Lectures on Ma-Trudinger-Wang curvature and regularity of optimal transport maps.

(The final version of this article is published in *Centre de Recherches Mathématiques CRM Proceedings and Lecture Notes volume* 56*, 2013, pp119–143.)*

Young-Heon Kim

Abstract. Optimal transportation concerns the phenomena when the cost of matching two mass distributions is minimized. Regarding the regularity of such optimal transport maps, a new notion of curvature, called MTW curvature, was found recently by Ma, Trudinger and Wang. In these lectures, we discuss MTW curvature and regularity of optimal transport, focusing the case when the transportation cost is given by the Riemannian distance squared.

Contents

1. Introduction

The present article is based on four lectures given in the SMS 2011 Summer School (50th Edition) "Metric Measure Spaces: Geometric and Analytic Aspects" June 27 – July 8, 2011, in CRM, Montreal. In these lecture notes, the goal is to explain the new curvature notion, called Ma-Trudinger-Wang curvature (or simply MTW curvature) that was discovered by Ma, Trudinger and Wang [**MTW**], in

Key words and phrases. Optimal transportation, Regularity, Curvature.

The author is supported partly by Canada NSERC discovery grant 371642-09. Any opinions, findings and conclusions or recommendations expressed in this material are those of authors and do not reflect the views of Natural Sciences and Engineering Research Council of Canada.

the context of the regularity theory of optimal transportation maps. We study the geometry and analysis involving this curvature, explaining why this is relevant to the optimal transportation theory, and giving some key examples, and discussing how to prove regularity, in particular, Hölder continuity of optimal transportation maps when we restrict the sign of MTW curvature to be nonnegative. Some of the topics in these lecture notes overlap with other expository articles and books such as [**FiV**, **Fi2**] (focusing more on geometric implications of conditions on the MTW curvature), [**GlM**, **V1**, **V2**] (more broader introduction to optimal transportation), however, in addition to giving different point of view, we focus more on the results in [**Km**, **KmM1**, **KmM3**, **FiKM1**]. Our aim is to give some highlights in the recent developments, so many results/facts are given without detailed proof. Throughout this article, we assume familiarity with basic Riemannian geometry such as exponential map, sectional curvature, cut locus, etc (c.f. [**ChEb**]).

2. The Ma, Trudinger and Wang curvature condition

2.1. Optimal transportation. Throughout the article we let M be a Riemannian manifold, and let Ω, $\bar\Omega$ be bounded open subsets in M. We consider two probability distributions $\rho = \rho(x)dx$, $\bar\rho = \bar\rho(\bar x)d\bar x$ with $\operatorname{supp}\rho \subset \Omega$, $\operatorname{supp}\bar\rho \subset \bar\Omega$. Here, we assume $\log\rho, \log\bar\rho \in L^\infty$, namely, the density functions $\rho(x), \bar\rho(\bar x)$ satisfy

$$0 < \lambda \le \rho(x), \bar\rho(\bar x) \le \Lambda \tag{2.1}$$

for some positive constant λ, Λ and for $x \in \Omega$ and $\bar x \in \bar\Omega$. Consider the transportation cost function

$$c(x, \bar x) = \operatorname{dist}^2(x, \bar x)/2,$$

where dist denotes the Riemannian distance function. This function c can be regarded a canonical cost function associated to a Riemannian manifold since $-D_x c(x,\bar x) = \exp_x^{-1}\bar x$, where D_x denotes the gradient in the x variable[1] and exp denotes the exponential map. Even though more general cost functions can be considered, **in this lecture we restrict ourselves to this Riemannian distance squared cost function for simplicity of discussion.** In general, due to the cut locus, the distance squared function is not smooth. *To simplify our discussion, we assume throughout these lectures that $c \in C^\infty(\Omega, \bar\Omega)$: in other words, $\Omega \times \bar\Omega \subset M \times M \setminus \mathrm{Cut}$, where* Cut *denotes the cut locus of M.*

By the fundamental result of Brenier [**B**] and later by McCann [**Mc**], the following holds: There exists a unique optimal map T, namely, a Borel measurable map that minimizes the transportation cost

$$\int_\Omega c(x, F(x))d\rho(x)$$

among all measurable maps F pushing ρ forward to $\bar\rho$, i.e. $F_\#\rho = \bar\rho$: here, $F_\#\rho(B) = \rho(F^{-1}B)$ for any Boral set B. We can simply denote this as

$$T = \operatorname*{argmin}_{F_\#\rho=\bar\rho} \int_\Omega c(x, F(x))d\rho(x)$$

This *optimal map* T is almost everywhere differentiable and it satisfies nice analytical properties we list below:

[1] One may consider D_x as the differential producing covectors, but we can use the Riemannian metric to identify those with tangent vectors and this is the convention we take in these lectures.

(1) $|\det D_x T(x)| =_{a.e.} \frac{\rho(x)}{\bar\rho(T(x))}$.
(2) There exists a function $\phi : \Omega \to \mathbb{R}$ called *c-potential* such that $T(x) =_{a.e.} \exp_x(D_x\phi(x))$, i.e. $-D_x c(x, T(x)) =_{a.e.} D_x\phi(x)$. Such a function ϕ is *c-convex*, namely, it has a dual function $\phi^c : \bar\Omega \to \mathbb{R}$ and

$$\phi(x) = \sup_{\bar x \in \bar\Omega} -c(x, \bar x) - \phi^c(\bar x); \tag{2.2}$$
$$\phi^c(\bar x) = \sup_{x \in \Omega} -c(x, \bar x) - \phi(x).$$

Due to boundedness of $\Omega, \bar\Omega$ and local Lipschitzness and semi-convexity of $-\operatorname{dist}^2$, one can verify that ϕ and ϕ^c are Lipschitz and semi-convex: see Section 2.2.
(3) The condition $T_\# \rho = \bar\rho$ (see (1)) forces the c-potential ϕ satisfy the *c-Monge-Ampère equation*:

$$\det(D^2_{xx}\phi(x) + D^2_{xx}c(x, T(x)) =_{a.e.} |\det D_x D_{\bar x} c(x, T(x))| \frac{\rho(x)}{\bar\rho(T(x)}.$$

Example 2.1. For $M = \mathbb{R}^n$, $c(x, \bar x) = |x - \bar x|^2/2$, the function $\phi(x) + |x|^2/2$ is convex and $T(x) = x + \nabla\phi(x)$, with

$$\det(D^2_{xx}\phi(x) + I) = \frac{\rho(x)}{\bar\rho(\nabla\phi(x) + x)}.$$

The paper [**CoMS**] by Cordero-Erausquin, McCann and Schmuckenschlaeger contains many useful results for optimal transport maps on Riemannian manifolds. See also the book [**V2**] by Villani.

From the c-Monge-Ampère equation, the question of regularity of its solution ϕ naturally arises:

Question 2.2. For $\log\rho, \log\bar\rho \in L^\infty/C^\infty$, is the optimal map $T \in C^0/C^\infty(\Omega)$ (i.e. $\phi \in C^1/C^\infty(\Omega)$)?

We discuss below key notions related to this question.

2.2. The subdifferential $\partial\phi$ and the c-subdifferential $\partial^c\phi$. Under our assumption that Ω and $\bar\Omega$ are bounded, one can show the following fact:

If ϕ is the c-potential defined in (2.2), then ϕ is Lipschitz and semi-convex.

A function is called (locally) semi-convex if it becomes convex by adding a quadratic function (in a local coordinate system). Note that semi-convexity allows the function not to be differentiable at a point. But, also note that if a semi-convex function is differentiable at every point, then it is C^1, i.e. the derivatives are continuous. Moreover, it is a well-known fact (due to A.D. Alexandrov, see e.g. [**V2**]) that for a (locally) semi-convex function the set of nondifferentiable points has Hausdorfff dimension less than or equal to $\dim M - 1$. In particular, such set has zero measure. Example of Lipschitz and semi-convex functions include $-dist^2$ for compact Riemannian manifolds. Note that dist^2 is on the other hand is semi-concave but not semi-convex (see e.g. [**CoMS**]).

For semi-convex functions we can define the subdifferential.

Definition 2.3 (subdifferential). The subdifferntial $\partial\phi(x)$ at $x \in \Omega$ is defined as the set in the tangent space T_xM given by

$$\partial\phi(x) = \{p \in T_xM \mid \phi(\exp_x v) - \phi(x) \geq \langle p, v\rangle + o(|v|), \quad \forall v \in T_xM \ \& \ |v| \ll 1\}$$

Here, $\langle,\rangle$ and $|\cdot|$ are the Riemannian metric and norm, repectively, and $o(|v|)$ denotes the usual small 'o' error, i.e. $\lim_{|v|\to 0} o(|v|)/|v| = 0$.

Notice that $\partial\phi(x) = \{\nabla\phi(x)\}$ (here $\nabla\phi$ is the gradient) if and only if ϕ is differentiable at x. Moreover, one can check that $\partial\phi(x)$ is a convex set in the affine space T_xM.

While the subdifferential at a point assigns a function a set of tangent vectors, the c-subdifferential gives a set of 'target' points:

Definition 2.4 (c-subdifferential).

$$\partial^c\phi(x) = \{\bar{x} \in \bar{\Omega} \mid \phi(\cdot) \geq \phi(x) - c(\cdot, \bar{x}) + c(x, \bar{x}) \ on \ \Omega\}$$

The expression $\partial^c\phi$ denotes the graph of this multi-valued map, i.e.

$$\partial^c\phi = \{(x, \bar{x}) \in \Omega \times \bar{\Omega} \ |\bar{x} \in \partial^c\phi(x)\}$$

To be more precise, one may add subscripts as $\partial^c_{\Omega,\bar{\Omega}}\phi$ since the c-subdifferential depends on the source and target domains $\Omega, \bar{\Omega}$. Here, the functions of the form $-c(\cdot, \bar{x}) + c(x, \bar{x}) + const$ are called c-supporting functions.

Regarding an optimal map T and its c-potential ϕ, we have inclusions between the graphs of T, $\partial^c\phi$ and the multi-valued map $\exp\partial\phi$ given by composing the subdifferential $\partial\phi$ with the exponential map, whose graph is defined (by abusing the notation) as

$$\exp\partial\phi = \{(x, \bar{x}) \in \Omega \times \bar{\Omega} \mid \bar{x} \in \exp_x \partial\phi(x)\}.$$

Fact 2.1.

$$\operatorname{graph} T \subset \partial^c\phi \subset \exp\partial\phi$$

Since T is defined a.e., the first inclusion should be understood in the a.e. sense.

One observes that if $\phi \in C^1$, then $\partial^c\phi = \exp\nabla\phi = \exp\partial\phi$. Regarding the equality between $\partial^c\phi$ and $\exp\partial\phi$ we define *Loeper's maximum principle* [**Lo1**], a principal notion in these lectures:

Definition 2.5 (Loeper's maximum principle). We say *Loeper's maximum principle* **(LMP)** holds if for any c-convex function ϕ,

$$\partial^c\phi = \exp\partial\phi.$$

The reason why this is called a maximum principle will be obvious from its another formulation (2.3) given in a later section.

We now state the first main theorem of these lectures, which is due to Loeper [**Lo1**] (such a result was also hinted by Ma, Trudinger and Wang [**MTW**, Sections 7.3 and 7.5]:

Theorem 2.6 (continuity of OT $\Rightarrow$ **LMP**). *Suppose that for* each $\log\rho, \log\bar{\rho} \in L^\infty$*, the corresponding optimal map T is continuous. Then, Loeper's maximum principle holds.*

In the following subsection we explain the reason why this theorem should hold.

