HAL
open science

A planning problem combining calculus of variations and optimal transport

Guillaume Carlier, Aimé Lachapelle

To cite this version:

Guillaume Carlier, Aimé Lachapelle. A planning problem combining calculus of variations and optimal transport. Applied Mathematics and Optimization, 2011, 63 (1), pp.1-9. 10.1007/s00245-010-9107-8 . hal-00432785

HAL Id: hal-00432785

https://hal.science/hal-00432785v1

Submitted on 18 Nov 2009

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A planning problem combining optimal control and optimal transport

G. Carlier, A. Lachapelle*

November 18, 2009

Abstract

We consider some variants of the classical optimal transport where not only one optimizes over couplings between some variables x and y but also over some control variables governing the (possibly stochastic) evolutions of these variables with time. Such a situation is motivated by an assignment problem of tasks with workers whose characteristics can evolve with time (and be controlled). We distinguish between the coupled and decoupled case. The coupled case is a standard optimal transport with the value of some optimal control problem as cost. The decoupled case is more involved since it is nonlinear in the transport plan. We fully treat the linear quadratic case and show that even in this case existence may fail for the decoupled problem.

Keywords: optimal transport, optimal control.

1 Introduction

In the classical optimal transport problem, a planner has to find some coupling or transport plan between two nonnegative measures of equal total mass (say 1) so as to minimize average cost. The Monge-Kantorovich problem is then the linear program consisting in finding a cost minimizing measure having prescribed marginals on variables, say x and y. Such problems have been very much studied in recent years in particular for their various applications in PDE's, geometry, probability theory (see the recent books of Villani [15], [16])... Interestingly, transport/assigment problems have their modern roots

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

in planning problems (optimally transporting coal from mines to steel factories for instance) and the pionneering works of Kantorovich in the USSR and Koopmans in the West were awarded by the Nobel prize in economics in 1975. Typical applications are optimal allocation of resources or assignment problems; for other economic applications, we refer to the recent work of Ekeland [5], [6] and Chiappori, McCann and Nesheim [4] relating matching problems and hedonic equilibria to optimal transport theory.

If, to fix ideas, we think of a firm having to assign N workers of different skills x to N tasks of different difficulties y, the assignment problem simply consists in assigning a task to each worker (for instance assigning most difficult tasks to most skilled workers). It is however reasonable to think that, in addition to designing such an assignment, the firm has the possibility to act directly on the workers' types using some (costly) control variable like internal training. The same of course applies to the task type y that may change with time if, for example, some investments are made to make the task easier or quicker to execute. This is where optimal control comes into play. The problems we consider in the present paper are precisely intended to deal with such situations where a planner or firm has at disposal as decision variables not only the initial assignment but also control variables that affect the (possibly stochastic) dynamics of the state variables x and y. More precisely, we consider the case where the planner chooses the initial coupling between x and y, initially distributed according to respective probabilities μ_0 and ν_0, and a control variable governing the evolutions of the characteristics x and y. It is assumed that a pair (worker/task) that is initially formed at time 0 with initial characteristics (x, y) will remain paired or married during a fixed period although its characteristics may evolve with time, this captures the idea that for instance workers are assigned a certain task during some minimal period.

We will consider two different cases that we respectively call the *coupled* and the *decoupled* case and will compare them. In the coupled case, one optimizes over initial transport plans and control for each initial pair (x, y). The coupled problem thus amounts to solve a standard optimal transport where the cost function simply is the value function of some optimal control problem. In the decoupled case, we will assume that y is constant in time and the control variable for x has to be chosen independently of y (for instance the planner offers the same training to all the workers of same skill, independently of the task they will be assigned). This case is much more involved since it is not linear (and not even lower semicontinuous) in the transport plan. Under quite general assumptions it turns out that the coupled and decoupled problems have the same value but that the latter may fail to have solutions

(a counter-example will be given in section 4).

The paper is organized as follows. In section 2, the coupled problem is introduced, an existence result is proved and some cases of uniqueness are given. In section 3, we study the decoupled problem which, contrary to the coupled case, gives rise to a nonlinear optimization problem over transport plans. Finally, section 4 is devoted to a linear quadratic deterministic case that we treat in details and for which we compare the coupled and the decoupled case, we also give some counterexamples.

2 The coupled case

2.1 The optimization problem

We consider a finite horizon framework on the period $[0, T]$. The state equation for the joint dynamics of $Z_t = (X_t, Y_t)$ is

$$dZ_t = b(Z_t, u_t)dt + \sigma(Z_t, u_t)dW_t, \ t \in [0, T] \tag{2.1}$$

where W_t is a standard $2d$-dimensional Brownian motion on some filtered probability space and $(u_t)_{t\in[0,T]}$ is an admissible control process i.e. it is a progressively measurable process with values in a control space U, assumed for simplicity to be a compact metric space. We will denote by $\mathcal{U}$ the set of admissible controls. We assume that $b \in C(\mathbb{R}^{2d}\times U, \mathbb{R}^{2d})$, $\sigma \in C(\mathbb{R}^{2d}\times U, M_{2d})$ (M_{2d} denoting the space of $2d \times 2d$ matrices with real entries) and that there is some constant K_1 such that for every $(z_1, z_2, u) \in \mathbb{R}^{2d} \times \mathbb{R}^{2d} \times U$ one has

$$|b(z_1, u) - b(z_2, u)| + |\sigma(z_1, u) - \sigma(z_2, u)| \leq K_1|z_1 - z_2|. \tag{2.2}$$

