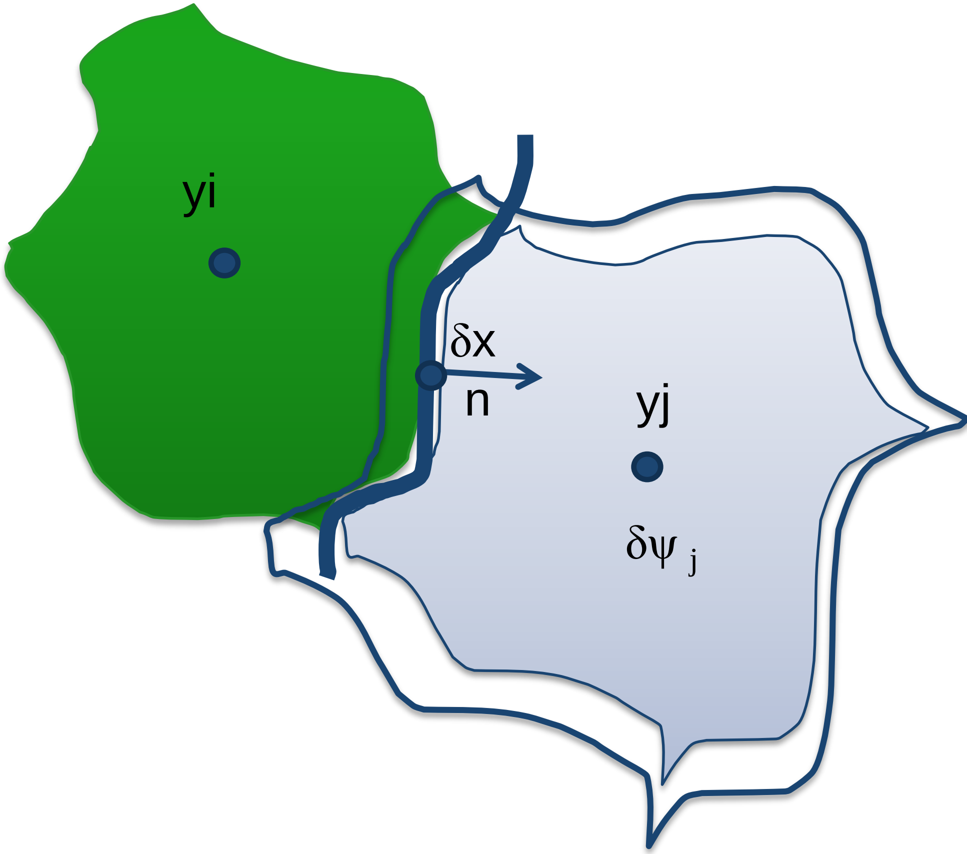
Reynold's thm:

$$\partial / \partial \psi_j \int_{Lag(yj)} d\mu(x) = \int_{\partial Lag(yj)} v.n \, d\mu(x)$$

$f_{ij}(x) = 0$

$c(x,y_i) - c(x,y_j) + \psi_j - \psi_i = 0$

$df_{ij} = \text{grad}_x(c(x,y_i) - c(x,y_j))dx + d\psi_j$



Part. 3 the Hessian

Reynold's thm:

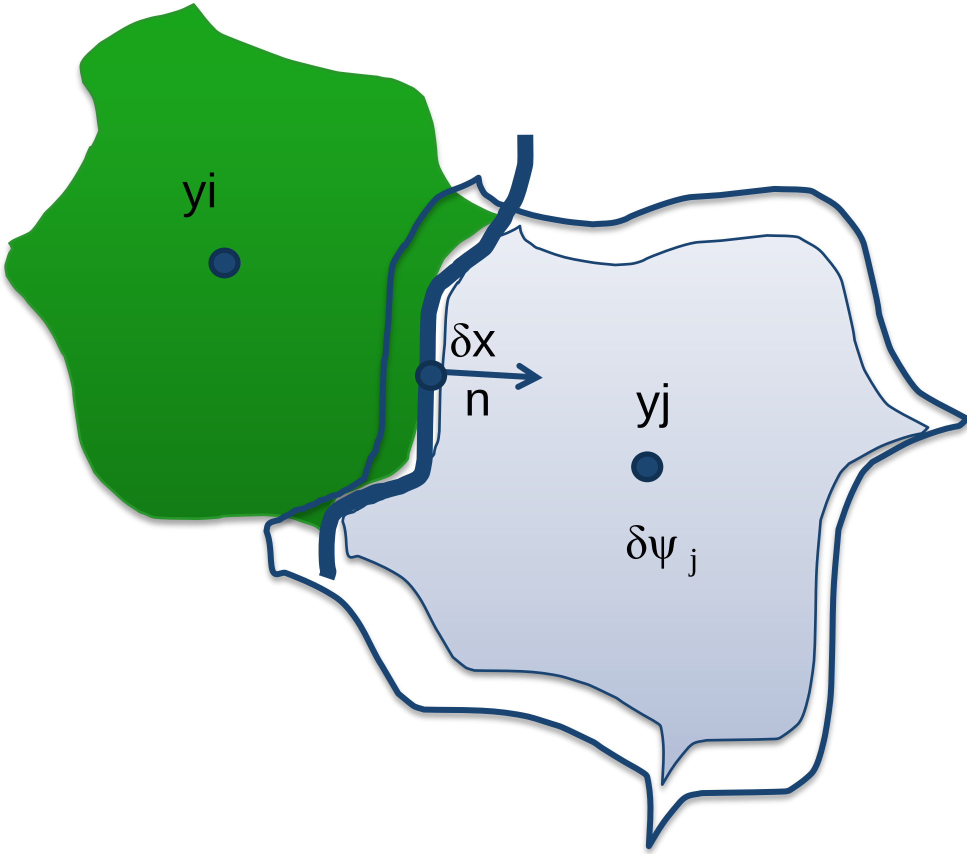
$$\partial / \partial \psi_j \int_{Lag(yj)} d\mu(x) = \int_{\partial Lag(yj)} v.n \, d\mu(x)$$

$f_{ij}(x) = 0$

$c(x,y_i) - c(x,y_j) + \psi_j - \psi_i = 0$

$df_{ij} = \text{grad}_x(c(x,y_i) - c(x,y_j))dx + d\psi_j$

$\delta x = \delta h \, n = \delta h \, \text{grad}_x f_{ij}(x) / \| \text{grad}_x f_{ij}(x) \|$



Part. 3 the Hessian

Reynold's thm:

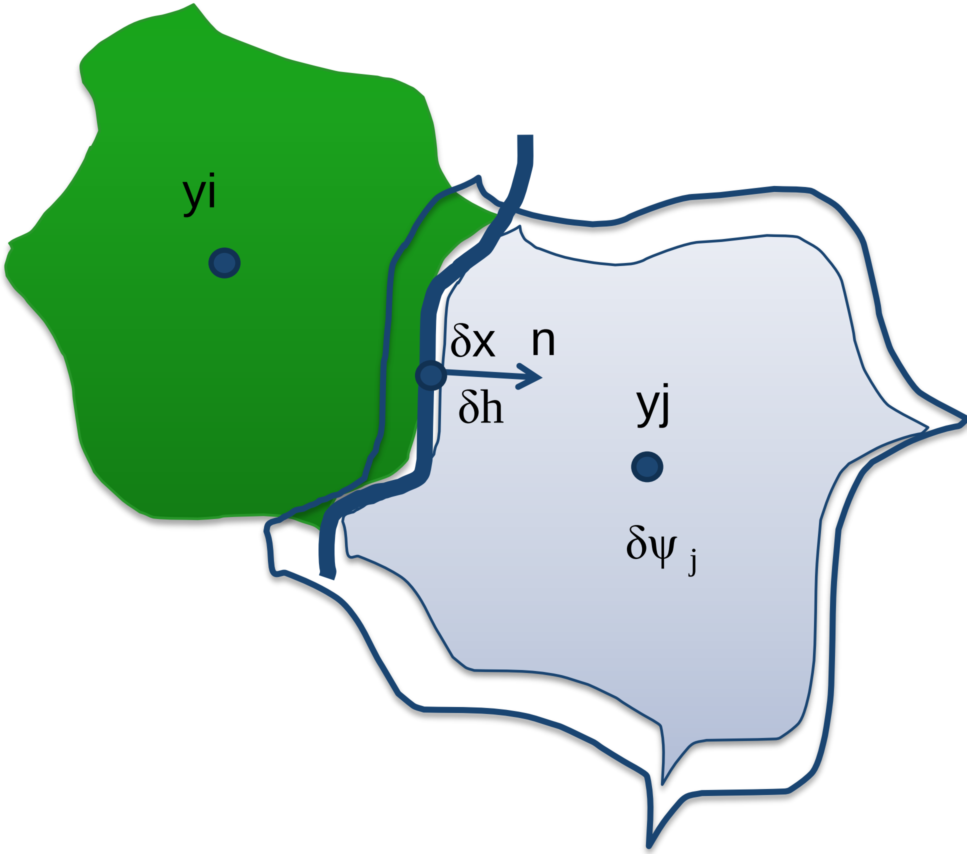
$$\partial / \partial \psi_j \int_{Lag(yj)} d\mu(x) = \int_{\partial Lag(yj)} v.n \, d\mu(x)$$

$f_{ij}(x) = 0$

$c(x,y_i) - c(x,y_j) + \psi_j - \psi_i = 0$

$df_{ij} = grad_x(c(x,y_i) - c(x,y_j))dx + d\psi_j$

$\delta x = \delta h \, n = \delta h \, grad_x \, f_{ij}(x) / \| grad_x \, f_{ij}(x)\|$



Part. 3 the Hessian

Reynold's thm:

$$\partial / \partial \psi_j \int_{Lag(yj)} d\mu(x) = \int_{\partial Lag(yj)} v.n \, d\mu(x)$$

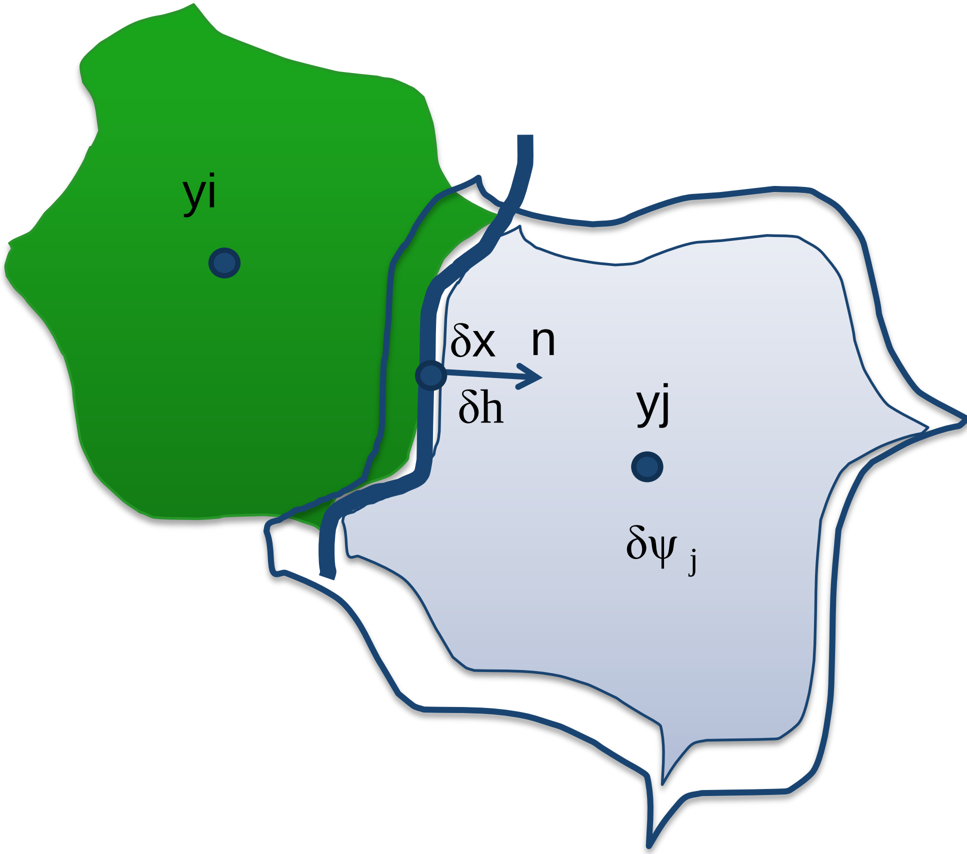
$f_{ij}(x) = 0$

$c(x,y_i) - c(x,y_j) + \psi_j - \psi_i = 0$

$df_{ij} = grad_x(c(x,y_i) - c(x,y_j))dx + d\psi_j$

$\delta x = \delta h \, n = \delta h \, grad_x f_{ij}(x) / \| grad_x f_{ij}(x) \|$

$$\partial h / \partial \psi_j = -1/ \| grad_x c(x,y_i) - grad_x c(x,y_j) \|$$



Part. 3 the Hessian

Reynold's thm:

$$\partial / \partial \psi_j \int_{Lag(yj)} d\mu(x) = \int_{\partial Lag(yj)} v.n \, d\mu(x)$$

