L^∞ -estimates in optimal transport for non quadratic costs

Published Version:
Gutierrez C.E., Montanari A. (2022). L^∞ -estimates in optimal transport for non quadratic costs. CALCULUS OF VARIATIONS AND PARTIAL DIFFERENTIAL EQUATIONS, 61(5), 1-21 [10.1007/s00526-022-02245-0].

Published:
DOI: http://doi.org/10.1007/s00526-022-02245-0

L^∞-ESTIMATES IN OPTIMAL TRANSPORT FOR NON QUADRATIC COSTS October 20, 2021

CRISTIAN E. GUTIÉRREZ AND ANNAMARIA MONTANARI

Abstract. For cost functions $c(x, y) = h(x - y)$ with $h \in C^2$ homogeneous of degree $p \geq 2$, we show L^∞-estimates of $Tx - x$ on balls, where T is an h-monotone map. Estimates for the interpolating mappings $T_t = t(T - I) + I$ are deduced from this.

Contents

1. Introduction

This note originates looking into the recent and very interesting paper by M. Goldman and F. Otto [GOdf] containing a new proof of the regularity of optimal maps for the Monge problem when the cost is quadratic. Our intention has been to investigate the validity of similar results for powers costs $|x - y|^p$ with $p \geq 2$, and in that endeavor we came up with local L^∞-estimates for monotone and interpolating maps relative to that cost, inequalities (2.5) and (3.7), respectively; these extend [GOdf, Lemma 3.1]. More generally, our estimates hold when the cost is given by a C^2 function that is homogeneous

October 20, 2021.

of degree p. Since we believe that these estimates may be useful to obtain regularity results for optimal transport when $p \neq 2$, and may have independent interest, it is our purpose to present them here. Moreover, we are able to show that these estimates suffice to prove, with modifications, several important steps in parallel with those carried out in [GOdf] toward the super-linear growth as in Prop. 3.3, eq. (3.15) of that paper; we will not provide these details in this note. However, a missing part is a replacement for $p \neq 2$ of the so called quasi-orthogonality property proved in [GOdf, Step 3, proof of Prop. 3.3]. Recent regularity results for general cost functions are considered in [OPRdf] but they do not include the case of non quadratic power costs, see Remark 3.1. We mention that global L^∞ estimates for optimal maps in terms of the p-Wasserstein distance are proved in [BJM].

The note is organized as follows. Section 2 contains a detailed proof of the L^∞-estimate (2.5) on general balls. In Section 3, we introduce a notion of monotonicity (3.1) that is equivalent to (2.2) and used it to prove in Section 3.1 the estimate (3.7) for interpolating maps. Section 3.2 shows, as a consequence, L^∞-estimates for the densities of the transport problem. Section 3.3 shows that the quantity on the right hand side of the L^∞-estimate (2.5) is comparable to an integral of a fluid flow. Section 4 is self-contained and shows an L^∞-estimate for monotone maps minus an arbitrary affine function, Lemma 4.1, which implies point-wise differentiability of locally integrable monotone maps, see Theorem 4.4 and Remark 4.8. Finally and for convenience, we include an appendix with the known formula (5.1) which is the starting point to prove the main estimate in Section 2.

Acknowledgements. We would like to thank Craig Evans for useful comments and for pointing out Krylov's work [Kry83]; see Remark 4.6. And we like to thank also Luigi Ambrosio for pointing out the connection between monotone maps and maps of bounded deformation, Remark 4.8, and useful comments. C.E.G was partially supported by NSF grant DMS–1600578, and A.M. was partially supported by a grant from GNAMPA of INdAM.

2. L^∞-estimates

If $c(x,y) : D \times D^* \to [0,+\infty)$ is a general cost function, then from optimal transport theory, the optimal map for the Monge problem is given by $T = \mathcal{N}_{c,\phi}$ where ϕ is c-concave and

$$\mathcal{N}_{c,\phi}(x) = \left\{m \in D^* : \phi(x) + \phi^c(m) = c(x,m)\right\}$$

with $\phi^c(m) = \inf_{x\in D}\left(c(x,m) - \phi(x)\right)$, see for example [GH09, Sect. 3.2]. This implies that

$$c\,(x, Tx) + c\,(y, Ty) \le c\,(x, Ty) + c\,(y, Tx) \tag{2.1}$$

assuming Tx is single valued for a.e. $x \in D$. In our analysis below we will only use that T satisfies (2.1); and *that T is optimal will not be used*.

We assume that the cost c has the form $c(x, y) = h(x - y)$ where *$h \ge 0$ is a C^2 convex function in $\mathbb{R}^n$*. What we have in mind is to obtain L^∞-estimates for $u(x) = Tx - x$, as in the paper by Goldman and Otto [GOdf, Lemma 3.1], but when *h is positively homogenous of degree p for some $1 < p < \infty$*. For this c, (2.1) obviously reads

$$h\,(x - Tx) + h\,(y - Ty) \le h\,(x - Ty) + h\,(y - Tx)\,, \tag{2.2}$$

that is, T is h-monotone, or equivalently

$$h\,(-u(x)) + h\,(-u(y)) \le h\,(x - y - u(y)) + h\,(y - x - u(x))\,. \tag{2.3}$$

Defining

$$G(a, b) = h(a - b) - h(a) - h(b),$$

and assuming that *h is even*, the inequality (2.3) reads

$$-\,G\,(x - y, u(y)) \le G\,(y - x, u(x)) + 2\,h(x - y). \tag{2.4}$$

Our purpose is then to prove the following local L^∞-estimate.

Theorem 2.1. *Suppose $h \in C^2(\mathbb{R}^n)$ is nonnegative, even, convex, positively homogeneous of degree p, for some $p \ge 2$, and $\min_{x\in S^{n-1}} h(x) = m > 0$. If T is a map satisfying the monotonicity condition (2.2) for a.e. $x, y \in \mathbb{R}^n$ and $u(x) = Tx - x$, then*

$$\sup_{y\in B_{\beta R}(x_0)} |u(y)| \le \begin{cases} L_1\, R^{n/(n+p)} \left(\fint_{B_R(x_0)} |u(x)|^p\, dx\right)^{1/(n+p)} & \text{if } \frac{1}{R^p} \fint_{B_R(x_0)} |u(x)|^p\, dx \le \left(\dfrac{1-\beta}{2}\right)^{n+p} \dfrac{(p-1)\, C_2}{(n+1)\, C_1\, \omega_n} \\ L_2 \left(R^{-1} \fint_{B_R(x_0)} |u(x)|^p\, dx\right)^{1/(p-1)} & \text{if } \frac{1}{R^p} \fint_{B_R(x_0)} |u(x)|^p\, dx \ge \left(\dfrac{1-\beta}{2}\right)^{n+p} \dfrac{(p-1)\, C_2}{(n+1)\, C_1\, \omega_n}, \end{cases} \tag{2.5}$$

for each $R > 0$, $x_0 \in \mathbb{R}^n$, and $0 < \beta < 1$ with positive constants C_1, C_2 depending only on p, n and h, with $\omega_n = |B_1|$; and with L_1 depending only on p, n and h, and L_2 depending only on p, n, h and β.

Proof. Our goal is to estimate the supremum of $|u|$ over a ball by the L^p-norm of u over a slightly larger ball. To do this, the idea is to use (5.1) and estimate the integrals by integrating (2.4) in x.

In fact, let us set $\omega = \dfrac{u(y)}{|u(y)|}$ and $r = \delta\,|u(y)|$, with $\delta > 0$ to be chosen; $u(y) \neq 0$. Applying the identity (5.1) with $v(x) \rightsquigarrow -G(x-y, u(y))$ and the ball $B_r(y) \rightsquigarrow B_r(y + r\,\omega)$ yields

$$\begin{aligned}
&v(y + r\,\omega) = -G(r\,\omega, u(y))\\
&= -\fint_{B_r(y+r\,\omega)} G\,(x-y, u(y))\,dx\\
&\qquad + \frac{n}{r^n}\int_0^r \rho^{n-1}\int_{|x-y-r\omega|\le\rho} (\Gamma(x-y-r\omega) - \Gamma(\rho))\,\Delta_x\,(-G\,(x-y, u(y)))\,dx\,d\rho\\
&= A + B. \qquad (2.6)
\end{aligned}$$

We first estimate the left hand side of (2.6) from below. Write

$$\begin{aligned}
&-G(r\,\omega, u(y))\\
&= -G(\delta u(y), u(y)) = h(\delta\,u(y)) + h(u(y)) - h\,(\delta\,u(y) - u(y))\\
&= \delta\left(\frac{h(\delta\,u(y))}{\delta} + \frac{h(u(y)) - h\,(\delta\,u(y) - u(y))}{\delta}\right)\\
&= \delta\left(\frac{h(\delta\,u(y))}{\delta} + \frac{h(-u(y)) - h\,(\delta\,u(y) - u(y))}{\delta}\right) \quad \text{since } h \text{ is even}\\
&= \delta\left(\frac{h(\delta\,u(y))}{\delta} + \frac{\nabla h\,(\xi)\cdot -\delta\,u(y)}{\delta}\right), \quad \text{with } \xi \text{ an intermediate point between } -u(y) \text{ and } \delta u(y) - u(y).
\end{aligned}$$

