Geometriae Dedicata **74:** 201–212, 1999.

A Remarkable Measure Preserving Diffeomorphism Between Two Convex Bodies in $\mathbb{R}^n$

S. ALESKER*, S. DAR and V. MILMAN**
Department of Mathematics, Tel-Aviv University, Ramat Aviv, 69978 Tel-Aviv, Israel;
e-mail: vitali@math.tau.ac.il

(Received: 26 January 1998; revised version: 16 March 1998)

Abstract. We prove that for any two convex open bounded bodies K and T there exists a diffeomorphism $f\colon K \to T$ preserving volume ratio (i.e. with constant determinant of the Jacobian) and such that the Minkowski sum $K + T = \{x + f(x) | x \in K\}$. As an application of this method, we prove some of the Alexandov–Fenchel inequalities.

Mathematics Subject Classification (1991): 52Axx.

Key words: Brenier map, Alexandrov–Fenchel inequalities.

1. Introduction

Let K and T be convex open bounded bodies in $\mathbb{R}^n$. We will discuss in this note C^1-diffeomorphic bijections $\Phi\colon K \to T$ with constant determinant of the Jacobian. So, if $\operatorname{vol} K = \operatorname{vol} T$ our maps are measure preserving. There are several such distinguished maps. For example, a Knothe map ([K]) $\Phi\colon K \to T$ has a triangular Jacobian with a positive diagonal and may be naturally associated with any fixed orthogonal basis in $\mathbb{R}^n$. So, there are many different such maps (see [MSch, App. 1], or [Bo] where a Knothe map was used to estimate integrals of polynomials over convex bodies). Recently, the so-called Brenier map [Br] was introduced and used successfully. We call $b\colon K \to T$ the Brenier map if there is a convex function $f \in C^2(K)$ defined on K such that $\nabla f = b\colon K \to T$ has the above properties (bijection which preserves ratio of volumes). Existence and uniqueness of such a map was proved in [Br] (see another proof in [McC1] where more general facts were also proved, and the proof of regularity in [C1-3]).

Naturally, the Jacobian matrix $\operatorname{Jac} b = \operatorname{Hess} f$ is a symmetric positive definite matrix. One of the (many) applications of such maps is in proving the Brunn–Minkowski inequality (and also in analyzing equality cases and their stability).

---

* Partially supported by a Minkowski Center grant.
** Research partially supported by the Israel Science Foundation founded by the Academy of Sciences and Humanities.

The scheme works in the following way. Instead of considering the Minkowski sum $K + T$ (and estimating its volume) we consider, a priori, a smaller set

$$S = \{x + \Phi(x) | x \in K\} \subset K + T,$$

and estimate its volume using the geometric-arithmetic mean inequality under the integral.

However, a priori, the set S may be strictly smaller than the Minkowski sum and then a conveniently described set S looses information with respect to an (inconveniently described) Minkowski sum.

So, the following question naturally arises:

For any K and T, does there exist a C^1-map Φ (with constant determinant of Jacobian) $\Phi: K \to T$ such that $K + T = \{x + \Phi(x) | x \in K\}$?

Below, we answer this question positively just by combining a few known (recent) results. In the proof we use an unpublished result on regularity by L. Caffarelli and we thank Professor Caffarelli for communicating his result to us and suggesting to add its proof in the Appendix of this note.

We also show one application of this fact by proving some cases of the Alexandrov–Fenchel inequality.

It was observed by R. McCann ([McC2, Remark 2.5]) that the Brenier map itself may not satisfy the above property, i.e. one can construct K, T and the Brenier map $b: K \to T$ such that $\{x + b(x) | x \in K\} \subsetneq K + T$.

His example is as follows: let us take K to be the Euclidean ball and T to be an ellipsoid of the same volume, then the Brenier map b is a gradient of some quadratic form, i.e. it is a linear map. Then $(Id + b)(K)$ is an ellipsoid, but $K + T$ is not necessarily an ellipsoid.

Let us state the existence theorem of the Brenier map in the form due to McCann [McC1] and discuss some of the properties of the Brenier map.

