THE OPTIMAL MASS TRANSPORT PROBLEM FOR RELATIVISTIC COSTS

JÉRÔME BERTRAND AND MARJOLAINE PUEL

Abstract. In this paper, we study the optimal mass transportation problem in $\mathbb{R}^d$ for a class of cost functions that we call relativistic cost functions. Consider as a typical example, the cost function $c(x, y) = h(x - y)$ being the restriction of a strictly convex and differentiable function to a ball and infinite outside this ball. We show the existence and uniqueness of the optimal map given a relativistic cost function and two measures with compact support, one of the two being absolutely continuous with respect to the Lebesgue measure. With an additional assumption on the support of the initial measure and for supercritical speed of propagation, we also prove the existence of a Kantorovich potential and study the regularity of this map. Besides these general results, a particular attention is given to a specific cost because of its connections with a relativistic heat equation as pointed out by Brenier in [18].

1. Setting of the problem

In the last ten years, the Monge-Kantorovich problem has been widely investigated. So far, there was no general result that could be compared to the Gangbo-McCann theorem for strictly convex and real-valued cost functions on $\mathbb{R}^d$ [29]. In this paper, we provide the first result of this kind for a large class of strictly convex functions that are infinite outside a bounded convex set. The cost functions we consider are not even continuous maps from $\mathbb{R}^d$ to $[0, +\infty]$. We show that under quite standard assumptions including the compactness of the supports, there exists a unique optimal transport map relative to such a cost function. We refer to Theorem 2.6 for a precise statement. Among all these cost functions, a particular attention is given to a cost function that we call relativistic heat cost function. This relativistic heat cost, introduced by Brenier in [18], is related to a relativistic heat equation as explained below. In [18], Brenier mentions the question of the existence of an optimal map for this cost function, putting the emphasis on the case where only a part of the mass can be transported. This cost function was our initial motivation to study this kind of problems. Let us now describe more precisely the problem we are interested in.

1.1. The general optimal transportation problem.

Let us first recall the Monge problem.

Given two probability measures $\mu_0 \in \mathcal{P}(\mathbb{R}^d)$ and $\mu_1 \in \mathcal{P}(\mathbb{R}^d)$, we consider the maps F, if any, that push μ_0 forward to μ_1, meaning the maps that satisfy $F_\# \mu_0 = \mu_1$. In terms of test functions, the equation $F_\# \mu_0 = \mu_1$ amounts to

$$\int_{\mathbb{R}^d} \varphi(F(x)) d\mu_0(x) = \int_{\mathbb{R}^d} \varphi(y) d\mu_1(y)$$

for any test function $\varphi \in C^0(\mathbb{R}^d)$.

Monge asked whether the infimum

$$\inf_{F_\# \mu_0 = \mu_1} \int_{\mathbb{R}^d} c(x, F(x)) d\mu_0(x)$$

is attained among the maps F such that $F_\# \mu_0 = \mu_1$. The function c above is called the cost function. Note that there might be no map that pushes μ_0 forward to μ_1 and that the Monge problem is nonlinear in general. This leads Kantorovich to consider a relaxation of the Monge problem, now called the Kantorovich problem [32]. In this problem, we consider transport plans instead of transport maps. A transport plan π is a probability measure on $\mathbb{R}^d \times \mathbb{R}^d$ with first marginal μ_0 and second marginal μ_1. This means that $\mu_0(U) = \pi(U \times \mathbb{R}^d)$ and $\mu_1(U) = \pi(\mathbb{R}^d \times U)$ for any Borel set $U \subset \mathbb{R}^d$. Note that a transport map F induces a transport plan $(Id, F)_\# \mu_0$. The set of transport plans between μ_0 and μ_1 is denoted by $\Pi(\mu_0, \mu_1)$.

The Kantorovich problem then consists in studying, given two probability measures μ_0 and μ_1 in $\mathcal{P}(\mathbb{R}^d)$, the optimal transport plans, namely the plans that achieve the minimum

$$\min_{\pi \in \Pi(\mu_0, \mu_1)} \int_{\mathbb{R}^d \times \mathbb{R}^d} c(x, y) d\pi(x, y).$$

Date: 2009 (v0).

Indeed, contrary to the Monge problem, the Kantorovich problem always admits solution as soon as the cost function is a lower semi continuous and non-negative map, see [43, 44] for a proof. We refer to [16, 17, 28, 29, 43, 44] for existence and uniqueness of solution to the Monge-Kantorovich problem in the quadratic (or L^2) case and for strictly convex cost functions and to [4, 27, 20, 7, 6, 24, 25, 21, 42] for the study of the L^1 case.

1.2. **The relativistic heat cost.** The original motivation of this paper comes from what we call the relativistic heat cost given by $c(x, y) = h(x - y)$ where

$$h(z) = \begin{cases} 1 - \sqrt{1 - |z|^2} & |z| \le 1 \\ \infty & |z| > 1. \end{cases} \tag{1}$$

The peculiarity of this cost function is that even if it is strictly convex, it is neither real-valued nor continuous and its gradient goes to infinity when $|z| \to 1$.

The name of this cost function comes from the fact that it prevents the displacement – or the velocity if we take the Lagrangian point of view – to be larger than one which corresponds to a relativistic behavior. This cost has been proposed by Brenier in [18] to obtain a relativistic heat equation as a gradient flow of the Boltzmann entropy (see [3, 5, 43] for references on the notion of gradient flows). Several Cauchy problems for PDE's has been studied via the JKO time discrete scheme [1, 31, 36]... For example, the heat equation can be obtained by computing the time discrete solution at the time step i in the following way

$$\rho^i = \arg\inf_{\rho \in P(\mathbb{R}^d)} \int \rho \log \rho + \varepsilon\, W^2_{2,\varepsilon}(\rho^{i-1}, \rho) \quad \text{with} \quad W^2_{2,\varepsilon}(\rho^{i-1}, \rho) = \min_{\pi \in \Pi(\rho^{i-1}, \rho)} \int \left(\frac{|x - y|}{\varepsilon} \right)^2 d\pi(x, y) \tag{2}$$

and by passing to the limit when the time step ε goes to zero. Analogously, a Cauchy result has been proved by R. McCann and the second author in [34] for the equation

$$\partial_t \rho = \operatorname{div}\Big(\rho \frac{\frac{\nabla \rho}{\rho}}{\sqrt{1 + |\frac{\nabla \rho}{\rho}|^2}}\Big) = \operatorname{div}(\rho \nabla h^*(\nabla(\log \rho))) \tag{3}$$

replacing $W^2_{2,\varepsilon}$ in (2) by

$$W_c = \min_{\pi \in \Pi(\rho^{i-1}, \rho)} \int h(\frac{x - y}{\varepsilon}) d\pi(x, y) \qquad \text{where } c(x, y) = h(x - y) \text{ is the relativistic heat cost.}$$

Equation (3) is called relativistic heat equation since it can be seen as a relativistic version of the heat equation interpreted [35, 38] as a transport of mass equation with velocity $\dfrac{\nabla \rho}{\rho}$. It has previously been investigated via more PDE's oriented method in the series of papers [9, 11, 10, 12, 13, 14, 15, 23]. The proof of the main result of [34] already requires the existence of an optimal map relative to the problem W_c with ρ^{i-1} and ρ^i. This map is obtained by a regularization argument applied to the cost function and based on the Yosida transform (we refer to [19] for the definition). However, the limiting process is strongly based on the double minimization involved in the scheme, in particular the compactness of the sequence of maps comes from the fact that the target measure ρ^i also minimizes the entropy. In this paper, we propose another method to prove the existence and uniqueness of the optimal map that goes beyond the particular setting described above.

The method we found applies to more general relativistic cost functions including the cost given by formula (1) and another important class of cost functions given by

$$c(x, y) = f(x - y) + \delta(x - y|\,\mathcal{C}) \tag{4}$$

where $\delta(\cdot|\,\mathcal{C})$ is the indicator function of a strictly convex body $\mathcal{C}$ containing 0, namely $\delta(x|\,\mathcal{C}) = 0$ if $x \in \mathcal{C}$ and $+\infty$ otherwise and $f : \mathbb{R}^d \to \mathbb{R}^+$ is a strictly convex and differentiable function such that $f(0) = 0$.

The question of the existence of an optimal map for cost functions as in (4) has been mentioned in [22]. In that paper, Carlier, De Pascale, and Santambrogio propose a different strategy based on the existence of a maximizer in the Kantorovich dual problem. As they point out, the existence of such a maximizer is a delicate question. We are nonetheless able to prove its existence under additional assumptions on the initial measure. Note that the case where $f = |\cdot|^2$ is also considered in [30] with the weaker assumption that $\mathcal{C}$ is a convex body.

1.3. **The parametrized optimal transportation problem for relativistic costs.** In this work, we introduce an additional parameter called the speed (of light) which takes into account the relativistic behaviour of the cost functions we consider. To explain this, it is convenient to use a Lagrangian point of view, namely given $c(x,y) = h(x-y)$ a relativistic cost function, observe that

$$c(x,y) = \inf_{C_u} \int_0^1 h(\gamma'(s))ds$$

where $C_u = \{\gamma : [0,1] \to \mathbb{R}^d; \gamma(0) = x, \gamma(1) = y \text{ and } \gamma \text{ is } C^1 \text{ on } (0,1)\}$.

When $h(x-y)$ is finite, we can only consider curves whose energy $\int_0^1 h(\gamma'(s))ds$ is finite, in other terms whose speed γ' is confined to $\mathcal{C}$. Thus, let us introduce the one-parameter family of cost functions

$$c_t(x,y) = h(\frac{x-y}{t})$$

for any positive number t.