The Lipschitz assumption (2.2) guarantees the existence and uniqueness of a strong solution to (2.1) for every initial condition $Z_0 = (x, y)$ and every admissible control process $u \in \mathcal{U}$. We will denote $Z_t^{x,y,u}$ this solution in the sequel.

To each initial condition (x, y) and each admissible control process $u \in \mathcal{U}$ is associated a cost

$$\mathbb{E}\Big(\int_0^T c(Z_s^{x,y,u}, u_s)ds + \phi(Z_T^{x,y,u})\Big).$$

We assume that both c and ϕ are nonnegative and continuous and that there exist a constant K_2 and an exponent $k \geq 1$ such that for every $(z_1, z_2, u) \in \mathbb{R}^{2d} \times \mathbb{R}^{2d} \times U$:

$$|c(z_1, u) - c(z_2, u)| + |\phi(z_1) - \phi(z_2)| \leq K_2|z_1 - z_2|(1 + |z_1|^{k-1} + |z_2|^{k-1}). \tag{2.3}$$

The last ingredient in the problem is a pair of Borel probability measures μ_0 and ν_0 on $\mathbb{R}^d$ giving respectively the distributions of the initial state for x and for y. The initial distribution γ for the joint process Z is chosen by the controller, it is a coupling between μ_0 and ν_0 and thus belongs to the set of transport plan $\Pi(\mu_0, \nu_0)$ consisting of all Borel probability measures γ on $\mathbb{R}^d \times \mathbb{R}^d$ having μ_0 and ν_0 as marginals. We shall finally assume the following on μ_0 and ν_0:

$$\int_{\mathbb{R}^d} |x|^k d\mu_0(x) + \int_{\mathbb{R}^d} |y|^k d\nu_0(y) < +\infty. \tag{2.4}$$

The controller's optimization problem then reads as

$$\inf_{\gamma \in \Pi(\mu_0,\nu_0)} \int_{\mathbb{R}^d \times \mathbb{R}^d} \Big(\inf_{u \in \mathcal{U}} \mathbb{E}\Big(\int_0^T c(Z_s^{x,y,u}, u_s) ds + \phi(Z_T^{x,y,u}) \Big) \Big) d\gamma(x,y) \tag{2.5}$$

which can simply be rewritten as the optimal transport problem:

$$\inf_{\gamma \in \Pi(\mu_0,\nu_0)} \int_{\mathbb{R}^d \times \mathbb{R}^d} v(0,x,y) d\gamma(x,y). \tag{2.6}$$

where v is the value function

$$v(t,z) := \inf_{u \in \mathcal{U}} \mathbb{E}\Big(\int_t^T c(Z_s^{z,u}, u_s) ds + \phi(Z_T^{z,u}) \mid Z_t^{z,u} = z \Big), \ \forall (t,z) \in [0,T] \times \mathbb{R}^{2d}.$$

We won't discuss whether the previous optimal control admits a solution (for each fixed z). This issue relies on whether the value function is regular enough so that one can define an optimal feedback control, this is of course tightly related with the regularity for the corresponding Hamilton-Jacobi-Bellman equation. Instead, we will focus on the optimal transport formulation (2.6).

2.2 Existence

It is now easy to check that the infimum in (2.6) is attained, although the arguments are very standard in optimal transport (see [15]), we give a proof for the sake of completeness.

Theorem 2.1. *Under the assumptions of paragraph 2.1, the value of (2.6) is finite and attained by some optimal transport plan.*

Proof. First, it follows from assumption (2.3) and standard arguments in control theory (see for instance [8], section IV.6) that $v(0,.)$ has polynomial growth and is continuous. More precisely, there is a constant K_3 such that

$$0 \leq v(0,x,y) \leq K_3(1 + |x|^k + |y|^k), \ \forall (x,y) \in \mathbb{R}^{2d} \tag{2.7}$$

and, for all $(z_1, z_2) \in \mathbb{R}^{2d} \times \mathbb{R}^{2d}$:

$$|v(0, z_1) - v(0, z_2)| \leq K_3 |z_1 - z_2|(1 + |z_1|^{k-1} + |z_2|^{k-1}). \tag{2.8}$$

It immediately follows from (2.7) and (2.4) that the value of (2.6) is finite. Since $\Pi(\mu_0, \nu_0)$ is tight and $v(0, .)$ is continuous, for each $n \in \mathbb{N}^*$, there is some $\gamma_n \in \Pi(\mu_0, \nu_0)$ such that

$$\int_{\mathbb{R}^{2d}} \min(v(0, .), n) d\gamma_n \leq \int_{\mathbb{R}^{2d}} \min(v(0, .), n) d\gamma, \ \forall \gamma \in \Pi(\mu_0, \nu_0). \tag{2.9}$$

