On Fluid mechanics formulation of Monge-Kantorovich Mass Transfer Problem

Kazufumi Ito

Center for Research in Scientific Computation
North Carolina State University
Raleigh, North Carolina 27695-8205

Abstract The Monge-Kantorovich mass transfer problem is equivalently formulated as an optimal control prblem for the mass transport equation. The equivalency of the two problems is establish using the Lax-Hopf formula and the optimal control theory arguments. Also, it is shown that the optimal solution to the equivalent control problem is given in a gradient form in terms of the potential solution to the Monge-Kantorovich problem. It turns out that the control formulation is a dual formulation of the Kantrovich distance problem via the Hamilton-Jacobi equations.

1 Introduction

Monge mass transfer problem is that given two probability density functions $\rho_0(x) \geq 0$ and $\rho_1(x) \geq 0$ of $x \in R^d$, find a coordinate map M such that

$$\int_A \rho_1(x)\,dx = \int_{M(x)\in A} \rho_0(x)\,dx \tag{1.1}$$

for all bounded subset A in R^n. If M is a smooth one-to-one map, then it is equivalent to

$$\det(\nabla M)(x)\rho_1(M(x)) = \rho_0(x) \tag{1.2}$$

where det denotes the determinant of Jacobian matrix of the map M. Clearly, this problem is underdetermined and it is natural to formulate a costfuctional for the optimal mass transfer. The so-called Kantorovich (or Wasserstain) distance between ρ_0 and ρ_1 is defined by

$$d(\rho_0,\rho_1) = \inf \int_{R^d} c(x - M(x))\rho_0(x)\,dx. \tag{1.3}$$

where c is a convex function and $c(x-y) = c(|x-y|)$ with $c(0) = 0$. For example $c(x-y) = \frac{1}{p}|x-y|^p$ is for the L^p Monge-Kantorovich problem (MKP). Whenever the infimum is attained by some map M, we say that M is an optimal transfer for the Monge-Kantorovich problem. The Kantorovich distance is the least action that is necessary to transfer ρ_0 into ρ_1.

The mass transport problems have attracted a lot of attentions in recent years and have found applications in many fields of mathematics such as statistics and fluid mechanics (e.g., see [2, 5] and [11] for extensive references). From a more scientific point of view the

Kantorovich distance provides a valuable quantitative informations to compare two different density functions and it has been used in various fields of applications [3].

It is shown e.g., in [1, 4, 8, 5] that the optimal map $\bar{M}$ is given by

$$Dc(x - \bar{M}(x)) = \nabla \bar{u}(x) \tag{1.4}$$

for a potential function $\bar{u}$, where Dc denotes the derivative of c. In fact $\bar{u}$ is the optimal solution to the Kantorovich dual problem (2.2). If c is uniformly convex, then we can solve (1.4) for $\bar{M}$ in terms of $\nabla \bar{u}$. For example for L^p MKP

$$x - \bar{M}(x) = |\nabla \bar{u}(x)|^{q-2}\, \nabla \bar{u}(x) \text{ with } \frac{1}{p} + \frac{1}{q} = 1.$$

For L^2 MKP, it follows from (1.2) and (1.4) that if $\psi = \frac{|x|^2}{2} - \bar{u}(x)$, then $\bar{M}(x) = x - \nabla \bar{u} = \nabla \psi$ and thus ψ satisfies the Monge-Ampere equation

$$\det(H\psi)(x)\rho_1(\nabla \psi) = \rho_0(x), \tag{1.5}$$

where $H\psi$ is the Hessian of ψ.

In [3] the L^2 MKP is equivalently reformulated as an optimal control problem:

$$d(\rho_0, \rho_1) = \min \left[\frac{1}{2} \int_0^1 \int_{R^d} \rho(t,x)\, |V(t,x)|^2 \, dx dt \quad \text{over vector field } V = V(t,x) \right] \tag{1.6}$$

subject to

$$\begin{aligned} &\rho_t + \nabla \cdot (\rho V) = 0 \\ &\rho(0,x) = \rho_0(x) \quad \text{and} \quad \rho(1,x) = \rho_1(x) \end{aligned} \tag{1.7}$$

Moreover if $\bar{V}(t,x)$ is an optimal solution to (1.6)–(1.7) and the Lagrange coordinate $\bar{X}(t;x)$ satisfies

$$\frac{d}{dt}\bar{X} = \bar{V}(t, \bar{X}(t;x)), \quad \bar{X}(0;x) = x,$$

then $\bar{M}(x) = \bar{X}(T;x)$.

The contribution of this paper is that we will show that the optimal vector field $\bar{V}$ to problem (1.6)–(1.7) is given by

$$\bar{V}(x,t) = \nabla_x \bar{\phi}(t,x) \tag{1.8}$$

where the potential function $\bar{\phi}$ satisfies the Hamilton-Jacobi equation

$$\bar{\phi}_t + \frac{1}{2} |\nabla \bar{\phi}|^2 = 0, \quad \bar{\phi}(0,x) = -\bar{u}(x) \tag{1.9}$$

and $\bar{u}$ determines the optimal map $\bar{M}$ in (1.4). Thus, (1.8)–(1.9) is an optimal feedback solution to control problem (1.6)–(1.7), i.e., given ρ_0, ρ_1 first we determine ψ by (1.5) and let $\bar{u} = \dfrac{|x|^2}{2} - \psi$ and then determine $\bar{V}$ by (1.8)–(1.9).

