OPTIMIZATION PROBLEMS IN MASS TRANSPORTATION THEORY

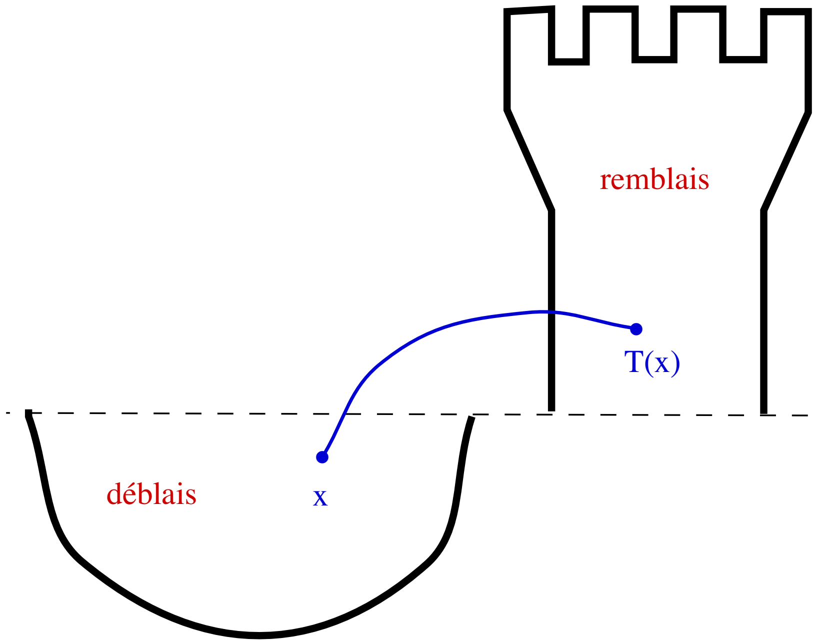
Giuseppe Buttazzo

Dipartimento di Matematica
Università di Pisa

buttazzo@dm.unipi.it
http://cvgmt.sns.it

School on Calculus of Variations
Roma July 4–8, 2005

Mass transportation theory goes back to Gaspard Monge (1781) when he presented a model in a paper on *Académie de Sciences de Paris*



The elementary work to move a particle x into $T(x)$ is given by $|x - T(x)|$, so that the total work is

$$\int_{\text{déblais}} |x - T(x)|\, dx \ .$$

A map T is called admissible transport map if it maps "déblais" into "remblais".

The Monge problem is then

$$\min \left\{ \int_{\text{déblais}} |x - T(x)| \, dx \ : \ T \text{ admiss.} \right\}.$$

It is convenient to consider the Monge problem in the framework of metric spaces:

- (X, d) is a metric space;
- f^+, f^- are two probabilities on X ($f^+ =$ "déblais", $f^- =$ "remblais");
- T is an admissible transport map if $T^{\#} f^+ = f^-$.

The Monge problem is then

$$\min \left\{ \int_X d\big(x, T(x)\big) \, dx \ : \ T\text{admiss.} \right\}.$$

In general the problem above does not admit a solution, when the measures f^+ and f^- are singular, since the class of admissible transport maps can be empty.

Example Take the measures

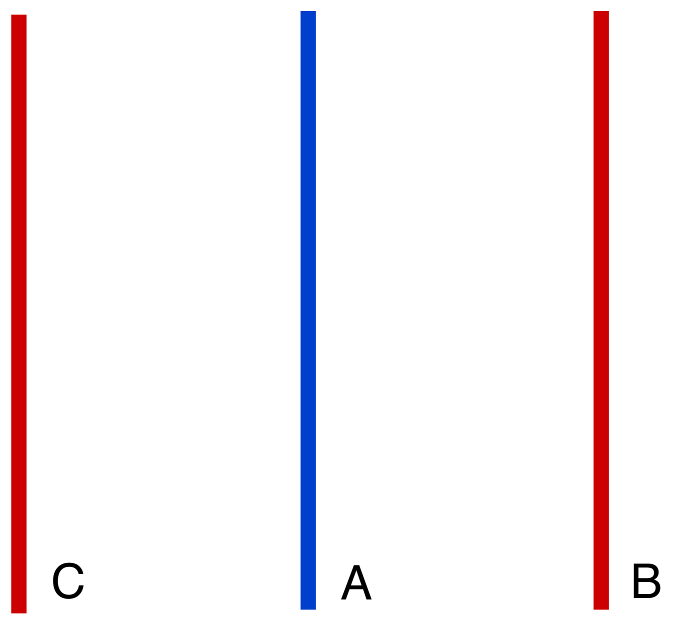
$$f^+ = \delta_A \quad \text{and} \quad f^- = \frac{1}{2}\delta_B + \frac{1}{2}\delta_C \ ;$$

it is clear that no map T transports f^+ into f^- so the Monge formulation above is in this case meaningless.

Example Take the measures in $\mathbf{R}^2$, still singular but nonatomic

$$f^+ = \mathcal{H}^1 \lfloor A \quad \text{and} \quad f^- = \frac{1}{2}\mathcal{H}^1 \lfloor B + \frac{1}{2}\mathcal{H}^1 \lfloor C$$

where A, B, C are the segments below.



Then the class of admissible transport maps is nonempty but the minimum in the Monge problem is not attained. Indeed, if the distance between the lines is L and the height is H it can be seen that the infimum of the Monge cost is HL, while every transport map T has a cost strictly greater than HL.

Example (book shifting) Consider in $\mathbf{R}$ the measures

$$f^+ = 1_{[0,a]}\mathcal{L}^1, \qquad f^- = 1_{[b,a+b]}\mathcal{L}^1.$$

Then the two maps

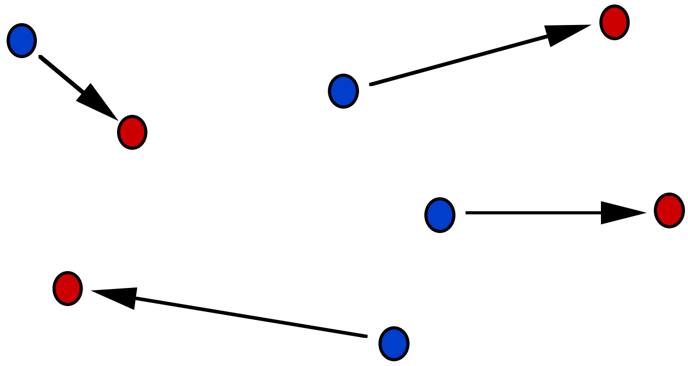
$$T_1(x) = b + x, \qquad T_2(x) = a + b - x$$

are both optimal; the map T_1 corresponds to a translation, while the map T_2 corresponds to a reflection. Indeed there are infinitely many optimal transport maps.

Example Take

$$f^{+} = \sum_{i=1}^{N} \delta_{p_i} \qquad f^{-} = \sum_{i=1}^{N} \delta_{n_i}.$$

Then the optimal Monge cost is given by the minimal connection of the p_i with the n_i.



Relaxed formulation (due to Kantorovich): consider measures γ on $X \times X$

- γ is an admissible transport plan if $\pi_1^{\#}\gamma = f^{+}$ and $\pi_2^{\#}\gamma = f^{-}$.
- Transport plans γ that are concentrated on γ-measurable graphs are actually transport maps $T : X \to X$, via the equality

$$\gamma = (Id \times T)^{\#} f^{+} \, .$$

Monge-Kantorovich problem:

$$\min\Big\{\int_{X\times X} d(x,y)\,d\gamma(x,y)\ :\ \gamma \text{ admiss.}\Big\}.$$

Theorem *There exists an optimal transport plan γ_{opt}; in the Euclidean case γ_{opt} is actually a transport map T_{opt} whenever f^+ and f^- are in L^1.*

The existence part is easy; indeed in the case X compact it follows from the weak* compactness of probabilities and the weak* continuity of the Monge-Kantorovich cost. In the general case the same holds by using the boundedness of the first moments. Note that, while in the Monge problem the cost is highly nonlinear with respect to the unknown T, in the Kantorovich formulation the cost is linear with respect to the unknown γ.

The Kantorovich formulation can be seen as the relaxation of the Monge problem; indeed, if f^+ is nonatomic we have

$$\min(\text{Kantorovich}) = \inf(\text{Monge}).$$

The existence of an optimal transport map is a rather delicate question. The first (incomplete) proof is due to Sudakov (1979); various different proofs have been given by Evans-Gangbo (1999), Trudinger-Wang (2001), Caffarelli-Feldman-McCann (2002), Ambrosio (2003).

Wasserstein distance of exponent p: replace the cost by $\left(\int_{X\times X} d^p(x,y)\,d\gamma(x,y)\right)^{1/p}$.

Most of the results remain valid for more general cost functions $c(x,y)$.

Shape optimization problems

There is a strong link between mass transportation and shape optimization problems. We present here this relation and we refer to the papers by Bouchitté-Buttazzo [BB] and Bouchitté-Buttazzo-Seppecher [BBS] for all details.

