Regularity for the optimal transportation problem with Euclidean distance squared cost on the embedded sphere

Jun Kitagawa and Micah Warren*

January 26, 2011

Abstract

We give a sufficient condition on initial and target measures supported on the sphere $\mathbb{S}^n$ to ensure the solution to the optimal transport problem with the cost $\frac{|x-y|^2}{2}$ is a diffeomorphism.

1 Introduction

In this paper, we will show the following theorem.

Theorem 1.1. *Suppose that we have two probability measures $\mu := e^f dVol$ and $\nu := e^g dVol$ on $\mathbb{S}^n \subseteq \mathbb{R}^{n+1}$, with f and g smooth functions. If*

$$\|Df\| + \|Dg\| < \frac{n-1}{2\pi} \tag{1.1}$$

then the optimal pairing of μ and ν under the optimal transport problem with cost given by the Euclidean distance squared is a diffeomorphism.

For the case when the cost is given by the geodesic distance squared on $\mathbb{S}^n$, regularity was shown by Loeper in [4]. The cost given by the Euclidean distance squared was first investigated by Gangbo and McCann in [2], where the authors show examples of measures given by smooth densities where the optimal pairing is not given by a map. Thus there is a need for some condition on the two measures.

In section 2 we give the setup of the problem. Since our cost does not satisfy the twist condition, we take one of the branches of the cost exponential function, then use that to define the elliptic equation (2.3). In section 3, we calculate the Ma-Trudinger-Wang tensor in a specific coordinate system. In section 4, we prove a gradient estimate for solutions of (2.3), while in section 5 we show that the contact set for such solutions can only consist of one point, both under

*M.W. is supported in part by NSF Grant DMS-0901644.

appropriate conditions on f and g. Finally, in section 6 we use the continuity method to prove our main theorem.

We provide here a reference table for the notation used in this paper.

Notation	Definition	Location
$\lvert\cdot\rvert$	Euclidean norm on $\mathbb{R}^{n+1}$	
$\langle \cdot, \cdot \rangle$	Euclidean inner product on $\mathbb{R}^{n+1}$	
$\mathring{g}(\cdot, \cdot)$	Canonical metric on $\mathbb{S}^n$	
$\lvert\cdot\rvert_{\mathbb{S}^n}$	Length of vectors and covectors on $\mathbb{S}^n$	
$\lVert Du \rVert$	$\sup_{x \in \mathbb{S}^n} \lvert Du(x) \rvert_{\mathbb{S}^n}$	
$c_{ij,kl}(x,y)$ etc.	$\frac{\partial}{\partial x_i} \frac{\partial}{\partial x_j} \frac{\partial}{\partial y_k} \frac{\partial}{\partial y_l} c(x,y)$ etc.	
$c^{i,j}$	inverse matrix of $c_{i,j}$	
$w_{ij}(x)$	$u_{ij}(x,t) + c_{ij}(x, T_u^+(x))$	(2.4)
$Y^+(x,p)$	Inverse of a branch of the map $y \mapsto D_x c(x,y) = p$	(2.6)
$T^+(x)$, $T_u^+(x)$	$Y^+(x, Du(x))$	before (2.3)
$Y^-(x,p)$	Inverse of a branch of the map $y \mapsto D_x c(x,y) = p$	(5.2)
$T_u^-(x)$	$Y^-(x, Du(x))$	(5.1)

2 Set up of problem

2.1 Monge and Kantorovich problems

For probability measure spaces $(\mathcal{X}, \mu)$ and $(\mathcal{Y}, \nu)$, let $\Pi(\mu, \nu)$ be the set of probability measures γ on $\mathcal{X} \times \mathcal{Y}$ such that

$$\gamma(E \times \mathcal{Y}) = \mu(E)$$
$$\gamma(\mathcal{X} \times \tilde{E}) = \nu(\tilde{E})$$

for all $E \subset \mathcal{X}$ and $\tilde{E} \subset \mathcal{Y}$ measurable. The cost is a measurable function

$$c : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}^+.$$

Definition 2.1. *A probability measure $\gamma_0 \in \Pi(\mu, \nu)$ is a Kantorovich solution to the optimal transportation problem between μ and ν with cost c if*

$$\int_{\mathcal{X} \times \mathcal{Y}} c(x,y) d\gamma_0(x,y) = \inf_{\gamma \in \Pi(\mu,\nu)} \int_{\mathcal{X} \times \mathcal{Y}} c(x,y) d\gamma(x,y).$$

Definition 2.2. *A measurable map $T : \mathcal{X} \to \mathcal{Y}$ is a Monge solution to the optimal transportation problem between μ and ν with cost c if*

$$\int_{\mathcal{X}} c(x, T(x)) d\mu(x) = \inf_{S_\# \mu = \nu} \int_{\mathcal{X}} c(x, S(x)) d\mu(x).$$

Definition 2.3. *A real valued function u defined on $\mathcal{X}$ is c-convex if for each $x_0 \in \mathcal{X}$, there exists a $\lambda_0 \in \mathbb{R}$ and $y_0 \in \mathcal{Y}$ such that*

$$\begin{aligned} u(x_0) &= -c(x_0, y_0) + \lambda_0 \\ u(x) &\geq -c(x, y_0) + \lambda_0, \ \forall x \in \mathcal{X}. \end{aligned}$$

If the second inequality is strict for $x \neq x_0$, we say the function is strictly c-convex. We call such a function $-c(\cdot, y_0) + \lambda_0$ a c-support function to u at x_0.