Now, given two probability measures μ_0 and μ_1 with compact support, and a relativistic cost function, we can always change the speed of light t so that the minimum in the Kantorovich problem relative to c_t is finite. The study of the variation of this minimum in terms of the speed also gives us some useful informations. Thus, let us define the total cost function

$$C(t) = \min_{\pi \in \Pi(\mu_0,\mu_1)} \int_{\mathbb{R}^{2d}} c(\frac{x-y}{t})\, d\pi(x,y).$$

The overall idea is that the total cost function should be infinite when t is small (at least if $\mu_0 \neq \mu_1$), whereas when t is large, the transportation plans should not be affected by the discontinuity of the cost function and we should get as in [29], existence and uniqueness of the optimal transport map. These properties are proved as part of a more general result in the next sections together with the existence of a critical speed of light T defined below.

Definition 1.1 (Critical speed)**.**

(5) There exists a speed T such that for any $t \geq T$, $C(t) < +\infty$ whereas $C(t) = +\infty$ otherwise.

2. SETTING OF THE RESULT

In order to state the main result of that paper, we need to introduce a few definitions and notations.

Definition 2.1. A probability measure $\mu = \rho dx$ on $\mathbb{R}^d$ is said to be regular if there exists a domain Ω (*i.e.* an open bounded connected set) on which μ is concentrated and such that

$$\exists m > 0; \quad \rho(x) \geq m$$

for almost every $x \in \Omega$.

Definition 2.2 (Directional derivatives)**.** Let $h : \mathbb{R}^d \to [0,+\infty]$ be a convex function such that $h^{-1}([0,+\infty)) = \mathcal{C}$ is the closure of an open bounded convex set. We set h' the directional derivative map

$$\begin{array}{rcl} h' : IDom & \longrightarrow & \mathbb{R} \cup \{-\infty\} \\ (x,w) & \longmapsto & h'(x;w) \end{array}$$

where $IDom = \{(x,w) \in \mathcal{C} \times \mathbb{R}^d; x + w \in \overset{\circ}{\mathcal{C}}\}$ and

$$h'(x;w) = \lim_{\lambda \downarrow 0} \frac{h(x+\lambda w) - h(x)}{\lambda}.$$

Remark 2.3. Observe that $IDom$ is an open bounded subset of $\mathcal{C} \times \mathbb{R}^d$ for the induced topology.

Definition 2.4 (Relativistic cost functions)**.** We say that $c(x,y) = h(x-y)$ is a relativistic cost function if $h : \mathbb{R}^d \to [0,+\infty]$ is a strictly convex function such that $h^{-1}([0,+\infty)) = \mathcal{C}$ is the closure of an open bounded and strictly convex set. We further assume that $h(0) = 0$ and $0 \in \overset{\circ}{\mathcal{C}}$. Last, we assume that h (respectively h') is continuous on $\mathcal{C}$ (respectively on the subset $(h')^{-1}(\mathbb{R}) \subset IDom$).

Remark 2.5. Notice that the condition on the directional derivatives amounts to h differentiable on $\overset{\circ}{\mathcal{C}}$ when the differential blows up on all the boundary of $\mathcal{C}$ as for the relativistic heat cost for instance.

Notations: we set

$$||c||_\infty = \sup_{x \in \mathcal{C}} c(x).$$

Moreover, using that $0 \in \overset{\circ}{\mathcal{C}}$, we set $r_m > 0$ a positive number such that the open ball of radius r_m centered at the origin satisfies

$$B(0, r_m) \subset \mathcal{C}. \tag{6}$$

Last, given two probability measures μ_0 and μ_1 on $\mathbb{R}^d$, we denote by

$$\Gamma_0^t(\mu_0, \mu_1)$$

the set of optimal transport plans relative to the cost function c_t where $t \geq T$ (T being defined in (5)).

The aim of this paper is the proof of the following theorem.

Theorem 2.6. *Let μ_0 and μ_1 be two probability measures with compact support on $\mathbb{R}^d$ and $c(x,y) = h(x-y)$ be a relativistic cost function. We assume that $\mu_0 = \rho_0 dx$ is absolutely continuous with respect to the Lebesgue measure. Then, for any critical or supercritical speed $t \geq T$, there exists a unique optimal transport plan π_0^t for the cost $c_t(x,y) = h(\frac{x-y}{t})$ and this plan is induced by a map*

$$\pi_0^t = (Id, F^t)_\sharp \mu_0.$$

Moreover, if we further assume that μ_0 is regular and that $t > T$, there exists a c_t-concave map ϕ^t such that F^t is defined μ_0 almost everywhere on $p_x(\{x - y \in t\overset{\circ}{\mathcal{C}}\} \cap \operatorname{supp} \pi_0^t)$ by the formula

$$F^t(x) = x - t\nabla h^*(|\tilde{\nabla}\phi^t(x)|) \tag{7}$$

where h^* denotes the Legendre transform of h and $\tilde{\nabla}\phi^t$ is the approximate gradient of ϕ^t (see Definition 5.14 for a definition) .

Remark 2.7. More can be said in the case of the relativistic heat cost by using that the differential of the cost blows up on the unit sphere. Indeed, Corollary 3.6 implies that for almost every supercritical speed, the map F^t is almost everywhere given by (7) since the subset made of pairs of points at maximal distance is a negligible set.

In the next section, we study the parametrized function $C(t)$ and we prove the existence of a critical speed of light T. Then, in section 4, for $t \geq T$ fixed, we use the notion of c-cyclical monotonicity to prove existence and uniqueness of the optimal transport map. In the last section, we give for supercritical speed, the expression of the optimal map in terms of the Kantorovich potential.

3. Properties of the parametrized total cost function

In this part, we are given two probability measures $\mu_0 \neq \mu_1 \in \mathcal{P}(\mathbb{R}^d)$ with compact support. Let us recall the parametrized cost function defined above

$$C(t) = \min_{\pi \in \Pi(\mu_0, \mu_1)} \int_{\mathbb{R}^{2d}} h(\frac{x-y}{t})\, d\pi(x,y). \tag{8}$$

Proposition 3.1 (Properties of $C(t)$)**.** *The total cost function satisfies:*

a) The speed $T = \inf\{s \,;\, C(s) < \infty\}$ is positive and finite. T is called the critical speed.
b) The function $C(t)$ is a decreasing function on $(T, +\infty)$.
c) The total cost $C(T)$ is finite.
d) The fonction $C(t)$ is right continuous on $[T, +\infty)$.

Proof. a) Let us first prove that $T > 0$. Since μ_0 and μ_1 are not equal, there exists a Borelian set B such that $\mu_0(B) \neq \mu_1(B)$. Without loss of generality we can assume that B is a compact set and

$$\mu_0(B) < \mu_1(B). \tag{9}$$

The proof of $T > 0$ is by contradiction. First of all, note that the assumption $C(t) < +\infty$ implies the two following inequalities

$$\mu_0(B) \leq \mu_1(B + t\mathcal{C}) \text{ and } \mu_0(B + t\mathcal{C}) \geq \mu_1(B). \tag{10}$$

Indeed when $C(t) < +\infty$, there exists a transport plan whose support is in $\{(x,y) \,;\, x - y \in t\mathcal{C}\}$. Therefore, if $C(t) < +\infty$ for any $t > 0$, we get by passing to the limit in $\mu_0(B + t\mathcal{C}) \geq \mu_1(B)$ that $\mu_0(B) \geq \mu_1(B)$. Here, we have used the fact that B is a closed subset and $\mathcal{C}$ is a bounded set. This contradicts (9).

Now, let us prove that $T < +\infty$. Since the supports of μ_0 and μ_1 are compact sets, the quantity $\sup\{|x-y|; (x,y) \in \operatorname{supp}\mu_0 \times \operatorname{supp}\mu_1\}$ is finite.

Thanks to (6), there exists t_0 such that $\{x-y; (x,y) \in \operatorname{supp}\mu_0 \times \operatorname{supp}\mu_1\} \subset t_0\,\mathcal{C}$. Consequently,

$$\int c_{t_0}\, d\mu_0 \otimes d\mu_1 < +\infty$$

and $C(t_0) < +\infty$.

b) Let $t_1 > t_2 > T$. By definition of T, $C(t_2) < +\infty$. Let $\pi_2 \in \Pi(\mu_0,\mu_1)$ be an optimal plan for the cost c_{t_2}. By definition of the cost function, we have $\operatorname{supp}\pi_2 \subset \{(x,y); x-y \in t_2\,\mathcal{C}\}$. Now, since $h(0)=0$ and h is strictly convex, for a fixed vector $u \neq 0 \in \mathcal{C}$ and $0 < a < b \leq 1$, we have $\frac{h(au)}{a} < \frac{h(bu)}{b}$. This yields for any $(x,y) \in \operatorname{supp}\pi_2$ such that $x \neq y$, $h(\dfrac{x-y}{t_1}) < h(\dfrac{x-y}{t_2})$. Therefore, since $\mu_0 \neq \mu_1$, we get

$$\int c_{t_1}\, d\pi_2 < \int c_{t_2}\, d\pi_2$$

hence the result, by definition of π_2.

c) Let us consider a sequence of speed $t_n \searrow T$ and π_n a sequence of optimal plans for c_{t_n}. By definition of T, the total cost for any t_n is finite and then for all $n \in \mathbb{N}$, $\pi_{t_n}(\{(x,y); x-y \in t_n\,\mathcal{C}\}) = 1$. Since π_{t_n} is a sequence of probability measures in $\Gamma(\mu_0,\mu_1)$, up to extracting a subsequence, we can assume that π_{t_n} converges weakly to $\pi_\infty \in \Pi(\mu_0,\mu_1)$ thanks to Prokhorov's theorem. Given $\varepsilon > 0$, by noticing that $\{(x,y); x-y \notin (T+\varepsilon)\,\mathcal{C}\}$ is an open set, we get

$$\pi_\infty(\{(x,y); x-y \notin (T+\varepsilon)\,\mathcal{C}\}) \leq \liminf_{n\to 0} \pi_{t_n}(\{(x,y); x-y \notin (T+\varepsilon)\,\mathcal{C}\}) = 0.$$