Shape optimization problem in elasticity:

given a force field f in $\mathbf{R}^n$ find the elastic body Ω whose "resistance" to f is maximal

Constraints:

- given volume, $|\Omega| = m$
- possible "design region" D given, $\Omega \subset D$
- possible support region Σ given, Dirichlet region

Optimization criterion: elastic compliance

More precisely, for every admissible domain Ω we consider the energy

$$\mathcal{E}(\Omega) = \inf \Big\{ \int_\Omega j(Du)\, dx - \langle f, u \rangle \ : \\ u \text{ smooth}, \ u = 0 \text{ on } \Sigma \Big\}$$

and the elastic compliance

$$\mathcal{C}(\Omega) = -\mathcal{E}(\Omega).$$

In the linear case, when j is a quadratic form, an integration by parts gives that the elastic compliance reduces to the work of external forces

$$\mathcal{C}(\Omega) = \frac{1}{2} \langle f, u_\Omega \rangle$$

being u_Ω the displacement of minimal energy in Ω. In linear elasticity, if $z^* = sym(z)$ and α, β are the Lamé constants,

$$j(z) = \beta |z^*|^2 + \frac{\alpha}{2} |tr z^*|^2.$$

Then the shape optimization problem is

$$\min\Big\{\mathcal{C}(\Omega)\ :\ \Omega\subset D,\ |\Omega|=m\Big\}.$$

A similar problem can be considered in the scalar case (optimal conductor), where f is a scalar function (the heat sources density) and

$$j(z)=\frac{1}{2}|z|^2.$$

Note that the minimum problem with respect to u depends on several elements $(\Omega,\Sigma,D,f\dots)$. Sometimes they are considered as given data, which for instance occurs in the usual problems of the calculus of variations, while sometimes they are unknown to be optimized for some given criterium (here the compliance).

The shape optimization problem above has in general no solution; in fact minimizing sequences may develop wild oscillations which give raise to limit configurations that are not in a form of a domain.

Therefore, in order to describe the behaviour of minimizing sequences a relaxed formulation is needed.

The literature on this problem is very wide; starting from the works by Murat and Tartar in the '70, it has been studied by many authors (Kohn and Strang, Lurie and Cherkaev, Allaire, Bendsøe, ...). There are strong links with homogenization theory.

A large parallel literature is present in engineering.

The complete relaxed form of the problem above is not known: several difficulties arise. For instance:

- Lack of coercivity; in most of the literature mixtures of two homogeneous isotropic materials are considered. Here one of the two materials is the empty space (outside Ω).

- Even for mixtures of two materials in the elasticity case the set of materials attainable by relaxation is not known.

- Possibility of nonlocal problems as limits of admissible sequences of domains.

- When D is unbounded, it is not clear that data with compact support provide minimizing sequences of domains that are equi-bounded.

To overcome some of these difficulties it is convenient to study a slightly different problem where we have the freedom to distribute a given quantity of material without the restriction of designing a domain Ω. They are called **Mass Optimization Problems** and consist in finding, given a force field f in $\mathbf{R}^n$ (which may possibly concentrate on sets of lower dimension), the mass distribution μ whose "resistance" to f is maximal.

Constraints:

- given mass, $\int d\mu = m$
- possible design region D given,

  with $\operatorname{spt} \mu \subset D$
- possible support region Σ given,

  Dirichlet region.

Optimization criterium:
elastic compliance

As before for every admissible mass distribution μ we consider the energy

$$\mathcal{E}(\mu) = \inf\Big\{ \int j(Du)\, d\mu - \langle f, u\rangle \ : \ u \text{ smooth}, \ u = 0 \text{ on } \Sigma \Big\}$$

and the elastic compliance $\mathcal{C}(\mu) = -\mathcal{E}(\mu)$. Then the mass optimization problem is

$$\min\Big\{ \mathcal{C}(\mu) \ : \ \operatorname{spt}\mu \subset D, \ \int d\mu = m \Big\}.$$

Elasticity: u and f vector-valued,

$$j(z) = \beta|z^*|^2 + \frac{\alpha}{2}|tr\, z^*|^2.$$

Conductivity: u and f scalar,

$$j(z) = \frac{1}{2}|z|^2.$$

In the mass optimization problem there is more freedom than in the shape optimization problem because we may use measures different from $\mu = 1_\Omega \, dx$, as it occurs in shape optimization problems.

On the other hand, we are not restricted only to forces in H^{-1} or to n or $n-1$ dimensional Dirichlet regions; for instance we may consider forces and Dirichlet regions concentrated in a single point, or in a finite number of points.

Of course, for a given f, we may have $\mathcal{E}(\mu) = -\infty$ for some measures μ; for instance this occurs when f is a Dirac mass and μ is the Lebesgue measure on a domain Ω, but these measures with infinite compliance are ruled out by the optimization criterion which consists in maximizing $\mathcal{E}(\mu)$.

There is a strong link between the mass optimization problem and the Monge-Kantorovich mass transfer problem.

This is described below in the scalar case, the elasticity case being still not completely understood.

When in the mass optimization problem we have $\Sigma = \emptyset$ and $D = \mathbf{R}^n$, then the associated distance in the transportation problem is the Euclidean one.

On the contrary, if a Dirichlet region Σ and design region D are present, the Euclidean distance $|x - y|$ has to be replaced by another distance $d(x, y)$ which measures the geodesic distance in D between the points x and y, counting for free the paths along Σ.

More precisely, we consider the geodesic distance on $D \times D$

$$d_D(x,y) = \min\Big\{ \int_0^1 |\gamma'(t)|\,dt \ : \\ \gamma(t) \in D,\ \gamma(0)=x,\ \gamma(1)=y \Big\} \\ = \sup\big\{|\varphi(x)-\varphi(y)| \ : \\ |D\varphi| \le 1 \text{ on } D\big\}$$

and its modification by Σ

$$d_{D,\Sigma}(x,y) = \inf\Big\{ d_D(x,y) \wedge \big(d_D(x,\xi_1) \\ + d_D(y,\xi_2)\big) \ : \ \xi_1,\xi_2 \in \Sigma \Big\} \\ = \sup\big\{|\varphi(x)-\varphi(y)| \ : \\ |D\varphi| \le 1 \text{ on } \Omega,\ \varphi = 0 \text{ on } \Sigma\big\}.$$

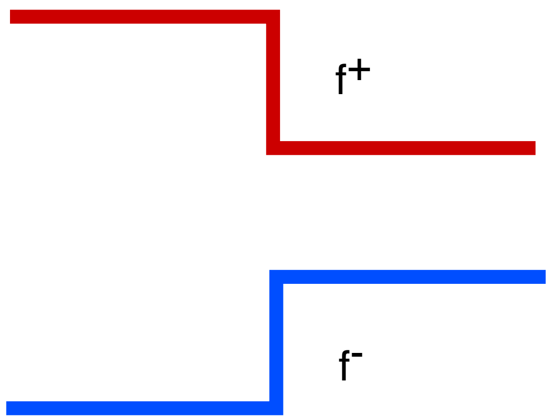
This semi-distance will play the role of cost function in the Monge-Kantorovich mass transport problem.

When $f = f^+ - f^-$ is a smooth function, $D = \mathbf{R}^n$, $\Sigma = \emptyset$, Evans and Gangbo [Mem. AMS 1999] have shown that the solution can be found by solving the PDE

$$\begin{cases} -\operatorname{div}\big(a(x)Du\big) = f \\ u \text{ is 1-Lipschitz} \\ |Du| = 1 \text{ a.e. on } \{a(x) > 0\} \end{cases}$$

where the coefficient a is a multiple of the optimal μ.

However, even simple cases with $f \in H^{-1}$ may produce optimal mass distributions which are singular measures. For instance



produces an optimal μ which has a onedimensional concentration in the vertical segment.

Then a theory of variational problems with respect to measures has to be developed. This has been done by Bouchitté, Buttazzo and Seppecher [Calc. Var. 1997]. The application of transport problems to mass optimization has been developed by Bouchitté and Buttazzo [JEMS 2001]. The key tools are:

- the tangent space $T_\mu(x)$ to a measure μ, defined for μ a.e. x;
- the tangential gradient D_μ;
- the Sobolev spaces $W^{1,p}_\mu$.

Then the transport problem becomes equivalent to the PDE

$$\begin{cases} -\operatorname{div}\big(\mu(x)D_\mu u\big) = f \quad \text{in } \mathbf{R}^n \setminus \Sigma \\ u \text{ is 1-Lipschitz on } D, \quad u = 0 \text{ on } \Sigma \\ |D_\mu u| = 1 \ \mu\text{-a.e. on } \mathbf{R}^n, \quad \mu(\Sigma) = 0. \end{cases}$$

Theorem. *For every measure f there exists a solution μ_{opt} of the mass optimization problem. Moreover, a measure μ solves the mass optimization problem if and only if (a multiple of) it solves the Monge-Kantorovich PDE.*

In order to describe the tangential gradients $D_\mu u$ which appear in the Monge-Kantorovich equation above we recall the main steps of the theory of the variational integrals w.r.t. a measure.