Moreover, it will be shown that

$$d(\rho_0,\rho_1) = \min \quad \left[\int_{R^d} (\rho_1(x)v(x) - \rho_0(x)\phi(0,x))\,dx \text{ over } v\right]$$

subject to

$$\phi_t + \frac{1}{2}\,|\nabla\phi|^2 = 0, \quad \phi(1,x) = v(x).$$

It is the other control formulation of the L^2 MKP and is an optimization problem over the potential fuction v subject to the Hamilton-Jacobi equation.

For the non-quadratic c case, the (generalized) optimal control problem is formulated as

$$\min \quad \int_0^1 \int_{R^d} \rho(t,x)\, c(V(t,x))\, dxdt \tag{1.10}$$

subject to (1.7). In this case the optimal vector field $\bar{V}$ is given by

$$\bar{V} = Dc^*(\nabla\bar{\phi}) \tag{1.11}$$

where c^* is the convex conjugate function of c defined by

$$c^*(y) = \sup_x \, \{x\cdot y - c(x)\}.$$

For L^p MKP

$$c(x) = \frac{1}{p}|x|^p, \ x \in R^d, \quad c^*(y) = \frac{1}{q}|y|^q, \ \ y \in R^d,$$

and

$$\bar{V}(t,x) = |\nabla_x\phi(t,x)|^{q-2}\,\nabla_x\phi(t,x)$$

where $\frac{1}{p}+\frac{1}{q} = 1$ and $p \in (1,\infty)$. The potential function $\phi = \phi(t,x)$ satisfies

$$\bar{\phi}_t + c^*(\nabla\bar{\phi}) = 0, \quad \bar{\phi}(0,x) = -\bar{u}(x). \tag{1.12}$$

If c is uniformly convex, then from (1.4)

$$\bar{M}(x) = x - Dc^*(\nabla\bar{u}(x)).$$

Thus, from (1.2) $\bar{u}$ satisfies

$$\det(\nabla\bar{M})(x)\rho_1(x - Dc^*(\nabla\bar{u}(x))) = \rho_0(x). \tag{1.13}$$

For L^2 MKP (1.13) is reduced to (1.5). Hence the optimal solution to (1.10) subject to (1.7) is given in the feedback form (1.11)-(1.13).

An outline of our presentation is as follows. In Section 2 the basic theoretical results concerning the MKP problem is reviewed following [5]. Then equivalent variational formulations (2.5) and (2.8) for the potential function are then derived using the duality and the Lax-Hoph formula. In Section 3 we present formal arguments that show the feedback solution (1.7)–(1.9) to (1.5)–(1.6). In Section 4 we validate the steps in Section 3 mathmatically for L^2 MKP. In Section 5 we present the proofs for the general case.

2 Variational Formulations

In order to present our treatment of the MKP problem, we first recall a basic theoretical result in this section. The following relaxed problem of (1.3) is introduced by Kantorovich. Let $\mathcal{M}$ be a class of random probability measures μ on $R^d \times R^d$ satisfying $\text{proj}_y \mu = \rho_0\, dx$ and $\text{proj}_x \mu = \rho_1\, dy$. Then we define the relaxed cost-functional

$$J(\mu) = \int_{R^d \times R^d} c(x-y)\, d\mu(x,y) \quad \text{over } \mathcal{M}. \tag{2.1}$$

Consider the dual problem of (2.1); maximize

$$\begin{gathered} \int_{R^d} u(x)\rho_0(x)\, dx + \int_{R^d} v(x)\rho_1(y)\, dy \\ \text{subject to } \ u(x) + v(y) \le c(x-y). \end{gathered} \tag{2.2}$$

The point of course is that the Lagrange multiplier associated with the inequality in (2.2) solves problem (2.1). The following theorem [1, 5, 4, 8] provides the solution to (2.2) and (1.3).

Theorem 2.1
(1) there exists a maximizer $(\bar{u}, \bar{v})$ of problem (2.2).

(2) $(\bar{u}, \bar{v})$ are dual c-conjugate functions, i.e.,

$$\bar{u}(x) = \inf_y \, (c(x-y) - \bar{v}(y))$$

$$\bar{v}(y) = \inf_x \, (c(x-y) - \bar{u}(x))$$

(3) $\bar{M}(x)$ satisfying $Dc(x - \bar{M}(x)) = \nabla \bar{u}(x)$ solves MKP problem.