Since h is smooth and homogenous of degree $p > 1$, i.e., $h(\lambda x) = \lambda^p\, h(x)$ for $\lambda > 0$, it follows that $\nabla h(\lambda x) = \lambda^{p-1}\, \nabla h(x)$ and so

$$\begin{aligned}
\frac{h(\delta\, u(y))}{\delta} + \frac{\nabla h\,(\xi) \cdot -\delta\, u(y)}{\delta} &= \frac{h\left(\delta |u(y)|\, \frac{u(y)}{|u(y)|}\right)}{\delta} - \nabla h\,(\xi) \cdot u(y) \\
&= \delta^{p-1}\, |u(y)|^p\, h\left(\frac{u(y)}{|u(y)|}\right) - \nabla h\left(|\xi| \frac{\xi}{|\xi|}\right) \cdot u(y) \\
&= \delta^{p-1}\, |u(y)|^p\, h\left(\frac{u(y)}{|u(y)|}\right) - |\xi|^{p-1} \nabla h\left(\frac{\xi}{|\xi|}\right) \cdot u(y) \\
&= \delta^{p-1}\, |u(y)|^p\, h\left(\frac{u(y)}{|u(y)|}\right) - |u(y)|^p\, \left(\frac{|\xi|}{|u(y)|}\right)^{p-1}\, \nabla h\left(\frac{\xi}{|\xi|}\right) \cdot \frac{u(y)}{|u(y)|} \\
&= |u(y)|^p\, \left(\delta^{p-1}\, h\left(\frac{u(y)}{|u(y)|}\right) - \left(\frac{|\xi|}{|u(y)|}\right)^{p-1}\, \nabla h\left(\frac{\xi}{|\xi|}\right) \cdot \frac{u(y)}{|u(y)|}\right) := |u(y)|^p\, f(\delta, y).
\end{aligned}$$

If $\delta \to 0^+$ we get $\xi \to -u(y)$ and

$$f(\delta, y) = \delta^{p-1}\, h\left(\frac{u(y)}{|u(y)|}\right) - \left(\frac{|\xi|}{|u(y)|}\right)^{p-1}\, \nabla h\left(\frac{\xi}{|\xi|}\right) \cdot \frac{u(y)}{|u(y)|} \to -\nabla h\left(\frac{-u(y)}{|u(y)|}\right) \cdot \frac{u(y)}{|u(y)|}.$$

Since h is convex, then for each x_0 and x we have $h(x) \geq h(x_0) + \nabla h(x_0) \cdot (x - x_0)$. Applying this inequality with $x_0 = \dfrac{-u(y)}{|u(y)|}$ and $x = 0$ yields

$$h(0) \geq h\left(\frac{-u(y)}{|u(y)|}\right) + \nabla h\left(\frac{-u(y)}{|u(y)|}\right) \cdot \frac{u(y)}{|u(y)|}$$

and since $h(0) = 0$,

$$h\left(\frac{-u(y)}{|u(y)|}\right) \leq -\nabla h\left(\frac{-u(y)}{|u(y)|}\right) \cdot \frac{u(y)}{|u(y)|}.$$

If h is strictly positive in the unit sphere, then

$$0 < m = \min_{x \in S^{n-1}} h(x) \leq M = \max_{x \in S^{n-1}} h(x)$$

by continuity. Therefore we get the inequality

$$0 < m \leq -\nabla h\left(\frac{-u(y)}{|u(y)|}\right) \cdot \frac{u(y)}{|u(y)|} \leq \max_{x \in S^{n-1}} |\nabla h(x)|.$$

We next show that $f(\delta, y) \to -\nabla h\left(\dfrac{-u(y)}{|u(y)|}\right) \cdot \dfrac{u(y)}{|u(y)|}$ as $\delta \to 0^+$ uniformly in $y \neq 0$. In fact,

$$f(\delta, y) + \nabla h\left(\frac{-u(y)}{|u(y)|}\right) \cdot \frac{u(y)}{|u(y)|} = \delta^{p-1}\, h\left(\frac{u(y)}{|u(y)|}\right) \\ - \left(\frac{|\xi|}{|u(y)|}\right)^{p-1} \nabla h\left(\frac{\xi}{|\xi|}\right) \cdot \frac{u(y)}{|u(y)|} + \nabla h\left(\frac{-u(y)}{|u(y)|}\right) \cdot \frac{u(y)}{|u(y)|} = D_1 + D_2.$$

We have $D_1 \leq M\,\delta^{p-1}$, and from the homogeneity of ∇h

$$D_2 = -\nabla h\left(\frac{\xi}{|u(y)|}\right) \cdot \frac{u(y)}{|u(y)|} + \nabla h\left(\frac{-u(y)}{|u(y)|}\right) \cdot \frac{u(y)}{|u(y)|},$$

so

$$|D_2| \leq \left| \nabla h\left(\frac{\xi}{|u(y)|}\right) - \nabla h\left(\frac{-u(y)}{|u(y)|}\right) \right|.$$

Since ξ is an intermediate point between $-u(y)$ and $\delta u(y) - u(y)$, $\xi = -u(y) + t\,\delta\, u(y)$ for some $0 < t < 1$, so $\left| \dfrac{\xi}{|u(y)|} - \dfrac{-u(y)}{|u(y)|} \right| < \delta$. Since ∇h is uniformly continuous in a neighborhood of S^{n-1} the uniform convergence of f follows.

Therefore, we get the following lower bound for the left hand side of (2.6): there exists $\delta_0 > 0$ depending only on h and independent of y such that

$$-\,G(r\,\omega, u(y)) \geq \frac{m}{2}\,\delta\,|u(y)|^p, \qquad \text{for } 0 < \delta < \delta_0, \tag{2.7}$$

with $\omega = u(y)/|u(y)|$ and $r = \delta\,|u(y)|$, for each y with $u(y) \neq 0$. On the other hand, if $\delta \geq \delta_0$, then $\dfrac{r}{|u(y)|} \geq \delta_0$, implying obviously that $|u(y)| \leq \dfrac{r}{\delta_0}$, and obtaining the bound $|u(y)| \leq \dfrac{\alpha}{\delta_0}$ for $0 < r \leq \alpha$.

We now turn to estimate the right hand side of (2.6). Let us first calculate $\Delta_z G(z, v)$:

$$\Delta_z G(z, v) = \Delta h(z - v) - \Delta h(z).$$

Hence

$$\Delta_x\left(-G\left(x - y, u(y)\right)\right) = -(\Delta_z G)(x - y, u(y)) = \Delta h(x - y) - \Delta h(x - y - u(y)),$$

and so

$$B = \frac{n}{r^n} \int_0^r \rho^{n-1} \int_{|x-y-r\omega| \leq \rho} \left(\Gamma(x - y - r\omega) - \Gamma(\rho)\right)\, \left(\Delta h(x - y) - \Delta h(x - y - u(y))\right)\, dx\, d\rho.$$

Let us analyze the inner integral

$$I(\rho, r, y) = \int_{|x-y-r\omega| \leq \rho} \left(\Gamma(x - y - r\omega) - \Gamma(\rho)\right)\, \left(\Delta h(x - y) - \Delta h(x - y - u(y))\right)\, dx.$$

Making the change of variables $z = x - y - r\,\omega$ yields

$$I(\rho, r, y) = \int_{|z|\leq\rho} (\Gamma(z) - \Gamma(\rho))\, (\Delta h(z + r\,\omega) - \Delta h\,(z + r\,\omega - u(y)))\, dz.$$

We have that Δh is homogenous of degree $p-2$ so

$$\Delta h(z + r\,\omega) = \Delta h\left(|z + r\,\omega|\, \frac{z + r\,\omega}{|z + r\,\omega|}\right) = |z + r\,\omega|^{p-2}\, \Delta h\left(\frac{z + r\,\omega}{|z + r\,\omega|}\right).$$

Write, with e_1 a fixed unit vector in S^{n-1},

$$\begin{aligned}
&\int_{|z|\leq\rho} (\Gamma(z) - \Gamma(\rho))\, \Delta h(z + r\,\omega)\, dz \\
&= \int_{|z|\leq\rho} (\Gamma(z) - \Gamma(\rho))\, |z + r\,\omega|^{p-2}\, \Delta h\left(\frac{z + r\,\omega}{|z + r\,\omega|}\right) dz \\
&= \int_{|v|\leq\rho} (\Gamma(v) - \Gamma(\rho))\, |Tv + r\,Te_1|^{p-2}\, \Delta h\left(\frac{Tv + r\,Te_1}{|Tv + r\,Te_1|}\right) dv, \text{ with } T \text{ rotation around 0 with } Te_1 = \omega \\
&= \int_{|v|\leq\rho} (\Gamma(v) - \Gamma(\rho))\, |v + r\,e_1|^{p-2}\, \Delta h\left(\frac{Tv + r\,Te_1}{|Tv + r\,Te_1|}\right) dv.
\end{aligned}$$

Similarly,

$$\begin{aligned}
&\int_{|z|\leq\rho} (\Gamma(z) - \Gamma(\rho))\, \Delta h\,(z + r\,\omega - u(y))\, dz \\
&= \int_{|z|\leq\rho} (\Gamma(z) - \Gamma(\rho))\, |z + r\,\omega - u(y)|^{p-2}\, \Delta h\left(\frac{z + r\,\omega - u(y)}{|z + r\,\omega - u(y)|}\right) dz \\
&= \int_{|v|\leq\rho} (\Gamma(v) - \Gamma(\rho))\, |Tv + r\,Te_1 - u(y)|^{p-2}\, \Delta h\left(\frac{Tv + r\,Te_1 - u(y)}{|Tv + r\,Te_1 - u(y)|}\right) dv, \text{ with } T \text{ rotation around 0 with } Te_1 = \omega \\
&= \int_{|v|\leq\rho} (\Gamma(v) - \Gamma(\rho))\, |Tv + r\,Te_1 - |u(y)|Te_1|^{p-2}\, \Delta h\left(\frac{Tv + r\,Te_1 - |u(y)|Te_1}{|Tv + r\,Te_1 - |u(y)|Te_1|}\right) dv \\
&= \int_{|v|\leq\rho} (\Gamma(v) - \Gamma(\rho))\, |v + (r - |u(y)|)\,e_1|^{p-2}\, \Delta h\left(\frac{Tv + (r - |u(y)|)Te_1}{|Tv + (r - |u(y)|)Te_1|}\right) dv,
\end{aligned}$$

since $\omega = u(y)/|u(y)|$. Then

$$I(\rho, r, y) = \int_{|v|\leq\rho} (\Gamma(v) - \Gamma(\rho)) \left(|v + r\,e_1|^{p-2}\, \Delta h\left(\frac{Tv + r\,Te_1}{|Tv + r\,Te_1|}\right) - |v + (r - |u(y)|)\,e_1|^{p-2}\, \Delta h\left(\frac{Tv + (r - |u(y)|)Te_1}{|Tv + (r - |u(y)|)Te_1|}\right)\right) dv.$$

