Optimal transportation with traffic congestion and Wardrop equilibria

G. Carlier, C. Jimenez [*], F. Santambrogio [†]

July 24, 2007

Abstract

In the classical Monge-Kantorovich problem, the transportation cost only depends on the amount of mass sent from sources to destinations and not on the paths followed by this mass. Thus, it does not allow for congestion effects. Using the notion of traffic intensity, we propose a variant taking into account congestion. This variant is a continuous version of a well-known traffic problem on networks that is studied both in economics and in operational resarch. The interest of this problem is in its relations with traffic equilibria of Wardrop type. What we prove in the paper is exactly the existence and the variational characterization of equilibria in a continuous space setting.

Keywords: optimal transportation, traffic congestion, Wardrop equilibria.

[*] CEREMADE, UMR CNRS 7534, Université Paris IX Dauphine, Pl. de Lattre de Tassigny, 75775 Paris Cedex 16, FRANCE `carlier@ceremade.dauphine.fr, jimenez@ceremade.dauphine.fr` .

[†] Scuola Normale Superiore, Classe di Scienze, Piazza dei Cavalieri 7, 56126, Pisa, ITALY `f.santambrogio@sns.it`.

1 Introduction

Researchers in the field of applied traffic modelling have long emphasized the role of congestion in networks. In the early 50's, Wardrop (see [14]) considered the situation where a large number of vehicles have to go from one location to another, connected by a finite number of different roads. Each vehicle has to choose one road (or a probability on the set of possible roads) to minimize some transportation cost which depends not only on the road chosen but also on the total flow of vehicles on this road. Some roads may be better than others, for instance because they are shorter or wider, but they are all affected by congestion effects: the "cost" (in terms of time, say) for a vehicle for following a road depends increasingly on the total number of vehicles that choose to use it. Wardrop gave a minimal stability requirement for transportation strategies: the cost of every actually used road should be equal or less than that which would be experienced by a single vehicle on any other road. In particular there is an equilibrium concept (all actually used road have the same cost, i.e. they compensate in term of congestion their differences given by length and other conditions), and a minimality one as well (those roads are minimal among all the possible ones). This natural equilibrium concept has been very popular since its introduction because of applications to networks of course but also due to the development of non-cooperative game theory in the meanwhile. From the best of our knowledge, the study of Wardrop equilibria have mainly been restricted to the case where admissible roads are given by a finite graph. In the present paper our main goal is to introduce Wardrop's concepts in a continuous state setting, to prove the existence of such equilibria and to relate it to the optimal transportation problem with congestion (Theorem 4.2).

The key point of the discrete theory about Wardrop equilibria is the fact that such an equilibrium problem may be linked to a variational one (see for instance [2] and the references therein). Suppose the cost for a road σ is given by $\int_\sigma g(i)$, where i denotes the traffic intensity along the road (it can be non-constant, since at branching points of different roads, the traffic splits according to Kirchhof's law) where $g \geq 0$ is an increasing function modelling congestion effects. We may say that $g(i)$ is the cost per unit length for a road where the traffic intensity is i. It is known that looking for an equilibrium is equivalent to solving a minimization problem with total congestion cost $\int_N H(i)$, where N represents the whole network and H is another increasing function which is linked to g. The relation which is necessary to ensure that minima of the variational problem coincide with solutions of the equilibrium one is $H' = g$. Notice that solving this problem does not amount in general to finding the configuration which minimizes the total cost paid by vehicles, since this quantity is represented instead by the integral $\int_N ig(i)$. The two functions $H(i)$ and $ig(i)$ are the same up to multiplicative constants in the case of power functions but otherwise they give rise to different optimization problems.

The unknown, both in the equilibrium and in the optimization problems, is the distribution of vehicles along the possible paths. In a continuous-setting language, this is a measure on the set of paths. The constraints on such a measure are given by the data: usually the total amount of vehicles commuting from a point x to a point y of the network is prescribed for every pair (x, y). This corresponds to fixing a measure γ on the set of sources-destinations pairs. As a possible alternative, one could prescribe the total quantity of vehicles leaving x and the total quantity reaching y, and look at all possible couplings between these "boundary data". In this case, the fixed data are two measures μ_0 and μ_1 modelling the distribution of sources and destinations separately. And the coupling or *transport plan* γ (i.e. a measure on the sources-destinations space whose projections on the two cooordinates are μ_0 and μ_1) is part of the unknowns.

In the present paper, we introduce a variational problem in a continuous setting, i.e. when the data on sources and destinations are arbitrary probability measures on a domain $\Omega \subset \mathbb{R}^2$ and the allowed paths are all possible Lipschitz curves connecting points of Ω. In this problem, the functional is built as a total cost arising from the congested transport problem. This means that we start from a congestion function g and we look at a minimization problem involving the function $i \mapsto H(i) = ig(i)$. Obviously, the concept of traffic intensity associated to a probability measure on a suitable set of paths has to be carefully defined in the continuous framework, which we do in paragraph 2.2. The definition of traffic intensity we provide is the path-dependent analogue of the well-known notion of *transport density* in Monge's problem (see Bouchitté, Buttazzo and Seppecher [4], Bouchitté and Buttazzo [3], Caffarelli, Feldman and McCann [7]).

After defining the cost and the constraint, we may forget about the origin of the function H and, as in the discrete case, look at the minimization problem: we are concerned with existence, finiteness of the minimum value, and optimality conditions. Interestingly, in the continuous case, this variational problem takes the form of a *path-dependent* optimal transportation problem.

We show, under some extra assumptions on H, that solutions of the variational problem exist and are characterized by two optimality conditions. One of them is nothing but the continuous counterpart of Wardrop equilibria: the quantity $H'(i) \geq 0$ defines to a metric on Ω and the paths that are actually used must be geodesics for this metric. The other optimality condition is peculiar to the case of non-fixed couplings, and is much more linked to Monge-Kantorovich optimal transport theory. In fact, once a solution is found, we can call $c(x,y)$ the minimal cost for commuting from x to y, according to the previously mentioned metric $H'(i)$. It turns out that, for an optimal solution, γ solves the optimal transport problem between μ_0 and μ_1 with respect to the cost c. The case of a non-fixed transport plan is the one which is developed in the paper because it is the richest from a mathematical point of view. Moreover, it allows for a direct comparison with optimal transport problems à la Monge-Kantorovich. A variant of the problem we study is the case where, instead of allowing any transport plan between μ_0 and μ_1, we prescribe a given convex and compact subset of couplings compatible with the data. All the results of this paper can be extended to such a variant, which includes the case of a single prescribed transport plan.

In the minimization problem we study, the functional is linked to a Monge transport cost, in the case of a non-uniform metric. It corresponds to a cost $\int d(x,y)\, d\gamma(x,y)$, where d is a distance which in this case is unknown as it depends on the traffic intensity itself. Other transportation costs, for instance squared distances, are very important for the applications of the Monge-Kantorovich theory, but are not directly linked to this equilibrium problem. Notice anyway that, from an individual point of view (i.e. the viewpoint of the equlibrium issue), minimizing the displacement cost or its square is the same.

The presentation of the model, the construction of the functional, and its links with Monge-Kantorovich theory follow the discrete case in its generality. Yet, most of the results require some additional assumptions on the function H. In particular H is required to behave like a power $H(i) = i^q$ with $1 < q < 2$. This is needed both for technical and feasibility reasons. Firstly, it ensures the validity of some crucial estimates giving continuity results providing well-defined transport costs $c(x,y)$. Secondly, it turns out that if γ is discrete (i.e. any path is allowed but the set of sources and destinations is finite) the condition for having a finite minimal cost is exactly $q < 2$. Moreover, if we come back to the congestion function g, this condition on H corresponds to g behaving like a concave power i^{q-1}, which is very natural from an economic point of view.

2 Optimal transportation with congestion

2.1 Notations

Given a Polish (i.e. metrizable, separable and complete) space X, we will denote respectively by $\mathcal{M}_+(X)$ and $\mathcal{M}^1_+(X)$ the set of positive and finite

Radon measures on X and the set of Radon probability measures on X. If X and Y are Polish spaces, $\mu \in \mathcal{M}_+^1(X)$, and $f : X \to Y$ is a Borel map we shall denote by $f\sharp\mu$ the push forward of μ through f i.e. the element of $\mathcal{M}_+^1(Y)$ defined by $f\sharp\mu(B) = \mu(f^{-1}(B))$ for every Borel subset B of Y.

In the sequel, $\mathcal{L}^d$ denotes the d-dimensional Lebesgue measure. If μ and ν are in $\mathcal{M}_+^1(\mathbb{R}^d)$ then $\frac{d\mu}{d\nu}$ denotes the Radon-Nikodym derivative of μ with respect to ν. We shall write $\mu << \nu$ to express that μ is absolutely continuous with respect to ν, in which case, slightly abusing notations, we will identify μ with the Radon-Nikodym derivative $\frac{d\mu}{d\nu}$.

The data of our problem are Ω (its closure $\overline{\Omega}$ modelling the city, say) which is some open bounded convex subset of $\mathbb{R}^2$, two probability measures, μ_0 and μ_1 in $\mathcal{M}_+^1(\overline{\Omega})$, giving respectively the distribution of residents and services in the city $\overline{\Omega}$. The set of transportation plans associated to μ_0 and μ_1 will be denoted $\Pi(\mu_0, \mu_1)$ it consists of the probability measures on $\overline{\Omega} \times \overline{\Omega}$ having μ_0 and μ_1 as marginals:

$$\Pi(\mu_0, \mu_1) := \{\gamma \in \mathcal{M}_+^1(\overline{\Omega} \times \overline{\Omega}) : \pi_0\sharp\gamma = \mu_0, \ \pi_1\sharp\gamma = \mu_1\} \tag{2.1}$$

where $(\pi_0(x, y), \pi_1(x, y)) := (x, y)$, stand for the canonical projections (x and y in $\overline{\Omega}$).

