An Introduction to Monge Ampère Equation

Nestor Guillen, notes taken by Mark Ma

March 13, 2025

This is set of mini lectures prof. Nestor Guillen gave the Columbia University in Spring 2025 for the Optimal Transport Reading Group. Thanks Nestor!

Contents

1 Introduction

Theorem 1.1 (Brenier)**.** *Given two probability measures μ, ν in $\mathbb{R}^n$ with finite second moments and absolutely continuouous w.r.t. Lebesgue measure*

$$\mu,\, \nu \ll dx$$

By Brenier, there exists $T(x)$ defined μ-a.e. s.t.

$$\int_{\mathbb{R}^n} \eta(T(x)) d\mu(x) = \int_{\mathbb{R}^n} \eta(y)\, d\nu(y) \qquad \forall\ \eta \in C_c^0(\mathbb{R}^n)$$

and there is a convex function $u(x)$ s.t.

$$T(x) = \nabla u(x) \qquad \forall\ \mu - a.e.$$

Remark 1.1. *Suppose we knew that*

$$\mu = f dx \qquad \nu = g dy$$

and suppose that $T(x) \in C^1$. If

$$D := \{f > 0\}$$
$$D^* := \{g > 0\}$$

and D, D^* bounded, with f, g continuous in D and D^*. Then one may apply the change of variables formula so that

$$\int_{\mathbb{R}^n} \eta(T(x)) f(x)\, dx = \int_{\mathbb{R}^n} \eta(T(x)) g(T(x)) \det(DT(x))\, dx \qquad \forall\ \eta \in C_c^0(\mathbb{R}^n)$$

This shows from continuity that

$$g(T(x)) \det(DT(x)) = f(x) \qquad \forall\ x \in D$$

Notice when $T(x) \in D^*$ for $x \in D$. Using

$$T = \nabla u$$

one has

$$\det(D^2 u(x)) = \frac{f(x)}{g(\nabla u(x))} \qquad \forall\ x \in D$$
$$\nabla u(D) = D^* \qquad \textit{Geometric Neumann boundary condition}$$

Understanding regularity of optimal transport maps gets to understanding regularity of Monge Ampere equations. In the case for ν with density 1, $g(\nabla(u))$ goes away. Now taking

$$g \equiv 1 \qquad \textit{in } D^*$$

one obtain

$$\det(D^2 u(x)) = \frac{f(x)}{g(\nabla u(x))} \qquad \forall\ x \in D$$
$$\nabla u(D) = D^* \qquad u \textit{ convex in } D$$

One has 2nd Boundary Value Problem (1st order Dirichlet BVP) with

$$u|_{\partial D} \qquad \textit{prescribed}$$

Comparing with

$$\Delta u = f$$
$$u = h \qquad \partial D$$

one has regularity estimates if

$$f \in L^p \implies u \in W^{2,p}$$

If moreover u is convex and $\lambda_k(u) \geq 0$, with $\sum_{i=1}^n \lambda_i \leq C$ then $|\lambda_i| \leq C$. But meanwhile, even in $n = 2$, for

$$\lambda_1 \cdot \lambda_2 \leq C$$

even if $\lambda_1 \geq 0$ and $\lambda_2 \geq 0$, one has no guarantee that

$$\lambda_1 \leq C \qquad \text{or} \qquad \lambda_2 \leq C$$

Remark 1.2 (Pogorelov). *There exists u convex solving in a weak sense*

$$\det(D^2u) = 1 \qquad \text{in } \mathbb{R}^3$$

and $u \notin C^2$. Between C^0 and C^2, all the fight happens.

