

# Convex Minimization Problems with Weak Constraint Qualifications

Christian Léonard

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

Received: October 15, 2007

Revised manuscript received: March 17, 2009

One revisits the standard saddle-point method based on conjugate duality for solving convex minimization problems. Our aim is to reduce or remove unnecessary topological restrictions on the constraint set. Dual equalities and characterizations of the minimizers are obtained with weak or without constraint qualifications.

The main idea is to work with intrinsic topologies which reflect some geometry of the objective function.

The abstract results of this article are applied in other papers to the Monge-Kantorovich optimal transport problem and the minimization of entropy functionals.

*Keywords:* Convex optimization, saddle-point, conjugate duality

*2000 Mathematics Subject Classification:* 46N10, 49J45, 28A35

## Contents

<b>1</b>	<b>Introduction</b>	<b>321</b>
<b>2</b>	<b>Statements of the results</b>	<b>324</b>
<b>3</b>	<b>Examples</b>	<b>330</b>
<b>4</b>	<b>Preliminary results</b>	<b>334</b>
<b>5</b>	<b>Proofs of the results of Section 2</b>	<b>340</b>

## 1. Introduction

An “*extension*” of the saddle-point method for solving a convex minimization problem is investigated. It is shown how to implement the standard saddle-point method in such a way that topological restrictions on the constraint sets (the constraint qualifications) may essentially be removed. With this aim in view, one works with topologies associated with gauge functionals of sets which are close to the level sets of the objective

function. These well-suited topologies partly reflect the geometry of the problem. At some point, one has to compute an *extended* dual problem. This is the price to pay for implementing this approach.

The method is based on conjugate duality as developed by R. T. Rockafellar in [13]. Dual equalities and characterizations of the minimizers are obtained with weak or without constraint qualification.

This paper is a companion of [10] and [11] where this extended saddle-point method is applied to the Monge-Kantorovich optimal transport problem and the minimization of entropy functionals.

### An abstract convex problem and related questions

Let  $\mathcal{U}$  be a vector space,  $\mathcal{V} = \mathcal{U}^*$  its algebraic dual space,  $\Phi$  a  $(-\infty, +\infty]$ -valued convex function on  $\mathcal{U}$  and  $\Phi^*$  its convex conjugate for the duality  $\langle \mathcal{U}, \mathcal{V} \rangle$ . Let  $\mathcal{Y}$  be another vector space,  $\mathcal{X} = \mathcal{Y}^*$  its algebraic dual space and  $T : \mathcal{V} \rightarrow \mathcal{X}$  is a linear operator. We consider the convex minimization problem

$$\text{minimize } \Phi^*(v) \text{ subject to } Tv \in C, \quad v \in \mathcal{V}, \quad (P)$$

where  $C$  is a convex subset of  $\mathcal{X}$ . As is well known, Fenchel's duality leads to the dual problem

$$\text{maximize } \inf_{x \in C} \langle y, x \rangle - \Phi(T^*y), \quad y \in \mathcal{Y}, \quad (D)$$

where  $T^*$  is the adjoint of  $T$ .

**Questions 1.1.** The usual questions related to (P) and (D) are

- the dual equality: Does  $\inf(P) = \sup(D)$  hold?
- the primal attainment: Does there exist a solution  $\hat{v}$  to (P)? What about the minimizing sequences?
- the dual attainment: Does there exist a solution  $\bar{y}$  to (D)?
- the representation of the primal solutions: Find an identity of the type:  $\hat{v} \in \partial\Phi(T^*\bar{y})$ .

In the case where the constraint set  $C = \{x\}$  is reduced to a single point, the value  $\sup(D)$  of the dual problem is

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

which is the convex conjugate of  $\Lambda(y) := \Phi(T^*y)$ ,  $y \in \mathcal{Y}$ .

We are going to answer the above questions in terms of some *extension*  $\bar{\Phi}$  of  $\Phi$  under the weak constraint qualification

$$C \cap \text{dom } \partial_{\mathcal{X}^*} \Lambda^* \neq \emptyset, \quad (1)$$

where  $\text{dom } \partial_{\mathcal{X}^*} \Lambda^* = \{x \in \mathcal{X}; \partial_{\mathcal{X}^*} \Lambda^*(x) \neq \emptyset\}$  is the subset of all vectors in  $\mathcal{X}$  at which  $\Lambda^*$  admits a nonempty subdifferential with respect to the algebraic dual pairing  $\langle \mathcal{X}, \mathcal{X}^* \rangle$  with  $\mathcal{X}^*$  the algebraic dual space of  $\mathcal{X}$ . Note that by the geometric version of Hahn-Banach theorem, the intrinsic core of  $\text{dom } \Lambda^*$ :  $\text{icordom } \Lambda^*$ , is included in  $\text{dom } \partial_{\mathcal{X}^*} \Lambda^*$ . Hence, a useful criterion to get (1) is

$$C \cap \text{icordom } \Lambda^* \neq \emptyset. \quad (2)$$

The drawback of such a general approach is that one has to compute the extension  $\bar{\Phi}$ . In specific examples, this might be a difficult task. In the case of the Monge-Kantorovich problem [10] it is immediate, but it requires some work in the case of entropy minimization [11].

The restriction (2) is very weak since the intrinsic core is the notion of interior which gives the largest possible set. As  $C \cap \text{dom } \Lambda^* = \emptyset$  implies that (P) has no solution, the only case where the problem remains open when  $\text{icordom } \Lambda^*$  is nonempty is the situation where  $C$  and  $\text{dom } \Lambda^*$  are tangent to each other. This is used in [11] to obtain general results for convex integral functionals. The representation of their minimizers, see (24), are obtained under the constraint qualification (2) which is much weaker than the usual constraint qualification:

$$\text{int } C \cap \text{dom } \Lambda^* \neq \emptyset \tag{3}$$

where  $\text{int } C$  is the interior of  $C$  with respect to some topology which is not directly connected to the “geometry” of  $\Lambda^*$ . In particular,  $\text{int } C$  must be nonempty; this is a considerable restriction.

In [8, 9], Combari, Laghdir and Thibault answered Questions 1.1 under the qualification condition that  $\mathbb{R}_+[C - \text{dom } \Lambda^*]$  is a closed subspace, which is also weaker than (3).

The Monge-Kantorovich optimal transport problem provides an interesting case where the specifications of the constraints never stand in  $\text{icordom } \Lambda^*$ , see [10], so that (2) is useless and (1) is the right assumption to be used.

### The strategy

A usual way to prove the dual attainment and obtain some representation of the primal solutions is to require that the constraint is qualified: a property which allows us to separate the convex constraint set  $T^{-1}C$  and the level sets of the objective function  $\Phi^*$ . The strategy of this article is different: one chooses suitable topologies so that the level sets have nonempty interiors. This also allows us to apply Hahn-Banach theorem, but this time the constraint set is not required to have a nonempty interior. We take the rule not to introduce arbitrary topological assumptions since (P) is expressed without any topological notion. Because of the convexity of the problem, one takes advantage of geometric easy properties: the topologies to be considered later are associated with seminorms which are gauges of level sets of the convex functions  $\Phi$  and  $\Phi^*$ . They are useful tools to work with the *geometry* of (P).

It appears that when the constraints are infinite-dimensional one can choose several different spaces  $\mathcal{Y}$  without modifying the value and the solutions of (P). Consequently, for a small space  $\mathcal{Y}$  the dual attainment is not the rule. As a consequence, we are facing the problem of finding an *extension* of (D) which admits solutions in generic cases and such that the representation of the primal solution is  $\hat{v} \in \partial\bar{\Phi}(T^*\bar{y})$  where  $\bar{\Phi}$  is some extension of  $\Phi$ .

We are going to

- use the standard saddle-point approach to convex problems based on conjugate duality as developed by Rockafellar in [13],
- work with topologies which reflect some of the geometric structure of the objective function.

These made-to-measure topologies are associated with the gauges of the level sets of  $\Phi$  and  $\Phi^*$ .

**Outline of the paper**

The results are stated without proof in Section 2. Their proofs are postponed to Section 5. Examples are introduced in Section 3 where we consider the Monge-Kantorovich transport and entropy minimization problems. These problems are investigated in [10, 11].

**Notation**

Let  $X$  and  $Y$  be topological vector spaces. 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.

Let  $f : X \rightarrow [-\infty, +\infty]$  be an extended-real-valued function. Its convex conjugate with respect to  $\langle X, Y \rangle$  is  $f^*(y) = \sup_{x \in X} \{\langle x, y \rangle - f(x)\} \in [-\infty, +\infty]$ ,  $y \in Y$ . Its subdifferential at  $x$  with respect to  $\langle X, Y \rangle$  is  $\partial_Y f(x) = \{y \in Y; f(x + \xi) \geq f(x) + \langle y, \xi \rangle, \forall \xi \in X\}$ . If no confusion occurs, one writes  $\partial f(x)$ .

One defines the level sets of  $f$  by  $\{f \leq a\}$ ,  $a \in \mathbb{R}$ . The function  $f$  is lower semicontinuous if and only if all its level sets are closed. One says that  $f$  is inf-compact if all its level sets are compact.

The lower semicontinuous regularization of  $f$  is denoted by  $ls f$ .

The intrinsic core of a subset  $A$  of a vector space is  $icor A = \{x \in A; \forall x' \in \text{aff } A, \exists t > 0, [x, x + t(x' - x)] \subset A\}$  where  $\text{aff } A$  is the affine space spanned by  $A$ .  $icordom f$  is the intrinsic core of the effective domain of  $f$   $\text{dom } f = \{x \in X; f(x) < \infty\}$ .

The indicator function of a subset  $A$  of  $X$  is defined by

$$\iota_A(x) = \begin{cases} 0, & \text{if } x \in A, \\ +\infty, & \text{otherwise,} \end{cases} \quad x \in X. \tag{4}$$

The support function of  $A \subset X$  is  $\iota_A^*(y) = \sup_{x \in A} \langle x, y \rangle$ ,  $y \in Y$ .

**2. Statements of the results**

The dual equality and the primal attainment are stated in Theorem 2.1; the dual attainment and the dual representation of the minimizers are stated in Theorems 2.2 and 2.3. Their proofs are postponed to Section 5.

**2.1. Basic diagram**

Let  $\mathcal{U}_o$  be a vector space,  $\mathcal{V}_o = \mathcal{U}_o^*$  its algebraic dual space,  $\Phi_o$  a  $(-\infty, +\infty]$ -valued convex function on  $\mathcal{U}_o$  and  $\Phi_o^*$  its convex conjugate for the duality  $\langle \mathcal{U}_o, \mathcal{V}_o \rangle$ :

$$\Phi_o^*(v) := \sup_{u \in \mathcal{U}_o} \{\langle u, v \rangle - \Phi_o(u)\}, \quad v \in \mathcal{V}_o.$$

Let  $\mathcal{Y}_o$  be another vector space,  $\mathcal{X}_o = \mathcal{Y}_o^*$  its algebraic dual space and  $T_o : \mathcal{V}_o \rightarrow \mathcal{X}_o$  be a linear operator. We consider the convex minimization problem

$$\text{minimize } \Phi_o^*(v) \text{ subject to } T_o v \in C_o, \quad v \in \mathcal{V}_o \tag{P_o}$$

where  $C_o$  is a convex subset of  $\mathcal{X}_o$ .

It is useful to define the constraint operator  $T_o$  by means of its adjoint  $T_o^* : \mathcal{Y}_o \rightarrow \mathcal{V}_o^*$  as follows. For all  $v \in \mathcal{V}_o$ ,

$$\langle y, T_o v \rangle_{\mathcal{Y}_o, \mathcal{X}_o} = \langle T_o^* y, v \rangle_{\mathcal{V}_o^*, \mathcal{V}_o}, \quad \forall y \in \mathcal{Y}_o.$$

We shall assume that the restriction

$$T_o^*(\mathcal{Y}_o) \subset \mathcal{U}_o \tag{5}$$

holds, where  $\mathcal{U}_o$  is identified with a subspace of  $\mathcal{V}_o^* = \mathcal{U}_o^{**}$ . It follows that the diagram

$$\begin{array}{ccc} \langle \mathcal{U}_o & , & \mathcal{V}_o \rangle \\ T_o^* \uparrow & & \downarrow T_o \\ \langle \mathcal{Y}_o & , & \mathcal{X}_o \rangle \end{array}$$

Diagram 0.

is meaningful.

### 2.2. Assumptions

Let us give the list of our main hypotheses.

- $(H_\Phi)$  1.  $\Phi_o : \mathcal{U}_o \rightarrow [0, +\infty]$  is  $\sigma(\mathcal{U}_o, \mathcal{V}_o)$ -lower semicontinuous, 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.  $\ker T_o^* = \{0\}$ .
- $(H_C)$   $C := C_o \cap \mathcal{X}$  is a convex  $\sigma(\mathcal{X}, \mathcal{Y})$ -closed subset of  $\mathcal{X}$ .

The definitions of the vector spaces  $\mathcal{X}$  and  $\mathcal{Y}$  which appear in the last assumption are stated below in Subsection 2.3. For the moment, let us only say that if  $C_o$  is convex and  $\sigma(\mathcal{X}_o, \mathcal{Y}_o)$ -closed, then  $(H_C)$  holds.