Thus, $\pi_\infty(\{(x,y); x-y \in (T+\varepsilon)\,\mathcal{C}\}) = 1$ for any $\varepsilon > 0$. Since $\mathcal{C}$ is a closed bounded set, this yields $\pi_\infty(\{(x,y); x-y \in T\,\mathcal{C}\}) = 1$ and, therefore, $\int c_T\, d\pi_\infty < +\infty$.

d) Fix $s \geq T$. Let us consider a sequence $t_n \searrow s$ and for each n, let $\pi_{t_n} \in \Pi(\mu_0,\mu_1)$ be an optimal transport plan for c_{t_n}. Thanks to Prokhorov's theorem, we can assume that π_{t_n} converges to $\pi_\infty \in \Pi(\mu_0,\mu_1)$. Since $t_n > s$,

$$\int_{\mathbb{R}^{2d}} h(\frac{x-y}{t_n}) d\pi_{t_n} \geq \int_{\{x-y\in s\,\mathcal{C}\}} h(\frac{x-y}{t_n}) d\pi_{t_n}.$$

Moreover, since $s\,\mathcal{C}$ is a compact set and h is assumed to be continuous on $\mathcal{C}$, for any $\varepsilon > 0$, there exists n_0 such that for $n \geq n_0$ and $x-y \in s\,\mathcal{C}$,

$$c_{t_n}(x,y) = h(\frac{x-y}{t_n}) \geq h(\frac{x-y}{s}) - \varepsilon.$$

By combining this together with the continuity of c_s when restricted to $\{(x,y); x-y \in s\,\mathcal{C}\}$, we get

$$\liminf_{t_n\to s} C(t_n) = \liminf_{t_n\to s} \int c_{t_n}\, d\pi_{t_n} \geq \int c_s\, d\pi_\infty.$$

On the other hand, since C is decreasing,

$$\liminf_{t_n\to s} C(t_n) \leq C(s).$$

Thus

$$C(s) \geq \lim_{n\to+\infty} C(t_n) = \int c_s\, d\pi_\infty \geq C(s)$$

and the statement is proved. □

Now, we give some properties relative to the critical speed T.

Lemma 3.2. *Let $\pi \in \Pi(\mu_0,\mu_1)$ such that $\int c_T\, d\pi < +\infty$. Then, this plan satisfies*

$$\operatorname{supp}\pi \cap \{(x,y); x-y \in T\partial\mathcal{C}\} \neq \emptyset.$$

Moreover, the critical speed is characterized by the equality

$$T = \min_{\pi\in\Pi(\mu_0,\mu_1)} \pi - ess\sup |x-y|_{\mathcal{C}}$$

where $|x|_{\mathcal{C}} = \inf\{\lambda; x \in \lambda\,\mathcal{C}\}$ and $\pi - ess\sup|x-y|_{\mathcal{C}} = \inf\{s; \pi(\{(x,y); |x-y|_{\mathcal{C}} > s\}) = 0\}$.

Remark 3.3. In the case of the relativistic heat cost (or any other radial relativistic cost function), $|\cdot|_{\mathcal{C}}$ is nothing but the Euclidean norm (up to scaling). Thus in this case, $\min_{\pi\in\Pi(\mu_0,\mu_1)} \pi - ess\sup|x-y|_{\mathcal{C}}$ coincides with the $\infty-$Wasserstein distance studied in [26].

Remark 3.4. This lemma means that for the critical speed T, the support of any optimal transport plan must intersect the critical set $\{(x,y)\ ;\ x-y\in T\partial\mathcal{C}\}$ but we do not know whether or not all the optimal transport plans give some mass to this critical set. See below for more on this set in the case of the relativistic heat cost.

Proof. We prove the first statement by contradiction. Suppose on the contrary that $\operatorname{supp}\pi\cap\{(x,y);x-y\in T\partial\mathcal{C}\}=\emptyset$. By assumption on π, we have $\operatorname{supp}\pi\subset\{(x,y);x-y\in T\,\mathcal{C}\}$. Thus, by compactness of the support of π, there exists $\delta>0$ such that

$$\operatorname{supp}\pi\subset\{(x,y);x-y\in(T-\delta)\,\mathcal{C}\}.$$

This implies $C(T-\delta)<+\infty$, a contradiction.

Now, we prove the second statement. Let us start with the following equalities.

$$\begin{aligned}
T &= \inf\{s;C(s)<+\infty\}\\
&= \inf\{s;\ \exists\pi\in\Pi(\mu_0,\mu_1)\text{ s.t. }\ \pi(\{(x,y);x-y\in s\,\mathcal{C}\})=1\}\\
&= \inf\{s;\ \exists\pi\in\Pi(\mu_0,\mu_1)\text{ s.t. }\ \pi(\{(x,y);|x-y|_{\mathcal{C}}\le s\})=1\}\\
&= \inf\{s;\ \exists\pi\in\Pi(\mu_0,\mu_1)\text{ s.t. }\ \pi(\{(x,y);|x-y|_{\mathcal{C}}> s\})=0\}\\
&= \inf_{\pi\in\Pi(\mu_0,\mu_1)}\inf\{s;\ \pi(\{(x,y);|x-y|_{\mathcal{C}}>s\})=0\}\\
&= \inf_{\pi\in\Pi(\mu_0,\mu_1)}\pi-ess\sup|x-y|_{\mathcal{C}}.
\end{aligned}$$

To conclude, let us prove that the infimum above is actually a minimum. To this aim, consider a minimizing sequence $(\pi_i)_{i\in\mathbb{N}}$ and assume thanks to Prokhorov's theorem, that $\pi_i\rightharpoonup\pi_\infty\in\Pi(\mu_0,\mu_1)$. We need the following fact.

Fact: $\forall(x,y)\in\operatorname{supp}\pi_\infty,\ \exists(x_i,y_i)_{i\in\mathbb{N}}\in\operatorname{supp}\pi_i;(x_i,y_i)\to(x,y)$.

A proof is given in [26, Lemma 2.3]. Thus, since $\mathcal{C}$ satisfies (6), this gives for any $\varepsilon>0$,

$$\operatorname{supp}\pi_\infty\subset\operatorname{supp}\pi_i+\varepsilon B(0,1)\subset\operatorname{supp}\pi_i+\frac{\varepsilon}{r_m}\,\mathcal{C}$$

for large i. This yields

$$\sup_{\operatorname{supp}\pi_\infty}|x-y|_{\mathcal{C}}\le\liminf_{i\to+\infty}\sup_{\operatorname{supp}\pi_i}|x-y|_{\mathcal{C}}$$

thus π_∞ is actually a minimizer. □

In the case of the relativistic heat cost, we obtain the following additional proposition and its straightforward corollary (similar statements are probably within reach for other relativistic cost functions).

Proposition 3.5. *In the case of the relativistic heat cost denoted by* $c_t(x,y)=h(\frac{|x-y|}{t})$, *the following property holds true for any* $t\ge T$

$$C'_+(t)>-\infty\Longrightarrow\forall\pi\in\Gamma_0^t(\mu_0,\mu_1),\ \ \pi(\{(x,y);|x-y|=t\})=0$$

where $C'_+(t)$ *denotes the right derivative of* C *at* t.

Proof. Recall that $T=\inf\{t\ ;\ c_t<\infty\}$. Let $\pi_t\in\Gamma_0^t(\mu_0,\mu_1)$. For any $s>t$, let us compute

$$\begin{aligned}
C(t)-C(s) &\ge \int h(\frac{|x-y|}{t})\,d\pi_t-\int h(\frac{|x-y|}{s})\,d\pi_t\\
&\ge \int_{\{(x,y);|x-y|=t\}}h(\frac{|x-y|}{t})-h(\frac{|x-y|}{s})\,d\pi_t\\
&\ge \int_{\{(x,y);|x-y|=t\}}h(1)-h(\frac{t}{s})\,d\pi_t\\
&\ge \frac{s-t}{s}h'(\frac{t}{s})\pi_t(\{|x-y|=t\}).
\end{aligned}$$

We get the result by letting s decrease to t. □

Since C is a monotone function, the proposition implies readily the following property on the critical set $\{(x,y);|x-y|=t\}$.

Corollary 3.6. *In the case of the relativistic heat cost, any optimal transport plan* $\pi_0^t \in \Gamma_0^t(\mu_0, \mu_1)$ *satisfies*

$$\pi_0^t(\{(x,y); |x-y| = t\}) = 0$$

for almost every $t \geq T$.

4. Existence and uniqueness of optimal map

In this section, we prove the uniqueness of the optimal transport plan and that this plan is induced by a map. Let us recall the definition of a relativistic cost function.

Definition 4.1 (Relativistic cost functions)**.** We say that $c(x,y) = h(x-y)$ is a relativistic cost function if $h : \mathbb{R}^d \to [0, +\infty]$ is a strictly convex function such that $h^{-1}([0,+\infty)) = \mathcal{C}$ is the closure of an open bounded and strictly convex set. We further assume that $h(0) = 0$ and $0 \in \overset{\circ}{\mathcal{C}}$. Last, we assume that h (respectively h') is continuous on $\mathcal{C}$ (respectively on the subset $(h')^{-1}(\mathbb{R}) \subset IDom$ where $IDom = \{(x,w) \in \mathcal{C} \times \mathbb{R}^d; x + w \in \overset{\circ}{\mathcal{C}}\}$).

Remark 4.2. In general, the function h' is an upper semi-continuous function on $IDom$ (this follows from a straightforward adaption of the proof of [37, Theorem 24.5]). Thus, the assumption above actually implies that h' is continuous on $IDom$. Note also that if $h'(x;w) = -\infty$ then $h'(x;w') = -\infty$ for all w' such that $(x,w') \in IDom$. Under our assumptions these points x are exactly those in $\mathcal{C}$ such that $\partial h(x) = \emptyset$ (we refer to [37, Theorem 23.3] for a proof of the last two statements).

The main result of this part is the following theorem.