2 Alexandrov's Perspective: Subdifferential and Convex Function

Now the goal is, for a convex

$$u : D \to \mathbb{R}$$

we want to construct a Monge-Ampère measure $\mathcal{M}_u$, which is a Borel measure s.t.

$$\mathcal{M}_u(E) = \int_E (\det(D^2 u))dx \qquad \forall\ u \in C^2$$

Definition 2.1 (Subdifferential, Convexity)**.** *From now on $D \subset \mathbb{R}^n$ open bounded, and*

$$u : \overline{D} \to \mathbb{R}$$

is continuous. Given $x \in D$, define the subdifferential of u at x as the set-valued map

$$\partial u(x) := \{p \in \mathbb{R}^n \mid u(y) \geq u(x) + p \cdot (y - x) \quad \forall\ y \in D\}$$

This is defined w.r.t. D. But of course this could be an empty set. The function

$$u : D \to \mathbb{R}$$

is convex in D if

$$\partial u(x) \neq \varnothing \qquad \forall\ x \in D$$

When the set D is not convex, the definition could be different from the classical definition of convexity. Our definition says any point in D have at least one supporting hyperplane.

Example 2.1. *For $u(x) = |x|$ with $n \geq 1$*

$$\partial u(0) = \overline{B_1(0)}$$

Example 2.2. *If u is convex, and differentiable at $x_0 \in D$, then the subdifferential of u at x_0*

$$\partial u(x_0) = \{\nabla u(x_0)\}$$

Remark 2.1.

$$(u(x) - u(x_0) - \nabla u(x_0) \cdot (x - x_0))$$

Proposition 2.1. *Now for u convex in D*

1. *For any x, $y \in D$ s.t.*
$$(1 - t)x + ty \in D \qquad \forall\ t \in [0, 1]$$
then
$$u((1 - t)x + ty) \leq (1 - t)u(x) + tu(y)$$

2. *If u is C^1 in D, then the monotone map*
$$(\nabla u(x) - \nabla u(y), x - y) \geq 0$$
for any x, y as in (i).

3. *If $u \in C^2$ near x, then*
$$D^2 u(x) \geq 0$$

Proof. For any fixed $z \in D$ s.t. $z \in [x, y]$, then

$$\begin{aligned}
u(x) &\geq u(z) + \nabla u(z) \cdot (x - z) \\
u(y) &\geq u(z) + \nabla u(z) \cdot (y - z) \\
\varepsilon u(x) &\geq \varepsilon u(z) + \varepsilon \nabla u(z) \cdot (x - z) \\
(1 - \varepsilon)u(y) &\geq (1 - \varepsilon)u(z) + (1 - \varepsilon)\nabla u(z) \cdot (y - z) \\
\varepsilon u(x) + (1 - \varepsilon)u(y) &\geq u(z) + \nabla u(z) \cdot (\varepsilon x + (1 - \varepsilon)y - z) \qquad \text{adding up}
\end{aligned}$$

now choosing $z = \varepsilon x + (1 - \varepsilon)y$ so the latter cancels we obtain (i). From (i), given x and $h \in \mathbb{R}^n$ small,

$$u(x + he) + u(x - he) \geq 2u(x)$$

Move to the Left we have the second derivative choose the direction $e \in \mathbb{S}^{n-1}$. $\square$

Remark 2.2. *Now any convex function away from the domain D is Lipschitz. In dimension 1 with $I = [a, b]$*

$$\partial u(x) \subset [-\frac{\operatorname{osc} u}{\varepsilon}, \frac{\operatorname{osc} u}{\varepsilon}]$$

where $\varepsilon = d(x, \partial I) = \max\{|x-a|, |x-b|\}$. Equivalently, if

$$K(y) := u(x) + \frac{\operatorname{osc} u}{\varepsilon}|y - x|$$

then

$$u(y) \leq K(y)$$

We need to use the fact that if $u \leq v$ in D at x_0, then

$$\partial u(x_0) \subset \partial v(x_0)$$

Lemma 2.1. *If u is convex in D, and there exists a ball $B_{R+\varepsilon}(x_0) \subset D$. Then we get*

$$\sup_{x,\, y \in B_R(x_0),\ x \neq y} \frac{|u(x) - u(y)|}{|x - y|} \leq \frac{\operatorname{osc}_D u}{\varepsilon}$$

