DUALITY FOR RECTIFIED COST FUNCTIONS

MATHIAS BEIGLBÖCK AND ALDO PRATELLI

ABSTRACT. It is well-known that duality in the Monge-Kantorovich transport problem holds true provided that the cost function $c : X \times Y \to [0, \infty]$ is lower semi-continuous or finitely valued, but it may fail otherwise. We present a suitable notion of *rectification* c_r of the cost c, so that the Monge-Kantorovich duality holds true replacing c by c_r. In particular, passing from c to c_r only changes the value of the primal Monge-Kantorovich problem. Finally, the rectified function c_r is lower semi-continuous as soon as X and Y are endowed with proper topologies, thus emphasizing the role of lower semi-continuity in the duality-theory of optimal transport.

1. Introduction

1.1. **Description of the main question.** We consider the *Monge-Kantorovich transport problem* for Borel probability measures μ, ν on Polish spaces X, Y. Standard references for the theory of optimal transportation are [Vil03, Vil09].

The set $\Pi(\mu, \nu)$ consists of all *transport plans*, that is, Borel probability measures on $X \times Y$ which have X-marginal μ and Y-marginal ν. The *transport cost* associated to a *cost function* $c : X \times Y \to [0, \infty]$ and a transport plan π is given by

$$\langle c, \pi \rangle = \iint_{X \times Y} c(x, y) \, d\pi(x, y). \tag{1}$$

The (primal) Monge-Kantorovich problem is then to determine the value

$$P_c := \inf\{\langle c, \pi \rangle : \pi \in \Pi(\mu, \nu)\}\,. \tag{2}$$

and to identify a primal optimizer $\hat{\pi} \in \Pi(\mu, \nu)$.

A natural condition which guarantees the existence of a primal optimizer is that the cost function c is lower semi-continuous. (See for instance [Vil09, Theorem 4.1].)

To formulate the dual problem, we let

$$\varphi \oplus \psi \, (x, y) := \varphi(x) + \psi(y)$$

for functions φ, ψ on X (resp. Y). The dual Monge-Kantorovich problem then consists in determining

$$D_c := \sup \left\{ \int \varphi \, d\mu + \int \psi \, d\nu : \varphi \in L^1_\mu(Y), \psi \in L^1_\nu(Y), \varphi \oplus \psi \leq c \right\}. \tag{3}$$

The first author acknowledges financial support from the Austrian Science Fund (FWF) under grant P21209. The work of the second author was partially supported by the ERC Starting Grant *Analysis of optimal sets and optimal constants: old questions and new results* and by the MEC of Spain Government through the 2008 project MTM2008-03541. The second author gratefully acknowleges the hospitality of the Wien University.

Given two functions φ, ψ which are integrable with respect to μ and ν respectively, and which satisfy $\varphi \oplus \psi \leq c$, and given a transport plan $\pi \in \Pi(\mu, \nu)$ we clearly have

$$\iint c\, d\pi \geq \iint \varphi \oplus \psi\, d\pi = \int \varphi\, d\mu + \int \psi\, d\nu\,,$$

hence it follows that $P_c \geq D_c$. The question if there actually is *equality*, i.e. whether Monge-Kantorovich duality $P_c = D_c$ holds true, has been intensively studied in the years by many authors, see for instance [Kan42, KR58, Dud76, Dud02, dA82, GR81, Fer81, Szu82, Mik06, MT06], and see also the bibliographical notes in [Vil09, p86, 87]. In particular, it is known that $P_c = D_c$ provided that the cost function c is *lower semi-continuous* (cf. [Kel84, Theorem 2.6] or [Vil09, Theorems 5.10] for a modern source), or merely measurable but *bounded* ([Kel84, Corollary 2.16]) or at least $\mu \otimes \nu$-*a.s. finitely valued* ([BS09, Theorem 1]). However, the duality does not hold in complete generality as simple examples show.

Example 1.1. *Let $X = Y = [0,1]$ and let $\mu = \nu$ be the Lebesgue measure. Define c on $X \times Y$ to be 0 below the diagonal, 1 on the diagonal and ∞ else, i.e.,*

$$c(x,y) = \begin{cases} 0, & \textit{for } 0 \leq y < x \leq 1\,, \\ 1, & \textit{for } 0 \leq x = y \leq 1\,, \\ \infty, & \textit{for } 0 \leq x < y \leq 1\,. \end{cases}$$

The only finite transport plan is concentrated on the diagonal, hence $P_c = 1$. On the other hand, if $\varphi : X \to [-\infty, \infty), \psi : Y \to [-\infty, \infty)$ satisfy $\varphi \oplus \psi \leq c$, one readily verifies that $\varphi(x) + \psi(x) > 0$ can hold true for at most countably many $x \in [0,1]$. Hence $D_c = 0$ so that there is a duality gap.

Let us discuss the example above a little bit. Strictly speaking, one should simply say that it presents a situation where the duality does not hold true. But on the other hand, one would like to say that in fact the duality should hold true, and it fails only because the cost function c takes the "wrong" value on the diagonal, while the "correct" cost function should be

$$c_r(x,y) = \begin{cases} 0 & \text{for } 0 \leq y \leq x \leq 1, \\ \infty & \text{for } 0 \leq x < y \leq 1. \end{cases} \tag{4}$$

In fact, in some sense, around the points in the diagonal there are "many" points where $c = 0$, hence it makes no sense to have $c = 1$ in the diagonal. Notice that with the cost function c_r duality holds, and in particular

$$P_{c_r} = D_{c_r} = D_c\,. \tag{5}$$

Basically, we are saying that in the above example the correct value of both the primal and the dual problem "should be" the same, namely 0, and it is not so only because the cost function c has been defined in a slightly meaningless way. In particular, the fact that $D_{c_r} = D_c$ is saying that the dual problem is less sensitive to the "mistakes" in the definition of c, while the primal problem is more sensitive and indeed $P_c > P_{c_r}$.

The aim of the present paper is to show that the situation is always the one described by means of the above simple example. More precisely, we will show that for any transport problem it is possible to define a meaningful *rectified* cost function $c_r \leq c$, and (5) always holds true. Roughly speaking, this means that duality in the Monge-Kantorovich problem *always* holds true, as soon as one considers the "correct" definitions of the cost functions c. Moreover, an "incorrect" definition may only affect the value of the primal problem, and can be corrected by passing to a suitable rectified cost function c_r.

