ESAIM: COCV 17 (2011) 682–704
DOI: 10.1051/cocv/2010013

ESAIM: Control, Optimisation and Calculus of Variations
www.esaim-cocv.org

A SADDLE-POINT APPROACH TO THE MONGE-KANTOROVICH OPTIMAL TRANSPORT PROBLEM

Christian Léonard[1]

Abstract. The Monge-Kantorovich problem is revisited by means of a variant of the saddle-point method without appealing to c-conjugates. A new abstract characterization of the optimal plans is obtained in the case where the cost function takes infinite values. It leads us to new explicit sufficient and necessary optimality conditions. As by-products, we obtain a new proof of the well-known Kantorovich dual equality and an improvement of the convergence of the minimizing sequences.

Mathematics Subject Classification. 46N10, 49J45, 28A35.

Received September 15, 2009.
Published online March 31, 2010.

1. Introduction

The Monge-Kantorovich problem is revisited by means of a variant of the saddle-point method derived in [12], without appealing to c-conjugates. A new abstract characterization of the optimal plans is obtained (Thm. 4.2) in the case where the cost function takes infinite values. It leads us to new explicit sufficient and necessary conditions of optimality which are stated in Theorems 1.8 and 1.9. As by-products, we obtain a new proof of the well-known Kantorovich dual equality and an improvement of the convergence of the minimizing sequences.

1.1. The Monge-Kantorovich transport problem

Let us take A and B two Polish (separable complete metric) spaces furnished with their respective Borel σ-fields, a lower semicontinuous (cost) function $c : A\times B \to [0,\infty]$ which may take infinite values and two probability measures $\mu \in \mathcal{P}_A$ and $\nu \in \mathcal{P}_B$ on A and B. We denote $\mathcal{P}_A$, $\mathcal{P}_B$ and $\mathcal{P}_{AB}$ the sets of all Borel probability measures on A, B and $A\times B$. The Monge-Kantorovich problem is

$$\textit{minimize } \pi \in \mathcal{P}_{AB} \mapsto \int_{A\times B} c(a,b)\,\pi(\mathrm{d}a\mathrm{d}b) \textit{ subject to } \pi \in P(\mu,\nu) \tag{MK}$$

where $P(\mu,\nu)$ is the set of all $\pi \in \mathcal{P}_{AB}$ with prescribed marginals $\pi_A = \mu$ on A and $\pi_B = \nu$ on B. Note that c is measurable since it is lower semicontinuous and the integral $\int_{A\times B} c\,\mathrm{d}\pi \in [0,\infty]$ is well-defined since $c \geq 0$.

For a general account on this active field of research, see the books of Rachev and Rüschendorf [15] and Villani [18,19].

Keywords and phrases. Convex optimization, saddle-point, conjugate duality, optimal transport.

[1] Modal-X, Université Paris Ouest, Bât. G, 200 av. de la République, 92001 Nanterre, France. christian.leonard@u-paris10.fr

Article published by EDP Sciences

© EDP Sciences, SMAI 2010

https://doi.org/10.1051/cocv/2010013 Published online by Cambridge University Press

The subset $\{c = \infty\}$ of $A \times B$ is a set of forbidden transitions. Optimal transport on the Wiener space [11] and on configuration spaces [8,9] provide natural infinite dimensional settings where c takes infinite values.

Let us denote C_A, C_B and C_{AB} the spaces of all continuous bounded functions on A, B and $A \times B$. The Kantorovich maximization problem:

$$\begin{cases} \text{maximize } \displaystyle\int_A f \,\mathrm{d}\mu + \int_B g \,\mathrm{d}\nu \text{ for all } f,\ g \text{ such that} \\ f \in C_A,\ g \in C_B \text{ and } f \oplus g \leq c \end{cases} \tag{K}$$

is the basic dual problem of (MK). Here and below, we denote $f \oplus g(a,b) = f(a) + g(b)$. Under our assumptions, we have

$$\inf (\mathrm{MK}) = \sup (\mathrm{K}) \in [0, \infty] \tag{1.1}$$

which is called the Kantorovich dual equality. For a proof of this well known result, see [19], Theorem 5.10, for instance.

Definitions 1.1 (plans)**.**

(1) Any probability measure in $P(\mu, \nu)$ is called a transport plan, or shorter: a *plan*.
(2) One says that $\pi \in P(\mu, \nu)$ is a *finite plan* if $\int_{A \times B} c \,\mathrm{d}\pi < \infty$. The set of all finite plans is denoted by $P(\mu, \nu, c)$.
(3) One says that π is an *optimal plan* if it is a finite plan and it minimizes $\gamma \mapsto \int_{A \times B} c \,\mathrm{d}\gamma$ on $P(\mu, \nu)$.

It is well-known that there exists at least an optimal plan if and only if $P(\mu, \nu, c)$ is nonempty; this will be found again in Theorem 3.1. Definition 1.1-(3) throws away the uninteresting case where $\int_{A \times B} c \,\mathrm{d}\pi = \infty$ for all $\pi \in P(\mu, \nu)$.

Since it is a convex but not a *strictly* convex problem, infinitely many optimal plans may exist.

Recently, Beiglböck *et al.* [2,3,17] have improved previous optimality criteria in several directions. Before stating some of their results, let us introduce the notion of strongly c-cyclically monotone plan.

Clearly, there is no reason for the dual equality (1.1) to be attained at *continuous* functions f and g. Suppose instead that there exist two $[-\infty, \infty)$-valued integrable functions $f \in \mathcal{L}^1(\mu)$ and $g \in \mathcal{L}^1(\nu)$ such that $f \oplus g \leq c$ everywhere and $\sup\{\int_A u \,\mathrm{d}\mu + \int_B v \,\mathrm{d}\nu;\, u \in \mathcal{L}^1(\mu), v \in \mathcal{L}^1(\nu), u \oplus v \leq c\} = \int_A f \,\mathrm{d}\mu + \int_B g \,\mathrm{d}\nu$. Let π be an optimal plan. We have

$$\sup (\mathrm{K}) \leq \sup_{u \in \mathcal{L}^1(\mu), v \in \mathcal{L}^1(\nu), u \oplus v \leq c} \int_A u \,\mathrm{d}\mu + \int_B v \,\mathrm{d}\nu = \int f \oplus g \,\mathrm{d}\pi \leq \int c \,\mathrm{d}\pi = \inf (\mathrm{MK})$$

and by (1.1), this is a series of equalities. In particular, the couple (f, g) is an *integrable dual optimizer*, $\int (c - f \oplus g) \,\mathrm{d}\pi = 0$ and since $c - f \oplus g \geq 0$, we see that $c = f \oplus g$, π-almost everywhere. This leads naturally to the following notion.

Definition 1.2 (strongly c-cyclically monotone plan [17])**.** A transport plan $\pi \in P(\mu, \nu)$ is called *strongly c-cyclically monotone* if there exist $[-\infty, \infty)$-valued measurable functions f on A and g on B such that

$$\begin{cases} f \oplus g \leq c & \text{everywhere} \\ f \oplus g = c & \pi\text{-almost everywhere.} \end{cases} \tag{1.2}$$

In the whole paper, measurable functions and sets are intended to be Borel measurable.

Note that it is not required in this definition that f and g are integrable. In fact, one can find examples where there is an optimal plan but no integrable dual optimizer, see [2], Example 4.5. Without integrability, one cannot write $\int_A f \,\mathrm{d}\mu + \int_B g \,\mathrm{d}\nu$ anymore. Nevertheless, the following result allows us to extend the notion of dual optimizer to measurable functions (f, g) as in Definition 1.2.

https://doi.org/10.1051/cocv/2010013 Published online by Cambridge University Press

Lemma 1.3 ([2], Lem. 1.1)**.** *Let $c : A\times B \to [0,\infty]$ be measurable, f and g be $[-\infty,\infty)$-valued measurable functions on A and B respectively such that $f\oplus g \le c$ everywhere. Then, for any $\pi,\ \tilde\pi \in P(\mu,\nu,c)$,*

$$\int_{A\times B} f\oplus g\,\mathrm{d}\pi = \int_{A\times B} f\oplus g\,\mathrm{d}\tilde\pi \in [-\infty,\infty).$$

With this lemma in hand, we are allowed to denote

$$\int_{A\times B} f\oplus g\,\mathrm{d}(\mu,\nu,c) := \int_{A\times B} f\oplus g\,\mathrm{d}\pi, \quad \pi\in P(\mu,\nu,c) \tag{1.3}$$

this common value. A natural extension of the dual problem (K) is

$$\begin{cases} \text{maximize } \displaystyle\int_{A\times B} f\oplus g\,\mathrm{d}(\mu,\nu,c) \text{ for all } f\in[-\infty,\infty)^A,\ g\in[-\infty,\infty)^B \\ f,\ g \text{ measurable such that } f\oplus g\le c \text{ everywhere.} \end{cases} \tag{$\widetilde{\mathrm{K}}$}$$

Of course, $\sup(\mathrm{K}) \le \sup(\widetilde{\mathrm{K}}) \le \inf(\mathbb{MK})$ so that (1.1) implies

$$\sup(\widetilde{\mathrm{K}}) = \inf(\mathbb{MK}). \tag{1.4}$$

A couple of functions (f,g) as in Definition 1.2 is clearly an optimizer of $(\widetilde{\mathrm{K}})$. We call it a *measurable dual optimizer*.

Some usual optimality criteria are expressed in terms of cyclical c-monotonicity.

Definition 1.4 (c-cyclically monotone plan)**.** A subset $\Gamma\subset A\times B$ is said to be *c-cyclically monotone* if for any integer $n\ge 1$ and any family $(a_1,b_1),\dots,(a_n,b_n)$ of points in Γ, $\sum_{i=1}^n c(a_i,b_i) \le \sum_{i=1}^n c(a_i,b_{i+1})$ with the convention $b_{n+1}=b_1$.

A probability measure $\pi\in\mathcal{P}_{AB}$ is said to be c-cyclically monotone if it is concentrated on a measurable c-cyclically monotone set Γ, *i.e.* $\pi(\Gamma)=1$.

This notion goes back to the seminal paper [16] by Rüschendorf where the standard cyclical monotonicity of convex functions introduced by Rockafellar has been extended in view of solving the Monge-Kantorovich problem.

One easily shows that a strongly c-cyclically monotone plan is c-cyclically monotone.

The main results of [2,3,17] are collected in the next two theorems.

Theorem 1.5 ([2,3])**.** *Let c be a measurable nonnegative function such that*

$$\mu\otimes\nu(\{c<\infty\}) = 1. \tag{1.5}$$

If there exists some $\pi^o\in P(\mu,\nu)$ such that $\int_{A\times B} c\,\mathrm{d}\pi^o<\infty$, then (1.4) holds true and for any $\pi\in P(\mu,\nu)$, the following three statements are equivalent:

(i) *π is an optimal plan;*
(ii) *π is c-cyclically monotone;*
(iii) *π is strongly c-cyclically monotone.*

This result is valid under a very weak regularity condition on c which is only assumed to be Borel measurable. It was first proved in [17] under the assumption that c is a lower semicontinuous nonnegative finitely-valued function. However, condition (1.5) is close to the assumption that c is finitely valued.

The next result is concerned with cost functions c which may take infinite values.

https://doi.org/10.1051/cocv/2010013 Published online by Cambridge University Press

Theorem 1.6 ([1,3,17]). *Let c be a lower semicontinuous $[0,\infty]$-valued function.*

(a) *Any optimal plan is c-cyclically monotone.*
(b) *If there exists some $\pi^o \in P(\mu,\nu)$ such that $\int_{A\times B} c\,\mathrm{d}\pi^o < \infty$, then any strongly c-cyclically monotone plan in $P(\mu,\nu)$ is an optimal plan.*

Let c be a measurable $[0,\infty]$-valued function.

(c) *Every finite c-cyclically monotone transport plan is optimal if there exist a closed set F and a $\mu\otimes\nu$-null set N such that $\{c=\infty\} = F\cup N$.*

Statement (a) is proved in Ambrosio and Pratelli's lecture notes [1] and (b) is taken from [17]. Statement (c) is proved in [3], it extends recent results of Pratelli [14].

Counterexamples 1.7. Otherwise stated, c is assumed to be lower semicontinuous.

(1) c is real valued.
  (a) [3], Example 1.3, c is not lower semicontinuous and no optimal plan exists.
  (b) [2], Example 4.5, c is the squared distance and no *integrable* dual optimizers exist.
(2) c takes infinite values on a non-null set.
  (a) [2], Example 4.1, c is not lower semicontinuous and the weak Kantorovich dual equality (1.4) fails to hold: we have $\sup(\widetilde{\mathrm{K}}) < \inf(\mathrm{MK})$.
  (b) [1], Example 3.5, a c-cyclically monotone plan which is not optimal.
  (c) [17], Example 1, a c-cyclically monotone plan which is not strongly c-cyclically monotone.
  (d) [3], Example 5.1, an optimal plan which is not strongly c-cyclically monotone.
  (e) [2], Example 4.2, an optimal plan which is not strongly c-cyclically monotone with c continuous.