Proof. Define cone

$$K(y) := u(x) + \frac{\operatorname{osc}_D u}{\varepsilon}|y - x| \qquad \forall\ y \in B_R(x_0)$$

By the remark in $n = 1$, we have

$$u(y) \leq K(y) \qquad \forall\ y \in B_R$$

Now we have

$$u(y) \leq u(x) + \frac{\operatorname{osc} u}{\varepsilon}|y - x|$$

by Rearranging

$$\frac{|u(y) - u(x)|}{|x - y|} \leq \frac{\operatorname{osc} u}{\varepsilon}$$

□

Theorem 2.1 (Rademacher's)**.**

$$\nabla u(x) \qquad \textit{exists a.e. } x \in D$$

and

$$|\nabla u(x)| \leq \frac{\operatorname{osc}_D u}{d(x, \partial D)} \qquad a.e. x \in D$$

Corollary 2.1. *For any $x \in D$*

$$\partial u(x) \subset \overline{B_{\frac{\operatorname{osc}_D u}{d(x,\partial D)}}(0)}$$

Now we have

$$\partial u(x) = \begin{cases} \{\nabla u(x)\} & a.e. x \\ \text{remains to discover} & \end{cases}$$

3 Alexandrov Solution and Monge-Ampère measure

Next we proceed to define

$$\partial u(E) := \bigcup_{x \in E} \partial u(x)$$

then we define the Monge-Ampere measure as

$$\mathcal{M}_u(E) := |\partial u(E)|$$

Now '$\mathcal{M}_u = \mu$' is to say $\det(D^2 u) = \mu$.

Definition 3.1 (Monge-Ampère measure)**.** *Given a Borel measure ν in $\mathbb{R}^n$, and $u : D \to \mathbb{R}$ convex, finite at every point (hence continuous), with $D \subset \mathbb{R}^n$ open and bounded, we define the Monge-Ampère measure associated to ν as*

$$\mathcal{M}_u^\nu(E) := \nu(\partial u(E)) \qquad \forall\ E \in \mathcal{B}$$

where

$$\partial u(E) := \bigcup_{x \in E} \partial u(x)$$

In particular if ν is the Lebesgue measure, we denote

$$\mathcal{M}_u(E)$$

as the Monge-Ampère measure of u.

Definition 3.2 (Alexandrov Solution)**.** *Given D and*

$$h : \partial D \to \mathbb{R}$$

continuous, and

$$f : D \to \mathbb{R}$$

where $f \in L^1(D)$. We say

$$u : \overline{D} \to \mathbb{R}$$

is an Alexandrov Solution of

$$\det(D^2 u(x)) = f(x)$$

if u is convex, continuous in $\overline{D}$ and

$$\mathcal{M}_u(E) = \int_E f(x)\, dx \qquad \forall\ E \in \mathcal{B} \qquad E \Subset D$$

In particular, one studies the Dirichlet problem: find u convex solving

$$\begin{aligned} \det(D^2 u) = f \qquad & D \\ u = h \qquad & \partial D \end{aligned}$$

Remark 3.1. *Suppose u is an Alexandrov Solution of the Dirichlet problem, i.e.*

$$\mathcal{M}_u(E) = \int_E f(x)\, dx$$

If u is $C^2(D)$ and

$$\det(D^2 u) > 0$$

Then u solves

$$\det(D^2 u(x)) = f(x) \qquad \textit{Lebesgue a.e.-}x \in D$$

Proof. Change of Variables. □

Remark 3.2. *Conversely, if*

$$u : D \to \mathbb{R}$$

is convex, $C^2(D)$ and for $f \in L^1(D)$

$$\det(D^2 u(x)) = f(x) \qquad \textit{Lebesgue a.e.-}x \in D$$

Then

$$\mathcal{M}_u(E) = \int_E f(x)\, dx$$

Theorem 3.1 (Existence of Alexandrov Solution)**.** *If D is a strictly convex domain, bounded, and*

$$h : \partial D \to \mathbb{R}$$

is continuous, and μ is a probability measure in D. Then there exists a unique Alexandrov Solution $\mathcal{M}_u = \mu$ to

$$\begin{aligned} \det(D^2 u) &= \mu \qquad D \\ u &= h \qquad \partial D \end{aligned}$$

