An optimization problem for mass transportation with congested dynamics

G. Buttazzo*, C. Jimenez †, E. Oudet ‡.

January 22, 2009

Abstract

Starting from the work by Brenier [10], where a dynamic formulation of mass transportation problems was given, we consider a more general framework, where different kind of cost functions are allowed. This seems relevant in some problems presenting congestion effects as for instance traffic on a highway, crowds moving in domains with obstacles, and in general in all cases where the transportation does not behave as in the classical Monge setting. We show some numerical computations obtained by generalizing to our framework the approximation scheme introduced in Benamou and Brenier [2].

Keywords: transport problems, functionals on measures, congested dynamics, movement of crowds.

MSC 2000: 49A50, 49Q20, 65M60, 60K30.

1 Introduction

Mass transportation theory received much attention in the mathematical community in the last years. Starting from the initial setting by Monge where, given two mass densities ρ_0 and ρ_1, a *transport map* $T : \mathbb{R}^d \to \mathbb{R}^d$

* Università di Pisa, Dipartimento di Matematica, Largo B. Pontecorvo, 5, 56127 Pisa, ITALY `buttazzo@dm.unipi.it`.

† Laboratoire de Mathématiques de Brest, UMR 6205, Université de Brest, 6 avenue le Gorgeu, CS 93837, F-29238 BREST Cedex 3 FRANCE `chloe.jimenez@univ-brest.fr`.

‡ Lama, Université de Savoie, Campus scientifique, 73376 Le Bourget du lac Cedex, FRANCE `edouard.oudet@univ-savoie.fr`.

was searched among the admissible maps transporting ρ_0 onto ρ_1 in order to minimize the total transportation cost

$$\int_{\mathbb{R}^d} |x - T(x)| \, d\rho_0(x) \ ,$$

several other equivalent formulations have been provided (see for instance [23], [18], [5]). In particular, the formulation given in [10] is the one which motivated our study: the goal in [10] was to introduce a "dynamic" formulation of the mass transportation problem providing a map $\rho : [0,1] \to \mathcal{P}(\overline{\Omega})$ which describes the motion of ρ_0 onto ρ_1 as a function of a parameter $t \in [0,1]$, where Ω is the space constraint that all the densities $\rho(t,\cdot)$ have to fulfill.

The set of applications of mass transportation theory is also very rich: many urban planning models have been studied, searching e.g. for the best design of public transportation networks (see [9], [11]), for the optimal pricing policies of their use (see [12]), for the best distribution of residential and working areas in a city (see [13]). We also mention the strict link between mass transportation theory and shape optimization in elasticity, as was shown in [7], [5].

The general framework we consider is the one of functionals defined on the space of measures acting on a time-space domain $\overline{Q} \subset \mathbb{R}^{1+d}$; the minimization problem we are interested in is then written in the form

$$\min \Big\{ \Psi(\sigma) \ : \ -\mathrm{div}\sigma = f \text{ in } \overline{Q}, \ \sigma \cdot \nu = 0 \text{ on } \partial Q \Big\} \tag{1.1}$$

where Ψ is an integral functional on the $\mathbb{R}^{1+d}$-valued measures defined on $\overline{Q}$. Writing $\sigma = (\rho, E)$ the classical Monge case is then related to the cost function

$$\Psi(\sigma) = \int_{\overline{Q}} d|E| \ ,$$

while the case considered by Brenier in [10] is represented by the cost function

$$\Psi(\sigma) = \int_{\overline{Q}} \Big| \frac{dE}{d\rho} \Big|^2 \, d\rho \ .$$

As shown in [10], [2], [21] all these cases are related to the Wasserstein distances $W_p(\rho_0, \rho_1)$, where each particle x in the source ρ_0 moves to its final point $T(x)$ in the target ρ_1 following a line segment, or a geodesic line in case the space constraint Ω is not convex. However, in many problems where a high number of particles (or a probability density) is involved, other effects are present which may deviate the trajectories from straight lines: in particular we are interested in the *congestion* effects that occur when the density $\rho(t,x)$ is high, slowing the ideal mass transportation and increasing the cost.

Modelling the congestion effects has been considered by several authors (see for instance [15], [22]); here we simply consider the Brenier formulation (1.1) assuming that the functional Ψ has a term which has a superlinear growth with respect to ρ.

In Sections 2 and 3 we discuss the general formulation (1.1) and its dual problem, with the primal-dual optimality conditions. In Section 4 we provide a numerical scheme to treat this kind of problems: the scheme is based on the one by Benamou and Brenier [2], adapted to include the congestion terms. In the cases we present the domain Ω is always nonconvex, having some obstacles at its interior, and the mass moves from ρ_0 onto ρ_1 according to:

- the Wasserstein distance W_2, so minimizing the cost $\int_{\overline{Q}} \left|\frac{dE}{d\rho}\right|^2 d\rho$;
- the Wasserstein distance W_2 with the addition of the congestion term $\int_{\overline{Q}} \rho^2 \, dt \, dx$;
- the Wasserstein distance W_2 with the addition of the constraint $\{\rho \leq M\}$ which for instance occurs when a crowd of individuals moves and two different individuals cannot stay too close.

2 The general setting

In this section we consider an open bounded subset Q of $\mathbb{R}^{d+1}$ ($d \geq 1$). We assume Q has a Lipschitz boundary and denote by $\nu(x)$ the outward pointing normal vector to x in the boundary ∂Q of Q, defined almost everywhere. Let $\mathcal{M}_b(\overline{Q}, \mathbb{R}^{d+1})$ be the space of vectorial Borel measures supported on $\overline{Q}$.
We also consider a functional Ψ on $\mathcal{M}_b(\overline{Q}, \mathbb{R}^{d+1})$ and we assume Ψ is lower semicontinuous for the weak* convergence of measures.
Let $f \in \mathcal{M}(\overline{Q})$ be a Borel measure of zero total mass that is $\int_{\overline{Q}} df = 0$. We deal with the following optimization problem:

$$\inf_{\sigma \in \mathcal{M}_b(\overline{Q}, \mathbb{R}^{d+1})} \Psi(\sigma) \tag{2.1}$$

with the constraint:

$$\begin{cases} -\text{div } \sigma &= f \quad \text{in } \overline{Q} \\ \sigma \cdot \nu &= 0 \quad \text{on } \partial Q. \end{cases} \tag{2.2}$$

The condition (2.2) is intended in the weak sense i.e. for every $\varphi \in \mathcal{C}^1(\overline{Q})$:

$$\int_{\overline{Q}} D\varphi \cdot d\sigma(x) = \int_{\overline{Q}} \varphi(y) \, df(y). \tag{2.3}$$

The following general existence result holds:

Theorem 2.1. *Let $\Psi : \mathcal{M}_b(\overline{Q}, \mathbb{R}^{d+1}) \to [0, +\infty]$ be lower semicontinuous for the weak* convergence and such that:*

$$\Psi(\sigma) \geq C|\sigma|(\overline{Q}) - \frac{1}{C} \quad \forall \sigma \in \mathcal{M}_b(\overline{Q}, \mathbb{R}^{d+1}) \tag{2.4}$$

for a suitable constant $C > 0$, where $|\sigma|$ denotes the total variation of σ. We assume that $\Psi(\sigma_0) < +\infty$ for at least one measure σ_0 satisfying (2.2). Then the problem

$$\min \left\{ \Psi(\sigma) : \ -\text{div } \sigma = f \ \textit{in } \overline{Q}, \ \sigma \cdot \nu = 0 \ \textit{on } \partial Q \right\} \tag{2.5}$$

admits a solution. Moreover if Ψ is strictly convex, this solution is unique.