We then get

$$B = \frac{n}{r^n} \int_0^r \rho^{n-1}\, I(\rho, r, y)\, d\rho = n \int_0^1 t^{n-1}\, I(r\,t, r, y)\, dt.$$

Now making the change of variables $v = r\zeta$ in the integral I yields

$$
\begin{aligned}
&I(r\,t, r, y)\\
&= \int_{|\zeta|\le t} (\Gamma(r\,\zeta) - \Gamma(r\,t)) \left(|r\,\zeta + r\,e_1|^{p-2}\,\Delta h\left(\frac{T(r\,\zeta) + r\,Te_1}{|T(r\,\zeta) + r\,Te_1|}\right) - |r\,\zeta + (r - |u(y)|)\,e_1|^{p-2}\,\Delta h\left(\frac{T(r\,\zeta) + (r - |u(y)|)Te_1}{|T(r\,\zeta) + (r - |u(y)|)Te_1|}\right)\right) r^n\,d\zeta\\
&= r^p \int_{|\zeta|\le t} (\Gamma(\zeta) - \Gamma(t)) \left(|\zeta + e_1|^{p-2}\,\Delta h\left(\frac{T(\zeta) + Te_1}{|T(\zeta) + Te_1|}\right) - |\zeta + (1 - |u(y)|/r)\,e_1|^{p-2}\,\Delta h\left(\frac{T(\zeta) + (1 - |u(y)|/r)Te_1}{|T(\zeta) + (1 - |u(y)|/r)Te_1|}\right)\right) d\zeta
\end{aligned}
$$

and now letting $r = \delta\,|u(y)|$ as before yields

$$
\begin{aligned}
B &= n\,|u(y)|^p\,\delta^p \int_0^1 t^{n-1} \int_{|\zeta|\le t} (\Gamma(\zeta) - \Gamma(t))\\
&\qquad \left(|\zeta + e_1|^{p-2}\,\Delta h\left(\frac{T(\zeta) + Te_1}{|T(\zeta) + Te_1|}\right) - |\zeta + (1 - (1/\delta))\,e_1|^{p-2}\,\Delta h\left(\frac{T(\zeta) + (1 - (1/\delta))Te_1}{|T(\zeta) + (1 - (1/\delta))Te_1|}\right)\right) d\zeta\,dt\\
&= n\,|u(y)|^p\,\delta^p \int_0^1 t^{n-1} \int_{|\zeta|\le t} (\Gamma(\zeta) - \Gamma(t))\ |\zeta + e_1|^{p-2}\,\Delta h\left(\frac{T(\zeta) + Te_1}{|T(\zeta) + Te_1|}\right) d\zeta\,dt\\
&\qquad - n\,|u(y)|^p\,\delta^2 \int_0^1 t^{n-1} \int_{|\zeta|\le t} (\Gamma(\zeta) - \Gamma(t))\ |\delta\,\zeta + (\delta - 1)\,e_1|^{p-2}\,\Delta h\left(\frac{\delta\,T(\zeta) + (\delta - 1)Te_1}{|\delta\,T(\zeta) + (\delta - 1)Te_1|}\right) d\zeta\,dt\\
&= n\,|u(y)|^p\,\delta\,F(\delta),
\end{aligned}
$$

where

$$
\begin{aligned}
F(\delta) = \delta^{p-1} &\int_0^1 t^{n-1} \int_{|\zeta|\le t} (\Gamma(\zeta) - \Gamma(t))\ |\zeta + e_1|^{p-2}\,\Delta h\left(\frac{T(\zeta) + Te_1}{|T(\zeta) + Te_1|}\right) d\zeta\,dt\\
&- \delta \int_0^1 t^{n-1} \int_{|\zeta|\le t} (\Gamma(\zeta) - \Gamma(t))\ |\delta\,\zeta + (\delta - 1)\,e_1|^{p-2}\,\Delta h\left(\frac{\delta\,T(\zeta) + (\delta - 1)Te_1}{|\delta\,T(\zeta) + (\delta - 1)Te_1|}\right) d\zeta\,dt.
\end{aligned}
$$

Since Δh is continuous, it is bounded in S^{n-1} and so $F(\delta) \to 0$ uniformly in y as $\delta \to 0^+$ when $p \ge 2$. Therefore there exists $\delta_1 > 0$ such that $F(\delta) \le \dfrac{m}{4\,n}$ for $0 < \delta \le \delta_1$ and so

$$B \le \frac{m}{4}\,|u(y)|^p\,\delta$$

for $0 < \delta \le \delta_1$. Combining this with (2.7) and (2.6) yields the inequality

$$\frac{m}{4}\,|u(y)|^p\,\delta \le A, \qquad \text{for } 0 < \delta < \bar{\delta} \tag{2.8}$$

with $\bar{\delta} = \min\{\delta_0, \delta_1\}$ independent of y -depending only on n, p and h- and with $r = \delta\,|u(y)|$.

We next estimate A from above. To do this will use (2.4). From (2.6)

$$
\begin{aligned}
A &= -\fint_{B_r(y+r\,\omega)} G\,(x-y,u(y))\;dx \le \fint_{B_r(y+r\,\omega)} \left(G\,(y-x,u(x)) + 2\,h(x-y)\right)\,dx\\
&= \fint_{B_r(y+r\,\omega)} G\,(y-x,u(x))\;dx + 2\fint_{B_r(y+r\,\omega)} h(x-y)\,dx\\
&= \fint_{B_r(y+r\,\omega)} \left(h(y-x-u(x)) - h(y-x) - h(u(x))\right)\,dx + 2\fint_{B_r(y+r\,\omega)} h(x-y)\,dx\\
&= \fint_{B_r(y+r\,\omega)} h(y-x-u(x))\,dx - \fint_{B_r(y+r\,\omega)} h(u(x))\,dx + \fint_{B_r(y+r\,\omega)} h(x-y)\,dx, \quad \text{since } h \text{ is even}\\
&\le \fint_{B_r(y+r\,\omega)} h(y-x-u(x))\,dx + \fint_{B_r(y+r\,\omega)} h(x-y)\,dx, \quad \text{since } h \ge 0\\
&= A_1 + A_2.
\end{aligned}
$$

Let us estimate A_i:

$$
\begin{aligned}
A_1 &= \fint_{B_r(y+r\,\omega)} h\left(|y-x-u(x)|\,\frac{y-x-u(x)}{|y-x-u(x)|}\right)dx\\
&= \fint_{B_r(y+r\,\omega)} |y-x-u(x)|^p\, h\left(\frac{y-x-u(x)}{|y-x-u(x)|}\right)dx\\
&\le \max_{x\in S^{n-1}} h(x) \fint_{B_r(y+r\,\omega)} |y-x-u(x)|^p\,dx\\
&\le M \fint_{B_r(y+r\,\omega)} 2^{p-1}\left(|y-x|^p + |u(x)|^p\right)\,dx\\
&= 2^{p-1}\,M \fint_{B_r(y+r\,\omega)} |y-x|^p\,dx + 2^{p-1}\,M \fint_{B_r(y+r\,\omega)} |u(x)|^p\,dx;
\end{aligned}
$$

$$
A_2 = \fint_{B_r(y+r\,\omega)} h\left(|x-y|\,\frac{x-y}{|x-y|}\right)dx = \fint_{B_r(y+r\,\omega)} |x-y|^p\,h\left(\frac{x-y}{|x-y|}\right)dx \le M \fint_{B_r(y+r\,\omega)} |x-y|^p\,dx.
$$

We then obtain

$$
A \le 2^{p-1}\,M \fint_{B_r(y+r\,\omega)} |u(x)|^p\,dx + \left(2^{p-1}+1\right)M \fint_{B_r(y+r\,\omega)} |x-y|^p\,dx,
$$

with $M = \max_{x \in S^{n-1}} h(x)$. We have

$$\fint_{B_r(y+r\,\omega)} |x-y|^p\,dx = \frac{1}{|B_r(0)|}\int_{|x-y-r\omega|\le r} |x-y|^p\,dx$$
$$= \frac{1}{|B_r(0)|}\int_{|z|\le 1} |r(z+\omega)|^p\,r^n\,dz \qquad \text{with } rz = x-y-r\omega$$
$$= r^p \fint_{B_1(0)} |z+\omega|^p\,dz \le 2^p\,r^p.$$

Let us now fix a ball $B_R(x_0)$, and suppose $y \in B_{\beta R}(x_0)$ with $0 < \beta < 1$, $R > 0$. Then $B_r(y + r\,\omega) \subset B_R(x_0)$ for $r \le \dfrac{1-\beta}{2}R$ and so

$$\fint_{B_r(y+r\,\omega)} |u(x)|^p\,dx \le \frac{1}{|B_r(0)|}\int_{B_R(x_0)} |u(x)|^p\,dx.$$