Theorem 3.2 (Convergence)**.** *Consider an infinite sequence*

$$u_n : D \to \mathbb{R}$$

of finite convex functions converging locally uniformly to

$$u : D \to \mathbb{R}$$

Then the measures

$$\mathcal{M}_{u_n} \rightharpoonup \mathcal{M}_u$$

i.e.

$$\lim_{n\to\infty} \int_D \phi(x) \mathcal{M}_{u_n}(dx) = \int_D \phi(x) \mathcal{M}_u(dx) \qquad \forall\ \phi \in C_c^0(D)$$

Remark 3.3 (On the limitations of the Alexandrov Solution)**.** *Not every solution to optimal transport is an alexandrov solution. Consider in $\mathbb{R}^2$, the measure*

$$\mu := \mathbb{1}_{B_1}\, dx$$

with

$$\nu := \mathbb{1}_{D^*}\, dx$$

where

$$D^* = (B_1 \cap (\{x_1 < 0\} - (1,0))) \cup (B_1 \cap (\{x_1 > 0\} + (1,0)))$$

Brenier's Theorem produces an Optimal Transport map which μ-a.e. satisfies

$$T(x) = \nabla u(x)$$

for a convex function u. And the Monge-Ampère Equation holds in the sense that

$$\int \eta(T(x)) \mu(dx) = \int \eta(y) \nu(dy)$$

Notice for

$$u(x) = \frac{1}{2}|x|^2 \qquad \forall\ x \in \mathbb{R}^2$$

one has $\nabla u(x) = x$. In this case u is given by

$$u(x) = \frac{1}{2}(x_1^2 + x_2^2) + |x_1|$$

now μ-a.e.

$$\nabla u(x) = \begin{cases} x - e_1 \\ x + e_1 \end{cases}$$

Claim : u is not an Alexandrov Solution of

$$\det(D^2u) = 1 \qquad D$$

Why? Since

$$\partial u(x) = \begin{cases} x - e_1 & \{x_1 < 0\} \\ x + [-1,1]e_1 & x_1 = 0 \\ x + e_1 & \{x_1 > 0\} \end{cases}$$

As a result, in B_1, given set

$$E = \{(0, x_2) \mid |x_2| \leq 1\}$$

then

$$|\partial u(E)| > 0$$

The whole region

$$\partial u(E) = [-1,1]^2$$

is filling everything inside. The support of ν isnnot a convex set. Alexandrov is too strong a solution that depends on the PDE. It is too sensitive. This was later fixed by Caffarelli.

3.1 Tools from Subdifferential

One has properties of $\partial u(x)$.

Lemma 3.1 (Lower Semi-continuity)**.** *If $x_k \to x_* \in D$, then*

$$\limsup_{k\to\infty} \partial u(x_k) \subset \partial u(x_*)$$

Proof. Suppose $p_k \in \mathbb{R}^n$ is sequence of points such that there exists $n_k \to \infty$ s.t.

$$p_k \in \partial u(x_{n_k})$$

i.e.,

$$u(x) \geq u(x_{n_k}) + p_k \cdot (x - x_{n_k}) \qquad \forall\ x \in D$$

Then we claim the sequence $\{p_k\}$ is bounded. This is implied by that

$$\partial u(x) \subset \overline{B_{\frac{\operatorname{osc}_D u}{d(x,\partial D)}}(0)}$$