Proof. Let $(\sigma_n)_{n\in\mathbb{N}}$ be a minimizing sequence for problem (2.5). By assumption (2.4), this sequence is bounded and by consequence it admits a subsequence $(\sigma_{n_k})_{k\in\mathbb{N}}$ which converges weakly* to a measure $\sigma \in \mathcal{M}_b(\overline{Q}, \mathbb{R}^{d+1})$. By writing the constraint (2.3) for any σ_{n_k} and passing to the limit as $k \to +\infty$, we get the admissibility of σ. Then, by the lower semicontinuity of Ψ, we get

$$\inf(2.5) = \lim_{k\to+\infty} \Psi(\sigma_{n_k}) \geq \Psi(\sigma)$$

which shows that σ is a solution of (2.5). □

In case Ψ is convex, problem (2.5) also admits a dual formulation. Indeed, if $A : \mathcal{C}(\overline{Q}) \to \mathcal{C}(\overline{Q}, \mathbb{R}^{d+1})$ denotes the operator given by:

$$A(\varphi) = D\varphi \text{ for all } \varphi \text{ in its domain } \mathcal{C}^1(\overline{Q}),$$

we have the convex analysis formula for the dual formulation of (2.5) (see [6]):

$$\begin{aligned} (\Psi^* \circ A)^*(f) &= \min_\sigma \left\{ \Psi(\sigma) : \ -\text{div } \sigma = f \text{ in } \overline{Q}, \ \sigma \cdot \nu = 0 \text{ on } \partial Q \right\} \\ &= \sup \left\{ \int_{\overline{Q}} \varphi(x) \ df(x) - \Psi^*(D\varphi) : \quad \varphi \in \mathcal{C}^1(\overline{Q}) \right\}. \end{aligned} \tag{2.6}$$

This formula holds if Ψ^* is continuous at least at a point of the image of A. For any set C, we denote by χ_C the function which is 0 inside C and $+\infty$ outside. The primal-dual optimality condition then reads as

$$\min \Psi(\sigma) + \chi_{\left\{ \begin{smallmatrix} -\text{div}\sigma = f \text{ in } \overline{Q} \\ \sigma \cdot \nu = 0 \text{ on } \partial Q \end{smallmatrix} \right\}} = \max \int \varphi \ df(x) - \Psi^*(D\varphi)$$

which, if a solution φ_{opt} of (2.6) exists, yields

$$\int D\varphi_{\text{opt}} \cdot d\sigma_{\text{opt}} = \Psi(\sigma_{\text{opt}}) + \Psi^*(D\varphi_{\text{opt}}) \tag{2.7}$$

where σ_{opt} is any solution of (2.5). The point is that, in general, the maximizers φ_{opt} in (2.6) are not in $\mathcal{C}^1(\overline{Q})$. As we will see in the next section, for a large class of cost functions Ψ, (2.6) can be relaxed so that the primal-dual optimality condition will be explicitly identified.

3 The Transportation model

In order to introduce a model for the description of the dynamics of a crowd in a given domain, it is convenient to particularize the framework above as follows:

$Q =$ $]0,1[\times\Omega$ where Ω is a bounded Lipschitz open subset of $\mathbb{R}^d$ with outward normal vector denoted by ν_Ω. The set Ω represents the domain the crowd is constrained to stay inside, including possible obstacles that cannot be crossed. The current variable in Q will be denoted by (t,x) ($t \in]0,1[$, $x \in \Omega$).

$\sigma =$ (ρ, E) where $\rho(t,x)$ represents the mass density at position x and time t and E is the flux at (t,x). In the usual mass transportation cases we have $E \ll \rho$ so that $E = \rho v$ being $v(t,x)$ the velocity field at (t,x). We assume the constraint $\rho \geq 0$ so that the set of admissible variables is:

$$\mathcal{D} := \{(\rho, E) : \ \rho \in \mathcal{M}_b(\overline{Q}, \mathbb{R}^+), E \in \mathcal{M}_b(\overline{Q}, \mathbb{R}^d)\}.$$

$f =$ $\delta_1(t) \otimes \rho_1(x) - \delta_0(t) \otimes \rho_0(x)$ where $\rho_0(x)$, $\rho_1(x)$ represent the crowd densities at $t = 0$ and $t = 1$ respectively, both prescribed as probabilities on $\overline{\Omega}$. Then equation (2.2) reads as:

$$\begin{cases} -\partial_t \rho - \text{div}_x E = 0 \text{ in } \overline{Q} \\ \rho(0,x) = \rho_0(x), \quad \rho(1,x) = \rho_1(x), \\ E \cdot \nu_\Omega = 0 \text{ on }]0,1[\times\partial\Omega \end{cases} \tag{3.1}$$

as it is easy to see using the weak formulation (2.3). Note that (3.1) is the continuity equation of our mass transportation model.

Our problem is then

$$\min\{\Psi(\rho, E) : \ (\rho, E) \text{ verifies } (3.1)\}$$

and we denote by $\mathcal{W}_\Psi(\rho_0, \rho_1)$ its minimal value.
We may deduce from (3.1) that for a.e. $t \in]0,1[$, $\rho(t,\cdot)$ is a probability on $\overline{\Omega}$. Indeed, disintegrating the measure ρ on $\overline{Q}$ we obtain

$$\rho(t,x) = m(t) \otimes \rho^t(x)$$

where m is the marginal of ρ with respect to t and $\rho^t(\cdot)$ is a probability for $m-$a.e. $t \in [0,1]$. Taking in (3.1) a test function $\alpha(t) \in \mathcal{C}^1_c(Q)$ depending only on t we have

$$0 = \int_{\overline{Q}} \alpha'(t)\ d\rho(t,x) = \int_0^1 \alpha'(t)\ dm(t)$$

which gives $m = cdt$ for a suitable constant c. Using the conservation of the mass gives that $c = 1$.

We now discuss the choice of Ψ. We may take for Ψ any local lower semicontinuous function on $\mathcal{M}_b(\overline{Q}, \mathbb{R}^{d+1})$. By the results that can be found in [3] and [4], these functions can be represented in the following form:

$$\Psi(\sigma) = \int_{\overline{Q}} \psi\left(\frac{d\sigma}{dm}\right)\ dm + \int_{\overline{Q}\setminus A_\sigma} \psi^\infty\left(\frac{d\sigma^s}{d|\sigma^s|}\right)\ d|\sigma^s| + \int_{A_\sigma} g(\sigma(x))\ d\sharp(x)$$

where

- m is a positive non-atomic Borel measure on Q;
- $d\sigma/dm$ is the Radon-Nikodym derivative of σ with respect to m;
- $\psi : \mathbb{R}^{d+1} \to [0, +\infty]$ is convex, lower semicontinuous and proper;
- ψ^∞ is the recession function $\psi^\infty(z) := \lim_{t\to+\infty} \frac{\psi(z_0+tz)}{t}$ (the limit is independent of the choice of z_0 in the domain of ψ);
- A_σ is the set of atoms of σ i.e. $A_\sigma := \{x :\ \sigma(x) := \sigma(\{x\}) \neq 0\}$;
- $g : \mathbb{R}^{d+1} \to [0, +\infty]$ is a lower semicontinuous subadditive function such that $g(0) = 0$ and $g_0(z) := \sup_{t>0} \frac{g(tz)}{t} = \psi^\infty(z)$;
- $\sharp$ is the counting measure.

In the sequel we assume the convexity of Ψ i.e. g is asked to be positively 1-homogeneous.

