ON THE COST-SUBDIFFERENTIALS OF COST-CONVEX FUNCTIONS

YOUNG-HEON KIM AND ROBERT J. MCCANN

Abstract. We are interested in the cost-convex potentials in optimal mass transport theory, and we show by direct and geometric arguments the equivalence between cost-subdifferentials and ordinary subdifferentials of cost-convex functions, under the assumptions $A_0$, $A_1$, $A_2$, and $A_3W$ on cost functions introduced by Ma, Trudinger, and Wang. The connectivity of contact sets of optimal transport maps follows as a direct corollary. Our approach is quite different from the previous result of Loeper which he obtained as the first step toward his Hölder regularity theory of potential functions, and which was based upon approximation using the regularity theory of Ma, Trudinger, and Wang. The result in this paper improves his result, by relaxing certain geometrical assumptions on the domains and targets; it also completes his Hölder regularity theory of potential functions on the round sphere, by making it self-contained.

1. Introduction

For a real valued function $u$ on $\mathbb{R}^n$, its subdifferential (or subgradient) $\partial u(x)$ at $x$ is defined locally as

$$\partial u(x) = \{ p \in T_x \mathbb{R}^n \mid u(z) \geq u(x) + p \cdot z + o(|x - z|) \text{ as } z \to x \}.$$ 

It is clear that each $\partial u(x_0)$ is a convex set and if $u$ is differentiable at $x_0$, $\partial u(x_0) = \{ \nabla u(x_0) \}$. If $u$ is a convex function then for $p \in \partial u(x_0)$ the hyperplane $H_p(x) = x \cdot p + u(x_0)$ supports $u$ globally, i.e. $u \geq H_p$ on $\mathbb{R}^n$. In this paper, motivated by optimal mass transport theory, we are interested in the generalization of this fact for generalized convex functions, which are called cost-convex (or simply $c$-convex) functions, defined as the generalized Legendre transformations with respect to the cost function $c = c(x, y)$ on some product space $\Omega \times \Lambda$. These are called $c$-Legendre transformations in Section 2.

Ordinary convexity is the special case $c(x, y) = -x \cdot y$, and Brenier [B] and others [RR, KnS, CuP, Ra, Mc1] showed that convex functions give potential functions for the optimal mass transport problem. Namely, given two mass distributions $\mu_+$, $\mu_-$, i.e. two Borel probability measures on $\mathbb{R}^n$, assuming the measure $\mu_+$ does not charge sets of lower dimension, the unique mapping $T_0$ which minimizes the average cost functional

$$C(T) = \int_{\mathbb{R}^n} |x - T(x)|^2 d\mu_+$$
among all Borel measurable maps which push-forward $\mu_+$ onto $\mu_-$, in the sense that $\mu_-(E) = \mu_+(T^{-1}(E))$ for all Borel set $E \subset \mathbb{R}^n$, is actually given as the gradient mapping of some convex function $\phi$, i.e. $T_0 = \nabla \phi$. Here $T_0$ and its uniqueness are defined in the $\mu_+$-almost everywhere sense. Brenier’s result was generalized to other cost functions on Euclidean space $\text{Ca}_2$ $\text{GMc}$, to the Riemannian distance squared cost on manifolds $\text{Mc}_2$ and to the Heisenberg group $\text{AR}$, by Caffarelli, Gangbo and McCann, McCann, and Ambrosio and Rigot respectively. Sufficient conditions for the map to be smooth were investigated by Delanoé (for $n = 2$) $\text{D1}$, Caffarelli (for $n \geq 2$) $\text{Ca}_1$ $\text{Ca}_3$, and by Urbas $\text{U}$, for the Euclidean distance squared, and then by Cordero-Erausquin and Delanoé for the Riemannian distance squared on the torus $\text{C1}$ and nearly flat manifolds $\text{D2}$, and finally for other cost functions in a series of papers by Ma, Trudinger, Wang $\text{MTW}$ $\text{TW}_1$, and Loeper $\text{L}$. We refer the reader to the book $\text{V}$ by Villani for an introduction to optimal mass transportation theory and its applications in partial differential equations and geometry.

For a $c$-convex function $\phi$ on the domain $\Omega$ one may define the $c$-subdifferential $\partial^c \phi(x_0)$ at $x_0$ to consist of the gradient vectors $\nabla M_y(x_0)$ of the functions $M_y(x) = -c(x,y) + \text{const.}$ which globally support $u$ at $x_0$, meaning $u(z) \geq M_y(z)$ for all $z \in \Omega$ plus the equality $u(x_0) = M_y(x_0)$. The mapping from $x_0$ to the contact set, namely the set of such points $y \in \Lambda$, is called the $c$-contact mapping of $\phi$, and generalizes Brenier’s subgradient mapping for ordinary convex functions. Unlike $\partial \phi$, $\partial^c \phi$ is a global notion depending on the whole domain of $c$, $\Omega \times \Lambda$, but for $\phi$ convex and $c(x,y) = -x \cdot y$, $\partial^c \phi = \partial \phi$ and one may naturally ask whether this indeed holds for other more general cost functions.

**Question 1.1.** For which cost functions $c$ on domain $\Omega \times \Lambda$, is there the equivalence between the $c$-subdifferential $\partial^c \phi$ and the ordinary differential $\partial \phi$ for $c$-convex functions $\phi$?

An answer was given by Loeper $\text{L}$: there is such an equivalence $\partial^c \phi = \partial \phi$ under certain conditions on the function $c$, introduced by Ma, Trudinger, and Wang $\text{MTW}$ $\text{TW}_1$, who called them $\text{A0}$, $\text{A1}$, $\text{A2}$, and $\text{A3W}$, and for bounded domains $\Omega, \Lambda \subset \mathbb{R}^n$ satisfying certain geometric restrictions with respect to $c$, $c^*$, where $c^*(x,y) = c(y,x)$, namely that $\Omega$ and $\Lambda$ are mutually uniformly strictly $c$, $c^*$-convex and $c$, $c^*$-bounded (see Section $\text{I}$). Loeper’s result is based upon an approximation argument which uses Trudinger and Wang’s $C^3$-regularity theory of potential functions of optimal transport for general cost functions $\text{TW}_1$, and those conditions on $c$ and on $\Omega, \Lambda$ which Loeper needed were actually the ones originally imposed by Trudinger and Wang when they applied a continuity method to the Monge-Ampère type equations satisfied by the potential functions. Loeper’s result on $c$-convex functions was actually the first step toward his Hölder regularity theory of potential functions for optimal transport with rough mass distributions, which need not be absolutely continuous with respect to the Lebesgue measure, under a strict version $\text{A3S}$ of the condition $\text{A3W}$ (see $\text{L}$). He also extended his Hölder
regularity theory to the standard $n$-dimensional sphere with cost given by its Riemannian distance squared, and to the reflector antenna problem, which helped to motivate Wang’s interest in the theory of optimal mass transportation \cite{W1, W2} and has also been studied by many others, e.g. \cite{CaGH, KoO}.

The sufficient condition $A3W$ of Ma, Trudinger, and Wang for regularity was shown by Loeper \cite{L} to be necessary for the continuity of optimal maps with smooth integrable mass distributions whose Radon-Nikodym derivatives are bounded away from 0 and $+\infty$. It is also necessary for the equivalence $\partial^c\phi = \partial\phi$ for $c$-convex functions.

