A NOTE ON DUALITY THEOREMS IN MASS TRANSPORTATION

PIETRO RIGO

Abstract. The duality theory of the Monge-Kantorovich transport problem is investigated in an abstract measure theoretic framework. Let $(\mathcal{X}, \mathcal{F}, \mu)$ and $(\mathcal{Y}, \mathcal{G}, \nu)$ be any probability spaces and $c : \mathcal{X} \times \mathcal{Y} \rightarrow \mathbb{R}$ a measurable cost function such that $f_1 + g_1 \leq c \leq f_2 + g_2$ for some $f_1, f_2 \in L_1(\mu)$ and $g_1, g_2 \in L_1(\nu)$. Define $\alpha(c) = \inf_P \int c\,dP$ and $\alpha^*(c) = \sup_P \int c\,dP$, where inf and sup are over the probabilities P on $\mathcal{F} \otimes \mathcal{G}$ with marginals μ and ν. Some duality theorems for $\alpha(c)$ and $\alpha^*(c)$, not requiring μ or ν to be perfect, are proved. As an example, suppose $\mathcal{X}$ and $\mathcal{Y}$ are metric spaces and μ is separable. Then, duality holds for $\alpha(c)$ (for $\alpha^*(c)$) provided c is upper-semicontinuous (lower-semicontinuous). Moreover, duality holds for both $\alpha(c)$ and $\alpha^*(c)$ if the maps $x \mapsto c(x, y)$ and $y \mapsto c(x, y)$ are continuous, or if c is bounded and $x \mapsto c(x, y)$ is continuous. This improves the existing results in [14] if c satisfies the quoted conditions and the cardinalities of $\mathcal{X}$ and $\mathcal{Y}$ do not exceed the continuum.

1. Introduction

Throughout, $(\mathcal{X}, \mathcal{F}, \mu)$ and $(\mathcal{Y}, \mathcal{G}, \nu)$ are probability spaces and

$$\mathcal{H} = \mathcal{F} \otimes \mathcal{G}$$

is the product σ-field on $\mathcal{X} \times \mathcal{Y}$. Further, $\Gamma(\mu, \nu)$ is the collection of probability measures P on $\mathcal{H}$ with marginals μ and ν, namely,

$$P(A \times \mathcal{Y}) = \mu(A) \quad \text{and} \quad P(\mathcal{X} \times B) = \nu(B) \quad \text{for all } A \in \mathcal{F} \text{ and } B \in \mathcal{G}.$$

For any probability space $(\Omega, \mathcal{A}, Q)$, we write $L_1(Q)$ to denote the class of $\mathcal{A}$-measurable and Q-integrable functions $\phi : \Omega \rightarrow \mathbb{R}$ (without identifying maps which agree Q-a.s.). We also write $Q(\phi) = \int \phi\,dQ$ for $\phi \in L_1(Q)$.

With a slight abuse of notation, for any maps $f : \mathcal{X} \rightarrow \mathbb{R}$ and $g : \mathcal{Y} \rightarrow \mathbb{R}$, we still denote by f and g the functions on $\mathcal{X} \times \mathcal{Y}$ given by $(x, y) \mapsto f(x)$ and $(x, y) \mapsto g(y)$. Thus, $f + g$ is the map on $\mathcal{X} \times \mathcal{Y}$ defined as

$$(f + g)(x, y) = f(x) + g(y) \quad \text{for all } (x, y) \in \mathcal{X} \times \mathcal{Y}.$$

In this notation, we let

$$L = \{f + g : f \in L_1(\mu),\, g \in L_1(\nu)\}.$$

Let $c : \mathcal{X} \times \mathcal{Y} \rightarrow \mathbb{R}$ be an $\mathcal{H}$-measurable function satisfying

$$(1) \qquad f_1 + g_1 \leq c \leq f_2 + g_2 \qquad \text{for some } f_1 + g_1 \in L \text{ and } f_2 + g_2 \in L.$$

2010 *Mathematics Subject Classification.* 60A10, 60E05, 28A35.

Key words and phrases. Duality theorem, Mass transportation, Perfect probability measure, Probability measure with given marginals, Separable probability measure.

For such a c, we define

$$\alpha(c) = \inf\big\{P(c) :\ P \in \Gamma(\mu,\nu)\big\},$$
$$\alpha^*(c) = \sup\big\{P(c) :\ P \in \Gamma(\mu,\nu)\big\},$$
$$\beta(c) = \sup\big\{\mu(f)+\nu(g) :\ f+g \in L,\ f+g \le c\big\},$$
$$\beta^*(c) = \inf\big\{\mu(f)+\nu(g) :\ f+g \in L,\ f+g \ge c\big\}.$$

It is not hard to see that

$$\beta(c) \le \alpha(c) \le \alpha^*(c) \le \beta^*(c).$$

A duality theorem (for both $\alpha(c)$ and $\alpha^*(c)$) is the assertion that

$$\alpha(c) = \beta(c) \quad \text{and} \quad \alpha^*(c) = \beta^*(c). \tag{2}$$

Indeed, duality theorems arise in a plenty of frameworks. The main one is possibly mass transportation, where $c(x, y)$ is regarded as the cost for moving a unit of good from $x \in \mathcal{X}$ into $y \in \mathcal{Y}$. However, duality results play a role even in risk theory, optimization problems and dependence modeling. See e.g. [1], [3], [4], [5], [6], [9], [11], [12], [13], [16], [17] and references therein.

Starting from Kantorovich himself [8], there is a long line of research on duality theorems; see again [3], [4], [17] and references therein. To our knowledge, *under the present assumptions on c*, the best result is due to Ramachandran and Ruschendorf [14]. According to the latter, one obtains both $\alpha(c) = \beta(c)$ and $\alpha^*(c) = \beta^*(c)$ provided c is $\mathcal{H}$-measurable, it satisfies condition (1), and at least one between μ and ν is perfect.

Now, some form of condition (1) can not be dispensed while removing measurability leads to involve inner and outer measures; see [9, Section 2]. Instead, whether the perfectness assumption can be dropped is still an *open problem.* Thus, if c is measurable and meets (1) but μ and ν are both non-perfect, it is currently unknown whether condition (2) is true or false. See points (2)-(3), page 355, of [15].

This paper provides duality theorems not requiring perfectness.