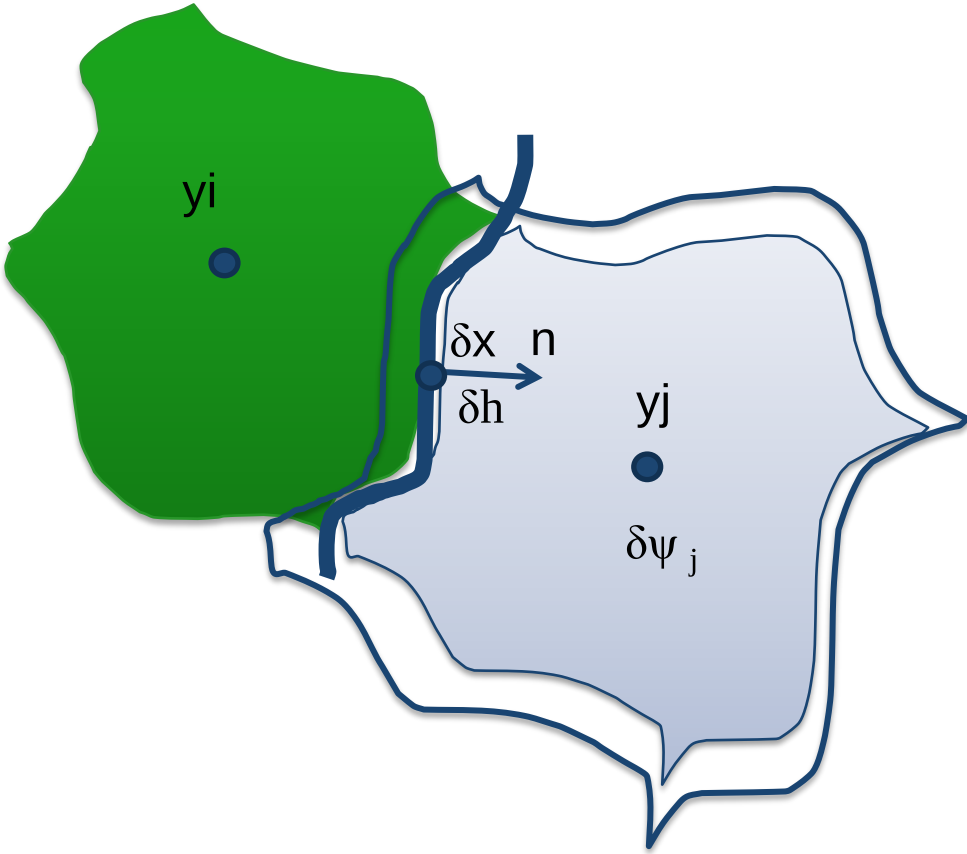
$f_{ij}(x) = 0$

$c(x,y_i) - c(x,y_j) + \psi_j - \psi_i = 0$

$df_{ij} = \mathrm{grad}_x(c(x,y_i) - c(x,y_j))dx + d\psi_j$

$\delta x = \delta h \, n = \delta h \, \mathrm{grad}_x f_{ij}(x) / \| \mathrm{grad}_x f_{ij}(x) \|$

$$\partial h / \partial \psi_j = -1 / \| \mathrm{grad}_x c(x,y_i) - \mathrm{grad}_x c(x,y_j) \|$$



$$\partial / \partial \psi_j \int_{Lag(yj)} d\mu(x) = \int_{Lag(yi) \cap Lag(yj)} -1/\| \mathrm{grad}_x c(x,y_i) - \mathrm{grad}_x c(x,y_j) \| d\mu(x)$$

Part. 3 the Hessian

$$\partial^2 / \partial\psi_i \partial\psi_j F = \int_{Lag(yi) \cap Lag(yj)} -1/\| \text{grad}_x c(x,y_i) - \text{grad}_x c(x,y_j) \| d\mu(x)$$

$$\partial^2 / \partial\psi_i^2 F = -\sum \partial^2 / \partial\psi_i \partial\psi_j$$

Part. 3 the Hessian

$$\partial^2 / \partial \psi_i \partial \psi_j F = \int_{Lag(yi) \cap Lag(yj)} -1/\| \text{grad}_x\, c(x,y_i) - \text{grad}_x\, c(x,y_j) \| d\mu(x)$$

$$\partial^2 / \partial \psi_i^2 F = - \sum \partial^2 / \partial \psi_i \partial \psi_j$$

$c(x,y) = \| x - y \|^2$

$$\partial^2 / \partial \psi_i \partial \psi_j F = \int_{Lag(yi) \cap Lag(yj)} 1 / \| xj - xi \| \, d\mu(x)$$

Part. 3 the Hessian

$$\partial^2 / \partial \psi_i \partial \psi_j F = \int_{Lag(yi) \cap Lag(yj)} -1/\| grad_x \, c(x,y_i) - grad_x \, c(x,y_j) \| d\mu(x)$$

$$\partial^2 / \partial \psi_i{}^2 F = - \sum \partial^2 / \partial \psi_i \partial \psi_j$$

$c(x,y) = \| x - y \|^2$

$$\partial^2 / \partial \psi_i \partial \psi_j F = \int_{Lag(yi) \cap Lag(yj)} 1 / \| xj - xi \| d\mu(x)$$

IP_1 FEM Laplacian (not a big surprise)

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:

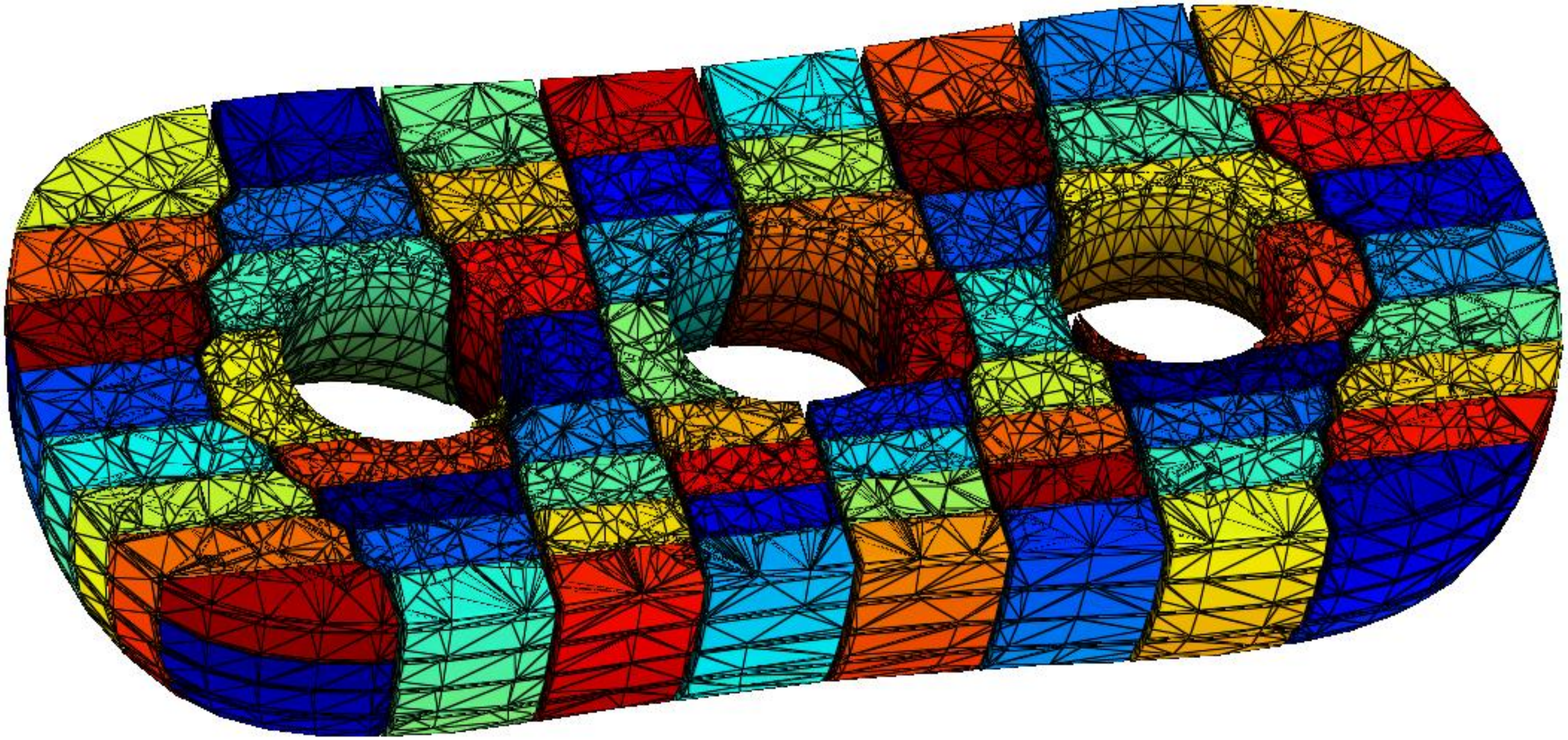
For each iterate $(s_1 \dots s_N)^{(k)}$:

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

Compute g, grad g, Hg

Damped Newton algorithm, [Kitagawa, Mérigot, Thibert]

Part. 3 Optimal Transport – the algorithm



Part. 3 Optimal Transport – the algorithm

The [AHA] paper summary:

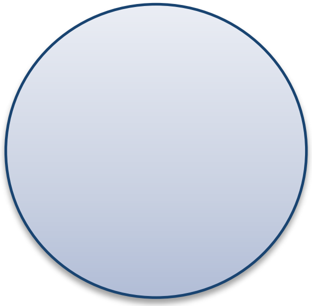
- The optimal weights minimize a convex function
- The gradient and Hessian of this convex function is easy to compute

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

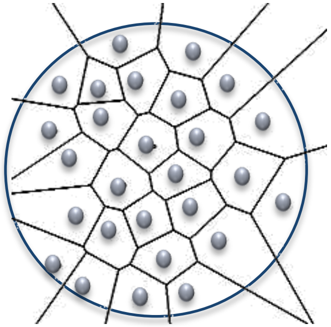
The algorithm:

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:

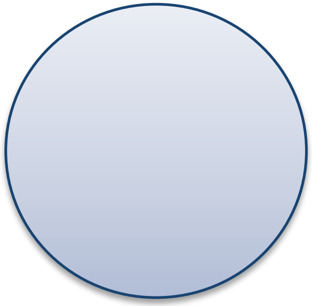
- The optimal weights minimize a convex function
- The gradient and Hessian of this convex function is easy to compute

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

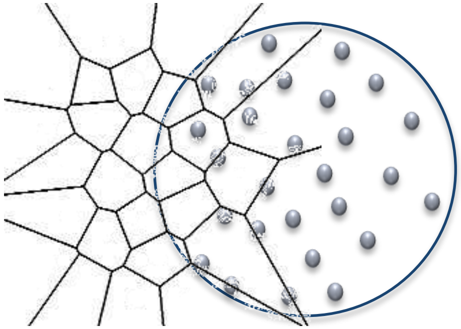
The algorithm:

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>

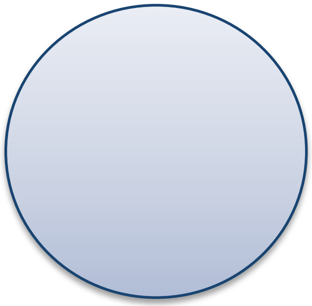
- The optimal weights minimize a convex function
- The gradient and Hessian of this convex function is easy to compute

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

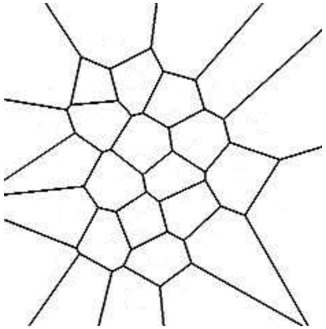
<u>The algorithm:</u>

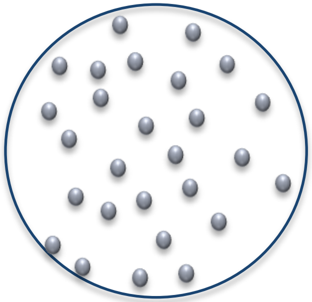
<u>Summary:</u>

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:

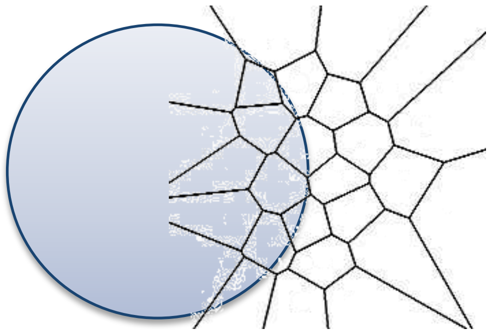
- The optimal weights minimize a convex function
- The gradient and Hessian of this convex function is easy to compute

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

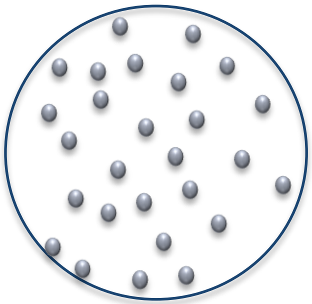
The algorithm:

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:

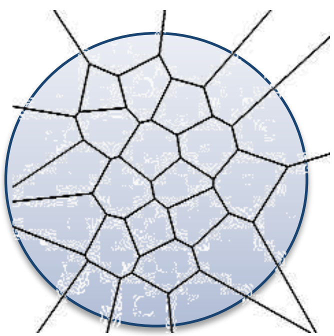
- The optimal weights minimize a convex function
- The gradient and Hessian of this convex function is easy to compute

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

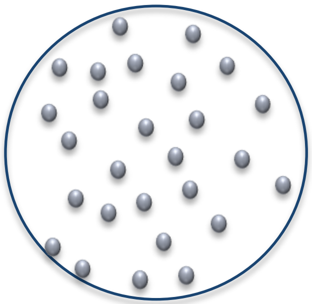
The algorithm:

Summary:

