Polar Factorization and Monotone Rearrangement of Vector-Valued Functions

YANN BRENIER

Université de Paris VI

Abstract

Given a probability space (X, μ) and a bounded domain Ω in $\mathbb{R}^d$ equipped with the Lebesgue measure $|\cdot|$ (normalized so that $|\Omega| = 1$), it is shown (under additional technical assumptions on X and Ω) that for every vector-valued function $u \in L^p(X, \mu; \mathbb{R}^d)$ there is a unique "polar factorization" $u = \nabla\psi \circ s$, where ψ is a convex function defined on Ω and s is a measure-preserving mapping from (X, μ) into $(\Omega, |\cdot|)$, provided that u is nondegenerate, in the sense that $\mu(u^{-1}(E)) = 0$ for each Lebesgue negligible subset E of $\mathbb{R}^d$.

Through this result, the concepts of polar factorization of real matrices, Helmholtz decomposition of vector fields, and nondecreasing rearrangements of real-valued functions are unified.

The Monge-Ampère equation is involved in the polar factorization and the proof relies on the study of an appropriate "Monge-Kantorovich" problem.

1. Introduction

1.1. Review of Some Well-Known Results

In this paper, several apparently unrelated classical results are unified through the concept of polar factorization of vector-valued functions. Let us review them.

Polar Coordinates in the Complex Plane

Any complex number can be written as $z = re^{i\theta}$, $r \geqq 0$, $\theta \in \mathbb{R}/2\pi\mathbb{Z}$.

Polar Factorization of Real Matrices

Any real matrix A can be written as the product RU of a symmetric non-negative matrix R by a real unitary matrix U. If A is regular ($\det A \neq 0$), then the factorization is unique.

Helmholtz Decomposition of Vector Fields

Let Ω be a smooth bounded connected open set in $\mathbb{R}^d$. Then, any smooth vector field z on Ω can be written, in a unique way, as $z = w + \nabla p$, where p is a smooth real function, defined on Ω up to an additive constant, and w is a smooth divergence free vector field, parallel to the boundary of Ω. This decomposition theorem is the simplest application of the Hodge theorem on differential forms; see [1].

Communications on Pure and Applied Mathematics, Vol. XLIV, 375–417 (1991)

CCC 0010-3640/91/040375-43$04.00

Nondecreasing Rearrangements of Real Functions

Before stating the last classical result considered in this introduction, let us first review some well-known definitions on probability spaces. A "measure-preserving mapping" from a probability space (X, μ) into another probability space (Y, ν) is a mapping $s : X \to Y$ such that:

$$\text{(1)} \qquad \begin{gathered} \text{for every } \nu\text{-measurable subset } A \text{ of } Y, \\ s^{-1}(A) \text{ is } \mu\text{-measurable and } \mu(s^{-1}(A)) = \nu(A). \end{gathered}$$

Or, equivalently,

$$\text{(2)} \qquad \begin{gathered} \text{for every } \nu\text{-integrable function } f, \\ f\circ s \text{ is } \mu\text{-integrable and } \int_X f\circ s \, d\mu = \int_Y f \, d\nu. \end{gathered}$$

Such a mapping is not necessarily one-to-one. Consider, for instance, the case: $(X, \mu) = (Y, \nu) = ([0, 1], |\cdot|)$, where $|\cdot|$ denotes the Lebesgue measure, and $s(x) = \min(2x, 2 - 2x)$. If there is a one-to-one measure-preserving mapping from (X, μ) into (Y, ν), then (X, μ) and (Y, ν) are said to be isomorphic. In this paper, only probability spaces that are isomorphic to $([0, 1], |\cdot|)$ will be considered. This restriction is not severe, since, for instance, any separable complete metric space X, equipped with a Borel measure μ such that $\mu(X) = 1$ and $\mu(\{x\}) = 0$, for all $x \in X$, is isomorphic to $([0, 1], |\cdot|)$; see [21].

If $p \in [1, +\infty[$ and $d = 1, 2, 3 \cdots$, we call rearrangement of $u \in L^p(X, \mu; \mathbb{R}^d)$ on (Y, ν), any $v \in L^p(Y, \nu; \mathbb{R}^d)$ such that:

$$\text{(3)} \qquad \begin{gathered} \int_X f(u(x)) \, d\mu(x) = \int_Y f(v(y)) \, d\nu(y). \\ \text{for each } f \in C(\mathbb{R}^d) \text{ such that } |f(\zeta)| \leqq \text{cst}(1 + \|\zeta\|^p). \end{gathered}$$

The last classical result deals with nondecreasing rearrangements, on the unit interval, of real-valued functions $u \in L^p(X, \mu; \mathbb{R})$ (which means $d = 1$ and $(Y, \nu) = ([0, 1], |\cdot|)$).

Let (X, μ) be a probability space isomorphic to $([0, 1], |\cdot|)$, and $p \in [1, +\infty[$. Then, for each $u \in L^p(X, \mu)$, there is a unique nondecreasing rearrangement $u^\# \in L^p(0, 1)$. Moreover, the mapping $u \to u^\#$ is continuous from $L^p(X, \mu)$ into $L^p(0, 1)$.

This result is well known; see [14], [5], and [18]. A recent, detailed review of these topics can be found in [2]. Let us briefly indicate some complementary properties:

$$\text{(4)} \qquad \|u_1^\# - u_2^\#\|_{L^p(0,1)} \leqq \|u_1 - u_2\|_{L^p(X,\mu)}.$$

$$(5) \qquad \int_0^1 u_1^\# u_2^\# \, dx \geqq \int_X u_1 u_2 \, d\mu, \qquad \forall u_1, u_2 \in L^2(X, \mu).$$

For each $u \in L^p(X, \mu)$,

(6) there is a measure-preserving mapping $s : (X, \mu) \to ([0, 1], |\cdot|)$,

such that $u = u^\# \circ s$.

This last statement, proved by Ryff in [22], is not very well known.

1.2. The Main Results

Our results can be essentially seen as the extension of the classical result on non-decreasing rearrangements of real functions to the case of vector-valued functions $u \in L^p(X, \mu; \mathbb{R}^d)$, for $d > 1$.

The unit interval $]0, 1[$ is now replaced by a bounded connected open set Ω in $\mathbb{R}^d$. The Lebesgue measure is replaced by a probability measure β on $\bar{\Omega}$ such that $\beta(\partial\Omega) = 0$. It is assumed that $d\beta(z) = \beta(z)\, dz$, for some Lebesgue integrable non-negative function β, bounded away from 0 on any compact subset of Ω. It follows that β has the same negligible sets as the Lebesgue measure, which means that β is absolutely continuous with respect to the Lebesgue measure and conversely. Notice that $(\bar{\Omega}, \beta)$ is isomorphic to $([0, 1], |\cdot|)$; see [21].

The weighted Sobolev space

$$(7) \qquad W^{1,p}(\Omega, \beta) = \{f \in L^p(\Omega, \beta), \nabla f \in L^p(\Omega, \beta; \mathbb{R}^d)\}$$

will play an important role and is assumed to be compactly embedded in $L^p(\Omega, \beta)$ (which is automatically enforced when Ω is smooth and β is bounded away from 0 and $+\infty$ on $\bar{\Omega}$). Under these assumptions, our main results are:

Rearrangements of Vector-Valued Functions

THEOREM 1.1. *For each $u \in L^p(X, \mu; \mathbb{R}^d)$, there is a unique rearrangement $u^\#$ in the class*

$$(8) \qquad K = \{\nabla\psi; \psi \in W^{1,p}(\Omega, \beta); \psi \text{ convex}\} \subset L^p(\Omega, \beta; \mathbb{R}^d)$$

and the mapping $u \to u^\#$ is continuous.

Polar Factorization of Vector-Valued Functions

THEOREM 1.2. *Let N be the set of all $u \in L^p(X, \mu; \mathbb{R}^d)$ for which the following "nondegeneracy" condition fails:*

(9) $\mu(u^{-1}(E)) = 0$ *for each Lebesgue negligible subset* E *of* $\mathbb{R}^d$.

Then, for each $u \in L^p(X, \mu; \mathbb{R}^d)\backslash N$, there is a unique pair $(u^\#, s)$, such that $u^\#$ belongs to K, s is a measure-preserving mapping from (X, μ) into $(\bar{\Omega}, \beta)$, and $u = u^\# \circ s$. Moreover,

(a) *$u^\#$ is the unique rearrangement of u in K (as in Theorem 1.1);*

(b) *s is the unique measure-preserving mapping that maximizes $\int_X u(x) \cdot s(x)\, d\mu(x)$;*

(c) *the mapping $u \to (u^\#, s)$ is continuous from $L^p(X, \mu; \mathbb{R}^d)\backslash N$ into $L^p(\Omega, \beta; \mathbb{R}^d) \times L^q(X, \mu; \mathbb{R}^d)$, for all $q \in [1, +\infty[$.*

Remark. We believe that the result is still true when Ω is unbounded, provided that $p > 1$ and $\int_\Omega \|z\|^q \beta(z)\, dz < +\infty$, where $1/q + 1/p = 1$.

Comments on the "Rearrangement Theorem"

Obviously, K is, with respect to $L^p(\Omega, \beta; \mathbb{R}^d)$, the generalization of the class of all nondecreasing functions in $L^p(0, 1)$. Thus Theorem 1.1 is a generalization of the classical result on nondecreasing rearrangements of real functions, which corresponds to the particular case $d = 1$, $(\bar{\Omega}, \beta) = ([0, 1], |\cdot|)$.

As in the case of real arrangements, the mapping $u \to u^\#$ is continuous. However, properties (4) and (5) are not true in general when $d > 1$. For instance, in the particular case $(X, \mu) = (\bar{\Omega}, \beta) = (\bar{\Omega}, |\cdot|)$, where $|\cdot|$ denotes the Lebesgue measure (normalized so that $|\Omega| = 1$), it will be shown in Section 2.4 that:

$$\int_\Omega u_1^\#(x) \cdot u_2^\#(x)\, dx \geqq \int_\Omega u_1(x) \cdot u_2(x)\, dx \tag{10}$$

is not true in general.

1.3. Recovery of the Classical Results

Our "rearrangement theorem" generalizes the classical results on nondecreasing rearrangements of real functions. Let us now check that the polar factorization of real matrices and the Helmholtz decomposition of vector fields are particular cases of the situation considered in our "factorization theorem".

Helmholtz Decomposition

The Helmholtz decomposition turns out to be the linearization of the polar factorization of vector-valued functions about the identity map, when $(X, \mu) = (\bar{\Omega}, \beta) = (\bar{\Omega}, |\cdot|)$.

Let us consider u as a smooth perturbation of the identity map:

$$u(x) = x + \epsilon z(x), \ \epsilon \ll 1, \qquad x \in \bar{\Omega}.$$

For ϵ small enough, u satisfies the nondegeneracy condition (9) (because the jacobian determinant is bounded away from 0) and u can be written: $u = \nabla\psi^{\#} \circ s$. It is natural to seek $\nabla\psi^{\#}$ and s as perturbations of the identity map:

$$\psi^{\#}(x) = \tfrac{1}{2}\|x\|^2 + \epsilon p(x) + O(\epsilon^2), \qquad s(x) = x + \epsilon w(x) + O(\epsilon^2).$$

Then, $z = \nabla p + w$ immediately follows. To keep $\psi^{\#}$ convex, p needs no special property (provided that p is smooth enough and ϵ is small enough). However, to keep s measure preserving, it is necessary to enforce

$$\int_\Omega f(s(x))\,dx = \int_\Omega f(x)\,dx$$

for each smooth function f defined on $\bar{\Omega}$. This leads to

$$\int_\Omega \nabla f(x)\cdot w(x)\,dx = 0,$$

which precisely means (in the weak sense) that w is divergence free on Ω and parallel to $\partial\Omega$.

So, the Helmholtz decomposition is the linearization of the polar factorization of vector-valued functions. At this point, a very intriguing question can be raised: Is it a general fact that a Hodge decomposition of differential forms (see [1]) can be seen as the linearization of some "generalized" polar factorization?

Polar Factorization of Matrices

Let us now show that the polar factorization of matrices is a particular case of the polar factorization of vector-valued functions. Here, $(X, \mu) = (\bar{\Omega}, \beta) = (\bar{\Omega}, |\cdot|)$ and Ω is a ball centered at the origin, u is a linear mapping: $u(x) = A\cdot x$, for some real $d \times d$ matrix A. The nondegeneracy condition (9) exactly means that A is nonsingular ($\det A \neq 0$). The polar factorization $A = RU$ corresponds to the polar factorization $u = \nabla\psi^{\#}\circ s$, where: $\psi^{\#}(x) = \frac{1}{2} Rx\cdot x$ is convex and $s(x) = U\cdot x$ is a measure-preserving mapping from $(\Omega, |\cdot|)$ into itself. Notice that, due to the special geometry of the ball, for a linear mapping $u(x) = A\cdot x$, each factor of the polar factorization $s(x) = U\cdot x$, $u^{\#}(x) = R\cdot x$ is a linear mapping. This cannot be true for a general geometry.

The case of complex numbers $z = re^{i\theta}$ can be treated in the same way. So our polar factorization theorem unifies all these classical results. Let us now show how our rearrangement theorem is strongly linked to the Monge-Ampère equation; see [9].

1.4. The Rearrangement Theorem and the Monge-Ampère Equation

Theorem 1.1 can be seen as an existence and uniqueness theorem of a "generalized" solution for the following Monge-Ampère problem.

Monge-Ampère Problem

Given $\alpha \in L^1(\mathbb{R}^d)$, $\alpha \geqq 0$, such that $\int (1 + \|y\|^p)\alpha(y)\, dy < +\infty$, find a (Lipschitz continuous) convex function $\phi : \mathbb{R}^d \to \mathbb{R}$ that satisfies (in a generalized sense to be precised):

$$\beta(\nabla\phi(y))\det D^2\phi(y) = \alpha(y), \quad \forall y \in \text{ support } (\alpha), \tag{11}$$

$$\nabla\phi \quad \text{maps the support of} \quad \alpha \quad \text{into} \quad \bar{\Omega}. \tag{12}$$

Usually, the Monge-Ampère equation is set on a bounded (often convex) domain with Dirichlet or Neumann boundary conditions; see [9]. Here, the "range condition" (12) replaces the usual boundary conditions.

