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 problem for the mass transport equation. The equivalency of the two problems is established 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 $x \in \mathbb{R}^d$, find a coordinate map $M$ such that

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

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

$$\det(\nabla M)(x)\rho_1(M(x)) = \rho_0(x)$$

where $\det$ denotes the determinant of Jacobian matrix of the map $M$. Clearly, this problem is underdetermined and it is natural to formulate a cost functional for the optimal mass transfer. The so-called Kantorovich (or Wasserstain) distance between $\rho_0$ and $\rho_1$ is defined by

$$d(\rho_0, \rho_1) = \inf \int_{\mathbb{R}^d} c(x - M(x))\rho_0(x) \, dx.$$
Kantorovich distance provides a valuable quantitative information 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

\[
(1.4) \quad Dc(x - \bar{M}(x)) = \nabla \bar{u}(x)
\]

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

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

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 \( \psi \) satisfies the Monge-Ampere equation

\[
(1.5) \quad \det(H\psi)(x)\rho_1(\nabla \psi) = \rho_0(x),
\]

where \( H\psi \) is the Hessian of \( \psi \).

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

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

subject to

\[
(1.7) \quad \rho_t + \nabla \cdot (\rho V) = 0
\]

\[
\rho(0, x) = \rho_0(x) \quad \text{and} \quad \rho(1, x) = \rho_1(x)
\]

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

\[
(1.8) \quad \bar{V}(x, t) = \nabla_x \bar{\phi}(t, x)
\]

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

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

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 \( \rho_0, \rho_1 \) first we determine \( \psi \) by (1.5) and let \( \bar{u} = \frac{|x|^2}{2} - \psi \) and then determine \( \bar{V} \) by (1.8)–(1.9).
Moreover, it will be shown that
\[ d(\rho_0, \rho_1) = \min \left[ \int_{\mathbb{R}^d} (\rho_1(x)v(x) - \rho_0(x)\phi(0, x)) \, dx \text{ over } v \right] \]
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 function \( v \) subject to the Hamilton-Jacobi equation.

For the non-quadratic \( c \) case, the (generalized) optimal control problem is formulated as
\[ \min \int_0^1 \int_{\mathbb{R}^d} \rho(t, x) c(V(t, x)) \, dx \, dt \]
subject to (1.7). In this case the optimal vector field \( \tilde{V} \) is given by
\[ \tilde{V} = Dc^*(\nabla \tilde{\phi}) \]
where \( c^* \) is the convex conjugate function of \( c \) defined by
\[ c^*(y) = \sup_x \left\{ x \cdot y - c(x) \right\}. \]
For \( L^p \) MKP
\[ c(x) = \frac{1}{p} |x|^p, \quad x \in \mathbb{R}^d, \quad c^*(y) = \frac{1}{q} |y|^q, \quad y \in \mathbb{R}^d, \]
and
\[ \tilde{V}(t, x) = |\nabla_x \phi(t, x)|^{q-2} \nabla_x \phi(t, x) \]
where \( \frac{1}{p} + \frac{1}{q} = 1 \) and \( p \in (1, \infty) \). The potential function \( \tilde{\phi} = \phi(t, x) \) satisfies
\[ \tilde{\phi}_t + c^*(\nabla \tilde{\phi}) = 0, \quad \tilde{\phi}(0, x) = -\tilde{u}(x). \]
If \( c \) is uniformly convex, then from (1.4)
\[ \tilde{M}(x) = x - Dc^*(\nabla \tilde{u}(x)). \]
Thus, from (1.2) \( \tilde{u} \) satisfies
\[ \det(\nabla \tilde{M})(x)\rho_1(x - Dc^*(\nabla \tilde{u}(x))) = \rho_0(x). \]
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-Hopf 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 mathematically 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 $\mu$ on $\mathbb{R}^d \times \mathbb{R}^d$ satisfying $\text{proj}_y \mu = \rho_0 \, dx$ and $\text{proj}_x \mu = \rho_1 \, dy$. Then we define the relaxed cost-functional

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

Consider the dual problem of (2.1); maximize

\[(2.2) \quad \int_{\mathbb{R}^d} u(x) \rho_0(x) \, dx + \int_{\mathbb{R}^d} v(y) \rho_1(y) \, dy \]

subject to $u(x) + v(y) \leq c(x - y)$.