Variational integrals w.r.t. a measure

Consider the functional

$$F(u) = \int f(x, Du)\, d\mu$$

defined for smooth functions u (and $+\infty$ elsewhere), where μ is a general nonnegative measure. On the integrand f we assume the usual convexity and p-growth conditions. We define the tangent fields

$$X_\mu^{p'} = \{\phi \in L_\mu^{p'}(\mathbf{R}^n; \mathbf{R}^n) \ : \\ \operatorname{div}(\phi\mu) \in L_\mu^{p'}(\mathbf{R}^n)\},$$

the tangent spaces

$$T_\mu^p(x) = \mu\text{-ess} \bigcup \{\phi(x) \ : \ \phi \in X_\mu^{p'}\},$$

the tangential gradient (for smooth functions)

$$D_\mu u(x) = P_\mu(x, Du(x)).$$

The operator D_μ is closable, i.e.

$$\begin{cases} u_h \to 0 \text{ weakly in } L^p_\mu \\ D_\mu u_h \to v \text{ weakly in } L^p_\mu \end{cases} \Rightarrow \; v = 0 \; \mu\text{a.e.}$$

and we still denote by D_μ its closure. The Sobolev space $W^{1,p}_\mu$ is then defined as the domain of the extension above, with norm

$$\|u\|_{1,p,\mu} = \|u\|_{L^p_\mu} + \|D_\mu u\|_{L^p_\mu} \; .$$

The relaxation of the integral functional above is given by

$$\overline{F}(u) = \int f_\mu(x, D_\mu u)\, d\mu \qquad u \in W^{1,p}_\mu$$

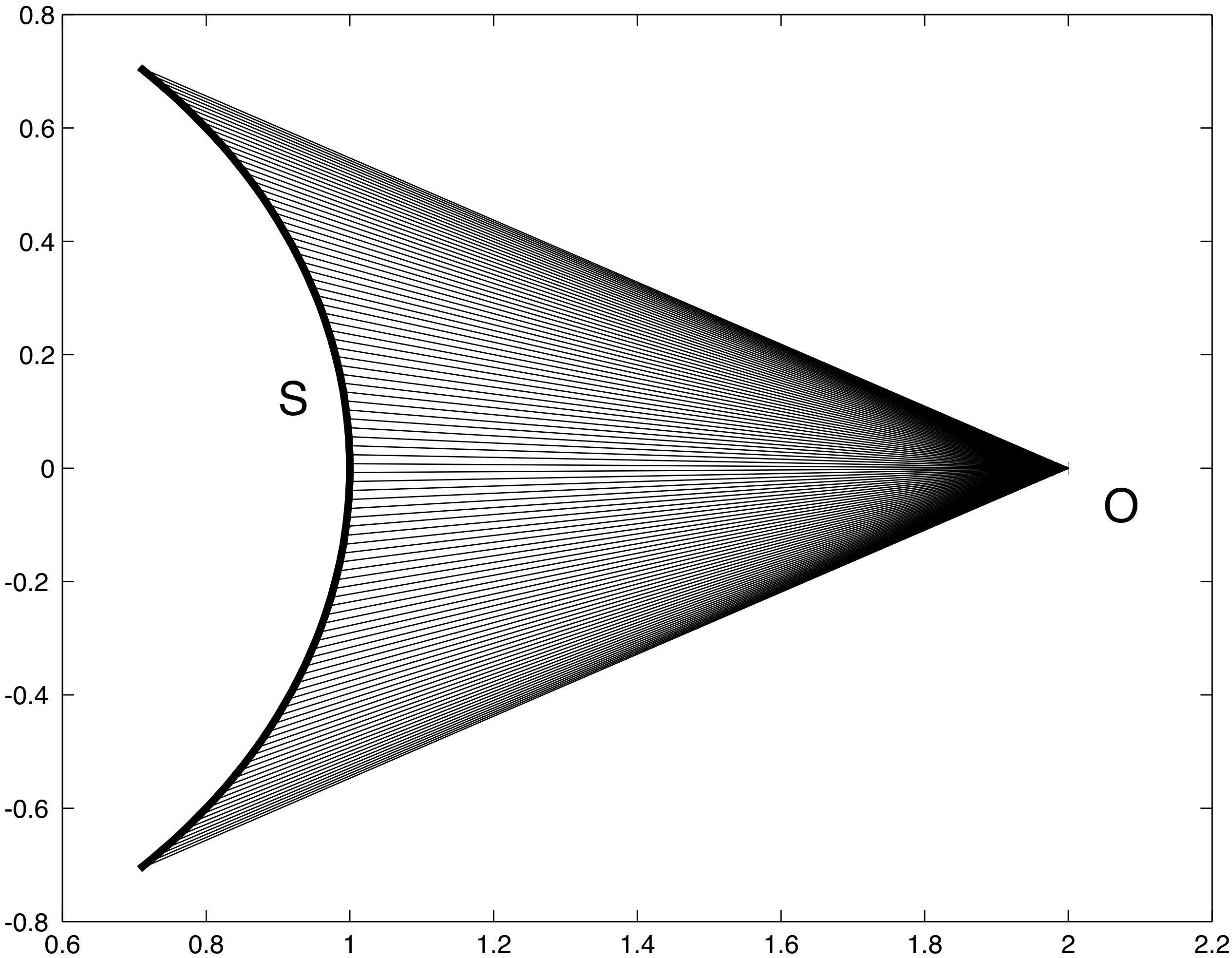
where

$$f_\mu(x,z) = \inf\left\{ f(x, z+\xi) \; : \; \xi \in \left(T^p_\mu(x)\right)^\perp \right\}.$$

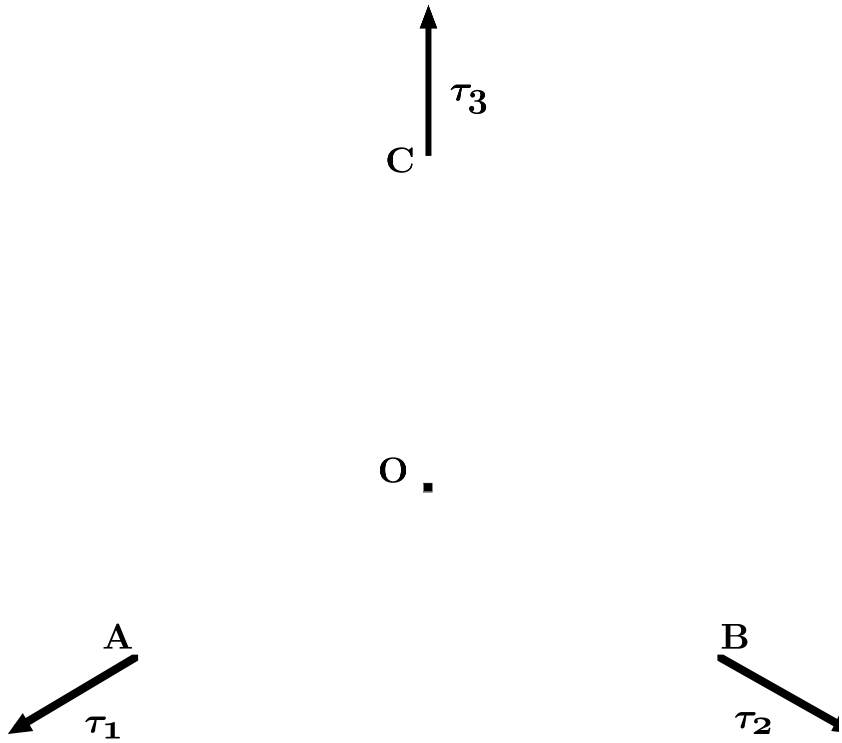
In particular,

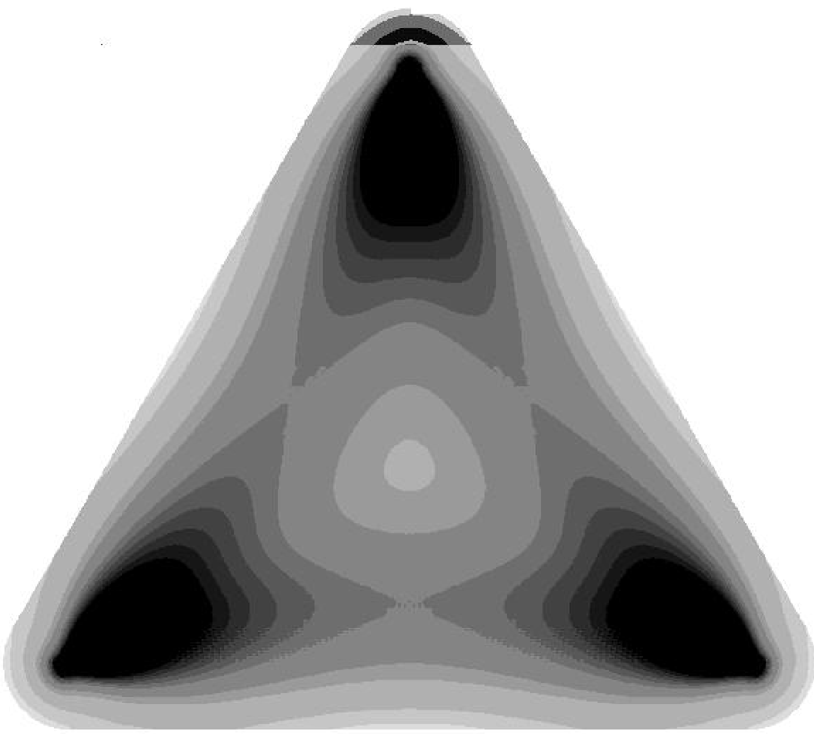
$$\int |Du|^2\, d\mu \quad \text{relaxes to} \quad \int |D_\mu u|^2\, d\mu \; .$$

Here are some cases where the optimal mass distribution can be computed by using the Monge-Kantorovich equation (see Bouchitté-Buttazzo [JEMS '01]).

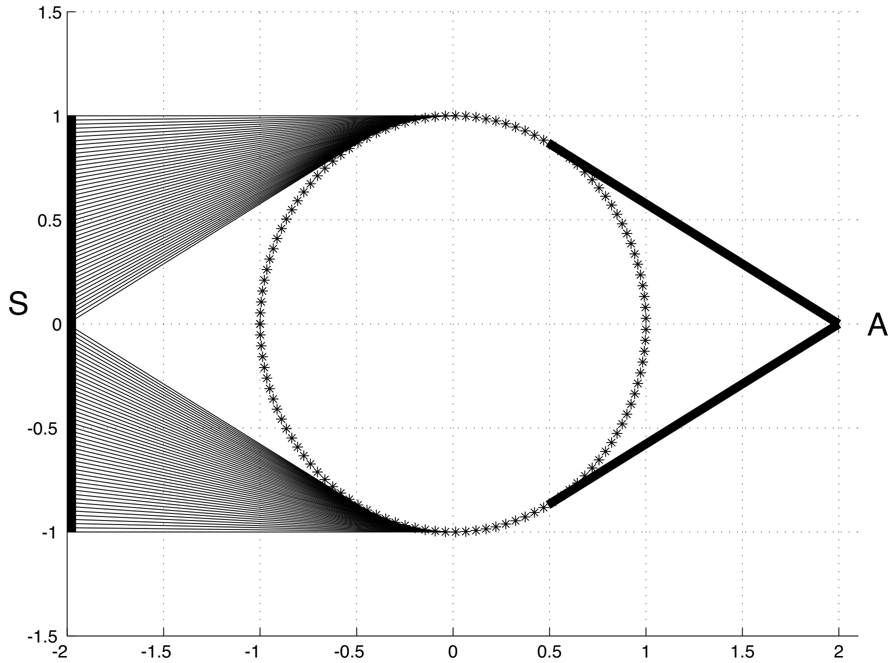


Optimal distribution of a conductor for heat sources $f = \mathcal{H}^1 \lfloor S - L\delta_O$.