Suppose $\mathcal{X}$ and $\mathcal{Y}$ are metric spaces and $\mathcal{F}$ and $\mathcal{G}$ the Borel σ-fields. Then, condition (2) is shown to be true if at least one of μ and ν is separable, c meets (1) and all the c-sections are continuous. Or else, condition (2) holds if μ and ν are both separable, c is bounded and measurable, and at least one of the c-sections is continuous. These results improve [14] when c satisfies the quoted assumptions and the cardinalities of $\mathcal{X}$ and $\mathcal{Y}$ do not exceed the continuum. Under the latter condition, in fact, a perfect probability measure is separable but not conversely. Note also that, if $\mathcal{X}$ and $\mathcal{Y}$ are separable metric spaces (so that separability of μ and ν is automatic) the scope of our results is to replace assumptions on μ or ν (required by [14]) with assumptions on c.

Various conditions for $\alpha(c) = \beta(c)$ *or* $\alpha^*(c) = \beta^*(c)$, but not necessarily for both, are given as well. For instance, if c meets (1) and at least one of μ and ν is separable, then $\alpha^*(c) = \beta^*(c)$ or $\alpha(c) = \beta(c)$ provided c is lower or upper semicontinuous. As another example, $\alpha^*(1_H) = \beta^*(1_H)$ if $H = \cup_n (A_n \times B_n)$ with $A_n \in \mathcal{F}$ and $B_n \in \mathcal{G}$. Further, $\alpha(1_H) = \beta(1_H)$ if $\mu(\limsup_n A_n) = 0$ or $\nu(\limsup_n B_n) = 0$. Without some extra condition, however, we do not know whether $\alpha(1_H) = \beta(1_H)$.

2. Preliminaries

For any topological space S, the Borel σ-field on S is denoted by $\mathcal{B}(S)$.

Let $(\Omega, \mathcal{A}, Q)$ be a probability space. Then, Q is *perfect* if, for any $\mathcal{A}$-measurable $\phi : \Omega \to \mathbb{R}$, there is $B \in \mathcal{B}(\mathbb{R})$ such that $B \subset \phi(\Omega)$ and $Q(\phi \in B) = 1$.

An important special case is Ω a metric space and $\mathcal{A} = \mathcal{B}(\Omega)$. In that case, Q is *separable* if $Q(A) = 1$ for some separable $A \in \mathcal{A}$ and Q is *tight* if $Q(A) = 1$ for some σ-compact $A \in \mathcal{A}$. Clearly, tightness implies separability but not conversely. Furthermore, tightness is equivalent to perfectness provided Ω satisfies the following condition:

> The power set of Ω does not support any 0-1-valued probability measure T such that $T\{\omega\} = 0$ for each $\omega \in \Omega$;

see [10, Theorem 3.2].

Two remarks are in order. First, the above condition on Ω is automatically true if $\text{card}(\Omega) \leq \text{card}(\mathbb{R})$. Thus, perfectness implies separability, but not conversely, if $\text{card}(\Omega) \leq \text{card}(\mathbb{R})$ (in particular, if Ω is a separable metric space). Second, it is consistent with the usual axioms of set theory (ZFC) that, for any metric space Ω, any probability measure on $\mathcal{B}(\Omega)$ is separable.

Note also that a simple example of non perfect probability measure is any non tight probability measure on the Borel sets of a separable metric space. For instance, take Q the outer Lebesgue measure on $\mathcal{B}(\Omega)$, where Ω is a subset of $[0, 1]$ with outer Lebesgue measure 1 and inner Lebesgue measure 0. Then, Q is not perfect.

Let us come back to duality theorems. Define

$$M = \{\mathcal{H}\text{-measurable functions } c : \mathcal{X} \times \mathcal{Y} \to \mathbb{R} \text{ satisfying condition (1)}\}$$

and note that

$$\alpha^*(c) = -\alpha(-c) \quad \text{and} \quad \beta^*(c) = -\beta(-c) \quad \text{for all } c \in M.$$

Thus, to get condition (2), it suffices to show $\alpha(c) = \beta(c)$ under some conditions which hold true for both c and $-c$.

Two preliminary lemmas are needed. The first is inspired to [7, Lemma 1].

Lemma 1. *Let $c \in M$. Then, $\beta^*(c) = \lim_n \beta^*(c_n)$ whenever $(c_n) \subset M$ is an increasing sequence such that $c_n \uparrow c$ pointwise.*

Proof. We first suppose $0 \leq c_n \leq c \leq k$ for some integer k. Under this assumption, for each n, there is $f_n + g_n \in L$ such that

$$f_n + g_n \geq c_n, \quad 0 \leq f_n, \, g_n \leq k, \quad \mu(f_n) + \nu(g_n) < \beta^*(c_n) + 1/n;$$

see e.g. [9, Lemma 1.8].

Since the sequences (f_n) and (g_n) are uniformly bounded, there are $f \in L_1(\mu)$, $g \in L_1(\nu)$ and a subsequence (m_n) such that

$$f_{m_n} \to f \text{ weakly in } L_1(\mu) \quad \text{and} \quad g_{m_n} \to g \text{ weakly in } L_1(\nu).$$

In turn, this implies the existence of a sequence (ϕ_n, ψ_n) such that $\phi_n \to f$ in $L_1(\mu)$, $\psi_n \to g$ in $L_1(\nu)$ and (ϕ_n, ψ_n) is a convex combination of $\{(f_{m_j}, g_{m_j}) : j \geq n\}$ for each n. By taking a further subsequence, it can be also assumed $\mu(\phi_n \to f) = \nu(\psi_n \to g) = 1$. Since (c_n) is increasing, $\phi_n + \psi_n \geq c_{m_n}$. Hence, after modifying f and g on null sets, one obtains $f + g \geq c$. On noting that $(\beta^*(c_n))$ is a monotone sequence, it follows that

$$\mu(f) + \nu(g) \geq \beta^*(c) \geq \lim_n \beta^*(c_n) = \lim_n \beta^*(c_{m_n})$$
$$= \lim_n \{\mu(f_{m_n}) + \nu(g_{m_n})\} = \mu(f) + \nu(g).$$

This concludes the proof if $0 \leq c_n \leq c \leq k$. To deal with the general case, fix $p+q \in L$ such that $p+q \leq c_1$ and define $b_n = c_n - (p+q)$ and $b = c - (p+q)$. Then, $0 \leq b_n \leq b$. Further, since $\beta^*(h+p+q) = \beta^*(h) + \mu(p) + \nu(q)$ for each $h \in M$, it suffices to show that $\beta^*(b) = \lim_n \beta^*(b_n)$.