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

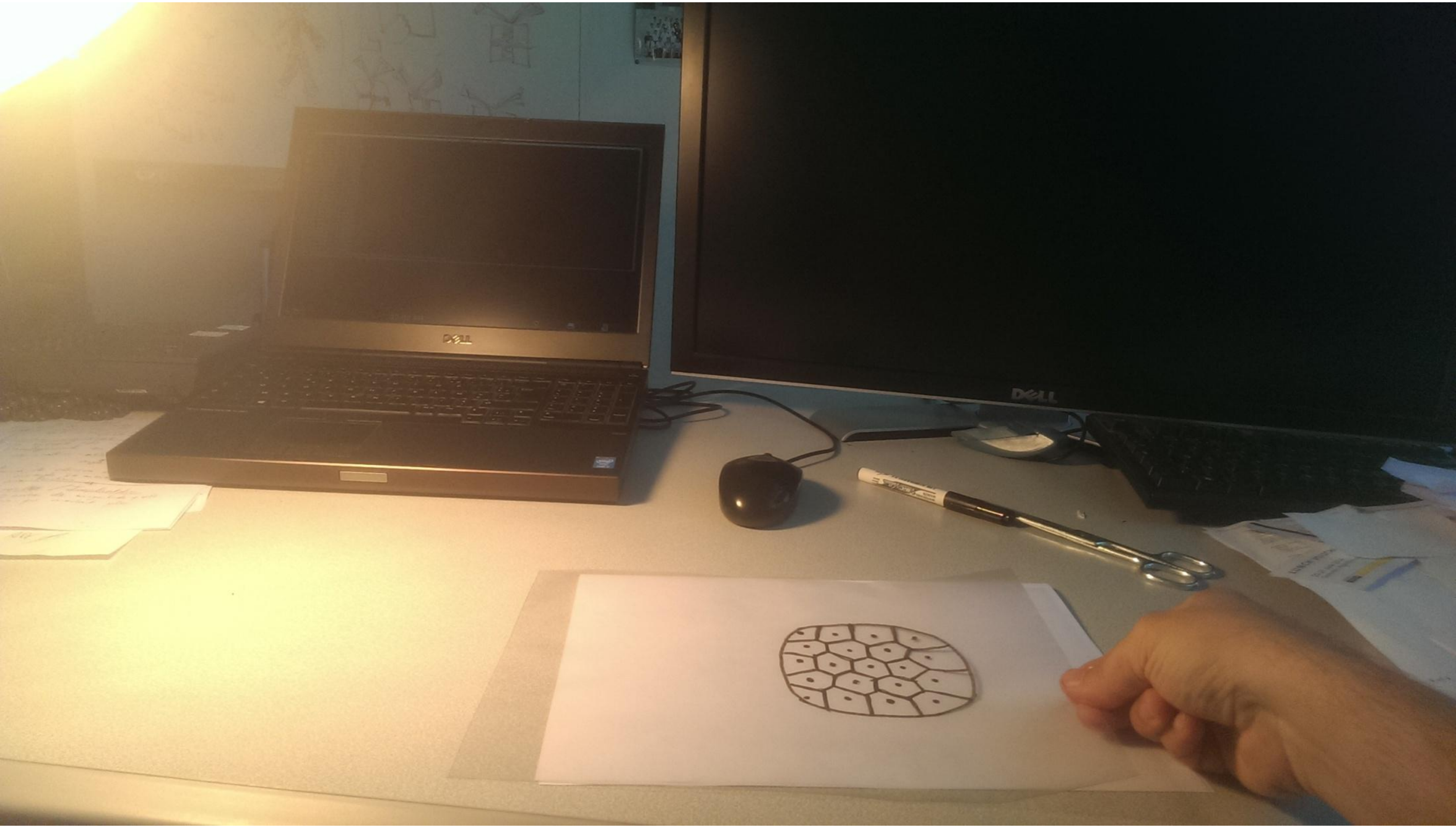


μ

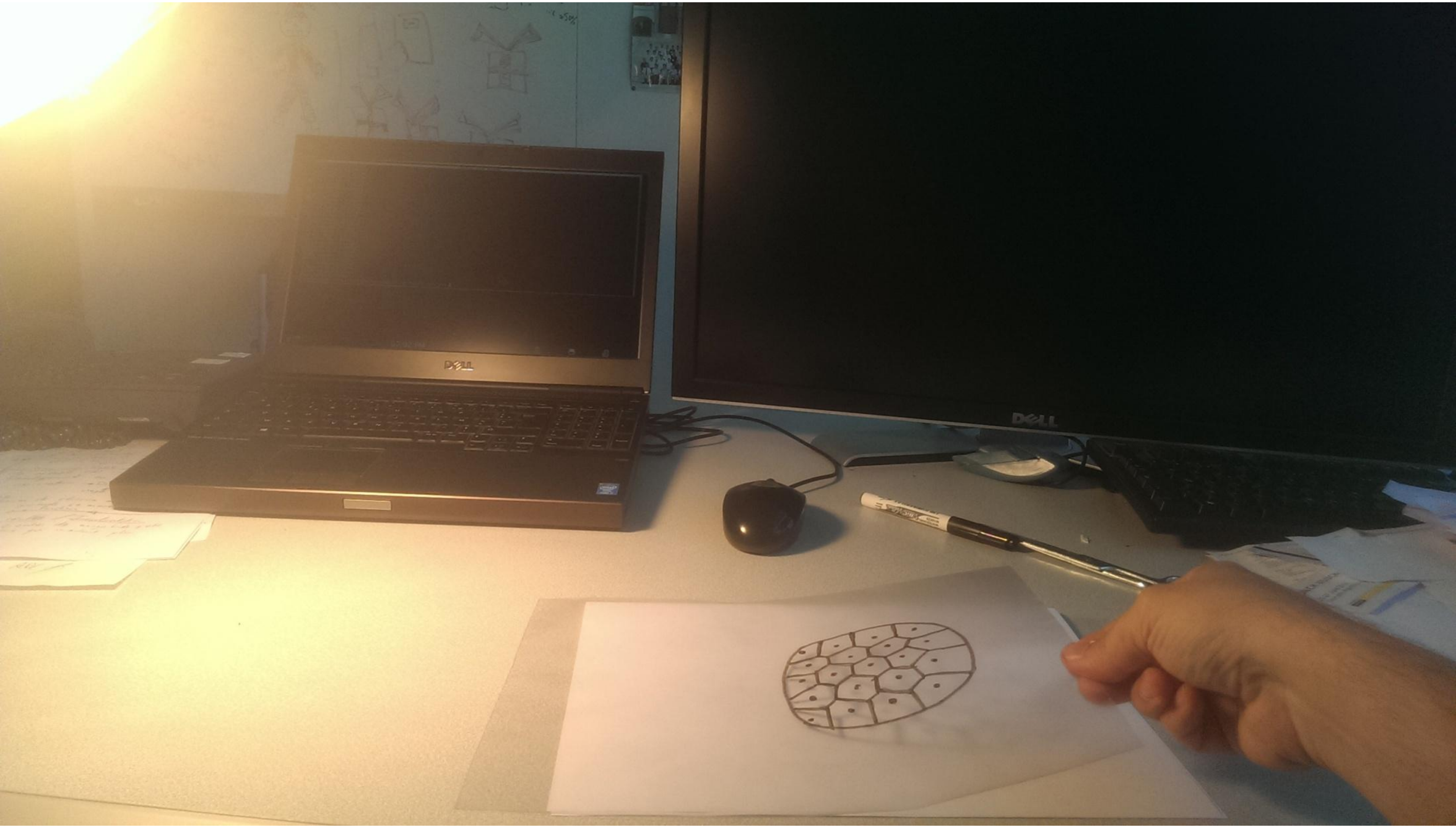


ν

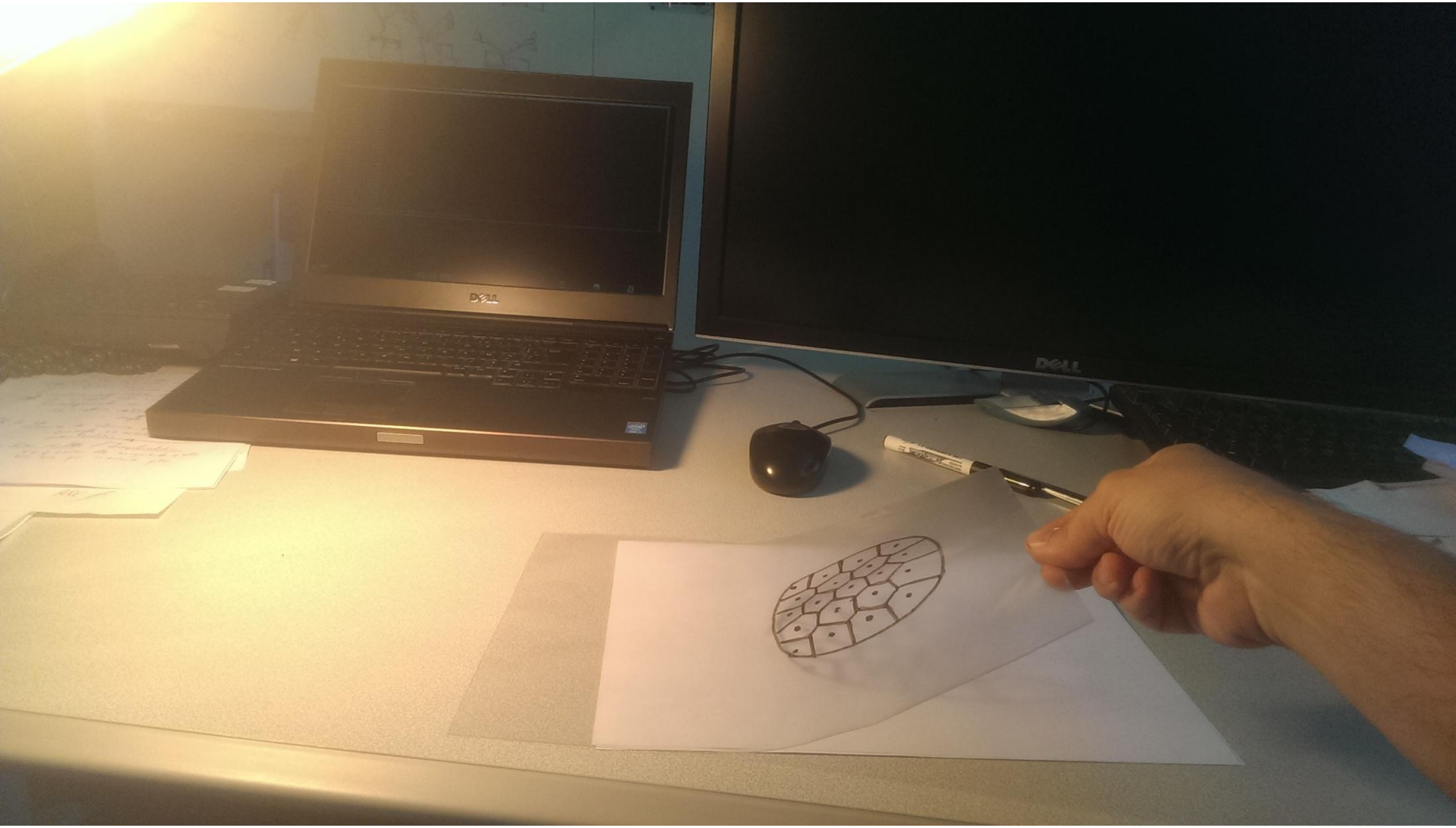
Part. 3 Power Diagrams & Transport



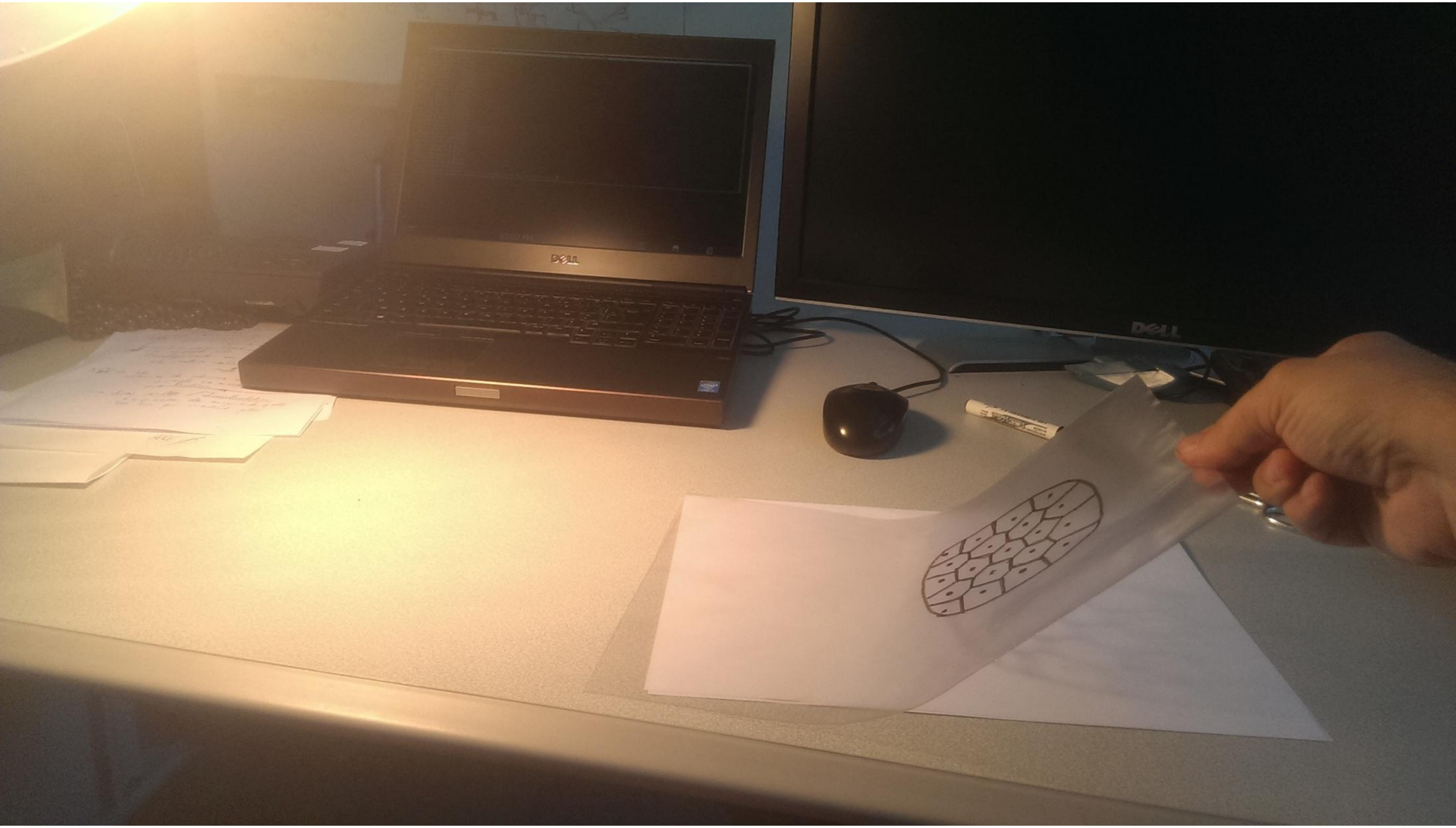
Part. 3 Power Diagrams & Transport



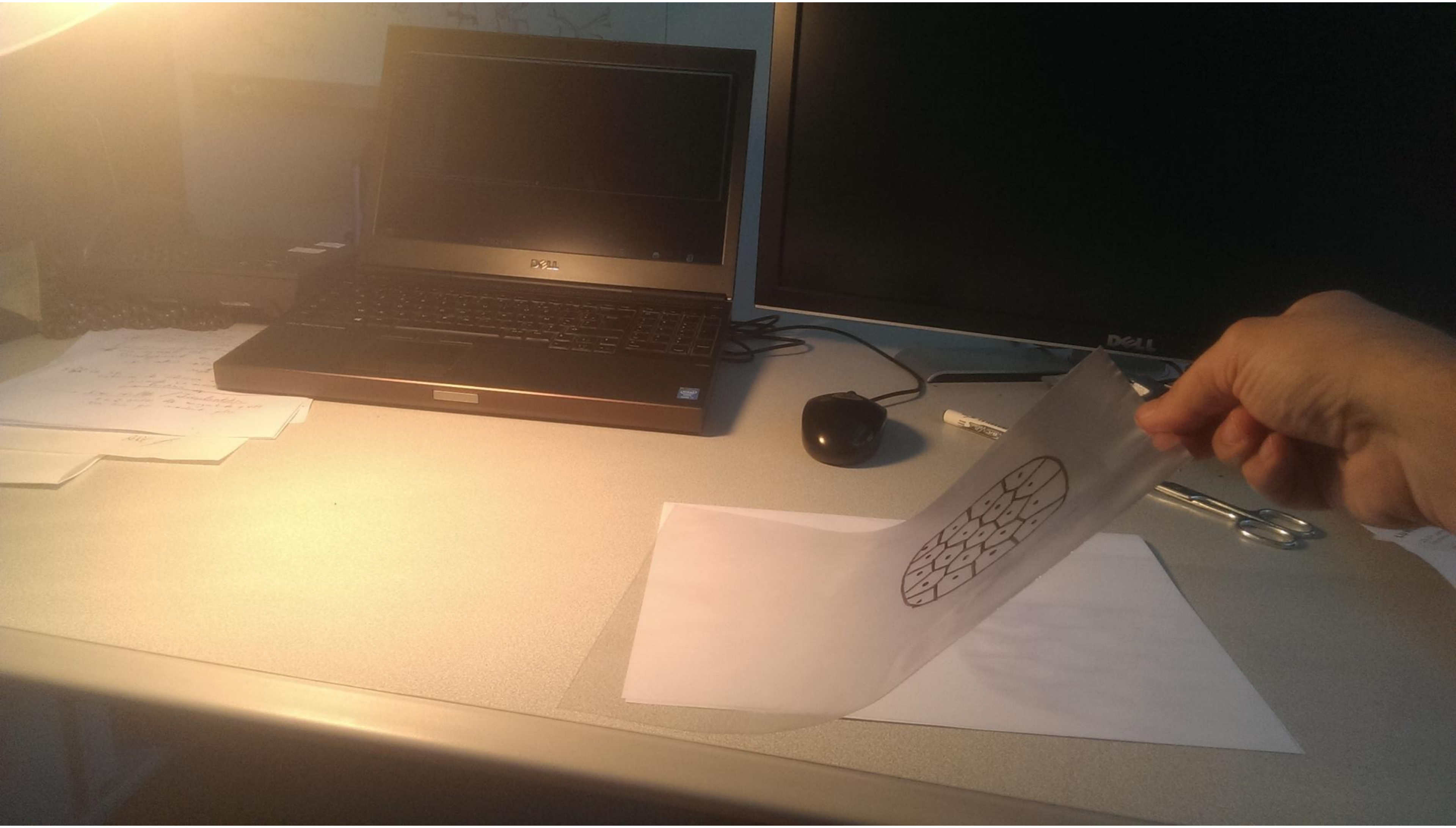
Part. 3 Power Diagrams & Transport



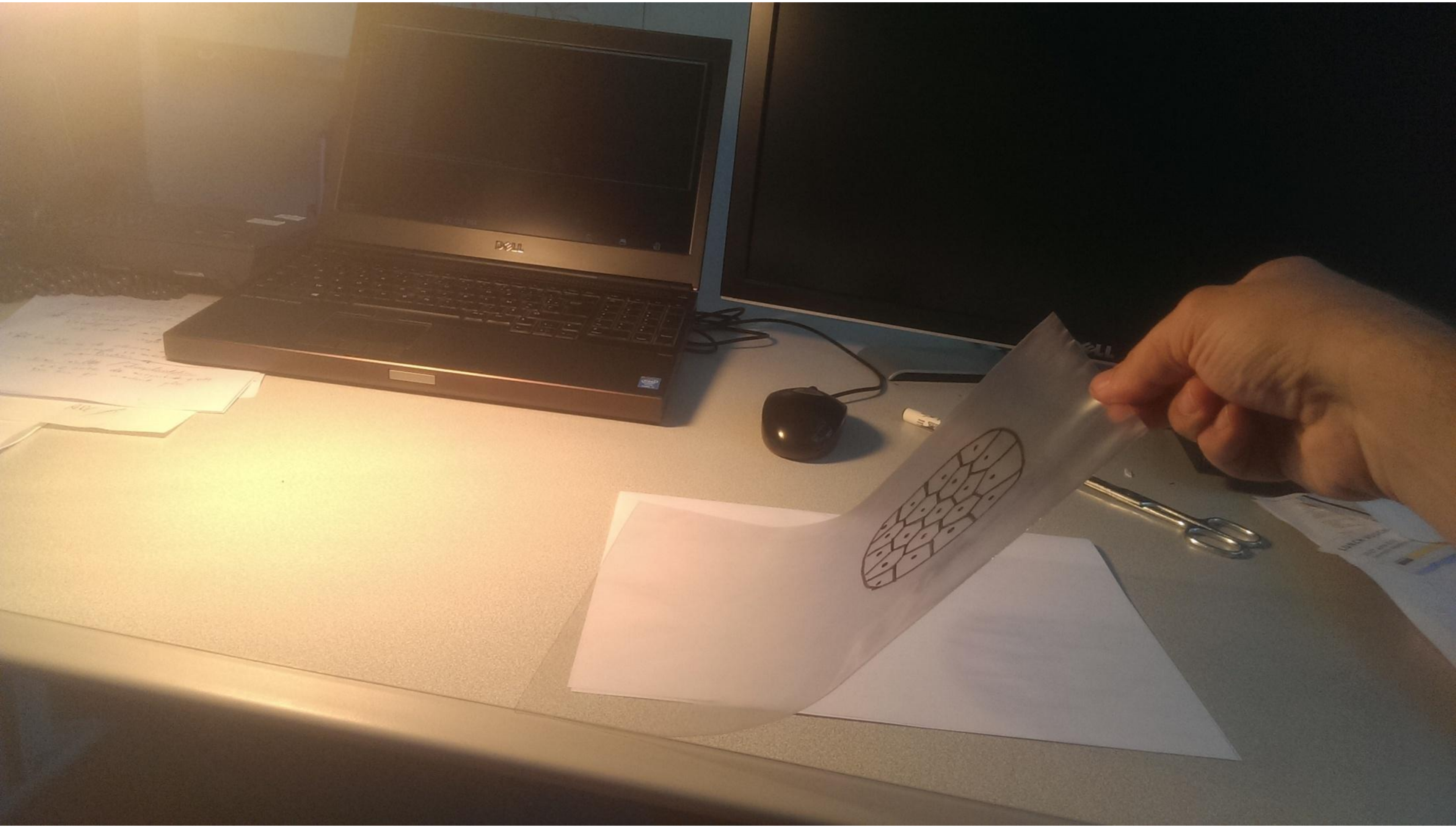
Part. 3 Power Diagrams & Transport



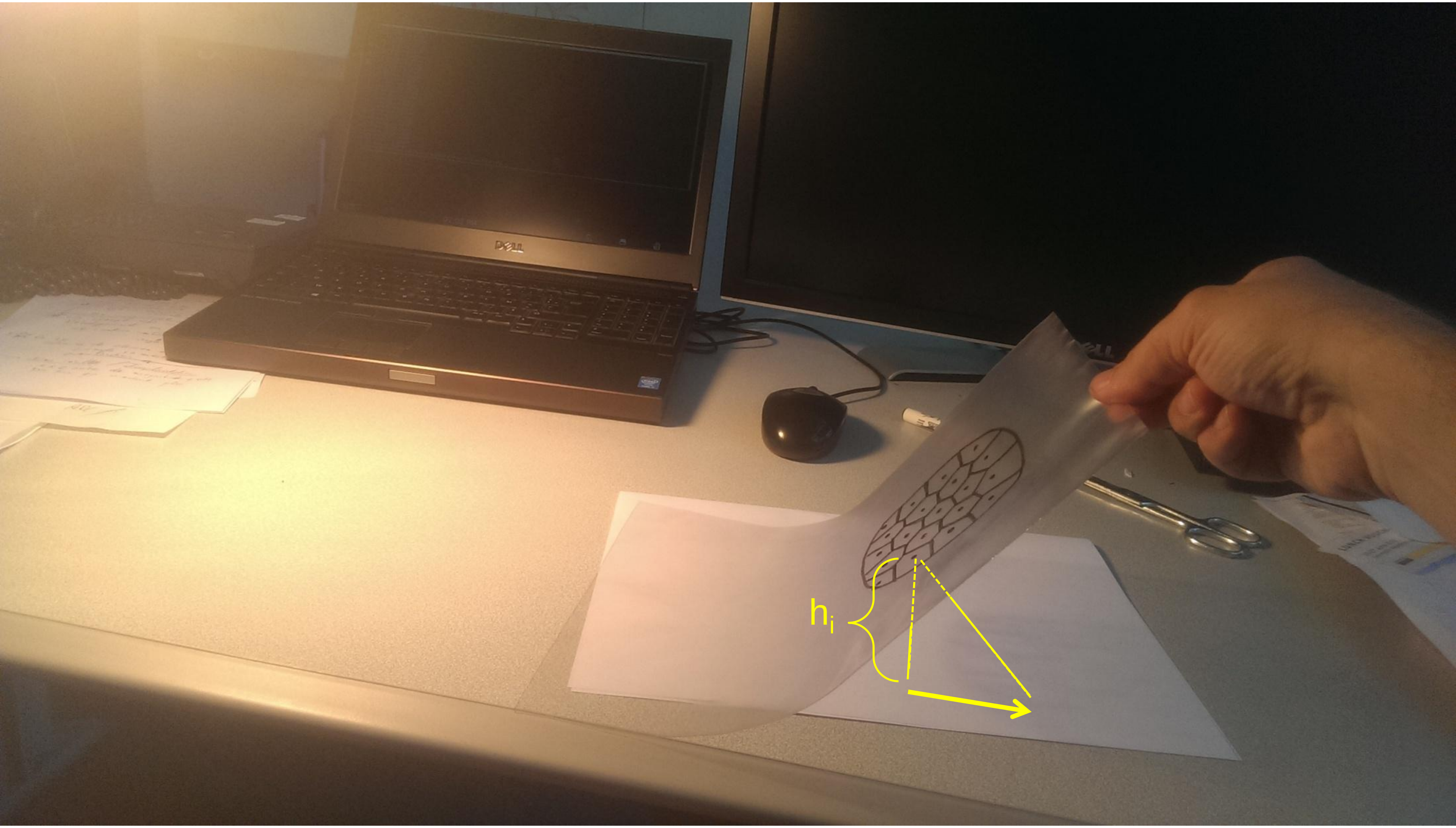