Combining these estimates with the lower bound (2.8) and the upper bound for A we obtain

$$\frac{m}{4}\,|u(y)|^p\,\delta \le \frac{M_1}{r^n}\int_{B_R(x_0)} |u(x)|^p\,dx + M_2\,r^p, \qquad \text{for } 0<\delta<\bar{\delta}$$

with $\bar{\delta}$ structural constant independent of y and with $r = \delta\,|u(y)|$, for $y \in B_{\beta R}(x_0)$ and $r \le (1-\beta)R/2$; $M_1 = 2^{p-1}M/\omega_n$, $M_2 = 2^p(2^{p-1}+1)M$. Therefore, if $y \in B_{\beta R}(x_0)$, $0 < r \le (1-\beta)R/2$, and $\delta = \dfrac{r}{|u(y)|} < \bar{\delta}$, then we obtain the bound

$$|u(y)|^{p-1} \le \frac{C_1}{r^{n+1}}\int_{B_R(x_0)} |u(x)|^p\,dx + C_2\,r^{p-1} := H(r),$$

with C_i constants depending only on p, n, and M/m; $C_1 = \dfrac{2^{p+1}}{\omega_n}(M/m)$, $C_2 = 2^{p+2}(2^{p-1}+1)(M/m)$. On the other hand, if $y \in B_{\beta R}(x_0)$, $0 < r \le (1-\beta)R/2$, and $\delta = \dfrac{r}{|u(y)|} \ge \bar{\delta}$, then

$$|u(y)| \le \frac{r}{\bar{\delta}} \le \frac{1-\beta}{2\,\bar{\delta}}R.$$

So for any $y \in B_{\beta R}(x_0)$ and any $0 < r \le (1-\beta)R/2$ we obtain

$$|u(y)| \le \max\left\{H(r)^{1/(p-1)}, \frac{r}{\bar{\delta}}\right\}.$$

Since the constant C_2 in the definition of $H(r)$ can be enlarged with the last estimate remaining to hold, we can take C_2 so that $C_2 \ge 1/\bar{\delta}^{p-1}$ and in this way $H(r)^{1/(p-1)} \ge \dfrac{r}{\bar{\delta}}$, and so $\max\left\{H(r)^{1/(p-1)}, \dfrac{r}{\bar{\delta}}\right\} = H(r)^{1/(p-1)}$. Therefore we obtain the estimate

$$\sup_{y \in B_{\beta R}(x_0)} |u(y)| \le \min_{0<r\le(1-\beta)R/2} H(r)^{1/(p-1)}. \tag{2.9}$$

Set

$$\Delta = \fint_{B_R(x_0)} |u(x)|^p\, dx,$$

so $H(r) = C_1\, \Delta\, r^{-(n+1)} + C_2\, r^{p-1}$. The minimum of H over $(0, \infty)$ is attained at

$$r_0 = \left(\frac{(n+1)\, C_1\, \Delta}{(p-1)\, C_2}\right)^{1/(n+p)},$$

H is decreasing in $(0, r_0)$ and increasing in (r_0, ∞), and

$$\min_{[0,\infty)} H(r) = H(r_0) = \left(\left(\frac{n+1}{p-1}\right)^{-(n+1)/(n+p)} + \left(\frac{n+1}{p-1}\right)^{(p-1)/(n+p)}\right) (C_1\, \Delta)^{(p-1)/(n+p)}\ C_2^{(n+1)/(n+p)}.$$

If $r_0 < (1-\beta)R/2$, then $\min_{0<r<(1-\beta)R/2} H(r) = H(r_0)$. On the other hand, if $r_0 > (1-\beta)R/2$, that is, $\Delta \geq \left(\frac{1-\beta}{2}R\right)^{n+p} \frac{(p-1)\, C_2}{(n+1)\, C_1} := \Delta_0$, then we have

$$\begin{aligned}
\min_{0<r<(1-\beta)R/2} H(r) = H\left(\frac{1-\beta}{2}R\right) &= C_1\, \Delta \left(\frac{1-\beta}{2}R\right)^{-(n+1)} + C_2 \left(\frac{1-\beta}{2}R\right)^{p-1} \\
&= C_1\, \Delta \left(\frac{1-\beta}{2}R\right)^{-(n+1)} + C_2\, \Delta\, \frac{1}{\Delta} \left(\frac{1-\beta}{2}R\right)^{p-1} \\
&\leq C_1\, \Delta \left(\frac{1-\beta}{2}R\right)^{-(n+1)} + \frac{n+1}{p-1}\, C_1\, \Delta \left(\frac{1-\beta}{2}R\right)^{-(n+1)} \\
&= C_1\, \frac{p+n}{p-1} \left(\frac{1-\beta}{2}R\right)^{-(n+1)} \Delta := K_2\, R^{-(n+1)}\, \Delta.
\end{aligned}$$

We then obtain the following estimate valid for all $0 < \beta < 1$

$$\sup_{y \in B_{\beta R}(x_0)} |u(y)|^{p-1} \leq \begin{cases} K_1\, \Delta^{(p-1)/(n+p)} & \text{if } \Delta \leq \Delta_0 \\ K_2\, R^{-(n+1)}\, \Delta & \text{if } \Delta \geq \Delta_0, \end{cases} \tag{2.10}$$

with $K_1 = \left(\left(\frac{n+1}{p-1}\right)^{-(n+1)/(n+p)} + \left(\frac{n+1}{p-1}\right)^{(p-1)/(n+p)}\right) C_1^{(p-1)/(n+p)}\, C_2^{(n+1)/(n+p)}$, $K_2 = C_1\, \frac{p+n}{p-1} \left(\frac{1-\beta}{2}\right)^{-(n+1)}$, and $\Delta = \fint_{B_R(x_0)} |u(x)|^p\, dx$.

This completes the proof of the theorem. $\square$

Remark 2.2. Suppose $x_0 \in \mathbb{R}^n$, $\lim_{R\to 0^+} \frac{1}{R^p} \fint_{B_R(x_0)} |u(x)|^p\, dx = 0$ and x_0 is a Lebesgue point of $|u(x)|^p$. Then (2.5) implies that $u(x)$ is Lipschitz at x_0. In fact, first notice that since x_0 is a Lebesgue point, the condition on the limit implies $u(x_0) = 0$. Now, pick for example $\beta = 1/2$. Then there exists $R_0 > 0$ such that

$$\frac{1}{R^p} \fint_{B_R(x_0)} |u(x)|^p \, dx \leq \left(\frac{1}{4}\right)^{n+p} \frac{(p-1)\, C_2}{(n+1)\, C_1\, \omega_n}, \quad \text{for } 0 < R < R_0$$

and so $\sup_{B_{R/2}(x_0)} |u(x)| \leq C_0 R$ from (2.5) for $0 < R < R_0$, with C_0 a positive constant depending only on n, p and h. If $y \in B_{R_0/2}(x_0)$ and $R = 2\,|y - x_0|$, then $|u(y)| \leq \sup_{B_{|y-x_0|}(x_0)} |u(x)| \leq 2\, C_0\, |y - x_0|$. In particular, this implies $|Ty - Tx_0| \leq C\, |y - x_0|$ for $y \in B_{R_0/2}(x_0)$.

3. Estimates for the displacement interpolating map

In order to prove the desired estimates we first give a condition equivalent to (2.2) resembling the classical notion of monotone map. In fact, from (2.2) we can write

$$\begin{aligned}
0 &\leq h(y - Tx) - h(y - Ty) - (h(x - Tx) - h(x - Ty)) \\
&= \int_0^1 \langle Dh(y - Ty + s(Ty - Tx)), Ty - Tx \rangle ds - \int_0^1 \langle Dh(x - Ty + s(Ty - Tx)), Ty - Tx \rangle ds \\
&= \int_0^1 \langle Dh(y - Ty + s(Ty - Tx)) - Dh((x - Ty + s(Ty - Tx)), Ty - Tx \rangle ds \\
&= -\int_0^1 \int_0^1 \langle D^2h(x - Ty + s(Ty - Tx) + t(x - y))(y - x), (Ty - Tx) \rangle dt\, ds \\
&= -\int_0^1 \int_0^1 \langle D^2h(y - Ty + s(Ty - Tx) + t(x - y))(x - y), (Ty - Tx) \rangle dt\, ds \\
&= \langle A(x, y)(x - y), Tx - Ty \rangle.
\end{aligned}$$

Therefore (2.2) is equivalent to

$$\langle A(x, y)(x - y), Tx - Ty \rangle \geq 0 \tag{3.1}$$

with

$$A(x, y) = \int_0^1 \int_0^1 D^2h(y - Ty + s(Ty - Tx) + t(x - y))\, dt\, ds. \tag{3.2}$$

Let us analyze the matrix $A(x, y)$. $A(x, y)$ is clearly symmetric, and satisfies $A(x, y) = A(y, x)$ by changing variables in the integral. If h is homogenous of degree p with $p \geq 2$, then $D^2h(z)$ is homogeneous of degree $p-2$, i.e., $D^2h(\mu z) = \mu^{p-2} D^2h(z)$ for all $\mu > 0$. In addition, if *h is strictly convex*, then $D^2h(x)$ is positive definite for each $x \in S^{n-1}$, i.e, there is a constant $\lambda > 0$ such that

$$\left\langle D^2h(x)\,\xi,\xi\right\rangle \geq \lambda\,|\xi|^2$$

for all $x \in S^{n-1}$ and all $\xi \in \mathbb{R}^n$. Since h is C^2, then there is also a positive constant Λ such that

$$\lambda\,|\xi|^2 \leq \left\langle D^2h(x)\xi,\xi\right\rangle \leq \Lambda\,|\xi|^2, \qquad \forall x \in S^{n-1}, \xi \in \mathbb{R}^n. \tag{3.3}$$

We then have

$$A(x,y) = \int_0^1\int_0^1 |y-Ty+s(Ty-Tx)+t(x-y)|^{p-2}D^2h\left(\frac{y-Ty+s(Ty-Tx)+t(x-y)}{|y-Ty+s(Ty-Tx)+t(x-y)|}\right)dt\,ds$$

and

$$\lambda\,\Phi(x,y)\,|\xi|^2 \leq \langle A(x,y)\,\xi,\xi\rangle \leq \Lambda\,\Phi(x,y)\,|\xi|^2 \quad \forall \xi \in \mathbb{R}^n, \tag{3.4}$$

with

$$\Phi(x,y) = \int_0^1\int_0^1 |y-Ty+s(Ty-Tx)+t(x-y)|^{p-2}dt\,ds. \tag{3.5}$$

We also have that $\Phi(x,y) = 0$ if and only if $y-Ty+s(Ty-Tx)+t(x-y) = 0$ for all $s,t \in [0,1]$. That is, $\Phi(x,y) = 0$ if and only if $y-Ty = 0$, $Ty-Tx = 0$ and $x-y = 0$. Therefore $\Phi(x,y) > 0$ if and only if $Ty \neq y$ or $Ty \neq Tx$ or $x \neq y$.