It follows from Theorem 2.1 that (2.3) is reduced to maximizing

$$J(u) = \int_{R^d} u(x)\rho_0(x)\, dx + \int_{R^d} v(y)\rho_1(y)\, dy \tag{2.4}$$

over functions u, where v is the c-conjugate function of u. The c-conjugate function of a function u is defined by

$$v(y) = \inf_x \, (c(x-y) - u(x)).$$

It is easy to show that the bi c-conjugate function $\tilde{u}$ of u satisfies $\tilde{u} \ge u$ a.e. and thus the maximizing pair (u, v) of (2.4) is automatically c-conjugate each other. Similarly, we have the equivalent problem of maximizing

$$J(v) = \int_{R^d} u(x)\rho_0(x)\, dx + \int_{R^d} v(y)\rho_1(y)\, dy \tag{2.5}$$

where

$$u(x) = \inf_x \,(c(x-y) - v(y)).$$

Let c^* be the convex conjugate of c, i.e.,

$$c^*(x) = \sup_y ((x,y) - c(y)).$$

By the Lax-Hopf formula [6], if ϕ is the viscosity solution to

$$\phi_t + c^*(\nabla\phi) = 0, \quad \phi(1,y) = v(y) \tag{2.6}$$

then

$$\phi(0,x) = \sup_y \,(v(y) - c(x-y)) = -u(x). \tag{2.7}$$

Thus, Problem (2.2) can be equivalently formulated as maximizing

$$J(v) = \int_{R^n} (\rho_1(x)v(x) - \rho_0(x)\phi(0,x))\,dx \tag{2.8}$$

subject to (2.6).

3 Derivation of Optimal Feedback Solution

The optimality condition of (2.8) subject (2.6) is formally derived as follows. We define the Lagrangian

$$L(\phi,\lambda) = J(\phi(1)) - \int_0^1 \int_{R^d} (\phi_t + c^*(\nabla\phi))\lambda\,dxdt. \tag{3.1}$$

By applying the Lagrange multiplier theory the necessary optimality is given by

$$\begin{aligned} L_\phi(\phi,\lambda)(h) &= \int_0^1 \int_{R^d} (\lambda_t + (Dc^*(\nabla\phi)\,\lambda)_x)h\,dxdt \\ &- \int_{R^d} (h(1,x)\lambda(1,x) - h(0,x)\lambda(0,x))\,dx + \int_{R^d} (h(1,x)\rho_1(x) - h(0,x)\rho_0(x))\,dx = 0 \end{aligned} \tag{3.2}$$

for all $h \in C_0^1([0,1]\times R^d)$. Hence the necessary optimality reduces to

$$\begin{aligned} &\lambda_t + (Dc^*(\nabla\bar{\phi})\,\lambda)_x = 0 \\ &\lambda(0,x) = \rho_0(x), \quad \lambda(1,x) = \rho_1(x). \end{aligned} \tag{3.3}$$

This implies that if we let $\bar{V}(t,x) = Dc^*(\nabla\bar{\phi}(t,x))$ in

$$\bar{\rho}_t + (\bar{V}\,\bar{\rho})_x = 0, \quad \bar{\rho}(0,x) = \rho_0,$$

then $\bar{\rho}(1,x) = \rho_1(x)$. Moreover, we can argue that

$$\int_{R^d} (\rho_0 \bar{u}(x) + \rho_1(x)\bar{v}(x))\,dx - \int_0^1 \int_{R^d} \bar{\rho}(t,x)c(\bar{V}(t,x))\,dxdt = 0 \tag{3.4}$$

since

$$c(\bar{V}) = (\nabla_x\bar{\phi}, Dc^*(\nabla_x\bar{\phi})) - c^*(\nabla_x\bar{\phi}).$$

It follows from (3.3)–(3.4) that $\bar{V} = Dc^*(\nabla\bar{\phi})$ is the optimal solution to (1.10) subject to (1.6). In fact, for sufficiently smooth pair (ρ, V) satisfying (1.6), we define Lagrange coordinate $X(t,x)$ by

$$\frac{d}{dt}X = V(t, X(t,x)), \quad X(0,x) = x,$$

Then for all test function f

$$\begin{aligned} &\int_0^1 \int_{R^d} f(t,x)\rho(t,x)\,dxdt = \int_0^1 \int_{R^d} f(t,X(t,x))\rho_0(x)\,dxdt. \\ &\int_0^1 \int_{R^d} f(t,x)\rho(t,x)V(t,x)\,dxdt = \int_0^1 \int_{R^d} V(t,x(t))f(t,X(t,x))\rho_0(x)\,dxdt. \end{aligned} \tag{3.5}$$

Note that (1.6) and (3.5) imply that $M(x) = X(1,x)$ satisfies condition (1.1). Letting $f = c(V)$ in (3.5), we have

$$\begin{aligned} I = \int_0^1 \int_{R^d} \rho(t,x)c(V(t,x))\,dxdt &= \int_0^1 \int_{R^d} c(V(t,X(t,x))\rho_0(x)\,dxdt. \\ = \int_0^1 \int_{R^d} c(\frac{d}{dt}X(t,x))\rho_0(x)\,dxdt &\geq \int c(X(1,x) - X(0,x))\rho_0(x)\,dx. \end{aligned}$$

where we used the Jessen's inequality. Since $\bar{M}(x)$ is the optimal solution to (1.3), it follows that

$$I \geq \int_{R^d} c(x - \bar{M}(x))\rho_0(x)\,dx = d(\rho_0, \rho_1). \tag{3.6}$$

From (3.4), (3.6) and Theorem 2.1

$$d(\rho_0,\rho_1) = \int_{R^d} (\rho_0\bar{u}(x) + \rho_1(x)\bar{v}(x))\,dx = \int_0^1 \int_{R^d} \bar{\rho}(t,x)c(\bar{V}(t,x))\,dxdt \leq I. \tag{3.7}$$

for all pair (ρ, V) satisfying (1.6). That is, $(\bar{\rho}, \bar{V})$ is optimal.

