Mathematical Analysis

ON OPTIMALITY OF c-CYCLICALLY MONOTONE TRANSFERENCE PLANS
SUR L'OPTIMALITÉ DES PLANS DE TRANSPORT c-CYCLIQUES MONOTONES

STEFANO BIANCHINI AND LAURA CARAVENNA

Abstract. This note deals with the equivalence between the optimality of a transport plan for the Monge-Kantorovich problem and the condition of c-cyclical monotonicity, as an outcome of the construction presented in [7]. We emphasize the measurability assumption on the hidden structure of linear preorder.
Résumé. Dans la présente note nous décrivons brièvement la construction introduite dans [7] à propos de l'équivalence entre l'optimalité d'un plan de transport pour le problème de Monge-Kantorovich et la condition de monotonie c-cyclique — ainsi que d'autres sujets que cela nous amène à aborder. Nous souhaitons mettre en évidence l'hypothèse de mesurabilité sur la structure sous-jacente de pré-ordre linéaire.

1. Introduction to the Problem

Optimal mass transportation has been an exceptionally prolific field in the very last decades, both in theory and applications. What we reconsider is though a basic question in the foundations.

Let μ, ν be two Borel probability measures on $[0,1]$ and $c : [0,1]^2 \to [0,+\infty]$ a cost function. The Monge-Kantorovich problem deals with the minimization of the *cost functional*

$$\mathcal{I}(\pi) := \int c(x,y)\pi(dxdy)$$

among the family of *transport plans* $\pi \in \Pi(\mu,\nu)$, which are defined as probability measures on $[0,1]^2$ having marginals respectively μ, ν: denoting with $\mathcal{B}$ the Borel σ-algebra,

$$\Pi(\mu,\nu) := \Big\{\pi \in \mathcal{P}([0,1]^2) : \ \pi(A \times [0,1]) = \mu(A), \ \pi([0,1] \times A) = \nu(A) \text{ for } A \in \mathcal{B}\Big\}.$$

We tacitly assumed that c is π-measurable for all $\pi \in \Pi(\mu,\nu)$.

When the cost c is l.s.c., then it is shown in [2] that a transport plan π, with $\mathcal{I}(\pi) < +\infty$, is optimal if it is concentrated on a *c-cyclically monotone* set Γ, meaning ([16]) that Γ satisfies the pointwise condition

$$\forall M \in \mathbb{N}, (x_i,y_i) \in \Gamma \quad c(x_0,y_0) + \cdots + c(x_M,y_M) \le c(x_1,y_0) + \cdots + c(x_M,y_{M-1}) + c(x_1,y_M).$$

This expresses that one cannot lower the cost of π by cyclic perturbations of the transport plan (see also below). They also provided the following counterexample showing that the condition is not sufficient in general.

Example 1.1 (Fig. 1). Consider $\mu = \nu = \mathcal{L}^1\llcorner_{[0,1]}$ and $\alpha \in [0,1] \setminus \mathbb{Q}$. Let

$$c(x,y) = \begin{cases} 1 & y = x \\ 2 & y = x + \alpha \mod 1 \\ +\infty & \text{otherwise} \end{cases}.$$

Being α irrational, the plan $(x, x+\alpha \mod 1)_\sharp \mathcal{L}^1$ is trivially c-cyclically monotone: the verification with $\Gamma := \{(x, x+\alpha \mod 1)\}_{x\in[0,1]}$ leads to $2M < +\infty$, $M \in \mathbb{N}$. However it is not optimal, since $(x,x)_\sharp \mathcal{L}^1$ has lower cost.

Figure 1: Level sets of the cost. Ensemble de niveau du c.

Since c-cyclical monotonicity is more handily verifiable, [17] rose the question of its equivalence with optimality for $c(x,y) = \|y-x\|^2$. Improvements of [2] were soon given, independently, in the case both of atomic marginals or continuous cost ([13]) and in the case of real valued, l.s.c. cost functions c ([14]), answering Villani question. Since then other cases have been covered ([15, 3, 6, 4, 5]). We briefly outline here the approach we pursued in [7].