Theorem 4.3. *Let* μ_0 *and* μ_1 *be two probability measures with compact support on* $\mathbb{R}^d$*. We assume that* $\mu_0 = \rho_0 dx$ *is absolutely continuous with respect to the Lebesgue measure. Then, for any critical or supercritical speed* $t \geq T$*, there exists a unique optimal transport plan* π_0^t *for the cost* c_t *and this plan is induced by a map*

$$\pi_0^t = (Id, F^t)_\sharp \mu_0.$$

The proof of this result is based on a method introduced by Champion, de Pascale and Juutinen [26]. A cornerstone of their proof is the following lemma.

Lemma 4.4. *Assume that* $\mu << \mathcal{L}^d$*. Let* $\pi \in \Pi_0^t(\mu,\nu)$ *and* Γ *a set on which* π *is concentrated. Then* π *is concentrated on a* $\sigma-$*compact set* $R(\pi) \subset \Gamma$ *such that for all* $(x,y) \in R(\pi)$ *the point* x *is a Lebesgue point of* $p_x(supp\ \mu_0 \times B(y,r) \cap \Gamma)$ *for all* $r > 0$.

This precise version is proved in [24, Lemma 4.3].

Definition 4.5. A couple $(x,y) \in \Gamma$ is said to be π-regular if x is a Lebesgue point of $p_x(\text{supp } \mu_0 \times B(y_0,r) \cap \Gamma)$ for any positive r.

Consequently, we get the corollary below.

Corollary 4.6. *Let* $(x_0,y_0) \in R(\pi)$*,* $r > 0$*,* $\alpha \in (0,1)$*,* $\xi \in \mathbb{S}^{d-1}$ *and* $\delta > 0$*. Then for* $\varepsilon > 0$ *sufficiently small, the set of points* $x \in p_x(supp\ \mu_0 \times B(y,r)) \cap \Gamma)$ *such that* $x \in \hat{C}(x_0,\xi,\delta) \cap (B(x_0,\varepsilon) \setminus B(x_0,\alpha\varepsilon))$ *has positive measure where*

$$\hat{C}(x_0,\xi,\delta) = \{x \in \mathbb{R}^d \setminus \{x_0\}; \frac{x-x_0}{|x-x_0|} \cdot \xi \geq 1 - \delta\}.$$

Before we start the proof of Theorem 4.3, let us recall the definition of $c-$cyclically monotone subset. This notion was introduced by Knott and Smith [41] in order to detect the optimality of a given transport plan.

Definition 4.7 (c-monotone set)**.** A subset $S \subset \mathbb{R}^d \times \mathbb{R}^d$ is called a c-cyclically monotone set if and only if for any positive integer m and any pairs $(x_1,y_1), \cdots, (x_m,y_m) \in S$, the following inequality is satisfied:

$$c(x_1,y_1) + \cdots + c(x_m,y_m) \leq c(x_1,y_2) + c(x_2,y_3) + \cdots + c(x_m,y_1).$$

It is well-known that an optimal plan is concentrated on a measurable c-cyclically monotone set when (the transport problem is well-posed and) the cost function $c : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^+ \cup \{+\infty\}$ is a lower semi-continuous map [7, Theorem 3.2].

4.1. **Proof of Theorem 4.3.** We are given $\pi_0^t \in \Pi_0^t(\mu_0, \mu_1)$. Thanks to Lemma 4.4 and the result above, there exists a measurable c-cyclically monotone set $\Gamma_0^t \subset \operatorname{supp} \pi_0^t$ made of regular points such that $\pi_0^t(\Gamma_0^t) = 1$. Following the approach in [26] and the subsequent papers, we are going to prove by contradiction that Γ_0^t is a graph.

Let us assume that (x_0, y_0) and (x_0, y_1) belong to Γ_0^t with $y_0 \neq y_1$ and set $u_0 = (x_0 - y_0)/t$ and $u_1 = (x_0 - y_1)/t$. Note that $u_0 \neq u_1$ and u_0 and u_1 belong to $\mathcal{C}$ since $t \geq T$. We also claim that

$$h'(u_0; \overrightarrow{u_0u_1}) + h'(u_1; \overrightarrow{u_1u_0}) < 0. \tag{11}$$

Indeed, $h'(u_0; \overrightarrow{u_0u_1}) + h'(u_1; \overrightarrow{u_1u_0}) = \theta'_+(0) - \theta'_-(1)$ where θ'_+ (respectively θ'_-) denotes the right derivative (respectively left derivative) of the convex real-valued function of a real variable $\theta(s) = h(u_0 + s\overrightarrow{u_0u_1})$. Now, since $\mathcal{C}$ is a strictly convex set, the open segment $(u_0, u_1) = \{tu_0 + (1-t)u_1; t \in (0,1)\}$ is actually within $\overset{\circ}{\mathcal{C}}$. Therefore, since $IDom$ is an open subset of $\mathcal{C} \times \mathbb{R}^d$ with respect to the induced topology, we get the existence of numbers $r > 0$ and $r' > 0$ such that for all $v \in B(u_0, r) \cap \mathcal{C}$, $w \in B(u_1, r) \cap \mathcal{C}$ and $z \in B(0, r')$

$$v + \overrightarrow{u_0u_1}/2 + z \in \overset{\circ}{\mathcal{C}} \text{ and } w - \overrightarrow{u_0u_1}/2 - z \in \overset{\circ}{\mathcal{C}}.$$

Let us recall that $y \to h'(x; y)$ is a positively homogeneous function. Thanks to the formula above, we get that

$$v + \lambda(\overrightarrow{u_0u_1}/2 + z) \in \overset{\circ}{\mathcal{C}} \text{ and } w - \lambda(\overrightarrow{u_0u_1}/2 - z) \in \overset{\circ}{\mathcal{C}}$$

whenever $v \in B(u_0, r) \cap \mathcal{C}$, $0 < \lambda \leq 1$ and $z \in B(0, r')$. For the sake of clarity, let us introduce a notation for some truncated cones:

$$C(0, \overrightarrow{u_0u_1}/2, r'; \eta) = \{\lambda x \in \mathbb{R}^d;\ \lambda > 0 \text{ and } x \in \overrightarrow{u_0u_1}/2 + B(0, r')\} \cap \overline{B}(0, \eta) \setminus B(0, \eta/2).$$

In particular, when η is small enough (and $r, r' << |\overrightarrow{u_0u_1}/2|$), we have

$$B(u_0, r) \cap \mathcal{C} + C(0, \overrightarrow{u_0u_1}/2, r'; \eta) \subset \overset{\circ}{\mathcal{C}} \text{ and } B(u_1, r) \cap \mathcal{C} - C(0, \overrightarrow{u_0u_1}/2, r'; \eta) \subset \overset{\circ}{\mathcal{C}}.$$

Now, since h is continuous on $\mathcal{C}$ and h' is continuous on $IDom$ (see Remark 4.2), for any $\varepsilon > 0$ there exist $0 < \overline{r} < r, 0 < \overline{r'} < r'$ and $\eta > 0$ small such that

$$\sup\left\{ \left| \frac{h(v+z) - h(v)}{|z|} - h'(u_0, \overrightarrow{u_0u_1}/|\overrightarrow{u_0u_1}|) \right|; v \in B(u_0, \overline{r}) \cap \mathcal{C}, z \in C(0, \overrightarrow{u_0u_1}/2, \overline{r'}; \eta) \right\} < \epsilon \tag{12}$$

and

$$\sup\left\{ \left| \frac{h(w-z) - h(w)}{|z|} - h'(u_1, -\overrightarrow{u_0u_1}/|\overrightarrow{u_0u_1}|) \right|; w \in B(u_1, \overline{r}) \cap \mathcal{C}, z \in C(0, \overrightarrow{u_0u_1}/2, \overline{r'}; \eta) \right\} < \varepsilon. \tag{13}$$

In the second part of the proof, we use the fact that (x_0, y_1) is a regular pair in the sense of Definition 4.5 to get a contradiction. Given $\varepsilon > 0$, let $\overline{r}, \overline{r'}, \eta > 0$ such that (12) and (13) hold. We also assume that $\eta < t\overline{r}/2$. Now, applying Corollary 4.6 to (x_0, y_1) with $r = t\overline{r}/2$, we get the existence of $(\tilde{x}, \tilde{y}) \in \Gamma_0^t$ such that $\tilde{x} \in \{x_0\} + C(0, \overrightarrow{u_0u_1}/2, \overline{r'}; \eta)$ and $\tilde{y} \in B(y_1, t\overline{r}/2)$ (up to shrinking η). By construction we have $(\tilde{x} - \tilde{y})/t \in B(u_1, \overline{r})$ and $(\tilde{x} - \tilde{y}) \in t\mathcal{C}$ since $t \geq T$. Now, by applying the estimates (12) and (13) with $v = u_0$, $w = (\tilde{x} - \tilde{y})/t$ and $z = (\tilde{x} - x_0)/t$, we get

$$\begin{aligned} h((\tilde{x} - y_0)/t) + h((x_0 - \tilde{y})/t) \quad \leq \quad & h((x_0 - y_0)/t) + h((\tilde{x} - \tilde{y})/t) \\ & + |(\tilde{x} - x_0)/t| \big(2\varepsilon + h'(u_0; \overrightarrow{u_0u_1}/|\overrightarrow{u_0u_1}|) + h'(u_1; \overrightarrow{u_1u_0}/|\overrightarrow{u_0u_1}|)\big). \end{aligned}$$

Combining this together with (11) contradicts the c-cyclical monotonicity of Γ_0^t for small ε. Thus, we have proved that any optimal plan is concentrated on a measurable graph.

The uniqueness proof is more classical. Consider two opimal plans, by what precedes these plans are induced by maps. The arithmetic mean of these plans is still an optimal plan but it cannot be induced by a map unless the two plans coincide.

5. Supercritical speed

This section is valid only for supercritical speed. The goal of this part is to obtain (7) linking the optimal map and the Kantorovich potential. We recall that a domain is an open bounded connected set.