Again, by tightness of $(\gamma_n)_n$, we may assume, taking a subsequence if necessary, that γ_n converges narrowly to some $\overline{\gamma} \in \Pi(\mu_0, \nu_0)$. We thus have, for every $M > 0$ and every $\gamma \in \Pi(\mu_0, \nu_0)$:

$$\begin{aligned}
\int_{\mathbb{R}^{2d}} \min(v(0, .), M) d\overline{\gamma} &= \lim_n \int_{\mathbb{R}^{2d}} \min(v(0, .), M) d\gamma_n \\
&\leq \limsup_n \int_{\mathbb{R}^{2d}} \min(v(0, .), n) d\gamma_n \\
&\leq \limsup_n \int_{\mathbb{R}^{2d}} \min(v(0, .), n) d\gamma \\
&= \int_{\mathbb{R}^{2d}} v(0, .) d\gamma
\end{aligned}$$

and then letting $M \to +\infty$ we conclude that $\overline{\gamma}$ solves (2.6).

□

2.3 Some cases of uniqueness

Our aim now is to exhibit some special cases where there is uniqueness of an optimal transport plan for (2.6) and this plan is given by a transport map, that is of the form $\gamma = \delta_{S(x)} \otimes \mu_0$ for some measure preserving map S between μ_0 and ν_0 (which corresponds to a deterministic optimal coupling).

It is indeed well-known (see [10], [12] [3]) that if the transportation cost $v(0, ., .)$ satisfies a certain structural condition called the *generalized Spence-Mirrlees* or *twist* condition and if μ_0 is regular enough (absolutely continuous for instance) then there is uniqueness of an optimal transportation plan that is in fact given by a transport map. We won't insist here on the twist condition and the regularity required on μ_0 in general. Let us indicate some particular cases instead:

- $d = 1$, $v(0, ., .)$ of class C^2 (say) and the cross-derivative $\partial_{xy} v(0, x, y)$ is everywhere strictly positive, in this case, as soon as μ_0 is atomless, there

is a unique transport plan that is given by the unique nondecreasing map (or monotone rearrangement) pushing μ_0 forward to ν_0,

- $v(0,x,y) = c_0(x-y)$ with c_0 strictly convex and differentiable (see [11]) and μ_0 absolutely continuous and compactly supported, then there is also a unique optimal transport plan and the latter is given by a transport map,

- of particular interest is the quadratic case $c_0(z) = |z|^2$ which was first solved by Brenier [2], in this case, uniqueness of an optimal plan, once again given by a transport map, holds as soon as μ_0 and ν_0 have second moments and μ_0 does not give mass to sets of Hausdorff dimension less than $d-1$ (see McCann [13]).

Uncontrolled case Let us first consider the uncontrolled case where $b = b(z)$ is a Lipschitz vector field, $\sigma = \sqrt{2}I_{2d}$, the running cost $c = c(z)$ and the terminal cost $\phi = \phi(z)$ satisfy the same assumptions as in section 2.1. In this case, it is well-known that the function

$$v(t,z) := \mathbb{E}\Big(\int_t^T c(Z_s)ds + \phi(Z_T) \ \mid Z_t = z\Big), \ (t,z) \in (0,T) \times \mathbb{R}^{2d}$$

is the unique solution with polynomial growth of the backward linear parabolic equation

$$\partial_t v + \Delta v + b \cdot \nabla v = -c, \ v|_{t=T} = \phi. \tag{2.10}$$

As seen in the previous paragraph, in the optimal transportation problem (2.6), the cost to take into account is $v(0,.)$, the value at time 0 of the solution of (2.10).

Let us treat in details the case $b = 0$ and $\phi = 0$. In this case, (2.10) is simply the heat equation

$$\partial_t v + \Delta v = -c, \ v|_{t=T} = 0,$$

so that

$$v(0,.) = \int_0^T (G_t \star c)(.)dt \text{ with } G_t(z) = \frac{1}{(4\pi t)^{d/2}} \exp\Big(-\frac{|z|^2}{4t}\Big). \tag{2.11}$$

Then the optimal transportation problem (2.6) reads as:

$$\inf_{\gamma \in \Pi(\mu_0,\nu_0)} \int_{\mathbb{R}^d \times \mathbb{R}^d} \Big(\int_0^T (G_t \star c)(x,y)dt\Big) d\gamma(x,y).$$

In the particular case of the quadratic transportation cost $c(x,y) = \frac{|x-y|^2}{2}$, a direct computation leads to:

$$v(0,x,y) = T\Big(\frac{|x-y|^2}{2} + 1\Big).$$

Consequently, the infimum of (2.6) is $T(W^2(\mu_0,\nu_0)+1)$ where W^2 stands for the squared-2-Wasserstein distance:

$$W^2(\mu_0,\nu_0) := \inf_{\gamma\in\Pi(\mu_0,\nu_0)} \int_{\mathbb{R}^d\times\mathbb{R}^d} |x-y|^2 d\gamma(x,y).$$

It is well-known that for the quadratic cost, if μ_0 does not *give mass to small sets* (i.e. Borel sets of Hausdorff dimension less than $d-1$) then there is a unique optimal transport plan that is in fact induced by a transport map S. Moreover S is characterized by $S = \nabla u$ with u convex (see [2], [9], [13], [11]).