2. Main Statement

Let $\bar\pi \in \Pi(\mu,\nu)$ a transference plan, with finite cost $\mathcal{I}(\bar\pi) < +\infty$, concentrated on a c-cyclically monotone subset Γ of $\{c < +\infty\}$. The aim is to give a concrete construction in order to test whether $\bar\pi$ is optimal by exploiting the c-cyclical monotonicity of Γ.

The answer we propose relies on the following intrinsic preorder on $[0,1]$. We recall that a *preorder* is a transitive relation $R \subset [0,1]^2$: whenever $xRx', x'Rx''$ then xRx''[(a)]. It reduces to a *partial order* when it is antisymmetric: if $xRx', x'Rx$ then $x = x'$. A relation is *linear* if every two elements are comparable: for $x, x' \in [0,1]$ either xRx' or $x'Rx$.

Definition 2.1. Define $x \preccurlyeq x'$ if there exists an *axial path* with finite cost connecting them:

$$\exists (x_i, y_i) \in \Gamma,\ i = 0,\dots,I,\ I \in \mathbb{N},\ x_0 = x,\ x_{I+1} = x' \ :$$
$$(x_{i+1}, y_i) \in \{c < +\infty\} \quad \forall i = 0,\dots,I.$$

Define $x \sim x'$ if there exists a *closed cycle* with finite cost connecting them:

$$\exists (x_i, y_i) \in \Gamma, i = 0,\dots,I,\ I \in \mathbb{N},\ \exists j \in \{0,\dots,I\},\ x_0 = x_{I+1} = x,\ x_j = x' \ :$$
$$(x_{i+1}, y_i) \in \{c < +\infty\} \quad \forall i = 0,\dots,I.$$

Lemma 2.2. *The relation $\preccurlyeq$ is a preorder. Moreover*

$$x \sim x' \qquad \textit{iff} \qquad x \preccurlyeq x' \textit{ and } x' \preccurlyeq x.$$

As a corollary, the relation $\{(x,x') : \ x \sim x'\}$ is an equivalence relation on a subset of $[0,1]$, easily extendable as $\{x \sim x'\} \cup \{x = x'\}$. The preorder $\preccurlyeq$ induces a partial order on the quotient space $[0,1]/\sim$, but its extension is more subtle. By the Axiom of Choice, every preorder is a subset of some linear preorder B — nevertheless, this set B in general does not satisfy any measurability condition. Let m a Borel probability measure.

Definition 2.3. A preorder P on $[0,1]$ is *Borel linearizable* if there exists a Borel linear order $B \supset P$ s.t. $B \cap B^{-1} = P \cap P^{-1} \cup \{x = y\}$, and *$m$-linearizable* if instead B is π'-measurable for all $\pi' \in \Pi(m,m)$.

Theorem 2.4. *If the preorder $\{x \preccurlyeq x'\}$ is μ-linearizable, then $\bar\pi$ minimizes $\mathcal{I}$.*

An analogous statement holds with ν instead of μ if one applies Definition 2.1 after inverting the coordinates; let $\propto$ and $\approx$ be the corresponding partial order and equivalence relation.

Since it may seem too abstract, before justifying why the statement holds we observe that it is not difficult to prove how the hypotheses are satisfied under the following requirement: there exists a countable family of Borel sets $A_i, B_i \subset [0,1]$, $i \in \mathbb{N}$, such that

$$\pi\Big(\bigcup_i A_i \times B_i\Big) = 1, \qquad \mu \otimes \nu\big(\cup_i (A_i \times B_i) \cap \{c = +\infty\}\big) = 0.$$

Indeed, this reduces the problem to countably many classes of $\sim$ and consequently to the linearization of a preorder on a discrete set, easily solvable by induction.

Notice moreover what happens in Example 1.1: $\sim$ is the trivial equivalence relation $\{y = x\}$ and $\preccurlyeq$ ends to the standard example of non linearizable Borel preorder — the Vitali one ([9]). If the preorder $\preccurlyeq$ is Borel, then either Theorem 2.4 holds or the preorder includes a copy of the Vitali preorder ([11]).

