# Gradient estimates for potentials of invertible gradient–mappings on the sphere*

Philippe Delanoë†, Grégoire Loeper‡

**Abstract**

McCann showed that, if the potential of a gradient-mapping, on a compact riemannian manifold, is c-convex, the length of its gradient cannot exceed the diameter of the manifold. We improve this bound in two different manners on the constant curvature spheres, under assumptions on the relative density of the image-measure of the riemannian volume. One proof, with the standard metric, relies on the Brenier–McCann optimal measure-transport property; the other, purely pde, ignores it.

## Introduction

Following Robert McCann [13], we consider *gradient-mappings* on a riemannian manifold: these are mappings of the form $m \mapsto \exp_m(\operatorname{grad}_m \phi) =: G(\phi)(m)$ where $\phi$ is a real function on the manifold, the "potential" of the gradient-mapping $G(\phi)$. Brenier, in the euclidean space [4, 5], then McCann on compact riemannian manifolds [13], established the optimal character of such mappings for measure-transport when the cost-integrand is the *squared* distance from the generic point to its image. For the mapping $G(\phi)$ this distance is nothing but the norm $|d\phi|$ of the gradient. On a compact manifold, granted the cost-convexity of $\phi$ (definition recalled below), which holds whenever $G(\phi)$ is a diffeomorphism [12, Proposition 2], that norm cannot exceed the *diameter* of the manifold [13, Lemma 2] (see also [12, Proposition 4]).

In the present note, using the Brenier–McCann optimal transport result [13], we improve that bound on the standard unit $n$-sphere in case the image-measure by $G(\phi)$ of the riemannian measure has essentially bounded relative density. Here is the idea of the proof: if the gradient-mapping $G(\phi)$ sends a point $m$ close to its antipodal point $m'$ (or equivalently, if $|d\phi|(m)$ is close to $\pi$), then the cost-convexity of the potential $\phi$ implies that $G(\phi)$ must actually send a large neighborhood of $m$ into a small cap around $m'$, contradicting the bound on the density of the image-measure (see section 1).

Under stronger regularity assumptions, but ignoring the Brenier–McCann result and dealing with a general constant curvature metric, we carry out further estimates of the gradient in terms of a relative density (section 2). The analysis here is not quite standard because we must cope with the exponential map involved in the operator $\phi \mapsto G(\phi)$. Such techniques will be required for higher

*2000 *Mathematics Subject Classification*: Primary 35J60,35B45; Secondary 28C99,49Q99
†supported by the CNRS, member of the EDGE network
‡research supported by an AC doctoral grant at Nice, then by EPFL (Lausanne)

order estimates (still open) which would provide (*via* the continuity method, see [12, section 5.1]) a regular potential anytime we are given a regular relative density. Given so, the *a posteriori* regularity of the Brenier-McCann cost-convex potential is not known, except on *flat* manifolds [6, 7, 8, 9], where a smooth potential is known to exist directly as well [12, Theorem 3] (see also [11, 15]). Progressing on that regularity problem was our present motivation.

# 1 Gradient estimate *via* optimal measure-transport

To begin with, let us recall the notion of cost-convexity on a riemannian manifold (see [13] and further references therein) formulated with the Brenier quadratic cost-function $c(m,p) = \frac{1}{2}\ d2(m,p)$.

**Definition 1** *Let $M$ be a compact riemannian manifold and $d(\cdot,\cdot)$, its riemannian distance-function. The c-transform $\phi^c$ of a function $\phi : M \mapsto \mathbb{R}$ is defined for all $p \in M$ by*

$$\phi^c(p) = \sup_{m \in M} \{-\frac{d2(m,p)}{2} - \phi(m)\}.$$

*The function $\phi$ is said to be c-convex if $(\phi^c)^c = \phi$.*

The c-transform is the riemannian counterpart of the Legendre transform in the euclidean space. If $\phi$ is c-convex, then $\phi$ is $D$-Lipschitz (with $D$ the diameter of $M$) [13, Lemma 2] and twice differentiable almost everywhere [10, Proposition 3.14] [1, 3]; in particular, the gradient-mapping $G(\phi)$ is (differentiable, hence) continuous at almost every point of $M$. Moreover, if the push-forward by $G(\phi)$ of the riemannian Lebesgue measure dVol is absolutely continuous with respect to dVol, then the gradient-mapping $G(\phi)$ is *invertible* almost everywhere on $M$ (see [13, Corollary 10]).

In this section, we take for $M$ the standard unit sphere $\mathbb{S}^n$. We aim at the following result:

**Theorem 1** *Let $\phi : \mathbb{S}^n \mapsto \mathbb{R}$ be a c-convex function such that*

$$G(\phi)_{\#}\mathrm{dVol} = \rho\,\mathrm{dVol}$$

*for some $\rho \in L^\infty(\mathbb{S}^n, \mathrm{dVol})$, where dVol stands for the canonical Lebesgue measure on $\mathbb{S}^n$. There exists $\epsilon > 0$ depending on $\phi$ only through $\|\rho\|_{L^\infty(\mathbb{S}^n,\mathrm{dVol})}$ such that $|d\phi| \le \pi - \epsilon$ almost everywhere on $\mathbb{S}^n$.*

This result will follow from a property of $G(\phi)$ called 2-monotonicity which we now define:

