J Glob Optim (2008) 42:401–412
DOI 10.1007/s10898-008-9301-2

Separation of sets and Wolfe duality

F. Giannessi · G. Mastroeni

Received: 10 March 2008 / Accepted: 13 March 2008 / Published online: 20 June 2008

Abstract Lagrangian duality can be derived from separation in the Image Space, namely the space where the images of the objective and constraining functions of the given extremum problem run. By exploiting such a result, we analyse the relationships between Wolfe and Mond-Weir duality and prove their equivalence in the Image Space under suitable generalized convexity assumptions.

Keywords Constrained optimization · Duality · Image space analysis · Separation

AMS Classifications 65K · 90C

1 Introduction

The image space analysis can be applied to study any kind of problem, say P, that can be expressed under the form of the impossibility of a parametric system [3,5].

The impossibility of such a system is reduced to the disjunction of two suitable subsets $\mathcal{K}$ and $\mathcal{H}$ of the Image Space (IS): $\mathcal{K}$ is defined by the images of the functions involved in P, while $\mathcal{H}$ is a convex cone that depends on the type of constraints which define P.

The disjunction between $\mathcal{K}$ and $\mathcal{H}$ is proved by showing that they lie in two disjoint level sets of a separating functional. When such a functional can be found linear, then we say that $\mathcal{K}$ and $\mathcal{H}$ admit linear separation. A suitable subclass of separating functionals is said to be regular, iff the separation obtained by means of a functional of such a subclass guarantees the disjunction between $\mathcal{K}$ and $\mathcal{H}$.

Exploiting separation arguments in the IS, several theoretical aspects can be developed, as duality, Lagrangian-type optimality conditions, regularity, and penalization.

F. Giannessi · G. Mastroeni (✉)
Department of Mathematics, University of Pisa, L.go B. Pontecorvo, 5, 56127 Pisa, Italy
e-mail: mastroen@dm.unipi.it

F. Giannessi
e-mail: gianness@dm.unipi.it

Duality arises from the existence of a linear separation between $\mathcal{K}$ and $\mathcal{H}$, while necessary optimality conditions arise from the linear separation between a convex conical approximation of $\mathcal{K}$ and $\mathcal{H}$.

Penalization turns out to be closely related to nonlinear separation [2,3].

The existence of a regular separating functional characterizes *strong duality* of P. In the paper, we focus our attention on Wolfe duality [10] and we show that the existence of a regular linear separation guarantees the equivalence between Wolfe and Mond-Weir duality [6].

Let us recall the main notations and definitions that will be used in the sequel.

$\mathbb{R}^n_+ := \{x \in \mathbb{R}^n : x \geq 0\}$. A set $\mathcal{H} \subseteq \mathbb{R}^n$ is said to be a cone iff $\lambda\mathcal{H} \subseteq \mathcal{H}$, with $\lambda \in \mathbb{R}_+ \backslash\{0\}$, and a convex cone iff, in addition, $\mathcal{H} + \mathcal{H} \subseteq \mathcal{H}$, where $\mathcal{H} + \mathcal{H} = \{h_1 + h_2 \in \mathbb{R}^n : h_1 \in \mathcal{H}, h_2 \in \mathcal{H}\}$. $cone\,\mathcal{H} = \{y \in \mathbb{R}^n : \; y = \lambda x, \quad \lambda > 0, \quad x \in \mathcal{H}\}$. The closure, the interior,the relative interior, the boundary, the relative boundary and the convex hull of a set $\mathcal{H}$ are denoted by $cl\mathcal{H}$, $int\mathcal{H}$, ri $\mathcal{H}$, bd $\mathcal{H}$, rbd $\mathcal{H}$, and conv $\mathcal{H}$, respectively. aff $\mathcal{H}$ and lin $\mathcal{H}$ will denote the smallest affine variety and the smallest subspace that contain $\mathcal{H}$, respectively. $\mathcal{H}^* := \{y \in \mathbb{R}^n : \langle x, y\rangle \geq 0, \quad \forall x \in \mathcal{H}\}$ is the positive polar of $\mathcal{H}$, where $\langle\cdot,\cdot\rangle$ is the scalar product in $\mathbb{R}^n$. $TC(\mathcal{H})$ denotes the (Bouligand) tangent cone to $\mathcal{H}$ at $0 \in \mathbb{R}^n$. Let $a, b \in \mathbb{R}^m$, $C \subset \mathbb{R}^m$ be a closed and convex cone; $a \geq_C b$ iff $a - b \in C$.