An interesting choice is the one of Benamou and Brenier (see [2], [10]):

$$\psi(r,e) = \begin{cases} \frac{|e|^2}{r} & \text{if } (r,e) \in]0,+\infty[\times\mathbb{R}^d, \\ 0 & \text{if } (r,e) = (0,0), \\ +\infty & \text{otherwise.} \end{cases}$$

This is a positively 1-homogeneous function so $\psi^\infty = \psi = g$ and Ψ does not depend on the choice of the measure m so that

$$\Psi(\rho,E) = \begin{cases} \displaystyle\int_{[0,1]\times\Omega} \psi(d\rho/dm, dE/dm)\ dm(t,x) & \text{if } \rho \geq 0 \\ +\infty & \text{otherwise.} \end{cases}$$

Note that since $\psi(0,e)$ is infinite for any $e \neq 0$, it holds:

$$\Psi(\rho,E) < +\infty \Rightarrow E \ll \rho \tag{3.2}$$

so for any (ρ,E) in the domain of Ψ, we may write:

$E(t,x) = v(t,x)\rho(t,x), \quad$ with $\rho(t,x) \in \mathcal{M}_b(\overline{Q},\mathbb{R}^+)$ and $v(t,x) \in \mathrm{L}^1_\rho(\overline{Q},\mathbb{R}^d)$.

The measure $\rho(t,x)$ can be viewed as the quantity of mass in time and space whereas $v(t,x)$ is the velocity of the mass transiting at x at time t. Moreover Ψ can be written in the simpler form:

$$\Psi(\rho,E) = \begin{cases} \displaystyle\int_{[0,1]\times\overline{\Omega}} \frac{|E|^2}{\rho} := \int_{[0,1]\times\overline{\Omega}} |v|^2\ d\rho(t,x) & \text{if } \rho \geq 0 \text{ and } E = v\rho, \\ +\infty & \text{otherwise.} \end{cases}$$

As shown in [10], in this case we have:

$$\mathcal{W}_\Psi(\rho_0,\rho_1) = (W_2(\rho_0,\rho_1))^2$$

where W_2 is the classical 2-Wasserstein distance (see for instance [24]). Indeed, in the formula above, the Wasserstein distance is intended as:

$$(W_2(\rho_0,\rho_1))^2 = \min\left\{\int_{\overline{\Omega}\times\overline{\Omega}} |x_1-x_2|^2\ d\gamma(x_1,x_2) :\ \gamma \text{ has marginals } \rho_0,\ \rho_1\right\}$$

when Ω is convex, while the Euclidean distance has to be replaced by the geodesic distance when Ω is not convex.

It has been proved in [21] that the same result can be reached with any p-Wasserstein distance ($p > 1$) by choosing the function:

$$\psi_p(r,e) = \begin{cases} \frac{|e|^p}{r^{p-1}} & \text{if } (r,e) \in]0,+\infty[\times\mathbb{R}^d, \\ 0 & \text{if } (r,e) = (0,0), \\ +\infty & \text{otherwise.} \end{cases} \tag{3.3}$$

In the case $p = 1$ we simply take $\psi(r, e) = |e|$.
As in the previous case (3.2) is satisfied, whenever $p \geq 1$, together with

$$\mathcal{W}_\Psi(\rho_0, \rho_1) = (W_p(\rho_0, \rho_1))^p$$

where W_p is the p-th Wasserstein distance:

$$(W_p(\rho_0, \rho_1))^p = \min \left\{ \int_{\overline{\Omega}\times\overline{\Omega}} |x_1 - x_2|^p \, d\gamma(x_1, x_2) : \ \gamma \text{ has marginals } \rho_0, \ \rho_1 \right\}.$$

An important remark is that, in this setting, a solution of problem (1.1) can be built using the idea that *masses should move along straight lines* when Ω is convex and *along geodesic curves* when Ω is not convex. More precisely, if we denote by $\gamma \in \mathcal{M}_b(\overline{\Omega}^2, \mathbb{R}^+)$ an optimal transport plan for W_p and by ξ_{x_1,x_2} a geodesic curve parametrized by $t \in [0, 1]$ joining x_1 to x_2 for γ-almost every (x_1, x_2), then, an optimal $\sigma = (\rho, E)$ is given by:

$$\begin{aligned} \int \varphi \, d\rho &= \int_{\overline{\Omega}^2} \int_0^1 \varphi(t, \xi_{x_1,x_2}(t)) \, dt \, d\gamma(x_1, x_2) \quad \forall \varphi \in \mathcal{C}(\overline{Q}) \\ \int \phi \cdot d\sigma &= \int_{\overline{\Omega}^2} \int_0^1 \phi(t, \xi_{x_1,x_2}(t)) \cdot (1, \dot{\xi}_{x_1,x_2}(t)) \, dt \, d\gamma(x_1, x_2) \ \forall \phi \in \mathcal{C}(\overline{Q})^{d+1}. \end{aligned} \tag{3.4}$$

Indeed, for this choice of σ, the decomposition $E = v\rho$ holds and we have:

$$\int_{\overline{Q}} \frac{|E|^p}{\rho^{p-1}} = \int_{\overline{Q}} |v|^p \, d\rho = \int_{\overline{\Omega}^2} \int_0^1 |\dot{\xi}_{x,y}(t)|^p \, dt \, d\gamma(x, y) = (W_p(\rho_0, \rho_1))^p .$$

Even if this is not the purpose of the paper, we notice that in general the condition $f_0 \ll dx$ does not imply in the case of p-Wasserstein distance (3.3) that the optimal σ is unique as the following example shows.

Example 3.1. Take Ω be the complement of a disc K, $f_0 = dx \llcorner S$ and $f_1 = \frac{1}{2}\delta_A + \frac{1}{2}\delta_B$ as in figure 1; where S is a disc of area 1 and A, B are two points at the same geodesic distance form P. It is clear that all geodesics joining a point of S to either A or B must pass through P. We denote by Γ a line whose points are at the same distance from P, which separates S in two parts S^+ and S^- with the same area. There are infinitely many transport plans γ between f_0 and f_1; in particular γ_1 which sends S^- to A and S^+ to B, and γ_2 which does the opposite. Formula (3.4) provides σ_1 and σ_2 associated to γ_1 and γ_2. Since every particle of S travels with constant speed and since they are at different distances form P, it is easy to see that the corresponding ρ_1 and ρ_2 cannot coincide. For instance there exists a time $\bar{t}$ such that the corresponding ρ_1 loads the geodesic from P to B but not the geodesic from P to A, while at the same time $\bar{t}$, the density ρ_2 does the opposite.

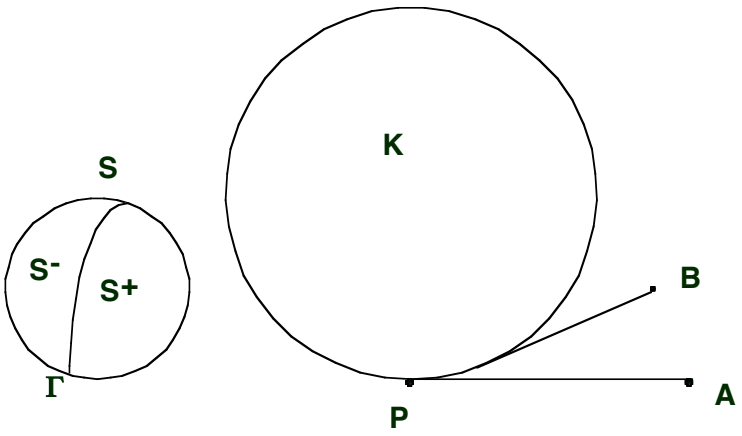

Figure 1: An example of non-uniqueness.