Let φ be a convex function on $\mathbb{R}^n$ (which might be infinite outside of some convex set). Let us denote by $\nabla\varphi(x)$ the subdifferential of φ at the point x, namely this is the set of all supporting hyperplanes of φ at x

$$\nabla\varphi(x) := \{z \in \mathbb{R}^n | \varphi(y) \geqslant \varphi(x) + \langle z, y - x\rangle \ \ \forall y\}.$$

THEOREM 1.1 ([McC1]). *Let μ, ν be probability measures on $\mathbb{R}^n$ such that μ vanishes on Borel subsets of $\mathbb{R}^n$ having Hausdorff dimension $n - 1$ (in particular, this condition is satisfied if μ is absolutely continuous with respect to the Lebesgue measure). Then there exists a convex function φ such that for every Borel subset $A \subset \mathbb{R}^n$ $\nu(A) = \mu(\nabla\varphi^{-1}(A))$, where*

$$\nabla\varphi^{-1}(A) = \{x \in \mathbb{R}^n | \nabla\varphi(x) \cap A \neq \emptyset\}.$$

In other words, $\nabla\varphi$ pushes forward μ to ν.

Remark. Under the assumptions on the measure μ, $\nabla\varphi(x)$ is a single point set for μ-almost every $x \in \mathbb{R}^n$ (see [AK]). So we may think of $\nabla\varphi$ as the usual map $\nabla\varphi: \mathbb{R}^n \to \mathbb{R}^n$.

The following proposition is certainly well-known, but since we have no exact reference, we outline its proof (under weaker conditions, a similar statement is explicitly given in Brenier's paper [Br]).

PROPOSITION 1.2. *Assume that μ, ν be probability measures on $\mathbb{R}^n$, which are absolutely continuous with respect to the Lebesgue measure. Then the Brenier map $\nabla\varphi$ has an inverse $(\nabla\varphi)^{-1}$ which is defined ν-almost everywhere, which is also a Brenier map (pushing forward ν to μ).*

Proof. In fact, we just outline the proof and refer the reader to the paper by McCann [McC1] for details. In order to distinguish between two copies of $\mathbb{R}^n$ we will denote each of them X and Y correspondingly.

In his proof of existence of the Brenier map, McCann shows that there exists a probability measure γ on $X \times Y$ with cyclically monotone support (see [McC1] for the definition and references), and μ, ν are its marginals, namely for every Borel subset $A, B \subset \mathbb{R}^n$

$$\gamma(A \times Y) = \mu(A), \qquad \gamma(X \times B) = \nu(B).$$

By the Rockafellar theorem [R] there exists a convex function $\varphi: X \to \mathbb{R} \cup \{+\infty\}$ such that

$$\operatorname{supp} \gamma \subset \partial\varphi, \tag{$*$}$$

where $\partial\varphi = \{(x, y): y \in \nabla\varphi(x)\}$. Since $\nabla\varphi$ is a single point set μ-almost everywhere (see [AK]), $\nabla\varphi$ is a well defined map $X \to Y$, and because of $(*)$ $\nabla\varphi$ pushes μ forward to ν. But since the condition of cyclic monotonicity of a subset of $\mathbb{R}^n \times \mathbb{R}^n$ is symmetric with respect to both coordinates, the same argument implies that there exists a convex function $\psi: Y \to \mathbb{R} \cup \{\infty\}$ such that

$$\operatorname{supp} \gamma \subset \partial\psi. \tag{$**$}$$

Again $\nabla\psi$ is a well defined map $Y \to X$ ν-almost everywhere and it pushes ν forward to μ. Conditions $(*)$ and $(**)$ easily imply that $\nabla\varphi, \nabla\psi$ are inverse to each other almost everywhere. □

We now describe a regularity result of Caffarelli [C4]. For completeness of our argument we provide in the Appendix the proof of this result which L. Caffarelli sketched for us. We thank him for permission to use it in this Appendix.