3. Sketch of the Proof

The result of Theorem 2.4 is based on a reduction argument. More precisely, we split the optimal transport problem within the classes of $\sim$ and to a problem of uniqueness in the quotient space $[0,1]/\sim$. This is formalized by means of the Disintegration Theorem, a very useful tool which decomposes a measure in a superposition of *conditional measures* concentrated on given subsets, thus 'localizing' it. Before presenting the reduction argument, we state what will solve the reduced problems.

Lemma 3.1. *If μ' is concentrated on a class of $\sim$, then each $\pi' \in \Pi(\mu',\nu')$ concentrated on Γ is a minimizer of the cost functional $\mathcal{I}$ in $\Pi(\mu',\nu')$.*

Sketch. Lemma 3.1 is a direct consequence of the definition of $\sim$. Indeed, for each equivalence class C, if a transport plan in $C \times \Gamma(C)$[(b)] is concentrated on Γ then by Rüschendorf formula one constructs optimal Kantorovich potentials and deduces the optimality. A different proof can be also found in [15]. ☐

[(a)] We are adopting the notation xRx' if $(x,x') \in R$.

[(b)] We use the notation of multivalued function $\Gamma(A) := \{y : \ \exists x \in A : (x,y) \in \Gamma\}$, as well as $\Gamma^{-1} := \{(y,x) : \ (x,y) \in \Gamma\}$.

Lemma 3.2. *Let $\breve\pi \in \Pi(\breve\mu, \breve\mu)$ be concentrated on a $\breve\mu \otimes \breve\mu$-measurable linear preorder L and let E the equivalence relation $L \cap L^{-1}$. Then $\breve\pi(E) = 1$.*

Sketch. Let $\trianglelefteq$ be the lexicographic ordering in $[0,1]^\alpha$, with $\alpha \in \omega_1$ countable ordinal. We first exhibit a Borel map $h_\alpha : [0,1] \to [0,1]^\alpha$ s.t., up to an $\breve\mu$-negligible set, xLx' if and only if $h_\alpha(x) \trianglelefteq h_\alpha(x')$. Then $\breve\pi(E) = 1$ is equivalent to $(h_\alpha \otimes h_\alpha)_\sharp \breve\pi(\{\alpha = \beta\}) = 1$. This follows proving that there is a unique transport plan from a measure to itself concentrated on $\{\alpha \trianglelefteq \beta\} \subset [0,1]^\alpha \times [0,1]^\alpha$, induced by the identity map.

The definition of h_α, by transfinite induction, is based on the Disintegration Theorem, introduced just below. The first component is $h_1(x) := \breve\mu(\{x' : x' \preccurlyeq x\})$; then we disintegrate $\breve\mu$ w.r.t. h_1 and define $h_2(x) = (h_1(x), \breve\mu_{h_1(x)}(\{x' : x' \preccurlyeq x\}))$, ... The sequence becomes constant in $|\alpha|$ steps. For L Borel, in [9] one finds a different Borel order preserving immersion in $(\{0,1\}^\alpha, \trianglelefteq)$ without reference measures. $\square$

We need now some basic technicality for explaining the reduction argument. Observe that one can associate, by the Axiom of Choice, an equivalence relation on $[0,1]$ with a map $q : [0,1] \to [0,1]$ whose level sets are the equivalence classes of the relation. One can then realize the quotient space as a subset of $[0,1]$. Given a Borel probability measure ξ on $[0,1]$, the *push forward* probability measure $\eta = q_\sharp \xi$ and the push forward σ-algebra of $\mathcal{B}$ are defined as

$$\begin{aligned} S \in q_\sharp \mathcal{B} \qquad &\Longleftrightarrow \qquad q^{-1}(S) \in \mathcal{B}, \\ \eta(S) := \xi(q^{-1}(S)) \qquad &\text{for } S \in q_\sharp \mathcal{B}. \end{aligned}$$