Remark 3.1. If $c(x,y) = |x-y|^p$, then $\nabla_{xy}c(x,y) = -p\,|x-y|^{p-2}\left(Id + (p-2)\left(\frac{x-y}{|x-y|}\otimes\frac{x-y}{|x-y|}\right)\right)$ and from the Sherman-Morrison formula it follows that $\det\nabla_{xy}c(x,y) = (p-1)\left(-p\,|x-y|^{p-2}\right)^n$. So condition [OPRdf, (C_4)] does not hold for $p \neq 2$.

Remark 3.2. To illustrate the notion of h-monotonicity, suppose T satisfies (3.1) and is C^1. Then writing $y = x+\delta\,\omega$ with $|\omega| = 1$ yields

$$A(x,x+\delta\,\omega) = \iint_{[0,1]^2} D^2h\,(x+\delta\,\omega - T(x+\delta\,\omega) + s(T(x+\delta\,\omega)-Tx) + t(-\delta\,\omega))\,dt\,ds \to D^2h\,(x-Tx)$$

as $\delta \to 0$ and

$$\langle A(x,x+\delta\,\omega)(-\delta\,\omega), Tx - T(x+\delta\,\omega)\rangle \geq 0.$$

Dividing the last expression by δ^2 and letting $\delta \to 0$ we obtain

$$\left\langle D^2h\,(x-Tx)\,\omega, \frac{\partial T}{\partial x}(x)\omega\right\rangle \geq 0,$$

where $\dfrac{\partial T}{\partial x}$ is the Jacobian matrix of T evaluated at x. Since h is C^2, the matrix D^2h is symmetric and we get

$$\left\langle \omega, D^2h\,(x-Tx)\,\frac{\partial T}{\partial x}(x)\omega \right\rangle \geq 0$$

for each unit vector ω. Therefore, if T is h-monotone and C^1, the matrix $D^2h\,(x-Tx)\,\dfrac{\partial T}{\partial x}(x)$ is positive semidefinite for each x; notice that $\dfrac{\partial T}{\partial x}(x)$ is not necessarily symmetric. In particular, when $n=1$, T is h-monotone if and only if T is non decreasing.

3.1. **L^∞-estimates of the interpolating map.** Let T be a h-monotone map, i.e., satisfies (2.2), and consider the interpolating map defined by

$$T_t x = t\,Tx + (1-t)\,x, \quad 0 \leq t \leq 1. \tag{3.6}$$

Theorem 3.3. *Suppose the assumptions of Theorem 2.1 hold and assume in addition that h is strictly convex. If the integral $\mathcal{E} = \int_{B_1(0)} |Tx-x|^p\,dx$ is sufficiently small, then given $0<\beta<1$ there exists $0<\beta<\bar\beta<1$ depending only on β and the ellipticity constants λ,Λ in (3.3) such that*

$$T_t^{-1}\left(B_\beta(0)\right) \subset B_{\bar\beta}(0) \quad \textit{for all } 0 \leq t \leq 1, \tag{3.7}$$

that is, $\bigcup_{0\leq t\leq 1} T_t^{-1}\left(B_\beta(0)\right) \subset B_{\bar\beta}(0)$.

Proof. The inclusion is obvious if $t=0$. Let $x \in T_t^{-1}\left(B_\beta(0)\right)$. If $|x|\leq\beta$, then we are done. Let $\beta<\beta_0<1$, consider the ball $B_{\beta_0}(0)$, and suppose that $|x|>\beta_0$. From (2.5) applied in $B_1(0)$, we will show that is not possible if $\mathcal{E}$ is sufficiently small, i.e., smaller than $\dfrac{\lambda}{2\,\Lambda}(\beta_0-\beta)$. We have $y=T_tx\in B_\beta(0)$, and $B_r(y)\subset B_{\beta_0}(0)$ with $r=\beta_0-\beta$. Let $[y,x]$ be the straight segment between y and x, and let $z\in\partial B_r(y)\cap[y,x]$. So $|z-y|=r$, and $|z|<\beta_0$. Applying (3.1) at x,z yields

$$\begin{aligned}
0 &\leq \langle A(x,z)(Tz-Tx), z-x\rangle = \langle A(x,z)(Tz-z), z-x\rangle + \langle A(x,z)\,(z-Tx)\,, z-x\rangle\\
&= \langle A(x,z)(Tz-z), z-x\rangle + \left\langle A(x,z)\left(\frac{1}{t}\,(z-y) + \left(1-\frac{1}{t}\right)(z-x)\right), z-x\right\rangle \quad \text{since } Tx = \frac{1}{t}\,y + \left(1-\frac{1}{t}\right)x\\
&= \langle A(x,z)(Tz-z), z-x\rangle + \frac{1}{t}\,\langle A(x,z)\,(z-y)\,, z-x\rangle + \left(1-\frac{1}{t}\right)\langle A(x,z)\,(z-x)\,, z-x\rangle\\
&= \Delta.
\end{aligned}$$

Since $x \neq z$, it follows from (3.5) that $\Phi(x,z) > 0$. Also notice that $\langle A(z-x), z-y\rangle$ is bounded above by a negative quantity, where we have set $A = A(x,z)$. In fact, since z is on the segment $[y,x]$, the vectors $z-x$ and $z-y$ have opposite directions. That is, there is $\mu < 0$ such that $z - y = \mu\,(z-x)$ and so $|z-y| = -\mu\,|z-x|$. Then

$$\begin{aligned}
&\langle A(z-x), z-y\rangle = \mu\,\langle A(z-x), z-x\rangle \\
&\le \lambda\,\mu\,\Phi(x,z)\,|z-x|^2 = \lambda\,\Phi(x,z)\,\mu\,|z-x|\,|z-x|, \quad \text{from (3.4)} \\
&= -\lambda\,\Phi(x,z)\,|z-y|\,|z-x| = -\lambda\,\Phi(x,z)\,r\,|z-x|.
\end{aligned}$$

If $0 < t \le 1$, then $1 - \dfrac{1}{t} \le 0$ and and once again from (3.4)

$$0 \le \Delta \le \Lambda\,\Phi(x,z)\,|Tz-z|\,|z-x| - \frac{1}{t}\,\lambda\,\Phi(x,z)\,r\,|z-x| + \left(1-\frac{1}{t}\right)\lambda\,\Phi(x,z)\,|z-x|^2.$$

Dividing this inequality by $\Lambda\,\Phi(x,z)$ we obtain

$$\begin{aligned}
0 \le \Delta &\le |Tz-z|\,|z-x| - \frac{1}{t}\,\frac{\lambda}{\Lambda}\,r\,|z-x| + \left(1-\frac{1}{t}\right)\frac{\lambda}{\Lambda}\,|z-x|^2 \\
&= |z-x|\left(|Tz-z| - \frac{1}{t}\,\frac{\lambda}{\Lambda}\,r + \left(1-\frac{1}{t}\right)\frac{\lambda}{\Lambda}\,|z-x|\right) \\
&\le |z-x|\left(\epsilon - \frac{1}{t}\,\frac{\lambda}{\Lambda}\,r + \left(1-\frac{1}{t}\right)\frac{\lambda}{\Lambda}\,|z-x|\right) \qquad \text{if } |Tz-z| \le \epsilon \text{ from (2.5) for } \mathcal{E} \text{ small} \\
&\le |z-x|\left(-\frac{1}{t}\,\frac{\lambda}{2\,\Lambda}\,r + \left(1-\frac{1}{t}\right)\frac{\lambda}{\Lambda}\,|z-x|\right) \qquad \text{if } \epsilon \le \frac{\lambda}{2\Lambda}r\left(\le \frac{\lambda}{t\,2\,\Lambda}r\right) \\
&\le |z-x|\left(-\frac{1}{t}\,\frac{\lambda}{2\,\Lambda}\,r\right) \qquad \text{since } 1-\frac{1}{t} \le 0.
\end{aligned}$$

Hence $|z-x| = 0$, and therefore $z = x$ obtaining $|x| < \beta_0$, a contradiction. □

We now use this to obtain an estimate for $T^{-1}x - x$, when T is the optimal map for the cost $c(x,y) = h(x-y)$. We have from the theory of optimal transport that $T^{-1}(Tx) = x$ for a.e. $x \in \mathbb{R}^n$. Then given $0 < \beta < 1$ we obtain

$$\begin{aligned}
\sup_{y \in B_\beta(0)} |T^{-1}y - y| &= \sup_{T^{-1}(B_\beta(0))} |x - Tx| \\
&\le \sup_{B_{\bar\beta}(0))} |x - Tx| \quad \text{from (3.7) with } t = 1 \\
&\le C\left(\int_{B_1(0)} |Tx - x|^p\,dx\right)^{1/(n+p)} \qquad \text{from (2.5)}
\end{aligned}$$

for $\mathcal{E}$ sufficiently small and with C a constant depending only on p, n and the structural constants of h.

