COMPLETENESS AND MAXIMAL MONOTONICITY OF MULTI-CONJUGATE CONVEX FUNCTIONS ON THE LINE

TONGSEOK LIM

ABSTRACT. A cornerstone in convex analysis is the crucial relationship between functions and their convex conjugate via the Young-Fenchel inequality. In this bivariate case, the maximal monotonicity of the contact set $\{(x,y) \mid f(x)+f^*(y) = \langle x,y\rangle\}$ is due to the involution $f^{**} = f$ holding for convex lower-semicontinuous functions defined on a Hilbert space.

We introduce and investigate the validity of a generalized notion of involution in multivariate convex analysis. As a result, we show that when the underlying space is the real line, the generalized involution, or completeness of convex conjugation, holds true. This results in the maximality of the contact set in $\mathbb{R}^N$ induced by N multiconjugate convex functions. We conclude with a remark whose resolution will allow the results to be extended into multidimensional underlying spaces.

Keywords: multivariate convex analysis, c-monotonicity, c-convex conjugacy, maximal monotonicity, Minty Theorem, multi-marginal, optimal transport.
MSC2020 Classification: Primary 47H05, 26B25; Secondary 49N15, 49K30, 52A01, 91B68.

1. INTRODUCTION

Let $\mathcal{H}$ represent a real Hilbert space equipped with an inner product $\langle\ ,\ \rangle$ on which functions are defined. Let $|x| = \sqrt{\langle x,x\rangle}$ denote the norm on $\mathcal{H}$. $\langle\ ,\ \rangle$ shall denote the usual dot product when $\mathcal{H} = \mathbb{R}^n$.

Over the last few decades, the theory and applications of (bivariate) convex analysis and monotone operator theory have advanced significantly ([5, 28]),

Date: July 29, 2022.
 The author wishes to express gratitude to the Korea Institute of Advanced Study (KIAS) AI research group and the director Hyeon, Changbong for their hospitality and support during his stay at KIAS in 2022, where parts of this work were performed. TL would also like to thank Robert McCann for the fruitful discussions.

studying bivariate $x, y \in \mathcal{H}$, a function f on $\mathcal{H}$, and its convex conjugate

$$f^*(y) := \sup_{x \in \mathcal{H}} \langle x, y \rangle - f(x) \tag{1.1}$$

which are linked by the celebrated Young-Fenchel inequality

$$f(x) + f^*(y) \geq \langle x, y \rangle, \quad x, y \in \mathcal{H}. \tag{1.2}$$

In particular, it is well known that if f and g are convex conjugate to each other, i.e., $f^* = g$ and $g^* = f$, then the following contact set

$$\begin{aligned} \Gamma &= \{(x, y) \mid f(x) + g(y) = \langle x, y \rangle\} \\ &= \{(x, y) \mid y \in \partial f(x)\} \\ &= \{(x, y) \mid x \in \partial g(y)\} \end{aligned}$$

is *maximally monotone* (Rockafellar [27]), which property turns out to be, by Minty's theorem (see [5, Theorem 21.1]), equivalent to

$$S(\Gamma) = \mathcal{H} \tag{1.3}$$

(and not a proper subset of $\mathcal{H}$), where $S(\Gamma) := \{x + y \mid (x, y) \in \Gamma\}$.

Note that if $A \subseteq \mathcal{H} \times \mathcal{H}$ is *monotone*, i.e.,

$$\langle x - y, u - v \rangle \geq 0 \ \text{ for any } \ (x, u) \in A, (y, v) \in A,$$

then $A \cap (\Delta^\perp + p)$ is either empty or a singleton for every $p \in \Delta$, where $\Delta = \{(x, x) \mid x \in \mathcal{H}\}$ is the diagonal subspace of $\mathcal{H} \times \mathcal{H}$. Thus (1.3) is a maximality assertion, saying there is no "hole" in the monotone set Γ. These results established by Minty, Rockafellar and many others now form important foundation for modern theory of nonlinear monotone operators.

Recently, S. Bartz, H.H. Bauschke, H.M. Phan, and X. Wang [2] provided a significant extension of the convex analysis theory into the multivariate, or multi-marginal (i.e., variables more than two), situation, stating that "a comprehensive multi-marginal monotonicity and convex analysis theory is still missing." We denote $\mathbf{x} = (x_1, ..., x_N) \in \mathcal{H}^N$, and define $c : \mathcal{H}^N \to \mathbb{R}$ by

$$c(\mathbf{x}) = \sum_{1 \leq i < j \leq N} \langle x_i, x_j \rangle. \tag{1.4}$$

An analogous statement of the bi-conjugacy — $f^* = g$ and $g^* = f$ — in the multivariate setting can be given as follows.

Definition 1.1 (c-conjugate tuple [2]). *For each $1 \le i \le N$, let $f_i : \mathcal{H} \to]-\infty, +\infty]$ be a proper function, i.e., not entirely $+\infty$. We say that $(f_1, \dots, f_N)$ is a* c-conjugate tuple *if for each $1 \le i_0 \le N$ and $x_{i_0} \in \mathcal{H}$,*

$$f_{i_0}(x_{i_0}) = \Big(\bigoplus_{i \neq i_0} f_i \Big)^c (x_{i_0}) := \sup_{i \neq i_0,\, x_i \in \mathcal{H}} c(x_1, ..., x_{i_0}, ..., x_N) - \sum_{i \neq i_0} f_i(x_i).$$

Now a culmination of the results in [2] (see Theorem 2.5 and 4.3 in [2]) is the following theorem, where $A_{i_0} : \mathcal{H} \to 2^{\mathcal{H}}$ is the set-valued mapping

$$A_{i_0}(x_{i_0}) = \left\{ \sum_{i \neq i_0} x_i \;\middle|\; (x_1, ..., x_{i_0}, ..., x_N) \in \Gamma \right\},$$

and $\square$ denotes the infimal convolution: $(f \square g)(x) = \inf_{y \in H} \big(f(y) + g(x - y) \big)$.