Introducing congestion naturally leads to consider spaces of paths, lengths of such paths and sets of probability measures on sets of paths. From now on, we shall denote:

- $C := W^{1,\infty}([0, 1], \overline{\Omega})$, viewed as a subset of $C^0([0, 1], \mathbb{R}^2)$, i.e. equipped with the uniform topology,
- $C^{x,y} := \{\sigma \in C : \ \sigma(0) = x, \ \sigma(1) = y\}$ ($x, \ y$ in $\overline{\Omega}$),
- $l(\sigma) := \int_0^1 |\dot{\sigma}(t)| \, dt$, the length of $\sigma \in C$,
- for $\sigma \in C$, $\widetilde{\sigma}$ denotes the constant speed reparameterization of σ belonging to C, hence $|\dot{\widetilde{\sigma}}(t)| = l(\sigma) = l(\widetilde{\sigma})$ for a.e. $t \in [0, 1]$,
- $\widetilde{C} := \{\sigma \in C \ : \ |\dot{\sigma}| \text{ is constant}\} = \{\widetilde{\sigma}, \ \sigma \in C\}$,
- slightly abusing notations, we will denote $Q \in \mathcal{M}_+^1(C)$, whenever $Q \in \mathcal{M}_+^1(C^0([0, 1], \mathbb{R}^2))$ and $Q(C) = 1$,
- for $Q \in \mathcal{M}_+^1(C)$, we define $\widetilde{Q} \in \mathcal{M}_+^1(\widetilde{C})$ as the push forward of Q through the map $\sigma \mapsto \widetilde{\sigma}$,

- for $\varphi \in C^0(\overline{\Omega}, \mathbb{R})$ and $\sigma \in C$, we define

$$L_\varphi(\sigma) := \int_0^1 \varphi(\sigma(t))|\dot{\sigma}(t)|dt = l(\sigma) \int_0^1 \varphi(\widetilde{\sigma}(t))dt,$$

- $e_0(\sigma) := \sigma(0)$, $e_1(\sigma) := \sigma(1)$, for all $\sigma \in C^0([0,1], \mathbb{R}^2)$.

2.2 Traffic congestion modelling

The classical Monge-Kantorovich optimal transportation problem for a given cost function $c \in C^0(\overline{\Omega} \times \overline{\Omega}, \mathbb{R})$ is:

$$\inf \left\{ \int_{\overline{\Omega}\times\overline{\Omega}} c(x,y)d\gamma(x,y) \ : \ \gamma \in \Pi(\mu_0, \mu_1) \right\}. \tag{2.2}$$

Note that, in the linear problem (2.2), the cost of transporting one unit of mass from x to y, $c(x,y)$, is given and does not depend on the path(s) followed by the mass from x to y. In order to take into account congestion effects, we explicitely introduce probabilities over $C^{x,y}$ as part of the optimization problem. More precisely, the overall transportation cost will depend not only on the transportation plan $\gamma \in \Pi(\mu_0, \mu_1)$ but also on the way travelers commuting from x to y use the different possible paths $\sigma \in C^{x,y}$. In the sequel, the way commuters from x to y are split according to the different paths will be given by a probability measure $p^{x,y}$ on $C^{x,y}$. Put differently, $p^{x,y}(\Sigma)$ is the proportion of travelers from x to y using a path $\sigma \in \Sigma \subset C^{x,y}$. This naturally leads to the following definition:

Definition 2.1. *A transportation strategy consists of a pair (γ, p) with $\gamma \in \Pi(\mu_0, \mu_1)$ and where $p = (p^{x,y})_{(x,y)\in\overline{\Omega}\times\overline{\Omega}}$ is a Borel family of probability measures on C (i.e. $(x,y) \to \int_C F(\sigma)dp^{x,y}(\sigma)$ is Borel for every bounded Borel function $F\colon C \to \mathbb{R}$) such that $p^{x,y}(C^{x,y}) = 1$ for γ-a.e. $(x,y) \in \overline{\Omega} \times \overline{\Omega}$.*

Thanks to Lemma 2.7 proved below, for every $\varphi \in C^0(\overline{\Omega}, \mathbb{R}_+)$, the map $\sigma \to L_\varphi(\sigma)$ is l.s.c. hence Borel on C (equipped with the uniform topology). For $Q \in \mathcal{M}_+^1(C)$ one can therefore define $\int_C L_\varphi(\sigma)dQ(\sigma)$ and this integral is finite for every $\varphi \in C^0(\overline{\Omega}, \mathbb{R}_+)$ whenever $\int_C l(\sigma)dQ(\sigma) < +\infty$ i.e. the average for the probability Q length is finite. Hence, there results, from the use of a transportation strategy (γ, p), an overall *traffic intensity* $I_{\gamma,p} \in \mathcal{M}_+(\overline{\Omega})$

defined by

$$\int_{\overline{\Omega}} \varphi(x) dI_{\gamma,p}(x) := \int_{\overline{\Omega}\times\overline{\Omega}} \left(\int_{C^{x,y}} \left(\int_0^1 \varphi(\sigma(t))|\dot{\sigma}(t)|dt \right) dp^{x,y}(\sigma) \right) d\gamma(x,y)$$
$$= \int_{\overline{\Omega}\times\overline{\Omega}} \left(\int_{C^{x,y}} L_\varphi(\sigma) dp^{x,y}(\sigma) \right) d\gamma(x,y), \quad \forall \varphi \in C^0(\overline{\Omega}, \mathbb{R}_+) \tag{2.3}$$

and an overall probability over paths $Q_{\gamma,p} \in \mathcal{M}^1_+(C)$ given by $Q_{\gamma,p} = p^{x,y} \otimes \gamma$, i.e.:

$$\int_C F(\sigma) dQ_{\gamma,p}(\sigma) = \int_{\overline{\Omega}\times\overline{\Omega}} \left(\int_{C^{x,y}} F(\sigma) dp^{x,y}(\sigma) \right) d\gamma(x,y) \quad \forall F \in C^0(C, \mathbb{R}). \tag{2.4}$$

One could consider the probability $Q_{\gamma,p}$ as if it represented the total number of travelers that use a path $\sigma \in \Sigma$ given the global transportation strategy (γ, p).

Let us remark that if we set $Q := Q_{\gamma,p} \in \mathcal{M}^1_+(C)$ then $I_{\gamma,p}$ only depends on Q, and can be written as $I_{\gamma,p} = i_Q \in \mathcal{M}_+(\overline{\Omega})$ where i_Q is defined for every $Q \in \mathcal{M}^1_+(C)$ by:

$$\int_{\overline{\Omega}} \varphi(x) di_Q(x) = \int_C L_\varphi(\sigma) dQ(\sigma), \quad \forall \varphi \in C^0(\overline{\Omega}, \mathbb{R}_+). \tag{2.5}$$

Let us also remark that since $L_\varphi(\sigma) = L_\varphi(\widetilde{\sigma})$ one has $i_Q = i_{\widetilde{Q}}$, for all $Q \in \mathcal{M}^1_+(C)$. Finally, let us note that the total mass of i_Q is the average length with respect to Q:

$$i_Q(\overline{\Omega}) = \int_C l(\sigma) dQ(\sigma). \tag{2.6}$$

If the probability Q is concentrated on injective curves, one can also express the measure i_Q through $\mathcal{H}^1-$integrals as explained in the next remark.

Remark 2.2. If a curve σ is injective, then one could also write $l(\sigma) = \mathcal{H}^1(\sigma([0,1]))$ and $L_\varphi(\sigma) = \int_{\sigma([0,1])} \varphi \, d\mathcal{H}^1$. Moreover, if for $\gamma-$a.e. (x,y) the probability $p^{x,y}$ is concentrated on the set of injectives curves from x to y, one could also define the measure $I_{\gamma,p}$ by replacing the integral with respect to $|\dot{\sigma}(t)|dt$ in (2.3) with an integral in $d\mathcal{H}^1$. Notice moreover that for every Borel subset $A \subset \overline{\Omega}$ one would have:

$$I_{\gamma,p}(A) = \int_{\overline{\Omega}\times\overline{\Omega}} \left(\int_{C^{x,y}} \mathcal{H}^1(A \cap \sigma) dp^{x,y}(\sigma) \right) d\gamma(x,y) = \int_C \mathcal{H}^1(A \cap \sigma) dQ_{\gamma,p}(\sigma).$$

If we imagine that for each $\sigma \in C^{x,y}$, the mass of travelers commuting on σ is uniformly distributed on σ, this means that $I_{\gamma,p}(A)$ represents the cumulative traffic through the region A. The same formula remains true, under no injectivity assumption, if we replace $\mathcal{H}^1(A\cap\sigma)$ with $L_{\mathbf{1}_A}(\sigma)$ and in this case the cumulative traffic takes into account the number of times a path σ passes through the points of A.

In the sequel, it will be convenient to formulate our optimization problem in terms of $Q = Q_{\gamma,p}$ rather than in the transportation strategy (γ, p). To that end, we shall use the following:

Lemma 2.3. *Let us define*

$$\mathcal{Q}(\mu_0,\mu_1) := \{Q_{\gamma,p}\ :\ (\gamma,p)\ \textit{transportation strategy}\}$$

then one has

$$\mathcal{Q}(\mu_0,\mu_1) = \{Q\in\mathcal{M}^1_+(C)\ :\ e_0\sharp Q=\mu_0,\ e_1\sharp Q=\mu_1\}.$$

Proof. If (γ, p) is a transportation strategy then $e_0\sharp Q_{\gamma,p} = \pi_0\sharp\gamma = \mu_0$, and $e_1\sharp Q_{\gamma,p} = \pi_1\sharp\gamma = \mu_1$. Now let $Q\in\mathcal{M}^1_+(C)$ be such that $e_0\sharp Q=\mu_0$, $e_1\sharp Q = \mu_1$. If we define $\gamma := (e_0, e_1)\sharp Q$, we have $\gamma\in\Pi(\mu_0,\mu_1)$. It then follows from the disintegration theorem (see [9]) that there exists $p = (p^{x,y})_{(x,y)\in\overline{\Omega}\times\overline{\Omega}}$ a Borel family of probability measures on C such that $p^{x,y}(C^{x,y}) = 1$ for γ-a.e. $(x, y)\in\overline{\Omega}\times\overline{\Omega}$ and $Q = p^{x,y}\otimes\gamma$. Hence $Q = Q_{\gamma,p}$ for a transportation strategy (γ, p). □

At this point, a natural way to model traffic congestion is, for a given transportation strategy (γ, p), to consider that the transportation cost per unit of mass between x and y is given by

$$c_{\gamma,p}(x,y) = \int_{C^{x,y}} L_{G_{I_{\gamma,p}}}(\sigma)dp^{x,y}(\sigma) \tag{2.7}$$

where $G_{I_{\gamma,p}}$ is a nonnegative function which depends (in a way that will be specified later on) on the traffic intensity $I_{\gamma,p}$. The optimal transportation with traffic congestion then takes the form (to be compared with the usual Monge-Kantorovich problem (2.2)):

$$\inf\left\{\int_{\overline{\Omega}\times\overline{\Omega}} c_{\gamma,p}(x,y)d\gamma(x,y)\ :\ (\gamma,p)\text{ transportation strategy}\right\}. \tag{2.8}$$