*Comments about the assumptions.*

- By construction,  $\Phi_o^*$  is a convex  $\sigma(\mathcal{V}_o, \mathcal{U}_o)$ -lower semicontinuous function, even if  $\Phi_o$  doesn't satisfy  $(H_{\Phi_1})$ . Assuming this property of  $\Phi_o$  is not a restriction.
- The assumption  $(H_{\Phi_1})$  also expresses that  $\Phi_o$  achieves its minimum at  $u = 0$  and that  $\Phi_o(0) = 0$ . This is a practical normalization requirement which will allow us to build a gauge functional associated with  $\Phi_o$ . Moreover,  $(H_{\Phi_1})$  implies that  $\Phi_o^*$  also shares this property. Gauge functionals related to  $\Phi_o^*$  will also appear later.
- With any convex function  $\tilde{\Phi}$  satisfying  $(H_{\Phi_2})$ , one can associate a function  $\Phi_o$  satisfying  $(H_{\Phi_1})$  in the following manner. Because of  $(H_{\Phi_2})$ , one has  $0 \in \text{icordom } \tilde{\Phi}$  and hence there exists  $v_o \in \mathcal{V}_o$  such that  $v_o \in \partial \tilde{\Phi}(0)$ . Then, the lower semicontinuous convex regularization  $\Phi_o$  of  $u \in \mathcal{U}_o \mapsto \tilde{\Phi}(u) - \langle v_o, u \rangle - \tilde{\Phi}(0)$ , satisfies  $(H_{\Phi_1})$  and  $\tilde{\Phi}^*(v) = \Phi_o^*(v - v_o) - \tilde{\Phi}(0)$ ,  $v \in \mathcal{V}_o$ .

- The hypothesis  $(H_{\Phi_3})$  is not a restriction. Indeed, assuming  $(H_{\Phi_1})$ , let us suppose that there exists a direction  $u_o \neq 0$  such that  $\Phi_o(tu_o) = 0$  for all real  $t$ . Then any  $v \in \mathcal{V}_o$  such that  $\langle v, u_o \rangle \neq 0$  satisfies  $\Phi_o^*(v) \geq \sup_{t \in \mathbb{R}} t \langle v, u_o \rangle = +\infty$  and cannot be a solution to  $(P)$ .
- The hypothesis  $(H_{T_2})$  is not a restriction either: If  $y_1 - y_2 \in \ker T_o^*$ , we have  $\langle T_o v, y_1 \rangle = \langle T_o v, y_2 \rangle$ , for all  $v \in \mathcal{V}_o$ . In other words, the spaces  $\mathcal{Y}_o$  and  $\mathcal{Y}_o / \ker T_o^*$  both specify the same constraint sets  $\{v \in \mathcal{V}_o; T_o v = x\}$ .

The effective assumptions are the following ones.

- The specific form of the objective function  $\Phi_o^*$  as a convex conjugate makes it a convex  $\sigma(\mathcal{V}_o, \mathcal{U}_o)$ -lower semicontinuous function.
- $(H_{\Phi_2})$  and  $(H_C)$  are geometric restrictions.
- $(H_{T_1})$  is a regularity assumption on  $T_o$ .

### 2.3. Variants of $(P)$ and $(D)$

These variants are expressed below in terms of new spaces and functions. Let us first introduce them.

#### The norms $|\cdot|_{\Phi}$ and $|\cdot|_{\Lambda}$

Let  $\Phi_{\pm}(u) = \max(\Phi_o(u), \Phi_o(-u))$ . By  $(H_{\Phi_1})$  and  $(H_{\Phi_2})$ ,  $\{u \in \mathcal{U}_o; \Phi_{\pm}(u) \leq 1\}$  is a convex absorbing balanced set. Hence its gauge functional which is defined for all  $u \in \mathcal{U}_o$  by  $|u|_{\Phi} := \inf\{\alpha > 0; \Phi_{\pm}(u/\alpha) \leq 1\}$  is a seminorm. Thanks to hypothesis  $(H_{\Phi_3})$ , it is a norm.

Taking  $(H_{T_1})$  into account, one can define

$$\Lambda_o(y) := \Phi_o(T_o^* y), \quad y \in \mathcal{Y}_o. \tag{6}$$

Let  $\Lambda_{\pm}(y) = \max(\Lambda_o(y), \Lambda_o(-y))$ . The gauge functional on  $\mathcal{Y}_o$  of the set  $\{y \in \mathcal{Y}_o; \Lambda_{\pm}(y) \leq 1\}$  is  $|y|_{\Lambda} := \inf\{\alpha > 0; \Lambda_{\pm}(y/\alpha) \leq 1\}$ ,  $y \in \mathcal{Y}_o$ . Thanks to  $(H_{\Phi})$  and  $(H_T)$ , it is a norm and

$$|y|_{\Lambda} = |T_o^* y|_{\Phi}, \quad y \in \mathcal{Y}_o. \tag{7}$$

#### The spaces

Let

$$\begin{aligned} \mathcal{U} &\text{ be the } |\cdot|_{\Phi}\text{-completion of } \mathcal{U}_o \text{ and let} \\ \mathcal{V} &:= (\mathcal{U}_o, |\cdot|_{\Phi})' \text{ be the topological dual space of } (\mathcal{U}_o, |\cdot|_{\Phi}). \end{aligned}$$

Of course, we have  $(\mathcal{U}, |\cdot|_{\Phi})' \cong \mathcal{V} \subset \mathcal{V}_o$  where any  $v$  in  $\mathcal{V}$  is identified with its restriction to  $\mathcal{U}_o$ . Similarly, we introduce

$$\begin{aligned} \mathcal{Y} &\text{ the } |\cdot|_{\Lambda}\text{-completion of } \mathcal{Y}_o \text{ and} \\ \mathcal{X} &:= (\mathcal{Y}_o, |\cdot|_{\Lambda})' \text{ the topological dual space of } (\mathcal{Y}_o, |\cdot|_{\Lambda}). \end{aligned}$$

We have  $(\mathcal{Y}, |\cdot|_{\Lambda})' \cong \mathcal{X} \subset \mathcal{X}_o$  where any  $x$  in  $\mathcal{X}$  is identified with its restriction to  $\mathcal{Y}_o$ . We also have to consider the algebraic dual spaces  $\mathcal{V}^*$  and  $\mathcal{X}^*$  of  $\mathcal{V}$  and  $\mathcal{X}$ .

**The operators  $T$  and  $T^*$**

It will be proved in Lemma 4.5 that

$$T_o\mathcal{V} \subset \mathcal{X}. \tag{8}$$

Let us denote  $T$  the restriction of  $T_o$  to  $\mathcal{V} \subset \mathcal{V}_o$ . By (8), we have  $T : \mathcal{V} \rightarrow \mathcal{X}$ . Let us define its adjoint  $T^* : \mathcal{X}^* \rightarrow \mathcal{V}^*$  for all  $\omega \in \mathcal{X}^*$  by:

$$\langle v, T^*\omega \rangle_{\mathcal{V}, \mathcal{V}^*} = \langle Tv, \omega \rangle_{\mathcal{X}, \mathcal{X}^*}, \quad \forall v \in \mathcal{V}.$$

This definition is meaningful, thanks to (8). It will be proved in Lemma 4.5 that

$$T^*\mathcal{Y} \subset \mathcal{U}. \tag{9}$$

We have the inclusions  $\mathcal{Y}_o \subset \mathcal{Y} \subset \mathcal{X}^*$ . The adjoint operator  $T_o^*$  is the restriction of  $T^*$  to  $\mathcal{Y}_o$ .

**Some modifications of  $\Phi_o$  and  $\Lambda_o$**

We introduce the following modifications of  $\Phi_o$ :

$$\begin{aligned} \Phi(u) &:= \sup_{v \in \mathcal{V}} \{ \langle u, v \rangle - \Phi_o^*(v) \}, \quad u \in \mathcal{U} \\ \bar{\Phi}(\zeta) &:= \sup_{v \in \mathcal{V}} \{ \langle v, \zeta \rangle - \Phi^*(v) \}, \quad \zeta \in \mathcal{V}^*. \end{aligned}$$

They are respectively  $\sigma(\mathcal{U}, \mathcal{V})$  and  $\sigma(\mathcal{V}^*, \mathcal{V})$ -lower semicontinuous convex functions. It is immediate to see that the restriction of  $\bar{\Phi}$  to  $\mathcal{U}$  is  $\Phi$ . As  $\mathcal{V} = \mathcal{U}'$ ,  $\Phi$  is also the  $|\cdot|_{\Phi}$ -lower semicontinuous convex regularization of  $\Phi_o$ . The function  $\bar{\Phi}$  is the extension which appears in the introductory Section 1. We also introduce

$$\begin{aligned} \Lambda(y) &:= \Phi(T^*y), \quad y \in \mathcal{Y} \\ \bar{\Lambda}(\omega) &:= \bar{\Phi}(T^*\omega), \quad \omega \in \mathcal{X}^* \end{aligned}$$

which look like the definition (6). Note that thanks to (9), the first equality is meaningful. Because of the previous remarks, the restriction of  $\bar{\Lambda}$  to  $\mathcal{Y}$  is  $\Lambda$ .

**The optimization problems**

Let  $\Phi_o^*$  and  $\Phi^*$  be the convex conjugates of  $\Phi_o$  and  $\Phi$  with respect to the dual pairings  $\langle \mathcal{U}_o, \mathcal{V}_o \rangle$  and  $\langle \mathcal{U}, \mathcal{V} \rangle$ :

$$\begin{aligned} \Phi_o^*(v) &:= \sup_{u \in \mathcal{U}_o} \{ \langle u, v \rangle - \Phi_o(u) \}, \quad v \in \mathcal{V}_o \\ \Phi^*(v) &:= \sup_{u \in \mathcal{U}} \{ \langle u, v \rangle - \Phi(u) \}, \quad v \in \mathcal{V} \end{aligned}$$

and  $\Lambda_o^*, \Lambda^*$  be the convex conjugates of  $\Lambda_o, \Lambda$  with respect to the dual pairings  $\langle \mathcal{Y}_o, \mathcal{X}_o \rangle$  and  $\langle \mathcal{Y}, \mathcal{X} \rangle$ :

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

Finally, denote  $C = C_o \cap \mathcal{X}$ .

The optimization problems to be considered are:

$$\begin{aligned}
 & \text{minimize } \Phi_o^*(v) \text{ subject to } T_o v \in C_o, \quad v \in \mathcal{V}_o & (P_o) \\
 & \text{minimize } \Phi^*(v) \text{ subject to } Tv \in C, \quad v \in \mathcal{V} & (P) \\
 & \text{minimize } \Lambda^*(x) \text{ subject to } x \in C, \quad x \in \mathcal{X} & (P^{\mathcal{X}}) \\
 & \text{maximize } \inf_{x \in C_o} \langle y, x \rangle - \Lambda_o(y), \quad y \in \mathcal{Y}_o & (D_o) \\
 & \text{maximize } \inf_{x \in C} \langle y, x \rangle - \Lambda(y), \quad y \in \mathcal{Y} & (D) \\
 & \text{maximize } \inf_{x \in C} \langle x, \omega \rangle - \bar{\Lambda}(\omega), \quad \omega \in \mathcal{X}^*. & (\bar{D})
 \end{aligned}$$

### 2.4. Statements

We are now ready to give answers to the Questions 1.1 related to (P) and (D).

**Theorem 2.1 (Primal attainment and dual equality).** *Assume that  $(H_{\Phi})$  and  $(H_T)$  hold.*

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

$$\inf\{\Phi_o^*(v); v \in \mathcal{V}_o, T_o v = x\} = \Lambda_o^*(x) \in [0, \infty]. \quad (10)$$

*Moreover, in restriction to  $\mathcal{X}$ ,  $\Lambda_o^* = \Lambda^*$  and  $\Lambda^*$  is  $\sigma(\mathcal{X}, \mathcal{Y})$ -inf-compact.*

(b) *The problems  $(P_o)$  and  $(P)$  are equivalent: they have the same solutions and  $\inf(P_o) = \inf(P) \in [0, \infty]$ .*

(c) *If  $C_o$  is convex and  $\sigma(\mathcal{X}_o, \mathcal{Y}_o)$ -closed, we have the dual equality*

$$\inf(P_o) = \sup(D_o) \in [0, \infty].$$

*Assume that  $(H_{\Phi})$ ,  $(H_T)$  and  $(H_C)$  hold.*

(d) *We have the dual equalities*

$$\inf(P_o) = \inf(P) = \sup(D) = \sup(\bar{D}) = \inf_{x \in C_o} \Lambda_o^*(x) = \inf_{x \in C} \Lambda^*(x) \in [0, \infty]. \quad (11)$$

(e) *If in addition  $\inf(P_o) < \infty$ , then  $(P_o)$  is attained in  $\mathcal{V}$ . Moreover, any minimizing sequence of  $(P_o)$  has  $\sigma(\mathcal{V}, \mathcal{U})$ -cluster points and every such cluster point solves  $(P_o)$ .*

(f) *Let  $\hat{v} \in \mathcal{V}$  be a solution to  $(P)$ , then  $\hat{x} := T_o \hat{v}$  is a solution to  $(P^{\mathcal{X}})$  and  $\inf(P) = \Phi^*(\hat{v}) = \Lambda_o^*(\hat{x})$ .*

**Theorem 2.2 (Dual attainment and representation. Interior convex constraint).** *Assume that  $(H_{\Phi})$ ,  $(H_T)$  and  $(H_C)$  hold.*

(1) *For any  $\hat{v} \in \mathcal{V}$  and  $\bar{\omega} \in \mathcal{X}^*$ ,*

$$\begin{cases}
 (a) & T\hat{v} \in C \\
 (b) & \langle \bar{\omega}, T\hat{v} \rangle \leq \langle \bar{\omega}, x \rangle \text{ for all } x \in C \\
 (c) & \hat{v} \in \partial_{\mathcal{V}} \bar{\Phi}(T^* \bar{\omega})
 \end{cases} \quad (12)$$



is equivalent to

$$\begin{cases} \hat{v} \text{ is a solution to } (P_o) \text{ and} \\ \bar{\omega} \text{ is a solution to } (\bar{D}). \end{cases}$$

(2) Suppose that in addition the interior constraint qualification

$$C_o \cap \text{icor}(T_o \text{dom } \Phi_o^*) \neq \emptyset \tag{13}$$

is satisfied. Then, the primal problem  $(P_o)$  is attained in  $\mathcal{V}$  and the dual problem  $(\bar{D})$  is attained in  $\mathcal{X}^*$ .

Note that (13) is equivalent to  $C_o \cap \text{icordom } \Lambda_o^* \neq \emptyset$ .