Linear control and zero transportation cost Let us now consider the linear drift $b(Z_t,u_t) = u_t$ with constant diffusion $\sigma = \sqrt{2}I_{2d}$ and the quadratic running cost $c(z,u) = c(u) = \frac{u^2}{2}$. We also assume that $\phi \in C^2(\mathbb{R}^{2d})$ and satisfies the same assumptions as before. The value function:

$$v(t,z) := \inf_{u\in\mathcal{U}} \mathbb{E}\Big(\int_t^T c(u_s)ds + \phi(Z_T) \ \mid Z_t = z\Big),\ (t,z)\in(0,T)\times\mathbb{R}^{2d}$$

is then the viscosity solution of the Hamilton-Jacobi-Bellman equation:

$$\partial_t v + \Delta v + \frac{|\nabla v|^2}{2} = 0,\ v|_{t=T} = \phi. \tag{2.12}$$

Performing the usual Hopf-Cole transformation $V(T-t,z) = e^{v(t,z)/2}$, $\Phi := e^{\phi/2}$, transforms (2.12) into the heat equation for V:

$$\partial_t V - \Delta V = 0,\ V|_{t=0} = \Phi. \tag{2.13}$$

As soon as Φ has polynomial growth, this gives the explicit expression

$$v(t,.) = 2\log(G_{T-t}\star\Phi(.))$$

where G is again the heat kernel as in (2.11). In the unidimensional case, one may thus check whether $v(0,.,.)$ satifies the cross-derivative criterion recalled above. Indeed, by a direct compuation, $\partial_{xy}v(0,.,.) > 0$ is equivalent to

$$(G_T\star\partial_{xy}\Phi)(G_T\star\Phi) > (G_T\star\partial_x\Phi)(G_T\star\partial_y\Phi).$$

Calculus of variations Let us now consider the simple deterministic case where the variable y is constant with respect to time, where the state equation for X is simply $\dot{X} = u$ and where the running and terminal costs respectively take the form

$$c(x-y) + F(u), \text{ and } \phi(x-y)$$

where c, ϕ and F are bounded from below, smooth and strictly convex functions, with $\Lambda \mathrm{id} \geq D^2 F \geq \lambda \mathrm{id}$ for some $\Lambda > 0$, $\lambda > 0$. Let us further assume for simplicity that μ_0 and ν_0 are compactly supported. In this case the value function is given by

$$v(0,x,y) = \inf_{X \in H^1,\, X(0)=x} \int_0^T \Big(c(X(t)-y) + F(\dot{X}(t)) \Big) dt + \phi(X(T)-y).$$

We then rewrite the value as $v(0,x,y) = w(x-y)$ where

$$w(z) = \inf_{X \in H^1,\, X(0)=z} \int_0^T \Big(c(X(t)) + F(\dot{X}(t)) \Big) dt + \phi(X(T)).$$

It is now easy to check that w is a C^1 and strictly convex function. Then, it is well-known (see [11], [9], [3]) that the corresponding optimal transport problem

$$\inf_{\gamma \in \Pi(\mu_0,\nu_0)} \int w(x-y) d\gamma(x,y)$$

possesses a unique solution that is given by an optimal transport map S as soon as μ_0 is absolutely continuous with respect to the Lebesgue measure (see [11]). In this simple case again, we then have existence and uniqueness of an optimal plan that is further given by a transport map. We will again use this example in the decoupled variant of section 3.

2.4 Alternative formulations

We now describe formally an alternative formulation of (2.5). Given $\gamma_0 \in \Pi(\mu_0, \nu_0)$ as initial coupling of X and Y and an admissible control that we assume to be of Markovian type i.e. of the form $u_t = u_t(X_t, Y_t)$, let $t \in [0,T]$ and γ_t^u be the joint probability law of (X_t, Y_t) (this is purely formal since the dependence of the Markovian control u on x, y may not be regular enough to define a flow to the SDE (2.1)):

$$\int_{\mathbb{R}^{2d}} \varphi d\gamma_t^u := \int_{\mathbb{R}^{2d}} \mathbb{E}(\varphi(Z_t^{x,y,u})) d\gamma_0(x,y), \ \forall \varphi \in C_c^\infty(\mathbb{R}^{2d}).$$

Then γ_t^u (again formally) solves the Fokker-Planck equation (controlled by u):

$$\partial_t \gamma - \frac{1}{2}\mathrm{tr}\Big(\sigma(z,u_t(z))\sigma(z,u_t(z))^* D^2\gamma\Big) + \mathrm{div}(b(z,u_t(z))\gamma) = 0,\ \gamma|_{t=0} = \gamma_0. \tag{2.14}$$

The total cost becomes linear with respect to γ^u and simply reads as:

$$\int_0^T \Big(\int_{\mathbb{R}^{2^d}} c(z,u_t(z))d\gamma_t^u(z)\Big)dt + \int_{\mathbb{R}^{2^d}} \phi(z)d\gamma_T^u(z) \tag{2.15}$$

which suggests as an alternative formulation of (2.5) to minimize (2.15) over initial conditions $\gamma_0 \in \Pi(\mu_0,\nu_0)$ and controls u, the dynamics of γ^u being governed by the PDE (2.14).