Given k, take $f_k + g_k \in L$ such that

$$f_k + g_k \geq b \wedge 2k \quad \text{and} \quad \mu(f_k) + \nu(g_k) < \beta^*\big(b \wedge 2k\big) + 1/k.$$

Take also $f+g \in L$ such that $f+g \geq b$ and note that

$$f\,1_{\{g>k\}} = f\,1_{\{f\leq k,g>k\}} + f\,1_{\{f>k,g>k\}} \leq g\,1_{\{g>k\}} + f\,1_{\{f>k\}}.$$

Similarly, $g\,1_{\{f>k\}} \leq g\,1_{\{g>k\}} + f\,1_{\{f>k\}}$. Hence,

$$\begin{aligned} b &\leq b\,1_{\{b\leq 2k\}} + (f+g)\,1_{\{f+g>2k\}} \\ &\leq f_k + g_k + (f+g)\big(1_{\{f>k\}} + 1_{\{g>k\}}\big) \\ &\leq f_k + g_k + 3f\,1_{\{f>k\}} + 3g\,1_{\{g>k\}}. \end{aligned}$$

Since $f_k + g_k + 3f\,1_{\{f>k\}} + 3g\,1_{\{g>k\}}$ belongs to L, it follows that

$$\begin{aligned} \beta^*(b) &\leq \mu(f_k) + \nu(g_k) + 3\mu\big[f\,1_{\{f>k\}}\big] + 3\nu\big[g\,1_{\{g>k\}}\big] \\ &< \beta^*(b \wedge 2k) + (1/k) + 3\mu\big[f\,1_{\{f>k\}}\big] + 3\nu\big[g\,1_{\{g>k\}}\big]. \end{aligned}$$

Fix $\epsilon > 0$ and take k such that $(1/k) + 3\mu\big[f\,1_{\{f>k\}}\big] + 3\nu\big[g\,1_{\{g>k\}}\big] < \epsilon$. By what already proved, $\beta^*(b \wedge 2k) = \lim_n \beta^*(b_n \wedge 2k)$. Therefore,

$$\beta^*(b) < \beta^*(b \wedge 2k) + \epsilon = \lim_n \beta^*(b_n \wedge 2k) + \epsilon \leq \lim_n \beta^*(b_n) + \epsilon.$$

This concludes the proof. □

In the second lemma, and in the rest of the paper, we write $\alpha(H) = \alpha(1_H)$ whenever $H \in \mathcal{H}$. The same notation is adopted for β, α^* and β^*.

Lemma 2. *Let $c \in M$. Then, condition* (2) *holds provided $\alpha(H) = \beta(H)$ for each $H \in \mathcal{H}$.*

Proof. It suffices to show $\alpha(c) = \beta(c)$. To this end, we first note that $\beta(c)$ is attained, i.e., $\beta(c) = \mu(f_1) + \nu(g_1)$ for some $f_1 + g_1 \in L$ such that $f_1 + g_1 \leq c$; see [14, Proposition 3]. Define $h = c - (f_1 + g_1)$ and fix $t > 1$ and $P \in \Gamma(\mu,\nu)$. Then,

$$\begin{aligned} P(h) &= P\big[h\,1_{\{h\leq t^{-1}\}}\big] + P\big[h\,1_{\{t^{-1}<h\leq 2t\}}\big] + P\big[h\,1_{\{h>2t\}}\big] \\ &\leq t^{-1} + 2t\,P(h > t^{-1}) + P\big[h\,1_{\{h>2t\}}\big]. \end{aligned}$$

Take $f_2 + g_2 \in L$ such that $f_2 + g_2 \geq c$ and define

$$f = f_2 - f_1 \quad \text{and} \quad g = g_2 - g_1.$$

Since $h \leq f + g$,

$$\begin{aligned} P\big[h\,1_{\{h>2t\}}\big] &\leq P\big[(f+g)\,1_{\{f+g>2t\}}\big] \leq P\big[(f+g)\,1_{\{f>t\}}\big] + P\big[(f+g)\,1_{\{g>t\}}\big] \\ &= \mu\big[f\,1_{\{f>t\}}\big] + \nu\big[g\,1_{\{g>t\}}\big] + P\big[f\,1_{\{g>t\}} + g\,1_{\{f>t\}}\big]. \end{aligned}$$

Arguing as in the proof of Lemma 1,

$$P\big[f\,1_{\{g>t\}} + g\,1_{\{f>t\}}\big] \leq 2\,P\big[f\,1_{\{f>t\}} + g\,1_{\{g>t\}}\big] = 2\mu\big[f\,1_{\{f>t\}}\big] + 2\nu\big[g\,1_{\{g>t\}}\big].$$

Hence,

$$P(h) \leq t^{-1} + 2t\,P(h > t^{-1}) + 3\,\big\{\mu\big[f\,1_{\{f>t\}}\big] + \nu\big[g\,1_{\{g>t\}}\big]\big\}.$$

Next, by Theorem 2.1.1 and Remark 2.1.2(b) of [13], there is a *finitely additive* probability Q on $\mathcal{H}$, with marginals μ and ν, such that $Q(c) = \beta(c)$. Since Q has marginals μ and ν, then $\beta(H) \le Q(H)$ for all $H \in \mathcal{H}$ and

$$Q(h) = Q(c) - Q(f_1 + g_1) = \beta(c) - \mu(f_1) - \nu(g_1) = 0.$$

Finally, since $h \ge 0$ and $\alpha(H) = \beta(H)$ for all $H \in \mathcal{H}$, one obtains

$$\alpha(h > t^{-1}) = \beta(h > t^{-1}) \le Q(h > t^{-1}) \le t\, Q(h) = 0.$$

Hence, there is $P_t \in \Gamma(\mu, \nu)$ such that $P_t(h > t^{-1}) < t^{-2}$. It follows that

$$\begin{aligned} \alpha(c) &\le P_t(c) = P_t(f_1 + g_1) + P_t(h) = \mu(f_1) + \nu(g_1) + P_t(h) \\ &\le \beta(c) + 3\left\{1/t + \mu\big[f\, 1_{\{f>t\}}\big] + \nu\big[g\, 1_{\{g>t\}}\big]\right\} \quad \text{for all } t > 1. \end{aligned}$$

Since $1/t + \mu\big[f\, 1_{\{f>t\}}\big] + \nu\big[g\, 1_{\{g>t\}}\big] \to 0$ as $t \to \infty$, this concludes the proof. □

3. Duality theorems without perfectness

It is convenient to distinguish two cases.