Let $\mathcal{X}$ be a convex subset of $\mathbb{R}^n$ and $f : \mathcal{X} \longrightarrow \mathbb{R}^m$. f is called $C - function$ on $\mathcal{X}$, iff $\forall x_1, x_2 \in \mathcal{X}$, we have

$$(1-\alpha)f(x_1) + \alpha f(x_2) - f((1-\alpha)x_1 + \alpha x_2)) \in C, \quad \forall\alpha \in [0, 1].$$

f is called $C - convexlike$ on $\mathcal{X}$, iff $\forall x_1, x_2 \in \mathcal{X}, \; \forall\alpha \in [0, 1]$,

$$\exists\hat{x} \in \mathcal{X} s.t. (1-\alpha)f(x_1) + \alpha f(x_2) - f(\hat{x}) \in C.$$

f is called $C - preconvexlike$ on $\mathcal{X}$, iff $\forall x_1, x_2 \in \mathcal{X}, \; \forall\alpha \in [0, 1]$,

$$\exists\rho > 0, \exists\hat{x} \in \mathcal{X} \; s.t. (1-\alpha)f(x_1) + \alpha f(x_2) - \rho f(\hat{x}) \in C.$$

2 Preliminaries on image space analysis

Let us consider a constrained extremum problem in the following format:

$$f^{\downarrow} := \min \; f(x), \qquad x \in R := \{x \in X : g(x) \in D\} \tag{1}$$

where $f : X \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$, $g : X \longrightarrow \mathbb{R}^m$, $D := 0_p \times \mathbb{R}^{m-p}_+$, so that p and $m - p$ (with $m \geq p$) are the numbers of constraining equalities and inequalities, respectively, and where $O_p := (0, \ldots, 0) \in \mathbb{R}^p$. We stipulate that $D = \mathbb{R}^m_+$ if $p = 0$ and $D = 0_m$ if $p = m$; $R = X$ if $m = 0$ so that D is not defined. What follows requires only formal changes for the case where X is a subset of a Banach space and f and g have finite dimensional images.

The dual problem of (1) is defined by:

$$\sup_{\lambda\in D^*} \inf_{x\in X} \; L(x;\lambda), \tag{2}$$

where

$$L(x;\lambda) := f(x) - \langle\lambda, g(x)\rangle$$

is the Lagrangian function associated with (1). By symmetry, (1) is associated with the problem

$$\inf_{x\in X} \sup_{\lambda\in D^*} \; L(x;\lambda), \tag{3}$$

which, under the assumption $R \neq \emptyset$, turns out to be equivalent to (1). (2) and (3) fulfil the inequality:

$$\sup_{\lambda \in D^*} \inf_{x \in X} L(x; \lambda) \leq \inf_{x \in X} \sup_{\lambda \in D^*} L(x; \lambda), \tag{4}$$

so that *weak duality* holds, whatever (1) may be. The difference between the right-hand side and left-hand side of (4) is called *duality gap*.

We remark that $\bar{x} \in R$ is a (global) minimum point of (1), iff the system (in the unknown x):

$$f(\bar{x}) - f(x) > 0, \quad g(x) \geq_D 0, \quad x \in X \tag{5}$$

is impossible. Let us introduce the sets:

$$\mathcal{K}_{\bar{x}} := \{(u, v) \in \mathbb{R} \times \mathbb{R}^m : u = f(\bar{x}) - f(x), \ v = g(x), \ x \in X\}.$$

$$\mathcal{H} := \{(u, v) \in \mathbb{R} \times \mathbb{R}^m : u > 0, \ v \geq_D 0\}$$

$$\mathcal{H}_u := \{(u, v) \in \mathbb{R} \times \mathbb{R}^m : u > 0, \ v = 0\}$$

$\mathcal{K}_{\bar{x}}$ is called the *image* associated with (1).

We observe that the impossibility of system (5) is equivalent to the condition

$$\mathcal{K}_{\bar{x}} \cap \mathcal{H} = \emptyset. \tag{6}$$

so that $\bar{x} \in R$ is a minimum point of (1) iff (6) holds.

To prove directly whether $\mathcal{K}_{\bar{x}} \cap \mathcal{H} = \emptyset$ or not is generally too difficult. Therefore, in order to show such a disjunction, it will be proved that the two sets, or the set $\mathcal{H}$ and an extension of the image depending on $\mathcal{H}$, lie in two disjoint level sets of a functional; when such a functional can be found linear, then $\mathcal{K}_{\bar{x}}$ and $\mathcal{H}$ will be called "linearly separable". A linear separation is said to be "proper" iff the two sets are not both contained in the separating hyperplane.

In general, the image of a generalized system is not convex even when the functions involved enjoy some convexity properties. To overcome this difficulty, we introduce a regularization of the image $\mathcal{K}_{\bar{x}}$, namely the extension with respect to the cone $cl\mathcal{H}$, denoted by $\mathcal{E}$:

$$\mathcal{E} := \mathcal{K}_{\bar{x}} - cl\mathcal{H} = \{(u, v) \in \mathbb{R} \times \mathbb{R}^m : u \leq f(\bar{x}) - f(x), v \leq_D g(x), x \in X\}$$

The importance of the conic extension is enforced by the following statement:

Proposition 2.1 *Condition (6) holds iff*

$$\mathcal{E} \cap \mathcal{H} = \emptyset. \tag{7}$$

or, equivalently,

$$\mathcal{E} \cap \mathcal{H}_u = \emptyset. \tag{8}$$

Proof In order to prove (7) it is enough to note that

$$\mathcal{H} + cl\mathcal{H} = \mathcal{H}. \tag{9}$$

We show that (7) and (8) are equivalent.