Let us also mention that it is well-known that (2.6) admits the dual formulation:

$$\sup\Big\{\int_{\mathbb{R}^d} \varphi d\mu_0 + \int_{\mathbb{R}^d} \psi d\nu_0\ :\ \varphi(x)+\psi(y) \le v(0,x,y)\Big\}. \tag{2.16}$$

3 The decoupled case

In the coupled problem considered previously, the situation is rather simple since the problem amounts first to compute the value function and then to solve the optimal transport problem where the cost is precisely this value function. However, this requires that the optimizer is able to design a control that depends on both state variables x and y. If one thinks of x as being workers' skills and y the difficulty of some task (constant in time to fix ideas), it may be reasonable to assume in certain cases that the control governing the evolution of x (training, education...) is the same for every worker of type x independently of the task y he will be assigned. In this case, one faces a somehow decoupled variant of the previous problem.

For the sake of simplicity, we will only consider in this section the deterministic case where in addition y is constant in time and $\dot{X} = u$ (calculus of variations). The running cost is given by some continuous Lagrangian $(x,u,y) \in \mathbb{R}^{3d} \mapsto L(x,u,y)$ that is convex with q-growth in the variable u for some $q > 1$, and the terminal cost is $(x,y) \in \mathbb{R}^{2d} \mapsto \phi(x,y)$. We also assume for simplicity that μ_0 and ν_0 are compactly supported. In this framework, the coupled problem considered previously reads as

$$(\mathcal{P}) \quad \inf_{\gamma\in\Pi(\mu_0,\nu_0)} I(\gamma) = \int_{\mathbb{R}^d\times\mathbb{R}^d} v(x,y)d\gamma(x,y)$$

where

$$v(x,y) := \inf_{X\in W^{1,q},\, X(0)=x} \int_0^T L(X(t),\dot{X}(t),y)dt + \phi(X(T),y).$$

Of course, $(\mathcal{P})$ possessses solutions as soon as v is continuous (or more generally lower semicontinuous). In the decoupled case, X is not allowed to depend on y but has instead to minimize the conditional average of the cost. More precisely, for $\gamma \in \Pi(\mu_0,\nu_0)$, let us denote the disintegration of γ with respect to its first marginal as $\gamma = \gamma^x \otimes \mu_0$ and then define for every x:

$$v_\gamma(x) := \inf_{X\in W^{1,q},\, X(0)=x} \int_{\mathbb{R}^d} \Big(\int_0^T L(X(t),\dot{X}(t),y)dt + \phi(X(T),y)\Big) d\gamma^x(y).$$

Then, the decoupled problem is:

$$(\mathcal{Q}) \quad \inf_{\gamma\in\Pi(\mu_0,\nu_0)} J(\gamma) = \int_{\mathbb{R}^d} v_\gamma(x) d\mu_0(x).$$

The novel feature of this decoupled problem is that, contrary to the optimal transport problem $(\mathcal{P})$, it is nonlinear. In fact, since $v_\gamma(x)$ is *concave* in γ^x (as an infimum of linear functionals of γ^x), $(\mathcal{Q})$ is a concave minimization problem over $\Pi(\mu_0,\nu_0)$. Existence of a minimizer is a real issue here since the criterion in $(\mathcal{Q})$ may not be weakly lower semicontinuous (and in fact even the measurability of $x \mapsto v_\gamma(x)$ is not obvious and a counterexample to lower semicontinuity will be given in the next section). Our aim now is not to study existence or nonexistence issues in general but simply to remark that $(\mathcal{P})$ turns out to be a kind of relaxation of $(\mathcal{Q})$ under quite weak assumptions:

Proposition 3.1. *If the value function v is continuous and μ_0 is atomless, then one has*

$$\inf(\mathcal{Q}) = \min(\mathcal{P}).$$

Moreover, γ solves $(\mathcal{Q})$ if and only if it solves $(\mathcal{P})$ and

$$v_\gamma(x) = \int_{\mathbb{R}^d} v(x,y) d\gamma^x(y) \ \ \mu_0\text{-a.e..}$$

In particular, if S is an optimal transport for $(\mathcal{P})$, i.e. $\delta_{S(x)} \otimes \mu_0$ solves $(\mathcal{P})$, then it also solves $(\mathcal{Q})$.