Definition 2.4. *For a c-convex function u, we define its c-transform u^c by*

$$u^c(y) := \sup_{x \in \mathcal{X}} (-c(x, y) - u(x)).$$

Remark 2.5. *If u is c-convex, then at any fixed x_0 where u and c are differentiable, we can see that for some $y_0 \in \mathcal{Y}$,*

$$\begin{aligned} u_i(x_0) &= -c_i(x_0, y_0) \\ u_{ij}(x_0) &\geq -c_{ij}(x_0, y_0) \end{aligned} \tag{2.1}$$

where the second inequality is in the sense of matrices.

2.1.1 Twisted case

Now suppose that for each x_0 the map $D_x c(x_0, \cdot) : \mathcal{Y} \to T^*_{x_0}\mathcal{X}$ is defined and injective. In this case we can implicitly define $T(x)$ from a c-convex function u as

$$-u_i(x) = c_i(x, T(x)).$$

Differentiating this and taking the determinant, one obtains

$$\det\left(u_{ij}(x) + c_{ij}(x, T(x))\right) = |\det c_{i,j}(x, T(x))| \det\left(\frac{\partial T^i}{\partial x_j}(x)\right)$$

and from this we can write down the elliptic optimal transport equation:

$$\det\left(u_{ij}(x) + c_{ij}(x, T(x))\right) = |\det c_{i,j}(x, T(x))| \frac{\mu(x)}{\nu(T(x))}, \tag{2.2}$$

here $\mu(x)$ and $\nu(y)$ are densities with respect to $dx^1 \cdots dx^n$ and $dy^1 \cdots dy^n$ for chosen coordinate systems (assuming the measures are absolutely continuous with respect to a coordinate volume form). This equation is (degenerate) elliptic for a c-convex solution u (see Remark 2.5), and is invariant under a change of coordinates. It is a standard result that c-convex solutions of the above equation determine uniquely the solution to the optimal transportation problem.

2.2 $\mathbb{S}^n$ with Euclidean cost

For the remainder of the paper, we specialize to the case

$$c(x, y) := \frac{|x - y|^2}{2},$$

where $|\cdot|$ is the Euclidean distance on $\mathbb{R}^{n+1}$.

In the twisted case, it is clear that each c-convex function determines a single-valued map $T(x)$. This is no longer the case for this cost function on the sphere. The main issue that we run into is the fact that for a fixed $x_0 \in \mathbb{S}^n$, the map $D_x c(x_0, \cdot) : \mathbb{S}^n \to T^*_{x_0}\mathbb{S}^n$ is not injective. In fact, if $p \in T^*_{x_0}\mathbb{S}^n$, $|p|_{\mathbb{S}^n} < 1$, there are exactly two points $y = y_1$ and $y = y_2$ such that $D_x c(x_0, y) = -p$. With this in mind, we make the following definition.

Definition 2.6. *We define the map $Y^+(x, p)$ implicitly by*

$$\begin{aligned} &D_x c(x, Y^+(x, p)) = -p \\ &\langle x, Y^+(x, p)\rangle > 0 \end{aligned}$$

for $p \in T^*_x\mathbb{S}^n$, $|p|_{\mathbb{S}^n} < 1$.

Now for u c-convex with $\|Du\| < 1$, we use Y^+ in place of the implicit definition of $T(x)$ in (2.2) to define an elliptic equation

$$\det w_{ij}(x) = |\det c_{i,j}(x, T^+(x))| \frac{\mu(x)}{\nu(T^+(x)))}, \tag{2.3}$$

where

$$w_{ij}(x) := u_{ij}(x) + c_{ij}(x, T^+(x)) \tag{2.4}$$

and

$$T^+(x) := Y^+(x, Du(x))$$

which is well-defined as long as $\|Du\| < 1$. Eventually, we will show that under appropriate conditions on the densities, solutions of this equation indeed determine single-valued solutions to the optimal transport problem.

Remark 2.7. *If u is c-convex and differentiable at x_0 with $|Du(x_0)|_{\mathbb{S}^n} < 1$, and (x_0, y_0) is a pair of points satisfying* (2.1)*, it is easy to see that either $\langle x_0, y_0\rangle > 0$ or there exists another point y_1 also satisfying* (2.1) *with $\langle x_0, y_1\rangle > 0$. In the latter case $c_{ij}(x_0, y_1) > c_{ij}(x_0, y_0)$. Hence, we see that:*

the pair $(x_0, T^+(x_0))$ satisfies (2.1)*,*
$w_{ij}(x_0)$ is positive semidefinite.