3.1. The abstract case.

Theorem 3. *Let $c \in M$. Then, condition (2) holds provided*

(*) *For each $\epsilon > 0$, there is a countable partition $\{A_0, A_1, \ldots\} \subset \mathcal{F}$ of $\mathcal{X}$ such that $\mu(A_0) = 0$ and*

$$\sup_{y \in \mathcal{Y}} |c(x, y) - c(z, y)| \le \epsilon \quad \textit{whenever } x,\, z \in A_i \textit{ and } i > 0.$$

Proof. Again, it suffices to show $\alpha(c) = \beta(c)$. Given $\epsilon > 0$, fix a point $x_i \in A_i$ for each $i > 0$, and define

$$\mathcal{F}_0 = \sigma\big(A_0 \cap A,\, A_i : A \in \mathcal{F},\, i > 0\big), \quad \mu_0 = \mu|\mathcal{F}_0,$$
$$c_0(x, y) = 1_{A_0}(x) c(x, y) + \sum_{i>0} 1_{A_i}(x) c(x_i, y).$$

Let $\Gamma(\mu_0, \nu)$ be the set of probability measures on $\mathcal{F}_0 \otimes \mathcal{G}$ with marginals μ_0 and ν.

Take $f_1 + g_1 \in L$ and $f_2 + g_2 \in L$ such that $f_1 + g_1 \le c \le f_2 + g_2$. Since $|c - c_0| \le \epsilon$, then $f_1 + g_1 - \epsilon \le c_0 \le f_2 + g_2 + \epsilon$. Further, $\sup_{A_i} f_1 < +\infty$ and $\inf_{A_i} f_2 > -\infty$ for each $i > 0$. Define

$$\phi_1 = -\epsilon + 1_{A_0} f_1 + \sum_{i>0} 1_{A_i} \Big(\sup_{A_i} f_1\Big) \quad \text{and} \quad \phi_2 = \epsilon + 1_{A_0} f_2 + \sum_{i>0} 1_{A_i} \Big(\inf_{A_i} f_2\Big).$$

Then, $\phi_1,\, \phi_2 \in L_1(\mu_0)$ and

$$\phi_1 + g_1 \le c_0 \le \phi_2 + g_2.$$

Because of such inequality and since c_0 is $\mathcal{F}_0 \otimes \mathcal{G}$-measurable, one can define

$$\alpha_0 = \inf_{T \in \Gamma(\mu_0, \nu)} T(c_0) \quad \text{and} \quad \beta_0 = \sup_{(f,g)} \big[\mu_0(f) + \nu(g)\big]$$

where sup is over the pairs (f, g) such that $f \in L_1(\mu_0)$, $g \in L_1(\nu)$ and $f + g \le c_0$.

Since μ_0 is an atomic probability measure, then μ_0 is perfect, which in turn implies $\alpha_0 = \beta_0$. Since $|c - c_0| \le \epsilon$, then $\beta_0 \le \beta(c) + \epsilon$. Hence, there is $T \in \Gamma(\mu_0, \nu)$ such that

$$T(c_0) < \alpha_0 + \epsilon = \beta_0 + \epsilon \le \beta(c) + 2\epsilon.$$

If T can be extended to a probability measure $P \in \Gamma(\mu, \nu)$, then

$$\alpha(c) \le P(c) \le \epsilon + P(c_0) = \epsilon + T(c_0) < 3\epsilon + \beta(c).$$

Thus, to conclude the proof, it suffices to show that T can be actually extended to a probability measure $P \in \Gamma(\mu, \nu)$.

For each i with $\mu(A_i) > 0$, define

$$\mu_i(A) = \mu(A \mid A_i) \quad \text{and} \quad \nu_i(B) = T\big(\mathcal{X} \times B \mid A_i \times \mathcal{Y}\big)$$

where $A \in \mathcal{F}$ and $B \in \mathcal{G}$. Define also

$$P = \sum_i \mu(A_i)\,(\mu_i \times \nu_i),$$

where $\mu_i \times \nu_i$ is the product measure of μ_i and ν_i (so that $\mu_i \times \nu_i$ is a probability measure on $\mathcal{H}$). It is straightforward to see that $P \in \Gamma(\mu, \nu)$. Fix $A \in \mathcal{F}_0$ and $B \in \mathcal{G}$. For $i > 0$, either $A \cap A_i = \emptyset$ or $A \cap A_i = A_i$, so that

$$P(A \times B) = \sum_i \mu(A_i)\mu_i(A)\nu_i(B) = \sum_i \mu(A \mid A_i) T(A_i \times B) = T(A \times B).$$

Therefore, $P = T$ on $\mathcal{F}_0 \otimes \mathcal{G}$.

□

In Theorem 3, clearly, the roles of μ and ν can be interchanged. Accordingly, condition (*) can be replaced by

(**) For each $\epsilon > 0$, there is a countable partition $\{B_0, B_1, \ldots\} \subset \mathcal{G}$ of $\mathcal{Y}$ such that $\nu(B_0) = 0$ and

$$\sup_{x \in \mathcal{X}} |c(x, y) - c(x, z)| \le \epsilon \quad \text{whenever } y,\, z \in B_i \text{ and } i > 0.$$

As an example, condition (*) holds (with $A_0 = \emptyset$) if $\mathcal{X}$ is a separable metric space and the function $x \mapsto c(x, y)$ is Lipschitz uniformly with respect to y, i.e.,

$$\sup_{y \in \mathcal{Y}} |c(x, y) - c(z, y)| \le u\, d(x, z) \quad \text{for all } x,\, z \in \mathcal{X} \tag{3}$$

where $u > 0$ is a constant and d the distance on $\mathcal{X}$. Fix in fact $\epsilon > 0$. Because of separability, $\mathcal{X}$ can be partitioned into sets $A_1, A_2, \ldots$ whose diameter is less than ϵ/u. Hence, condition (*) follows trivially from (3). Similarly, condition (**) holds if $\mathcal{Y}$ is a separable metric space and $y \mapsto c(x, y)$ is Lipschitz uniformly with respect to x. Further, as shown in the proof of Theorem 7, separability of $\mathcal{X}$ (of $\mathcal{Y}$) can be weakened into separability of μ (of ν).