Theorem 1.2 ([2]). *For $1 \le i \le N$, let $f_i : \mathcal{H} \to]-\infty, +\infty]$ be convex, lower semicontinuous, and proper, satisfying $\displaystyle\sum_{i=1}^N f_i(x_i) \ge c(\mathbf{x})$ for all $\mathbf{x} \in \mathcal{H}^N$. Let*

$$\Gamma = \left\{ \mathbf{x} \in \mathcal{H}^N \;\middle|\; \sum_{i=1}^N f_i(x_i) = c(\mathbf{x}) \right\} \text{ and } S(\Gamma) = \left\{ \sum_{i=1}^N x_i \;\middle|\; (x_1, ..., x_N) \in \Gamma \right\}.$$

Then the following assertions are equivalent:

(i) *There exist $1 \le i_0 \le N$ such that A_{i_0} is maximally monotone;*
(ii) *There exist $1 \le i_0 \le N$ such that $A_{i_0} = \partial f_{i_0}$;*
(iii) *$A_i = \partial f_i$ for each $1 \le i \le N$;*
(iv) *$\mathrm{Prox}_{f_1} + \cdots + \mathrm{Prox}_{f_N} = \mathrm{Id}$ where $\mathrm{Id}(x) = x$;*
(v) *$e_{f_1^*} + \cdots + e_{f_N^*} = q$ where $q(x) = \frac{1}{2}|x|^2$ and $e_f = f \square q$;*
(vi) *$\Gamma + \Delta^\perp = \mathcal{H}^N$;*
(vii) *$S(\Gamma) = \mathcal{H}$.*

In this case, $(f_1, \dots, f_N)$ is a c-conjugate tuple, and Γ determines $(f_1, \dots, f_N)$ uniquely up to an additive constant tuple $(\rho_1, \dots, \rho_N)$ such that $\sum_{i=1}^N \rho_i = 0$.

Here (iv) represents the partition of the identity into a sum of firmly non-expansive mappings, and (v) represents Moreau's decomposition of the quadratic function into envelopes in the multivariate settings, in which Prox_f and

e_f are the *proximal mapping* and the *Moreau envelope* of f respectively; see [2, 5] for more details. And $\Delta = \{(x, \dots, x) \,|\, x \in H\}$ is the diagonal subspace of $\mathcal{H}^N$. In addition, [2] shows Γ is maximally c-monotone (and, consequently, maximally c-cyclically monotone) if any of the assertions (i)–(vii) hold.

In the bivariate case $N = 2$, if f_1, f_2 are biconjugate ($f_1^* = f_2$ and $f_2^* = f_1$), then $S(\Gamma) = \mathcal{H}$, hence the statements (i)–(vii) hold. Thus it is natural to ask whether the multi-conjugacy of $(f_1, ..., f_N)$ conversely yields the assertions as well, in particular, (1.3). In this regard, [2] provides the following result.

Theorem 1.3 ([2])**.** *Let $n \in \mathbb{N}$, $N = 3$ and $\mathcal{H} = \mathbb{R}^n$. Let $g, h : \mathbb{R}^n \to]-\infty, +\infty]$ be proper, lower semicontinuous and convex functions. Suppose that $f = (g \oplus h)^c$ (in particular if (f, g, h) is a c-conjugate triple) and that f is essentially smooth. Let Γ be as in Theorem 1.2 generated by (f, g, h). Then assertions (i)–(vii) in Theorem 1.2 hold and Γ is maximally c-monotone.*

While Theorem 1.3 provides the first and affirmative answer toward the converse of Theorem 1.2, it is limited in three aspects: $\mathcal{H}$ must be Euclidean, N must be 3, and one of the c-conjugate convex functions needs to be smooth. This paper will prove the converse for any $N \in \mathbb{N}$ and without smoothness assumption on f_i's when the underlying space $\mathcal{H}$ is the real line. The author hopes that this will contribute to a better understanding of multi-conjugate convex functions and the maximal monotonicity of Γ (1.3) generated by them.

In order to better understand the fundamental relationship between multi-convex conjugacy and maximal monotonicity, let us go back to the case $N = 2$ where we observe that (1.3) is essentially a consequence of the following fundamental statement

$$f^{**} = f \text{ if and only if } f \text{ is convex lower-semicontinuous on } \mathcal{H}. \tag{1.5}$$

This inspires us to find and prove an analogous statement for the multivariate case $N \geq 3$. To this end, observe that if f and g are convex lower-semicontinuous functions, (1.5) is equivalent to the following statement

$$f^* = g \text{ implies } g^* = f. \tag{1.6}$$

(1.6) suggests that the multivariate version of (1.5) could be stated as follows.

Definition 1.4 (Completeness of convex conjugation)**.** *We say that a real Hilbert space $\mathcal{H}$ is* complete in convex conjugation *if for any $N \in \mathbb{N}$ and for any convex, lower semicontinuous and proper functions $\{f_i\}_{i=1,...,N}$ on $\mathcal{H}$, satisfying*

$$f_i(x_i) = \Big(\bigoplus_{j\neq i} f_j\Big)^c(x_i) \ \ \textit{for every} \ \ i = 2, ..., N, \tag{1.7}$$

it holds that $(f_1, \dots, f_N)$ is a c-conjugate tuple, that is, f_1 also satisfies

$$f_1(x_1) = \Big(\bigoplus_{j\neq 1} f_j\Big)^c(x_1). \tag{1.8}$$

In essence, this paper shows that the converse of Theorem 1.2 holds if $\mathcal{H}$ is complete in convex conjugation, and that the real line $\mathbb{R}$ is indeed complete; see Proposition 2.4, Theorem 2.6, and the remarks that follow.