**Definition 2** *A map $A : \mathbb{S}^n \mapsto \mathbb{S}^n$ is called 2-monotone if it satisfies identically the inequality:*

$$d2[A(m), m] + d2[A(p), p] \le d2[A(m), p] + d2[A(p), m]\ .$$

*It is called a.e. 2-monotone if the inequality holds only almost-everywhere in $\mathbb{S}^n \times \mathbb{S}^n$.*

The following property holds for 2-monotone maps:

**Theorem 2** *Let $A : \mathbb{S}^n \mapsto \mathbb{S}^n$ be a measurable a.e. 2-monotone map. If $A_{\#}\mathrm{dVol} = \rho\,\mathrm{dVol}$ for some $\rho \in L^\infty(\mathbb{S}^n, \mathrm{dVol})$, there exists $\epsilon > 0$ (depending on $\|\rho\|_{L^\infty(\mathbb{S}^n,\mathrm{dVol})}$) such that $d[m, A(m)] \le \pi - \epsilon$ at almost every $m \in \mathbb{S}^n$.*

Before proving Theorem 2, let us show how it implies Theorem 1, relying on the Brenier–McCann's optimal transportation property [13, Theorem 8] recalled here for completeness:

**Theorem 3 (McCann)** *If a function $\psi : \mathbb{S}^n \mapsto \mathbb{R}$ is c-convex, then the gradient-mapping $G(\psi)$ minimizes the quadratic cost*

$$\int_{\mathbb{S}^n} d2[m, A(m)]\ \mathrm{dVol}(m)$$

*among all measurable maps $A : \mathbb{S}^n \mapsto \mathbb{S}^n$ satisfying $A_{\#}\mathrm{dVol} = G(\psi)_{\#}\mathrm{dVol}$.*

This key-property implies the following lemma from which Theorem 2 yields at once Theorem 1:

**Lemma 1** *If a function $\psi : \mathbb{S}^n \mapsto \mathbb{R}$ is c-convex, then the gradient-mapping $G(\psi)$ is a.e. 2-monotone.*

**Proof of Lemma 1.** Let $m_1, m_2$ be distinct points of continuity for $A = G(\psi)$. Consider an ambiant rotation in the plane $(m_1, 0, m_2)$ that sends $m_1$ to $m_2$, leaving 0 unchanged. The rotation induces a measure-preserving map $R$ of $\mathbb{S}^n$. Let now $R_\epsilon$ be defined as follows (setting $B(m,r)$ for the riemannian ball of $\mathbb{S}^n$ given by $d(m,p) < r$):

$$\begin{aligned} &\forall m \in \mathbb{S}^n \setminus [B(m_1,\epsilon) \cup B(m_2,\epsilon)]\ \ R_\epsilon(m) = m;\\ &\forall m \in B(m_1,\epsilon),\ \ R_\epsilon(m) = R(m);\\ &\forall m \in B(m_2,\epsilon),\ \ R_\epsilon(m) = R^{-1}(m). \end{aligned}$$

It is easily checked that $R_\epsilon$ is also measure-preserving, satisfying $(A \circ R_\epsilon)_{\#}\mathrm{dVol} = A_{\#}\mathrm{dVol}$. From Theorem 3, we have (setting $V_\epsilon$ for the volume of geodesic balls of radius $\epsilon$ in $\mathbb{S}^n$):

$$\begin{aligned} 0 \ &\ge\ \frac{1}{V_\epsilon}\int_{\mathbb{S}^n} [d2(A(m), m) - d2((A \circ R_\epsilon)(m), m)]\ \mathrm{dVol}\\ &=\ \frac{1}{V_\epsilon}\left\{\int_{B(x_1,\epsilon)} [d2(A(m), m) - d2((A \circ R)(m), m)]\ \mathrm{dVol}\right.\\ &\quad\left. + \int_{B(x_2,\epsilon)} \left[d2(A(m), m) - d2((A \circ R^{-1})(m), m)\right]\ \mathrm{dVol}\right\} \end{aligned}$$

and, due to the continuity of $A$ at $m_1$ and $m_2$, the latter readily goes to

$$d2(A(m_1), m1) + d2(A(m_2), m_2) - [d2(A(m_1), m2) + d2(A(m_2), m_1)]$$

as $\epsilon$ goes to 0, proving the lemma.

**Proof of Theorem 2.**

For $(m_1, m_2, p) \in (\mathbb{S}^n)3$, let us consider the function:

$$F(m_1, m_2, p) = \frac{d2}{2}(m_2, p) - \frac{d2}{2}(m_1, p).$$

The couple $(m_1, m_2)$ being fixed, we have

$$\text{grad}_p F(m_1, m_2, p) = \exp_p^{-1}(m_1) - \exp_p^{-1}(m_2)$$

(see [12, p.152]). This gradient is thus defined everywhere except at $m'_1$ and $m'_2$, the antipodal points respectively to $m_1$ and $m_2$. Moreover, the map

$$V \in T_p\mathbb{S}^n \mapsto \exp_p(V) \in \mathbb{S}^n$$

is uniformly Lipschitz, so there exists a positive constant $\theta$ independent of $(m_1, m_2)$ such that:

$$\left|\text{grad}_p F(m_1, m_2, p)\right| \geq \theta \; d(m_1, m_2) \; . \tag{1}$$

Now it is easily seen that $p \mapsto F(m_1, m_2, p)$ reaches its infimum at $p = m'_1$ and nowhere else.