Definition 3.3. The disintegration of a Borel probability measure ξ on $[0,1]$ *strongly consistent* with a map $q : [0,1] \to [0,1]$ is a family of Borel probability measures $\{\xi_\alpha\}_{\alpha \in [0,1]}$, the *conditional probabilities*, such that $\alpha \mapsto \int_S \xi_\alpha(O)$ is η-measurable, where $\eta := q_\sharp \xi$, and

$$\xi(O \cap h^{-1}(S)) = \int_S \xi_\alpha(O)\eta(d\alpha) \qquad \text{for all } O \in \mathcal{B},\ S \in q_\sharp \mathcal{B}, \tag{3.1a}$$

$$\xi_\alpha(X_\alpha) = 1 \qquad \text{for } \eta\text{-a.e. } \alpha \in [0,1]. \tag{3.1b}$$

Theorem 3.4 (Disintegration Theorem). *If $q_\sharp \mathcal{B} \supset \mathcal{B}$, then there exists a unique disintegration strongly consistent with q, meaning that $\xi_\alpha = \xi'_\alpha$ for η-a.e. α, for any other family $\{\xi'_\alpha\}_{\alpha \in [0,1]}$ satisfying* (3.1a).

The reduction argument consists in disintegrating μ in probability measures $\{\mu_\alpha\}_{\alpha \in [0,1]}$ on the equivalence classes of $\sim$, ν in probabilities $\{\nu_\alpha\}_{\alpha \in [0,1]}$ on the classes of $\approx$ and every transport plan $\pi \in \Pi(\mu, \nu)$ with finite cost in transport plans $\pi_\alpha \in \Pi(\mu_\alpha, \nu_\alpha)$. From $\bar\pi(\Gamma) = 1$ one would obtain $\bar\pi_\alpha \in \Pi(\mu_\alpha, \nu_\alpha)$ concentrated on Γ. Lemma 3.1 would ensure that $\mathcal{I}(\bar\pi_\alpha) \leq \mathcal{I}(\pi_\alpha)$ for m-a.e. α. Therefore in this case, by the disintegration formula and the optimality within the classes, we would get $\mathcal{I}(\bar\pi) \leq \mathcal{I}(\pi)$:

$$\begin{aligned} \int c(x,y)\bar\pi(dxdy) &\overset{(3.1)}{=} \int \bigg\{ \int c(x,y)\bar\pi_\alpha(dxdy) \bigg\} m(d\alpha) \\ &\overset{\text{L.3.1}}{\leq} \int \bigg\{ \int c(x,y)\pi_\alpha(dxdy) \bigg\} m(d\alpha) \overset{(3.1)}{=} \int c(x,y)\pi(dxdy). \end{aligned}$$

We are thus left with justifying Consider the linear order condition of Theorem 2.4. The proof of Lemma 3.2 incidentally provides also a Borel quotient map $q_1 := h \circ h_\alpha : [0,1] \to [0,1]$ for $\sim$, where h is a Borel injection from $[0,1]^\alpha$ to $[0,1]$.

Corollary 3.5. *Under the hypothesis of Theorem 2.4, μ has a disintegration strongly consistent with $\sim$.*

Now it is worth noticing that the nontrivial equivalence classes of $\approx$ are of the form

$$\Gamma(A) = \big\{ y : (x,y) \in \Gamma \text{ for some } x \in A \big\} \qquad \text{with } A = \{x' : x' \sim x\},\ x \in [0,1]. \tag{3.2}$$

As a consequence, one can define the quotient projection w.r.t. $\approx$ by setting $q_2(\Gamma(x)) := q_1(x)$ and also ν has a disintegration strongly consistent with $\approx$, since q_1 is Borel. By (3.2) then the quotient probability spaces $([0,1], \mu)/\sim$ and $([0,1], \nu)_{/\approx}$ can be identified with a Borel probability space $([0,1], m)$.