1.1. **Connection with theory of multi-marginal optimal transport.** Let $(X_1, \mu_1), \dots, (X_N, \mu_N)$ be Borel probability spaces, and $X := X_1 \times \cdots \times X_N$. Denote $\Pi(X)$ by the set of all Borel probability measures π on X whose marginals are the μ_i's [29, 30]. Given a cost function $c : X \to \mathbb{R}$, the optimal transport problem refers to the following optimization problem:

$$P_c := \min_{\pi \in \Pi(X)} \int_X c(x) d\pi(x). \tag{1.9}$$

To distinguish it from the two-marginal case, the problem is commonly referred to as multi-marginal optimal transport when $N \geq 3$. In this problem, the optimal transport cost P_c, as well as the geometry and structure of optimal transport plans – the solutions to (1.9) – are sought.

Because (1.9) is an infinite-dimensional linear programming problem, it has a dual problem whose formulation turns out to have the following form:

$$D_c := \max_{\substack{f_i \in L_1(\mu_i), \\ \sum_{1\leq i\leq N} f_i(x_i) \leq c(\mathbf{x})}} \sum_{1\leq i\leq N} \int_{X_i} f_i(x_i) d\mu_i(x_i). \tag{1.10}$$

Kellerer's [13] generalization of the Kantorovich duality states that, under mild assumptions on the marginals $\mu_1, ..., \mu_N$ and cost function c, it holds

$$P_c = D_c. \tag{1.11}$$

This has the important implication that every optimal transport π solving (1.9) is concentrated on the contact set

$$\Gamma = \left\{ \mathbf{x} \in X \;\middle|\; \sum_{1 \le i \le N} f_i(x_i) = c(\mathbf{x}) \right\} \tag{1.12}$$

where $(f_1, \dots, f_N)$ is a solution to the dual problem (1.10). This provides a critical foundation for investigating the geometry of optimal transport plans.

The interaction between the optimal transport and its dual problems is what makes the theory surprisingly powerful for many applications in fields such as analysis, geometry, PDEs, probability, statistics, economics, data sciences, and many researchers have helped to advance the field [10, 12, 20, 25, 26, 29, 30]. Regarding the geometry of optimal transport for $N = 2$, arguably one of the most well-known and widely applied result is the Brenier's theorem [6]: given marginals $\mu_1, \mu_2 \in \mathcal{P}_2(\mathbb{R}^n)$ and cost function $c(x_1, x_2) = |x_1 - x_2|^2$, there is a convex function φ such that for any solution π to (1.9), it holds

$$y \in \partial\varphi(x) \quad \pi - a.e.\,(x, y), \text{ moreover, } y = \nabla\varphi(x) \text{ if } \mu_1 \text{ has density.} \tag{1.13}$$

Because the geometry of the subdifferential $\partial\varphi$ is well understood by studies in convex analysis, Brenier's theorem could yield important further results. Likewise, a better understanding of multivariate convex analysis should also have a significant impact on the theory of multi-marginal optimal transportations, their geometry, and applications. This is a motivation of this paper.

Recent advances in the theory of multi-marginal optimal transport and its geometrical structures have been rapid and fruitful, yielding a plethora of further research directions and open problems [1–4, 7–11, 14–16, 18, 19, 21–24]. In light of two-marginal optimal transport theory, Brenier's theorem and their consequences, it is clear that understanding the geometry of the contact set (1.12) is critical, much of which falls within the scope of the multivariate convex analysis. Given that most practical and relevant problems frequently involve real-valued random variables with marginal distributions on the real line, the author hopes that this paper will contribute to better understanding of the geometry of multi-marginal optimal transport in this case.

2. Results

In the sequel, lsc stands for "lower semi-continuous". And $q(x) := \frac{1}{2}|x|^2$.

Lemma 2.1. *Let f, g, h be proper functions, satisfying*

$$f(x) = \sup_{y \in \operatorname{dom} g} h(x+y) - g(y).$$

If h is lsc and λ-strongly convex, i.e., $h - \lambda q$ is convex, then so is f.

Proof. Let $k = h - \lambda q$, which is convex lsc. Hence $k = k^{**}$. We calculate

$$\begin{aligned}
f(x) &= \sup_{y \in \operatorname{dom} g} h(x+y) - g(y) \\
&= \sup_{y \in \operatorname{dom} g} \sup_{z \in \operatorname{dom} k^*} \lambda q(x+y) + \langle x+y, z\rangle - k^*(z) - g(y) \\
&= \lambda q(x) + \sup_{z \in \operatorname{dom} k^*} \left\{ \langle x, z\rangle - k^*(z) + \sup_{y \in \operatorname{dom} g} \{\langle y, \lambda x + z\rangle - (g - \lambda q)(y)\}\right\} \\
&= \lambda q(x) + \sup_{z \in \operatorname{dom} k^*} \langle x, z\rangle - k^*(z) + (g - \lambda q)^*(\lambda x + z) \\
&=: \lambda q(x) + \xi(x).
\end{aligned}$$

Observe that as a supremum of convex lsc functions, ξ is convex lsc. $\square$

Recall that $\Delta = \{(x, x, ..., x) \in \mathcal{H}^N\}$ denote the diagonal subspace of $\mathcal{H}^N$.