**Lemma 2** *For all $(m_1, m_2, p) \in (\mathbb{S}^n)3$, with $m_1 \neq m_2$, we have*

$$d(p, m'_1) \leq 2\pi \; \frac{F(m_1, m_2, p) - F(m_1, m_2, m'_1)}{\theta 2 \, d2(m_1, m_2)} \; ,$$

*where $m'_1$ stands for the antipodal point to $m_1$.*

**Proof of Lemma 2.** Fixing $(m_1, m_2)$, let us consider on $\mathbb{S}^n$ the steepest descent equation:

$$\dot{p}(t) = -\text{grad}_p F[m_1, m_2, p(t)].$$

Using (1), any solution $p(t)$ satisfies

$$\frac{d}{dt}F[m_1, m_2, p(t)] = -\left|\text{grad}_p F[m_1, m_2, p(t)]\right|^2 \leq -\theta 2 \; d2(m_1, m_2) \; .$$

Therefore, starting from $p(0) = p_0$ distinct from $m'_1$ and $m'_2$, the minimum of $p \mapsto F(m_1, m_2, p)$ is reached (necessarily at $p = m'_1$) within some finite time $T$ estimated by:

$$T \leq \frac{F(m_1, m_2, p_0) - F(m_1, m_2, m'_1)}{\theta 2 \; d2(m_1, m_2)} \; .$$

Since $|\dot{p}(t)|$ is bounded above by $2\pi$, we also have $d(p_0, m'_1) \leq 2\pi T$. Combining both inequalities we obtain the lemma.

Back to the proof of Theorem 2, the assumption '$A$ is a.e. 2-monotone' reads equivalently: for almost all $(m_1, m_2) \in \mathbb{S}^n \times \mathbb{S}^n$,

$$F[m_1, m_2, A(m_2)] \leq F[m_1, m_2, A(m_1)] \; . \tag{2}$$

Combining this with Lemma 2, we get

$$d[m_1', A(m_2)] \leq 2\pi \, \frac{F[m_1, m_2, A(m_1)] - F(m_1, m_2, m_1')}{\theta 2 \; d2(m_1, m_2)},$$

hence also, since $p \mapsto F(m_1, m_2, p)$ is $2\pi$-Lipschitz,

$$d[m_1', A(m_2)] \leq \frac{4\pi^2}{\theta 2} \frac{d[m_1', A(m_1)]}{d2(m_1, m_2)}.$$

Fixing $\delta > 0$, and a point $m_1$ such that (2) holds at almost all points $m_2$, let $E_\delta$ denote the set of points $m_2 \in \mathbb{S}^n$ such that: $d(m_1, m_2) \geq \delta$ and (2) holds at $(m_1, m_2)$. On the one hand, from the preceding inequality, which holds almost everywhere, we infer

$$\mathrm{Vol}[A(E_\delta)] \leq \mathrm{Vol}[B(m_1', \epsilon)]$$

with $\epsilon = \dfrac{4\pi^2}{\theta 2} \dfrac{d[A(m_1), m_1']}{\delta 2}$. On the other hand, the definition of $d\mu = A_\# \mathrm{dVol}$ implies:

$$\mathrm{Vol}(E_\delta) \leq \mu[A(E_\delta)] \; .$$

Altogether, we thus obtain:

$$\frac{\mathrm{Vol}(E_\delta)}{\mathrm{Vol}[B(m_1', \epsilon)]} \leq \frac{\mu[A(E_\delta)]}{\mathrm{Vol}[A(E_\delta)]} \equiv \frac{1}{\mathrm{Vol}[A(E_\delta)]} \int_{A(E_\delta)} \rho \; \mathrm{dVol} \leq \|\rho\|_{L^\infty(\mathbb{S}^n)} \; .$$

With $\delta > 0$ fixed, the left-hand side goes to infinity as $\epsilon$ or $d[A(m_1), m_1']$ goes to 0, whereas $\|\rho\|_{L^\infty(\mathbb{S}^n)}$ stays finite: so there must exist $\epsilon_0$ such that $d[A(m_1), m_1'] \geq \epsilon_0$, which is the desired result.

**Remark.** Sticking to the standard metric, but given an arbitrary regular positive measure $d\mu = M \,\mathrm{dVol}$ on $\mathbb{S}^n$ with same total mass as dVol, if $\phi$ is c-convex satisfying $G(\phi)_\# d\mu = \rho \, d\mu$ with $\rho \in L^\infty$, the simplest way to again derive a bound on $|d\phi|$ sharper than $\pi$ goes by noting that, if $\tau \leq M \leq 1/\tau$ for some $\tau \in (0, 1)$ and if $G(\phi)_\# \mathrm{dVol} =: f \,\mathrm{dVol}$, then $\|f\|_{L^\infty(\mathbb{S}^n)} \leq \tau^{-2} \|\rho\|_{L^\infty(\mathbb{S}^n)}$.

## 2 Gradient estimates *via* a classical approach

This section contains a pde approach to gradient estimates for potentials of $C1$ gradient-diffeomorphisms on the sphere equipped with a constant curvature metric. It is less straightforward and requires stronger regularity assumptions than the preceding approach. But it shows how to deal with more general metrics than just round ones and to cope without McCann's theorem (Theorem 3 above), relying on a careful use of the Jacobi equation which is encoded in the gradient-rearrangement operator.

Fixing a real $K > 0$, let us work on the $n$-sphere equipped with a metric $g = g_K$ of constant curvature $K$, denoted by $\mathbb{S}^n_K$. For each $C1$ real function $\psi$, we set $G_K(\psi)$ for the gradient-mapping $m \mapsto \exp_m(\mathrm{grad}_m \psi)$ built on $\mathbb{S}^n_K$ and drop the subscript $K$ unless necessary.