When Ω is convex, $p > 1$ and $f_0 \ll dx$, there is only one optimal transport plan $\bar{\gamma}$ and the unique $\bar{\sigma}$ associated to $\bar{\gamma}$ by use of (3.4) is the only solution of problem (1.1). Let us give a quick scheme of a proof of this uniqueness. Take $\sigma = (\rho, v\rho)$ another solution. Using a result by Ambrosio, Gigli and Savaré (see [1], Theorem 8.2.1.), we can write σ as a superposition of generalized curves. More precisely, it exists some probability measure Γ on the set of absolutely continuous curves $G := W^{1,1}([0,1], \mathbb{R}^N)$ such that:

$$\begin{aligned}
\int \varphi \, d\rho &= \int_G \int_0^1 \varphi(t, \alpha(t)) \, dt \, d\Gamma(\alpha) \quad \forall \varphi \in \mathcal{C}(\overline{Q}) \\
\int \phi \cdot d\sigma &= \int_G \int_0^1 \phi(t, \alpha(t)) \cdot (1, \dot{\alpha}(t)) \, dt \, d\Gamma(\alpha) \; \forall \phi \in \mathcal{C}(\overline{Q})^{N+1} \\
v(t, \alpha(t)) &= \dot{\alpha}(t) \text{ a.e. } t \in [0,1].
\end{aligned}$$

The measure Γ is associated to a transport plan π by the following formula:

$$\int_{\overline{\Omega}^2} \varphi(x,y) d\pi(x,y) := \int_G \varphi(\alpha(0), \alpha(1)) d\Gamma(\alpha).$$

Now, it can easily be seen that the optimality of σ implies the optimality of π and that, Γ-almost everywhere, $\alpha([0,1])$ is the straight line $[\alpha(0), \alpha(1)]$. By uniqueness of the optimal transport plan we have $\pi = \bar{\gamma}$ which yields that σ and $\overline{\sigma}$ coincides.

This does not remain true for $p = 1$. In this case the uniqueness of the optimal transport plan is not insured. Moreover, not all the solutions of

problem (1.1) are of the type (3.4). Actually the following measure (teletransport) happens to be optimal too:

$$\rho(t,x) = (1-t)\rho_0 + t\rho_1,$$

$$\int \Phi\cdot dE = \int_0^1 \int_{\overline{\Omega}^2} \int_{[x_0,x_1]} \Phi(t,x)\cdot \frac{x_1 - x_0}{|x_1 - x_0|}\, d\mathcal{L}^1(x) d\gamma(x_0,x_1) dt,\ \forall \Phi \in \mathcal{C}(\overline{\Omega})^{d+1}.$$

However, the choice of Benamou and Brenier does not take into account congestion effects which are crucial in problems of crowd dynamics. Indeed there is a wide choice (see also [17]) for the cost function Ψ, the congestion effect being due to the superlinear terms. For instance the following are prototypical examples:

- $\psi(r,e) = \frac{|e|^p}{pr^{p-1}} + kr^2$ $(k > 0)$ which gives the cost

  $$\Psi(\rho, E) = \int_0^1 \int_\Omega \left[\frac{|E|^p}{p\rho^{p-1}} + k\rho^2 \right] dt\, dx$$

  intending that $\Psi(\rho, E) = +\infty$ if ρ is not absolutely continuous with respect to $dt \otimes dx$ or ρ is not positive. In this case the high concentrations of ρ are penalized providing a lower congestion during the mass transportation from ρ_0 to ρ_1. Note that, in this case, as ψ is strictly convex in r, the optimal ρ is unique without any other assumption. If, in addition, we have $p > 1$, then E is of the form $E = v\rho$ and, the functional being also strictly convex in v, we have the uniqueness of the optimal measure (E, ρ).

- $\psi(r,e) = \frac{|e|^p}{pr^{p-1}} + \chi_{\{0 \le r \le M\}}(r)$ which gives the cost:

  $$\Psi(\rho, E) = \begin{cases} \displaystyle\int_0^1 \int_\Omega \frac{|E|^p}{p\rho^{p-1}}\, dt\, dx & \text{if } 0 \le \rho \le M, \\ +\infty & \text{otherwise,} \end{cases}$$

  In this case the density ρ is constrained to remain below M, which is for instance the case when the model takes into account that two different individuals of the crowd cannot get too close.

We now study the dual problem (2.6) for general functionals $\Psi(\sigma)$ of the form above.

For the computation of Ψ^* we use a result by Bouchitté and Valadier (Theorem 1 of [8]) on the interchange between sup and integral; we get:

$$\Psi^*(\phi) = \begin{cases} \displaystyle\int_{\overline{Q}} \psi^*(\phi)\ dm & \text{if } (\psi^\infty)^*(\phi(t,x)) + g^*(\phi(t,x)) = 0 \quad \forall (t,x) \in \overline{Q}, \\ +\infty & \text{otherwise,} \end{cases}$$

for all $\phi \in \mathcal{C}(\overline{Q}, \mathbb{R}^{d+1})$, so that (2.6) writes as:

$$\sup_{\varphi \in \mathcal{C}^1(\overline{Q})} \left\{ \int \varphi\ df - \int_{\overline{Q}} \psi^*(D\varphi)\ dm : \ (\psi^\infty)^*(D\varphi) + g^*(D\varphi) = 0 \right\}. \tag{3.5}$$

Note that, as g and Ψ^∞ are positively 1-homogeneous, the constraint of (3.5) can be reformulated saying that $D\varphi(t,x)$ belongs to a convex set K:

$$K := \{u \in \mathbb{R}^{d+1} : \ u \cdot z \leq \min(g(z), \psi^\infty(z))\ \forall z \in \mathbb{R}^{d+1} \text{ such that } |z| = 1\}.$$

As we have already said, Problem (3.5) has to be relaxed in order to make the primal-dual optimality condition meaningful.

To that aim, we need to choose an appropriate space for the dual variable φ and give a sense to the gradient $D\varphi$ appearing in (3.5) and in (2.7) which will write as:

$$\begin{aligned} \int D\varphi_{\text{opt}} \cdot d\sigma_{\text{opt}} \ &= \ \int_{\overline{Q}} \psi\left(\frac{d\sigma_{\text{opt}}}{dm}\right) dm + \int_{\overline{Q}\setminus A_\sigma} \psi^\infty\left(\frac{\sigma^s_{\text{opt}}}{|\sigma^s_{\text{opt}}|}\right)\ d|\sigma^s_{\text{opt}}| \\ &\quad + \int_{A_\sigma} g(\sigma_{\text{opt}}(x))\ d\sharp(x) + \int_{\overline{Q}} \psi^*(D\varphi_{\text{opt}}) dm \\ &\text{with the constraint } \ (\psi^\infty)^*(D\varphi_{\text{opt}}) + g^*(D\varphi_{\text{opt}}) = 0. \end{aligned} \tag{3.6}$$

The space X of the dual variables φ and its topology must be chosen according to the properties of ψ. Then the idea will be to approach φ by a sequence of regular functions $(\varphi_n)_n$ tending to φ. The problem is that the vectorial function η obtained as the limit – in a weak sense – of the sequence $(D\varphi_n)_n$ is not unique in the sense that it depends on the choice of the sequence $(\varphi_n)_n$. Uniqueness can be recovered by making locally the projection of $D\psi_n(t,x)$ on an appropriate tangent space to a measure μ at (t,x) (see [6] and [7]) which has to be chosen in a proper way. In the following, we give some references for some particular cases.