It appears that Theorem 1.5 is the best possible result under the assumption (1.5) but that its extension to the general case where c takes infinite values on a non-null set is still open. The Counterexample (2)-(b) tells us that one must drop the notion of c-cyclically monotone plan and with the Counterexamples (2)-(d) and (2)-(e), one sees that even the notion of strongly c-cyclically monotone plan is not enough to characterize optimality.

Because of the dual gap when c is not lower semicontinuous (Counterexample (2)-(a), it is reasonable to assume that c is lower semicontinuous in the general case. This will be assumed from now on.

1.2. The aim of this paper

This paper is aimed at implementing the saddle-point method for solving (MK) in the general case where c might take infinite values, without appealing to c-conjugates, see [18,19]. After all, (MK) is a convex minimization problem and one may wonder what the standard approach based on convex duality could yield. It appears that an extended version of the saddle-point is necessary to investigate (MK) without imposing strong qualifications of the constraints such as the standard requirement: $\int_A c_A\,\mathrm{d}\mu + \int_B c_B\,\mathrm{d}\nu < \infty$ where $c \le c_A \oplus c_B$, see [18,19]. The present paper relies on an extension of the standard saddle-point method which has been developed in [12].

- A new proof of the Kantorovich dual equality (1.1) is given in Theorem 3.1 providing at the same time an improved result about the behavior of the minimizing sequences of (MK).
- An abstract characterization of the optimal plans is given in Theorem 4.2. It expresses a saddle-point property and extends the notion of strong c-cyclical monotonicity.
- It leads easily to a sufficient condition of optimality in Theorem 1.8 below which is an improvement of Theorem 1.6-(b). Its proof is given in Section 5.
- Finding good necessary conditions for optimal plans with a genuinely $[0,\infty]$-valued cost function c is still an open problem. Theorem 1.9 below, which is a corollary of Theorem 4.2, goes one step in this direction. Its proof is given in Section 6.

One says that f is μ-measurable if there exists a measurable set N such that $\mu(N) = 0$ and $\mathbf{1}_N f$ is a measurable function. A property holds $P(\mu,\nu,c)$-almost everywhere if it holds true outside a measurable set N

https://doi.org/10.1051/cocv/2010013 Published online by Cambridge University Press

such that $\gamma(N)=0$, for all $\gamma\in P(\mu,\nu,c)$. With these definitions in hand, we can state our sufficient condition of optimality.

Theorem 1.8. *Let $\pi\in P(\mu,\nu,c)$ be any finite plan. If there exist a μ-measurable function f on A and a ν-measurable function g on B which satisfy*

$$\begin{cases} f\oplus g\le c & P(\mu,\nu,c)\text{-almost everywhere}\\ f\oplus g=c & \pi\text{-almost everywhere,}\end{cases} \tag{1.6}$$

then π is optimal.

The Counterexamples 1.7-(d,e) are optimal plans which are not strongly c-cyclically monotone but they both satisfy the weaker property (1.6). See Section 5.3 for more details.

The following result is our necessary condition of optimality.

Theorem 1.9. *Take any optimal plan π, $\epsilon>0$ and p any probability measure on $A\times B$ such that $\int_{A\times B}c\,\mathrm{d}p<\infty$. Then, there exist functions $\varphi\in L_1(\pi+p)$, u and v bounded continuous on A and B respectively and a measurable subset $D\subset A\times B$ such that:*

(1) $\varphi=c$, *π-almost everywhere on $A\times B\setminus D$;*
(2) $\int_D(1+c)\,\mathrm{d}\pi\le\epsilon$;
(3) $-c/\epsilon\le\varphi\le c$, *$(\pi+p)$-almost everywhere;*
(4) $-c/\epsilon\le u\oplus v\le c$, *everywhere;*
(5) $\|\varphi-u\oplus v\|_{L_1(\pi+p)}\le\epsilon$.

An important thing to notice in this result is the appearance of the probability measure p in items (3) and (5). One can read (3)–(5) as an approximation of $f\oplus g\le c$, $(\pi+p)$-almost everywhere. Since it is required that $\int_{A\times B}c\,\mathrm{d}p<\infty$, one can choose p in $P(\mu,\nu,c)$, and the properties (1)–(5) of this theorem are an approximation of (1.6) where $P(\mu,\nu,c)$-a.e. is replaced by $(\pi+p)$-a.e.

In [4], Beiglböck *et al.* investigate the same problem with a different approach which is still based on a saddle point method. A characterization of optimal plans which is more explicit than Theorem 4.2 is obtained under some restrictive assumptions. In particular, the problem of considering simultaneously all the measures $p\in P(\mu,\nu,c)$ in Theorem 1.9 is handled by means of a projective limit argument.

1.3. Notation

Let X and Y be topological vector spaces and $f:X\to[-\infty,+\infty]$ be an extended-real valued function.

- The algebraic dual space of X is X^*, the topological dual space of X is X'.
- The topology of X weakened by Y is $\sigma(X,Y)$ and one writes $\langle X,Y\rangle$ to specify that X and Y are in separating duality.
- The convex conjugate of f with respect to $\langle X,Y\rangle$ is $f^*(y)=\sup_{x\in X}\{\langle x,y\rangle-f(x)\}\in[-\infty,+\infty]$, $y\in Y$.
- The subdifferential of f at x with respect to $\langle X,Y\rangle$ is $\partial_Y f(x)=\{y\in Y;\ f(x+\xi)\ge f(x)+\langle y,\xi\rangle,\ \forall\xi\in X\}$. If no confusion occurs, one writes $\partial f(x)$.
- The effective domain of f is $\mathrm{dom}\,f=\{x\in X;\ f(x)<\infty\}$.
- One denotes $\mathrm{icordom}\,f$ the intrinsic core of the effective domain of f $\mathrm{dom}\,f$. Recall that the intrinsic core of a subset A of a vector space is $\mathrm{icor}\,A=\{x\in A;\ \forall x'\in\mathrm{aff}\,A,\ \exists t>0,\ [x,x+t(x'-x)[\subset A\}$ where $\mathrm{aff}\,A$ is the affine space spanned by A.
- The domain of ∂f is $\mathrm{dom}\partial f=\{x\in X;\partial f(x)\neq\emptyset\}$.
- The indicator function of a subset A of X is defined by

$$\iota_A(x)=\begin{cases}0, & \text{if } x\in A\\ +\infty, & \text{otherwise,}\end{cases}\qquad x\in X.$$

- The support function of $A\subset X$ is $\iota_A^*(y)=\sup_{x\in A}\langle x,y\rangle$, $y\in Y$.
- The Dirac measure at a is denoted δ_a.

https://doi.org/10.1051/cocv/2010013 Published online by Cambridge University Press

2. The abstract convex minimization problem

The Monge-Kantorovich problem is a particular instance of an abstract convex minimization problem which is solved in [12] by means of an extension of the saddle-point method. This extension allows to remove standard topological restrictions on the constraint sets (the so-called constraint qualifications) at the price of solving an arising new problem. Namely, one has to compute a lower semicontinuous convex extension of the convex conjugate of the objective function. This may be a rather difficult task in some situations, but it is immediate in the case of the Monge-Kantorovich problem.

Let us recall the main results of [12].

2.1. The minimization problem

Let $\mathcal{U}_o$, $\mathcal{Y}_o$ be two vector spaces,

$$\Phi_o : \mathcal{U}_o \to [0, \infty]$$

an extended-nonnegative function on $\mathcal{U}_o$, and

$$T_o^* : \mathcal{Y}_o \to \mathcal{U}_o$$

a linear mapping from $\mathcal{Y}_o$ to $\mathcal{U}_o$. Throughout this section, it is assumed that

(H_Φ) (1) $\Phi_o : \mathcal{U}_o \to [0, +\infty]$ is convex and $\Phi_o(0) = 0$;
(2) $\forall u \in \mathcal{U}_o,\ \exists \alpha > 0,\ \Phi_o(\alpha u) < \infty$;
(3) $\forall u \in \mathcal{U}_o,\ u \neq 0,\ \exists t \in \mathbb{R},\ \Phi_o(tu) > 0$;
(H_T) (1) $T_o^*(\mathcal{Y}_o) \subset \mathcal{U}_o$;
(2) $\ker T_o^* = \{0\}$.

Let $\mathcal{L}_o$ and $\mathcal{X}_o$ be the *algebraic* dual vector spaces of $\mathcal{U}_o$ and $\mathcal{Y}_o$. The convex conjugate of Φ_o with respect to the dual pairing $\langle \mathcal{U}_o, \mathcal{L}_o \rangle$ is

$$\Phi_o^*(\ell) = \sup_{u \in \mathcal{U}_o} \{\langle u, \ell \rangle - \Phi_o(u)\} \in [0, \infty], \quad \ell \in \mathcal{L}_o$$

and the adjoint of T_o^* is $T_o : \mathcal{L}_o \to \mathcal{X}_o$ defined for all $\ell \in \mathcal{L}_o$ by $\langle y, T_o \ell \rangle_{\mathcal{Y}_o, \mathcal{X}_o} = \langle T_o^* y, \ell \rangle_{\mathcal{U}_o, \mathcal{L}_o}$. We obtain the diagram

$$\begin{array}{ccc} \left\langle \mathcal{U}_o, & & \mathcal{L}_o \right\rangle \\ T_o^* \uparrow & & \downarrow T_o \\ \left\langle \mathcal{Y}_o, & & \mathcal{X}_o \right\rangle. \end{array} \qquad \text{(Diagram 0)}$$

For each $x \in \mathcal{X}_o$, the optimization problem to be considered is

$$\text{minimize } \Phi_o^*(\ell) \text{ subject to } T_o \ell = x, \qquad \ell \in \mathcal{L}_o. \qquad (P_o^x)$$

2.2. A dual problem

Let us define

$$\begin{aligned} \Lambda_o(y) &:= \Phi_o(T_o^* y), \quad y \in \mathcal{Y}_o \\ |u|_\Phi &:= \inf \left\{ a > 0;\ \max \left(\Phi_o(-u/a), \Phi_o(u/a) \right) \leq 1 \right\}, \quad u \in \mathcal{U}_o \\ |y|_\Lambda &:= \inf \left\{ a > 0;\ \max \left(\Lambda_o(-y/a), \Lambda_o(y/a) \right) \leq 1 \right\}, \quad y \in \mathcal{Y}_o. \end{aligned}$$

Under our assumptions, $|\cdot|_\Phi$ and $|\cdot|_\Lambda$ are norms. One introduces

– $\mathcal{U}$ and $\mathcal{Y}$ the completions of $(\mathcal{U}_o, |\cdot|_\Phi)$ and $(\mathcal{Y}_o, |\cdot|_\Lambda)$;
– The *topological* dual spaces of $\mathcal{U}$ and $\mathcal{Y}$ are $\mathcal{U}' = \mathcal{L}$ and $\mathcal{Y}' = \mathcal{X}$;

https://doi.org/10.1051/cocv/2010013 Published online by Cambridge University Press

– The *algebraic* dual spaces of $\mathcal{L}$ and $\mathcal{X}$ are denoted by $\mathcal{L}^*$ and $\mathcal{X}^*$;
– T is the restriction of T_o to $\mathcal{L} \subset \mathcal{L}_o$;
– Φ^* is the restriction of Φ_o^* to $\mathcal{L} \subset \mathcal{L}_o$.

It is proved in [12] that $T\mathcal{L} \subset \mathcal{X}$. This allows one to define the algebraic adjoint $T^* : \mathcal{L}^* \to \mathcal{X}^*$. We have the following diagram

$$\begin{array}{ccc} \Big\langle\ \mathcal{L}, & & \mathcal{L}^*\ \Big\rangle \\ T\Big\downarrow & & \Big\uparrow T^* \\ \Big\langle\ \mathcal{X}, & & \mathcal{X}^*\ \Big\rangle. \end{array} \qquad \text{(Diagram 1)}$$

Also consider

$$\left\{ \begin{array}{lcll} \overline{\Phi}(\zeta) & := & \sup_{\ell\in\mathcal{L}}\{\langle\zeta,\ell\rangle - \Phi^*(\ell)\}, & \zeta \in \mathcal{L}^* \\ \overline{\Lambda}(\omega) & := & \overline{\Phi}(T^*\omega), & \omega \in \mathcal{X}^*. \end{array} \right. \tag{2.1}$$

The maximization problem

$$\textit{maximize } \langle\omega, x\rangle - \overline{\Lambda}(\omega), \quad \omega \in \mathcal{X}^* \tag{$\overline{D}^x$}$$

is the dual problem associated with

$$\textit{minimize } \Phi^*(\ell) \textit{ subject to } T\ell = x, \qquad \ell \in \mathcal{L} \tag{P^x}$$

with respect to the usual Fenchel perturbation.

2.3. **Statements of the abstract results**

Let us introduce

$$\Lambda^*(x) := \sup_{y\in\mathcal{Y}_o}\{\langle y, x\rangle - \Lambda_o(y)\}, \quad x \in \mathcal{X}$$

the convex conjugate of $\Lambda_o + \iota_{\mathcal{Y}_o}$ with respect to the dual pairing $\langle\mathcal{Y}, \mathcal{X}\rangle$.