Part. 3 Power Diagrams & Transport



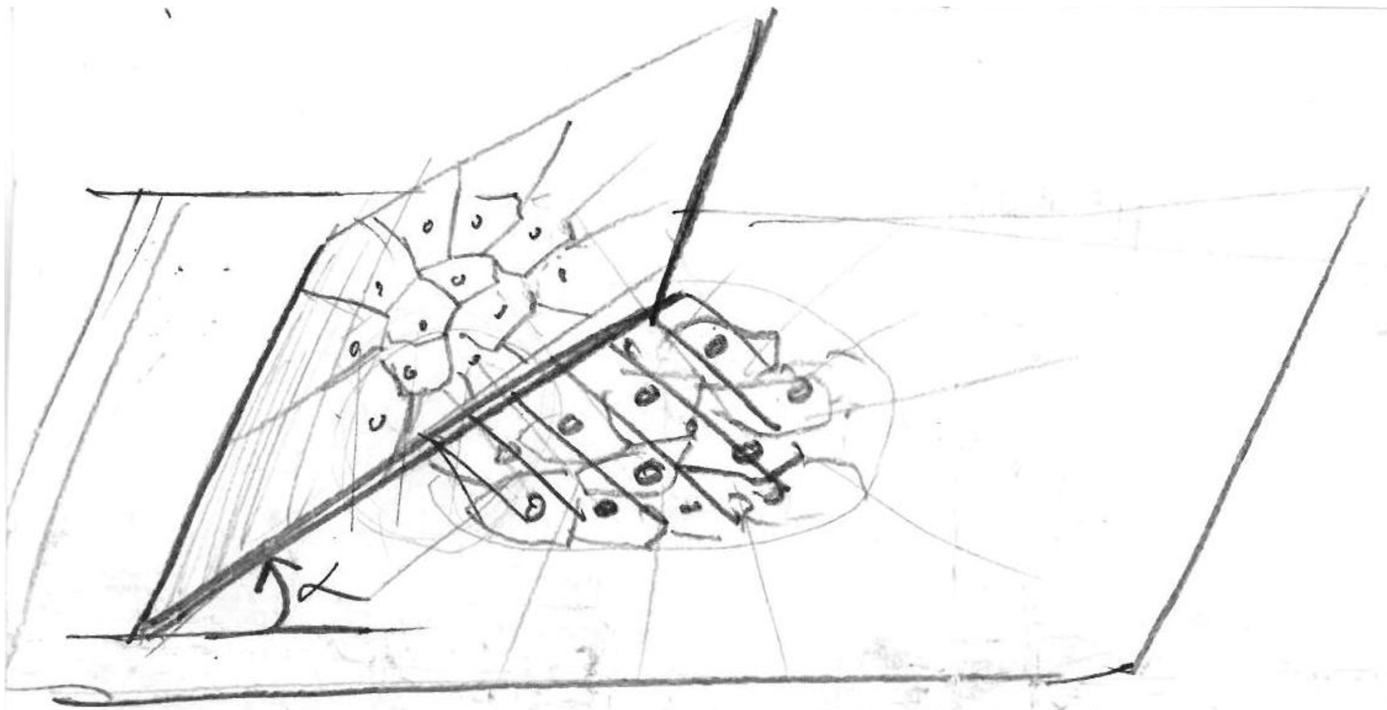
Part. 3 Power Diagrams & Transport



Part. 3 Power Diagrams & Transport



Part. 3 Power Diagrams & Transport

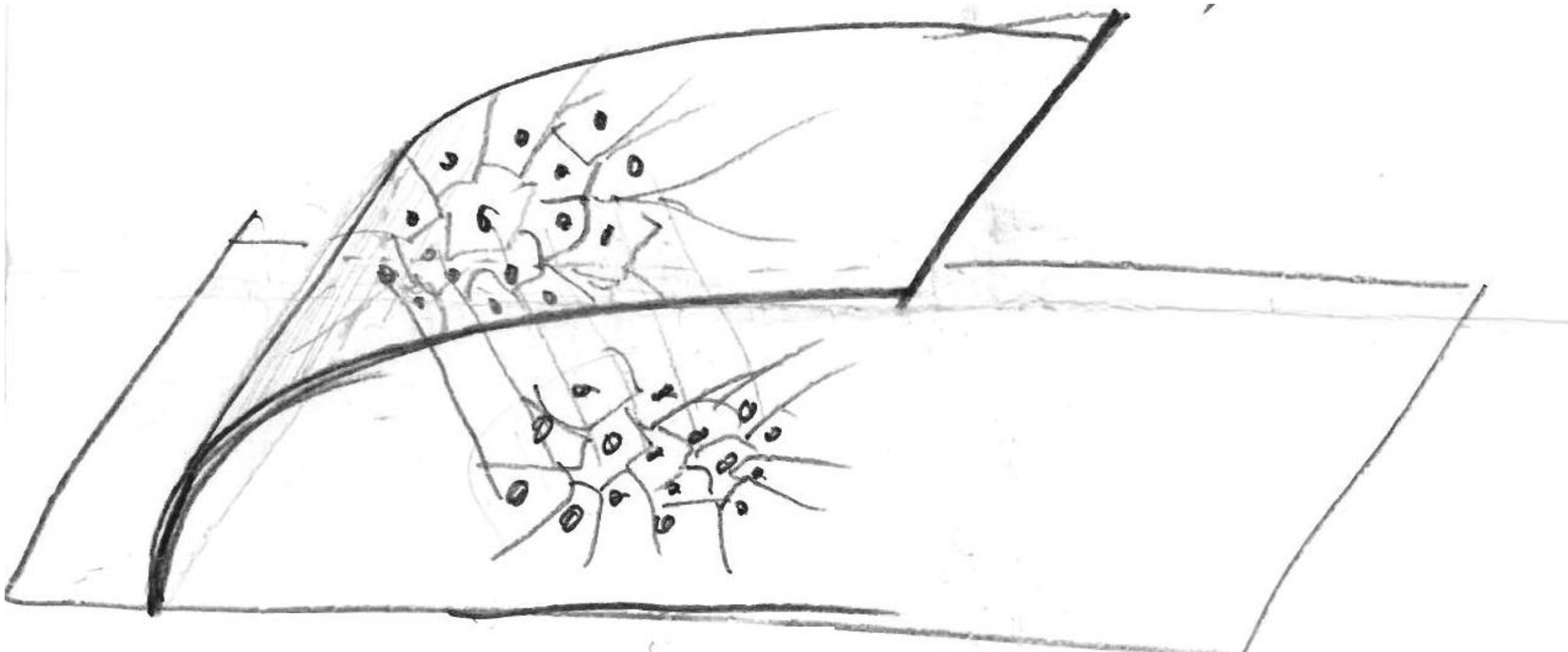


Translating a Voronoi diagram.

1st Try : linear lifting

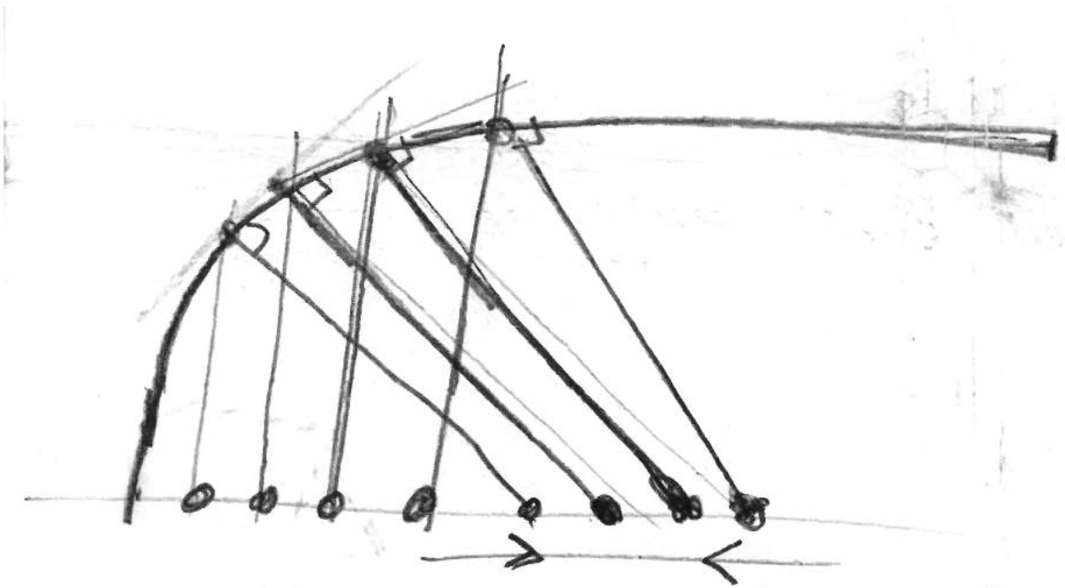
(FAIL : scales by 1/cos(α))

Part. 3 Power Diagrams & Transport



2nd Try : Curved lifting

Part. 3 Power Diagrams & Transport



"converging beams" can compensate the 1/cos(α) expansion by "re-concentrating" the points

Part. 3 Power Diagrams & Transport

$$d^2(p_i, q) \begin{array}{l} + h_i^2 \\ - w_i \end{array} < d^2(p_j, q) \begin{array}{l} + h_i^2 \\ - w_j \end{array} \quad \forall j$$

$$d^2(p_i, q - T) < d^2(p_j, q - T) \quad \forall j$$

$$(p_i - q + T)^2 < (p_j - q + T)^2 \quad \forall j$$

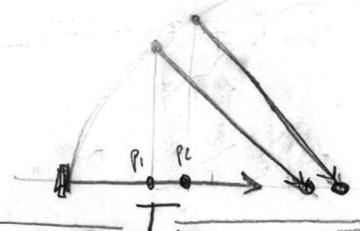
$$d^2(p_i, q) + 2T \cdot (p_i - q) + T^2 < d^2(p_j, q) + 2T(p_j - q) + T^2 \quad \forall j$$

