



KANTOROVICH'S MASS TRANSPORT PROBLEM FOR CAPACITIES

Sorin G. GAL¹, Constantin P. NICULESCU²

¹University of Oradea, Department of Mathematics and Computer Science, University Street 1, Oradea, 410087, Romania

²University of Craiova, Department of Mathematics, A. I. Cuza Street 13, Craiova, 200585, Romania

Corresponding author: Constantin P. NICULESCU, E-mail: cpniculescu@gmail.com

Abstract: The aim of the present paper is to extend Kantorovich's mass transport problem to the framework of upper/lower continuous capacities and to prove the cyclic monotonicity of the supports of optimal supermodular plans. As in the probabilistic case, this easily yields the corresponding extension of the Kantorovich duality.

Key words: Choquet integral, comonotonic additivity, Monge-Kantorovich transportation problem, upper/lower continuous capacity, random variable.

1. INTRODUCTION

Kantorovich's mass transport problem was formulated in 1942, when Kantorovich [14] published a note containing the following explicit description of it in probabilistic terms. Suppose that X and Y are two compact metric spaces and $c : X \times Y \rightarrow [0, \infty)$ is a Borel-measurable map referred to as a *cost* function. Given two Borel probability measures μ and ν , defined respectively on the Borel σ -algebras $\mathcal{B}(X)$ and $\mathcal{B}(Y)$, a Borel probability measure π on $\mathcal{B}(X \times Y) = \mathcal{B}(X) \otimes \mathcal{B}(Y)$ is called a *transport plan* for μ and ν if μ is the projection of π on X and ν is the projection of π on Y , that is,

$$\pi(A \times Y) = \mu(A) \text{ for all } A \in \mathcal{B}(X)$$

and

$$\pi(X \times B) = \nu(B) \text{ for all } B \in \mathcal{B}(Y).$$

Under the above conditions one said that μ and ν are the *marginals* of π . The set $\Pi(\mu, \nu)$, of all transport plans for μ and ν , is always nonempty (it contains at least the product measure $\mu \otimes \nu$) and also convex and weak star compact as a subset of $C(X \times Y)^*$ (the dual of the Banach space $C(X \times Y)$ of all continuous functions $f : X \times Y \rightarrow \mathbb{R}$). This is a combination of the Banach-Alaoglu theorem and the Riesz representation theorem. Based on this fact, Kantorovich has noticed that the functional

$$\text{Cost}(\pi) = \int_{X \times Y} c(x, y) d\pi(x, y), \quad \pi \in \Pi(\mu, \nu), \quad (1)$$

attains its infimum in the case of continuous cost functions. The plans π at which the infimum is attained are called *optimal transport plans*.

Kantorovich's problem is a relaxation of the Monge problem on "excavation and embankments" (see [16]), which refers to the minimization of the *transference cost*,

$$\int_X c(x, T(x)) d\mu,$$

over all Borel measurable mappings $T : X \rightarrow Y$ that *push forward* μ to ν (that is, $\nu(A) = \mu(T^{-1}(A))$ for every $A \in \mathcal{B}(Y)$). The transport plan associated to such a mapping T is

$$\pi_T(B) = \mu(\{x : (x, T(x)) \in B\}) \quad \text{for every } B \in \mathcal{B}(X \times Y).$$

As a consequence, any optimal transport plan can be viewed as a generalized solution for Monge's problem.

Kantorovich [15] realized this connection in 1948 and since then one speaks on the Monge-Kantorovich problem, a fusion of the two problems into a vast subject with deep applications in economics, dynamical systems, probability and statistics, information theory etc. Details are covered by a number of excellent surveys and fine books published by Ambrosio [2], Evans [9], Galichon [11], Gangbo and McCann [12], Rachev [18], Rachev and Rüschendorf [19], Santambrogio [21] and Villani [23, 24], just to cite a few.

The aim of the present paper is to extend Kantorovich's mass transport problem to the framework of upper/lower continuous capacities and to prove the cyclical monotonicity of the supports of optimal supermodular plans. As in the probabilistic case, this easily yields the corresponding extension of the Kantorovich duality.

The concept of capacity (a kind of monotone set function not necessarily additive) and the integral associated to it were introduced by Choquet [4, 5] in the early 1950s, motivated by some problems in potential theory. Nowadays they also become powerful tools in decision making under risk and uncertainty, game theory, ergodic theory, pattern recognition, interpolation theory etc. See Adams [1], Denneberg [6], Föllmer and Schied [10], Wang and Klir [25] and Wang and Yan [26], as well as the references therein.