To see the link with this Monge-Ampère problem, we need a different version of our rearrangement theorem that relies on the following observation: What really matters in Theorem 1.1 is the probability measure α associated with $u \in L^p(X, \mu; \mathbb{R}^d)$ and defined by

$$\int_{\mathbb{R}^d} f(y)\, d\alpha(y) = \int_X f(u(x))\, d\mu(x) \tag{13}$$

for each compactly supported $f \in C(\mathbb{R}^d)$. Notice that $(1 + \|y\|^p)\, d\alpha(y)$ is a tight positive measure, which means

$$\lim_{r\to\infty} \int_{\|y\| \geqq r} (1 + \|y\|^p)\, d\alpha(y) = 0,$$

and (13) still holds for each $f \in C(\mathbb{R}^d)$ such that $|f(y)| \leqq \text{cst}(1 + \|y\|^p)$. In addition, if a sequence (u_n) converges to u in $L^p(X, \mu; \mathbb{R}^d)$, then the corresponding probability measures α_n satisfy

$$\lim_{r\to\infty} \sup_n \int_{\|y\| \geqq r} (1 + \|y\|^p)\, d\alpha_n(y) = 0,$$

and, therefore, converge to α in the following sense: $\int f\, d\alpha_n \to \int f\, d\alpha$ for every $f \in C(\mathbb{R}^d)$ such that $|f(y)| \leqq \text{cst}(1 + \|y\|^p)$.

So, it is not hard to see that Theorem 1.1 is a corollary of:

THEOREM 1.3. *For each probability measure α on $\mathbb{R}^d$ satisfying $\int (1 + \|y\|^2)\, d\alpha(y) < +\infty$, there is a unique $u^\# = \nabla\psi^\#$ in K such that*

$$\int f(y)\, d\alpha(y) = \int_\Omega f(\nabla\psi^\#(z))\, d\beta(z), \tag{14}$$

for each $f \in C(\mathbb{R}^d)$ such that $|f(y)| \leqq \text{cst}(1 + \|y\|^p)$. Moreover, if $\int f\, d\alpha_n \to \int$

$f\,d\alpha$, for a sequence of such probability measures α_n and any $f \in C(\mathbb{R}^d)$ such that $|f(y)| \leqq \mathrm{cst}(1 + \|y\|^p)$, then

$$\psi_n^\# \to \psi^\# \text{ in } W^{1,p}(\Omega, \beta)/\mathbb{R}.$$

A link can be established with the Monge-Ampère problem in the case when α is absolutely continuous with respect to the Lebesgue measure

$$d\alpha(y) = \alpha(y)\,dy, \quad \alpha \in L^1(\mathbb{R}^d), \quad \alpha \geqq 0,$$

which exactly means that u is nondegenerate in the sense of (9) (by Lebesgue-Nikodym theorem). In Theorem 1.3, $\psi^\#$ is convex and $\nabla\psi^\#$ maps $\bar{\Omega}$ into the support of α, as follows from (14). Let us consider the Legendre-Fenchel transform $\phi^\#$ of $\psi^\#$ (cf. [13]), defined by

$$\phi^\#(y) = \sup_{z \in \Omega} \{y \cdot z - \psi^\#(z)\}, \qquad y \in \mathbb{R}^d. \tag{15}$$

It will be shown (see Section 3) that $\phi^\#$ is Lipschitz continuous on $\mathbb{R}^d$, $\nabla\phi^\#$ is well defined almost everywhere on $\mathbb{R}^d$, and

$$\nabla\phi^\#(\nabla\psi^\#(z)) = z, \quad \text{for a.e.} \quad z \in \Omega,$$

$$\nabla\psi^\#(\nabla\phi^\#(y)) = y, \quad \text{for} \quad \alpha - \text{a.e. } y \in \mathbb{R}^d.$$

Thus, $\nabla\psi^\#$ and $\nabla\phi^\#$ are reciprocal and, formally, one gets for every compactly supported $f \in C(\mathbb{R}^d)$:

$$\int_\Omega f(\nabla\psi^\#(z))\beta(z)\,dz = \int_{\mathrm{supp}\alpha} f(y)\beta(\nabla\phi^\#(y))\det D^2\phi^\#(y)\,dy, \tag{16}$$

by using the change of variable $y = \nabla\psi^\#(z)$, $z = \nabla\phi^\#(y)$. By definition of $\psi^\#$, one has

$$\int_\Omega f(\nabla\psi^\#(z))\beta(z)\,dz = \int f(y)\alpha(y)\,dy,$$

and, thus, $\phi^\#$ satisfies (in a generalized sense) the Monge-Ampère equation (11) together with the range condition defined by (12).

1.5. Origin of the Results

The polar factorization theorem was motivated by the study of the following "projection problem" (introduced in [6]). Here $(X, \mu) = (\bar{\Omega}, \beta) = (\bar{\Omega}, |\cdot|)$ and $p = 2$.

Let us denote by H the Hilbert space $L^2(\Omega; \mathbb{R}^d)$ and by S the set of all measure-preserving mapping from $(\Omega, |\cdot|)$ into itself, which means

$$\text{for every} \quad f \in L^1(\Omega),$$

$$f \circ s \in L^1(\Omega) \quad \text{and} \quad \int_\Omega f \circ s(x)\, dx = \int_\Omega f(x)\, dx. \tag{17}$$

This set S is a closed (nonconvex, noncompact) bounded subset of H (cf. Section 2). Notice that S is contained in a sphere centered at the origin. Indeed, from (17), one gets: $\|s\|^2 = \int_\Omega \|s(x)\|^2\, dx = \int_\Omega \|x\|^2\, dx = \text{cst}$. Let us now consider:

The Projection Problem

Find $s \in S$ that minimizes $\|u - s\|^2 = \int_\Omega \|u(x) - s(x)\|^2\, dx$, or, equivalently, that maximizes $((u, s)) = \int_\Omega u(x) \cdot s(x)\, dx$.

As mentioned in Theorem 1.2, when the nondegeneracy condition (9) is satisfied, the factor s in the polar factorization of $u = u^{\#} \circ s$, for $u \in L^2(\Omega; \mathbb{R}^d)$, is exactly the unique maximizer in S of $\int_\Omega u(x) \cdot s(x)\, dx$ and, therefore, is the unique Hilbert projection of u on S. The projection problem is a key to understanding the concept of polar factorization. For instance, the set K, here defined by

$$K = \{\nabla\psi;\ \psi \in W^{1,2}(\Omega);\ \psi \text{ convex}\} \subset L^2(\Omega; \mathbb{R}^d),$$

is closely linked to the "projection problem". Indeed, it can be shown (cf. Section 2) that K is exactly the set of all $u \in H$ for which the identity map e is a Hilbert projection of u on S:

$$K = \{u \in H;\ \|u - e\| \leqq \|u - s\|,\ \forall s \in S\},$$

or, equivalently,

$$K = \{u \in H;\ ((u, e - s)) \geqq 0,\ \forall s \in S\}.$$

Notice that, in terms of convex analysis, K is the "polar cone" of S (or, equivalently, of its convex hull). This "geometrical" description of the polar factorization can be made even more precise, by noticing that S is a semigroup with respect to the composition rule

$$s_1 \circ s_2 \in S, \quad \text{whenever} \quad s_1, s_2 \in S,$$

and the identity map e is the neutral element. If S were a group (which is definitely not true since, for instance, in the case $(\bar{\Omega}, |\cdot|) = ([0, 1], |\cdot|)$, $s(x) = \min(2x,$

$2 - 2x$) is not invertible in S), then the "polar factorization" theorem would follow from the study of the "projection problem".

Indeed, if $u \in H$ has a Hilbert projection s on S, where s is invertible in S, then u admits a polar factorization $u = k \circ s$, where k is defined by $k = u \circ s^{-1}$ and clearly belongs to K , since, for any $\sigma \in S$, because of property (17), one has

$$\|k - \sigma\| = \|u - \sigma \circ s\| \geqq \|u - s\| = \|k - e\|.$$

There is a general answer for the projection problem, given by Edelstein's theorem (see [4]), which asserts that "almost every" $u \in H$, in the sense of Baire, has a unique projection $s = \pi(u)$ on S. This is due to the fact that S is a closed, bounded subset of a Hilbert space. Unfortunately, in our case, since S is not a group, there is no reason for s to be in general invertible and, therefore, there is no direct way to recover the factor k of the polar decomposition of u. We have not been able to overcome this difficulty to obtain our main results directly and a more involved proof has been used, through the study of the "Monge-Kantorovich" problem (cf. Section 1.6). Note that it is not a good idea to substitute for S the group G of all invertible measure-preserving mappings in S, since G is not closed in H, which makes impossible the use of Edelstein's theorem. Even if this "geometrical" approach does not yield a proof of our main results, it seems to us that the discussion of the polar factorization in terms of group and Hilbert projection is of interest (especially in view of possible generalizations). This is why the second section of this paper is mainly devoted to this approach, in a rather abstract framework: we are given a Hilbert space H, a closed bounded subset S, and a composition rule $*$ on $S \times H$ such that $(S, *)$ is a group and $\|s*u\| = \|u\|$, for all $u \in H$, $s \in S$. Then it is proved, under additional assumptions, that for almost every $u \in H$ there is a unique polar factorization $u = s*k$, $s \in S$, $k \in K$. Moreover, it is shown that $u \to s$ is the gradient of the Lipschitz continuous convex function $J(u) = \sup_{s \in S}((u, s))$ (which is the Legendre-Fenchel transform of the indicator function of S).

Section 2 can be read independently of the remainder of the paper and can be ignored by those who are interested only in the proofs of our main results.

1.6. A Proof Using the Monge-Kantorovich Problem

Our proof is based on the observation that the "projection problem" is a variant of the "mass transference problem" introduced by Monge, in the eighteenth century (see [17], cf. [3]), and generalized by Kantorovich in the 1940s; see [15]. In modern terms, the Monge problem (MP) and the Monge-Kantorovich problem (MKP) can be described as follows.

MP

Given two compact metric probability spaces (X, μ), (Y, ν) and a continuous "cost" function $c : X \times Y \to \mathbb{R}_+$, find a one-to-one measure-preserving mapping from (X, μ) into (Y, ν) that minimizes the "transportation cost" $\int_X c(x, s(x))\, d\mu(x)$.

MKP

Given two compact metric probability spaces (X, μ), (Y, ν) and a continuous "cost" function $c : X \times Y \to \mathbb{R}_+$, find a probability measure γ on $X \times Y$ with marginals μ and ν, which means, for each $f \in C(X)$ and each $g \in C(Y)$:

$$\int_X f(x)\, d\mu(x) = \int_{X \times Y} f(x)\, d\gamma(x, y),$$

$$\int_Y g(y)\, d\nu(y) = \int_{X \times Y} g(y)\, d\gamma(x, y),$$

that minimizes the "generalized transport cost" $\int c(x, y)\, d\gamma(x, y)$.

The MKP is a "relaxed" version of the MP, in the sense that any admissible solution s to the MP yields an admissible solution γ to the MKP, defined by

$$d\gamma(x, y) = \delta(y - s(x))\, d\mu(x).$$

The MKP has many applications and a very complete review can be found in a recent paper by Rachev; see [19]. The MKP is an infinite dimensional linear program and a key point of the analysis developed by Kantorovich is the study of the dual linear program:

$$I = \inf\left\{\int f\, d\mu + \int g\, d\nu,\ f \in C(X),\ g \in C(Y),\right.$$

$$\left.\text{such that} \quad f(x) + g(y) \leqq c(x, y)\right\}.$$

It can be proved (see [19]) that the MKP has a solution γ that satisfies the "strong duality" relation $I = \int c(x, y)\, d\gamma(x, y)$. A more refined result, due to Sudakov (see [19]), asserts that under more specific conditions (that are satisfied when the cost function is a power of some distance function), the solution γ is actually of the form $d\gamma(x, y) = \delta(y - s(x))\, d\mu(x)$, where s is a one-to-one measure-preserving mapping, which means that the Monge problem has a solution.

Our projection problem corresponds to the case when $(X, \mu) = (Y, \nu) = (\bar{\Omega}, |\cdot|)$ and $c(x, y) = \|u(x) - y\|^2$. We look for a measure-preserving mapping, not necessarily one-to-one, s that minimizes the cost $\int_\Omega c(x, s(x))\, dx$. In the corresponding MKP problem, we look for a "doubly stochastic" probability measure on $\bar{\Omega} \times \bar{\Omega}$, which means that the marginals of γ are $|\cdot|$, that minimizes the cost $I = \int \|u(x) - y\|^2\, d\gamma(x, y)$. Notice that the cost function is not necessarily continuous, since u is given in $L^2(\Omega; \mathbb{R}^d)$. However, the MKP still makes sense because

c is automatically integrable for any doubly stochastic probability measure (this can be easily seen).

In the earlier version of this paper (see [6]), we proved, with the help of the strong duality relation and some well-known properties of convex functions, that, when u is nondegenerate in the sense of (9), this MKP has a unique solution of the form

$$d\gamma(x, y) = \delta(y - s(x))\, dx$$

for some measure-preserving mapping s, that turns out to be the unique solution of our projection problem. Moreover, the solution of the dual program is linked to a pair of convex conjugate functions $(\psi^{\#}, \phi^{\#})$ and $\nabla\psi^{\#}$ is precisely the rearrangement of u in K.

In this paper, a slightly different and more efficient approach is used, by considering a different MKP that generalizes the projection problem in a different way. We look for a probability measure p on $\bar{\Omega} \times \bar{\Omega}$, with the following marginals:

$$\int f(u(x))\, dx = \int f(y)\, dp(x, y),$$

$$\int f(x)\, dx = \int f(x)\, dp(x, y),$$

for each $f \in C(\bar{\Omega})$, that minimizes $I = \int \|x - y\|^2\, dp(x, y)$. Clearly, any admissible solution s of the projection problem yields an admissible solution p of this new MKP, defined by

$$\int f(x, y)\, dp(x, y) = \int f(s(x), u(x))\, dx$$

for each $f \in C(\bar{\Omega} \times \bar{\Omega})$. The main advantage of the new MKP is that the cost function $c(x, y) = \|x - y\|^2$ is now simpler and smoother.

1.7. References to Related Works

Our polar factorization theorem was introduced, in a weaker form (for instance, u was supposed to be bounded and Riemann integrable), in [6] and the proof (using the Monge-Kantorovich problem (MKP)) was sketched. Our motivation was the numerical study of the motion of perfect incompressible fluids. In [7], a "Lagrangian" scheme was introduced, where the "projection problem" plays an essential role. Our proof relies on the study of the MKP and Rachev's paper, [19], was very useful. We would like to thank G. Strang for communicating it to us.

Since our first paper, we have heard (thanks to P.-L. Lions, J. Mossino, J. Norbury, and S. Rachev) of two papers that anticipated our "rearrangement theorem". Both are due to British scientists, one in statistics and the other in atmospheric sciences. In the second one, by Cullen and Purser (see [10], followed by [11]), a numerical scheme is designed, using a quasi-geostrophic model, for weather forecasting. The solution is described by a vector field that evolves in time. At each time step, this field is rearranged as the gradient of some convex potential, for some stability reasons linked to the physical model. A numerical procedure to compute this rearrangement is also described. This procedure is reminiscent of the method that Pogorelov introduced to prove the existence of generalized solutions for the Monge-Ampère equation; see [9].

The first paper, recently pointed out to us by S. Rachev, is due to Knott and Smith (see [16]) and deals with the following problem in statistics.