Hence WLOG $p_k \to p_\infty$. Then for any $x \in D$

$$u(x) \geq \lim_{k\to\infty} \{u(x_{n_k}) + p_k \cdot (x - x_{n_k})\}$$
$$u(x) \geq u(x_*) + p_\infty \cdot (x - x_*) \qquad \forall\ x$$

Hence

$$p_\infty \in \partial u(x_*)$$

□

Corollary 3.1. *If*

$$u : D \to \mathbb{R}$$

is convex and $x_0 \in D$ is such that

$$\partial u(x_0) = \{p_0\}$$

Then u is C^1 at x_0 and

$$\nabla u(x_0) = p_0$$

Proof. Let's restate the corollary. For $e \in \mathbb{R}^n$, we want to show

$$\frac{u(x+he)-u(x)}{h} \overset{h\to 0}{\to} e\cdot p_0$$

this limit exists. On one hand, by assumption

$$u(x_0+he) \geq u(x_0) + hp_0 \cdot e$$
$$\frac{u(x_0+he)-u(x_0)}{h} \geq e\cdot p_0 \qquad \forall\ e \in \mathbb{R}^n \qquad \forall\ |h| \text{ small}$$

To check the limit exists, it suffices to check for any discrete sequence $h_n \to 0$. Given a sequence $\{h_n\}_n$ s.t. $\lim\limits_{n\to\infty} h_n = 0$. Choose for each n an element of the subdifferential

$$p_n \in \partial u(x_0 + h_n e)$$

Then

$$u(x_0) \geq u(x_0+h_n e) + (x_0 - (x_0+h_n e))\cdot p_n$$

Bounding from above can also be achieved by touching from below!

$$u(x_0) \geq u(x_0+h_n e) - h_n e\cdot p_n$$
$$e\cdot p_n \geq \frac{u(x_0+h_n e)-u(x_0)}{h_n}$$
$$e\cdot p_n \geq \frac{u(x_0+h_n e)-u(x_0)}{h_n} \geq e\cdot p_0$$

But since

$$\partial u(x_0) = \{p_0\}$$

the only possible limiting point is p_0. WLOG the sequence $\{p_n\}$ is Cauchy up to a subsequence (using subdifferential is bounded away from the boundary), then

$$p_n \to p_0$$

This shows that

$$\limsup_{n\to\infty} e\cdot p_n = e\cdot p_0$$

By lower-semicontinuity Lemma 3.1, the two-sided inequality yields

$$\lim_{h_n\to 0} \frac{u(x_0+h_n e)-u(x_0)}{h_n} = e\cdot p_0$$

□

3.2 $\mathcal{M}_u$ is Borel measure

Next, we aim to show

Theorem 3.3. *$\mathcal{M}_u$ is indeed a Borel measure. Moreover it is finite on compact sets.*

Let's begin with some lemma.

Lemma 3.2 (Subdifferential preserves Compactness)**.** *If $E \subset D$ is compact, then $\partial u(E)$ is also compact.*

Proof. Since E is compact, $\operatorname{dist}(E, \partial D)$ is positive. Thus there exists $M > 0$ s.t.

$$\partial u(E) \subset B_M(0)$$

as in Corollary 2.1. Consider

$$\{p_n\} \subset \partial u(E)$$

Then $\{p_n\} \subset B_M(0)$. Then by compactness of $\overline{B_M(0)}$ there exits a subsequence $\{p_n'\}$ and p_∞ s.t.

$$p_n' \to p_\infty$$

For each n pick $x_n \in E$ s.t.

$$p'_n \in \partial u(x_n)$$

Now $\{x_n\} \subset E$ and again by compactness, there is a subsequence $\{p''_n\}$ and $\{x'_n\}$ and x_∞ s.t.

$$\begin{aligned} x_\infty &\in E \\ p''_n &\in \partial u(x'_n) \\ x'_n &\to x_\infty \end{aligned}$$