$$d^2(p_i, q) + 2T \cdot p_i < d^2(p_j, q) + 2T \cdot p_j$$

$$\begin{array}{l} h_i^2 \\ w_i \end{array} \Big| = -2T \cdot p_i + cte$$

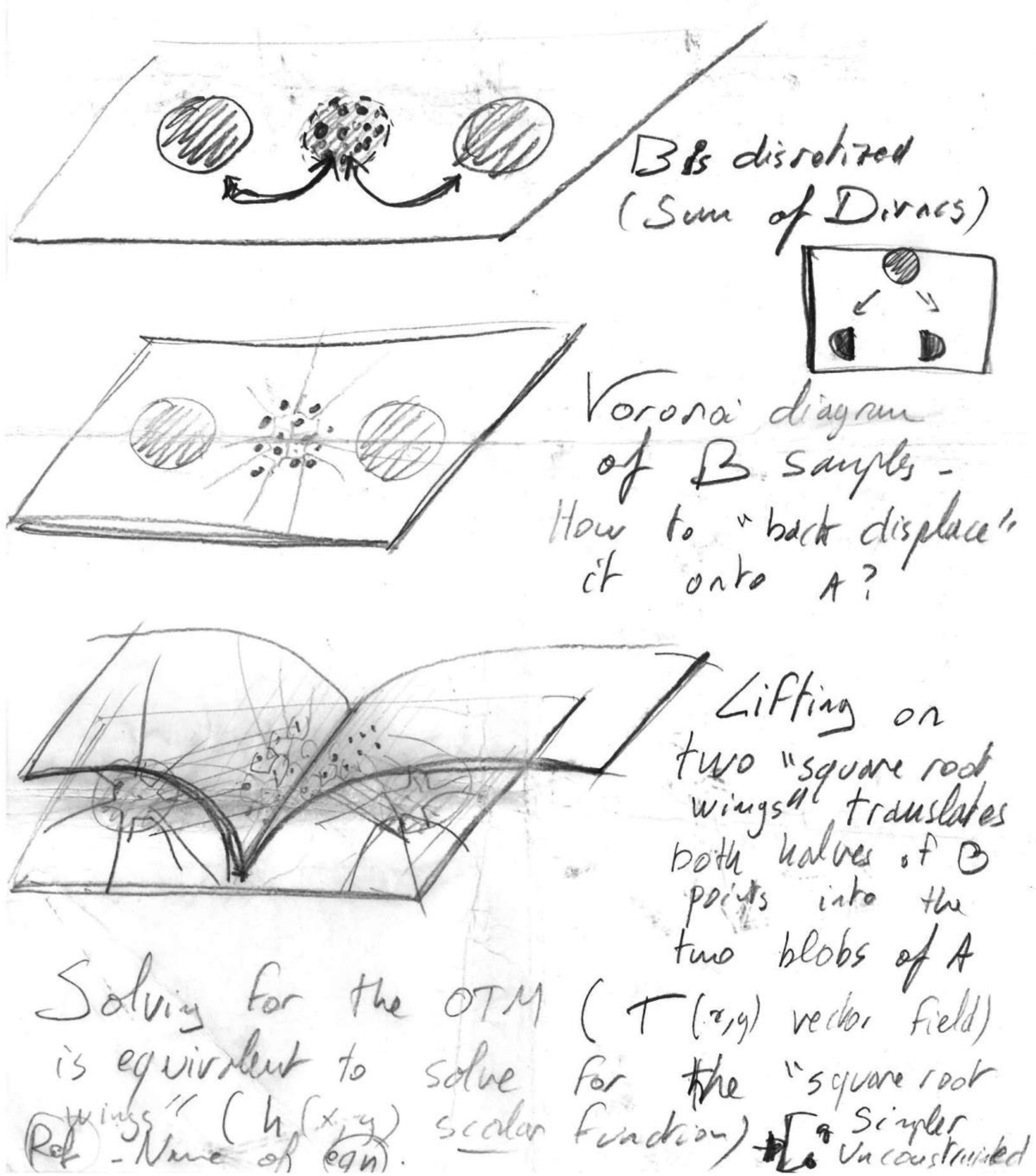
$$h_i^2 = (2T \cdot p_i + Cte)$$

$$h_i = \sqrt{2T \cdot p_i - \min_i (T \cdot p_i)}$$



Translation d'un diagramme de Voronoï sectionnel - Relèvement en racine carrée -

Part. 3 Power Diagrams and Transport



Part. 3 Optimal Transport – 2D examples

Numerical Experiment: *A disk becomes two disks*

4

Optimal Transport

applications in computational physics

Part. 4 Optimal Transport – 3D examples

Numerical Experiment: *A sphere becomes two spheres*

Part. 4 Optimal Transport – 3D examples

Numerical Experiment: *Armadillo to sphere*

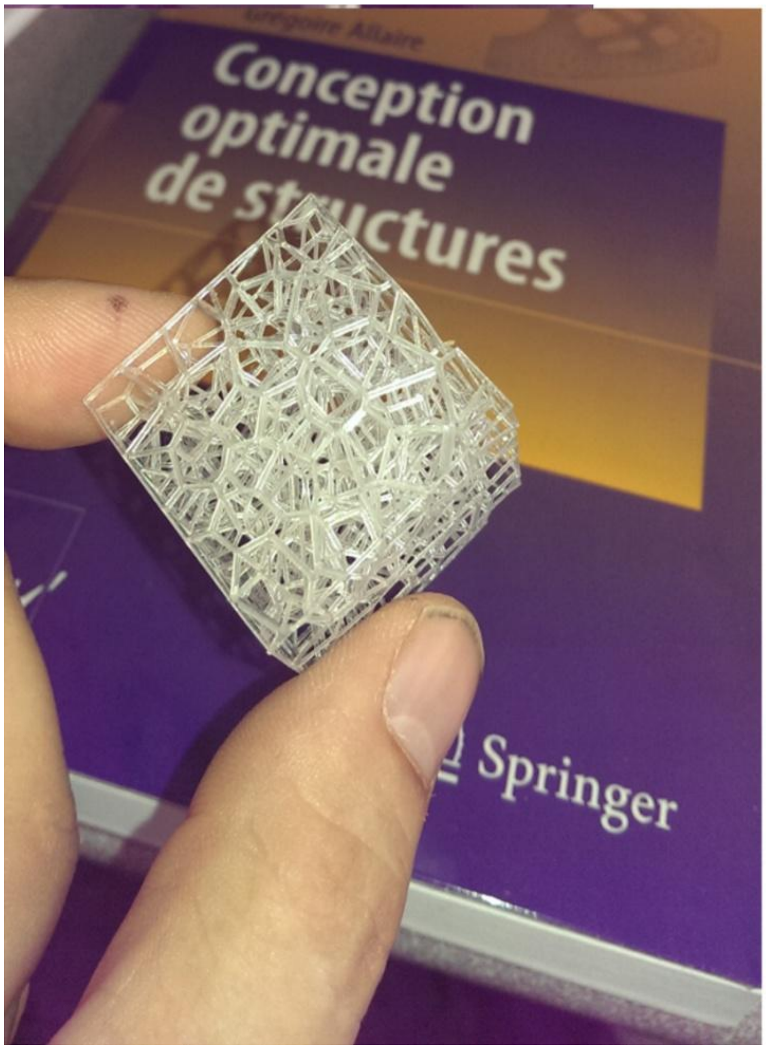
Part. 4 Optimal Transport – 3D examples

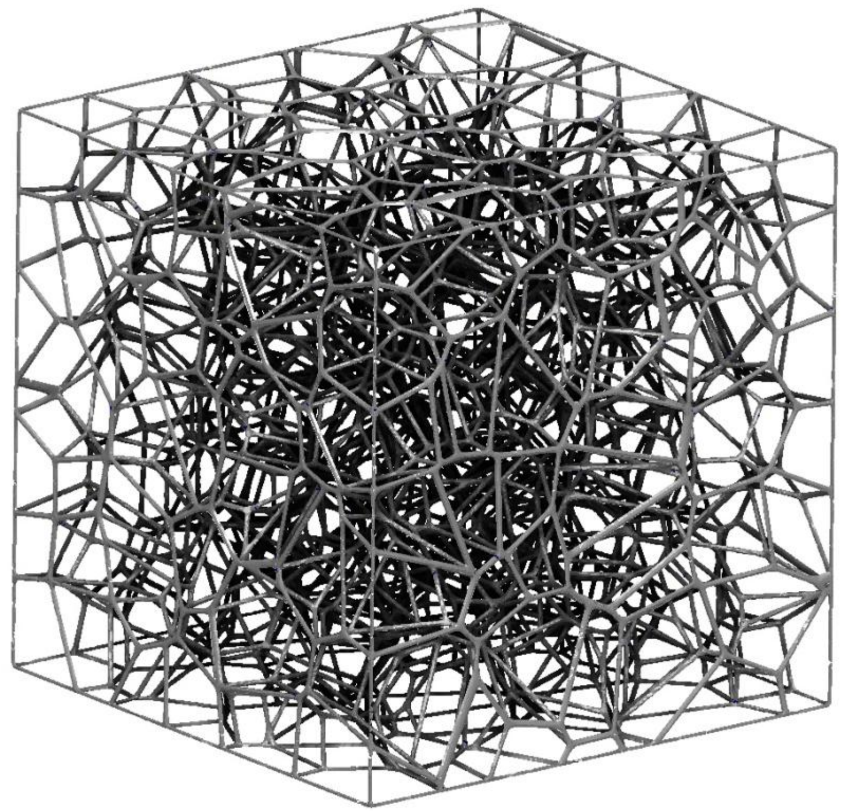
Numerical Experiment: *Other examples*

Computing singularity set…

Part. 4 Optimal Transport – 3D examples

Numerical Experiment: *Other examples*





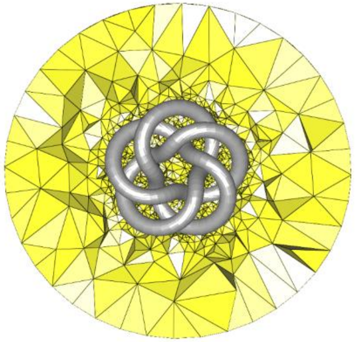
Inria

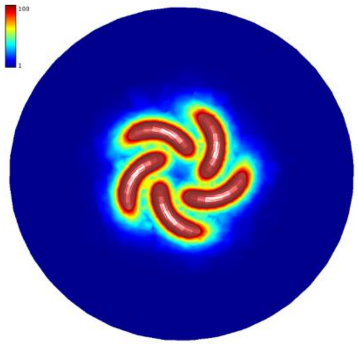
Part. 4 Optimal Transport – 3D examples

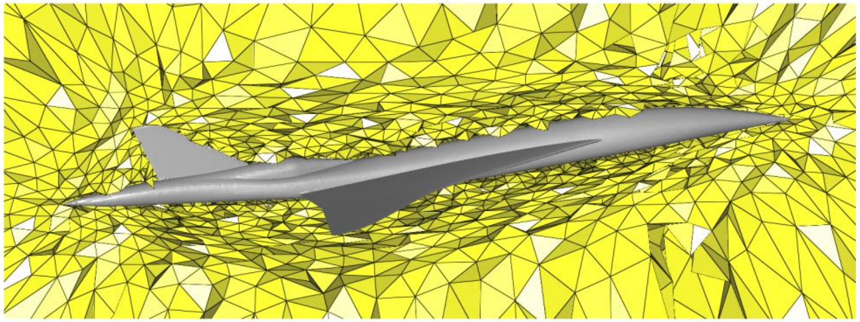
Numerical Experiment: *Varying density*

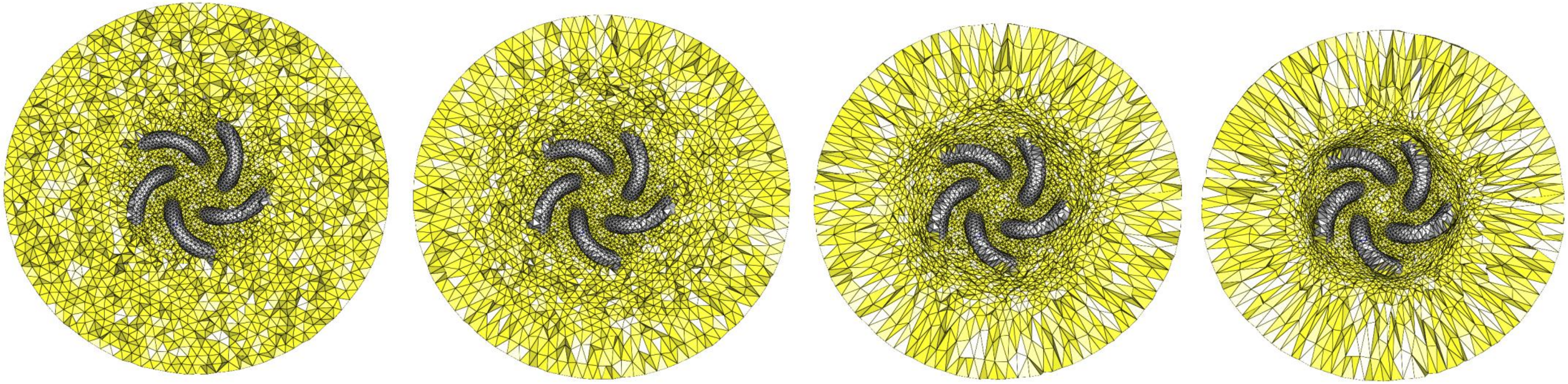
Part. 4 Optimal Transport – 3D examples

Numerical Experiment: *Varying density*

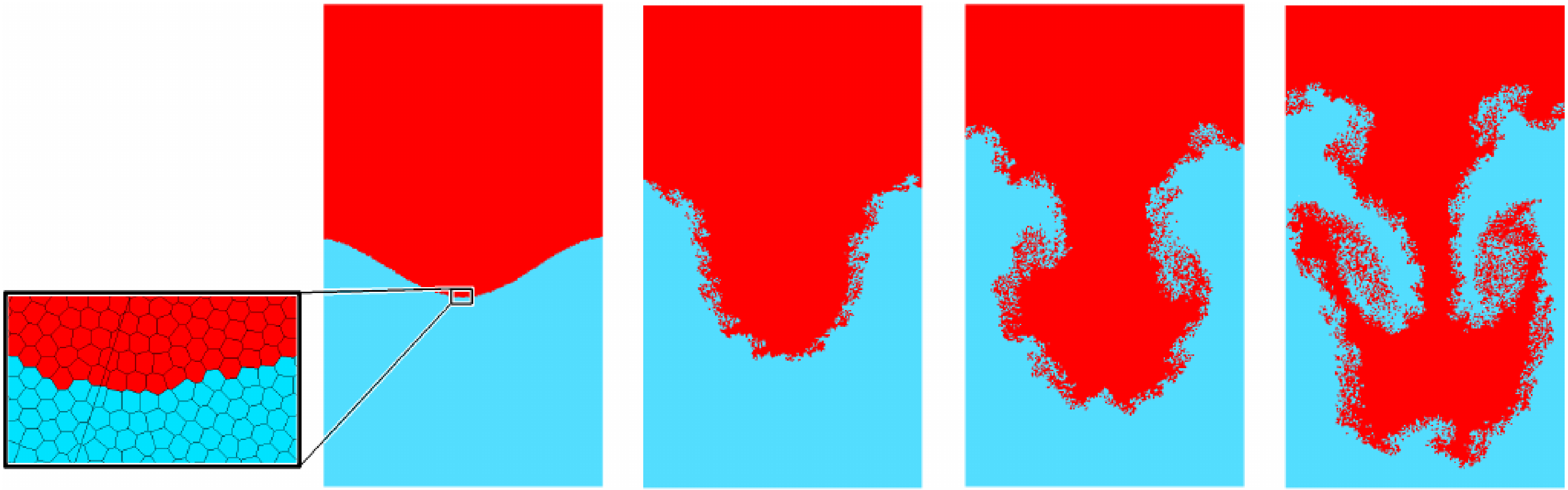








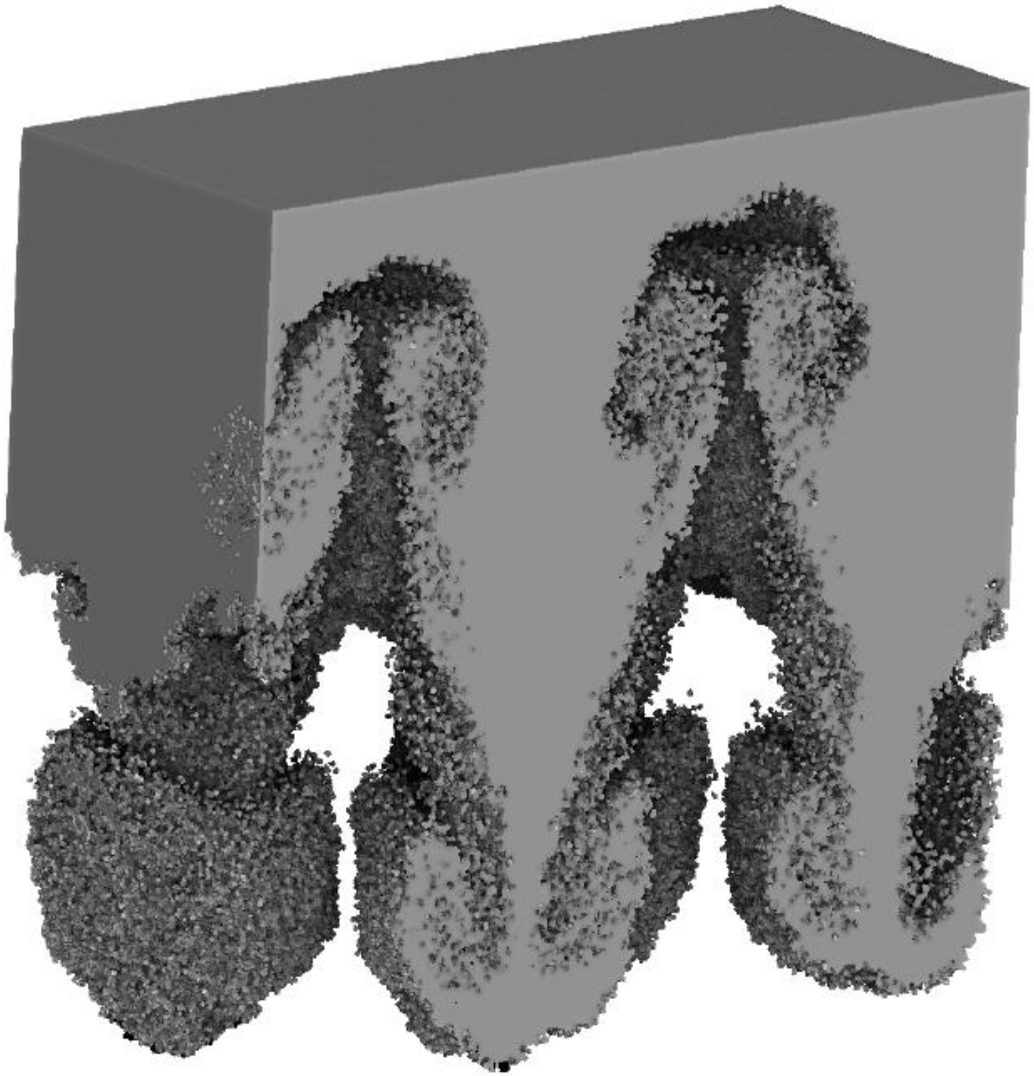
Part. 4 Optimal Transport – Fluids



Le schéma [Mérigot-Gallouet]
Applications en graphisme: [De Goes et.al] (power particles)

Part. 4 Optimal Transport – Fluids

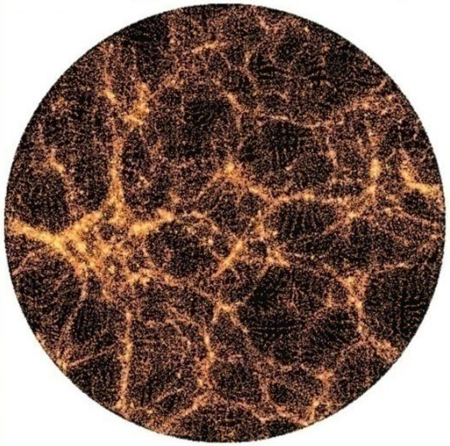
Part. 4 Optimal Transport – Fluids



Part. 4 Optimal Transport – Fluids

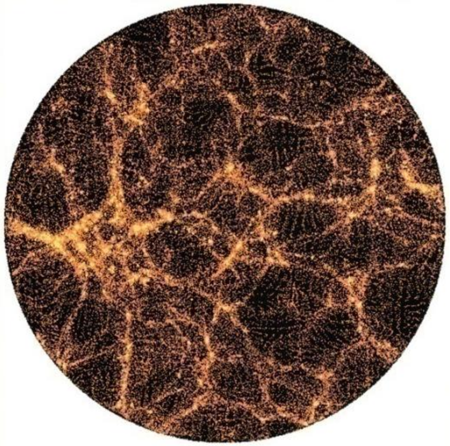
Part. 4 Optimal Transport – 3D examples

Numerical Experiment: *Performances*



2002: several hours of supercomputer time were needed for computing OT with a few thousand Dirac masses, with a com algorithm in $O(n^2 \log(n))$

Part. 4 Optimal Transport – 3D examples
Numerical Experiment: *Performances*



2002: several hours of supercomputer time were needed
for computing OT with a few thousand Dirac masses, with a com
algorithm in $O(n^2 \log(n))$

2015:

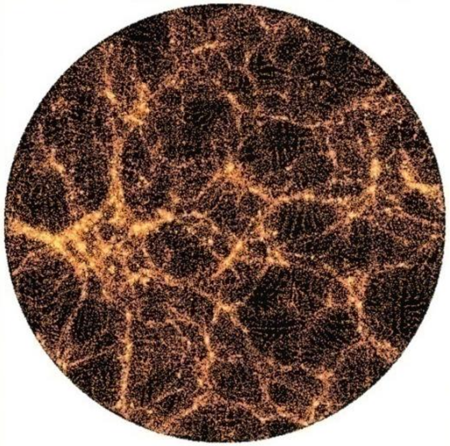
nb masses	1000	2000	5000	10000	30000	50000	10^5	3×10^5	5×10^5	10^6
time (s)	1.45	3.2	7.3	17.3	55	154	187	671	1262	2649

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.

3D version of [Mérigot] (multilevel + BFGS) + several tricks [**L** 2015]

Part. 4 Optimal Transport – 3D examples

Numerical Experiment: *Performances*



2002: several hours of supercomputer time were needed for computing OT with a few thousand Dirac masses, with a com algorithm in O(n^2log(n))

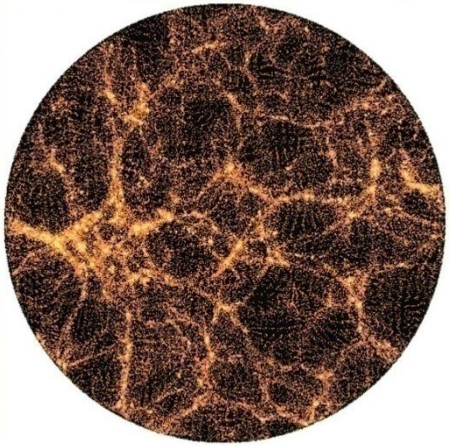
2015:

nb masses	1000	2000	5000	10000	30000	50000	10^5	3×10^5	5×10^5	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.

3D version of [Mérigot] (multilevel + BFGS) + several tricks [**L** 2015]