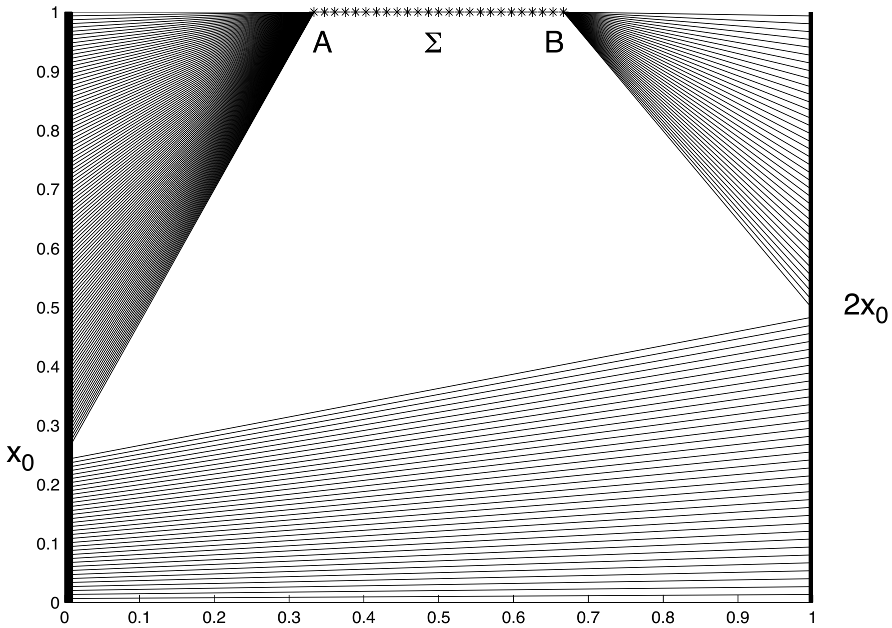




Optimal distribution of an elastic material when the forces are as above.



Optimal distribution of a conductor, with an obstacle, for heat sources $f = \mathcal{H}^1 \lfloor S - 2\delta_A$.



Optimal distribution of a conductor for heat sources $f = 2\mathcal{H}^1 \lfloor S_0 - \mathcal{H}^1 \lfloor S_1$ and Dirichlet region Σ.

We present now some further optimization problems related to mass transportation theory.

Given a metric space (X, d) and two probabilities f^+ and f^- on X, it is convenient to denote by $MK(f^+, f^-, d)$ the minimum value of the transportation cost in the Monge-Kantorovich problem, that is

$$MK(f^+, f^-, d) = \min\Big\{\int_{X\times X} d(x,y)\,d\gamma(x,y)$$
$$: \ \gamma \text{ has marginals } f^+, f^-\Big\}.$$

We will consider some data as fixed and we will let the remaining ones vary in some suitable admissible classes; the goal is to optimize some given total costs which include a term related to the mass transportation.

Problem 1 - Optimal Networks

We consider the following model for the optimal planning of an urban transportation network (Buttazzo-Brancolini COCV 2005).

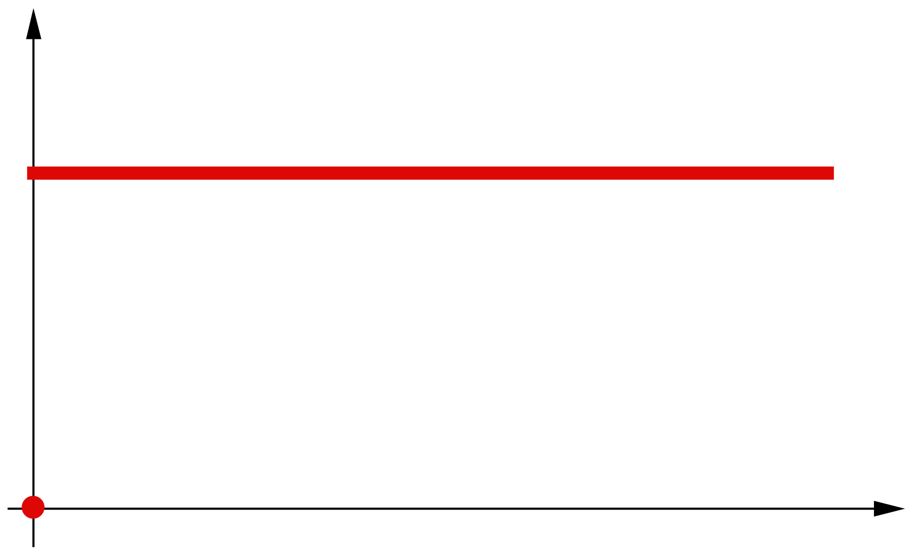
- Ω *the geographical region or urban area*
  a compact regular domain of $\mathbf{R}^N$
- f^+ *the density of residents*
  a probability measure on Ω
- f^- *the density of working places*
  a probability measure on Ω
- Σ *the transportation network*
  a closed connected 1-dimensional subset of Ω, the unknown.

The goal is to introduce a cost functional $F(\Sigma)$ and to minimize it on a class of admissible choices.

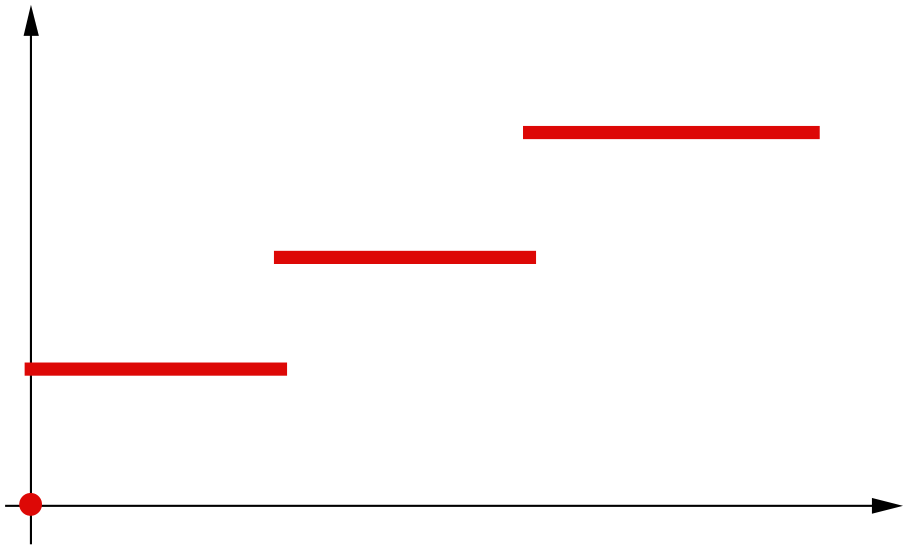
Consider two functions:

$A : \mathbf{R}^+ \to \mathbf{R}^+$ continuous and increasing; $A(t)$ represents the cost to cover a length t by one's own means (walking, time consumption, car fuel, ...);

$B : \mathbf{R}^+ \to \mathbf{R}^+$ l.s.c. and increasing; $B(t)$ represents the cost to cover a length t by using the transportation network (ticket, time consumption, ...).



Small town policy: only one ticket price



Large town policy: several ticket prices

We define

$$d_\Sigma(x,y) = \inf \Big\{ A\big(\mathcal{H}^1(\Gamma \setminus \Sigma)\big)$$
$$+ B\big(\mathcal{H}^1(\Gamma \cap \Sigma)\big) \;:\; \Gamma \text{ connects } x \text{ to } y \Big\}.$$

The cost of the network Σ is defined via the Monge-Kantorovich functional:

$$F(\Sigma) = MK(f^+, f^-, d_\Sigma)$$

and the admissible Σ are simply the closed connected sets with $\mathcal{H}^1(\Sigma) \leq L$.

Therefore the optimization problem is

$$\min\Big\{F(\Sigma)\ :\ \Sigma \text{ cl. conn.}, \mathcal{H}^1(\Sigma)\le L\Big\}.$$

Theorem *There exists an optimal network Σ_{opt} for the optimization problem above.*

In the special case $A(t) = t$ and $B \equiv 0$ (communist model) some necessary conditions of optimality on Σ_{opt} have been derived (Buttazzo-Oudet-Stepanov 2002 and Buttazzo-Stepanov 2003). For instance:

- no closed loops;
- at most triple point junctions;
- $120°$ at triple junctions;
- no triple junctions for small L;
- asymptotic behavior of Σ_{opt} as $L \to +\infty$ (Mosconi-Tilli JCA 2005);
- regularity of Σ_{opt} is an open problem.

It is interesting to study the optimization problem above if we drop the assumption that admissible Σ are connected. This could be for instance of interest in the case of residents spread over a large area. We take as admissible Σ all rectifiable sets with $\mathcal{H}^1(\Sigma) \le L$ (paper [BPSS] in preparation).