3 Calculation of MTW tensor

In this section, we utilize a coordinate system specialized to this problem on the sphere to calculate various quantities involving c, most notably the MTW tensor of [5].

We will define a coordinate system centered around a point $x_0 \in \mathbb{S}^n$ as follows. First, rotate x_0 so it is given by e_{n+1}, then take coordinates for the upper hemisphere by representing it as the graph $(x, \beta(x))$ over $B_1(0) \subset \mathbb{R}^n$, where

$$\beta(x) := \sqrt{1 - |x|^2}.$$

Note that we leave ourselves the freedom to further rotate $\mathbb{S}^n$ as long as the e_{n+1} direction remains unchanged.

Definition 3.1. *[5, Section 2] Given V, $W \in T_x\mathbb{S}^n$ and η, $\zeta \in T_x^*\mathbb{S}^n$, define*

$$(MTW)_{ij}^{kl}(x,y)V^iW^j\eta_k\zeta_l := -(c_{ij,pq} - c_{ij,r}c^{r,s}c_{s,pq})c^{p,k}c^{q,l}(x,y)V^iW^j\eta_k\zeta_l.$$

We will say that a cost c has the property (A3s) *at $(x, y) \in \mathbb{S}^n \times \mathbb{S}^n$ with constant $\delta_0 > 0$ if*

$$(MTW)_{ij}^{kl}(x,y)V^iV^j\eta_k\eta_l \geq \delta_0|V|^2_{\mathbb{S}^n}|\eta|^2_{\mathbb{S}^n}. \tag{A3s}$$

The following is widely known, but we carry out the calculations in our coordinate system here for later reference.

Proposition 3.2. *The cost $\frac{|x-y|^2}{2}$ satisfies* (A3s) *with a uniform constant $\delta_0 = 1$ for all $(x, y) \in \mathbb{S}^n \times \mathbb{S}^n$ such that $\langle x, y\rangle > 0$.*

Proof. We will utilize the coordinate system indicated above, and calculate various derivatives of c. First,

$$\begin{aligned}
c(x,y) &= \frac{|(x, \beta(x)) - (y, \beta(y))|^2}{2} \\
&= \frac{|x-y|^2 + (\beta(x) - \beta(y))^2}{2} \\
&= \frac{|x|^2 + |y|^2 - 2\langle x, y\rangle + (1 - |x|^2 + 1 - |y|^2 - 2\beta(x)\beta(y))}{2} \\
&= 1 - \langle x, y\rangle - \beta(x)\beta(y).
\end{aligned}$$

Thus we calculate, at generic x and y,

$$\begin{aligned}
c_i &= -y_i - \beta_i(x)\beta(y) \\
c_{ij} &= -\beta_{ij}(x)\beta(y) \\
c_{ij,k} &= -\beta_{ij}(x)\beta_k(y) \\
c_{i,k} &= -\delta_{ik} - \beta_i(x)\beta_k(y) \\
c_{i,kl} &= -\beta_i(x)\beta_{kl}(y) \\
c_{ij,kl} &= -\beta_{ij}(x)\beta_{kl}(y)
\end{aligned} \tag{3.1}$$

and

$$\beta_i(x) = -\frac{x_i}{\beta(x)}$$
$$\beta_{ij}(x) = -\frac{\delta_{ij}}{\beta(x)} - \frac{x_i x_j}{\beta(x)^3}.$$

Thus we find at $x = 0$,

$$\begin{aligned}
c_i &= -y_i \\
c_{ij} &= \delta_{ij}\beta(y) \\
c_{ij,k} &= -\delta_{ij}\frac{y_k}{\beta(y)} \\
c_{i,k} &= -\delta_{ik} \\
c_{i,kl} &= 0 \\
c_{ij,kl} &= -\delta_{ij}\left(\frac{\delta_{kl}}{\beta(y)} + \frac{y_k y_l}{\beta(y)^3}\right).
\end{aligned} \tag{3.2}$$

Now we can calculate for $V \in T_0\mathbb{S}^n$, $\eta \in T_0^*\mathbb{S}^n$,

$$\begin{aligned}
(MTW)_{ij}^{kl} V^i V^j \eta_k \eta_l &= -(c_{ij,pq} - c_{ij,r}c^{r,s}c_{s,pq})c^{p,k}c^{q,l}V^iV^j\eta_k\eta_l \\
&= \delta_{ij}\left(\frac{\delta_{pq}}{\beta(y)} + \frac{y_p y_q}{\beta(y)^3}\right)(-\delta^{pk})(-\delta^{ql})V^iV^j\eta_k\eta_l \\
&= \frac{|V|^2}{\beta(y)}\left(\delta_{kl} + \frac{y_k y_l}{\beta(y)^2}\right)\eta_k\eta_l \\
&\geq |V|^2|\eta|^2 \\
&= |V|^2_{\mathbb{S}^n}|\eta|^2_{\mathbb{S}^n}
\end{aligned}$$

since

$$\sum_{k,l} y_k\eta_k y_l\eta_l = \langle y, \eta\rangle^2 \geq 0,$$
$$\beta(y) \leq 1$$

for any y and η. The last equality is seen from calculation of the metric in our coordinates, shown below. □