Theorem 5.1. *Let μ_0 and μ_1 be two probability measures with compact support on $\mathbb{R}^d$. We assume that $\mu_0 = \rho_0 dx$ is regular and concentrated on a domain Ω_0 in the sense of Definition 2.1. Then, for any supercritical* *speed* $t > T$*, there exists a* c_t*-concave map* ϕ^t *such that the map* F^t *defined by* $\pi_0^t = (Id, F^t)_\sharp \mu_0$ *is defined almost everywhere on* $p_x(\{x - y \in t\overset{\circ}{\mathcal{C}}\} \cap \Gamma_0^t)$ *by the formula*

$$F^t(x) = x - t\nabla h^*(\tilde{\nabla}\phi^t(x)).$$

This section is divided into three parts. In the first one, we establish a technical result on the volume growth of set enlargements. This result is a key result in the study of special maps related to the transport problem (namely c_t-concave maps) that is performed in the second part. In the last part, we prove the theorem above.

5.1. **Volume growth estimate for set enlargements.** The goal of this part is to establish the technical proposition below.

Proposition 5.2. *Let* Ω *be a domain and* $t > 0$*. For any constants* $0 < M_1 < M_2 < |\Omega|$*, there exists a positive constant* $\epsilon = \epsilon(\Omega, t, M_1, M_2)$ *depending on* Ω*,* t*,* M_1*, and* M_2 *such that*

$$\inf_{\mathcal{S}} \left(|(E + t\,\mathcal{C}) \cap \Omega| - |E \cap \Omega|\right) \geq \epsilon$$

where $\mathcal{S} = \{E \text{ compact subset of } \mathbb{R}^d; M_1 \leq |E \cap \Omega| \leq M_2\}$*.*

Moreover, there exists a constant $0 < \kappa < |\Omega|$ *depending on* t *and* Ω *such that for any compact subset* E *of* $\mathbb{R}^d$ *satisfying* $|E \cap \Omega| > \kappa$*, the following equality holds*

$$|(E + t\,\mathcal{C}) \cap \Omega| = |\Omega|.$$

This result is proved by contradiction. The first step of the proof is to reduce the result to a similar statement involving the De Giorgi perimeter. Then, we use a compactness result for functions of bounded variation.

5.1.1. *Notions of boundary area for a nonsmooth set.* The first notion is the (relative) Minkowski content of a set. This is the notion we are mainly interested in. Its definition is the following.

Definition 5.3 (Relative Minkowski content)**.** Let E be a compact subset of $\mathbb{R}^d$. The (relative) Minkowski content of E is defined as

$$Mink_{\mathcal{C}}(E \,|\, \Omega) = \liminf_{r \downarrow 0} \frac{|(E + r\,\mathcal{C}) \cap \Omega| - |E \cap \Omega|}{r}.$$

The second notion is the De Giorgi perimeter. We refer to the book [2] for a detailed exposition.

Definition 5.4 (Relative perimeter)**.** Let E be a mesurable subset of $\mathbb{R}^d$. The perimeter of E relative to Ω is defined as

$$Per(E \,|\, \Omega) = \sup\left\{\int_E \operatorname{div} \varphi \, dx; \varphi \in \mathcal{C}_c^1(\Omega, B)\right\}$$

where $B = B(0,1)$ is the unit Euclidean ball centered at the origin in $\mathbb{R}^d$.

This definition is not the only one but is certainly the most elementary. Note that when $E \subset \mathbb{R}^d$ is a bounded open set with smooth boundary, the divergence theorem implies $Per(E \,|\, \Omega) = \mathcal{H}^{d-1}(\partial E \cap \Omega)$. Moreover, with this definition, it is easy to see that the perimeter is a lower semi-continuous functional with respect to the L^1-topology (also called topology of sets convergence) whose definition is recalled below.

Definition 5.5 (Set of finite perimeters and convergence)**.** We set $\mathcal{M}_{<+\infty}$ the set of measurable subsets E of $\mathbb{R}^d$ whose relative perimeter $Per(E \,|\, \Omega)$ is finite. The set $\mathcal{M}_{<+\infty}$ is endowed with the L^1-topology. Namely, a sequence $(E_n)_{n\in\mathbb{N}}$ converges to E in $\mathcal{M}_{<+\infty}$ if, by definition, the (restriction to Ω of the) characteristic functions χ_{E_n} converges to χ_E in $L^1(\Omega)$.

Lemma 5.6. *Given a domain* Ω *in* $\mathbb{R}^d$*, the relative perimeter is a lower semi-continuous map*

$$Per(\cdot \,|\, \Omega) : \mathcal{M}_{<+\infty} \longrightarrow \mathbb{R}^+$$

when the set $\mathcal{M}_{<+\infty}$ *is endowed with the* L^1*-topology.*

Proof. Given $\varphi \in \mathcal{C}_c^1(\Omega, B)$, the map $E \mapsto \int_E \operatorname{div} \varphi \, dx$ is clearly continuous for the L^1-topology. Therefore, $Per(\cdot \,|\, \Omega)$ is lower semi-continuous as the supremum of continuous functions. □

Our goal is to compare the relative Minkowski content to the De Giorgi perimeter. To this aim, we first reduce this comparison result to the case when $\mathcal{C} = B$. Recall that, by definition of $\mathcal{C}$, there exists a constant $r_m > 0$ such that the open ball $B(0, r_m)$ satisfies

$$B(0, r_m) \subset \mathcal{C}. \tag{14}$$

Thus, we get the inequality

$$Mink_{\mathcal{C}}(E\,|\,\Omega) \geq r_m \liminf_{r\downarrow 0} \frac{|(E+rB(1))\cap\Omega| - |E\cap\Omega|}{r} = r_m Mink_B(E\,|\,\Omega). \tag{15}$$

Complementary to this result is the following proposition.

Proposition 5.7. *Let E be a compact subset of $\mathbb{R}^d$. Then, the following inequality holds*

$$Mink_B(E\,|\,\Omega) \geq Per(E\,|\,\Omega). \tag{16}$$

Proof. We first rewrite

$$\frac{|(E+rB)\cap\Omega| - |E\cap\Omega|}{r} = \frac{1}{r}\int_{E_r\setminus E}\chi_\Omega$$

where χ_Ω denotes the characteristic function of Ω and $E_r = E + rB$. Now, we set d_E the Euclidean distance funtion to the set E. More precisely for any $x \in \mathbb{R}^d$,

$$d_E(x) = \min_{y\in E}|x-y|.$$

We have the following equalities

$$\begin{aligned}\frac{1}{r}\int_{E_r\setminus E}\chi_\Omega &= \frac{1}{r}\int_{E_r\setminus E}\chi_\Omega|\nabla d_E|\\ &= \frac{1}{r}\int_0^r \mathcal{H}^{d-1}(\partial E_s\cap\Omega)\,ds\\ &= \int_0^1 \mathcal{H}^{d-1}(\partial E_{sr}\cap\Omega)\,ds.\end{aligned}$$

where the second equality follows from the coarea formula. Now, thanks to [2, Proposition 3.62], we get for any measurable subset A of $\mathbb{R}^d$

$$\mathcal{H}^{d-1}(\partial A\cap\Omega) \geq Per(A\,|\,\Omega).$$

Therefore, this yields

$$\frac{|(E+rB(1))\cap\Omega| - |E\cap\Omega|}{r} \geq \int_0^1 Per(E_{sr}\,|\,\Omega)\,ds.$$

To conclude, note that E_{sr} converges to $E_0 = E$ with respect to the L^1-convergence. Thus, we get the result by combining Fatou's lemma together with Lemma 5.6. ☐

Remark 5.8. The proof above is inspired by [8].

5.1.2. *An intermediate result.* Thanks to the properties of De Giorgi perimeter, we can prove the following result.

Proposition 5.9. *Let Ω be a domain of $\mathbb{R}^d$. For any constants $0 < M_1 < M_2 < |\Omega|$, there exists a positive constant $\epsilon = \epsilon(\Omega, M_1, M_2)$ depending on Ω, M_1, and M_2 such that*

$$\inf_{\mathcal{M}} Per(E\,|\,\Omega) \geq \epsilon$$

where $\mathcal{M} = \{E \text{ measurable subset of } \Omega; M_1 \leq |E| \leq M_2\}$

Proof. We prove the result by contradiction by assuming that the infimum above is 0. Let $(E_i)_{i\in\mathbb{N}}$ be a minimizing subsequence. Thanks to [2, Theorem 3.39 p.145], there exists a subsequence still denoted by $(E_i)_{i\in\mathbb{N}}$ which converges to a measurable subset E_∞ of Ω in $L^1(\Omega)$. By definition of this convergence, E_∞ belongs to $\mathcal{M}$ and Lemma 5.6 yields

$$Per(E_\infty\,|\,\Omega) = 0.$$

Consequently, up to modifying E_∞ on a negligible subset, χ_{E_∞} is constant on the connected set Ω (see [2, Proposition 3.2 p118] for a proof). This contradicts the fact that $E_\infty \in \mathcal{M}$ and the proof is complete. ☐

By combining (15), Proposition 5.7 together with the proposition above, we get

Proposition 5.10. *Let Ω be a domain of $\mathbb{R}^d$. For any constants $0 < M_1 < M_2 < |\Omega|$, there exists $\varepsilon = \varepsilon(\Omega, M_1, M_2) > 0$ depending on Ω, M_1, and M_2 such that*

$$\inf_{\mathcal{S}} Mink_{\mathcal{C}}(E\,|\,\Omega) \geq \varepsilon.$$

5.1.3. *Proof of Proposition 5.2.* Both statements are proved by contradiction. Let us remark that the quantity $|(E+t\mathcal{C})\cap\Omega| - |E\cap\Omega|$ is unchanged if we modify E outside a large ball $B(R)$ (whose radius R depends on the diameters of Ω and $\mathcal{C}$). Thus in this proof, we only consider compact sets into $B(R)$. Let $(E_n)_{n\in\mathbb{N}}$ be a sequence of elements in $\mathcal{S}$ such that

$$|(E_n+t\mathcal{C})\cap\Omega| \leq |E_n\cap\Omega| + \frac{1}{n} \tag{17}$$

and $(E_n)_{n\in\mathbb{N}}$ converges to the compact set $E_\infty \subset B(R)$ with respect to the Hausdorff metric (see [40] for a definition). By definition of this metric, we get

$$\limsup_{n\to+\infty} |E_n\cap\Omega| \leq |E_\infty\cap\Omega|. \tag{18}$$

Now, recall that $B(0,r_m)\subset\mathcal{C}$ (14). For large n, we have

$$E_\infty \subset E_n + \frac{r_m t}{2}B$$

by definition of the Hausdorff metric, therefore

$$E_\infty + \frac{t}{2}\mathcal{C} \subset E_n + t\mathcal{C}$$

and

$$|E_\infty\cap\Omega| \leq |(E_\infty + \frac{t}{2}\mathcal{C})\cap\Omega| \leq |(E_n+t\mathcal{C})\cap\Omega| \tag{19}$$

for large n.