Theorem 2.1 (primal attainment and dual equality [12])**.** *Assume that* (H_Φ) *and* (H_T) *hold.*

(a) *For all* x *in* $\mathcal{X}$*, we have the dual equality*

$$\inf\{\Phi^*(\ell);\ \ell \in \mathcal{L},\ T\ell = x\} = \inf(P_o^x) = \inf(P^x) = \Lambda^*(x) \in [0, \infty]. \tag{2.2}$$

Moreover, Λ^* *is* $\sigma(\mathcal{X}, \mathcal{Y})$*-inf-compact.*

(b) *If* $x \notin \mathcal{X}$*,* (P_o^x) *admits no solution. Otherwise,* P_o^x *and* P^x *admit the same (possibly empty) set of solutions.*

(c) *If in addition*

$$x \in \operatorname{dom}\Lambda^*,$$

then P_o^x *(or equivalently* P^x*) is attained in* $\mathcal{L}$*. Moreover, any minimizing sequence for* P_o^x *has* $\sigma(\mathcal{L},\mathcal{U})$*-cluster points and every such cluster point solves* P_o^x*.*

Denote the "algebraic" subdifferentials

$$\begin{array}{rcl} \partial_{\mathcal{L}}\overline{\Phi}(\zeta_o) & = & \{\ell \in \mathcal{L};\ \overline{\Phi}(\zeta_o + \zeta) \geq \overline{\Phi}(\zeta_o) + \langle\zeta, \ell\rangle,\ \forall\zeta \in \mathcal{L}^*\}, \quad \zeta_o \in \mathcal{L}^* \\ \partial_{\mathcal{X}^*}\Lambda^*(x_o) & = & \{\omega \in \mathcal{X}^*;\ \Lambda^*(x_o + x) \geq \Lambda^*(x_o) + \langle\omega, x\rangle,\ \forall x \in \mathcal{X}\}, \quad x_o \in \mathcal{X} \end{array}$$

and define

$$\operatorname{dom}\partial\Lambda^* = \{x \in \mathcal{X};\ \partial_{\mathcal{X}^*}\Lambda^*(x) \neq \emptyset\}$$

the subset of the constraint specifiers $x \in \mathcal{X}$ such that $\partial_{\mathcal{X}^*}\Lambda^*(x)$ is not empty.

https://doi.org/10.1051/cocv/2010013 Published online by Cambridge University Press

Theorem 2.2 (dual attainment and representation [12]). *Let us assume that* (H_Φ) *and* (H_T) *hold.*

(1) *For any* $\ell \in \mathcal{L}$ *and* $\omega \in \mathcal{X}^*$,

$$\begin{cases} T\ell = x \\ \ell \in \partial_{\mathcal{L}} \overline{\Phi}(T^*\omega) \end{cases} \tag{2.3}$$

is equivalent to

$$\begin{cases} \ell \text{ is a solution to } P^x, \\ \omega \text{ is a solution to } \overline{D}^x \text{ and} \\ \text{the dual equality } \inf(P^x) = \Lambda^*(x) \text{ holds.} \end{cases}$$

(2) *Suppose that in addition the constraint qualification*

$$x \in \mathrm{dom}\partial\Lambda^*, \tag{2.4}$$

is satisfied. Then, the primal problem P^x *is attained in* $\mathcal{L}$, *the dual problem* $\overline{D}^x$ *is attained in* $\mathcal{X}^*$ *and every couple of solutions* (ℓ, ω) *to* P^x *and* $\overline{D}^x$ *satisfies* (2.3).

By the geometric version of Hahn-Banach theorem, $\mathrm{icordom}\,\Lambda^*$: the intrinsic core of $\mathrm{dom}\,\Lambda^*$, is included in $\mathrm{dom}\partial\Lambda^*$. But, as will be seen at Remark 4.3 below, the Monge-Kantorovich problem provides us with a situation where $\mathrm{icordom}\,\Lambda^*$ is empty. This is one of the main difficulties to be overcome when applying the saddle-point method to solve (MK).

It is a well-known result of convex conjugacy that the representation formula

$$\ell \in \partial_{\mathcal{L}} \overline{\Phi}(T^*\omega) \tag{2.5}$$

is equivalent to

$$T^*\omega \in \partial_{\mathcal{L}^*} \Phi^*(\ell)$$

and also equivalent to Fenchel's identity

$$\Phi^*(\ell) + \overline{\Phi}(T^*\omega) = \langle \omega, T\ell \rangle. \tag{2.6}$$

Proposition 2.3 ([12]). *Assume that* (H_Φ) *and* (H_T) *hold. Any solution* $\omega \in \mathcal{X}^*$ *of* $\overline{D}^x$ *shares the following properties:*

(a) ω *is in the* $\sigma(\mathcal{X}^*, \mathcal{X})$*-closure of* $\mathrm{dom}\,\Lambda$;
(b) $T^*\omega$ *is in the* $\sigma(\mathcal{L}^*, \mathcal{L})$*-closure of* $T^*(\mathrm{dom}\,\Lambda)$.

If in addition the level sets of Φ_o *are* $|\cdot|_\Phi$*-bounded, then*

(a') ω *is in* $\mathcal{Y}''$. *More precisely, it is in the* $\sigma(\mathcal{Y}'', \mathcal{X})$*-closure of* $\mathrm{dom}\,\Lambda$;
(b') $T^*\omega$ *is in* $\mathcal{U}''$. *More precisely, it is in the* $\sigma(\mathcal{U}'', \mathcal{L})$*-closure of* $T^*(\mathrm{dom}\,\Lambda)$

where $\mathcal{Y}''$ *and* $\mathcal{U}''$ *are the topological bidual spaces of* $(\mathcal{U}, |\cdot|_\Phi)$ *and* $(\mathcal{Y}, |\cdot|_\Lambda)$.

3. Kantorovich dual equality and existence of optimal plans

We apply the results of Section 2 to the Monge-Kantorovich problem (MK).

3.1. Statement of the results

The set of all probability measures π on $A \times B$ such that $\int_{A\times B} c\,\mathrm{d}\pi < \infty$ is denoted $\mathcal{P}_c$. Hence, $P(\mu, \nu, c) = \mathcal{P}_c \cap P(\mu, \nu)$. It is immediate to see that

$$\sup\,(\mathrm{K}) \le \sup(\widetilde{\mathrm{K}}) \le \inf\,(\mathrm{MK}) \tag{3.1}$$

where for the last inequality, we state $\inf \emptyset = +\infty$ as usual.

https://doi.org/10.1051/cocv/2010013 Published online by Cambridge University Press

In the next theorem, $\mathcal{P}_c$ will be endowed with the weak topology $\sigma(\mathcal{P}_c, \mathcal{C}_c)$ where $\mathcal{C}_c$ is the space of all continuous functions u on $A \times B$ such that

$$\lim_{t\to\infty} \sup\{|u(x)|/(1+c(x));\ x \in A\times B,\ |u(x)| \geq t\} = 0$$

with the convention that $\sup \emptyset = 0$ which implies that the space C_{AB} of bounded continuous functions on $A \times B$ satisfies $C_{AB} \subset \mathcal{C}_c$.

Theorem 3.1 (dual equality and primal attainment)**.**

(1) *The dual equality for* (MK) *is*

$$\inf (\mathrm{MK}) = \sup (\mathrm{K}) = \sup (\widetilde{\mathrm{K}}) \in [0, \infty]. \tag{3.2}$$

(2) *Assume that there exists some* π^o *in* $P(\mu, \nu)$ *such that* $\int_{A\times B} c \, \mathrm{d}\pi^o < \infty$. *Then:*
   (a) *there is at least one optimal plan;*
   (b) *any minimizing sequence is relatively compact for the topology* $\sigma(\mathcal{P}_c, \mathcal{C}_c)$ *and all its cluster points are optimal plans.*

Proof. It directly follows from Theorem 2.1 with Proposition 3.5-(3) and Proposition 3.7 which are stated and proved below. About statement (1), note that the dual equality (2.2) is $\inf (\mathrm{MK}) = \sup (\mathrm{K})$ and conclude with (3.1). □

Except for the appearance of the dual problem $(\widetilde{\mathrm{K}})$ and for the statement (2)-(b) with the topology $\sigma(\mathcal{P}_c, \mathcal{C}_c)$, this result is well-known. The dual equality (3.2) is the Kantorovich dual equality. The proof of Theorem 3.1 will be an opportunity to make precise the abstract material $\Phi_o, \mathcal{U}_o, T_o$ and so on in terms of the Monge-Kantorovich problem.

3.2. **The spaces and functions**

Let us particularize the spaces and functions of Section 2 which are necessary for solving the Monge-Kantorovich problem.

3.2.1. *Description of* $\mathcal{U}_o$, $\mathcal{L}_o$, Φ_o *and* Φ_o^*

Define

$$\mathcal{S} := \{(a, b) \in A\times B;\ c(a, b) < \infty\}$$

the effective domain of c and consider the equivalence relation on C_{AB} defined for any u, v in C_{AB} by

$$u \sim v \Leftrightarrow u_{|\mathcal{S}} = v_{|\mathcal{S}},$$

which means that u and v match on $\mathcal{S}$. Take

$$\mathcal{U}_o := C_{AB}/\sim$$

the set of all equivalence classes and

$$\Phi_o(u) := \iota_\Gamma(u) = \begin{cases} 0, & \text{if } u \leq c \\ +\infty, & \text{otherwise,} \end{cases} \qquad u \in \mathcal{U}_o$$

the convex indicator function of

$$\Gamma = \{u \in \mathcal{U}_o;\ u \leq c\}.$$

https://doi.org/10.1051/cocv/2010013 Published online by Cambridge University Press

Assumption $(H_{\Phi 1})$ is obviously satisfied and it is easy to see that $(H_{\Phi 3})$ also holds. Note that, without factorizing by the equivalence relation in the definition of $\mathcal{U}_o$, $(H_{\Phi 3})$ would fail if c is infinite on some nonempty open set. For $(H_{\Phi 2})$ to be satisfied, assume that

$$c \geq 1.$$

Doing this, one doesn't loose any generality since changing c into $c+1$ raises $\inf(\mathbb{MK})$ by $+1$ without changing the minimizers.

The algebraic dual space $\mathcal{L}_o = \mathcal{U}_o^*$ of $\mathcal{U}_o$ is identified as a subset of C_{AB}^* as follows. Any $\ell \in C_{AB}^*$ is in $\mathcal{L}_o$ if and only if for all $u, v \in C_{AB}$ such that $u_{|\mathcal{S}} = v_{|\mathcal{S}}$, we have $\ell(u) = \ell(v)$. One shows at Proposition 3.3 below that any $\ell \in C_{AB}^*$ which is in $\mathcal{L}_o$ has its support included in the closure $\operatorname{cl}\mathcal{S}$ of $\mathcal{S}$ in the sense of the following:

Definition 3.2. For any linear form ℓ on the space of all (possibly unbounded) continuous functions on $A\times B$, we define the *support* of ℓ as the subset of all $(a,b) \in A\times B$ such that for any open neighborhood G of (a,b), there exists some function u in C_{AB} satisfying $\{u \neq 0\} \subset G$ and $\langle u, \ell\rangle \neq 0$. It is denoted by $\operatorname{supp}\ell$.

Proposition 3.3. *For any $\ell \in \mathcal{L}_o$, we have* $\operatorname{supp}\ell \subset \operatorname{cl}\mathcal{S}$.

Proof. As $\ell \in \mathcal{L}_o$, for all $u \in C_{AB}$ such that $u_{|\mathcal{S}} = 0$, we have $\langle u, \ell\rangle = 0$. Take $x \notin \operatorname{cl}\mathcal{S}$. There exists an open neighborhood G of x such that $G \cap \mathcal{S} = \emptyset$. For all $u \in C_{AB}$ such that $\{u \neq 0\} \subset G$, we have $u_{|\mathcal{S}} = 0$ so that $\langle u, \ell\rangle = 0$. This proves that $x \notin \operatorname{supp}\ell$. □

Denote $C_{|\mathcal{S}}$ the space of the restrictions $u_{|\mathcal{S}}$ of all the functions u in C_{AB}.

Remark 3.4. Beware, $C_{|\mathcal{S}}$ is smaller than the space $C_{\mathcal{S}}$ of all bounded and continuous functions on $\mathcal{S}$. It consists of the functions in $C_{\mathcal{S}}$ which admit a continuous and bounded extension to $A\times B$.

Identifying

$$\mathcal{U}_o \cong C_{|\mathcal{S}}, \tag{3.3}$$

one sees with Proposition 3.3 that for all $\ell \in \mathcal{L}_o$,

$$\Phi_o^*(\ell) := \sup\{\langle u, \ell\rangle;\ u \in C_{|\mathcal{S}},\ u \leq c\}.$$

3.2.2. *Description of $\mathcal{Y}_o$, $\mathcal{X}_o$, T_o^* and T_o*

Consider

$$\mathcal{Y}_o := (C_A \times C_B)/\sim$$

where C_A, C_B are the spaces of bounded continuous functions on A, B and $\sim$ is a new equivalence relation on $C_A \times C_B$ defined for all $f, f' \in C_A$, $g,\ g' \in C_B$ by

$$(f,g) \sim (f',g') \overset{\text{def}}{\Leftrightarrow} f \oplus g \sim f' \oplus g' \Leftrightarrow f \oplus g_{|\mathcal{S}} = f' \oplus g'_{|\mathcal{S}}.$$

The algebraic dual space $\mathcal{X}_o = \mathcal{Y}_o^*$ of $\mathcal{Y}_o$ is identified as a subset of $C_A^* \times C_B^*$ as follows. Any $(k_1, k_2) \in C_A^* \times C_B^*$ is in $\mathcal{X}_o$ if and only if for all $f, f' \in C_A$, $g, g' \in C_B$ such that $f \oplus g_{|\mathcal{S}} = f' \oplus g'_{|\mathcal{S}}$, we have $\langle f, k_1\rangle + \langle g, k_2\rangle = \langle f', k_1\rangle + \langle g', k_2\rangle$.