Given two random variables X and Y, with prescribed laws, find a correspondence $Y = Z(X)$ that maximizes the expectation of $X \cdot Y$. It is shown that the optimality condition for Z is to be the subdifferential of some convex function. Behind this problem, it is not hard to recognize the situation considered in our rearrangement theorem. The solution of this problem, in the one-dimensional case ($d = 1$), goes back to Fréchet and Hoeffding (see [20]).

So, in these earlier works, the concept of rearrangement for vector-valued functions is more or less explicitly present, although the notion of polar factorization is missing.

1.8. Organization of the Paper

Section 2 is devoted to the abstract concept of "polar factorization". A definition and some examples are considered in Section 2.1. In particular, the polar factorization of vector-valued functions (in the case $p = 2$, $(X, \mu) = (\bar{\Omega}, |\cdot|)$) is discussed. In Section 2.2, an abstract polar factorization theorem is proved and relies on a group property that is not satisfied in the case of vector-valued functions. Sections 2.3 and 2.4 give complementary results in the case of vector-valued functions. Section 2 is completely independent and can be ignored by readers who are only interested in the proof of our main results. These proofs are given in Section 3 and essentially rely on the study of the Monge-Kantorovich problem (MKP). For simplicity the proof is given in the case $p = 1$, which is less restrictive (indeed (X, μ) is a probability space and, therefore, all spaces $L^p(X, \mu; \mathbb{R}^d)$ are contained in $L^1(X, \mu; \mathbb{R}^d)$). In Section 3.1, the MKP is presented and two essential *a priori* results are stated (Propositions 3.1 and 3.2). Their proofs are given in Section 3.2. Then, our rearrangement and polar factorization theorems are deduced in Section 3.3. Finally, the existence proof for the MKP is given in Section 3.4.

2. The Abstract Concept of Polar Factorization

2.1. Definition and Examples

In this section, the notion of "polar factorization" is defined in a general framework that involves:

(a) a real Hilbert space H, with Hilbert product $((\cdot, \cdot))$ and norm $\|\cdot\|$;

(b) a closed bounded subset S of H;

(c) a composition rule $* : S \times H \to H$ such that $(S, *)$ is a semigroup, with neutral element e, which means:

$$s_1 * s_2 \in S \quad \text{whenever} \quad s_1, s_2 \in S, \; s*e = e*s = s, \quad \text{for all} \quad s \in S.$$

It is assumed that

$$\begin{aligned} &s*(\alpha u + \beta v) = \alpha s*u + \beta s*v, \quad \forall s \in S, \quad \forall u, v \in H, \quad \forall \alpha, \beta \in \mathbb{R} \\ (18)\qquad &\forall u \in H, s \in S \to s*u \quad \text{is continuous and} \quad : \|s*u\| = \|u\|. \end{aligned}$$

Then, the "polar cone"

$$(19) \qquad K = \{u \in H; ((u, e - s)) \geqq 0, \forall s \in S\}$$

is introduced and the following definition is stated.

Abstract Polar Factorization

It is said that there is a polar factorization of the Hilbert space H by the semigroup S and the polar cone K if there is a negligible subset N of H (in the sense of Baire: $H \backslash N$ contains a dense countable intersection of open sets), such that

(i) for every $u \in H \backslash N$, there is a unique factorization $u = s*k$, $(s, k) \in S \times K$,

(ii) $u \to (s, k)$ is continuous on $H \backslash N$.

Let us review the examples of polar factorization already considered in the introduction.

Example 1. Polar factorization of complex numbers. Here $H = \mathbb{C} = \mathbb{R}^2$ ($\mathbb{C}$ is considered as a real Hilbert space), $\|z\| = |z|$, $S = \{e^{i\theta}, \theta \in \mathbb{R}/2\pi\mathbb{Z}\}$, and the composition rule is the usual multiplication of complex numbers. Then, one gets $K = \mathbb{R}_+$, $N = \{0\}$.

Example 2. Polar factorization of real $d \times d$ matrices. Here

$$H = \{A = (a_{ij})_{i,j=1,d}\} = \mathbb{R}^{d \times d},$$

$$\|A\|^2 = \sum_{i,j=1}^{d} a_{ij}^2,$$

$$S = \{U \in H \quad \text{such that} \quad U^T U = U U^T = I\},$$

and the composition rule is the usual multiplication of real matrices. Then, the "exceptional" set N is the set of all degenerate matrices and K is the convex cone of all symmetric non-negative matrices.

Example 3. Polar factorization of vector-valued functions. Let us compare our factorization theorem, as stated in the introduction, and our abstract definition of polar factorization. For simplicity, we assume $(X, \mu) = (\bar{\Omega}, \beta) = (\bar{\Omega}, |\cdot|)$ and $p = 2$. Our Hilbert space H now is the space $L^2(\Omega; \mathbb{R}^d)$ and $\|\cdot\|$ is the L^2-norm. The set S is the set of all measure-preserving mappings s from $(\bar{\Omega}, |\cdot|)$ into itself, which means

$$(20) \qquad \forall f \in L^1(\Omega), \quad f\circ s \in L^1(\Omega), \quad \text{and} \quad \int_\Omega f\circ s = \int_\Omega f.$$

The composition law $*$ is the usual composition law for mappings $* = \circ$ and is well defined on $S \times H$. Indeed, since each $s \in S$ preserves the Lebesgue negligible sets and each $u \in H$ is a Borel mapping up to a possible change on a Lebesgue negligible set, $u \circ s$ is a well defined Borel mapping. It is a fairly elementary exercise to check that property (18) is satisfied (because of property (20)), $(S, \circ)$ is a semigroup, and S is a closed bounded subset of H. So, our "factorization theorem" 1.2 would exactly fit into the framework of the abstract definition, if one could prove that

(a) the set N of all $u \in L^2(\Omega; \mathbb{R}^d)$ for which the nondegeneracy condition (9) fails is negligible in the sense of Baire (i.e., contained in a countable union of closed subsets with empty interior);

(b) the set $\{\nabla\psi;\ \psi \in W^{1,2}(\Omega),\ \psi \text{ convex}\}$ is the polar cone defined by (19), namely:

$$(21) \qquad K = \{u \in L^2(\Omega; \mathbb{R}^d);\ \int_\Omega u(x)\cdot(x - s(x))\, dx \geqq 0, \quad \forall s \in S\}.$$

We did not succeed in proving the first statement (although it is easy to check that $H \backslash N$ is dense). However we were able to prove:

PROPOSITION 2.1 (Characterization of the cone K). *The cone K defined by* (21) *exactly is* $\{\nabla\psi;\ \psi \in W^{1,2}(\Omega),\ \psi \text{ convex}\}$.

The proof is at the end of Section 2.3. Let us now get back to the abstract framework of the polar factorization concept.

2.2. An Abstract Polar Factorization Theorem

We do not know exactly which conditions H, S, and $*$ must satisfy to enforce the polar factorization of H by S and K. However, we can prove:

THEOREM 2.1 (Abstract polar factorization). *If $(S, *)$ is a group, then the polar factorization of H by S and K holds. Moreover, for each $u \in H\backslash N$:*

(a) *there is a unique Hilbert projection $\pi(u)$ on S;*

(b) *the factor $s \in S$ in the polar factorization $u = s * k$, $(s, k) \in S \times K$ precisely is $\pi(u)$;*

(c) $\{\pi(u)\} = \partial J(u)$, *where J is the Lipschitz continuous convex function $J(u) = \sup_{s \in S}((u, s))$.*

Notice that J is the Legendre-Fenchel transform of the indicator function of S, $I_S(u) = 0$ if $u \in S$, $= +\infty$ otherwise.

The proof is essentially based on the group property (the semigroup property is not sufficient) and Edelstein's theorem [4] that asserts the following:

THEOREM 2.2 (Edelstein's theorem). *Let S be a closed-bounded subset of a real Hilbert space H. Then, the set of all $u \in H$ for which there is a unique Hilbert projection $\pi(u)$ on S contains a dense countable intersection of open sets $H \backslash N$, defined by*

$$(22)\qquad \begin{aligned} H\backslash N = \{u \in H;\ &\forall \epsilon > 0,\ \exists \delta > 0 \text{ such that} \\ &\|s_1 - s_2\| \leqq \epsilon,\ \forall s_1, s_2 \in S, \text{ whenever} \\ &\|s_i - u\| \leqq \operatorname{dist}(u, S) + \delta,\ i = 1, 2\}. \end{aligned}$$

Moreover π is continuous from $H \backslash N$ into H.

The proof given in [4] follows from the Ekeland-Lebourg theorem on local "ϵ-differentiability". Let us sketch a more direct proof.

Proof of Edelstein's theorem: Let us show that, for a fixed $\epsilon > 0$, the set T_ϵ defined by

$$(23)\qquad \begin{aligned} T_\epsilon = \{u \in H;\ &\exists \delta > 0 \text{ such that} \\ &\|s_1 - s_2\| \leqq \epsilon,\ \forall s_1, s_2 \in S, \text{ whenever} \\ &\|s_i - u\| \leqq \operatorname{dist}(u, S) + \delta,\ i = 1, 2\} \end{aligned}$$

is a dense open subset of the Hilbert space H. The rest of the theorem then easily follows. The fact that T_ϵ is open is almost obvious. Indeed, assume that u belongs to T_ϵ and take any $v \in H$ such that $\|u - v\| \leqq \delta/3$. Now, consider the set $A = \{s \in S;\ \|s - v\| \leqq \operatorname{dist}(v, S) + \delta/3\}$. If s_i, $i = 1, 2$, belong to A, then

$$\|s_i - u\| \leqq \|s_i - v\| + \|v - u\| \leqq \operatorname{dist}(v, S) + \delta/3 + \|v - u\| \leqq \operatorname{dist}(u, S) + \delta$$

which implies $\|s_1 - s_2\| \leqq \epsilon$, since u belongs to T_ϵ. It follows that v also belongs to T_ϵ, which shows that T_ϵ is open. To prove that T_ϵ is dense in H, let us consider $u \in H$ and set $d = \operatorname{dist}(u, S)$. If $d = 0$ (i.e., u belongs to S), then u obviously belongs to T_ϵ. Indeed, the diameter of $\{s \in S;\ \|s-u\| \leqq \operatorname{dist}(u, S)+\epsilon/2\}$ is less than ϵ. If $d>0$, it is not restrictive to assume $u=0$, $d=1$. Then, there is $s_0 \in S$ such that $\|s_0\|=R$ for some $R \geqq 1$ arbitrarily close to 1. Let us introduce $u_r = rs_0/R$, where $r \in]0,1[$ is arbitrarily small, and show that u_r belongs to T_ϵ, if R is appropriately chosen. Let us consider the set $A=\{s \in S;\ \|s-u_r\| \leqq \operatorname{dist}(u_r, S)+\delta\}$ and find $\delta>0$ so that the diameter of A is smaller than ϵ. We have $\operatorname{dist}(u_r, S) \leqq \|u_r - s_0\| = \|rs_0/R - s_0\| = R-r$. Thus, A is contained in the ball $B(u_r, R-r+\delta)$. Since $d=\operatorname{dist}(0, S)=1$, S is contained in $\{v \in H;\ \|v\| \geqq 1\}$ and, therefore, A is also contained in $\{v \in H;\ \|v\| \geqq 1\}$. For any $v \in A$, let us estimate $\|v-s_0\|^2 = \|v\|^2 - 2((v, s_0)) + R^2$. Because $\|v-u_r\| \leqq R-r+\delta$ and $u_r = rs_0/R$, we get $\|v\|^2 - 2r/R((v, s_0)) + r^2 \leqq (R-r+\delta)^2$ and $\|v-s_0\|^2 \leqq \|v\|^2(1-R/r) + R^2 + R/r[(R-r+\delta)^2 - r^2]$. Since $R>r$ and $\|v\| \geqq 1$, we deduce $\|v-s_0\|^2 \leqq 1 - R/r + R^2 + R/r[(R-r+\delta)^2 - r^2] = (R/r-1)(1-R^2+2\delta R) + \delta^2 R/r$. By choosing first R, close enough to 1, and, then, $\delta>0$ sufficiently small, the right-hand side becomes smaller than $\epsilon^2/4$, which proves that the diameter of A is smaller than ϵ and shows that u_r belong to T_ϵ. Since $\|u - u_r\| = r$ is arbitrarily small, this proves that T_ϵ is dense in H and achieves the main part of the proof of Edelstein's theorem.

Let us now get back to the polar factorization framework. In the case considered in Theorem 2.1, S is a bounded closed subset of H and, therefore, Edelstein's theorem can be applied. Moreover, S is contained in a sphere centered at the origin. Indeed,

$$\|s\| = \|e*s\| = \|e\| = \text{cst}, \qquad \forall s \in S$$

immediately follows from (18). This allows us to use the following characterization of the projection operator π.

PROPOSITION 2.2 (Characterization of π). *Let S be a closed subset of a sphere centered at the origin in a real Hilbert space H. Then, the projection operator $\pi : H \to S$ can be characterized as the gradient of the Lipschitz continuous function*

$$J(u) = \sup_{s \in S}((u, s)). \tag{24}$$

More precisely, one has $\partial J(u) = \{\pi(u)\}$, for all $u \in H\backslash N$, *where $H\backslash N$ is defined by* (22).

Before proving this proposition, let us first deduce Theorem 2.1.

Step 1 (existence of a polar factorization). By Edelstein's theorem and Proposition 2.2, we know that, for every $u \in H\backslash N$ (where $H\backslash N$ is defined by (22)), there is a unique projection $s=\pi(u) \in S$. Since S is a group, $k = s^{-1}*u$ is well defined in H and $u = s*k$. Thus, to prove the existence of a polar factorization of

u, it is enough to show that k belongs to the polar cone K. This can be easily deduced from the following calculations: for every $\sigma \in S$,

$$((k, e - \sigma)) = ((s*k, s - s*\sigma)) \text{ (by property (18))}$$

$$= ((u, s - s*\sigma)) \text{ (by definition of } k)$$

$$= \tfrac{1}{2}\{\|u - s*\sigma\|^2 - \|u - s\|^2\} \text{ (because } S \text{ is contained in a sphere)}$$

$$\geqq 0 \text{ (since } s = \pi(u)).$$

Step 2 (uniqueness of the polar factorization). Let $u \in H\backslash N$. Assume that there is $s \in S$ and $k \in K$ such that $u = s*k$. Since $k = s^{-1}*u$ (because of the group property), it is sufficient to prove that s is unique and, more precisely, that s is the unique projection $\pi(u)$ of u on S, that is $((u, s)) \geqq ((u, \sigma))$, for all $\sigma \in S$, or, equivalently, $((s*k, s)) \geqq ((s*k, \sigma))$. By property (18); this exactly means $((k, e)) \geqq ((k, s^{-1}*\sigma))$ and is always true, since k is assumed to belong to K.