As can be seen in [10], the Monge-Kantorovich problem provides an example where no constraint is interior. In order to solve it, we are going to consider the more general situation (1) where the constraint is said to be a *subgradient constraint*. This means that  $\hat{x}$  belongs to

$$\begin{aligned} \text{dom } \partial_{\mathcal{X}^*} \Lambda_o^* &= \{x \in \mathcal{X}; \partial_{\mathcal{X}^*} \Lambda_o^*(x) \neq \emptyset\}, \quad \text{where} \\ \partial_{\mathcal{X}^*} \Lambda_o^*(x) &= \{\omega \in \mathcal{X}^*; \Lambda_o^*(x') \geq \Lambda_o^*(x) + \langle x' - x, \omega \rangle, \forall x' \in \mathcal{X}\}. \end{aligned}$$

Two new optimization problems to be considered are

$$\begin{aligned} \text{minimize } \Phi_o^*(v) \text{ subject to } T_o v = \hat{x}, \quad v \in \mathcal{V}_o & \tag{P^{\hat{x}}} \\ \text{maximize } \langle \hat{x}, \omega \rangle - \bar{\Lambda}(\omega), \quad \omega \in \mathcal{X}^* & \tag{D^{\hat{x}}} \end{aligned}$$

where  $\hat{x} \in \mathcal{X}_o$ . This corresponds to the simplified case where  $C_o$  is reduced to the single point  $\hat{x}$ .

**Theorem 2.3 (Dual attainment and representation. Subgradient affine constraint).** *Let us assume that  $(H_\Phi)$  and  $(H_T)$  hold.*

(1) For any  $\hat{v} \in \mathcal{V}$  and  $\bar{\omega} \in \mathcal{X}^*$ ,

$$\begin{cases} (a) \quad T\hat{v} = \hat{x} \\ (b) \quad \hat{v} \in \partial_{\mathcal{V}} \bar{\Phi}(T^* \bar{\omega}) \end{cases} \tag{14}$$

is equivalent to

$$\begin{cases} \hat{v} \text{ is a solution to } (P^{\hat{x}}) \text{ and} \\ \bar{\omega} \text{ is a solution to } (\bar{D}^{\hat{x}}). \end{cases}$$

(2) Suppose that in addition the subgradient constraint qualification

$$\hat{x} \in \text{dom } \partial_{\mathcal{X}^*} \Lambda_o^*, \tag{15}$$

is satisfied. Then, the primal problem  $(P^{\hat{x}})$  is attained in  $\mathcal{V}$ , and the dual problem  $(\bar{D}^{\hat{x}})$  is attained in  $\mathcal{X}^*$ .

It is well-known that the representation formula (12(c)) or (14(b)):

$$\hat{v} \in \partial_{\mathcal{V}} \bar{\Phi}(T^* \bar{\omega}) \tag{16}$$

is equivalent to

$$T^*\bar{\omega} \in \partial_{\mathcal{V}^*}\Phi^*(\hat{v})$$

and also equivalent to Fenchel's identity

$$\Phi^*(\hat{v}) + \bar{\Phi}(T^*\bar{\omega}) = \langle \bar{\omega}, T\hat{v} \rangle = \Lambda^*(\hat{x}) + \bar{\Lambda}(\bar{\omega}). \tag{17}$$

Formula (16) can be made a little more precise by means of the following regularity result.

**Theorem 2.4.** *Assume that  $(H_\Phi)$ ,  $(H_T)$  and  $(H_C)$  hold. Any solution  $\bar{\omega}$  of  $(\bar{D})$  or  $(\bar{D}^{\hat{x}})$  shares the following properties*

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

If in addition the level sets of  $\Phi : \{\Phi \leq a\}, a \in \mathbb{R}$ , are  $|\cdot|_\Phi$ -bounded, then

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

where  $\mathcal{Y}''$  and  $\mathcal{U}''$  are the topological bidual spaces of  $(\mathcal{Y}, |\cdot|_\Lambda)$  and  $(\mathcal{U}, |\cdot|_\Phi)$ . This occurs when  $\Phi$ , and therefore  $\Phi^*$ , is an even function.

### 3. Examples

The abstract results of Section 2 are exemplified by means of the Monge-Kantorovich optimal transport problem and the problem of minimizing entropy functionals on convex sets.

#### 3.1. The Monge-Kantorovich optimal transport problem

Denote  $P_A, P_B$  and  $P_{AB}$  the sets of all probability measures on the spaces  $A, B$  and  $A \times B$ . Let  $c : A \times B \rightarrow [0, \infty)$  a nonnegative (cost) function and two probability measures  $\mu \in P_A$  and  $\nu \in P_B$  on  $A$  and  $B$ . The Monge-Kantorovich problem is

$$\text{minimize } \pi \in P_{AB} \mapsto \int_{A \times B} c(a, b) \pi(dadb) \text{ subject to } \pi \in P(\mu, \nu) \tag{MK}$$

where  $P(\mu, \nu)$  is the set of all  $\pi \in P_{AB}$  with prescribed marginals  $\pi_A = \mu$  on  $A$  and  $\pi_B = \nu$  on  $B$ . Any solution of (MK) is called an optimal plan. For a general account on this active field of research, see Villani's books [15, 16].

Without going into the details, let us indicate how this problem enters the present framework. Assume that  $A$  and  $B$  are topological spaces and denote  $C_A, C_B$  and  $C_{AB}$  the spaces of all continuous bounded functions on  $A, B$  and  $A \times B$ . The function  $\Phi_o$  is defined on the space  $\mathcal{U}_o = C_{AB}$  by

$$\Phi_o(u) = \iota_{\{u \leq c\}}, \quad u \in C_{AB}$$

see (4). The marginal constraint  $\pi \in P(\mu, \nu)$  is obtained by choosing  $\mathcal{Y}_o = C_A \times C_B$  and

$$T_o^*(f, g) = f \oplus g, \quad f \in C_A, \quad g \in C_B$$

with

$$f \oplus g(a, b) := f(a) + g(b), \quad a \in A, \quad b \in B,$$

see Section 3.3 below. This gives  $\Lambda_o(f, g) = \iota_{\{f \oplus g \leq c\}}$  and the dual equality (10) is the well-known Kantorovich dual equality

$$\begin{aligned} & \inf \left\{ \int_{A \times B} c(a, b) \pi(dadb); \pi \in P(\mu, \nu) \right\} \\ &= \sup \left\{ \int_A f(a) \mu(da) + \int_B g(b) \nu(db); f \in C_A, g \in C_B : f \oplus g \leq c \right\}. \end{aligned}$$

In [10], cost functions  $c$  which may take infinite values are considered and Theorem 2.3 is used to derive a necessary condition of optimality, yielding a new result on this well-known problem.

### 3.2. Entropy minimization

The problem is sketched in this section and studied in further details in [11].

#### Entropy

Let  $R$  be a positive measure on a space  $\mathcal{Z}$  and take a  $[0, \infty]$ -valued jointly measurable function  $\gamma^*$  on  $\mathcal{Z} \times \mathbb{R}$  such that  $\gamma^*(z, \cdot) := \gamma_z^*$  is convex and lower semicontinuous for all  $z \in \mathcal{Z}$ . Denote  $M_{\mathcal{Z}}$  the space of all signed measures  $Q$  on  $\mathcal{Z}$ . The entropy functional to be considered is defined by

$$I(Q) = \begin{cases} \int_{\mathcal{Z}} \gamma_z^* \left( \frac{dQ}{dR}(z) \right) R(dz) & \text{if } Q \prec R, \\ +\infty & \text{otherwise,} \end{cases} \quad Q \in M_{\mathcal{Z}} \tag{18}$$

where  $Q \prec R$  means that  $Q$  is absolutely continuous with respect to  $R$ . Assume that for each  $z$  there exists a unique  $m(z)$  which minimizes  $\gamma_z^*$  with  $\gamma_z^*(m(z)) = 0$ . Then,  $I$  is  $[0, \infty]$ -valued, its unique minimizer is  $mR$  and  $I(mR) = 0$ .

As for each  $z$ ,  $\gamma_z^*$  is lower semicontinuous convex, it is the convex conjugate of some lower semicontinuous convex function  $\gamma_z$ . Defining

$$\lambda(z, s) = \gamma(z, s) - m(z)s, \quad z \in \mathcal{Z}, \quad s \in \mathbb{R},$$

one sees that for  $R$ -a.e.  $z$ ,  $\lambda_z$  is a nonnegative convex function and it vanishes at 0. A favorable choice for  $\mathcal{U}_o$  is the space of all measurable functions  $u$  on  $\mathcal{Z}$  such that

$$\int_{\mathcal{Z}} \lambda(z, \alpha u(z)) R(dz) < \infty, \quad \text{for all } \alpha \in \mathbb{R}. \tag{19}$$

With

$$\lambda_o(z, s) = \max[\lambda(z, s), \lambda(z, -s)] \in [0, \infty], \quad z \in \mathcal{Z}, \quad s \in \mathbb{R},$$

(19) is equivalent to  $u$  belongs to

$$\mathcal{U}_o = \mathcal{E}_{\lambda_o} := \left\{ u; \int_{\mathcal{Z}} \lambda_o(z, \alpha u(z)) R(dz) < \infty, \forall \alpha > 0 \right\}$$

the “small” Orlicz space associated with the Young function  $\lambda_\circ$ . Taking

$$\Phi_\circ(u) = \int_{\mathcal{Z}} \lambda_z(u(z)) R(dz) \in [0, \infty], \quad u \in \mathcal{E}_{\lambda_\circ} \tag{20}$$

leads to

$$I(Q) = \Phi_\circ^*(Q - mR), \quad Q \in M_{\mathcal{Z}}. \tag{21}$$

This identity is a consequence of general results of Rockafellar on conjugate duality for integral functionals [12]. Moreover, the effective domain of  $I$  is included in the space

$$M_{\mathcal{Z}}^{\lambda_\circ} := \left\{ Q \in M_{\mathcal{Z}}; \int_{\mathcal{Z}} |u| d|Q| < \infty, \forall u \in \mathcal{E}_{\lambda_\circ} \right\}.$$

**Constraint**

In order to define the constraint, take  $\mathcal{X}_\circ$  a vector space and a function  $\theta : \mathcal{Z} \rightarrow \mathcal{X}_\circ$ . One wants to give some meaning to the formal constraint  $\int_{\mathcal{Z}} \theta dQ = x$  with  $Q \in M_{\mathcal{Z}}^{\lambda_\circ}$  and  $x \in \mathcal{X}_\circ$ . Suppose that  $\mathcal{X}_\circ$  is the algebraic dual space of some vector space  $\mathcal{Y}_\circ$  and define for all  $y \in \mathcal{Y}_\circ$ ,

$$T_\circ^*y(z) := \langle y, \theta(z) \rangle_{\mathcal{Y}_\circ, \mathcal{X}_\circ}, \quad z \in \mathcal{Z}. \tag{22}$$

Assuming that

$$T_\circ^*y \in \mathcal{E}_{\lambda_\circ}, \quad \forall y \in \mathcal{Y}_\circ$$

allows us to define the constraint operator

$$T_\circ Q := \int_{\mathcal{Z}} \theta dQ, \quad Q \in M_{\mathcal{Z}}^{\lambda_\circ}$$

by

$$\left\langle y, \int_{\mathcal{Z}} \theta dQ \right\rangle_{\mathcal{Y}_\circ, \mathcal{X}_\circ} = \int_{\mathcal{Z}} \langle y, \theta(z) \rangle_{\mathcal{Y}_\circ, \mathcal{X}_\circ} Q(dz), \quad \forall y \in \mathcal{Y}_\circ.$$

**Minimization problem**

The entropy minimization problem to be considered is

$$\text{minimize } I(Q) \text{ subject to } \int_{\mathcal{Z}} \theta dQ \in C_\circ, \quad Q \in M_{\mathcal{Z}}^{\lambda_\circ},$$

where  $C_\circ$  is a convex subset of  $\mathcal{X}_\circ$ .

**Results**

Applying the abstract results of the present paper, in [11] are obtained the following results. Let  $\Gamma^*(x) = \sup_{y \in \mathcal{Y}_\circ} \{ \langle y, x \rangle - \int_{\mathcal{Z}} \gamma_z(\langle y, \theta(z) \rangle) R(dz) \}$ ,  $x \in \mathcal{X}_\circ$ . The dual equality is  $\inf\{I(Q); Q \in M_{\mathcal{Z}}^{\lambda_\circ}, \int_{\mathcal{Z}} \theta dQ \in C_\circ\} = \inf_{C_\circ} \Gamma^*$  and under the assumption

$$C_\circ \cap \text{icordom } \Gamma^* \neq \emptyset, \tag{23}$$

the characterization of the minimizer  $\widehat{Q}$  is as follows. Defining  $\widehat{x} \triangleq \int_{\mathcal{Z}} \theta d\widehat{Q}$  in the weak sense with respect to the duality  $\langle \mathcal{Y}_\circ, \mathcal{X}_\circ \rangle$ ,  $\widehat{Q}$  is a minimizer if and only if there exists

some linear form  $\bar{y}$  on  $\mathcal{X}_o$  such that  $\langle \bar{y}, \theta(\cdot) \rangle$  is measurable,  $\int_{\mathcal{Z}} \lambda_o(z, \alpha_o \langle \bar{y}, \theta(z) \rangle) R(dz) < \infty$  for some  $\alpha_o > 0$  and

$$\begin{cases} (a) & \hat{x} \in C_o \cap \text{dom } \Gamma^* \\ (b) & \langle \bar{y}, \hat{x} \rangle \leq \langle \bar{y}, x \rangle, \forall x \in C_o \cap \text{dom } \Gamma^* \\ (c) & \widehat{Q}(dz) = \gamma'_z(\langle \bar{y}, \theta(z) \rangle) R(dz). \end{cases} \tag{24}$$

where  $\gamma'_z(s) = \frac{\partial}{\partial s} \gamma(z, s)$ .