Setting $Q = Q_{\gamma,p}$ and using formally (2.5), we see that the total transportation cost in (2.8) can be rewritten as:

$$\int_{\overline{\Omega}\times\overline{\Omega}} c_{\gamma,p}(x,y)d\gamma(x,y) = \int_C L_{G_{i_Q}}(\sigma)dQ(\sigma) = \int_{\overline{\Omega}} G_{i_Q}(x)di_Q(x).$$

Hence using lemma 2.3, we can reformulate (2.8) in terms of Q only:

$$\inf\left\{\int_{\overline{\Omega}} G_{i_Q}(x)di_Q(x)\ :\ Q\in\mathcal{Q}(\mu_0,\mu_1)\right\}. \tag{2.9}$$

Note that in the definition (2.7), it is required that $G_{I_{\gamma,p}}$ is continuous (or at least l.s.c.) whereas the form (2.9) allows for more general forms of congestion through $i\mapsto G_i$. From now on, we assume that G has the following local form:

$$G_i(x)=g\left(\frac{di}{d\mathcal{L}^2}(x)\right), \tag{2.10}$$

where $\frac{di}{d\mathcal{L}^2}$ is the Radon-Nikodym derivative of i with respect to the Lebesgue measure and g is a nondecreasing function $\mathbb{R}_+\to\mathbb{R}_+$ such that the function H defined by $H(z)=zg(z)$ for all $z\in\mathbb{R}_+$ is convex and superlinear (i.e. $\lim_{z\to+\infty} g(z)=+\infty$).

The optimization problem we shall study now then reads as:

$$\inf_{Q\in\mathcal{Q}(\mu_0,\mu_1)}\mathcal{H}(i_Q)\ \text{where}\ \ \mathcal{H}(i)=\begin{cases}\int_\Omega H(i(x))dx \text{ if } i<<\mathcal{L}^2,\\ +\infty \text{ otherwise.}\end{cases} \tag{2.11}$$

In the sequel, we shall say that a transportation strategy (γ,p) is optimal if $Q_{\gamma,p}$ solves (2.11).

Remark 2.4. It will be clear in the sequel that the probability $Q_{\gamma,p}$ associated to an optimal transportation strategy (γ,p) will be concentrated on injective curves, so that the interpretation in terms of $\mathcal{H}^1-$integrals (see Remark 2.2) may apply.

2.3 Existence of minimizers

From now on, we make the following assumptions:

- H is convex and nondecreasing on $\mathbb{R}_+$ with $H(0)=0$,
- there exists $q>1$, and positive constants a and b such that $az^q\leq H(z)\leq b(z^q+1)$ for all $z\in\mathbb{R}_+$,
- H is differentiable on $\mathbb{R}_+$, and there exists a positive constant c such that $0\leq H'(z)\leq c(z^{q-1}+1)$, for all $z\in\mathbb{R}_+$,
- the following set

  $$\mathcal{Q}^q(\mu_0,\mu_1):=\{Q\in\mathcal{Q}(\mu_0,\mu_1)\ :\ i_Q\in L^q\} \tag{2.12}$$

  is nonempty (in the definition of $\mathcal{Q}^q(\mu_0,\mu_1)$ we intend of course both $i_Q<<\mathcal{L}^2$ and $\frac{di_Q}{d\mathcal{L}^2}\in L^q$).

These assumptions enable us to simply rewrite (2.11) as:

$$\inf_{Q \in \mathcal{Q}^q(\mu_0,\mu_1)} \int_\Omega H(i_Q(x))dx. \tag{2.13}$$

Remark 2.5. Let us discuss the assumption that $\mathcal{Q}^q(\mu_0,\mu_1) \neq \emptyset$ which, at first glance, may seem difficult to check. In order to have the existence of a $Q \in \mathcal{Q}(\mu_0,\mu_1)$ such that $i_Q \in L^q$ it is sufficient that μ_0 and μ_1 are in L^q. This result, which is not obvious, follows from the regularity results of De Pascale and Pratelli (see [10] and [11]) who proved that L^q regularity of μ_0 and μ_1 implies that for γ solving the Monge-Kantorovich problem (2.2) with $c(x,y) = |x-y|$ and $p^{x,y} = \delta_{[x,y]}$ (the Dirac mass at the segment $[x,y]$) for every x and y the corresponding traffic density $I_{\gamma,p}$ is L^q.

Remark 2.6. It is not necessary however that μ_0 and μ_1 are absolutely continuous for the assumption to be satisfied. Indeed, in the discrete case i.e. when μ_0 and μ_1 have finite support, one can easily prove that $\mathcal{Q}^q(\mu_0,\mu_1) \neq \emptyset$ for every $q \in (1,2)$. Finally, let us consider for instance the case where $\overline{\Omega} = [0,1]^2$ and μ_0 and μ_1 are respectively the one-dimensional Hausdorff measures of the segments $[(0,0),(0,1)]$ and $[(1,0),(1,1)]$. If we define $\gamma := (\mathrm{id}, \mathrm{id} + (1,0))\sharp\mu_0$ and $p^{x,y} = \delta_{[x,y]}$ then a straightforward computation shows that $I_{\gamma,p}$ is uniform on $[0,1]^2$.

Under the assumptions above, we are going prove that (2.13) admits a solution. The proof of existence involves some preliminary lemmas.

Lemma 2.7. *For any $\varphi \in C^0(\overline{\Omega}, \mathbb{R}_+)$, L_φ is lower semi-continuous on C for the uniform topology, indeed for any $\sigma \in C$, one has:*

$$L_\varphi(\sigma) = \sup\Big\{ \sum_{i=1}^n \Big(\inf_{[t_i,t_{i+1}]} (\varphi\circ\sigma) \Big) |\sigma(t_{i+1}) - \sigma(t_i)| : \\ ([t_i,t_{i+1}])_i \text{ is a subdivision of } [0,1] \Big\}. \tag{2.14}$$

Proof. For any subdivision $([t_i,t_{i+1}])_{i=1,\dots n}$, we have:

$$\begin{aligned} L_\varphi(\sigma) &= \sum_{i=1}^n \int_{t_i}^{t_{i+1}} \varphi(\sigma(t))|\dot\sigma(t)|\, dt \\ &\geq \sum_{i=1}^n \inf_{[t_i,t_{i+1}]} (\varphi\circ\sigma) \int_{t_i}^{t_{i+1}} |\dot\sigma(t)|\, dt \\ &\geq \sum_{i=1}^n \inf_{[t_i,t_{i+1}]} (\varphi\circ\sigma) |\sigma(t_{i+1}) - \sigma(t_i)|. \end{aligned}$$

Taking the supremum over all such divisions, we get:

$$L_\varphi(\sigma) \geq \sup\Big\{ \sum_{i=1}^{n} \inf_{[t_i,t_{i+1}]}(\varphi\circ\sigma)|\sigma(t_{i+1})-\sigma(t_i)| : ([t_i,t_{i+1}])_i \text{ is a subdivision of } [0,1]\Big\}.$$

Let us prove the converse inequality. Let $\varepsilon > 0$, since $\varphi \circ \sigma$ is uniformly continuous, there is a $\delta > 0$ such that:

$$\forall t, t' \in [0,1]^2, \quad (|t-t'| \leq \delta \Rightarrow |\varphi(\sigma(t)) - \varphi(\sigma(t'))| \leq \varepsilon)\,.$$

For any subdivision $([t_i,t_{i+1}])_{i=1,\ldots n}$ such that $|t_i - t_{i+1}| \leq \delta$ for all i, we have:

$$\begin{aligned}
L_\varphi(\sigma) &\leq \sum_{i=1}^{n} (\inf_{[t_i,t_{i+1}]}(\varphi\circ\sigma)+\varepsilon)\int_{t_i}^{t_{i+1}} |\dot\sigma(t)|\, dt \\
&= \sum_{i=1}^{n} (\inf_{[t_i,t_{i+1}]}(\varphi\circ\sigma)+\varepsilon) \sup\Big\{ \sum_j |\sigma(\tau_j)-\sigma(\tau_{j+1})| : \\
&\qquad\qquad ([\tau_j,\tau_{j+1}])_j \text{ is a subdivision of } [t_i,t_{i+1}]\Big\} \\
&\leq \sup\Big\{ \sum_i \sum_j (\inf_{[\tau_j,\tau_{j+1}]}(\varphi\circ\sigma)+\varepsilon)|\sigma(\tau_j)-\sigma(\tau_{j+1})| : \\
&\qquad\qquad ([\tau_j,\tau_{j+1}])_j \text{ is a subdivision of } [t_i,t_{i+1}]\Big\} \\
&= \sup\Big\{ \sum_{i=1}^{n} (\inf_{t\in[t_i,t_{i+1}]}(\varphi\circ\sigma)+\varepsilon)|\sigma(t_{i+1})-\sigma(t_i)| : \\
&\qquad\qquad ([t_i,t_{i+1}])_i \text{ is a subdivision of } [0,1]\Big\}.
\end{aligned}$$

As this last inequality is true for any $\varepsilon > 0$ we get (2.14). The lower semi-continuity is then obvious since, by (2.14), L_φ is the supremum of family of lower semi-continuous functions on $C^0([0,1],\overline{\Omega})$. ☐

Lemma 2.8. *Let $(Q_n)_n \in \mathcal{M}^1_+(C^0([0,1],\mathbb{R}^2))^{\mathbb{N}}$ such that $Q_n(C) = 1$ for all n and there exists a constant $M > 0$ such that:*

$$\sup_n \int_C l(\sigma)\, dQ_n(\sigma) \leq M.$$

Then the sequence $(\widetilde{Q}_n)_n$ is tight and admits a subsequence that converges weakly $*$ to a probability Q such that $Q(C) = 1$.

