

On the c -concavity with Respect to the Quadratic Cost on a Manifold

Federico Glaudo*

January 28, 2018

Abstract

Pushing a little forward an approach proposed by Villani [9], we are going to prove that in the Riemannian setting the condition $\nabla^2 f < g$ implies that f is c -concave with respect to the quadratic cost as soon as it has a sufficiently small C^1 -norm. From this, we deduce a sufficient condition for the optimality of transport maps.

1 Introduction

Let us briefly recall the optimal transport problem on \mathbb{R}^n with quadratic cost $c(x, y) = \frac{1}{2}d^2(x, y)$. Given two probability measures $\mu, \nu \in \mathcal{P}_2(\mathbb{R}^n)$ we want to find a map $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$ such that $T_{\#}\mu = \nu$ and the quantity $\frac{1}{2} \int |x - T(x)|^2 d\mu(x)$ is minimized.

It is very well-known (see [4]) that, as soon as μ and ν are absolutely continuous with respect to the Lebesgue measure, an optimal map always exists. Moreover a map such that $T_{\#}\mu = \nu$ is optimal if and only if T is the gradient of a convex function. This striking characterization is a peculiarity of the quadratic cost.

The optimal transport problem has been thoroughly studied in the last three decades (see the monographs [9, 1, 8]) and for instance a fruitful generalization was the replacement of the space \mathbb{R}^n with a generic Riemannian manifold. In the Riemannian setting the turning point of the theory is given by McCann's Theorem [6], that generalizes the theorem of Brenier to a general compact Riemannian manifold. Denoting with M the manifold, McCann proves that there exists an optimal map $T : M \rightarrow M$ and furthermore the map T can be written as $\exp(-\nabla f)$,

*Scuola Normale Superiore, federico.glaudo@sns.it.

MSC-2010: 49K, 49Q20, 53C21.

Keywords: optimal transport, Riemannian manifold, c -concavity, McCann theorem.

where $f : M \rightarrow \mathbb{R}$ is a suitable c -concave function. A c -concave function, on a general space X endowed with a symmetric cost function $c : X \times X \rightarrow \mathbb{R}$, is a function $f : X \rightarrow \mathbb{R}$ such that there exists a family of $(x_i, \alpha_i)_{i \in I} \subseteq X \times \mathbb{R}$ such that it holds

$$f(x) = \inf_{i \in I} c(x, y_i) - \alpha_i.$$

Of course this is very similar to the definition of concave functions as the infimum of linear functions. Indeed in the Euclidean space, with cost $c = \frac{1}{2}|x - y|^2$, a function f is c -concave if and only if $\frac{1}{2}x^2 - f(x)$ is convex. It is therefore very natural to ask ourselves whether this equivalence can be proven also in the Riemannian setting. The natural generalization, on a compact Riemannian manifold (M, g) , should look like:

False Theorem (Naïve Statement). A function $f : M \rightarrow \mathbb{R}$ is c -concave if and only if $\nabla^2 f \leq g$.

Though, this statement does not take into account the fact that the manifold is curved and indeed it turns out being false. We show a counterexample in the last section of this document.

On the other hand, the only result known in literature that goes in this direction is the following, stated by Villani in his monograph [9] as Theorem 13.5:

Theorem (Villani). *Let M be a compact Riemannian manifold. Then, there is $\epsilon > 0$ such that any function $\psi \in C_c^2(M)$ satisfying $\|\psi\|_{C_c^2} \leq \epsilon$ is $d^2/2$ -concave.*

It is immediately clear that such a statement seems not optimal as instead of $\nabla^2 f \leq g$ it asks that $\nabla^2 f$ is very small. Our main goal is proving a *true* version of the false theorem stated above. The approach is exactly the same as the one adopted by Villani, but instead of using compactness arguments, we deduce explicit inequalities that depend on natural quantities associated to the manifold (curvature, injectivity radius and diameter). The exact statement of our main theorem is:

Theorem 1.1 (Main Theorem). *Let (M, g) be a compact Riemannian manifold with sectional curvature bounded from above by $K \geq 0$. Then there exists a constant $C_* = C_*(\text{inj}(M), K, \text{diam}(M)) > 0$ such that, for any $\epsilon > 0$, if $f \in C^2(M, \mathbb{R})$ is a function with*

$$\|\nabla f\|_\infty \leq \min \left(\frac{\epsilon}{3K \text{diam}(M)}, C_* \right) \quad \text{and} \quad \nabla^2 f \leq (1 - \epsilon)g$$

then f is c -concave.

The way this theorem should be intended is that if $g - \nabla^2 f \lesssim \|\nabla f\|_\infty \cdot g$, then f is c -concave.

As a consequence of the main theorem, we give a sufficient condition for a map $T : M \rightarrow M$ to be optimal from μ to $T_\# \mu$. The need for such an optimality condition arose naturally while trying to simplify the approach to the random matching problem proposed in [3]. The simplified approach will appear in the forthcoming paper [2] coauthored with Ambrosio.

Acknowledgment. The author warmly thanks professor L. Ambrosio for constructive criticism of the manuscript and for several comments and suggestions.

2 Notations

Given a compact Riemannian manifold (M, g) , let us fix the following notation:

- Let $d : M \times M \rightarrow [0, \infty)$ the Riemannian distance on M and $\exp : TM \rightarrow M$ the exponential map.
- Let K be the supremum of the positive part of the sectional curvature.
- Let $\text{inj}(M) > 0$ be the injectivity radius of the exponential map on M .
- Let $\text{diam}(M)$ be the diameter of the manifold M . We will use that for any Lipschitz function $f : M \rightarrow \mathbb{R}$, it holds

$$\sup f - \inf f \leq \text{diam}(M) \|\nabla f\|_\infty.$$

The constant $\text{diam}(M)$ is sharp, indeed if $f(\cdot) = d(\cdot, \bar{x})$ equality can be attained for some $\bar{x} \in M$.

Throughout this notes we will implicitly assume that the cost c is given by $c(x, y) = \frac{1}{2}d^2(x, y)$. Hence, when we say that a function f is c -concave we mean that it can be written as the infimum of functions of the form $x \mapsto \frac{1}{2}d^2(x, x_i) - a_i$, where $x_i \in X$, $a_i \in \mathbb{R}$ and i varies in a suitable set of indexes.

3 Main Theorem

Exploiting the technical results that we will prove later, the proof of the main theorem becomes straight-forward.

Proof of the Main Theorem. This is an easy consequence of [Proposition 4.5](#) (Technical Version of the Main Theorem). \square

Remark 3.1. As said in the introduction, the previous theorem should be seen as a generalization of the trivial fact that if a function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ satisfies $D^2f \leq \mathbf{1}$ then $x \mapsto \frac{1}{2}x^2 - f$ is convex (indeed on the Euclidean space this convexity is equivalent to the c -concavity of f).

For a couple of reasons such a statement is harder to prove on a Riemannian manifold. First of all the exponential map need not to be injective globally and that is why we need an additional bound on the gradient. Furthermore, in positive curvature, the Hessian of the square of the distance can be strictly smaller than the metric and consequently we will need to ask a stricter condition on the Hessian of the function itself (i.e. we need $\nabla^2 f$ *strictly* smaller than the metric).

Remark 3.2. The theorem can be easily extended to the case where f is compactly supported (and the manifold is non-compact).

Remark 3.3. Under the hypothesis that $\|\nabla f\|_\infty$ is small enough, the requirement $\nabla^2 f < (1 + \epsilon)g$ is necessary for f to be c -concave. This is a byproduct of the proof of [Proposition 4.5](#) (Technical Version of the Main Theorem). Indeed, using the notation of that proof, if the global oscillation of f is small enough, the choice of $x^* = \exp_x(-\nabla f(x))$ is mandatory (as we will see in the last section). Therefore it must hold $\nabla^2 h(x) \geq 0$, and that implies the desired bound on the Hessian.