2.3. Loeper's Maximum Principle (LMP). In this section, we explain (not a proof) why Loeper's maximum principle **(LMP)** is a necessary condition to ensure continuity of optimal maps for each $\log\rho, \log\bar\rho \in L^\infty$. This will be done by considering the following important example, which shows that if for each $\log\rho, \log\bar\rho \in L^\infty$ the corresponding optimal map is continuous, then Loeper's Maximum Principle (**LMP**) should hold.

EXAMPLE 2.7 (Heuristic explanation why (continuity of OT $\Rightarrow$ **LMP**)). (See [**Lo1**, Proposition 4.4] and also [**MTW**, Section 7.3] for a similar example) Fix a point $x \in \Omega$ and two points $\bar x_0, \bar x_1 \in \bar\Omega$. Let

$$m_i(\cdot) = -c(\cdot, \bar x_i) + c(x, \bar x_i), \qquad i = 0, 1$$
$$\phi(\cdot) = \max[m_0, m_1]$$

as functions on Ω. Notice that the c-subdifferential $\partial^c\phi$ pushes forward the uniform measure 1_Ω to the sum of two Dirac measures, $\bar\rho_0 := c_0\delta_{\bar x_0} + c_1\delta_{\bar x_1}$, where $c_0, c_1 \in \mathbb{R}$ are some appropriate constants. We write this as

$$(\partial^c\phi)_\# 1_\Omega = \bar\rho_0.$$

We consider a smooth target probability density $\bar\rho_\varepsilon \in C^\infty(\bar\Omega)$ which is positive on $\bar\Omega$ and converges (weakly as measure) to the measure $\bar\rho_0$ as $\varepsilon \to 0$. Between 1_Ω and $\bar\rho_\varepsilon$, consider the corresponding optimal map with potential ϕ_ε. Namely,

$$(\partial^c\phi_\varepsilon)_\# 1_\Omega = \bar\rho_\varepsilon$$

Suppose "continuity of optimal transport", namely, for every $\log\rho, \log\bar\rho \in L^\infty$ *the corresponding optimal map is continuous.* We will see that **LMP** is a consequence of this assumption, thus a necessary condition. From this assumption we see that $\phi_\varepsilon \in C^1$, since the ensities 1_Ω and $\bar\rho_\varepsilon$ are bounded from above and below. Now, because $\phi_\varepsilon \in C^1$, it immediately holds

$$\partial^c\phi_\varepsilon = \exp\nabla\phi_\varepsilon = \exp\partial\phi_\varepsilon$$

Then, one can show (see Loeper [**Lo1**])) by taking limit $\varepsilon \to 0$, that

$$\partial^c\phi = \exp\partial\phi$$

The last equality is what is required by Loeper's maximum principle **(LMP)** for the c-convex function ϕ. One can in fact show that this special case implies **LMP** for general c-convex functions (see Loeper [**Lo1**])).

2.4. Geometric interpretation of Loeper's Maximum Principle (LMP). To understand a more geometric meaning of **LMP**, we first need a notion called c-segment that extends the notion of geodesic. (This definition is due to Ma, Trudinger and Wang [**MTW**].)

DEFINITION 2.8 (c-segment). Fix $x \in \Omega$. Let $p : t \in [0, 1] \to p(t) \in T_xM$ be a line segment, i.e. $p''(t) = 0$. Then, the curve $\bar x(t) = \exp_x p(t)$ is called a *c-segment* with respect to x. As a special case, if the line segment $p(t)$ passes through the origin in T_xM, the corresponding c-segment is a geodesic passing through x.

One can define similarly a c-segment $x(t)$ with respect to $\bar x \in \bar\Omega$.

DEFINITION 2.9 (sliding mountain). Let $\bar x(t)$ be a c-segment. Let $\bar x_i = \bar x(i)$, $i = 0, 1$. Define the function

$$m_t(\cdot) = -c(\cdot, \bar x(t)) + c(x, \bar x(t))$$

We can call this type of functions a *sliding mountain*.

One sees that

$$\{\nabla m_t(x)\}_{0\le t\le 1} = \big(\partial \max[m_0, m_1]\big)(x).$$

Now, Loeper's maximum principle **(LMP)** can be stated as the following: For all $x \in \Omega$ and for any c-segment $\{\bar{x}(t)\}_{0\le t\le 1}$ with respect to x,

$$\textbf{LMP:} \qquad m_t \le max[m_0, m_1] \qquad \forall\, 0 \le t \le 1 \;\; on \;\; \Omega; \tag{2.3}$$

$$\textbf{local LMP:} \qquad m_t \le max[m_0, m_1] \qquad \forall\, 0 \le t \le 1 \;\; \text{on a neighborhood of } x.$$

Thus, **LMP** prevents the function $m_t(z) : [0,1] \to \mathbb{R}$ (for fixed $z \in \Omega$) from having (in fact, local) maximum in the interior of the interval $[0,1]$.

2.5. Ma-Trudinger-Wang curvature condition and examples. There is an infinitesimal version of Loeper's maximum principle, called the Ma-Trudinger-Wang curvature condition [**MTW**]. As we will see in Section 3.1, this condition strengthens the sectional curvature nonnegativity condition. Recall $c(x,\bar{x}) = \mathrm{dist}^2(x,\bar{x})/2$.

Consider a pair $(x,\bar{x}) \notin \mathrm{Cut}$, i.e. c is C^∞ near $(x,\bar{x})$. Consider two curves $\{x(s)\}_{s\in[-1,1]} \in \Omega$, $\{\bar{x}(t)\}_{t\in[-1,1]} \in \bar{\Omega}$ where either of them is a c-segment with respect to $\bar{x}$, x, respectively. Let $x(0) = x$ and $\bar{x}(0) = \bar{x}$. Let us use the terminology *MTW-curvature* to describe the tensor quantity of Ma, Trudinger and Wang [**MTW**] called c-curvature by Loeper [**Lo1**], or called cross-curvature in [**KmM3**, Definition 1.1] [**KmM1**, (2.2) and Lemma 4.5].

Definition 2.10 (MTW-curvature). Let $(p,\bar{p}) \in T_xM \oplus T_{\bar{x}}M$.

$$\mathrm{MTW}_{(x,\bar{x})}(p,\bar{p}) = -\frac{d^4}{ds^2dt^2}\Big|_{(s,t)=(0,0)} c(x(s),\bar{x}(t))$$

Remark 2.11. The MTW-curvature is indeed a curvature since it is induced by the Riemannian curvature tensor of a pseudo-metric defined on the product space $\Omega \times \bar{\Omega}$ as found by McCann and the author [**KmM1**]. Furthermore, with Warren [**KmMW**] they extended this result to define another pseudo-metric (a conformal perturbation of the one in [**KmM1**]) and showed that the graph of the optimal map T in the product space $\Omega \times \bar{\Omega}$ gives a volume maximizing special Lagrangian submanifold, thus finding a connection to symplectic geometry.

Definition 2.12 (**MTW** condition). We say that **MTW** condition is satisfied if for all $(x,\bar{x}) \notin \mathrm{Cut}$,

$$\mathrm{MTW}_{(x,\bar{x})}(p,\bar{p}) \ge 0, \qquad \forall\, \langle p, (D\exp_x)^{-1}\bar{p}\rangle = 0.$$

Here, the last inner product is with respect to the Riemannian metric and can also be written as $p^i D_{x^i} D_{\bar{x}^j} c\, \bar{p}^j = 0$.

We say that $\mathbf{MTW}_+$ is satisfied if **MTW** is satisfied and if $p = 0$ or $\bar{p} = 0$ in case of equality in the inequality in Definition 2.12.

These **MTW**, $\mathbf{MTW}_+$ conditions were originally called **A3w**, **A3**, respectively by Ma, Trudinger and Wang [**MTW**, **TW1**].

We say that **NNCC** (nonnegative cross curvature) is satisfied if for all $(x,\bar{x}) \notin$ Cut,

$$\mathrm{MTW}_{(x,\bar{x})}(p,\bar{p}) \ge 0.$$

Notice that $\mathbf{MTW}_+ \Longrightarrow \mathbf{MTW}$ and $\mathbf{NNCC} \Longrightarrow \mathbf{MTW}$, but neither $\mathbf{MTW}_+$ or **NNCC** implies the other.

REMARK 2.13.
- The $\mathbf{MTW}_+$ condition was introduced in [**MTW**] to get a priori estimates for Monge-Ampère type equations for optimal transport problems, and the method goes back to the work of Wang on reflector antenna problems [**Wn1**, **Wn2**]. Loeper [**Lo2**] verified $\mathbf{MTW}_+$ for the cost function arising in the far-field reflector antenna problem and then showed regularity of the solution (see also a previous work of Caffarelli, Gutierrez and Huang [**CaGH**] for a different approach). More recently, there is a work by Karakhanyan and Wang [**KaW**] that uses a variant of $\mathbf{MTW}_+$ condition to give a rather complete solution to the regularity of the general (near-field) reflector antenna problem.
- It is a folk-lore conjecture among experts that for $c = \mathrm{dist}^2/2$ on Riemannian manifold, **MTW** implies **NNCC**.
- The condition **NNCC** has an unexpected application to principal-agent problem in microeconomics theory [**FiKM3**] (see also [**GlM**] for an exposition). On the other hand, Sei [**S**] found applications to statistical problems. In both applications, it was used that under **NNCC**, the set of c-convex functions is convex, that is, if ϕ_0 and ϕ_1 are c-convex then $(1-t)\phi_0 + t\phi_1$ is c-convex, too [**FiKM3**, **S**].

It is immediate to see that for the Euclidean space $(\mathbb{R}^n, g_0)$, the MTW-curvature completely vanishes: MTW $\equiv 0$.

Loeper [**Lo1**] found a connection of MTW-curvature to the Riemannian curvature.

THEOREM 2.14 (Loeper [**Lo1**, **Lo2**]).
(1) *For $x = \bar{x}$, $\mathrm{MTW}_{(x,x)}(p, \bar{p}) = \frac{4}{3}K(p \wedge \bar{p})$, where K denotes the Riemannian sectional curvature. In particular,* **MTW** *implies nonnegative sectional curvature.*
(2) **MTW** $\Longleftrightarrow$ **local LMP** *(see [**KmM1**] for an elementary geometric proof for $\Rightarrow$.)*
(3) *If the sectional curvature K is negative somewhere on Ω, then there exists a discontinuous optimal map $T : \Omega \to \bar{\Omega}$, with smooth densities $\rho, \bar{\rho}$. (Here, the domains $\Omega, \bar{\Omega}$ can have any nice properties (smoothness, convexity, etc).*
(4) *The round sphere (S^n, g_0) satisfies* $\mathbf{MTW}_+$ *and* **LMP** *(see also [**KmM1**, **KmM2**]).*

REMARK 2.15. Villani [**V3**] showed that the **MTW** condition is stable under Gromo-Hausdorff convergence, under suitable additional assumptions that give an equivalence between **MTW** and **LMP**. Note that **LMP** is more suitable for synthetic formulation than the **MTW** involving fourth order derivatives.

The statement (3) in this theorem is basically due to (1) and (2) and the fact that **LMP** is a necessary condition for regularity of optimal maps. In fact, one can also show that even positive curvature restriction is not enough for **MTW**, so for the regularity of optimal maps.

THEOREM 2.16 ([**Km**]). *There are positively curved manifolds that do not satisfy* **MTW**.

This is the first result showing the **MTW** condition is stronger than the positive curvature. Since then, there appeared other examples in this spirit [**LoV**, Appendix D] [**FiRV2**]. Especially, in [**FiRV2**], it is shown that two dimensional ellipsoid surfaces in $\mathbb{R}^3$, when they are thin enough in one direction, do not satisfy **MTW** condition.