It is immediate to see that the operator T_o^* defined by

$$T_o^*(f,g) := f \oplus g_{|\mathcal{S}}, \quad (f,g) \in \mathcal{Y}_o$$

satisfies the assumptions (H_T). In particular, one sees that $T_o^*(\mathcal{Y}_o) \subset C_{|\mathcal{S}}$.

Define the A and B-marginal projections $\ell_A \in C_A^*$ and $\ell_B \in C_B^*$ of any $\ell \in C_{AB}^*$ by

$$\langle f, \ell_A\rangle := \langle f \otimes 1, \ell\rangle \quad \text{and} \quad \langle g, \ell_B\rangle := \langle 1 \otimes g, \ell\rangle$$

https://doi.org/10.1051/cocv/2010013 Published online by Cambridge University Press

for all $f \in C_A$ and all $g \in C_B$. Since for all $(f,g) \in \mathcal{Y}_o$, $\langle T_o^*(f,g), \ell\rangle_{C_{|\mathcal{S}},\mathcal{L}_o} = \langle f \oplus g, \ell\rangle_{C_{|\mathcal{S}},\mathcal{L}_o} = \langle f, \ell_A\rangle + \langle g, \ell_B\rangle = \langle (f,g), (\ell_A,\ell_B)\rangle$, one obtains that for all $\ell \in \mathcal{L}_o \subset C_{AB}^*$,

$$T_o(\ell) = (\ell_A, \ell_B) \in \mathcal{X}_o \subset C_A^* \times C_B^*.$$

3.2.3. *Description of $\mathcal{U}$ and $\mathcal{Y}$*

For each $u \in C_{|\mathcal{S}}$, $\Phi_o^{\pm}(u) := \max(\Phi_o(u), \Phi_o(-u)) = \iota_{\{|u| \le c\}}$. Hence, $|u|_\Phi$ is the norm

$$\|u\|_c := \sup_{\mathcal{S}} |u|/c, \quad u \in C_{|\mathcal{S}}.$$

Recall that it is assumed that $c \ge 1$. Let us denote the Banach space

$$\mathcal{U} := U_c$$

which is the completion of $C_{|\mathcal{S}}$ with respect to the norm $\|\cdot\|_c$.

Proposition 3.5.

(1) *U_c is a space of functions on $\mathcal{S}$ which are continuous in restriction to each level set $\{c \le \alpha\}$, $\alpha \ge 1$. In particular, any $u \in U_c$ is a measurable function.*
(2) *Even if c is finite everywhere, U_c may contain functions which are not continuous.*
(3) $\mathcal{C}_c \subset U_c$.

Proof.

• Proof of (1). Let $(u_n)_{n\ge1}$ be a $\|\cdot\|_c$-Cauchy sequence in $C_{|\mathcal{S}}$. For all $x \in \mathcal{S}$, $\lim_{n\to\infty} u_n(x) := u(x)$ exists since $(u_n(x))_{n\ge1}$ is a Cauchy sequence in $\mathbb{R}$. For each $\epsilon > 0$, and there are large enough n, m such that $\sup_{x\in\mathcal{S}} |u_n(x) - u_m(x)|/c(x) \le \epsilon$ for all $x \in \mathcal{S}$. Letting m tend to infinity leads us to $|u_n(x) - u(x)|/c(x) \le \epsilon$ for all $x \in \mathcal{S}$, which gives us

$$\lim_{n\to\infty} \|u_n - u\|_c = 0$$

where the definition $\|v\|_c := \sup_{\mathcal{S}} |v|/c$ still holds for any function v on $\mathcal{S}$. As for any x, $y \in \mathcal{S}$,

$$\frac{u(x) - u(y)}{c(x) + c(y)} \le 2\|u - u_n\|_c + \frac{u_n(x) - u_n(y)}{c(x) + c(y)} \le 2\|u - u_n\|_c + |u_n(x) - u_n(y)|,$$

the announced continuity result follows from the above limit. The measurability statement follows from this continuity result and the measurability of c.

• Proof of (2). To see this, take:

* $A = B = \mathbb{R}$;
* $c(a,b) = 1 + 1/a^2$ for any (a,b) with $a \ne 0$, and $c(0,b) = 1$ for all real b;
* $0 \le u_n \le 1$ with $u_n(a,b) = 0$ if $a \le -1/n$, $u_n(a,b) = 1$ if $a \ge 1/n$ and $u_n(0,b) = 1/2$ for all real b.

Then, $(u_n)_{n\ge1}$ admits the limit u in U_c with $u(a,b) = 0$ if $a < 0$, $u(a,b) = 1$ if $a > 0$ and $u(0,b) = 1/2$.

• Proof of (3). By the very definition of $\mathcal{C}_c$, one sees that for any $u \in \mathcal{C}_c$, defining $u_n := (-n) \vee u \wedge n \in C_{AB}$, the sequence $(u_n)_{n\ge1}$ converges to u in U_c. □

The norm on $\mathcal{Y}_o$ is given by

$$|(f,g)|_\Lambda := \|f \oplus g\|_c = \sup_{\mathcal{S}} |f \oplus g|/c, \quad (f,g) \in \mathcal{Y}_o.$$

Identifying

$$\mathcal{Y}_o \cong (C_A \oplus C_B)_{|\mathcal{S}}$$

by means of $(f,g) \in \mathcal{Y}_o \mapsto f \oplus g_{|\mathcal{S}} \in (C_A \oplus C_B)_{|\mathcal{S}} \subset C_{|\mathcal{S}}$, the $|\cdot|_\Lambda$-completion $\mathcal{Y}$ of $\mathcal{Y}_o$ is identified with the $\|\cdot\|_c$-closure of $(C_A \oplus C_B)_{|\mathcal{S}}$ in U_c.

https://doi.org/10.1051/cocv/2010013 Published online by Cambridge University Press

3.2.4. *Description of* $\mathcal{L}$, $\mathcal{X}$, Λ *and* Λ^*

The topological dual space of U_c

$$\mathcal{L} := U_c'$$

is equipped with the dual norm

$$\|\ell\|_c^* = \sup\{\langle u, \ell\rangle;\ u \in U_c,\ \|u\|_c \le 1\}, \quad \ell \in U_c'.$$

The function Λ is given by

$$\Lambda(f,g) = \iota_{\{f\oplus g \le c\}}, \quad f \oplus g \in (C_A \oplus C_B)_{|\mathcal{S}}$$

and the corresponding norm is

$$|(f,g)|_\Lambda = \|f \oplus g\|_c, \quad f \oplus g \in (C_A \oplus C_B)_{|\mathcal{S}}.$$

The topological dual space of $\mathcal{Y}$ is

$$\mathcal{X} = \{(k_1,k_2) \in \mathcal{X}_o;\ |(k_1,k_2)|_\Lambda^* < \infty\}$$

with

$$\begin{aligned} |(k_1,k_2)|_\Lambda^* &= \sup\{\langle f,k_1\rangle + \langle g,k_2\rangle;\ (f,g) \in \mathcal{Y}_o,\ \|f\oplus g\|_c \le 1\} \\ &= \inf\{\|\ell\|_c^*;\ \ell \in U_c' : \ell_A = k_1,\ \ell_B = k_2\} \end{aligned}$$

where this last equality is a direct consequence of the dual equality (2.2). Hence, $k = (k_1,k_2) \in C_A^* \times C_B^*$ belongs to $\mathcal{X}$ if and only if there exists some $\ell(k) \in \mathcal{L}$ such that $\langle f,k_1\rangle + \langle g,k_2\rangle = \langle f\oplus g, \ell(k)\rangle$ for all $(f,g) \in \mathcal{Y}_o$ and $\lim_{n\to\infty}\langle u_n, \ell(k)\rangle = 0$ for any sequence $(u_n)_{n\ge1}$ in C_{AB} such that $\lim_{n\to\infty}\|u_n\|_c = 0$.

The function Λ^* is given by

$$\Lambda^*(k_1,k_2) = \sup\{\langle f,k_1\rangle + \langle g,k_2\rangle;\ (f,g) \in \mathcal{Y}_o,\ f\oplus g \le c\}, \quad k = (k_1,k_2) \in \mathcal{X}.$$

3.2.5. *Description of* $\mathcal{X}^*$ *and* T^*

Seeing $\mathcal{X}$ as a subspace of $C_A^* \times C_B^*$, $\mathcal{X}^*$ is a subspace of $(C_A^* \times C_B^*)^*$. By the axiom of choice, there exists a subspace $\mathcal{Z}$ of $C_A^* \times C_B^*$ in direct sum with $\mathcal{X} : C_A^* \times C_B^* = \mathcal{X} \oplus \mathcal{Z}$. To any $\omega \in \mathcal{X}^*$, one associates its extension $\bar\omega$ to $C_A^* \times C_B^*$ characterized by $\bar\omega_{|\mathcal{Z}} = 0$ and $\bar\omega_{|\mathcal{X}} = \omega$. This permits us to define the marginal projections ω_A and ω_B and their tensor sum $\omega_A \oplus \omega_B \in U_c'^*$ as follows. For all $(k_1,k_2) \in \mathcal{X}$,

$$\begin{aligned} \langle \omega, (k_1,k_2)\rangle_{\mathcal{X}^*,\mathcal{X}} &= \langle \bar\omega, (k_1,k_2)\rangle_{(C_A^*\times C_B^*)^*, C_A^*\times C_B^*} \\ &= \langle \bar\omega, (k_1,0) + (0,k_2)\rangle_{(C_A^*\times C_B^*)^*, C_A^*\times C_B^*} \\ &:= \langle \omega_A, k_1\rangle_{C_A^{**}, C_A^*} + \langle \omega_B, k_2\rangle_{C_B^{**}, C_B^*} \\ &:= \langle \omega_A \oplus \omega_B, \ell(k)\rangle_{U_c'^*, U_c'}. \end{aligned}$$

For any $\omega \in \mathcal{X}^*$ and $\ell \in U_c'$, $\langle T^*\omega, \ell\rangle_{U_c'^*, U_c'} = \langle \omega, (\ell_A,\ell_B)\rangle_{\mathcal{X}^*,\mathcal{X}} := \langle \omega_A \oplus \omega_B, \ell\rangle_{U_c'^*, U_c'}$ which means that

$$T^*\omega = \omega_A \oplus \omega_B \in U_c'^*, \quad \omega \in \mathcal{X}^*. \tag{3.4}$$

3.3. **The connection with (MK)**

The connection with the Monge-Kantorovich problem is given at Proposition 3.7 below. The modified primal problem is

$$\textit{minimize } \Phi^*(\ell) \textit{ subject to } \ell_A = \mu \textit{ and } \ell_B = \nu, \quad \ell \in U_c' \tag{P_1}$$

https://doi.org/10.1051/cocv/2010013 Published online by Cambridge University Press

where

$$\Phi^*(\ell) = \sup\{\langle u, \ell\rangle;\ u \in U_c,\ u \le c\}, \quad \ell \in U_c'.$$

Definition 3.6.

(a) One says that $\ell \in U_c'$ *acts as a probability measure* if there exists $\tilde{\ell} \in \mathcal{P}_{AB}$ such that $\operatorname{supp}\tilde{\ell} \subset \operatorname{cl}\mathcal{S}$ and for all $u \in C_{AB}$, $\langle u_{|\mathcal{S}}, \ell\rangle = \int_{\mathcal{S}} u \,\mathrm{d}\tilde{\ell}$. In this case, we write: $\ell \in \mathcal{P}_{\mathcal{S}}$.

(b) One says that $\ell \in U_c'$ stands in $\mathcal{P}_c$ if there exists $\tilde{\ell} \in \mathcal{P}_c$ such that $\operatorname{supp}\tilde{\ell} \subset \operatorname{cl}\mathcal{S}$ and for all $u \in U_c$, $\langle u_{|\mathcal{S}}, \ell\rangle = \int_{\mathcal{S}} u \,\mathrm{d}\tilde{\ell}$. In this case, we write: $\ell \in \mathcal{P}_c$.

Of course, if there exists $\tilde{\ell}$ satisfying (a), it belongs to $\mathcal{P}_c$ and is unique since any probability measure on a metric space is determined by its values on the continuous bounded functions. This explains why the notation $\ell \in \mathcal{P}_c$ in (b) is not misleading.

Note also that any probability measure $\tilde{\ell} \in \mathcal{P}_c$ has a support included in $\operatorname{cl}\mathcal{S}$. Since $A\times B$ is a metric space, for any $\ell \in \mathcal{P}_c$ acting as a measure, $\operatorname{supp}\ell$ in the sense of Definition 3.2 matches with the usual support of the measure $\tilde{\ell}$.