Our paper is, to the best of our knowledge, the first to investigate up to what extent the mass transport theory extends to a context marked by uncertainty and incomplete knowledge. The necessary background on capacities and Choquet integral makes the subject of Section 2. The main result of Section 3 is Theorem 2, which asserts the existence of optimal transport plans. The critical ingredient in the proof is a nonlinear version of the Riesz representation theory, due to Epstein and Wang (see Theorem 1). The fact that the support of any supermodular optimal transport plan is necessarily a c -cyclically monotone set is proved in Theorem 4, Section 4. Actually, for a given pair of marginals, there is a c -cyclically monotone set including the supports of all supermodular optimal transport plans. See Corollary 2. The paper ends by noticing that this fact together with a previous result due to Rüschendorf [20] and Smith and Knott [22], easily yield the extension of the Kantorovich duality in the framework of capacities.

2. PRELIMINARIES ON CAPACITIES AND CHOQUET INTEGRAL

For the convenience of the reader we will briefly recall some basic facts about capacities and Choquet integral.

Let (X, \mathcal{A}) be an arbitrarily fixed measurable space, consisting of a nonempty abstract set X and a σ -algebra \mathcal{A} of subsets of X .

Definition 1. A set function $\mu : \mathcal{A} \rightarrow [0, 1]$ is called a capacity if it verifies the following two conditions:

- (a) $\mu(\emptyset) = 0$ and $\mu(X) = 1$;
- (b) $\mu(A) \leq \mu(B)$ for all $A, B \in \mathcal{A}$, with $A \subset B$.

In applications some additional properties are useful.

A capacity μ is called upper continuous (or continuous by descending sequences) if

$$\lim_{n \rightarrow \infty} \mu(A_n) = \mu\left(\bigcap_{n=1}^{\infty} A_n\right)$$

for every nonincreasing sequence $(A_n)_n$ of sets in \mathcal{A} ; μ is called lower continuous (or continuous by ascending sequences), if $\lim_{n \rightarrow \infty} \mu(A_n) = \mu(\cup_{n=1}^{\infty} A_n)$ for every nondecreasing sequence $(A_n)_n$ of sets in \mathcal{A} .

The upper/lower continuity of a capacity is a generalization of countable additivity of an additive measure. Indeed, if μ is an additive capacity, then upper/lower continuity is the same with countable additivity.

A capacity μ is called *supermodular* if

$$\mu(A) + \mu(B) \leq \mu(A \cup B) + \mu(A \cap B) \text{ for all } A, B \in \mathcal{A};$$

μ is called *submodular* if the last inequality works in the opposite direction.

A simple way to construct nontrivial examples of upper continuous supermodular capacities is to start with a probability measure $P : \mathcal{A} \rightarrow [0, 1]$ and to consider any nondecreasing, convex and continuous function $u : [0, 1] \rightarrow [0, 1]$ such that $u(0) = 0$ and $u(1) = 1$; for example, one may chose $u(t) = t^\alpha$ with $\alpha > 1$. Then the *distorted probability* $\mu = u(P)$ is an upper continuous supermodular capacity on the σ -algebra \mathcal{A} .

Suppose that (X, \mathcal{A}) and (Y, \mathcal{B}) are two measurable spaces. Any (upper continuous) capacity $\mu : \mathcal{A} \rightarrow [0, 1]$ and any measurable mapping $T : X \rightarrow Y$ induce a (upper continuous) capacity $T\#\mu$ called the *push-forward* of μ through T and defined by the formula

$$(T\#\mu)(B) = \mu(T^{-1}(B)) \text{ for all } B \in \mathcal{B}.$$

The main feature of this kind of capacities is the following change of variables formula

$$(C) \int_Y g(y) d(T\#\mu) = (C) \int_X g(T(x)) d\mu \quad (2)$$

which works for all nonnegative bounded random variables $g : Y \rightarrow \mathbb{R}$. The concept of integrability with respect to a capacity refers to the whole class of random variables, that is, to all functions $f : X \rightarrow \mathbb{R}$ verifying the condition of \mathcal{A} -measurability ($f^{-1}(A) \in \mathcal{A}$ for every set $A \in \mathcal{B}(\mathbb{R})$).

Definition 2. The Choquet integral of a random variable f with respect to the capacity μ is defined as the sum of two Riemann improper integrals,

$$(C) \int_X f d\mu = \int_0^{+\infty} \mu(\{x \in X : f(x) \geq t\}) dt + \int_{-\infty}^0 [\mu(\{x \in X : f(x) \geq t\}) - 1] dt,$$

Accordingly, f is said to be Choquet integrable if both integrals above are finite.