Regarding (4) in Theorem 2.14, in fact, a stronger result holds, from which we can produce a lot of **NNCC** (thus **MTW**) examples of Riemannian manifolds:

THEOREM 2.17 (see [**KmM3**]).
(1) *The round sphere* (S^n, g_0) *satisfies* **NNCC**.
(2) *For the Riemannian submersion* $\pi : M \to B$, *if* M *satisfies* **MTW**, **MTW**$_+$, **NNCC**, **LMP**, *respectively, then* B *satisfies the corresponding conditions, respectively. In particular, the complex projective space* $\mathbf{CP}^n$ *with Fubini-Study metric satisfies all these conditions (because* (S^{2n+1}, g_0) *does).*
(3) *Let* M_i, $i = 1, 2$ *satisfy* **NNCC** *then the Riemannian product* $M_1 \times M_2$ *satisfies* **NNCC**.

REMARK 2.18.
- For example, $S^{n_1}(r_1) \times \cdots \times S^{n_j}(r_j) \times \mathbf{CP}^{l_1} \times \cdots \times \mathbf{CP}^{l_k} \times \mathbb{R}^m$ satisfies **NNCC**. It is shown that [**KmM1**, **KmM3**] this example satisfies **LMP**. (See also [**LoV**] [**V2**, Ch. 12] [**FiRV1**] where they extended the method in [**KmM1**] for deriving **LMP** to handle more general cases.)
- In [**KmM3**], an O'Neill type inequality, which says Riemannian submersions increase curvature (see [**ChEb**]), is obtained regarding the MTW-curvature (Definition 2.10), and this yields the statement (2) for **MTW**, **MTW**$_+$ and **NNCC** in Theorem 2.17.
- Regarding the statement (2) in Theorem 2.17, there are examples of **MTW**$_+$ obtained by perturbing the round sphere as shown by Delanoë-Ge [**DG1**], Figalli, Rifford and Villani [**FiR**, **FiRV1**, **FiRV3**]. In particular, Delanoë and Ge showed that small perturbations of the two dimensional round sphere are **NNCC** [**DG2**]: this is not known in higher dimensions.
- Regarding the statement (3) in Theorem 2.17, the same result does not hold for **MTW**: see [**KmM3**].

- There are examples of costs satisfying **MTW** condition, originated from mechanical action, found by Lee and McCann [**LeM**], and also from a modified distance function on the *hyperbolic* space found by Lee and Li [**LeLi**].

3. Geometry of MTW curvature condition

The goal of this section is twofold: first, we will explain the result in Theorem 2.16, second, we will explain a quantitative version of **LMP** for **MTW**$_+$ (which is due to Loeper [**Lo1**]), and as an application will show continuity of optimal maps on the round sphere. The latter result was actually shown with Hölder continuity [**Lo1**].

3.1. MTW is finer than positive sectional curvature. In this subsection, we explain why Theorem 2.16 holds: see [**Km**] for more details.

Consider a surface M such that in a small neighbourhood, say $\mathcal{N}$, of a point $x_0 \in M$, the sectional curvature K is positive and outside the neighbourhood,

$K = 0$. Here the curvature is bounded to be sufficiently small. One can construct such a surface by rounding about the vertex of a sufficiently thin cone. We will show that such surface does not satisfy **local LMP**, thus not **MTW** (see Theorem 2.14). Of course, this surface is not positively curved, but one can slightly perturb the surface to get a positively curved example that does not satisfy **MTW**.

Fix a point $x \in M$ outside $\mathcal{N}$. Consider the exponential map $\exp_x : T_xM \to M$. For simplicity we assume that $\exp_x$ is injective everywhere, but in general, under the sufficiently small bound on the curvature, one can find a neighbourhood such that all relevant points in the following discussion are within the injectivity radius (see [**Km**]). Find a line segment $\{\bar{p}_t\}_{0\leq t\leq 1}$in T_xM such that (i) the curve $\bar{x}(t) = \exp_x \bar{p}_t$ contains x_0, i.e. there is $t_0 \in (0,1)$ such that $x_0 = \bar{x}(t_0)$, (ii) moreover, its two end points $\bar{x}(0)$ an $\bar{x}(1)$ are outside $\mathcal{N}$, and (iii) there exists a geodesic segment γ outside $\mathcal{N}$ passing through the points x and $\bar{x}(0)$ such that it is orthogonal to the segment $\{\bar{x}(t)\}_{0\leq t\leq 1}$ at the point $\bar{x}(0)$. Let y denote a point in γ close to x but farther from $\bar{x}(0)$ than x. Let $p \in T_xM$ such that $\exp_x p = y$. We are now ready for the following argument: Define $f(t), \tilde{f}(t)$ as

$$\begin{aligned} f(t) &= -\operatorname{dist}^2(y, \bar{x}(t)) + \operatorname{dist}^2(x, \bar{x}(t)) \\ \tilde{f}(t) &= -|p - \bar{p}_t|^2 + |\bar{p}_t|^2 \end{aligned}$$

We will violate **local LMP** (see (2.3)) by showing that $f(t_0) > \max[f(0), f(1)]$. This will be done by comparing the two functions $f(t)$ and $\tilde{f}(t)$. Notice that $f(0) = \tilde{f}(0)$ and $f(1) = \tilde{f}(1)$. Moreover, $\tilde{f}(t) \equiv \tilde{f}(0) = \tilde{f}(1)$ for all $t \in [0,1]$. (This is due to the orthogonality of γ and $\{\bar{x}(t)\}$ at $\bar{x}(0)$.) Now, use the well-known Toponogov theorem (see Cheeger and Ebin [**ChEb**]),

Theorem 3.1. **(Toponogov's comparison theorem)** *Let M be a complete Riemannian manifold with sectional curvature $K_M \geq H$, and let M^H be the simply connected 2-dimensional space of constant curvature H. Let $\gamma_i : [0,1] \to M$ and $\bar{\gamma}_i : [0,1] \to M^H$, $i = 1,2$, be minimal geodesic segments, i.e. they are unique geodesic segments connecting their end points. Suppose that $\gamma_1(0) = \gamma_2(0)$, $\bar{\gamma}_1(0) = \bar{\gamma}_2(0)$; $\measuredangle(\dot{\gamma}_1(0), \dot{\gamma}_2(0)) = \measuredangle(\dot{\bar{\gamma}}_1(0), \dot{\bar{\gamma}}_2(0)) < \pi$, where $\measuredangle$ denotes the angle between tangent vectors. Assume $L[\gamma_i] = L[\bar{\gamma}_i]$, $i = 1,2$, where L denotes arc-length. Then*

$$\operatorname{dist}(\gamma_1(1), \gamma_2(1)) \leq \operatorname{dist}(\bar{\gamma}_1(1), \bar{\gamma}_2(1)), \tag{3.1}$$

where dist *denotes the Riemannian distance. Moreover, if there exists a point z on $\gamma_1 \cup \gamma_2 \subset M$ such that $K_M(z) > H$, then the inequality* (3.1) *is strict.*

Applying this theorem, we see that

$$|\bar{p}_{t_0} - p| > \operatorname{dist}(y, \bar{x}(t_0))$$

because the sectional curvature of M, $K \geq 0$. Here, the strict inequality is due to the condition $K(\bar{x}(t_0)) > 0$. Therefore, we see $f(t_0) > \tilde{f}(t_0) = f(0) = f(1)$, and this violates **local LMP**, thus **MTW**. Thus, the surface M does not satisfy **MTW**. This finishes the discussion of this subsection.

3.2. An open question. The result in the previous subsection motivates us to consider finer relation between the sectional curvature and the MTW curvature. Regarding this, the following question is raised by Trudinger:

Question 3.2. Do there exist appropriate norms $\|\cdot\|_1$, $\|\cdot\|_2$ and an $\epsilon > 0$ such that $\|\nabla R\|_1 \leq \epsilon \|R\|_2$ implies **MTW**?

Here, R denotes the Riemannian curvature tensor, and ∇R is its covariant derivative (thus a tensor itself). See [**FiR**, **DG1**] for partial results in this direction, in the case of perturbations of the round sphere. Answers to Question 3.2 are useful to obtain a robust method of finding Riemannian manifolds that satisfy **MTW**, especially since it is very hard to verify the **MTW** condition in general. Note that there are a few works for finding effective criteria of **MTW** condition (see [**FiRV2**, **Le**]). Also, there is a relation between the **MTW** condition and the convexity of tangential injectivity domains [**LoV**, **FiR**, **FiRV1**, **FiRV3**].

3.3. Loeper's quantitative maximum principle (LQMP) under $\mathbf{MTW}_+$.

We now state a quantitative version of Loeper's maximum principle and its applications to the continuity of optimal maps. The original results in this subsection are due to Loeper [**Lo1**].

THEOREM 3.3 (**LQMP**). *Assume that the cost c satisfies* $\mathbf{MTW}_+$. *Namely,* $\mathrm{MTW}(p, \bar p) \geq K_0 |p|^2 |\bar p|^2$ *for* $\langle p, D \exp^{-1} \bar p\rangle = 0$. *Let $\bar x(t)$ be a c-segment with respect to x. Define*

$$m_t(\cdot) = -c(\cdot, \bar x(t)) + c(x, \bar x(t)).$$

Then, there exist r_0 and K_1 (both depending on the cost function c, especially on K_0) such that $\forall 0 < r \leq r_0$ *and* $\forall z \in B_r(x)$, $\forall 0 \leq t \leq 1$,

$$\begin{aligned} m_t(z) \leq \quad & \max[m_0(z), m_1(z)] \\ & - K_1 t(1-t) \operatorname{dist}^2(z,x) \operatorname{dist}^2(\bar x(0), \bar x(1)) + \|c\|_{C^3} \operatorname{dist}^3(z,x). \end{aligned}$$

This theorem was originally proved in [**Lo1**]: see [**KmM3**, Appendix] for a different proof. Also, see [**LoV**] and [**FiRV1**] for an improved version of this result.

Loeper's quantitative maximum principle (**LQMP**) shows that the gap between the double mountain like function $\max[m_0, m_1]$ and the sliding mountain like function m_t, is quadratic in the distance near the point x along where they coincide. To have this estimate turns out to be very useful for showing regularity (continuity) of optimal maps as we see below.

3.3.1. *Application of Loeper's quantitative maximum principle (LQMP).* Loeper's quantitative maximum principle is a powerful tool for proving regularity of optimal maps on positively curved domains with $\mathbf{MTW}_+$. In particular, one can show Hölder continuity of optimal maps on S^n [**Lo1**, **Lo2**], $\mathbf{CP}^n$ [**KmM3**], and perturbations of S^n [**LoV**, **FiR**, **FiRV1**] for source and target measures bounded from above and below, i.e. $\log \rho, \log \bar\rho \in L^\infty$.

We show the following, originally due to Loeper, as an example.

THEOREM 3.4 ([**Lo1**, **Lo2**] (see also [**KmM2**])). *Let $\Omega = \bar\Omega = S^n$, the round sphere. Assume that the source and target measures satisfy $0 < \lambda \leq \rho, \bar\rho \leq \Lambda$ where λ, Λ are constants. Let T be the optimal map $T_\# \rho = \bar\rho$. Then, $T \in C^0(\Omega)$ (in fact $T \in C^\alpha(\Omega)$ for some $0 < \alpha < 1$, depending on the dimension n).*

We note here that a sharp Hölder exponent $\alpha = \frac{n+1}{2n^2+n-1}$ is obtained by Liu [**Li**]. Also, if further $\rho, \bar\rho \in C^\infty$, then by applying [**MTW**], $T \in C^\infty$ as in [**Lo2**]. The continuity method of [**MTW**] also applies to get smooth optimal maps on perturbations of the round sphere [**DG1**].

Proof of Theorem 3.4. In this proof, we ignore the technical problem that the cost fucntion dist^2 is not smooth for antipodal pairs in S^n: these points where dist^2 fails to be smooth is called *cut-locus*: we discuss this issue in Section 7.