- In [6] a relaxation result is given in case ψ satisfies the following assumption for some $p \in]1, +\infty[$:

$$c_1|(r,e)|^p - \frac{1}{c_1} \leq \psi(r,e) \leq c_2(|(r,e)|^p + 1) \quad \forall (r,e) \in \mathbb{R}^{d+1}$$

for suitable $c_1,\ c_2 > 0$. Therefore for a fixed measure $m \in \mathcal{M}_b(\overline{Q}, \mathbb{R}^+)$ the functionals Ψ and Ψ^* are:

$$\Psi(\sigma) = \begin{cases} \displaystyle\int_{\overline{Q}} \psi\left(\frac{d\sigma}{dm}\right) dm & \text{if } \sigma \ll m \\ +\infty & \text{otherwise,} \end{cases}$$

$$\Psi^*(\phi) = \int_{\overline{Q}} \psi^*(\phi)\ dm,$$

where in the definition of Ψ, we have taken $g \equiv +\infty$. The dual variable φ then belongs to the Sobolev space $W^{1,p'}_m(\overline{Q})$ with $1/p + 1/p' = 1$ made with respect to the measure m (see [6]). Following [6] and [7], the gradient $D_m\varphi(t,x)$ has to be intended as an element of the tangent space $T^{p'}_m(t,x)$ for m-almost every (t,x). Then, as shown in [6], the relaxed dual problem can be expressed as:

$$\sup_{\varphi \in W^{1,p'}_m(Q)} \left\{ \int \varphi df - \int_{\overline{Q}} \psi^*_m(D_m\varphi)\ dm \right\}$$

where

$$\psi^*_m(r,e) = \inf\{\psi^*(r, e+\eta) :\ \eta \in (T^{p'}_m(r,e))^\perp\}.$$

Finally the primal-dual optimality condition reads as:

$$\begin{cases} \displaystyle\int D_m\varphi_{\text{opt}} \cdot d\sigma_{\text{opt}} = \int_{\overline{Q}} \psi\left(\frac{d\sigma_{\text{opt}}}{dm}\right) dm + \int_{\overline{Q}} \psi^*_m(D_m\varphi_{\text{opt}}) dm \\ \sigma_{\text{opt}} \ll m. \end{cases}$$

- In case $\psi(r,e) = \frac{|e|^p}{pr^{p-1}}$ with $p \geq 1$ (see [21]), the functional Ψ^* becomes:

$$\Psi^*(\phi) = \begin{cases} 0 & \text{if } \phi_1 + \frac{|(\phi_2,\dots\phi_{d+1})|^{p'}}{p'} \leq 0 \text{ a.e.} \\ +\infty & \text{otherwise} \end{cases}$$

where p' is such that $1/p + 1/p' = 1$. For $p = 1$, $\frac{|(\phi_2,\dots\phi_{d+1})|^{p'}}{p'}$ has to be intended as $\chi_{\{|(\phi_2,\dots\phi_{d+1})| \leq 1\}}$. The dual variable then is Lipschitz continuous and the relaxed dual problem becomes:

$$\sup_{\varphi \text{ Lipschitz}} \left\{ \int \varphi\ df :\ \partial_t\varphi(t,x) + \frac{|\nabla_x\varphi(t,x)|^{p'}}{p'} \leq 0 \quad \text{a.e. } (t,x) \right\}.$$

In case $p > 1$, the primal-dual optimality condition can be written as:

$$\begin{cases} \displaystyle\int D_{\rho_{\rm opt}}\varphi_{\rm opt}\cdot(1,v_{\rm opt}(t,x))d\rho_{\rm opt} = \int_{\overline{Q}} \frac{|v_{\rm opt}(t,x)|^p}{p}\, d\rho_{\rm opt}(t,x), \\ \partial_t\varphi_{\rm opt}(t,x) + \frac{|\nabla_x\varphi_{\rm opt}(t,x)|^{p'}}{p'} \leq 0 \quad \text{a.e. } (t,x), \end{cases} \tag{3.7}$$

where the gradient $D_{\rho_{\rm opt}}\varphi(t,x) = (\partial_{(\rho_{\rm opt},t)}\varphi(t,x), \nabla_{(\rho_{\rm opt},x)}\varphi(t,x))$ is an element of the tangent space $T^{\infty}_{\rho_{\rm opt}}(t,x)$ for $\rho_{\rm opt}$-almost every (t,x). As it can be seen in [21], we have

$$D_{\rho_{\rm opt}}\varphi_{\rm opt}(t,x) - D\varphi_{\rm opt}(t,x) \in T^{\perp}_{\rho_{\rm opt}}(t,x) \quad \rho_{\rm opt} - \text{a.e.}$$

and thanks to (3.1):

$$(1, v_{\rm opt}(t,x)) \in T_{\rho_{\rm opt}}(t,x) \quad \rho_{\rm opt} - \text{a.e.}$$

so that the inequality (3.7) gives:

$$\begin{aligned} D_{\rho_{\rm opt}}\varphi_{\rm opt}\cdot(1,v_{\rm opt}) &= D\varphi_{\rm opt}\cdot(1,v_{\rm opt}) \\ &\leq -\frac{|\nabla_{(\rho_{\rm opt},x)}\varphi_{\rm opt}|^{p'}}{p'} + v_{\rm opt}(t,x)\cdot\nabla_{(\rho_{\rm opt},x)}\varphi_{\rm opt} \\ &\leq \sup_{\omega\in\mathbb{R}^d}\left\{v_{\rm opt}(t,x)\cdot\omega - \frac{|\omega|^{p'}}{p'}\right\} = \frac{|v_{\rm opt}|^p}{p}. \end{aligned}$$

Then, the equality in (3.7) gives that all the previous inequalities happen to be equalities that is to say

$$\begin{aligned} &\partial_{(\rho_{\rm opt},t)}\varphi_{\rm opt} = -\frac{|\nabla_{(\rho_{\rm opt},x)}\varphi_{\rm opt}|^{p'}}{p'}, \\ &\nabla_{(\rho_{\rm opt},x)}\varphi_{\rm opt}(x,t) \in \operatorname{argmax}\left\{\omega \mapsto v_{\rm opt}(t,x)\cdot\omega - \frac{|\omega|^{p'}}{p'}\right\}. \end{aligned} \tag{3.8}$$

By making an easy computation, we get:

$$\begin{aligned} &\nabla_{(\rho_{\rm opt},x)}\varphi_{\rm opt} = |v_{\rm opt}|^{p-2}v_{\rm opt}, \\ &\partial_{(\rho_{\rm opt},t)}\varphi_{\rm opt} = \frac{-|v_{\rm opt}|^p}{p'}. \end{aligned} \tag{3.9}$$

If $p = 1$, we make the computation in the similar way by writing $(\rho_{\rm opt}, E_{\rm opt})$ as

$$(\rho_{\rm opt}(t,x), E_{\rm opt}(t,x)) = (h_{\rm opt}(t,x), v_{\rm opt}(t,x))d\mu_{\rm opt}$$

where $\mu_{\text{opt}} \in \mathcal{M}_b(\overline{Q}, \mathbb{R}^+)$ and $(h_{\text{opt}}, v_{\text{opt}}) \in \mathrm{L}^1_{\mu_{\text{opt}}}(\overline{Q}) \times \mathrm{L}^1_{\mu_{\text{opt}}}(\overline{Q}, \mathbb{R}^d)$. Then the primal-dual optimality condition writes as:

$$\begin{cases} \displaystyle\int D_{\mu_{\text{opt}}}\varphi_{\text{opt}} \cdot (h_{\text{opt}}(t,x), v_{\text{opt}}(t,x)) d\mu_{\text{opt}} = \int_{\overline{Q}} |v_{\text{opt}}(t,x)| \; d\mu_{\text{opt}}(t,x), \\ \partial_t \varphi_{\text{opt}}(t,x) \leq 0 \text{ and } |\nabla_x \varphi_{\text{opt}}(t,x)| \leq 1 \quad \text{a.e. } (t,x), \end{cases} \tag{3.10}$$

which leads to:

$$\begin{aligned} &\nabla_{(\rho_{\text{opt}},x)}\varphi_{\text{opt}} = \frac{v_{\text{opt}}}{|v_{\text{opt}}|}, \\ &\partial_{(\rho_{\text{opt}},t)}\varphi_{\text{opt}} = 0. \end{aligned} \tag{3.11}$$

4 Numerical computation

We describe in the present section an algorithm to approximate problem (1.1). This method is directly adapted from the augmented Lagrangian method presented in [2]. For the reader convenience, we recall below in our formalism the main steps of this algorithm.