Let us now describe another important feature of our rectification procedure. Consider a simple variant of Example 1.1, where the value $+\infty$ in the definition of c is replaced by some number $1 < M \in \mathbb{R}$. In this case the cost is finite, then according to the classical results we know that the duality holds. However, the transport problem has now another drawback, namely that there are no optimal transport plans. In fact, the infimum of the costs of the transport plans is now 0, but every transport plan has a strictly positive cost. In particular, every optimizing sequence of transport plans converges to the plan concentrated on the diagonal, which has cost 1. Clearly, also this bad behaviour disappears if one passes to the rectified cost function c_r, which has value 0 in the diagonal and coincides with c outside.

We will show that also this pleasant feature of the rectification process holds in general, that is, the transport problem with the rectified cost c_r always admits optimal transport plans. We can say something even stronger, namely, that for any sequence of plans π_n weakly converging to π, the liminf inequality for the costs holds, that is,

$$\pi_n \rightharpoonup \pi \qquad \Longrightarrow \qquad \langle c_r, \pi\rangle \leq \liminf_{n\to\infty} \langle c_r, \pi_n\rangle \,. \tag{6}$$

Before concluding this introductory description, it is important to underline here two things. First of all, one is easily lead to guess that the correct rectification c_r is simply the lower semi-continuous envelope of c. In fact, c_r coincides with the l.s.c. envelope of c in the two examples that we presented above, and moreover for a l.s.c. function the property (6) is clearly always true. However, it is also easy to realize that the l.s.c. of c *does not* work as we want. To see this, it is enough to consider the following example.

Example 1.2. *Let $X = Y = [0,1]$, let $\mu = \nu$ be the Lebesgue measure, and define*

$$c(x,y) = \begin{cases} 0 & \textit{if } (x,y) \in \mathbb{Q}\times\mathbb{Q}\,, \\ 1 & \textit{otherwise}\,. \end{cases}$$

In this case, the value of the cost function is almost surely 1*, so the problem is perfectly equivalent to the trivial problem with $c \equiv 1$, hence the duality already holds, the minimum is already attained, and there is no need to change anything. But on the other hand, the lower semi-continuous envelope of c is costantly* 0*.*

Looking at the example above, one easily understands what is wrong with the l.s.c. envelope. Roughly speaking, one needs to have $c_r(x,y) < c(x,y)$ if there are "many" points around (x,y) with a low value of the cost function, while the lower semi-continuous envelope goes down even if there are only "few", but infinitely close, such points. We will make a better discussion later, with Example 3.2.

The second thing that we want to underline is, whether or not the rectification c_r of c depends on the measures μ and ν. On one hand, it seems quite reasonable, and it would be of course much better, if it is not the case and c_r depends only on c. But on the other hand, it is also easy to realize that this is not possible in general. In fact, if for instance μ and ν are concentrated on two points $\bar{x} \in X$ and $\bar{y} \in Y$, then the value of c out of $(\bar{x}, \bar{y})$ does not play any role and it cannot affect the definition of c_r. More precisely, we can observe that the fact whether there are "many" of "few" points around $(x, y) \in X \times Y$ of course depends on the measures μ and ν. In fact, we will show (see Remark 3.1) that the rectification c_r of c only depends on the class of negligible sets with respect to μ and ν, which is the best one could hope in view of the above considerations.

1.2. **Formal statement of our result.** In this section, we can give the formal definition of the rectification c_r of c and the correct statement of our main result. First of all, we need to introduce the following notion.

Definition 1.3. *A set $A \subseteq X \times Y$ is called* L-negligible *if there exist two sets $M \subseteq X$ and $N \subseteq Y$ with $\mu(M) = \nu(N) = 0$ such that*

$$A \subseteq (M \times Y) \cup (X \times N)\,.$$

Accordingly, if a property holds on the complement of an L-negligible set, than we sey that it holds L-almost surely.

It is trivial but fundamental to observe that the transport problem is not affected if the cost function is changed on an L-negligible set.

We can now give our definition of the rectified cost function.

Definition 1.4. *Let $c : X \times Y \to [0, \infty]$ be measurable. A function $c_r : X \times Y \to [0, +\infty]$ is said to be the* rectification of c *if the following holds:*

(i) *for all Borel functions $\varphi, \psi : [0, 1] \to [-\infty, \infty)$ satisfying $\varphi \oplus \psi \leq c$ we have $\varphi \oplus \psi \leq c_r$ L-almost surely;*
(ii) *c_r is minimal subject to (i), i.e. if d is another function satisfying (i) then L-almost surely $c_r \leq d$.*

It is clear from (ii) that every cost function has at most one rectification, while the existence is not obvious. We can now state our result.

Theorem 1. *Take two Polish spaces X and Y, two probability measures μ and ν on X and Y respectively, and a Borel measurable cost function $c : X \times Y \to [0, \infty]$. Then the following holds.*

(**A**) *There exists a (L-almost surely) unique rectification c_r of c. Moreover*
  (**A1**) *one has L-almost surely $c_r \leq c$;*
  (**A2**) *if c is lower semi-continuous, then L-almost surely $c_r = c$;*
  (**A3**) *for the transport problem associated to c_r duality holds, in particular*

$$P_{c_r} = D_{c_r} = D_c\,.$$

(**B**) *The transport problem associated to c_r admits a solution (i.e., an optimal transport plan). Moreover*

(**B1**) *for any transport plan π and for any sequence of transport plans $\pi_n \rightharpoonup \pi$ one has*
$$\iint_{X\times Y} c_r\, d\pi \le \liminf_{n\to\infty} \iint_{X\times Y} c_r\, d\pi_n\,;$$
(**B2**) *for any transport plan π there is a suitable sequence of measures $\pi_n \rightharpoonup \pi$ so that*
$$\iint_{X\times Y} c_r\, d\pi = \lim_{n\to\infty} \iint_{X\times Y} c\, d\pi_n\,.$$
(**C**) *There exist Polish topologies τ_X, τ_Y on X resp. Y which refine the original topologies, lead to the same Borel sets and are so that c_r is lower semi-continuous w.r.t. $\tau_X \otimes \tau_Y$.*

Remark 1.5. *We underline that another way of "solving" the situations where the duality does not hold has been given in* [BLS09]*. For $\varepsilon > 0$, define the $1-\varepsilon$* partial transportation problem *(considered for instance in* [Fig10, CM10]*) by*
$$P_c^\varepsilon := \inf\left\{ \iint c\, d\pi : P_X\pi \le \mu, P_Y \le \nu, \|\pi\| \ge 1-\varepsilon \right\}.$$
Then [BLS09, Theorem 1.2] *asserts that*
$$D_c = P_c^{relaxed} := \lim_{\varepsilon\downarrow 0} P_c^\varepsilon\,.$$

2. Proof of the main result

In this section, we prove our theorem. We start with one of the main ingredients of the proof, namely, to show the existence of a rectification c_r corresponding to the cost function c. In fact, we can show something more precise.