- For general functions A and B an optimal network Σ_{opt} may not exist; the optimum has to be searched in a relaxed sense among measures.

- If A and B are concave, then an optimal network Σ_{opt} exists; however there could also be other optima which are measures.

- If A and B are concave, one of them is strictly concave, and $B'_+(0) < A'_-(\operatorname{diam}\Omega)$, then all optima (also in a relaxed sense) are rectifiable networks.

Problem 2 - Optimal Pricing Policies

With the notation above, we consider the measures f^+, f^- fixed, as well as the transportation network Σ. The unknown is the pricing policy the manager of the network has to choose through the l.s.c. monotone increasing function B. The goal is to maximize the total income, a functional $F(B)$, which can be suitably defined (Buttazzo-Pratelli-Stepanov 2004) by means of the Monge-Kantorovich transport plans.

Of course, a too low ticket price policy will not be optimal, but also a too high ticket price policy will push customers to use their own transportation means, decreasing the total income of the company.

The function B can be seen as a control variable and the corresponding transport plan as a state variable, so that the optimization problem we consider:

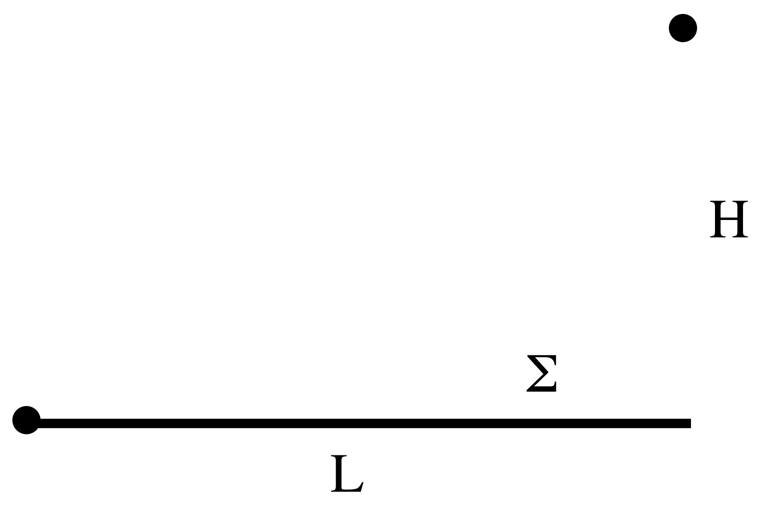
$$\min \left\{ F(B) \ : \ B \text{ l.s.c. increasing, } B(0) = 0 \right\}$$

can be seen as an optimal control problem.

Theorem *There exists an optimal pricing policy B_{opt} solving the maximal income problem above.*

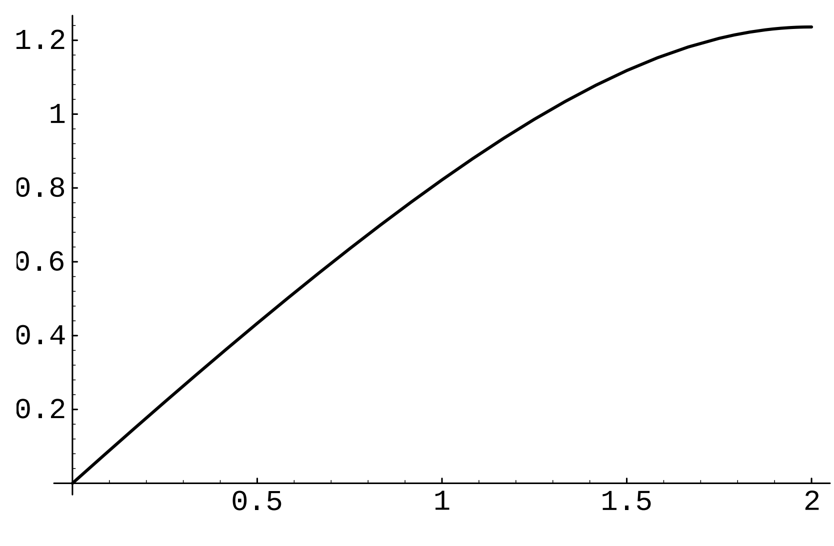
Also in this case some necessary conditions of optimality can be obtained. In particular, the function B_{opt} turns out to be continuous, and its Lipschitz constant can be bounded by the one of A (the function measuring the own means cost).

Here is the case of a service pole at the origin, with a residence pole at (L, H), with a network Σ. We take $A(t) = t$.



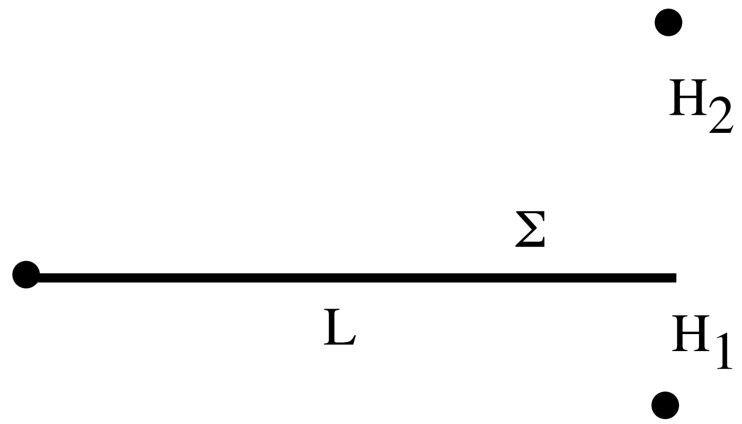
The optimal pricing policy $B(t)$ is given by

$$B(t) = (H^2 + L^2)^{1/2} - (H^2 + (L-t)^2)^{1/2}.$$



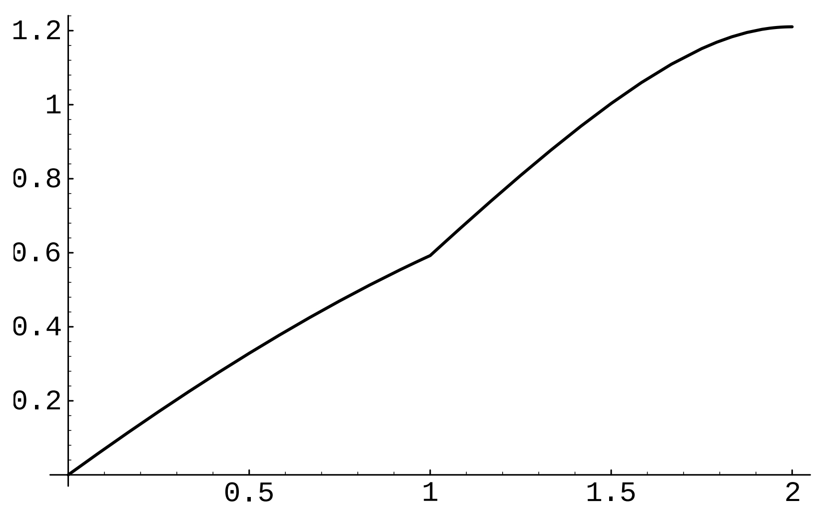
The case $L = 2$ and $H = 1$.

Here is another case, with a single service pole at the origin, with two residence poles at (L, H_1) and (L, H_2), with a network Σ.



The optimal pricing policy $B(t)$ is then

$$B(t) = \begin{cases} B_2(t) & \text{in } [0, T] \\ B_2(T) - B_1(T) + B_1(t) & \text{in } [T, L] \end{cases}.$$



The case $L = 2$, $H_1 = 0.5$, $H_2 = 2$.

Problem 3 - Optimal City Structures

We consider the following model for the optimal planning of an urban area (Buttazzo-Santambrogio SIAM M.A. 2005).

- Ω *the geographical region or urban area*
  a compact regular domain of $\mathbf{R}^N$
- f^+ *the density of residents*
  a probability measure on Ω
- f^- *the density of services*
  a probability measure on Ω.

Here the distance d in Ω is fixed (for simplicity we take the Euclidean one) while the unknowns are f^+ and f^- that have to be determined in an optimal way taking into account the following facts:

- there is a transportation cost for moving from the residential areas to the services poles;
- people desire not to live in areas where the density of population is too high;
- services need to be concentrated as much as possible, in order to increase efficiency and decrease management costs.

The transportation cost will be described through a Monge-Kantorovich mass transportation model; it is indeed given by a p-Wasserstein distance ($p \geq 1$) $W_p(f^+, f^-)$, being $p = 1$ the classical Monge case.

The total unhappiness of citizens due to high density of population will be described by a penalization functional, of the form

$$H(f^+) = \begin{cases} \int_\Omega h(u)\,dx & \text{if } f^+ = u\,dx \\ +\infty & \text{otherwise,} \end{cases}$$

where h is assumed convex and superlinear (i.e. $h(t)/t \to +\infty$ as $t \to +\infty$). The increasing and diverging function $h(t)/t$ then represents the unhappiness to live in an area with population density t.

Finally, there is a third term $G(f^-)$ which penalizes sparse services. We force f^- to be a sum of Dirac masses and we consider $G(f^-)$ a functional defined on measures, of the form studied by Bouchitté-Buttazzo (Nonlinear An. 1990, IHP 1992, IHP 1993):

$$G(f^-) = \begin{cases} \sum_n g(a_n) & \text{if } f^- = \sum_n a_n \delta_{x_n} \\ +\infty & \text{otherwise,} \end{cases}$$

where g is concave and with infinite slope at the origin. Every single term $g(a_n)$ in the sum represents here the cost for building and managing a service pole of dimension a_n, located at the point $x_n \in \Omega$.