The goal of this paper is to provide a direct and geometric proof of the equivalence $\partial^c\phi = \partial\phi$ between the cost-subdifferential and the ordinary subdifferential for cost-convex functions $\phi$, under Ma, Trudinger, and Wang’s conditions $A0$, $A1$, $A2$, and $A3W$. Our result only assumes the mutual (not necessarily strict) $c$ and $c^*$-convexity of $\Omega$ and $\Lambda$. For example we do not assume them to be bounded, and this allows us to deal with such domains as $\mathbb{R}^n$. In fact we even assume slightly less (see Remark 6.1), and our result can easily be applied to domains in the round sphere $S^n$, including $S^n$ itself (see Example 7.5 and Example 7.6), thus it completes Loeper’s Hölder regularity theory of potential functions for optimal mass transport maps on domains in $\mathbb{R}^n$ or in $S^n$ without relying on Trudinger and Wang’s regularity result or any other method. As a consequence, it strengthens the results of Caffarelli, Gutierrez, and Huang \cite{CaGH}, by weakening their assumptions on the incoming intensity in the reflector antenna problem.

In this research announcement, we adopt the point of view that our domain $\Omega$ and target $\Lambda$ are either subdomains of Euclidean space, or else manifolds equipped with Riemannian metrics which allow us to identify their tangent and cotangent spaces. Unless otherwise stated, computations of higher derivatives are always carried out with respect to Riemannian exponential coordinates centered at the point where the derivatives are evaluated. In a subsequent paper, we abandon this artifice and give a fully covariant development of the results for general cost functions on smooth manifolds \cite{KiMc}. Aside from providing conceptual clarity and geometric insight, this point of view offers the possibility of extending the regularity theory of Loeper (and presumably that of Ma, Trudinger and Wang) to other geometries.

The present manuscript is organized as the following: In Section 2 we list the basic definitions and assumptions we need in this paper; in Section 3 we give the precise statements of main results; in Section 4 we summarize some of the relevant prior results by Ma, Trudinger, Wang, and Loeper; throughout Section 5 and Section 6 we prove our main results; in Section 7 we discuss some examples of interest.

**Acknowledgment.** We thank Gregoire Loeper, Neil Trudinger, and Xu-Jia Wang for their helpful discussions and comments, and their recent preprints. In particular, while writing this article, we learned from Trudinger and Wang that they had independently relaxed some of the restrictions on their earlier theory using a different method \cite{TW2}. There they indicated the desirability of a direct approach
to such questions, which we believe the present manuscript provides. We are grateful to Adrian Nachman and the 2006-07 participants of Fields Analysis Working Group, for the stimulating environment which they helped to create. This research was supported in part by Natural Sciences and Engineering Research Council of Canada Grant 217006-03 and United States National Science Foundation Grant DMS-0354729.

2. Definitions

In this section we list some (modified) definitions introduced by Ma, Trudinger, Wang, and Loeper [MTW][TW1][L]. In many cases, we follow the notation of [L], despite the fact that it is at odds with the notation of [GMc]. Let \( \Omega, \Lambda \) be two (not necessarily bounded) domains in Riemannian manifolds \( M \) and \( N \) respectively; for example \( M = N = \mathbb{R}^n \) with the Euclidean metric.

**Definition 2.1.** A cost function \( c : \Omega \times \Lambda \to \mathbb{R} \) is said to satisfy

\( (A_0) \) if \( c \in C^4_{\text{loc}}(\Omega \times \Lambda) \),

\( (A_1) \) if for all \( x \in \Omega \) the mapping \( y \to -\nabla_x c(x, y) \) is injective on \( \Lambda \) and for all \( y \in \Lambda \) the mapping \( x \to -\nabla_y c(x, y) \) is injective on \( \Omega \),

\( (A_2) \) if \( \det D^2_{xy} c(x, y) \neq 0 \), for all \( x \in \Omega \) and \( y \in \Lambda \).

**Definition 2.2.** (**c-Exponential Maps**) Under the assumption \( A_1 \), we define the \( c \) and \( c^* \)-exponential maps and their domains by

\[ c^{\text{-Exp}}_x \ p = y \text{ if } p = -\nabla_x c(x, y), \]
\[ \text{dom}(c^{\text{-Exp}}_x) = \{ p \in T_x \Omega \mid p = -\nabla_x c(x, y) \text{ for some } y \in \Lambda \}. \]

and

\[ c^*^{\text{-Exp}}_y \ q = x \text{ if } q = -\nabla_y c(x, y), \]
\[ \text{dom}(c^*^{\text{-Exp}}_y) = \{ q \in T_y \Lambda \mid q = -\nabla_y c(x, y) \text{ for some } x \in \Omega \}. \]

The assumption \( A_2 \) implies that \( c \) and \( c^* \)-exponential maps are local diffeomorphisms.

Throughout this paper we will assume \( A_0, A_1, A_2 \) for the cost function \( c : \Omega \times \Lambda \to \mathbb{R} \) and we define \( c^* \) by \( c^*(y, x) = c(x, y) \). The following simple observation will be useful.

**Lemma 2.1.** (**Symmetry**) For \( \eta \in T_x \Omega, \xi \in T_y \Lambda \),

\[ (D^2_{xy} c(x, y))^{-1} \eta \cdot \xi = \eta \cdot (D^2_{yx} c(x, y))^{-1} \xi \]

i.e.

\[ [d c^{\text{-Exp}}_x(c^{\text{-Exp}}_x)^{-1}(\eta)](\xi) = \eta \cdot [d c^*^{\text{-Exp}}_y(c^*^{\text{-Exp}}_y)^{-1}(\xi)](\eta) \]

where ‘\( \cdot \)’ means the Riemannian inner product.

**Proof.** The proof is straightforward.
Definition 2.3. \((c\text{-sectional curvature and A3W, A3S)}\) Define the c-section曲率 for \((x, y) \in \Omega \times \Lambda, \eta, \xi \in T_x \Omega,\) by
\[
\mathcal{S}_c(x, y)(\eta, \xi) := \frac{d^2}{dt^2} \bigg|_{t=0} [-D^2_{xx} c](x, c\text{-Exp}_x(p + t\xi)) \eta \eta
\]
for \(p = -\nabla_x c(x, y) = [c\text{-Exp}_x]^{-1}(y).\) We say that \(c\) satisfies A3W (resp. A3S) if there exists a constant \(C_0 \geq 0\) (resp. \(> 0\)) such that for all \((x, y) \in \Omega \times \Lambda,\)
\[
(2.1) \quad \mathcal{S}_c(x, y)(\eta, \xi) \geq C_0 |\eta|^2 |\xi|^2 \quad \text{for all } \eta \perp \xi.
\]

Definition 2.4. (c-convexity) We say \(\Lambda\) is c-convex with respect to \(x \in \Omega\) if \(c\text{-Exp}_x\) is convex as a set in the Euclidean space \(T_x \Omega.\) We say \(\Lambda\) is c-convex with respect to \(\Omega\) if it is c-convex with respect to all \(x \in \Omega.\) We define similarly the c*-convexity of \(\Omega.\)

Definition 2.5. (c-convex function indexed by a set) We define the c-Legendre transformation \(L^c c\) of \(v : \Lambda \to \mathbb{R} \cup \{+\infty\}\) to be the function on \(\Omega\) given by
\[
L^c v(x) = \sup_{y \in \Lambda} -c(x, y) - v(y).
\]
We define a function \(\phi\) on \(\Omega\) to be a c-convex function indexed by \(\Lambda,\) if
\[
(2.2) \quad \phi = L^c v \text{ for some function } v \text{ on } \Lambda
\]
and
\[
(2.3) \quad \text{for each } x \in \Omega \text{ there is } y \in \Lambda \text{ such that } \phi(x) + v(y) = -c(x, y).
\]
Setting \(c^*(x, y) = c(y, x)\) we may define the c*-Legendre transformation \(L^{c^*}\) from functions on \(\Omega\) to the functions on \(\Lambda,\) and c*-convex functions indexed by \(\Omega,\) similarly.