Lemma 2.1. *There exists a unique rectification $c_r : X \times Y \to [0,\infty]$ of c. Moreover, there exist two sequences of measurable and bounded functions $\varphi_n : X \to \mathbb{R}$ and $\psi_n : Y \to \mathbb{R}$ such that $\varphi_n \oplus \psi_n \le c$ for all n, and*
$$c_r = \sup_{n\ge 1} \varphi_n \oplus \psi_n\,.$$

In the proof of this result, we will use the following characterization of L-negligible sets.

Lemma 2.2. *A Borel set $A \subseteq X \times Y$ is L-negligible if and only $\pi(A) = 0$ for every transport plan $\pi \in \Pi(\mu,\nu)$.*

Proof. If A is L-negligible, then clearly $\pi(A) = 0$ for every transport plan $\pi \in \Pi(\mu,\nu)$. The other direction is more difficult and was first established by Kellerer as a consequence of the Duality Theorem for bounded cost functions [Kel84, Proposition 3.5]. See also [BLS09, Appendix A] for a more direct proof. □

Proof of Lemma 2.1. As already noticed, the uniqueness of a rectification is trivial by property (ii) of Definition 1.4, hence we have only to show the existence. For simplicity, we will divide the proof of the lemma in some steps.

Step I. Reduction to the case of a bounded cost function c.
We start the proof by reducing to the case of a bounded function c. First of all, for any function τ and any $n \in \mathbb{N}$ let us set

$$\tau^{(n)} := \max\Big(\min(\tau, n), -n\Big).$$

It is now immediate to notice that, for any two functions $\varphi : X \to [-\infty, \infty)$ and $\psi : Y \to [-\infty, \infty)$, the sequence

$$n \mapsto \varphi^{(n)} \oplus \psi^{(n)}$$

is increasing where $\varphi \oplus \psi$ is positive, and it is pointwise converging to $\varphi \oplus \psi$. Hence, a function $d : X \times Y \to [0, \infty]$ satisfies L-almost surely the inequality $\varphi \oplus \psi \le d$ if and only if $\varphi^{(n)} \oplus \psi^{(n)} \le d$ for all $n \in \mathbb{N}$. As a consequence, if a function d satisfies property (i) of Definition 1.4 for all the pairs of functions which are bounded, then it already satisfies (i) in full generality. We are then ready to show the claim of this step. Indeed, let us assume that the lemma has been already established for all bounded cost functions, and pick a generic cost function c. By assumption, for any n we know that $c^{(n)}$ admits a rectification $c_r^{(n)} = \sup_{j \in \mathbb{N}} \varphi_{n,j} \oplus \psi_{n,j}$. We claim then that

$$c_r := \sup_{n,j} \varphi_{n,j} \oplus \psi_{n,j} = \sup_n c_r^{(n)}$$

is a rectification of c. Concerning property (i), for all Borel functions φ, ψ we have

$$\begin{aligned}\varphi \oplus \psi \le c \quad &\Longrightarrow \quad \varphi^{(n)} \oplus \psi^{(n)} \le c \quad \forall\, n \quad \Longrightarrow \quad \varphi^{(n)} \oplus \psi^{(n)} \le c^{(2n)} \quad \forall\, n \\ &\Longrightarrow \quad \varphi^{(n)} \oplus \psi^{(n)} \le c_r^{(2n)} \le c_r \quad \forall\, n \quad \Longrightarrow \quad \varphi \oplus \psi \le c_r\,.\end{aligned}$$

On the other hand, concerning property (ii), let $d : X \times Y \to [0, +\infty]$ satisfy (i), and let $n \in \mathbb{N}$. Since we assume the validity of the lemma for $c^{(n)}$, which is bounded, from the fact that

$$\varphi \oplus \psi \le c^{(n)} \quad \Longrightarrow \quad \varphi \oplus \psi \le c \quad \Longrightarrow \quad \varphi \oplus \psi \le d\,,$$

we immediately deduce that $c_r^{(n)} \le d$. Hence, clearly $c_r = \sup_n c_r^{(n)} \le d$.
Step II. The bounded case: definition of c_r and property (ii).
In view of Step I, let us now concentrate on the case of a bounded cost function c, say $c : X \times Y \to [0, M]$. Consider the set

$$V := \left\{ (f, g) :\ f : X \to [0,1],\ g : Y \to [0,1], \int f\, d\mu = \int g\, d\nu \right\},$$

and pick a family $\{(f_n, g_n)\}_{n \in \mathbb{N}} \subseteq V$ which is dense in V in the sense that for all $(f, g) \in V$ and $\varepsilon > 0$ there are f_n, g_n satisfying $\|f - f_n\|_1 + \|g - g_n\|_1 \le \varepsilon$. Here, and in the following, by the sake of shortness for each $h \in L^1_\mu(X)$ (resp. $k \in L^1_\nu(Y)$), we write $\|h\|_1$ (resp. $\|k\|_1$) to denote $\|h\|_{L^1_\mu(X)}$ (resp. $\|k\|_{L^1_\nu(Y)}$). For each $n \in \mathbb{N}$, let us now take a pair of functions (φ_n, ψ_n) which are optimal for the dual problem, hence such that $\varphi_n \oplus \psi_n \le c$ and

$$\int \varphi_n\, d(f_n \mu) + \int \psi_n\, d(g_n \nu) = \inf \left\{ \iint c\, d\gamma : P_X \gamma = f_n \nu, P_Y \gamma = g_n \mu \right\}. \tag{7}$$

This is possible thanks to the known duality results for bounded cost functions ([Kel84, Theorem 2.21]).