We have then the optimization problem

$$\min\Big\{W_p(f^+,f^-)+H(f^+)+G(f^-)\ :\ f^+,f^- \text{ probabilities on } \Omega\Big\}.$$

Theorem *There exists an optimal pair* (f^+,f^-) *solving the problem above.*

Also in this case we obtain some necessary conditions of optimality. In particular, if Ω is sufficiently large, the optimal structure of the city consists of a finite number of disjoint subcities: circular residential areas with a service pole at the center.

Problem 4 - Optimal Riemannian Metrics

Here the domain Ω and the probabilities f^+ and f^- are given, whereas the distance d is supposed to be conformally flat, that is generated by a coefficient $a(x)$ through the formula

$$d_a(x,y) = \inf\Big\{ \int_0^1 a\big(\gamma(t)\big)|\gamma'(t)|\,dt \ : \\ \gamma \in \mathrm{Lip}(]0,1[;\Omega),\ \gamma(0)=x,\ \gamma(1)=y\Big\}\,.$$

We can then consider the cost functional

$$F(a) = MK(f^+, f^-, d_a).$$

The goal is to prevent as much as possible the transportation of f^+ onto f^- by maximizing the cost $F(a)$ among the admissible coefficients $a(x)$. Of course, increasing $a(x)$ would increase the values of the distance d_a

and so the value of the cost $F(a)$. The class of admissible controls is taken as

$$\mathcal{A} = \Big\{a(x) \text{ Borel measurable } \quad : \\ \alpha \le a(x) \le \beta, \int_\Omega a(x)\,dx \le m\Big\}.$$

In the case when $f^+ = \delta_x$ and $f^- = \delta_y$ are Dirac masses concentrated on two fixed points $x, y \in \Omega$, the problem of maximizing $F(a)$ is nothing else than that of proving the existence of a conformally flat Euclidean metric whose geodesics joining x and y are as long as possible.

This problem has several natural motivations; indeed, in many concrete examples, one can be interested in making as difficult as possible the communication between some masses f^+ and f^-. For instance, it is easy to imagine that this situation may arise in economics, or in medicine, or simply in traffic planning, each time the connection between two *"enemies"* is undesired. Of course, the problem is made non trivial by the integral constraint $\int_\Omega a(x)\,dx \leq m$, which has a physical meaning: it prescribes the quantity of material at one's disposal to solve the problem.

The analogous problem of minimizing the cost functional $F(a)$ over the class $\mathcal{A}$, which corresponds to favor the transportation of f^+ into f^-, is trivial, since

$$\inf\Big\{F(a)\ :\ a\in\mathcal{A}\Big\}=F(\alpha).$$

The existence of a solution for the maximization problem

$$\max\Big\{F(a)\ :\ a\in\mathcal{A}\Big\}$$

is a delicate matter. Indeed, maximizing sequences $\{a_n\} \subset \mathcal{A}$ could develop an oscillatory behavior producing only a relaxed solution. This phenomenon is well known; basically what happens is that the class $\mathcal{A}$ is not closed with respect to the natural convergence

$$a_n \to a \qquad \Longleftrightarrow \qquad d_{a_n} \to d_a \text{ uniformly}$$

and actually it can be proved that $\mathcal{A}$ is dense in the class of all geodesic distances (in particular, in all the Riemannian ones).

Nevertheless, we were able to prove the following existence result.

Theorem *The maximization problem above admits a solution in $a_{opt} \in \mathcal{A}$.*

Several questions remain open:

- Under which conditions is the optimal solution unique?

- Is the optimal solution of bang-bang type? In other words do we have $a_{opt} \in \{\alpha, \beta\}$ or intermediate values (homogenization) are more performant?

- Can we characterize explicitely the optimal coefficient a_{opt} in the case $f^+ = \delta_x$ and $f^- = \delta_y$?

Some Numerical Computations

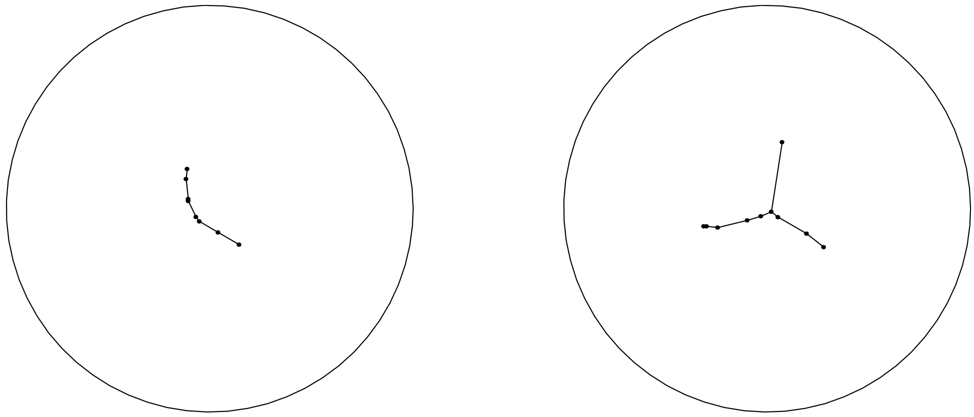
Here are some numerical computations performed (Buttazzo-Oudet-Stepanov 2002) in the simpler case of the so-called problem of optimal irrigation.

This is the optimal network Problem 1 in the case $f^- \equiv 0$, where customers only want to minimize the averaged distance from the network.

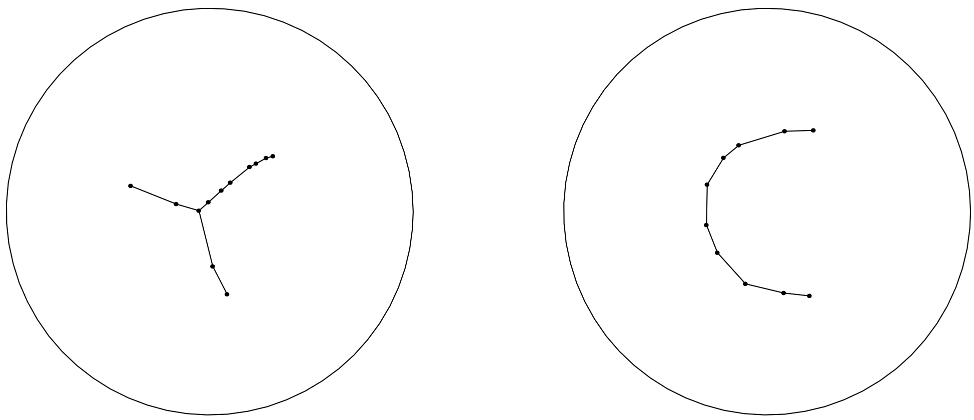
In other words, the optimization criterion becomes simply

$$F(\Sigma) = \int_\Omega \mathrm{dist}(x, \Sigma)\, df^+(x) \, .$$

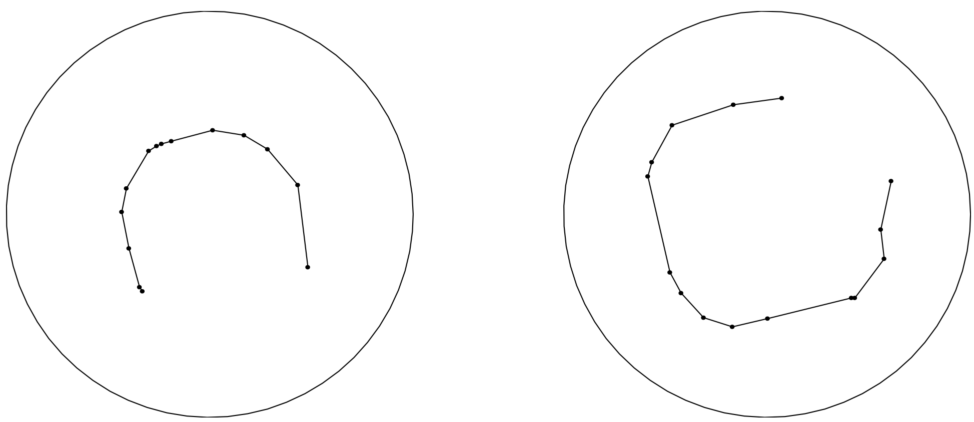
Due to the presence of many local minima the method is based on a genetic algorithm.



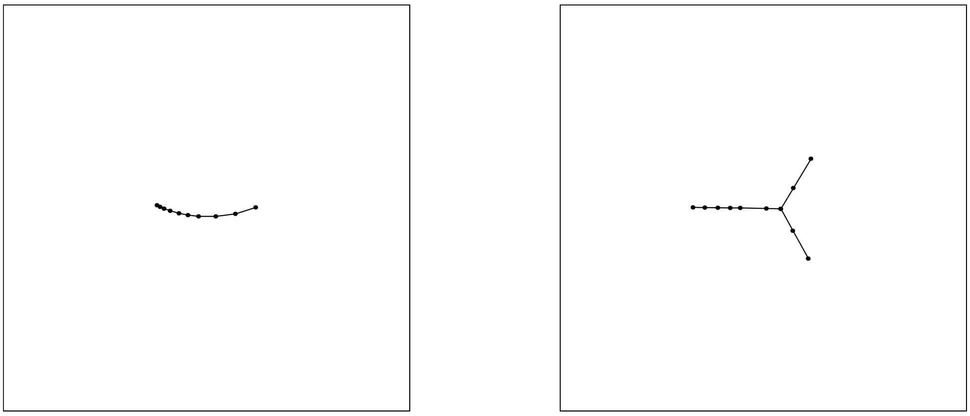
Optimal sets of length 0.5 and 1 in a unit disk