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

\[(2.4) \quad J(u) = \int_{\mathbb{R}^d} u(x) \rho_0(x) \, dx + \int_{\mathbb{R}^d} v(y) \rho_1(y) \, dy \]

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 $\tilde{u} \geq u$ 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

\[(2.5) \quad J(v) = \int_{\mathbb{R}^d} u(x) \rho_0(x) \, dx + \int_{\mathbb{R}^d} v(y) \rho_1(y) \, dy\]
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 \( \phi \) is the viscosity solution to
\[
\phi_t + c^*(\nabla \phi) = 0, \quad \phi(1, y) = v(y)
\]
then
\[
\phi(0, x) = \sup_y (v(y) - c(x - y)) = -u(x).
\]

Thus, Problem (2.2) can be equivalently formulated as maximizing
\[
J(v) = \int_{\mathbb{R}^n} (\rho_1(x)v(x) - \rho_0(x)\phi(0, x)) \, dx
\]
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
\[
L(\phi, \lambda) = J(\phi(1)) - \int_0^1 \int_{\mathbb{R}^d} (\phi_t + c^*(\nabla \phi)) \lambda \, dx dt.
\]

By applying the Lagrange multiplier theory the necessary optimality is given by
\[
L_{\phi}(\phi, \lambda)(h) = \int_0^1 \int_{\mathbb{R}^d} (\lambda_t + (Dc^*(\nabla \phi) \lambda)_x) h \, dx dt
\]
\[
- \int_{\mathbb{R}^d} (h(1, x)\lambda(1, x) - h(0, x)\lambda(0, x)) \, dx + \int_{\mathbb{R}^d} (h(1, x)\rho_1(x) - h(0, x)\rho_0(x)) \, dx = 0
\]
for all \( h \in C^1_0([0,1] \times \mathbb{R}^d) \). Hence the necessary optimality reduces to
\[
\lambda_t + (Dc^*(\nabla \phi) \lambda)_x = 0
\]
\[
\lambda(0, x) = \rho_0(x), \quad \lambda(1, x) = \rho_1(x).
\]

This implies that if we let \( \tilde{V}(t, x) = Dc^*(\nabla \tilde{\phi}(t, x)) \) in
\[
\tilde{\rho}_t + (\tilde{V} \tilde{\rho})_x = 0, \quad \tilde{\rho}(0, x) = \rho_0,
\]

\[ 5 \]
then $\bar{\rho}(1, x) = \rho_1(x)$. Moreover, we can argue that
\begin{equation}
\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)) \, dx dt = 0
\end{equation}
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 $\bar{V} = Dc^*(\nabla_x \bar{\phi})$ is the optimal solution to (1.10) subject to (1.6). In fact, for sufficiently smooth pair ($\rho, V$) 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 $f$
\[
\int_0^1 \int_{R^d} f(t, x) \rho(t, x) \, dx dt = \int_0^1 \int_{R^d} f(t, X(t, x)) \rho_0(x) \, dx dt.
\]
(3.5)
\[
\int_0^1 \int_{R^d} f(t, x) \rho(t, x)V(t, x) \, dx dt = \int_0^1 \int_{R^d} V(t, x(t)) f(t, X(t, x)) \rho_0(x) \, dx dt.
\]
Note that (1.6) and (3.5) imply that $\bar{M}(x) = X(1, x)$ satisfies condition (1.1). Letting $f = c(V)$ in (3.5), we have
\[
I = \int_0^1 \int_{R^d} \rho(t, x)c(V(t, x)) \, dx dt = \int_0^1 \int_{R^d} c(V(t, X(t, x))) \rho_0(x) \, dx dt.
\]
(3.6)
\[
= \int_0^1 \int_{R^d} c\left(\frac{d}{dt} X(t, x)\right) \rho_0(x) \, dx dt \geq \int c(X(1, x) - X(0, x)) \rho_0(x) \, dx.
\]