4 Proof of (3.3)–(3.4)

In this section we give a proof for the steps of deriving the optimality condition (3.3) and equality (3.4) in the case when $p=2$, i.e., $c(|x-y|)=\frac{1}{2}\,|x-y|^2$. Suppose $v\in W^{1,\infty}(R^d)$ and v is semi-convex. Then it follows from the Lax-Hopf formula

$$\phi(t,x)=\sup_y\,\{v(y)-(1-t)c(\frac{x-y}{1-t})\}$$

(e.g., see [6]) that (2.6) has a unique solution $\phi\in W^{1,\infty}([0,1]\times R^d)$ with

$$|\phi(t)|_{W^{1,\infty}}\le|v|_{W^{1,\infty}}\quad\text{and}\quad|\phi_t(t)|_{L^\infty}\le\frac{1}{2}|v|^2_{W^{1,\infty}}\tag{4.1}$$

and

$$\phi(t,x+z)-2\phi(t,x)+\phi(t,x-z)\ge -C\,|z|^2\text{ for all }t\in[0,1]\text{ and }x,z\in R^d.\tag{4.2}$$

where we assumed $v+\dfrac{C}{2}|x|^2$ is convex. Let ϕ^τ be the solution to (2.6) with $\phi^\tau(1)=v+\tau\,h$ for $h\in C^2_0(R^d)$. Assume $\phi^\tau(1)+\dfrac{C}{2}|x|^2$ be convex for $|\tau|\le 1$ and thus (4.2) holds for ϕ^τ.

<u>Step 1</u> Since $y\to c(x-y)-v(y)$ is coersive, for each $x\in R^d$ there exist $y,\ y^\tau\in R^d$ such that

$$\phi(0,x)=v(y)-c(x-y),\quad \phi^\tau(0,x)=(v+\tau\,h)(y^\tau)-c(x-y^\tau).$$

Thus

$$\phi(0,x)\ge v(y^\tau)-c(x-y^\tau)=(x+\tau\,h)(y^\tau)-c(x-y^\tau)+\tau\,h(y^\tau)$$

and

$$\phi(0,x)-\phi^\tau(0,x)\ge\tau\,h(y^\tau)$$

Similarly

$$\phi^\tau(0,x)-\phi(0,x)\ge\tau\,h(y).$$

Hence

$$|\phi^\tau(0,\cdot)-\phi(0,\cdot)|_\infty\le\tau\,|h|_\infty.\tag{4.3}$$

<u>Step 2</u> Note that

$$(\phi^\tau-\phi)_t+\frac{1}{2}(\nabla\phi^\tau+\nabla\phi)\cdot(\nabla\phi^\tau-\nabla\phi)=0.\tag{4.4}$$

Let $\eta_\epsilon,\ \epsilon>0$ be the standard molifier. Then

$$|\nabla(\eta_\epsilon*\phi)|_\infty\le|\nabla\phi|_\infty\tag{4.5}$$

and

$$\nabla(\eta_\epsilon * \phi) \to \nabla\phi \quad \text{a.e. as } \epsilon \to 0^+. \tag{4.6}$$

Moreover (4.2) implies

$$D^2(\eta_\epsilon * \phi) \geq -C.$$

Thus

$$(\nabla(\eta_\epsilon * \phi), \nabla\psi) \leq dC \int_{R^d} \psi\, dx$$

for $\psi \in W^{1,1}(R^d)$ and $\psi \geq 0$ a.e. in R^d. Thus from (4.5)–(4.6) and the Lebesgue dominated convergence theorem, letting $\epsilon \to 0^+$

$$(\nabla\phi, \nabla\psi) \leq dC \int_{R^d} \psi\, dx \tag{4.7}$$

for $\psi \in W^{1,1}(R^d)$ and $\psi \geq 0$ a.e. in R^d. It now follows from (4.4) and (4.7) that

$$\frac{d}{dt}|\phi^\tau(t,\cdot) - \phi(t,\cdot)|_1 \geq -dC\,|\phi^\tau(t,\cdot) - \phi(t,\cdot)|_1, \quad \phi^\tau(1) = \phi(1) + \tau\, h$$

and thus

$$|\phi^\tau(0,\cdot) - \phi(0,\cdot)|_1 \leq \tau\, e^{dC}\,|h|_1. \tag{4.8}$$

Since from (2.6)

$$\int_0^1 \int_{R^d} \frac{1}{2}|\nabla\phi^\tau|^2\, dxdt = \int_{R^d} (\phi^\tau(0,x) - v(x))\, dx,$$

we have

$$\int_0^1 \int_{R^d} |\nabla\phi^\tau|^2\, dxdt \to \int_0^1 \int_{R^d} |\nabla\phi|^2\, dxdt.$$

as $\tau \to 0$. Since $L^2((0,1) \times R^d)$ is a Hilbert space, this implies that

$$\int_0^1 \int_{R^d} |\nabla\phi^\tau - \nabla\phi|^2\, dxdt \to 0 \tag{4.9}$$

as $\tau \to 0$.