Proof. Since $v_\gamma(x) \geq \int_{\mathbb{R}^d} v(x,y)d\gamma^x(y)$, one immediately deduces that $\inf(\mathcal{Q}) \geq \min(\mathcal{P})$. Now, let us remark that if γ is given by a transport map that is $\gamma = \delta_{S(x)} \otimes \mu_0$ then $v_\gamma(x) = v(x,S(x))$ for μ_0-a.e. x and then in this case

one has $\int_{\mathbb{R}^d} v_\gamma d\mu_0 = \int_{\mathbb{R}^d} v(x, S(x)) d\mu_0(x) = \int_{\mathbb{R}^d \times \mathbb{R}^d} v d\gamma$. Since μ_0 is atomless the set of transport plans induced by transport maps is weakly-$*$ dense in $\Pi(\mu_0, \nu_0)$ (see [1]). In particular, since v is continuous, there is a minimizing sequence for $(\mathcal{P})$ of the form $\gamma_n = \delta_{S_n} \otimes \mu_0$ where $S_n \sharp \mu_0 = \nu_0$ for every n. One then has

$$\inf(\mathcal{Q}) \leq \int v_{\gamma_n} d\mu_0 = \int v(x, S_n(x)) d\mu_0(x)$$

and passing to the limit one gets $\inf(\mathcal{Q}) \leq \min(\mathcal{P})$. Let $\gamma \in \Pi(\mu_0, \nu_0)$, γ solves $(\mathcal{Q})$ if and only if

$$J(\gamma) = \inf(\mathcal{Q}) = \min(\mathcal{P}) \geq I(\gamma)$$

that is γ solves $(\mathcal{P})$ and $I(\gamma) = J(\gamma)$ i.e.

$$v_\gamma(x) = \int_{\mathbb{R}^d} v(x, y) d\gamma^x(y) \ \mu_0\text{-a.e..}$$

□

If we consider again the case

$$L(x, u, y) = c(x - y) + F(u), \ \phi(x, y) = \phi(x - y)$$

under the same convexity, growth and regularity assumptions as in the example of subsection 2.3, then we immediately deduce that if μ_0 is absolutely continuous with respect to the Lebesgue measure, then the corresponding decoupled problem admits a unique solution that is given by a transport map.

4 The linear-quadratic deterministic case

In this paragraph, we treat a special case where the state variable Y remains constant and the control on X is $u = \dot{X}$. We will consider the case of a quadratic running cost and zero terminal cost:

$$c(x, y, u) = \frac{1}{2}|x - y|^2 + \frac{1}{2}|u|^2, \phi(x, y) = 0$$

and we will distinguish the coupled case (where u is allowed to depend on both x and y) and the decoupled case where u depends on the initial condition x but has to be the same for every y. Of course, in this quadratic framework, we will also assume that μ_0 and ν_0 have finite second-order moments.

4.1 The coupled case

In the coupled case, one first solves for every (x, y):

$$v(x,y) := \inf_{X \in H^1(0,T,\mathbb{R}^d),\ X(0)=x} \int_0^T \Big(\frac{1}{2}|X(t)-y|^2 + \frac{1}{2}|\dot{X}(t)|^2\Big)dt.$$

By standard arguments, this problem has unique solution given by:

$$X(t,x,y) := a(t)(x-y)+y, \quad \text{with } a(t) = \frac{e^{t-2T}+e^{-t}}{1+e^{-2T}},\ t \in [0,T]. \tag{4.1}$$

Replacing in v then leads to the optimal transportation problem:

$$\frac{1}{2}\int_0^T (a^2+\dot{a}^2) \inf_{\gamma \in \Pi(\mu_0,\nu_0)} \int_{\mathbb{R}^d \times \mathbb{R}^d} |x-y|^2 d\gamma(x,y) \tag{4.2}$$

which is simply the optimal transportation problem with quadratic cost. Denoting by $W^2(\mu_0, \nu_0)$ the squared-2-Wasserstein distance between μ_0 and ν_0, we then have the following expression for the value of (4.2):

$$\frac{1}{2}\|a\|^2_{H^1} W^2(\mu_0,\nu_0) \text{ with } \|a\|^2_{H^1} = \frac{1-e^{-2T}}{1+e^{-2T}}. \tag{4.3}$$

4.2 The decoupled case

In the decoupled case, one has to minimize with respect to $\gamma \in \Pi(\mu_0, \nu_0)$ and $(t, x) \mapsto X(t, x)$ (independent of y) such that $X(0, x) = x$, the quantity

$$\int_{\mathbb{R}^d \times \mathbb{R}^d} \Big(\int_0^T \Big(\frac{1}{2}|X(t,x)-y|^2 + \frac{1}{2}|\dot{X}(t,x)|^2\Big)dt\Big)d\gamma(x,y). \tag{4.4}$$

By the disintegration theorem, every $\gamma \in \Pi(\mu_0, \nu_0)$ admits a disintegration with respect to its first marginal μ_0 i.e. can be written as $\gamma = \gamma^x \otimes \mu_0$ where $x \mapsto \gamma^x$ is a Borel family of probability measures (naturally interpreted as conditional probability of y given x). We will denote by $g_\gamma(x)$ the average of γ^x:

$$g_\gamma(x) := \int_{\mathbb{R}^d} y d\gamma^x(y), \text{ for } \mu_0\text{-a.e. } x \in \mathbb{R}^d.$$