Obviously, (7) implies (8), since $\mathcal{H}_u \subseteq \mathcal{H}$. To prove the reverse implication, ab absurdo, suppose that (8) holds and there exists $(\bar{u}, \bar{v}) \in \mathcal{E} \cap \mathcal{H}$. Since

$$\mathcal{E} - cl\mathcal{H} = \mathcal{K}_{\bar{x}} - (cl\mathcal{H} + cl\mathcal{H}) = \mathcal{E}, \tag{10}$$

then $(\bar{u}, \bar{v}) - (0, \bar{v}) = (\bar{u}, 0) \in \mathcal{E}$. Since $(\bar{u}, 0) \in \mathcal{H}_u$, this leads to a contradiction. □

The optimality condition of $\bar{x}$ can be expressed in terms of the disjunction between the sets $\mathcal{E}$ and $\mathcal{H}$.

Proposition 2.2 *$\bar{x} \in R$ is a (global) minimum point for (1) iff*

$$\mathcal{E} \cap \mathcal{H} = \emptyset. \tag{11}$$

Exploiting the fact that $\mathcal{H}$ is a cone, we obtain a further characterization of the optimality condition in the IS.

Corollary 2.1 *(11) is equivalent to*

$$cone(\mathcal{E}) \cap \mathcal{H} = \emptyset. \tag{12}$$

3 Linear separation in the image space

Definition 3.1 The sets $\mathcal{K}_{\bar{x}}$ and $\mathcal{H}$ admit a linear separation, iff there exist $\theta^* \geq 0$ and $\lambda^* \in D^*$ with $(\theta^*, \lambda^*) \neq 0$, such that:

$$\theta^*[f(\bar{x}) - f(x)] + \langle \lambda^*, g(x) \rangle \leq 0, \quad \forall x \in X. \tag{13}$$

If in (13) $\theta^* \neq 0$, then the separation is said to be regular.

Next result shows that a linear functional separates $\mathcal{K}_{\bar{x}}$ and $\mathcal{H}$, iff it separates $\mathcal{E}$ and $\mathcal{H}$.

Proposition 3.1 *Let $(\theta^*, \lambda^*) \in \mathcal{H}^* \setminus \{0\}$. Then the following conditions are equivalent*

(i) $\theta^* u + \langle \lambda^*, v \rangle \leq 0, \quad \forall (u, v) \in \mathcal{K}_{\bar{x}}$.
(ii) $\theta^* u + \langle \lambda^*, v \rangle \leq 0, \quad \forall (u, v) \in \mathcal{E}$.

Proof Suppose that (i) holds. Let $(h_1, h_2) \in \mathcal{H}$. Since $\theta^*(-h_1) + \langle \lambda^*, -h_2 \rangle \leq 0$, then

$$\theta^*(u - h_1) + \langle \lambda^*, v - h_2 \rangle \leq 0, \qquad \forall (u, v) \in \mathcal{K}_{\bar{x}},$$

and (ii) holds.
(ii) $\Rightarrow$ (i) is obvious, since $\mathcal{K}_{\bar{x}} \subseteq \mathcal{E}$. □

The previous result allows us to consider condition (11) which can be proved by showing that $\mathcal{E}$ and $\mathcal{H}$ lie in two disjoint level sets of a suitable functional: when such a functional can be found linear we say that $\mathcal{E}$ and $\mathcal{H}$ admit linear separation.

Theorem 3.1 *$\mathcal{K}_{\bar{x}}$ and $\mathcal{H}$ are properly linearly separable iff*

$$ri\ conv(\mathcal{E}) \cap ri\ \mathcal{H} = \emptyset. \tag{14}$$

Lemma 3.1 Let $B \subseteq \mathbb{R}^m$ be a convex set and $S \subseteq \mathbb{R}^m$. Then,

$$\text{conv}\,(S + B) = \text{conv}(S) + B.$$

Proof ($\subseteq$) Let $\mu_i \geq 0$ with $\sum_{i=1}^{p} \mu_i = 1$. If $s_i \in S$, $b_i \in B$, $i = 1, \ldots, p$, then

$$\sum_{i=1}^{p} \mu_i (s_i + b_i) = \sum_{i=1}^{p} \mu_i s_i + \sum_{i=1}^{p} \mu_i b_i \in \text{ conv}(S) + B,$$

since B is convex.
($\supseteq$) If $s_i \in S$, $i = 1, \ldots, p$ and $b \in B$, then

$$\sum_{i=1}^{p} \mu_i s_i + b = \sum_{i=1}^{p} \mu_i (s_i + b) \in conv(S + B),$$

and the inclusion is proved. □

The previous result allows us to consider the following equivalent formulation of Theorem 3.1.

