WEAK FORMULATION OF THE MTW CONDITION AND CONVEXITY PROPERTIES OF POTENTIALS

GRÉGOIRE LOEPER¹, NEIL TRUDINGER²

¹Monash University, School of Mathematical Sciences
²The Australian National University, Mathematical Sciences Institute

Abstract. We simplify the geometric interpretation of the weak Ma-Trudinger-Wang condition for regularity in optimal transport and provide a geometric proof of the global c-convexity of locally c-convex potentials when the cost function c is only assumed twice differentiable.

1. Introduction

We consider a cost function $c$ defined on the product $\Omega \times \Omega^*$ of two domains $\Omega, \Omega^*$ in Euclidean space $\mathbb{R}^n$. For a mapping $\phi : \Omega \rightarrow \mathbb{R}$ we define its c-transform $\phi^c : \Omega^* \rightarrow \mathbb{R}$ by

$$\forall y \in \Omega^*, \phi^c(y) = \sup_{x \in \Omega} \{-\phi(x) - c(x, y)\}.$$  

Conversely we define the $c^*$-transform of $\psi : \Omega^* \rightarrow \mathbb{R}$. A c-convex potential has at every point $x \in \Omega$ a c-support, i.e., there exists $y \in \Omega^*, \psi = \psi(y) \in \mathbb{R}$ such that

$$\forall x' \in \Omega, \phi(x') \geq -\psi(y) - c(x', y),$$

with equality at $x' = x$. It follows from this definition that

$$\phi(x) = \sup_{y \in \Omega^*} \{-\psi(y) - c(x, y)\}$$

and that $\phi$ can be obtained as the $c^*$ transform of $\psi : \Omega^* \rightarrow \mathbb{R}$. It then turns out that $\psi = \phi^c$. For $\phi$ a c-convex potential, and $\phi^c$ its c-transform, we define as in [2] the contact set as a set valued map $G_{\phi}$ given by

$$G_{\phi}(y) = \{x : \phi(x) + \phi^c(y) = -c(x, y)\}.$$  

for $y \in \Omega^*$. We will also use the notions of c-segment, c-convexity of domains. Whenever needed, we will refer to the conditions $A1, A2, A3, A3w$ that have been introduced in [5, 6]. One of the main features of this paper is that we will assume throughout that the cost function

E-mail address: gregoire.loeper@monash.edu.
Date: July 7, 2020.
c is globally $C^2(\Omega \times \Omega^*)$, without any further explicit smoothness hypotheses. As usual we will use subscripts to denote partial derivatives of c with respect to variables $x \in \Omega$ and subscripts preceded by a comma to denote partial derivatives with respect to $y \in \Omega^*$, so that in particular $c_x, c_{i,y}, c_{i,j}, c_{i,j}$ denote the partial derivatives of c with respect to $x, x_i, y, y_j, x_i y_j$. We also use $c^{i,j} = [c_{i,j}]$ to denote the inverse of the matrix $c_{x,y} = [c_{i,j}]$. We further assume throughout the paper that c satisfies the assumptions $A_1, A_2$ of [5], that is for all $x \in \Omega$ the mapping $y \rightarrow -c_x(x,y)$ is injective, that the dual counterpart holds and the matrix $c_{x,y}$ is not singular. We also introduce what will be a weak form of assumption $A_3w$:

**Definition 1.1.** The cost function satisfies $A_3v$ if: for all $x, x_1 \in \Omega$ and $y_0, y_1 \in \Omega^*$, for all $\theta \in (0, 1)$, with

$$c_x(x, y_0) = \theta c_x(x, y_1) + (1 - \theta)c_x(x, y_0),$$

there holds

$$\max\{-c(x, y_0) + c(x_0, y_0), -c(x, y_1) + c(x_0, y_1)\} \geq -c(x, y_0) + c(x_0, y_0) + o(|x - x_0|^2),$$

where the term $o(|x - x_0|^2)$ may depend on $\theta$.

From [2] it is known that when the cost function is $C^4$, $A_3v$ is equivalent to $A_3w$.