Definition 2.2 (Δ-convex envelope)**.** *Let $\mathbf{x} = (x_1, ..., x_N) \in \mathcal{H}^N$, $S(\mathbf{x}) = \sum_{i=1}^N x_i$. Let $f : \mathcal{H}^N \to\ ]-\infty, +\infty]$ be proper, satisfying $f(\mathbf{x}) \geq \langle S(\mathbf{x}), y\rangle + b$ for some $y \in \mathcal{H}$, $b \in \mathbb{R}$. Then g is called the* Δ-convex envelope *of f if g is the largest convex lsc function on $\mathcal{H}$ satisfying $f(\mathbf{x}) \geq g(S(\mathbf{x}))$.*

Lemma 2.3. *Let f satisfy the condition in Definition 2.2. Then g is the Δ-convex envelope of f if and only if $g^*(y) = \sup_{\mathbf{x}} \langle S(\mathbf{x}), y\rangle - f(\mathbf{x})$.*

Proof. Let h be any convex lsc function. The following equivalence

$$\begin{aligned}
& f(\mathbf{x}) \geq h(S(\mathbf{x})) \quad \text{for every } \mathbf{x} \in \mathcal{H}^N \\
\iff & f(\mathbf{x}) \geq \langle S(\mathbf{x}), y\rangle - h^*(y) \quad \text{for every } \mathbf{x} \in \mathcal{H}^N, y \in \mathcal{H} \\
\iff & h^*(y) \geq \sup_{\mathbf{x}} \langle S(\mathbf{x}), y\rangle - f(\mathbf{x}) \quad \text{for every } y \in \mathcal{H}
\end{aligned}$$

along with the fact that maximality of g corresponds to minimality of g^* yields the lemma. $\square$

Proposition 2.4. *Let f, g, h be convex, lower-semicontinuous and proper functions on $\mathbb{R}$. Assume that (f, g) are h-conjugate, that is*

$$f(x) = \sup_y \{h(x+y) - g(y)\}, \quad g(y) = \sup_x \{h(x+y) - f(x)\}. \tag{2.1}$$

Assume further that f and h are continuous, and h is λ-strongly convex for some $\lambda > 0$. Then h is the Δ-convex envelope of $f \oplus g(x, y) := f(x) + g(y)$.

Proof. By Lemma 2.1, f and g are λ- stongly convex. In particular, $\bigcup_{x\in\mathbb{R}} \partial f(x) = \bigcup_{y\in\mathbb{R}} \partial g(y) = \mathbb{R}$. Firstly, we claim that the following set $\mathcal{I}$ is dense in $\mathbb{R}$:

$$\begin{aligned} \mathcal{I} = \{s \in \mathbb{R} \mid s = x + y \text{ such that } \partial f(x) \cap \partial g(y) \neq \emptyset, \text{ and} \\ \text{either } f \text{ is differentiable at } x \text{ or } g \text{ is differentiable at } y\}. \end{aligned} \tag{2.2}$$

Let us prove the claim later. Let H denote the Δ-convex envelope of $f \oplus g$. Then $H \geq h$ since $f(x) + g(y) \geq h(x + y)$ by (2.1). The proposition asserts $H = h$. To prove this, we claim that it is sufficient to prove the following tightness: For any $x_0, y_0 \in \mathbb{R}$ such that $\partial f(x_0) \cap \partial g(y_0) \neq \emptyset$ and either f is differentiable at x_0 or g is differentiable at y_0, we have

$$f(x_0) + g(y_0) = h(x_0 + y_0). \tag{2.3}$$

The sufficiency is because (2.3) implies $H = h$ on $\mathcal{I}$, and thus for any $s \in \mathbb{R}$, by the first claim, there exists a sequence s_n in $\mathcal{I}$ such that $\lim s_n = s$, and

$$H(s) \leq \liminf H(s_n) = \liminf h(s_n) = h(s)$$

as desired, thanks to the continuity of h.

Now to verify (2.3), by translation, we may assume without loss of generality that $x_0 = y_0 = 0$. Moreover we may assume that $0 \in \partial f(0) \cap \partial g(0)$. To see why this can be assumed, let $a \in \partial f(0) \cap \partial g(0)$. Consider $\tilde{f}(x) = f(x) - \langle a, x\rangle - f(0)$, $\tilde{g}(y) = g(y) - \langle a, y\rangle - g(0)$, and $\tilde{h}(z) = h(z) - \langle a, z\rangle - f(0) - g(0)$. Then f and g are h-conjugate if and only if $\tilde{f}$ and $\tilde{g}$ are $\tilde{h}$-conjugate. And since $\min \tilde{f} = \tilde{f}(0) = 0$ and $\min \tilde{g} = \tilde{g}(0) = 0$, (2.3) holds if and only if $\tilde{h}(0) = 0$. Our discussion so far indicates that it is sufficient to prove the claim (2.3) under the assumption that $x_0 = y_0 = 0$, f and g are h-conjugate, $f(0) = 0 = \min f$, $g(0) = 0 = \min g$, and either f or g is differentiable at 0; and the goal is to show that $h(0) = 0$.

Now to derive a contradiction, suppose $h(0) < 0$, so that $m = \min h < 0$. Assume f is differentiable at 0 (the proof will be the same in the case g is differentiable at 0, by switching the role of f and g). Let $K = \{x \mid h(x) \leq m/2\}$. Since h is strongly convex, there exists $\delta > 0$ such that

$$x \in \mathbb{R} \setminus K \text{ and } z \in \partial h(x) \text{ implies } |z| \geq \delta. \tag{2.4}$$

By (2.1), given $\epsilon > 0$, there exists y_ϵ such that $-\epsilon < h(y_\epsilon) - g(y_\epsilon) \leq 0 = f(0)$. Now $f(x) \geq h(x + y_\epsilon) - g(y_\epsilon)$ for all x and $\nabla f(0) = 0$ imply that for all sufficiently small ϵ, we must have $y_\epsilon \in K$ by (2.4). However, we then have

$$-\epsilon < h(y_\epsilon) - g(y_\epsilon) \leq h(y_\epsilon) \leq m/2,$$

a contradiction for small ϵ.