Theorem 3.2 *$\mathcal{E}$ and $\mathcal{H}$ are properly linearly separable iff*

$$0 \notin ri\ conv(\mathcal{E}) \tag{15}$$

Proof We recall [8] that for any couple of convex sets A and B in $\mathbb{R}^m$ and for any $\gamma, \mu \in \mathbb{R}$ we have that

$$ri(\gamma A + \mu B) = \gamma ri(A) + \mu ri(B), \qquad ri(A) = ri[cl(A)].$$

Therefore (14) is equivalent to

$$0 \notin ri(conv(\mathcal{E})) - ri(\mathcal{H}) = ri(conv(\mathcal{E})) - ri(cl\mathcal{H}) = ri(conv(\mathcal{E}) - cl\mathcal{H}).$$

By Lemma 2.1

$$conv(\mathcal{E}) - cl\mathcal{H} = conv(\mathcal{E} - cl\mathcal{H}) = conv(\mathcal{E})$$

which completes the proof. □

Remark 3.1 We observe that $0 \notin ri\ conv(\mathcal{E})$ iff 0 and $\mathcal{E}$ are properly linearly separable, which means that $\mathcal{E}$ admits a supporting hyperplane at the origin that does not contain $\mathcal{E}$.

Let

$$\mathcal{L}(\theta; \lambda, x) := \theta f(x) - \langle \lambda, g(x) \rangle$$

be the generalized Lagrangian function associated with (1).

Proposition 3.2 *Suppose that $\bar{x} \in R$. Then $\mathcal{E}$ and $\mathcal{H}$ admit a linear separation, iff there exist $\theta^* \geq 0$ and $\lambda^* \in D^*$ with $(\theta^*, \lambda^*) \neq 0$, such that $(\lambda^*, \bar{x})$ is a saddle point for $\mathcal{L}(\theta^*; \lambda, x)$ on $D^* \times X$.*

Proof Suppose that $\mathcal{E}$ and $\mathcal{H}$ admit a linear separation. Since $\bar{x} \in R$, then $\langle \lambda^*, g(\bar{x}) \rangle \geq 0$, so that (13) implies $\langle \lambda^*, g(\bar{x}) \rangle = 0$. Therefore (13) is equivalent to the inequality:

$$\mathcal{L}(\theta^*; \lambda^*, \bar{x}) \leq \mathcal{L}(\theta^*; \lambda^*, x), \quad \forall x \in X.$$

The inequality

$$\mathcal{L}(\theta^*; \lambda^*, \bar{x}) \geq \mathcal{L}(\theta^*; \lambda, \bar{x}), \quad \forall \lambda \in D^*$$

is equivalent to

$$0 \leq \langle \lambda, g(\bar{x}) \rangle, \quad \forall \lambda \in D^*$$

which is fulfilled since $g(\bar{x}) \in D$. □

We recall the next result, Theorem 2.2.7 of [2], that will allow us to characterize the regularity of a linear separation.

Theorem 3.3 *Let $C \subset \mathbb{R}^n$ be a nonempty convex cone with $0 \notin C$, such that $C + clC = C$ and F be any face of C. Let $S \subseteq \mathbb{R}^n$ be nonempty with $0 \in clS$ and such that $S - clC$ is convex. Then F is contained in every hyperplane which separates C and S if and only if*

$$F \subseteq TC(S - clC).$$

By the obvious remark that C and S are linearly separable if and only if so are C and $convS$, we obtain the following generalization of Theorem 3.3:

Theorem 3.4 *Let $C \subset \mathbb{R}^n$ be a nonempty convex cone with $0 \notin C$, such that $C + clC = C$ and F be any face of C. Let $S \subseteq \mathbb{R}^n$ be nonempty with $0 \in clS$. Then F is contained in every hyperplane which separates C ans S if and only if*

$$F \subseteq TC(conv(S - clC)). \tag{16}$$

Proof Applying Theorem 3.3, replacing S by $convS$, we obtain that F is contained in every hyperplane which separates C ans S if and only if

$$F \subseteq TC(conv(S) - clC).$$

By Lemma 3.1, where we put $B := -clC$, we achieve (16). □

Specializing the previous theorem to the case of the separation between $\mathcal{E}$ and $\mathcal{H}$ we obtain the following regularity condition.

Theorem 3.5 *$\mathcal{E}$ and $\mathcal{H}$ admit a separating hyperplane of equation*

$$\theta^* u + \langle \lambda^*, v \rangle = 0, \quad with\ \theta^* > 0,\ \lambda^* \in D^* \tag{17}$$

iff

$$TC(conv(\mathcal{E})) \cap \mathcal{H}_u = \emptyset \tag{18}$$

Proof Let $S := \mathcal{E}$, $C := \mathcal{H}$ and $F := \mathcal{H}_u$. Observe that, because of (10), condition (18) is equivalent to the negation of (16). Applying Theorem 3.4, we obtain that is equivalent to the fact that there exists a hyperplane that separates $\mathcal{E}$ and $\mathcal{H}$ which does not contain $\mathcal{H}_u$, so that its equation has the form (17). □

We recall the following result due to Lemarechal and Hiriart-Urruty [4].