Recall that the round sphere satisfies $\mathbf{MTW}_+$, thus both **LMP** and **LQMP**. The following argument can be called "Sausage-Meat Ball Argument".

Let ϕ be the c-potential function for the optimal map T, i.e. $\mathrm{graph}\, T \subset \partial^c\phi$. Suppose by contradiction that $T \notin C^0$. Then there exists a point x such that the c-subdifferential $\partial^c\phi(x)$ has two distinct points $\bar{x}_0, \bar{x}_1$ ($\bar{x}_0 \neq \bar{x}_1$). Let $\bar{x}(t)$ be a c-segment between $\bar{x}_0 = \bar{x}(0)$ and $\bar{x}_1 = \bar{x}(1)$ with respect to x. Define for each $\delta > 0$, a tubular neighborhood

$$N_\delta = \{\bar{z} \mid \mathrm{dist}(z, \bar{x}(t)) \leq \delta, \quad \frac{1}{3} \leq t \leq \frac{2}{3}\}.$$

We will use the following result which we will show later:

Claim 3.5. Recall r_0 from **LQMP** (Theorem 3.3). There exists $C_1 > 0$ (depending only on the cost c and $\mathrm{dist}(\bar{x}(0), \bar{x}(1))$) small enough such that for all $r < r_0$, if $\delta = C_1 r$, then $N_\delta \subset \partial\phi(B_r(x))$.

Choose r and δ as in this claim. Now, observe that

$$\bar{\rho}(\partial^c\phi(B_r(x)) = \bar{\rho}(T(B_r(x)) = \rho(B_r(x))$$

Use Claim 3.5 and $\lambda \leq \rho, \bar{\rho} \leq \Lambda$, to see that

$$\bar{\rho}(\partial^c\phi(B_r(x)) \geq \bar{\rho}(N_\delta) \gtrsim \delta^{n-1} \sim r^{n-1}.$$

But, on the other hand, $\rho(B_r(x)) \sim r^n$. Thus, comparing these, we have $r^n \gtrsim r^{n-1}$. Letting $r \to 0$, we get a contradiction. This shows Theorem 3.4 □

Proof of Claim 3.5. Recall $m_t(\cdot) = -c(\cdot, \bar{x}(t)) + c(x, \bar{x}(t))$. Define

$$m_{\bar{z}}(\cdot) = -c(\cdot, \bar{z}) + c(x, \bar{z}).$$

The following simple estimates will be useful:

$$\begin{aligned} m_{\bar{z}}(z) - m_t(z) \leq &\|D_x D_{\bar{x}} c\| \,\mathrm{dist}(z, x)\,\mathrm{dist}(\bar{z}, \bar{x}(t)) \\ &+ \text{higher order terms of } \mathrm{dist}^2(z, x), \text{ etc.} \end{aligned} \tag{3.2}$$

We will show that there exists $C_1 > 0$ that for each $r \leq r_0$, $\delta = C_1 r$, $\bar{z} \in N_\delta$,

$$m_{\bar{z}}(z) \leq \phi(z) \qquad \forall z \in \partial B_r(x). \tag{3.3}$$

This then will imply, by comparison principle and the definition of $\partial^c\phi$, that $\bar{z} \in \partial^c\phi(B_r)$, completing the proof. Now it remains to show (3.3). Notice that $\phi(\cdot) \geq \max[m_0(\cdot), m_1(\cdot)]$. Let $z \in B_r(x)$, $r \leq r_0$. Then, by (3.2) and **LQMP**,

$$\begin{aligned} m_{\bar{z}}(z) &\leq m_t(z) + \|D_x D_{\bar{x}} c\|\,\mathrm{dist}(z, x)\,\mathrm{dist}(\bar{z}, \bar{x}(t)) \\ &\leq \max[m_0(z), m_1(z)] - K_1 t(1-t) r^2\,\mathrm{dist}^2(\bar{x}(0), \bar{x}(1)) + \|c\|_{C^2} C_1 r^2 + \text{higher order in } r. \end{aligned}$$

By choosing C_1 small enough, we see that for $1/3 \leq t \leq 2/3$, the last line is bounded above by $\max[m_0(z), m_1(z)]$, thus by $\phi(z)$. We showed (3.3). □

4. Hölder continuity of optimal transport maps under MTW without LQMP.

Under only **MTW** (without **MTW**$_+$), the Loeper's quantitative maximum principle (**LQMP**) is not available anymore, making the analysis of optimal maps more difficult. Nevertheless, we can show (interior) regularity of optimal maps in this case [**FiKM1**]. Notice that the special case $c(x, \bar{x}) = -x \cdot \bar{x}$ (equivalently $c(x, \bar{x}) = \frac{1}{2}|x - \bar{x}|^2$) in $\mathbb{R}^n$, is addressed in the pioneering work of Delanoë [**D**], Urbas [**U**] and Caffarelli [**Ca1**, **Ca2**, **Ca3**, **Ca4**, **Ca5**] on the regularity of Monge-Ampère equation. Especially, in this case, Caffarelli has obtained Hölder continuity of optimal maps with measurable data, namely, assuming only L^∞ (upper and lower) bounds (2.1) on the source and target densities ρ, $\bar{\rho}$. One can view the results below (Theorem 4.3) as an extension of Caffarelli's methods and results to more general cost functions. One of the novelties here is that one now can handle domains in Riemannian manifolds: e.g. products of round spheres [**FiKM2**]. We remark that Liu, Trudinger and Wang have obtained higher regularity results with more regular data, using continuity methods [**TW1**, **LTW**] (see also [**LiT**]). Such continuity method is not available for merely measurable data, and we need more geometric arguments for the analysis.

4.1. (Interior) Hölder continuity of optimal transport maps. Throughout this section we consider domains Ω', $\bar{\Omega}$ in an n-dimensional Riemannian manifold M with $(\Omega' \times \bar{\Omega}) \cap \mathrm{Cut} = \emptyset$ so that the cost function $c(x, y) = \mathrm{dist}^2(x, y)/2$ is smooth on $\Omega' \times \bar{\Omega}$. Here, the domain Ω' is an open set containing the source domain Ω of the optimal transportation.

Definition 4.1 ((strong) c-convexity of $\bar{\Omega}$ with respect to Ω' [**MTW**]). We say that $\bar{\Omega}$ is (strongly) c-convex with respect to Ω', if for all $x \in \Omega'$, the inverse image $\exp_x^{-1} \bar{\Omega}$ is (strongly) convex as a subset in the tangent space $T_x\Omega$: use the Riemannian metric in $T_x\Omega$ to measure how strong the set is convex. Recall that a set in $\mathbb{R}^n$ is strongly convex if the set is an intersection of balls of uniformly upper bounded radius. The smaller this bound, the stronger the convexity.

Similary, we define (strong) c-convexity of Ω' with respect to $\bar{\Omega}$.

Remark 4.2. The c-convexity on the domain $\bar{\Omega}$ is a necessary condition for regularity theory of optimal maps. For instance, for the case $c(x, \bar{x}) = -x \cdot \bar{x}$, Caffarelli showed a counterexample [**Ca3**] to regularity (in fact, continuity) where the target domain is not convex (thus not c-convex). Ma, Trudinger and Wang [**MTW**][Section 7.3] showed similar example for more general cost functions.

We now state the main theorem of this section:

Theorem 4.3. *([**FiKM1**, Theorem 2.1]) Assume that*
- *Ω', $\bar{\Omega}$ are strongly c-convex with respect to each other;*
- *$\log \bar{\rho} \in L^\infty(\bar{\Omega})$, $\log \rho \in L^\infty(\Omega)$, $\Omega \subset \Omega'$ is an open set. The set Ω is not necessarily c-convex;*
- *T is an optimal map with $T_\# \rho = \bar{\rho}$;*
- *the cost $c = \mathrm{dist}^2/2$ satisfies* **MTW**.

Then,
1. *$T \in C^\alpha_{loc}(\Omega)$.*
2. *the restriction $T\big|_\Omega$ of T to Ω, is one-to-one.*

REMARK 4.4.
- In two dimensions, Figalli and Loeper [**FiL**] obtained continuity of optimal maps without assuming that the source measure ρ is bounded from below. Their method goes back to the classical work of Alexandrov [**A**].
- The Hölder exponent α depends only on the dimension n and the upper and lower bounds of ρ and $\bar{\rho}$, in particular not on the specific cost c: see [**FiKM1**, Section 9].
- When $\rho, \bar{\rho} \in C^\infty$, one can apply the above injectivity (Theorem 4.3 (2)) to the result of [**LTW**] (see also [**LiT**]) to get $T \in C^\infty(\Omega)$.

4.2. Tools for regularity of optimal transport maps. In this subsection, we explain a few tools for the regularity of optimal maps as in Theorem 4.3, which are available without the **MTW** assumption.

4.2.1. *(weak) c-Monge-Ampère equation.* Let $T_\# \rho = \bar{\rho}$ be the optimal map push-forwarding ρ onto $\bar{\rho}$ with the corresponding c-potential ϕ. Assume that ρ, $\bar{\rho}$ are bounded away from zero and infinity, namely, $\log \rho \in L^\infty(\Omega)$ and $\log \bar{\rho} \in L^\infty(\bar{\Omega})$. Then, it is well-known that ϕ satisfies the following weak form of the c-Monge-Ampère equation: see for example, [**FiKM1**, Lemma 3.1 (e)]. Namely, there exists a constant $\lambda > 0$ such that

$$(\mathbf{MA}_\lambda) \cdots\cdots \lambda |B| \leq |\partial^c \phi(B)| \leq \frac{1}{\lambda}|B| \qquad \forall \text{ Borel susbset } B \subset \Omega$$

Here, $\partial^c\phi(B) = \cup_{x \in B} \partial^c \phi(x)$ and $|B| = \int_B d\,\mathrm{vol}$.

The above condition $\mathbf{MA}_\lambda$ can be denoted simply as $|\partial^c\phi| \sim 1_\Omega$, since it says the c-Monge-Ampère measure $|\partial^c\phi|$, defined as $|\partial^c\phi|(B) = |\partial^c\phi(B)|$, is equivalent to the uniform measure on Ω.

EXAMPLE 4.5. Let ϕ be the c-cone on $\mathbb{R}^n$ (with $c(x,y) = |x-y|^2/2$),

$$\phi(x) = \sup_{y \in B_1(0)} -|x-y|^2 + |y|^2$$

Then, one can see that $|\partial^c\phi| \sim \delta_0$, the Dirac-delta measure at 0, because $\partial^c\phi(B) = B_1(0)$ for any B containing 0. Thus ϕ does not satisfy the above c-Monge-Ampère equation $\mathbf{MA}_\lambda$.

4.2.2. *Interior-not-to-boundary result for optimal maps.* Another important tool is a lemma that assures that the optimal transport map does not mix the interior points with the boundary, at least if the domains satisfy appropriate convexity conditions. More precisely,

THEOREM 4.6 (Interior-not-to-boundary). (See [**FiKM1**, Theorem 5.1]) *Assume that*
- $(\Omega' \times \bar{\Omega}) \cap \mathrm{Cut} = \emptyset$.
- $\Omega', \bar{\Omega}$ *strongly c-convex with respect to each other;*
- $\Omega \subset \Omega'$*;*
- $\partial^c\phi(\Omega) \subset \bar{\Omega}$*;*
- $|\partial^c\phi| \sim 1_\Omega$.

Then,
(1) $\partial^c\phi(\mathrm{int}\,\Omega) \cap \partial\bar{\Omega} = \emptyset$*;*
(2) $\partial^c\phi(\partial\Omega') \cap \mathrm{int}\,\bar{\Omega} = \emptyset$

REMARK 4.7.
- Notice that the **MTW** condition is not assumed in this theorem.