For technical reasons it will be convenient to take also the pair of functions $\varphi_0 \equiv 0, \psi_0 \equiv 0$ into account.

Let us now define $c_r := \sup_{n\geq 0} \varphi_n \oplus \psi_n$. The proof will be obtained by checking that c_r is a rectification of c. Let us start with the minimality property (ii), which is straightforward. Indeed, let d be a function which satisfies (i). For any $n \geq 0$, then, by construction we have $\varphi_n \oplus \psi_n \leq c$, and by (i) this implies $\varphi_n \oplus \psi_n \leq d$. Passing to the supremum, we obtain $c_r = \sup_n \varphi_n \oplus \psi_n \leq d$, then the required minimality property (ii).

Step III. The bounded case: proof of (i).

In view of the preceding steps, we still only have to check that the function c_r defined above verifies (i). Striving for a contradiction, we assume that there exist functions $\varphi : X \to \mathbb{R}, \psi : Y \to \mathbb{R}, \varphi \oplus \psi \leq c$ such that the set $\{\varphi \oplus \psi > c_r\}$ is not L-negligible.

Pick, by Lemma 2.2, a transport plan $\pi_0 \in \Pi(\mu, \nu)$ so that $\pi_0(\{\varphi \oplus \psi > c_r\}) > 0$. As

$$\{\varphi \oplus \psi > c_r\} = \bigcup_{a,b,\delta\in\mathbb{Q},\delta>0} \Big\{(x, y) : \varphi(x) > a, \psi(y) > b, a + b > c_r(x, y) + \delta\Big\},$$

there exist $a, b \in \mathbb{R}, \delta > 0$ and a Borel set $\Gamma \subseteq X \times Y$ so that

$$\begin{aligned} &\pi_0(\Gamma) > 0 && \\ &a < \varphi && \text{on } A := P_X\Gamma, \\ &b < \psi && \text{on } B := P_Y\Gamma, \\ &c_r < a + b - \delta && \text{on } \Gamma. \end{aligned}$$

Let now

$$\gamma_0 := \pi_0\restriction \Gamma\,, \qquad f := \frac{d(P_X\gamma_0)}{d\mu}\,, \qquad g := \frac{d(P_Y\gamma_0)}{d\nu}\,,$$

where the first definition means that for any Borel set Δ one has

$$\gamma_0(\Delta) = \pi_0\big(\Gamma \cap \Delta\big)\,.$$

Since $(f, g) \in V$, we can pick $n \geq 1$ so that f_n, g_n satisfy

$$\|f - f_n\|_1 + \|g - g_n\|_1 < \frac{\delta\|\gamma_0\|}{2M}\,. \tag{8}$$

Take now a plan $\gamma \in \Pi(f_n\mu, g_n\nu)$, and notice that

$$\begin{aligned} \gamma\big(A \times Y\big) &= f_n\mu(A) \geq f\mu(A) - \|f - f_n\|_1 = \gamma_0\big(A \times Y\big) - \|f - f_n\|_1 \\ &= \|\gamma_0\| - \|f - f_n\|_1\,, \end{aligned}$$

so that

$$\begin{aligned} \gamma\big(A \times B\big) &= \gamma\big(A \times Y\big) - \gamma\big(A \times (Y \setminus B)\big) \geq \gamma\big(A \times Y\big) - \gamma\big(X \times (Y \setminus B)\big) \\ &\geq \|\gamma_0\| - \|f - f_n\|_1 - g_n\nu(Y \setminus B) \\ &\geq \|\gamma_0\| - \|f - f_n\|_1 - \Big(g\nu(Y \setminus B) + \|g - g_n\|_1\Big) \\ &= \|\gamma_0\| - \|f - f_n\|_1 - \|g - g_n\|_1\,. \end{aligned}$$

As a consequence, recalling that $c \geq \varphi \oplus \psi > a+b$ on $A \times B$, we can estimate

$$
\begin{aligned}
\inf_{\gamma\in\Pi(f_n\mu,g_n\nu)} \iint_{X\times Y} c\,d\gamma &\geq \inf_{\gamma\in\Pi(f_n\mu,g_n\nu)} \iint_{A\times B} c\,d\gamma \\
&\geq (a+b)\Big(\|\gamma_0\| - \|f-f_n\|_1 - \|g-g_n\|_1\Big)\,.
\end{aligned} \tag{9}
$$

On the other hand, set

$$
\alpha := \frac{d(f_n\mu)}{d(f\mu)} \wedge 1\,, \qquad \beta := \frac{d(g_n\nu)}{d(g\nu)} \wedge 1\,, \qquad \tilde\gamma_0 = \big(\alpha\wedge\beta\big)\gamma_0 \leq \gamma_0\,,
$$

and notice that

$$
\begin{aligned}
\|\gamma_0\| - \|\tilde\gamma_0\| = \iint 1 - \big(\alpha\wedge\beta\big)\,d\gamma_0 \leq \iint 1-\alpha\,d\gamma_0 + \iint 1-\beta\,d\gamma_0 \\
\leq \|f-f_n\|_1 + \|g-g_n\|_1\,.
\end{aligned}
$$

We can then call

$$
\tilde f := \frac{d(P_X\tilde\gamma_0)}{d\mu}\,, \quad \tilde g := \frac{d(P_Y\tilde\gamma_0)}{d\nu}\,, \quad f_r := f_n - \tilde f \geq 0\,, \quad g_r := g_n - \tilde g \geq 0\,,
$$

observe that

$$
\|f_r\mu\| = \|g_r\nu\| \leq \|f-f_n\|_1 + \|g-g_n\|_1\,, \tag{10}
$$

thus getting to evaluate

$$
\begin{aligned}
&\int \varphi_n\,d(f_n\mu) + \int \psi_n\,d(g_n\nu) \\
&\quad= \int \varphi_n\,d(\tilde f\mu) + \int \psi_n\,d(\tilde g\nu) + \int \varphi_n\,d(f_r\mu) + \int \psi_n\,d(g_r\nu) \\
&\quad= \iint \varphi_n\oplus\psi_n\,d\tilde\gamma_0 + \iint \varphi_n\oplus\psi_n\,d\bigg(\frac{(f_r\mu)\otimes(g_r\nu)}{\|f_r\mu\|}\bigg) \\
&\quad\leq \iint c_r\,d\tilde\gamma_0 + M\|f_r\mu\| \leq \iint c_r\,d\gamma_0 + M\|f_r\mu\| \\
&\quad\leq \big(a+b-\delta\big)\|\gamma_0\| + M\Big(\|f-f_n\|_1 + \|g-g_n\|_1\Big)\,,
\end{aligned}
$$