If $f \geq 0$, then the last integral in the formula appearing in Definition 2 is 0.

The inequality sign \geq in the above two integrands can be replaced by $>$; see [25], Theorem 11.1, p. 226.

Every bounded random variable is Choquet integrable. The Choquet integral coincides with the Lebesgue integral when the underlying set function μ is a σ -additive measure.

The Choquet integral of a function $f : X \rightarrow \mathbb{R}$ over a set $A \in \mathcal{A}$ is defined via the formula

$$(C) \int_A f d\mu = \mu(A) \cdot (C) \int_X f d\mu_A,$$

where μ_A is the capacity defined by $\mu_A(B) = \mu(B \cap A) / \mu(A)$ for all $B \in \mathcal{A}$.

We next summarize some basic properties of the Choquet integral.

Remark 1. (a) If $\mu : \mathcal{A} \rightarrow [0, 1]$ is a capacity and $A \in \mathcal{A}$, then the associated Choquet integral is a functional on the space of all bounded random variables defined on A such that:

$$\begin{aligned} f \geq 0 & \text{ implies } (C) \int_A f d\mu \geq 0 \quad (\text{positivity}) \\ f \leq g & \text{ implies } (C) \int_A f d\mu \leq (C) \int_A g d\mu \quad (\text{monotonicity}) \\ (C) \int_A a f d\mu & = a \cdot \left((C) \int_A f d\mu \right) \text{ for } a \geq 0 \quad (\text{positive homogeneity}) \\ (C) \int_A 1 \cdot d\mu & = \mu(A) \quad (\text{calibration}); \end{aligned}$$

see [6], p. 64, Proposition 5.1 (ii), for the proof of positive homogeneity.

(b) In general, the Choquet integral is not additive but, if the bounded random variables f and g are comonotonic (that is, $(f(\omega) - f(\omega')) \cdot (g(\omega) - g(\omega')) \geq 0$, for all $\omega, \omega' \in A$), then

$$(C) \int_A (f + g) d\mu = (C) \int_A f d\mu + (C) \int_A g d\mu.$$

This is usually referred to as the property of comonotonic additivity. An immediate consequence is the property of translation invariance,

$$(C) \int_A (f + c) d\mu = (C) \int_A f d\mu + c \cdot \mu(A)$$

for all $c \in \mathbb{R}$ and all bounded random variables f . See [6], Proposition 5.1 (vi), p. 65.

(c) If μ is an upper continuous capacity, then the Choquet integral is upper continuous in the sense that

$$\lim_{n \rightarrow \infty} \left((C) \int_A f_n d\mu \right) = (C) \int_A f d\mu,$$

whenever $(f_n)_n$ is a nonincreasing sequence of bounded random variables that converges pointwise to the bounded variable f . This is a consequence of the Beppo Levi monotone convergence theorem from the theory of Lebesgue integral (see [7], Theorem 2, p. 133).

(d) If μ is a supermodular capacity, then the associated Choquet integral is a superadditive functional, that is

$$(C) \int_A f d\mu + (C) \int_A g d\mu \leq (C) \int_A (f + g) d\mu$$

for all bounded random variables f and g . It is also a supermodular functional in the sense that

$$(C) \int_A \sup\{f, g\} d\mu + (C) \int_A \inf\{f, g\} d\mu \geq (C) \int_A f d\mu + (C) \int_A g d\mu$$

for all bounded random variables f and g .

In this paper we are interested in a special kind of measurable spaces, those of the form $(X, \mathcal{B}(X))$, where X is a compact metric space and $\mathcal{B}(X)$ is the σ -algebra of all Borel subsets of X . We will denote by $\text{Ch}(X)$ the class of all upper continuous capacities $\mu : \mathcal{B}(X) \rightarrow [0, 1]$.

The following analogue of the Riesz representation theorem is due to L. G. Epstein and T. Wang. See [8], Theorem 4.2. See also Zhou [27], Theorem 1 and Lemma 3, for a simple (and more general) argument.

THEOREM 1. *Suppose that $I : C(X) \rightarrow \mathbb{R}$ is a comonotonically additive and monotone functional and $I(1) = 1$. Then it is also upper continuous and there exists a unique upper continuous capacity $\mu : \mathcal{B}(X) \rightarrow [0, 1]$ such that I coincides with the Choquet integral associated to it.*

On the other hand, according to Remark 1, the Choquet integral associated to any upper continuous capacity is a comonotonically additive, monotone and upper continuous functional.