Let $\Gamma \subset \mathbb{R}^n$ be a convex bounded open set. Let $f(x)$ be a probability density on $\mathbb{R}^n$, $g(y)$ be a probability density on Γ so that

$$\mathrm{d}\mu = f \cdot \mathrm{d}m, \qquad \mathrm{d}\nu = g \cdot \mathrm{d}m,$$

where m is the Lebesgue measure. Assume that

(i) $f(x)$ is locally bounded and bounded away from zero on compact sets, i.e. for every $R > 0$

$$0 < C(R) \leqslant f(x) \leqslant C(R) \quad \text{for } |x| \leqslant R;$$

(ii) there exist positive constants $\Lambda, \lambda > 0$ such that for every $y \in \Gamma$, $\lambda \leqslant g(y) \leqslant \Lambda$.

THEOREM 1.3 (Caffarelli [C4]). (a) *Under the conditions* (i) *and* (ii) *the Brenier map*

$$\nabla\varphi : (\mathbb{R}^n, f\,\mathrm{d}x) \to (\mathbb{R}^n, g\,\mathrm{d}x)$$

is continuous. Moreover, $\nabla\varphi$ belongs locally to the Hölder class C^α for some $\alpha >$ 0.

(b) *If f, g are locally Hölder, then the solution φ belongs to $C^{2,\alpha}$ for some $\alpha > 0$ (locally).*

2. The Main Result

THEOREM 2.1. *Let K and T be open convex bounded subsets of $\mathbb{R}^n$ of volume* 1. *Then there exists a C^1-diffeomorphism $F: K \to T$ preserving the Lebesgue measure such that, for every $\lambda > 0$,*

$$\{x + \lambda F(x) | x \in K\} = K + \lambda T.$$

Before proving this result let us recall a theorem which can be found in [Gr] (in fact, under weaker assumptions on the potential f).

PROPOSITION 2.2 [Gr]. (i) *Let $f: \mathbb{R}^n \to \mathbb{R}$ be a C^2-smooth convex function with strictly positive Hessian. Then the image of the gradient map $K = \operatorname{Im} \nabla f$ is an open convex set.*

(ii) *Furthermore, if $f_1, f_2: \mathbb{R}^n \to \mathbb{R}$ are two such functions and $K_i = \operatorname{Im}(\nabla f_i)$, $i = 1, 2$, then*

$$\operatorname{Im}(\nabla f_1 + \nabla f_2) = K_1 + K_2.$$

Proof of Theorem 2.1. Consider on $\mathbb{R}^n$ any probability measure with smooth strictly positive density ρ.

Consider the Brenier maps

$$\nabla f_1 \colon (\mathbb{R}^n, \rho\, \mathrm{d}x) \to (K, \mathrm{d}x),$$
$$\nabla f_2 \colon (\mathbb{R}^n, \rho\, \mathrm{d}x) \to (T, \mathrm{d}x),$$

where dx denotes the Lebesgue measure. They are C^1-smooth by Theorem 1.3.

By Proposition 2.2 for every $\lambda > 0$

$$K + \lambda T = \{\nabla f_1(x) + \lambda \nabla f_2(x) | x \in \mathbb{R}^n\}.$$

Consider $F \colon K \to T$ defined by $F = \nabla f_2 \circ (\nabla f_1)^{-1}$. Obviously it satisfies all the conditions of the theorem. □

To prove Proposition 2.2 one needs

LEMMA 2.3. *Suppose $f \colon \mathbb{R}^n \to \mathbb{R}$ is a convex C^1-function defined on the whole domain $\mathbb{R}^n$ and $f(x) \geqslant 0$ for every $x \in \mathbb{R}^n$. Then $\inf_{x\in\mathbb{R}^n} \|\nabla f(x)\| = 0$, where $\|\cdot\|$ denotes the Euclidean norm on $\mathbb{R}^n$.*

Proof. Let $\alpha = \frac{1}{2} \inf_{x\in\mathbb{R}^n} \|\nabla f(x)\|$ and suppose $\alpha > 0$. Then for every $x \in \mathbb{R}^n$ there exists some $y \neq x$ such that $f(x) - f(y) \geqslant \alpha \cdot \|x - y\|$.