Using the disintegration of γ with respect to μ_0 and the conditional moment g_γ, it is convenient to rewrite the cost (4.4) as:

$$\int_{\mathbb{R}^d} \Big(\int_0^T \Big(\frac{1}{2}|X(t,x)|^2 - X(t,x)\cdot g_\gamma(x) + \frac{1}{2}|\dot{X}(t,x)|^2\Big)dt\Big)d\mu_0(x) + \frac{T}{2}\int_{\mathbb{R}^d} |y|^2 d\nu_0(y).$$

Minimizing in $X(.,x)$ for each fixed x then yields

$$X(t,x) = a(t)(x - g_\gamma(x)) + g_\gamma(x) \tag{4.5}$$

where a is as in (4.1). Replacing in (4.4), then yields

$$v_\gamma(x) = \frac{1}{2}\|a\|_{H^1}^2 |x - g_\gamma(x)|^2 + \frac{T}{2}\int_{\mathbb{R}^d} |y - g_\gamma(x)|^2 d\gamma^x(y). \tag{4.6}$$

We thus find that the least cost, as a function of $\gamma_0 \in \Pi(\mu_0,\nu_0)$ only, is:

$$\begin{aligned} &\frac{1}{2}\|a\|_{H^1}^2 \int_{\mathbb{R}^d} |x|^2 d\mu_0(x) + \frac{T}{2}\int_{\mathbb{R}^d} |y|^2 d\nu_0(y) \\ &+\frac{1}{2}\Big(\|a\|_{H^1}^2 - T\Big)\int_{\mathbb{R}^d} |g_\gamma|^2 d\mu_0 - \|a\|_{H^1}^2 \int_{\mathbb{R}^d} x\cdot g_\gamma d\mu_0. \end{aligned} \tag{4.7}$$

Recalling that $\|a\|_{H^1}^2$ is given by (4.3), one easily gets $\|a\|_{H^1}^2 - T < 0$ for every $T > 0$ (which is consistent with the concavity property of the decoupled problem as noticed in section 3) and thus finding an optimal transport plan amounts to:

$$\inf_{\gamma\in\Pi(\mu_0,\nu_0)} J(\gamma) = -A(T)\int_{\mathbb{R}^d} |g_\gamma(x)|^2 d\mu_0(x) - 2\int_{\mathbb{R}^d\times\mathbb{R}^d} x\cdot y d\gamma(x,y) \tag{4.8}$$

where $A(T)$ is the positive constant:

$$A(T) := \frac{T}{\|a\|_{H^1}^2} - 1 = \frac{e^{-2T}(T+1) + T - 1}{1 - e^{-2T}}.$$

Of course when μ_0 is absolutely continuous, there is an optimal transport for the quadratic cost and then by proposition 3.1 it is also the unique solution of (4.8). The particular and simple form of the functional J in (4.8) enables us to deduce various counterexamples and some particular simple cases of non existence when μ_0 is atomless but may charge small sets.

The functional J is not lower semicontinuous: To see this, let us take μ_0 and ν_0 equal and uniform on the segment $[-1,1]$, $\gamma_n = \delta_{S_n}\otimes\mu_0$ where S_n is a sequence of μ_0-measure preserving maps such that γ_n weakly-$*$ converges to the independent coupling $\mu_0\otimes\mu_0$, then $\int |g_\gamma|^2 d\mu_0 = 0$ and then

$$\limsup_n \int |g_{\gamma_n}|^2 d\mu_0 = \limsup \|S_n\|_{L^2}^2 > \int |g_\gamma|^2 d\mu_0$$

since otherwise S_n would converge to 0 in L^2 and then one would have $\gamma = \delta_0\otimes\mu_0$.

A counterexample to existence: In view of the previous example, one should in general suspect non existence for the minimization of J over $\Pi(\mu_0, \nu_0)$. Let us consider the two-dimensional case where μ_0 is uniformly distributed on the segment $[0,1] \times \{0\}$ (it thus charges a small set but still is atomless) and let ν_0 be uniform on the square $[0,1]^2$. Assume that there exists a minimizer γ of the decoupled problem. In this case, since μ_0 is atomless, using proposition 3.1, the explicit values of $v(x,y)$ and $v_\gamma(x)$ given by (4.6), we would then have for μ_0-a.e. x

$$\begin{aligned} v_\gamma(x) &= \frac{1}{2}\|a\|_{H^1}^2 |x - g_\gamma(x)|^2 + \frac{T}{2}\int_{\mathbb{R}^d} |y - g_\gamma(x)|^2 d\gamma^x(y) \\ &= \int_{\mathbb{R}^d} v(x,y) d\gamma^x(y) = \frac{1}{2}\|a\|_{H^1}^2 \int_{\mathbb{R}^d} |x-y|^2 d\gamma^x(y) \\ &= \frac{1}{2}\|a\|_{H^1}^2 \Big(|x - g_\gamma(x)|^2 + \int_{\mathbb{R}^d} |y - g_\gamma(x))|^2 d\gamma^x(y) \Big) \ . \end{aligned}$$