Proposition 3.3 *Let B be a convex set in $\mathbb{R}^m$ and $x \in rbd(B)$. Then B admits a supporting hyperplane at x and its normal vector belongs to aff($B - x$).*

Proof See Lemma 4.2.1 and Remark 4.2.2 of Chapter III in [4]. □

We are now in the position to prove a generalization of the classic Slater constraint qualification.

Theorem 3.6 *Suppose that $\mathcal{E}$ and $\mathcal{H}$ admit a proper linear separation and, furthermore,*

$$0 \in ri\ conv(g(X) - D). \tag{19}$$

Then there exist $\theta^* > 0$ and $\lambda^* \in D^*$, such that

$$\theta^* u + \langle \lambda^*, v \rangle \leq 0, \quad \forall (u, v) \in \mathcal{E} \tag{20}$$

Proof By Theorem 3.2, proper linear separation is equivalent to

$$0 \notin ri\ conv(\mathcal{E}).$$

Since $0 \in \mathcal{E}$ then $0 \in rbd[conv(\mathcal{E})]$.

Applying Proposition 3.3, we obtain that there exist $(\theta^*, \lambda^*) \in aff[conv(\mathcal{E})]$ such that (20) holds. Since $0 \in \mathcal{E}$ then $(\theta^*, \lambda^*) \in lin[conv(\mathcal{E})]$.

Ab absurdo, suppose that $\theta^* = 0$. Then

$$\lambda^* \in lin[(conv(g(X) - D)], \tag{21}$$

and (20) implies

$$\langle \lambda^*, v \rangle \leq 0, \quad \forall v \in conv(g(X) - D). \tag{22}$$

By (19), there exists a neighbourhood U of $0 \in \mathbb{R}^m$ such that

$$V := U \cap lin[conv(g(X) - D)] \subseteq conv[g(X) - D].$$

Taking into account (21), we obtain that $\gamma\lambda^* \in V$ for $|\gamma| < \epsilon$, sufficiently small. Since $V \subseteq conv[g(X) - D]$, by (22),

$$\gamma \langle \lambda^*, \lambda^* \rangle \leq 0, \quad \forall \gamma : \ |\gamma| < \epsilon,$$

which is impossible, for $\lambda^* \neq 0$. □

Remark 3.2 When $int D \neq \emptyset$, then (19) collapses to the classic Slater condition

$$\exists \bar{x} \in X \quad s.t. \quad g(\bar{x}) \in int D. \tag{23}$$

4 The Wolfe and Mond-Weir duals

In the IS, (2) turns out to be equivalent to the following problem:

$$\inf_{\lambda \in D^*} \sup_{(u,v) \in \mathcal{K}_{\bar{x}}} [u + \langle \lambda, v \rangle] \tag{24}$$

Observe that the following relations hold:

$$\inf_{\lambda \in D^*} \sup_{(u,v) \in \mathcal{K}_{\bar{x}}} [u + \langle \lambda, v \rangle] \geq \sup_{(u,v) \in \mathcal{K}_{\bar{x}}} \inf_{\lambda \in D^*} [u + \langle \lambda, v \rangle] = \sup_{\substack{(u,v) \in \mathcal{K}_{\bar{x}} \\ v \in D}} (u) \tag{25}$$

Because of the equality, the second side in (25) will be called *image primal problem*. The difference between the first and the second side in (25) is the *image duality gap*.

Let $S \subseteq \mathbb{R}^{1+m}$, and consider the function

$$\delta^*(\lambda, S) := \sup_{(u,v) \in S} [u + \langle \lambda, v \rangle], \qquad \lambda \in D^*, \tag{26}$$

and observe that, if in (26) there were $(\theta, \lambda) \in \mathbb{R} \times \mathbb{R}^m$ (instead of $(\theta, \lambda) \in \{1\} \times D^*$), then δ^* should be the support function of S; therefore, (26) is its restriction to the subset $\{1\} \times D^*$. The elements of the set

$$\mathcal{Z}(\lambda, S) := \operatorname{argsup}_{(u,v)\in S}[u + \langle \lambda, v\rangle], \qquad \lambda \in D^*, \tag{27}$$

are the arguments of the supremum in (26); they may have one or more components equal to $+\infty$ ($-\infty$ iff $S = \emptyset$); if the supremum is finite, but not maximum, and $(u, v) \in \mathcal{Z}(\lambda, S)$ is finite, then (u, v) is called *Weierstrass point*. In the sequel, S will be either $\mathcal{K}_{\bar{x}}$ or $\mathcal{E}$.