Proposition 3.7. *For all* $\ell \in U_c'$,

(a) $\Phi^*(\ell) < \infty \Rightarrow \ell \ge 0$;

(b) $\Phi^*(\ell) < \infty \Rightarrow \operatorname{supp}\ell \subset \operatorname{cl}\mathcal{S}$;

(c) $[\ell \ge 0,\ \operatorname{supp}\ell \subset \operatorname{cl}\mathcal{S},\ \ell_A = \mu \text{ and } \ell_B = \nu] \Rightarrow \ell \in \mathcal{P}_{\mathcal{S}}$; *and*

(d) *for all* $\ell \in \mathcal{P}_{\mathcal{S}}$, $\Phi^*(\ell) = \int_{\mathcal{S}} c \,\mathrm{d}\ell \in [0,\infty]$.

It follows that $\operatorname{dom}\Phi^* \subset \mathcal{P}_c$ *and the problems* (MK) *and* (P_1) *share the same values and the same minimizers.*

Proof. Clearly, the last statement follows from the first part of the proposition.

• Proof of (a). Suppose that $\ell \in U_c'$ is not in the nonnegative cone. This means that there exists $u_o \in U_c$ such that $u_o \ge 0$ and $\langle u_o, \ell\rangle < 0$. Since u_o satisfies $\lambda u_o \le 0 \le c$ for all $\lambda < 0$, we have $\Phi^*(\ell) \ge \sup_{\lambda<0}\langle \lambda u_o, \ell\rangle = +\infty$. Hence, $\Phi^*(\ell) < \infty$ implies that $\ell \ge 0$ and one can restrict our attention to the nonnegative ℓ's.

• Proof of (b). Suppose *ad absurdum* that $\operatorname{supp}\ell \not\subset \operatorname{cl}\mathcal{S}$. Then, there exists a nonnegative function $u_o \in C_{AB}$ such that $\{u_o > 0\} \cap \mathcal{S} = \emptyset$ and $\langle u_o, \ell\rangle > 0$. As $\lambda u_o \le c_{|A\times B\setminus\mathcal{S}} \equiv \infty$ for all $\lambda > 0$, $\Phi^*(\ell) \ge \sup_{\lambda>0}\langle \lambda u_o, \ell\rangle = +\infty$.

• Proof of (c). Let us take $\ell \ge 0$ such that $\operatorname{supp}\ell \subset \operatorname{cl}\mathcal{S}$, $\ell_A = \mu$ and $\ell_B = \nu$. It is clear that $\langle 1, \ell\rangle = 1$. It remains to check that for any $\ell \in U_c'$

$$[\ell \ge 0,\ \operatorname{supp}\ell \subset \operatorname{cl}\mathcal{S},\ \ell_A = \mu \text{ and } \ell_B = \nu] \Rightarrow \ell \text{ is } \sigma\text{-additive}, \tag{3.5}$$

rather than only additive. Since $A\times B$ is a metric space, one can apply an extension of the construction of Daniell's integrals ([13], Prop. II.7.2) to see that ℓ acts as a measure if and only if for any decreasing sequence (u_n) of continuous functions such that $0 \le u_n \le 1$ for all n and $\lim_{n\to\infty} u_n = 0$ pointwise, we have $\lim_{n\to\infty}\langle u_n, \ell\rangle = 0$. This insures the σ-additivity of ℓ.

Unfortunately, this pointwise convergence of (u_n) is weaker than the uniform convergence with respect to which any $\ell \in U_c'$ is continuous. Except if $A\times B$ is compact, since in this special case, any decreasing sequence of continuous functions which converges pointwise to zero also converges uniformly on the compact space $\operatorname{cl}\mathcal{S}$.

So far, we have only used the fact that $A\times B$ is a metric space. We now rely on the Polishness of A and B to get rid of this compactness restriction. It is known that any probability measure P on a Polish space X is tight (*i.e.* a Radon measure): for all $\epsilon > 0$, there exists a compact set $K_\epsilon \subset X$ such that $P(X \setminus K_\epsilon) \le \epsilon$ ([13], Prop. II.7.3). As in addition a Polish space is completely regular, there exists a continuous function f_ϵ with a compact support such that $0 \le f_\epsilon \le 1$ and $\int_X (1 - f_\epsilon)\,\mathrm{d}P \le \epsilon$. This is true in particular for the probability measures $\mu \in \mathcal{P}_A$ and $\nu \in \mathcal{P}_B$ which specify the constraint in (MK). Hence, there exist $f_\epsilon \in C_A$ and $g_\epsilon \in C_B$ with compact supports such that $0 \le f_\epsilon, g_\epsilon \le 1$ and $0 \le \int_A (1 - f_\epsilon)\,\mathrm{d}\mu$, $\int_B (1 - g_\epsilon)\,\mathrm{d}\nu \le \epsilon$. It follows with the elementary fact: $0 \le 1 - st \le 2 - s - t$ for all $0 \le s, t \le 1$, that any nonnegative $\ell \in U_c'$ with $\ell_A = \mu$ and $\ell_B = \nu$

https://doi.org/10.1051/cocv/2010013 Published online by Cambridge University Press

satisfies $0 \leq \langle (1 - f_\epsilon \otimes g_\epsilon), \ell \rangle \leq 2\epsilon$. With the following easy estimate $0 \leq \langle u_n, \ell \rangle \leq 2\epsilon + \langle u_n (f_\epsilon \otimes g_\epsilon), \ell \rangle$ and the compactness of the support of $f_\epsilon \otimes g_\epsilon$, one concludes that $\lim_{n\to\infty} \langle u_n, \ell \rangle = 0$ which proves (3.5).

• Proof of (d). Let us take $\ell \in \mathcal{P}_\mathcal{S}$. As c is bounded below and lower semicontinuous, there exists a sequence $(c_n)_{n\geq 1}$ in C_{AB} such that $0 \leq c_n \leq c$ converges pointwise and increasingly to c. Therefore,

$$\int_\mathcal{S} c \, \mathrm{d}\ell \overset{(a)}{=} \lim_{n\to\infty} \int_\mathcal{S} c_n \, \mathrm{d}\ell \overset{(b)}{=} \lim_{n\to\infty} \langle c_n, \ell \rangle \leq \Phi^*(\ell) \tag{3.6}$$

where equality (a) holds by monotone convergence and equality (b) holds since c_n belongs to C_{AB}.

Let us show the converse inequality: $\Phi^*(\ell) \leq \int_\mathcal{S} c \, \mathrm{d}\ell$. Let us first assume that $\ell \in \operatorname{dom} \Phi^*$. It is proved at Lemma 3.8 below that for any $u \geq 0$ in U_c, $\langle u, \ell \rangle = \int_\mathcal{S} u \, \mathrm{d}\ell$. It follows that $\Phi^*(\ell) = \sup\{\langle u, \ell \rangle;\ u \in U_c,\ u \leq c\} \leq \int_\mathcal{S} c \, \mathrm{d}\ell$.

If $\Phi^*(\ell) = \infty$, there exists a sequence $(u_n)_{n\geq 1}$ in C_{AB} such that $0 \leq u_n \leq c$ and $\sup_n \int_\mathcal{S} u_n \, \mathrm{d}\ell = \infty$. Therefore, $\int c \, \mathrm{d}\ell \geq \sup_n \int u_n \, \mathrm{d}\ell = \infty$. □

Lemma 3.8. *For any $u \geq 0$ in U_c and any $\ell \in \mathcal{P}_\mathcal{S} \cap U_c'$ such that $\Phi^*(\ell) < \infty$, we have $\langle u, \ell \rangle = \int_\mathcal{S} u \, \mathrm{d}\ell$.*

Proof. There exists a sequence $(u_n)_{n\geq 1}$ in C_{AB} such that $\lim_{n\to\infty} u_n = u$ in U_c. As ℓ belongs to U_c',

$$\lim_{n\to\infty} \int_\mathcal{S} u_n \, \mathrm{d}\ell = \lim_{n\to\infty} \langle u_n, \ell \rangle = \langle u, \ell \rangle. \tag{3.7}$$

On the other hand, $|u_n - u| \leq \|u_n - u\|_c c$ implies that $|u_n| \leq [\|u\|_c + \|u_n - u\|_c]c$. Hence, for some large enough n_o,

$$|u_n| \leq (1 + \|u\|_c)c, \quad \forall n \geq n_o.$$

Together with (3.6), the assumption that $\Phi^*(\ell) < \infty$ and the dominated convergence theorem, this entails $\lim_{n\to\infty} \int_\mathcal{S} u_n \, \mathrm{d}\ell = \int_\mathcal{S} u \, \mathrm{d}\ell$. This and (3.7) lead us to $\langle u, \ell \rangle = \int_\mathcal{S} u \, \mathrm{d}\ell$ which is the desired result. □

4. An abstract characterization of optimality

The abstract characterization of optimality is stated in Theorem 4.2. It will allow us to obtain as corollaries, an explicit sufficient condition in Theorem 1.8 and an explicit necessary condition in Theorem 1.9. To prove it, one has to compute the extension $\overline{\Phi}$. As it is the greatest convex $\sigma(U_c'^*, U_c')$-lower semicontinuous extension of $\Phi = \iota_\Gamma$ and $\Gamma = \{u \in \mathcal{U}_o;\ u \leq c\}$ is a convex subset of $\mathcal{U}_o$, we have

$$\overline{\Phi}(\xi) = \iota_{\overline{\Gamma}}(\xi), \quad \xi \in U_c'^* \tag{4.1}$$

where $\overline{\Gamma}$ is the $\sigma(U_c'^*, U_c')$-closure of Γ. By (3.4), this gives

$$\overline{\Lambda}(\omega) = \iota_{\overline{\Gamma}}(\omega_A \oplus \omega_B), \quad \omega \in \mathcal{X}^*$$

and the extended problem $(\overline{D}^x)$ is

$$\textit{maximize } \langle \omega_A, \mu \rangle + \langle \omega_B, \nu \rangle, \quad \omega \in \mathcal{X}^* \textit{ such that } \omega_A \oplus \omega_B \in \overline{\Gamma}. \tag{$\overline{\mathrm{K}}$}$$

Note that for this dual problem to be meaningful, it is necessary that $(\mu, \nu) \in \mathcal{X}$. This is realized if $(\mu, \nu) \in \operatorname{dom} \Lambda^*$ or equivalently if $\inf (\mathrm{MK}) < \infty$.

Applying the first part of Theorem 2.2, taking into account the dual equality (3.2) and Proposition 3.7 gives the following:

https://doi.org/10.1051/cocv/2010013 Published online by Cambridge University Press

Lemma 4.1. *Any $(\pi,\omega)\in\mathcal{P}_{AB}\times\mathcal{X}^*$ is a solution to* (MK,$\overline{\text{K}}$) *if and only if*

$$\left\{\begin{array}{ll}(a) & \pi\in P(\mu,\nu,c);\\ (b) & \pi\in\partial_{U_c'}\overline{\Phi}(\eta)\ \textit{where}\\ (c) & \eta=T^*\omega.\end{array}\right. \tag{4.2}$$

As Φ^* and $\overline{\Phi}$ are mutually convex conjugates, (4.2)-(b) is equivalent to

$$\eta\in\partial_{U_c'^*}\Phi^*(\pi) \tag{4.3}$$

and also equivalent to Fenchel's identity

$$\Phi^*(\pi)+\overline{\Phi}(\eta)=\langle\eta,\pi\rangle \tag{4.4}$$

and by Proposition 3.7 this is also equivalent to

$$\left\{\begin{array}{l}\overline{\Phi}(\eta)=0\\ \langle\eta,\pi\rangle=\int_{A\times B}c\mathrm{d}\pi.\end{array}\right.$$

Therefore, with Theorem 2.2 one obtains:

Theorem 4.2.

(1) *Any $(\pi,\omega)\in P(\mu,\nu,c)\times\mathcal{X}^*$ is a solution to* (MK,$\overline{\text{K}}$) *if and only if*

$$\left\{\begin{array}{ll}(a) & \omega_A\oplus\omega_B\in\overline{\Gamma}\\ (b) & \langle\omega,(\mu,\nu)\rangle=\int_{A\times B}c\mathrm{d}\pi.\end{array}\right. \tag{4.5}$$

(2) *Suppose in addition that*

$$(\mu,\nu)\in\mathrm{dom}\partial\Lambda^*. \tag{4.6}$$

Then, (MK,$\overline{\text{K}}$) *admits a solution in $P(\mu,\nu,c)\times\mathcal{X}^*$.*

(3) *Writing $\eta=\omega_A\oplus\omega_B$,* (4.2)-(b), (4.3), (4.4) *and* (4.5) *are equivalent statements.*

One was allowed to apply the second part of Theorem 2.2 under the constraint qualification (2.4) = (4.6). Let us give some details about this abstract requirement.