Remark 2.1. (2.3) is not explicitly assumed for c-convexity in [MTW] [TW1] [L], since the domains there \(\Omega\) and \(\Lambda\) are assumed to be bounded, and hence (2.3) is automatically satisfied due to compactness of \(\Omega \times \Lambda\) and continuity of \(c.\)

Definition 2.6. (c-contact mapping) Let \(u\) be a c-convex function on \(\Omega\) indexed by \(\Lambda.\) We may continuously extend it to the closure \(\overline{\Omega}\) of \(\Omega.\) We define the set-valued mapping \(G_u\) by
\[
G_u(x) := \{y \in \Lambda | u(x) + L^{c^*}(u)(y) = -c(x, y)\} \text{ for each } x \in \overline{\Omega}.
\]
\(G_u\) is called the c-contact mapping (or simply the contact mapping) of \(u,\) and the set \(G_u(x)\) is called the contact set of \(u\) at \(x.\) We can also define in a similar way the contact map \(G_v\) from \(\Lambda\) to \(\overline{\Omega},\) for a c*-convex function \(v\) on \(\Lambda\) indexed by \(\Omega.\) Then, it is clear that for any \(x \in \overline{\Omega}, \) \(y \in \Lambda\) if \(v = L^{c^*}(u),\)
\[
G_u(x) = (G_v)^{-1}(x), \quad G_v(y) = (G_u)^{-1}(y).
\]

Definition 2.7. (subdifferential and c-subdifferential) Assume that \(\Omega\) is given a Riemannian inner product \(\cdot\) and the distance function \(d\) induced by this metric.
Let $u$ be a $\mathbb{R} \cup \{+\infty\}$-valued function on $\Omega$. The subdifferential $\partial u$ is defined locally for each $x \in \Omega$ by

$$\partial u(x) = \{ p \in T_x \Omega \mid u(z) \geq u(x) + p \cdot \exp^{-1} z + o(d(x, z)) \text{ as } z \to x \}.$$ 

Note that the subdifferential is always a convex set though it may possibly be empty. However, for a $c$-convex function $u$ indexed by $\Lambda$, this set is nonempty, and we may further define the $c$-subdifferential $\partial^c u$ for each $x_0 \in \Omega$ by

$$\partial^c u(x_0) := \{ p \in T_{x_0} \Omega \mid p = -\nabla_{x=x_0} c(x, y), \ y \in G_u(x_0) \}.$$ 

One should note that $\partial^c u(x) \subset \partial u(x)$ for any $x \in \Omega$ and that $\partial^c u$ depends globally on the sets $\Omega, \Lambda$.

3. Main Results

Using the definitions given in the previous section, we can now precisely state our main result. It was formulated by Loeper [L], who proved it under various technical hypotheses described in the section to follow using a result of Trudinger and Wang [TW1].

**Theorem 3.1. (A3W $\iff$ CSIS)** Let $\Omega, \Lambda$ be two (not necessarily bounded) domains in Riemannian manifolds $M$ and $N$, resp. (e.g. $M = N = \mathbb{R}^n$, with Euclidean metric). Assume that the cost $c : \Omega \times \Lambda \to \mathbb{R}$ satisfies $A0$, $A1$, $A2$ and that $\Lambda$ is $c$-convex with respect to $\Omega$, and $\Omega$ is $c^*$-convex with respect to $\Lambda$. Then the following are equivalent.

1. (A3W) $c$ satisfies $A3W$.

2. (CSIS: $c$-subdifferential is subdifferential) For any $c$-convex function $\phi$ on $\Omega$ indexed by $\Lambda$,

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

From this theorem, we can deduce the connectivity of contact sets of $c$-convex functions when we assume A3W.

**Corollary 3.2. (connectivity of contacts sets)** Under the assumption of Theorem 3.1 if further A3W holds, then for any $c$-convex functions $\phi$ on $\Omega$ indexed by $\Lambda$, the contact set $G_\phi(x) \subset \Lambda$ of $\phi$ is $c$-convex (and so connected) for all $x \in \Omega$.

The following corollary is useful, especially when $\Omega = \mathbb{R}^n$ (see Example 7.1 and Example 7.3), and it is parallel with Caffarelli’s result [Ca1] for Euclidean distance squared cost, which required the convexity only of the target domain $\Lambda$.

**Corollary 3.3. (arbitrary subdomains)** Put the same assumption on $c$, $\Omega$, $\Lambda$ as in Theorem 3.1 Then for any arbitrary subdomain $\Omega_1 \subset \Omega$ and for a subset $\Lambda_1 \subset \Lambda$ $c$-convex with respect to $\Omega_1$, and for any $c$-convex function $\phi$ on $\Omega_1$ indexed by $\Lambda_1$, we have $\partial^\phi \phi = \partial \phi$.

**Proof.** The proof is given at the end of Section 5. □
4. Previous results: Ma, Trudinger, Wang, and Loeper

We now discuss the previous results of Ma, Trudinger, Wang, and Loeper. The following conditions were critical in their initial studies \cite{MTW}, \cite{TW1}, \cite{L}. Our approach avoids many of these restrictions, so that our arguments in Section \ref{section5} and Section \ref{section6} only assume (not necessarily uniformly strict) $c, c^*$-convexity — an improvement which, like Corollary \ref{corollary3.3}, was attained independently in the parallel work of Trudinger and Wang \cite{TW2}.

We denote
\begin{equation}
A(x,p) = -D^2_{xx}c(x,y) \text{ for } y \text{ such that } -\nabla_x c(x,y) = p.
\end{equation}

**Definition 4.1. (Uniform strict $c, c^*$-convexity)** We say that $\Omega$ is uniformly strictly $c^*$-convex with respect to $\Lambda$ if its inverse $c^*$-exponential image $[c^*-\text{Exp}_y]^{-1}(\Omega)$ is uniformly strictly convex with respect to $y \in \Lambda$, i.e., if $\Omega$ is $c^*$-convex with respect to $\Lambda$, $\partial\Omega \in C^2$ and there exists a constant $\delta_0 > 0$ such that
\begin{equation}
[D_i \gamma_j(x) - D_{pk} A_{ij}(x,p) \gamma_k] \tau_i \tau_j \geq \delta_0
\end{equation}
for all $x \in \partial\Omega$, $p \in -\nabla_x c(x,\Lambda) \subset T_x \Omega$, unit tangent vector $\tau$ and outer unit normal vector $\gamma$. Here $A$ is given in \eqref{4.1}. We similarly define the uniform strict $c$-convexity of $\Lambda$ with respect to $\Omega$.