This result allows us to identify $\text{Ch}(X)$ with the set of all functionals I on $C(X)$ which are comonotonically additive, monotone and verify $I(1) = 1$. An immediate consequence is as follows:

COROLLARY 1 (O'Brien, W. Vervaat [17]). *$\text{Ch}(X)$ is compact and metrizable with respect to the weak topology on $\text{Ch}(X)$ induced by the duality*

$$\langle \cdot, \cdot \rangle : C(X) \times \text{Ch}(X) \rightarrow \mathbb{R}, \quad \langle f, \mu \rangle = (C) \int_X f d\mu.$$

A direct argument for Corollary 1 makes the objective of Theorem 2 in [27], were it is noticed that the weak

convergence on $\text{Ch}(X)$ is equivalent with the convergence with respect to the metric

$$d_{\text{Ch}(X)}(\mu, \nu) = \sum_{j=1}^{\infty} \frac{1}{2^j} \left| (C) \int_X f_j d\mu - (C) \int_X f_j d\nu \right|,$$

associated to an arbitrary sequence $(f_j)_j$ dense in the unit sphere of $C(X)$.

Remark 2. Under the assumptions of Theorem 1, the functional I is supermodular if and only if the capacity μ is supermodular. For details see [3], Theorem 13(c). Notice also that the subset $\text{Ch}^*(X)$ of all supermodular capacities $\mu \in \text{Ch}(X)$ is closed with respect to the weak topology (and thus it is compact, according to Corollary 1).

3. THE EXISTENCE OF OPTIMAL TRANSPORT PLAN

The framework used in this section parallels that of Borel probability measures, but the details are based on the integral representation of comonotonically additive and monotone functionals provided by Theorem 1.

Let X and Y be two compact metric spaces on which there are given the upper continuous capacities $\mu \in \text{Ch}(X)$ and respectively $\nu \in \text{Ch}(Y)$. A *transport plan* for μ and ν is any capacity $\pi \in \text{Ch}(X \times Y)$ with marginals $\mu = \text{pr}_X \# \pi$ and $\nu = \text{pr}_Y \# \pi$, that is, such that

$$\pi(A \times Y) = \mu(A) \text{ for all } A \in \mathcal{B}(X)$$

and

$$\pi(X \times B) = \nu(B) \text{ for all } B \in \mathcal{B}(Y).$$

The set of all such transport plans will be denoted $\Pi_{\text{Ch}}(\mu, \nu)$. $\Pi_{\text{Ch}}(\mu, \nu)$ is a nonempty set. For example, if P and Q are probability measures, then $(P \otimes Q)^\alpha$ is a transport plan for the distorted probabilities $\mu = P^\alpha$ and $\nu = Q^\alpha$, whenever $\alpha \geq 1$.

Given a Borel measurable cost function $c : X \times Y \rightarrow [0, \infty)$, the cost of a transport plan $\pi \in \Pi_{\text{Ch}}(\mu, \nu)$ is defined by a formula similar to formula (1):

$$\text{Cost}(\pi) = (C) \int_{X \times Y} c(x, y) d\pi(x, y).$$

The existence of the optimal transport plans is motivated by the following result:

THEOREM 2. *If the cost function $c : X \times Y \rightarrow [0, \infty)$ is continuous, then there exists $\pi_0 \in \Pi_{\text{Ch}}(\mu, \nu)$ such that*

$$\text{Cost}(\pi_0) = \inf_{\pi \in \Pi_{\text{Ch}}(\mu, \nu)} \text{Cost}(\pi).$$

Proof. We will denote by $\pi \rightarrow I_\pi$ the bijection stated by Theorem 1, that makes possible to identify the set $\text{Ch}(X \times Y)$, of all upper continuous capacities on $X \times Y$, with the set Φ , of all comonotonically additive and monotone functionals I on $C(X \times Y)$ that verify $I(1) = 1$. By this bijection, the set $\Pi_{\text{Ch}}(\mu, \nu)$ corresponds to the subset $\Phi(\mu, \nu)$ of Φ , consisting of those functionals $I : C(X \times Y) \rightarrow \mathbb{R}$ such that

$$I(\bar{u}) = I_\mu(u) \text{ for all } u \in C(X) \tag{3}$$

and

$$I(\bar{w}) = I_\nu(w) \text{ for all } w \in C(Y), \tag{4}$$

where \bar{u} and respectively \bar{w} represent the extensions of u and w to $X \times Y$ via the formulas

$$\bar{u}(x, y) = u(x) \text{ and } \bar{w}(x, y) = w(y) \text{ for all } (x, y) \in X \times Y, \tag{5}$$