where we used the Jessen’s inequality. Since $\bar{M}(x)$ is the optimal solution to (1.3), it follows that
\[
I \geq \int_{R^d} c(x - \bar{M}(x)) \rho_0(x) \, dx = d(\rho_0, \rho_1).
\]
From (3.4), (3.6) and Theorem 2.1
\[
d(\rho_0, \rho_1) = \int_{R^d} (\rho_0 \bar{u}(x) + \rho_1(x) \bar{v}(x)) \, dx = \int_0^1 \int_{R^d} \bar{\rho}(t, x)c(\bar{V}(t, x)) \, dx dt \leq I.
\]
(3.7)
for all pair ($\rho, V$) 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 [3]) that (2.6) has a unique solution \( \phi \in W^{1,\infty}([0,1] \times \mathbb{R}^d) \) with

\[
|\phi(t)|_{W^{1,\infty}} \leq |v|_{W^{1,\infty}} \quad \text{and} \quad |\phi_t(t)|_{L^\infty} \leq \frac{1}{2} |v|_{W^{1,\infty}}^2
\]

and

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

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

**Step 1** Since \( y \to c(x - y) - v(y) \) is coercive, for each \( x \in \mathbb{R}^d \) there exist \( y, y^\tau \in \mathbb{R}^d \) 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

\[
\phi(0, x) \geq v(y^\tau) - c(x - y^\tau) = (x + \tau h)(y^\tau) - c(x - y^\tau) + \tau h(y^\tau)
\]

and

\[
\phi(0, x) - \phi^\tau(0, x) \geq \tau h(y^\tau)
\]

Similarly

\[
\phi^\tau(0, x) - \phi(0, x) \geq \tau h(y).
\]

Hence

\[
|\phi^\tau(0, \cdot) - \phi(0, \cdot)|_{\infty} \leq \tau |h|_{\infty}.
\]

**Step 2** Note that

\[
(\phi^\tau - \phi)_t + \frac{1}{2}(\nabla \phi^\tau + \nabla \phi) \cdot (\nabla \phi^\tau - \nabla \phi) = 0.
\]

Let \( \eta_\epsilon, \epsilon > 0 \) be the standard mollifier. Then

\[
|\nabla (\eta_\epsilon \ast \phi)|_{\infty} \leq |\nabla \phi|_{\infty}
\]

and

\[
\nabla (\eta_\epsilon \ast \phi) \to \nabla \phi \quad \text{a.e. as } \epsilon \to 0^+.
\]

Moreover (4.2) implies

\[
D^2(\eta_\epsilon \ast \phi) \geq -C.
\]

Thus

\[
(\nabla (\eta_\epsilon \ast \phi), \nabla \psi) \leq dC \int_{\mathbb{R}^d} \psi \, dx
\]
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 \( \epsilon \to 0^+ \)

\[
(4.7) \quad (\nabla \phi, \nabla \psi) \leq dC \int_{R^d} \psi \, dx
\]

for \( \psi \in W^{1,1}(R^d) \) and \( \psi \geq 0 \) a.e. in \( R^d \). 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

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

Since from (2.6)

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

we have

\[
\int_0^1 \int_{R^d} |\nabla \phi^\tau|^2 \, dx \, dt \to \int_0^1 \int_{R^d} |\nabla \phi|^2 \, dx \, dt.
\]

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

\[
(4.9) \quad \int_0^1 \int_{R^d} |\nabla \phi^\tau - \nabla \phi|^2 \, dx \, dt \to 0
\]

as \( \tau \to 0 \).

**Step 3** For \( \epsilon > 0 \) let us consider

\[
\lambda_t + (\nabla \rho_t)_x = \epsilon \Delta \lambda, \quad \lambda(0) = \rho_0
\]