Theorem 4.1 *(i) Suppose that $\bar{x} \in R$ and that $\delta^*(\lambda; \mathcal{E}) < +\infty$ for at least one $\lambda \in D^*$. Then (24) is finite and equals both*:

$$\inf_{\substack{(u,v)\in\mathcal{Z}(\lambda;\mathcal{K}_{\bar{x}})\\ \lambda\in D^*}} [u + \langle \lambda, v\rangle], \tag{28}$$

and

$$\inf_{\substack{(u,v)\in\mathcal{Z}(\lambda;\mathcal{E})\\ \lambda\in D^*}} [u + \langle \lambda, v\rangle]. \tag{29}$$

(ii) Suppose that (18) is fulfilled. Then the infimum in (24) is achieved, equals the 2nd and 3rd sides of (25) and both:

$$\min_{\substack{(u,v)\in\mathcal{Z}(\lambda;\mathcal{K}_{\bar{x}})\\ \lambda\in D^*}} [u + \langle \lambda, v\rangle], \tag{30}$$

and

$$\min_{\substack{(u,v)\in\mathcal{Z}(\lambda;\mathcal{E})\\ \lambda\in D^*}} [u + \langle \lambda, v\rangle]. \tag{31}$$

Proof (i) The nonemptyness of R implies the existence of $(\hat{u}, \hat{v}) \in \mathcal{K}_{\bar{x}} \cap (\mathbb{R} \times D)$. Therefore, we have:

$$\hat{u} + \langle \lambda, \hat{v}\rangle \geq \hat{u}, \quad \forall \lambda \in D^*.$$

It follows that, if $\lambda \in D^*$, then the supremum in (24) is $\geq \hat{u}$. This inequality and the assumption on δ^* imply that the infimum in (24) is finite. To complete the proof it is enough to note that, similarly to what happens for Proposition 3.1, the infimum in (24) does not change if in (26) we replace $(u, v) \in \mathcal{K}_{\bar{x}}$ with $(u, v) \in \mathcal{E}$.

(ii) By Theorem 3.5, (18) is equivalent to the existence of a regular linear separation between $\mathcal{E}$ and $\mathcal{H}$, so that (13) holds with $\theta^* > 0$. With no loss of generality we can suppose $\theta^* = 1$.

By Proposition 3.2, linear separation between $\mathcal{E}$ and $\mathcal{H}$ is equivalent to the existence of a saddle point of the generalized Lagrangian $\mathcal{L}(1; \lambda, x)$ on $D^* \times X$, which can be expressed by the relation

$$\sup_{\lambda\in D^*} \inf_{x\in X} L(x; \lambda) = \inf_{x\in X} \sup_{\lambda\in D^*} L(x; \lambda) \tag{32}$$

and, moreover, the supremum and the infimum are attained in the first and second side of (32), respectively. By changing the sign and adding $f(\bar{x})$ in (32), and taking into account the definition of $\mathcal{K}_{\bar{x}}$, we obtain

$$\inf_{\lambda\in D^*} \sup_{(u,v)\in\mathcal{K}_{\bar{x}}} [u + \langle \lambda, v\rangle] = \sup_{(u,v)\in\mathcal{K}_{\bar{x}}} \inf_{\lambda\in D^*} [u + \langle \lambda, v\rangle].$$

To complete the proof, it is enough to observe that, by Proposition 3.1, the set of multipliers $(1, \lambda)$ that define the gradient of a separating hyperplane between $\mathcal{K}_{\bar{x}}$ and $\mathcal{H}$ or $\mathcal{E}$ and $\mathcal{H}$, coincide. □

The dual problem (24) can be associated with two bounds, by splitting the sign of the scalar product $\langle \lambda, v \rangle$.

Proposition 4.1 *The dual problem (24) enjoys the inequalities*:

$$\inf_{\lambda \in D^*} \sup_{\substack{(u,v) \in \mathcal{K}_{\bar{x}} \\ \langle \lambda, v \rangle \geq 0}} (u) \leq \inf_{\lambda \in D^*} \sup_{(u,v) \in \mathcal{K}_{\bar{x}}} [u + \langle \lambda, v \rangle] \leq \inf_{\lambda \in D^*} \sup_{\substack{(u,v) \in \mathcal{K}_{\bar{x}} \\ \langle \lambda, v \rangle \leq 0}} (u) \tag{33}$$

Proof It is an immediate consequence of the obvious relations:

$$\langle \lambda, v \rangle \geq 0 \Leftrightarrow u \leq u + \langle \lambda, v \rangle; \quad \langle \lambda, v \rangle \leq 0 \Leftrightarrow u \geq u + \langle \lambda, v \rangle.$$

□

Theorem 4.2 *Suppose that (18) is fulfilled. Then*

$$\min_{\substack{(u,v) \in \mathcal{Z}(\lambda; \mathcal{K}_{\bar{x}}) \\ \lambda \in D^*}} [u + \langle \lambda, v \rangle] = \min_{\substack{(u,v) \in \mathcal{Z}(\lambda; \mathcal{K}_{\bar{x}}) \\ \langle \lambda, v \rangle \leq 0 \\ \lambda \in D^*}} (u), \tag{34}$$