where we have used (10) and the fact that $c_r \geq 0$, which immediately comes from the definition of φ_0 and ψ_0. Finally, inserting the last inequality and (9) into (7), and recalling that by construction $a+b \leq \sup c \leq M$, we readily obtain

$$
\|f-f_n\|_1 + \|g-g_n\|_1 \geq \frac{\delta\|\gamma_0\|}{2M}\,,
$$

which together with (8) provides the searched contradiction. □

We can now come to the proof of Theorem 1.

Proof of Theorem 1. Let us start from Property (**C**). Our argument will be based on [Kec95, Theorem 13.1]: if Z is a Polish space and $B_1, B_2, \dots$ are Borel sets in Z, then there exists a Polish topology τ on Z so that τ refines the original topology, τ generates the same Borel sets as the original topology and all sets $B_n, n \in \mathbb{N}$ are open in τ.

A useful application is that a Borel function can be viewed as continuous function on a modified space. More precisely, if $f : Z \to \mathbb{R}$ is Borel, then we can apply the above-mentioned result to the sets $B_n := f^{-1}(U_n), n \in \mathbb{N}$, being $\{U_n\}_{n\in\mathbb{N}}$ a neighborhood basis of $\mathbb{R}$, to obtain a Polish topology τ on Z which refines the original topological and in which all sets B_n are open. Consequently, f is a continuous function on the space (Z, τ).

By Lemma 2.1, there exist two sequences of measurable functions $\varphi_n : X \to \mathbb{R}$ and $\psi_n : Y \to \mathbb{R}$ such that $\sup_{n\geq 1} \varphi_n \oplus \psi_n = c_r$. Using the just explained argument, we can find topologies τ_X and τ_Y on X resp. Y so that, for every $n \in \mathbb{N}$, φ_n and ψ_n are continuous functions on (X, τ_X) resp. (Y, τ_Y). As a consequence, the functions $\varphi_n \oplus \psi_n, n \in \mathbb{N}$ are continuous on $(X \times Y, \tau_X \otimes \tau_Y)$ and hence $c_r = \sup_{n\geq 1} \varphi_n \oplus \psi_n$ is l.s.c. with respect to $\tau_X \otimes \tau_Y$, so that Property (**C**) follows.

Let us now consider Property (**A**). The existence and uniqueness of a rectification have been already established with Lemma 2.1. Concerning Property (**A1**), it is clear from the definition of the rectification.

We pass then to consider Property (**A2**). Pick families of open sets $\{U_n\}_{n\in\mathbb{N}}$ and $\{V_m\}_{m\in\mathbb{N}}$ which form bases of the topologies of X resp. Y. For $n, m \in \mathbb{N}$, set

$$e_{n,m} := \inf_{(x,y)\in U_n\times V_m} c(x, y) \leq +\infty$$

and define $\varphi_{n,m,j} : X \to \mathbb{R}$ and $\psi_{n,m,j} : Y \to \mathbb{R}$ so that

$$\begin{aligned} &\varphi_{n,m,j} \oplus \psi_{n,m,j} = e_{n,m} \wedge j && \text{on } U_n \times V_m \\ &\varphi_{n,m,j} \oplus \psi_{n,m,j} \leq 0 && \text{otherwise.} \end{aligned}$$

Since by construction $\varphi_{n,m,j} \oplus \psi_{n,m,j} \leq c$, by definition we have $\varphi_{n,m,j} \oplus \psi_{n,m,j} \leq c_r$, and hence also

$$c_r \geq \sup_{n,m,j} \varphi_{n,m,j} \oplus \psi_{n,m,j} \,.$$

Finally, if c is l.s.c. then the latter supremum coincides with c itself, so L-a.s. one has $c_r \geq c$, which together with Property (**A1**) concludes the searched equality.

Finally, we consider Property (**A3**). First of all, we can observe that c and c_r have the same dual problem, that is, $D_c = D_{c_r}$. To do so, take two functions φ, ψ, integrable with respect to μ and ν respectively, and such that $\varphi \oplus \psi \leq c$. By definition, there exist sets $M \subseteq X, N \subseteq Y, \mu(M) = \nu(N) = 0$ so that $(\varphi \oplus \psi)(x, y) \leq c_r(x, y)$ for all $x \in X \setminus M, y \in Y \setminus N$, hence for $\tilde{\varphi} := \varphi - I_M, \tilde{\psi} := \psi - I_N$ we have $\tilde{\varphi} \oplus \tilde{\psi} \leq c_r$ and $\int \varphi \, d\mu = \int \tilde{\varphi} \, d\mu, \int \psi \, d\nu = \int \tilde{\psi} \, d\nu$. This shows $D_{c_r} \geq D_c$. The other inequality is identical. Indeed, if φ, ψ are integrable and $\varphi \oplus \psi \leq c_r$, by Property (**A1**) there are again two sets $M \subseteq X, N \subseteq Y, \mu(M) = \nu(N) = 0$ so that $(\varphi \oplus \psi)(x, y) \leq c_r(x, y) \leq c(x, y)$ for all $x \in X \setminus M, y \in Y \setminus N$, so exactly as before we get $D_c \geq D_{c_r}$, and in particular we have $D_c = D_{c_r}$.

Moreover, having already established Property (**C**), the equality $D_{c_r} = P_{c_r}$ comes directly from the standard duality theorem for l.s.c. cost functions (notice that a change of the topology which does not change the Borel sets does not effect neither the primal nor the dual problem).

We are then finally left with Property (**B**). First of all, the existence of an optimal transport plan with respect to c_r is obvious by Property (**C**).

Indeed, it is well-known that a transport problem with a l.s.c. cost admits an optimal transport plan, and the optimality of a plan is again not effected, of course, by a change of the topology.