3.2. **L^∞-estimates of densities.** We recall that the function $F(A) = \log(\det A)$ is concave over the set of matrices A that are positive definite, i.e.,

$$F((1-t)A + tB) \geq (1-t)F(A) + tF(B), \quad 0 \leq t \leq 1.$$

Exponentiating this yields

$$\det((1-t)A + tB) \geq (\det A)^{1-t}(\det B)^t, \quad 0 \leq t \leq 1. \tag{3.8}$$

Let T be a measure preserving map (ρ_0, ρ_1), and let $T_t = tT + (1-t)Id$ be the interpolating map. Assuming the Jacobian matrix ∇T is positive definite[1], we get from (3.8) that

$$\det(\nabla T_t)(x) \geq (\det \nabla T(x))^t. \tag{3.9}$$

Let ρ_t be the measure defined by $\rho_t = (T_t)_\# \rho_0$, that is, $\rho_t(E) = \int_{(T_t)^{-1}(E)} \rho_0(x)\,dx$. Assuming invertibility of the matrices involved, changing variables yields

$$\int_{(T_t)^{-1}(E)} \rho_0(x)\,dx = \int_E \rho_0\left((T_t)^{-1}z\right) \frac{1}{\det\left((\nabla T_t)\left((T_t)^{-1}z\right)\right)}\,dz.$$

That is, the measure ρ_t has density

$$\begin{aligned} \rho(t,z) &= \rho_0\left((T_t)^{-1}z\right) \frac{1}{\det\left((\nabla T_t)\left((T_t)^{-1}z\right)\right)} \\ &\leq \rho_0\left((T_t)^{-1}z\right) \frac{1}{\left(\det\left((\nabla T)\left((T_t)^{-1}z\right)\right)\right)^t} \end{aligned} \tag{3.10}$$

from (3.9). On the other hand, since T is measure preserving

$$\rho_0(x) = \det(\nabla T(x))\,\rho_1(Tx)$$

which combined with the previous inequality yields

$$\begin{aligned} \rho(t,z) &\leq \rho_0\left((T_t)^{-1}z\right)\left(\frac{\rho_1\left(T(T_t)^{-1}z\right)}{\rho_0\left((T_t)^{-1}z\right)}\right)^t \\ &= \rho_0\left((T_t)^{-1}z\right)^{1-t} \rho_1\left(T(T_t)^{-1}z\right)^t. \end{aligned}$$

From (2.5), $T(B_{r_1}(0)) \subset B_{r_2}(0)$ for $0 < r_1 < r_2 < 1$, when $\mathcal{E} = \int_{B_1(0)} |Tx - x|^p\,dx$ is sufficiently small. And, from (3.7), $T_t^{-1}\left(B_\beta(0)\right) \subset B_{\bar\beta}(0)$ for some $0 < \beta < \bar\beta < 1$ uniform for $0 \leq t \leq 1$.

---

[1] A proof of this may be given along the lines of [Agu02, Section 5.2, Theorem 5.2.1] and [GvN07, Remark 2.9], see also [San15, Theorem 7.28, pp. 272-273] when the differentiability of c, c^* at zero is not assumed. Notice also that if h is homogenous of degree p, then h^* is homogenous of degree q with $1/p + 1/q = 1$.

Hence $T(T_t)^{-1}\left(B_\beta(0)\right) \subset B_{\beta''}(0)$ for some $0 < \beta < \bar{\beta} < \beta'' < 1$. Therefore, assuming that $\rho_0(0) = \rho_1(0) = 1$ and ρ_0, ρ_1 are Hölder continuous of order α, we obtain

$$\rho_0\left((T_t)^{-1} z\right) = 1 + \rho_0\left((T_t)^{-1} z\right) - 1 \le 1 + [\rho_0]_{\alpha,1}$$

and

$$\rho_1\left(T(T_t)^{-1} z\right) = 1 + \rho_1\left(T(T_t)^{-1} z\right) - 1 \le 1 + [\rho_1]_{\alpha,1}$$

for all $z \in B_\beta(0)$. Consequently

$$\sup_{z\in B_\beta(0)} \rho(t,z) \le \left(1 + [\rho_0]_{\alpha,1}\right)^{1-t} \left(1 + [\rho_1]_{\alpha,1}\right)^{t};$$

where $[\rho_i]_{\alpha,1} = \sup_{x,y\in B_1(0), x\neq y} \dfrac{|\rho_i(x) - \rho_i(y)|}{|x-y|^\alpha}$.

3.3. **Connection with fluids.** The connection between the Monge problem and fluid flows was discovered in [BB00] for quadratic costs. It can be seen that it holds also for general cost functions $h(x-y)$ as above. Suppose ρ_i, $i = 1,2$ are given, $v : \mathbb{R}^n \times [0,1] \to \mathbb{R}^n$ is a smooth field, and let $\rho(x,t)$ be a smooth solution of the continuity equation

$$\partial_t \rho + \operatorname{div}_x(\rho\, v) = 0 \quad \text{for } (x,t) \in \mathbb{R}^n \times [0,1] \text{ with } \rho(x,i) = \rho_i(x), i = 0,1.$$

Let T be the optimal map of the Monge problem with cost h. Given the interpolating map $T_t x = t\,Tx + (1-t)\,x$, $0 \le t \le 1$, consider the field

$$v(x,t) = (T - Id)\left(T_t^{-1} x\right),$$

and let $\rho(x,t)$ be solution to the continuity equation above with this v. Define

$$j(x,t) = \rho(x,t)\ (T - Id)\left(T_t^{-1} x\right). \tag{3.11}$$

Then

$$\begin{aligned}
\int_0^1 \int_{B_\beta} \frac{1}{\rho(x,t)^{p-1}} |j(x,t)|^p \,dxdt &= \int_0^1 \int_{B_\beta} \left|(T - Id)\left(T_t^{-1} x\right)\right|^p \rho(x,t)\,dxdt \\
&= \int_0^1 \int_{T_t^{-1}(B_\beta)} |Tz - z|^p\ \rho(T_t z, t)\,|\det \nabla T_t z|\,dzdt \\
&= \int_0^1 \int_{T_t^{-1}(B_\beta)} |Tz - z|^p\ \rho_0(z)\,dzdt \quad \text{from (3.10)} \\
&\le \int_0^1 \int_{B_{\beta'}} |Tz - z|^p\ \rho_0(z)\,dzdt \quad \text{from (3.7) for } \beta < \beta' < 1
\end{aligned}$$

assuming $\mathcal{E} = \int_{B_1(0)} |Tx - x|^p \, dx$ is sufficiently small. Here we have assumed that $\rho_0(1) = 1$ and $\rho_0 \approx 1$ in B_1.

On the other hand, if $\beta'' < \beta$ it follows from (2.5) that

$$\sup_{|x| \le \beta''} |T_t x| \le \beta'' + \sup_{|x| \le \beta''} |Tx - x| \le \beta'' + \mathcal{E}^{\text{power}>0} < \beta,$$

for $\mathcal{E}$ sufficiently small and therefore

$$\int_0^1 \int_{B_{\beta''}} |Tz - z|^p \ \rho_0(z) \, dz dt \le \int_0^1 \int_{B_\beta} \frac{1}{\rho(x,t)^{p-1}} |j(x,t)|^p \, dx dt \le \int_0^1 \int_{B_{\beta'}} |Tz - z|^p \ \rho_0(z) \, dz dt,$$

for j in (3.11).

4. Differentiability of Monotone maps

When T is monotone in the standard sense, the idea used in the proof of Theorem 2.1 can be implemented in a simpler way to obtain the following estimates for T minus a general affine function.

Lemma 4.1. *Let $A \in \mathbb{R}^{n \times n}$, $b \in \mathbb{R}^n$, T a monotone operator, $0 < \beta < 1$, and $u(x) = Tx - Ax - b$. Then there are positive constants C_1, C_2 depending only on the dimension n such that*

(a) *for $A \neq 0$ we have*

$$\sup_{y \in B_{\beta R}(x_0)} |u(y)| \le C_1 \, (\|A\| \, R)^{n/(n+1)} \left(⨍_{B_R(x_0)} |u(x)| \, dx \right)^{1/(n+1)}$$

if

$$\frac{1}{R} ⨍_{B_R(x_0)} |u(x)| \, dx \le C_2 \, \|A\| \left(\frac{1-\beta}{2} \right)^{n+1} ;$$

and

$$\sup_{y \in B_{\beta R}(x_0)} |u(y)| \le C_1 \left(\left(\frac{2}{1-\beta} \right)^n ⨍_{B_R(x_0)} |u(x)| \, dx + (1-\beta) \, R \, \|A\| \right)$$

if

$$\frac{1}{R} ⨍_{B_R(x_0)} |u(x)| \, dx \ge C_2 \, \|A\| \left(\frac{1-\beta}{2} \right)^{n+1} .$$

(b) *if $A = 0$, then*

$$\sup_{y \in B_{\beta R}(x_0)} |u(y)| \le C_1 \left(\frac{2}{1-\beta} \right)^n ⨍_{B_R(x_0)} |u(x)| \, dx.$$

Proof. By monotonicity of T,

$$(u(x) - u(y)) \cdot (x - y) \geq -\langle A(x - y), x - y\rangle, \quad \text{for a.e. } x, y, \tag{4.12}$$

which implies

$$f(x) := u(y) \cdot (x - y) \leq u(x) \cdot (x - y) + \langle A(x - y), x - y\rangle.$$