**Definition 4.2. ($c, c^*$-boundedness)** We say that $\Omega$ is $c^*$-bounded, with respect to $\Lambda$ if there exists some function $\varphi \in C^2(\Omega)$, satisfying for some constant $\delta_1 > 0$,
\begin{equation}
[D_i^2 \varphi(x) - D_{pk} A_{ij}(x,p) D_k \varphi] \xi_i \xi_j \geq \delta_1 |\xi|^2
\end{equation}
for all $x \in \Omega$ and for all $p \in -\nabla_x c(x,\Lambda) \subset T_x \Omega$. Here $A$ is given in \eqref{4.1}. We similarly define the $c$-boundedness of $\Lambda$.

**Remark.** In \cite{TW1}, Trudinger and Wang mean by $c$-convexity and $c^*$-boundedness of $\Omega$, the $c^*$-convexity and $c^*$-boundedness in this paper. We follow the convention given in \cite{L}.

**Remark.** (\cite{TW1}, p19) Any sufficiently small ball $\Lambda$ is $c$-bounded with respect to a bounded set $\Omega$.

In \cite{MTW}, \cite{TW1}, Ma, Trudinger and Wang obtained the following regularity result for $c$-convex functions as solutions to optimal transport problems.

**Theorem 4.1. (Ma, Trudinger, and Wang. \cite{MTW} \cite{TW1} see also \cite{L}, p13)** Let $\Omega, \Lambda$ be two bounded domains in $\mathbb{R}^n$. Let $c$ satisfy $A_0, A_1, A_2,$ and $A_3 W$ on $\Omega \times \Lambda$. Assume that $\Omega, \Lambda$ are strictly uniformly $c, c^*$-convex with respect to each other, either $\Lambda$ is $c$-bounded with respect to $\Omega$ or $\Omega$ is $c^*$-bounded with respect to $\Lambda$. Then for two probability measures $\mu_0, \mu_1$ on $\Omega, \Lambda$ having densities $\rho_1 \in C^2(\overline{\Omega})$, $\rho_2 \in C^2(\overline{\Lambda})$ which are bounded away from 0, and for any $c$-convex function $\phi$ on $\Omega$, with $G_\phi(\Omega) \subset \overline{\Lambda}$, $(G_\phi)\#\mu_0 = \mu_1$, we have $\phi \in C^3(\overline{\Omega})$.

Using this regularity result, Loeper showed the equivalence between $c$-subdifferentials and ordinary subdifferentials (under the same assumptions of Trudinger and Wang), by approximating $c$-convex functions by smooth ones.
Theorem 4.2. (Looper [L.13, Theorem 3.1.]) Under the same assumption on the cost $c$, domains $\Omega$, $\Lambda$ as in Theorem 4.1 (especially with $c$ or $c^*$-boundedness), we have for any $c$-convex function $\phi$ on $\Omega$, with $G_\phi(\Omega) \subset \overline{\Lambda}$, that
\[ \partial^c \phi(x) = \partial \phi(x) \text{ for all } x \in \Omega, \]
and so the set $\partial^c \phi(x) \in T_x \Omega$ is convex for all $x \in \Omega$.

5. Proof of Theorem 3.1

In this section, we illustrate a blueprint of the proof of Theorem 3.1. We first start with the geometric setting of the proof. Pick $x_m \in \Omega$ and two different points $y_0, y_1 \in \Lambda$ and their inverse $c$-exponential images $p_0, p_1 \in T_{x_m} \Omega$, $p_0 \neq p_1$. Let $p_\theta = (1 - \theta)p_0 + \theta p_1$, $\theta \in \mathbb{R}$, be the line in $T_{x_m} \Omega$ passing through $p_0$ with direction $p_1 - p_0$. For such $p_\theta \in \text{dom}(c\text{-Exp}_{x_m})$, consider its $c$-exponential image $y_\theta$, i.e.
\[ p_\theta = -\nabla_x c(x_m, y_\theta) \in T_{x_m} \Omega, \quad y_\theta = c\text{-Exp}_{x_m}(p_\theta). \]
Then, by the condition A2, we see that $\frac{\partial}{\partial \theta} y_\theta \neq 0$. If such $y_\theta$ is defined at least for $\theta \in [0, 1]$, then we denote by $[y_0, y_1]_{x_m}$ the curve $\{y_\theta\}_{0 \leq \theta \leq 1}$ and call it the $c$-segment with respect to $x_m$ joining $y_0$ to $y_1$. Such $[y_0, y_1]_{x_m}$ exists, for example if $\Lambda$ is $c$-convex with respect to $x_m$.

Definition 5.1. (Sliding and double mountains) Given $x_m \in \Omega$ and $y_0, y_1 \in \Lambda$, define the function $f_\theta$ on $\Omega$, which we call the sliding mountain between $y_0, y_1$ centered at $y_\theta$ and normalized at $x_m$, by
\[ f_\theta(x) = -c(x, y_\theta) + c(x_m, y_\theta) \]
where $y_\theta$ is the point corresponding to $\theta \in [0, 1]$ in the $c$-segment $[y_0, y_1]_{x_m}$. We call $\max\{f_0, f_1\}$ the double mountain of $y_0, y_1$, normalized at $x_m$.

Definition 5.2. (The set $S_\theta$ and $S_\theta^+$) We define for the sliding mountain $f_\theta$,
\[ \text{the level set } S_\theta := \{x \in \Omega \mid \frac{\partial}{\partial \theta} f_\theta(x) = 0\} \text{ and} \]
\[ \text{the super-level set } S_\theta^+ := \{x \in \Omega \mid \frac{\partial}{\partial \theta} f_\theta(x) \geq 0\}. \]
In this notation we have the following key observation whose proof is straightforward.

Lemma 5.1. (DASM $\leftrightarrow$ Monotonicity) For $x_m \in \Omega$, suppose that $y_0, y_1 \in \Lambda$ is joined by the $c$-segment $[y_0, y_1]_{x_m}$ with respect to $x_m$. Then (1) is equivalent to (2), and (1') is equivalent to (2'):

(1) (DASM: Double Mountain Above Sliding Mountains)
\[ f_\theta(x) \leq \max\{f_0, f_1\}(x), \text{ for all } (\theta, x) \in [0, 1] \times \Omega. \]

(2) (Monotonicity) The super-level set $S_\theta^+$ monotonically increases in $\theta$, i.e.
\[ S_{\theta_1}^+ \subset S_{\theta_2}^+, \text{ for any } 0 \leq \theta_1 \leq \theta_2 \leq 1. \]

(1') (Strict DASM) \[ (5.3) \text{ holds, and the inequality is strict unless } x = x_m. \]

(2') (Strict Monotonicity) \[ (5.4) \text{ holds and } S_{\theta_1} \cap S_{\theta_2} = \{x_m\} \text{ if } \theta_1 < \theta_2. \]
Remark 5.1. In this lemma, we do not assume \( c, c^*-\)convexity on \( \Omega, \Lambda \).