Let us now consider Property (**B1**). To show its validity, we first recall the very well-known fact that, whenever d is a l.s.c. function and γ_n is a sequence of measures weakly converging to γ, one has

$$\int d\, d\gamma \leq \liminf_{n\to\infty} \int d\, d\gamma_n\,.$$

Let us now take a transport plan π and a sequence $\{\pi_n\}$ such that $\pi_n \rightharpoonup \pi$. Thanks to Lemma 2.3 below (and to the discussion of the following Remark 2.4), we obtain that the weak convergence $\pi_n \rightharpoonup \pi$ holds also with respect to the new topology $\tau_X \otimes \tau_Y$. But then, since c_r is l.s.c. with respect to this topology, we immediately get the searched liminf inequality.

Finally, we are left with Property (**B2**). Keeping in mind Property (**B1**), it is sufficient to show that for every $\pi \in \Pi(\mu,\nu)$ there exists a sequence $\pi_n \rightharpoonup \pi$ so that

$$\limsup_{n\to\infty} \iint c\, d\pi_n \leq \iint c_r\, d\pi\,. \tag{11}$$

We present the proof for the case $X = Y = [0,1]$ and $\mu = \nu = \lambda$, because the argument becomes much simpler to read, but at the end it will be clear that the proof of the general case is equivalent, and just more notationally uncomfortable.

If $\iint c_r\, d\pi = \infty$ there is nothing to prove, so assume that $\iint c_r\, d\pi < \infty$. Fix $n \in \mathbb{N}$ and $l, m \in \{1, \dots, 2^n\}$. Set

$$D^n_{l,m} := \big(\tfrac{l-1}{2^n}, \tfrac{l}{2^n}\big] \times \big(\tfrac{m-1}{2^n}, \tfrac{m}{2^n}\big]\,.$$

Denote by $\mu^n_{l,m}, \nu^n_{l,m}$ the marginals of $\pi \restriction D^n_{l,m}$. For $\mu^n_{l,m}$-, resp. $\nu^n_{l,m}$-integrable functions $\varphi : \big(\frac{l-1}{2^n}, \frac{l}{2^n}\big] \to \mathbb{R}, \psi : \big(\frac{m-1}{2^n}, \frac{m}{2^n}\big] \to \mathbb{R}$ satisfying $\varphi \oplus \psi \leq c$ we have

$$\int \varphi\, d\mu^n_{l,m} + \int \psi\, d\nu^n_{l,m} \leq \iint_{D^n_{l,m}} c_r\, d\pi < \infty\,.$$

Hence the optimal dual value corresponding to the cost function c and the spaces $\Big(\big(\frac{l-1}{2^n}, \frac{l}{2^n}\big], \mu^n_{l,m}\Big)$, $\Big(\big(\frac{m-1}{2^n}, \frac{m}{2^n}\big], \nu^n_{l,m}\Big)$ is finite. By Remark 1.5 (resp. [BLS09, Theorem 1.2]), there exist $\mu^n_{l,m}$-, resp. $\nu^n_{l,m}$-integrable functions $\varphi^n_{l,m} : \big(\frac{l-1}{2^n}, \frac{l}{2^n}\big] \to \mathbb{R}, \psi^n_{l,m} : \big(\frac{m-1}{2^n}, \frac{m}{2^n}\big] \to \mathbb{R}$ and a measure $\pi^n_{l,m}$ on $D^n_{l,m}$ so that

$$P_X\pi^n_{l,m} \leq \mu^n_{l,m}\,, \quad P_Y\pi^n_{l,m} \leq \nu^n_{l,m}\,, \quad \|\pi^n_{l,m}\| \geq \|\mu^n_{l,m}\| - \tfrac{1}{n4^n}\,, \tag{12}$$

$$\varphi^n_{l,m} \oplus \psi^n_{l,m} \leq c\,, \tag{13}$$

$$\int \varphi^n_{l,m}\, d\mu^n_{l,m} + \int \psi^n_{l,m}\, d\nu^n_{l,m} \geq \iint c\, d\pi^n_{l,m} - \tfrac{1}{n4^n}\,. \tag{14}$$

Define a measure π_n on $X \times Y$ by the requiring that $\pi_n \restriction D^n_{l,m} = \pi^n_{l,m}$ for all $l, m \in \{1, \dots, 2^n\}$. It follows from (12) that

$$\lim_{k\to\infty} \pi_k(D^n_{l,m}) = \pi(D^n_{l,m})$$

for all $l,m,n \in \mathbb{N}, l,m \leq 2^n$. Consequently, $(\pi_k)_{k\geq 1}$ converges weakly to π. From (13) and (14) we deduce

$$\begin{aligned}\iint c\, d\pi_n - \tfrac{1}{n} &= \sum_{l,m\leq 2^n} \left(\iint c\, d\pi^n_{l,m} - \tfrac{1}{n4^n}\right)\\ &\leq \sum_{l,m\leq 2^n} \int \varphi^n_{l,m}\, d\mu^n_{l,m} + \int \psi^n_{l,m}\, d\nu^n_{l,m}\\ &= \sum_{l,m\leq 2^n} \iint_{D^n_{l,m}} \varphi^n_{l,m} \oplus \psi^n_{l,m}\, d\pi\\ &\leq \sum_{l,m\leq 2^n} \iint_{D^n_{l,m}} c_r\, d\pi = \iint c_r\, d\pi\,.\end{aligned}$$

Letting n tend to ∞ in the last inequality, we obtain (11). $\square$

In the above proof, we needed to use the following technical topological result.

Lemma 2.3. *Let X, Y be Polish spaces, let μ, ν be Borel probability measures on X, Y respectively, and take a measure π and measures π_n, $n \in \mathbb{N}$, on $X \times Y$ such that*

$$P_X\pi \leq \mu\,, \quad P_X\pi_n \leq \mu\,, \qquad P_Y\pi \leq \nu\,, \quad P_Y\pi_n \leq \nu\,. \tag{15}$$

Then, π_n weakly converges to π if and only if $\pi_n(A\times B) \to \pi(A\times B)$ for all Borel sets $A \subseteq X$, $B \subseteq Y$.