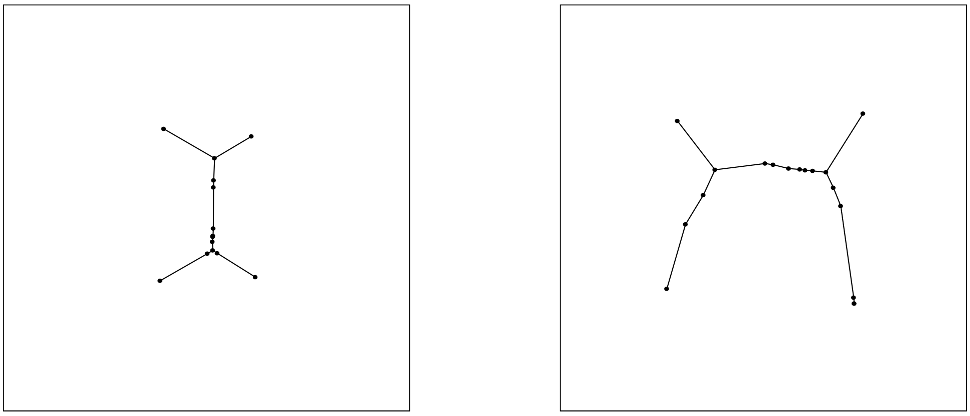
Optimal sets of length 1.25 and 1.5 in a unit disk



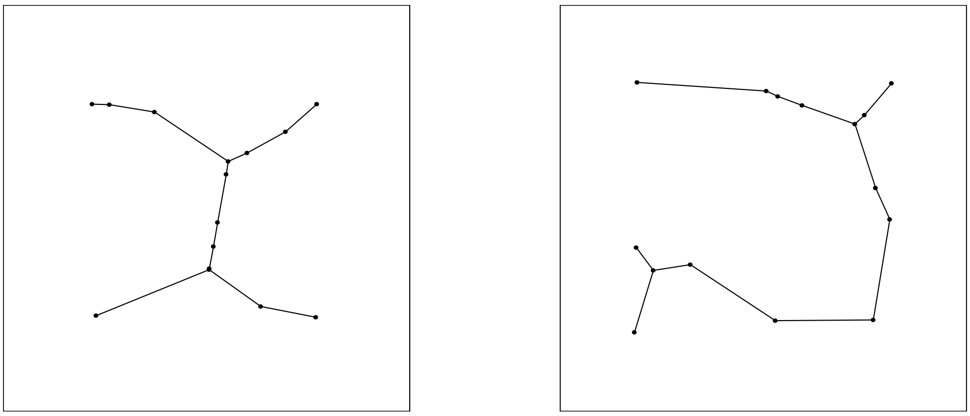
Optimal sets of length 2 and 3 in a unit disk



Optimal sets of length 0.5 and 1 in a unit square

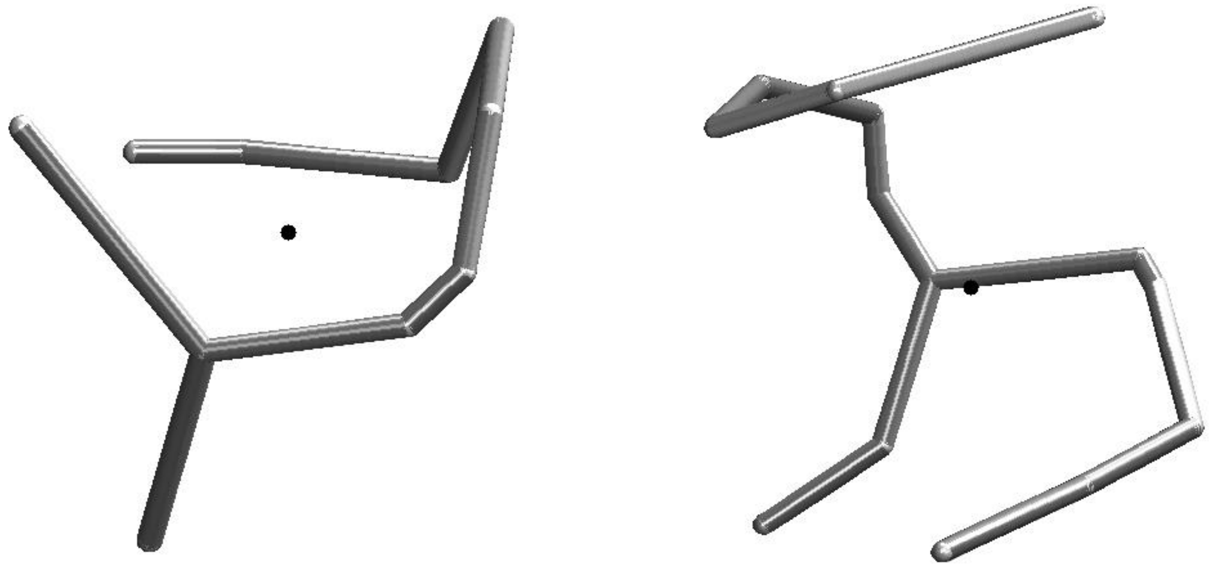


Optimal sets of length 1.5 and 2.5 in a unit square



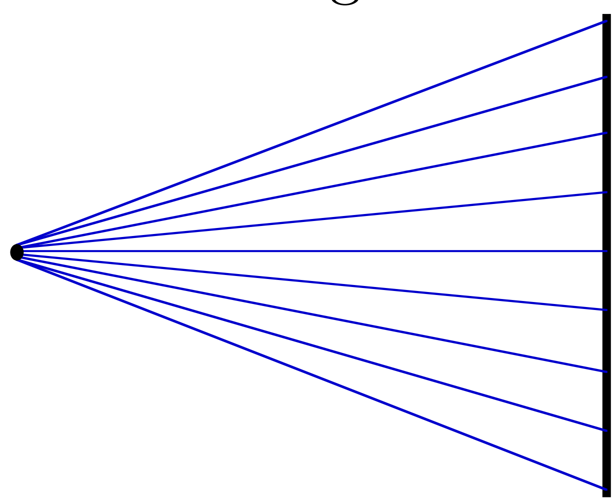
Optimal sets of length 3 and 4 in a unit square

Optimal sets of length 1 and 2 in the unit ball of $\mathbf{R}^3$



Optimal sets of length 3 and 4 in the unit ball of $\mathbf{R}^3$

Now we give a model that takes into account what occurs in several natural structures (see figures above). Indeed, they present some interesting features that should be interpreted in terms of mass transportation. However the usual Monge-Kantorovich theory does not provide an explaination to the reasons for which the structures above exist, and it cannot be taken as a mathematical model for them. For instance, if the source is a Dirac mass and the target is a segment, as in figure below, the Monge-Kantorovich theory provides a wrong behaviour.



The Monge transport rays.

Various approaches have been proposed to give more appropriated models:

- Q. Xia (Comm. Cont. Math. 2003) through the minimization of a suitable functional defined on currents;
- V. Caselles, J. M. Morel, S. Solimini, ... (Preprint 2003 `http://www.cmla.ens-cachan.fr/Cmla/`, Interfaces and Free Boundaries 2003, PNLDE **51** 2002) through a kind of analogy of fluid flow in thin tubes.
- A. Brancolini, G. Buttazzo, E. Oudet, E. Stepanov, (see `http://cvgmt.sns.it`) through a variational model for irrigation trees.

Here we propose a different approach based on a definition of path length in a Wasserstein space.

We also give a model where the opposite feature occurs: instead of favouring the concentration of transport rays, the variational functional prevents it giving a lower cost to diffused measures.

The mathematical model will be considered in an abstract metric space framework where both behaviours (favouring concentration and diffusion) are included; we do not know of natural phenomena where the diffusive behaviour occurs.

The results presented here are contained in

A. Brancolini, G. Buttazzo, F. Santambrogio: *Path functionals over Wasserstein spaces.* Preprint 2004, available at `http://cvgmt.sns.it`

We consider an abstract framework where a metric space X with distance d is given. We assume that closed bounded subsets of X are compact. Fix two points x_0 and x_1 in X and consider the path functional

$$\mathcal{J}(\gamma) = \int_0^1 J(\gamma(t))|\gamma'|(t)\,dt$$

where $\gamma : [0,1] \to X$ ranges among all Lipschitz curves such that $\gamma(0) = x_0$ and $\gamma(1) = x_1$. Here:

- $J : X \to [0,+\infty]$ is a given mapping on X;
- $|\gamma'|(t)$ is the metric derivative of γ at the point t, i.e.

$$|\gamma'|(t) = \limsup_{s\to t} \frac{d\big(\gamma(s),\gamma(t)\big)}{|s-t|}\,.$$

Theorem *Assume:*

- J *is lower semicontinuous in* X*;*
- $J \geq c$ *with* $c > 0$*, or more generally* $\int_0^{+\infty} \big(\inf_{B(r)} J\big)\, dr = +\infty$.

Then, for every $x_0, x_1 \in X$ *there exists an optimal path for the problem*

$$\min \Big\{ \mathcal{J}(\gamma) \ : \ \gamma(0) = x_0, \ \gamma(1) = x_1 \Big\}$$

provided there exists a curve γ_0*, connecting* x_0 *to* x_1*, such that* $\mathcal{J}(\gamma_0) < +\infty$.

The application of the theorem above consists in taking as X a Wasserstein space $\mathcal{W}_p(\Omega)$ where Ω is a compact metric space equipped with a distance function c and a positive finite non-atomic Borel measure m (usually a compact of $\mathbf{R}^N$ with the Euclidean distance and the Lebesgue measure).