We also make a few calculations for later use. In the above coordinates,

$$\begin{aligned}
\mathring{g}_{ij} &= \delta_{ij} + \beta_i(x)\beta_j(x) \\
(\mathring{g}_{ij})_k &= \beta_{ik}(x)\beta_j(x) + \beta_i(x)\beta_{jk}(x) \\
(\mathring{g}_{ij})_{kl} &= \beta_{ikl}(x)\beta_j(x) + \beta_{ik}(x)\beta_{jl}(x) + \beta_{il}(x)\beta_{jk}(x) + \beta_i(x)\beta_{jkl}(x)
\end{aligned}$$

and at $x = 0$ we have

$$\begin{aligned}
\mathring{g}_{ij} &= \delta_{ij} \\
(\mathring{g}_{ij})_k &= 0 \\
(\mathring{g}_{ij})_{kl} &= \delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk}.
\end{aligned}$$

Hence at $x = 0$,

$$\begin{aligned}
\mathring{g}^{ij} &= \delta_{ij} \\
(\mathring{g}^{ij})_k &= 0 \\
(\mathring{g}^{ij})_{kl} &= -\delta_{ik}\delta_{jl} - \delta_{il}\delta_{jk}.
\end{aligned} \tag{3.3}$$

Remark 3.3. *Suppose we have a C^1, c-convex function u. Then, when written in a coordinate system chosen as above centered at x_0, using the implicit relation in Definition 2.6 we notice that*

$$\left(T^+(x_0)\right)^i = u_i(x_0).$$

4 Gradient estimate

In this section, we will prove an *a priori* gradient estimate for a solution u of our elliptic equation (2.3). A gradient bound of the form $\|Du\| < 1 - \varepsilon$ will ensure that Proposition 3.2 is applicable at $(x, T^+(x))$, hence we can use the MTW theory to to obtain *a priori* second derivative estimates for u. The method we use is similar to that of Delanoë and Loeper in [1], where the cost is the geodesic distance squared. In (2.3) we will take

$$\begin{aligned}
\mu(x) &:= e^{f(x)}\sqrt{\det \mathring{g}_{ij}(x)} \\
\nu(y) &:= e^{g(y)}\sqrt{\det \mathring{g}_{ij}(y)}.
\end{aligned}$$

Theorem 4.1. *Let $n \geq 2$. Suppose u is a C^2 solution to equation* (2.3) *such that*

$$\|Du\| < 1.$$

Then,

$$\|Du\| \leq \left(\frac{1}{n-1}\right)\left(\frac{\max e^f}{\min e^g}\right)^{\frac{1}{n-1}} \left(\|Df\| + \|Dg\|\right).$$

Proof. Define

$$\phi(x) := \frac{|Du(x)|^2_{\mathbb{S}^n}}{2}$$

where u is a solution to the equation (2.3). Now we let x_0 be the point where ϕ achieves its maximum on $\mathbb{S}^n$, and take the coordinate system defined in section 3 centered at this point. We take dx^1 in the direction of Du, and rotate the remaining $n-1$ directions so that u_{ij} for $1 < i, j \le n$ is diagonal at $x_0 =$. Note that at $x = 0$,

$$\begin{aligned} 0 &= \phi_i \\ &= \frac{(\mathring{g}^{pq})_i u_p u_q + 2\mathring{g}^{pq} u_p u_{qi}}{2} \\ &= u_1 u_{1i} \end{aligned}$$

and

$$\begin{aligned} \phi_{ij} &= \frac{(\mathring{g}^{pq})_{ij} u_p u_q + 2(\mathring{g}^{pq})_i u_p u_{qj} + 2(\mathring{g}^{pq})_j u_p u_{qi} + 2\mathring{g}^{pq} u_p u_{qij} + 2\mathring{g}^{pq} u_{pj} u_{qi}}{2} \\ &= u_1^2 \frac{(\mathring{g}^{11})_{ij}}{2} + u_1 u_{1ij} + \sum_p u_{pi} u_{pj} \\ &= -u_1^2 \delta_{i1} \delta_{j1} + u_1 u_{1ij} + \sum_p u_{pi} u_{pj}. \end{aligned}$$

Here we have used (3.3). Now if $u_1(0) = 0$, that implies that $\|Du\| = 0$ and u is constant. Thus we may assume $u_1(0) \neq 0$ and hence

$$u_{1i}(0) = 0$$

for all $1 \le i \le n$. In particular, the whole matrix u_{ij} is diagonal at 0.