**Theorem 4** *Let $\phi : \mathbb{S}^n \to \mathbb{R}$ be a $C3$ function such that $G_K(\phi)$ is a diffeomorphism and let $\rho : \mathbb{S}^n \to \mathbb{R}$ be the positive $C1$ function defined by:*

$$G_K(\phi)_\# \mathrm{dVol}_K = \rho \, \mathrm{dVol}_K \; .$$

*There exists $\epsilon > 0$ depending on $\phi$ and $K$ only through the quantity*

$$\frac{\pi}{\sqrt{K}} \max_{\mathbb{S}^n} |d(\log \rho)|$$

*such that the following estimate holds:*

$$\max_{\mathbb{S}^n} \left(\sqrt{K}|d\phi|\right) \leq \pi - \epsilon \; . \tag{3}$$

*Furthermore, letting*

$$c = \max_{\mathbb{S}^n} \left[ \frac{\sqrt{K}|d\phi|}{\sin(\sqrt{K}|d\phi|)} \right]$$

*we also have:*

$$\max_{\mathbb{S}^n} \left(\sqrt{K}|d\phi|\right) \leq \frac{c}{\sqrt{K}} \max_{\mathbb{S}^n} \left| d\left[\rho^{-1/(n-1)}\right] \right| \; . \tag{4}$$

**Remarks.** $(i)$ The estimates (3) and (4) are *dilation-invariant.* Indeed, the polar factorization [13] [12, Remark 2] on $\mathbb{S}^n_K$ of a gradient-diffeomorphism $G_1(\phi_1)$ built on $\mathbb{S}^n_1$ yields $G_1(\phi_1) = G_K(\phi_K)$ with $\phi_K = \phi_1/K$. Now, one easily verifies that $\sqrt{K}|d\phi_K|_{g_K} \equiv |d\phi_1|_{g_1}$. Furthermore, the density $\rho$ is readily independent of $K$; finally $|df|_{g_K}/\sqrt{K}$ is independent of $K$, for any function $f$ on $\mathbb{S}^n$.
$(ii)$ The estimate (3) ensures, in terms of the density $\rho$, that the generic point $m$ and its image $G_K(\phi)(m)$ are uniformly non-antipodal on $\mathbb{S}^n_K$. The estimate (4) specifies rather how *close* the points $m$ and $G_K(\phi)(m)$ must be when the density $\rho$ is *slowly varying*; in particular, we recover the implication $\rho = 1 \Rightarrow G_K(\phi) = I$ (see [12, Remark 4]).
$(iii)$ Let us provide a motivation for using more general constant curvature metrics than just the round ones. Given a regular positive measure $d\mu$ on the sphere, with total mass equal to the one of a round metric $g_0$, there exists a diffeomorphism $\psi$ pulling back the $g_0$-measure to $d\mu$ [14]. By naturality $d\mu$ coincides with the Lebesgue measure of the pulled-back metric $\psi^* g_0$; the latter has the same curvature as $g_0$ but, in general, it is no more round.

**Proof.** Consider the scalar second order differential operator $u \mapsto F(u)$ defined on $\mathbb{S}^n$ by $G(u)^* \mathrm{dVol} = F(u)\, \mathrm{dVol}$. It is elliptic at $\phi$ [12, Proposition 3], satisfying

$$F(\phi) = \frac{1}{\rho \circ G(\phi)} \; . \tag{5}$$

Fix a point $m_0 \in \mathbb{S}^n$ where $|d\phi|$ assumes its *maximum* and a Fermi chart $\left(x^1, \ldots, x^n\right)$ at $m_0$ (see *e.g.* [2, p.15]) such that $t \mapsto (0, \ldots, 0, t\, r_0)$ represents the geodesic $t \mapsto \exp_{m_0}(t \, \mathrm{grad}_{m_0} \phi)$. Set $r_0 = |d\phi|(m_0)$ and $p_0 = G(\phi)(m_0)$; recall $\sqrt{K} r_0 \leq \pi$ [13, Lemma 2] [12, Proposition 4]. The critical condition at $m_0$ for $w := |d\phi|^2/2$ simply reads (setting $u_i = \frac{\partial u}{\partial x^i}$, and so on):

$$\forall i \in \{1, \ldots, n\}, \; \phi_{in}(0) = 0 \, , \tag{6}$$

while differentiating twice $w$ at $m_0$ yields:

$$w_{ij}(0) = r_0\, \phi_{nij} + \sum_{k=1}^{n} \phi_{ik}\, \phi_{jk} + r_0 2K(\delta_{ij} - \delta_{in}\delta_{jn})\,. \tag{7}$$

Henceforth, we further select the Fermi chart such that the sub-matrix $\phi_{\alpha\beta}(0)$, for $1 \leq \alpha, \beta \leq n-1$, is *diagonal.* Now we write the maximum condition for $w$ at $m_0$ using the linearization of the operator $F$ at $\phi$; specifically we write:

$$0 \geq dF(\phi)(w)(m_0)\,. \tag{8}$$

**Lemma 3** *In our Fermi chart, we have at $m_0$:*

$$dF(\phi)(w)(m_0) = Kr_0 2 \left( \sum_{\alpha<n} F^{\alpha\alpha} \right) - r_0 \left( \frac{\rho_n}{\rho 2}(0, \ldots, 0, r_0) \right) + \sum_{\alpha<n} F^{\alpha\alpha}\, (\phi_{\alpha\alpha})\, 2$$