But then by Lower Semi-continuity 3.1 we get

$$p_\infty \in \partial u(x_\infty)$$

More precisely, for every n and for every $x \in D$ we have

$$u(x) \geq u(x'_n) + p''_n \cdot (x - x'_n)$$

Since

$$\begin{aligned} \lim_{n\to\infty} u(x'_n) &= u(x_\infty) \\ \lim_{n\to\infty} p''_n &= p_\infty \\ \lim_{n\to\infty} x'_n &= x_\infty \end{aligned}$$

Thus

$$\begin{aligned} &u(x) \geq u(x_\infty) + p_\infty \cdot (x - x_\infty) \\ &p_\infty \in \partial u(x_\infty) \end{aligned}$$

Thus $p_\infty \in E$ and so $\partial u(E)$ is compact. □

In the following we need to understand how big is the set s.t.

$$\{p \mid \text{there exists } x_1,\, x_2 \in D \text{ s.t. } x_1 \neq x_2 \text{ and } p \in \partial u(x_1) \cap \partial u(x_2)\} \tag{1}$$

It is surprising that this is Lebesgue measure 0! Hence it is measurable. We want to show this set corresponds to non-differentiability of the function and conclude by Rademacher 2.1.

Definition 3.3 (Legendre-Transform)**.** *Given $u : D \to \mathbb{R}$ $u \in C(\overline{D})$ convex, we define*

$$u^*(y) := \sup_{x\in D}\{x \cdot y - u(x)\}$$

Note u^* may not be finite for certain values of y.

Remark 3.4. *If $y \in \partial u(D)$, then $u^*(y)$ is finite and*

$$u^*(y) := \sup_{x\in D}\{x \cdot y - u(x)\} = \max_{x\in D}\{x \cdot y - u(x)\}$$

Remark 3.5 (Duality)**.** *$x_0 \in \partial u^*(y_0)$ implies $y_0 \in \partial u(x_0)$. Now by Rademacher's 2.1, $\partial u^*(y)$ is a singleton for all y outside a set of Lebesgue measure zero.*

Corollary 3.2 (The set has measure zero)**.** *The set (1) has Lebesgue measure zero.*

Now let's go!

Lemma 3.3. *The set*

$$S := \{E \subset D \mid \partial u(E) \text{ is Lebesgue measurable}\}$$

is a σ-algebra containing $\mathcal{B}(D)$.

Proof. We already know

$$E \subset S$$

if E is compact. Now we need to show S is a σ-algebra. Consider an infinite sequence

$$\{E_n\} \subset S$$

Then want to see

$$\partial u(\bigcup_n E_n) = \bigcup_n \partial u(E_n) \qquad \text{is measurable}$$

But the set on the RHS is the countable union of Lebesgue-measurable sets, thus it is measurable and thus its union $\bigcup_n E_n$ is also in S. Our previous argument implies that $D \in S$. What's left is if $E \in S$, we want to show $D \setminus E \in S$. But this is

$$\partial u(D \setminus E) = (\partial u(D) \setminus \partial u(E)) \cup (\partial u(D \setminus E) \cap \partial u(E))$$

But the right-most set $\partial u(D \setminus E) \cap \partial u(E) \subset (1)$ hence it has Lebesgue measure zero. Now we check the part $\partial u(D) \setminus \partial u(E)$ is measurable. But due to E, $D \in S$, indeed it is measurable, hence $D \setminus E \in S$. ☐