Remark 4.3. Note that if $A\times B$ is an uncountable set, $(\mu,\nu)\notin\mathrm{icordom}\,\Lambda^*$ for all $\mu\in\mathcal{P}_A$, $\nu\in\mathcal{P}_B$. Indeed, for all $\pi\in P(\mu,\nu,c)$ one can find (a_o,b_o) such that with $\delta_{(a_o,b_o)}$ the Dirac measure at (a_o,b_o), $\ell_t:=t\delta_{(a_o,b_o)}+(1-t)\pi\not\geq 0$ for all $t<0$, so that $\Phi_1^*(\ell_t)=+\infty$ (Prop. 3.7-(a)). This shows that $[\ell_0,\ell_1]=[\pi,\delta_{(a_o,b_o)}]\subset\mathrm{dom}\,\Phi_1^*$ while $\ell_t\notin\mathrm{dom}\,\Phi_1^*$ for all $t<0$. Hence, $(\mu,\nu)\notin\mathrm{icordom}\,\Lambda^*$ and one has to consider the assumption (4.6) on (μ,ν) rather than $(\mu,\nu)\in\mathrm{icordom}\,\Lambda^*$.

Remarks 4.4 (some remarks about $\Phi^*(\ell)$, $\Phi^*(|\ell|)$, $\Lambda^*(k)$ and $\Lambda^*(|k|)$)**.** Remark 4.3 shows that $\mathrm{icordom}\,\Lambda^*$ is empty in general. The following remarks are motivated by the problem of circumventing this restrictive property which stops us from applying Theorem 4.2 with an easy sufficient condition for (4.6).

(a) As U_c' is a Riesz space (it is the topological dual space of the Riesz space $C_{|\mathcal{S}}$), any ℓ in U_c' admits an absolute value $|\ell|=\ell_+ +\ell_-$ and one can consider the convex and real-valued function $\Phi^*(|\ell|)=\Phi^*(\ell_+)+\Phi^*(\ell_-)$ on U_c'.

(b) Rather than the positively homogeneous sublinear function $\Lambda^*(k)$ one could think of a *real-valued* positively homogeneous function of the type $\Lambda^*(|k|)$, since its icordom is nonempty. But, unlike U_c', $\mathcal{X}$ is not a Riesz space and $\Lambda^*(|k|)$ is meaningless.

https://doi.org/10.1051/cocv/2010013 Published online by Cambridge University Press

(c) The dual equality is $\Lambda^*(k) = \inf\{\Phi^*(\ell);\ \ell \in U_c',\ T_o\ell = k\}$ and unlike Φ^*, the effective domain of $\Phi^*(|\ell|)$ is the whole space. It is natural to think of the function $J(k) := \inf\{\Phi^*(|\ell|);\ \ell \in U_c',\ T_o\ell = k\}$ instead of Λ^*. The corresponding dual equality is $J = \iota^*_{\Upsilon_o}$ where $\Upsilon_o := \{(f', g') \in (C_A \oplus C_B)_{|\mathcal{S}};\ -c \leq f' \oplus g' \leq c\}$.It follows that J is a positively homogeneous sublinear function. But it is not true that J and Λ^* match on $\operatorname{dom}\Lambda^*$. We have $J \leq \Lambda^*$ and this inequality can be strict. To see this, let us consider the following example. Take $A = \{a, \alpha\}$, $B = \{b, \beta\}$, $c(a, b) = c(a, \beta) = c(\alpha, b) = 0$ and $c(\alpha, \beta) = 1$. Clearly $\Lambda^*(\delta_\alpha, \delta_\beta) = c(\alpha, \beta) = 1$ while $J(\delta_\alpha, \delta_\beta) = \Phi^*(|\delta_{(a,\beta)} + \delta_{(\alpha,b)} - \delta_{(a,b)}|) = 0$.

Theorem 4.2 is the core of the extended saddle-point method applied to the Monge-Kantorovich problem. To prove a practical optimality criterion one still has to translate these abstract properties.

5. A sufficient condition of optimality

The aim of this section is to prove Theorem 1.8.

5.1. Any finite strongly c-cyclically monotone plan is optimal

Next lemma gives a characterization of the closed convex hull $\operatorname{cl}\operatorname{cv} A$ of a set A in terms of its support functional ι^*_A.

Lemma 5.1. *Let X and Y be two topological vector spaces in duality. For any subset A of X, one has*

$$x \in \operatorname{cl}\operatorname{cv}(A) \Leftrightarrow \langle x, y\rangle \leq \iota^*_A(y),\ \forall y \in Y$$

where $\iota^*_A(y) = \sup_{z\in A}\langle z, y\rangle$, $y \in Y$.

Proof. The biconjugate ι^{**}_A of the indicator function ι_A is its closed convex envelope which is also the indicator function $\iota_{\operatorname{cl}\operatorname{cv} A}$ of $\operatorname{cl}\operatorname{cv} A$. Therefore,

$$\begin{aligned} x \in \operatorname{cl}\operatorname{cv}(A) &\Leftrightarrow \iota^{**}_A(x) = 0 \\ &\Leftrightarrow \iota^{**}_A(x) \leq 0 \\ &\Leftrightarrow \langle x, y\rangle \leq \iota^*_A(y),\ \forall y \in Y \end{aligned}$$

where the second equivalence follows from $\iota^{**}_A(x) = \sup_y\langle x, y\rangle - \iota^*_A(y) \geq \langle x, 0\rangle - \iota^*_A(0) = 0$, for all $x \in X$. □

Proposition 5.2. *Any finite strongly c-cyclically monotone plan is optimal.*

This result is a restatement of [17], Theorem 2, see Theorem 1.6-(b).

Proof. This is a corollary of the first (easy) part of Theorem 4.2.

Let π be a finite strongly c-cyclically monotone plan: There exist two measurable $[-\infty, \infty)$-valued functions f and g such that $f \oplus g$ satisfies (1.2):

$$\begin{cases} f \oplus g \leq c & \text{everywhere} \\ f \oplus g = c & \pi\text{-almost everywhere.} \end{cases}$$

Let us first check that (μ, ν) belongs to $\mathcal{X}$. Indeed, $|(\mu, \nu)|^*_\Lambda = \inf\{\|\ell\|^*_c;\ \ell \in U_c',\ \ell_A = \mu,\ \ell_B = \nu\} \leq \|\pi\|^*_c \leq \int c\,\mathrm{d}\pi < \infty$. Let E_o be the vector subspace of $\mathcal{X}$ spanned by (μ, ν) and by means of Lemma 1.3, define the linear form on E_o

$$\langle \omega_o, t(\mu, \nu)\rangle := t \int f \oplus g\,\mathrm{d}(\mu, \nu, c), \quad t \in \mathbb{R}.$$

https://doi.org/10.1051/cocv/2010013 Published online by Cambridge University Press

As $f \oplus g = c$, π-almost everywhere and $\pi \in P(\mu,\nu,c)$, we have

$$\langle \omega_o, (\mu,\nu)\rangle = \int c \,\mathrm{d}\pi. \tag{5.1}$$

On the other hand, thanks to (2.2) and Theorem 3.1-(1), $\Lambda^*(\mu,\nu) = \sup(\overline{K}) \geq \int f \oplus g \,\mathrm{d}(\mu,\nu,c) = \langle \omega_o, (\mu,\nu)\rangle$. It follows that

$$\langle \omega_o, k\rangle \leq \Lambda^*(k), \quad \forall k \in E_o.$$

Denoting

$$\Upsilon := \{(f',g') \in (C_A \oplus C_B)_{|\mathcal{S}};\ f' \oplus g' \leq c\},$$

one has $\Lambda^*(k) = \iota^*_\Upsilon(k)$, $k \in \mathcal{X}$. In particular, it is a $[0,\infty]$-valued positively homogeneous convex function on $\mathcal{X}$. By the analytic form of the Hahn-Banach theorem (see Rem. 5.3 below), there exists an extension ω of ω_o to $\mathcal{X}$ which satisfies

$$\langle \omega, k\rangle \leq \iota^*_\Upsilon(k),\ \forall k \in \mathcal{X}.$$

By Lemma 5.1, this means that ω belongs to the $\sigma(\mathcal{X}^*,\mathcal{X})$-closure $\overline{\Upsilon}$ of Υ. It is clear that $T_o^*\Upsilon \subset \Gamma$ and one sees that $T^*\overline{\Upsilon} \subset \overline{\Gamma}$ because of the $\sigma(\mathcal{X}^*,\mathcal{X})$-$\sigma(\mathcal{L}^*,\mathcal{L})$-continuity of $T^* : \mathcal{X}^* \to \mathcal{L}^*$, see [12], Lemma 4.13. Since $T^*\omega = \omega_A \oplus \omega_B$, we have

$$\omega_A \oplus \omega_B \in \overline{\Gamma}.$$

Together with (5.1), this allows us to apply Theorem 4.2-(1) to obtain the desired result. □

Remark 5.3. We have used an unusual form of the Hahn-Banach theorem where the positively homogeneous sublinear function $\Lambda^*(k)$ is $(-\infty,+\infty]$-valued instead of real-valued. For a proof of this variant, rewrite without any change the proof the analytic version of Hahn-Banach theorem based on Zorn's lemma; see for instance [6], Theorem 1.1.

The reason for not considering this extended version in the literature might be the following: to obtain separation by *closed* hyperplane, the positively homogeneous sublinear function of interest is the gauge of an *open* convex neighborhood of zero which is finite.

5.2. Measurability considerations. Strong c-cyclical monotonicity revisited

Definitions 5.4. Let γ be a probability measure and Γ be a set of probability measures on some measurable space.

(1) A set N is said to be γ-negligible if it is measurable and $\gamma(N) = 0$.
(2) A set is said to be Γ-negligible if it is γ-negligible for all $\gamma \in \Gamma$.
(3) A property holds Γ-almost everywhere if it holds everywhere except on a Γ-negligible set.
(4) A function h is said to be γ-measurable if there exists a γ-negligible set N such that $\mathbf{1}_{N^c}h$ is measurable.
(5) A function h is said to be Γ-measurable if there exists a Γ-negligible set N such that $\mathbf{1}_{N^c}h$ is measurable.

In particular the spaces $L_1(\mu)$ and $L_1(\nu)$ of all μ and ν-integrable functions on A and B consist of classes with respect to the μ and ν-almost everywhere equalities of μ and ν-measurable functions.

Definitions 5.5 ((μ,ν)-measurability)**.** These notions are meaningful only for a measurable product space $A\times B$.

(1) A subset N of $A\times B$ is said to be (μ,ν)-*negligible* if there exist two measurable sets $N_A \subset A$ and $N_B \subset B$ such that $\mu(N_A) = \nu(N_B) = 0$ and $N \subset (N_A \times B) \cup (A \times N_B)$.
(2) A property holds (μ,ν)-almost everywhere if it holds everywhere except on a (μ,ν)-negligible set.
(3) A function φ on $A\times B$ is said to be (μ,ν)-measurable if there exists a (μ,ν)-negligible measurable set N such that $\mathbf{1}_{N^c}\varphi$ is measurable on $A\times B$.

https://doi.org/10.1051/cocv/2010013 Published online by Cambridge University Press

Proposition 5.6.

(1) *Any (μ,ν)-negligible set is $P(\mu,\nu)$-negligible.*
Hence, any (μ,ν)-measurable function is $P(\mu,\nu)$-measurable.
(2) *Let f and g be functions on A and B. The following statements are equivalent.*
(a) *f is μ-measurable and g is ν-measurable;*
(b) *$f\oplus g$ is (μ,ν)-measurable;*
(c) *$f\oplus g$ is $P(\mu,\nu)$-measurable;*
(d) *$f\oplus g$ is $\mu\otimes\nu$-measurable.*

Proof.
• Proof of (1). Let $\pi\in P(\mu,\nu)$ and $N=(N_A\times B)\cup(A\times N_B)$ be a (μ,ν)-negligible measurable set. We have $\pi(N)\le\pi(N_A\times B)+\pi(A\times N_B)=\mu(N_A)+\nu(N_B)=0$.

Let us prove (2).

• Proof of $(a)\Rightarrow(b)$. Let N_A and N_B be negligible sets such that $f\mathbf{1}_{N_A^c}$ and $g\mathbf{1}_{N_B^c}$ are measurable. Of course, $N=(N_A\times B)\cup(A\times N_B)$ is (μ,ν)-negligible and $\mathbf{1}_{N^c}f\oplus g$ is measurable.

• Proof of $(b)\Rightarrow(c)$. This follows from (1).

• Proof of $(c)\Rightarrow(d)$. Immediate.

• Proof of $(d)\Rightarrow(a)$. Let S be a measurable subset of $A\times B$ such that $\mu\otimes\nu(S)=1$ and $\mathbf{1}_S f\oplus g$ is measurable. For all $a\in A$, denote $S_a=\{b\in B;(a,b)\in S\}$ the a-section of S. It is measurable and by Fubini's theorem

$$1=\mu\otimes\nu(S)=\int_A\left[\int_B\mathbf{1}_S(a,b)\,\nu(\mathrm{d}b)\right]\mu(\mathrm{d}a)=\int_A\nu(S_a)\,\mu(\mathrm{d}a).$$

Therefore, $\nu(S_a)=1$, μ-a.e. and there exists some $a_o\in A$ such that $\nu(S_{a_o})=1$. As a section of a measurable function, the function $b\mapsto\mathbf{1}_{S_{a_o}}(b)(f(a_o)+g(b))$ is measurable. It follows that $\mathbf{1}_{S_{a_o}}g$ is measurable: this proves that g is ν-measurable. A similar proof works for f. □

It is immediate from the Definition 1.2 that a transport plan $\pi\in P(\mu,\nu)$ is strongly c-cyclically monotone if and only if there exist a μ-measurable function f on A and a ν-measurable function g on B such that

$$\begin{cases} f\oplus g\le c & (\mu,\nu)\text{-almost everywhere}\\ f\oplus g=c & \pi\text{-almost everywhere.}\end{cases}\tag{5.2}$$

The underlying measurability properties of $f\oplus g$ which are required by (5.2) are insured by Proposition 5.6.