and $I_\mu : C(X) \rightarrow \mathbb{R}$ and $I_\nu : C(Y) \rightarrow \mathbb{R}$ are the unique comonotonically additive and monotone functionals generated by μ and ν via the Choquet integral. Indeed, assuming that $I = I_\pi$, since μ and ν are the marginals of π , we have

$$\pi(\{(x, y) \in X \times Y : \bar{u}(x, y) \geq \alpha\}) = \mu(\{x \in X : u(x) \geq \alpha\}),$$

and

$$\pi(\{(x, y) \in X \times Y : \bar{w}(x, y) \geq \alpha\}) = \nu(\{y \in Y : w(y) \geq \alpha\}),$$

for all $\alpha \in \mathbb{R}$, whence, by the definition of the Choquet integral, we infer that

$$I_\pi(\bar{u}) = I_\mu(u) \text{ and } I_\pi(\bar{w}) = I_\nu(w).$$

We will use the above remark to prove that $\Phi(\mu, \nu)$ is a closed subset of Φ (and thus compact, according to Corollary 1). Since the topology of Φ is metrizable, this reduces to the fact that $\Phi(\mu, \nu)$ is closed under the operation of taking countable limits. For this, let $(\gamma_n)_n$ be a sequence of elements of $\Phi(\mu, \nu)$ converging to a capacity $\gamma \in \Phi$ and put $I_n = I_{\gamma_n}$ and $I = I_\gamma$ in order to simplify the notation. By our assumptions, the functionals I_n and I are comonotonically additive and monotone and

$$I_n(f) \rightarrow I(f) \quad \text{for all } f \in C(X \times Y). \quad (6)$$

The membership of I_n to $\Phi(\mu, \nu)$ translates into the formulas

$$I_n(\bar{u}) = I_\mu(u) \quad \text{for all } u \in C(X) \quad (7)$$

and

$$I_n(\bar{w}) = I_\nu(w) \quad \text{for all } w \in C(Y), \quad (8)$$

where \bar{u} and respectively \bar{w} represent the extensions of u and w to $X \times Y$ via the formulas (5). Combining this fact with (6), one easily conclude that I verifies similar formulas and thus the capacity γ that generates I belongs to $\Phi(\mu, \nu)$. In order to end the proof let's choose a sequence $(\pi_n)_n$ of elements of $\Pi_{\text{Ch}}(\mu, \nu)$ which minimizes the cost function, that is, such that

$$\lim_{n \rightarrow \infty} \text{Cost}(\pi_n) = m = \inf_{\gamma \in \Pi_{\text{Ch}}(\mu, \nu)} \text{Cost}(\gamma).$$

Since $\Pi_{\text{Ch}}(\mu, \nu)$ is a compact set, we may assume (by passing to a subsequence if necessary) that $(\pi_n)_n$ converges to some $\pi \in \Pi_{\text{Ch}}(\mu, \nu)$. Then $\text{Cost}(\pi) = m$, which means that π is an optimal transport plan.

Remark 3. It is a simple exercise to prove that if X is a compact metric space, then every functional I on $C(X)$ which is comonotonically additive and monotonic is also lower continuous. Combining this fact with Proposition 17 (i)&(ii) and Corollary 18 in [3], one can easily obtain the validity of the "lower" version of Theorem 1.

As concerns the case of supermodular capacities, let us notice first that the existence of a supermodular transport plan $\pi \in \Pi_{\text{Ch}}(\mu, \nu)$ imposes that both μ and ν are supermodular. See [13], Lemma 4, p. 281. An inspection of the argument of Theorem 2 easily yields the following result:

THEOREM 3. *If the cost function $c : X \times Y \rightarrow [0, \infty)$ is continuous and μ and ν are two supermodular upper continuous capacities such that*

$$\Pi_{\text{Ch}^*}(\mu, \nu) = \{\pi \in \Pi_{\text{Ch}}(\mu, \nu) : \pi \text{ supermodular}\}$$

is nonempty, then there exists $\pi_0 \in \Pi_{\text{Ch}^}(\mu, \nu)$ such that*

$$\text{Cost}(\pi_0) = \inf_{\pi \in \Pi_{\text{Ch}^*}(\mu, \nu)} \text{Cost}(\pi).$$

4. A NECESSARY CONDITION FOR THE OPTIMALITY OF A TRANSPORT PLAN

The aim of this section is to prove that optimal transport plans have c -cyclically monotone supports, a fact that relates them to the theory of c -concave functions. For this, we need some preparation.

As above, X and Y are compact metric spaces and $c : X \times Y \rightarrow [0, \infty)$ is a cost function.