Consider the operator

$$Lv := w^{ij} v_{ij} \tag{4.1}$$

which is the second order part of the linearization of the natural logarithm of (2.3). Taking $v = \phi$ and $x = 0$, and applying the maximum principle we find that

$$\begin{aligned} 0 &\ge L\phi(0) \\ &= w^{ij} \phi_{ij} \\ &= w^{ij}(-u_1^2 \delta_{i1} \delta_{j1} + u_1 u_{1ij} + \sum_p u_{pi} u_{pj}) \\ &= -u_1^2 w^{11} + u_1 w^{ij} u_{1ij} + w^{ij} \sum_p u_{pi} u_{pj} \\ &= -\frac{u_1^2}{\beta(Du)} + \sum_\alpha w^{\alpha\alpha}(u_1 u_{1\alpha\alpha} + u_{\alpha\alpha}^2). \end{aligned} \tag{4.2}$$

Here we have used (3.2) and the fact that u_{ij} is diagonal at 0.

By differentiating the implicit relation

$$u_i(x) + c_i(x, T^+(x)) = 0$$

we find that

$$u_{ij} + c_{ij} = -c_{i,k}(T^{+})^k_j.$$

Thus from (3.2) and Remark 3.3,

$$\begin{aligned}(T^{+})^i_j(0) &= u_{ij}(0) + \delta_{ij}\beta\left(T^{+}(0)\right)\\ &= u_{ij}(0) + \delta_{ij}\beta\left(Du(0)\right).\end{aligned}$$

In particular, $(T^{+})^i_j(0)$ is diagonal. Note additionally, from this we find that

$$w^{ij}(0) = \frac{1}{u_{ij}(0) + \delta_{ij}\beta\left(Du(0)\right)}. \tag{4.3}$$

Also, from (3.1) we calculate

$$c_{ij1}(0,y) = -\beta_{ij1}\left(0\right)\beta\left(y\right) = 0.$$

By the calculations at the end of section 3, we have

$$\begin{aligned}e^{f(x)}\sqrt{\det \mathring{g}_{ij}(x)} &= \frac{e^{f(x)}}{\beta\left(x\right)}\\ e^{g(T^{+}(x))}\sqrt{\det \mathring{g}_{ij}(T^{+}(x))} &= \frac{e^{g(T^{+}(x))}}{\beta\left(T^{+}(x)\right)}.\end{aligned}$$

We now differentiate equation (2.3) (after taking logarithms again) in the x_1 direction to obtain, at $x = 0$,

$$\begin{aligned}0 &= w^{ij}(u_{ij1} + c_{ij1} + c_{ij,k}(T^{+})^k_1) - c^{i,j}(c_{i1,j} + c_{i,jk}(T^{+})^k_1)\\ &\qquad - D_{x_1}(f - \log\beta) + D_{x_1}(g\circ T^{+} - \log\beta\left(T^{+}\right))\\ &= \sum_{\alpha}[w^{\alpha\alpha}(u_{\alpha\alpha 1} + c_{\alpha\alpha,1}(T^{+})^1_1) + c_{\alpha 1,\alpha}]\\ &\qquad - f_1 + \frac{\beta_1}{\beta} + (g_k\circ T^{+})(T^{+})^k_1 - \frac{\beta_k\left(T^{+}\right)(T^{+})^k_1}{\beta\left(T^{+}\right)}\\ &= \sum_{\alpha}[w^{\alpha\alpha}(u_{\alpha\alpha 1} + c_{\alpha\alpha,1}(T^{+})^1_1) + c_{\alpha 1,\alpha}]\\ &\qquad - f_1 + \frac{\beta_1}{\beta} + (g_1\circ T^{+})(T^{+})^1_1 - \frac{\beta_1\left(T^{+}\right)(T^{+})^1_1}{\beta\left(T^{+}\right)}\\ &= \sum_{\alpha}[w^{\alpha\alpha}(u_{\alpha\alpha 1} - u_1)] - \frac{u_1}{\beta\left(Du\right)}\\ &\qquad - f_1 + g_1(Du)\beta\left(Du\right) + \frac{u_1}{\beta\left(Du\right)}.\end{aligned}$$

Substituting into (4.2), we obtain

$$
\begin{aligned}
0 &\geq -\frac{u_1^2}{\beta\left(Du\right)} + \sum_\alpha w^{\alpha\alpha} u_{\alpha\alpha}^2 + u_1\left(\sum_\alpha w^{\alpha\alpha} u_1 + f_1 - g_1(Du)\beta\left(Du\right)\right) \\
&= -\frac{u_1^2}{\beta\left(Du\right)} + \sum_\alpha w^{\alpha\alpha} u_{\alpha\alpha}^2 + u_1^2 \sum_\alpha w^{\alpha\alpha} + u_1 f_1 - u_1 g_1(Du)\beta\left(Du\right) \\
&\geq u_1^2 \sum_\alpha w^{\alpha\alpha} + u_1 f_1 - u_1 g_1(Du)\beta\left(Du\right) - \frac{u_1^2}{\beta\left(Du\right)} \\
&= u_1^2 \sum_{\alpha>1} w^{\alpha\alpha} + u_1 f_1 - u_1 g_1(Du)\beta\left(Du\right).
\end{aligned}
$$