Proof By Theorem 3.5, (18) implies the existence of $\lambda^* \in D^*$ such that

$$u + \langle \lambda^*, v \rangle \leq 0, \quad \forall (u, v) \in \mathcal{K}_{\bar{x}}. \tag{35}$$

Therefore the image of the point $\bar{x}$, say $(0, \bar{v})$, belongs to $\mathcal{Z}(\lambda^*; \mathcal{K}_{\bar{x}})$. We also observe that, since $(0, \bar{v}) \in \mathcal{K}_{\bar{x}}$ and $\bar{v} \in D$, then

$$u + \langle \lambda, v \rangle \geq 0, \quad \forall (u, v) \in \mathcal{Z}(\lambda; \mathcal{K}_{\bar{x}}),\ \lambda \in D^*.$$

Therefore

$$\min_{\substack{(u,v) \in \mathcal{Z}(\lambda; \mathcal{K}_{\bar{x}}) \\ \lambda \in D^*}} [u + \langle \lambda, v \rangle] \geq 0.$$

By (35) we have that the point $(u, v, \lambda) := (0, \bar{v}, \lambda^*)$ is a global minimum point of the 1st problem in (34). It is obvious that

$$\min_{\substack{(u,v) \in \mathcal{Z}(\lambda; \mathcal{K}_{\bar{x}}) \\ \lambda \in D^*}} [u + \langle \lambda, v \rangle] \leq \min_{\substack{(u,v) \in \mathcal{Z}(\lambda; \mathcal{K}_{\bar{x}}) \\ \langle \lambda, v \rangle \leq 0 \\ \lambda \in D^*}} [u + \langle \lambda, v \rangle] \leq \min_{\substack{(u,v) \in \mathcal{Z}(\lambda; \mathcal{K}_{\bar{x}}) \\ \langle \lambda, v \rangle \leq 0 \\ \lambda \in D^*}} (u), \tag{36}$$

Since $(0, \bar{v}, \lambda^*)$ fulfils the inequality

$$\langle \lambda^*, \bar{v} \rangle \leq 0,$$

then by (35) we have that such a point is feasible for the 2nd problem in (34). From the inequalities (36) we derive that (34) is fulfilled. □

Remark 4.1 Suitable generalized convexity assumptions on the function $F(x) := (f(x), -g(x))$ turn out to be equivalent to the convexity of the extended image $\mathcal{E}$ or of $cone(\mathcal{E})$. In particular,

(i) $\mathcal{E}$ is convex iff F is $(\mathbb{R}_+ \times D)$-convexlike on X [9];
(ii) $cone(\mathcal{E})$ is convex iff $(f - f(\bar{x}), -g)$ is $(\mathbb{R}_+ \times D)$-preconvexlike on X [11].

Theorem 4.3 *Let $\bar{x}$ be a global minimum point to (1). Suppose that the function $(f, -g)$ is $(\mathbb{R} \times D)$-convexlike on X and that (19) is fulfilled. Then*

$$f(\bar{x}) = \max_{\lambda \in D^*} \inf_{x \in X} L(x; \lambda) = \max_{\lambda \in D^*} \inf_{\substack{x \in X \\ \langle \lambda, g(x) \rangle \leq 0}} f(x), \tag{37}$$

Proof From Remark 4.1 we have that $\mathcal{E}$ is convex so that (13) holds. Moreover, by Theorem 3.6, since (19) is fulfilled, then we can suppose that $\theta^* > 0$. Recalling that

$$f(\bar{x}) = \min_{x \in X} \sup_{\lambda \in D^*} L(x; \lambda),$$

by Proposition 3.2 we obtain that the first equality in (37) holds. The second equality follows from Theorem 4.2, observing that the existence of a regular linear separation implies that (18) is fulfilled. □

Now, for every $\lambda \in D^*$, consider the set of points $\hat{x} \in X$ such that

$$f(\hat{x}) - \langle \lambda, g(\hat{x}) \rangle = \inf_{x \in X} [f(x) - \langle \lambda, g(x) \rangle], \tag{38}$$

and let

$$M := \{(\hat{x}, \lambda) \in X \times D^* \ s.t. \ (38) \text{ is fulfilled}\}.$$

Observe that the point $(u, v) \in \mathcal{K}_{\bar{x}}$ defined by

$$u = f(\bar{x}) - f(\hat{x}), \quad v = g(\hat{x}),$$

belongs to the set $\mathcal{Z}(\lambda, \mathcal{K}_{\bar{x}})$.

This position leads us to prove the following result.