Define now $L = \{x \in \mathbb{R}^n;\ f(0) - f(x) \geqslant \alpha \cdot \|x\|\}$. Then L is a nonempty compact set. Hence, f achieves its minimum on L at some point $x_0 \in L$.

But then there exists some $y_0 \neq x_0$ such that $f(x_0) - f(y_0) \geqslant \alpha \cdot \|x_0 - y_0\|$ and then

$$\begin{aligned} f(0) - f(y_0) &= [f(0) - f(x_0)] + [f(x_0) - f(y_0)] \\ &\geqslant \alpha \cdot \|x_0\| + \alpha \cdot \|x_0 - y_0\| \geqslant \alpha \cdot \max\{\|x_0 - y_0\|, \|y_0\|\}. \end{aligned}$$

So

$$y_0 \in L \quad \text{and} \quad f(y_0) \leqslant f(x_0) - \alpha \cdot \|x_0 - y_0\| < f(x_0)$$

and we get a contradiction. □

We prove here part (ii) of Proposition 2.2 only. Obviously

$$\begin{aligned} \nabla(f_1 + f_2)(\mathbb{R}^n) &= \{\nabla f_1(x) + \nabla f_2(x) | x \in \mathbb{R}^n\} \\ &\subset \nabla f_1(\mathbb{R}^n) + \nabla f_2(\mathbb{R}^n) = K_1 + K_2. \end{aligned}$$

On the other hand, given $u_1 \in K_1$, $u_2 \in K_2$, there exist $x_1, x_2 \in \mathbb{R}^n$ such that $u_i = \nabla f_i(x_i)$, $i = 1, 2$. Then for every $x \in \mathbb{R}^n$ we have

$$\begin{aligned} &(f_1 + f_2)(x) - \langle u_1 + u_2, x\rangle - \\ &\quad -[f_1(x_1) + f_2(x_2) - \langle u_1, x_1\rangle - \langle u_2, x_2\rangle] \\ &= [f_1(x) - f_1(x_1) - \langle \nabla f_1(x_1), x - x_1\rangle] + \\ &\quad +[f_2(x) - f_2(x_2) - \langle \nabla f_2(x_2), x - x_2\rangle] \geqslant 0. \end{aligned}$$

Thus, $\inf_{x\in\mathbb{R}^n}[(f_1+f_2)(x)-\langle u_1+u_2,x\rangle] > -\infty$. It easily follows from Lemma 2.3 that

$$u_1+u_2 \in \overline{\nabla(f_1+f_2)(\mathbb{R}^n)}.$$

Thus $\nabla(f_1+f_2)(\mathbb{R}^n) \subset K_1+K_2 \subset \overline{\nabla(f_1+f_2)(\mathbb{R}^n)}$.

Since f_1, f_2 are C^2-smooth and have strictly positive Hessian, then the image of $\nabla(f_1+f_2)$ is an open convex set by part (i) of Proposition 2.2 (which we do not prove here). Hence

$$K_1+K_2 = \nabla(f_1+f_2)(\mathbb{R}^n). \qquad \square$$

3. Application to the Alexandrov–Fenchel Inequalities

Using the above method let us prove some of the Alexandrov–Fenchel inequalities.

THEOREM 3.1. *Let $K_1,\dots,K_n$ be convex compact subsets in $\mathbb{R}^n$. Then*

$$V(K_1,\dots,K_n) \geqslant \left(\prod_{r=1}^{n}|K_r|\right)^{1/n},$$

where $V(K_1,\dots,K_n)$ denotes the mixed volume of $K_1,\dots,K_n$.