Proof. The tightness of $(\widetilde{Q}_n)_n$ easily follows from the inequality:

$$\begin{aligned}
\widetilde{Q}_n\left(\{\sigma \in \widetilde{C} :\ |\dot\sigma| > K\}\right) &= Q_n\left(\{\sigma \in C :\ l(\sigma) > K\}\right) \\
&\leq \frac{1}{K}\int_C l(\sigma)\, dQ_n(\sigma). \qquad (2.15)
\end{aligned}$$

By Prokhorov theorem, we may therefore assume, passing to a subsequence if necessary, that $(\widetilde{Q}_n)_n$ converges weakly $*$ to $Q \in \mathcal{M}^1_+(C^0([0,1],\mathbb{R}^2)$. It remains to show that $Q(C) = 1$. For $K > 0$ let us define $C_K := \{\sigma \in C \; : \; |\dot{\sigma}| \leq K\}$, then inequality (2.15) and the fact that the measures $\widetilde{Q}_n$ are concentrated on $\widetilde{C}$ yield

$$\sup_n \widetilde{Q}_n(C \backslash C_K) = \sup_n \widetilde{Q}_n(\widetilde{C} \backslash C_K) \leq \frac{M}{K},$$

for every $K > 0$, which implies

$$\begin{aligned} 1 = \limsup_n \widetilde{Q}_n(C) &\leq \limsup_n \widetilde{Q}_n(C_K) + \limsup_n \widetilde{Q}_n(C \setminus C_K) \\ &\leq Q(C_K) + \frac{M}{K}. \end{aligned}$$

Letting K tend to ∞, we then get $Q(C) = \sup_K Q(C_K) = 1$. □

Lemma 2.9. *Let $(Q_n)_n$ be a sequence in $\mathcal{M}^1_+(C)$ that converges weakly $*$ to some $Q \in \mathcal{M}^1_+(C)$. If there exists $i \in \mathcal{M}_+(\overline{\Omega})$ such that i_{Q_n} converges weakly $*$ to i in $\mathcal{M}_+(\overline{\Omega})$ then we have $i_Q \leq i$.*

Proof. Let $\varphi \in C^0(\overline{\Omega}, \mathbb{R}_+)$, we first have:

$$\int_{\overline{\Omega}} \varphi di = \lim_n \int_{\overline{\Omega}} \varphi di_{Q_n} = \lim_n \int_C L_\varphi dQ_n$$

it easily follows from lemma 2.7 that $Q \mapsto \int_C L_\varphi dQ$ is l.s.c. for the weak $*$ topology of $\mathcal{M}^1_+(C)$, we then have:

$$\int_{\overline{\Omega}} \varphi di \geq \int_C L_\varphi dQ = \int_{\overline{\Omega}} \varphi di_Q.$$ □

Now, we are in position to prove:

Theorem 2.10. *The minimization problem (2.13) admits a solution.*

Proof. Our assumptions imply that the value of (2.13) is finite. Let $(Q_n)_n$ be some minimizing sequence of (2.13). From the identity $i_Q = i_{\widetilde{Q}}$, we may assume $Q_n = \widetilde{Q}_n$ for all n. We deduce from our growth condition on H, that $(i_{Q_n})_n$ is bounded in L^q. On the one hand, extracting a subsequence if necessary, we may therefore assume that $(i_{Q_n})_n$ converges weakly in L^q to some i. On the other hand, since i_{Q_n} is bounded in L^q and hence in L^1 we have

$$\sup_n \int_C l(\sigma) dQ_n(\sigma) = \sup_n \int_\Omega i_{Q_n} < +\infty.$$

Moreover $Q_n = \widetilde{Q}_n$ and we deduce from lemma 2.8 that (up to some subsequence) $(Q_n)_n$ weakly $*$ converges to some Q in $\mathcal{M}^1_+(C)$. Since $\mathcal{Q}(\mu_0, \mu_1)$ is obviously weakly $*$ closed, we have $Q \in \mathcal{Q}(\mu_0, \mu_1)$ and lemma 2.9 implies that $i_Q \leq i$ (consequently to this inequality i_Q is absolutely continuous). From the monotonicity and convexity of H we then have:

$$\int_\Omega H(i_Q(x))dx \leq \int_\Omega H(i(x))dx \leq \liminf_n \int_\Omega H(i_{Q_n}(x))dx$$

which proves that Q solves (2.13). $\square$

Let us remark that if H is furthermore assumed to be strictly convex then if Q_1 and Q_2 solves (2.13) then $i_{Q_1} = i_{Q_2}$ so that the optimal traffic intensity is unique (of course, this does not imply in general that $Q_1 = Q_2$ or that the corresponding optimal transportation strategy is unique).

3 Characterization of the minimizers

In the sequel, we shall denote by q^* the conjugate exponent of q, given by $q^* = q/(q-1)$.

3.1 Optimality conditions

The variational inequalities characterizing solutions of the convex problem (2.13) can be expressed as follows

Proposition 3.1. *$\overline{Q} \in \mathcal{Q}^q(\mu_0, \mu_1)$ solves (2.13) if and only if*

$$\int_\Omega \overline{\xi} i_{\overline{Q}} = \inf \left\{ \int_\Omega \overline{\xi} i_Q \ : \ Q \in \mathcal{Q}^q(\mu_0, \mu_1) \right\} \ \text{with } \overline{\xi} := H'(i_{\overline{Q}}) \in L^{q^*}. \quad (3.1)$$

Proof. Assume that $\overline{Q}$ solves (2.13), then for every $Q \in \mathcal{Q}^q(\mu_0, \mu_1)$, one has:

$$\begin{aligned} 0 &\leq \lim_{\varepsilon \to 0^+} \frac{1}{\varepsilon}[\mathcal{H}(i_{\overline{Q}+\varepsilon(Q-\overline{Q})}) - \mathcal{H}(i_{\overline{Q}})] = \lim_{\varepsilon \to 0^+} \frac{1}{\varepsilon}[\mathcal{H}(i_{\overline{Q}} + \varepsilon(i_Q - i_{\overline{Q}})) - \mathcal{H}(i_{\overline{Q}})] \\ &= \int_\Omega H'(i_{\overline{Q}})(i_Q - i_{\overline{Q}}) = \int_\Omega \overline{\xi}(i_Q - i_{\overline{Q}}) \end{aligned}$$

which proves (3.1). Conversely, if $\overline{Q} \in \mathcal{Q}^q(\mu_0, \mu_1)$ satisfies (3.1), then by convexity of H, for every every $Q \in \mathcal{Q}^q(\mu_0, \mu_1)$, one has:

$$\mathcal{H}(i_Q) - \mathcal{H}(i_{\overline{Q}}) \geq \int_\Omega \overline{\xi}(i_Q - i_{\overline{Q}}) \geq 0. \qquad \square$$

The next paragraphs will be devoted to investigate the precise meaning of (3.1). Before going further, let us do some formal manipulations to give a formal interpretation of (3.1) in terms of optimal transportation strategy. Let us assume that $\overline{Q}$ solves (2.13) and let us write $\overline{Q} = Q_{\overline{\gamma},\overline{p}}$ for some (optimal) transportation strategy $(\overline{\gamma}, \overline{p})$ and define $\overline{\xi} := H'(i_{\overline{Q}})$, then (3.1) formally can be rewritten as:

$$\begin{aligned}
\int_{\Omega} \overline{\xi} i_{\overline{Q}} &= \int_C L_{\overline{\xi}}(\sigma) d\overline{Q}(\sigma) \\
&= \int_{\overline{\Omega}\times\overline{\Omega}} \left(\int_{C^{x,y}} L_{\overline{\xi}}(\sigma) d\overline{p}^{x,y}(\sigma) \right) d\overline{\gamma}(x,y) \\
&= \inf_{(\gamma,p) \text{ transp. strategy}} \int_{\overline{\Omega}\times\overline{\Omega}} \left(\int_{C^{x,y}} L_{\overline{\xi}}(\sigma) dp^{x,y}(\sigma) \right) d\gamma(x,y) \\
&= \inf_{\gamma\in\Pi(\mu_0,\mu_1)} \int_{\overline{\Omega}\times\overline{\Omega}} \left(\inf_{p\in\mathcal{M}_+^1(C^{x,y})} \int_{C^{x,y}} L_{\overline{\xi}}(\sigma) dp(\sigma) \right) d\gamma(x,y) \\
&= \inf_{\gamma\in\Pi(\mu_0,\mu_1)} \int_{\overline{\Omega}\times\overline{\Omega}} \left(\inf_{\sigma\in C^{x,y}} L_{\overline{\xi}}(\sigma) \right) d\gamma(x,y)
\end{aligned}$$

defining (again formally) the transportation cost:

$$c_{\overline{\xi}}(x,y) = \inf_{\sigma\in C^{x,y}} L_{\overline{\xi}}(\sigma),$$

we then firstly have:

$$\int_{\overline{\Omega}\times\overline{\Omega}} c_{\overline{\xi}}(x,y) d\overline{\gamma}(x,y) \le \int_C L_{\overline{\xi}} d\overline{Q} = \inf_{\gamma\in\Pi(\mu_0,\mu_1)} \int_{\overline{\Omega}\times\overline{\Omega}} c_{\overline{\xi}}(x,y) d\gamma(x,y)$$

so that $\overline{\gamma}$ solves the Monge-Kantorovich problem:

$$\inf_{\gamma\in\Pi(\mu_0,\mu_1)} \int_{\overline{\Omega}\times\overline{\Omega}} c_{\overline{\xi}}(x,y) d\gamma(x,y).$$

Secondly:

$$\begin{aligned}
\int_C L_{\overline{\xi}}(\sigma) d\overline{Q}(\sigma) &= \int_{\overline{\Omega}\times\overline{\Omega}} c_{\overline{\xi}}(x,y) d\overline{\gamma}(x,y) \\
&= \int_C c_{\overline{\xi}}(\sigma(0),\sigma(1)) d\overline{Q}(\sigma)
\end{aligned}$$

and since $L_{\overline{\xi}}(\sigma) \ge c_{\overline{\xi}}(\sigma(0),\sigma(1))$, we get

$$L_{\overline{\xi}}(\sigma) = c_{\overline{\xi}}(\sigma(0),\sigma(1)) \quad \text{for } \overline{Q}\text{-a.e. } \sigma.$$

or, in an equivalent way, for $\overline{\gamma}$-a.e. (x, y) one has:

$$L_{\overline{\xi}}(\sigma) = c_{\overline{\xi}}(x, y) \qquad \text{for } \overline{p}^{x,y}\text{-a.e. } \sigma.$$

Since $\overline{\xi}$ is only L^{q^*}, $L_{\overline{\xi}}$ and $c_{\overline{\xi}}$ are not well-defined and the previous arguments are purely formal. In paragraph 3.2, we will extend the definition of c_ξ to the case where ξ is only L^{q^*} under the additional assumption $q < 2$. This will enable us to make the formal argument above rigorous and to characterize optimal transportation strategies in paragraph 3.3.