Proof of Theorem 3.3. We show $\mathcal{M}_u$ is a measure in S and finite on compact sets. We show the σ-additivity. Consider countable family of sets $\{E_n\} \subset S$ pairwise disjoint. Set

$$F_m := \partial u(E_m)$$

All we need to show is that

$$|\bigcup_m F_m| = \sum_m |F_m|$$

We only need to know F_m is pairwise disjoint. Notice that by Corollary 3.2

$$|F_m \cap F_n| = 0$$

if $m \neq n$. So let's write

$$\bigcup_m F_m = F_1 \cup (F_2 \setminus F_1) \cup (F_3 \setminus (F_1 \cup F_2)) \cup \cdots (F_m \setminus (F_1 \cup \cdots \cup F_{m-1}))$$

as a disjoint union. So

$$|\bigcup_m F_m| = \sum_m |F_m \setminus (F_1 \cup \cdots \cup F_{m-1})|$$

But notice

$$F_m = (F_m \setminus (F_1 \cup \cdots \cup F_{m-1})) \cup (F_m \cap (F_1 \cup \cdots \cup F_{m-1}))$$

where the latter set has measure zero by Corollary 3.2. So

$$|F_m| = |F_m \setminus (F_1 \cup \cdots \cup F_{m-1ma})|$$

☐

4 The Brenier Perspective

We do the delicate part of the theory first. We start with the Dual of the Kantorovich Problem.

Definition 4.1 (Dual of the Kantorovich Problem)**.** *The problem is: Given data*

1. *two μ, ν measures in $\mathbb{R}^n$ compactly supported,*
2. *and $D \Subset \operatorname{supp}(\mu)$*
3. *$D^* \Subset \operatorname{supp}(\nu)$*

We want to maximize over pairs of function (ϕ, ψ) the quantity

$$\int \phi(x)\mu(dx) + \int \psi(y)\nu(dy)$$

subject to the condition

$$\phi \in L^1(\mu) \qquad \psi \in L^1(\nu) \qquad \phi(x) + \psi(y) \leq \frac{1}{2}|x-y|^2$$

The Maximum to this problem is equal to the minimum to the original problem.

Theorem 4.1. *There is a solution (ϕ_0, ψ_0) to Dual of Kantorovich Problem 4.1, which are pairs of convex function and Jegendre duals of each other, i.e.,*

$$\phi_0(x) = \inf_{y \in D^*}\{\frac{1}{2}|x-y|^2 - \psi_0(y)\}$$
$$\psi_0(y) = \inf_{x \in D}\{\frac{1}{2}|x-y|^2 - \phi_0(x)\}$$

and (recall $\mu \ll dx$) for any $\eta \in C_c^\infty$, we have the 'Brenier Condition'

$$\int \eta(y)\,\nu(dy) = \int \eta(y(x))\mu(dx) \qquad \textit{where } y(x) := \nabla_x(\frac{1}{2}|x|^2 - \phi_0(x)) \textit{ is convex} \tag{2}$$

Remark 4.1. (2) *is the weak formulation of the Euler-Lagrange Equation of the Dual problem.*

Proof of 'Brenier Condition' (2). Take ϕ_0, ψ_0 optimal. Take any $\eta \in C_c^\infty$ and we define

$$\psi_s(y) := \psi_0(y) + s\eta(y)$$

We pick

$$\phi_s(x) := \inf_y\{\frac{1}{2}|x-y|^2 - \psi_s(y)\}$$

One may check ϕ_s and ψ_s is an admissible pair for our problem. Indeed

$$\phi_s(x) := \inf_y\{\frac{1}{2}|x-y|^2 - \psi_s(y)\} \leq \frac{1}{2}|x-y|^2 - \psi_s(y) \qquad \forall\ y$$

One should also check that

$$s \mapsto \Theta(s) := \int \phi_s(x)\mu(dx) + \int \psi_s(y)\nu(dy)$$

has a global maximum at $s = 0$. The next thing we do is take a derivative.