Which implies that γ^x should be a Dirac mass for μ_0-a.e. x, it would thus imply that γ is induced by a transport map $S = (S_1, S_2)$. At this stage it is not possible to conclude directly since there exist measure preserving maps between μ_0 and ν_0, and even continuous ones such as Peano's function (see for instance [14]). To derive the desired contradiction, one has to use further proposition 3.1 to deduce that S is also an *optimal* transport map between μ_0 and ν_0 for the quadratic cost. Thus, by well-known arguments (see [11], [15], [16]) there exists a convex function $(x,y) \in [0,1]^2 \mapsto \psi(x,y)$ such that $(S_1(x,0), S_2(x,0)) \in \partial\psi(x,0)$ for (Lebesgue)-a.e. $x \in [0,1]$ which easily implies that $S_1(.,0)$ is nondecreasing and since $S_1(.,0)$ preserves the Lebesgue measure this yields $S_1(x,0) = x$. Now since $S(x,0) = (x, S_2(x,0))$ is Borel and measure-preserving between μ_0 and ν_0 we should have

$$\int_0^1 \varphi(x, S_2(x,0)) dx = \int_{[0,1]^2} \varphi(x,y) dx dy$$

for every Borel and bounded $\varphi : [0,1]^2 \to \mathbb{R}$, so that taking $\varphi(x,y) = (y - S_2(x,0))^2$ yields the desired contradiction.

The discrete case, different values for the coupled and decoupled problems: Let us now consider the case where both μ_0 and ν_0 have finite support, in this case (4.8) is in fact a finite-dimensional problem and it is easy to check that it possesses solutions. However, because of atoms, proposition 3.1 does not apply and it is easy to find a case where the values are different. To see this, let us choose $d = 1$, $\mu_0 = \delta_0$ and $\nu_0 = \frac{1}{2}\delta_{-1} + \frac{1}{2}\delta_1$. The set of transport plans reduces to the single element $\gamma = \frac{1}{2}\delta_{(0,1)} + \frac{1}{2}\delta_{(0,-1)}$. Obviously,

the transport plan γ is not induced by any transport map. In this example, the solution of both the coupled and the decoupled problem is the transport plan γ. A simple computation leads to find different values for the two problems. More precisely, for a fixed horizon $T > 0$, one gets the difference: $\inf(\mathcal{Q}) - \inf(\mathcal{P}) = \frac{1}{2}\Big(T - ||a||^2_{H^1}\Big) > 0$. Note that in this very simple case, the difference between the decoupled and coupled problem values is increasing in the horizon T.

References

[1] L. Ambrosio, *Lecture notes on optimal transport problems*, Lecture Notes in Math., 1812, Springer, Berlin, (2003).

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

[3] G. Carlier, *Duality and existence for a class of mass transportation problems and economic applications*, Adv. Math. Econ., 5, 1–21 (2003).

[4] P.A. Chiappori, R.J. Mc Cann, L. Nesheim. *Hedonic price equilibria, stable matching, and optimal transport: equivalence, topology, and uniqueness*, to appear in Econom. Theory.

[5] I. Ekeland, *An optimal matching problem*, ESAIM Control Optim. Calc. Var., 11, no. 1, 57–71(2005).

[6] I. Ekeland, *Existence, uniqueness and efficiency of equilibrium in hedonic markets with multidimensional types*, to appear in Econom. Theory.

[7] L.C. Evans, *Partial Differential Equations*, Grad. Stud. Math., AMS, 1998.

[8] W. H. Fleming, H.M. Soner, *Controlled Markov Processes and Viscosity Solutions*, Graduate Studies in Mathematics, 58, Springer-Verlag, Applications of Mathematics, (1993).

[9] W. Gangbo, *An elementary proof of the polar factorization of vector-valued functions*, Arch. Rational Mech. Anal., 128, 380–399 (1994).

[10] W. Gangbo. Habilitation thesis. Université de Metz, 1995.

[11] W. Gangbo, R.J. McCann, *The Geometry of Optimal Transportation*, Acta Math., 177 , 113–161 (1996).

[12] V.L. Levin, *Abstract cyclical monotonicity and Monge solutions for the general MongeKantorovich problem*, Set-valued Anal., 7, 7–32 (1999).

[13] R.J. McCann, *Existence and uniqueness of monotone measure-preserving maps*, Duke Math. J., 80, 309–323 (1995).

[14] A.H. Stone, *Measure-preserving maps in General topology and its relations to modern analysis and algebra*, IV (Proc. Fourth Prague Topological Sympos., Prague, 1976), Part A, , Lecture Notes in Math., 609, Springer, Berlin, pp. 205–210 1977.

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

[16] C. Villani, *Optimal transport. Old and new*, Springer-Verlag, Berlin, (2009).