Let $r > 0$ and $z_r \in \mathbb{R}^n$ both to be determined, and consider the ball $B_r(z_r)$. The function f is harmonic in all space so integrating the last inequality for x over $B_r(z_r)$ and applying the mean value theorem yields

$$\begin{aligned} u(y) \cdot (z_r - y) &\leq \fint_{B_r(z_r)} u(x) \cdot (x - y)\, dx + \fint_{B_r(z_r)} \langle A(x - y), x - y\rangle\, dx \\ &\leq \fint_{B_r(z_r)} |u(x)|\, |x - y|\, dx + \|A\| \fint_{B_r(z_r)} |x - y|^2\, dx \\ &= B + C. \end{aligned}$$

Fix x_0, $R > 0$, and pick $r > 0$, $z_r = y + r\,\dfrac{u(y)}{|u(y)|}$ such that $B_r(z_r) \subset B_R(x_0)$; $u(y) \neq 0$. If $y \in B_{\beta R}(x_0)$, then the inclusion holds if $r < (1 - \beta)\, R/2$. Also, if $x \in B_r(z_r)$, then $|x - y| \leq 2r$. Hence

$$B \leq \frac{2\, r}{\omega_n\, r^n} \int_{B_R(x_0)} |u(x)|\, dx, \qquad C \leq 4\, \|A\|\, r^2,$$

and consequently

$$|u(y)| \leq \frac{2}{\omega_n\, r^n} \int_{B_R(x_0)} |u(x)|\, dx + 4\, \|A\|\, r := F(r) \qquad \forall y \in B_{\beta R}(x_0); \quad r \leq (1 - \beta)\, R/2.$$

We then obtain

$$\sup_{y \in B_{\beta R}(x_0)} |u(y)| \leq \min \{F(r) : 0 < r \leq (1 - \beta)\, R/2\} := m.$$

Suppose $A \neq 0$. Set $\Delta = \dfrac{2}{\omega_n} \int_{B_R(x_0)} |u(x)|\, dx$, so $F(r) = \dfrac{1}{r^n}\, \Delta + 4\, \|A\|\, r$. We have $F'(r) = -n\, r^{-(n+1)} \Delta + 4\, \|A\| = 0$ for $r = r_0 := \left(\dfrac{n\,\Delta}{4\,\|A\|}\right)^{1/(n+1)}$. So

$$\begin{aligned} &\min\{F(r) : 0 < r < +\infty\} = F(r_0) \\ &= \left(\frac{4\|A\|}{n\,\Delta}\right)^{n/(n+1)} \Delta + 4\, \|A\| \left(\frac{n\,\Delta}{4\|A\|}\right)^{1/(n+1)} \\ &= C_n\, \|A\|^{n/(n+1)} \left(\int_{B_R(x_0)} |u(x)|\, dx\right)^{1/(n+1)}. \end{aligned}$$

If $r_0 < \dfrac{1-\beta}{2}R$, then $m \le F(r_0)$ and we obtain

$$\sup_{y\in B_{\beta R}(x_0)} |u(y)| \le C_n \, (\|A\|\, R)^{n/(n+1)} \left(\fint_{B_R(x_0)} |u(x)|\, dx \right)^{1/(n+1)} \tag{4.13}$$

when $C_n \dfrac{1}{\|A\|\, R} \fint_{B_R(x_0)} |u(x)|\, dx \le \left(\dfrac{1-\beta}{2}\right)^{n+1}$; in such a case we get

$$\sup_{y\in B_{\beta R}(x_0)} |u(y)| \le C_n\, (1-\beta)\, \|A\|\, R.$$

On the other hand, if $\dfrac{1-\beta}{2}R \le r_0$, then $m = F\left(\dfrac{1-\beta}{2}R\right)$ and we get

$$\sup_{y\in B_{\beta R}(x_0)} |u(y)| \le C_n \left(\frac{2}{1-\beta}\right)^n \fint_{B_R(x_0)} |u(x)|\, dx + C_n\, (1-\beta)\, R\, \|A\|$$

when $C_n \dfrac{1}{\|A\|\, R} \fint_{B_R(x_0)} |u(x)|\, dx \ge \left(\dfrac{1-\beta}{2}\right)^{n+1}$.

If $A = 0$, then $F(r) = \dfrac{1}{r^n}\,\Delta$ is decreasing and so

$$\sup_{y\in B_{\beta R}(x_0)} |u(y)| \le m = C_n \left(\frac{2}{1-\beta}\right)^n \fint_{B_R(x_0)} |u(x)|\, dx.$$

□

Using part (b) of this lemma we will show strong differentiability of monotone maps. Following Calderón and Zygmund [CZ61], see also [Zi89, Sect. 3.5], we recall the notion of differentiability in L^p-sense.

Definition 4.2. *Let $1 \le p \le \infty$, k is a positive integer and $f \in L^p(\Omega)$, with $\Omega \subset \mathbb{R}^n$ open, and let $x_0 \in \Omega$. We say that $f \in T^{k,p}(x_0)$ $\left(f \in t^{k,p}(x_0)\right)$ if there exists a polynomial P_{x_0} of degree $\le k-1$ (P_{x_0} of degree $\le k$) such that*

$$\left(\fint_{B_r(x_0)} |f(x) - P_{x_0}(x)|^p\, dx \right)^{1/p} = O(r^k) \quad \text{as } r \to 0$$

$$\left(\left(\fint_{B_r(x_0)} |f(x) - P_{x_0}(x)|^p\, dx \right)^{1/p} = o(r^k) \quad \text{as } r \to 0 \right);$$

when $p = \infty$ the averages are replaced by $\operatorname{ess\,sup}_{x\in B_r(x_0)} |f(x) - P_{x_0}(x)| = \|f - P_{x_0}\|_{L^\infty(B_r(x_0))}$.

We mention the following landmark result of Calderón and Zygmund [CZ61, Thm. 5], see also [Zi89, Thm. 3.8.1] or [St70, Chap. VIII, Sect. 6.1]:

Theorem 4.3. *If $1 < p \leq \infty$ and $f \in T^{k,p}(x_0)$ for all $x_0 \in E$ with $E \subset \mathbb{R}^n$ measurable, then $f \in t^{k,p}(x_0)$ for almost all $x_0 \in E$; emphasizing that the orders of magnitude are not necessarily uniform in x_0*[2].

The case when $p = \infty$ is a famous theorem of Stepanov which combined with Lemma 4.1(b) yields immediately the following point-wise differentiability of monotone maps.

Theorem 4.4. *Let T be a monotone map that is locally in $L^1(\mathbb{R}^n)$*[3] *satisfying*

$$\fint_{B_R(x_0)} |Tx - b|\, dx = O(R) \quad \text{as } R \to 0 \tag{4.14}$$

for some vector $b = b_{x_0}$, i.e, $Tx \in T^{1,1}(x_0)$ for all x_0 in a measurable set E. Then

$$\|Tx - A(x - x_0) - Tx_0\|_{L^\infty(B_R(x_0))} = o(R) \quad \text{as } R \to 0$$

for almost all $x_0 \in E$, i.e., $Tx \in t^{1,\infty}(x_0)$ for a.e. $x_0 \in E$.

Proof. For each $x_0 \in E$ there exist constants $M(x_0) \geq 0$, $R_0 > 0$ and $b \in \mathbb{R}^n$ such that

$$\fint_{B_R(x_0)} |Tx - b|\, dx \leq M(x_0)\, R$$

for all $0 < R < R_0$, i.e., $Tx \in T^{1,1}(x_0)$. Since T is monotone, from Lemma 4.1(b)

$$\sup_{B_{\beta R}(x_0)} |Tx - b| \leq C(n, \beta) \fint_{B_R(x_0)} |Tx - b|\, dx \leq C(n, \beta)\, M(x_0)\, R$$

for $0 < R < R_0$. This means $\sup_{B_R(x_0)} |Tx - b| = O(R)$ as $R \to 0$ for all $x_0 \in E$, i.e., $Tx \in T^{1,\infty}(x_0)$. By Stepanov's theorem [St70, Chap. VIII, Thm. 3, p. 250] this implies that Tx is differentiable for a.e. $x_0 \in E$, i.e., $Tx \in t^{1,\infty}(x_0)$ for a.e. $x_0 \in E$.

□

[2] Whether this result holds when $p = 1$ does not seem available in the literature.

[3] In general, T is a multivalued map. However, the monotonicity implies that Tx is a singleton for a.e. $x \in \mathbb{R}^n$. Denote dom $T = \{x \in \mathbb{R}^n : Tx \neq \emptyset\}$. From [RW98, Corollary 12.38], a maximal monotone mapping T is locally bounded at $\bar{x}$ if and only if $\bar{x}$ is not a boundary point of $\overline{\text{dom}\, T}$. Also from [RW98, Thm. 12.63], if T is maximal monotone, then T is continuous at $\bar{x}$ if and only if T is single valued at $\bar{x}$, in which case necessarily $\bar{x} \in \text{int}\ (\text{dom}\, T)$. For a clear and in depth presentation of the properties of monotone maps we recommend the comprehensive book [RW98].

Remark 4.5. Notice that $\fint_{B_R(x_0)} |Tx - Ax - b|\,dx = o(R)$ (or $Tx \in t^{1,1}(x_0)$) implies (4.14) because if x_0 is a Lebesgue point, then $b = Tx_0 - Ax_0$ and

$$\begin{aligned}
\fint_{B_R(x_0)} |Tx - c|\,dx &= \fint_{B_R(x_0)} |Tx - Ax - b + Ax + b - c|\,dx \\
&\leq \fint_{B_R(x_0)} |Tx - Ax - b|\,dx + \fint_{B_R(x_0)} |Ax + b - c|\,dx \\
&= o(R) + \fint_{B_R(x_0)} |A(x - x_0)|\,dx, \quad \text{if } c = Tx_0 \\
&\leq o(R) + \|A\|\,R = O(R).
\end{aligned}$$

Remark 4.6. When T is a monotone map that is maximal, the differentiability of T a.e. was proved by Mignot [Mig76, Thm. 3.1] using Sard's Theorem; see also the more recent and perhaps simpler proof of Alberti and Ambrosio [AA99, Thm. 3.2]. When T is monotone and bounded the differentiability is proved in [Kry83, Thm. 2.2].