- This theorem was a necessary ingredient in the paper of Figall and Loeper [**FiL**] where they showed the same result in the two dimensions without the lower bound of the source measure ρ.

For the proof of the above theorem, we recall the c-monotonicity of $\partial^c\phi$. When $c(x,y) = -x \cdot y$ on $\mathbb{R}^n$ where $\partial^c\phi = \partial\phi$ (the subdifferential) and c-convex functions are nothing but convex functions, it reads as

$$(x,\bar{x}),(z,\bar{z}) \in \partial\phi \quad \Longrightarrow \quad \langle z-x, \bar{z}-\bar{x}\rangle \geq 0.$$

For more general case, it reads as

$$(x,\bar{x}),(z,\bar{z}) \in \partial^c\phi \quad \Longrightarrow \quad -c(z,\bar{z}) + c(x,\bar{z}) + c(z,\bar{x}) - c(x,\bar{x}) \geq 0.$$

The c-monotonicity roughly says that infinitesimally, the (multi-valued) map $\partial^c\phi$ is irrotational.

Idea of proof of Theorem 4.6. We will only show the assertion (1). This is enough for showing the idea. For simplicity, we present only the case when $M = \mathbb{R}^2$ and $c(x,y) = -x\cdot y$. Then, the function ϕ is convex (thus $\partial^c\phi = \partial\phi$) and $\bar{\Omega}$ is a strongly convex set.

Suppose by contradiction that there is a pair $(x,\bar{x})$ with $x \in \operatorname{int}\Omega$ and $\bar{x} \in \partial\phi(x)\cap\partial\bar{\Omega}$. Find a vector v such that the normal plane $N_v(\bar{x}) = \{z \mid \langle z-\bar{x}, v\rangle = 0\}$ at $\bar{x}$ has the unique intersection $\bar{x}$ with $\bar{\Omega}$. Moreover, by the strong convexity of $\bar{\Omega}$, the boundary $\partial\bar{\Omega}$ looks like the graph of a quadradic function over $N_v(\bar{x})$: we can give coordinates $(\bar{x}^1,\bar{x}^2) \in N_v(\bar{x})\times\mathbb{R}$, such that $\bar{\Omega}$ is above the graph of the function $\bar{x}^2 = C|\bar{x}^1|^2$, and that $v = (0,-1)$. Now, consider for $\theta,\epsilon > 0$ small enough, the conical set $E_{\theta,\epsilon} \subset \Omega$ defined as

$$E_{\theta,\epsilon} = \{z \mid \langle z-x, v\rangle \geq \cos\theta |z-x||v| \ \&\ |z-x| \leq \epsilon\}$$

Note that $|E_{\theta,\epsilon}| \sim \theta\epsilon^2$. Define

$$\bar{E}_\theta = \big\{\bar{z} \mid \langle \bar{z}-\bar{x}, v\rangle \geq \cos\big(\frac{\pi}{2}+\theta\big)\big\} \cap \bar{\Omega}\big\}.$$

One can compute using the quadratic function $\bar{x}^2 = C|\bar{x}^1|^2$ as above, that $|\bar{E}_\theta| \lesssim \theta^3$.

But, by monotonicity, $\partial\phi(E_{\theta,\epsilon}) \subset \bar{E}_\theta$. Thus, the desired contradiction follows by comparing volumes:

$$\begin{aligned}\theta\epsilon^2 \sim |E_{\theta,\epsilon}| &\sim |\partial\phi(E_{\theta,\epsilon})| \qquad (\text{by } |\partial^c\phi| \sim 1_\Omega)\\ &\leq |\bar{E}_\theta| \lesssim \theta^3.\end{aligned}$$

Let $\theta \to 0$ while fixing ϵ, then we get a contradiction. This finishes the proof. □

4.3. Why is MTW good for regularity of optimal transport maps? Appeared convexity! To see how the **MTW** condition affects the geometry of c-convex functions, we consider the following transformation of the coordinates and the functions.

We first assume that all relevant points are outside the cut-locus so that the cost function is smooth, and moreover the exponential map is invertible on such points.

Fix $\bar{x}_0 \in M$. Give correspondence between $q \in T_{\bar{x}_0}M$ and $x \in M$ as

$$x(q) = \exp_{\bar{x}_0} q, \qquad\qquad q(x) = \exp_{\bar{x}_0}^{-1} x.$$

Now modify the cost function and a c-convex function ϕ in the q variable as

$$\tilde{c}(q,\bar{x}) := c(x(q),\bar{x}) - c(x(q),\bar{x}_0),$$
$$\tilde{\phi}(q) := \phi(x(q)) + c(x(q),\bar{x}_0)$$

Notice that $\tilde{\phi}$ is $\tilde{c}$-convex and if $B = \exp_{\bar{x}_0}\tilde{B}$ then $\partial^c\phi(B) = \partial^{\tilde{c}}\tilde{\phi}(\tilde{B})$, namely the image of c-subdifferential is not changed under this transformation of c and ϕ.

This transformation is very useful under the **MTW** condition because of the following result [**FiKM1**, **Li**]:

THEOREM 4.8 (appeared convexity). *Assume Ω, $\bar{\Omega}$ be c-convex with respect to each other. Let ϕ be c-convex in the variable $x \in \Omega$.*
(1) *If* **MTW** *holds for c, then $\tilde{\phi}$ is level set convex in the variable $q \in \exp_{\bar{x}_0}^{-1}\Omega$.*
(2) *If* **NNCC** *holds for c, then $\tilde{\phi}$ is convex in the variable $q \in \exp_{\bar{x}_0}^{-1}\Omega$.*
Here, level-set convexity means that each sub-level set $\{\tilde{\phi} \le k\}$, $k \in \mathbb{R}$, is a convex set.

REMARK 4.9. Notice that the above statement (1) is a direct consequence of Loeper's maximum principle **LMP** 2.3.

EXAMPLE 4.10. , Recall that in $\mathbb{R}^n$ with the quadratic cost $(x,\bar{x}) = |x-\bar{x}|^2/2$, the cross curvature in 2.10 vanishes identically, thus **NNCC** holds. In this special case, the exponential map $x(q) = \exp_{\bar{x}_0} q = q + \bar{x}_0$, thus, the above transform becomes

$$\begin{aligned}\tilde{c}(q,\bar{x}) &= \frac{1}{2}|q+\bar{x}_0-\bar{x}|^2 - \frac{1}{2}|q|^2 \\ &= -q\cdot(\bar{x}-\bar{x}_0) + \frac{1}{2}|\bar{x}-\bar{x}_0|^2,\end{aligned}$$

which is linear (thus convex) in q. Note that in this case c-convex functions are not convex; for example, consider the c-convex function $\phi(x) = \max[-\frac{1}{2}|x-\bar{x}_1|^2, -\frac{1}{2}|x-\bar{x}_2|^2]$ for two fixed $\bar{x}_0 \neq \bar{x}_1 \in \mathbb{R}^n$. However, the transformed function $\tilde{\phi}(q) = \phi(x(q)) + \frac{1}{2}|x(q),\bar{x}_0|^2$ is convex, for example,

$$\begin{aligned}&\max[-\frac{1}{2}|x-\bar{x}_1|^2, -\frac{1}{2}|x-\bar{x}_2|^2] + \frac{1}{2}|x(q)-\bar{x}_0|^2 \\ &= \max\big[-q\cdot(\bar{x}_1-\bar{x}_0) - \frac{1}{2}|\bar{x}_1-\bar{x}_0|^2,\ q\cdot(\bar{x}_2-\bar{x}_0) - \frac{1}{2}|\bar{x}_1-\bar{x}_0|^2\big].\end{aligned}$$

5. Alexandrov type estimates

The appeared convexity (see Theorem 4.8) tells us that under the **MTW** condition c-convex functions can be transformed to level-set convex functions in appropriate exponential coordinates. For this observation to be useful in applications, we extend the Alexandrov estimates well-known for convex functions, to level-set convex c-convex functions.

First consider a basic tool in convex analysis, the so-called Fritz John's ellipsoid lemma:

THEOREM 5.1 (John's ellipsoid lemma [**J**]). *Let $Z \subset \mathbb{R}^n$ be an open bounded convex set. Then, there exists an ellipsoid E, such that*

$$E \subset Z \subset nE \tag{5.1}$$

where nE is the dilation of E with respect to its centre. Equivalently, there exists an invertible affine map $L : \mathbb{R}^n \to \mathbb{R}^n$ such that

$$B_1(0) \subset L^{-1}(Z) \subset B_n(0),$$

for the balls $B_1(0)$, $B_n(0)$ centred at the origin with radius 1 and n, respectively.

The following theorem in [**FiKM1**] extends the classical Alexandrov estimates for convex functions to the c-convex functions under the **MTW** condition.

THEOREM 5.2 (Alexandrov upper and lower bound [**FiKM1**]). *Use the notation in the previous section and Theorem 4.8. Assume that*

- *the functioin $-\tilde{c}$ is level set convex (which holds under* **MTW***);*
- *$\tilde{\phi}$ is $\tilde{c}$-convex (thus it is also level-set convex since $-\tilde{c}$ is level set convex);*
- *$\frac{1}{\lambda} \geq |\partial^{\tilde{c}} \tilde{\phi}| \geq \lambda > 0$;*
- *the set $Z := \{z \in \exp_{\bar{x}_0}^{-1} \Omega \mid \tilde{\phi} < 0\} \subset\subset \exp_{\bar{x}_0}^{-1} \Omega$.*

Then, we have the following:

(1) *(Alexandrov lower bound) There exists a constant $C(n, \lambda) > 0$ such that*

$$|Z|^2 \leq C(n, \lambda) \Big(\sup_Z |\tilde{\phi}| \Big)^n. \tag{5.2}$$

(2) *(Alexandrov upper bound) If $\operatorname{diam} Z \ll 1$ and Z is sufficiently far from $\partial \exp_{\bar{x}_0}^{-1} \Omega$, then*

$$|\tilde{\phi}(q_t)|^n \lesssim (1-t)^{\frac{1}{2^{n-1}}} |Z|^2 \tag{5.3}$$

where $q_t \in t\partial Z$, for $0 < t < 1$. Here, $t\partial Z$ is the dilation of ∂Z by the factor t with respect to the centre of the ellipsoid for the convex set Z as in the John's lemma (Theorem 5.1).

An important idea behind Alexadrov type estimates is that convex functions behaves like quadratic functions, and for quadratic functions such estimates are straightforward:

EXAMPLE 5.3. For $x = (x_1, x_2) \in \mathbb{R}^2$, let $\tilde{\phi}(x) = a_1 x^2 + a_2 x_2^2 - b$ for $0 < a, b \in \mathbb{R}$. Assume $|\det D^2 \tilde{\phi}| \sim 1$, i.e. $a_1 a_2 \sim 1$.

Now, for $Z = \{x \mid \tilde{\phi}(x) \leq 0\}$,

$$\begin{aligned} |Z|^2 &\sim \frac{b^2}{a_1 a_2} \sim b^2 \qquad (\text{since } |\det D^2 \tilde{\phi}| \sim 1) \\ &= \sup_Z |\tilde{\phi}|^2. \end{aligned}$$

This gives the Alexandrov lower bound.

On the other hand we see for $q_t \in t\partial Z$,

$$|\tilde{\phi}(q_t)|^2 = |t^2 b - b|^2 = (1-t)^2 (1+t)^2 |b|^2 \sim (1-t)^2 |Z|^2$$

which gives the Alexandrov upper bound.

The point is that the estimates (5.2) (5.3) hold regardless the shape of the convex set Z, i.e. it can be very thin. This latter case is unavoidable, since if $\tilde{\phi}$ is merely $C^{1,\alpha}$ (which is the optimal regularity for potentials of optimal maps with merely measurable source and target densities), there is no control on how thin the section can be. Of course, for C^2 (uniformly) convex functions, there is a uniform control on the shape of the sections.