It remains to prove the denseness of $\mathcal{I}$. For any $s \in \mathbb{R}$, recall that there exists $x, y \in \mathbb{R}$ such that $s \in \partial f(x) \cap \partial g(y)$. By translation and subtracting affine functions as before, we may assume $0 \in \partial f(0) \cap \partial g(0)$. Then notice the desired denseness will follow if we can show that for any $r > 0$, there exists $x, y \in \mathbb{R}$ such that $|x| < r$, $|y| < r$, $\partial f(x) \cap \partial g(y) \neq \emptyset$, and either f is differentiable at x or g is differentiable at y.

Now to show the claim, assume that neither f nor g is differentiable at 0, since otherwise there is nothing to prove. $\partial f(0)$ is a compact interval, say $[a, b]$, since f is continuous. And $W := \partial g((-r, r))$ is an open interval since g is strongly convex, thus g^* is differentiable and $\partial g^* = (\partial g)^{-1}$ is continuous.

Suppose $b \in W$. Then for any $\epsilon > 0$, there exists $x \in [0, \epsilon)$ such that f is differentiable at x and $\nabla f(x) \in [b, b + \epsilon)$ since f is differentiable a.e.. Thus $\nabla f(x) \in W$ for small ϵ, and this implies $\nabla f(x) \in \partial g(y)$ for some $y \in (-r, r)$, proving the claim. Likewise, the claim holds in the case $a \in W$. Finally, if $\{a, b\} \cap W = \emptyset$, then $W \subseteq (a, b)$, yielding that g is Lipschitz in $(-r, r)$. By the same argument using g, we find that $\nabla g(y) \in \partial f(x)$ for some $x, y \in (-r, r)$. This proves the desired denseness of $\mathcal{I}$ in $\mathbb{R}$, hence the proposition. $\square$

Remark 2.5. *The above proof shows that Proposition 2.4 will continue to hold for functions defined on a general Hilbert space $\mathcal{H}$ if the corresponding set $\mathcal{I}$ is dense in $\mathcal{H}$ given the mutual conjugacy* (2.1).

Theorem 2.6. *Let $c : \mathbb{R}^N \to \mathbb{R}$ be given by (1.4). For $1 \le i \le N$, let f_i be convex, lower semicontinuous and proper functions defined on $\mathbb{R}$, satisfying*

$$f_i(x_i) = \Big(\bigoplus_{j \neq i} f_j\Big)^c(x_i) \quad \text{for every} \quad i = 2, ..., N. \tag{2.5}$$

Then $(f_1, \dots, f_N)$ is a c-conjugate tuple, that is, f_1 also satisfies

$$f_1(x_1) = \Big(\bigoplus_{j \neq 1} f_j\Big)^c(x_1). \tag{2.6}$$

In this case, the contact set

$$\Gamma = \Big\{ \mathbf{x} \in \mathbb{R}^N \;\Big|\; \sum_{j=1}^N f_j(x_j) = c(\mathbf{x}) \Big\}$$

is maximally c-monotone, in the sense that

$$S(\Gamma) = \mathbb{R}. \tag{2.7}$$

In particular, the assertions (i)–(vii) in Theorem 1.2 hold true if $\mathcal{H} = \mathbb{R}$.

Proof. If $N = 2$, (2.6) is equivalent to the well-known statement $f^{**} = f$ if and only if f is convex lsc. And (2.7) is also well known; Rockafellar [27] showed the subdifferential of a lower semicontinuous proper convex function is a maximal monotone operator, and Minty showed (see [5, Theorem 21.1]) an operator A is maximally monotone if and only if $\mathrm{ran}(\mathrm{Id} + A) = \mathcal{H}$.

When $\mathcal{H} = \mathbb{R}$, we will first extend these results (2.6), (2.7) for the case $N = 3$. However, since most of the proof presented below will be valid for general Hilbert space $\mathcal{H}$, in the sequel, we will denote $\mathcal{H}$ (rather than $\mathbb{R}$) by the underlying space, though we will assume $\mathcal{H} = \mathbb{R}$ (see also Remark 2.7).

To begin, we have $f(x) + g(y) + h(z) \ge \langle x, y\rangle + \langle y, z\rangle + \langle z, x\rangle$, hence

$$\varphi(x, y) := f(x) + g(y) - \langle x, y\rangle \ge h^*(x + y). \tag{2.8}$$

We will show h^* is the Δ-convex envelope of φ. Notice (2.8) is equivalent to

$$F(x) + G(y) \ge H(x + y) \tag{2.9}$$

where $F = f + q$, $G = g + q$, $H = h^* + q$, and $q(s) = \frac{1}{2}|s|^2$. Observe that the conjugacy assumption (2.5), which reads

$$f(x) = \sup_y \{h^*(x+y) + \langle x, y\rangle - g(y)\},\ \ g(y) = \sup_x \{h^*(x+y) + \langle x, y\rangle - f(x)\}$$

implies that F and G are also H-conjugate, that is

$$F(x) = \sup_y \{H(x+y) - G(y)\} \quad \text{and} \quad G(y) = \sup_x \{H(x+y) - F(x)\}.$$

We claim that $H(x+y)$ is the Δ-convex envelope of $F(x)+G(y)$. Observe that this implies (2.6) for $N = 3$, because if $H(x+y)$ is the Δ-convex envelope of $F(x)+G(y)$, then $h^*(x+y)$ must be the Δ-convex envelope of $\varphi(x, y)$ due to the equivalence of (2.8) and (2.9). But this precisely means that $h = (f \oplus g)^c$ by Lemma 2.3, as desired.