Definition 3. A subset $S \subset X \times Y$ is called c -cyclically monotone if for every finite number of points $(x_i, y_i) \in S$, $i = 1, \dots, n$, and any permutation σ of $\{1, \dots, n\}$, we have

$$\sum_{i=1}^n c(x_i, y_i) \leq \sum_{i=1}^n c(x_{\sigma(i)}, y_i).$$

By definition, the support of an upper continuous capacity $\gamma : \mathcal{B}(X \times Y) \rightarrow [0, 1]$ is the set $\text{supp}(\gamma)$ of all points (x, y) in $X \times Y$ for which every open neighborhood \mathcal{W} verifies $\gamma(\mathcal{W}) > 0$. The support is a closed set since its complement is the union of the open sets of capacity zero.

We are now in a position to prove the following result.

THEOREM 4. *If the cost function $c : X \times Y \rightarrow [0, \infty)$ is continuous, then every supermodular optimal transport plan $\pi \in \Pi_{\text{Ch}}(\mu, \nu)$ has a c -cyclically monotone support.*

Proof. Our argument is close to that used by Ambrosio [2], Theorem 2.2, in the probabilistic framework. If $\text{supp}(\pi)$ were not c -cyclically monotone, then there would exist points $(\bar{x}_1, \bar{y}_1), \dots, (\bar{x}_n, \bar{y}_n)$ in $\text{supp}(\pi)$ and a permutation σ of $\{1, \dots, n\}$ such that

$$\sum_{i=1}^n c(\bar{x}_i, \bar{y}_i) > \sum_{i=1}^n c(\bar{x}_{\sigma(i)}, \bar{y}_i).$$

Choose ε such that

$$0 < \varepsilon < \frac{1}{2n} \left(\sum_{i=1}^n c(\bar{x}_i, \bar{y}_i) - \sum_{i=1}^n c(\bar{x}_{\sigma(i)}, \bar{y}_i) \right).$$

Since c is a continuous function, there exists compact neighborhoods U_i of \bar{x}_i and V_i of \bar{y}_i such that $c(x_i, y_i) > c(\bar{x}_i, \bar{y}_i) - \varepsilon$ for all $(x_i, y_i) \in U_i \times V_i$ and $c(x_i, y_i) < c(\bar{x}_{\sigma(i)}, \bar{y}_i) + \varepsilon$ for all $(x_i, y_i) \in U_{\sigma(i)} \times V_i$. We have $\pi(U_i \times V_i) > 0$, due to the fact that $(x_i, y_i) \in \text{supp}(\pi)$. Then $\alpha = (1/n) \min_i \pi(U_i \times V_i) > 0$ and we can consider the upper continuous capacities π_i defined by

$$\pi_i(\mathcal{W}) = \frac{\pi(\mathcal{W} \cap (U_i \times V_i))}{\pi(U_i \times V_i)} \quad \text{for every } \mathcal{W} \in \mathcal{B}(X \times Y).$$

The marginals of π_i are $\mu_i = \text{pr}_X \# \pi_i$ and $\nu_i = \text{pr}_Y \# \pi_i$, where pr_X and pr_Y are the canonical projections of $X \times Y$ respectively on X and Y . The set function

$$\gamma = \pi - \alpha \sum_{i=1}^n \pi_i + \alpha \sum_{i=1}^n \mu_i \otimes \nu_i$$

is nonnegative since $\pi - \alpha \sum_{i=1}^n \pi_i \geq 0$, according to the choice of α . Moreover, $\gamma(\emptyset) = 0$ and $\gamma(X \times Y) = 1$. In order to prove the monotonicity of γ , let's consider two Borel subsets \mathcal{V} and \mathcal{W} of $X \times Y$ such that $\mathcal{V} \subset \mathcal{W}$. The inequality $\gamma(\mathcal{V}) \leq \gamma(\mathcal{W})$ is a consequence of the fact that

$$\pi(\mathcal{W}) - \pi(\mathcal{V}) \geq \alpha \sum_{i=1}^n \frac{\pi(\mathcal{W} \cap (U_i \times V_i)) - \pi(\mathcal{V} \cap (U_i \times V_i))}{\pi(U_i \times V_i)}.$$