Step 3 For $\epsilon > 0$ let us consider

$$\lambda_t + (\nabla\phi\,\lambda)_x = \epsilon\,\Delta\lambda, \quad \lambda(0) = \rho_0$$

Since ϕ is Lipschitz on $[0,1] \times R^d$,

$$t \to \int_{R^d} [\epsilon\,(\nabla\lambda, \nabla\psi) - (\nabla\phi(t,\cdot)\lambda, \nabla\psi)]\, dx$$

defines an integrable, bounded, coersive form on $H^1(R^d)\times H^1(R^d)$ and thus it follows from the parabolic equation theory (e.g., see [12, 10]) that there exits a unique solution $\lambda_\epsilon \in H^1(0,1;L^2(R^d))\cap L^2(0,1;H^2(R^d))$ provided that $\rho_0\in H^1(R^d)\cap L^1(R^d)\cap L^\infty(R^d)$. Moreover

$$|\lambda_\epsilon(t)|_1 \le |\rho_0|_1,$$

$$\frac{1}{2}\left(|\lambda_\epsilon(1)|_2^2-|\rho_0|_2^2\right)+\epsilon\int_0^1 |\nabla\lambda_\epsilon(t)|_2^2\,dt=0 \tag{4.10}$$

$$|\lambda_\epsilon(t)|_\infty \le e^{Ct}\,|\rho_0|_\infty.$$

For the last estimate we have from (4.7)

$$\frac{1}{p}\frac{d}{dt}\,|\lambda_\epsilon|^p \le \frac{p-1}{p}(\nabla\phi,\nabla|\lambda_\epsilon|^p)\le \frac{(p-1)dC}{p}\,|\lambda_\epsilon|^p.$$

for $p\ge 1$ and thus

$$|\lambda_\epsilon|_p \le e^{\frac{p-1}{p}dCt}\,|\rho_0|_p.$$

Thus λ_ϵ is uniformly bounded in $L^2((0,1)\times R^d)$. Hence there exists a $\lambda\in L^\infty(0,1;L^1(R^d)\cap L^\infty(R^d))$ and subsequaence of λ_ϵ (denoted by the same) such that λ_ϵ converges weakly to λ in $L^2((0,1)\times R^d)$ and $\lambda_\epsilon(1)\to\lambda(1)$ in $L^2(\Omega)$. Since for $\psi\in C_0^1([0,1]\times R^d)$

$$\int_0^1\int_{R^d}(\lambda_\epsilon\psi_t+\lambda_\epsilon\nabla\phi\cdot\nabla\psi-\epsilon\,\nabla\lambda_\epsilon\cdot\nabla\psi)\,dxdt=\int_{R^d}(\rho_0\psi(0)-\lambda_\epsilon(1)\psi(1))\,dx \tag{4.11}$$

it follows from (4.10)–(4.11) that letting $\epsilon\to 0^+$

$$\int_0^1\int_{R^d}(\lambda\psi_t+\lambda\nabla\phi\cdot\nabla\psi)\,dxdt=\int_{R^d}(\rho_0\psi(0)-\lambda(1)\psi(1))\,dx. \tag{4.12}$$

Hence λ is a weak solution to

$$\lambda_t+(\lambda\,\nabla\phi)=0,\quad \lambda(0)=\rho_0. \tag{4.13}$$

Next we show that (4.13) has the weak unique solution in $L^\infty(0,1;L^1(R^d)\cap L^1(R^d))$. Let $\eta_\epsilon,\ \epsilon>0$ be the standard molifier and consider the adjoint equation

$$\psi_t+\nabla(\eta_\epsilon*\phi)\cdot\nabla\psi=f\in C_0^\infty([0,1]\times R^d),\quad \psi(1)=0. \tag{4.14}$$

Then, (4.14) has a smooth unique sulution ψ and $J=|\nabla\psi|$ satisfies

$$J_t+\nabla(\eta_\epsilon*\phi)\cdot\nabla J+D^2(\eta_\epsilon*\phi)J=\nabla f,\quad J(1)=0. \tag{4.15}$$

Since ψ has compact support, J has a positive maximum over $[0,1]\times R^d$ at some point (t_0,x_0). If $0\le t_0<1$, then from (4.15)

$$J_t(t_0,x_0)\le 0 \text{ and } \nabla J(t_0,x_0)=0.$$

Thus,

$$D^2(\eta_\epsilon * \phi)J(t_0,x_0) \geq \nabla f(t_0,x_0)$$

Since from (4.2) $D^2(\eta_\epsilon * \phi) \leq -C$, this implies

$$|\nabla\psi|_\infty = J(t_0,x_0) \leq \frac{|\nabla f|_\infty}{C} \tag{4.16}$$

Let $\lambda,\ \tilde{\lambda}$ is two weak solutions to (4.13). Then, it folows from (4.12) and (4.14) that

$$\int_0^1 \int_{R^d} (\lambda-\tilde{\lambda})f\,dxdt = \int_0^1 \int_{R^d} (\lambda-\tilde{\lambda})(\nabla(\eta_\epsilon * \phi) - \nabla\phi)\nabla\psi\,dxdt.$$

By letting $\epsilon \to 0^+$, it follows from (4.5)–(4.6), (4.16) and the Lebesgue dominated convergence theorem that

$$\int_0^1 \int_{R^d} (\lambda-\tilde{\lambda})f\,dxdt = 0$$

for all $f \in C_0^\infty([0,1]\times R^d)$ and therfore $\lambda = \tilde{\lambda}$.