Here we have used that $w^{\alpha\alpha}(0) \geq 0$ and $w^{11}(0) = \frac{1}{\beta(Du(0))}$ by (4.3). Thus we find that

$$
\begin{aligned}
\left(\|Df\| + \|Dg\|\right) &\geq |f_1| + |g_1(Du)\beta\left(Du\right)| \\
&\geq u_1 \sum_{\alpha>1} w^{\alpha\alpha} \\
&\geq u_1(n-1)\Big(\prod_{\alpha>1} w^{\alpha\alpha}\Big)^{\frac{1}{n-1}} \\
&= (n-1)u_1\left(\frac{e^{g(Du)}}{e^f}\right)^{\frac{1}{n-1}} \\
&\geq (n-1)u_1\left(\frac{\min e^g}{\max e^f}\right)^{\frac{1}{n-1}} \\
&= (n-1)\left(\frac{\min e^g}{\max e^f}\right)^{\frac{1}{n-1}} \|Du\|.
\end{aligned}
$$

Hence

$$
\|Du\| \leq \left(\frac{1}{n-1}\right)\left(\frac{\max e^f}{\min e^g}\right)^{\frac{1}{n-1}} \left(\|Df\| + \|Dg\|\right).
$$

□

5 Nonsplitting

First we define the c-subdifferential of a c-convex function u at a point x by

Definition 5.1.

$\partial_c u(x) := \{y \in \mathbb{S}^n | -c(\cdot, y) + \lambda$ *is a c-support function to u at x for some* $\lambda \in \mathbb{R}\}$.

We also define

$$
T_u^-(x) := Y^-(x, Du(x)) \tag{5.1}
$$

where Y^- is characterized by

$$\begin{aligned} D_x c(x, Y^-(x,p)) &= -p \\ \langle x, Y^-(x,p)\rangle &< 0 \end{aligned} \tag{5.2}$$

for $p \in T^*_x\mathbb{S}^n$, $|p|_{\mathbb{S}^n} < 1$ (compare Definition 2.6). We add the subscript u to emphasize the dependency on the potential function u.

Now we show a pointwise estimate on $|Du(x)|_{\mathbb{S}^n}$ if $\partial_c u(x)$ is more than one point for any x.

Lemma 5.2. *Suppose u is a C^2, c-convex solution u to* (2.3)*, and for some $x_0 \in \mathbb{S}^n$,*

$$\partial_c u(x_0) = \left\{T_u^+(x_0), T_u^-(x_0)\right\}.$$

Then,

$$\frac{e^{f(x_0)}}{e^{g(T^+(x_0))}} \geq 2^n \left(1 - |Du(x_0)|^2_{\mathbb{S}^n}\right)^{\frac{n-1}{2}}.$$

Proof. We may assume $x_0 = e_{n+1}$ by rotation. Now, we use a coordinate system on the lower hemisphere, where the point $(y, -\beta(y)) \in \mathbb{S}^n$ is represented by $y \in B_1(0) \subset \mathbb{R}^n$, and the coordinate directions are rotated so the matrix (u_{ij}) is diagonal at $x_0 = 0$. Using this we calculate, letting $y^\pm = T_u^\pm(0)$

$$c_{ij}(0, y^-) = -\delta_{ij}\beta\left(y^-\right) = -\delta_{ij}\beta\left(Du(0)\right),$$

and since $y^- \in \partial_c u(0)$ the matrix $u_{ij}(0) + c_{ij}(0, y^-)$ is positive semidefinite. Since $u_{ij}(0)$ is diagonal in these coordinates, we have

$$u_{\alpha\alpha}(0) + c_{\alpha\alpha}(0, y^-) \geq 0$$

or

$$u_{\alpha\alpha}(0) \geq \beta\left(Du(0)\right)$$

for all $1 \leq \alpha \leq n$. However, since u satisfies (2.3) we find at 0

$$\begin{aligned} \frac{e^f(0)}{e^{g(T_u^+(0))}}\beta\left(Du(0)\right) &= |\det(c_{i,j}(0, T_u^+(0))| \frac{e^{f(0)}\sqrt{\det \mathring{g}_{ij}(0)}}{e^{g(T_u^+(0))}\sqrt{\det \mathring{g}_{ij}(T_u^+(0))}} \\ &= \prod_{\alpha}^{n}(u_{\alpha\alpha}(0) + c_{ij}(0, y^+)) \\ &\geq \prod_{\alpha}^{n} 2\beta\left(Du(0)\right) \\ &= 2^n\beta\left(Du(0)\right)^n \end{aligned}$$

giving the desired estimate. ∎

6 Proof of Main Theorem

We will now use the continuity method to show the existence of a Monge solution to our problem, proving our main theorem.

We will actually show a sharper theorem from which our main theorem will follow.