5.1. Alexandrov lower bound. Let us first discuss the proof of the Alexandrov lower bound (5.2). Instead of giving the full proof of (5.2), we show the well-known special case $\tilde{c}(x,y) = -x \cdot y$ and for convex $\tilde{\phi}$ so that $\partial^{\tilde{c}}\tilde{\phi} = \partial\tilde{\phi}$. We will then discuss the more general case.

5.1.1. *Proof of the Alexandrov lower bound* (5.2) *for the case* $\tilde{c}(x,y) = -x \cdot y$ *and* $\tilde{\phi}$ *is convex.* This is a standard proof and one can find it elsewhere (for example, in the book of Gutierrez [**Gt**]). We give the proof here for user's convenience and also for discussion of the more general case.

We first renormalize the set Z. Find an affine map $L : \mathbb{R}^n \to \mathbb{R}^n$ such that the set $Z^* := L^{-1}Z$ is comparable to $B_1(0)$ i.e. $B_1(0) \subset Z^* \subset B_n(0)$ and $|Z| \sim |\det L|$. Let

$$\tilde{\phi}^*(z^*) := \frac{1}{(\det L)^{2/n}}\tilde{\phi}(Lz^*).$$

Then, for all Borel set B,

$$\left|\partial\tilde{\phi}^*(L^{-1}(B)\right| = \frac{1}{|\det L|}|\partial\phi(B)|.$$

We also see that

$$|\partial\tilde{\phi}| \sim 1 \Longleftrightarrow |\partial\tilde{\phi}^*| \sim 1.$$

Pick any $p \in \partial\tilde{\phi}^*(\frac{1}{2}Z^*)$, where the set $\frac{1}{2}Z^*$ is the dilation of the set Z^* by factor $\frac{1}{2}$ with respect to the origin 0. Then, we see by convexity of $\tilde{\phi}^*$ and the fact that Z^* is in shape comparable to the unit ball,

$$|p| \lesssim h^*$$

where $h^* := \sup_{Z^*} |\tilde{\phi}^*|$. This shows that

$$\partial\tilde{\phi}^*(\frac{1}{2}Z^*) \subset B_{Ch^*}(0) \qquad \text{for some constant } C > 0,$$

and thus,

$$|\partial\tilde{\phi}^*(\frac{1}{2}Z^*)| \lesssim |h^*|^n.$$

Now the lefthand side is the same as

$$\begin{aligned}
&\frac{1}{|\det L|}\left|\partial\tilde{\phi}(\frac{1}{2}Z)\right| \\
&\gtrsim \frac{1}{|\det L|}|Z| \qquad (\text{since } |\partial\tilde{\phi}| \gtrsim \lambda > 0) \\
&\sim \frac{1}{|Z|}|Z| = 1. \qquad (\text{since } |Z| \sim |\det L|).
\end{aligned}$$

On the other hand, the righthand side is

$$\frac{1}{(\det L)^2}(\sup_Z |\tilde{\phi}|)^n \sim \frac{(\sup_Z |\tilde{\phi}|)^n}{|Z|^2}$$

Comparing these, we see

$$|Z|^2 \leq (\sup_Z |\tilde{\phi}|)^n$$

as desired.

5.1.2. *Discussion for more general case.* Let us discuss the more general case.

Under **NNCC** condition, the corresponding inequality reduces to the case $\tilde{c}(x,y) = -x \cdot y$, mainly because of the two reasons below:

- $\tilde{\phi}$ is convex (due to Theorem 4.8 (2)) ;
- $|\partial^c \tilde{\phi}| \lesssim |\partial \tilde{\phi}|$.

The second assertion holds since (ignoring differentiability),

$$\begin{aligned} |\det \mathrm{Jac}\partial^{\tilde{c}} \tilde{\phi}| &= |\det(D_x D_{\bar{x}} \tilde{c})|^{-1} \det(D^2_{xx}\tilde{\phi} + D^2_{xx}\tilde{c}) \\ &\lesssim |\det D^2_{xx}\phi| = |\det \mathrm{Jac}\partial\phi| \end{aligned}$$

where the inequality holds because $\det(D_x D_{\bar{x}}\tilde{c})$ is bounded and **NNCC** condition implies $D^2_{xx}\tilde{c} \leq 0$ thus $D^2_{xx}\tilde{\phi} + D^2_{xx}\tilde{c} \leq D^2_{xx}\tilde{\phi}$ as matrices.

On the other hand, the **MTW** case is much more difficult, because both of above key properties under **NNCC** do not hold anymore.

One may argue that a renormalization method as in the proof for the classical affine cost case $\tilde{c}(x,y) = -x \cdot y$, would work, especially letting the relevant set smaller and smaller so to make the cost function more close to the affine cost (when x and y are close any smooth cost looks like an affine cost asymptotically). But, there is a serious difficulty in this argument. Namely, for more general cost function $\tilde{c}$, the term $D_{\bar{x}} D^2_{xx}\tilde{c}$ (which measures how much the local behaviour of the cost function differs from that of the affine cost) may blow-up under the renormalization, if the set Z before normalization is very thin.

Thus, it does not seem reasonable to use the renormalization method to treat general **MTW** case. But, by the appeared convexity (Theorem 4.8 (1)) we can still use John's ellipsoid lemma (Theorem 5.1) to treat a convex body geometrically as an ellipsoid. The actual proof is not so simple, and we refer the reader to the paper [**FiKM1**].

5.2. Alexandrov upper bound. We now discuss the proof of the Alexandrov upper bound (5.3). As an auxiliary result, the following lemma is proved in [**FiKM1**, Lemma 6.10], whose proof manipulates the fact that the cost $\tilde{c}$ is close to the linear cost in small scale.

Lemma 5.4 (see [**FiKM1**]). *Use the same assumption and notation as in Theorem 5.2. Let Π^+, Π^- be two parallel hyperplanes contained in $T_{\bar{x}_0}M \setminus Z$ and touching ∂Z from two opposite sides. If $\operatorname{diam} Z \ll 1$ and Z is sufficiently far from $\partial \exp^{-1}_{\bar{x}_0} \Omega$, then*

$$|\tilde{\phi}(\tilde{q})|^n \lesssim \frac{\min\{\operatorname{dist}(\tilde{q},\Pi^+), \operatorname{dist}(\tilde{q},\Pi^-)\}}{\ell_{\Pi^+}} |\partial^{\tilde{c}}\tilde{\phi}|(Z)\mathscr{L}^n(Z),$$

where ℓ_{Π^+} denotes the maximal length among all the segments obtained by intersecting Z with a line orthogonal to Π^+.

This lemma is enough to show the injectivity and continuity of T [**FiKM1**, Seciton 7 & 8]. But, for this to be applicable to the Hölder continuity of T [**FiKM1**][Seciton 9] (this section uses the method of Gutierrez and Huang [**GH**] and Forzani and Maldonado [**FoM**]), it is important to know that for any Z as in the Lemma 5.4, one can choose parallel hyperplanes Π^+, Π^- in such way that the ratio $\frac{\min\{\operatorname{dist}(\tilde{q},\Pi^+), \operatorname{dist}(\tilde{q},\Pi^-)\}}{\ell_{\Pi^+}} \to 0$ gets close to zero as q is close to the boundary ∂Z, in a 'uniform' way independent of the particular shape of Z. In other words, we need an estimate which quantifies the dimensional dependence of the claim that corresponding to any (interior) point near the boundary of a convex set, is a supporting hyperplane much closer than the thickness of the set in the orthogonal direction. Such estimate is obtained in the following new result in convex geometry, whose proof is elementary but quite nontrivial:

THEOREM 5.5 (Convex bodies and supporting hyperplanes [**FiKM4**]). *Let $\tilde{Q} \subset \mathbb{R}^n$ be a convex body (with nonempty interior) such that* (5.1) *holds for some ellipsoid E centered at the origin. Fix $0 \leq s \leq \frac{1}{2n}$. To each $y \in (1-s)\partial\tilde{Q}$ corresponds at least one line ℓ through the origin and hyperplane Π supporting $\tilde{Q}$ such that: Π is orthogonal to ℓ and*

$$\mathrm{dist}(y, \Pi) \leq c(n) s^{1/2^{n-1}} \,\mathrm{diam}(\ell \cap \tilde{Q}). \tag{5.4}$$

Here, $c(n)$ is a constant depending only on n.

We refer the reader to the paper [**FiKM4**] for more discussions about this estimate and its proof.

Lemma 5.4 and Theorem 5.5 implies (5.3).

6. How to prove injectivity of optimal transport maps under MTW

To illustrate how the previous results are used, we explain as an example, how to show injectivity of the optimal map T under the **MTW** condition and conditions on the source and the target domains. (A similar method can be used to show the continuity of T.) Here, we use the same assumptions as given in Theorem 4.3.

DEFINITION 6.1 (Contact set). For each $\bar{x} \in \bar{\Omega} = \mathrm{int}\bar{\Omega}$, the contact set $S(\bar{x})$ for $\bar{x}$ is the set

$$S(\bar{x}) = \{x \in \Omega \mid \partial^c \phi(x) = \bar{x}\}.$$

The injectivity of T is equivalent to that $S(\bar{x})$ is singleton for all $\bar{x} \in \bar{\Omega}$.

Let us briefly explain the idea how to show $S(\bar{x})$ is singleton under some technical conditions. For simplicity of exposition, let us consider the case $c(x, \bar{x}) = -x \cdot \bar{x}$, which is due to Cafferelli: [**Ca1**,**Ca3**,**Ca5**]. The more general case as in Theorem 4.3 is a bit more complicated: see [**FiKM1**]. We emphasize here that the reason why we can carry out Caffarelli's idea is because we now have

- **MTW**, in particular, Appeared convexity (Theorem 4.8),
- Interior-not-to-boundary (Theorem 4.6),
- Alexandrov type estimates (Theorem 5.2).

Now, let us explain Caffarelli's localization argument in [**Ca1**]: the expository article [**Ca5**] is very useful. We include it here to demonstrate how the previous results in Sections 4 and 5 are used. In the case $c(x, \bar{x}) = -x \cdot \bar{x}$, the c-potential function ϕ is convex on $\mathbb{R}^n$, the c-convex domains Ω, $\bar{\Omega}$ are convex in $\mathbb{R}^n$, and the strong c-convexity of $\bar{\Omega}$ implies strong convexity. We assume the Monge-Ampère equation $|\partial\phi| \sim 1$. (Of course, for more general cost functions, this is replaced by $|\partial^c\phi| \sim 1$.)

Now, let us show the injectivity of T, i.e. $\#S(\bar{x}_0) = 1$ for each $\bar{x} \in \bar{\Omega}$. Suppose $\#S(\bar{x}_0) > 1$ for some $\bar{x}_0 \in \bar{\Omega}$. The goal is to contradict this.

Step 1: By Theorem 4.6 (Interior-not-to-boundary), we see

$$S(\bar{x}_0) \subset \mathrm{int}\Omega, \qquad \text{since } \partial^c\phi(\partial\Omega) \cap \mathrm{int}\bar{\Omega} = \emptyset.$$

REMARK 6.2. In fact, for $c(x,\bar{x}) = -x \cdot \bar{x}$, this inteor-not-to-boundary result is not necessary, since in this case one can extend the convex function ϕ to the whole $\mathbb{R}^n$. Here, the case to exclude is when the contact set contains an infinite line. If this occurs, then as pointed out in **[Ca3]** the convex function has zero Monge-Ampère measure, i.e. $|\partial\phi(B)| = 0$ for any Borel set B, thus this case is excluded by our assumption $|\partial\phi| \sim 1$.