Question. Is the assumption $\|\nabla f\|_\infty = \mathcal{O}(\epsilon)$ optimal?

The following corollary, deeply linked to McCann's Theorem (see [\[6\]](#)), is the reason behind our investigation of c -concave functions.

Corollary 3.4 (Optimality Condition). *Let M be a compact Riemannian manifold. If $f \in C^2(M)$ satisfies the requirements of [Theorem 1.1](#) (Main Theorem) then, for any probability measure $\mu \in \mathcal{P}(M)$, the map $T = \exp(-\nabla f)$ is optimal from μ to $T\#\mu$ with respect to the quadratic cost $c = \frac{1}{2}d^2$.*

Proof. This is a consequence of the strategy adopted in the proof of [Theorem 1.1](#) (Main Theorem). Indeed we will prove that

$$\frac{1}{2}d^2(x, \exp_x(-\nabla f(x))) - f(x) \leq \frac{1}{2}d^2(y, \exp_x(-\nabla f(x))) - f(y)$$

for any $x, y \in M$. Given a set of n points $(x_i)_{1 \leq i \leq n}$ and a permutation $\sigma \in S_n$, summing n times the latter inequality we obtain

$$\sum_{i=1}^n \frac{1}{2}d^2(x_i, \exp_{x_i}(-\nabla f(x_i))) \leq \sum_{i=1}^n \frac{1}{2}d^2(x_i, \exp_{x_{\sigma(i)}}(-\nabla f(x_{\sigma(i)})))$$

that proves the c -monotonicity of the graph of T and therefore the optimality of T . \square

Remark 3.5. The previous corollary could also be deduced directly from the c -concavity of the function f using an approach similar to the one used to prove McCann theorem. Anyway we have chosen to give a simpler proof that exploits the equivalence between optimality and c -monotonicity.

4 Technical Propositions and Proofs

Let us start stating a well-known characterization of c -concavity.

Lemma 4.1. *Let X be a metric space and let $c : X \times X \rightarrow \mathbb{R}$ be a symmetric lower-semicontinuous cost. A function $f : X \rightarrow \mathbb{R}$ is c -concave if and only if for any $x \in X$ there exists $x^* \in X$ such that $x \in \arg \min\{c(x^*, \cdot) - f(\cdot)\}$.*

Proof. It is an easy consequence of the fact that f is c -concave if and only if for any $x \in X$ there exists $x^* \in X$ such that $f(x) + f^*(x^*) = c(x, x^*)$. \square

The two following statements are rather known results in Riemannian geometry. The first one is a version of the Hessian comparison that compares a manifold with the constant curvature model, whereas the second is a lower-bound for the convexity radius of a manifold. As a corollary of the Hessian comparison we will obtain a *quantitative* estimate on the Hessian of the square of the distance.

It is not restrictive to assume that $K > 0$, and indeed we are going to do it in the following statements, since when $K = 0$ all the results can be recovered through a limit procedure.

Theorem 4.2 (Hessian Comparison). *Let us fix a point $x \in M$ and define $r(y) := d(x, y)$. At any point $y \in M$ such that $d(x, y) \leq \min(\frac{\pi}{\sqrt{K}}, \text{inj}(M))$ it holds*

$$\nabla^2 r \geq \frac{\sqrt{K} \cos(\sqrt{K}r)}{\sin(\sqrt{K}r)} (g - dr \otimes dr) .$$

Proof. An equivalent, albeit not completely identical, statement can be found at [7, p. 342]. \square