*with*

$$\rho(p_0) F^{\alpha\alpha} = \frac{\sin(\sqrt{K} r_0)}{\sqrt{K} r_0} \left[ \cos(\sqrt{K} r_0) + \phi_{\alpha\alpha}(0) \frac{\sin(\sqrt{K} r_0)}{\sqrt{K} r_0} \right]^{-1} .$$

Differing the proof of the lemma, we infer from it (since, by ellipticity, the $F^{\alpha\alpha}$'s are positive) first of all, the vanishing of $r_0$ (thus, the constancy of $\phi$) if $\rho$ is constant (which we exclude from now on); secondly, the inequalities:

$$\forall \alpha < n,\ \ F^{\alpha\alpha}\, (\phi_{\alpha\alpha})\, 2 \leq \frac{\pi}{\sqrt{K}} \frac{\rho_n}{\rho 2}(0, \ldots, 0, r_0)$$

hence also

$$\forall \alpha < n,\ \ \frac{\sin(\sqrt{K} r_0)}{\sqrt{K} r_0}\, (\phi_{\alpha\alpha})\, 2 \leq C_1 \left[ \cos(\sqrt{K} r_0) + \phi_{\alpha\alpha}(0) \frac{\sin(\sqrt{K} r_0)}{\sqrt{K} r_0} \right]$$

with $C_1 = \frac{\pi}{\sqrt{K}} \max_{\mathbb{S}^n} |d(\log \rho)| \neq 0$. So $X := \phi_{\alpha\alpha}(0)$ satisfies the inequality:

$$\frac{\sin(\sqrt{K} r_0)}{\sqrt{K} r_0} (X2 - C_1 X) - C_1 \cos(\sqrt{K} r_0) \leq 0\,,$$

with $\left[ \sin(\sqrt{K} r_0)/\sqrt{K} r_0 \right] > 0$ (by ellipticity); therefore the discriminant of the left-hand quadratic polynomial must be non-negative, which reads:

$$\frac{\sin(\sqrt{K} r_0)}{\sqrt{K} r_0} + \frac{4}{C_1} \cos(\sqrt{K} r_0)\ \geq 0\ .$$

The latter implies (3) as routinely verified.

Besides, we also infer from (8) and Lemma 3 the bound:

$$K r_0 \leq \left( \sum_{\alpha<n} F^{\alpha\alpha} \right)^{-1} \frac{\rho_n}{\rho 2}(0, \ldots, 0, r_0)\ .$$

On the one hand, the arithmetic-geometric inequality provides:

$$\sum_{\alpha<n} F^{\alpha\alpha} \geq (n-1)\left(\prod_{\alpha<n} F^{\alpha\alpha}\right)^{1/(n-1)} \quad ;$$

on the other hand, we will check (below) the following relation at $m_0$:

$$\prod_{\alpha<n} F^{\alpha\alpha} = \left[\frac{\sin(\sqrt{K}r_0)}{\sqrt{K}r_0}\right]^{n-1} \rho^{2-n}(0,\ldots,0,r_0)\ . \tag{9}$$

Altogether, we thus obtain:

$$Kr_0 \leq \frac{\sqrt{K}r_0}{\sin(\sqrt{K}r_0)}\ \frac{1}{(n-1)}\ \rho^{-\frac{1}{n-1}-1}\ \rho_n(0,\ldots,0,r_0)$$

which immediately yields (4). So we are left with the proof of Lemma 3, in the course of which (9) will be checked.

**Proof of Lemma 3.** We will proceed stepwise.

**Step 1.** Given $\phi$ and $\rho$ as in Theorem 4, we fix a generic point $m \in \mathbb{S}^n$ and, recalling that $m$ and $G(\phi)(m)$ are not antipodal [12, Corollary 1], we take a chart $p \in \mathbb{S}^n \mapsto x(p) = \left(x^1,\ldots,x^n\right) \in \mathbb{R}^n$ whose domain contains $m$ and $G(\phi)(m)$, with $x(m)=0$. In such a chart, setting $\mathrm{dVol}(p) = v[x(p)]\, dx1\ldots dx^n$ and $G^i[x(p)] = x^i[G(\phi)(p)]$, we readily find for $F(\phi)$ the local expression:

$$F(\phi)(p) = \frac{v\{x[G(\phi)(p)]\}}{v[x(p)]}\ \det\left(\frac{\partial G^i}{\partial x^j}\right)\ , \tag{10}$$

where $\left(\frac{\partial G^i}{\partial x^j}\right)$ stands for the matrix of the *jacobian map* $JG(\phi)$ in our chart (see [12]).

**Step 2.** First, let us specify what is $G^i(x)$ equal to. It is the value taken at $t=1$ by the solution $X^i(t)$ of the geodesic Cauchy problem, namely (setting $\dot{X}^i = \frac{dX^i}{dt}$, and so on):

$$\ddot{X}^i + \Gamma^i_{ab}[X(t)]\dot{X}^a\dot{X}^b = 0\ ,\ X^i(0) = x^i\ ,\ \dot{X}^i(0) = g^{ir}(x)\phi_r(x)\ ,$$

where the $\Gamma^i_{ab}$'s are the Christoffel symbols of the round metric $g$ in our chart, and Einstein's convention is used. Then, let us compute $\frac{\partial G^i}{\partial x^j}(x)$: it is the value at $t=1$ taken by the solution $X^i_j(t)$ of the previous problem differentiated once with respect to the parameter $x^j$, that is to say, the problem

$$\ddot{X}^i_j + \left(\Gamma^i_{ab}\right)_r X^r_j \dot{X}^a\dot{X}^b + 2\Gamma^i_{ab}\dot{X}^a\dot{X}^b_j = 0\ ,\ X^i_j(0) = \delta^i_j\ ,\ \dot{X}^i_j(0) = \left(g^{ir}\right)_j \phi_r + g^{ir}\phi_{rj}$$

(we recognize the *Jacobi equation* along the geodesic $X(t)$).