3.2 The transportation cost $\overline{c}_\xi$ when ξ is L^{q^*}

For a non-negative function $\xi \in C^0(\overline{\Omega}, \mathbb{R}_+)$ we define

$$c_\xi(x, y) = \inf\{L_\xi(\sigma) \, : \, \sigma \in C^{x,y}\}.$$

Proposition 3.2. *Let us assume that $q < 2$ and define $\alpha := 1 - 2/q^*$, then there exists a non-negative constant C such that for every $\xi \in C^0(\overline{\Omega}, \mathbb{R}_+)$ and every $(x_1, y_1, x_2, y_2) \in \Omega^4$, one has:*

$$|c_\xi(x_1, y_1) - c_\xi(x_2, y_2)| \le C\|\xi\|_{L^{q^*}(\Omega)} \left(|x_1 - x_2|^\alpha + |y_1 - y_2|^\alpha\right). \qquad (3.2)$$

Consequently, if $(\xi_n)_n \in C^0(\overline{\Omega}, \mathbb{R}_+)^{\mathbb{N}}$ is bounded in L^{q^*}, then $(c_{\xi_n})_n$ admits a subsequence that converges in $C^0(\overline{\Omega} \times \overline{\Omega}, \mathbb{R}_+)$.

Proof. Let $\xi \in C^0(\overline{\Omega}, \mathbb{R}_+)$ and $x, \ y \in \Omega^2$. For $k > 0$ let $\sigma_k \in C^{x,y}$ be such that

$$\int_0^1 \xi(\sigma_k(t))|\dot{\sigma}_k(t)| \, dt \le c_\xi(x, y) + \frac{1}{k}.$$

Then for all $\varepsilon > 0$ and z such that $y + \varepsilon z \in \Omega$ and $t_0 \in (0, 1)$ we consider the following element of $C^{x,y+\varepsilon z}$:

$$\sigma_{k,t_0}(t) := \begin{cases} \sigma_k\left(\frac{t}{t_0}\right) & \text{if } t \in [0, t_0] \\ y + \left(\frac{t-t_0}{1-t_0}\right)\varepsilon z & \text{if } t \in [t_0, 1]. \end{cases}$$

We then have, for all $k > 0$:

$$\begin{aligned} c_\xi(x, y + \varepsilon z) &\le \int_0^1 \xi(\sigma_{k,t_0}(t))|\dot{\sigma}_{k,t_0}(t)| \, dt \\ &= \int_0^1 \xi(\sigma_k(t))|\dot{\sigma}_k(t)| \, dt + \int_0^1 \xi(y + t\varepsilon z)\varepsilon|z| \, dt \\ &\le c_\xi(x, y) + \varepsilon|z| \int_0^1 \xi(y + t\varepsilon z)dt + \frac{1}{k}. \end{aligned}$$

Now we let k tend to ∞ and we get

$$\frac{c_\xi(x, y+\varepsilon z) - c_\xi(x,y)}{\varepsilon} \leq |z| \int_0^1 \xi(y + t\varepsilon z)\ dt,$$

and by a similar argument

$$\frac{c_\xi(x,y) - c_\xi(x, y+\varepsilon z)}{\varepsilon} \leq |z| \int_0^1 \xi(y + (1-t)\varepsilon z)\ dt.$$

This implies that $c_\xi(x,.) \in W^{1,\infty}$ and:

$$|\nabla_y c_\xi(x,.)| \leq |\xi(.)|, \quad \text{for all } x. \tag{3.3}$$

By symmetry we also have

$$|\nabla_x c_\xi(.,y)| \leq |\xi(.)|, \quad \text{for all } y. \tag{3.4}$$

Since $q^* > 2$, we deduce from (3.3), (3.4) and Morrey's Theorem (see [6], Chapter IX), that there is a constant C such that:

$$|c_\xi(x,y_1) - c_\xi(x,y_2)| \leq C\|\xi\|_{L^{q^*}} |y_1 - y_2|^\alpha, \text{ for all } x, y_1, y_2 \text{ in } \Omega,$$
$$|c_\xi(x_1,y) - c_\xi(x_2,y)| \leq C\|\xi\|_{L^{q^*}} |x_1 - x_2|^\alpha, \text{ for all } x_1, x_2, y \text{ in } \Omega.$$

This proves (3.2). The second claim in the proposition then follows from (3.2), the identity $c_{\xi_n}(x,x) = 0$ and Ascoli's theorem. □

From now on, we further assume that $q < 2$. For a non-negative function $\xi \in L^{q^*}(\Omega)$ we then define

$$\overline{c}_\xi(x,y) = \sup\{c(x,y)\ :\ c \in \mathcal{A}(\xi)\},$$

where

$$\mathcal{A}(\xi) = \left\{\lim_n c_{\xi_n} \text{ in } C^0(\overline{\Omega}\times\overline{\Omega})\ :\ (\xi_n)_n \in C^0(\overline{\Omega}),\ \xi_n \geq 0,\ \xi_n \to \xi \text{ in } L^{q^*}\right\}.$$

Remark 3.3. The definition of $\overline{c}_\xi$ is unchanged if one replaces $\xi_n \to \xi$ in L^{q^*} by $\xi_n \rightharpoonup \xi$ in L^{q^*} in the definition of $\mathcal{A}(\xi)$. Indeed, if we do so, we obviously obtain a function which is larger than $\overline{c}_\xi$. Now, let us assume that $\xi_n \rightharpoonup \xi$ in L^{q^*}, and c_{ξ_n} converges to c in $C^0(\overline{\Omega}\times\overline{\Omega})$, using Mazur's Lemma there exists a sequence η_n which converges strongly to ξ and such that each η_n is in the convex hull of $\{\xi_k,\ k \geq n\}$. It is clear that for fixed x, y, $\xi \to c_\xi(x,y)$ is concave hence $c(x,y) = \lim c_{\xi_n}(x,y) \leq \limsup\ c_{\eta_n}(x,y) \leq \overline{c}_\xi(x,y)$.

When ξ is continuous, one has:

Lemma 3.4. *If ξ is continuous and non-negative, then $\overline{c}_\xi = c_\xi$.*

Proof. The inequality $\overline{c}_\xi \geq c_\xi$ is obvious, as one can always choose the constant sequence $\xi_n = \xi$ in the definition of $\overline{c_\xi}$. Let us show now the opposite inequality. Take $x, y \in \Omega$, $\varepsilon > 0$ and $\sigma \in C^{x,y}$ such that $L_\xi(\sigma) < c_\xi(x, y) + 1/k$. We can choose σ so that it is piecewise linear, by density of this kind of curves and using the continuity of ξ. Let $(S_i)_{i=0,\ldots,m-1}$ be the segments which compose σ with $S_i = x_i x_{i+1}$, $x_0 = x$ and $x_m = y$. Let us fix, moreover, a sequence $\xi_n \to \xi$ such that $c_{\xi_n} \to c$. Now, we want to prove $c \leq c_\xi$. Fix a small number $\delta > 0$ and for any $\alpha \in [0, \delta]$ let us define a curve σ^α in the following way: let R be the clockwise 90 degrees rotation in the plane; let $x_i'(\alpha)$ and $x_i''(\alpha)$ be the only points such that $x_i x_i'(\alpha) = \alpha R e_i$ and $x_{i+1} x_i''(\alpha) = \alpha R e_i$, where e_i is the tangent unit vector to σ in the S_i part; define σ^α by linking any point $x_i'(\alpha)$ to $x_i''(\alpha)$ by some segments $S_i'(\alpha)$ and $x_i''(\alpha)$ to $x_{i+1}'(\alpha)$ by some arcs $A_{i+1}(\alpha)$ with center x_{i+1} and radius α. In this way we have $\sigma^\alpha \in C^{x_\alpha, y_\alpha}$, where $x_\alpha = x_0'(\alpha)$ and $y_\alpha = x_m''(\alpha)$. Let $R_i(\delta)$ be the rectangle whose vertices are the points x_i, $x_i'(\delta)$, $x_i''(\delta)$ and x_{i+1} and let $B_i(\delta)$ be the circular sector centered at x_i and whose vertices are $x_{i-1}''(\delta)$ and $x_i'(\delta)$.

If we compute $\int_0^\delta L_{\xi_n}(\sigma_\alpha)\, d\alpha$ it is not difficult to see that we get

$$\int_0^\delta L_{\xi_n}(\sigma_\alpha)\, d\alpha = \sum_{i=0}^{m-1} \left(\int_{R_i(\delta)} \xi_n d\mathcal{L}^2 \right) + \sum_{i=1}^{m-1} \left(\int_{B_i(\delta)} \xi_n d\mathcal{L}^2 \right).$$

Moreover it holds $c_{\xi_n}(x^\alpha, y^\alpha) \leq L_{\xi_n}(\sigma^\alpha)$, hence we get

$$\int_0^\delta c_{\xi_n}(x^\alpha, y^\alpha)\, d\alpha \leq \sum_{i=0}^{m-1} \left(\int_{R_i(\delta)} \xi_n d\mathcal{L}^2 \right) + \sum_{i=1}^{m-1} \left(\int_{B_i(\delta)} \xi_n d\mathcal{L}^2 \right).$$

If we pass to the limit as $n \to \infty$ we get, by using the uniform convergence of c_{ξ_n} to c on the left hand side and the L^{q^*} convergence of ξ_n to ξ on the right hand side,

$$\int_0^\delta c(x^\alpha, y^\alpha)\, d\alpha \leq \sum_{i=0}^{m-1} \left(\int_{R_i(\delta)} \xi d\mathcal{L}^2 \right) + \sum_{i=1}^{m-1} \left(\int_{B_i(\delta)} \xi d\mathcal{L}^2 \right).$$

Then we divide by δ and we pass to the limit as $\delta \to 0$. Using the fact that c is continuous we have

$$\lim_{\delta \to 0} \frac{1}{\delta} \int_0^\delta c(x^\alpha, y^\alpha)\, d\alpha = c(x, y).$$

On the other side, we may notice that the areas of the sectors $B_i(\delta)$ may be estimated by $C\delta^2$ and hence we have, for $\delta \to 0$,

$$\frac{1}{\delta}\sum_{i=1}^{m-1}\left(\int_{B_i(\delta)} \xi d\mathcal{L}^2\right) \leq mC|\xi|\delta \to 0.$$

On the contrary the integrals over $R_i(\delta)$, when divided by δ, converge on the integrals on the segments S_i, which give exactly the integral over the curve σ, i.e. $L_\xi(\sigma)$. We have consequently

$$\lim_{\delta\to 0}\frac{1}{\delta}\left(\sum_{i=0}^{m-1}\left(\int_{R_i(\delta)} \xi d\mathcal{L}^2\right) + \sum_{i=1}^{m-1}\left(\int_{B_i(\delta)} \xi d\mathcal{L}^2\right)\right) = L_\xi(\sigma) < c_\xi(x,y) + \frac{1}{k}.$$