Step 2: Now, by appeared convexity (Theorem 4.8), if we let $S = \exp_{\bar{x}_0}^{-1}(S(\bar{x}_0)) \subset T_{\bar{x}_0}\bar{\Omega}$, then

$$S \text{ is convex and bounded.}$$

Of course, in the current special case, this convexity immediately follows from the convexity of ϕ.
Step 3: We now find an exposed point, say x_e of S. Exposed point is by definition, such a point where a hyperplane touches the convex set only at that point.
Step 4: For a family $\bar{x}_\theta \in \bar{\Omega}$, $0 \le \theta \le 1$, (thus $\bar{x}_\theta = \bar{x}_0$ for $\theta = 0$) and a point x_0, define.

$$m_\theta(\cdot) := -c(\cdot, \bar{x}_\theta) + c(x_0, \bar{x}_\theta)$$

Let $\phi_\theta := \phi - m_\theta$ and let $Z_\theta := \{z \mid \phi_\theta < 0\}$. Since x_e is an exposed point of $S = \{z \mid \partial^c\phi(z) = \bar{x}_0\}$, we can choose the point $x_0 \in S$, nearby x_e, and the family $\bar{x}_\theta \in \bar{\Omega}$, so that

$$Z_\theta \to S \text{ and } \mathrm{dist}(x_e, \partial Z_\theta) \to 0, \text{ as } \theta \to 0$$

Step 5: One can also show that for $\theta \ll 1$,

$$\inf_{Z_\theta} |\phi_\theta| \sim \phi_\theta(x_e).$$

Step 6: From the assumption $|\partial\phi| \sim 1$, we can apply the Alexandrov estimates (see Theorem 5.2).

$$|Z_\theta|^2 \lesssim (\inf_{Z_\theta} |\phi_\theta|)^n$$

$$|\phi_\theta(x_e)|^n \lesssim \eta(\theta)|Z_\theta|^2 \qquad \text{(some function } \eta(\theta) \text{ such that } \lim_{\theta\to 0+} \eta(\theta) = 0)$$

Step 7: Now apply $\theta \to 0$ in Step 6 and Step 5, and we get a contradiction $1 \lesssim 0$. This shows that $\#S(\bar{x}_0) = 1$ for each $\bar{x} \in \bar{\Omega}$, thus, the injectivity of T.

7. Regularity of optimal maps on global domains

We now give a few remarks on regularity of optimal transportation on global domains. Here, by a global domain, we mean a closed manifold M, $\Omega = \bar{\Omega} = M$, with $\log\rho, \log\bar{\rho} \in L^\infty(M)$. Loeper gave the first such regularity result ($T \in C^\alpha/C^\infty$ for $\log\rho, \log\bar{\rho} \in L^\infty(M)/C^\infty(M)$) by showing it on the round sphere S^n **[Lo2]**. It was then followed by work of many researchers including the author.

First, as a necessary condition for regularity of optimal maps, Loeper's maximum principle **LMP** needs to be verified. It was first shown by Loeper **[Lo1]** on domains in $\mathbb{R}^n$ (for cost functions satisfying the **MTW** condition) using the regularity results of Trudinger and Wang **[TW1]**. To treat more global manifold domains (e.g. products of round spheres), an elementary method for deriving **LMP** from **MTW** and appropriate geometric conditions, was introduced by McCann and the author **[KmM1]** (see the work of Trudinger and Wang **[TW2]** for other approach, obtained independently from **[KmM1]**), which later was strengthened by Loeper, Figalli, Rifford and Villani [**LoV**, **FiR**, **FiRV1**, **V2**]. Up to now, the examples of domains with **LMP** includes Riemannian distance squared costs on the products $S^{n_1}(r_1)\times\cdots\times S^{n_j}(r_j)\times \mathbf{CP}^{l_1}\times\cdots\times\mathbf{CP}^{l_k}\times\mathbb{R}^m$ of Euclidean spaces, round spheres, complex projective spaces, and their appropriate quotient spaces (this product example trivially includes flat tori) [**KmM3**], as well as perturbations of the round sphere and its discrete quotients [**LoV**,**FiRV1**]. See e.g. [**FiRV3**] for more detailed list.

As we mentioned before, for manifold domains, a problem arises due to nonsmoothness of the cost function dist^2 along the cut-locus. So, a key step is to show that the optimal map T stays away from the cut locus, so that one can assume that the cost function is smooth. Namely,

QUESTION 7.1 (Stay Away from Cut-locus). Fix a Riemannian manifold M. Suppose $\log\rho, \log\bar\rho \in L^\infty(M)$. Is

$$T(x)\cap \mathrm{cut}(x) = \emptyset \qquad \text{for each } x\in M \text{ and its cut-locus } \mathrm{cut}(x)? \tag{7.1}$$

The property (7.1) is a necessary condition for higher regularity (e.g. C^1, C^∞) of optimal maps. Such *stay-away* result was obtained affirmatively for the case of the round sphere S^n by Delanoë and Loeper [**DL**], its perturbation by Delanoë and Ge [**DG1**] (but with further restriction on $\rho, \bar\rho$ depending on the perturbation), and the product of round spheres $S^{n_1}(r_1)\times\cdots\times S^{n_k}(r_k)$ (of arbitrary dimensions and size) by Figalli, McCann and the author [**FiKM2**]. The latter result [**FiKM2**] gives the first regularity result for optimal transport ($T\in C^\alpha/C^\infty$ for $\log\rho, \log\bar\rho\in L^\infty(M)/C^\infty(M)$) on global domains that are not positively curved and not totally flat. This case differs significantly from the known regularity results on positivley curved domains, e.g. $\mathbf{CP}^n$ [**KmM3**], RP^n and its perturbation [**LoV**], perturbations of the round sphere and their discrete quotients [**DG1**] [**FiR**,**FiRV1**], where Loeper's quantified maximum principle **LQMP** (or a strong a priopri estimates of Ma, Trudinger and Wang [**MTW**]) can be applied. Note that on the flat tori, Cordero-Erausquiun [**Co**] showed regularity of the optimal map T by lifting the situation to the universal covering space $\mathbb{R}^n$, where Caffarelli's regularity theory [**Ca1**, **Ca2**, **Ca3**, **Ca4**, **Ca5**] applies.

This *stay-away-from-cut-locus* problem is not well understood. For example, we do not yet have such result (for $\rho, \bar\rho$ independent on the perturbation) for the perturbation of the round sphere. Note that even without such stay-away-from-singularity it is still possible to show that T is continuous on the perturbation of the sphere as in [**FiR**] [**FiRV3**]. However, the stay-away result will, if it holds, show Hölder continuity and higher regularity.[2]

8. Additional remarks on the literature

We close these lectures with a few remarks on the literature about some directions involving the Ma-Trudinger-Wang curvature which are not mentioned in the above discussion.

[2]Notice that Delanoë and Ge [**DG1**] showed smoothness of optimal maps on perturbations of the round sphere, however, their perturbation of the domain is restricted by the source and target measures $\rho, \bar\rho$, or in other words, ρ and $\bar\rho$ has to be chosen appropriately depending on the perturbation. Higher regularity of optimal maps on a fixed small perturbation of the sphere, but with arbitrary smooth $\rho, \bar\rho$, is still an open problem.

8.1. Parabolic optimal transport. The parabolic problem of optimal transportation theory is considered by Street, Warren and the author [**KmSW**] under $\mathbf{MTW}_+$ condition on manifold domains (e.g $S^n, \mathbf{CP}^n$) and Kitagawa [**Kt**] under **MTW** condition on domains in $\mathbb{R}^n$ with appropriate geometric assumptions. They considered the parabolic Monge-Ampère type equation

$$\frac{\partial u}{\partial t} = \ln\det(D^2_{xx}u + D^2_{xx}c(x,T(x))) - \ln\rho(x) + \ln\bar\rho(T(x)) - \ln\det|D^2_{x\bar x}c(x,T(x))|,$$

and have obtained long-time existence results and convergence (exponential convergence under $\mathbf{MTW}_+$ [**KmSW**]) to the solution to optimal transportation problem. This parabolic approach, in particular, gives a natural algorithm for finding optimal maps.

8.2. Multivalued optimal maps. In general, optimal transportation between two measures gives not a single-valued map but a multivalued map. The analysis of such multivalued maps is not well understood, though there are results by Gangbo and McCann [**GaM**] and McCann and Sosio [**McS**] who considered the multivalued (bivalent) optimal maps for the cost given by the Euclidian distance sqaured restricted to the round sphere $S^{n-1} \subset \mathbb{R}^n$. McCann and Sosio used techniques involving Loeper's quantitative maximum principle **LQMP** (see Section 3.3) to study Hölder continuity of such bivalent maps. Note that however, even in this case, if the source and target measure $\rho, \bar\rho$ are sufficiently close to constant densities, Kitagawa and Warren [**KtWr**] showed that the optimal transportation is given by a single-valued map and smooth.

8.3. Regularity/partial regularity without MTW or convexity assumptions on the domain. As we have discussed in these lecture notes, we now have counterexamples to continuity of optimal maps when either **MTW** [**Lo1**] (see Theorem 2.6) or appropriate convexity assumptions [**Ca3**, **MTW**] are not satisfied. Notice that the known necessary (sometimes sufficient) conditions do not involve other key players in the transportation problem, namely the source and target distributions ρ and $\bar\rho$. In particular, there still is a possibility to have regularity of the optimal map T without **MTW** condition by imposing further restrictions on ρ, $\bar\rho$. For example, it is obvious that if $\rho = \bar\rho$ or if $\bar\rho$ is a Dirac-delta measure, then, T is the identity map or a trivial map, respectively, thus C^∞ on the support of ρ. In this spirit, Warren [**Wr**] obtained regularity of optimal maps between *narrow enough* Gaussian measures, regardless of **MTW** condition. It is an important wide open question to find a necessary and sufficient condition for regularity of optimal transport, which contains all the relevant data ρ,$\bar\rho$, the cost function c, and the geometry of source and target domains. A guess is that one may try to find some curvature condition for the pseudometric in [**KmMW**].

A related outstanding open problem is to get partial regularity of optimal maps without **MTW** or appropriate convexity assumptions on the domains. It is well-known that the singular set (the set of discontinuity) of the optimal map (for *cost* $= \mathrm{dist}^2/2$) has Hausdorff dimension less than or equal to $n-1$ in n-dimensional domain. So, the point is to get a *sharper* description of the singular set. No such result has been known regarding the violation of **MTW** condition, however, for the Euclidean distance squared (cross $\equiv 0$, thus **MTW**), a partial regularity is known by Figalli and the author [**FiK**] when the convexity assumption on the domains is violated, extending the two dimensional results of Yu [**Y**] and

Figalli [**Fi1**]. In [**FiK**], the singular set is shown to be contained in a measure zero *closed* set, thus for example, excluding the case it to be dense. However, this result lacks of the more precise description on the structure of the singular set as given in two dimensions in [**Y**, **Fi1**].

Acknolwegement

The author thanks the SMS school and the organizers, Galia Dafni and Alina Stancu, for their hospitality, Robert McCann for introducing him to the world of optimal transportation, and also Tongseok Lim (the author's PhD student) and the referee for useful comments about the manuscript. The present lecture notes and the author's work presented here, were not possible without the sincere support, encouragement and love of his wife, Dong Soon Shim, and the two daughters, Joyce and Ashley.

References

[A] A. D. Alexandrov, *Smoothness of the convex surface of bounded Gaussian curvature.* C. R. (Doklady) Acad. Sci. URSS (N.S.), 36 (1942), 195–199.

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

[Ca1] L.A. Caffarelli, *A localization property of viscosity solutions to the Monge-Ampère equation and their strict convexity.* Ann. of Math., 131 (1990), 129-134.

[Ca2] L.A. Caffarelli, *Some regularity properties of solutions of Monge Ampère equation.* Comm. Pure Appl. Math., 44 (1991), no. 8-9, 965-969.