Theorem 6.1. *Suppose that*

$$(\|Df\| + \|Dg\|)^2 < (n-1)^2 \left(\left(\frac{\min e^g}{\max e^f} \right)^{\frac{2}{n-1}} - 2^{\frac{-2n}{n-1}} \right)$$

and

$$(\|Df\| + \|Dg\|)^2 < (n-1)^2 \left(\left(\frac{\min e^f}{\max e^g} \right)^{\frac{2}{n-1}} - 2^{\frac{-2n}{n-1}} \right).$$

Then, there exists a C^2, c-convex solution u to the equation (2.3), *and moreover*

$$\partial_c u(x) = \{T_u^+(x)\}$$

for every $x \in \mathbb{S}^n$.

Proof. Assume $n \geq 2$, as the case $n = 1$ is vacuous.

We will apply the continuity method to the equations

$$\det\left(u_{ij}(x) + c_{ij}(x, T_u^+(x))\right) = |\det c_{i,j}(x, T_u^+(x))| \frac{e^{f_t(x)}\sqrt{\det \mathring{g}_{ij}(x)}}{e^{g_t(T_u^+(x))}\sqrt{\det \mathring{g}_{ij}(T_u^+(x))}} \tag{6.1}$$

and

$$\det\left(v_{ij}(y) + c_{ij}(T_v^+(y), y)\right) = |\det c_{i,j}(T_v^+(y), y)| \frac{e^{g_t(y)}\sqrt{\det \mathring{g}_{ij}(y)}}{e^{f_t(T_v^+(y))}\sqrt{\det \mathring{g}_{ij}(T_v^+(y))}} \tag{6.2}$$

where, for $t \in [0, 1]$:

$$f_t(x) := \log\left((1-t) + tf(x)\right)$$
$$g_t(y) := \log\left((1-t) + tg(y)\right).$$

Let

$$I := \{t \in [0,1] \mid \text{(6.1) has a smooth solution } u \text{ such that } \|Du\| < 1\}.$$

Using that $\min e^g \leq \max e^f$, by a routine calculation we find for any $t \in [0, 1]$,

$$(\|Df_t\| + \|Dg_t\|)^2 \leq (n-1)^2 \left(\left(\frac{\min e^{g_t}}{\max e^{f_t}} \right)^{\frac{2}{n-1}} - 2^{\frac{-2n}{n-1}} \right) \tag{6.3}$$

and a similar inequality holds with f_t and g_t reversed.

Then for $t \in I$, and u a solution to (6.1), by Theorem 4.1 we have

$$\begin{aligned}
\|Du\|^2 &\le \left(\frac{1}{n-1}\right)^2 \left(\frac{\max e^{f_t}}{\min e^{g_t}}\right)^{\frac{2}{n-1}} \left(\|Df_t\| + \|Dg_t\|\right)^2 \\
&< \left(\frac{1}{n-1}\right)^2 \left(\frac{\max e^{f_t}}{\min e^{g_t}}\right)^{\frac{2}{n-1}} (n-1)^2 \left(\left(\frac{\min e^{g_t}}{\max e^{f_t}}\right)^{\frac{2}{n-1}} - 2^{\frac{-2n}{n-1}}\right) \\
&= 1 - 2^{\frac{-2n}{n-1}} \left(\frac{\max e^{f_t}}{\min e^{g_t}}\right)^{\frac{2}{n-1}} \\
&\le 1 - 2^{\frac{-2n}{n-1}}.
\end{aligned}$$

Thus, from Proposition 3.2 and by the MTW maximum principle calculation in [5, Section 4], an *a priori* second derivative estimate for u follows. Higher order estimates follow by the Evans-Krylov Theorem and standard elliptic theory, thus I is closed.

To show openness, we set up the implicit function theorem as in [3, Theorem 17.6], by taking

$$G : \left\{u \in C^{2,\alpha}(S^n) : \int u dV_{\mathring{g}} = 0\right\} \times [0,1] \to \left\{v \in C^{0,\alpha}(S^n) : \int v dV_{\mathring{g}} = 0\right\}$$

to be defined as

$$G(u,t) = e^{f_t} \sqrt{\det \mathring{g}_{ij}} \left(\frac{\det\left(u_{ij} + c_{ij}(\cdot, T_u^+)\right)}{\det(-c_{i,j}(\cdot, T_u^+))} \frac{e^{g_t(T_u^+)} \sqrt{\det \mathring{g}_{ij}(T_u^+)}}{e^{f_t}\sqrt{\det \mathring{g}_{ij}}} - 1\right).$$

At a solution $u(x)$ at some time t_0, we have that $G(u, t_0) = 0$. Now the linearized operator on the first factor (whose principal part is a multiple of the operator L in (4.1)), is an elliptic operator with no zeroth order terms. Since the linearized operator has index zero, the maximum principle guarantees it is a bijection, and openness follows. Since $e^{f_0} = e^{g_0} = 1$, we may take $u = 0$ at $t = 0$ and apply the continuity method to infer the existence of smooth solutions u to (6.1) for all $t \in [0,1]$. Similarly, we obtain smooth solutions v to (6.2) for all $t \in [0,1]$.