Another example is the following. Let $\mathcal{R}$ be the field of subsets of $\mathcal{X} \times \mathcal{Y}$ generated by the measurable rectangles. Each $R \in \mathcal{R}$ can be written as $R = \cup_{i=1}^n (A_i \times B_i)$ for some $n \ge 1$ and $A_i \in \mathcal{F}$, $B_i \in \mathcal{G}$ such that $A_i \cap A_j = \emptyset$ for $i \ne j$. Thus, when $c = 1_R$ with $R \in \mathcal{R}$, condition (*) is trivially true and Theorem 3 yields $\alpha(R) = \beta(R)$ and $\alpha^*(R) = \beta^*(R)$. We next prove duality of certain sets related to $\mathcal{R}$.

Theorem 4. *Let $H = \cup_n R_n$ and $K = \cap_n R_n$ where $R_n \in \mathcal{R}$ for each n. Then,*

$$\alpha^*(H) = \beta^*(H) \quad \text{and} \quad \alpha(K) = \beta(K).$$

In addition, $\alpha(H) = \beta(H)$ provided H can be written as $H = \cup_n (A_n \times B_n)$ with $A_n \in \mathcal{F}$, $B_n \in \mathcal{G}$, and

$$\mu(\limsup_n A_n) = 0 \quad \text{or} \quad \nu(\limsup_n B_n) = 0.$$

(Here, $\limsup_n A_n = \cap_n \cup_{j>n} A_j$ *and* $\limsup_n B_n = \cap_n \cup_{j>n} B_j$*).*

Proof. Let $H_n = \cup_{i=1}^n R_i$. Since $\alpha^*(H_n) = \beta^*(H_n)$, Lemma 1 implies

$$\alpha^*(H) = \sup_{P \in \Gamma(\mu,\nu)} P(H) = \sup_{P \in \Gamma(\mu,\nu)} \sup_n P(H_n) = \sup_n \sup_{P \in \Gamma(\mu,\nu)} P(H_n)$$
$$= \sup_n \alpha^*(H_n) = \sup_n \beta^*(H_n) = \beta^*(H).$$

Thus, α^* and β^* agree on countable unions of elements of $\mathcal{R}$. Since $R_n^c \in \mathcal{R}$, this implies

$$\alpha(K) = 1 - \alpha^*(K^c) = 1 - \alpha^*\big(\cup_n R_n^c\big)$$
$$= 1 - \beta^*\big(\cup_n R_n^c\big) = 1 - \beta^*(K^c) = \beta(K).$$

Next, suppose $H = \cup_n (A_n \times B_n)$ and $\mu(\limsup_n A_n) = 0$. Let $V_n = \cup_{i=1}^n (A_i \times B_i)$. Given $\epsilon > 0$, take $n \geq 1$ such that $\mu\big(\cup_{i>n} A_i\big) < \epsilon$, and then take $P \in \Gamma(\mu,\nu)$ satisfying $P(V_n) < \alpha(V_n) + \epsilon$. Since $\alpha(V_n) = \beta(V_n)$, one obtains

$$\alpha(H) \leq P(H) \leq P(V_n) + P\big(\cup_{i>n}(A_i \times B_i)\big) \leq P(V_n) + \mu\big(\cup_{i>n} A_i\big)$$
$$< \alpha(V_n) + 2\epsilon = \beta(V_n) + 2\epsilon \leq \beta(H) + 2\epsilon.$$

The proof is exactly the same if $\nu(\limsup_n B_n) = 0$. □

Because of Theorem 4, a (classical) question raised by Arveson [2] admits a positive answer for countable unions of measurable rectangles.

Arveson's problem: *If $H \in \mathcal{H}$ satisfies $P(H) = 0$ for all $P \in \Gamma(\mu,\nu)$, are there $A \in \mathcal{F}$ and $B \in \mathcal{G}$ such that $\mu(A) = \nu(B) = 0$ and $H \subset (A \times \mathcal{Y}) \cup (\mathcal{X} \times B)$?*

Indeed, it is not hard to see that $\beta^*(H) = \mu(A) + \nu(B)$ for some $A \in \mathcal{F}$ and $B \in \mathcal{G}$ with $H \subset (A \times \mathcal{Y}) \cup (\mathcal{X} \times B)$; see e.g. [7, Lemma 1]. If H is a countable union of measurable rectangles, Theorem 4 implies $\beta^*(H) = \alpha^*(H) = 0$ so that $\mu(A) = \nu(B) = 0$.

In addition, exploiting Theorem 4, duality for $H = \cup_n (A_n \times B_n)$ can be obtained under various conditions. One such condition is $\mu(\limsup_n A_n) = 0$ or $\nu(\limsup_n B_n) = 0$. A similar condition is that H^c is also a countable union of measurable rectangles. In this case, in fact, $\alpha^*(H^c) = \beta^*(H^c)$ or equivalently $\alpha(H) = \beta(H)$. A last condition is

(4) for each $n \geq 1$ there is a measurable function $\phi_n : \mathcal{X} \to \mathcal{Y}$ such that

$$\nu = \mu \circ \phi_n^{-1} \quad \text{and} \quad \mu\big\{x : (x, \phi_n(x)) \in H\big\} < 1/n.$$

Define in fact $P_n(U) = \mu\big\{x : (x, \phi_n(x)) \in U\big\}$ for each $n \geq 1$ and $U \in \mathcal{H}$. Then, $P_n \in \Gamma(\mu,\nu)$ and $\alpha(H) \leq P_n(H) < 1/n$. Thus $\alpha(H) = 0$, which in turn implies $\alpha(H) = \beta(H)$. Here is a simple example.