Our main result is the following:

**Theorem 1.2.** Let $c : \Omega \times \Omega^* \rightarrow \mathbb{R}$ be a $C^2$ cost-function satisfying $A_1, A_2$ with $\Omega, \Omega^*$ $c$-convex with respect to each other. Assume that

(i) $c$ satisfies $A_3v$.

Then

(ii) for all $y_0, y_1 \in \Omega^*$, $\sigma \in \mathbb{R}$, the set $U = \{x \in \Omega : c(x, y_0) - c(x, y_1) \leq \sigma\}$ is $c$-convex with respect to $y_0$.

(iii) for all $\phi$ $c$-convex, $x \in \Omega$, $y \in \Omega^*$, the contact set $G_\phi(y)$ and its dual $G_{\phi'}(x)$ are connected,

(iv) any locally $c$-convex function in $\Omega$ is globally $c$-convex.

**Remark.** The novelty of the result lies in the way it is obtained; at no point do we have to differentiate the cost function $c$. Hence the computations from previous proofs [1, 4, 8], in the case when $c \in C^4$, do not have to be reproduced. The proof will be based on a purely geometric interpretation of condition $A_3v$.

**2. Proof of Theorem 1.2**

In what follows we will use the term $c$–exponential ($c$-exp), as in [2], to denote the mapping in condition $A_1$, that is

$$y = c$-exp_x(p) \iff -c_x(x, y) = p.$$
We recall also that
\[ D_p(\text{c-exp}_x) = -c^{x,y}. \]

The core of the proof lies in the following two lemmas,

**Lemma 2.1 (c-hyperplane lemma).** Let \( x_0 \in \Omega, y_0, y_1 \in \Omega^* \) and let \( y_\theta = \text{c-exp}_{x_0} p_\theta \) where \( p_\theta = (1 - \theta)c_x(x_0, y_0) + \theta c_x(x_0, y_1), \ 0 \leq \theta \leq 1 \), denote a point on the c-segment from \( y_0 \) to \( y_1 \), with respect to \( x_0 \).

Consider
\[ S_\theta = \{ x \in \Omega : c(x, y_0) - c(x_0, y_0) \leq c(x, y_\theta) - c(x_0, y_\theta) \} \]
Then as \( \theta \) approaches 0, \( \partial S_\theta \cap \Omega \) converges to \( H_0 \), the \( c^* \)-hyperplane with respect to \( y_0 \), passing through \( x_0 \), with \( c \)-normal vector \( p_1 - p_0 \), given by
\[ H_0 = \{ x \in \Omega : -c^{x,y}(x_0, y_0)(p_1 - p_0) \cdot [c_y(x, y_0) - c_y(x_0, y_0)] = 0 \} \]

**Proof.** Locally around \( \theta = 0 \), the equation of \( \partial S_\theta \) reads
\[ [c_y(x, y_0) - c_y(x_0, y_0)] \cdot (y_\theta - y_0) = o(\theta). \]
Passing to the limit as \( \theta \) goes to 0, we obtain
\[ [c_y(x, y_0) - c_y(x_0, y_0)] \cdot \partial_\theta y_\theta = 0, \]
which gives the desired result, since
\[ \partial_\theta y_\theta = -c^{x,y}(x_0, y_0)(p_1 - p_0). \]

**Remark.** We call \( H_0 \) a c-hyperplane with respect to \( y_0 \) because if we express \( x \) as \( c^*-\exp_{y_0}(q) \) then
\[ H_0 = c^*-\exp_{y_0}(\tilde{H}_0), \]
or equivalently
\[ \tilde{H}_0 = -c_y(\cdot, y_0)(H_0), \]
where
\[ \tilde{H}_0 = \{ q \in c_y(\cdot, y_0)(\Omega) : c^{x,y}(x_0, y_0)(p_1 - p_0) \cdot (q - q_0) = 0 \}, \quad q_0 = -c_y(x_0, y_0) \]
Therefore, \( H_0 \) is the image by \( c^*-\exp_{y_0} \) of a hyperplane.