Remark 2.4. *It is important to underline a consequence of the above result, namely, that the weak convergence of transport plans only depends on the Borel structure of X and Y, instead of the topological one. In other words, replacing the topologies on X and Y with other topologies generating the same Borel structures, the weak convergence of the sequences of plans is left unchanged.*

Proof of Lemma 2.3. Suppose first that $\pi_n \rightharpoonup \pi$, take two Borel sets $A \subseteq X$ and $B \subseteq Y$, and fix $\varepsilon > 0$. It is then possible to select two open sets $\widetilde{A} \supseteq A$ and $\widetilde{B} \supseteq B$, as well as two compact sets $\widehat{A} \subseteq A$ and $\widehat{B} \subseteq B$, in such a way that

$$\mu\Big(\widetilde{A} \setminus \widehat{A}\Big) < \varepsilon\,, \qquad \nu\Big(\widetilde{B} \setminus \widehat{B}\Big) < \varepsilon\,.$$

By the standard semi-continuity properties of the weak convergence, and recalling (15), one then has

$$\begin{aligned}\liminf_{n\to\infty} \pi_n\big(\widetilde{A}\times\widetilde{B}\big) &\geq \pi\big(\widetilde{A}\times\widetilde{B}\big) \geq \pi\big(A\times B\big) \geq \pi\big(\widehat{A}\times\widehat{B}\big)\\ &\geq \limsup_{n\to\infty} \pi_n\big(\widehat{A}\times\widehat{B}\big) \geq \limsup_{n\to\infty} \pi_n\big(\widetilde{A}\times\widetilde{B}\big) - 2\varepsilon\,,\end{aligned}$$

which clearly implies $\pi_n\big(A\times B\big) \longrightarrow \pi\big(A\times B\big)$.

On the other side, assuming the convergence of $\pi_n\big(A\times B\big)$ to $\pi\big(A\times B\big)$ for all Borel sets $A \subseteq X$, $B \subseteq Y$, we have to prove that $\pi_n \rightharpoonup \pi$. Notice that the sequence $\{\pi_n\}$ is relatively sequentially weakly compact, by (15) and by Prokhorov Theorem (this is standard, see for instance [Vil09, p55-57]).

Hence, one can find a measure $\tilde{\pi}$ and extract a (not relabelled) subsequence such that $\pi_n \rightharpoonup \tilde{\pi}$. Of course, we are done once we check that $\tilde{\pi} = \pi$. But in fact, by assumption and by the first half of the proof we know that

$$\tilde{\pi}\big(A \times B\big) \xleftarrow[n\to\infty]{} \pi_n\big(A \times B\big) \xrightarrow[n\to\infty]{} \pi\big(A \times B\big)$$

for all Borel sets $A \subseteq X$, $B \subseteq Y$. Hence, π and $\tilde{\pi}$ agree on all the Borel rectangles $A \times B$, and since Borel rectangles are a basis for the Borel sets of $X \times Y$, we deduce that $\tilde{\pi} = \pi$. □

3. Remarks and Examples

In this last section, we collect some examples and remarks about the rectification. First of all, we show that the rectification of a cost function c is not independent of μ and ν, but in fact it depends only on which are the negligible sets with respect to μ and ν. As already discussed in the introduction, this is the strongest possible result in this sense.

Remark 3.1. *Given a cost function $c : X \times Y \to [0, +\infty]$, the rectification c_r of c is not independent of μ and ν. However, it only depends on the class of the μ- (resp. ν-) negligible sets in X (resp. Y). This is immediate from Definitions 1.4 and 1.3, since everything depends on which sets are L-negligible, and in turn this only depends on the μ- and ν-negligible sets.*

A second observation is needed, concerning the lower semi-continuity properties of c_r. In fact, we have already seen, with Example 1.2 in the introduction, that c_r is not the l.s.c. envelope of c. However, that example may leave the impression that c_r can still be defined as the l.s.c. envelope of c, made after a modification of c on an L-negligible set. This is indeed true in Example 1.2, as well as in many other situations. We can show with the example below that this is not in general the case.

Example 3.2. *Let $(X, \mu) = (Y, \nu) = ([0, 1], \lambda)$. Let $(q_n)_{n\geq 1}$ be an enumeration of the rationals in $[0, 1]$. Pick α so that*

$$\Gamma := [0, 1] \setminus \bigcup_{n\geq 1} (q_n - \alpha/2^n, q_n + \alpha/2^n)$$

has Lebesgue-measure $1/2$. Set $\varphi = I_\Gamma, \psi \equiv 0$ and $c \equiv \varphi \oplus \psi$. Then $P_c = D_c = 1/2$, and $c_r = c$. Nevertheless, any lower semi-continuous function $g : X \times Y \to [0, \infty]$ which is L-almost surely smaller than c necessarily satisfies $g \leq 0$ L-almost surely.

We can now discuss a little bit the situation concerning the boundedness of the cost function c. In fact, as we already mentioned in the introduction, if c is bounded then the duality $P_c = D_c$ already holds true, so one could think that $c_r = c$ whenever c is bounded. We already know that this is not true, thanks to the variant of the Example 1.1 where $+\infty$ is replaced by $M > 1$. But as we discussed, in that situation c had the drawback that no optimal transport plans existed, while there are always optimal transport plans for c_r. It is then interesting to see an example where c is bounded and there are optimal transport plans for c, but still c_r does not coincide L-almost surely with c.

Example 3.3. *Let us again consider the setting where $X = Y = [0,1]$ and $\mu = \nu$ is the Lebesgue measure, and consider the cost function given by*

$$c(x,y) := \begin{cases} 0 & \text{if } x - y \in \mathbb{Q}\,, \\ 1 & \text{otherwise}\,. \end{cases}$$

Then, it is very easy to see that $P_c = D_c = 0$, and the identity is an optimal transport plan. However, it is also simple to observe that, whenever $\varphi \oplus \psi \leq c$, then in fact $\varphi \oplus \psi \leq 0$, and this readily yields that $c_r \equiv 0$, so c_r does not coincide with c L-almost surely.

The above examples show in particular that Property (**A**) of Theorem 1 is not enough to characterize the rectification c_r. We can show, with another example, that also Property (**B**) does not characterize the rectification (this means that, if a function $\tilde{c}_r$ verifies Property (**B**), then not necessarily $\tilde{c}_r$ coincides with c_r).