Proof. By homogeneity we can normalize the volumes of $K_1,\dots,K_n$ to equal 1. Let us fix a probability measure on $\mathbb{R}^n$ with strictly positive smooth density ρ. It is enough for our purposes to consider, say, the standard Gaussian measure on $\mathbb{R}^n$. Consider the Brenier maps

$$\nabla f_r\colon (\mathbb{R}^n,\rho\,\mathrm{d}x) \to (K_r,\mathrm{d}x), \quad r=1,\dots,n.$$

They are C^1-smooth. For all $t_r>0$

$$\det\left[\sum_{r=1}^{n} t_r\left(\frac{\partial^2 f_r}{\partial x_i\partial x_j}\right)\right] = \sum_{r_1,\dots,r_n} t_{r_1}\dots t_{r_n} D\left(\frac{\partial^2 f_{r_1}}{\partial x_i\partial x_j},\dots,\frac{\partial^2 f_{r_n}}{\partial x_i\partial x_j}\right),$$

where $D(\cdot,\dots,\cdot)$ is the mixed discriminant of the corresponding matrices (see [A] or [Sch] or [H]). Note that

$$\mathrm{Det}\left(\frac{\partial^2 f_r}{\partial x_i\partial x_j}\right)(x) = \rho(x),$$

because ∇f_r is a measure preserving map. We obtain

$$\left|\sum_{r=1}^n t_r K_r\right| = \int_{\mathbb{R}^n} \det\left[\sum_r t_r \left(\frac{\partial^2 f_r(x)}{\partial x_i \partial x_j}\right)\right] \mathrm{d}x$$
$$= \sum_{r_1,\dots,r_n} t_{r_1}\dots t_{r_n} \int_{\mathbb{R}^n} D\left(\frac{\partial^2 f_{r_1}(x)}{\partial x_i \partial x_j},\dots,\frac{\partial^2 f_{r_n}(x)}{\partial x_i \partial x_j}\right) \mathrm{d}x.$$

Hence, we obtain the following expression for the mixed volumes

$$V(K_1,\dots,K_n) = \int_{\mathbb{R}^n} D\left(\frac{\partial^2 f_1(x)}{\partial x_i \partial x_j},\dots,\frac{\partial^2 f_n(x)}{\partial x_i \partial x_j}\right) \mathrm{d}x.$$

Let us apply the following particular case of the Alexandrov inequalities for mixed discriminants, which states that if $A_1,\dots,A_n$ are positive definite matrices, then

$$D(A_1,\dots,A_n) \geqslant \left(\prod_{r=1}^n \det A_r\right)^{1/n}$$

(we refer to [A], [Bu, Theorem 7.17] or [H, pp. 64–65]). Hence,

$$V(K_1,\dots,K_n) \geqslant \int_{\mathbb{R}^n} \left(\prod_{r=1}^n \det\left(\frac{\partial^2 f_r(x)}{\partial x_i \partial x_j}\right)\right)^{1/n} \mathrm{d}x$$
$$= \int_{\mathbb{R}^n} \rho(x)\,\mathrm{d}x = 1. \qquad \square$$

Remark 1. R. Schneider pointed out to us that the above method, in combination with a description of the equality cases in Alexandrov's inequalities for mixed discriminants of positive definite matrices, gives a description of the equality cases in the inequalities of Theorem 3.1. Namely, if $K_1,\dots,K_n$ are non-degenerate convex bodies (i.e. with non-empty interior) and if

$$V(K_1,\dots,K_n) = \left(\prod_{r-1}^n |K_r|\right)^{1/n},$$

then the bodies $K_1,\dots,K_n$ are pairwise homothetic. This follows from the fact that if $A_1,\dots,A_n$ are positive definite matrices and

$$D(A_1,\dots,A_n) = \left(\prod_{r=1}^n \det A_r\right)^{1/n},$$

then $A_1,\dots,A_n$ are pairwise proportional (see [A]). Then the above argument (with the normalization $|K_r| = 1$) implies that the matrices $(\partial^2 f_r(x)/\partial x_i \partial x_j)$ are equal one to the other for every x (since ∇f_r are measure preserving) and, hence, ∇f_r coincide up to a linear functional, i.e. K_r coincide up to translation.

Note that the classical proof of Theorem 3.1 shows that equality in this inequality implies equality in the Minkowski inequality for any two of the bodies, which in turn implies that the bodies must be homothetic (see [Bu, p. 50], or [Sch]).