Remark 5.2. (DASM \( \iff \) CSIS) Note that (5.3) is equivalent to the convexity of the \( c \)-subdifferential \( \partial^c \max [f_0, f_1](x_m) \) at \( x_m \), and so to \( \partial^c \max [f_0, f_1](x_m) = \partial \max [f_0, f_1](x_m) \). In fact, it is easy to see as observed by Loeper (see Proposition 2.12 in [L]) that for \( \Lambda \) being \( c \)-convex with respect to \( \Omega \), the property (5.3) holds for every \( x_m, x \in \Omega \), and for all \( c \)-segments \( [y_0, y_1] \subset \Lambda \) if and only if \( \partial^c \phi = \partial \phi \) for all \( c \)-convex functions \( \phi \) on \( \Omega \) indexed by \( \Lambda \).

Remark 5.3. (DASM \( \Rightarrow \) A3W) Loeper used Taylor expansion to prove the important result that

\[
\text{(5.3) holds locally near } x_m \Rightarrow \text{A3W}.
\]

(See [L] pp20–21.)

From Lemma 5.1, Remark 5.2, and Remark 5.3, it is clear that to prove Theorem 5.1 it suffices to prove either (DASM) or (Monotonicity) for all \( x_m, x, [y_0, y_1] \subset \Lambda \), while assuming A3W. We will do this by proving the following theorem in Section 6.

Theorem 5.2. (A3 \( \Rightarrow \) Monotonicity) Under the same assumption on \( c, \Omega, \Lambda \) as in Theorem 3.1, if \( c \) satisfies A3W then (Monotonicity) (5.4) holds for all \( x_m \in \Omega \) and \( y_0, y_1 \in \Lambda \). Similarly A3S implies that (Strict Monotonicity) holds for all \( x_m \in \Omega \) and \( y_0, y_1 \in \Lambda \).

We emphasize that our proof of Theorem 5.2, thus the direction (1) \( \Rightarrow \) (2) in Theorem 3.1, is elementary and geometric. It is not based on the previous results of Ma, Trudinger, Wang, and Loeper.

We finish this section by giving the proof of Corollary 3.3.

Proof of Corollary 3.3. It is clear that the property (DASM) (5.3) has monotonicity with respect to the domain: For \( \Omega_1 \subset \Omega, \Lambda_1 \subset \Lambda \) with \( \Lambda, \Lambda_1 \) being \( c \)-convex with respect to \( \Omega, \Omega_1 \), respectively, if (5.3) holds for every \( x_m, x \in \Omega \), and for all \( c \)-segments \( [y_0, y_1] \subset \Lambda \), then the same holds for the pair \( \Omega_1, \Lambda_1 \). Therefore we can conclude using Remark 5.2 and Theorem 3.1 that \( \partial^c \phi = \partial \phi \) for all \( c \)-convex functions \( \phi \) on \( \Omega_1 \) indexed by \( \Lambda_1 \).

6. Proof of Theorem 5.2 (A3 \( \Rightarrow \) Monotonicity)

We divide the proof into three steps. We will first set up the necessary notation and show some preliminary geometric facts in the subsection 6.1. Then, we will show in the subsection 6.2 that the rate of expansion of \( S^+_\theta \) along \( S_\theta \) (see Definition 5.3 and Definition 5.2) has the same sign as the quantity \( \frac{d^2}{d\theta^2} f_0 \) on \( S_\theta \). In the subsection 6.3 we will show the non-negativity of this expansion rate to conclude the desired monotone dependence of the rising region \( S^+_\theta \) on \( \theta \in [0, 1] \) (see Proposition 6.2), and it will finish our proof of Theorem 5.2.
6.1. Geometric preliminaries. We first recall the definitions of $p_\theta, y_\theta$ from Section 5. For each $y_\theta$, we define

\begin{equation}
q_\theta = -\nabla_y c(x_m, y_\theta) \in T_{y_\theta} \Lambda, \quad x_m = c^*\text{-Exp}_y(q_\theta)
\end{equation}

We denote $\dot{y}_\theta = \frac{d}{dt}y_\theta$, $\ddot{y}_\theta = \frac{d^2}{dt^2}y_\theta$, and we also use $\cdot$ to denote the Riemannian inner product. Our calculations will be performed in geodesic normal coordinates of $x \in \Omega$ and $y_\theta \in \Lambda$.

Now, let’s discuss the geometric properties of $S_\theta$. We first note that

\begin{equation}
D_x \frac{d}{dt} f_\theta \cdot \xi = \left( - D_{y,x}^2 c(x, y_\theta) \right) \dot{y}_\theta, \quad \text{for any tangent vector } \xi \text{ at } x,
\end{equation}

and by A2 that $D_x \frac{d}{dt} f_\theta$ is nowhere vanishing on $\Omega$. Thus we see the zero level set $S_\theta$ of $\frac{d}{dt} f_\theta$ is a smooth $(n - 1)$-dimensional submanifold in the $n$-dimensional domain $\Omega$. It is clear that $x_m \in S_\theta$ and that

\begin{equation}
S_\theta = c^*-\text{Exp}_{y_\theta}((q_\theta + W_\theta) \cap \text{dom}(c^*-\text{Exp}_y)),
\end{equation}

for the $(n - 1)$-dimensional subspace

\begin{equation}
W_\theta = \{ w \in T_{y_\theta} \Lambda \mid w \cdot \dot{y}_\theta = 0 \} \subset T_{y_\theta} \Lambda.
\end{equation}

We parametrize by $w$ the points in $W_\theta \subset T_{y_\theta} \Lambda$ and for each $w$ let $x_w$ be the point in $S_\theta \subset \Omega$ given by

\begin{equation}
w = -D_y c(x_w, y_\theta) + q_\theta.
\end{equation}

At each point $x_w \in S_\theta$, denote by $p_\theta(x_w)$ the inverse-c-exponential image of $y_\theta$ given by

\begin{equation}
p_\theta(x_w) = -D_x c(x_w, y_\theta),
\end{equation}

and denote

\begin{equation}
\dot{p}_\theta(x_w) = \frac{d}{dt} p_\theta(x_w).
\end{equation}

Note that $w \cdot \dot{y}_\theta = 0$, and by Lemma 2.1, $\dot{p}_\theta(x_w)$ is orthogonal to $S_\theta$ at each $x_w$, i.e.

\begin{equation}
\dot{p}_\theta(x_w) \cdot D_w x_w = 0,
\end{equation}

when the derivative $D_w x_w$ is taken for any direction of $w$. In particular, we see that $S_\theta$ is orthogonal to $p_1 - p_0$ at $x$ for all $\theta$.