Corollary 4.3. *With the same assumptions of Theorem 4.2 (Hessian Comparison), if we also have $r < \frac{1}{\sqrt{K}}$, then it holds*

$$\frac{1}{2} \nabla^2 (r^2) \geq (1 - Kr^2)g .$$

Proof. Applying the usual calculus rules and Theorem 4.2 (Hessian Comparison) we get

$$\frac{1}{2} \nabla^2 (r^2) = dr \otimes dr + r \nabla^2 r \geq \alpha(\sqrt{K}r)g + (1 - \alpha(\sqrt{K}r)) dr \otimes dr ,$$

where $\alpha : [0, \infty) \rightarrow \mathbb{R}$ is the function given by $\alpha(t) = \frac{\cos(t)t}{\sin(t)}$. The identity $|\nabla r| = 1$ implies $dr \otimes dr \leq g$, therefore we can continue the chain of inequalities and obtain

$$\frac{1}{2}\nabla^2(r^2) \geq \left(\alpha(\sqrt{K}r) - |1 - \alpha(\sqrt{K}r)|\right)g.$$

Hence the thesis follows if we show that for any $0 \leq t < 1$ it holds

$$|1 - \alpha(t)| \leq \frac{t^2}{2}.$$

We leave the proof of this elementary inequality to the reader. \square

Theorem 4.4 (Convexity Radius Lower-Bound). *For any point $x \in M$ and $\delta \leq \min\left(\frac{\text{inj}(M)}{2}, \frac{\pi}{2\sqrt{K}}\right)$, the ball $B(x, \delta)$ is geodesically convex¹.*

Proof. It can be found at [5, p. 404]. \square

Proposition 4.5 (Technical Version of the Main Theorem). *Given a C^2 function $f : M \rightarrow \mathbb{R}$, let us denote $\delta = \frac{\text{inj}(M)}{\sqrt{2 \text{diam}(M) \|\nabla f\|_\infty + \|\nabla^2 f\|_\infty}}$. If $\delta \leq \min\left(\frac{\text{inj}(M)}{2}, \frac{1}{\sqrt{k}}\right)$ and $\nabla^2 f \leq (1 - K\delta^2)g$, then f is c -concave.*

Proof. The proof of this proposition is *heavily* inspired by the proof of Theorem 3.15 [9]. Indeed, we are making quantitative the approach proposed by Villani with the help of the Hessian Comparison Theorem.

Let us fix $x \in M$ and define $x^* = \exp_x(-\nabla f(x))$. We are going to prove that x is a minimizer of the function $h(y) := \frac{1}{2}d^2(x^*, y) - f(y)$. The c -concavity follows thanks to [Lemma 4.1](#).

More specifically we will show that the three following claims hold:

1. If $y \notin B(x^*, \delta)$ then $h(y) \geq h(x)$.
2. It holds $\nabla h(x) = 0$.
3. In the ball $B(x^*, \delta)$ the function h is convex (i.e. $\nabla^2 h \geq 0$).

These three claims imply that x is a global minimizer as we can restrict ourselves in the ball $B(x^*, \delta)$ thanks to 1. and then x is a critical point of a convex function in a convex domain. The convexity of $B(x^*, \delta)$ is a consequence of the assumption on δ thanks to [Theorem 4.4](#) (Convexity Radius Lower-Bound).

Let us prove the three claims separately:

¹A domain $D \subseteq M$ is geodesically convex if for any $x, y \in D$ there exists a geodesic of length $d(x, y)$ that connects the two points and is completely contained in D .

Proof of 1. If $d(x^*, y) \geq \delta$ we have

$$\begin{aligned} h(y) - h(x) &\geq \frac{1}{2}d^2(x^*, y) - (\sup f - \inf f) - \frac{1}{2}\|\nabla f\|_\infty^2 \\ &\geq \frac{1}{2}\delta^2 - \text{diam}(M)\|\nabla f\|_\infty - \frac{1}{2}\|\nabla f\|_\infty^2 \geq 0. \end{aligned}$$

Proof of 2. The function $d^2(x^*, \cdot)$ is smooth in x since $d(x, x^*) = |\nabla f(x)| \leq \delta < \text{inj}(M)$. Hence also the function h is smooth and its gradient is

$$\nabla h(x) = \gamma'(1) - \nabla f(x).$$

where $\gamma : [0, 1] \rightarrow M$ is the constant speed geodesic from x^* to x . From the definition of x^* it follows that $\gamma'(1) = \nabla f(x)$ and therefore $\nabla h(x) = 0$.