Remark 2. A particular case of Theorem 3.1 is the case when one considers only two bodies K_1, K_2 and the mixed volume of K_1 taken k times with K_2 taken $n-k$ times, which is denoted by $V_k(K_1, K_2)$. Then

$$V_k(K_1, K_2) \geqslant |K_1|^{k/n}|K_2|^{(n-k)/n}.$$

Here one should use Alexandrov's inequalities for mixed discriminants of two positive definite matrices only. Namely, if A_1, A_2 are two positive definite matrices, then

$$D_k(A_1, A_2) \geqslant (\det A_1)^{k/n}(\det A_2)^{(n-k)/n},$$

where the left-hand side denotes the mixed discriminant of A_1 taken k times with A_2 taken $n-k$ times. This fact is much simpler and has a proof based on the arithmetic-geometric mean inequality only. Indeed, two positive definite matrices can be diagonalized simultaneously by a linear transformation with determinant one. So one may assume A_1, A_2 to be diagonal. Let

$$A_1 = \operatorname{diag}(\alpha_1, \dots, \alpha_n), \qquad A_2 = \operatorname{diag}(\beta_1, \dots, \beta_n).$$

Then, for every t,

$$\begin{aligned}\det(A_1 + tA_2) &= \prod_{r=1}^{n}(\alpha_r + t\beta_r)\\ &= \sum_{k=0}^{n}\left(\sum_{\substack{I\subset\{1,\dots,n\}\\ |I|=k}}\left(\prod_{i\in I}\alpha_i\right)\left(\prod_{i\notin I}\beta_i\right)\right)t^{n-k}.\end{aligned}$$

Then by the definition of the mixed discriminant

$$D_k(A_1, A_2) = \binom{n}{k}^{-1}\sum_{\substack{I\in\{1,\dots,n\}\\ |I|=k}}\left(\prod_{i\in I}\alpha_i\right)\left(\prod_{i\notin I}\beta_i\right).$$

By the arithmetic-geometric mean inequality the right-hand side of the last expression is at least

$$\left(\prod_{\substack{I\subset\{1,\dots,n\}\\|I|=k}}\left(\prod_{i\in I}\alpha_i\right)\left(\prod_{i\notin I}\beta_i\right)\right)^{\binom{n}{k}^{-1}}$$

$$=\left(\left(\prod_{i=1}^n\alpha_i\right)^{\binom{n-1}{k-1}}\cdot\left(\prod_{i=1}^n\beta_i\right)^{\binom{n-1}{k}}\right)^{\binom{n}{k}^{-1}}$$

$$=\left(\prod_{i=1}^n\alpha_i\right)^{k/n}\cdot\left(\prod_{i=1}^n\beta_i\right)^{(n-k)/n}=(\det A_1)^{k/n}(\det A_2)^{(n-k/n}.$$

Remark 3. By the above method one can also prove the Brunn–Minkowski inequality for two compact convex sets K_1, K_2. Namely,

$$|K_1+K_2|^{1/n}\geqslant|K_1|^{1/n}+|K_2|^{1/n}.$$

Moreover, if one of the sets, say, K_1 has nonempty interior, the equality is achieved if and only if K_2 is homothetic to K_1 (or is a point).

One should apply the following inequality for two positive definite matrices A_1, A_2

$$(\det(A_1+A_2))^{1/n}\geqslant(\det A_1)^{1/n}+(\det A_2)^{1/n}.$$

As in Remark 2 it is sufficient to prove the last inequality only for diagonal matrices. Then it takes the form

$$\prod_{r=1}^n(\alpha_r+\beta_r)^{1/n}\geqslant\left(\prod_{r=1}^n\alpha_r\right)^{1/n}+\left(\prod_{r=1}^n\beta_r\right)^{1/n},$$

which is known as Minkowski's inequality. The equality happens for $\alpha_r,\beta_r>0$ iff the vectors $(\alpha_1,\dots,\alpha_n)$ and $(\beta_1,\dots,\beta_n)$ are proportional. This implies also the equality cases in the Brunn–Minkowski inequality. Indeed, if $|K_2|=0$, then it is easy to see that K_2 must be a point. If K_2 is non-degenerate, then the above argument works.