Now, let λ^τ be the solution to (4.13) associated with ϕ^τ. Since λ^τ is uniformly bounded in $L^\infty(0,1;L^1(R^d)\cap L^\infty(R^d))$, there exists a $\lambda^* \in L^\infty(0,1;L^1(R^d)\cap L^\infty(R^d))$ such that λ^τ converges weakly to λ^* in $L^2((0,1)\times R^d)$ and $\lambda^\tau(1)$ converges weakly to $\lambda^*(1)$ in $L^2(R^d)$ as $\tau \to 0$. Note that

$$\int_0^1 \int_{R^d} (\lambda^\tau\,\psi_t + (\lambda^\tau\nabla\phi + \lambda^\tau(\nabla\phi^\tau - \nabla\phi)\cdot\nabla\psi)\,dxdt = \int_{R^d} (\rho_0\psi(0) - \lambda^\tau(1)\psi(1))\,dx. \tag{4.17}$$

Since λ^τ is uniformly bounded in $L^\infty(0,1;L^1(R^d)\cap L^\infty(R^d))$, it follows from (4.9) and (4.17) that λ^* is the weak solution to (4.13). Since (4.13) has the unique weak solution, we conclude λ^τ converges weakly to λ in $L^2((0,1)\times R^d)$ and $\lambda^\tau(1)$ weakly to $\lambda(1)$ in $L^2(R^d)$ as $\tau \to 0$.

Step 4 Note that

$$\int_0^1 \int_{R^d} \lambda^\tau(\psi_t + \nabla\phi^\tau\cdot\nabla\psi)\,dxdt = \int_{R^d} (\psi(0)\rho_0 - \psi(1)\lambda^\tau(1))\,dx. \tag{4.18}$$

for all $\psi \in W^{1,\infty}((0,1)\times R^d)$. Since

$$(\phi^\tau - \phi)_t + \nabla\phi^\tau\cdot\nabla(\phi^\tau - \phi) - \frac{1}{2}|\nabla\phi^\tau - \nabla\phi|^2 = 0$$

by setting $\psi = \phi^\tau - \phi$ in (4.18), we obtain

$$\int_{R^d} ((\phi^\tau(0) - \phi(0))\rho_0(x) - \tau\,\lambda^\tau(1)h(x))\,dx = \frac{1}{2}\int_0^1 \int_{R^d} \lambda^\tau\,|\nabla\phi^\tau - \nabla\phi|^2\,dxdt$$

Similarly, since

$$(\phi^\tau - \phi)_t + \nabla\phi\cdot\nabla(\phi^\tau - \phi) + \frac{1}{2}|\nabla\phi^\tau - \nabla\phi|^2 = 0,$$

we have

$$\int_{R^d} ((\phi^\tau(0) - \phi(0))\rho_0(x) - \tau\,\lambda(1)h(x))\,dx = -\frac{1}{2}\int_0^1 \int_{R^d} \lambda\,|\nabla\phi^\tau - \nabla\phi|^2\,dxdt.$$

From (4.3) and (4.8) there exists a subsequence of $\dfrac{\phi^\tau(0) - \phi}{\tau}$ that converges weakly in $L^2(R^d)$ as $\tau \to 0$. Since $\lambda,\ \lambda^\tau \geq 0$ a.e. in $(0,1) \times R^d$ and $\lambda^\tau(1)$ converges weakly to $\lambda(1)$ as $\tau \to 0$, we have

$$\lim_{\tau\to 0} \int_{R^d} \frac{\phi^\tau(0) - \phi(0)}{\tau}\rho_0(x)\,dx = \int_{R^d} \lambda(1)h(x)\,dx.$$

Hence

$$J'(v)(h) = \int_{R^d} (\rho_1 - \lambda(1))h(x)\,dx. \tag{4.19}$$

Step 5 Assume $\bar{v}$ attains the minimum of $J(v)$ in (2.8) and $\bar{v}$ is Lipshitz and semi-convex. Then $J'(\bar{v})(h) = 0$ for all $h \in C_0^2(R^d)$ and thus from (4.19) $\bar{\lambda}(1) = \rho_1$ a.e., where $\bar{\lambda}$ is the weak solution to (4.13) with $\phi = \bar{\phi}$. Thus, (3.3) holds with $\rho = \bar{\lambda}$. Since

$$\phi_t + \nabla\phi \cdot \nabla\phi - \frac{1}{2}|\nabla\phi|^2 = 0,$$

it follows from (4.12) with $\psi = \phi$ that

$$\int_{R^d} (\rho_0(x)\bar{u}(x) + \rho_1(x)\bar{v}(x))\,dx = \int_0^1 \int_{R^d} \frac{1}{2}\,\bar{\lambda}(t,x)|\bar{V}(t,x)|^2\,dxdt.$$

which shows (3.4).

5 General Case

In this section we prove (3.3)–(3.4) for the general case $c(x - y) = \dfrac{1}{p}|x - y|^p,\ 1 < p < \infty$. Assume v is Lipschitz and semi-convex. It the follows from [5] that

$$\phi_t + \frac{1}{q}|\nabla\phi|^q = 0, \quad \phi(1,x) = v(x) \tag{5.1}$$

has a unique viscosity solution $\phi \in W^{1,\infty}((0,1) \in R^d)$ satisfying (4.2).