**Remark 3.1.** A usual form of constraint qualification required for this representation is  $\text{int } C_o \cap \text{dom } \Gamma^* \neq \emptyset$  where  $\text{int } C_o$  is the interior of  $C_o$  with respect to some topology which is not directly connected to the “geometry” of  $\Gamma^*$ . In particular,  $\text{int } C_o$  must be nonempty; this is a considerable restriction. The constraint qualification  $C_o \cap \text{icordom } \Gamma^* \neq \emptyset$  is much weaker.

**Literature about entropy minimization**

Entropy minimization problems appear in many areas of applied mathematics and sciences. The literature about the minimization of entropy functionals under convex constraints is considerable: many papers are concerned with an engineering approach, working on the implementation of numerical procedures in specific situations. In fact, entropy minimization is a popular method to solve ill-posed inverse problems.

Surprisingly, rigorous general results on this topic are quite recent. Let us cite, among others, the main contribution of Borwein and Lewis: [1, 2, 3, 4, 5, 6] together with the paper [14] by Teboulle and Vajda. In these papers, topological constraint qualifications are required: it is assumed that the constraints stand in some topological interior of the domain of  $I$ . Such restrictions are removed in [11].

**3.3. Some examples of constraints**

Let us consider the two standard constraints which are the moment constraints and the marginal constraints.

**Moment constraints**

Let  $\theta = (\theta_k)_{1 \leq k \leq K}$  be a measurable function from  $\mathcal{Z}$  to  $\mathcal{X}_o = \mathbb{R}^K$ . The moment constraint is defined by

$$\int_{\mathcal{Z}} \theta dQ = \left( \int_{\mathcal{Z}} \theta_k dQ \right)_{1 \leq k \leq K} \in \mathbb{R}^K,$$

for each  $Q \in M_{\mathcal{Z}}$  which integrates all the real valued measurable functions  $\theta_k$ .

**Marginal constraints**

Let  $\mathcal{Z} = A \times B$  be a product space,  $M_{AB}$  be the space of all *bounded* signed measures on  $A \times B$  and  $U_{AB}$  be the space of all measurable bounded functions  $u$  on  $A \times B$ . Denote  $v_A = v(\cdot \times B)$  and  $v_B = v(A \times \cdot)$  the marginal measures of  $v \in M_{AB}$ . The constraint of prescribed marginal measures is specified by

$$\int_{A \times B} \theta dv = (v_A, v_B) \in M_A \times M_B, \quad v \in M_{AB}$$

where  $M_A$  and  $M_B$  are the spaces of all bounded signed measures on  $A$  and  $B$ . The function  $\theta$  which gives the marginal constraint is

$$\theta(a, b) = (\delta_a, \delta_b), \quad a \in A, \quad b \in B$$

where  $\delta_a$  is the Dirac measure at  $a$ . Indeed,  $(v_A, v_B) = \int_{A \times B} (\delta_a, \delta_b) v(dadb)$ .

More precisely, let  $U_A, U_B$  be the spaces of measurable functions on  $A$  and  $B$  and take  $\mathcal{Y}_o = U_A \times U_B$  and  $\mathcal{X}_o = U_A^* \times U_B^*$ . Then,  $\theta$  is a measurable function from  $\mathcal{Z} = A \times B$  to  $\mathcal{X}_o = U_A^* \times U_B^*$ . It is easy to see that the adjoint of the marginal operator

$$T_o v = (v_A, v_B) \in U_A^* \times U_B^*, \quad v \in \mathcal{Y}_o = U_{AB}^*$$

where  $\langle f, v_A \rangle := \langle f \otimes 1, v \rangle$  and  $\langle g, v_B \rangle := \langle 1 \otimes g, v \rangle$  for all  $f \in U_A$  and  $g \in U_B$ , is given by

$$T_o^*(f, g) = f \oplus g \in U_{AB}, \quad f \in U_A, \quad g \in U_B \tag{25}$$

where  $f \oplus g(a, b) := f(a) + g(b)$ ,  $a \in A, b \in B$ .

#### 4. Preliminary results

In this section, one introduces notation and proves preliminary technical results for the proofs of the results of Section 2.

##### 4.1. The saddle-point method (for fixing notation)

We are going to apply the general results of the Lagrangian approach to the minimization problem  $(P_o)$ . To quote easily and precisely some well-known results of convex minimization while proving our results, we give a short overview of the approach to convex minimization problems by means of conjugate duality as developed in Rockafellar's monograph [13].

Let  $A$  be a vector space and  $f : A \rightarrow [-\infty, +\infty]$  an extended-real-valued convex function. We consider the following convex minimization problem

$$\text{minimize } f(a), \quad a \in A. \tag{\mathcal{P}}$$

Let  $Q$  be another vector space. The perturbation of the objective function  $f$  is a function  $F : A \times Q \rightarrow [-\infty, +\infty]$  such that for  $q = 0 \in Q$ ,  $F(\cdot, 0) = f(\cdot)$ . The problem  $(\mathcal{P})$  is imbedded in a parameterized family of minimization problems

$$\text{minimize } F(a, q), \quad a \in A. \tag{\mathcal{P}_q}$$

The value function of  $(\mathcal{P}_q)_{q \in Q}$  is

$$\varphi(q) := \inf(\mathcal{P}_q) = \inf_{a \in A} F(a, q) \in [-\infty, +\infty], \quad q \in Q.$$

Let us assume that the perturbation is chosen such that

$$F \text{ is jointly convex on } A \times Q. \tag{26}$$

Then,  $(\mathcal{P}_q)_{q \in Q}$  is a family of convex minimization problems and the value function  $\varphi$  is convex.

Let  $B$  be a vector space in dual pairing with  $Q$ . This means that  $B$  and  $Q$  are locally convex topological vector spaces in separating duality such that their topological dual spaces  $B'$  and  $Q'$  satisfy  $B' = Q$  and  $Q' = B$  up to some isomorphisms. The Lagrangian associated with the perturbation  $F$  and the duality  $\langle B, Q \rangle$  is

$$K(a, b) := \inf_{q \in Q} \{ \langle b, q \rangle + F(a, q) \}, \quad a \in A, \quad b \in B. \tag{27}$$

Under (26),  $K$  is a convex-concave function. Assuming in addition that  $F$  is chosen such that

$$q \mapsto F(a, q) \text{ is a closed convex function for any } a \in A, \tag{28}$$

one can reverse the conjugate duality relation (27) to obtain

$$F(a, q) = \sup_{b \in B} \{ K(a, b) - \langle b, q \rangle \}, \quad \forall a \in A, \quad q \in Q. \tag{29}$$

Recall that for a convex function to be closed, it is enough that it is lower semicontinuous and it doesn't take the value  $-\infty$ .

Introducing another vector space  $P$  in separating duality with  $A$  we define the function

$$G(b, p) := \inf_{a \in A} \{ K(a, b) - \langle a, p \rangle \}, \quad b \in B, \quad p \in P. \tag{30}$$

This formula is analogous to (29). Going on symmetrically, one interprets  $G$  as the concave perturbation of the objective concave function

$$g(b) := G(b, 0), \quad b \in B$$

associated with the concave maximization problem

$$\text{maximize } g(b), \quad b \in B \tag{\mathcal{D}}$$

which is the dual problem of  $(\mathcal{P})$ . It is imbedded in the family of concave maximization problems  $(\mathcal{D}_p)_{p \in P}$

$$\text{maximize } G(b, p), \quad b \in B \tag{\mathcal{D}_p}$$

whose value function is

$$\gamma(p) := \sup_{b \in B} G(b, p), \quad p \in P.$$

Since  $G$  is jointly concave,  $\gamma$  is also concave. We have the following diagram

$$G(b, p) \begin{array}{ccc} \left\langle \begin{array}{c} \gamma(p) \\ P \end{array}, \begin{array}{c} f(a) \\ A \end{array} \right\rangle \\ \left\langle \begin{array}{c} B \\ g(b) \end{array}, \begin{array}{c} Q \\ \varphi(q) \end{array} \right\rangle \end{array} K(a, b) \quad F(a, q)$$

The concave conjugate of the function  $f$  with respect to the dual pairing  $\langle Y, X \rangle$  is  $f^*(y) = \inf_x \{ \langle y, x \rangle - f(x) \}$  and its superdifferential at  $x$  is  $\widehat{\partial}f(x) = \{ y \in Y; f(x') \leq f(x) + \langle y, x' - x \rangle \}$ .

**Theorem 4.1.** We assume that  $\langle P, A \rangle$  and  $\langle B, Q \rangle$  are topological dual pairings.

- (a) We have  $\sup(\mathcal{D}) = \varphi^{**}(0)$ . Hence, the dual equality  $\inf(\mathcal{P}) = \sup(\mathcal{D})$  holds if and only if  $\varphi(0) = \varphi^{**}(0)$ .
- (b) In particular,

$$\left. \begin{array}{l} \bullet F \text{ is jointly convex} \\ \bullet \varphi \text{ is lower semicontinuous at } 0 \\ \bullet \sup(\mathcal{D}) > -\infty \end{array} \right\} \Rightarrow \inf(\mathcal{P}) = \sup(\mathcal{D}).$$

- (c) If the dual equality holds, then

$$\operatorname{argmax} g = -\partial\varphi(0).$$

Let us assume in addition that (26) and (28) are satisfied.

- (a') We have  $\inf(\mathcal{P}) = \gamma^{\hat{*}}(0)$ . Hence, the dual equality  $\inf(\mathcal{P}) = \sup(\mathcal{D})$  holds if and only if  $\gamma(0) = \gamma^{\hat{*}}(0)$ .
- (b') In particular,

$$\left. \begin{array}{l} \bullet \gamma \text{ is upper semicontinuous at } 0 \\ \bullet \inf(\mathcal{P}) < +\infty \end{array} \right\} \Rightarrow \inf(\mathcal{P}) = \sup(\mathcal{D}).$$

- (c') If the dual equality holds, then

$$\operatorname{argmin} f = -\widehat{\partial}\gamma(0).$$

**Definition 4.2 (Saddle-point).** One says that  $(\bar{a}, \bar{b}) \in A \times B$  is a saddle-point of the function  $K$  if

$$K(\bar{a}, b) \leq K(\bar{a}, \bar{b}) \leq K(a, \bar{b}), \quad \forall a \in A, b \in B.$$

**Theorem 4.3 (Saddle-point theorem and KKT relations).** The following statements are equivalent.

- (1) The point  $(\bar{a}, \bar{b})$  is a saddle-point of the Lagrangian  $K$ .
- (2)  $f(\bar{a}) \leq g(\bar{b})$ .
- (3) The following three statements hold
  - (a) we have the dual equality:  $\sup(\mathcal{D}) = \inf(\mathcal{P})$ ,
  - (b)  $\bar{a}$  is a solution to the primal problem  $(\mathcal{P})$  and
  - (c)  $\bar{b}$  is a solution to the dual problem  $(\mathcal{D})$ .

In this situation, one also gets

$$\sup(\mathcal{D}) = \inf(\mathcal{P}) = K(\bar{a}, \bar{b}) = f(\bar{a}) = g(\bar{b}). \tag{31}$$

Moreover,  $(\bar{a}, \bar{b})$  is a saddle-point of  $K$  if and only if it satisfies

$$\partial_a K(\bar{a}, \bar{b}) \ni 0 \tag{32}$$

$$\widehat{\partial}_b K(\bar{a}, \bar{b}) \ni 0 \tag{33}$$

where the subscript  $a$  or  $b$  indicates the unfixed variable.



### 4.2. Gauge functionals associated with a convex function

The following result is probably well-known, but since I didn't find a reference for it, I give its short proof.

Let  $\theta : S \rightarrow [0, \infty]$  be an extended nonnegative convex function on a vector space  $S$ , such that  $\theta(0) = 0$ . Let  $S^*$  be the algebraic dual space of  $S$  and  $\theta^*$  the convex conjugate of  $\theta$ :

$$\theta^*(r) := \sup_{s \in S} \{\langle r, s \rangle - \theta(s)\}, \quad r \in S^*.$$

It is easy to show that  $\theta^* : S^* \rightarrow [0, \infty]$  and  $\theta^*(0) = 0$ . We denote  $C_\theta := \{\theta \leq 1\}$  and  $C_{\theta^*} := \{\theta^* \leq 1\}$  the unit level sets of  $\theta$  and  $\theta^*$ . The gauge functionals to be considered are

$$j_\theta(s) := \inf\{\alpha > 0; s \in \alpha C_\theta\} = \inf\{\alpha > 0; \theta(s/\alpha) \leq 1\} \in [0, \infty], \quad s \in S.$$

$$j_{\theta^*}(r) := \inf\{\alpha > 0; r \in \alpha C_{\theta^*}\} = \inf\{\alpha > 0; \theta^*(r/\alpha) \leq 1\} \in [0, \infty], \quad r \in S^*.$$

As 0 belongs to  $C_\theta$  and  $C_{\theta^*}$ , one easily proves that  $j_\theta$  and  $j_{\theta^*}$  are positively homogeneous. Similarly, as  $C_\theta$  and  $C_{\theta^*}$  are convex sets,  $j_\theta$  and  $j_{\theta^*}$  are convex functions.

**Proposition 4.4.** *For all  $r \in S^*$ , we have*

$$\frac{1}{2}j_{\theta^*}(r) \leq \iota_{C_\theta}^*(r) := \sup_{s \in C_\theta} \langle r, s \rangle \leq 2j_{\theta^*}(r).$$

We also have

$$\text{cone dom } \theta^* = \text{dom } j_{\theta^*} = \text{dom } \iota_{C_\theta}^*$$

where  $\text{cone dom } \theta^*$  is the convex cone (with vertex 0) generated by  $\text{dom } \theta^*$ .

**Proof.** • Let us first show that  $\iota_{C_\theta}^*(r) \leq 2j_{\theta^*}(r)$  for all  $r \in S^*$ . For all  $s \in C_\theta$  and  $\alpha > j_{\theta^*}(r)$ , we have  $\langle r, s \rangle = \langle r/\alpha, s \rangle \alpha \leq [\theta(s) + \theta^*(r/\alpha)]\alpha \leq (1 + 1)\alpha$ . Then, optimize both sides of this inequality.