Appendix

In this appendix we present the proof of the regularity theorem (Theorem 1.3) of the Brenier map due to L. Caffarelli [C4], which was communicated to us by him.

We will need several lemmas.

LEMMA A.1. *Let $\nabla\varphi$ be the Brenier map as above. Then for every x_0 $\nabla\varphi(x_0)\cap\overline{\Gamma}\neq\emptyset$.*

Proof. It is easy to see that for any convex function U defined in an open convex domain and for every x, $\nabla U(x)$ is a convex compact set. Moreover, ∇U is locally bounded, and for every point x_0 and for every neighborhood V_1 of $\nabla U(x_0)$ there exists a neighborhood V_2 of x_0 such that for all $x \in V_2$ $\nabla U(x) \subset V_1$. Hence, if $\nabla\varphi(x_0) \cap \overline{\Gamma} = \emptyset$, there exist a neighborhood V_1 of $\nabla\varphi(x_0)$ and a neighborhood V_2 of x_0 such that

$$V_1 \bigcap \overline{\Gamma} = \emptyset \quad \text{and} \quad \nabla\varphi(x) \subset V_1 \quad \text{for every } \; x \in V_2.$$

So

$$0 < \int_{V_2} f(x)\,\mathrm{d}x \leqslant \int_{(\nabla\varphi)^{-1}(V_1)} f(x)\,\mathrm{d}x = \int_{V_1} g(y)\,\mathrm{d}y = 0.$$

We get a contradiction. □

LEMMA A.2. *Let μ and ν satisfy the conditions of Theorem* 1.3. *Then* $\det D^2\varphi$ *is locally bounded away from zero and infinity in the Alexandrov sense, i.e. for every ball B_R of radius R there exist $\Lambda(R), \lambda(R) > 0$ such that for every compact subset $K \subset B_R$*

$$\lambda(R)m(K) \leqslant m(\nabla\varphi(K)) \leqslant \Lambda(R)m(K),$$

where m is the Lebesque measure.

Proof. Because of the invertibility of $\nabla\varphi$ (Proposition 1.2) $\mu(K) = \nu(\nabla\varphi(K))$, thus

$$\int_K f(x)\,\mathrm{d}x = \int_{\nabla\varphi(K)} g(y)\,\mathrm{d}y.$$

The lemma follows from conditions (i),(ii) immediately. □

The following result was proved in [C1], Theorem 1 (see also [C2], where the definitions are given).

PROPOSITION A.3. *Assume that a convex function U in an open region, satisfies $0 < \lambda_1 \leqslant \det D^2U \leqslant \lambda_2$ in the Alexandrov sense as above* (*or more generally, U is a viscosity solution of these inequalities*). *Then for every point x_0 and every supporting hyperplane $\ell_{x_0}(x)$ to the graph of U at x_0 either the set $\{x|U(x) = \ell_{x_0}(x)\}$ consists of x_0 only or has no extremal points in the interior of the domain of definition.*

DEFINITION A.4. A convex function U is called *strictly convex* at a point x_0 with a strict 'modulus of convexity' $\sigma(\rho)$, if for every $\rho > 0$

$$\operatorname{diam}\{U \leqslant \ell_{x_0} + \rho\} \leqslant \sigma(\rho),$$

and $\lim_{\rho\to 0}\sigma(\rho)=0$; ℓ_{x_0} denotes any supporting hyperplane at x_0.

It is easy to see that if U is such a convex function that for every x_0 from the domain of definition and for every supporting plane ℓ_{x_0} at x_0

$$\{x|\ell_{x_0}(x)=U(x)\}=\{x_0\},$$

then on every compact subset ν of the domain of definition there exists a strict modulus of convexity (defined uniformly for all points from this subset).

PROPOSITION A.5 ([C3]). *Let φ be a strictly convex solution of* $0<\lambda_1\leqslant \det\nabla\varphi\leqslant\lambda_2$ *in the Alexandrov sense. Then $\nabla\varphi$ belongs locally to the Hölder class C^α for some* $\alpha>0$ (*i.e. α can be chosen the same for any compact subset of the domain of definition of* φ).