As our key ingredients, we consider for each $\theta$ the functions (on $\Omega$)

\begin{align*}
d f_\theta(x) &= \left( - D_y c(x, y_\theta) + D_y c(x_m, y_\theta) \right) \dot{y}_\theta, \\
\frac{d^2}{dt^2} f_\theta(x) &= \left( - D_{y,x}^2 c(x, y_\theta) + D_{y,x}^2 c(x_m, y_\theta) \right) \ddot{y}_\theta + \left( - D_y c(x, y_\theta) + D_y c(x_m, y_\theta) \right) \dddot{y}_\theta.
\end{align*}

Note that at $x_m$, $f_\theta(x_m)$, $\frac{d}{dt} f_\theta(x_m)$, $\frac{d^2}{dt^2} f_\theta(x_m) = 0$, and for all $x \in T_{x_m} \Omega$,

\begin{align*}
D_x \frac{d^2}{dt^2} f_\theta(x_m) \xi &= -D_x D_{y,x}^2 c(x_m, y_\theta) \xi \dot{y}_\theta \\
&\quad - D_{y,x}^2 c(x_m, y_\theta) \xi \ddot{y}_\theta.
\end{align*}
And, by differentiating \((1 - \theta)p_0 + \theta p_1 = -D_x c(x_m, y_\theta)\) twice with respect to \(\theta\), we see
\[
0 = -D^2_{y\theta} D_x c(x_m, y_\theta) \dot{y}_\theta \dot{y}_\theta - D^2_{x\theta} c(x_m, y_\theta) \ddot{y}_\theta
\]
and that
\[
D_x \frac{d^2}{dy^2} f_\theta(x_m) = 0.
\]

6.2. The rate of expansion of \(S^+_\theta\) along \(S_\theta\) has the same sign as the quantity \(\frac{d^2}{dy^2} f_\theta\) on \(S_\theta\). This result comes form the following lemma, by letting \(g_t = \frac{d}{dt} f_\theta \big|_{t=0} \).

**Lemma 6.1.** Let \(g : (-\varepsilon, \varepsilon) \times U \to \mathbb{R}\) be a \(C^2\) function, where \(U\) is an open domain in \(\mathbb{R}^n\). Denote \(g_t(x) = g(t, x)\) and suppose the gradient \(\nabla_x g_t\) is non-vanishing on \(U\) for \(t \in (-\varepsilon, \varepsilon)\). Then the moving level set \(L_t := \{x \in U \mid g_t(x) = 0\}\) is locally parametrized for small \(t\) by \(X(t, x), x \in L_0\), where we define \(X(t, x)\) by solving the following first order ODE:
\[
\begin{align*}
\frac{d}{dt} X(t, x) &= -\frac{\hat{g}_t(X(t, x))}{\|\nabla_x \hat{g}_t(X(t, x))\|} \nabla_x \hat{g}_t(X(t, x)) \\
X(0, x) &= x,
\end{align*}
\]
where we denote \(\hat{g}_t := \frac{\partial}{\partial t} g_t\). Moreover, the expansion rate of the super level set \(L^+_t := \{x \in U \mid g_t(x) \geq 0\}\) along \(L_t\) in this parametrization is given by
\[
-\frac{\nabla_x \hat{g}_t(x)}{\|\nabla_x \hat{g}_t(x)\|} \cdot \frac{d}{dt} X(t, x) = \frac{\hat{g}_t(x)}{\|\nabla_x \hat{g}_t(x)\|}
\]

**Proof.** Since
\[
\frac{d}{dt} g_t(X(t, x)) = \hat{g}_t(X(t, x)) + \nabla_x \hat{g}_t(X(t, x)) \cdot \frac{d}{dt} X(t, x) = 0,
\]
it is clear that this \(X(t, x)\) is indeed the desired parametrization of \(L_t\). \(\square\)

6.3. The positivity of \(\frac{d^2}{dp^2} f_\theta\big|_{S_\theta}\). Now we will show that \(\frac{d^2}{dp^2} f_\theta\big|_{S_\theta} \geq 0\) under **A3W**. Under **A3S** condition, the equality holds only at \(x_m\), so it is the only common point of the expanding boundaries \(S_\theta\). This result together with the result in the subsection 6.2 will finish the proof of Theorem 5.2.

**Proposition 6.2.** (Positivity of \(\frac{d^2}{dp^2} f_\theta\) on \(S_\theta\)) Suppose that the cost function \(c\) satisfies **A0**, **A1**, & **A2**. Assume further that \(\Omega\) is \(c^\ast\)-convex with respect to \(\Lambda\) and that the \(c\)-segment \(\{y_\theta\}_0 \leq \theta \leq 1\) with respect to \(x_m\) is in \(\Lambda\). Then we have:

1. under **A3W**, \(\frac{d^2}{dp^2} f_\theta\big|_{S_\theta} \geq 0\).
2. under **A3S**, \(\frac{d^2}{dp^2} f_\theta\big|_{S_\theta}(x) > 0\) for \(x \neq x_m\).

To prove this proposition, let’s consider the following key lemma whose proof is given at the end of this section.
Lemma 6.3. (Sliding mountain and \(c\)-curvature) We recall the \(c\)-sectional curvature \(S_c\) in Definition 2.3 and the parametrization \(w\) of the set \(W_\theta\) of tangent vectors in \(T_{y_\theta}\Lambda\) orthogonal to \(\dot{y}_\theta\) as given in (6.3). Let \(x_w\) be the points in \(S_\theta \subset \Omega\) defined as in (6.4), and let \(p_\theta, \dot{p}_\theta\) be the vectors defined in (6.5) and (6.6). Fix a tangent vector \(\eta \in T_wW_\theta\), and let \(D_w\) be the directional derivative in \(w\) of the direction \(\eta\) and let \(D^2_ww\) be the second iteration of \(D_w\). Then

\[
D^2_ww \frac{d^2}{d\theta^2} f_\theta(x_w) = S_c(x_w, y_\theta)(D_wx_w, \dot{p}_\theta(x_w)).
\]

Proof of Proposition 6.2. Denote

\[ V_\theta := (q_\theta + W_\theta) \cap \text{dom}(c^*-\text{Exp}_{y_\theta}) \]

and note that \(S_\theta = c^*-\text{Exp}_{y_\theta}(V_\theta)\). We may regard the function \(\frac{d^2}{dt^2} f_\theta|_{S_\theta}\) as a function, say \(\tilde{f}\), on \(V_\theta\) by pulling it back by \(c^*-\text{Exp}_{y_\theta}\). We apply Lemma 6.3, then under \(A3W\) and the orthogonality (6.7), we see that the function \(\tilde{f}\) is convex.

Now the \(c^*\)-convexity of \(\Omega\) at each \(y_\theta\) implies that \(V_\theta\) is convex, and so \(\tilde{f}\) is a convex function on a convex domain. Then we see by (6.9) that the function \(\frac{d^2}{dt^2} f_\theta|_{S_\theta}\) has the minimum value 0 at \(x_m\) so implying (1), and this minimum is unique under \(A3S\) when \(\tilde{f}\) is strictly convex so implying (2). \(\square\)

Remark 6.1. The proof of Proposition 6.2 is the only place in the proof of Theorem 5.2 where we use the \(c^*\)-convexity of \(\Omega\) with respect to \(\Lambda\). What we actually need here is that a dense subset of the set \((q_\theta + W_\theta) \cap \text{dom}(c^*-\text{Exp}_{y_\theta})\) can be reached by line segments from \(q_\theta\) in \(\text{dom}(c^*-\text{Exp}_{y_\theta})\). We can use this slight weaker condition, to deal with such examples as \(c(x, y) = -\frac{1}{2} \log |x - y|^2\) in \(\mathbb{R}^n \times \mathbb{R}^n \setminus \{x = y\}\) (Example 7.3); \(c(x, y) = \pm \frac{1}{p} |x - y|^p, p \neq 0\) for \(p = -2\) or \(-\frac{1}{2} \leq p < 1\) (only) (Example 7.4); the cost \(c(x, y) = -\frac{1}{p} \log |x - y|^2\) restricted on \(S^{n-1} \subset \mathbb{R}^n\) which appears in the reflector antenna problem (Example 7.6).