**Remark.** We will define in the same way the section \( S_\theta, \theta' \) for \( \theta' \in (\theta, 1) \) and the \( c^* \)-hyperplane \( H_\theta \).

The following lemma is then the second main ingredient of the proof: it says that the \( c \)-convexity of \( S_\theta \) is non-decreasing with respect to \( \theta \); (note that the previous lemma asserts that the \( c \)-convexity of \( S_\theta \) vanishes at \( \theta = 0 \)).

**Lemma 2.2.** Assume that \( c \) satisfies A3v. Then the second fundamental form of \( S_\theta \) at \( x_0 \) is non-decreasing with respect to \( \theta \), for \( \theta \) in \((0, 1]\).
Proof. Consider
\[ h_\theta = c(x, y_\theta) - c(x, y_0) - c(x_0, y_0) + c(x_0, y_\theta). \]
Note that \( h_\theta \) is a defining function for \( S_\theta \) in the sense that \( S_\theta = \{ x \in \Omega : h_\theta \leq 0 \} \).

Note also that at \( x = x_0 \) we have \( h_\theta(x_0) = 0 \) for all \( \theta \) and the set
\[ \{ \partial_x h_\theta | x = x_0, \theta \in [0, 1] \} \]
is a line. Therefore all the sets \( \partial S_\theta \) contain \( x_0 \) and have the same unit normal at \( x_0 \).

Then we note that property \( A_3v \) is equivalent to the following: locally around \( x_0 \) we have
\[ h_\theta \leq \max\{h_1, 0\} + o(|x - x_0|^2). \tag{1} \]
(To see this, we just subtract \( c(x_0, y_0) - c(x, y_0) \) from both sides of the inequality \( A_3v \)).

Then (1) implies that the second fundamental form of \( S_\theta \) cannot strictly dominate the second fundamental form of \( S_1 \) in any tangential direction at \( x_0 \). By changing \( y \) into \( y_\theta' \) for \( \theta' \geq \theta \), this implies that the second fundamental form of \( S_\theta \) is non-decreasing with respect to \( \theta \).

\[ \square \]

Remark. We remark that analytically the conclusion of Lemma 2.2 can be expressed as a co-dimension one convexity of the matrix
\[ A(x, p) = -c_{xx}(x, c\text{-exp}_{x_0}(p)) \]
with respect to \( p \), in the sense that the quadratic form \( A\xi, \xi \) is convex on line segments in \( p \) orthogonal to \( \xi \) or more explicitly:
\[ \left[ A_{ij}(x, p_\theta) - (1 - \theta)A_{ij}(x, p_0) - \theta A_{ij}(x, p_1) \right] \xi_i \xi_j \leq 0, \tag{2} \]
for all \( \xi \in \mathbb{R}^n \) such that \( \xi \cdot (p_1 - p_0) = 0 \), which, for arbitrary \( y_0, y_1 \in \Omega^* \), is clearly equivalent to \( A_3w \) when \( c \in C^4 \).

We now deduce assertion \( (ii) \) in Theorem 1.2 from \( A_3v \); this will be done in several steps.

Step 1. Uniform boundedness of the section’s curvature (including \( c \)-hyperplanes)
From the previous corollary, it follows that \( \theta \to c_{xx}(x_0, y_\theta)\xi_i \xi_j \) is convex and therefore Lipschitz, and for a.e. \( \theta \in [0, 1] \),
\[ A = \partial_\theta c_{x_i x_j}(x_0, c\text{-exp}_{x_0}(p_\theta))\xi_i \xi_j \]
exists and is equal to \( \lim_{\theta' \to \theta} B(\theta, \theta') \) where
\[ B(\theta, \theta') = \frac{(c_{x_i x_j}(x_0, c\text{-exp}_{x_0}(p_{\theta'})) - c_{x_i x_j}(x_0, c\text{-exp}_{x_0}(p_\theta))) \xi_i \xi_j}{\theta' - \theta}. \]
The first term would be the curvature of \( H_\theta \) if it exists. The second term in the limit is the curvature of \( S_{\theta,\theta'} \). We can deduce right away that the curvature of \( S_{\theta,\theta'} \) remains uniformly bounded at \( x_0 \) thanks to (2). Now this reasoning can be extended to any point \( x_1 \in \partial S_{\theta,\theta'} \), although the \( c \)-segment between \( y_0 \) and \( y_0' \) will be with respect to \( x_1 \), but the conclusion that the curvature of \( S_{\theta,\theta'} \) at \( x_1 \) is uniformly bounded remains. Therefore the curvature of all sections is uniformly bounded as the uniform limit of \( S_{\theta,\theta'} \), \( H_\theta \) is a \( C^{1,1} \) hypersurface, and therefore has a curvature a.e. given by \( A \).