**Step 3.** We pause to complete the calculation of $\frac{\partial G^i}{\partial x^j}(0)$ when $m = m_0$ (*cf. supra*) and the chart is our above Fermi chart at $m_0$. When so, $X^i_j$ (like $Y^i$ below) satisfies the normalized Jacobi equation along the geodesic from $m_0$ to $G(\phi)(m_0)$, namely:

$$\ddot{Y}^i + R_{ninq}(0,\ldots,0,tr_0)Y^q r_0 2 = 0 \tag{11}$$

here with $R_{ninq} \equiv K(\delta_{iq} - \delta_{in}\delta_{qn})$, together with the initial conditions:

$$X^i_j(0) = \delta^i_j \, , \, \dot{X}^i_j(0) = \phi_{ij}(0) \; .$$

Therefore we routinely find (*cf. supra*, in particular equation (6)) that the jacobian matrix $\frac{\partial G^i}{\partial x^j}(0)$ is *diagonal*, with $\frac{\partial G^n}{\partial x^n}(0) = 1$ and:

$$\forall \alpha < n, \; \frac{\partial G^\alpha}{\partial x^\alpha}(0) = \cos(\sqrt{K} r_0) + \phi_{\alpha\alpha}(0) \, \frac{\sin(\sqrt{K} r_0)}{\sqrt{K} r_0} \; . \tag{12}$$

Moreover, noting that $v(0) = v(0, \dots, 0, r_0) = 1$, we get from (10):

$$F(\phi)(m_0) = \prod_{\alpha < n} \left[ \cos(\sqrt{K} r_0) + \phi_{\alpha\alpha}(0) \frac{\sin(\sqrt{K} r_0)}{\sqrt{K} r_0} \right] \, ,$$

from which equation (9) readily follows, using (5). Besides, we also get:

$$\frac{\partial}{\partial x^n}[\rho \circ G(\phi)](m_0) = \sum_{i=1}^{n} \rho_i(0, \dots, 0, r_0) \, \frac{\partial G^i}{\partial x^n}(0) \equiv \rho_n(0, \dots, 0, r_0) \; . \tag{13}$$

**Step 4.** Now, back to the fixed generic point $m$, varying the potential $\phi$ in the direction of the above function $w$, we proceed to compute:

$$\frac{d}{d\lambda} \left[ F(\phi + \lambda w) \right]_{\lambda = 0} (m) \; .$$

To begin with, we do so on $X^i(t)$, setting $\widetilde{X^i}(t)$ for the resulting solution of:

$$\ddot{\widetilde{X^i}} + \left(\Gamma^i_{ab}\right)_r \widetilde{X^r} \dot{X}^a \dot{X}^b + 2\Gamma^i_{ab} \dot{X}^a \dot{\widetilde{X^b}} = 0 \, , \; \widetilde{X^i}(0) = 0 \, , \; \dot{\widetilde{X^i}}(0) = g^{ir} w_r$$

and also on $X^i_j(t)$, setting $\widetilde{X^i_j}(t)$ for the resulting solution of:

$$\begin{aligned} \ddot{\widetilde{X^i_j}} \quad &+ \quad \left(\Gamma^i_{ab}\right)_{rs} \widetilde{X^s} X^r_j \dot{X}^a \dot{X}^b \\ &+ \quad \left(\Gamma^i_{ab}\right)_r \left( \widetilde{X^r_j} \dot{X}^a \dot{X}^b + 2 X^r_j \dot{X}^a \dot{\widetilde{X^b}} + 2 \widetilde{X^r} \dot{X}^a \dot{X}^b_j \right) \\ &+ \quad 2\Gamma^i_{ab} \left( \dot{\widetilde{X^a}} \dot{X}^b_j + \dot{X}^a \dot{\widetilde{X^b_j}} \right) = 0 \, , \\ \widetilde{X^i_j}(0) \quad &= \quad 0 \; , \; \dot{\widetilde{X^i_j}}(0) \; = \; \left(g^{ir}\right)_j w_r + g^{ir} w_{rj} \; . \end{aligned}$$

**Step 5.** We simplify the preceding calculations by taking for $m$ the point $m_0$ where $w$ is maximum and by using our above Fermi chart at $m_0$. Then $\widetilde{X^i} \equiv 0$ and $\widetilde{X^i_j}(t)$ satisfies the normalized equation (11) with the initial conditions:

$$\widetilde{X^i_j}(0) = 0, \; \dot{\widetilde{X^i_j}}(0) = w_{ij}(0) \; .$$

Recalling (7), we routinely find for $\widetilde{X^i_j}$ at $t = 1$ the expressions:

$$\begin{aligned} \forall j = 1, \dots, n, \; \widetilde{X^n_j}(1) \quad &= \quad r_0 \, \phi_{nnj}(0) \;\; ; \\ \forall \alpha < n, \; \widetilde{X^\alpha_\alpha}(1) \quad &= \quad [r_0 \phi_{n\alpha\alpha}(0) + \phi 2_{\alpha\alpha}(0) + r_0 2K] \, \frac{\sin(\sqrt{K} r_0)}{\sqrt{K} r_0} \;\; ; \\ \forall \alpha < n, \, \forall j \neq \alpha, \; \widetilde{X^\alpha_j}(1) \quad &= \quad r_0 \, \phi_{n\alpha j}(0) \, \frac{\sin(\sqrt{K} r_0)}{\sqrt{K} r_0} \;\; . \end{aligned}$$

