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. Denote by $\mathcal{B}$ the Borel σ-algebra. 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 Borel probability measures on $[0,1]^2$ having marginals μ, ν:

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

We tacitly assume that c is π-measurable for all $\pi \in \Pi(\mu,\nu)$ and that $\mathcal{I}(\pi) < +\infty$ for some $\pi \in \Pi(\mu,\nu)$.

When the cost c is l.s.c., then by [2] any optimal transport plan π must be concentrated on a *c-cyclically monotone* (briefly c-monotone) set Γ, meaning ([16]) that Γ satisfies the pointwise condition

$$\forall M \in \mathbb{N}, (x_i,y_i) \in \Gamma \quad c(x_0,y_0) + \cdots + c(x_M,y_M) \le c(x_1,y_0) + \cdots + c(x_M,y_{M-1}) + c(x_1,y_M). \tag{1.1}$$

This expresses that one cannot lower the cost of π by cyclic perturbations of the transport plan (see also below). In [2] it is 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

$$c(x,y) = \begin{cases} 1 & y = x \\ 2 & y = x + \alpha \mod 1 \ , \ \alpha \in [0,1] \setminus \mathbb{Q}. \\ +\infty & \text{otherwise} \end{cases}$$

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.

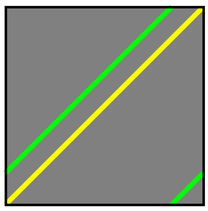

Figure 1. Level sets of the cost. Ensemble de niveau de c.

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's 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 $\mathcal{I}(\bar\pi)$, concentrated on a c-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-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

[(a)] We denote xRx' if $(x,x') \in R$.

$xRx', x'Rx$ then $x = x'$. A relation is *linear* if every two elements are comparable: for $x, x' \in [0,1]$ either xRx' or $x'Rx$. If R is a linear preorder, notice also that $R \cap R^{-1}$ is a natural equivalence relation associated to it.

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

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

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

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

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

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

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. Let m be 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. $B \cap B^{-1} = P \cap P^{-1} \cup \{x = y\}$. It is *$m$-linearizable* if instead B is π'-measurable for all $\pi' \in \Pi(m, m)$.

The following result do not rely on the Axiom of Choice. However, under it any preorder is a subset of a linear preorder B—but B in general fails to be $m \otimes m$-measurable ([7], C.12): μ-linearizability is then a measurability condition.

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.

Corollary 2.5. *If there exists a countable family of Borel sets $A_i, B_i \subset [0,1]$, $i \in \mathbb{N}$, s.t.*

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

then the preorder $\{x \preccurlyeq x'\}$ is μ-linearizable and therefore $\bar\pi$ minimizes $\mathcal{I}$.

The corollary covers the presently known cases where c-monotonicity implies optimality: there are countably many classes of $\sim$ and consequently the linearization of the preorder, on a discrete set, is easily solvable by induction.

Notice moreover what happens in Example 1.1 choosing the diagonal as Γ: $\sim$ is the trivial equivalence relation $\{y = x\}$ and $\preccurlyeq$ ends to the standard example of non linearizable Borel preorder—the Vitali one ([9]); with $\Gamma = \{c < +\infty\}$ instead the quotient is the Vitali set and there is no disintegration of $\mathcal{L}^1$ strongly consistent with $\sim$.

If the preorder $\preccurlyeq$ is Borel, then either Theorem 2.4 holds or the preorder embeds 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. We recall the Disintegration Theorem in the last section.

The solution of the reduced transport problems will follow from the next property of the equivalence classes.

Lemma 3.1. *Let μ', ν' be Borel probability measures on $[0,1]$. 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. Optimality is obtained constructing by Rüschendorf's formula optimal Kantorovich potentials in each equivalence class C, for any transport plan concentrated on [b] $\Gamma \cap (C \times \Gamma(C))$. A different proof is also in [15]. □

The reduction argument, performed just below, 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: the perturbations occur within the equivalence classes. 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

$$\begin{aligned}\mathcal{I}(\bar\pi) = \int c(x,y)\bar\pi(dxdy) &\overset{(A.1)}{=} \int \bigg\{ \int c(x,y)\bar\pi_\alpha(dxdy) \bigg\} m(d\alpha) \\ &\overset{L.3.1}{\leq} \int \bigg\{ \int c(x,y)\pi_\alpha(dxdy) \bigg\} m(d\alpha) \overset{(A.1)}{=} \int c(x,y)\pi(dxdy) = \mathcal{I}(\pi).\end{aligned}$$

If $\preccurlyeq$ is μ-linearizable, the next theorem provides a μ-measurable quotient map $q_1 : [0,1] \to [0,1]$ for $\sim$.