Proof of 3. Our assumptions on δ are exactly those needed to apply [Corollary 4.3](#), hence, denoting $r(y) = d(x^*, y)$, we get

$$\nabla^2 h = \frac{1}{2}\nabla^2(r^2) - \nabla^2 f \geq (1 - Kr^2)g - (1 - K\delta^2)g \geq 0$$

that is exactly what we had to show. □

5 A Counterexample to the Naïve Statement

In this section we show that the condition $\nabla^2 f \leq g$ is not sufficient of the c -concavity of f . We will find a counterexample when the Riemannian manifold is the 2-dimensional sphere \mathbb{S}^2 .

Let us start by giving a necessary condition for being c -concave.

Proposition 5.1. *There exists a constant $\delta = \delta(M)$ such that, for any $f \in C^1(M)$ with $\text{osc}(f) \leq \delta$, the following statements are equivalent:*

1. f is c -concave;
2. for any $x \in M$, it holds $x \in \arg \min \frac{1}{2}d^2(x^*, y) - f(y)$ where $x^* = \exp_x(-\nabla f)$.

Proof. The implication 2. \implies 1. is a straightforward consequence of [Lemma 4.1](#).

For the other implication, let us use again [Lemma 4.1](#) to get that for any $x \in M$ there exists $x^* \in M$ such that $x \in \arg \min \frac{1}{2}d^2(x^*, y) - f(y)$. Thanks to our hypothesis on $\|f\|_\infty$ we can easily get that any such x^* must be near to x and therefore the distance from x^* has to be smooth at x . Hence, given that in x the function $\frac{1}{2}d^2(x^*, y) - f(y)$ has a minimum, its gradient must be null. Therefore, computing the gradient of the distance function it is easy to prove that it must hold $x^* = \exp_x(-\nabla f)$. □

Our strategy is now to negate that $x \in \arg \min \frac{1}{2}d^2(x^*, y) - f(y)$ looking at the Hessian. Indeed we will build a function such that $\nabla^2 f \leq g$ but the Hessian of $\frac{1}{2}d^2(x^*, y) - f(y)$ is not positive-definite at $y = x$.

From now on we will always work on \mathbb{S}^2 (of course g will denote the Riemannian metric on \mathbb{S}^2). The main reason behind this choice is that in this setting the inequality stated in [Theorem 4.2](#) (Hessian Comparison) becomes an identity². Let us state explicitly the said identity:

Proposition 5.2 (Hessian on the Sphere). *Let us fix a point $x \in S^2$ and define $r(y) := d(x, y)$. At any point $y \in S^2$ such that $0 < d(x, y) < \pi$, it holds*

$$\nabla^2 r = \cot(r) (g - dr \otimes dr)$$

and

$$\nabla^2 \left(\frac{1}{2}r^2 \right) = r \cot(r) \cdot g + (1 - r \cot(r)) dr \otimes dr.$$

Proposition 5.3. *Let $f : \mathbb{S}^2 \rightarrow \mathbb{R}$ be a $C^2(\mathbb{S}^2)$ function such that $\|f\|_\infty$ is as small as asked in [Proposition 5.1](#). If there exists an $x \in \mathbb{S}^2$ such that $\nabla f(x) \neq 0$ and $\nabla^2 f(x) = g$, then f is not c -concave.*