In the rest of the section, we prove Lemma 6.3.

Proof of Lemma 6.3 (Sliding mountain and \(c\)-curvature). We will use the same notation of Lemma 6.3 and perform our calculations in the geodesic normal coordinates of \(x_w\) and of \(y_\theta\). Differentiating (6.4) twice with respect to \(w\) we see that

\[
0 = -D^2_{xx} D_y c(x_w, y_\theta) D_w x_w \quad D_w x_w - D^2_{yy} c(x_w, y_\theta) D^2_{ww} x_w.
\]
We compute by using (6.8) and (6.11)
\[
D^2_{ww} \frac{d^2}{d\theta^2} f_\theta(x_w)
\]
\[
= -D^2_{xx} D^2_{yy} c(x_w, y_0) D_{w, x_w} D_{w, x_w} \dot{\gamma}_0 \cdot \dot{\gamma}_0
- D_x D^2_{yy} c(x_w, y_0) D^2_{ww, x_w} \dot{\gamma}_0 \cdot \dot{\gamma}_0
+ \left[- D^2_{xx} D_y c(x_w, y_0) D_{w, x_w} D_{w, x_w} - D^2_{yy} c(x_w, y_0) D^2_{ww, x_w} \right] \cdot \ddot{\gamma}_0
\]
(6.12)
\[
= -D^2_{xx} D^2_{yy} c(x_w, y_0) D_{w, x_w} D_{w, x_w} \dot{\gamma}_0 \cdot \dot{\gamma}_0
- D_x D^2_{yy} c(x_w, y_0) \left[ \left( D^2_{yy, c(x_w, y_0)} \right)^{-1} \left( D^2_{xx} \right) D_y c(x_w, y_0) D_{w, x_w} D_{w, x_w} \right] \dot{\gamma}_0 \cdot \dot{\gamma}_0,
\]
To compute \( \xi_c(x_w, y_0) (D_{w, x_w, \bar{p}_o(x_w)}) \), we consider the line segment \( p_0(x_w) + \psi \bar{p}_o(x_w), \psi \in \mathbb{R} \), in \( T_{x_o, \Omega} \), and its \( \epsilon \)-exponential image \( y(\psi) \) as
\[
(6.13)
\]
\[
\bar{p}_o(x_w) + \psi \bar{p}_o(x_w) = - D_x c(x_w, y(\psi)).
\]
Note that \( y(\psi)\big|_{\psi=0} = y_0 \), and by differentiating (6.13) with respect to \( \psi \) that
\[
\dot{\bar{p}}_o(x_w) = - D^2_{xy} c(x_w, y_0) \left. \frac{d}{d\psi} y(\psi) \right|_{\psi=0},
\]
therefore, we see
\[
\dot{\gamma}_0 = \left. \frac{d}{d\psi} y(\psi) \right|_{\psi=0}
\]
and by differentiating (6.13) twice with respect to \( \psi \) that
\[
(6.14)
0 = -D^2_{yy} D_x c(x_w, y(\psi)) \ddot{\gamma}_0 \dot{\gamma}_0 - D^2_{xy} c(x_w, y(\psi)) \left. \frac{d^2}{d\psi^2} y(\psi) \right|_{\psi=0}
\]
Now we compute using (6.14)
\[
\xi_c(x_w, y_0) (D_{w, x_w, \bar{p}_o(x_w)})
\]
\[
= \left. \frac{d^2}{d\psi^2} \right|_{\psi=0} - D^2_{xx} c(x_w, y(\psi)) D_{w, x_w} D_{w, x_w}
- D^2_{yy} D^2_{xx} (x_w, y_0) D_{w, x_w} D_{w, x_w} \ddot{\gamma}_0 \ddot{\gamma}_0
- D_x D^2_{xx} c(x_w, y_0) D_{w, x_w} D_{w, x_w} \left. \frac{d^2}{d\psi^2} y(\psi) \right|_{\psi=0}
\]
(6.15)
\[
= -D^2_{yy} D^2_{xx} (x_w, y_0) D_{w, x_w} D_{w, x_w} \ddot{\gamma}_0 \ddot{\gamma}_0
- D_x D^2_{xx} c(x_w, y_0) D_{w, x_w} D_{w, x_w} \left[ \left( D^2_{yy} c(x_w, y_0) \right)^{-1} \left( -D^2_{yy} D_x c(x_w, y_0) \ddot{\gamma}_0 \dot{\gamma}_0 \right) \right]
\]
Therefore, comparing (6.12) and (6.15) by interchanging the order of the derivatives and by Lemma 2.1, we see that
\[
D^2_{ww} \frac{d^2}{d\theta^2} f_\theta(x_w) = \xi_c(x_w, y_0) (D_{w, x_w, \bar{p}_o(x_w)}).
\]
This finishes the proof of Lemma 6.3. \( \square \)
7. Examples

In the following, we demonstrate our result in some examples of cost functions among the ones given by Ma, Trudinger, and Wang [MTW], [TW1], and by Loeper [L]. A common feature of these examples is that they only require the $c$-convexity of the target domain $\Lambda$ with respect to the source domain $\Omega$, not the other direction, for the equivalence $\partial^c\phi = \partial\phi$ of $c$-subdifferential and ordinary subdifferential of $c$-convex functions $\phi$ on $\Omega$ indexed by $\Lambda$.

**Example 7.1.** $c(x, y) = |x - y|^2 + |f(x) - g(y)|^2$ satisfies $A0$, $A1$, $A2$, $A3W$ (resp. $A3S$) for $\Omega = \Lambda = \mathbb{R}^n$, if $f, g : \mathbb{R}^n \to \mathbb{R}$ convex (resp. strictly convex) and $|\nabla f|, |\nabla g| < 1$. Such $\Omega$ and $\Lambda$ are $c, c^*$-convex with respect to the other, since the domains of $c, c^*$ exponential maps are the whole space $\mathbb{R}^n$ (One can verify this for example, by showing that the map $-\nabla x c(x, \cdot)$ is a proper map.) So by Corollary 6.3 we see that for any domains $\Omega_1, \Lambda_1 \subset \mathbb{R}^n$, $\Omega_1$ being $c$-convex with respect to $\Omega_1$, and for any $c$-convex function $\phi$ on $\Omega_1$ indexed by $\Lambda_1$, we have $\partial^c\phi = \partial\phi$ without the $c^*$-convexity of $\Omega_1$. This result is in fact easy to show by directly computing $\frac{d^2}{dx} f_\theta$ and checking its positivity.