Example 5. Suppose $(\mathcal{X},\mathcal{F})=(\mathcal{Y},\mathcal{G})$ and $\mu=\nu$, with $\mathcal{X}$ a separable metric space and $\mathcal{F}=\mathcal{B}(\mathcal{X})$. (Up to some technicalities, separability of $\mathcal{X}$ could be weakened into separability of μ). Let $\Delta=\{(x,x):x\in\mathcal{X}\}$ be the diagonal and H a countable union of measurable rectangles. Then, *duality holds for $H\cap\Delta^c$, and it holds for $H\cap\Delta$ provided μ vanishes on singletons.* In fact, $H\cap\Delta^c$ is a countable union of measurable rectangles and $\mu\{x:(x,x)\in H\cap\Delta^c\}=0$. Letting $\phi_n(x)=x$, Theorem 4 and condition (4) yield

$$\alpha(H\cap\Delta^c)=\beta(H\cap\Delta^c)\quad\text{and}\quad\alpha^*(H\cap\Delta^c)=\beta^*(H\cap\Delta^c).$$

To deal with $H\cap\Delta$, suppose μ null on singletons and define $P_1=\mu\times\mu$ and $P_2(U)=\mu\{x:(x,x)\in U\}$ for each $U\in\mathcal{H}$. Then, $P_1,\,P_2\in\Gamma(\mu,\mu)$. Since μ is null on singletons, $\alpha(H\cap\Delta)\le P_1(H\cap\Delta)\le P_1(\Delta)=0$, which in turn implies $\alpha(H\cap\Delta)=\beta(H\cap\Delta)$. Finally, writing H as $H=\cup_n(A_n\times B_n)$, one obtains

$$\alpha^*(H\cap\Delta)\le\beta^*(H\cap\Delta)\le\mu\big(\cup_n(A_n\cap B_n)\big)=P_2(H\cap\Delta)\le\alpha^*(H\cap\Delta).$$

We close this Subsection with two remarks. The first (stated as a lemma) suggests a possible strategy for proving a general duality theorem.

Lemma 6. *Let $\mathcal{H}_0=\big\{H\in\mathcal{H}:\alpha(H)=\beta(H)$ and $\alpha^*(H)=\beta^*(H)\big\}$. Then, condition (2) holds for each $c\in M$ if and only if*

$$\text{(5)}\qquad H_n\in\mathcal{H}_0 \text{ and } H_n\subset H_{n+1} \text{ for each } n \implies \alpha\big(\cup_n H_n\big)=\beta\big(\cup_n H_n\big).$$

Proof. By Lemma 2, it suffices to show that $\mathcal{H}_0=\mathcal{H}$. In turn, since $\mathcal{R}\subset\mathcal{H}_0$, it suffices to see that $\mathcal{H}_0$ is a monotone class. Also, since $\mathcal{H}_0$ is closed under complements, it is enough to prove that $H\in\mathcal{H}_0$ provided H is the union of an increasing sequence of elements of $\mathcal{H}_0$. Let $H=\cup_n H_n$ where $H_n\in\mathcal{H}_0$ and $H_n\subset H_{n+1}$ for each n. For such H, arguing as in the proof of Theorem 4, one obtains $\alpha^*(H)=\beta^*(H)$. Thus, under (5), $\mathcal{H}_0$ is actually a monotone class. □

The second remark briefly compares the arguments underlying Theorem 3 and the usual duality theorems. The latter are summarized into the result by Ramachandran and Ruschendorf [14].

For definiteness, we aim to prove $\alpha(c)=\beta(c)$. By (1) and since $\beta(c)$ is attained, it can be assumed $c\ge 0$ and $\beta(c)=0$. As noted in the proof of Lemma 2, there is a finitely additive probability Q on $\mathcal{H}$, with marginals μ and ν, satisfying $Q(c)=\beta(c)$. Since $c\ge 0$ and $\beta(c)=0$, it must be $Q(c>\epsilon)=0$ for each $\epsilon>0$. A basic intuition in [14] is that, if one of μ and ν is perfect, then Q is σ-additive on $\mathcal{R}$; see [13, Theorem 2.1.3] and recall that $\mathcal{R}$ is the field generated by the measurable rectangles. Hence, there is $P\in\Gamma(\mu,\nu)$ such that $P=Q$ on $\mathcal{R}$. With such a P, one obtains

$$P\big(\cup_i R_i\big)=\sup_n P\big(\cup_{i=1}^n R_i\big)=\sup_n Q\big(\cup_{i=1}^n R_i\big)\le Q\big(\cup_i R_i\big)$$

provided $R_i\in\mathcal{R}$ for all i. Hence, $P(c>\epsilon)\le Q(c>\epsilon)=0$ if the set $\{c>\epsilon\}$ is a countable union of measurable rectangles. Up to some technicalities, suitable versions of this argument work even if $\{c>\epsilon\}$ fails to be a countable union of measurable rectangles. This provides a rough sketch of the proof of $\alpha(c)=\beta(c)$ under the assumption that one of μ and ν is perfect. We now turn to Theorem 3. Here, instead of proving that Q is σ-additive on $\mathcal{R}$, one requires that c can be suitably approximated by $\mathcal{R}$-simple functions. For instance, conditions (*)-(**) are trivially true if c is the uniform limit of a sequence of $\mathcal{R}$-simple functions, and in this case no assumptions on μ or ν are needed. Apparently, conditions (*)-(**) are too restrictive to be useful in real problems. Instead, they allow to get duality in various situations, including Theorem 4, Example 5, and the results in the next subsection.

3.2. **The metric case.** In this subsection, $\mathcal{X}$ and $\mathcal{Y}$ are metric spaces, $\mathcal{F} = \mathcal{B}(\mathcal{X})$ and $\mathcal{G} = \mathcal{B}(\mathcal{Y})$. The *sections of* c are the functions $x \mapsto c(x, y)$ and $y \mapsto c(x, y)$, with y fixed in the first map and x fixed in the second.

A remark is in order. All claims made so far are still valid, even if c is not $\mathcal{H}$-measurable, provided $c\,1_{A\times B}$ is $\mathcal{H}$-measurable for some $A \in \mathcal{F}$ and $B \in \mathcal{G}$ with $\mu(A) = \nu(B) = 1$. In fact, $\alpha(c) = \alpha(c\,1_{A\times B})$ whenever $\alpha(c)$ is defined in the obvious way, i.e.

$$\alpha(c) = \inf_{P\in\Gamma(\mu,\nu)} \overline{P}(c) \quad \text{where } \overline{P} \text{ is the completion of } P.$$

Similarly, $\alpha^*(c) = \alpha^*(c\,1_{A\times B})$, $\beta(c) = \beta(c\,1_{A\times B})$ and $\beta^*(c) = \beta^*(c\,1_{A\times B})$.