Step 3 (continuity of the polar factorization). Let $u_n = s_n*k_n \in H\backslash N$ that converges to $u = s*k \in H\backslash N$. By Edelstein's theorem, $s_n = \pi(u_n)$ converges to $s = \pi(u)$. Let us prove that k_n converges to k. By (18) we have:

$$\|k_n - k\| = \|s_n^{-1}*u_n - s^{-1}*u\| = \|u_n - s_n*s^{-1}*u\|.$$

Since s and u are fixed, $s_n*s^{-1}*u$ converges to $s*s^{-1}*u = u$ (by (18)). Since u_n converges to u (by assumption), it follows that k_n converges to k. This achieves the proof of Theorem 2.1, provided that we prove Proposition 2.2.

Proof of Proposition 2.2: Since S is contained in a sphere centered at the origin, for each $u \in H$, $s \in S$ is the Hilbert projection of u on S if and only if s maximizes $((u, s))$ (indeed, $\|u - s\|^2 = \|u\|^2 + \text{cst} - 2((u, s)))$. It follows from the definition of J (24) that

$$J(u) = ((u, \pi(u))), \qquad \forall u \in H\backslash N. \tag{25}$$

To prove that π is the gradient of J, we shall use two elementary lemmas of convex analysis.

LEMMA 2.1. *For any $u \in H$, $\partial J(u)$ is contained in the closure of the convex hull of S in H.*

LEMMA 2.2. *For any $u \in H$ and any $p \in \partial J(u)$, $((p, u)) = J(u)$.*

Before proving these lemmas, let us deduce Proposition 2.2 from them by showing $\partial J(u) = \{\pi(u)\}$ for every $u \in H\backslash N$.

Step 1. It is easy to check that $\pi(u)$ belongs to $\partial J(u)$, which means $J(v) \geqq J(u) + ((\pi(u), v - u))$, for all $v \in H$. By (25), this is equivalent to $J(v) \geqq ((\pi(u), v))$, which immediately follows from the definition of J (24).

Step 2. Let us prove that $p = \pi(u)$ for any $p \in \partial J(u)$. By Lemmas 2.1 and 2.2, we know that $((p, u)) = J(u)$ and p belongs to the closure of the convex hull of S. Since u belongs to $H \backslash N$, by definition (22) of $H \backslash N$ in Edelstein's theorem, for each fixed $\epsilon > 0$, there is $\delta > 0$ such that every $s \in S$

$$\|s - u\|^2 \leqq \|\pi(u) - u\|^2 + \delta \Longrightarrow \|\pi(u) - s\| \leqq \epsilon/2. \tag{26}$$

Let us choose $\gamma > 0$ so that $\gamma \|u\| \leqq \frac{1}{2}\delta^2$. Because p belongs to the closure of the convex hull of S, there is a convex combination

$$\sum_{i \in I} \theta_i s_i,\ \theta_i \geqq 0, \quad \sum_{i \in I} \theta_i = 1, \quad s_i \in S$$

such that

$$\|p - \sum_{i \in I} \theta_i s_i\| \leqq \gamma,$$

which implies

$$\sum_{i \in I} \theta_i((s_i, u)) \geqq ((p, u)) - \gamma\|u\| \geqq J(u) - \tfrac{1}{2}\delta^2.$$

Let us introduce

$$a_i = 2J(u) - 2((s_i, u)) = 2((\pi(u), u)) - 2((s_i, u)) = \|u - s_i\|^2 - \|u - \pi(u)\|^2.$$

We have $a_i \geqq 0$ and $\sum_{i \in I} \theta_i a_i \leqq \delta^2$. Thus, by Chebyshev's inequality, there is a subset I' of I such that $a_i \leqq \delta$ for each $i \in I'$ and $\sum_{i \in I'} \theta_i \geqq 1 - \delta$. By (26) we deduce that $\|\pi(u) - s_i\| \leqq \epsilon/2$ for each $i \in I'$. It follows that

$$\begin{aligned}\|\pi(u) - p\| \leqq \|\pi(u) - \sum_{i \in I'} \theta_i s_i\| + \|\sum_{i \in I \backslash I'} \theta_i s_i\| \\ + \|p - \sum_{i \in I} \theta_i s_i\| \leqq \epsilon/2 + \delta\|e\| + \gamma\end{aligned}$$

(since $\sum_{i \in I \backslash I'} \theta_i \leqq \delta$ and $\|s_i\| = \|e\|$). Because γ and δ can be chosen so that $\gamma + \delta\|e\| \leqq \epsilon/2$ where ϵ is arbitrarily small, we conclude that $p = \pi(u)$, which achieves the proof of Proposition 2.2

Let us now prove Lemmas 2.1 and 2.2. Lemma 2.1 is a direct consequence of the definition (24) of J (use the Hahn-Banach theorem, for instance).

Proof of Lemma 2.2

If $p \in \partial J(u)$, we have $J(v) \geqq J(u) + ((p, v - u))$, for all $v \in H$ and, in particular for $v = 0$, $0 = J(0) \geqq J(u) - ((p, u))$. Thus $J(u) \leqq ((p, u))$. By Lemma 2.1, p belongs to the closure of the convex hull of S. Thus $((p, u)) \leqq \sup\{((s, u)); s \in S\}$. Since this supremum precisely is $J(u)$, it follows that $((p, u)) = J(u)$, which achieves the proof.

Finally, the proof of Theorem 2.1 is completed.

2.3. Proof of Proposition 2.1

Let us split the proof into three steps.

Step 1 (each $u \in K$ is the gradient of a function in $W^{1,2}(\Omega)$). An easy way to build a family of Lebesgue measure-preserving mappings from $\bar{\Omega}$ into itself is to integrate a smooth, compactly supported in Ω, divergence free vector field w; see [1]. The corresponding flowmap $t \in \mathbb{R} \to g(t) = \exp(tw)$ gives, for each fixed t, a smooth Lebesgue measure-preserving mapping $g(t) \in S$ (indeed, $g(t)$ is a diffeomorphism from $\bar{\Omega}$ into itself and the jacobian determinant $\det(D_x g(t, x))$ is identically equal to 1). Moreover, $g(t) = e + tw + O(t^2)$, $t \to 0$, where e denotes the identity map. By definition (21) of K, for each $u \in K$ and each $s \in S$, $\int_\Omega u(x) \cdot (x - s(x))\, dx \geqq 0$. Thus

$$\int_\Omega u(x) \cdot (-tw(x) + O(t^2))\, dx \geqq 0$$

follows, by setting $s = g(t)$ and, therefore, $\int_\Omega u(x) \cdot w(x)\, dx$ must vanish for any smooth compactly supported divergence free vector field w defined on Ω. This implies (see [12], Chapter 9A) that there is a distribution ψ such that $u = \nabla\psi$ in the sense of distributions and, since $u \in L^2(\Omega; \mathbb{R}^d)$, ψ belongs to the Sobolev space $W^{1,2}(\Omega)$.

Step 2 (u is monotone and ψ is convex). To prove that the potential ψ is convex, it is enough to show that

$$(27) \qquad (u(x_1) - u(x_2)) \cdot (x_1 - x_2) \geqq 0, \quad \text{for a.e.} \quad x_1, x_2 \in \Omega.$$

Because u is Lebesgue integrable, almost every point $x \in \Omega$ is a Lebesgue point, which means

$$(28) \qquad u(x) = \lim_{\epsilon \to 0} |B|^{-1} \int_B u(x + \epsilon y)\, dy,$$

where B denotes the unit ball in $\mathbb{R}^d$.

Let us consider a pair (x_1, x_2) of such points. When ϵ is small enough, then the following mapping s_ϵ is Lebesgue measure-preserving from Ω into itself

$$
s_\epsilon(x) = \begin{cases} x - x_1 + x_2 & \text{if} \quad x \in B(x_1, \epsilon) \\ x - x_2 + x_1 & \text{if} \quad x \in B(x_2, \epsilon) \\ x, & \text{otherwise.} \end{cases} \tag{29}
$$

Since u belongs to K, we have $\int_\Omega u(x) \cdot (x - s_\epsilon(x))\, dx \geqq 0$, that is

$$
(x_1 - x_2) \cdot \left[\int_B u(x_1 + \epsilon y)\, dy - \int_B u(x_2 + \epsilon y)\, dy \right] \geqq 0,
$$

which immediately leads to (27) and shows that $u = \nabla \psi$ is monotone and, therefore, ψ is convex.

Notice that this result could have been obtained in just one step by using the following characterization of the subdifferential u of a convex function ψ, due to Rockefaeller [8],

$$
\sum_{i=1,n} u(x_i) \cdot (x_i - x_{i-1}) \geqq 0
$$

for (almost) all finite sequences $x_1, \cdots, x_n = x_0$ of points in Ω. This inequality would have been directly obtained by using the following volume-preserving mapping

$$
s_\epsilon(x) = \begin{cases} x - x_{i+1} + x_i & \text{if} \quad x \in B(x_i, \epsilon) \\ x, & \text{otherwise,} \end{cases} \tag{30}
$$

for ϵ small enough and for any sequences of Lebesgue points $x_1, \cdots, x_n = x_0$.

Step 3. To achieve the proof, let us show that $\nabla \psi$ belongs to K for each convex function ψ in $W^{1,2}(\Omega)$. This immediately follows from the convexity property, since

$$
\psi(s(x)) \geqq \psi(x) + \nabla \psi(x) \cdot (s(x) - x), \quad \text{a.e. in} \quad \Omega
$$

holds for every Lebesgue measure-preserving mapping s and, after integrating over Ω,

$$
\int_\Omega \nabla \psi(x) \cdot (x - s(x))\, dx \geqq \int_\Omega [\psi(x) - \psi(s(x))]\, dx = 0,
$$

which means that $\nabla \psi$ belongs to K.

2.4. Proof of a Comment on the Rearrangement Theorem

In the introduction, we asserted that property

$$(31) \qquad \int_\Omega u_1^\#(x) \cdot u_2^\#(x)\, dx \geqq \int_\Omega u_1(x) \cdot u_2(x)\, dx$$

is not true for all u_1, u_2 in $L^2(\Omega; \mathbb{R}^d)$, where $u_i^\# = \nabla\psi_i$, $i = 1, 2$, are the rearrangements of u_1, u_2, when the dimension d is larger than 1 (for $d = 1$ and $\Omega =]0, 1[$, this property always holds). Let us now justify this statement. If property (31) were true, it would follow that

$$(32) \qquad \int_\Omega \nabla\psi_1(x) \cdot \nabla\psi_2(x)\, dx \geqq \int_\Omega \nabla\psi_1(x) \cdot \nabla\psi_2(s(x))\, dx$$

holds for every pair (ψ_1, ψ_2) of convex functions in $W^{1,2}(\Omega)$ and any measure-preserving mapping s. By using exactly the same argument as in the previous subsection (where we proved that any $u \in K$ is monotone), we would deduce

$$(33) \qquad [\nabla\psi_1(x_1) - \nabla\psi_1(x_2)] \cdot [\nabla\psi_2(x_1) - \nabla\psi_2(x_2)] \geqq 0, \quad \text{a.e.} \quad x_1, x_2 \in \Omega.$$

It is easy to fine a counter example. Take, for instance,

$$\psi_1(x) = \tfrac{1}{2} A_1^x x, \qquad \psi_2(x) = \tfrac{1}{2} A_2^x x,$$

where A_1, A_2 are two symmetric non-negative real $d \times d$ matrices. Then, property (33) would exactly mean that $A_1 A_2 + A_2 A_1$ is symmetric non-negative, which cannot be true in general. Indeed, if

$$A_1 = \begin{pmatrix} 1 & 0 \\ 0 & \delta \end{pmatrix} \qquad A_2 = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix},$$

then $A_1 A_2 + A_2 A_1$ is not non-negative for $\delta > 0$ sufficiently small.

3. Proof of the Main Results

3.1. The Monge-Kantorovich Problem

Our proof is based on the study of a particular Monge-Kantorovich problem (MKP) (see [19] for a general review). We are given a probability measure α on $\mathbb{R}^d$ such that $\int (1 + \|y\|)\, d\alpha(y) < +\infty$ and the three following problems are considered.

The Primal MKP

Find $\phi \in C(\mathbb{R}^d) \cap L^1(\mathbb{R}^d, \alpha)$ and $\psi \in C(\Omega) \cap L^1(\Omega, \beta)$ that minimize $\int \phi \, d\alpha$ and satisfy the following conditions.

$$\int \psi \, d\beta = 0$$

$$\phi(y) + \psi(z) \geqq y \cdot z, \qquad \forall (y, z) \in \mathbb{R}^d \times \Omega.$$

The Dual MKP

Find a probability measure p on $\mathbb{R}^d \times \bar{\Omega}$ that maximizes $\int y \cdot z \, dp(y, z)$ under the following conditions: $\int \|y\| \, dp(y, z) < +\infty$, α and β are the marginals of p on $\mathbb{R}^d$ and $\bar{\Omega}$, which means

$$\int f(y) \, dp(y, z) = \int f(y) \, d\alpha(y)$$

for each $f \in C(\mathbb{R}^d)$ such that $|f(y)| \leqq \mathrm{cst}(1 + \|y\|)$, and

$$\int g(z) \, dp(y, z) = \int g(z) \beta(z) \, dz, \qquad \forall g \in C(\bar{\Omega}).$$

The Mixed MKP

Find $\phi \in C(\mathbb{R}^d) \cap L^1(\mathbb{R}^d, \alpha)$, $\psi \in C(\Omega) \cap L^1(\Omega, \beta)$ and a probability measure p on $\mathbb{R}^d \times \bar{\Omega}$ such that

$$\phi(y) + \psi(z) \geqq y \cdot z, \qquad \forall (y, z) \in \mathbb{R}^d \times \Omega;$$

$\int \|y\| \, dp(y, z) < +\infty$, α and β are the marginals of p on $\mathbb{R}^d$ and $\bar{\Omega}$, $\int \psi \, d\beta = 0$, and $\int \phi \, d\alpha \leqq \int y \cdot z \, dp(y, z)$.

Notice that, in the dual and the mixed MKP, p necessarily is a tight probability measure on $\mathbb{R}^d \times \bar{\Omega}$ (since $\int \|y\| \, dp(y, z) < +\infty$) and $p(\mathbb{R}^d \times \partial\Omega) = \beta(\partial\Omega)$ (since β is the marginal of p) $= 0$ (by assumption, cf. Introduction).

Our study of these three problems involves in an essential way the subset K_0 of the Sobolev space $W^{1,1}(\Omega, \beta)$, defined by:

$$K_0 = \{\psi \in W^{1,1}(\Omega, \beta) \cap C(\Omega); \int \psi \, d\beta = 0, \tag{34}$$

$$\exists \tilde{\psi} : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}, \text{ convex, l.s.c.,} \quad \text{such that} \quad \psi = \tilde{\psi} \text{ on } \Omega\}.$$

An existence proof will be given in Section 3.4 for the mixed MKP. From this existence proof, our main results (as listed in the Introduction) will be obtained in Section 3.3, with the help of the following "*a priori*" results.