Consequently thanks to (17) we get $\liminf_{n\to+\infty}|E_n\cap\Omega| \geq |E_\infty\cap\Omega|$ and, thanks to (18), $E_\infty\in\mathcal{S}$. Moreover, (19) also yields

$$|(E_\infty + \frac{t}{2}\mathcal{C})\cap\Omega| - |E_\infty\cap\Omega| = 0. \tag{20}$$

Let us prove that the equality above leads to a contradiction. We set θ the non decreasing and right continuous map defined by

$$\theta(s) = |(E_\infty + s\,\mathcal{C})\cap\Omega|.$$

In particular, this map is differentiable for almost every s_0 and the first derivative equals $Mink_{\mathcal{C}}(E_\infty + s_0\,\mathcal{C}\,|\,\Omega)$. Moreover, since θ is right continuous and non decreasing, the set $(E_\infty + s_0\,\mathcal{C})$ is in $\mathcal{S}$ when s_0 is small (up to slightly increase M_2). Therefore, Proposition 5.10 implies that for almost every small s, $\theta'(s)\geq\varepsilon$ and this contradicts (20). The first part of the proposition is proved.

The proof of the second part is very similar so we only sketch it. Consider a sequence $(E_n)_{n\in\mathbb{N}}$ of compact subsets of $B(R)$ such that $|E_n\cap\Omega|\to|\Omega|$ and $|(E_n+t\mathcal{C})\cap\Omega| < |\Omega|$. Since $B(R)$ is compact, we can assume that E_n converges to a compact set E_∞ with respect to the Hausdorff metric. Thanks to (18), we get $|E_\infty\cap\Omega| = |\Omega|$. Now, using (14), we get

$$E_n + t\,\mathcal{C} \supset E_n + tr_m B \supset E_\infty$$

for large n. Thus,

$$|(E_n+t\,\mathcal{C})\cap\Omega| \geq |E_\infty\cap\Omega| = |\Omega|$$

for large n, hence a contradiction.

5.2. **Regularity of Kantorovich potentials.** Throughout this section, we fix $t>T$ a supercritical speed and the dependence on t may be omitted at some points. Let us recall the notion of c-concave function.

Definition 5.11 (c-concave function and c-subdifferential)**.** Let $\phi:\Omega_0\to\mathbb{R}\cup\{-\infty\}$ be a function. We define $\phi^c:\operatorname{supp}\mu_1\to\mathbb{R}\cup\{-\infty,+\infty\}$, the c-transform of ϕ by the formula

$$\phi^c(y) = \inf_{x\in\Omega_0} c(x,y)-\phi(x).$$

Such a function ϕ is said to be c-concave if for all $y\in\operatorname{supp}\mu_1$, $\phi^c(y)<+\infty$ (so that $(\phi^c)^c$ is well-defined) and $(\phi^c)^c=\phi$ (where usually, the role of x and y is unchanged because of the possible asymmetry of the cost function; note that in the rest of the paper, we will write ϕ^{cc} for simplicity). The c-subdifferential of a c-concave function ϕ is defined by the formula $\partial_c\phi(\Omega_0)=\cup_{x\in\Omega_0}\partial_c\phi(x)$ where

$$\partial_c\phi(x) = \{y\in\operatorname{supp}\mu_1;\ \phi(x)+\phi^c(y)=c(x,y)\}$$

Remark 5.12. The c-subdifferential of a c-convave function is a c-monotone set, see for instance [44].

In order to motivate the study of (certain) c_t-concave maps, we start with a Rockafellar-type theorem.

Theorem 5.13. *Let μ_0 and μ_1 be two probability measures on $\mathbb{R}^d$ with compact support. We also assume that $\mu_0 = \rho_0 dx$ is regular and concentrated on a domain Ω_0 in the sense of Definition 2.1. Then for $t > T$, there exists a c_t-concave map ϕ^t such that*
$$\pi_0^t((Id, \partial_{c_t}\phi^t)) = 1$$
where $(Id, \partial_{c_t}\phi^t) = \{(x,y) \in \Omega_0 \times \operatorname{supp}\mu_1; \phi^t(x) + {\phi^t}^c(y) = c_t(x,y)\}$. Such a map is called a Kantorovich potential.

Proof. Let Γ_0^t be the set introduced in the proof of Theorem 4.3. We shall prove
$$\Gamma_0^t \subset (Id, \partial_{c_t}\phi^t).$$
We fix $(x_0, y_0) \in \Gamma_0^t$ and define
$$\phi^t(x) =$$
$$\inf_{m\in\mathbb{N}} \inf\{c_t(x_1,y_0)+\cdots+c_t(x_m,y_{m-1})+c_t(x,y_m)-(c_t(x_0,y_0)+\cdots+c_t(x_m,y_m)); \forall i \in \{1,\cdots,m\}\ (x_i,y_i) \in \Gamma_0^t\}.$$

Since t is a supercritical speed, $\operatorname{supp}\pi_0^t \subset \{c_t < +\infty\}$. Thus, for $(x_1,y_1)\cdots(x_m,y_m) \in \Gamma_0^t$, the term inside the brackets above is in $\mathbb{R}\cup\{+\infty\}$ and the map ϕ^t is well-defined. Moreover, $\phi^t(\Omega_0) \subset \mathbb{R}\cup\{\pm\infty\}$. Now we shall prove that for all $x \in \Omega_0$, $\phi^t(x) < +\infty$. To this aim, let us recall that Γ_0^t is a c_t-monotone set. Taking $m = 0$ in the definition above leads to the inequality $\phi^t(x_0) \le 0$, the c_t-monotonicity of Γ_0^t gives us the reverse inequality, thus $\phi^t(x_0) = 0$. As a particular case, consider $x \in \{y_0\} + t\,\mathcal{C}$. By definition of the cost function, it is clear that $\phi^t(x) < +\infty$ for such a x. For arbitrary x, we shall show that there exists a finite chain of points linking x_0 to x such that, loosely speaking, the previous condition is satisfied. More precisely, we shall prove that there exists $k \in \mathbb{N}\setminus\{0\}$ such that for any $x \in \Omega_0$, there exists $(x_i,y_i)_{i=1}^k \in \Gamma_0^t$ such that
$$\begin{aligned} c_t(x_1,y_0) &< +\infty,\\ \forall i \in \{1,\cdots,k-1\}\ c_t(x_{i+1},y_i) &< +\infty,\\ c_t(x,y_k) &< +\infty. \end{aligned} \tag{21}$$
To this aim, let us notice that by Proposition 3.1, we have the property $C(T) < +\infty$ which implies, see (10), that for all closed set A,
$$\begin{aligned} \mu_0(A + T\,\mathcal{C}) &\ge \mu_1(A),\\ \mu_1(A + T\,\mathcal{C}) &\ge \mu_0(A). \end{aligned} \tag{22}$$

Now, we fix t' such that $T < t' < t$ and define by induction the following sequence of sets:
$$\begin{aligned} A_0 &= \{y_0\},\\ A_{i+1} &= \overline{p_y(p_x^{-1}(A_i + t'\,\mathcal{C}) \cap \Gamma_0^t)}. \end{aligned}$$
where p_x and p_y stand for the projections on the x and y coordinates respectively, and $\overline{A}$ is the closure of A (in $\mathbb{R}^d$). Formally speaking, Property (21) is equivalent to the existence of an integer k such that $A_k + t'\,\mathcal{C} \supset \Omega_0$ with $t' = t$. However, since the sets $p_y(p_x^{-1}(A_i + t\,\mathcal{C}) \cap \Gamma_0^t)$ are not necessarily closed sets, we have to require a little more to get the claim. Note that the existence of such a k also implies
$$\forall x \in \Omega_0, \quad \phi^t(x) \le ||c||_\infty (k+1). \tag{23}$$
Let us prove that $A_k + t'\,\mathcal{C} \supset \Omega_0$ for large k.