Part. 4 Optimal Transport – 3D examples
Numerical Experiment: *Performances*



2002: several hours of supercomputer time were needed for computing OT with a few thousand Dirac masses, with a com algorithm in $O(n^2 \log(n))$

2015:

nb masses	1000	2000	5000	10000	30000	50000	10^5	3×10^5	5×10^5	10^6
time (s)	1.45	3.2	7.3	17.3	55	154	187	671	1262	2649

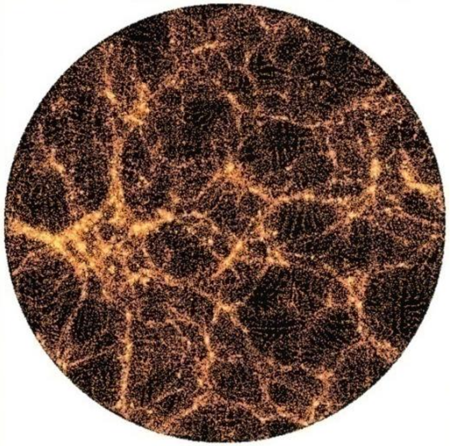
TABLE 4. Statistics for the Armadillo → sphere optimal transport with varying number of masses (see third row of Figure 12). Timings are given in seconds. The multi-level algorithm with BRIO pre-ordering and degree 2 regressions is used.

3D version of [Mérigot] (multilevel + BFGS) + several tricks [**L** 2015]

2016: Damped Newton [Mérigot, Thibert] + several tricks for 3D:
1 million Dirac masses in 240 seconds

Part. 4 Optimal Transport – 3D examples

Numerical Experiment: *Performances*



2002: several hours of supercomputer time were needed for computing OT with a few thousand Dirac masses, with a com algorithm in $O(n^2 \log(n))$

2015:

nb masses	1000	2000	5000	10000	30000	50000	10^5	3×10^5	5×10^5	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.

3D version of [Mérigot] (multilevel + BFGS) + several tricks [**L** 2015]

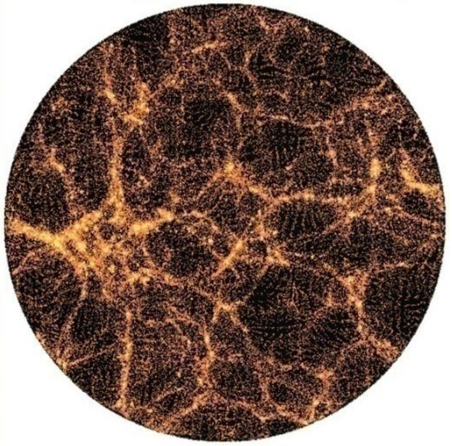
2016: Damped Newton [Mérigot, Thibert] + several tricks for 3D:

1 million Dirac masses in 240 seconds

10 million Dirac masses in 90 minutes

Part. 4 Optimal Transport – 3D examples

Numerical Experiment: *Performances*



2002: several hours of supercomputer time were needed for computing OT with a few thousand Dirac masses, with a com algorithm in O(n^{2}log(n))

2015:

nb masses	1000	2000	5000	10000	30000	50000	10^5	3×10^5	5×10^5	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.

3D version of [Mérigot] (multilevel + BFGS) + several tricks [**L** 2015]

2016: Damped Newton [Mérigot, Thibert] + several tricks for 3D:

1 million Dirac masses in 240 seconds

10 million Dirac masses in 90 minutes

Semi-discrete OT is now scalable ! (new tool in Num. Ana. Toolbox)

Conclusions – Open questions

* **Connections with physics, Legendre transform and entropy ?**

[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

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

Conclusions - 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)

Filippo Santambrogio – Optimal Transport for Applied Mathematician, Calculus of Variations, PDEs and Modeling – Jan 15, 2015

Gabriel Peyré, Marco Cuturi, Computational Optimal Transport, 2018

Online resources

All the sourcecode/documentation available from:
http://alice.loria.fr/software/geogram
Demo: www.loria.fr/~levy/GLUP/vorpaview

* **L.**, A numerical algorithm for semi-discrete L2 OT in 3D, ESAIM Math. Modeling and Analysis, 2015

* **L.** and E. Schwindt, Notions of OT and how to implement them on a computer, Computer and Graphics, 2018.