• Let us show that  $j_{\theta^*}(r) \leq 2\iota_{C_\theta}^*(r)$ . If  $\iota_{C_\theta}^*(r) = \infty$ , there is nothing to prove. So, let us suppose that  $\iota_{C_\theta}^*(r) < \infty$ . As  $0 \in C_\theta$ , we have  $\iota_{C_\theta}^*(r) \geq 0$ .

*First case:*  $\iota_{C_\theta}^*(r) > 0$ . For all  $s \in S$  and  $\epsilon > 0$ , we have  $s/[j_\theta(s) + \epsilon] \in C_\theta$ . It follows that  $\langle r/\iota_{C_\theta}^*(r), s \rangle = \langle r, s/[j_\theta(s) + \epsilon] \rangle \frac{j_\theta(s) + \epsilon}{\iota_{C_\theta}^*(r)} \leq \iota_{C_\theta}^*(r) \frac{j_\theta(s) + \epsilon}{\iota_{C_\theta}^*(r)} = j_\theta(s) + \epsilon$ . Therefore,  $\langle r/\iota_{C_\theta}^*(r), s \rangle \leq j_\theta(s)$ , for all  $s \in S$ .

If  $s$  doesn't belong to  $C_\theta$ , then  $j_\theta(s) \leq \theta(s)$ . This follows from the the assumptions on  $\theta$ : convex function such that  $\theta(0) = 0 = \min \theta$  and the positive homogeneity of  $j_\theta$ . Otherwise, if  $s$  belongs to  $C_\theta$ , we have  $j_\theta(s) \leq 1$ . Hence,  $\langle r/\iota_{C_\theta}^*(r), s \rangle \leq \max(1, \theta(s)), \forall s \in S$ . On the other hand, there exists  $s_o \in S$  such that  $\theta^*(r/[2\iota_{C_\theta}^*(r)]) \leq \langle r/[2\iota_{C_\theta}^*(r)], s_o \rangle - \theta(s_o) + 1/2$ . The last two inequalities provide us with  $\theta^*(r/[2\iota_{C_\theta}^*(r)]) \leq \frac{1}{2} \max(1, \theta(s_o)) - \theta(s_o) + \frac{1}{2} \leq 1$  since  $\theta(s_o) \geq 0$ . We have proved that  $j_{\theta^*}(r) \leq 2\iota_{C_\theta}^*(r)$ .

*Second case:*  $\iota_{C_\theta}^*(r) = 0$ . We have  $\langle r, s \rangle \leq 0$  for all  $s \in C_\theta$ . As  $\text{dom } \theta$  is a subset of the cone generated by  $C_\theta$ , we also have for all  $t > 0$  and  $s \in \text{dom } \theta$ ,  $\langle tr, s \rangle \leq 0$ . Hence  $\langle tr, s \rangle - \theta(s) \leq 0$  for all  $s \in S$  and  $\theta^*(tr) \leq 0$ , for all  $t \geq 0$ . As  $\theta^* \geq 0$ , we have  $\theta^*(tr) = 0$ , for all  $t \geq 0$ . It follows that  $j_{\theta^*}(r) = 0$ . This completes the proof of the equivalence of  $j_{\theta^*}$  and  $\iota_{C_\theta}^*$ .

• Finally, this equivalence implies that  $\text{dom } j_{\theta^*} = \text{dom } \iota_{C_\theta}^*$  and as  $\theta^*(0) = 0$  we have  $0 \in \text{dom } \theta^*$  which implies that  $\text{cone dom } \theta^* = \text{dom } j_{\theta^*}$ .  $\square$

### 4.3. Preliminary technical results

Recall that  $|u|_\Phi = \inf\{\alpha > 0; \Phi_\pm(u/\alpha) \leq 1\}$  with  $\Phi_\pm(u) = \max(\Phi_o(u), \Phi_o(-u))$ . Its associated dual uniform norm is

$$|v|_\Phi^* := \sup_{u, |u|_\Phi \leq 1} |\langle u, v \rangle|, \quad v \in \mathcal{V}.$$

The topological dual space of  $(\mathcal{V}, |\cdot|_\Phi^*)$  is denoted by  $\mathcal{U}''$ : the bidual space of  $(\mathcal{U}, |\cdot|_\Phi)$ . Similarly, recall that  $|y|_\Lambda = \inf\{\alpha > 0; \Lambda_\pm(y/\alpha) \leq 1\}$  with  $\Lambda_\pm(y) = \max(\Lambda_o(y), \Lambda_o(-y))$ . Its associated dual uniform norm is

$$|x|_\Lambda^* := \sup_{y, |y|_\Lambda \leq 1} |\langle y, x \rangle|, \quad x \in \mathcal{X}.$$

The topological dual space of  $(\mathcal{X}, |\cdot|_\Lambda^*)$  is denoted by  $\mathcal{Y}''$ : the bidual space of  $(\mathcal{Y}, |\cdot|_\Lambda)$ .

**Lemma 4.5.** *Let us assume  $(H_\Phi)$  and  $(H_T)$ . Then the following assertions hold true:*

- (a)  $\text{dom } \Phi_o^* \subset \mathcal{V}$  and  $\text{dom } \Lambda_o^* \subset \mathcal{X}$ .
- (b)  $T_o(\text{dom } \Phi_o^*) \subset \text{dom } \Lambda_o^*$  and  $T_o\mathcal{V} \subset \mathcal{X}$ .
- (c)  $T_o$  is  $\sigma(\mathcal{V}_o, \mathcal{U}_o)$ - $\sigma(\mathcal{X}_o, \mathcal{Y}_o)$ -continuous.
- (d)  $T^* : \mathcal{X}^* \rightarrow \mathcal{V}^*$  is  $\sigma(\mathcal{X}^*, \mathcal{X})$ - $\sigma(\mathcal{V}^*, \mathcal{V})$ -continuous.
- (e)  $T : \mathcal{V} \rightarrow \mathcal{X}$  is  $|\cdot|_\Phi^*$ - $|\cdot|_\Lambda^*$ -continuous.
- (f)  $T^*\mathcal{Y}'' \subset \mathcal{U}''$  where  $\mathcal{Y}''$  and  $\mathcal{U}''$  are the topological bidual spaces of  $\mathcal{Y}$  and  $\mathcal{U}$ .
- (g)  $T^*\mathcal{Y} \subset \mathcal{U}$  and  $T^* : \mathcal{Y} \rightarrow \mathcal{U}$  is  $\sigma(\mathcal{Y}, \mathcal{X})$ - $\sigma(\mathcal{U}, \mathcal{V})$ -continuous.
- (h)  $T : \mathcal{V} \rightarrow \mathcal{X}$  is  $\sigma(\mathcal{V}, \mathcal{U})$ - $\sigma(\mathcal{X}, \mathcal{Y})$ -continuous.

**Proof.** • Proof of (a). For all  $v \in \mathcal{V}_o$  and  $\alpha > 0$ , Fenchel's inequality yields:  $\langle u, v \rangle = \alpha \langle v, u/\alpha \rangle \leq [\Phi_o(u/\alpha) + \Phi_o^*(v)]\alpha$ , for all  $u \in \mathcal{U}_o$ . Hence, for any  $\alpha > |u|_\Phi$ ,  $\langle u, v \rangle \leq [1 + \Phi_o^*(v)]\alpha$ . It follows that  $\langle u, v \rangle \leq [1 + \Phi_o^*(v)]|u|_\Phi$ . Considering  $-u$  instead of  $u$ , one gets

$$|\langle u, v \rangle| \leq [1 + \Phi_o^*(v)]|u|_\Phi, \quad \forall u \in \mathcal{U}_o, v \in \mathcal{V}_o. \tag{34}$$

It follows that  $\text{dom } \Phi_o^* \subset \mathcal{V}$ . One proves  $\text{dom } \Lambda_o^* \subset \mathcal{X}$  similarly.

• Proof of (b). It is easy to show that  $\Lambda_o^*(T_o v) \leq \Phi_o^*(v)$ , for all  $v \in \mathcal{V}_o$ . It follows immediately that  $T_o(\text{dom } \Phi_o^*) \subset \text{dom } \Lambda_o^*$ .

Let us consider  $|\cdot|_{\Phi_\pm^*}$  and  $|\cdot|_{\Lambda_\pm^*}$  the gauge functionals of the level sets  $\{\Phi_\pm^* \leq 1\}$  and  $\{\Lambda_\pm^* \leq 1\}$ . As above,

$$\Lambda_\pm^*(T_o v) \leq \Phi_\pm^*(v), \quad \forall v \in \mathcal{V}_o. \tag{35}$$

Therefore,  $T_o(\text{dom } \Phi_\pm^*) \subset \text{dom } \Lambda_\pm^*$ . On the other hand, by Proposition 4.4, the linear space spanned by  $\text{dom } \Phi_\pm^*$  is  $\text{dom } |\cdot|_{\Phi_\pm^*}$  and the linear space spanned by  $\text{dom } \Lambda_\pm^*$  is  $\text{dom } |\cdot|_{\Lambda_\pm^*}$ . But,  $\text{dom } |\cdot|_{\Phi_\pm^*} = \text{dom } |\cdot|_\Phi^* = \mathcal{V}$  and  $\text{dom } |\cdot|_{\Lambda_\pm^*} = \text{dom } |\cdot|_\Lambda^* = \mathcal{X}$  by Proposition 4.4 again. Hence,  $T_o\mathcal{V} \subset \mathcal{X}$ .

• Proof of (c). To prove that  $T_o$  is continuous, one has to show that for any  $y \in \mathcal{Y}_o$ ,  $v \in \mathcal{V}_o \mapsto \langle y, T_o v \rangle \in \mathbb{R}$  is continuous. We get  $v \mapsto \langle y, T_o v \rangle = \langle T_o^* y, v \rangle$  which is continuous since  $(H_{T1})$  gives  $T_o^* y \in \mathcal{U}_o$ .

- Proof of (d). It is a direct consequence of  $T_o\mathcal{V} \subset \mathcal{X}$ . See the proof of (c).
- Proof of (e). We know by Proposition 4.4 that  $|\cdot|_{\Phi_{\pm}^*} \sim |\cdot|_{\Phi}^*$  and  $|\cdot|_{\Lambda_{\pm}^*} \sim |\cdot|_{\Lambda}^*$  are equivalent norms on  $\mathcal{V}$  and  $\mathcal{X}$  respectively. For all  $v \in \mathcal{V}$ ,  $|Tv|_{\Lambda}^* \leq 2|Tv|_{\Lambda_{\pm}^*} = 2 \inf\{\alpha > 0; \Lambda_{\pm}^*(Tv/\alpha) \leq 1\} \leq 2 \inf\{\alpha > 0; \Phi_{\pm}^*(v/\alpha) \leq 1\}$ . This last inequality follows from (35). Going on, we get  $|Tv|_{\Lambda}^* \leq 2|v|_{\Phi_{\pm}^*} \leq 4|v|_{\Phi}^*$ , which proves that  $T$  shares the desired continuity property with  $\|T\| \leq 4$ .
- Proof of (f). Let us take  $\omega \in \mathcal{Y}''$ . For all  $v \in \mathcal{V}$ ,  $|\langle T^*\omega, v \rangle_{\mathcal{V}^*, \mathcal{V}}| = |\langle \omega, Tv \rangle_{\mathcal{Y}'', \mathcal{X}}| \leq \|\omega\|_{\mathcal{Y}''} |Tv|_{\Lambda}^* \leq \|\omega\|_{\mathcal{Y}''} \|T\| |v|_{\Phi}^*$  where  $\|T\| < \infty$ , thanks to (e). Hence,  $T^*\omega$  stands in  $\mathcal{U}''$ .
- Proof of (g). Take  $y \in \mathcal{Y}$ . Let us show that  $T^*y$  is the strong limit of a sequence in  $\mathcal{U}_o$ . Indeed, there exists a sequence  $(y_n)$  in  $\mathcal{Y}_o$  such that  $\lim_{n \rightarrow \infty} y_n = y$  in  $(\mathcal{Y}, |\cdot|_{\Lambda})$ . Hence, for all  $v \in \mathcal{V}$ ,  $|\langle T^*y_n - T^*y, v \rangle_{\mathcal{V}^*, \mathcal{V}}| = |\langle y_n - y, Tv \rangle_{\mathcal{Y}, \mathcal{X}}| \leq \|T\| |y_n - y|_{\Lambda} |v|_{\Phi}^*$  and  $\sup_{v \in \mathcal{V}, |v|_{\Phi}^* \leq 1} |\langle T^*y_n - T^*y, v \rangle| \leq \|T\| |y_n - y|_{\Lambda}$  tends to 0 as  $n$  tends to infinity, where  $T^*y_n$  belongs to  $\mathcal{U}_o$  for all  $n \geq 1$  by  $(H_{T1})$ . Consequently,  $T^*y \in \mathcal{U}$ . The continuity statement now follows from (d).
- Proof of (h). By (b),  $T$  maps  $\mathcal{V}$  into  $\mathcal{X}$  and because of (g):  $T^*\mathcal{Y} \subset \mathcal{U}$ . Hence, for all  $y \in \mathcal{Y}$ ,  $v \mapsto \langle Tv, y \rangle_{\mathcal{X}, \mathcal{Y}} = \langle v, T^*y \rangle_{\mathcal{V}, \mathcal{U}}$  is  $\sigma(\mathcal{V}, \mathcal{U})$ -continuous. This completes the proof of Lemma 4.5. □

Recall that  $\Phi_o^*$ ,  $\Lambda_o^*$  and  $\Lambda^*$  are the convex conjugates of  $\Phi_o$ ,  $\Lambda_o$  and  $\Lambda$  for the dual pairings  $\langle \mathcal{U}_o, \mathcal{V}_o \rangle$ ,  $\langle \mathcal{Y}_o, \mathcal{X}_o \rangle$  and  $\langle \mathcal{Y}, \mathcal{X} \rangle$ .