By using the optimal transport plan π_0^t, we obtain for any integer i
$$\mu_1(A_{i+1}) \ge \mu_0(A_i + t'\,\mathcal{C}).$$
Property (22) above gives us
$$\begin{aligned} \mu_1(A_{i+1}) &\ge \mu_1(A_i) + \mu_0((A_i + t'\,\mathcal{C}) \setminus (A_i + T\,\mathcal{C}))\\ &\ge \mu_1(A_i) + \operatorname{essinf}(\rho_0)\,(|(A_i + t'\,\mathcal{C}) \cap \Omega_0| - |(A_i + T\,\mathcal{C}) \cap \Omega_0|). \end{aligned} \tag{24}$$
Now, we claim that there exists $M_1 > 0$ such that for any integer $i \ge 1$,
$$|(A_i + T\,\mathcal{C}) \cap \Omega_0| \ge M_1. \tag{25}$$
To this aim, first notice that
$$\mu_0(A_0 + t'\,\mathcal{C}) = \mu_0(\{y_0\} + t'\,\mathcal{C}) > 0. \tag{26}$$
This follows from the existence of a small ε such that
$$\{y_0\} + t'\,\mathcal{C} \supset \{y_0\} + T\,\mathcal{C} + \overline{B}(\varepsilon)$$

together with an application of (22):

$$\mu_0(\overline{B}(y_0,\varepsilon)+T\,\mathcal{C}) \geq \mu_1(\overline{B}(y_0,\varepsilon)) > 0$$

since y_0 belongs to $\operatorname{supp}\mu_1$. Consequently, we get

$$\begin{array}{rcl}\mu_0(A_1+T\,\mathcal{C}) & \geq & \mu_1(A_1)\\ & \geq & \mu_0(A_0+t'\,\mathcal{C})\end{array}$$

where the first inequality follows from (22) and the second one from the definition of A_1. Since $(A_i)_{i\in\mathbb{N}}$ is a non decreasing sequence with respect to the inclusion and by assumption on μ_0, we get (25). Now, we claim that the non decreasing sequence $|(A_i+T\,\mathcal{C})\cap\Omega_0| \to |\Omega_0|$ when i goes to infinity. Otherwise, the sequence $|(A_i+T\,\mathcal{C})\cap\Omega_0|$ would be bounded away from 0 and $|\Omega_0|$, and using the first statement of Proposition 5.2, we would get the existence of a positive number ε' such that for all $i\geq 1$,

$$|(A_i+T\,\mathcal{C}+(t'-T)\,\mathcal{C})\cap\Omega_0| - |(A_i+T\,\mathcal{C})\cap\Omega_0| \geq \varepsilon'.$$

This property combined with (24) clearly contradicts the fact that μ_1 is a probability measure and the claim is proved. With this property in hand, we can now apply the second statement of Proposition 5.2 and obtain that for large i, $|(A_i+t'\,\mathcal{C})\cap\Omega_0| = |\Omega_0|$. This implies that for large i, $\mu_0(A_i+t'\,\mathcal{C})=1$ which in return gives $\mu_1(A_{i+1})=1$ by definition of A_{i+1}. As a consequence,

$$\operatorname{supp}\mu_1 \subset A_{i+1}.$$

Since t' is a supercritical speed, $C(t')<+\infty$. Consequently, the following inclusion holds

$$\Omega_0 \subset \operatorname{supp}\mu_1 + t'\,\mathcal{C}.$$

By combining these last two properties, we get (21).

Using the same approach, we shall show that $\phi^t > -\infty$ on $p_x(\Gamma_0^t)$ hence μ_0-almost everywhere. Let us mention that this last property entails that $\phi^{t^{c_t}}(\operatorname{supp}\mu_1) \subset \mathbb{R}\cup\{-\infty\}$ as required in the definition of c_t-concave map. Once again, consider as a particular case a pair $(x,y)\in\Gamma_0^t$ such that $c_t(x_0,y)<+\infty$. By definition of ϕ^t, we have for any z

$$\phi^t(z) \leq \phi^t(x) + c_t(z,y) - c_t(x,y). \tag{27}$$

Choosing $z=x_0$ in (27) together with the fact that $\phi^t(x_0)=0$ gives the result. To prove the general case, let us define $C_0 = \overline{p_y(p_x^{-1}(\{x\})\cap\Gamma_0^t)}$ (which is a non empty set by assumption on (x,y)) and $C_{i+1} = \overline{p_y(p_x^{-1}(C_i+t'\,\mathcal{C})\cap\Gamma_0^t)}$. The same computations as above give $\mu_1(C_k)=1$ for k large enough, and consequently, $x_0\in C_k+t'\,\mathcal{C}$. This entails $\phi^t(x)>-\infty$, and more precisely

$$\forall x\in\, p_x(\Gamma_0^t),\;\; \phi^t(x) \geq -||c||_\infty(k+1). \tag{28}$$

It remains to check that $\phi^{t^{c_t c_t}} = \phi^t$ and that $\Gamma_0^t \subset (Id,\partial_{c_t}\phi^t)$. The latter property follows easily from (27). Let us prove the other one. We start by renaming y_m as y,

$$\phi^t(x) = \inf_{y\in\operatorname{supp}\mu_1} c_t(x,y) - u(y)$$

where $u(y)=-\infty$ if $y\notin p_y(\Gamma_0^t)$ and

$$u(y) \;=\; \sup_m \sup\{c_t(x_0,y_0)+\cdots+c_t(x_m,y) - (c_t(x_1,y_0)+\cdots+c_t(x_m,y_{m-1}));\; (x_1,y_1),\cdots(x_m,y)\;\in\;\Gamma_0^t\}$$

otherwise.

By (28), $u(p_y(\Gamma_0^t))\subset\mathbb{R}$, that is to say $\phi^t=u^c$ with $u:\operatorname{supp}\mu_1\to\mathbb{R}\cup\{-\infty\}$. Now, for arbitrary admissible map α, we have

$$\alpha \leq \alpha^{cc} \tag{29}$$

by definition of the c-transform. Applying the c-transform again, we get

$$\alpha^c \geq \alpha^{ccc}.$$

On the other hand, (29) applied to $\alpha=u^c$ gives the reverse inequality. Therefore, $u^{ccc}=u^c$ which exactly means $\phi^t=\phi^{t^{cc}}$ as required. □

To state our result concerning the regularity of the Kantorovich potential, we need to recall the notion of approximate differential.

Definition 5.14. We say that $f : X \mapsto \mathbb{R} \cup \{-\infty\}$ has an approximate differential at a point x if $f(x) \in \mathbb{R}$ and there exists a map $g : X \mapsto \mathbb{R}$ differentiable at x such that the set $\{f \neq g\}$ has density 0 at x and $f(x) = g(x)$. The approximate gradient of f at x is denoted by $\tilde{\nabla} f(x)$.

Theorem 5.15. *Under the assumptions of Theorem 5.13, the map* $\phi^t_{|p_x(\{x-y\in t\mathring{\mathcal{C}}\}\cap\Gamma_0^t)}$ *is measurable (up to a modification on a negligible set) and belongs to* $\mathcal{L}^\infty(\Omega_0)$*. Moreover,* ϕ^t *is approximately differentiable Lebesgue almost everywhere on* $p_x(\{x - y \in t\mathring{\mathcal{C}}\} \cap \Gamma_0^t)$.

Proof. First, notice that $|\phi^t|$ is bounded apart from a negligible set thanks to (23) and (28). For any positive integer n, we set

$$\Theta_n = \left\{x \in \overline{\Omega_0}; \exists\, y\,; (x, y) \in \Gamma_0^t \text{ and } x - y \in (1 - 1/n)t\,\mathcal{C}\right\}.$$

Thus, the following equality holds

$$\Theta := p_x(\{x - y \in t\mathring{\mathcal{C}}\} \cap \Gamma_0^t) = \bigcup_{n\in\mathbb{N}\setminus\{0\}} \Theta_n. \tag{30}$$

Note that Θ is measurable as an analytic set. We claim that for any n, $\phi^t|_{\Theta_n}$ is a Lipschitz map. Indeed, for any $x \in \Theta_n$ and any $z \in B(x, \frac{tr_m}{2n}) \cap \Theta_n$ (see (14) for the definition of r_m), let us denote by y a point such that $(x, y) \in \Gamma_0^t$ and $x - y \in (1 - \frac{1}{n})t\,\mathcal{C}$. We get by (27),

$$\phi^t(z) - \phi^t(x) \leq c_t(z, y) - c_t(x, y)$$

which implies the existence of a positive constant $k(n)$

$$\phi^t(z) - \phi^t(x) \leq k(n)|x - z|$$

by assumption on the cost function. Thus ϕ^t is locally Lipschitz and since it is bounded, it is a Lipschitz map on Θ_n.

Now using Kirszbraun's theorem, we set ϕ^t_n a Lipschitz function defined on Θ and which coincides with ϕ^t on Θ_n. For all n, ϕ^t_n is measurable on Θ. Moreover, using (30), we get that ϕ^t_n converges pointwise to ϕ^t on Θ. This implies the measurability of $\phi^t_{|\Theta}$. To conclude the proof, let us remark that if $n \leq m$, then $\Theta_n \subset \Theta_m$, and consequently $\phi^t_n|_{\Theta_n} = \phi^t_m|_{\Theta_n}$. Using Rademacher's Theorem together with Lebegue's density Theorem, we get

$$|A_n| = 0 \text{ and } |\tilde{A}_n| = 0$$

where

$$A_n = \{x \in \Theta; \phi^t_n \text{ is not differentiable at } x\}$$

and

$$\tilde{A}_n = \{x \in \Theta_n; x \text{ is not a point of density 1 of } \Theta_n\}.$$

Note that since $\Theta_n \subset \Theta_{n+1}$ for all n, a point of density 1 of Θ_n is a point of density 1 of Θ as well. We claim that ϕ^t is approximately differentiable on $Z = \Theta \setminus \left(\cup_{n\in\mathbb{N}\setminus\{0\}}(A_n \cup \tilde{A}_n)\right)$.

Indeed, given $x \in Z$, there exists n_0 such that $x \in \Theta_{n_0}$ and $\phi^t_{n_0}$ is differentiable at x. Now since by definition of Z, x is a point of density 1 of Θ_{n_0} and $\Theta_{n_0} \subset \{\phi^t = \phi^t_{n_0}\}$, ϕ^t is approximatively differentiable at x.