For this it suffices to show that

$$\pi(\mathcal{W}) - \pi(\mathcal{V}) \geq \pi(\mathcal{W} \cap (U_i \times V_i)) - \pi(\mathcal{V} \cap (U_i \times V_i)),$$

for all indices i . We claim that actually

$$\pi(\mathcal{W}) + \pi(\mathcal{V} \cap \mathcal{H}) \geq \pi(\mathcal{V}) + \pi(\mathcal{W} \cap \mathcal{H})$$

for every Borel subset $\mathcal{H} \subset X \times Y$. Indeed,

$$\chi_{\mathcal{W}} + \chi_{\mathcal{V} \cap \mathcal{H}} \geq \chi_{\mathcal{V}} + \chi_{\mathcal{W} \cap \mathcal{H}}$$

and because $\chi_{\mathcal{W}}$ and $\chi_{\mathcal{V} \cap \mathcal{H}}$ are comonotonic and π is monotone and supermodular we have

$$\begin{aligned} \pi(\mathcal{W}) + \pi(\mathcal{V} \cap \mathcal{H}) &= (C) \int_{X \times Y} \chi_{\mathcal{W}} d\pi + (C) \int_{X \times Y} \chi_{\mathcal{V} \cap \mathcal{H}} d\pi \\ &= (C) \int_{X \times Y} (\chi_{\mathcal{W}} + \chi_{\mathcal{V} \cap \mathcal{H}}) d\pi \geq (C) \int_{X \times Y} (\chi_{\mathcal{V}} + \chi_{\mathcal{W} \cap \mathcal{H}}) d\pi \\ &= (C) \int_{X \times Y} \chi_{\mathcal{V}} d\pi + (C) \int_{X \times Y} \chi_{\mathcal{W} \cap \mathcal{H}} d\pi \\ &= \pi(\mathcal{V}) + \pi(\mathcal{W} \cap \mathcal{H}). \end{aligned}$$

This ends the proof of the monotonicity of γ . Clearly, γ is upper continuous and its marginals are μ and ν . The cost of the transport plan γ is less than the cost of π because

$$\begin{aligned} &\text{Cost}(\pi) - \text{Cost}(\gamma) \\ &= \alpha \sum_{i=1}^n \left((C) \int_{X \times Y} c(x, y) d\pi_i \right) - \alpha \sum_{i=1}^n \left((C) \int_{X \times Y} c(x, y) d\mu_i \otimes \nu_i \right) \\ &\geq \alpha \sum_{i=1}^n (c(\bar{x}_i, \bar{y}_i) - \varepsilon) - \alpha \sum_{i=1}^n (c(\bar{x}_{\sigma(i)}, \bar{y}_i) + \varepsilon) \\ &= \alpha \sum_{i=1}^n \left(c(\bar{x}_i, \bar{y}_i) - \sum_{i=1}^n c(\bar{x}_{\sigma(i)}, \bar{y}_i) - 2n\varepsilon \right) > 0. \end{aligned}$$

Since this contradicts the optimality of π , we conclude that $\text{supp}(\pi)$ is c -cyclically monotone.

COROLLARY 2. *Under the assumption of Theorem 3, there is a c -cyclically monotone subset of $X \times Y$ containing the supports of all supermodular optimal transport plans.*

Proof. Indeed, $\Pi_{\text{Ch}^*}(\mu, \nu)$ is a convex set, a fact which easily implies that the union of all supports $\text{supp}(\pi)$ with $\pi \in \Pi_{\text{Ch}^*}(\mu, \nu)$ is a c -cyclically monotone set.

As was noticed by Rüschemdorf [20] and Smith and Knott [22], the notion of c -cyclically monotone set is intimately related to the theory of c -concave functions (concavity relative to a cost function). One important fact in this connection is the existence for each c -cyclically monotone subset S of $X \times Y$ of a pair of continuous functions $\varphi : X \rightarrow \mathbb{R}$ and $\psi : Y \rightarrow \mathbb{R}$ such that:

- (CM1) $\varphi(x) = \inf \{c(x, y) - \psi(y) : y \in Y\}$ for all x ;
- (CM2) $\psi(y) = \inf \{c(x, y) - \varphi(x) : x \in X\}$ for all y ;
- (CM3) $S \subset \{(x, y) \in X \times Y : \varphi(x) + \psi(y) = c(x, y)\}$.

Combining this remark with Theorem 3 one can easily deduce the following extension of Kantorovich

duality to the framework of supermodular upper continuous capacities:

$$\min_{\gamma \in \Pi_{\text{Ch}^*}(\mu, \nu)} \text{Cost}(\gamma) = \sup \left\{ (C) \int_X \varphi d\mu + (C) \int_Y \psi d\nu : (\varphi, \psi) \in C(X) \times C(Y), \varphi(x) + \psi(y) \leq c(x, y) \right\}.$$

Remark 4. Theorem 3 and Theorem 4 also work in the context of lower continuous supermodular capacities.