First, solving problem (1.1) is equivalent to solve the saddle point problem:

$$\min_{\sigma} \max_{\varphi \in \mathcal{C}(Q)} L(\sigma, \varphi) \tag{4.1}$$

where $L(\sigma, \varphi)$ is the Lagrangian defined by:

$$L(\sigma, \varphi) = \Psi(\sigma) - \int D\varphi \cdot d\sigma + \int \varphi \; df.$$

Following [2], for all $r > 0$, we introduce the augmented Lagrangian

$$L_r(\sigma, \sigma^*, \varphi) := \Psi^*(\sigma^*) + \int (D\varphi - \sigma^*) \cdot d\sigma - \int \varphi \; df + \frac{r}{2} \int |D\varphi - \sigma^*|^2 dy.$$

Using the identity $\Psi^*(\sigma^*) + \Psi(\sigma) = \int \sigma^* \cdot d\sigma$ it can easily be established that the saddle point problem (4.1) is equivalent to the new problem:

$$\max_{\sigma} \min_{\sigma^*, \varphi} L_r(\sigma, \sigma^*, \varphi). \tag{4.2}$$

As reported in [2], the simple algorithm ALG2 (see [19]), which is a classical relaxation of Uzawa's method, can be used to approximate problem (4.2). Let us recall with our notation this iterative process:

- let $(\sigma_n, \sigma^*_{n-1}, \varphi_{n-1})$ be given;

- Step A: find φ_n such that:

$$L_r(\sigma_n, \sigma^*_{n-1}, \varphi_n) \leq L_r(\sigma_n, \sigma^*_{n-1}, \varphi), \quad \forall \varphi \in \mathcal{C}^1(\overline{Q});$$

- Step B: find σ^*_n such that:

$$L_r(\sigma_n, \sigma^*_n, \varphi_n) \leq L_r(\sigma_n, \sigma^*, \varphi_n), \quad \forall \sigma^* \in \mathcal{C}(\overline{Q}, \mathbb{R}^{d+1});$$

- Step C: set $\sigma_{n+1} = \sigma_n + r(D\varphi_n - \sigma^*_n)$;
- go back to Step A.

Note that the variables (ν, q) in [2] are renamed (σ, σ^*) in the previous description of the algorithm. Let us now underline the two main differences of our approach.
First, Step A consists in solving the Euler-Lagrange equation:

$$\int D\varphi \cdot d\sigma_n - \int \varphi \, df + r \int D\varphi(-\sigma^*_{n-1} + D\varphi_n) \, dy = 0, \qquad \forall \varphi \in \mathcal{C}^1(\overline{Q}).$$

This variational formulation is nothing else than the weak form of the partial differential equation:

$$\begin{cases} -r\Delta\varphi_n = \operatorname{div}(\sigma_n - r\sigma^*_{n-1}) + f & \text{in } \overline{Q} \\ r\frac{\partial \varphi_n}{\partial n} = (\sigma_n - r\sigma^*) \cdot \nu & \text{on } \partial Q. \end{cases}$$

The resolution of the previous PDE has been achieved with the very efficient software freeFEM3D (see [20] and [16]) provided by S. Del Pino and O. Pironneau. As in [2], for computational stability, we perturbed the previous Laplace equation in:

$$-r\Delta\varphi_n + r\varepsilon\varphi_n = \operatorname{div}(\sigma_n - r\sigma^*_{n-1}) + f$$

with $\varepsilon = 10^{-4}$.
Second, since in our general framework, Ψ^* is not always a characteristic function, step B consists in minimizing the following quantity with respect to σ^*:

$$\Psi^*(\sigma^*) + \int (D\varphi_n - \sigma^*) \cdot d\sigma_n + \frac{r}{2} \int |D\varphi_n - \sigma^*|^2 dy.$$

In all the test cases presented below, it has been possible to solve this problem analytically. Indeed, this pointwise optimization problem reduces to the numerical computation of the roots of a polynomial with real coefficients.

Example 4.1. We consider here a transportation domain $\Omega = [-1,1]^2$ in which there are spatial obstacles that the mass cannot cross. This is for instance the case of a subway gate that a mass of individuals has to cross to reach a final destination. In this first example, the transportation is described simply by the Wasserstein distance W_2 whichturns out, setting $\sigma = (\rho, E)$, to consider the convex function

$$\Psi(\sigma) = \int_Q \frac{|E|^2}{2\rho}$$

in the sense precised in Section 3. The Fenchel transform Ψ^* can be easily computed and we have:

$$\Psi^*(\Phi) = \begin{cases} 0 & \text{if } \Phi_1 + \frac{|(\Phi_2,\Phi_3)|^2}{2} \leq 0 \text{ a.e.} \\ +\infty & \text{otherwise.} \end{cases}$$

Notice that, since Ψ is homogeneous of degree 1, the function Ψ^* is the indicator of a convex set. Here below, we plot the mass density ρ_t at various instants of time. The initial configuration ρ_0 is taken as a Gaussian distribution centered at the point $(-0.65, 0)$ and the final measure ρ_1 is taken as $\rho_1(x_1, x_2) = \rho_0(x_1 - 1.3, x_2)$. Notice that without the obstacle gate, the mass density $\rho(t, \cdot)$ would simply be the translation $\rho(t, x_1, x_2) = \rho_0(x_1 - 1.3t, x_2)$. In general, in presence of obstacles, the mass density ρ will follow the geodesic paths and by consequence the supports of all $\rho(t, \cdot)$ have to be contained in the geodesic envelope of ρ_0 and ρ_1; this is why most of the mass passes through the central gate. Our computation done on a regular grid of $70 \times 70 \times 70$ (from which cells corresponding to the obstacles have been removed) and presented in Figure 2 is in agreement with that observation. Convergence with respect to the criterium proposed in [2] has been achieved in 150 iterations.

Example 4.2. We consider the same geometrical configuration as in the previous example. In this case, we add a diffusion term in order to penalize mass congestion which is described in our case by high values of ρ. The function Ψ we consider is:

$$\int_0^1 \int_\Omega \frac{|E|^2}{2\rho} + c\rho^2 \; dt \; dx$$

with $c = 0.1$. The Fenchel transform is given by:

$$\Psi^*(\Phi) = \frac{1}{2c} \int_Q \left(\left(\Phi_1 + \frac{|(\Phi_2, \Phi_3)|^2}{2} \right)^+ \right)^2 (y) \, dy.$$

Notice that, due to the addition of the diffusion term, the dual function Ψ^* is now finite everywhere. This fact could explain the improvement in the convergence of the iteration scheme: in that example, convergence is reached in only 50 iterations.
As expected (see Figure 3), the mass crosses the obstacle by using several gates.

Example 4.3. In our last example we consider again the same geometrical configuration and a new term which takes congestion into account. More precisely, we consider the cost functional

$$\int_0^1 \int_\Omega \frac{|E|^2}{2\rho} + \chi_{\rho \leq 1}.$$

The Fenchel transform is given by:

$$\Psi^*(\Phi) = \int_Q \left(\Phi_1 + \frac{|(\Phi_2, \Phi_3)|^2}{2} \right)^+ (y)\, dy.$$

At a first glance (see Figure 4 where level lines are plotted), the result seems to be very similar to our first situation where the congestion effect was not considered. Again, most of the mass passes through the central gate, but contrary to the first case the density in the front gate is spread all over the channel and not only near the boundaries of the obstacles.

Acknowledgements: This work has been done during a visit of Edouard Oudet to Pisa in the framework of a GNAMPA program for scientific visits. Chloé Jimenez gratefully acknowledges the Scuola Normale Superiore di Pisa for the post-doc fellowship provided.

References

[1] L. Ambrosio, N. Gigli, G. Savaré: *Gradient flows in metric spaces and in the space of probability measures.* Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, 2005. viii+333 pp.

[2] J.-D. Benamou, Y. Brenier: *A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem.* Numer. Math., **84** (3) (2000), 375–393.