□

5.3. **Proof of Theorem 5.1.**

Proof. In this short section, we combine previous results in order to prove Theorem 5.1. Let $(x, y) \in \{x - y \in t\mathring{\mathcal{C}}\} \cap \Gamma_0^t \cap Z \times \mathbb{R}^d$. In particular, we have

$$\phi^t(x) + {\phi^t}^c(y) = c_t(x, y).$$

Therefore by definition of the c-transform and since $c_t(x, y) = h(\frac{x-y}{t})$ with h a strictly convex function, we get for z close to x

$$\phi^t(z) - \phi^t(x) \leq h(\frac{z - y}{t}) - h(\frac{x - y}{t}) < +\infty. \tag{31}$$

Let g be a differentiable map at x such that the set $\{\phi^t = g\}$ has density one at x and $g(x) = \phi^t(x)$. We get that for any $z \in \{\phi^t = g\}$,

$$h(\frac{z - y}{t}) - h(\frac{x - y}{t}) \geq \tilde{\nabla}\phi^t(x) \cdot (z - x) + o(|z - x|)$$

and since $y - x \in t\overset{\circ}{\mathcal{C}}$, h is continuous at $\frac{x-y}{t}$. Therefore, the inequality above yields

$$\tilde{\nabla}\phi^t(x) \in \partial h(\frac{x-y}{t})$$

which gives the result since h is stricly convex. □

Acknowledgements. The authors would like to thank professors J.-B. Hiriart-Urruty, F. Morgan and A. Pratelli for fruitful discussions. M. Puel acknowledges support from the project No. BLAN07-2 212988 entitled "QUATRAIN" and funded by the Agence Nationale de la Recherche. J. Bertrand is partially supported by the ANR through the project Evol.

References

[1] Agueh, Martial. *Existence of solutions to degenerate parabolic equations via the Monge-Kantorovich theory*, Adv. Differential Equations 10 (2005), no. 3, 309–360.

[2] Ambrosio, Luigi; Fusco, Nicola; Pallara, Diego. *Functions of bounded variation and free discontinuity problems*, Oxford Mathematical Monographs, The Clarendon Press Oxford University Press, New York, 2000.

[3] Ambrosio, Luigi. *Steepest descent flows and applications to spaces of probability measures* Lectures notes, Santander, (July 2004).

[4] Ambrosio, Luigi. *Lecture notes on optimal transport problems. Mathematical aspects of evolving interfaces (Funchal, 2000)*, 1–52, Lecture Notes in Math., 1812, Springer, Berlin, (2003).

[5] Ambrosio, Luigi; Gigli, Nicola; Savaré, Giuseppe. *Gradient flows in metric spaces and in the spaces of probability measures* Lectures in Mathematics ETH Zurich, Birkauser Verlag, Basel, (2005).

[6] Ambrosio, Luigi; Kirchheim, Bernd; Pratelli, Aldo.*Existence of optimal transport maps for crystalline norms.* Duke Math. J. 125 (2004), no. 2, 207–241.

[7] Ambrosio, Luigi; Pratelli, Aldo. *Existence and stability results in the L^1 theory of optimal transportation.* Lecture Notes in Math., 1813, 123–160, Springer, Berlin, (2003).

[8] Ambrosio, Luigi; Colesanti, Andrea; Villa, Elena. *Outer Minkowski content for some classes of closed sets*, Math. Ann. **342** (2008), no. 4, 727–748.

[9] Andreu, Fuensanta; Caselles, Vicent; Mazón, José M.*Existence and uniqueness of solution for a parabolic quasilinear problem for linear growth functionals with L^1 data.* Math. Ann., 322, 139–206, (2002).

[10] Andreu, Fuensanta; Caselles, Vicent; Mazón, José M. *Parabolic Quasilinear Equations Minimizing Linear Growth Functionals.* Progress in Mathematics vol 223, Birkhäuser Verlag (2004)

[11] Andreu, Fuensanta; Caselles, Vicent; Mazón, José M.*A strongly degenerate quasilinear equation: the elliptic case.* Ann. Sc. Norm. Super. Pisa Cl. Sci. (5) 3 (2004), no. 3, 555–587.

[12] Andreu, Fuensanta; Caselles, Vicent; Mazón, José M. *A strong degenerate quasilinear equation: the parabolic case* Arch Rational Mech. Anal., (2005).

[13] Andreu, Fuensanta; Caselles, Vicent; Mazón, José M. *A strongly degenerate quasilinear elliptic equation.* Non Linear analysis, 61, 637–669, (2005).

[14] Andreu, Fuensanta; Caselles, Vicent; Mazón, José M.*The Cauchy Problem for a Strong Degenerate Quasilinear Equation.* J. Europ. Math.Soc. **7** (2005), 361–393.

[15] Andreu, Fuensanta; Caselles, Vicent; Mazón, José M.; Moll, Salvador *Finite propagation speed for limited flux diffusion equations.* Arch. Ration. Mech. Anal. 182 (2006), no. 2, 269–297.

[16] Brenier, Yann *Décomposition polaire et réarrangement monotone des champs de vecteurs.* C.R. Acad. Sci. Paris Sér. I Math. 305, 19 (1987), 805–808.

[17] Brenier, Yann *Polar factorization and monotone rearrangement of vector-valued functions.* Comm. Pure Appl. Math. 44, 4 (1991), 375–417.

[18] Brenier, Yann *Extended Monge-Kantorovich theory. Optimal transportation and applications* (Martina Franca, 2001), 91–121, Lecture Notes in Math., 1813, Springer, Berlin, (2003).

[19] Brezis Haim *Opérateurs maximaux monotones et semi-groupes de contractions dans les espaces de Hilbert.* (French) North-Holland Mathematics Studies, No. 5. Notas de Matemàtica (50). North-Holland Publishing Co., Amsterdam-London; American Elsevier Publishing Co., Inc., New York, (1973)

[20] Caffarelli, Luis A; Feldman Mikhail; McCann, Robert J *Constructing optimal maps for Monge's transport problem as a limit of strictly convex costs.* J. Amer. Math. Soc. 15 (2002), no. 1, 1–26.

[21] Caravenna, Laura. *A proof of Sudakov theorem with strictly convex norms.* Math. Z. 268 (2011), no. 1-2, 371–407.

[22] Carlier, Guillaume; De Pascale, Luigi; Santambrogio, Filippo *A Strategy for non-strictly convex transport costs and the example of $||x-y||^p$ in $\mathbb{R}^2$.* Commun. Math. Sci. 8 (2010), no. 4, 931–941.

[23] Caselles, Vicent*Convergence of the "relativistic" heat equation to the heat equation as $c \to \infty$.* Publicaciones Matemàtiques. Vol 51, Nùm. 1.(2007), 121-142.

[24] Champion, Thierry; De Pascale, Luigi *The Monge problem for strictly convex norms in $\mathbb{R}^d$.* J. Eur. Math. Soc. (JEMS) 12 (2010), no. 6, 1355–1369.

[25] Champion, Thierry; De Pascale, Luigi. *The Monge problem in $\mathbb{R}^d$.* Duke Math. J. 157 (2011), no. 3, 551–572.

[26] Champion, Thierry; De Pascale, Luigi; Juutinen, Petri *The ∞-Wasserstein distance: local solutions and existence of optimal transport maps.* SIAM J. Math. Anal. 40 (2008), no. 1, 1–20.

[27] Evans, Lawrence; Gangbo Wilfrid. *Differential equation methods for the Monge kantorovich mass transfer.* Memoirs AMS, 653, (1999).

[28] Gangbo, Wilfrid; McCann, Robert J. *Optimal maps in Monge's mass transport problem.* C. R. Acad. Sci. Paris Sér. I Math. 321 (1995), no. 12, 1653–1658.
[29] Gangbo, Wilfrid; McCann, Robert J. *The geometry of optimal transportation.* Acta Math. 177 (1996), no. 2, 113–161.
[30] Jimenez, Chloé; Santambrogio, Filippo *Optimal transportation for a quadratic cost with convex constraints and applications*, accepted in Journal de Mathématiques Pures et Appliquées.
[31] Jordan, Richard; Kinderlehrer, David; Otto, Felix *The variational formulation of the Fokker-Planck equation.* SIAM J. Math. Anal. 29 (1998), no. 1, 1–17
[32] L. Kantorovitch, *On the translocation of masses*, C. R. (Doklady) Acad. Sci. URSS (N.S.) **37** (1942), 199–201.
[33] McCann, Robert J.*A convexity principle for interacting gases.*Adv. Math. **128** (1997) 153–179.
[34] McCann, Robert and Puel, Marjolaine *Constructing a relativistic heat flow by transport time steps.* Ann. Inst. H. Poincaré Anal. Non Linéaire 26 (2009), no. 6, 2539–2580.
[35] Mihalas D.; Mihalas B. *Foundations of radiation hydrodynamics.* Oxford University Press, (1984).
[36] Otto, Felix *Doubly degenerate diffusion equations as steepest descent* , (Preprint 1996).
[37] Rockafellar R.T. , *Convex analysis*, Princeton Mathematical Series, No. 28, Princeton University Press, Princeton, N.J., 1970.
[38] Rosenau Philip *Tempered diffusion: A Transport Process with Propagating fronts and Initial Delay.* Phys. Review A **46** (1992), 7371–7374.
[39] Rüschendorf, Ludger *On c-optimal random variables* Statistic & probability letters, 27 (1996), 267-270.
[40] Schneider R., *Convex bodies: the Brunn-Minkowski theory*, Encyclopedia of Mathematics and its Applications, vol. 44, Cambridge University Press, Cambridge, 1993.
[41] Cyril S. Smith and Martin Knott. Note on the optimal transportation of distributions. *J. Optim. Theory Appl.*, 52(2):323–329, 1987.
[42] Trudinger, Neil S. and Wang, Xu-Jia. *On the Monge mass transfer problem.* Calc. Var. Partial Differential Equations 13 (2001), no. 1, 19–31.
[43] Villani Cédric, *Topics in Optimal Transportation.* Graduate studies in Math, 58, AMS, (2003).
[44] ———, *Optimal transport*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 338, Springer-Verlag, Berlin, 2009, Old and new.

Institut de Mathématiques de Toulouse, UMR CNRS 5219, Université Toulouse III, 31062 Toulouse cedex 9, France

E-mail address: jerome.bertrand@math.univ-toulouse.fr

Institut de Mathématiques de Toulouse, UMR CNRS 5219, Université Toulouse III, 31062 Toulouse cedex 9, France

E-mail address: puel@math.univ-toulouse.fr