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

Theorem 3.2. *Let $\breve\pi \in \Pi(\breve\mu,\breve\mu)$ be concentrated on a linear preorder L which is π'-measurable for all $\pi \in \Pi(\mu,\nu)$. Let E be the equivalence relation $L \cap L^{-1}$. Then $\breve\pi(E) = 1$ and the disintegration of $\breve\mu$ is strongly consistent.*

Sketch. Let $\trianglelefteq$ be the lexicographic ordering in $[0,1]^\alpha$, with $\alpha \in \omega_1$ countable ordinal. We 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 $h_\alpha(x) \trianglelefteq h_\alpha(x')$. Since $([0,1]^\alpha,\mathcal{B})$ can be measurably injected in $([0,1],\mathcal{B})$, the existence of h_α ensures on one hand that the disintegration of $\breve\mu$ w.r.t. E is strongly consistent (Theorem A.2). On the other hand, $\breve\pi(E) = 1$ becomes equivalent to $(h_\alpha \otimes h_\alpha)_\sharp \breve\pi(\{\alpha = \beta\}) = 1$: indeed, one proves by transfinite induction on α 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. $\square$

Now it is worth noticing the *crosswise structure* of the relations: the nontrivial classes of $\approx$ are of the form

$$\Gamma(A) = \big\{y : (x,y) \in \Gamma \text{ for some } x \in A\big\} \qquad \text{with } A = \{x' : x' \sim x\},\ x \in [0,1]. \tag{3.1}$$

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 $\approx$, by $\bar\pi$'s marginal conditions. By (3.1) 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 3.3. *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 linear order which is π'-measurable for every $\pi' \in \Pi(m,m)$. Applying Theorem 3.2 one obtains then that $n = (\mathbb{I},\mathbb{I})_\sharp m$ for every plan of finite cost π. As a consequence, any $\pi \in \Pi(\mu,\nu)$ 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 crosswise structure yields then the following statement.

Lemma 3.4. *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 Theorem 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]. 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

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

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, one can 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 Souslin 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. This problem is intrinsic in the 'pointwise' definition of c-monotonicity. Another question is what happens when there is no such set Γ such that

[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.

$\preccurlyeq$ is μ-linearizable. Examples show a crazy behavior. 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. $\cup_{M\in\mathbb{N}}\big(\{x \preccurlyeq y\} \cup \{x \preccurlyeq y\}^{-1}\big)^M$ 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. A necessary condition is $\pi(\Gamma) = 1$ for a Γ, say universally measurable, such that the r.h.s. of (1.1) is always $+\infty$ when $c = \chi_\Gamma/$ resp. $c = \chi_A$—this condition defines Γ *acyclic*/*A-acyclic*. In the first case we precisely recover the Borel Countable Limb Condition in [10]. The second case comes from the problem of uniqueness in the quotient space described above. Considering the same axial preorder $\preccurlyeq$ defined above, but with Γ/A in place of $\{c < +\infty\}$, the statement becomes

Theorem 4.1. *If Γ is acyclic/A-acyclic and $\preccurlyeq$ is μ-linearizable, then π is extremal/the unique transport on A.*

Appendix A. The Disintegration Theorem

Consider an equivalence relation $\approx$ on $[0,1]$ and its quotient map $q : [0,1] \to [0,1]/\approx$. Given a Borel probability measure ξ on $[0,1]$, let Θ_ξ be the σ-algebra of ξ-measurable sets. The *push forward* of a σ-algebra of $\mathcal{A} \subset \Theta_\xi$ and the push forward probability measure $\eta = q_\sharp\xi$ are defined on $[0,1]/\approx$ as

$$S \in q_\sharp\mathcal{A} \quad \Longleftrightarrow \quad q^{-1}(S) \in \mathcal{A}, \qquad\qquad \eta(S) := \xi(q^{-1}(S)) \quad \text{for } S \in q_\sharp\Theta_\xi.$$

Definition A.1. The disintegration of a Borel probability measure ξ on $[0,1]$ *strongly consistent* with a map $q : [0,1] \to [0,1]/\approx$ 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 for all $S \in q_\sharp\mathcal{B}$, where $\eta := q_\sharp\xi$, and

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

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

It is *unique* if for any other family $\{\xi'_\alpha\}_{\alpha\in[0,1]}$ satisfying (A.1a) then $\xi_\alpha = \xi'_\alpha$ for η-a.e. α.

Theorem A.2 (Disintegration Theorem)**.** *If there exists a measurable injection $([0,1]/\approx, q_\sharp\Theta_\xi) \hookrightarrow ([0,1],\mathcal{B})$, then there exists a unique disintegration of ξ strongly consistent with q.*

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. Wahrsch. 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.

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`