Lemma 4.1. $\Theta(s)$ *is differentiable at $s = 0$ and*

$$\Theta'(0) = \int -\eta(y(x))\mu(dx) + \int \eta(y)\nu(dy) \tag{3}$$

Proof. Look at

$$\frac{\Theta(s)-\Theta(0)}{s} = \int \frac{\phi_s(x)-\phi_0(x)}{s}\mu(dx) + \int \eta(y)\,\nu(dy)$$

Now the map $y(x)$ is going to be important. The goal is to find

$$\lim_{s\to 0} \frac{\phi_s(x)-\phi_0(x)}{s} \qquad \mu\text{-a.e. } x$$

Notice

$$\begin{aligned}
\phi_s(x) &= \inf_y\{\frac{1}{2}|x-y|^2 - \psi(y) - s\eta(y)\} \\
&\leq \inf_y\{\frac{1}{2}|x-y|^2 - \psi(y)\} + s\,\|\eta\|_\infty \\
\phi_s(x) &\geq \inf_y\{\frac{1}{2}|x-y|^2 - \psi(y)\} - s\,\|\eta\|_\infty \\
|\frac{\phi_s(x)-\phi_0(x)}{s}| &\leq \|\eta\|_\infty
\end{aligned}$$

Now one may apply DCT. If ϕ_0 is differentiable at a point x_0, then this means that the function

$$y \mapsto \frac{1}{2}|x_0-y|^2 - \psi(y)$$

can only have one minimum in $\overline{D^*}$. Indeed if y_0 achieves the minimum, then $\nabla\phi(x_0) = x_0 - y$ there is only one choice. This defines for a.e. x the function (due to unique correspondence of y for each x)

$$y(x) = x - \nabla\phi(x)$$

i.e., in the domain of $y(x)$, we have

$$y(x) = \text{argmin}\{\frac{1}{2}|x-y|^2 - \psi_0(y)\}$$

Then we take x where ϕ_0 is differentiable. We consider any sequence $s_n \to 0$ as $n\to\infty$. Then for each n we choose y_n s.t.

$$y_n \in \underset{y}{\text{argmin}}\{\frac{1}{2}|x-y|^2 - \psi_{s_n}(y)\}$$

Then we write

$$\phi_{s_n}(x) - \phi_0(x) = \left(\frac{1}{2}|x-y_n|^2 - \psi_0(y_n)\right) - s_n\eta(y_n) - \left(\frac{1}{2}|x-y_0|^2 - \psi_0(y_0)\right)$$

We make an observation: Since

$$y \mapsto |x-y|^2 - \psi_0(y)$$

has a unique minimum at y_0, and it is differentiable since

$$x \in \text{Domain}(\nabla\phi)$$

One can take

$$\left(\frac{1}{2}|x-y_n|^2 - \psi(y_n)\right) - \left(\frac{1}{2}|x-y_0|^2 - \psi(y_0)\right) = o(|y_n - y_0|)$$

Later we'll see $o(|y_n-y_0|) \leq o(s_n)$. Then

$$\frac{1}{s_n}\left(\phi_{s_n}(x) - \phi_0(x)\right) = o(|y_n-y_0|) - \eta(y_n)$$

An exercise is to check that

$$y_n \to \underset{y}{\text{argmin}}\{\frac{1}{2}|x-y|^2 - \psi(y)\} = \{y(x_0)\}$$

Then we have proved that for a.e. $x \in D$

$$\lim_{s \to 0} \frac{\phi_s(x) - \phi_0(x)}{s} = -\eta(y(x))$$

And finally by Dominated Convergence we have (3)

$$\Theta'(0) = \int -\eta(y(x))\mu(dx) + \int \eta(y)\nu(dy)$$

□

And since $\Theta'(0) = 0$ we get that

$$\int \eta(x - \nabla\phi_0(x))\mu(dx) = \int \eta(y)\nu(dy) \qquad \forall\ \eta \in C_c^\infty$$

□

Some suggested readings:

1. Proof of Rockefeller's Theorem
2. Displacement convexity/interpolation

$$\mu_0,\ \mu_1 \qquad y(x) = \text{Brenier Map}$$

   and

$$\mu_t := ((1-t)x + ty(x))_{\#}\mu_0$$

3. The Wasserstein metric.
4. The JKO Scheme.