Then as argued in the proof of Proposition 2.4, it is sufficient to prove the claim under the assumption that F and G are H-conjugate, $F(0) = 0 = \min F$, $G(0) = 0 = \min G$, and the goal is to show $H(0) = 0 = \min H$.

To this end, for each $R > 0$, define

$$\begin{aligned}
G_R(y) &= G(y) \ \text{ if } \ |y| \le R, \quad G_R(y) = +\infty \ \text{ if } \ |y| > R,\\
h_R^*(x) &= \sup_{|y|\le R} \langle x, y\rangle - h(y), \quad H_R = h_R^* + q,\\
F_R(x) &= \sup_y \{H_R(x+y) - G_R(y)\},\\
F_R^H(y) &= \sup_x \{H_R(x+y) - F_R(x)\},\\
K_R(y) &= \sup_x \{H_R(x+y) - F(x)\}.
\end{aligned}$$

Assume that R is large enough such that $\min G_R = \min G$. We observe

(2.10) h_R^* has Lipschitz constant at most R,

(2.11) F_R is locally Lipschitz and monotone increasing to F as $R \to \infty$,

(2.12) F_R^H converges pointwise to G.

(2.11) is because H_R and $-G_R$ monotonically increase in R, and F_R is locally Lipschitz because the supremum defining F_R is taken over $|y| \le R$ only. And (2.12) is because $G_R \ge F_R^H \ge K_R$, G_R decreases to G, and K_R increases to G

in R. Now since F_R and F_R^H are H_R-conjugate, by Proposition 2.4, we have

$$H_R(x+y) \text{ is the } \Delta\text{-convex envelope of } F_R(x) + F_R^H(y). \tag{2.13}$$

Recall that our goal is to show $\min H = \min F + \min G$. It is clear that $\min H \le \min F + \min G$. To show the reverse, since $H_R \le H$ and $\min H_R = \min F_R + \min F_R^H$ by (2.13), and also $G_R \ge F_R^H \ge K_R$, it is enough to show

$$\min F_R \nearrow \min F \quad \text{and} \quad \min K_R \nearrow \min G \quad \text{as} \quad R \to \infty. \tag{2.14}$$

$\min F = F(0) = 0$ and the fact that F is of the form $F = f+q$ implies $F \ge q$, thus $F^* \le q$. Since F_R is increasing to F, F_R^* is decreasing in R bounded below by F^*. Let F_∞ denote the limit of F_R^*. We claim that $F_\infty = F^*$. To see this, note that F_∞ is convex as a limit of convex functions, and it is real-valued, i.e., $\operatorname{dom} F_\infty = \mathcal{H}$. Meanwhile, $F^* = (F_\infty)^{**}$ by [5, Proposition 13.47]. Now if $\mathcal{H}$ is Euclidean, i.e., $\mathcal{H} = \mathbb{R}^n$, then F_∞ is continuous, yielding $F_\infty = (F_\infty)^{**} = F^*$ as claimed. This implies in particular,

$$\min F = -F^*(0) = -F_\infty(0) = -\lim_{R\to\infty} F_R^*(0) = \lim_{R\to\infty} \min F_R.$$

Similarly, $\min K_R \nearrow \min G$. This proves $\min H = H(0) = 0$, and thus (2.6).

Now to prove (2.7), fix any $s \in \mathcal{H}$. By (2.7) holding for $N = 2$, there exists $z \in \mathcal{H}$ such that $h(z) + h^*(s-z) = \langle z, s-z \rangle$. This yields $z \in \partial h^*(s-z)$, implying $s \in \partial H(s-z)$. Since F, G, H are of the form $F = f+q$, $G = g+q$, $H = h^* + q$, there exist unique $x, y, u \in \mathcal{H}$ such that $s \in \partial F(x) \cap \partial G(y) \cap \partial H(u)$. Our previous proof yields $F(x)+G(y) = H(x+y)$ and $s \in \partial H(x+y)$. Then the uniqueness of u implies $s - z = x + y$, or $s = x+y+z$. Finally,

$$\begin{aligned}
&F(x) + G(y) = H(x+y) \\
\iff &f(x) + g(y) = h^*(x+y) + \langle x, y \rangle \\
\iff &f(x) + g(y) = h^*(s-z) + \langle x, y \rangle \\
\iff &f(x) + g(y) = \langle z, s-z \rangle - h(z) + \langle x, y \rangle \\
\iff &f(x) + g(y) + h(z) = \langle x, y \rangle + \langle y, z \rangle + \langle z, x \rangle.
\end{aligned}$$

This proves the maximality (2.7), hence the theorem for $N = 3$.