**Step 6.** Let us compute $dF(\phi)(w)(m_0)$ from (10), using the preceding expressions $\widetilde{X^i_j}(1)$ and $\widetilde{X^i} \equiv 0$. The latter yields 0 at $m_0$ for the linearization of $v\{x[G(\phi)]\}$. We thus obtain, using (12) to invert the jacobian matrix then taking the trace:

$$\begin{aligned} dF(\phi)(w)(m_0) &= F(\phi)(m_0)\left\{\widetilde{X^n_n}(1) + \sum_{\alpha<n}\left[\frac{\partial G^\alpha}{\partial x^\alpha}(0)\right]^{-1}\widetilde{X^\alpha_\alpha}(1)\right\} \\ &\equiv \frac{r_0}{\rho(p_0)}\phi_{nnn}(0) + \sum_{\alpha<n} F^{\alpha\alpha}\left[r_0\,\phi_{n\alpha\alpha}(0) + \phi^2_{\alpha\alpha}(0) + Kr_0 2\right] \ , \end{aligned}$$

with $F^{\alpha\alpha}$ as defined in Lemma 3.

**Step 7.** It remains only to treat the *third derivatives* of $\phi$ occuring in the preceding expression of $dF(\phi)(w)(m_0)$, namely to prove the equality:

$$(14) \qquad r_0\left[\frac{1}{\rho(p_0)}\phi_{nnn}(0) + \sum_{\alpha<n} F^{\alpha\alpha}\phi_{n\alpha\alpha}(0)\right] = -r_0\frac{\rho_n}{\rho^2}(0,\ldots,0,r_0) \ .$$

To do so, we first go back to the end of step 2 and differentiate the initial value problem once again, with respect to the parameter $x^k$, getting:

$$\begin{aligned} \ddot{X}^i_{jk} &+ \left(\Gamma^i_{ab}\right)_{rs} X^s_k X^r_j \dot{X}^a \dot{X}^b \\ &+ \left(\Gamma^i_{ab}\right)_r \left(X^r_{jk}\dot{X}^a\dot{X}^b + 2X^r_j\dot{X}^a\dot{X}^b_k + 2X^r_k\dot{X}^a\dot{X}^b_j\right) \\ &+ 2\Gamma^i_{ab}\left(\dot{X}^a_k\dot{X}^b_j + \dot{X}^a\dot{X}^b_{jk}\right) = 0 \ , \end{aligned}$$

with the initial conditions:

$$X^i_{jk}(0) = 0\,,\ \dot{X}^i_{jk}(0) = \left(g^{ir}\right)_{jk}\phi_r + \left(g^{ir}\right)_j\phi_{rk} + \left(g^{ir}\right)_k\phi_{rj} + g^{ir}\phi_{rjk} \ .$$

Then we take $m = m_0$ and $k = n$ in our Fermi chart. From step 2 we have $X^s_n \equiv \delta^s_n$ due to (6); it implies the relation

$$\left(\Gamma^i_{ab}\right)_{rs} X^s_n \dot{X}^a\dot{X}^b = r_0 2\left(\Gamma^i_{nn}\right)_{rn}$$

the right-hand side of which *vanishes* because, along $Ox^n$ (the geodesic from $m_0$ to $G(\phi)(m_0)$), we have:

$$\left(\Gamma^i_{nn}\right)_r \equiv R_{ninr} = K\left(\delta_{ir} - \delta_{in}\delta_{rn}\right) \ .$$

Still from $X^b_n \equiv \delta^b_n$, we infer $\dot{X}^b_n = 0$; and from $X^r_n = \delta^r_n$ we also have:

$$2\left(\Gamma^i_{ab}\right)_r X^r_n\dot{X}^a = 2r_0\left(\Gamma^i_{nb}\right)_n \equiv 0\,.$$

Altogether, the above equation for $X^i_{jn}(t)$ thus reduces to the *Jacobi* equation, namely (11). Moreover, the initial conditions become $X^i_{jn}(0) = 0$ and

$$\dot{X}^i_{jn}(0) = \phi_{ijn}(0)$$

(noting that $\left(g^{ir}\right)_{jn}\phi_r = r_0\left(g^{in}\right)_{jn} \equiv 0$, since $\left(g^{in}\right)_j \equiv 0$ along $Ox^n$). In particular at $t=1$, we find $X^n_{nn}(1) = \phi_{nnn}(0)$ and:

$$\forall\alpha < n,\ X^\alpha_{\alpha n}(1) = \phi_{n\alpha\alpha}(0)\,\frac{\sin(\sqrt{K}r_0)}{\sqrt{K}r_0}\ .$$

Finally, we differentiate equation (5) at $m_0$ with respect to $x^n$ (in our Fermi chart), using the generic expression (10). Since $v \equiv 1$ along $Ox^n$, it yields:

$$\begin{aligned}[1/\rho\circ G(\phi)]_n &= F(\phi)(m_0)\left\{X^n_{nn}(1) + \sum_{\alpha<n}\left[\frac{\partial G^\alpha}{\partial x^\alpha}(0)\right]^{-1} X^\alpha_{\alpha n}(1)\right\}\\ &\equiv \frac{1}{\rho(p_0)}\,\phi_{nnn}(0) + \sum_{\alpha<n} F^{\alpha\alpha}\phi_{n\alpha\alpha}(0)\ ,\end{aligned}$$

and recalling (13), the proof of (14) is complete.