[Ca3] L.A. Caffarelli, *The regularity of mapping with a convex potential.* J. Amer. Math. Soc., 5 (1992), 99-104.

[Ca4] L. A. Caffarelli, *Boundary regularity of maps with convex potentials II.* Ann. of Math., 144 (1996), 453-496.

[Ca5] L. A. Caffarelli, *A priori estimates and the geometry of the Monge Ampère equation.* Nonlinear partial differential equations in differential geometry (Park City, UT, 1992), 5–63, IAS/Park City Math. Ser., 2, Amer. Math. Soc., Providence, RI 1996

[CaGH] L. A. Caffarelli, C. Gutierrez and Q. Huang, *On the Regularity of Reflector Antennas.* Annals of Math. 167 (2008), 299 – 323.

[ChEb] J. Cheeger and D. G. Ebin, Comparision theorems in Riemannian geometry. North-Holland Mathematical Library, vol. 9, North-Holland, Amsterdam; American Elsevier, New York, 1975.

[Co] D. Cordero-Erausquin, *Sur le transport de mesures périodiques.* C. R. Acad. Sci. Paris Sèr. I Math., 329 (1999), 199-202.

[CoMS] D. Cordero-Erausquin, R. J. McCann and M. Schmuckenschlaeger *A Riemannian interpolation inequality a la Borell, Brascamp and Lieb.* Invent. Math. 146 (2001) 219-257

[D] P. Delanoë, *Classical solvability in demension two of the second boundary value problem associated with the Monge-Ampère operator.* Ann. Inst. Henri Poincaré-Anal. Non Lin., 8 (1991), 443-457.

[DG1] P. Delanoë and Y. Ge, *Regularity of optimal transportation maps on compact, locally nearly spherical, manifolds.* J. Reine Angew. Math. 646, 65-115 (2010).

[DG2] P. Delanoë and Y. Ge, *Locally nearly spherical surfaces are almost-positively curved.* To appear in Methods and Applications of Analysis.

[DL] P. Delanoë and G. Loeper, *Gradient estimates for potentials of invertible gradient mappings on the sphere.* Calc. Var. Partial Differential Equations, 26 (2006), no. 3, 297-311.

[Fi1] A. Figalli, *Regularity properties of optimal maps between nonconvex domains in the plane.* Comm. Partial Differential Equations, 35 (2010), no. 3, 465 – 479.

[Fi2] A. Figalli, *Regularity of optimal transport maps (after Ma-Trudinger-Wang and Loeper).* Séminaire Bourbaki. Vol. 2008/2009. Exposés 997-1011. Astérisque, 332 (2010), Exp. No. 1009.

[FiK] A. Figalli and Y-H. Kim, *Partial regularity of Brenier solutions of the Monge-Ampère equation.* Discrete Contin. Dyn. Syst. Series A, 28 (2010), no. 2, 559–565.

[FiKM1] A. Figalli, Y.-H. Kim and R. J. McCann, *Hölder continuity and injectivity of optimal maps.* Preprint at arXiv:1107.1014.

[FiKM2] A. Figalli, Y.-H. Kim and R. J. McCann, *Regularity of optimal transport maps on multiple products of spheres.* Preprint at arXiv:1006.1957. To appear in J. Eur. Math. Soc. (JEMS).

[FiKM3] A. Figalli, Y.-H. Kim and R. J. McCann, *When is multidimensional screening a convex program?* J. Econom. Theory, 146 (2011), 454-478.

[FiKM4] A. Figalli, Y.-H. Kim and R. J. McCann, *On supporting hyperplanes to convex bodies.* Preprint at arXiv:1107.1016.

[FiL] A. Figalli and G. Loeper, *C^1-regularity of solutions of the Monge-Ampère equation for optimal transport in dimension two.* Calc. Var. Partial Differential Equations, 35 (2009), no. 4, 537-550.

[FiR] A. Figalli and L. Rifford, *Continuity of optimal transport maps on small deformations of $\mathbb{S}^2$.* Comm. Pure Appl. Math., 62 (2009), no. 12, 1670-1706.

[FiRV1] A. Figalli, L. Rifford, and C. Villani, *Nearly round spheres look convex.* To appear in Amer. J. Math.

[FiRV2] A. Figalli, L. Rifford, and C. Villani, *On the Ma-Trudinger-Wang curvature on surfaces.* Calc. Var. Partial Differential Equations, 39 (2010), no. 3-4, 307 – 332.

[FiRV3] A. Figalli, L. Rifford, and C. Villani, *Necessary and sufficient conditions for continuity of optimal transport maps on Riemannian manifolds.* Tohoku Math. J., to appear.

[FiV] A. Figalli and C. Villani, *Optimal Transport and Curvature.* Nonlinear PDE's and Applications. Lectures from the C.I.M.E. Summer School held in Cetraro, June 23-28, 2008. Lecture Notes in Mathematics, to appear.

[FoM] L. Forzani and D. Maldonado, *Properties of the solutions to the Monge-Ampère equation.* Nonlinear Analysis 57 (2004) 815–829.

[GaM] W. Gangbo and R.J. McCann, *Shape recognition via Wasserstein distance.* Quart. Appl. Math. 58 (2000), 705-737.

[GlM] N. Gullen and R. J. McCann, *Five lectures on optimal transportation: geometry, regularity and applications.* Preprint.

[Gt] C. Gutiérrez, *The Monge-Ampère Equation.* Birkhauser, 2001.

[GH] C. Gutiérrez and Q. Huang, *Geometric properties of the sections of solutions to the Monge-Ampère equation.* Trans. Amer. Math. Soc. 352 (2000), no. 9, 4381–4396.

[J] F. John, *Extremum problems with inequalities as subsidiary conditions.* Studies and Essays Presented to R. Courant on his 60th Birthday, January 8, 1948, 187-204. Interscience Publishers, Inc., New York, N. Y., 1948.

[KaW] A. Karakhanyan and X.-J. Wang, *On the reflector shape design.* J. Differential Geom. 84, no. 3 (2010), 561–610.

[Km] Y.-H. Kim, *Counterexamples to continuity of optimal transportation on positively curved Riemannian manifolds.* Int. Math. Res. Not., Art. ID rnn120:1-15, 2008.

[KmM1] Y.-H. Kim and R. J. McCann, *Continuity, curvature, and the general covariance of optimal transportation.* J. Eur. Math. Soc. (JEMS), 12:1009-1040, 2010.

[KmM2] Y.-H. Kim and R.J. McCann. Appendices to original version of Continuity, curvature, and the general covariance of optimal transportation. Preprint at arXiv:math/0712.3077v1.

[KmM3] Y.-H. Kim and R. J. McCann, *Towards the smoothness of optimal maps on Riemannian submersions and Riemannian products (of round spheres in particular).* Preprint at arXiv:math/0806.0351v1. To appear in J. Reine Angew. Math.

[KmMW] Y.-H. Kim, R. J. McCann and M. Warren, *Pseudo-Riemannian geometry calibrates optimal transportation.* Math. Res. Lett. 17 (2010) 1183-1197

[KmSW] Y.-H. Kim, J. Streets and M. Warren, *Parabolic optimal transport equations on manifolds.* Preprint at arXiv:1008.3824v1. To appear in International Mathematics Research Notices.

[Kt] J. Kitagawa, *A parabolic flow toward solutions of the optimal transportation problem on domains with boundary.* To appear in J. Reine und Angew. Math.

[KtWr] J. Kitagawa and M. Warren, *Regularity of optimal transport with Euclidean distance squared cost on the embedded sphere.* Preprint at arXiv:1101.5146.

[Le] P. W. Y. Lee, *New computable necessary conditions for the regularity theory of optimal transportation.* SIAM J. Math. Anal., 42 (2010), no.6, 3054-3075.

[LeM] P. W. Y. Lee and R. J. McCann, *The Ma-Trudinger-Wang curvature for natural mechanical actions.* Calc. Var. Partial Differential Equations 41 (2010), no.1-2, 285-299.

[LeLi] P.W. Y. Lee and J. Li, *New examples on spaces of negative sectional curvature satisfying Ma-Trudinger-Wang conditions.*arXiv: 0911.3978, (2009), 17pp. To appear in SIAM J. Math. Anal.

[Li] J. Liu, *Hölder regularity of optimal mappings in optimal transportation.* Calc. Var. Partial Differential Equations 34 (2009), no. 4, 435–451.

[LiT] J. Liu and N. Trudinger, *On Pogorelov estimates for Monge-Ampère type equations.* Discrete And Continuous Dynamical Systems (2010) 28 no. 3, 1121-1135.

[LTW] J. Liu, N. Trudinger and X.-J. Wang, *Interior $C^{2,\alpha}$-regularity for potential functions in optimal transportation*, Comm. Partial Differential Equations 35 (2010), no. 1, 165-184.

[Lo1] G. Loeper, *On the regularity of solutions of optimal transportation problems.* Acta Math. 202, no. 2 (2009), 241 – 283.

[Lo2] G. Loeper, *Regularity of optimal maps on the sphere: The quadratic cost and the reflector antenna.* Arch. Ration. Mech. Anal. 199, no. 1 (2011), 269 – 289,

[LoV] G. Loeper and C. Villani, *Regularity of optimal transport in curved geometry: the nonfocal case.* Duke Math. J. 151, no. 3 (2010), 431– 485.

[Mc] R. J. McCann, *Polar factorization of maps on Riemannian manifolds.* Geom. Funct. Anal., 11 (2001), 589-608.

[McS] R.J. McCann and M. Sosio, *Hoelder continuity for optimal multivalued mappings.* SIAM J. Math. Anal. 43 (2011) 1855-1871

[MTW] X.-N. Ma, N. Trudinger and X.-J. Wang, *Regularity of potential functions of the optimal transport problem.* Arch. Ration. Mech. Anal., 177(2): 151–183, 2005.

[S] T. Sei, *A Jacobian inequality for gradient maps on the sphere and its application to directional statistics.* Preprint at arXiv:0906.0874.

[TW1] N. Trudinger and X.-J. Wang, *On the second boundary value problem for Monge-Ampère type equations and optimal transportation.* Ann. Sc. Norm. Super. Pisa Cl. Sci. (5) 8 (2009), no. 1, 143-174.

[TW2] N. Trudinger and X.-J. Wang, *On strict convexity and continuous differentiablity of potential functions in optimal transportation.* Arch. Ration. Mech. Anal. 192 (2009), no. 3, 403-418.

[U] J. Urbas, *On the second boundary value problem for equations of Monge-Ampère type.* J. Reine Angew. Math., 487 (1997), 115-124.

[V1] C. Villani, *Topics in Optimal Transportation*, Graduate Studies in Mathematics, vol. 58. American Mathematical Society, Providence, RI, 2003.

[V2] C. Villani, *Optimal Transport, Old and New*, Grundlehren des mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], Vol. 338, Springer-Verlag, Berlin-New York, 2009.

[V3] C. Villani, *Stability of a 4th-order curvature condition arising in optimal transport theory.* J. of Functional Analysis 255 (2008) 2683 – 2708.

[Wn1] X.-J. Wang. On the design of a reflector antenna. *Inverse Problems* 12 (1996) 351–375.

[Wn2] X.-J. Wang, *On the design of a reflector antenna II.* Calc. Var. PDE, (2004), 329-341.

[Wr] M. Warren, *Regularity for a log-concave to log-concave mass transfer problem with near Euclidean cost.* Comm. Anal. Geom. 19 (2011), no. 1, 191–208.

[Y] Y. Yu, *Singular set of a convex potential in two dimensions.* Comm. Partial Differential Equations 32 (2007), no. 10–12, 1883–1894.

Department of Mathematics, Mathematics Dept., University of British Columbia, Vancouver BC Canada V6T 1Z2

Email address: `yhkim@math.ubc.ca`