REFERENCES

1. D.R. ADAMS, *Choquet integrals in potential theory*, Publicacions Matemàtiques, **42**, pp. 3-66, 1998.
2. L. AMBROSIO, *Lecture notes on optimal transport problems*, In vol. *Mathematical Aspects of Evolving Interfaces*, Lecture Notes in Mathematics, **1812**, pp. 1-52, 2003, Springer-Verlag, Berlin.
3. S. CERREIA-VIOGLIO, F. MACCHERONI, M. MARINACCI, L. MONTRUCCHIO, *Signed integral representations of comonotonic additive functionals*, J. Math. Anal. Appl., **385**, pp. 895-912, 2012.
4. G. CHOQUET, *Theory of capacities*, Annales de l' Institut Fourier, **5**, pp. 131-295, 1954.
5. G. CHOQUET, *La naissance de la théorie des capacités: réflexion sur une expérience personnelle*, Comptes rendus de l' Académie des sciences, Série générale, La Vie des sciences, **3**, pp. 385-397, 1986.
6. D. DENNEBERG, *Non-additive measure and integral*, Kluwer Academic Publisher, Dordrecht, 1994.
7. N. DINCULEANU, *Vector measures*, Pergamon Press, New York, 1967.
8. L.G. EPSTEIN, T. WANG, "Beliefs about beliefs" without probabilities, *Econometrica*, **64**, 6, pp. 1343-1373, 1996.
9. L.C. EVANS, *Partial differential equations and Monge-Kantorovich mass transfer*, in vol. *Current Developments in Mathematics, 1997* (ed. S.T. Yau), International Press, 1999.
10. H. FÖLLMER, A. SCHIED, *Stochastic Finance*, Fourth revised and extended edition, De Gruyter, 2016.
11. A. GALICHON, *Optimal transport methods in economics*, Princeton University Press, 2016.
12. W. GANGBO, R.J. McCANN, *The geometry of optimal transportation*, Acta Math., **177**, pp. 113-161, 1996.
13. P. GHIRARDATO, *On independence for non-additive measures, with a Fubini theorem*, Journal of economic theory **73**, pp. 261-291, 1997.
14. L.V. KANTOROVICH, *On the translocation of masses*, Journal of Mathematical Sciences, **133**, pp. 1381-1382, 2006. Translation after Russian original, Dokl. Akad. Nauk SSSR, **37**, pp. 199-201, 1942.
15. L.V. KANTOROVICH, *On a problem of Monge*, Journal of Mathematical Sciences, **133**, p. 1383, 2006. Translation after Russian original, Uspekhi Mat. Nauk., **3**, pp. 225-226, 1948.
16. G. MONGE, *Mémoire sur la théorie des déblais et des remblais*, Histoire de l' Académie Royale des Sciences de Paris, avec les Mémoires de Mathématique et de Physique pour la même année, pp. 666-704, 1781.
17. G.L. O'BRIEN, W. VERVAAT, *Capacities, large deviations and loglog laws*, in vol. *Stable Processes and Related Topics* (eds. S. Cambanis, G. Samorodnitsky, M.S. Taquu), Birkhäuser, Boston, 1991, pp. 43-83.
18. S.T. RACHEV, *The Monge-Kantorovich mass transference problem and its stochastic applications*, Theory of Probability and Its Applications, **19**, pp. 647-676, 1985.
19. S.T. RACHEV, L. Rüschemdorf, *Mass Transportation Problems*, Vol. I: "Theory" and Vol. 2: "Applications", Springer-Verlag, Berlin, 1998.
20. L. RÜSCHEMDORF, *On c-optimal random variables*, Statistics and Probability Letters, **27**, pp. 267-270, 1996.
21. F. SANTAMBROGIO, *Optimal transport for applied mathematicians*, Birkhäuser, New York, 2015.
22. C. SMITH, M. KNOTT, *On Hoeffding-Fréchet bounds and cyclic monotone relations*, Journal of Multivariate Analysis, **40**, pp. 328-334, 1992.
23. C. VILLANI, *Topics in optimal transportation*, Graduate Studies in Mathematics, **58**, American Mathematical Society, 2003.
24. C. VILLANI, *Optimal transport: old and new*, Grundlehren der mathematischen Wissenschaften, **338**, 2009, Springer-Verlag, Berlin.
25. Z. WANG, G.J. KLIR, *Generalized measure theory*, Springer-Verlag, New York, 2009.
26. Z. WANG, J.-A. YAN, *A selective overview of applications of Choquet integrals*, Advanced Lectures in Mathematics, pp. 484-515, Higher Educational Press and International Press, 2007.
27. L. ZHOU, *Integral representation of continuous comonotonically additive functionals*, Trans. Amer. Math. Soc., **350**, pp. 1811-1822, 1998.

Received on January 8, 2019