In the next result, $c\,1_{A\times B}$ is actually $\mathcal{H}$-measurable for some $A \in \mathcal{F}$ and $B \in \mathcal{G}$ such that $\mu(A) = \nu(B) = 1$ (with possibly $A = \mathcal{X}$ or $B = \mathcal{Y}$).

Theorem 7. *Suppose c satisfies condition* (1)*, the map $x \mapsto c(x, y)$ is continuous for each $y \in \mathcal{Y}$ and the map $y \mapsto c(x, y)$ is $\mathcal{G}$-measurable for each $x \in \mathcal{X}$. Then,*

(i) *$\alpha^*(c) = \beta^*(c)$ if c is bounded below and μ is separable;*
(ii) *$\alpha(c) = \beta(c)$ if c is bounded above and μ is separable;*
(iii) *$\alpha(c) = \beta(c)$ and $\alpha^*(c) = \beta^*(c)$ if c is bounded and μ is separable;*
(iv) *$\alpha(c) = \beta(c)$ and $\alpha^*(c) = \beta^*(c)$ if all the sections of c are continuous and at least one of μ and ν is separable.*

Proof. Since (ii) and (iii) are consequences of (i), it suffices to prove (i) and (iv).

Let μ and c be as in (i). Since μ is separable, there is a separable set $A \in \mathcal{F}$ with $\mu(A) = 1$. Since $x \mapsto c(x, y)$ is continuous, $y \mapsto c(x, y)$ is Borel measurable and A is separable, the restriction of c on $A \times \mathcal{Y}$ is measurable with respect to $\mathcal{B}(A) \otimes \mathcal{B}(\mathcal{Y})$. Therefore, $c\,1_A$ is $\mathcal{H}$-measurable.

Take a countable set $D \subset A$ such that $\overline{D} = \overline{A}$ and define

$$c_n(x, y) = \inf_{z\in D} \big\{ n\, d(x, z) + c(z, y) \big\}$$

where $n \geq 1$, $(x, y) \in \mathcal{X} \times \mathcal{Y}$ and d is the distance on $\mathcal{X}$.

Since c is bounded below, c_n is real-valued, and a direct calculation shows that

$$\sup_{y\in\mathcal{Y}} |c_n(x, y) - c_n(z, y)| \leq n\, d(x, z) \quad \text{for all } x,\, z \in \mathcal{X}. \tag{6}$$

Since D is countable, $y \mapsto c_n(x, y)$ is Borel measurable. Hence, $c_n\,1_A$ is $\mathcal{H}$-measurable. In addition, $c_n \leq c_{n+1}$ and c_n meets condition (1) (since c meets (1) and is bounded below). Finally, since $x \mapsto c(x, y)$ is continuous, one obtains

$$c(x, y) = \sup_n c_n(x, y) \quad \text{for all } (x, y) \in \overline{A} \times \mathcal{Y}.$$

Next, $\overline{D} = \overline{A}$ implies $A \subset \bigcup_{x\in D} B(x, \delta)$ for each $\delta > 0$, where $B(x, \delta)$ is the $\mathcal{X}$-ball of radius δ around x. Given $n \geq 1$ and $\epsilon > 0$, it follows that A can be partitioned into sets $A_1, A_2, \ldots \in \mathcal{F}$ whose diameter is less than ϵ/n. Hence, c_n meets condition (*) (with $A_0 = A^c$) because of (6). By Lemma 1 and Theorem 3,

$$\alpha^*(c) = \alpha^*(c\,1_A) \leq \beta^*(c\,1_A) = \lim_n \beta^*(c_n\,1_A)$$
$$= \lim_n \alpha^*(c_n\,1_A) \leq \alpha^*(c\,1_A) = \alpha^*(c).$$

This concludes the proof of (i).

Let us turn to (iv). Suppose that all the c-sections are continuous. Since the sections of $-c$ are continuous as well, it suffices to prove $\alpha^*(c) = \beta^*(c)$. We first assume μ separable.

By (1), there are $\psi \in L_1(\mu)$ and $g \in L_1(\nu)$ such that $\psi + g \leq c$. Define

$$f(x) = \inf_{y\in\mathcal{Y}} \big\{c(x,y) - g(y)\big\}, \quad x \in \mathcal{X},$$

and note that $f \in L_1(\mu)$, $f + g \leq c$ and f is upper-semicontinuous. Define also

$$c_n(x,y) = \inf_{z\in\mathcal{X}} \big\{n\,d(x,z) + c(z,y) - f(z)\big\}.$$

Again, condition (6) holds, c_n meets condition (1) (since $g \leq c_n \leq c - f$) and $y \mapsto c_n(x,y)$ is Borel measurable (it is in fact upper-semicontinuous). On noting that $x \mapsto c(x,y) - f(x)$ is lower-semicontinuous, it is not hard to see that $c_n \uparrow c - f$ pointwise as $n \to \infty$. Because of (6) and μ separable, c_n meets condition (*). Thus,

$$\beta^*(c) - \mu(f) = \beta^*(c - f) = \lim_n \beta^*(c_n)$$
$$= \lim_n \alpha^*(c_n) \leq \alpha^*(c - f) = \alpha^*(c) - \mu(f).$$

Hence, $\alpha^*(c) = \beta^*(c)$ if μ is separable.

Finally, if ν is separable, it suffices to let

$$c_n(x,y) = \inf_{z\in\mathcal{Y}} \big\{n\,\rho(y,z) + c(x,z) - g(z)\big\}$$

where now g is upper-semicontinuous and ρ is the distance on $\mathcal{Y}$. Arguing as above and using separability of ν, it follows that c_n meets condition (**) and $c_n \uparrow c - g$ pointwise as $n \to \infty$. Hence, $\alpha^*(c) = \beta^*(c)$ and this concludes the proof. □

Once again, the roles of μ and ν can be interchanged in Theorem 7.