**Step 2. Local convexity** Wherever \( A \) is well defined, the curvature of \( H_\theta \) is equal to \( A \). Moreover, for \( \theta' > \theta \), the second fundamental form of \( S_{\theta,\theta'} \) dominates a.e. the one of \( H_\theta \).

Let us define the hypersurfaces

\[ P_m = \{ x \in \Omega, c(x, y_0) - c(x, y_1) = m \}, m \in \mathbb{R} \]

By standard measure theoretical arguments, the previous result implies the following:

**Lemma 2.3.** For a.e. \( y_0, y_1, m \) there holds at \( \mathcal{H}^{n-1} \) every point \( x_0 \) on \( P_m(y_0, y_1) \), that

- the second fundamental form (SFF) of \( H_0(x_0, y_0, y_1) \) at \( x_0 \) is well defined, let us call it \( A \), equivalently \( H_0(x_0, y_0, y_1) \) is twice differentiable (as a hypersurface)
- \( A \) is dominated by the SFF of \( \partial S_1(x_0, y_0) \)
- going back to the tangent space (i.e. composing with \( c \cdot y) \)), the second fundamental form of \( c \cdot y) \partial S_1(x_0, y_0) \) dominates the null form.

We now conclude the local convexity. Starting from a point \( x_0 \) where \( H_0 \) and \( S_0 \) are tangent. Both are defined by \( x_0, y_0, y_1 \). We let \( p = c_\pi(x_0, y_0) - c_\pi(x_0, y_1) \). Representing \( S_0 \) and \( H_0 \) as graphs over \( \mathbb{R}^{n-1} \), and we denote by \( h_0 \) and \( s_0 \) the corresponding functions. We assume \( x_0 = 0 \), and that both graphs have a flat gradient at \( 0 \). For \( x \in \mathbb{R}^{n-1} \) we have

\[ h(x') = |x'|^2 \int_0^1 \partial_{\nu}h(\theta x')(1 - \theta)d\theta \]

and the same holds for \( s \) (\( \nu \) is the appropriate unit vector). By the definition of \( H_0 \), at a given point \( z = (x', h_0(x')) \), \( H_0 \) is tangent to

\[ S_z = S(z, y_0, c - c_\pi(z, y_0) + p) \]

For almost every choice of \( x_0 \) there will hold for a.e. \( x' \) that

\[ \partial_{\nu}h(x') \leq \partial_{\nu}s_z(x') \leq \partial_{\nu}s_0(x') + \varepsilon(x' - 0) \]
with \( \lim_0 \varepsilon = 0 \), depending on the continuity of \( c_{xx}, c_{x,y} \). Therefore

\[
h_0(x') \leq |x'|^2 \left( \int_0^1 \partial_{xx} s_0(\theta x')(1 - \theta)d\theta + \varepsilon(x') \right)
\]

\[
\leq s_0(x') + \varepsilon(x' - x_0)|x'|^2.
\]

Going now in the tangent space, for \( q' \) in a well chosen \( n - 1 \) subspace, and \( \pi \) the projection on \( \{x_n = 0\} \), we call \( x(q') = \pi(c^* - \exp(y_0, q')) \) and we have

\[
h_0(x(q')) \leq s_0(x(q')) + \varepsilon(x(q') - x_0)|x(q')|^2,
\]

\( h_0(x(q')) \) is an affine function, \( s_0(x(q')) \) defines the image of \( S_0 \) by \( c_{y} \) and \( \varepsilon(x(q'))|x(q')|^2 \leq \varepsilon(q')|q'|^2 \) for some \( \varepsilon' \). For a.e. choice of \( x_0 \), this holds for a.e. \( q' \). More importantly the \( \varepsilon' \) is (locally) uniform. This implies the convexity through the following lemma.