From now on, we extend the proof for $N \geq 4$. We proceed by an induction on N. Suppose the theorem holds for $N-1$. Define

$$\tilde{\mathbf{x}} = (x_3, ..., x_N), \quad S(\tilde{\mathbf{x}}) = \sum_{i=3}^{N} x_i, \quad c(\tilde{\mathbf{x}}) = \sum_{3 \leq i < j \leq N} \langle x_i, x_j \rangle,$$

$$\psi(\tilde{\mathbf{x}}) = \sum_{i=3}^{N} f_i(x_i) - c(\tilde{\mathbf{x}}), \quad \varphi(x_1, x_2) = f_1(x_1) + f_2(x_2) - \langle x_1, x_2 \rangle,$$

$$g = (f_1 \oplus f_2)^c, \text{ i.e., } g(y) = \sup_{x_1, x_2} \big\{ \langle y, x_1 + x_2 \rangle - \varphi(x_1, x_2) \big\}.$$

From the inequality $\sum_{i=1}^{N} f_i(x_i) \geq c(\mathbf{x})$, we have

$$\psi(\tilde{\mathbf{x}}) \geq g\big(S(\tilde{\mathbf{x}})\big) = \sup_{x_1, x_2} \big\{ \langle S(\tilde{\mathbf{x}}), x_1 + x_2 \rangle - \varphi(x_1, x_2) \big\}. \tag{2.15}$$

Now comes the crux of the observation: (2.15), the induction hypothesis, and the conjugacy (2.5) (i.e., each of the $f_3, ..., f_N$ is the smallest convex function satisfying (2.15) given others) combine to imply that

$$(f_3, ..., f_N, g^*) \text{ are } c(\tilde{\mathbf{x}}, y) - \text{conjugate, where } c(\tilde{\mathbf{x}}, y) = c(\tilde{\mathbf{x}}) + \langle S(\tilde{\mathbf{x}}), y \rangle.$$

In other words, $g\big(S(\tilde{\mathbf{x}})\big)$ is the Δ-convex envelope of $\psi(\tilde{\mathbf{x}})$, by Lemma 2.3. This in turn implies $f_2 = (f_1 \oplus g)^c$, because

$$\begin{aligned}
f_2(x_2) &= \Big(\bigoplus_{j \neq 2} f_j \Big)^c (x_2) \\
&= \sup_{x_1, \tilde{\mathbf{x}}} \big\{ \langle x_1, x_2 \rangle + \langle x_1 + x_2, S(\tilde{\mathbf{x}}) \rangle - f_1(x_1) - \psi(\tilde{\mathbf{x}}) \big\} \\
&= \sup_{x_1} \big\{ \langle x_1, x_2 \rangle - f_1(x_1) + \sup_{\tilde{\mathbf{x}}} \{ \langle x_1 + x_2, S(\tilde{\mathbf{x}}) \rangle - \psi(\tilde{\mathbf{x}}) \} \big\} \\
&= \sup_{x_1, y} \big\{ \langle x_1, x_2 \rangle - f_1(x_1) + \langle x_1 + x_2, y \rangle - g(y) \big\} \\
&= \sup_{x_1, y} \big\{ \langle x_2, x_1 + y \rangle - (f_1(x_1) + g(y) - \langle x_1, y \rangle) \big\} \\
&= (f_1 \oplus g)^c (x_2)
\end{aligned}$$

where the fourth equality is because g is the Δ-convex envelope of ψ. The induction hypothesis (or the theorem we established for $N = 3$) now implies

$$f_1 = (f_2 \oplus g)^c = \Big(\bigoplus_{j \neq 1} f_j \Big)^c (x_1)$$

by an analogous calculation given above. This completes the proof of (2.6).

Finally, we establish the maximality $S(\Gamma) = \mathcal{H}$. Recall (2.15), that is

$$\psi(\tilde{\mathbf{x}}) \geq g(S(\tilde{\mathbf{x}})) \geq \langle y, S(\tilde{\mathbf{x}})\rangle - g^*(y) \quad \text{for every } x_3, ..., x_N, y.$$

Fix any $s \in \mathcal{H}$. By the conjugacy of $(f_3, ..., f_N, g^*)$ and the induction hypothesis on the maximality, there exists $\tilde{\mathbf{x}}^s = (\tilde{x}_3^s, ..., \tilde{x}_N^s)$ such that

$$\psi(\tilde{\mathbf{x}}^s) = g(S(\tilde{\mathbf{x}}^s)) = \langle s - S(\tilde{\mathbf{x}}^s), S(\tilde{\mathbf{x}}^s)\rangle - g^*(s - S(\tilde{\mathbf{x}}^s)). \tag{2.16}$$

Similarly, the conjugacy of (f_1, f_2, g) yields

$$\varphi(x_1, x_2) \geq g^*(x_1 + x_2) \geq \langle z, x_1 + x_2\rangle - g(z) \quad \text{for every } x_1, x_2, z,$$

and for the same s, there exists $x_1^s, x_2^s \in \mathcal{H}$ such that

$$\varphi(x_1^s, x_2^s) = g^*(x_1^s + x_2^s) = \langle s - x_1^s - x_2^s, x_1^s + x_2^s\rangle - g(s - x_1^s - x_2^s). \tag{2.17}$$

However, the pair $u, v \in \mathcal{H}$ that satisfies $u + v = s$ and $g(u) + g^*(v) = \langle u, v\rangle$ is unique. Hence (2.16), (2.17) implies $S(\tilde{\mathbf{x}}^s) = s - x_1^s - x_2^s$, or $s = \sum_{i=1}^N x_i^s$. From this identity, by adding the identities (2.16), (2.17), we obtain

$$\begin{aligned} &\varphi(x_1^s, x_2^s) + \psi(\tilde{\mathbf{x}}^s) = \langle S(\tilde{\mathbf{x}}^s), x_1^s + x_2^s\rangle \\ \iff &(x_1^s, ..., x_N^s) \in \Gamma = \Big\{ \mathbf{x} \ \Big| \ \sum_{j=1}^N f_j(x_j) = c(\mathbf{x}) \Big\}. \end{aligned}$$