An analogous proof works for the backward transport equation (which reads $F(\phi)=\rho$). It yields the following result (with related remarks as above):

**Theorem 5** *Let $\phi : \mathbb{S}^n \to \mathbb{R}$ be a $C3$ function such that $G_K(\phi)$ is a diffeomorphism and let $\rho : \mathbb{S}^n \to \mathbb{R}$ be the positive $C1$ function defined by:*

$$G_K(\phi)^*\mathrm{dVol}_K = \rho\ \mathrm{dVol}_K\ .$$

*There exists $\epsilon > 0$ depending on $\phi$ and $K$ only through the quantity*

$$\frac{\pi}{\sqrt{K}}\max_{\mathbb{S}^n}|d(\log\rho)|$$

*such that the following estimate holds:*

$$\max_{\mathbb{S}^n}\left(\sqrt{K}|d\phi|\right) \le \pi - \epsilon\ . \tag{15}$$

*Furthermore, letting*

$$c = \max_{\mathbb{S}^n}\left[\frac{\sqrt{K}|d\phi|}{\sin(\sqrt{K}|d\phi|)}\right]$$

*we also have:*

$$\max_{\mathbb{S}^n}\left(\sqrt{K}|d\phi|\right) \le \frac{c}{\sqrt{K}}\ \max_{\mathbb{S}^n}\left|\,d\left[\rho^{1/(n-1)}\right]\right|\ . \tag{16}$$

Finally, recalling that the diffeomorphism inverse of $G(\phi)$ is nothing but $G(\phi^c)$ [12, Corollary 3] and thus, that the equation $G(\phi)_\#\mathrm{dVol}_K = \rho\mathrm{dVol}_K$ is equivalent to $G(\phi^c)^*\mathrm{dVol}_K = \rho\mathrm{dVol}_K$ (with the same density-function $\rho$), noting moreover the double identity $|d\phi(m)| \equiv |d\phi^c(p)| \equiv d_K(m,p)$ where $p = G(\phi)(m)$ and $d_K$ stands for the distance-function of $g_K$, we may combine the estimates (4) (applied to $\phi$) and (16) (applied to $\phi^c$) and get the sharper bound:

$$\max_{\mathbb{S}^n}\left(\sqrt{K}|d\phi|\right) \le \frac{c}{\sqrt{K}}\ \min\left\{\max_{\mathbb{S}^n}\left|\,d\left[\rho^{-1/(n-1)}\right]\right|, \max_{\mathbb{S}^n}\left|\,d\left[\rho^{1/(n-1)}\right]\right|\right\}\ .$$

**Acknowledgment:** the authors thank Yann Brenier for introducing each to the other and for stimulating conversations.

## References


[1] A. D. Aleksandrov, Almost everywhere existence of the second differential of a convex function and some properties of convex surfaces connected with it, *Uchenye Zapiski Leningrad. Gos. Univ.* **6**:37 (1939) 3–35 (in russian).

[2] Th. Aubin, *Nonlinear analysis on manifolds. Monge–Ampère equations*, Grundlehren der math. Wissensch. **252** (Springer, New-York, 1982).

[3] V. Bangert, Analytische Eigenschaften konvexer Funktionen auf Riemannschen Manigfaltigkeiten, *J. Reine Angew. Math.* **307** (1979) 309–324.

[4] Y. Brenier, Décomposition polaire et réarrangement monotone des champs de vecteurs, *C. R. Acad. Sci. Paris* série I **305** (1987) 805–808.

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

[6] Y. Brenier & G. Loeper, A geometric approximation to the Euler equations: the Vlasov–Monge-Ampère equation , (preprint Nice, 2003) to appear.

[7] L. Caffarelli, Some regularity properties of the solutions of the Monge–Ampère equation, *Comm. Pure Appl. Math.* **44** (1991) 965–969.

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

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

[10] D. Cordero-Erausquin, R. McCann & M. Schmuckenschläger, A riemannian interpolation inequality à la Borell, Brascamp and Lieb, *Inventiones math.* **146** (2001) 219–257.

[11] Ph. Delanoë, Classical solvability in dimension two of the second boundary-value problem associated with the Monge–Ampère operator, *Ann. Inst. Henri Poincaré Anal. non lin.* **8** (1991) 443–457.

[12] Ph. Delanoë, Gradient rearrangement for diffeomorphisms of a compact manifold, *Diff. Geom. Appl.* **20**:2 (2004) 145-165.

[13] R. McCann, Polar factorization of maps on riemannian manifolds, *Geom. Funct. Anal.* **11** (2001) 589–608.

[14] J. Moser, On the volume elements on a manifold, *Trans. Amer. Math. Soc.* **120** (1965) 286–294.

[15] J. Urbas, On the second boundary-value problem for equations of Monge–Ampère type, *J. reine angew. Math.* **487** (1997) 115–124.



Ph. Delanoë
Université de Nice-Sophia Antipolis
Mathématiques, Parc Valrose
F-06108 Nice CEDEX 2
e-mail: delphi@math.unice.fr

G. Loeper
EPFL, SB, IMA
10015 Lausanne
e-mail: gregoire.loeper@epfl.ch