Proof. Let us assume that f is c -concave. Hence we know from [Proposition 5.1](#) that the function $y \mapsto \frac{1}{2}d^2(x^*, y) - f(y)$ has a global minimum at $y = x$. Hence it must hold that the Hessian of that function is positive semi-definite at x . Therefore it must hold

$$\nabla^2 \left(\frac{1}{2}d^2(x^*, y) \right) \Big|_{y=x} \geq g.$$

However, thanks to [Proposition 5.2](#) (Hessian on the Sphere), it is easy to see that such inequality does not hold if $x^* \neq x$ and that shows the contradiction as $x^* = \exp_x(-\nabla f)$ and $\nabla f(x) \neq 0$. \square

It remains to build a function $f : \mathbb{S}^2 \rightarrow \mathbb{R}$ such that:

- $\|f\|_\infty$ is arbitrarily small.
- $\nabla f(N) \neq 0$, where N is the north pole of the sphere.
- $\nabla^2 f(N) = g$.
- Everywhere it holds $\nabla^2 f \leq g$.

²We don't really need to work in dimension 2 instead of general dimension, but we believe it is much easier to follow a reasoning on \mathbb{S}^2 than on a higher dimensional sphere.

Such a function would be a counterexample to the “naïve theorem” stated in the introduction as it could not be c -concave thanks to [Proposition 5.3](#). We construct our example as a linear combination $f = f_1 + \epsilon f_2$ where $\epsilon > 0$ is a sufficiently small constant and f_1, f_2 are such that:

- $\|f_1\|_\infty$ is arbitrarily small.
- $\nabla^2 f_1 \leq g$ with equality only in N .
- $\nabla f_1(N) = 0$.
- $\nabla f_2(N) \neq 0$, $\nabla^2 f_2(N) = 0$ and $\nabla^2 f_2 \leq 0$ in a neighbourhood of N .

It is obvious that if ϵ is sufficiently small then $f = f_1 + \epsilon f_2$ satisfies all our requirements.

We are left to prove the existence of f_1 and f_2 with the said properties. What we ask on f_2 is almost nothing and therefore we leave it to the reader to convince himself that such a function exists. A good choice for f_1 is given by $y \mapsto \rho(\frac{1}{2}d^2(N, y))$ where $\rho : \mathbb{R} \rightarrow \mathbb{R}$ satisfies $\rho(t) = t$ in a neighbourhood of 0 and ρ becomes constant as soon as $t > \epsilon$ for a certain $\epsilon > 0$. We are not going to perform the computations, but we remark that they are pretty easy exploiting once again that the inequality given by the Hessian comparison is an identity on the sphere.

References

- [1] L. AMBROSIO, N. GIGLI, AND G. SAVARÉ, *Gradient flows in metric spaces and in the space of probability measures*, Lectures in Mathematics ETH Zürich, Birkhäuser Verlag, Basel, second ed., 2008.
- [2] L. AMBROSIO AND F. GLAUDO, *In preparation*.
- [3] L. AMBROSIO, F. STRA, AND D. TREVISAN, *A pde approach to a 2-dimensional matching problem*, ArXiv Mathematics e-prints, (2016).
- [4] Y. BRENIER, *Polar factorization and monotone rearrangement of vector-valued functions*, Comm. Pure Appl. Math., 44 (1991), pp. 375–417.
- [5] I. CHAVEL, *Riemannian Geometry: A Modern Introduction*, vol. 98 of Cambridge Studies in Advanced Mathematics, Cambridge University Press, second ed., 2006.
- [6] R. J. MCCANN, *Polar factorization of maps on riemannian manifolds*, Geometric and Functional Analysis, 11 (2001), pp. 589–608.

- [7] P. PETERSEN, *Riemannian Geometry*, vol. 171 of Graduate Texts in Mathematics, Springer-Verlag, New York, 1998.
- [8] F. SANTAMBROGIO, *Optimal Transport for Applied Mathematicians*, Progress in Nonlinear Differential Equations and Their Applications, Birkhäuser Basel, 2015.
- [9] C. VILLANI, *Optimal transport, old and new*, Springer Verlag, 2008.