**Example 7.2.** We can conclude the same as in Example 7.1 for the cost function $c(x, y) = \sqrt{1 + |x - y|^2}$ on $\mathbb{R}^n \times \mathbb{R}^n$.

In the following examples, we will use the notation of Section 5. These examples concern cost functions that have singularities.

**Example 7.3.** $c(x, y) = -\frac{1}{2} \log |x - y|^2$ on $\mathbb{R}^n \times \mathbb{R}^n \setminus \{x = y\}$ satisfies $A3S$. In this example, one can show that the curve $\{y_\theta\}_{\theta \in \mathbb{R}}$ is a circle (possibly with infinite radius) passing through $x_m$, tangent to $p_1 - p_0$ at $x_m$, and that $S_\theta$'s are $(n-1)$-dimensional spheres (with different radius and centers, possibly with infinite radius) passing through $y_\theta$ and $x_m$, with normal vector $p_1 - p_0$ at $x_m$. This geometric configuration shows directly the monotonicity of $S_\theta^+$, as in Theorem 5.2, and thus either for $\Omega = \Lambda = \mathbb{R}^n$ or for any $\Omega \subset \mathbb{R}^n$, and $\Lambda \subset \mathbb{R}^n$, $c$-convex with respect to $\Omega$, we have $\partial^c\phi = \partial\phi$ for all $c$-convex functions $\phi$ on $\Omega$ indexed by $\Lambda$. The same result can be obtained applying our general theory, especially Theorem 5.2 and Remark 6.1. Note that in this case the domains $\text{dom}(c\text{-}\text{Exp}_x)$, $\text{dom}(c^*\text{-}\text{Exp}_y)$ (for $\Omega = \Lambda = \mathbb{R}^n$) are $\mathbb{R}^n \setminus \{0\}$ and so

$$(q_\theta + W_\theta) \cap \text{dom}(c^*\text{-}\text{Exp}_y) = (q_\theta + W_\theta) \setminus \{0\},$$

(see (6.2) [6.3] for the definition of $W_\theta$). The set $(q_\theta + W_\theta) \setminus \{0\}$ is convex except when $\{0\} \subset q_\theta + W_\theta$. In the latter case, still a dense subset of it can be reached by line segments from $q_\theta$ within it, if the dimension $n \geq 3$, as we needed in Remark 6.1. The case $n = 1$ is trivial, and the case $n = 2$ can be handled by an approximation argument.

One may be concerned with that when $p_0, 0, p_1$ are collinear in the tangent space at $x_m$, the corresponding $c$-segment is no longer differentiable when $p_0$ passes through the origin 0 since the point $y_\theta$ in the $c$-segment goes to infinity. This problem can be bypassed by approximating the segment $[p_0, p_1]$ by a line segment...
which does not pass through the origin. Then our desired DASM survives under the approximation, and so it holds for \( \Omega = \Lambda = \mathbb{R}^n \).

**Example 7.4.** \( c(x, y) = \pm \frac{1}{p} |x - y|^p, p \neq 0 \) satisfies A3W for \( p = \pm 2 \) or \( p = -\frac{1}{2} \) (only) and A3S for \(-\frac{1}{2} < p < 1 \) (only). The case \( p = 2 \) is well known and the same conclusion as in Example 7.1 holds. For other \( p \), the cost is not differentiable for \( x = y \), but this case is similar to Example 7.3 and by the same approximation method using Remark 6.1 we get the same conclusion.

**Example 7.5.** (The round sphere) We consider the cost \( c(x, y) = \frac{1}{2} d^2(x, y) \) on \( S^n \), with \( d \) the Riemannian distance function of the standard round metric of \( S^n \), i.e. with the diameter \( \pi \). Loeper showed [L] that \( c \) satisfies A3S by directly computing the cost-sectional curvature \( \xi_x \). In this case, \( \text{dom}(c^*-\text{Exp}_{y_0}) = B_{\pi}(0) \subset T_{y_0}S^n \) (for \( \Omega = S^n \)) and so \( (q_\theta + W_\theta) \cap \text{dom}(c^*-\text{Exp}_{y_0}) \) is convex (see (6.2) for the definition of \( W_\theta \)) and we can apply Proposition 5.2 for those points where the function \( c(\cdot, y_\theta) \) is differentiable, to get **Strict DASM** on the set \( S^n \setminus \bigcup_{0 \leq \theta < 1} \hat{y}_\theta \), where we denote by \( \hat{y}_\theta \) the antipodal point of \( y_\theta \). Now, we use the continuity of the distance function to extend the **Strict DASM** to the whole \( S^n \). From this result together with Remark 5.2 we can conclude that if either \( \Omega = \Lambda = S^n \) or \( \Lambda \subset S^n \) is \( c \)-convex with respect to \( \Omega \subset S^n \) then \( \partial^c \phi = \partial \phi \) for all \( c \)-convex functions \( \phi \) on \( S^n \) indexed by \( \Lambda \).

**Example 7.6.** (Reflector antenna problem) The cost \( c(x, y) = -\frac{1}{2} \log |x - y|^2 \) restricted on the unit sphere \( S^{n-1} \subset \mathbb{R}^n \) satisfies A3S on \( S^{n-1} \times S^{n-1} \setminus \{x = y\} \). In this case, for \( \Omega = S^{n-1} \),
\[
\text{dom}(c^*-\text{Exp}_y) = T_y S^{n-1} \cong \mathbb{R}^{n-1}
\]
and \( c^*-\text{Exp}_y(T_y S^{n-1}) = S^{n-1} \setminus \{y\} \), \( c^*-\text{Exp}_y(0) = \hat{y} \), where we denote by \( \hat{y} \) the antipodal point of \( y \) in \( S^{n-1} \). So by our general theory (Theorem 5.2 and Remark 6.1) as applied in Example 7.3 we can make the conclusion that if either \( \Omega = \Lambda = S^n \) or \( \Lambda \subset S^n \) is \( c \)-convex with respect to \( \Omega \subset S^n \), then \( \partial^c \phi = \partial \phi \) for all \( c \)-convex functions \( \phi \) on \( S^n \) indexed by \( \Lambda \).

**References**

[AR] L. Ambrosio and S. Rigot, Optimal mass transportation in the Heisenberg group, J. Funct. Anal., to appear.

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

[Ca1] L.A. Caffarelli, The regularity of mapping with a convex potential. J. Amer. Math. Soc. 5 (1992), 99–104.

[Ca2] L.A. Caffarelli, Allocation maps with general cost functions, in ‘Partial Differential Equations and Applications’, (P. Marcellini, G. Talenti, and E. Vesintini eds.), pp29–35. Lecture Notes in Pure and Appl. Math., 177. Dekker, New York, 1996.

[Ca3] L.A. Caffarelli, Boundary regularity of maps with convex potentials II, Ann. Math. 144 (1996), 453–496.

[CaGH] L.A. Caffarelli, C. Gutierrez, and Q. Huang, On the regularity of reflector antennas, Annals of Math., to appear.
Department of Mathematics, University of Toronto, Toronto, Ontario Canada M5S 2E4

E-mail address: yhkim@math.toronto.edu, mccann@math.toronto.edu