Step 1 For $h \in C_0^2(R^d)$ let $v^\tau = v + \tau\,h$. If for $x \in R^d$, $y^\tau = y^\tau(x) \in R^d$ attains the maximum of $y \to v^\tau(y) - c(x - y)$, then

$$x - y^\tau = -|\nabla v^\tau(y^\tau)|^{q-2}\nabla v^\tau(y^\tau)$$

Since $v^\tau \in W^{1,\infty}(R^d)$, $|y^\tau(x)-x| \le \alpha$ for some α uniformly in x and τ. Since as shown in Section 4

$$\tau\, h(y) \le \pi^\tau(0,x)-\phi(0,x) \le -\tau h(y^\tau),$$

$$\int_{R^d} |\phi^\tau(0,x)-\phi(0,x)|\,dx = \tau \int_{R^d} (|h(y(x))|+|h(y^\tau(x))|)\,dx.$$

Since h is compactly supported, it follws that there exists a constant M (depends on h) such that

$$\int_{R^d} |\phi^\tau(0,x)-\phi(0,x)|\,dx \le M\,\tau. \tag{5.2}$$

Since from (5.1)

$$\int_0^1 \int_{R^d} \frac{1}{q}\,|\nabla\phi^\tau|^q\,dxdt = \int_{R^d} (\phi^\tau(0,x)-\phi^\tau(1,x))\,dx,$$

we have

$$\int_0^1 \int_{R^d} (|\nabla\phi^\tau|^q - |\nabla\phi|^q)\,dx \to 0$$

as $\tau \to 0$. Hence

$$\int_0^1 \int_{R^d} ||\nabla\phi^\tau|^{\frac{q-2}{2}}\nabla\phi^\tau|^2\,dxdt \to \int_0^1 \int_{R^d} ||\nabla\phi|^{\frac{q-2}{2}}\nabla\phi|^2\,dxdt$$

as $\tau \to 0$ and therefore

$$|\nabla\phi^\tau|^{\frac{q-2}{2}}\nabla\phi^\tau \to |\nabla\phi|^{\frac{q-2}{2}}\nabla\phi \quad \text{in } L^2((0,1)\times R^d).$$

Moreover, there exists a subsequence (denoted by the same) of τ such that $\nabla\phi^\tau(x) \to \nabla\phi(x)$ a.e. in R^d. Since by the Lebesgue dominated convergence theorem

$$|\nabla\phi^\tau|^{q-2}\nabla\phi^\tau \to |\nabla\phi|^{q-2}\nabla\phi \quad \text{in } L^2((0,1)\times R^d). \tag{5.3}$$

Step 2 We assume that for $p>2$ (i.e., $q<2$) $|\nabla v(\cdot)|^2 \ge c > 0$ a.e. in R^d. Then it follows from (5.1) that $|\nabla\phi(t,\cdot)|^2 \ge c$ a.e. in R^d for $t \in [0,1]$. Let

$$J_\epsilon = |\nabla(\eta_\epsilon * \phi)|^{q-2}\nabla(\eta_\epsilon * \phi)$$

Then,

$$\nabla\cdot J_\epsilon = |\nabla(\eta_\epsilon*\phi)|^{q-4}(|\nabla(\eta_\epsilon*\phi)|^2\,\Delta(\eta_\epsilon*\phi) + (q-2)\,\nabla(\eta_\epsilon*\phi)^t[D^2(\eta_\epsilon*\phi)]\nabla(\eta_\epsilon*\phi))$$

Since from (4.2) $D^2(\eta_\epsilon*\phi) \ge -C$, there exists a positive constant C_q such that $\nabla\cdot J_\epsilon \ge -C_q$ and thus

$$(J_\epsilon, \nabla\psi) \le C_q \int_{R^d} \psi\,dx$$

for $\psi \in W^{1,1}(R^d)$ and $\psi \geq 0$ a.e. in R^d. It thus follows from (4.5)–(4.6) and the Lebesgue dominated convergence theorem that

$$(|\nabla\phi|^{q-2}\nabla\phi, \nabla\psi) \leq C_q \int_{R^d} \psi\, dx. \tag{5.4}$$

for $\psi \in W^{1,1}(R^d)$ and $\psi \geq 0$ a.e. in R^d, by letting $\epsilon \to 0^+$.