Since \( \phi \) 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, coercive 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

\[
|\lambda_\epsilon(t)|_1 \leq |\rho_0|_1,
\]

\[
(4.10) \quad \frac{1}{2} (|\lambda_\epsilon(1)|_2^2 - |\rho_0|_2^2) + \epsilon \int_0^1 |\nabla \lambda_\epsilon(t)|_2^2 \, dt = 0
\]

\[
|\lambda_\epsilon(t)|_\infty \leq e^{Ct} |\rho_0|_\infty.
\]
For the last estimate we have from (4.7)
\[
\frac{1}{p} \frac{d}{dt} |\lambda_\epsilon|^p \leq \frac{p-1}{p} (\nabla \phi, \nabla |\lambda_\epsilon|^p) \leq \frac{(p-1)dC}{p} |\lambda_\epsilon|^p.
\]
for \(p \geq 1\) and thus
\[
|\lambda_\epsilon|^p \leq e^{\frac{p-1}{p}dCt} |\rho_0|^p.
\]
Thus \(\lambda_\epsilon\) 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 subsequence of \(\lambda_\epsilon\) (denoted by the same) such that \(\lambda_\epsilon\) converges weakly to \(\lambda\) in \(L^2((0,1) \times R^d)\) and \(\lambda_\epsilon(1) \to \lambda(1)\) in \(L^2(\Omega)\). Since for \(\psi \in C_0^1([0,1] \times R^d)\)
\[
(4.11) \quad \int_0^1 \int_{R^d} (\lambda_\epsilon \psi_t + \lambda_\epsilon \nabla \phi \cdot \nabla \psi - \epsilon \nabla \lambda_\epsilon \cdot \nabla \psi) \, dx \, dt = \int_{R^d} (\rho_0 \psi(0) - \lambda_\epsilon(1) \psi(1)) \, dx
\]
it follows from (4.10)–(4.11) that letting \(\epsilon \to 0^+\)
\[
(4.12) \quad \int_0^1 \int_{R^d} (\lambda \psi_t + \lambda \nabla \phi \cdot \nabla \psi) \, dx \, dt = \int_{R^d} (\rho_0 \psi(0) - \lambda(1) \psi(1)) \, dx.
\]
Hence \(\lambda\) is a weak solution to
\[
(4.13) \quad \lambda_t + (\lambda \nabla \phi) = 0, \quad \lambda(0) = \rho_0.
\]

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 mollifier and consider the adjoint equation
\[
(4.14) \quad \psi_t + \nabla (\eta_\epsilon \ast \phi) \cdot \nabla \psi = f \in C_0^\infty([0,1] \times R^d), \quad \psi(1) = 0.
\]
Then, (4.14) has a smooth unique solution \(\psi\) and \(J = |\nabla \psi|\) satisfies
\[
(4.15) \quad J_t + \nabla (\eta_\epsilon \ast \phi) \cdot \nabla J + D^2(\eta_\epsilon \ast \phi) J = \nabla f, \quad J(1) = 0.
\]
Since \(\psi\) has compact support, \(J\) has a positive maximum over \([0,1] \times R^d\) at some point \((t_0, x_0)\). If \(0 \leq t_0 < 1\), then from (4.15)
\[
J(t_0, x_0) \leq 0 \text{ and } \nabla J(t_0, x_0) = 0.
\]
Thus,
\[
D^2(\eta_\epsilon \ast \phi) J(t_0, x_0) \geq \nabla f(t_0, x_0)
\]
Since from (4.2) \(D^2(\eta_\epsilon \ast \phi) \leq -C\), this implies
\[
(4.16) \quad |\nabla \psi|_\infty = J(t_0, x_0) \leq \frac{|\nabla f|_\infty}{C}
\]
Let \(\lambda, \tilde{\lambda}\) is two weak solutions to (4.13). Then, it follows from (4.12) and (4.14) that
\[
\int_0^1 \int_{R^d} (\lambda - \tilde{\lambda}) f \, dx \, dt = \int_0^1 \int_{R^d} (\lambda - \tilde{\lambda})(\nabla (\eta_\epsilon \ast \phi) - \nabla \phi) \nabla \psi \, dx \, dt.
\]
By letting $\epsilon \to 0^+$, it follows from (4.5)–(4.6), (4.16) and the Lebesgue dominated convergence theorem that
\[
\int_0^1 \int_{R^d} (\lambda - \tilde{\lambda}) f \, dx \, dt = 0
\]
for all $f \in C_0^\infty([0,1] \times R^d)$ and therefore $\lambda = \tilde{\lambda}$.