Example 3.4. *Let us consider another variant of Example 1.1, with the cost function c given by*

$$c(x,y) = \begin{cases} 2, & \text{for } 0 \leq y < x \leq 1\,, \\ 1, & \text{for } 0 \leq x = y \leq 1\,, \\ \infty, & \text{for } 0 \leq x < y \leq 1\,. \end{cases}$$

*We know that $c_r = c$, since c is l.s.c.. However, let us define $\tilde{c}_r$ replacing 2 by 0 under the diagonal (or, more in general, define $\tilde{c}_r$ to be 1 on the diagonal, $+\infty$ above it, and any other measurable function below it). Then, since all the transport plans have an infinite cost except for the identity, which has cost 1, it is immediate to observe that the function $\tilde{c}_r$ verifies Property (**B**), though it is false that $\tilde{c}_r = c_r$ L-a.s..*

We conclude by introducing the following stronger variant (**B'**) of Property (**B**), and showing that in fact (**B'**) uniquely characterizes the rectification.

(**B'**) for any measure π on $X \times Y$, $P_X\pi \leq \mu, P_Y\pi \leq \nu$ and any sequence $\pi_n \rightharpoonup \pi$, $P_X\pi_n \leq \mu, P_Y\pi_n \leq \nu$ one has

$$\int_{X\times Y} c_r\,d\pi \leq \liminf_{n\to\infty} \int_{X\times Y} c_r\,d\pi_n \leq \liminf_{n\to\infty} \int_{X\times Y} c\,d\pi_n\,,$$

and moreover for any such measure π there is a suitable sequence $\pi_n \rightharpoonup \pi$ such that the above inequalities are equalities.

The only difference between Properties (**B**) and (**B'**) is that in the first case one only considers transport plans π, hence $P_X\pi = \mu$, $P_Y\pi = \nu$, while in the latter one considers the more general case when $P_X\pi \leq \mu$ and $P_Y\pi \leq \nu$. This could seem a slight difference, but on the contrary it makes Property (**B'**) strong enough to characterize the rectification, as the next remark underlines.

Remark 3.5. *It is pretty simple to realize, from the proof of Property (**B**) in Theorem 1, that also the stronger Property (**B'**) holds true. On the other hand, given any function $\tilde{c}_r$ which satisfies the Property (**B'**), we claim that $\tilde{c}_r$ coincides with c_r L-almost surely.*

*To see this, take any measure π on $X \times Y$ satisfying $P_X\pi \le \mu, P_Y\pi \le \nu$, and observe that Property (**B'**) implies that*

$$\iint c_r \, d\pi = \iint \tilde{c}_r \, d\pi \, .$$

It follows that $\pi(\{c_r < \tilde{c}_r\}) = \pi(\{c_r > \tilde{c}_r\}) = 0$ for any such measure π, so in particular for any $\pi \in \Pi(\mu,\nu)$. Our claim then is immediately obtained thanks to Lemma 2.2.

*In particular, it is easy to notice that the function $\tilde{c}_r$ of Example 3.4 verifies Property (**B**) but not Property (**B'**).*

References

[BLS09] M. Beiglböck, C. Léonard, and W. Schachermayer. A general duality theorem for the Monge-Kantorovich transport problem. *submitted*, 2009.

[BS09] M. Beiglböck and W. Schachermayer. Duality for Borel measurable cost functions. *Trans. Amer. Math. Soc.*, to appear, 2009.

[CM10] L. A. Caffarelli and R. J. McCann. Free boundaries in optimal transport and Monge-Ampère obstacle problems. *Ann. of Math. (2)*, 171(2):673–730, 2010.

[dA82] A. de Acosta. Invariance principles in probability for triangular arrays of B-valued random vectors and some applications. *Ann. Probab.*, 10(2):346–373, 1982.

[Dud76] R. M. Dudley. *Probabilities and metrics.* Matematisk Institut, Aarhus Universitet, Aarhus, 1976. Convergence of laws on metric spaces, with a view to statistical testing, Lecture Notes Series, No. 45.

[Dud02] R. M. Dudley. *Real analysis and probability*, volume 74 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2002. Revised reprint of the 1989 original.

[Fer81] X. Fernique. Sur le théorème de Kantorovich-Rubinstein dans les espaces polonais. In *Seminar on Probability, XV (Univ. Strasbourg, Strasbourg, 1979/1980) (French)*, volume 850 of *Lecture Notes in Math.*, pages 6–10. Springer, Berlin, 1981.

[Fig10] A. Figalli. The optimal partial transport problem. *Arch. Ration. Mech. Anal.*, 195(2):533–560, 2010.

[GR81] N. Gaffke and L. Rüschendorf. On a class of extremal problems in statistics. *Math. Operationsforsch. Statist. Ser. Optim.*, 12(1):123–135, 1981.

[Kan42] L. Kantorovich. On the translocation of masses. *C. R. (Doklady) Acad. Sci. URSS (N.S.)*, 37:199–201, 1942.

[Kec95] A. S. Kechris. *Classical descriptive set theory*, volume 156 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1995.

[Kel84] H. G. Kellerer. Duality theorems for marginal problems. *Z. Wahrsch. Verw. Gebiete*, 67(4):399–432, 1984.

[KR58] L. V. Kantorovič and G. Š. Rubinšteĭn. On a space of completely additive functions. *Vestnik Leningrad. Univ.*, 13(7):52–59, 1958.

[Mik06] T. Mikami. A simple proof of duality theorem for Monge-Kantorovich problem. *Kodai Math. J.*, 29(1):1–4, 2006.

[MT06] T. Mikami and M. Thieullen. Duality theorem for the stochastic optimal control problem. *Stochastic Process. Appl.*, 116(12):1815–1835, 2006.

[Szu82] A. Szulga. On minimal metrics in the space of random variables. *Teor. Veroyatnost. i Primenen.*, 27(2):401–405, 1982.

[Vil03] C. Villani. *Topics in optimal transportation*, volume 58 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2003.

[Vil09] C. Villani. *Optimal Transport. Old and New*, volume 338 of *Grundlehren der mathematischen Wissenschaften*. Springer, 2009.

Fakultät für Mathematik, Universität Wien, Nordbergstrasse 15, A-1090 Wien (Austria)
E-mail address: mathias.beiglboeck@gmail.com

Dipartimento di Matematica, via Ferrata 1, 27100 Pavia (Italy)
E-mail address: aldo.pratelli@unipv.it