Thus, Proposition A.5 would imply Theorem 1.3 if we show that, in our situation, every supporting plane to the graph of the potential φ of the Brenier map has a single touching point. To check this, let us observe that if this condition is not satisfied, then for some x_0 a convex set $\{x|\ell_{x_0}(x)=\varphi(x)\}$ has no extremal points, hence it contains a line N (since φ is defined on all $\mathbb{R}^n$).

Hence, any other supporting hyperplane to the graph of φ has to be parallel to N. Hence, for every $x\in\mathbb{R}^n$ we get $\nabla\varphi(x)\subset N^\perp$ – a contradiction (since Γ is nondegenerate). This proves Theorem 1.3(a).

Let us now prove (b) of Theorem 1.3.

Assume f, g belong locally to the Hölder class. By Theorem 1.3, $\nabla\varphi\in C^\alpha$ for some $\alpha>0$ (locally). Note that $\det D^2\varphi$ satisfies (is the Alexandrov sense) an equation

$$g(\nabla\varphi(x))(\det D^2\varphi)(x)=f(x).$$

But since $g(\nabla\varphi(x))$ is a usual function, we have $\det D^2\varphi=h$, where $h=f(x)/g(\nabla\varphi(x))$.

Since φ is a solution of the equation $(***)$ and h is locally Hölder, then the result follows from [C2, Theorem 2, p. 40], which states that if in equation $(***)$ $h\in C^\alpha$ then $\varphi\in C^{2,\alpha}$. □

References

[A] Alexandrov, A. D.: Zur Theorie gemischter Volumina von konvexen Körpern IV, Die gemischten Diskriminanten und die gemischten Volumina, *Mat. Sb. SSSR* 3 (1938), 227–251.

[AK] Anderson, R. D. and Klee, V. L.: Convex functions and upper semicontinuous collections. *Duke Math. J.* **19** (1952), 349–357.

[Bo] Bourgain, J.: On the distribution of polynomials on high dimensional convex sets, in: *GAFA Seminar Notes '89–'90*, Lecture Notes in Math. 1469, Springer-Verlag, New York, 1991, pp. 127–137.

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

[BZ] Burago, Yu. D. and Zalgaller, V. A.: *Geometric Inequalities*, Springer Ser. Soviet Math., Springer-Verlag, Berlin, 1988.

[Bu] Busemann, H.: *Convex Surfaces*, Interscience, New York, 1958.

[C1] Caffarelli, L. A.: A localization property of viscosity solutions to the Monge–Ampère equation and their strict convexity, *Ann. Math.* **131** (1990), 129–134.

[C2] Caffarelli, L. A.: *A-priori Estimates and the Geometry of the Monge-Ampère Equation*, Park City/IAS Math. Ser. II, Amer. Math. Soc., Providence, 1992.

[C3] Caffarelli, L. A.: Some regularity properties of solutions of Monge–Ampère equation, *Comm. Pure Appl. Math.* **44** (1991), 965–969.

[C4] Caffarelli, L. A.: Private communication.

[Gr] Gromov, M.: Convex sets and Kähler manifolds, in *Advances in Differential Geometry and Topology*, World Scientific, Teaneck, NJ, 1990, pp. 1–38.

[H] Hörmander, L.: *Notions of Convexity*, Progress in Math. 127, Birkhäuser, Boston, 1994.

[McC1] McCann, R. J.: Existence and uniqueness of monotone measure-preserving maps, *Duke Math. J.* **80**(2) (1995), 309–323.

[McC2] McCann, R. J.: A convexity principle for interacting gases, *Adv. in Math.* **128**(1) (1997), 153–179.

[MSch] Milman, V. D. and Schechtman, G.: Asymptotic theory of finite dimensional normed spaces (with appendix by M. Gromov), Lecture Notes in Math. 1200, Springer-Verlag, New York, 1986.

[R] Rockafellar, R. T.: *Convex Analysis*, Princeton University Press, Princeton, 1972.

[Sch] Schneider, R.: Convex bodies: the Brunn–Minkowski theory, *Encyclopedia Math. Appl.* 44, Cambridge University Press, Cambridge, 1993.