Now, let $\lambda^\tau$ be the solution to (4.13) associated with $\phi^\tau$. Since $\lambda^\tau$ 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 $\lambda^\tau$ converges weakly to $\lambda^*$ 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
\[
\int_0^1 \int_{R^d} (\lambda^\tau \psi_t + (\lambda^\tau \nabla \phi + \lambda^\tau (\nabla \phi^\tau - \nabla \phi) \cdot \nabla \psi)) \, dx \, dt = \int_{R^d} (\rho_0 \psi(0) - \lambda^\tau(1) \psi(1)) \, dx.
\]
Since $\lambda^\tau$ 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 $\lambda^*$ is the weak solution to (4.13). Since (4.13) has the unique weak solution, we conclude $\lambda^\tau$ converges weakly to $\lambda$ 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
\[
\int_0^1 \int_{R^d} \lambda^\tau \psi_t + \nabla \phi^\tau \cdot \nabla \psi) \, dx \, dt = \int_{R^d} (\rho_0 \psi(0) - \psi(1) \lambda^\tau(1)) \, dx.
\]
for all $\psi \in W^{1,\infty}((0,1) \times R^d)$. Since
\[
(\phi^\tau - \phi)_t + \nabla \phi^\tau \cdot \nabla (\phi^\tau - \phi) - \frac{1}{2} \nabla \phi^\tau - \nabla \phi \right)^2 = 0
\]
by setting $\psi = \phi^\tau - \phi$ in (4.18), we obtain
\[
\int_{R^d} ((\phi^\tau(0) - \phi(0)) \rho_0(x) - \tau \lambda^\tau(1) h(x)) \, dx = \frac{1}{2} \int_0^1 \int_{R^d} \lambda^\tau \left| \nabla \phi^\tau - \nabla \phi \right|^2 \, dx \, dt.
\]
Similarly, since
\[
(\phi^\tau - \phi)_t + \nabla \phi \cdot \nabla (\phi^\tau - \phi) + \frac{1}{2} \nabla \phi^\tau - \nabla \phi \right)^2 = 0,
\]
we have
\[
\int_{R^d} ((\phi^\tau(0) - \phi(0)) \rho_0(x) - \tau \lambda(1) h(x)) \, dx = -\frac{1}{2} \int_0^1 \int_{R^d} \lambda \left| \nabla \phi^\tau - \nabla \phi \right|^2 \, dx \, dt.
\]
From (4.3) and (4.8) there exists a subsequence of $\frac{\phi^\tau(0) - \phi}{\tau}$ 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
\[
\lim_{\tau \to 0} \int_{R^d} \frac{\phi^\tau(0) - \phi}{\tau} \rho_0(x) \, dx = \int_{R^d} \lambda(1) h(x) \, dx.
\]
Hence
\[
J'(v)(h) = \int_{R^d} (\rho_1 - \lambda(1)) h(x) \, dx.
\]
Step 5 Assume $\bar{v}$ attains the minimum of $J(v)$ in (2.8) and $\bar{v}$ is Lipschitz and semi-convex. Then $J'(\bar{v})(h) = 0$ for all $h \in C_0^2(R^d)$ and thus from (4.19) $\lambda(1) = \rho_1$ a.e., where $\lambda$ is the weak solution to (4.13) with $\phi = \bar{v}$. Thus, (3.3) holds with $\rho = \lambda$. Since
\[
\phi_t + \nabla \phi \cdot \nabla \phi - \frac{1}{2} |\nabla \phi|^2 = 0,
\]
et 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 $c(x - y) = \frac{1}{p}|x - y|^p$, $1 < p < \infty$. Assume $v$ is Lipschitz and semi-convex. It follows from \[5\] that
\[
(5.1) \quad \phi_t + \frac{1}{q}|\nabla \phi|^q = 0, \quad \phi(1, x) = v(x)
\]
has a unique viscosity solution $\phi \in W^{1,\infty}((0, 1) \subset R^d)$ 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 $y \rightarrow v^\tau(y) - c(x - y)$, then
\[
x - y^\tau = -|\nabla v^\tau(y^\tau)|^{q-2}\nabla v^\tau(y^\tau)
\]
Since $v^\tau \in W^{1,\infty}(R^d)$, $|y^\tau(x) - x| \leq \alpha$ for some $\alpha$ uniformly in $x$ and $\tau$. Since as shown in Section 4
\[
\tau h(y) \leq \pi^\tau(0, x) - \phi(0, x) \leq -\tau h(y^\tau),
\]
\[
\int_{R^d} |\phi^\tau(0, x) - \phi(0, x)|\,dx = \tau \int_{R^d} (|h(y(x))| + |h(y^\tau(x))|)\,dx.
\]
Since $h$ is compactly supported, it follows that there exists a constant $M$ (depends on $h$) such that
\[
(5.2) \quad \int_{R^d} |\phi^\tau(0, x) - \phi(0, x)|\,dx \leq M \tau.
\]
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
\[
\int_0^1 \int_{R^d} (|\nabla \phi^\tau|^q - |\nabla \phi|^q)\,dx \to 0
\]
11
as $\tau \to 0$. Hence