This completes the proof of the theorem. □

Remark 2.7. *As demonstrated by its proof, Theorem 2.6 holds true for a general Hilbert space domain $\mathcal{H}$ if two points can also be verified on the same domain: the first and most important is Proposition 2.4. The second is the lower-semicontinuity of the function $F_\infty = \lim_{R\to\infty} F_R^*$, which was shown to be continuous when $\mathcal{H} = \mathbb{R}^n$ in the proof of Theorem 2.6. Thus, if both claims are shown to be true for $\mathcal{H}$, then Theorem 2.6 is also true for $\mathcal{H}$. Alternatively, if Proposition 2.4 is verified for $\mathbb{R}^n$, Theorem 2.6 is also verified for $\mathbb{R}^n$. We also demonstrated that Proposition 2.4 holds true if the set $\mathcal{I}$ in* (2.2) *is shown to be dense in $\mathcal{H}$, for which the conjugacy condition* (2.1) *must be critical.*

References

[1] M. Agueh and G. Carlier, Barycenters in the Wasserstein Space, *SIAM Journal on Mathematical Analysis* 43 (2011), 904–924.

[2] S. Bartz, H.H. Bauschke, H.M. Phan and X. Wang, Multi-marginal maximal monotonicity and convex analysis, *Mathematical programming* 185 (2021), 385–408.

[3] S. Bartz, H.H. Bauschke and X. Wang, The resolvent order: a unification of the orders by Zarantonello, by Loewner, and by Moreau, *SIAM Journal on Optimization* 27 (2017), 466–477.

[4] S. Bartz, H.H. Bauschke and X. Wang, A class of multi-marginal c-cyclically monotone sets with explicit c-splitting potentials, *Journal of Mathematical Analysis and Applications* 461 (2018), 333–348.

[5] H.H. Bauschke and P.L. Combettes, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*, 2nd edition, Springer, 2017.

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

[7] G. Carlier, On a class of multidimensional optimal transportation problems, *Journal of Convex Analysis* 10 (2003), 517–529.

[8] G. Carlier and B. Nazaret, Optimal transportation for the determinant, *ESAIM: Control, Optimisation and Calculus of Variations* 14 (2008), 678–698.

[9] C. Cotar, G. Friesecke and C. Klüppelberg, Density functional theory and optimal transportation with Coulomb cost, *Commun. Pure Appl. Math.* 66, No. 4 (2013) 548–599.

[10] W. Gangbo and R.J. McCann, The geometry of optimal transportation, *Acta Mathematica* 177 (1996), 113–161.

[11] W. Gangbo and A. Swiech, Optimal maps for the multidimensional Monge-Kantorovich problem, *Communications on Pure and Applied Mathematics* 51 (1998), 23–45.

[12] N. Ghoussoub and A. Moameni, Symmetric Monge-Kantorovich problems and polar decompositions of vector fields, *Geometric and Functional Analysis* 24 (2014), 1129–1166.

[13] H.G. Kellerer, Duality theorems for marginal problems, *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete* 67 (1984), 399–432.

[14] Y.-H. Kim and B. Pass, A general condition for Monge solutions in the multi-marginal optimal transport problem, *SIAM Journal on Mathematical Analysis* 46 (2014), 1538–1550.

[15] Y.-H. Kim and B. Pass, Multi-marginal optimal transport on Riemannian manifolds, *Amer. J. Math.*, 137 (2015) 1045–1060.

[16] Y.-H. Kim and B. Pass, Wasserstein Barycenters over Riemannian manifolds, *Adv. Math.*, 307 (2017) 640–683.

[17] M. Knott and C.S. Smith, On the optimal mapping of distributions, *Journal of Optimization Theory and Applications* 43 (1984), 39–49.

[18] S. Di Marino, L. De Pascale and M. Colombo, Multimarginal optimal transport maps for 1-dimensional repulsive costs, *Canadian Journal of Mathematics* 67 (2015), 350–368.

[19] S. Di Marino, A. Gerolin and L. Nenna, Optimal transportation theory with repulsive costs, *Topological Optimization and Optimal Transport: In the Applied Sciences* 9 (2017), 204–256.

[20] R.J. McCann, A convexity principle for interacting gases, *Adv. Math.* 128, (1997) 153–179.

[21] L. Nenna, Numerical methods for multi-marginal optimal transportation, *PhD diss., Université Paris sciences et lettres*, 2016.

[22] B. Pass, On the local structure of optimal measures in the multi-marginal optimal transportation problem, *Calculus of Variations and Partial Differential Equations* 43 (2012), 529–536.

[23] B. Pass, Multi-marginal optimal transport: theory and applications, *ESAIM: Mathematical Modelling and Numerical Analysis* 49 (2015), 1771–1790.

[24] B. Pass and A. Vargas-Jiménez, Multi-marginal optimal transportation problem for cyclic costs, *SIAM J. Math. Anal.*, 53 (2021) 4386–4400.

[25] S.T. Rachev and L. Rüschendorf, *Mass transportation problems. Vol. I. Theory*, Probability and its Applications (New York). Springer-Verlag, New York, 1998.

[26] S.T. Rachev and L. Rüschendorf, *Mass transportation problems. Vol. II. Applications*, Probability and its Applications (New York). Springer-Verlag, New York, 1998.

[27] R.T. Rockafellar, On the maximal monotonicity of subdifferential mappings, *Pacific Journal of Mathematics* 33(1) (1970) 209–216.

[28] R.T. Rockafellar and R.J-B. Wets, *Variational Analysis*, Springer, 1998.

[29] F. Santambrogio, *Optimal Transport for Applied Mathematicians*, Birkhäuser, 2015.

[30] C. Villani, *Optimal Transport: Old and New*, Springer, 2009.

Tongseok Lim: Krannert School of Management
Purdue University, West Lafayette, Indiana 47907, USA
Email address: lim336@purdue.edu