PROPOSITION 3.1. *Assume that the mixed MKP has a solution (ψ, ϕ, p). Then*:

(i) *ψ belongs to K_0 (defined by (34)) and $\|\nabla\psi\|_{L^1(\Omega,\beta)} = \int \|y\|\, d\alpha(y)$;*

(ii) *$\phi = \psi^*$, α-almost everywhere on $\mathbb{R}^d$, where ψ^* is the Legendre transform of ψ defined by*

$$\psi^*(y) = \sup_{z \in \Omega} \{y \cdot z - \psi(z)\}, \qquad \forall y \in \mathbb{R}^d;$$

(iii) *$dp(y, z) = \delta(y - \nabla\psi(z))\beta(z)\, dz$;*

(iv) *$\int \phi\, d\alpha = \int y \cdot z\, dp(y, z)$ and $\int \psi\, d\beta = 0$;*

(v) *(ψ, ϕ) is the unique solution to the primal MKP and p is the unique solution of the dual MKP.*

Moreover, if α is absolutely continuous with respect to the Lebesgue measure, then

(i) *$z = \nabla\psi^*(y)$, $y = \nabla\psi(z)$, p-almost everywhere on $\mathbb{R}^d \times \bar{\Omega}$;*

(ii) *$\nabla\psi^*(\nabla\psi(z)) = z$, β-almost everywhere on $\bar{\Omega}$ and $\nabla\psi(\nabla\psi^*(y)) = y$, α-almost everywhere on $\mathbb{R}^d$;*

(iii) *$dp(y, z) = \delta(y - \nabla\psi^*(y))\, d\alpha(y)$.*

PROPOSITION 3.2. *Let (α_n) be a sequence of probability measures on $\mathbb{R}^d$ such that $\int f\, d\alpha_n \to \int f\, d\alpha$, for any $f \in C(\mathbb{R}^d)$ such that $|f(y)| \leq \mathrm{cst}(1 + \|y\|)$. If (ψ_n, ϕ_n, p_n) is a solution to the mixed MKP corresponding to α_n, then the mixed MKP corresponding to α has a unique solution (ψ, ϕ, p) and $\phi_n \to \psi$ uniformly on any compact subset of $\mathbb{R}^d$, $\psi_n \to \psi$ in $W^{1,1}(\Omega, \beta)$, $\int f\, dp_n \to \int f\, dp$, for each $f \in C(\mathbb{R}^d \times \bar{\Omega})$ such that $|f(y, z)| \leq \mathrm{cst}(1 + \|y\|)$.*

From the existence result (proved in Section 3.4) and Propositions 3.1 and 3.2, we finally get:

THEOREM 3.1 (Solution of the MKP). *The mixed MKP has a unique solution (ψ, ϕ, p), (ψ, ϕ) is the unique solution of the primal MKP, and p is the unique solution of the dual MKP. All the properties listed in Propositions* 3.1 *and* 3.2 *are satisfied.*

3.2. Proof of Propositions 3.1 and 3.2

Notice that, since Ω is bounded, $\bar{\Omega}$ is contained in the ball $B(0, r)$ for some $r > 0$. To prove Propositions 3.1 and 3.2, we shall use a preliminary result on the set K_0 defined by (34).

PROPOSITION 3.3 (Main properties of the set K_0).

(a) *For each* ψ *in* K_0, *the following properties hold*:

(i) $-2rM \leqq \psi(z)$, *for all* $z \in \Omega$, *where* $M = \int \|\nabla\psi\|\, d\beta$,

(ii-a) $\psi(z) \leqq 2r\|\nabla\psi(z)\|$, *almost everywhere on* Ω,

(ii-b) $\psi(z) \leqq C(\delta)rM$, *for all* $z \in \Omega$, *where* $\delta = \operatorname{dist}(z, \partial\Omega)$,

(iii) $\int |\psi|\, d\beta \leqq 4rM$.

(iv) *The Legendre-Fenchel transform* ψ^* *of* ψ, *defined by*

$$\psi^*(y) = \sup_{z \in \Omega}\{y \cdot z - \psi(z)\}, \qquad \forall y \in \mathbb{R}^d, \tag{35}$$

is Lipschitz continuous on $\mathbb{R}^d$ *and satisfies* $-r\|y\| \leqq \psi^*(y) \leqq r\|y\| + 2rM$, *for all* $y \in \mathbb{R}^d$, $\operatorname{Lip}(\psi^*) \leqq r$ *and* $\psi(z) = \psi^{**}(z) = \sup_{y \in \mathbb{R}^d}\{y \cdot z - \psi^*(y)\}$, *for all* $z \in \Omega$.

(b) *Let* (ψ_n) *be a sequence in* K_0 *such that* $\int \|\nabla\psi_n\|\, d\beta \leqq M$. *Then, there is a subsequence, still labelled by* n, *and a pair* (ψ, ϕ) *such that*

$$\psi \in C(\Omega) \cap L^1(\Omega, \beta),\ \phi \in C(\mathbb{R}^d),$$

$$\phi(y) + \psi(z) \geqq y \cdot z,\ \forall y \in \mathbb{R}^d,\ \forall z \in \Omega,$$

$\psi_n \to \psi$ *in* $L^1(\Omega, \beta)$ *and uniformly on any compact subset of* Ω,

$\psi_n^* \to \phi$ *uniformly on any compact subset of* $\mathbb{R}^d$

and $|\psi_n^*(y)| \leqq r(2M + \|y\|)$.

Proof of Proposition 3.3. Let us first recall that Ω is supposed to be contained in the ball $B(0, r)$ for some $r > 0$ and that $W^{1,1}(\Omega, \beta)$ is supposed to be compactly embedded into $L^1(\Omega, \beta)$.

Step 1. Let $\psi \in K_0$. Since ψ is continuous on Ω and $\int \psi\, d\beta = 0$, there is $z_0 \in \Omega$ such that $\psi(z_0) = 0$. Since ψ belongs to $W^{1,1}(\Omega, \beta)$, $\nabla\psi$ is defined almost everywhere on Ω (with respect to both β and the Lebesgue measure, since these measures have the same negligible sets, by assumption). By convexity, one gets $\psi(z_0) \geqq \psi(z) + \nabla\psi(z)\cdot(z_0 - z)$, for almost every $z \in \Omega$, and, thus, $\psi(z) \leqq -\nabla\psi(z)\cdot(z_0 - z) \leqq 2r\|\nabla\psi(z)\|$ (since $z, z_0 \in \Omega \subset B(0, r)$). We also have for almost every $z, \tilde{z} \in \Omega$, $\psi(z) \geqq \psi(\tilde{z}) + \nabla\psi(\tilde{z})\cdot(z - \tilde{z})$, and, hence, after integrating this inequality over Ω with respect to $\tilde{z}$,

$$\psi(z)\int d\beta \geqq \int \psi(\tilde{z})\beta(\tilde{z})\, d\tilde{z} + \int \nabla\psi(\tilde{z})\cdot(z - \tilde{z})\beta(\tilde{z})\, d\tilde{z}.$$

Since β is a probability measure and $\int \psi\, d\beta = 0$ (because ψ belongs to K_0), it follows that

$$\psi(z) \geqq \int \nabla\psi(\tilde{z})\cdot(z-\tilde{z})\beta(\tilde{z})\, d\tilde{z} \geqq -2r\int \|\nabla\psi\|\, d\beta$$

holds for almost every $z \in \Omega$. So far, we have proven $-2rM \leqq \psi(z) \leqq 2r\|\nabla\psi(z)\|$, almost everywhere on Ω.

By integrating this inequality over Ω, we get

$$\int |\psi|\, d\beta \leqq 2r\int \|\nabla\psi\|\, d\beta + 2rM\int d\beta = 4rM.$$

Let us now fix $z \in \Omega$ and set $\delta = \operatorname{dist}(z, \partial\Omega) > 0$. If B denotes the unit ball in $\mathbb{R}^d$, then Ω contains both $z + \delta/2B$ and $\bar{\omega} = \{\tilde{z} \in \Omega; \operatorname{dist}(\tilde{z}, \partial\Omega) \geqq \delta/2\}$. By assumption (cf. Introduction), β is essentially bounded away from 0 on $\bar{\omega}$ by some constant $\rho(\delta) > 0$. Thus, ψ is Lebesgue integrable on $\bar{\omega}$ and

$$\int_{\bar{\omega}} |\psi(\tilde{z})|\, d\tilde{z} \leqq \rho(\delta)^{-1}\int |\psi|\, d\beta \leqq 4rM\rho(\delta)^{-1}.$$

By convexity,

$$\psi(z) \leqq \left|z + \frac{\delta}{2}B\right|^{-1} \int_{z+\delta/2B} \psi(\tilde{z})\, d\tilde{z}.$$

Thus (since $\bar{\omega}$ contains $z + \delta/2B$),

$$\psi(z) \leqq \left|z + \frac{\delta}{2}B\right|^{-1} \rho(\delta)^{-1}4rM = |B|^{-1}\left(\frac{\delta}{2}\right)^{-d} \rho(\delta)^{-1}4rM,$$

which means that there is a constant $C(\delta)$ such that $\psi(z) \leqq C(\delta)rM$, for all $z \in \Omega$, where $\delta = \operatorname{dist}(z, \partial\Omega)$.

Step 2. By definition (35), ψ^* is well defined and satisfies $\psi^*(y) \geqq z_0\cdot y - \psi(z_0) = z_0\cdot y$ (by definition of z_0), $\geqq -r\|y\|$, and

$$\psi^*(y) \leqq r\|y\| - \inf_{\Omega}\psi \leqq r(\|y\| + 2M).$$

Moreover ψ^* is Lipschitz continuous on $\mathbb{R}^d$ (indeed in definition (35) the supremum is taken over $\Omega \subset B(0, r)$) and $\operatorname{Lip}(\psi^*) \leqq r$.

Since $\psi = \tilde{\psi}$ on Ω, where ψ is a convex l.s.c function $\mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$, we also have $\psi^{**} = \psi$ on Ω (this is a classical result in convex analysis; see [13]).

Step 3. Let (ψ_n) be a sequence in K_0 such that $\int \|\nabla\psi_n\|\, d\beta \leqq M$. From the former estimates, it follows that

(i) (ψ_n) is a bounded sequence in $W^{1,1}(\Omega, \beta)$ and, therefore, by assumption (cf. Introduction); (ψ_n) has a convergent subsequence, still labelled by n, in $L^1(\Omega, \beta)$. Moreover, since $-2rM \leqq \psi_n(z) \leqq C(\delta)rM$, where δ is $\mathrm{dist}(z, \partial\Omega)$, for any $z \in \Omega$, the ψ_n are uniformly bounded and, thus, uniformly Lipschitz continuous (cf. [13]) on any compact subset of Ω;

(ii) (ψ_n^*) is uniformly Lipschitz continuous on $\mathbb{R}^d$ and

$$|\psi_n^*(y)| \leqq r(2M + \|y\|), \qquad \forall y \in \mathbb{R}^d.$$

Thus, there is a subsequence, still labelled by n, and a pair $\phi \in C(\mathbb{R}^d)$, $\psi \in C(\Omega) \cap L^1(\Omega, \beta)$, such that

$\psi_n^* \to \phi$, uniformly on any compact subset of $\mathbb{R}^d$,

$\psi_n \to \psi$ in $L^1(\Omega, \beta)$ and uniformly on any compact subset of Ω.

Moreover ϕ satisfies $|\phi(y)| \leqq r(2M + \|y\|)$. Notice that

$$\phi(y) + \psi(z) \geqq y \cdot z, \qquad \forall z \in \Omega,\ y \in \mathbb{R}^d$$

immediately follows from the definition of ψ_n^* and the convergence properties. This achieves the proof of Proposition 3.3.

Proof of Proposition 3.1: The proof of Proposition 3.1 which, in our opinion, is the most important of our intermediary results, relies on the following well-known property of convex conjugate functions ψ, ψ^*, namely $\psi(z) + \psi^*(y) = y \cdot z$, if and only if $z \in \partial\psi^*(y)$ and $y \in \partial\psi(z)$. The uniqueness and the precise characterization of the solution to the mixed MKP (whenever it exists) follow from this elementary property.

Step 1. Let (ϕ, ψ, p) be any solution to the mixed MKP. Let us first show that $\psi \in K_0$ and (ϕ, ψ, p) are linked together by the following relations

$$\phi = \psi^*,\ \alpha - \text{a.e.}, \quad \int \phi\, d\alpha = \int y \cdot z\, dp(y, z) \quad \text{and}$$

$$dp(y, z) = \delta(y - \nabla\psi(z))\beta(z)\, dz.$$

Let us first introduce

$$\tilde{\phi}(y) = \sup_{z \in \Omega} \{y \cdot z - \psi(z)\}, \qquad \forall y \in \mathbb{R}^d. \tag{36}$$

Since (ϕ, ψ, p) is a solution to the mixed MKP, we have

$$\phi(y) + \psi(z) \geqq y \cdot z, \qquad \forall z \in \Omega,\ y \in \mathbb{R}^d.$$

Thus $\tilde{\phi}(y) \leqq \phi(y)$, for all $y \in \mathbb{R}^d$. For some fixed $z_0 \in \Omega$, one has $\tilde{\phi}(y) \geqq y \cdot z_0 - \psi(z_0)$, for all $y \in \mathbb{R}^d$. So $\tilde{\phi}$ is finite everywhere and, because Ω is contained in $B(0, r)$, it follows from definition (36) that $\tilde{\phi}$ is convex and Lipschitz continuous on $\mathbb{R}^d$ with $\mathrm{Lip}(\tilde{\phi}) \leqq r$.