We recall that, in the case Ω compact, $\mathcal{W}_p(\Omega)$ is the space of all Borel probability measures μ on Ω equipped with the p-Wasserstein distance

$$w_p(\mu_1, \mu_2) = \inf \Big(\int_{\Omega\times\Omega} c(x,y)^p \, \lambda(dx, dy) \Big)^{1/p}$$

where the infimum is taken on all transport plans λ between μ_1 and μ_2, that is on all probability measures λ on $\Omega \times \Omega$ whose marginals $\pi_1^{\#}\lambda$ and $\pi_2^{\#}\lambda$ coincide with μ_1 and μ_2 respectively.

In order to define the functional

$$\mathcal{J}(\gamma) = \int_0^1 J(\gamma(t))|\gamma'|(t) \, dt$$

it remains to fix the "coefficient" J. We take a l.s.c. functional on the space of measures, of the kind considered by Bouchitté and Buttazzo (Nonlinear Anal. 1990, Ann. IHP 1992, Ann. IHP 1993):

$$J(\mu) = \int_\Omega f(\mu^a)dm + \int_\Omega f^\infty(\mu^c) + \int_\Omega g(\mu(x))d\#$$

where

- $\mu = \mu^a \cdot m + \mu^c + \mu^\#$ is the Lebesgue-Nikodym decomposition of μ with respect to m, into absolutely continuous, Cantor, and atomic parts;
- $f : \mathbf{R} \to [0, +\infty]$ is convex, l.s.c. and proper;
- f^∞ is the recession function of f;
- $g : \mathbf{R} \to [0, +\infty]$ is l.s.c. and subadditive, with $g(0) = 0$;
- $\#$ is the counting measure;
- f and g verify the compatibility condition

$$\lim_{t\to+\infty} \frac{f(ts)}{t} = \lim_{t\to 0^+} \frac{g(ts)}{t}\ .$$

In this way the functional J is l.s.c. for the weak* convergence of measures. If we further assume that $f(s) > 0$ for $s > 0$ and $g(1) > 0$, then we have $J \geq c > 0$. The existence theorem then applies and, given two probabilities μ_0 and μ_1, we obtain at least a minimizing path for the problem

$$\min\left\{\mathcal{J}(\gamma)\ :\ \gamma(0) = \mu_0,\ \gamma(1) = \mu_1\right\}.$$

provided $\mathcal{J}$ is not identically $+\infty$.

We now study two special cases ($\Omega \subset \mathbf{R}^N$ compact and $m = dx$):

Concentration $f \equiv +\infty$, $g(z) = |z|^r$ with $r \in]0,1[$. We have then

$$J(\mu) = \int_\Omega |\mu(x)|^r \, d\# \qquad \mu \text{ atomic}$$

Diffusion $f(z) = |z|^q$ with $q > 1$, $g \equiv +\infty$. We have then

$$J(\mu) = \int_\Omega |u(x)|^q \, dx \qquad \mu = u \cdot dx,\ u \in L^q.$$

Concentration case. The following facts in the concentration case hold:

- If μ_0 and μ_1 are convex combinations (also countable) of Dirac masses, then they can be connected by a path $\gamma(t)$ of finite minimal cost $\mathcal{J}$.

- If $r > 1 - 1/N$ then every pair of probabilities μ_0 and μ_1 can be connected by a path $\gamma(t)$ of finite minimal cost $\mathcal{J}$.

- The bound above is sharp. Indeed, if $r \leq 1 - 1/N$ there are measures that cannot be connected by a finite cost path (for instance a Dirac mass and the Lebesgue measure).

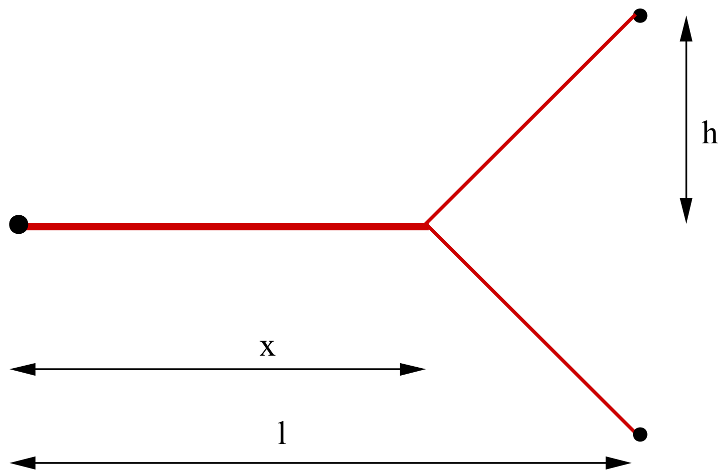
Example. (Y-shape versus V-shape). We want to connect (concentration case $r < 1$ fixed) a Dirac mass to two Dirac masses (of weight $1/2$ each) as in figure below, l and h are fixed. The value of the functional $\mathcal{J}$ is given by

$$\mathcal{J}(\gamma) = x + 2^{1-r}\sqrt{(l-x)^2 + h^2}.$$

Then the minimum is achieved for

$$x = l - \frac{h}{\sqrt{4^{1-r} - 1}}.$$

When $r = 1/2$ we have a Y-shape if $l > h$ and a V-shape if $l \leq h$.



A Y-shaped path for $r = 1/2$.

Diffusion case. The following facts in the diffusion case hold:

- If μ_0 and μ_1 are in $L^q(\Omega)$, then they can be connected by a path $\gamma(t)$ of finite minimal cost $\mathcal{J}$. The proof uses the displacement convexity (McCann 1997) which, for a functional F and every μ_0, μ_1, is the convexity of the map $t \mapsto F(T(t))$, being $T(t) = [(1-t)\mathrm{Id} + tT]^{\#}\mu_0$ and T an optimal transportation between μ_0 and μ_1.
- If $q < 1 + 1/N$ then every pair of probabilities μ_0 and μ_1 can be connected by a path $\gamma(t)$ of finite minimal cost $\mathcal{J}$.
- The bound above is sharp. Indeed, if $q \geq 1 + 1/N$ there are measures that cannot be connected by a finite cost path (for instance a Dirac mass and the Lebesgue measure).

The previous existence results were based on two important assumptions:

- the compactness of Wasserstein spaces $\mathcal{W}_p(\Omega)$ for Ω compact and $1 \leq p < +\infty$;
- the estimate like $F_q \geq c > 0$, that can be obtained when $|\Omega| < +\infty$.

Both the facts do not hold when Ω is unbounded (indeed, the corresponding Wasserstein spaces are not even locally compact). The same happens when Ω is compact but we consider the space $\mathcal{W}_\infty(\Omega)$.
Here is a more refined abstract setting which adapts to the unbounded case.

Notice that in the unbounded case the Wasserstein spaces $\mathcal{W}_p(\Omega)$ do not contain all the probabilities on Ω but only those with finite momentum of order p., that is

$$\int_\Omega |x|^p \, \mu < +\infty \, .$$

Theorem *Let* (X, d, d') *be a metric space with two different distances. Assume that:*

- $d' \leq d$;
- *all* d*-bounded sets in* X *are relatively compact with respect to* d';
- *the mapping* $d : X \times X \to \mathbf{R}^+$ *is a lower semicontinuous function with respect to the distance* $d' \times d'$.

Consider a functional $J : X \to [0, +\infty]$ *and assume that:*

- J *is* d'*-l.s.c.;*
- $\int_0^{+\infty} \left(\inf_{B_d(r)} J\right) dr = +\infty$.

Then the functional

$$\mathcal{J}(\gamma) = \int_0^1 J(\gamma(t))|\gamma'|_d(t)\, dt$$

has a minimizer on the set of $d-$*Lipschitz curves connecting two given points* x_0 *and* x_1*, provided it is not identically* $+\infty$.

As an application of the theorem above we consider the diffusion case when $\Omega = \mathbf{R}^N$. The concentration case in the unbounded setting still presents some extra difficulties that we did not yet solve. We take:

- d the Wasserstein distance;
- d' the distance metrizing the weak* topology on probabilities:

$$d'(\mu, \nu) = \sum_{k=1}^{\infty} \frac{2^{-k}}{1 + c_k} \left| \langle \phi_k, \mu - \nu \rangle \right|$$

  where (ϕ_k) is a dense sequence of Lipschitz functions in the unit ball of $C_b(\Omega)$ and c_k are their Lipschitz constants.

We may prove that the assumptions of the abstract scheme are fulfilled, and we have (for the diffusion case):

- if $q < 1 + 1/N$ for every μ_0 and μ_1 there exists a path giving finite and minimal value to the cost $\mathcal{J}$;
- if $q \geq 1 + 1/N$ there exist measures μ_0 (actually a Dirac mass) such that $\mathcal{J} = +\infty$ on every nonconstant path starting from μ_0.

Open Problems

- linking two L^q measures in the diffusion-unbounded case;
- concentration case in unbounded setting;
- Ω unbounded but not necessarily the whole space;
- working with the space $\mathcal{W}_\infty(\Omega)$;
- comparing this model to the ones by Xia and by Morel, Solimini, ...;
- numerical computations;
- evolution models?

References

(all available at http://cvgmt.sns.it)

A. Brancolini, G. Buttazzo: *Optimal networks for mass transportation problems.* Esaim COCV (to appear).

G. Buttazzo, A. Davini, I. Fragalà, F. Macià: *Optimal Riemannian distances preventing mass transfer.* J. Reine Angew. Math., (to appear).

G. Buttazzo, E. Oudet, E. Stepanov: *Optimal transportation problems with free Dirichlet regions.* In PNLDE **51**, Birkhäuser Verlag, Basel (2002), 41–65.

G. Buttazzo, A. Pratelli, E. Stepanov: *Optimal pricing policies for public transportation networks.* Preprint Dipartimento di Matematica Università di Pisa, Pisa (2004).

G. Buttazzo, F. Santambrogio: *A Model for the Optimal Planning of an Urban Area.* Preprint Dipartimento di Matematica Università di Pisa, Pisa (2003).