**Lemma 4.6.** *Under the hypotheses  $(H_{\Phi})$  and  $(H_T)$ , we have*

- (a)  $\Phi_o = \Phi$  on  $\mathcal{U}_o$ ;
- (b)  $\Lambda_o = \Lambda$  on  $\mathcal{Y}_o$ ;
- (c)  $\Phi_o^* = \Phi^*$  on  $\mathcal{V}$ .

**Proof.** (a) follows directly from Lemma 4.5(a) and the assumption that  $\Phi_o$  is lower semicontinuous and convex. (b) follows from (a).

Let us show (c). As  $\mathcal{U}_o$  is a dense subspace of  $\mathcal{U}$ , we obtain that  $\Phi$  is the convex  $\sigma(\mathcal{U}, \mathcal{V})$ -lower semicontinuous regularization of  $\Phi_o + \iota_{\mathcal{U}_o}$  where  $\iota_{\mathcal{U}_o}$  is the convex indicator function of  $\mathcal{U}_o$ . Since the convex conjugate of a function and the convex conjugate of its convex lower semicontinuous regularization match, this implies that  $\Phi_o^* = \Phi^*$  on  $\mathcal{V}$ . □

**Lemma 4.7.** *Under the hypothesis  $(H_{\Phi})$ ,*

- (a)  $\Phi_o^*$  is  $\sigma(\mathcal{V}_o, \mathcal{U}_o)$ -inf-compact and
- (b)  $\Phi^*$  is  $\sigma(\mathcal{V}, \mathcal{U})$ -inf-compact.

**Proof.** • Proof of (b). Recall that we already obtained at (34) that  $|\langle u, v \rangle| \leq [1 + \Phi_o^*(v)]|u|_{\Phi}$ , for all  $u \in \mathcal{U}_o$  and  $v \in \mathcal{V}_o$ . By completion, one deduces that for all  $v \in \mathcal{V}$  and  $u \in \mathcal{U}$ ,  $|\langle u, v \rangle| \leq [1 + \Phi^*(v)]|u|_{\Phi}$  ( $\Phi_o^* = \Phi^*$  on  $\mathcal{V}$ , Lemma 4.6(c).) Hence,  $\Phi^*(v) \leq A$  implies that  $|v|_{\Phi}^* \leq A + 1$ . Therefore, the level set  $\{\Phi^* \leq A\}$  is relatively  $\sigma(\mathcal{V}, \mathcal{U})$ -compact.

By construction,  $\Phi^*$  is  $\sigma(\mathcal{V}, \mathcal{U})$ -lower semicontinuous. Hence,  $\{\Phi^* \leq A\}$  is  $\sigma(\mathcal{V}, \mathcal{U})$ -closed and  $\sigma(\mathcal{V}, \mathcal{U})$ -compact.

• Proof of (a). As  $\Phi_o^* = \Phi^*$  on  $\mathcal{V}$  (Lemma 4.6(c)),  $\text{dom } \Phi_o^* \subset \mathcal{V}$  (Lemma 4.5(a)) and  $\mathcal{U}_o \subset \mathcal{U}$ , it follows from the  $\sigma(\mathcal{V}, \mathcal{U})$ -inf-compactness of  $\Phi^*$  that  $\Phi_o^*$  is  $\sigma(\mathcal{V}_o, \mathcal{U}_o)$ -inf-compact. □

### 5. Proofs of the results of Section 2

The results of Section 2 are a summing up of Proposition 5.2, Lemma 5.3, Proposition 5.4, Corollary 5.5, Lemma 5.6, Proposition 5.7, Corollary 5.8, Proposition 5.10 and Proposition 5.12.

#### 5.1. A first dual equality

In this section we only consider the basic spaces  $\mathcal{U}_o, \mathcal{V}_o, \mathcal{Y}_o$  and  $\mathcal{X}_o$ . Let us begin by recalling Section 4.1 with  $\langle P, A \rangle = \langle \mathcal{U}_o, \mathcal{V}_o \rangle$  and  $\langle B, Q \rangle = \langle \mathcal{Y}_o, \mathcal{X}_o \rangle$  and the topologies are the weak topologies  $\sigma(\mathcal{V}_o, \mathcal{U}_o), \sigma(\mathcal{U}_o, \mathcal{V}_o), \sigma(\mathcal{X}_o, \mathcal{Y}_o)$  and  $\sigma(\mathcal{Y}_o, \mathcal{X}_o)$ . The function to be minimized is  $f(v) = \Phi_o^*(v) + \iota_{C_o}(T_o v)$ ,  $v \in \mathcal{V}_o$ . The perturbation  $F$  of  $f$  is Fenchel's one:

$$F_0(v, x) = \Phi_o^*(v) + \iota_{C_o}(T_o v + x), \quad v \in \mathcal{V}_o, x \in \mathcal{X}_o.$$

We assume  $(H_{T_1})$ :  $T_o^* \mathcal{Y}_o \subset \mathcal{U}_o$ , so that the duality diagram is

$$\begin{array}{ccc} \langle \mathcal{U}_o & , & \mathcal{V}_o \rangle \\ T_o^* \uparrow & & \downarrow T_o \\ \langle \mathcal{Y}_o & , & \mathcal{X}_o \rangle \end{array}$$

Diagram 0.

The analogue of  $F$  for the dual problem is

$$\begin{aligned} G_0(y, u) &:= \inf_{v, x} \{ \langle y, x \rangle - \langle u, v \rangle + F_0(v, x) \} \\ &= \inf_{x \in C_o} \langle y, x \rangle - \Phi_o(T_o^* y + u), \quad y \in \mathcal{Y}_o, u \in \mathcal{U}_o. \end{aligned}$$

The corresponding value functions are

$$\begin{aligned} \varphi_0(x) &= \inf \{ \Phi_o^*(v); v \in \mathcal{V}_o : T_o v \in C_o - x \}, \quad x \in \mathcal{X}_o \\ \gamma_0(u) &= \sup_{y \in \mathcal{Y}_o} \{ \inf_{x \in C_o} \langle y, x \rangle - \Phi_o(T_o^* y + u) \}, \quad u \in \mathcal{U}_o. \end{aligned}$$

The primal and dual problems are  $(P_o)$  and  $(D_o)$ .

**Lemma 5.1.** *Assuming  $(H_\Phi)$  and  $(H_{T_1})$ , if  $C_o$  is a  $\sigma(\mathcal{X}_o, \mathcal{Y}_o)$ -closed convex set,  $F_0$  is jointly lower semicontinuous convex on  $\mathcal{V}_o \times \mathcal{X}_o$ .*

**Proof.** As  $T_o$  is linear continuous (Lemma 4.5(c)) and  $C_o$  is closed convex,  $\{(v, x); T_o v + x \in C_o\}$  is closed convex in  $\mathcal{V}_o \times \mathcal{X}_o$ . As  $\Phi_o^*$  is lower semicontinuous convex on  $\mathcal{V}_o$ , its epigraph is closed convex in  $\mathcal{V}_o \times \mathbb{R}$ . It follows that  $\text{epi } F_0 = (\mathcal{X}_o \times \text{epi } \Phi_o^*) \cap [\{(x, v); T_o v +$

$x \in C_o\} \times \mathbb{R}]$  is closed convex, which implies that  $F_0$  is convex and closed. As it is nowhere equal to  $-\infty$  (since  $\inf F_0 \geq \inf \Phi_o^* > -\infty$ ),  $F_0$  is also a lower semicontinuous convex function.  $\square$

Therefore, assuming that  $C_o$  is a  $\sigma(\mathcal{X}_o, \mathcal{Y}_o)$ -closed convex set, one can apply the general theory of Section 4.1 since the perturbation function  $F_0$  satisfies the assumptions (26) and (28).

**Proposition 5.2.** *Let us assume that  $(H_\Phi)$  and  $(H_T)$  hold. If  $C_o$  is convex and  $\sigma(\mathcal{X}_o, \mathcal{Y}_o)$ -closed, we have the dual equality*

$$\inf(P_o) = \sup(D_o) \in [0, \infty]. \tag{36}$$

*In particular, for all  $x$  in  $\mathcal{X}_o$ , we have the simple dual equality*

$$\inf\{\Phi_o^*(v); v \in \mathcal{V}_o, T_o v = x\} = \Lambda_o^*(x) \in [0, \infty]. \tag{37}$$

**Proof.** The identity (37) is a special case of (36) with  $C_o = \{x\}$ .

To prove (36), we consider separately the cases where  $\inf(P_o) < +\infty$  and  $\inf(P_o) = +\infty$ .

*Case where  $\inf(P_o) < +\infty$ .* Thanks to Theorem 4.1( $b'$ ), it is enough to prove that  $\gamma_0$  is upper semicontinuous at  $u = 0$ . We are going to prove that  $\gamma_0$  is continuous at  $u = 0$ . Indeed, for all  $u \in \mathcal{U}_o$ ,

$$-\gamma_0(u) = \inf_y \{\Phi_o(T_o^* y + u) - \inf_{x \in C_o} \langle y, x \rangle\} \leq \Phi_o(u)$$

where the inequality is obtained by taking  $y = 0$ . The norm  $|\cdot|_\Phi$  is designed so that  $\Phi_o$  is bounded above on a  $|\cdot|_\Phi$ -neighbourhood of zero. By the previous inequality, so is the convex function  $-\gamma_0$ . Therefore,  $-\gamma_0$  is  $|\cdot|_\Phi$ -continuous on  $\text{icordom}(-\gamma_0) \ni 0$ . As it is convex and  $\mathcal{V} = (\mathcal{U}_o, |\cdot|_\Phi)'$ , it is also  $\sigma(\mathcal{U}_o, \mathcal{V})$ -lower semicontinuous and a fortiori  $\sigma(\mathcal{U}_o, \mathcal{V}_o)$ -lower semicontinuous at 0, since  $\mathcal{V} \subset \mathcal{V}_o$ .

*Case where  $\inf(P_o) = +\infty$ .* Note that  $\sup(D_o) \geq -\Phi_o(0) = 0 > -\infty$ , so that we can apply Theorem 4.1( $b$ ). It is enough to prove that

$$\text{ls } \varphi_0(0) = +\infty$$

in the situation where  $\varphi_0(0) = \inf(P_o) = +\infty$ . Recall that  $\text{ls } \varphi_0$  is the lower semicontinuous regularization of  $\varphi_0$ : we have  $\text{ls } \varphi_0(0) = \sup_{W \in \mathcal{N}(0)} \inf\{\Phi_o^*(v); v : T_o v \in C_o + W\}$  where  $\mathcal{N}(0)$  is the set of all the  $\sigma(\mathcal{X}_o, \mathcal{Y}_o)$ -open neighbourhoods of  $0 \in \mathcal{X}_o$ . Now, we prove ad absurdum that  $\text{ls } \varphi_0(0) = +\infty$ . Suppose that  $\text{ls } \varphi_0(0) < +\infty$ . Let  $\epsilon > 0$ . By the above equality of  $\text{ls } \varphi_0(0)$ , for all  $W \in \mathcal{N}(0)$ , there exists  $v \in \mathcal{V}_o$  such that  $T_o v \in C_o + W$  and  $\Phi_o^*(v) \leq \text{ls } \varphi_0(0) + \epsilon$ . This implies that

$$T_o(\{\Phi_o^* \leq \text{ls } \varphi_0(0) + \epsilon\}) \cap (C_o + W) \neq \emptyset, \quad \forall W \in \mathcal{N}(0). \tag{38}$$

On the other hand,  $\inf(P_o) = +\infty$  is equivalent to:  $T_o(\text{dom } \Phi_o^*) \cap C_o = \emptyset$  and we have a fortiori

$$T_o(\{\Phi_o^* \leq \text{ls } \varphi_0(0) + \epsilon\}) \cap C_o = \emptyset.$$

As  $\Phi_o^*$  is inf-compact (Lemma 4.7(a)) and  $T_o$  is continuous (Lemma 4.5(c)),  $T_o(\{\Phi_o^* \leq \text{ls } \varphi_o(0) + \epsilon\})$  is a  $\sigma(\mathcal{X}_o, \mathcal{Y}_o)$ -compact subset of  $\mathcal{X}_o$ . Clearly, it is also convex. But  $C_o$  is assumed to be closed and convex, so that by Hahn-Banach theorem,  $C_o$  and  $T_o(\{\Phi_o^* \leq \text{ls } \varphi_o(0) + \epsilon\})$  are *strictly* separated. This contradicts (38), considering open neighbourhoods  $W$  of the origin in (38) which are open half-spaces. Consequently,  $\text{ls } \varphi_o(0) = +\infty$ . This completes the proof of the proposition.  $\square$

### 5.2. Primal attainment and dual equality

We are going to consider the following duality diagram, see Section 2.3:

$$\begin{array}{ccc} \langle \mathcal{U} & , & \mathcal{V} \rangle \\ T^* \uparrow & & \downarrow T \\ \langle \mathcal{Y} & , & \mathcal{X} \rangle \end{array}$$

Diagram 1.

Note that the inclusions  $T\mathcal{V} \subset \mathcal{X}$  and  $T^*\mathcal{Y} \subset \mathcal{U}$  which are stated in Lemma 4.5 are necessary to validate this diagram.

Let  $F_1, G_1$  and  $\gamma_1$  be the analogues of  $F_0, G_0$  and  $\gamma_0$ . Denoting  $\varphi_1$  the primal value function, we obtain

$$\begin{aligned} F_1(v, x) &= \Phi^*(v) + \iota_C(Tv + x), \quad v \in \mathcal{V}, x \in \mathcal{X} \\ G_1(y, u) &= \inf_{x \in C} \langle y, x \rangle - \Phi(T^*y + u), \quad y \in \mathcal{Y}, u \in \mathcal{U} \\ \varphi_1(x) &= \inf\{\Phi^*(v); v \in \mathcal{V} : Tv \in C - x\}, \quad x \in \mathcal{X} \\ \gamma_1(u) &= \sup_{y \in \mathcal{Y}} \{\inf_{x \in C} \langle y, x \rangle - \Phi(T^*y + u)\}, \quad u \in \mathcal{U}. \end{aligned}$$

It appears that the primal and dual problems are (P) and (D).