Let us now introduce

$$\tilde{\psi}(z) = \sup_{y \in \mathbb{R}^d} \{y \cdot z - \tilde{\phi}(y)\}, \qquad \forall z \in \mathbb{R}^d. \tag{37}$$

Here $\tilde{\psi}$ is a well defined convex l.s.c. function from $\mathbb{R}^d$ into $\mathbb{R} \cup \{+\infty\}$ and $\tilde{\phi}$, $\tilde{\psi}$ are convex conjugates (see [13]): $\tilde{\psi} = \tilde{\phi}^*$, $\tilde{\phi} = \tilde{\psi}^*$. Moreover, one gets from definitions [36] and [37]

$$\tilde{\psi}(z) \leqq \psi(z), \tilde{\psi}(z) \geqq -\tilde{\phi}(0) > -\infty, \qquad \forall z \in \Omega.$$

Since we know that $\int (1 + \|y\|)\, d\alpha(y) < +\infty$, $\int (1 + |\psi(z)|)\beta(z)\, dz < +\infty$, it follows from the bounds on $\tilde{\psi}$ and $\tilde{\phi}$ that they respectively belong to $L^1(\Omega, \beta)$ and $L^1(\mathbb{R}^d, \alpha)$. In addition,

$$\int \tilde{\phi}\, d\alpha \leqq \int \phi\, d\alpha, \int \tilde{\psi}\, d\beta \leqq \int \psi\, d\beta,$$

and, by definition of $\tilde{\psi}$, $\tilde{\phi}(y) + \tilde{\psi}(z) \geqq y \cdot z$, for all $y, z \in \mathbb{R}^d$. Thus, since α and β are the marginals of p, one deduces

$$\begin{aligned} 0 &\leqq \int [\tilde{\phi}(y) + \tilde{\psi}(z) - y \cdot z]\, dp(y, z) \\ &= \int \tilde{\phi}\, d\alpha + \int \tilde{\psi}\, d\beta - \int y \cdot z\, dp(y, z) \\ &\leqq \int \phi\, d\alpha + \int \psi\, d\beta - \int y \cdot z\, dp(y, z). \end{aligned}$$

The right-hand side of the last inequality is not larger than 0, since (ϕ, ψ, p) is a solution to the mixed MKP. It follows that these inequalities actually are equalities and (since $\tilde{\phi} \leqq \phi$, $\tilde{\psi} \leqq \psi$)

$$\begin{gathered} \tilde{\phi} = \phi, \alpha - \text{a.e.}, \qquad \tilde{\psi} = \psi, \beta - \text{a.e.}, \\ \tilde{\phi}(y) + \tilde{\psi}(z) = y \cdot z, \qquad p - \text{a.e. on } \mathbb{R}^d \times \bar{\Omega}, \\ \int \tilde{\phi}\, d\alpha = \int \phi\, d\alpha = \int y \cdot z\, dp(y, z), \qquad \int \tilde{\psi}\, d\beta = \int \psi\, d\beta = 0. \end{gathered} \tag{38}$$

By definition of the mixed MKP, ψ belongs to $C(\Omega)$. Since the convex function $\tilde{\psi}$ is equal to the continuous function ψ, β-almost everywhere on Ω and β has the same negligible sets as the Lebesgue measure, it follows that $\tilde{\psi}$ is locally Lipschitz continuous on Ω (cf. [13]) and, therefore, equal to ψ everywhere on Ω. Moreover $\nabla\psi$ is well defined (up to a Lebesgue negligible set) as a Borelian mapping from Ω into $\mathbb{R}^d$ and $\partial\psi(z) = \{\nabla\psi(z)\}$ holds almost everywhere on Ω (here ∂ denotes the subdifferential of a convex function; see [13]).

According to a well-known result in convex analysis (see [13]), it follows from properties (38) that

$$z \in \partial\tilde{\phi}(y),\ y \in \partial\psi(z) p - \text{a.e. } (y, z) \in \mathbb{R}^d \times \bar{\Omega}.$$

Since α and β are the marginals of p on $\bar{\Omega}$ and $\mathbb{R}^d$, and because β has the same negligible sets as the Lebesgue measure, we deduce

$$p(\{(y, z) \in \mathbb{R}^d \times \bar{\Omega}; \qquad \partial\psi(z) \neq \{\nabla\psi(z)\}\}) = 0$$

(after noticing that $p(\mathbb{R}^d \times \partial\Omega) = \beta(\partial\Omega) = 0$, by assumption (cf. Introduction). By combining these properties, we get

$$y = \nabla\psi(z), p - \text{a.e. } (y, z) \in \mathbb{R}^d \times \bar{\Omega}. \tag{39}$$

Let us now consider an arbitrarily chosen function $f \in C(\mathbb{R}^d \times \bar{\Omega})$ such that

$$|f(y, z) \leqq \text{cst}(1 + \|y\|), \qquad \forall (y, z) \in \mathbb{R}^d \times \bar{\Omega}.$$

Since (ϕ, ψ, p) is a solution to the mixed MKP, $\int (1 + \|y\|)\, dp(y, z)$ must be finite, and, therefore f is p-integrable. Because of property (39), we have

$$\int f(y, z)\, dp(y, z) = \int f(\nabla\psi(z), z)\, dp(y, z)$$

and the right-hand side of this equality is $\int f(\nabla\psi(z), z)\beta(z)\, dz$, because β is the marginal of p on $\bar{\Omega}$. This shows that $dp(y, z) = \delta(y - \nabla\psi(z))\beta(z)\, dz$. Moreover, in the particular case when $f(y, z) = \|y\|$, one gets

$$\int \|\nabla\psi(z)\|\beta(z)\, dz = \int \|y\|\, d\alpha(y) < +\infty,$$

which shows that ψ belongs to the Sobolev space $W^{1,1}(\Omega, \beta)$ (we already know that ψ belongs to $C(\Omega) \cup L^1(\Omega, \beta)$). Finally, since $\psi = \tilde{\psi}$ on Ω, we see that ψ belongs to the set K_0, $\|\nabla\psi\|_{L^1(\Omega;\beta)} = \int \|y\|\, d\alpha(y)$.

So, the proof of the first four statements of Proposition 3.1 is completed.

Step 2. Let us now prove the fifth statement of Proposition 3.1 by showing that, for any solution (ϕ, ψ, p) to the mixed MKP, (ϕ, ψ) is the unique solution to the primal MKP and p is the unique solution to the dual MKP.

Let us consider a solution (ϕ, ψ, p) to the mixed MKP, any solution (ϕ_1, ψ_1) to the primal MKP, and any solution p_1 to the dual MKP. It is straightforward to check that (ϕ, ψ, p_1) and (ϕ_1, ψ_1, p) solve the mixed MKP. Indeed,

$$\int y \cdot z \, dp(y, z) \leqq \int y \cdot z \, dp_1(y, z)$$

(because p_1 solves the dual MKP)

$$\leqq \int \{\phi_1(y) + \psi_1(z)\} \, dp_1(y, z)$$

(because $\phi_1(y) + \psi_1(z) \geqq y \cdot z$ for every $(y, z) \in \mathbb{R}^d \times \Omega$ and $p_1(\mathbb{R}^d \times \partial\Omega) = \beta(\partial\Omega) = 0$)

$$= \int \phi_1 \, d\alpha + \int \psi_1 \, d\beta$$

(since α and β are the marginals of p_1)

$$= \int \phi_1 \, d\alpha \leqq \int \phi \, d\alpha$$

(because (ϕ_1, ψ_1) solves the primal MKP)

$$= \int y \cdot z \, dp(y, z)$$

(as just shown). Thus, all these inequalities become equalities. It follows that (ϕ_1, ψ_1, p) and (ϕ, ψ, p_1) also solve the mixed MKP. According to the first step of the proof, it follows that

$$\psi_1 \in K_0, \quad \phi_1 = \psi_1^*, \quad \alpha - \text{a.e.},$$

$$dp_1(y, z) = \delta(y - \nabla\psi(z))\beta(z) \, dz = dp(y, z),$$

$$y = \nabla\psi_1(z), \quad p - \text{a.e.} \quad \text{on} \quad \mathbb{R}^d \times \bar{\Omega},$$

$$y = \nabla\psi(z), \quad p - \text{a.e.} \quad \text{on} \quad \mathbb{R}^d \times \bar{\Omega}.$$

Thus $p_1 = p$ and $\nabla\psi_1 = \nabla\psi$, almost everywhere on $\bar{\Omega}$ (since β is the marginal of p on $\bar{\Omega}$ and β has the same negligible sets as the Lebesgue measure). Since ψ_1, ψ belong to K_0 and Ω is supposed to be connected, we deduce $\psi_1 = \psi$ and $\phi_1 = \psi_1^* = \psi^* = \phi$, α-almost everywhere.

This shows that (ϕ, ψ) is the unique solution to the primal MKP and p is the unique solution to the dual MKP (notice, however, that ϕ is uniquely defined up to some α-negligible subset of $\mathbb{R}^d$).

Step 3. Let us now consider the particular case when α is absolutely continuous with respect to the Lebesgue measure.

We know that ψ^* is a Lipschitz continuous convex function defined on $\mathbb{R}^d$, and, therefore, $\nabla\psi^*$ is a well defined (up to some Lebesgue negligible set) Borel mapping from $\mathbb{R}^d$ into itself. Since α is absolutely continuous with respect to the Lebesgue measure, the set $\{y \in \mathbb{R}^d;\ \partial\psi^*(y) \neq \{\nabla\psi^*(y)\}\}$ is α-negligible. It follows that $\partial\psi^*(y) = \{\nabla\psi^*(y)\}$ is true for p-almost everywhere $(y, z) \in \mathbb{R}^d \times \bar{\Omega}$, since α is the marginal of p on $\mathbb{R}^d$.

We already know that $z \in \partial\psi^*(y)$ and $y = \nabla\psi(z)$ for p-almost everywhere (y, z). Thus, $z = \nabla\psi^*(y)$ and $y = \nabla\psi(z)$ holds for p-almost everywhere (y, z), and, since α and β are the marginals of p, we get $z = \nabla\psi^*(\nabla(\psi(z))$ for β-almost every $z \in \bar{\Omega}$ and $y = \nabla\psi(\nabla\psi^*(y))$ for α-almost every $y \in \mathbb{R}^d$.

Thus $\nabla\psi$ and $\nabla\psi^*$ are reciprocal. Moreover $dp(y, z) = \delta(z - \nabla\psi^*(y))\, d\alpha(y)$ immediately follows, which achieves the proof of Proposition 3.1.

Proof of Proposition 3.2: Let (α_n) be a sequence of probability measures on $\mathbb{R}^d$ such that

$$\int f\, d\alpha_n \to \int f\, d\alpha, \quad \forall f \in C(\mathbb{R}^d) \quad \text{such that} \quad |f(y)| \leqq \mathrm{cst}(1 + \|y\|).$$

Let (ϕ_n, ψ_n, p_n) be a solution to the mixed MKP corresponding to (α_n). From Proposition 3.1, we know that

$$\psi_n \in K_0 \quad \phi_n = \psi_n^*, \quad \alpha_n - \text{a.e.},$$

$$dp_n(y, z) = \delta(y - \nabla\psi_n(z))\beta(z)\, dz \tag{40}$$

$$\int \|\nabla\psi_n\|\, d\beta = \int \|y\|\, d\alpha_n(y) \to \int \|y\|\, d\alpha(y).$$

From Proposition 3.3, we deduce that, for a subsequence labelled by m, there is a pair (ϕ, ψ) such that

$$\psi \in C(\Omega) \cap L^1(\Omega, \beta), \qquad \phi \in C(\mathbb{R}^d),$$

$$\phi(y) + \psi(z) \geqq y \cdot z, \qquad \forall (y, z) \in \mathbb{R}^d \times \Omega.$$

$\psi_m \to \psi$ in $L^1(\Omega, \beta)$ and uniformly on any compact subset of Ω,

$\phi_m \to \phi$ uniformly on any compact subset of $\mathbb{R}^d$,

and $|\phi_m(y)| \leqq \mathrm{cst}(1 + \|y\|)$.

Because (ϕ_n, ψ_n, p_n) is a solution to the mixed MKP, we have

$$\int \phi_m \, d\alpha_m \leqq \int y \cdot z \, dp_m(y, z), \qquad \int \psi_m \, d\beta = 0.$$

Thus $\int \psi \, d\beta = 0$ and $\int \phi \, d\alpha \leqq \lim \int \phi_m \, d\alpha_m$. (This follows from (i) the uniform convergence of (ϕ_m) on every compact subset of $\mathbb{R}^d$, (ii) the uniform bounds

$$|\phi_m(y)| \leqq \text{cst}(1 + \|y\|), \qquad \int [1 + \|y\|] \, d\alpha_m(y) \leqq \text{cst})$$

Let us now consider the sequence (p_m). The marginals of p_m are α_m and β. Moreover, by assumption,

$$\int [1 + \|y\|] \, dp_m(y, z) = \int [1 + \|y\|] \, d\alpha_m(y) \to \int [1 + \|y\|] \, d\alpha(y).$$

Thus, for a subsequence, still labelled by m, there is a positive measure p on $\mathbb{R}^d \times \Omega$, such that

$$\int f(y, z) \, dp_m(y, z) \to \int f(y, z) \, dp(y, z) \tag{41}$$

for any compactly supported continuous function f on $\mathbb{R}^d \times \bar{\Omega}$, and

$$\int [1 + \|y\|] \, dp_m(y, z) \to \int [1 + \|y\|] \, d\alpha(y).$$

This implies that (i) p is a (tight) probability measure on $\mathbb{R}^d \times \bar{\Omega}$ with marginals α, β and that (ii) equation (41) holds for any continuous function on $\mathbb{R}^d \times \bar{\Omega}$ such that

$$|f(y, z)| \leqq \text{cst}(1 + \|y\|), \qquad \forall (y, z) \in \mathbb{R}^d \times \bar{\Omega}. \tag{42}$$

In particular,

$$\int y \cdot z \, dp_m(y, z) \to \int y \cdot z \, dp(y, z).$$

So, $\int y \cdot z \, dp(y, z) \geqq \int \phi \, d\alpha$, $\int \psi \, d\beta = 0$ and, therefore, (ϕ, ψ, p) satisfies all the conditions required to solve the mixed MKP. By Proposition 3.1, $(\phi = \psi^*, \psi)$ is the unique solution to the primal MKP, p is the unique solution to the dual MKP, and

$$dp(y, z) = \delta(y - \nabla\psi(z))\beta(z) \, dz. \tag{43}$$

Now, it is easy to deduce that (ψ_m) strongly converges to ψ in $W^{1,1}(\Omega, \beta)$ (instead of $L^1(\Omega, \beta)$, as already obtained!). Indeed, we get from equations (41), (40), and (43)

$$\int f(\nabla\psi_m(z), z)\,\beta(z)\,dz \to \int f(\nabla\psi(z), z)\beta(z)\,dz$$

for any continuous function satisfying (42), which is possible only if ψ_m strongly converges to ψ in $W^{1,1}(\Omega, \beta)$.

Because of the uniqueness of the solution (ϕ, ψ, p) to the MKP, the whole sequence (ψ_n) converges to (ψ) in $W^{1,1}(\Omega, \beta)$. In the same way, $\phi_n \to \phi$ uniformly on any compact subset of $\mathbb{R}^d$ and p_n converges to p (in the sense of (41) and (42)).

This achieves the proof of Proposition 3.2.

3.3. Proof of the Main Theorems

Let us recall that the Rearrangement Theorem 1.1 is a corollary of Theorem 1.3 that we are going to prove with the help of Theorem 3.1 and the following characterization result.

PROPOSITION 3.4 (Characterization of $\nabla\psi$).