By Proposition 5.6-1, if a property holds true (μ,ν)-almost everywhere, then it is still true $P(\mu,\nu)$-almost everywhere and *a fortiori* $P(\mu,\nu,c)$-almost everywhere. Therefore,

$$\begin{cases} f\oplus g\le c & P(\mu,\nu,c)\text{-almost everywhere}\\ f\oplus g=c & \pi\text{-almost everywhere}\end{cases}\tag{5.3}$$

is weaker than the strong c-cyclical monotonicity. Without changing a word to the proof of Proposition 5.2, one obtains the following sufficient *condition* of optimality.

Theorem 1.8. *Let $\pi\in P(\mu,\nu,c)$ be any finite plan. If there exist a μ-measurable function f on A and a ν-measurable function g on B which satisfy (5.3), then π is optimal.*

5.3. The Counterexamples 1.7-(d, e)

They are optimal plans which are not strongly c-cyclically monotone but they both satisfy the weaker property (5.3).

https://doi.org/10.1051/cocv/2010013 Published online by Cambridge University Press

– Counterexample 1.7-(d). Let $A = B = [0,1]$ both equipped with the Lebesgue measure $\lambda = \mu = \nu$. Define c to be ∞ above the diagonal and $1 - \sqrt{a-b}$ for $b \leq a$. The set $P(\mu,\nu,c)$ is reduced to the uniform probability measure π on the diagonal which is *a fortiori* optimal. It is shown in [3] that π is not strongly c-cyclically monotone. But (5.3) is trivially satisfied.
– Counterexample 1.7-(e). Let $A = B = \mathbb{N} \cup \{\omega\}$ where ω is a "number" larger than all $n \in \mathbb{N}$. Equip A and B with the discrete topology and define $\mu = \nu$ with a full support. Define $c(a,b) = \begin{cases} \infty & \text{for } a < b \\ 1 & \text{for } a = b \\ 0 & \text{for } a > b \end{cases}$ for each $a \in A$ and $b \in B$. Again, the set $P(\mu,\nu,c)$ is reduced to a single probability measure π which is *a fortiori* optimal and (5.3) is trivially satisfied. Nevertheless, it is proved in [2] that π is not strongly c-cyclically monotone.

6. A necessary condition of optimality

The aim of this section is to prove Theorem 1.9 which states that an optimal plan satisfies approximately (5.3).

6.1. A first approach to the necessary condition

We sketch a direct approach to the necessary condition and emphasize some problems which remain to be solved. The Kantorovich dual equality (3.2) yields a maximizing sequence $\{(f_n, g_n)\}_{n\geq 1}$ in $C_A \times C_B$. Assume that $\inf(\mathrm{MK}) < \infty$ and define $c_n := f_n \oplus g_n$ so that $c_n \leq c$ and $\lim_{n\to\infty} \int c_n \,\mathrm{d}\pi^* = \inf(\mathrm{MK}) = \int c \,\mathrm{d}\pi^*$ for any optimal plan π^*. Clearly, $c - c_n \geq 0$ and $\lim_{n\to\infty} \int (c - c_n)\,\mathrm{d}\pi^* = 0$. Hence, c_n converges to c in $L^1(\pi^*)$ and one can extract a subsequence, denoted $(c_k)_{k\geq 1}$, which converges to c pointwise π^*-almost everywhere. By a result of Borwein and Lewis [5], Corollary 3.4, $(c_k)_{k\geq 1}$ converges pointwise π^*-almost everywhere to some sum function $f \oplus g$. Therefore, $c = f \oplus g$, π^*-almost everywhere.

The remaining problem of extending f and g such that $f \oplus g \leq c$ everywhere is not obvious. By Tykhonov's theorem, one can extract a *subnet* from $(c_k)_{k\geq 1}$ which converges pointwise to a $[-\infty,+\infty]$-valued function $\tilde{c}$ such that $\tilde{c} \leq c$ everywhere and $\tilde{c} = c$, π^*-almost everywhere. Unfortunately, a subnet limit is not enough to insure that $\tilde{c}$ is measurable. In addition, one cannot apply the above cited Borwein-Lewis convergence result since $[-\infty,+\infty]$ is not a group. Consequently, one cannot assert that $\tilde{c}$ is of the sum form $f \oplus g$.

6.2. A necessary condition

The idea in this section is to approximate $\Phi = \iota_\Gamma$ by the sequence $(\Phi_k)_{k\geq 1}$ with $\Phi_k = \iota_{\Gamma_k}$ the convex indicator function of

$$\Gamma_k := \{u \in \mathcal{U}_o; -kc \leq u \leq c\}, \quad k \geq 1.$$

We define

$$\Phi_k^*(\ell) = \|\ell^+\|_c^* + k\|\ell^-\|_c^*, \quad \ell \in U_c' \tag{6.1}$$

its convex conjugate, where $\|\ell\|_c^* = \sup\{\langle u, \ell\rangle;\ u \in U_c : \|u\|_c \leq 1\}$ is the dual norm on U_c'. Note that as U_c' is a Riesz space, the positive and negative part ℓ^+ and ℓ^- are meaningful for all $\ell \in U_c'$.

The restriction of Φ_k^* to the set $Q(\mu,\nu,c)$ of all $q \in U_c'$ with first and second marginals μ and ν is denoted

$$F_k(q) = \|q^+\|_c^* + k\|q^-\|_c^*, \quad q \in Q(\mu,\nu,c) := \{q \in U_c';\ q_A = \mu,\ q_B = \nu\}.$$

It is clear that $(F_k)_{k\geq 1}$ converges pointwise and increasingly to the restriction F of Φ^* to $Q(\mu,\nu,c)$. By Proposition 3.7-(c), the effective domain of F is $P(\mu,\nu,c)$. Hence, F is the objective function of (MK):

$$F(q) = \begin{cases} \int_{A\times B} c\,\mathrm{d}q, & \text{if } q \in P(\mu,\nu,c), \\ +\infty, & \text{otherwise,} \end{cases} \qquad q \in Q(\mu,\nu,c).$$

https://doi.org/10.1051/cocv/2010013 Published online by Cambridge University Press

Lemma 6.1.

(a) *The following* Γ*-convergence result*

$$\Gamma\text{-}\lim_{k\to\infty} F_k = F$$

holds true for both the $*$*-weak topology* $\sigma(Q(\mu,\nu,c),U_c)$ *and the strong topology associated with* $\|\cdot\|_c^*$ *on* $Q(\mu,\nu,c)$;

(b) $\min F = \lim_{k\to\infty} \inf F_k = \sup_k \min F_k$.

Proof. Since $Q(\mu,\nu,c)$ is $\sigma(U_c',U_c)$-closed, one sees with the Banach-Alaoglu theorem that F_k is $\sigma(Q(\mu,\nu,c),U_c)$-inf-compact for all $k\geq 1$. It is *a fortiori* lower semicontinuous with respect to $\sigma(Q(\mu,\nu,c),U_c)$ and $\|\cdot\|_c^*$. As the sequence $(F_k)_{k\geq 1}$ is increasing, it is also $\sigma(Q(\mu,\nu,c),U_c)$-equicoercive.

Now, (a) and (b) follow respectively from [7], Proposition 5.4 and Remark 5.5, and [7], Theorem 7.8. □

Let us denote $M_\epsilon(F_k)$ the set of the ϵ-minimizers of F_k:

$$M_\epsilon(F_k) = \{q\in Q(\mu,\nu,c);\ F_k(q)\leq \inf F_k + \epsilon\},\quad \epsilon>0,\ k\geq 1$$

and $M(F)$ the set of all the minimizers of F.

Lemma 6.2. *Clearly, the set of all the optimal plans is* $M(F)$.

Assume that $P(\mu,\nu,c)$ *is nonempty. Then* $M(F)$ *is nonempty and*

$$M(F) = \bigcap_{\epsilon>0,k\geq 1} \overline{M_\epsilon}(F_k)$$

where $\overline{M_\epsilon}(F_k)$ *is the closure of* $M_\epsilon(F_k)$ *with respect to* $\|\cdot\|_c^*$.

Proof. Lemma 6.1 and [7], Theorem 7.19, give us $M(F)=\bigcap_\epsilon K\text{-}\liminf_{k\to\infty} M_\epsilon(F_k)$ where $K\text{-}\liminf_{k\to\infty} M_\epsilon(F_k)$ is the liminf in the sense of Kuratowski associated with the topology generated by $\|\cdot\|_c^*$. Since $(F_k)_{k\geq 1}$ is increasing, we obtain $K\text{-}\liminf_{k\to\infty} M_\epsilon(F_k) = \bigcap_{k\geq 1}\overline{M_\epsilon}(F_k)$ which completes the proof. □

This result invites us to learn more about $M_\epsilon(F_k)$. We denote U_c'' the strong bidual of U_c and denote $\|\cdot\|_c$ its norm. Let $\overline{\Sigma}_k$ be the $\sigma(U_c'',U_c')$-closure of

$$\Sigma_k = \{f\oplus g\in (C_A\oplus C_B)_{|\mathcal{S}};\ -kc\leq f\oplus g\leq c\}$$

and $\overline{\Phi}_k$ the largest convex $\sigma(U_c'',U_c')$-lower semicontinuous extension of Φ_k to U_c''. We denote Λ_k^* the analogue of Λ^* where Φ is replaced by Φ_k.

Lemma 6.3. *For each* $\epsilon>0$, $k\geq 1$ *and* $q\in M_\epsilon(F_k)$, *there exist* $\bar\eta\in\overline{\Sigma}_k$, $\widetilde{\eta}\in U_c''$ *and* $\tilde\ell\in U_c'$ *such that* $\|\tilde\ell - q\|_c^*\leq\sqrt{\epsilon}$, $\|\widetilde{\eta}-\bar\eta\|_c\leq\sqrt{\epsilon}$ *and* $\tilde\ell\in\partial\overline{\Phi}_k(\widetilde{\eta})$.

Proof. Since the effective domain of Φ_k^* is the whole space U_c', the dual equality insures that the effective domain of Λ_k^* is the whole space $\mathcal{X}$. In particular,

$$\partial\Lambda_k^*(\mu,\nu)\neq\emptyset.$$

This non-emptiness is crucial. The approximation F_k of F was introduced to circumvent the problem of knowing whether $\partial\Lambda^*(\mu,\nu)$ is empty or not. One is allowed to apply Theorem 2.2 and Proposition 2.3.

Let $\bar q$ be a minimizer of F_k. We have $\bar q\in Q(\mu,\nu,c)$ and $\partial\Phi_k^*(\bar q)$ is nonempty. More precisely, with $\bar\omega\in\partial\Lambda_k^*(\mu,\nu)$ and $\bar\eta=\bar\omega_A\oplus\bar\omega_B$ we have $\bar\eta\in\partial\Phi_k^*(\bar q)$ and

$$\langle\bar\eta,\bar q\rangle = \langle\bar\eta,q\rangle,\quad \forall q\in Q(\mu,\nu,c). \tag{6.2}$$

https://doi.org/10.1051/cocv/2010013 Published online by Cambridge University Press

Fenchel's equality is $\Phi_k^*(\bar q) = \langle \bar\eta, \bar q\rangle - \overline{\Phi}_k(\bar\eta)$. But, $\overline{\Phi}_k(\bar\eta)$ is equal to 0 or $+\infty$ and $\Phi_k^*(\bar q) > -\infty$. Hence, $\overline{\Phi}_k(\bar\eta) = 0$ and

$$\begin{cases} \bar\eta \in \overline{\Sigma}_k \\ \Phi_k^*(\bar q) = \langle \bar\eta, \bar q\rangle. \end{cases}$$

As the level sets of Φ_k are $\|\cdot\|_c$-bounded, by Proposition 2.3-(b') we have $\bar\eta \in U_c''$. Now, we consider the topological duality $\langle U_c', U_c''\rangle$.