Theorem 4.4 *Let $\bar{x}$ be a global minimum point to (1) and let $A \subseteq M$ be such that there exists $\bar{\lambda} \in D^*$ with $(\bar{x}, \bar{\lambda}) \in A$. Then*

$$f(\bar{x}) = \max_{(x,\lambda) \in A} L(x; \lambda) \tag{39}$$

Proof From (4) it follows that

$$\inf_{x \in X} L(x; \lambda) \leq \inf_{x \in R} f(x) = f(\bar{x}), \quad \forall \lambda \in D^*,$$

which implies

$$L(x; \lambda) \leq f(\bar{x}), \quad \forall (x, \lambda) \in M. \tag{40}$$

Computing (40) for $(x, \lambda) := (\bar{x}, \bar{\lambda})$, we obtain

$$-\langle \bar{\lambda}, g(\bar{x}) \rangle \leq 0.$$

Since $\bar{x} \in R$, the previous inequality leads to the complementarity condition

$$\langle \bar{\lambda}, g(\bar{x}) \rangle = 0, \tag{41}$$

and, in turn, to

$$L(\bar{x}; \bar{\lambda}) = f(\bar{x}),$$

which completes the proof. □

In the hypothesis where $(f, -g)$ is a differentiable $(\mathbb{R}_+ \times D)$-function and X is an open set, setting

$$A := \{(x, \lambda) \in X \times D^* : \nabla_x L(x; \lambda) = 0\}$$

the problem that appears in the right-hand side of (39) collapses to the Wolfe dual. Setting

$$A := \{(x, \lambda) \in X \times D^* : \nabla_x L(x; \lambda) = 0,\ \langle \lambda, g(x) \rangle \leq 0\}$$

then by Theorem 4.4 we can easily recover the Mond-Weir duality. We recall that the Mond-Weir dual [6] is defined by

$$\max_{(x,\lambda)\in A} f(x).$$

Corollary 4.1 *Let X be an open set in $\mathbb{R}^n$ and let $\bar{x}$ be a global minimum point to (1). Suppose that the function $(f, -g)$ is a differentiable $(\mathbb{R}_+ \times D)$-function on X and that (19) is fulfilled. Then*

$$f(\bar{x}) = \max_{(x,\lambda)\in A} f(x). \tag{42}$$

Proof Since $(f, -g)$ is a differentiable $(\mathbb{R}_+ \times D)$-function then it is also $(\mathbb{R}_+ \times D)$-convexlike on X so that $\mathcal{E}$ is convex (see Remark 4.1); moreover, $L(x; \lambda)$ is a differentiable convex function with respect to x, for every $\lambda \in D^*$. Taking into account Proposition 3.2 and Theorem 3.6, we obtain that there exists $\bar{\lambda} \in D^*$ such that

$$L(x; \bar{\lambda}) \geq L(\bar{x}, \bar{\lambda}), \quad \forall x \in X,$$

which implies that $(\bar{x}, \bar{\lambda}) \in A \subseteq M$. Since

$$L(x; \lambda) \geq f(x), \quad \forall (x, \lambda) \in A,$$

then, by Theorem 4.4, we have that

$$f(\bar{x}) = \max_{(x,\lambda)\in A} L(x; \lambda) \geq \max_{(x,\lambda)\in A} f(x),$$

which leads to (42). □

5 Concluding remarks

Most of duality theories existing in the literature may have as a common theoretical background the linear separation between two sets in the Image Space. We have considered the Wolfe and Mond-Weir duals, providing their interpretation in the Image Space. In particular, we have shown that the existence of a regular linear separation is a sufficient conditions for the equivalence of such dual problems.

References

1. Frenk, J.B.G., Kassay, G.: On classes of generalized convex functions, Gordan-Farkas type theorems, and Lagrangian duality. J. Optim. Theory Applic. **102**, 315–343 (1999)
2. Giannessi, F.: Constrained Optimization and Image Space Analysis, vol. 1: Separation of Sets and Optimality Conditions. Springer (2005)
3. Giannessi, F.: Theorems of the alternative and optimality conditions. J. Optimiz. Theory Applic. **42**, 331–365 (1984)
4. Hiriart-Urruty, J.B., Lemarechal, C.: Convex Analysis and Minimization Algorithms, vol. 1. Springer Verlag, Berlin (1993)
5. Mastroeni, G., Rapcsák, T.: On convex generalized systems. J. Optimiz. Theory Applic. **104**, 605–627 (2000)
6. Mond, B., Weir T.: Generalized concavity and duality, in Generalized Concavity in Optimization, pp. 263–279. Academic Press (1981)
7. Rapcsák, T.: Smooth Nonlinear Optimization in $\mathbb{R}^n$, Series Nonconvex Optimization and its Applications, vol. 19, Kluwer (1997)
8. Rockafellar, R.T.: Convex Analysis. Princeton University Press, Princeton (1970)
9. Tardella, F.: On the Image of a Constrained Extremum Problem and Some Applications to the Existence of a Minimum. J. Optim. Theory Applic. **60**, 93–104 (1989)
10. Wolfe, P.: A duality theorem for nonlinear programming. Qu. Appl. Math. **19**, 239–244 (1961)
11. Zeng, R., Caron, R.J.: Generalized Motzkin alternative theorems and vector optimization problems. J. Optim. Theory Applic. **131**, 281–299 (2006)