For any plan $\pi \in \Pi(\mu, \nu)$ its quotient measure n w.r.t. the product equivalence relation $q_1 \otimes q_2$ belongs consequently to $\Pi(m, m)$. If π has finite cost, n is clearly concentrated on $q_1 \otimes q_2(\{c < +\infty\})$.

Lemma 3.6. *The set $q_1 \otimes q_2(\{c < +\infty\})$ is the partial order $q_1 \otimes q_1(\{x \preccurlyeq x'\})$ in the quotient space.*

The assumption of Theorem 2.4 grants that this partial order can be extended to a $m \otimes m$-measurable linear order. Applying Lemma 3.2 one obtains then that $n = (\mathbb{I}, \mathbb{I})_\sharp m$ for every plan of finite cost π. As a consequence, π admits the strongly consistent disintegration $\pi = \int \pi_\alpha m(d\alpha)$ w.r.t. the partition $\{(q_1 \otimes q_2)^{-1}(\alpha)\}_{\alpha \in [0,1]}$. The reasoning is finally concluded by the following lemma.

Lemma 3.7. *By the marginal conditions, $\pi_\alpha \in \Pi(\mu_\alpha, \nu_\alpha)$ for m-a.e. α.*

4. Mention of side Studies and Remarks

Our basic tool has been the Disintegration Theorem. The main references for our review on that has been [8, 1]. We also applied it to a family of equivalence relations closed under countable intersection, establishing that there is an element of the family which is the finest partition, in a measure theoretic sense. In particular, in the construction of the immersion h_α of Lemma 3.2 we ended up with such a family, which was not closed under uncountable intersection because uncountable intersections of sets generally are not measurable. Having a finest element, we could value the projection in the Polish space $[0,1]^\alpha$, $\alpha \in \omega_1$, instead of $[0,1]^{\omega_1}$.

As briefly sketched, optimality holds in the equivalence classes basically by Kantorovich duality, the equivalence relation is indeed chosen for having real valued optimal Kantorovich potentials by Rüschendorf's formula.

Moreover, the necessity of c-cyclical monotonicity with co-analytic cost functions — clearly assuming that the optimal cost is finite — is a corollary of the general duality in [12][c]. We notice that it follows just by the fact that there is no cyclic perturbation λ of the optimal plan π such that $\mathcal{I}(\pi+\lambda) < \mathcal{I}(\pi)$, where cyclic perturbations of π are defined as nonzero measures λ with Jordan decomposition $\lambda = \lambda^+ - \lambda^-$ satisfying $\lambda^- \leq \pi$ and which can be written, for some $m_I \in \mathcal{M}^+([0,1]^{2I})$, $I \in \mathbb{N}$, as

$$\lambda^+ = \sum_I \frac{1}{I}\int_{[0,1]^{2I}} \sum_{i=1}^{I} \delta_{P_{(2i-1,2i)}w} m(dw), \quad \lambda^- = \sum_I \frac{1}{I}\int_{[0,1]^{2I}} \sum_{i=1}^{I} \delta_{P_{(2i+1 \mod 2n, 2i)}w} m(dw).$$

Observe that if $\preccurlyeq$ is μ-linearizable, each transport plan of finite cost is concentrated on $(q_1 \otimes q_2)^{-1}(\{\alpha = \beta\})$; as a separate observation based on Von Neumann's Selection Theorem, we construct optimal potentials for the cost which is $+\infty$ out of that set, gluing the ones in the classes.

As a final remark on the topic, we observe the following asymmetry: for universally measurable cost functions, c-cyclically monotone transference plans are optimal under the universally measurable linear preoder condition; however, in this case the necessity of c-cyclical monotonicity is not proven, since duality is provided in [12] for analytic functions, corresponding to co-analytic costs.