**Lemma 2.4.** Let \( s \) be \( C^1 \). Assume that for some continuous \( \varepsilon(\cdot) \) with \( \varepsilon(0) = 0 \), there holds for almost every \( x_0, x \)

\[
s(x) \geq l_{x_0}(x) - \varepsilon(x - x_0)|x - x_0|^2
\]

\( l_{x_0} \) being the tangent function at \( x_0 \), then \( s \) is convex.

**Proof.** Elementary, both sides of the inequality are continuous in \( x, x_0 \), so this holds in fact everywhere.

\[\square\]

Global convexity To complete the proof of assertion (ii), we need to show that the set \( \hat{S}_1 \) is connected. The proof goes as follows, and it is very close to the argument of [6], Section 2.5. Let \( \sigma \) be a constant, and assuming that the set

\[
\{ c(x, y_0) - c(x, y_1) \leq \sigma \}
\]

has two disjoint components, we let \( \sigma \) increase until the two components touch in a \( C^1 \) \( c \)-convex subdomain \( \Omega' \subset \subset \Omega \). From the local convexity property this can only happen on the boundary of \( \Omega' \). At this point, say \( x_1 \) there holds locally that

\[
c(x, y_0) - c(x, y_1) \leq \sigma
\]

on \( \partial \Omega' \) and for \( x^\varepsilon = x_1 - \varepsilon \nu, \nu \) the outer unit normal to \( \Omega' \),

\[
c(x^\varepsilon, y_0) - c(x^\varepsilon, y_1) > h.
\]

This implies that

\[
c(x, y_0) - c(x, y_1) \geq h
\]

is locally \( c \)-convex around \( x_1 \), a contradiction, and from this we deduce that \( S_1 \) can have at most one component. Since a connected locally convex set in Euclidean space must be globally convex, we thus deduce that \( S_1 \) is globally \( c \)-convex.
2.1. **An analytical proof for a smooth cost function.** If a $C^2$ domain $\Omega$ is defined locally by $\varphi > 0$, its local $c$-convexity with respect to $y_0$, for $c \in C^3$, is expressed by
\[
\left[ \varphi_{ij} + c_{ij,k} k^{l} (\cdot, y_0) \partial_l \varphi \right] \tau_i \tau_j \geq 0,
\]
or equivalently
\[
\left[ \varphi_{ij} + \partial_p A_{ij} \partial_p \varphi \right] \tau_i \tau_j \geq 0
\]
for all $\tau \in \partial \Omega$ [8]. Plugging $\varphi(x) = c(x, y_0) - c(x, y_1) - h$ into this inequality, we obtain immediately from (2) that $S_1$ is locally $c$-convex with respect to $y_0$. More generally this argument proves Theorem 1.2 when we assume additionally that the form $A_\xi \cdot \xi$ is differentiable with respect to $p$ in directions orthogonal to $\xi$.

2.2. **Connectedness of the contact set.** This new characterization implies right away the $c$-convexity of the global $c$-sub-differential. We prove now that (i) implies (iii).