This gives

$$c(x,y) < c_\xi(x,y) + \frac{1}{k}$$

and, k being arbitrary, we also get $c \leq c_\xi$ and the thesis. □

Lemma 3.5. *Let us assume that $q < 2$ and let ξ be non-negative function belonging to L^{q^*}, then there exists a sequence $(\xi_n)_n \in C^0(\Omega)$, $\xi_n \geq 0$, $\xi_n \to \xi$ in L^{q^*}, such that c_{ξ_n} converges uniformly to $\overline{c}_\xi$ on $\Omega \times \Omega$.*

Proof. It is easy to see that for every $(x,y) \in \Omega^2$ there exists a sequence of non-negative continuous functions $(\xi_n)_n$ converging to ξ in $L^{q^*}(\Omega)$ such that c_{ξ_n} converges in C^0 and $\overline{c}_\xi(x,y) = \lim_n c_{\xi_n}(x,y)$. Let I be a finite set, $(x_i,y_i) \in \Omega^2$ for all $i \in I$ and for every i, let (ξ_n^i) be a sequence of non-negative continuous functions converging to ξ in $L^{q^*}(\Omega)$ such that $\overline{c}_\xi(x_i,y_i) = \lim_n c_{\xi_n^i}(x_i,y_i)$. Let us set $\xi_n := \max_{i\in I}\xi_n^i$, we then have $(\xi_n)_n$ converging to ξ in $L^{q^*}(\Omega)$, and

$$\overline{c}_\xi(x_i,y_i) \leq \liminf_n\ c_{\xi_n}(x_i,y_i) \leq \limsup_n\ c_{\xi_n}(x_i,y_i) \leq \overline{c}_\xi(x_i,y_i).$$

We thus have $\overline{c}_\xi(x_i,y_i) = \lim_n c_{\xi_n}(x_i,y_i)$ for every $i \in I$. Now, let $(x_i,y_i)_{i\in\mathbb{N}}$ be a dense sequence of points of Ω^2. From what preceeds, for every n, there exists a continuous non-negative ξ_n such that

$$\|\xi_n - \xi\|_{L^{q^*}} \leq \frac{1}{n},\ |\overline{c}_\xi(x_k,y_k) - c_{\xi_n}(x_k,y_k)| \leq \frac{1}{n},\ \forall k \leq n.$$

By the Hölder estimate of proposition 3.2 and Ascoli's theorem, passing to a subsequence if necessary, we may assume that c_{ξ_n} converges in C^0 to some c. Since obviously $c(x_k,y_k) = \overline{c}_\xi(x_k,y_k)$ for all k , we deduce $c = \overline{c}_\xi$ and the desired result follows. □

The next lemma enables us to extend L_ξ in some sense when $\xi \geq 0$ is only L^{q^*}:

Lemma 3.6. *Let us assume that $q < 2$. Let $Q \in \mathcal{Q}^q(\mu_0, \mu_1)$, ξ be a non-negative element of L^{q^*}, and $(\xi_n)_n$ be a sequence of non-negative continuous functions that converges to ξ in L^{q^*}, then we have the following:*

(i) $(L_{\xi_n})_n$ converges strongly in $L^1(C, Q)$ to some limit which is independent of the approximating sequence $(\xi_n)_n$ and which will again be denoted L_ξ.

(ii) The following equality holds:

$$\int_\Omega \xi(x) i_Q(x)\ dx = \int_C L_\xi(\sigma)\ dQ(\sigma). \tag{3.5}$$

(iii) The following inequality holds for Q-a.e. $\sigma \in C$:

$$L_\xi(\sigma) \geq \overline{c}_\xi(\sigma(0), \sigma(1)). \tag{3.6}$$

Proof. For all n and m in $\mathbb{N}$ we have:

$$\begin{aligned} \int_C |L_{\xi_n}(\sigma) - L_{\xi_m}(\sigma)|\ dQ(\sigma) &= \int_C \left| \int_0^1 (\xi_n(\sigma(t)) - \xi_m(\sigma(t)))\, |\dot\sigma(t)|\ dt \right| dQ(\sigma) \\ &\leq \int_\Omega |\xi_n(x) - \xi_m(x)| i_Q(x)\ dx \\ &\leq \|\xi_n - \xi_m\|_{L^{q^*}} \|i_Q\|_{L^q}. \end{aligned}$$

This implies that $(L_{\xi_n})_n$ is a Cauchy sequence in $L^1(C, Q)$ and it is obvious, from the previous inequality, that its $L^1(C, Q)$ limit does not depend on the approximating sequence $(\xi_n)_n$.

The proof of (ii) follows from (i):

$$\begin{aligned} \int_\Omega \xi(x) i_Q(x)\ dx &= \lim_n \int_\Omega \xi_n(x) i_Q(x)\ dx \\ &= \lim_n \int_C L_{\xi_n}(\sigma)\ dQ(\sigma) \\ &= \int_C L_\xi(\sigma)\ dQ(\sigma). \end{aligned}$$

To prove (iii) we choose an approximating sequence $(\xi_n)_n$ as in lemma 3.5 and pass to the limit in

$$L_{\xi_n}(\sigma) \geq c_{\xi_n}(\sigma(0), \sigma(1)).$$

□

Remark 3.7. The condition $q < 2$ may seem restrictive, however if for instance μ_0 and μ_1 are discrete (and $\mu_0 \neq \mu_1$) this condition is in fact sharp. Indeed, in this case, it can easily be proved that $\mathcal{Q}^q(\mu_0, \mu_1) \neq \emptyset$ for $q \in (1, 2)$ but $\mathcal{Q}^2(\mu_0, \mu_1) = \emptyset$. To see that $\mathcal{Q}^2(\mu_0, \mu_1) = \emptyset$, assume on the contrary that there is a $Q \in \mathcal{Q}(\mu_0, \mu_1)$ such that $i_Q \in L^2$, and define the vector measure $\vec{i}_Q$ by:

$$\int_\Omega X(x) d\vec{i}_Q(x) = \int_C \left(\int_0^1 X(\sigma(t)) \cdot \dot{\sigma}(t) dt \right) dQ(\sigma), \ \forall X \in C^0(\overline{\Omega}, \mathbb{R}^2).$$

It is easy to check:

$$\operatorname{div}(\vec{i}_Q) = \mu_0 - \mu_1, \text{ and } \|\vec{i}_Q\|_{L^2} \leq \|i_Q\|_{L^2} < +\infty.$$

But, since $\mu_0 - \mu_1 \notin H^{-1}(\Omega)$, we get the desired contradiction (and in fact we have proved that $\mathcal{Q}^2(\mu_0, \mu_1) = \emptyset$ as soon as $\mu_0 - \mu_1 \notin H^{-1}(\Omega)$). In other words, if $q \geq 2$, the congestion effect is so strong that the total congested cost in (2.13) is always $+\infty$ as soon as $\mu_0 - \mu_1 \notin H^{-1}(\Omega)$.

Remark 3.8. We have assumed throughout the paper that the ambient dimension is 2, and we have seen in this case that one can extend the definition of c_ξ and L_ξ to the case $\xi \in L^{q^*}$, $\xi \geq 0$ provided $q < 2$. In dimension $d \geq 2$, it is easy to see that the Hölder's estimate of Proposition 3.2 still holds for $q^* > d$ i.e. $q < d/(d-1)$. All the results of the paper in fact extend to this more general case.

3.3 Characterization of optimal transport strategies

In this paragraph, our aim is to make the formal arguments of paragraph 3.1 rigorous in order to characterize optimal transport strategies. This can be done under the additional assumption that H is strictly convex. First, we relate the optimality condition (3.1) to the Monge-Kantorovich problem with cost $\overline{c}_{\overline{\xi}}$:

Proposition 3.9. *Let us assume that $q < 2$ and that H is strictly convex. If $\overline{Q}$ solves (2.13) and $\overline{\xi} := H'(i_{\overline{Q}})$ then we have:*

$$\int_\Omega \overline{\xi} i_{\overline{Q}} = \inf_{Q \in \mathcal{Q}^q(\mu_0, \mu_1)} \int_\Omega \overline{\xi} i_Q = \inf_{\gamma \in \Pi(\mu_0, \mu_1)} \int_{\overline{\Omega} \times \overline{\Omega}} \overline{c}_{\overline{\xi}}(x, y) d\gamma(x, y). \tag{3.7}$$

Proof. Let us recall that from proposition 3.1, we have:

$$\int_\Omega \overline{\xi} i_{\overline{Q}} = \inf_{Q \in \mathcal{Q}^q(\mu_0, \mu_1)} \int_\Omega \overline{\xi} i_Q. \tag{3.8}$$

Let ξ be a non-negative element of L^{q^*} and let $Q \in \mathcal{Q}^q(\mu_0, \mu_1)$, using Lemma 3.6 and the definition of $\mathcal{Q}^q(\mu_0, \mu_1)$ yields:

$$\begin{aligned}\int_\Omega \xi i_Q = \int_C L_\xi dQ &\geq \int_C \overline{c}_\xi(\sigma(0)), \sigma(1)) dQ(\sigma) \\ &\geq \inf_{\gamma \in \Pi(\mu_0,\mu_1)} \int_{\overline{\Omega}\times\overline{\Omega}} \overline{c}_\xi(x,y) d\gamma(x,y).\end{aligned}$$

We then have, for all $\xi \in L^{q^*}$, $\xi \geq 0$:

$$\inf_{Q\in\mathcal{Q}^q(\mu_0,\mu_1)} \int_\Omega \xi i_Q \geq \inf_{\gamma\in\Pi(\mu_0,\mu_1)} \int_{\overline{\Omega}\times\overline{\Omega}} \overline{c}_\xi(x,y) d\gamma(x,y). \tag{3.9}$$

By a similar argument and using Lemma 3.4, we also have, for all $\xi \in C^0(\overline{\Omega}, \mathbb{R}_+)$:

$$\inf_{Q\in\mathcal{Q}(\mu_0,\mu_1)} \int_{\overline{\Omega}} \xi di_Q \geq \inf_{\gamma\in\Pi(\mu_0,\mu_1)} \int_{\overline{\Omega}\times\overline{\Omega}} \overline{c}_\xi(x,y) d\gamma(x,y). \tag{3.10}$$