In general, $\preccurlyeq$ can be μ-linearizable for some c-cyclically monotone set Γ, $\pi(\Gamma) = 1$, and not for others, and we do not see how to choose a best one — which in the case of continuous cost would be the support. Another question is what happens when there is no such set Γ such that $\preccurlyeq$ is μ-linearizable. Examples show a crazy behavior indicating that this construction, built trying to encode all the informations given just by c-cyclical monotonicity, is maybe not suitable to answer. As already mentioned, for Borel sets [11] states that in this case there is a situation analogous to Example 1.1 and the quotient projection w.r.t. $\{x \preccurlyeq y\} \cup \{x \preccurlyeq y\}^{-1}$ is not universally measurable — but, however, since we have fixed measures optimality could still hold.

We conclude noticing that in [7] we study with the same approach the problems of establishing if a plan $\pi \in \Pi(\mu,\nu)$ is extremal and if it is the unique plan in $\Pi(\mu,\nu)$ concentrated on a given set A, say universally measurable. In the first case we precisely recover the condition in [10]. The second case comes from the problem of uniqueness in the quotient space described above, and we answer by a μ-linearizability condition.

References

[1] L. Ambrosio, N. Fusco, and D. Pallara. *Functions of bounded variation and free discontinuity problems.* Oxford, 2000.

[2] L. Ambrosio and A. Pratelli. Existence and stability results in the L^1 theory of optimal transportation. In Springer, editor, *Optimal Transportation and Applications*, volume 1813 of *Lecture notes in mathematics*, 2001.

[3] M. Beiglböck, M. Goldstern, G. Maresch, and W. Schachermayer. Optimal and better transport plans. *J. Funct. Anal.*, 256(6):1907–1927, 2009.

[4] M. Beiglböck, C. Leonard, and W. Schachermayer. A general duality theorem for the Monge–Kantorovich transport problem.

[5] M. Beiglböck, C. Leonard, and W. Schachermayer. On the duality theory for the Monge–Kantorovich transport problem.

[6] M. Beiglböck and W. Schachermayer. Duality for Borel measurable cost function. Preprint.

[7] S. Bianchini and L. Caravenna. On the extremality, uniqueness and optimality of transference plans. *Bull. Inst. Math. Acad. Sin. (N.S.)*, 4(4):353–455, 2009.

[8] D. H. Fremlin. *Measure theory*, volume 1-4. Torres Fremlin, Colchester, 2001-2003.

[9] L. Harrington, D. Marker, and S. Shelah. Borel orderings. *Trans. Amer. Math. Soc.*, 310(1):293–302, 1988.

[10] K. Hestir and S. C. Williams. Supports of doubly stochastic measures. *Bernoulli*, 1(3):217–243, 1995.

[11] V. Kanovei. When a partial Borel order is linearizable. *Fund. Math.*, 155(3):301–309, 1998.

[12] H. G. Kellerer. Duality theorems for marginals problems. *Z. Wahrcsh. Verw. Gebiete*, 67(4):399–432, 1984.

[13] A. Pratelli. On the sufficiency of c-cyclical monotonicity for optimality of transport plans. *Math. Z.*, 2007.

[14] W. Schachermayer and J. Teichmann. Solution of a problem in Villanis book. Preprint, 2005.

[15] W. Schachermayer and J. Teichmann. Characterization of optimal transport plans for the Monge-Kantorovich problem. *Proc. Amer. Math. Soc.*, 137(2):519–529, 2009.

[16] C. Smith and M. Knott. On Hoeffding-Fréchet bounds and cyclic monotone relations. *J. Multivariate Anal.*, 40(2):328–334, 1992.

[17] C. Villani. *Topics in Optimal Transportation.* Providence, RI American Mathematical Society, AMS, Graduate Studies in Mathematics vol. 58, 2003.

[c] Necessity was proven in [2] by Kantorovich duality for l.s.c. cost functions, and extended to the Borel case in [3] with Kellerer results. We provide in [7] a different proof, still based on Keller duality, for co-analytic cost functions.

SISSA, via Beirut 2, IT-34014 Trieste (ITALY)
E-mail address: bianchin@sissa.it

CRM De Giorgi, Collegio Puteano, Scuola Normale Superiore, Piazza dei Cavalieri 3, I-56100 PISA (ITALY)
E-mail address: laura.caravenna@sns.it