**Lemma 5.3.** *Assuming  $(H_\Phi)$  and  $(H_T)$ , the problems  $(P_o)$  and (P) are equivalent: they have the same solutions and  $\inf(P_o) = \inf(P) \in [0, \infty]$ .*

**Proof.** It is a direct consequence of  $\text{dom } \Phi_o^* \subset \mathcal{V}$ ,  $T_o\mathcal{V} \subset \mathcal{X}$  and  $\Phi_o^* = \Phi^*$  on  $\mathcal{V}$ , see Lemma 4.5(a), (b) and Lemma 4.6(c).  $\square$

**Proposition 5.4 (Primal attainment and dual equality).** *Assume that  $(H_\Phi)$  and  $(H_T)$  hold.*

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

$$\inf\{\Phi_o^*(v); v \in \mathcal{V}_o, T_o v = x\} = \Lambda^*(x) \in [0, \infty]. \tag{39}$$

*Assume that in addition  $(H_C)$  holds.*

(b) *We have the dual equalities*

$$\inf(P_o) = \sup(D) \in [0, \infty] \tag{40}$$

$$\inf(P_o) = \inf(P) = \inf_{x \in C} \Lambda^*(x) \in [0, \infty]. \tag{41}$$

- (c) If in addition  $\inf(P_o) < \infty$ , then  $(P_o)$  is attained in  $\mathcal{V}$ .
- (d) Let  $\hat{v} \in \mathcal{V}$  be a solution to  $(P)$ , then  $\hat{x} := T\hat{v}$  is a solution to  $(P^X)$  and  $\inf(P) = \Phi^*(\hat{v}) = \Lambda^*(\hat{x})$ .

**Proof.** • We begin with the proof of (40). As  $\inf(P_o) = \inf(P)$  by Lemma 5.3, we have to show that  $\inf(P) = \sup(D)$ . We consider separately the cases where  $\inf(P) < +\infty$  and  $\inf(P) = +\infty$ .

*Case where  $\inf(P) < +\infty$ .* Because of  $(H_C)$ ,  $F_1$  is jointly convex and  $F_1(v, \cdot)$  is  $\sigma(\mathcal{X}, \mathcal{Y})$ -lower semicontinuous convex for all  $v \in \mathcal{V}$ . As  $T^*\mathcal{Y} \subset \mathcal{U}$  (Lemma 4.5), one can apply the approach of Section 4.1 to the duality Diagram 1. Therefore, by Theorem 4.1(b'), the dual equality holds if  $\gamma_1$  is  $\sigma(\mathcal{U}, \mathcal{V})$ -upper semicontinuous at 0. As in the proof of Proposition 5.2, we have  $-\gamma_1(u) \leq \Phi(u)$ , for all  $u \in \mathcal{U}$ . But  $\Phi$  is the  $\sigma(\mathcal{U}, \mathcal{V})$ -lower semicontinuous regularization of  $\Phi_o + \iota_{\mathcal{U}_o}$  on  $\mathcal{U}$  and  $\Phi_o$  is bounded above by 1 on the ball  $\{u \in \mathcal{U}_o; |u|_\Phi < 1\}$ . As  $\mathcal{V} = (\mathcal{U}, |\cdot|_\Phi)'$ ,  $\Phi$  is also the  $|\cdot|_\Phi$ -regularization of  $\Phi_o + \iota_{\mathcal{U}_o}$ . Therefore,  $\Phi$  is bounded above by 1 on  $\{u \in \mathcal{U}; |u|_\Phi < 1\}$ , since  $\{u \in \mathcal{U}_o; |u|_\Phi < 1\}$  is  $|\cdot|_\Phi$ -dense in  $\{u \in \mathcal{U}; |u|_\Phi < 1\}$ . As  $-\gamma_1(\leq \Phi)$  is convex and bounded above on a  $|\cdot|_\Phi$ -neighbourhood of 0, it is  $|\cdot|_\Phi$ -continuous on  $\text{icordom}(-\gamma_1) \ni 0$ . Hence, it is  $\sigma(\mathcal{U}, \mathcal{V})$ -lower semicontinuous at 0.

*Case where  $\inf(P) = +\infty$ .* This proof is a transcription of the second part of the proof of Proposition 5.2, replacing  $T_o$  by  $T$ ,  $C_o$  by  $C$ , all the subscripts 0 by 1 and using the preliminary results:  $\Phi^*$  is inf-compact (Lemma 4.7) and  $T$  is weakly continuous (Lemma 4.5(h)). This completes the proof of (40).

- The identity (39) is simply (40) with  $C = \{x\}$ .
- Let us prove (c). By Lemma 4.5(h),  $T$  is  $\sigma(\mathcal{V}, \mathcal{U})$ - $\sigma(\mathcal{X}, \mathcal{Y})$ -continuous. Since  $C$  is  $\sigma(\mathcal{X}, \mathcal{Y})$ -closed,  $\{v \in \mathcal{V}; Tv \in C\}$  is  $\sigma(\mathcal{V}, \mathcal{U})$ -closed. As  $\Phi^*$  is  $\sigma(\mathcal{V}, \mathcal{U})$ -inf-compact (Lemma 4.7), it achieves its infimum on the closed set  $\{v \in \mathcal{V}; Tv \in C\}$  if  $\inf(P) = \inf(P_o) < \infty$ .
- Let us prove (41). The dual equality (40) gives us for all  $x_o \in C$ ,  $\inf(P) = \sup_{y \in \mathcal{Y}} \{\inf_{x \in C} \langle y, x \rangle - \Lambda(y)\} \leq \sup_{y \in \mathcal{Y}} \{\langle x_o, y \rangle - \Lambda(y)\} = \Lambda^*(x_o)$ . Therefore

$$\inf(P) \leq \inf_{x \in C} \Lambda^*(x). \tag{42}$$

In particular, equality holds instead of inequality if  $\inf(P) = +\infty$ . Suppose now that  $\inf(P) < \infty$ . From statement (c), we already know that there exists  $\hat{v} \in \mathcal{V}$  such that  $\hat{x} := T\hat{v} \in C$  and  $\inf(P) = \Phi^*(\hat{v})$ . Clearly  $\inf(P) \leq \inf\{\Phi^*(v); Tv = \hat{x}, v \in \mathcal{V}\} \leq \Phi^*(\hat{v})$ . Hence,  $\inf(P) = \inf\{\Phi^*(v); Tv = \hat{x}, v \in \mathcal{V}\}$ . By the simple dual equality (39) we have  $\inf\{\Phi^*(v); Tv = \hat{x}, v \in \mathcal{V}\} = \Lambda^*(\hat{x})$ . Finally, we have obtained  $\inf(P) = \Lambda^*(\hat{x})$  with  $\hat{x} \in C$ . Together with (42), this leads us to the desired identity:  $\inf(P) = \inf_{x \in C} \Lambda^*(x)$ .

- Finally, (d) is a by-product of the proof of (41). □

**Corollary 5.5.** We have  $\text{dom } \Lambda^* \subset \text{dom } \Lambda_o^*$ ,  $\text{dom } \Lambda^* \subset \mathcal{X}$  and in restriction to  $\mathcal{X}$ ,  $\Lambda_o^* = \Lambda^*$ .

**Proof.** The first part is already proved in Lemma 4.5(a). The matching  $\Lambda_o^* = \Lambda^*$  follows from (37) and (39). □

**Lemma 5.6.** *Under the hypotheses  $(H_\Phi)$  and  $(H_T)$ ,  $\Lambda^*$  is  $\sigma(\mathcal{X}, \mathcal{Y})$ -inf-compact.*

**Proof.** By (39):  $\inf\{\Phi^*(v); v \in \mathcal{V}, Tv = x\} = \Lambda^*(x)$  for all  $x \in \mathcal{X}$  (note that  $\Phi_o^* = \Phi^*$  on  $\mathcal{V}$  by Lemma 4.6(c).) As  $T$  is continuous (Lemma 4.5(h)) and  $\Phi^*$  is inf-compact (Lemma 4.7), it follows that  $\Lambda^*$  is also inf-compact.  $\square$

### 5.3. Dual attainment

We now consider the following duality diagram

$$\begin{array}{ccc} \langle \mathcal{V} & , & \mathcal{V}^* \rangle \\ T \downarrow & & \uparrow T^* \\ \langle \mathcal{X} & , & \mathcal{X}^* \rangle \end{array}$$

Diagram 2.

where the topologies are the respective weak topologies. The associated perturbation functions are

$$\begin{aligned} F_2(v, x) &= \Phi^*(v) + \iota_C(Tv + x), \quad v \in \mathcal{V}, x \in \mathcal{X}, \\ G_2(\zeta, \omega) &= \inf_{x \in C} \langle x, \omega \rangle - \bar{\Phi}(T^*\omega + \zeta), \quad \zeta \in \mathcal{V}^*, \omega \in \mathcal{X}^* \end{aligned}$$

As  $F_2 = F_1$ , the primal problem is  $(P)$  and its value function is  $\varphi_1$ :

$$\varphi_1(x) = \inf_{x' \in C-x} \Lambda^*(x'), \quad x \in \mathcal{X} \tag{43}$$

where we used (39). The dual problem is  $(\bar{D})$ .

**Proposition 5.7 (Dual attainment).** *Assume that  $(H_\Phi)$ ,  $(H_T)$  and  $(H_C)$  hold. Suppose that*

$$C_o \cap \text{icordom } \Lambda^* \neq \emptyset. \tag{44}$$

*Then the dual problem  $(\bar{D})$  is attained in  $\mathcal{X}^*$ .*

**Proof.** As  $F_2 = F_1$ , one can apply the approach of Section 4.1 to the duality Diagram 2. Let us denote  $\varphi_1^{**}$  the  $\sigma(\mathcal{X}, \mathcal{Y})$ -lower semicontinuous regularization of  $\varphi_1$  and  $\varphi_2^{**}$  its  $\sigma(\mathcal{X}, \mathcal{X}^*)$ -lower semicontinuous regularization. Since  $\mathcal{X}$  separates  $\mathcal{Y}$ , the inclusion  $\mathcal{Y} \subset \mathcal{X}^*$  holds. It follows that  $\varphi_1^{**}(0) \leq \varphi_2^{**}(0) \leq \varphi_1(0)$ . But we have (40) which is  $\varphi_1^{**}(0) = \varphi_1(0)$ . Therefore, one also obtains  $\varphi_2^{**}(0) = \varphi_1(0)$  which is the dual equality

$$\inf(P) = \sup(\bar{D}) \tag{45}$$

and one can apply Theorem 4.1(c) which gives

$$\text{argmax}(\bar{D}) = -\partial\varphi_1(0). \tag{46}$$

It remains to show that the value function  $\varphi_1$  given at (43) is such that

$$\partial\varphi_1(0) \neq \emptyset. \tag{47}$$



As the considered dual pairing  $\langle \mathcal{X}, \mathcal{X}^* \rangle$  is the saturated algebraic pairing, for (47) to be satisfied, by the geometric version of Hahn-Banach theorem, it is enough that  $0 \in \text{icordom } \varphi_1$ . But this holds provided that the constraint qualification (44) is satisfied.  $\square$

Supposing that  $\inf(P_o) < \infty$  one knows by Proposition 5.4(d) that  $(P^{\mathcal{X}})$  admits at least a solution  $\hat{x} = T\hat{v}$  where  $\hat{v}$  is a solution to  $(P)$ . Let us consider the following new minimization problem

$$\text{minimize } \Phi^*(v) \text{ subject to } Tv = \hat{x}, \quad v \in \mathcal{V}. \tag{P^{\hat{x}}}$$

Of course  $\hat{v}$  is a solution to  $(P)$  if and only if it is a solution to  $(P^{\hat{x}})$  where  $\hat{x} = T\hat{v}$ . Since our aim is to derive a representation formula for  $\hat{v}$ , it is enough to build our duality schema upon  $(P^{\hat{x}})$  rather than upon  $(P)$ . The associated perturbation functions are

$$\begin{aligned} F_2^{\hat{x}}(v, x) &= \Phi^*(v) + \iota_{\{\hat{x}\}}(Tv + x), \quad v \in \mathcal{V}, \quad x \in \mathcal{X} \\ G_2^{\hat{x}}(\zeta, \omega) &= \langle \hat{x}, \omega \rangle - \bar{\Phi}(T^*\omega + \zeta), \quad \zeta \in \mathcal{V}^*, \quad \omega \in \mathcal{X}^*. \end{aligned}$$

As  $F_2^{\hat{x}}$  is  $F_1$  with  $C = \{\hat{x}\}$ , the primal problem is  $(P^{\hat{x}})$  and its value function is

$$\varphi_1^{\hat{x}}(x) = \Lambda^*(\hat{x} - x), \quad x \in \mathcal{X}.$$

The dual problem is

$$\text{maximize } \langle \hat{x}, \omega \rangle - \bar{\Lambda}(\omega), \quad \omega \in \mathcal{X}^*. \tag{D^{\hat{x}}}$$

**Corollary 5.8 (Dual attainment).** *Assume that  $(H_{\Phi})$  and  $(H_T)$  hold. Suppose that  $C_o \cap \text{dom } \Lambda^* \neq \emptyset$ . Then,  $\inf(P_o) < \infty$  and we know (see Proposition 5.4(d)) that  $(P^{\mathcal{X}})$  admits at least a solution. If in addition, there exists a solution  $\hat{x}$  to  $(P^{\mathcal{X}})$  such that*

$$\hat{x} \in \text{dom } \partial_{\mathcal{X}^*} \Lambda^*, \tag{48}$$

then the dual problem  $(D^{\hat{x}})$  is attained in  $\mathcal{X}^*$ .