Let $\psi \in K_0$ such that

$$\int_\Omega f(\nabla\psi(z))\beta(z)\,dz = \int f(y)\,d\alpha(y), \tag{44}$$

for each $f \in C(\mathbb{R}^d)$ such that $|f(y)| \leqq \mathrm{cst}(1 + \|y\|)$. Then $(\phi = \psi^*, \psi, p)$ is the unique solution to the MKP corresponding to α, where

$$\phi(y) = \psi^*(y) = \sup_{z \in \Omega} \{y\cdot z - \psi(z)\}, \qquad \forall y \in \mathbb{R}^d,$$

$$dp(y, z) = \delta(y - \nabla\psi(z))\beta(z)\,dz. \tag{45}$$

Before proving this result, let us first deduce Theorem 1.3 from Theorem 3.1 and Proposition 3.4. The first part of Theorem 1.3 directly follows from these results. Let us briefly prove the second part. Let α_n be a sequence of probability measures such that $\int f\,d\alpha_n \to \int f\,d\alpha$, for every $f \in C(\mathbb{R}^d)$ such that $|f(y)| \leqq \mathrm{cst}(1 + \|y\|)$. From the results on the MKP (Proposition 3.2), we know that the MKP corresponding to α_n has a unique solution $(\phi_n = \psi_n^*, \psi_n, p_n)$ that converges to the unique solution $(\phi = \psi^*, \psi, p)$ of the MKP corresponding to α. In particular, ψ_n converges to ψ in $W^{1,1}(\Omega, \beta)$, which achieves the proof.

Proof of Proposition 3.4: By assumption, ψ belongs to K_0, thus $\nabla\psi$ belongs to $L^1(\Omega, \beta; \mathbb{R}^d)$ and can be considered as a Borel mapping (up to a possible modification on a Lebesgue negligible set) from $\bar{\Omega}$ into $\mathbb{R}^d$ (recall that β has the same negligible sets as the Lebesgue measure and $\beta(\partial\Omega) = 0$). So p is well defined as a (tight) probability measure on $\mathbb{R}^d \times \bar{\Omega}$ by equation (45) and

$$\int f(y, z)\, dp(y, z) = \int f(\nabla\psi(z), z)\, d\beta(z)$$

holds for all $f \in C(\mathbb{R}^d \times \bar{\Omega})$ such that $|f(y, z)| \leqq \mathrm{cst}(1 + \|y\|)$. Indeed, since ψ belongs to K_0, $\int (1 + \|y\|)\, dp(y, z) = \int (1 + \|\nabla\psi(z)\|)\beta(z)\, dz < +\infty$. By assumption (44) and definition (45), the marginals of p precisely are α and β on $\mathbb{R}^d$ and $\bar{\Omega}$. Since $\psi \in K_0$, ψ is convex, locally Lipschitz continuous on Ω, and

$$\psi(z) + \nabla\psi(z)\cdot(\tilde{z} - z) \leqq \psi(\tilde{z}) \tag{46}$$

holds for any $\tilde{z} \in \Omega$ and every $z \in \Omega\backslash E$, where E is a Lebesgue negligible subset of Ω. By definition of p, there is a p-negligible subset F of $\mathbb{R}^d \times \bar{\Omega}$, such that $y = \nabla\psi(z)$ for any $(y, z) \in \mathbb{R}^d \times \bar{\Omega}\backslash F$. So we deduce from convexity property (46),

$$\psi(z) + y\cdot(\tilde{z} - z) \leqq \psi(\tilde{z}), \quad \forall \tilde{z} \in \Omega, \quad \forall (y, z) \in A, \tag{47}$$

where $A = (\mathbb{R}^d \times (\Omega\backslash E))\backslash F$ has p-measure 1, since

$$1 - p(A) \leqq p(\mathbb{R}^d \times (\partial\Omega \cup E)) + p(F) = \beta(\partial\Omega \cup E) = 0.$$

This shows that

$$\phi(y) = \psi^*(y) = \sup_{\tilde{z} \in \Omega} \{y\cdot\tilde{z} - \psi(\tilde{z})\}, \quad \forall y \in \mathbb{R}^d$$

satisfies $\phi(y) + \psi(z) \leqq y\cdot z$, for all $(y, z) \in A$, that is p-almost everywhere on $\mathbb{R}^d \times \bar{\Omega}$. Since the marginals of p are α and β on $\mathbb{R}^d$ and $\bar{\Omega}$, it follows that

$$\int \phi\, d\alpha + \int \psi\, d\beta \leqq \int y\cdot z\, dp(y, z).$$

By definition, ϕ satisfies $\phi(y) + \psi(z) \geqq y\cdot z$, for all $(y, z) \in \mathbb{R}^d \times \Omega$. So $(\phi = \psi^*, \psi, p)$ is a solution to the mixed MKP corresponding to α and, by Proposition 3.1, is the unique solution to the MKP.

This completes the proof of Theorem 1.3 and Theorem 1.1 which is just a corollary.

Proof of the polar factorization theorem:

Step 1 (existence of a polar factorization). Let $u \in L^1(X, \mu; \mathbb{R}^d)$. The fact that u satisfies the nondegeneracy condition (9) exactly means that the probability measure α defined by

$$\int f(y)\, d\alpha(y) = \int_X f(u(x))\, d\mu(x), \qquad \forall f \in C_c(\mathbb{R}^d),$$

is absolutely continuous with respect to the Lebesgue measure, $d\alpha(y) = \alpha(y)\, dy$, where α is a non-negative Lebesgue integrable function on $\mathbb{R}^d$. By using our results on the MKP, Proposition 3.1 in particular, we deduce that there is $\psi \in K_0$, a Lipschitz continuous convex function ϕ defined by

$$\phi(y) = \psi^*(y) = \sup_{z \in \Omega} \{y \cdot z - \psi(z)\}, \qquad \forall y \in \mathbb{R}^d,$$

such that

$$(48) \quad z = \nabla\phi(\nabla\psi(z)), \quad \beta - \text{a.e. } z \in \bar{\Omega}, \quad y = \nabla\psi(\nabla\phi(y)), \quad \alpha - \text{a.e. } y \in \mathbb{R}^d,$$

and a probability measure p defined on $\mathbb{R}^d \times \bar{\Omega}$ by

$$(49) \qquad dp(y, z) = \delta(y - \nabla\psi(z))\beta(z)\, dz = \delta(z - \nabla\phi(y))\alpha(y)\, dy.$$

Since u and $\nabla\phi$ can be considered as Borel mappings (up to a possible modification on a negligible set) respectively from (X, μ) into $\mathbb{R}^d$ and from $\mathbb{R}^d$ into itself,

$$s(x) = \nabla\phi(u(x)), \qquad x \in X,$$

defines a Borel mapping from (X, μ) into $\mathbb{R}^d$. This mapping is a measure-preserving mapping from (X, μ) into $(\bar{\Omega}, \beta)$. Indeed, for any $f \in C_c(\mathbb{R}^d)$, one has

$$\int f(s(x))\, d\mu(x) = \int f(\nabla\phi(u(x)))\, d\mu(x) \quad \text{(by definition of } s\text{)}$$

$$= \int f(\nabla\phi(y))\alpha(y)\, dy \quad \text{(by definition of } \alpha\text{)}$$

$$= \int f(z)\, dp(y, z) \quad \text{(by definition (49) of } p\text{)}$$

$$= \int f(z)\beta(z)\, dz \quad \text{(since } \beta \text{ is the marginal of } p \text{ on } \bar{\Omega}\text{)}$$

and this can be extended to any $f \in L^1(\Omega, \beta)$. Thus, s belongs to the class S of all measure-preserving mapping from (X, μ) into $(\bar{\Omega}, \beta)$. To prove the existence of the polar factorization of u, it is now enough to show that

$$(50) \qquad u(x) = \nabla\psi(s(x)), \qquad \mu - \text{a.e. } x \in X.$$

Let us consider the set

$$M = \{x \in X; u(x) \neq \nabla\psi(s(x))\}$$

and prove that it is μ-negligible. We have

$$M = \{x \in X; u(x) \neq \nabla\psi(\nabla\phi(u(x)))\} \quad \text{(by definition of } s\text{)}$$

$$= u^{-1}(\{y \in \mathbb{R}^d; y \neq \nabla\psi(\nabla\phi(y))\}).$$

Thus,

$$\mu(M) = \alpha(\{y \in \mathbb{R}^d; y \neq \nabla\psi(\nabla\phi(y))\}) \quad \text{(by definition of } \alpha\text{)}$$

$$= p(\{(y, z) \in \mathbb{R}^d \times \bar{\Omega}; y \neq \nabla\psi(\nabla\phi(y))\}) \quad (\alpha \text{ is the marginal of } p)$$

$$= 0 \quad \text{(by property (48)), which exactly is 50.}$$

This completes the proof of the existence part of the Polar Factorisation Theorem 1.2.

Step 2 (uniqueness of the polar factorization). Let us assume that there is a different way to write $u \in L^1(X, \mu; \mathbb{R}^d)\backslash N$ as $u = \nabla\psi' \circ s'$ where $s' \in S$, $\psi' \in K_0$ and show that, actually, $\psi' = \psi$, $s' = s = \nabla\phi \circ u$, μ-almost everywhere on X.

For any $f \in C(\mathbb{R}^d)$ such that $|f(y)| \leqq \text{cst.}(1 + \|y\|)$, we get

$$\int f(y)\alpha(y)\, dy = \int f(u(x))\, d\mu(x) \quad \text{(by definition of } \alpha\text{)}$$

$$= \int f(\nabla\psi'(s'(x)))\, d\mu(x) \quad \text{(by assumption)}$$

$$= \int f(\nabla\psi'(z))\beta(z)\, dz \quad \text{(since } s' \text{ is measure-preserving from } (X, \mu) \text{ into } (\bar{\Omega}, \beta)\text{)}.$$

Thus, it follows from Proposition 3.4 that $\psi' = \psi$.

Let us now show that $s' = \nabla\phi \circ u = s$, μ-almost everywhere on X. To do that, it is enough to show $s' = \nabla\phi \circ \nabla\psi \circ s'$ (indeed, by assumption, $\nabla\psi \circ s' = u$). This is clear, since

$$\mu(\{x \in X; s'(x) \neq \nabla\phi(\nabla\psi(s'(x)))\})$$

$$= \beta(\{z \in \bar{\Omega}; z \neq \nabla\phi(\nabla\psi(z))\}) \quad \text{(because } s \text{ is measure-preserving)}$$

$$= p(\{(y, z) \in \mathbb{R}^d \times \bar{\Omega}; z \neq \nabla\phi(\nabla\psi(z))\}) \quad (\beta \text{ is the marginal of } p))$$

$$= 0 \quad \text{(by property (48))}.$$

Step 3 (continuity of the polar factorization). From the rearrangement theorem, we already know that $u \to \psi$ is continuous from $L^1(X, \mu; \mathbb{R}^d)$ into $W^{1,1}(\Omega, \beta)$. Let us now show that $u \to s$ is continuous from $L^1(X, \mu; \mathbb{R}^d)\backslash N$ into $L^1(X, \mu; \mathbb{R}^d)$ by considering a sequence $(u_n = \nabla\psi_n \circ s_n)$ in $L^1(X, \mu; \mathbb{R}^d)\backslash N$ that converges to $u \in L^1(X, \mu; \mathbb{R}^d)\backslash N$ in $L^1(X, \mu; \mathbb{R}^d)$. By Proposition 3.4, (ψ_n^*, ψ_n, p_n) is the unique solution to the MKP corresponding to the probability measure α_n associated with u_n. By Proposition 3.2, we deduce that

$$\int f(\nabla\psi_n(z), z)\beta(z)\, dz \to \int f(\nabla\psi(z), z)\beta(z)\, dz$$

for any compactly supported continuous function f on $\mathbb{R}^d \times \bar{\Omega}$. Since s and each s_n are measure-preserving from (X, μ) into $(\bar{\Omega}, \beta)$ and $u_n = \nabla\psi_n \circ s_n$, $u = \nabla\psi \circ s$, this is equivalent to

$$\int f(u_n(x), s_n(x))\, d\mu(x) \to \int f(u(x), s(x))\, d\mu(x)$$

and implies (since u_n converges to u in $L^1(X, \mu; \mathbb{R}^d)$)

$$\int f(u(x), s_n(x))\, d\mu(x) \to \int f(u(x), s(x))\, d\mu(x).$$

This property can be extended by density to any function f of the form $f(y, z) = g(y)h(z)$, where $g \in L^1(\mathbb{R}^d, \alpha)$ and $h \in C(\bar{\Omega})$. Indeed, if g_ϵ is a smooth approximation to g and $f_\epsilon(y, z) = g_\epsilon(y)h(z)$, then

$$\left| \int f_\epsilon(u(x), s_n(x))\, d\mu(x) - \int f(u(x), s_n(x))\, d\mu(x) \right|$$

$$\leqq \sup|h| \int |g_\epsilon(u(x)) - g(u(x))|\, d\mu(x)$$

$$= \sup|h| \int |g_\epsilon(y) - g(y)|\, d\alpha(y) \to 0.$$

In particular, for $f(y, z) = \nabla\phi(y)\cdot z$, we get

$$\int \nabla\phi(u(x))\cdot s_n(x)\,d\mu(x) \to \int \nabla\phi(u(x))\cdot s(x)\,d\mu(x),$$

that is, since $s = \nabla\phi\circ u$, μ-almost everywhere on X,

$$\int s(x)\cdot s_n(x)\,d\mu(x) \to \int s(x)\cdot s(x)\,d\mu(x).$$

Since s_n and s are measure-preserving from (X, μ) into $(\bar{\Omega}, \beta)$,

$$\int \|s_n(x)\|^2\,d\mu(x) = \int \|z\|^2\beta(z)\,dz = \int \|s(x)\|^2\,d\mu(x),$$

and, thus,

$$\int \|s_n(x) - s(x)\|^2\,d\mu(x) \to 0,$$

which proves that s_n converges to s in $L^2(X, \mu; \mathbb{R}^d)$. Because Ω is contained in a ball $B(0, r)$, all measure-preserving mappings from (X, μ) into $(\bar{\Omega}, \beta)$ are μ-essentially bounded by the same constant r. Moreover, since (X, μ) is a probability space, the spaces $L^p(X, \mu; \mathbb{R}^d)$ are decreasingly embedded in $L^1(X, \mu; \mathbb{R}^d)$. It follows that s_n converges to s in all $L^p(X, \mu; \mathbb{R}^d)$, for $1 \leqq p < +\infty$.

Step 4 (characterization of the factors of the polar factorization). We already know, by the rearrangement theorem, the characterization of $\nabla\psi$ as the unique rearrangement of u in the class $\{\nabla\psi, \psi \in K_0\}$.