[3] G. Bouchitté, G. Buttazzo: *New lower semicontinuity results for nonconvex functionals defined on measures.* Nonlinear Anal., **15** (7) (1990), 679–692.

[4] G. Bouchitté, G. Buttazzo: *Integral representation of nonconvex functionals defined on measures.* Ann. Inst. H. Poincaré Anal. Non Linéaire, **9** (1) (1992), 101–117.

[5] G. Bouchitté, G. Buttazzo: *Characterization of optimal shapes and masses through Monge-Kantorovich equation.* J. Eur. Math. Soc., **3** (2) (2001). 139–168.

[6] G. Bouchitté, G. Buttazzo, P. Seppecher: *Energies with respect to a measure and applications to low dimensional structures.* Calc. Var., **5** (1997), 37–54.

[7] G. Bouchitté, G. Buttazzo, P. Seppecher: *Shape optimization solutions via Monge-Kantorovich equation.* C. R. Acad. Sci. Paris Sér. I Math., **324** (10) (1997), 1185–1191.

[8] G. Bouchitté, M. Valadier: *Integral representation of convex functionals on a space of measures.* J. Funct. Anal. 80 (1988), no. 2, 398–420.

[9] A. Brancolini, G. Buttazzo: *Optimal networks for mass transportation problems.* ESAIM Control Optim. Calc. Var., **11** (1) (2005), 88–101.

[10] Y. Brenier: *Extended Monge-Kantorovich theory.* In "Optimal Transportation and Applications" (Martina Franca 2001), Lecture Notes in Math. **1813**, Springer-Verlag, Berlin (2003), 91–121.

[11] G. Buttazzo, A. Pratelli, S. Solimini, E. Stepanov: *Optimal urban networks via mass transportation.* Lecture Notes in Math. **1961**, Springer-Verlag, (in print).

[12] G. Buttazzo, A. Pratelli, E. Stepanov: *Optimal pricing policies for public transportation networks.* SIAM J. Optim., **16** (3) (2006), 826–853.

[13] G. Buttazzo, F. Santambrogio: *A model for the optimal planning of an urban area.* SIAM J. Math. Anal., **37** (2) (2005), 514–530.

[14] G. Carlier, P. Cardaliaguet, C. Jimenez: *Optimal transport with convex constraints.* Work in progress.

[15] G. Carlier, C. Jimenez, F. Santambrogio: *Optimal transportation with traffic congestion and Wardrop equilibria.* SIAM J. Control Optim. 47 (2008), no. 3, 1330–1350.

[16] S. Del Pino, O. Pironneau: *A Fictitious domain based general PDE solver.* In "Numerical Methods for Scientific Computing", conf. METSO-ECCOMAS, E. Heikkola ed., CIMNE, Barcelona (2003).

[17] J. Dolbeault, B. Nazaret, G. Savaré: *A new class of transport distances between measures.* Calc. Var. PDE, **34** (2) (2009), 193-231.

[18] L. C. Evans, W.Gangbo: *Differential equation methods for the Monge-Kantorovich mass transfer problem.* Mem. Amer. Math. Soc., **658** (1999).

[19] M. Fortin, R. Glowinski: *Augmented Lagrangian Methods. Applications to the Numerical Solution of Boundary Value Problems.* Stud. Math. Appl. **15**, North-Holland, Amsterdam (1983).

[20] freeFEM3D: available at `http://www.freefem.org/ff3d`.

[21] C. Jimenez: *Dynamic formulation of optimal transport problems.* Dynamic formulation of optimal transport problems. J. Convex Anal. 15 (2008), no. 3, 593–622.

[22] B. Maury, J. Venel: *Un modèle de mouvements de foule.* Paris-Sud Working Group on Modelling and Scientific Computing 2006–2007, 143–152, ESAIM Proc., 18, EDP Sci., Les Ulis, 2007.

[23] L. V. Kantorovich: *On the transfer of masses.* Dokl. Akad. Nauk. SSSR, **37** (1942), 227–229.

[24] C. Villani: *Topics in Optimal Transportation.* Grad. Stud. Math. **58**, Amer. Math. Soc., Providence (2003).

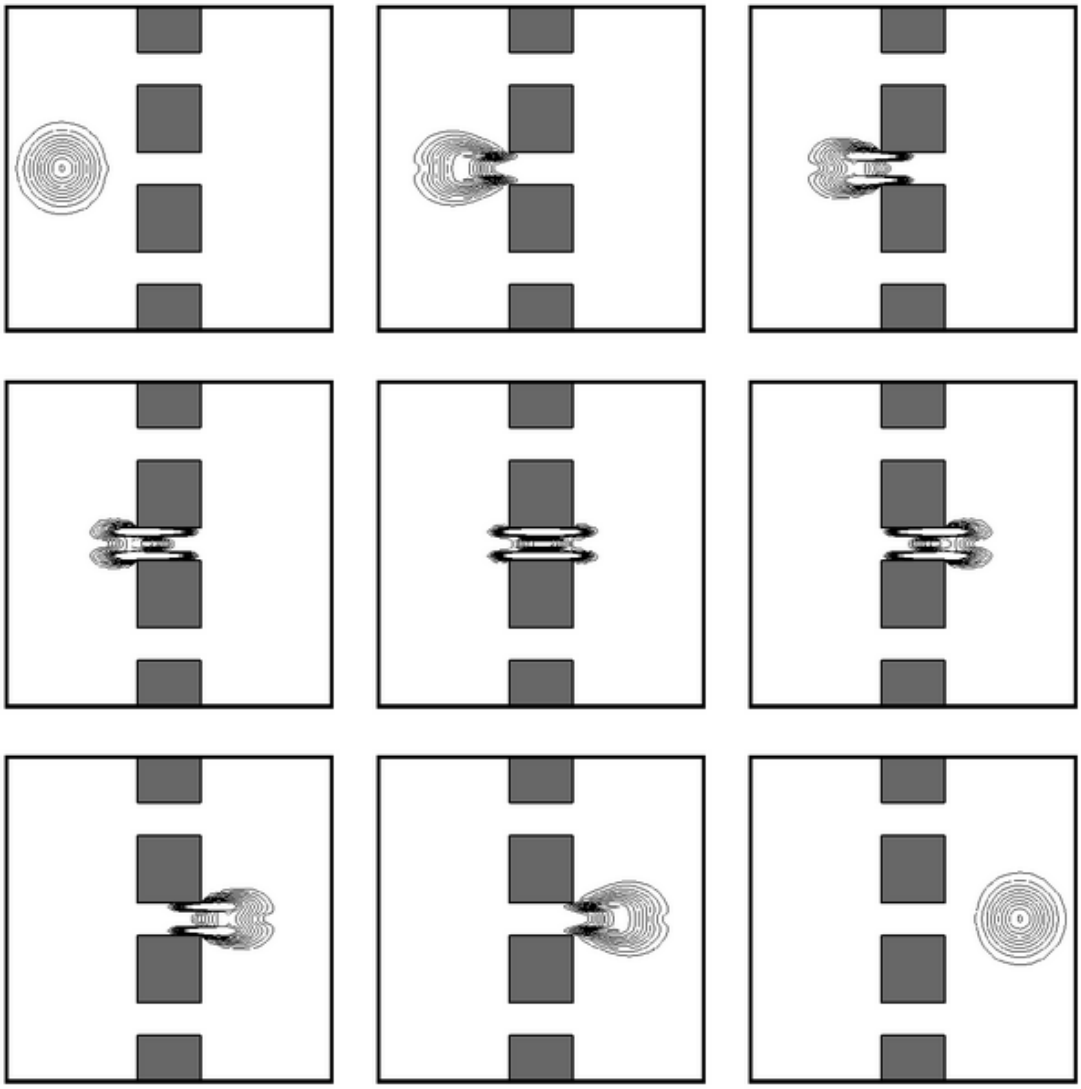

Figure 2: Plot of $\rho(t, \cdot)$ for 9 values of t.