**Proof.** Let us specialize Proposition 5.7(d) to the special case where  $C = \{\hat{x}\}$ . The dual equality (45) becomes

$$\inf(P^{\hat{x}}) = \sup(D^{\hat{x}}) \tag{49}$$

and (47) becomes  $\partial\varphi_1^{\hat{x}}(0) \neq \emptyset$  which is implied by (48).  $\square$

**Remark 5.9.** Let us denote the extended-real-valued functions on  $\mathcal{X}^*$

$$\begin{aligned} \Lambda_1 &:= \Lambda + \iota_{\mathcal{Y}}, \\ \Lambda_2 &:= \bar{\Lambda}. \end{aligned}$$

We also denote  $\Lambda_1^*, \Lambda_2^*$  their convex conjugates with respect to  $\langle \mathcal{X}, \mathcal{X}^* \rangle$  and  $\tilde{\Lambda}_1, \tilde{\Lambda}_2$  their convex  $\sigma(\mathcal{X}^*, \mathcal{X})$ -lower semicontinuous regularizations. Clearly,

$$\Lambda_1^* = \Lambda^*$$

and the dual equality (49) is

$$\Lambda_1^* = \Lambda_2^* \tag{50}$$

which implies the identity

$$\tilde{\Lambda}_1 = \tilde{\Lambda}_2. \tag{51}$$

Usual results about convex conjugation tell us that  $\Lambda_1^*(\hat{x}) = \sup_{\omega \in \mathcal{X}^*} \{\langle \hat{x}, \omega \rangle - \tilde{\Lambda}_1(\omega)\} = \sup(\overline{D}^{\hat{x}})$  and the above supremum is attained at  $\bar{\omega}$  if and only if  $\bar{\omega} \in \partial_{\mathcal{X}^*} \Lambda^*(\hat{x})$ . This is the attainment statement in Corollary 5.8.

### 5.4. Dual representation of the minimizers

We keep the framework of Diagram 2 and derive the KKT relations in this situation. The Lagrangian associated with  $F_2 = F_1$  and Diagram 2 is for any  $v \in \mathcal{V}, \omega \in \mathcal{X}^*$ ,

$$\begin{aligned} K_2(v, \omega) &:= \inf_{x \in \mathcal{X}} \{\langle x, \omega \rangle + \Phi^*(v) + \iota_C(Tv + x)\}, \\ &= \Phi^*(v) - \langle Tv, \omega \rangle + \inf_{x \in C} \langle x, \omega \rangle. \end{aligned}$$

**Proposition 5.10 (Dual representation).** *Assume that  $(H_\Phi)$ ,  $(H_T)$  and  $(H_C)$  hold. For any  $\hat{v} \in \mathcal{V}$  and  $\bar{\omega} \in \mathcal{X}^*$ ,*

$$\begin{cases} (a) & T\hat{v} \in C_o \\ (b) & \langle \bar{\omega}, T\hat{v} \rangle \leq \langle \bar{\omega}, x \rangle \text{ for all } x \in C \\ (c) & \hat{v} \in \partial_{\mathcal{V}} \overline{\Phi}(T^*\bar{\omega}) \end{cases} \tag{52}$$

is equivalent to

$$\begin{cases} \hat{v} \text{ is a solution to } (P_o), \\ \bar{\omega} \text{ is a solution to } (\overline{D}) \text{ and} \\ \text{the dual equality (40) holds.} \end{cases} \tag{53}$$

It is well-known that the representation formula (52(c)):

$$\hat{v} \in \partial_{\mathcal{V}} \overline{\Phi}(T^*\bar{\omega}) \tag{54}$$

is equivalent to

$$T^*\bar{\omega} \in \partial_{\mathcal{V}^*} \Phi^*(\hat{v})$$

and also equivalent to Fenchel's identity

$$\Phi^*(\hat{v}) + \overline{\Phi}(T^*\bar{\omega}) = \langle \bar{\omega}, T\hat{v} \rangle. \tag{55}$$

**Proof.** This proof is an application of Theorem 4.3. Under the general assumptions  $(H_\Phi)$ ,  $(H_T)$  and  $(H_C)$ , we have seen in Proposition 5.7 that the dual equalities (45) and (49) hold true. Hence, (53) is equivalent to  $(\hat{v}, \bar{\omega})$  is a saddle-point. All we have to do now is to show that (52) is a translation of the KKT relations (32) and (33).

With  $K_2$  as above, (32) and (33) are  $\partial_v K_2(\hat{v}, \bar{\omega}) \ni 0$  and  $\partial_\omega (-K_2)(\hat{v}, \bar{\omega}) \ni 0$ . Since  $-\langle Tv, \omega \rangle$  is locally weakly upper bounded as a function of  $\omega$  around  $\bar{\omega}$  and as a function of  $v$  around  $\hat{v}$ , one can apply (Rockafellar, [13], Theorem 20) to derive  $\partial_v K_2(\hat{v}, \bar{\omega}) = \partial \Phi^*(\hat{v}) - T^*\bar{\omega}$  and  $\partial_\omega (-K_2)(\hat{v}, \bar{\omega}) = \partial(-\inf_{x \in C} \langle x, \cdot \rangle) + T\hat{v}$ . Therefore the KKT relations are

$$T^*\bar{\omega} \in \partial \Phi^*(\hat{v}) \tag{56}$$

$$-T\hat{v} \in \partial(\iota_{-C}^*)(\bar{\omega}) \tag{57}$$

where  $\iota_{-C}^*$  is the convex conjugate of the convex indicator function of  $-C$ . As a convex conjugate,  $\Phi^*$  is a lower semicontinuous convex functions. Its convex conjugate is  $\bar{\Phi}$ . Therefore (56) is equivalent to the following equivalent statements

$$\begin{aligned} \hat{v} &\in \partial \bar{\Phi}(T^*\bar{\omega}) \\ \Phi^*(\hat{v}) + \bar{\Phi}(T^*\bar{\omega}) &= \langle \hat{v}, T^*\bar{\omega} \rangle. \end{aligned}$$

Similarly, as a convex conjugate  $\iota_{-C}^*$  is a lower semicontinuous convex function. Its convex conjugate is  $\iota_{-\bar{C}}$  where  $\bar{C}$  stands for the  $\sigma(\mathcal{X}, \mathcal{X}^*)$ -closure of  $C$ . Of course, as  $C$  is  $\sigma(\mathcal{X}, \mathcal{Y})$ -closed by hypothesis  $(H_C)$ , it is a fortiori  $\sigma(\mathcal{X}, \mathcal{X}^*)$ -closed, so that  $\bar{C} = C$ . Therefore (57) is equivalent to

$$\iota_C(T\hat{v}) + \iota_{-C}^*(\bar{\omega}) = \langle -T\hat{v}, \bar{\omega} \rangle. \tag{58}$$

It follows from (58) that  $\iota_C(T\hat{v}) < \infty$  which is equivalent to  $T\hat{v} \in C$ . Now (58) is  $-\langle T\hat{v}, \bar{\omega} \rangle = \iota_{-C}^*(\bar{\omega}) = -\inf_{x \in C} \langle x, \bar{\omega} \rangle$  which is  $\langle T\hat{v}, \bar{\omega} \rangle = \inf_{x \in C} \langle x, \bar{\omega} \rangle$ . This completes the proof.  $\square$

**Remark 5.11.** Thanks to Proposition 5.4(d), (55) leads us to

$$\Lambda^*(\hat{x}) + \bar{\Lambda}(\bar{\omega}) = \langle \hat{x}, \bar{\omega} \rangle \tag{59}$$

for all  $\hat{x} \in \text{dom } \Lambda^*$  and all  $\bar{\omega} \in \mathcal{X}^*$  solution to  $(\bar{D}^{\hat{x}})$ . By Fenchel's inequality:  $\Lambda_2^*(\hat{x}) + \tilde{\Lambda}_2(\bar{\omega}) \geq \langle \hat{x}, \bar{\omega} \rangle$  and the identities (50) and (59), we see that  $\tilde{\Lambda}_2(\bar{\omega}) \geq \bar{\Lambda}(\bar{\omega})$ . But, the reversed inequality always holds true. Therefore, we have  $\tilde{\Lambda}_2(\bar{\omega}) = \bar{\Lambda}(\bar{\omega})$ . This proves that  $\bar{\Lambda} = \tilde{\Lambda}_2$  on  $\text{dom } \bar{\Lambda} : \bar{\Lambda}$  is  $\sigma(\mathcal{X}^*, \mathcal{X})$ -lower semicontinuous on its effective domain.

**Proposition 5.12.** Assume that  $(H_\Phi)$ ,  $(H_T)$  and  $(H_C)$  hold. Any solution  $\bar{\omega}$  of  $(\bar{D})$  or  $(\bar{D}^{\hat{x}})$  shares the following properties

- (a)  $\bar{\omega}$  is in the  $\sigma(\mathcal{X}^*, \mathcal{X})$ -closure of  $\text{dom } \Lambda$ ;
- (b)  $T^*\bar{\omega}$  is in the  $\sigma(\mathcal{V}^*, \mathcal{V})$ -closures of  $T^*(\text{dom } \Lambda)$  and  $\text{dom } \Phi_o$ .

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

- (a')  $\bar{\omega}$  is in  $\mathcal{Y}''$ . More precisely, it is in the  $\sigma(\mathcal{Y}'', \mathcal{X})$ -closure of  $\text{dom } \Lambda$ ;
- (b')  $T^*\bar{\omega}$  is in  $\mathcal{U}''$ . More precisely, it is in the  $\sigma(\mathcal{U}'', \mathcal{V})$ -closures of  $T^*(\text{dom } \Lambda)$  and  $\text{dom } \Phi_o$

where  $\mathcal{Y}''$  and  $\mathcal{U}''$  are the topological bidual spaces of  $\mathcal{Y}$  and  $\mathcal{U}$ . This occurs if  $\Phi$ , and therefore  $\Phi^*$ , is an even function.

**Proof.** • Proof of (a). Because of (59), we have  $\bar{\omega} \in \text{dom } \bar{\Lambda}$ . As  $\tilde{\Lambda}_2 \leq \bar{\Lambda}$  and  $\tilde{\Lambda}_1 = \tilde{\Lambda}_2$  (see (51)), we obtain  $\bar{\omega} \in \text{dom } \tilde{\Lambda}_1$  which implies that  $\bar{\omega}$  is in the  $\sigma(\mathcal{X}^*, \mathcal{X})$ -closure of  $\text{dom } \Lambda$ .

• Proof of (b). By Lemma 4.5(d),  $T^*$  is continuous from  $\mathcal{X}^*$  to  $\mathcal{V}^*$ . It follows from (a) that  $T^*\bar{\omega}$  is in the  $\sigma(\mathcal{V}^*, \mathcal{V})$ -closure of  $T^*(\text{dom } \Lambda)$ .

On the other hand,  $T^*\bar{\omega} \in \text{dom } \bar{\Phi}$  and  $\bar{\Phi}$  is the  $\sigma(\mathcal{V}^*, \mathcal{V})$ -lower semicontinuous convex regularization of  $\Phi_o$ . It follows that  $T^*\bar{\omega}$  is in the  $\sigma(\mathcal{V}^*, \mathcal{V})$ -closure of  $\text{dom } \Phi_o$ .

- Proof of (a'). Because of (a),  $\bar{\omega}$  is the  $\sigma(\mathcal{X}^*, \mathcal{X})$ -limit of a generalized sequence  $\{y_\alpha\}$  in  $\text{dom } \Lambda$ . Our additional assumption allows us to take  $\{y_\alpha\}$  in a  $|\cdot|_\Phi$ -ball: it is an equicontinuous set. It follows from [7, Cor. of Prop. III.5] that  $\bar{\omega}$  is continuous on  $\mathcal{X}$ .
- Proof of (b'). Similar to (b)'s proof using (a') and Lemma 4.5(f). □

## References

- [1] J. M. Borwein, A. S. Lewis: Convergence of best entropy estimates, *SIAM J. Optim.* 1 (1991) 191–205.
- [2] J. M. Borwein, A. S. Lewis: Duality relationships for entropy-like minimization problems, *SIAM J. Control Optimization* 29 (1991) 325–338.
- [3] J. M. Borwein, A. S. Lewis: On the convergence of moment problems, *Trans. Amer. Math. Soc.* 325 (1991) 249–271.
- [4] J. M. Borwein, A. S. Lewis: Decomposition of multivariate functions, *Can. J. Math.* 44(3) (1992) 463–482.
- [5] J. M. Borwein, A. S. Lewis: Partially-finite programming in  $l_1$  and the existence of the maximum entropy estimates, *SIAM J. Optim.* 3 (1993) 248–267.
- [6] J. M. Borwein, A. S. Lewis, R. D. Nussbaum: Entropy minimization, *DAD* problems and doubly stochastic kernels, *J. Funct. Anal.* 123 (1994) 264–307.
- [7] N. Bourbaki: *Espaces Vectoriels Topologiques*. Chapitres 1 à 5, Masson, Paris (1981).
- [8] C. Combari, M. Laghdir, L. Thibault: A note on subdifferentials of convex composite functionals, *Arch. Math.* 67(3) (1996) 239–252.
- [9] C. Combari, M. Laghdir, L. Thibault: On subdifferential calculus for convex functions defined on locally convex spaces, *Ann. Sci. Math. Qué.* 23(1) (1999) 23–36.
- [10] C. Léonard: A saddle-point approach to the Monge-Kantorovich optimal transport problem, preprint.
- [11] C. Léonard: Minimization of entropy functionals, *J. Math. Anal. Appl.* 346 (2008) 183–204.
- [12] R. T. Rockafellar: Integrals which are convex functionals, *Pacific J. Math.* 24(3) (1968) 525–539.
- [13] R. T. Rockafellar: *Conjugate Duality and Optimization*, Regional Conference Series in Applied Mathematics 16, SIAM, Philadelphia (1974).
- [14] M. Teboulle, I. Vajda: Convergence of best  $\phi$ -entropy estimates, *IEEE Trans. Inf. Theory* 39 (1993) 297–301.
- [15] C. Villani: *Topics in Optimal Transportation*, Graduate Studies in Mathematics 58, American Mathematical Society, Providence (2003).
- [16] C. Villani: *Optimal Transport. Old and New*, Grundlehren der mathematischen Wissenschaften 338, Springer, Berlin (2009).