Let us now show that s is the unique maximizer in S of $\int s(x)\cdot u(x)\,d\mu(x)$. It is easy to see that s is a maximizer, by using a straightforward convexity argument. Indeed, since ψ is convex and locally Lipschitz continuous on Ω, we get, for every measure-preserving mapping s' from (X, μ) into $(\bar{\Omega}, \beta)$,

$$\psi(s'(x)) \geqq \psi(s(x)) + \nabla\psi(s(x))\cdot(s'(x) - s(x)), \qquad \mu - \text{a.e. } x \in X$$

(here the fact that the Lebesgue negligible subsets of Ω are mapped back by both s' and s into μ-negligible subsets of X is used). Thus, after integrating this inequality over X, we deduce

$$\int \psi\circ s'\,d\mu \geqq \int \psi\circ s\,d\mu + \int \nabla\psi\circ s\cdot(s' - s)\,d\mu.$$

Since both s' and s are measure-preserving from (X, μ) into $(\bar{\Omega}, \beta)$,

$$\int \psi\circ s'\,d\mu = \int \psi\circ s\,d\mu = \int \psi(z)\beta(z)\,dz,$$

which leads to

$$\int \nabla\psi \circ s \cdot (s' - s)\, d\mu \leqq 0,$$

that is $\int u \cdot (s' - s)\, d\mu \leqq 0$ since $u = \nabla\psi \circ s$. So s maximizes $\int s(x) \cdot u(x))\, d\mu(x)$.

Let us now show there is no other maximizer $s' \in S$. Let us introduce a probability measure p' on $\mathbb{R}^d \times \bar{\Omega}$ defined by

$$\int f(y, z)\, dp'(y, z) = \int f(u(x), s'(x))\, d\mu(x)$$

for any $f \in C(\mathbb{R}^d \times \bar{\Omega})$ such that $|f(y, z)| \leqq \mathrm{cst}(1 + \|y\|)$. We claim that p' is the unique solution to the dual MKP associated with α and therefore is equal to p defined by $dp(y, z) = \delta(y - \nabla\psi(z))\beta(z)\, dz$.

First, from the definition of p', we deduce that α and β are the marginals on $\mathbb{R}^d$ and $\bar{\Omega}$, since

$$\int f(y)\, dp'(y, z) = \int f(u(x))\, d\mu(x) = \int f(y)\, d\alpha(y)$$

(by definition of α) for each $f \in C(\mathbb{R}^d)$ such that $|f(y)| \leqq \mathrm{cst}(1 + \|y\|)$, and

$$\int f(z)\, dp'(y, z) = \int f(s'(x))\, d\mu(x) = \int f(z)\, d\beta(z)$$

(since s' is measure-preserving from (X, μ) into $(\bar{\Omega}, \beta)$) for each $f \in C(\bar{\Omega})$.

Then, we check that p' maximizes $\int y \cdot z\, dp'(y, z)$. Indeed

$$\int y \cdot z\, dp'(y, z) = \int u(x) \cdot s'(x)\, d\mu(x) \quad \text{(by definition of } p')$$

$$= \int u(x) \cdot s(x)\, d\mu(x) \quad \text{(since } s' \text{ is a maximizer, by assumption)}$$

$$= \int \nabla\psi(s(x)) \cdot s(x)\, d\mu(x) \quad \text{(because of the polar factorization)}$$

$$= \int \nabla\psi(z) \cdot z\beta(z) dz \quad \text{(since } s \text{ is measure-preserving)}$$

$$= \int y \cdot z\, dp(y, z) \quad \text{(by definition of } p\text{, the solution of the MKP).}$$

So p' is a maximizer and, therefore, $p' = p$, which shows that

$$\int f(u(x, s'(x))\, d\mu(x) = \int f(y, z)\, dp(y, z) = \int f(u(x, s(x))\, d\mu(x)$$

for each $f \in C(\mathbb{R}^d \times \bar{\Omega})$ such that $|f(y, z)| \leqq \text{cst}(1 + \|y\|)$. By using the same argument as in the third step of the proof, this equality also holds for $f(y, z) = \nabla\phi(y) \cdot z$, which shows

$$\int \nabla\phi(u(x)) \cdot s'(x)\, d\mu(x) = \int \nabla\phi(u(x)) \cdot s(x)\, d\mu(x),$$

that is $\int s(x) \cdot s'(x)\, d\mu(x) = \int s(x) \cdot s(x)\, d\mu(x)$, since we have $s = \nabla\phi \circ u$. Because s and s' are measure-preserving, we also have

$$\int \|s'(x)\|^2\, d\mu(x) = \int \|s(x)\|^2\, d\mu(x)$$

and $\int \|s(x) - s'(x)\|^2\, d\mu(x) = 0$ finally follows, which completes the proof.

This achieves the proof of the polar factorization theorem.

3.4. Existence of a Solution to the Mixed MKP

To prove that the mixed MKP has at least a solution (ϕ, ψ, p), we proceed in two steps:

(i) the case when α is compactly supported, in some ball $B(0, R)$;
(ii) the general case.

The Compact Case

From Rachev's paper (see [19]) or from classical results on convex analysis (see [13]), one gets:

PROPOSITION 3.5 (Strong Duality Principle). *There is a probability measure p on $B(0, R) \times \bar{\Omega}$, with marginals α and β that satisfies $\int y \cdot z\, dp(y, z) = I$, where*

$$I = \inf\left\{\int \phi\, d\alpha + \int \psi\, d\beta;\quad \phi \in C(B(0, R)),\quad \psi \in C(\bar{\Omega}) \right.$$
$$\left. \phi(y) + \psi(z) \geqq y \cdot z,\quad \forall (y, z) \in B(0, R) \times \bar{\Omega}\right\}.$$

Proof (sketch): Let us briefly sketch the proof by using Theorem 4.1 and Remark 4.2 in [13]. First set

$$V = C(B(0, R)) \times C(\bar{\Omega}), \qquad Y = C(B(0, R) \times \bar{\Omega}),$$

then define

$$\Lambda : V \to Y, \qquad \Lambda(\phi, \psi)(y, z) = \phi(y) + \psi(z), (y, z) \in B(0, R) \times \bar{\Omega};$$

$$F : V \to \mathbb{R}, \qquad F(\phi, \psi)(y, z) = \int \phi \, d\alpha + \int \psi \, d\beta;$$

$$G : Y \to \mathbb{R} \cup \{+\infty\},$$

$$G(\theta) = 0 \quad \text{if} \quad \theta(y, z) \geqq y \cdot z, \quad \forall (y, z); \quad +\infty \quad \text{otherwise.}$$

After checking that the conditions of Theorem 4.1 (in [13]) are satisfied, one gets

$$\inf_{(\phi,\psi) \in V} [F(\phi, \psi) + G(\Lambda(\phi, \psi))] = \max_{p \in Y^*} [-F^*(\Lambda^* p) - G^*(-p)],$$

which is exactly Proposition 3.5, since

$$Y^* = C(B(0, R) \times \bar{\Omega})',$$

$$G^*(-p) = -\int y \cdot z \, dp(y, z), \quad \text{if} \quad p \geqq 0, \quad +\infty \quad \text{otherwise,}$$

$F^*(\Lambda^* p) = 0$ if α and β are the marginals of p, $+\infty$ otherwise. This achieves the proof of Proposition 3.5.

In this duality result, it is not clear that the infimum is reached by some pair (ϕ, ψ), but if there is such an optimal pair, then it is clear that the mixed MKP has a solution.

Let us consider a minimizing sequence (ϕ_n, ψ_n), $\phi_n \in C(B(0, R))$, $\psi_n \in C(\bar{\Omega})$, that satisfies

$$\phi_n(y) + \psi_n(z) \geqq y \cdot z, \qquad \forall (y, z) \in B(0, R) \times \bar{\Omega}, \tag{51}$$

$$\int \phi_n \, d\alpha + \int \psi_n \, d\beta \to I. \tag{52}$$

It is not restricted to assume

$$\min_{B(0,R)} \phi_n = 0. \tag{53}$$

(Indeed, conditions (51) and (52) are unchanged when a constant is added to ϕ_n and subtracted from ψ_n.)

A natural regularization of such a minimizing sequence (often used for the general MKP; see [19]) is provided by

$$\tilde{\psi}_n(z) = \sup_{y \in B(0,R)} \{y \cdot z - \phi_n(y)\}, \qquad \forall z \in \mathbb{R}^d, \tag{54}$$

$$\tilde{\phi}_n(y) = \sup_{z \in \bar{\Omega}} \{y \cdot z - \tilde{\psi}_n(z)\}, \qquad \forall y \in \mathbb{R}^d. \tag{55}$$

The new sequence $(\tilde{\phi}_n, \tilde{\psi}_n)$ turns out to be (i) still a minimizing sequence, (ii) compact for the uniform convergence topology on $B(0, R) \times \bar{\Omega}$.

Let us first prove that $\tilde{\phi}_n$, $\tilde{\psi}_n$ is uniformly Lipschitz continuous on $\mathbb{R}^d$. From property (53) and definition (54), we get

$$\tilde{\psi}_n(0) = 0, \qquad \mathrm{Lip}(\tilde{\psi}_n) \leqq R \tag{56}$$

(since in definition (54), the supremum is taken over the ball $B(0, R)$). Now, from definition (55), we deduce

$$0 = -\tilde{\psi}_n(0) \leqq \tilde{\phi}_n(y) \leqq r\|y\| - \inf_{\Omega} \tilde{\psi}_n$$

(since Ω is contained in the ball $B(0, r)$) $\leqq r\|y\| + Rr$ (because of property (56)). So

$$0 \leqq \tilde{\phi}_n(y) \leqq r(\|y\| + R), \quad \forall y \in \mathbb{R}^d.$$

Moreover, $\tilde{\phi}_n$ is Lipschitz continuous on $\mathbb{R}^d$ and $\mathrm{Lip}(\tilde{\phi}_n) \leqq r$. By Ascoli's theorem, there is a pair (ϕ, ψ) of Lipschitz continuous functions such that $\phi_n \to \phi$, $\psi_n \to \psi$, uniformly on $B(0, R)$ and $\bar{\Omega}$. Since β and α are compactly supported, we have

$$\psi \in C(\Omega) \cap L^1(\Omega, \beta), \qquad \phi \in C(\mathbb{R}^d) \cap L^1(\mathbb{R}^d, \alpha).$$

From definition (55), we easily get

$$\phi(y) + \psi(z) \geqq y \cdot z, \qquad \forall (y, z) \in \mathbb{R}^d \times \bar{\Omega}.$$

Moreover, from definitions (54) and (55), it follows that

$$\tilde{\phi}_n(y) \leqq \phi_n(y), \qquad \forall y \in B(0, R),$$

$$\tilde{\psi}_n(z) \leqq \psi_n(z), \qquad \forall z \in \bar{\Omega}.$$

Thus $(\tilde{\phi}_n, \tilde{\psi}_n)$ still is a minimizing sequence and

$$\int \phi \, d\alpha + \int \psi \, d\beta \leqq I.$$

By adding a suitable constant to ϕ and subtracting the same constant from ψ, it is not restrictive to set $\int \psi \, d\beta = 0$. Finally, (ϕ, ψ, p) solves the mixed MKP which completes the existence proof in the compact case.

The Noncompact Case

The existence of a solution to the mixed MKP directly follows from Proposition 3.2. Indeed, it is possible to approximate α by a sequence (α_n) of compactly supported probability measured on $\mathbb{R}^d$, (for which we just proved the existence of a solution to the mixed MKP) defined, for n large enough, by

$$\int f \, d\alpha_n = C_n \int_{\|y\| \leqq n} f(y) \, d\alpha, \qquad \forall f \in C_c(\mathbb{R}^d),$$

where $C_n = \alpha(B(0, n))^{-1}$. It is elementary to check that

$$\int f \, d\alpha_n \to \int f \, d\alpha$$

for all $f \in C(\mathbb{R}^d)$ such that $|f(y)| \leqq \text{cst}(1 + \|y\|)$. This allows us to use Proposition 3.2 that asserts the existence of a solution to the mixed MKP corresponding to α.

This completes the proofs of our main results.

Acknowledgment. This research was partly done during the author's visit to the Department of Mathematics, UCLA, Los Angeles, March 1989.

Bibliography

[1] Abraham, R., Marsden, J. E., and Ratiu, T., *Manifolds, Tensor Analysis, and Applications*, Reading, MA, Addison-Wesley, 1983.

[2] Alvino, A., Lions, P.-L., and Trombetti, G., *On optimization problems with prescribed rearrangements*, preprint.

[3] Appell, P., *Le problème géometrique de déblais et des remblais*, Mémoriale de l'Académie des Sciences Mathématiques, Fascicule XXVII, Paris, 1928.

[4] Aubin, J.-P., *Mathematical Methods of Game and Economic Theory*, North-Holland, Amsterdam, 1979.

[5] Bandle, C., *Isoperimetric Inequalities*, Pitman, London, 1980.

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

[7] Brenier, Y., *A combinatorial algorithm for the Euler equation of incompressible flows*, 8th International Conf. on Computing Methods, Versailles, 1987.

[8] Brezis, H., *Opérateurs Maximaux Monotones*, North-Holland, Amsterdam, 1973.

[9] Caffarelli, L., Nirenberg, L., and Spruck, J., *The Dirichlet problem for nonlinear second-order elliptic equations I: Monge-Ampère equation*, Comm. Pure Appl. Math. 37, 1984, pp. 369–402.

[10] Cullen, M. J. P., and Purser, R. J., *An extended Lagrangian model of semi-geostrophic frontogenesis*, J. Atmos. Sci. 41, 1984, pp. 1477–1497.

[11] Cullen, M. J. P., Norbury, J., Purser, R. J., and Shutts, G. J., *Modelling the quasi-equilibrium dynamics of the atmosphere*, QJR Meteorol. Soc. 113, 1987, pp. 735–757.

[12] Dautray, R., and Lions, J.-L., *Analyse mathématique et calcul numérique pour les sciences et les techniques*, Masson, Paris, 1985.

[13] Ekeland, I., and Temam, R., *Analyse Convexe et Problèmes Varitionnels*, Dunod, Paris, 1974.

[14] Hardy, G. H., Littlewood, J. E., and Polya, G., *Inequalities*, Cambridge Univ. Press, 1952.

[15] Kantorovich, L. V., *On the transfer of masses*, Dokl. Akad. Nauk SSSR 37, 1942, pp. 227–229.

[16] Knott, M., and Smith, C. S., *On the optimal mapping of distributions*, J. Opt. Th. Appl. 43, 1984, pp. 39–49.

[17] Monge, G., *Mémoire sur la théorie des déblais et de remblais*, Mémoires de l'Académie des Sciences, 1781.

[18] Mossino, J., *Inégalités Isopérimétriques et Applications en Physique*, Hermann, Paris, 1985.

[19] Rachev, S. T., *The Monge-Kantorovich mass transference problem*, Th. Prob. Appl. 49, 1985, pp. 647–676.

[20] Rachev, S. T., *Recent results in the theory of probability metrics*, to appear.

[21] Royden, H. L., *Real Analysis*, MacMillan, New York, 1963.

[22] Ryff, J. V., *Orbits of L^1 functions under doubly stochastic transformations*, Trans. AMS 117, 1965, pp. 92–100.

Received January 1990.