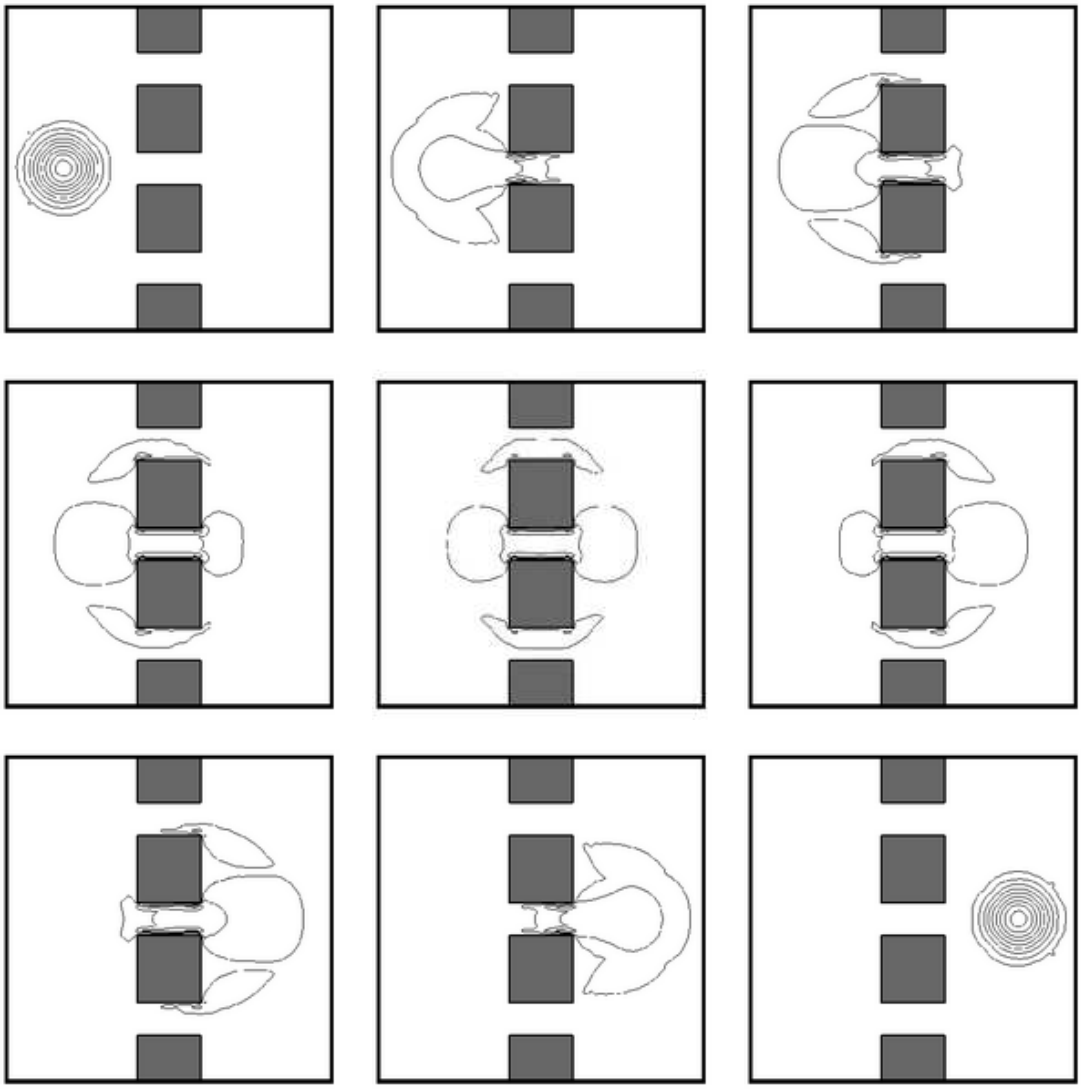

Figure 3: Plot of $\rho(t, \cdot)$ for 9 values of t.

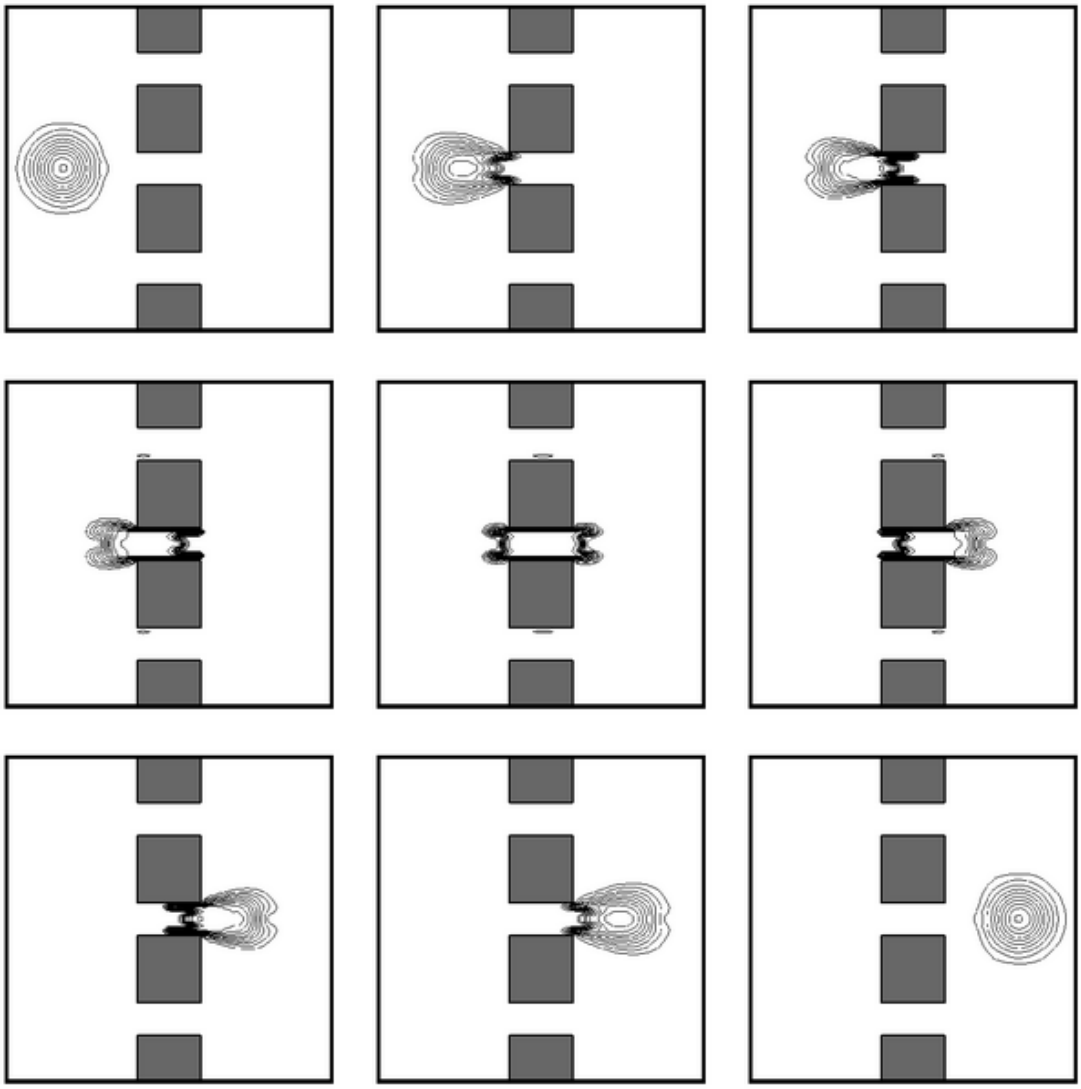

Figure 4: Plot of $\rho(t, \cdot)$ for 9 values of t.