$$
\int_0^1 \int_{R^d} |\nabla \phi^\tau|^2 \, dx \, dt \to \int_0^1 \int_{R^d} |\nabla \phi|^2 \, dx \, dt
$$
as $\tau \to 0$ and therefore

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

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

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

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

$$
J_\epsilon = |\nabla (\eta_\epsilon \ast \phi)|^q \nabla (\eta_\epsilon \ast \phi)
$$

Then,

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

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

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

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

$$
\lambda_t + (|\nabla \phi^\tau|^q \nabla \phi^\tau, \lambda)_x = 0
$$
has the unique weak solution $\lambda^\tau \in L^\infty(0, 1; L^1(R^d) \cap L^\infty(R^d))$, i.e.,

$$
(5.5) \quad \int_0^1 \int_{R^d} \lambda^\tau(\psi_t + |\nabla \phi^\tau|^q \nabla \phi^\tau \cdot \nabla \psi) \, dx \, dt = \int_{R^d} (\psi(0)\lambda(0) - \psi(1)\lambda(1)) \, dx.
$$
for all $\psi \in W^{1,\infty}((0, 1) \times R^d)$. Moreover $\lambda^\tau$ converges weakly to $\lambda$ 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

$$
(\phi^\tau - \phi)_t + |\nabla \phi^\tau|^q \nabla \phi^\tau \cdot \nabla (\phi^\tau - \phi) + I_1 = 0
$$
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
\[ \int_{\mathbb{R}^d} ((\phi^\tau(0) - \phi(0)) \rho_0(x) - \tau \lambda^\tau(1) h(x)) \, dx = - \int_{\mathbb{R}^d} \lambda^\tau I_1 \, dx \]

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

Since \( \lambda, \lambda^\tau \geq 0 \) a.e. in \((0,1) \times \mathbb{R}^d\) and \( \lambda^\tau(1) \) converges weakly to \( \lambda(1) \) as \( \tau \to 0 \), it follows that
\[ \lim_{\tau \to 0} \int_{\mathbb{R}^d} \frac{\phi^\tau(0) - \phi(0)}{\tau} \rho_0(x) \, dx = \int_{\mathbb{R}^d} \lambda(1) h(x) \, dx. \]

Thus
\[ (5.6) \quad J'(v)(h) = \int_{\mathbb{R}^d} (\rho_1 - \lambda(1)) h(x) \, dx. \]

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 \( h \in C^2_0(\mathbb{R}^d) \) and therefore from (5.6) \( \bar{\lambda}(1) = \rho_1 \) where \( \bar{\lambda} \) is the weak solution to
\[ (5.7) \quad \lambda_t + (\lambda |\nabla \phi|^{q-2} \nabla \phi)_x = 0, \quad \lambda(0) = \rho_0 \]

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

which shows (3.4).
References

[1] Y. Brenier, Decomposition 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-Kantorovich mass transfer problem, Numer. Math., 84 (2000), 375-393.

[4] L. Caffarelli, Allocation maps with general cost functions, in Partial 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 Differential 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).