Step 3 Using the same arguments as in Step 3 in Section 4,

$$\lambda_t + (|\nabla\phi^\tau|^{q-2}\nabla\phi^\tau\, \lambda)_x = 0$$

has the unique weak solution $\lambda^\tau \in L^\infty(0,1; L^1(R^d) \cap L^\infty(R^d))$, i.e.,

$$\int_0^1 \int_{R^d} \lambda^\tau(\psi_t + |\nabla\phi^\tau|^{q-2}\nabla\phi^\tau \cdot \nabla\psi)\, dxdt = \int_{R^d} (\psi(0)\rho_0 - \psi(1)\lambda^\tau(1))\, dx. \tag{5.5}$$

for all $\psi \in W^{1,\infty}((0,1)\times R^d)$. Moreover λ^τ converges weakly to λ in $L^2((0,1)\times R^d)$ and $\lambda^\tau(1)$ converges weakly to $\lambda(1)$ in $L^2(R^d)$ as $\tau \to 0$. Since

$$(\phi^\tau - \phi)_t + |\nabla\phi^\tau|^{q-2}\nabla\phi^\tau \cdot \nabla(\phi^\tau - \phi) + I_1 = 0$$

where

$$I_1 = \frac{1}{q}|\nabla\phi^\tau|^q - \frac{1}{q}|\nabla\phi|^q - |\nabla\phi^\tau|^{q-2}\nabla\phi^\tau \cdot (\nabla\phi^\tau - \nabla\phi) \leq 0$$

By setting $\psi = \phi^\tau - \phi$ in (4.5), we obtain

$$\int_{R^d} ((\phi^\tau(0) - \phi(0))\rho_0(x) - \tau\, \lambda^\tau(1)h(x))\, dx = -\int_{R^d} \lambda^\tau\, I_1\, dx$$

Similarly, since

$$(\phi^\tau - \phi)_t + |\nabla\phi|^{q-2}\nabla\phi \cdot \nabla(\phi^\tau - \phi) + I_2 = 0$$

where

$$I_2 = \frac{1}{q}|\nabla\phi^\tau|^q - \frac{1}{q}|\nabla\phi|^q - |\nabla\phi|^{q-2}\nabla\phi \cdot (\nabla\phi^\tau - \nabla\phi) \geq 0,$$

we have

$$\int_{R^d} ((\phi^\tau(0) - \phi(0))\rho_0(x) - \tau\, q(1)h(x))\, dx = -\int \lambda\, I_2\, dx.$$

Since $\lambda,\ \lambda^\tau \geq 0$ a.e. in $(0,1)\times R^d$ and $\lambda^\tau(1)$ converges weakly to $\lambda(1)$ as $\tau \to 0$, it follws that

$$\lim_{\tau\to 0} \int_{R^d} \frac{\phi^\tau(0) - \phi(0)}{\tau} \rho_0(x)\, dx = \int_{R^d} \lambda(1)h(x)\, dx.$$

Thus

$$J'(v)(h) = \int_{R^d} (\rho_1 - \lambda(1))h(x)\, dx. \tag{5.6}$$

Assume $\bar{v}$ attains the minimum of $J(v)$ in (2.8) and $\bar{v}$ is Lipshitz and semi-convex. Then, $J'(\bar{v})(h) = 0$ for all $h \in C_0^2(R^d)$ and therrrfore from (5.6) $\bar{\lambda}(1) = \rho_1$ where $\bar{\lambda}$ is the weak solution to

$$\lambda_t + (\lambda\, |\nabla\bar{\phi}|^{q-2}\nabla\phi)_x = 0, \quad \lambda(0) = \rho_0 \tag{5.7}$$

Thus, (3.3) holds with $\rho = \bar{\lambda}$. Since

$$\phi_t + |\nabla\phi|^{q-2}\nabla\phi\cdot\nabla\phi - \frac{1}{p}|\nabla\phi|^q = 0,$$

it follows from (5.5) with $\psi = \phi$ that

$$\int_{R^d} (\rho_0(x)\bar{u}(x) + \rho_1(x)\bar{v}(x))\,dx = \int_0^1 \int_{R^d} \frac{1}{p}\,\bar{\lambda}(t,x)|\bar{V}(t,x)|^p\,dxdt.$$

which shows (3.4).

References

[1] Y. Brenier, Decomposposition polaire et rearrangement monotone des champs de vecteurs, C.R. Acad. Sci. Paris, Series I, Math. 305 (1987), 805-808.

[2] Y. Brenier, The dual least action problem for an ideal, compressible fluid, Arch. Rat. Mech. Analysis 122 (1993), 322-351.

[3] J.D. Benamou and Y. Brenier, A computational fluid flow the Monge-Kantrovich mass transfer problem, Numer. Math., 84 (2000), 375-393.

[4] L. Caffarelli, Allocation maps with general cost functions, in Partiall Differential Equations and Applications, (P. Marcellini, G. Talenti and E. Vesintini eds), Lecture Notes in Pure and appl. Math., 177 (1996), 29-35.

[5] L.C. Evans, Partial Differential Equations and Monge–Kantorovich Mass Transfer, preprint.

[6] L.C. Evans, *Partial Differntial Equations*, AMS, Providence (1998).

[7] I. Ekeland and T. Turnbull, *Infinite-dimensional Optimization and Convexity*, University of Chicago Press, Chicago (1983).

[8] W.Gangbo and R.J. McCann, Optimal maps in Monge's mass transport problem, C.R. Acad. Sci. Paris, Series I, Math. 321 (1995), 1653-1658.

[9] K. Ito, An Optimal Optical Flow, SIAM J. Control and Optimization (2003), submitted.

[10] K.Ito and F. Kappel *Evolution Equations and Approximations*, World Scientific, Singapore (2002).

[11] S.T.Rachev and L. Rüschendorf, *Mass transfer problems, Vol I and II, Probability and Its Applications*, Springer-Velag, New York (1998).

[12] H.Tanabe, *Equations of Evolution*, Pitman, London (1979).