Let us take q in $M_\epsilon(F_k)$. With (6.2) one sees that $\langle \bar\eta, \bar q\rangle = \langle \bar\eta, q\rangle$. It follows that $\Phi_k^*(q) \le \Phi_k^*(\bar q) + \epsilon = \langle \bar\eta, q\rangle + \epsilon = \langle \bar\eta, q\rangle - \overline{\Phi}_k(\bar\eta) + \epsilon$. This means that q stands in the ϵ-subdifferential $\partial_\epsilon \overline{\Phi}_k(\bar\eta)$ of $\overline{\Phi}_k$ at $\bar\eta$. One completes the proof applying Ekeland's principle [10], Theorem 6.2. □

Lemma 6.4. *Let π be an optimal plan. There exist a sequence $(\tilde\ell_k)_{k\ge1}$ in U_c' and two sequences $(\widetilde\eta_k)_{k\ge1}$ and $(\bar\eta_k)_{k\ge1}$ in U_c'' such that:*

(i) $\lim_{k\to\infty} \|\tilde\ell_k - \pi\|_c^* = 0$
and for each $k \ge 1$:
(ii) $\widetilde\eta_k \in \partial\Phi_k^*(\tilde\ell_k)$;
(iii) $\|\widetilde\eta_k - \bar\eta_k\|_c \le 1/k$ *and* $\bar\eta_k \in \overline{\Sigma}_k$;
(vi) $\|\tilde\ell_k^-\|_c^* \le (\inf(\mathbb{MK}) + 2)/k$.

Proof. By Lemma 6.2, there exists a sequence $(q_k)_{k\ge1}$ such that

$$\lim_{k\to\infty} \|q_k - \pi\|_c^* = 0 \tag{6.3}$$

and $q_k \in M_{\epsilon_k}(F_k)$ for each $k \ge 1$, where $\lim_{k\to\infty}\epsilon_k = 0$.

By Lemma 6.3, for each $k \ge 1$, there exist $\tilde\ell_k \in U_c'$, $\bar\eta \in \overline{\Sigma}_k$ and $\widetilde\eta_k \in U_c'$ such that $\|\tilde\ell_k - q_k\|_c^* \le \sqrt{\epsilon_k}$, $\|\widetilde\eta_k - \bar\eta_k\|_c \le \sqrt{\epsilon_k}$ and $\widetilde\eta_k \in \partial\Phi_k^*(\tilde\ell_k)$.

Since Φ_k^* is k-Lipschitz, we have

$$\begin{aligned} \Phi_k^*(\tilde\ell_k) &\le \Phi_k^*(q_k) + k\|\tilde\ell_k - q_k\|_c \\ &\le \inf(\mathbb{MK}) + \epsilon_k + k\sqrt{\epsilon_k} \\ &\le \inf(\mathbb{MK}) + 2 \end{aligned}$$

taking

$$\epsilon_k = 1/k^2.$$

It follows from (6.1) that $\|\tilde\ell_k^-\|_c^* \le \Phi_k^*(\tilde\ell_k)/k \le (\inf(\mathbb{MK}) + 2)/k$. One concludes with $\|\tilde\ell_k - \pi\|_c^* \le \|\tilde\ell_k - q_k\|_c^* + \|q_k - \pi\|_c^* \le 1/k + \|q_k - \pi\|_c^*$ and (6.3). □

Lemma 6.5. *Let $\eta \in U_c''$ be in the $\sigma(U_c'', U_c')$-closure of Σ_k for some $k \ge 1$. For each probability measure $p \in \mathcal{P}_c$, there exists a function $\widehat\eta$ in $L_1(p)$ such that*

$$\begin{cases} (a) \quad \langle \eta, h.p\rangle = \int_{A\times B} h\widehat\eta \,\mathrm{d}p, \quad \forall h \in L_\infty(p) \\ (b) \quad \widehat\eta \in \mathrm{cl}_{L_1(p)}(\Sigma_k) \end{cases} \tag{6.4}$$

where $\mathrm{cl}_{L_1(p)}(\Sigma_k)$ is the closure of Σ_k in $L_1(p)$ with respect to its usual strong topology.

Proof. For any bounded measurable function h on $A\times B$ we have

$$|\langle \eta, h.p\rangle| \le \|\eta\|_{U_c''} \int_{A\times B} |h|c \,\mathrm{d}p \tag{6.5}$$

since $|\langle \eta, h.p\rangle| \le \|\eta\|_{U_c''}\|h.p\|_{U_c'}$ with $\|h.p\|_{U_c'} = \int_{A\times B}|h|c\,\mathrm{d}p$. Let $(h_k)_{k\ge1}$ be a sequence of bounded measurable functions such that the sequence $(|h_k|)_{k\ge1}$ decreases pointwise to zero. By (6.5) and dominated convergence,

https://doi.org/10.1051/cocv/2010013 Published online by Cambridge University Press

we have $\lim_{k\to\infty}\langle\eta, h_k.p\rangle = 0$. This means that $h \mapsto \langle\eta, h.p\rangle$ is σ-additive: there exists a measure ρ such that $\langle\eta, h.p\rangle = \int_{A\times B} h\,d\rho$, $\forall h$. By (6.5), ρ is absolutely continuous with respect to p: there exists $\widehat{\eta}$ in $L_1(p)$ such that (6.4-a) holds.

Let $\{u_\alpha\}_\alpha$ be a net in Σ_k such that $\lim_\alpha u_\alpha = \eta$. For all h in $L_\infty(p)$,

$$\int_{A\times B} h\widehat{\eta}\,\mathrm{d}p = \langle\eta, h.p\rangle = \lim_\alpha\langle u_\alpha, h.p\rangle = \lim_\alpha \int_{A\times B} hu_\alpha\,\mathrm{d}p.$$

This means that $\widehat{\eta}$ is in the $\sigma(L_1(p), L_\infty(p))$-closure of Σ_k. As Σ_k is convex, this implies (6.4)-(b). □

We are now ready to give the proof of Theorem 1.9.

Theorem 1.9. *Let π be any optimal plan and let us take $\epsilon > 0$ and $p \in \mathcal{P}_c$. Then, there exist $\varphi \in L_1(\pi + p)$, $f \in C_A$, $g \in C_B$ and a measurable subset $D \subset A\times B$ such that:*

(1) $\varphi = c$, $\mathbf{1}_{(\mathcal{S}\setminus D)}.\pi$*-almost everywhere;*
(2) $\int_D (1 + c)\,\mathrm{d}\pi \le \epsilon$;
(3) $-c/\epsilon \le \varphi \le c$, $(\pi + p)$*-almost everywhere;*
(4) $-c/\epsilon \le f \oplus g \le c$, *everywhere;*
(5) $\|\varphi - f \oplus g\|_{L_1(\pi+p)} \le \epsilon$.

Proof. By Lemma 6.4, for any $\epsilon > 0$, there exist k large enough, $\tilde{\ell}$ in U_c' and $\widetilde{\eta}, \bar{\eta}$ in U_c'' such that

$$\begin{cases} \text{(i)} & \|\tilde{\ell} - \pi\|_c^* \le \epsilon \text{ and } \|\tilde{\ell}^-\|_c^* \le \epsilon \\ \text{(ii)} & \widetilde{\eta} \in \partial\Phi_k^*(\tilde{\ell}) \\ \text{(iii)} & \bar{\eta} \in \overline{\Sigma}_k \\ \text{(iv)} & \|\widetilde{\eta} - \bar{\eta}\|_c \le \epsilon. \end{cases}$$

Let $\widehat{\widetilde{\eta}}$ and $\widehat{\bar{\eta}}$ be the functions in $L_1(\pi+p)$ which are built from $\widetilde{\eta}$ and $\bar{\eta}$ as in Lemma 6.5 and satisfy: $\langle\widetilde{\eta}, h.(\pi+p)\rangle = \int_{A\times B} h\widehat{\widetilde{\eta}}\,\mathrm{d}(\pi + p)$ and $\langle\bar{\eta}, h.(\pi+p)\rangle = \int_{A\times B} h\widehat{\bar{\eta}}\,\mathrm{d}(\pi + p)$ for all $h \in L_\infty(\pi + p)$. By (iii) and (6.4)-(b), there exist $f \in C_A$ and $g \in C_B$ such that $-kc \le f \oplus g \le c$ and

$$\|\widehat{\bar{\eta}} - f \oplus g\|_{L_1(\pi+p)} \le \epsilon.$$

With the notation of Lemma 6.5, we have $\|\widehat{\eta}\|_{L_1(p)} \le [\int_{A\times B} c\,\mathrm{d}p]\|\eta\|_c$ for all $\eta \in U_c''$. Taking $\varphi = \widehat{\widetilde{\eta}}$, this and (iv) give $\|\varphi - \widehat{\bar{\eta}}\|_{L_1(\pi+p)} \le \epsilon \int_{A\times B} c\,\mathrm{d}(\pi + p)$. Therefore,

$$\|\varphi - f \oplus g\|_{L_1(\pi+p)} \le \left(1 + \int_{A\times B} c\,\mathrm{d}(\pi + p)\right)\epsilon.$$

By (ii), we have $\widetilde{\eta} \in \operatorname{dom}\overline{\Phi}_k = \Sigma_k$. Hence, $-kc \le \varphi = \widehat{\widetilde{\eta}} \le c$, $(\pi + p)$-a.e.

Since $\mathcal{U}_o$ is identified with $C_{|\mathcal{S}}$ which is a subspace of $C_{\mathcal{S}}$, see (3.3), any $\ell \in U_c'$ is the restriction to $C_{|\mathcal{S}}$ of a continuous linear form on $C_{\mathcal{S}}$. Therefore it can be uniquely decomposed as $\ell = \ell^a + \ell^s$ where ℓ^a is a measure and ℓ^s is singular with $\|\ell\|_c^* = \|\ell^a\|_c^* + \|\ell^s\|_c^*$. By (i), we have $\|\tilde{\ell}^+ - \pi\|_c^* \le 2\epsilon$. Since π is a measure, its singular part vanishes: $\pi^s = 0$. It follows that $\|\tilde{\ell}^{+a} - \pi\|_c^* \le 2\epsilon$ where $\tilde{\ell}^{+a}$ is the measure part of $\tilde{\ell}^+$. Therefore, the nonnegative measures $\tilde{\ell}^{+a}$ and π are concentrated on the same set except for a measurable set D such that $\int_D c\,\mathrm{d}(\tilde{\ell}^{+a} + \pi) \le 2\epsilon$.

Finally, (ii) implies that $c = \varphi$, $\tilde{\ell}^{+a}$-almost everywhere (and also that $-kc = \varphi$, $\tilde{\ell}^{-a}$-almost everywhere). One completes the proof, putting everything together and replacing ϵ by ϵ/C with an appropriate constant C. □

Acknowledgements. The author thanks Stefano Bianchini for very useful comments.

https://doi.org/10.1051/cocv/2010013 Published online by Cambridge University Press

References

[1] L. Ambrosio and A. Pratelli, *Existence and stability results in the L^1-theory of optimal transportation – CIME Course*, in *Lecture Notes in Mathematics* **1813**. Springer Verlag (2003) 123–160.
[2] M. Beiglböck and W. Schachermayer, Duality for Borel measurable cost functions. *Trans. Amer. Math. Soc.* (to appear).
[3] M. Beiglböck, M. Goldstern, G. Maresh and W. Schachermayer, Optimal and better transport plans. *J. Funct. Anal.* **256** (2009) 1907–1927.
[4] M. Beiglböck, C. Léonard and W. Schachermayer, *A general duality theorem for the Monge-Kantorovich transport problem*. Preprint (2009).
[5] J.M. Borwein and A.S. Lewis, Decomposition of multivariate functions. *Can. J. Math.* **44** (1992) 463–482.
[6] H. Brezis, *Analyse fonctionnelle – Théorie et applications*. Masson, Paris (1987).
[7] G. Dal Maso, *An Introduction to* Γ*-Convergence*. *Progress in Nonlinear Differential Equations and Their Applications* **8**. Birkhäuser (1993).
[8] L. Decreusefond, Wasserstein distance on configuration space. *Potential Anal.* **28** (2008) 283–300.
[9] L. Decreusefond, A. Joulin and N. Savy, Upper bounds on Rubinstein distances on configuration spaces and applications. *Communications on Stochastic Analysis* (to appear).
[10] I. Ekeland and R. Témam, *Convex Analysis and Variational Problems*, *Classics in Applied Mathematics* **28**. SIAM (1999).
[11] D. Feyel and A.S. Üstünel, Monge-Kantorovitch measure transportation and Monge-Ampère equation on Wiener space. *Probab. Theory Relat. Fields* **128** (2004) 347–385.
[12] C. Léonard, Convex minimization problems with weak constraint qualifications. *Journal of Convex Analysis* **17** (2010) 312–348.
[13] J. Neveu, *Bases mathématiques du calcul des probabilités*. Masson, Paris (1970).
[14] A. Pratelli, On the sufficiency of the c-cyclical monotonicity for optimality of transport plans. *Math. Z.* **258** (2008) 677–690.
[15] S. Rachev and L. Rüschendorf, *Mass Transportation Problems. Vol. **I**: Theory, Vol. **II**: Applications*. Springer-Verlag, New York (1998).
[16] L. Rüschendorf, On c-optimal random variables. *Statist. Probab. Lett.* **27** (1996) 267–270.
[17] W. Schachermayer and J. Teichman, Characterization of optimal transport plans for the Monge-Kantorovich problem. *Proc. Amer. Math. Soc.* **137** (2009) 519–529.
[18] C. Villani, *Topics in Optimal Transportation*, *Graduate Studies in Mathematics* **58**. American Mathematical Society, Providence (2003).
[19] C. Villani, *Optimal Transport – Old and New*, *Grundlehren der mathematischen Wissenschaften* **338**. Springer (2009).

https://doi.org/10.1051/cocv/2010013 Published online by Cambridge University Press