We now prove the c-convexity of u_t, solutions to (6.1). It is clear that the set $I' := \{t \in [0,1] \mid u_t \text{ is strictly } c\text{-convex}\}$ is relatively open and contains 0. Now take any $t \in I'$. By Remark 2.7 Theorem 4.1, Lemma 5.2, and (6.3) we find that $\partial_c u_t(x) = \{T_{u_t}^+(x)\}$ for all $x \in \mathbb{S}^n$, that is, $\partial_c u_t$ is a single valued map. The strict c-convexity of u_t implies that $\partial_c u_t$, hence $T_{u_t}^+$ is injective. By (6.1), the Jacobian determinant of $T_{u_t}^+$ is nonzero, so by an open-closed argument we see that $T_{u_t}^+$ is surjective, and hence a bijection, in fact, a diffeomorphism. Thus we see that for any y,

$$u_t((T_{u_t}^+)^{-1}(y)) + c((T_{u_t}^+)^{-1}(y), y) = -(u_t)^c(y)$$

where $(u_t)^c$ is the c-transform from Definition 2.4, clearly differentiable by the above relation. Differentiating this relation twice and taking determinants of both sides, we see that $(u_t)^c$ satisfies equation (6.2). Thus, after normalizing v_t by adding an appropriate constant, we find that $v_t = (u_t)^c$.

If $I' \neq [0,1]$, let $t_0 := \inf([0,1] \setminus I') > 0$. Then u_{t_0} is c-convex but not strictly c-convex. By the uniform convergence of u_t and v_t as $t \to t_0$, we can see that $v_{t_0} = u_{t_0}^c$. As above, $\partial_c u_{t_0}^c(y) = \partial_c v_{t_0}(y) = \{T_{v_{t_0}}^+(y)\}$ for all y, which implies in turn that u_{t_0} is strictly c-convex, a contradiction. Thus, $I' = [0,1]$ and u_t is strictly c-convex for all $t \in [0,1]$.

In particular, when $t = 1$, $u := u_t$ satisfies (2.3), is c-convex, and $\partial_c u(x) = \{T_u^+(x)\}$ for all $x \in \mathbb{S}^n$. Thus we may apply [6, Theorem 5.10(ii) (replace u with $-u$)] to conclude that the measure $(\boldsymbol{Id} \times T_u^+)_\# \mu$ is a Kantorovich solution to the optimal transportation problem between μ and ν. However, since T_u^+ is a diffeomorphism, we see that it is actually a Monge solution. It is unique due to the uniqueness of the Kantorovich solution proven in [2, Theorem 2.6]. $\square$

We can now prove our main theorem, Theorem 1.1. Since μ and ν are probability measures, f and g each equal 0 at least once. Then, by assumption (1.1) on $(\|Df\| + \|Dg\|)$ we obtain that for some $0 \leq \lambda \leq 1$,

$$\begin{aligned} \min g &\geq -\lambda \left(\frac{n-1}{2}\right) \\ \max f &\leq (1-\lambda)\left(\frac{n-1}{2}\right) \\ \frac{\min e^g}{\max e^f} &\geq e^{-\lambda\left(\frac{n-1}{2}\right)-(1-\lambda)\left(\frac{n-1}{2}\right)} = e^{-\frac{n-1}{2}}. \end{aligned}$$

Hence,

$$\begin{aligned} (n-1)^2\left(\left(\frac{\min e^g}{\max e^f}\right)^{\frac{2}{n-1}} - 2^{\frac{-2n}{n-1}}\right) &\geq (n-1)^2\left(\left(e^{-\frac{n-1}{2}}\right)^{\frac{2}{n-1}} - 2^{\frac{-2n}{n-1}}\right) \\ &= (n-1)^2\left(e^{-1} - 2^{\frac{-2n}{n-1}}\right) \\ &\geq \frac{(n-1)^2}{4\pi^2} \\ &\geq (\|Df\| + \|Dg\|)^2. \end{aligned}$$

The second to last inequality can be verified by direct calculation. A calculation symmetric in f and g shows that we may now apply Theorem 6.1, proving Theorem 1.1.

References

[1] P. Delanoë and G. Loeper. Gradient estimates for potentials of invertible gradient-mappings on the sphere. *Calc. Var. Partial Differential Equations*, 26(3):297–311, 2006.

[2] W. Gangbo and R. J. McCann. Shape recognition via Wasserstein distance. *Quart. Appl. Math.*, 58(4):705–737, 2000.

[3] David Gilbarg and Neil S. Trudinger. *Elliptic partial differential equations of second order*. Classics in Mathematics. Springer-Verlag, Berlin, 2001. Reprint of the 1998 edition.

[4] G. Loeper. Regularity of optimal maps on the sphere: the quadratic cost and the reflector antenna. *Archive for Rational Mechanics and Analysis*, pages 1–21, 2010. 10.1007/s00205-010-0330-x.

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

[6] Cédric Villani. *Optimal transport : old and new*, volume 338. Springer, Berlin, 2009. Cédric Villani.; Includes bibliographical references (p. 915-956) and index.; Grundlehren der mathematischen Wissenschaften ; 338.