Remark 4.7. If ϕ is a convex function in $\mathbb{R}^n$, then from [EG92, Thm. 3, p. 240] $\nabla\phi \in BV_{\text{loc}}(\mathbb{R}^n)$. Therefore, from [EG92, Thm. 1, p. 228] $\nabla\phi$ is $L^{n/(n-1)}$-differentiable a.e., that is $\nabla\phi \in t^{1,n/(n-1)}(x)$ a.e., and since $\nabla\phi$ is monotone, it follows from Remark 4.5 and Theorem 4.4 that $\nabla\phi \in t^{1,\infty}(x)$ a.e.

Remark 4.8. Following [ACDM97], a locally integrable mapping $u : \mathbb{R}^n \to \mathbb{R}^n$ is of bounded deformation ($u \in BD$) if the symmetrized gradient in the sense of distributions $\nabla u + (\nabla u)^t$ is a Radon measure. If $T = (T_1, \cdots, T_n)$ is a monotone map in $L^1_{\text{loc}}(\mathbb{R}^n)$, it then follows from the definitions of monotonicity and distributional derivative that $T \in BD$. Because all distributional derivatives $\dfrac{\partial T_i}{\partial x_j}$ are non negative and therefore they are Radon measures. From [ACDM97, Theorem 7.4], if $T \in BD$, then $T \in t^{1,1}(x_0)$ for a.e. $x_0 \in \mathbb{R}^n$. Therefore from Remark 4.5, condition (4.14) holds for any locally integrable monotone map.

Remark 4.9. For completeness we also prove the following known fact: if $f \in L^p_{\text{loc}}(\mathbb{R}^n)$, with $p \geq 1$, then

$$\lim_{r\to 0} \left(\fint_{B_r(x_0)} |f(x) - f(x_0)|^p\,dx \right)^{1/p} = 0 \quad \text{for a.e. } x_0.$$

Define

$$\Lambda f(x_0) = \limsup_{r\to 0} \left(\fint_{B_r(x_0)} |f(x) - f(x_0)|^p\,dx \right)^{1/p}.$$

We have $0 \leq \Lambda f(x_0) \leq \limsup_{r\to 0} \left(\fint_{B_r(x_0)} |f(x)|^p\,dx \right)^{1/p} + |f(x_0)| \leq (M(|f|^p)(x_0))^{1/p} + |f(x_0)|$ with M the Hardy-Littlewood maximal function. Since $f \in L^p_{\text{loc}}(\mathbb{R}^n)$, the right hand side of the last inequality is finite for a.e. x_0 and so $\Lambda f(x_0)$ is finite for a.e. x_0. In addition, Λ is sub-linear: $\Lambda(f+g)(x_0) \leq \Lambda f(x_0) + \Lambda g(x_0)$ and $\Lambda g(x_0) = 0$ for each g continuous at x_0. By localizing f with a compact support function we may assume $f \in L^p(\mathbb{R}^n)$. Given $\varepsilon > 0$ there exists $g \in C(\mathbb{R}^n)$ such that $\|f - g\|_p \leq \varepsilon$. For each $\alpha > 0$ we then have

$$\begin{aligned}\{x : \Lambda f(x) > \alpha\} &\subset \{x : \Lambda(f-g)(x) > \alpha/2\} \cup \{x : \Lambda g(x) > \alpha\} = \{x : \Lambda(f-g)(x) > \alpha/2\} \\ &\subset \{x : (M(|f-g|^p)(x))^{1/p} > \alpha/4\} \cup \{x : |f(x) - g(x)| > \alpha/4\}\end{aligned}$$

and so

$$\begin{aligned}|\{x : \Lambda f(x) > \alpha\}| &\leq |\{x : M(|f-g|^p)(x) > (\alpha/4)^p\}| + |\{x : |f(x)-g(x)| > \alpha/4\}| \\ &\leq \frac{C_n}{\alpha^p}\|f-g\|_p^p + \frac{4^p}{\alpha^p}\|f-g\|_p^p \leq \frac{C}{\alpha^p}\,\varepsilon^p.\end{aligned}$$

Since ε is arbitrary, we obtain $\Lambda f(x) = 0$ for a.e. x.

5. Appendix

Recall that $\Gamma(x) = \dfrac{1}{n\omega_n(2-n)}|x|^{2-n}$, with $n > 2$ where ω_n is the volume of the unit ball in $\mathbb{R}^n$, and the Green's representation formula

$$v(y) = \int_{\partial\Omega} \left(v(x)\,\frac{\partial \Gamma}{\partial \nu}(x-y) - \Gamma(x-y)\,\frac{\partial v}{\partial \nu}(x) \right) d\sigma(x) + \int_\Omega \Gamma(x-y)\,\Delta v(x)\,dx$$

where ν is the outer unit normal and $y \in \Omega$. If $\Omega = B_\rho(y)$, then $\dfrac{\partial \Gamma}{\partial \nu}(x-y) = \dfrac{1}{n\,\omega_n}\,|x-y|^{1-n}$ and so the representation formula reads

$$\begin{aligned}v(y) &= \fint_{|x-y|=\rho} v(x)\,d\sigma(x) - \Gamma(\rho) \int_{|x-y|=\rho} \frac{\partial v}{\partial \nu}(x)\,d\sigma(x) + \int_{|x-y|\leq\rho} \Gamma(x-y)\,\Delta v(x)\,dx \\ &= \fint_{|x-y|=\rho} v(x)\,d\sigma(x) + \int_{|x-y|\leq\rho} \left(\Gamma(x-y) - \Gamma(\rho)\right)\,\Delta v(x)\,dx\end{aligned}$$

from the divergence theorem. Multiplying the last identity by ρ^{n-1} and integrating over $0 \leq \rho \leq r$ yields

$$v(y) = \fint_{|x-y|\leq r} v(x)\,dx + \frac{n}{r^n}\int_0^r \rho^{n-1} \int_{|x-y|\leq\rho} \left(\Gamma(x-y) - \Gamma(\rho)\right)\,\Delta v(x)\,dx\,d\rho. \tag{5.1}$$

References

[Agu02] Martial Agueh, *Existence of solutions to degenerate parabolic equations via the Monge-Kantorovich theory*. PhD thesis, Georgia Tech, April 2002.

[AA99] G. Alberti and L. Ambrosio, *A geometrical approach to monotone functions in* $\mathbb{R}^n$, Mathematische Zeitschrift **230** (1999), 259–316.

[ACDM97] L. Ambrosio, A. Coscia and G. Dal Maso, *Fine properties of functions with bounded deformation*, Arch. Rational Mech. Anal., **139** (1997), 201–238.

[BB00] Jean-David Benamou and Yann Brenier, A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem. *Numerische Matematik*, 84(3):375–393, 2000.

[BJM] G. Bouchitté, C. Jimenez, and R. Mahadevan, A new L^∞ estimate in optimal mass transport. *Proc. Amer. Math. Soc.*, 135(11):3525–3535, 2007.

[CZ61] A. P. Calderón and A. Zygmund, *Local Properties of Solutions of Elliptic Partial Differential Equations*, Studia Math. **XX** (1961), 171–225.

[EG92] L. C. Evans and R. F. Gariepy, *Measure Theory and Fine Properties of Functions*. CRC Press, 1992.

[GH09] C. E. Gutiérrez and Qingbo Huang, The refractor problem in reshaping light beams. *Arch. Rational Mech. Anal.*, 193(2):423–443, 2009.

[GOdf] Michael Goldman and Felix Otto, A variational proof of partial regularity for optimal transportation maps. *Ann. Scient. Ec. Norm. Sup.*, 53:1209–1233, 2020, https://arxiv.org/pdf/1704.05339.pdf.

[GvN07] Cristian E. Gutiérrez and Truyen van Nguyen, On Monge–Ampère type equations arising in optimal transportation problems. *Calculus of Variations and Partial Differential Equations*, 28(3):275–316, 2007.

[Mig76] F. Mignot, *Contrôle dans les Inéquations Variationelles Elliptiques*, J. Funct. Anal. **22** (1976), 130–185.

[Kry83] N. Krylov, *Properties of monotone mappings*, Lithuanian Mathematical Journal, 1983, **22**(2), 140–145.

[OPRdf] Felix Otto, Maxime Prod'Homme, and Tobias Ried, Variational approach to regularity of optimal transport maps: general cost functions. September 2020, https://arxiv.org/pdf/2007.11549.pdf.

[RW98] R. T. Rockafellar and R. J-B. Wets, *Variational Analysis*, Grundlehren der mathematischen Wissenschaften, Vol. 317, Springer, 1998, Corrected 3rd. printing 2009. https://sites.math.washington.edu/~rtr/papers/rtr169-VarAnalysis-RockWets.pdf

[San15] Filippo Santambrogio, *Optimal transport for applied mathematicians*, volume 87 of *Progress in Nonlinear Differential Equations and Their Applications*. Birkhäuser, Boston, MA, 2015.

[St70] E.M. Stein, *Singular Integrals and Differentiability Properties of Functions*, Princeton University Press, Princeton, NJ, 1970.

[Zi89] W.P. Ziemer, *Weakly Differentiable Functions*, Graduate Texts in Mathematics, vol. 120, Springer-Verlag, 1989.

Department of Mathematics, Temple University, Philadelphia, PA 19122

Email address: cristian.gutierrez@temple.edu

Dipartimento di Matematica, Piazza di Porta San Donato 5, Università di Bologna, 40126 Bologna, Italy

Email address: annamaria.montanari@unibo.it