Let $\xi \in C^0(\overline{\Omega}, \mathbb{R}_+)$, and $\varepsilon > 0$, for every x and y in $\overline{\Omega}$, there exists $\sigma_\varepsilon^{x,y} \in C^{x,y}$ such that $(x,y) \mapsto \sigma_\varepsilon^{x,y}$ is measurable (see for instance [8]) and by Lemma 3.4

$$L_\xi(\sigma_\varepsilon^{x,y}) \leq c_\xi(x,y) + \varepsilon = \overline{c}_\xi(x,y) + \varepsilon. \tag{3.11}$$

Let $\gamma \in \Pi(\mu_0, \mu_1)$ and let us define the element of $\mathcal{Q}(\mu_0, \mu_1)$: $Q_\varepsilon := \delta_{\sigma_\varepsilon^{x,y}} \otimes \gamma$, we then have:

$$\int_{\overline{\Omega}} \xi di_{Q_\varepsilon} = \int_{\overline{\Omega}\times\overline{\Omega}} L_\xi(\sigma_\varepsilon^{x,y}) d\gamma(x,y) \leq \int_{\overline{\Omega}\times\overline{\Omega}} \overline{c}_\xi(x,y) d\gamma(x,y) + \varepsilon.$$

Since γ and ε are abitrary, using (3.10) we obtain

$$\inf_{Q\in\mathcal{Q}(\mu_0,\mu_1)} \int_{\overline{\Omega}} \xi di_Q = \inf_{\gamma\in\Pi(\mu_0,\mu_1)} \int_{\overline{\Omega}\times\overline{\Omega}} \overline{c}_\xi(x,y) d\gamma(x,y), \ \forall \xi \in C^0(\overline{\Omega}, \mathbb{R}_+). \tag{3.12}$$

In what follows, for every $\mu \in \mathcal{M}_+(\overline{\Omega})$, we extend μ by 0 outside $\overline{\Omega}$. Let $(\rho_n)_n$ be a standard mollifying sequence. For $n \in \mathbb{N}^*$ let us consider the regularized problem:

$$\inf_{Q\in\mathcal{Q}(\mu_0,\mu_1)} \int_{\mathbb{R}^2} H(\rho_n \star i_Q). \tag{3.13}$$

The existence of a solution $\overline{Q}_n$ of (3.13) can be obtained by similar arguments as in theorem 2.10 (using lemma 2.8 and the fact that the L^1 norm of $\rho_n \star i_Q$

equals the total mass of i_Q). Proceeding as in proposition 3.1 and defining $j_n := \rho_n \star i_{\overline{Q}_n}$, $\xi_n := H'(j_n)$, $\eta_n := \rho_n \star \xi_n$, we have:

$$\int_{\mathbb{R}^2} H'(\rho_n \star i_{\overline{Q}_n})(\rho_n \star i_{\overline{Q}_n}) = \int_{\mathbb{R}^2} \xi_n j_n = \int_{\mathbb{R}^2} \eta_n d i_{\overline{Q}_n} = \inf_{Q \in \mathcal{Q}(\mu_0,\mu_1)} \int_{\mathbb{R}^2} \eta_n d i_Q. \tag{3.14}$$

With (3.12), we then get:

$$\int_{\mathbb{R}^2} \eta_n d i_{\overline{Q}_n} = \int_{\mathbb{R}^2} \xi_n j_n = \inf_{\gamma \in \Pi(\mu_0,\mu_1)} \int_{\overline{\Omega}\times\overline{\Omega}} \overline{c}_{\eta_n}(x,y) d\gamma(x,y). \tag{3.15}$$

By convexity of H, we also have:

$$\int_{\mathbb{R}^2} H(j_n) \leq \int_{\mathbb{R}^2} H(\rho_n \star i_{\overline{Q}}) \leq \int_{\mathbb{R}^2} \rho_n \star H(i_{\overline{Q}}) \tag{3.16}$$

which implies that j_n is bounded in L^q. Passing to subsequences, we may therefore assume:

$$j_n \rightharpoonup j \text{ in } L^q,\ \xi_n \rightharpoonup \xi \text{ in } L^{q^*},\ \eta_n \rightharpoonup \xi \text{ in } L^{q^*}. \tag{3.17}$$

Since the total mass of $i_{\overline{Q}_n}$ is the same of j_n and j_n is bounded in L^q (and hence in L^1), we get a bound on $i_{\overline{Q}_n}(\Omega)$ and, from lemma 2.8, we may also assume:

$$\overline{Q}_n \overset{*}{\rightharpoonup} Q \text{ in } \mathcal{M}_+(C),\ i_{\overline{Q}_n} \overset{*}{\rightharpoonup} i \text{ in } \mathcal{M}_+(\overline{\Omega}). \tag{3.18}$$

It is obvious that $i = j$ and lemma 2.9 implies $j \geq i_Q$. With (3.16) and the monotonicity of H, we then get:

$$\int_{\Omega} H(i_Q) \leq \int_{\Omega} H(j) \leq \liminf_n \int_{\mathbb{R}^2} H(j_n) \leq \int_{\Omega} H(i_{\overline{Q}}). \tag{3.19}$$

With the strict convexity of H and the optimality of $\overline{Q}$, this also yields

$$i_{\overline{Q}} = i_Q = j \in L^q \text{ and } \liminf_n \int_{\mathbb{R}^2} H(j_n) = \int_{\mathbb{R}^2} H(i_{\overline{Q}}). \tag{3.20}$$

Up to some subsequence (as $\overline{\xi} = H'(i_{\overline{Q}}) \in L^{q^*}$), this also implies

$$H(j_n) - H(i_{\overline{Q}}) - \overline{\xi}(j_n - i_{\overline{Q}}) \to 0 \text{ a.e. and in } L^1$$

and using the strict convexity of H, we deduce that j_n converges a.e. to $i_{\overline{Q}}$. This implies that ξ_n converges a.e. to $\overline{\xi} = H'(i_{\overline{Q}})$ and that $\xi = \overline{\xi}$. With

Fatou's Lemma and (3.15), we therefore obtain:

$$\begin{aligned}
\int_\Omega \overline{\xi} i_{\overline{Q}} &= \int_{\mathbb{R}^2} H'(i_{\overline{Q}}) i_{\overline{Q}} \\
&\leq \liminf_n \int_{\mathbb{R}^2} \xi_n j_n \\
&= \liminf_n \inf_{\gamma \in \Pi(\mu_0,\mu_1)} \int_{\overline{\Omega}\times\overline{\Omega}} \overline{c}_{\eta_n}(x,y) d\gamma(x,y).
\end{aligned}$$

Using $\eta_n \rightharpoonup \overline{\xi}$ and remark 3.3, from the uniform convergence of c_{η_n} to a cost $c \leq c_{\overline{\xi}}$, we get

$$\int_\Omega \overline{\xi} i_{\overline{Q}} \leq \inf_{\gamma \in \Pi(\mu_0,\mu_1)} \int_{\overline{\Omega}\times\overline{\Omega}} \overline{c}_{\overline{\xi}}(x,y) d\gamma(x,y)$$

together with (3.8) and (3.9), this completes the proof. □

The characterization of optimal transport strategies then reads as:

Theorem 3.10. *Let us assume that $q < 2$ and that H is strictly convex. A transportation strategy $(\overline{\gamma}, \overline{p})$ is optimal if and only if, setting $\overline{Q} := Q_{\overline{\gamma},\overline{p}}$ and $\overline{\xi} := H'(i_{\overline{Q}})$, one has:*

1. *$\overline{\gamma}$ solves the Monge-Kantorovich problem:*

$$\inf_{\gamma \in \Pi(\mu_0,\mu_1)} \int_{\overline{\Omega}\times\overline{\Omega}} \overline{c}_{\overline{\xi}}(x,y) d\gamma(x,y), \tag{3.21}$$

2. *for $\overline{Q}$-a.e. $\sigma \in C$, one has:*

$$L_{\overline{\xi}}(\sigma) = \overline{c}_{\overline{\xi}}(\sigma(0), \sigma(1)). \tag{3.22}$$

Proof. Let us assume first that the transportation strategy $(\overline{\gamma}, \overline{p})$ is optimal and set $\overline{Q} := Q_{\overline{\gamma},\overline{p}}$ and $\overline{\xi} := H'(i_{\overline{Q}})$. From Proposition 3.9 and Lemma 3.6, we get:

$$\begin{aligned}
\int_{\overline{\Omega}\times\overline{\Omega}} \overline{c}_{\overline{\xi}}(x,y) d\overline{\gamma}(x,y) &= \int_C \overline{c}_{\overline{\xi}}(\sigma(0),\sigma(1)) d\overline{Q}(\sigma) \\
&\leq \int_C L_{\overline{\xi}} d\overline{Q} = \int_\Omega \overline{\xi} i_{\overline{Q}} \\
&= \inf_{\gamma \in \Pi(\mu_0,\mu_1)} \int_{\overline{\Omega}\times\overline{\Omega}} \overline{c}_{\overline{\xi}}(x,y) d\gamma(x,y)
\end{aligned}$$

this proves that $\overline{\gamma}$ solves (3.21) and implies that the inequalities above are equalities. We therefore deduce (3.22) from the inequality $\overline{c}_{\overline{\xi}}(\sigma(0), \sigma(1)) \leq L_{\overline{\xi}}(\sigma)$.

Conversely, assume that the transportation strategy $(\overline{\gamma}, \overline{p})$ satisfies the two conditions of the theorem. Condition (3.22) firstly yields:

$$\int_\Omega \overline{\xi} i_{\overline{Q}} = \int_C L_{\overline{\xi}} d\overline{Q} = \int_C \overline{c}_{\overline{\xi}}(\sigma(0), \sigma(1)) d\overline{Q}(\sigma) = \int_{\overline{\Omega}\times\overline{\Omega}} \overline{c}_{\overline{\xi}}(x, y) d\overline{\gamma}(x, y).$$

Secondly, if $Q = Q_{\gamma,p} \in \mathcal{Q}^q(\mu_0, \mu_1)$, one has:

$$\int_\Omega \overline{\xi} i_Q = \int_C L_{\overline{\xi}} dQ \geq \int_C \overline{c}_{\overline{\xi}}(\sigma(0), \sigma(1)) dQ(\sigma) = \int_{\overline{\Omega}\times\overline{\Omega}} \overline{c}_{\overline{\xi}}(x, y) d\gamma(x, y)$$

and since $\overline{\gamma}$ solves (3.21), we finally have

$$\int_\Omega \overline{\xi} i_{\overline{Q}} \leq \int_\Omega \overline{\xi} i_Q, \ \forall Q \in \mathcal{Q}^q(\mu_0, \mu_1)$$

which, with proposition 3.1, proves that $(\overline{\gamma}, \overline{p})$ is optimal.