Theorem 8. *Suppose c satisfies condition* (1)*, the map $x \mapsto c(x,y)$ is $\mathcal{F}$-measurable for each $y \in \mathcal{Y}$ and the map $y \mapsto c(x,y)$ is continuous for each $x \in \mathcal{X}$. Then,*

(j) *$\alpha^*(c) = \beta^*(c)$ if c is bounded below and ν is separable;*
(jj) *$\alpha(c) = \beta(c)$ if c is bounded above and ν is separable;*
(jjj) *$\alpha(c) = \beta(c)$ and $\alpha^*(c) = \beta^*(c)$ if c is bounded and ν is separable.*

Note that, if μ and ν are both separable, then $\alpha(c) = \beta(c)$ and $\alpha^*(c) = \beta^*(c)$ provided c is bounded, $\mathcal{H}$-measurable, and at least one of the c-sections is continuous. Further, the argument underlying Theorems 7-8 yields other similar results. As an example, we state (without a proof) the following.

Theorem 9. *Suppose c satisfies condition* (1) *and at least one of μ and ν is separable. Then, $\alpha^*(c) = \beta^*(c)$ if c is lower-semicontinuous (with respect to the product topology on $\mathcal{X} \times \mathcal{Y}$) and $\alpha(c) = \beta(c)$ if c is upper-semicontinuous.*

Finally, we list some consequences of Theorems 7-9. Indeed, they unify and slightly improve some known results.

- Theorems 7-8 improve [14], the result by Ramachandran and Ruschendorf, provided c satisfies some conditions and
$$\text{card}(\mathcal{X}) \le \text{card}(\mathbb{R}) \quad \text{and} \quad \text{card}(\mathcal{Y}) \le \text{card}(\mathbb{R}).$$
Under such cardinality assumption, in fact, perfectness implies separability but not conversely; see Section 2.
- As an example, suppose $c \in M$ and $\mathcal{X}$ and $\mathcal{Y}$ are separable metric spaces (so that μ and ν are both separable and the cardinality assumption is satisfied). Then, [14] implies $\alpha(c) = \beta(c)$ and $\alpha^*(c) = \beta^*(c)$ provided at least one between μ and ν is perfect. Instead, Theorems 7-8 lead to the same conclusions whenever all the c-sections are continuous, or whenever c is bounded and at least one of the c-sections is continuous.
- By Theorems 7-8, it is consistent with the usual axioms of set theory (ZFC) that condition (2) holds for every $c \in M$ with continuous sections, or for every bounded $c \in M$ with at least one continuous section. In fact, as noted in Section 2, it is consistent with ZFC that any Borel probability on any metric space is separable.
- Let $\mathcal{X} = \mathcal{Y}$ and $c = d$, where d is the distance on $\mathcal{X}$. Suppose d measurable with respect to $\mathcal{B}(\mathcal{X}) \otimes \mathcal{B}(\mathcal{X})$ and
$$\int d(x, x_0)\, \mu(dx) + \int d(x, x_0)\, \nu(dx) < \infty \quad \text{for some } x_0 \in \mathcal{X}.$$
Then, $\alpha(d)$ reduces to Wasserstein distance between μ and ν while $\beta(d)$ can be written as
$$\beta(d) = \sup_f |\mu(f) - \nu(f)|$$
where sup is over the 1-Lipschitz functions $f : \mathcal{X} \to \mathbb{R}$. In this case, it is well known that $\alpha(d) = \beta(d)$ if $\mathcal{X}$ is separable; see e.g. [9, page 400]. This known fact is generalized by Theorems 7-9 under two respects: separability of $\mathcal{X}$ can be weakened into separability of at least one of μ and ν, and d can be replaced by any upper-semicontinuous function or by any function with continuous sections.
- By Theorem 9, Arveson's question has a positive answer if H is open and one of μ and ν is separable.

References

[1] Ambrosio L., Gigli N., Savare' G. (2008) *Gradient flows*, Second Edition, Birkhauser, Basel.
[2] Arveson W. (1974) Operator algebras and invariant subspaces, *Ann. of Math.*, 100, 433-533.
[3] Beiglbock M., Schachermayer W. (2011) Duality for Borel measurable cost functions, *Trans. Amer. Math. Soc.*, 363, 4203-4224.
[4] Beiglbock M., Leonard C., Schachermayer W. (2012) A general duality theorem for the Monge-Kantorovich transport problem, *Studia Math.*, 209, 151-167.
[5] Berti P., Pratelli L., Rigo P. (2015) Two versions of the fundamental theorem of asset pricing, *Electronic J. Probab.*, 20, 1-21.
[6] Berti P., Pratelli L., Rigo P., Spizzichino F. (2015) Equivalent or absolutely continuous probability measures with given marginals, *Dependence Modeling*, 3, 47-58.
[7] Haydon R., Shulman V. (1996) On a measure-theoretic problem of Arveson, *Proc. Amer. Math. Soc.*, 124, 497-503.
[8] Kantorovich L. (1942) On the translocation of masses, *C. R. (Doklady) Acad. Sci. URSS (N.S.)*, 37, 199-201.

[9] Kellerer H.G. (1984) Duality theorems for marginal problems, *Z. Wahrscheinlichkeitstheorie Verw. Geb.*, 67, 399-432.

[10] Koumoullis G. (1981) On perfect measures, *Trans. Amer. Math. Soc.*, 264, 521-537.

[11] Puccetti G., Wang R. (2015) Extremal dependence concepts, *Statist. Science*, 30, 485-517.

[12] Puccetti G., Rigo P., Wang B., Wang R. (2019) Centers of probability measures without the mean, *J. Theore. Probab.*, 32, 1482-1501.

[13] Rachev S.T., Ruschendorf L. (1998) *Mass transportation problems, Volume I: Theory*, Springer, New York.

[14] Ramachandran D., Ruschendorf L. (1995) A general duality theorem for marginal problems, *Prob. Theo. Relat. Fields*, 101, 311-319.

[15] Ramachandran D., Ruschendorf L. (2000) On the Monge-Kantorovitch duality theorem, *Theory Probab. Appl.*, 45, 350-356.

[16] Ruschendorf L. (2013) *Mathematical Risk Analysis: Dependence, Risk Bounds, Optimal Allocations and Portfolios*, Springer, Heidelberg.

[17] Villani C. (2009) *Optimal transport, old and new*, Springer, New York.

Pietro Rigo, Dipartimento di Matematica "F. Casorati", Universita' di Pavia, via Ferrata 1, 27100 Pavia, Italy

E-mail address: pietro.rigo@unipv.it