For $\phi$ $c$-convex, we have
\[
\phi(x) = \sup_y \{-\phi^c(y) - c(x, y)\},
\]
\[
\phi^c(y) = \sup_x \{-\phi(x) - c(x, y)\}.
\]
Then
\[
\{\phi(x) \leq -c(x, y_0) + h\} = \cap_y \{x : -\phi^c(y) - c(x, y) \leq -c(x, y_0) + h\}
\]
\[
= \cap_y \{x : c(x, y_0) \leq c(x, y) - h + \phi^c(y)\}.
\]
Therefore $\{\phi(x) \leq -c(x, y_0) + h\}$ is an intersection of $c$-convex sets and hence also $c$-convex. We then have
\[
G_{\phi}(y) = \{x, \phi(x) = -c(x, y) - \phi^c(y)\}
\]
\[
= \{x, \phi(x) \leq -c(x, y) - \phi^c(y)\},
\]
and hence $G_{\phi}(y)$ is a $c$-convex set. To show the dual conclusion, we may rewrite assertion (ii) as: for all $y, y_1 \in \Omega^*$, $x_0, x_1 \in \Omega^*$ and $\theta \in (0, 1)$, with
\[
c_\theta(x_\theta, y) = \theta c_\theta(x_1, y) + (1 - \theta)c_\theta(x_0, y),
\]
there holds
\[
\max\{-c(x_0, y) + c(x_0, y_0), -c(x_1, y) + c(x_1, y_0)\}
\]
\[
\geq -c(x_\theta, y) + c(x_\theta, y_0).
\]
Since this shows in particular that $A_{3v}$ is invariant under duality we complete the proof of assertion (iii). Moreover as a byproduct of this argument we also see that the sets $S_\theta$ are non-increasing with respect to $\theta$ and that $A_{3v}$ holds without the term $o(|x - x_0|^2)$. □
2.3. **Local implies global.** We prove that (ii) implies (iv). We consider $\phi$ a locally $c$-convex function, i.e., $\phi$ has at every point a local $c$-support. Locally, $\phi$ can be expressed as

$$
\phi(x) = \sup_{y \in \omega} \{-\psi(y) - c(x, y)\},
$$

for some $\omega(x) \subseteq \Omega^*$ (if $\phi$ was globally $c$-convex there would hold that $\omega \equiv \Omega^*$ and $\psi$ would be equal to $\phi^c$). It follows that the level sets

$$
S_{m, y_0} = \{ x : \phi(x) + c(x, y_0) \leq m \}
$$

are locally $c$-convex with respect to $y_0$ for any $y_0$. We obtain that $-\partial_y c(S_{m, y_0}, y_0)$ is locally convex. Reasoning again as in the proof of the global convexity in point (ii) (i.e. increasing $m$ until two components touch), we obtain that, for $\phi$ locally $c$-convex, $-\partial_y c(S_{m, y_0}, y_0)$ is globally convex for all $y_0$. This implies in turn the global $c$-convexity of $\phi$, following Proposition 2.12 of [2]. As already mentioned, this part is very similar to the argument of [8], section 2.5.

Finally we remark that the arguments in this paper extend to generating functions as introduced in [7] and also provide as a byproduct an alternative geometric proof of the invariance of condition $A3\omega$ under duality to the more complicated calculation in [7]. The resultant convexity theory is presented in [3].

□

**References**

[1] 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.

[2] G. Loeper. On the regularity of solutions of optimal transportation problems. *Acta Mathematica*, 202(2):241–283, 2009.

[3] G Loeper and Neil S. Trudinger. preprint.

[4] G. Loeper and C. Villani. Regularity of optimal transport in curved geometry: the non-focal case. *Duke Math. Journal*, 151(3):431–485, 2010.

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

[6] N. S. Trudinger and X.-J. Wang. On the second boundary value problem for Monge-Ampère type equations and optimal transportation. *Ann. Scuola Norm. Sup. Pisa Cl. Sci.*, 5(8):143–174, 2009.

[7] Neil S. Trudinger. On the local theory of prescribed Jacobian equations. *Discrete Contin. Dyn. Syst.*, 34(4):1663–1681, 2014.

[8] Neil S Trudinger and Xu-Jia Wang. On strict convexity and continuous differentiability of potential functions in optimal transportation. *Archive for rational mechanics and analysis*, 192(3):403–418, 2009.

Gregoire Loeper
Monash University
School of Mathematics
9 Rainforest Walk
3800 CLAYTON VIC, AUSTRALIA

email: gregoire.loeper@monash.edu