□

Let us see, through an easy example, an application of Theorem 3.10.

Exemple 3.11. Suppose that $\overline{\Omega}$ contains the two segments $A = \{0\} \times [0, 1]$ and $B = \{1\} \times [0, 1]$ and the square $S = [0, 1] \times [0, 1]$ which is their convex hull. Set $\mu_1 = \mathcal{H}^1 \llcorner A$ and $\mu_2 = \mathcal{H}^1 \llcorner B$ and denote by $T : A \to C$ the map that associates to every point $(0, x) \in A$ the curve $T(x)$ given by $T(x)(t) = (t, x)$, i.e. the horizontal segment from A to B starting from x. Set $Q = T \sharp \mu_1$. It is clear that Q comes from an admissible transportation strategy linking μ_1 to μ_2 and it is not difficult to see that the traffic intensity i_Q has constant density 1 on S and 0 elsewhere. We consider two particular cases only: we claim that Q is optimal if $\Omega =]0, 1[\times]0, 1[$ while it is not if S is compactly contained in Ω. Indeed, if $\overline{\Omega} = S$, the metric induced by i_Q is the euclidean metric, the paths $T(x)$ are geodesic and the transport plan induced by Q is optimal according to this metric. On the other hand, if Ω is larger than S, then all the segments $T(x)$ that are very close to the upper or lower boundary of S are not geodesic accoding to this metric, because they could be improved by non-straight line paths which arrive up to zone $\Omega \setminus S$ where $i_Q = 0$ and the transportation is cheaper. In the former case, consequently, the sufficient optimality conditions are satisfied, while in the latter the geodesic conditions on the paths (Wardrop condition, see the next section) is not and prevents optimality.

Remark 3.12. Let us remark that $\overline{\xi} = H'(i_{\overline{Q}})$ (with $\overline{Q}$ solving (2.13)) solves the following (dual of (2.13)) problem:

$$\sup_{\xi \in L^{q^*},\, \xi \geq 0} W(\xi) - \int_\Omega H^*(\xi(x))dx \tag{3.23}$$

where H^* is the Fenchel transform of H and:

$$W(\xi) := \inf_{\gamma \in \Pi(\mu_0, \mu_1)} \int_{\overline{\Omega} \times \overline{\Omega}} \overline{c}_\xi(x, y) d\gamma(x, y).$$

Indeed, it follows from proposition 3.9 and $\overline{\xi} = H'(i_{\overline{Q}})$ that

$$W(\overline{\xi}) - \int_\Omega H^*(\overline{\xi}(x))dx = \int_\Omega (\overline{\xi} i_{\overline{Q}} - H^*(\overline{\xi})) = \int_\Omega H(i_{\overline{Q}}).$$

If $\xi \in L^{q^*}$ is non-negative, we deduce from (3.9) and Young's inequality:

$$W(\xi) \leq \int_\Omega \xi i_{\overline{Q}} \leq \int_\Omega H(i_{\overline{Q}}) + \int_\Omega H^*(\xi)$$

which proves that $\overline{\xi}$ solves (3.23).

Remark 3.13. It is very natural to investigate numerical schemes for the primal problem (2.13). We believe that the dual problem (3.23) offers a tractable and convenient way to address this numerical issue. We are not developing this point further here and let the numerical approximation of (2.13) and (3.23) for future research.

4 Application to equilibria of Wardrop type

In this final section, we relate the results of the previous sections to some concepts of equilibria of Wardrop type. Modelling congestion as in paragraph 2.2 enables us to extend the concept of Wardrop equilibrium to a continuous setting.

Let us consider a congestion function $g : \mathbb{R}_+ \to \mathbb{R}_+$ which is continuous increasing and satisfies $az^{q-1} \leq g(z) \leq b(z^{q-1} + 1)$ for all $z \in \mathbb{R}_+$ and some $q \in (1, 2)$ and non-negative constants a and b. Then for any transportation strategy (γ, p) such that $I_{\gamma,p}$ (defined by (2.3)) belongs to L^q, the transportation cost function resulting from the strategy (γ, p) is $\overline{c}_\xi$ for $\xi := g \circ I_{\gamma,p} \in L^{q^*}$. Roughly speaking, an equilibrium is then a transportation strategy (γ, p) that satisfies Wardrop stability condition (i.e. $Q_{\gamma,p}$ gives full mass to the set of geodesics for the metric $\xi = g \circ I_{\gamma,p}$) and the additional requirement that γ is an optimal transportation plan between μ_0 and μ_1 for the cost resulting from (γ, p). This leads to the following

Definition 4.1. *A transportation strategy* $(\overline{\gamma}, \overline{p})$ *is said to be an equilibrium if* $I_{\overline{\gamma},\overline{p}} \in L^q$ *and, setting* $\overline{\xi} := g \circ I_{\overline{\gamma},\overline{p}}$ *one has*

1. $L_{\overline{\xi}}(\sigma) = \overline{c}_{\overline{\xi}}(\sigma(0), \sigma(1))$ *for* $Q_{\overline{\gamma},\overline{p}}$*-a.e.* $\sigma \in C$,

2. $\overline{\gamma}$ *solves the Monge-Kantorovich problem:*

$$\inf_{\gamma \in \Pi(\mu_0, \mu_1)} \int_{\overline{\Omega} \times \overline{\Omega}} \overline{c}_{\overline{\xi}}(x, y) d\gamma(x, y).$$

Only the first condition above is linked to Wardrop's original equilibrium concept. Imagine that some social planner chooses the transportation plan γ, then the second equilibrium condition expresses that γ is optimal for the transportation cost resulting from γ itself and the traveler's individual behavior. Our notion of equilibrium can therefore be viewed as a refinement of the Wardrop equilibrium or its generalization to the case where the transportation plan is not given a priori.

Defining the function $H_g : \mathbb{R}_+ \to \mathbb{R}$ by

$$H_g'(z) = g(z), \ \forall z \in \mathbb{R}_+, \ H_g(0) = 0, \tag{4.1}$$

a direct application of theorems 2.10 and 3.10 then gives the existence of equilibria together with a variational characterization:

Theorem 4.2. *If* $\mathcal{Q}^q(\mu_0, \mu_1) \neq \emptyset$*, there exists an equilibrium. Moreover* $(\overline{\gamma}, \overline{p})$ *is an equilibrium if and only if* $\overline{Q} := Q_{\overline{\gamma},\overline{p}}$ *solves the minimization problem:*

$$\inf_{Q \in \mathcal{Q}^q(\mu_0, \mu_1)} \int_{\Omega} H_g(i_Q(x)) dx \tag{4.2}$$

where H_g *is defined by (4.1).*

Remark 4.3. We have always assumed that the congestion function g only depends on the intensity z. It is of course straightforward to extend all our results to the case of a congestion function $(x, z) \to g(x, z)$ that also depends on x. In this case, one finds equilibria by solving:

$$\inf_{Q \in \mathcal{Q}^q(\mu_0, \mu_1)} \int_{\Omega} H_g(x, i_Q(x)) dx$$

where H_g is defined by $H_g(x, 0) = 0$ and $\partial_z H_g(x, z) = g(x, z)$.

Remark 4.4. A slightly different situation, which can be relevant in some applications, occurs when the transportation plan $\overline{\gamma} \in \Pi(\mu_0, \mu_1)$ is fixed and not only the marginals μ_0 and μ_1. In this case, one defines equilibria as the set of $\overline{p}$'s such that $(\overline{\gamma}, \overline{p})$ satisfies the first condition (Wardrop) of definition 4.1. If the set:

$$\mathcal{Q}^q(\overline{\gamma}) := \{Q \in \mathcal{M}^1_+(C) : (e_0, e_1)\sharp Q = \overline{\gamma},\ i_Q \in L^q\}$$

is nonempty, then slightly adapting our arguments, we have existence of equilibria and $\overline{p}$ is an equilibrium if and only if $\overline{Q} := Q_{\overline{\gamma},\overline{p}}$ solves the minimization problem:

$$\inf_{Q \in \mathcal{Q}^q(\overline{\gamma})} \int_\Omega H_g(i_Q(x))dx.$$

References

[1] L. Ambrosio, *Lecture Notes on Optimal Transport Problems, Mathematical aspects of evolving interfaces,* CIME Summer School in Madeira, vol. 1812, Springer, 2003.

[2] J.-B. Baillon, R. Cominetti, Markovian Traffic Equilibrium, *Math. Prog.*, **111**, no. 1-2, 33–56, 2008.

[3] G. Bouchitté, G. Buttazzo, Characterization of Optimal Shapes and Masses through Monge-Kantorovich Equation, *J. Eur. Math. Soc.*, **3**, no. 2, 139–168, 2001.

[4] G. Bouchitté, G. Buttazzo, P. Seppecher, Shape Optimization solutions via Monge-Kantorovich equation, *C.R. Acad. Sci. Paris*, **324** , no. 10, 1185–1191, 1997.

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

[6] H. Brézis, *Analyse Fonctionnelle*, Masson, Paris, 1983.

[7] L. Caffarelli, M. Feldman, R. J. McCann, Constructing optimal maps for Monge's transport problem as a limit of strictly convex costs, *J. Amer. Math. Soc.*, **15** , no. 1, 1–26, 2002.

[8] C. Castaing, M. Valadier, *Convex Analysis and Measurable Multifunctions*, Lecture Notes in Mathematics, **580**, Springer-Verlag, Berlin 1977.

[9] C. Dellacherie, P.-A. Meyer, *Probabilities and Potential*, Mathematical Studies **29**, North-Holland, 1978.

[10] L. De Pascale and A. Pratelli, Regularity properties for Monge transport density and for solutions of some shape optimization problem. *Calc. Var. Partial Differential Equations* **14**, no. 3, 249–274, 2002.

[11] L. De Pascale and A. Pratelli, Sharp summability for Monge Transport density via Interpolation. *ESAIM, Control Opt. Calc. Var.* **10**, no. 4, 549-552, 2004.

[12] G. Monge, Mémoire sur la théorie des déblais et des remblais, *Hist. de l'Acad. des Sci. de Paris* 1781.

[13] C. Villani, *Topics in optimal transportation*, Graduate Studies in Mathematics, 58, American Mathematical Society, Providence, RI, 2003.

[14] J.G. Wardrop, Some theoretical aspects of road traffic research, *Proc. Inst. Civ. Eng.*, **2**, no.2, 325-378, 1952.