Journal of Global Optimization **13:** 337–347, 1998.

Conjugate Functions for Convex and Nonconvex Duality

WERNER OETTLI and DIRK SCHLÄGER
Lehrstuhl für Mathematik VII, Universität Mannheim, Mannheim, Germany

(Accepted in final form 30 June 1998)

Abstract. We study conjugate duality with arbitrary coupling functions. Our only tool is a certain support property, which is automatically fulfilled in the two most widely used special cases, namely the case where the underlying space is a topological vector space and the coupling functions are the continuous linear ones, and the case where the underlying space is a metric space and the coupling functions are the continuous ones. We obtain thereby a simultaneous axiomatic extension of these two classical models. Also included is a condition for global optimality, which requires only the mentioned support property.

Key words: Conjugate functions, Duality, Global optimality, Subdifferential

1. Introduction

Fenchel's theory of conjugate convex functions is a classical source of duality in nonlinear optimization [7, 14, 15]. In Fenchel's setting, the conjugate f^* of a function $f \in \mathcal{A}$ is defined as

$$f^*(c) := \sup_{x \in X}(c(x) - f(x)) \quad \forall\, c \in \mathcal{C}.$$

Here X is a locally convex topological vector space, $\mathcal{A}$ is the family of all functions $f : X \to \mathbb{R} \cup \{+\infty\}$ which are proper, convex, and lower semicontinuous, and $\mathcal{C}$ is the family of all functions $c : X \to \mathbb{R}$ which are linear and continuous. We call this setting 'case A'.

It was soon realized [11] that part of Fenchel's results remains valid in a more general setting. Of particular interest is the case where X is a metric space, $\mathcal{A}$ is the family of all functions $f : X \to \mathbb{R} \cup \{+\infty\}$ which are lower semicontinuous, and $\mathcal{C}$ is the family of all functions $c : X \to \mathbb{R}$ which are continuous [1, 4, 13]. We call this situation 'case B'.

Until recently, in studying conjugate duality for case B, it was customary to use some ad hoc assumptions, not needed for case A, for instance the assumption that $f \in \mathcal{A}$ is bounded from below on X by some $c \in \mathcal{C}$. This assumption is not needed for case B, as follows from a classical result of Hahn [9]. To our knowledge, it was Flores-Bazán [8] who for the first time was able to treat case B without such ad hoc assumptions, mainly by reproving Hahn's result [8, Lemma 2.9]. Hahn's result states the following:

If X is a metric space, $f: X \to \mathbb{R} \cup \{+\infty\}$ is lower semicontinuous, $g: X \to \mathbb{R} \cup \{-\infty\}$ is upper semicontinuous, and $g(x) \le f(x)$ for all $x \in X$, then there exists a continuous $c: X \to \mathbb{R}$ such that $g(x) \le c(x) \le f(x)$ for all $x \in X$.

A self-contained proof can be found in [16, pp. 132–134]; for possible extensions see [2, 6, p. 61].

In what follows our only structural requirement will be a certain support property which is automatically fulfilled in case A as well as in case B, its validity in case B being a straightforward consequence of Hahn's result. This support property is formulated as hypothesis (A) in Section 3 below. We discuss some results of conjugation, global optimality, and duality, which need only this hypothesis, hence are valid simultaneously in both settings. In this way we obtain a joint axiomatic extension of cases A and B.

2. Preliminaries

Let X be a nonempty set. Let $\mathcal{C}$ be a nonempty family of real-valued functions defined on X. For any function $f: X \to \overline{\mathbb{R}}$ and $\varepsilon \ge 0$ we define the *ε-subdifferential* of f at $\overline{x} \in X$ as

$$\partial_\varepsilon f(\overline{x}) := \{c \in \mathcal{C} \mid f(x) - c(x) \ge f(\overline{x}) - c(\overline{x}) - \varepsilon \quad \forall x \in X\}.$$

For any function $h: \mathcal{C} \to \overline{\mathbb{R}}$ and $\varepsilon \ge 0$ we define the ε-subdifferential of h at $\overline{c} \in \mathcal{C}$ as

$$\partial_\varepsilon h(\overline{c}) := \{x \in X \mid h(c) - c(x) \ge h(\overline{c}) - \overline{c}(x) - \varepsilon \quad \forall c \in \mathcal{C}\}.$$

For any $f: X \to \overline{\mathbb{R}}$ the *conjugate* $f^*: \mathcal{C} \to \overline{\mathbb{R}}$ and the *biconjugate* $f^{**}: X \to \overline{\mathbb{R}}$ are defined as

$$f^*(c) := \sup_{x \in X}(c(x) - f(x)), \qquad f^{**}(x) := \sup_{c \in \mathcal{C}}(c(x) - f^*(c)).$$

We list some properties of the conjugate and biconjugate (see also [11]).

PROPOSITION 1. *Let $f: X \to \overline{\mathbb{R}}$ be given. The following properties hold:*

(i) *for every $g: X \to \overline{\mathbb{R}}$, $f \le g \implies f^* \ge g^* \implies f^{**} \le g^{**}$;*
$f \not\equiv +\infty \implies f^*(c) \ne -\infty\ \forall c \in \mathcal{C}$;
(ii) *$f^{**} \le f$;*
(iii) *for every $x \in X$, $f^{**}(x) = \sup\{\vartheta(x) \mid \vartheta \in \Theta, \vartheta \le f\}$,*[⋆] *where $\Theta := \{\vartheta: X \to \mathbb{R} \mid \vartheta = c + r,\ c \in \mathcal{C},\ r \in \mathbb{R}\}$;*
(iv) *for every $x \in X$ and $c \in \mathcal{C}$, $f(x) + f^*(c) \ge c(x)$;*[⋆⋆]

⋆ Here $\sup \emptyset := -\infty$.

⋆⋆ For reasons of symmetry, given $a, b \in \overline{\mathbb{R}}$, $r \in \mathbb{R}$, we write $a + b \ge r$ iff $a \ge r - b$. The analogous convention applies to $a + b \le r$.

(v) *for every* $\overline{x} \in X, \overline{c} \in \mathcal{C}, \varepsilon \geq 0$, *there holds*

$$\overline{c} \in \partial_\varepsilon f(\overline{x}) \iff f(\overline{x}) + f^*(\overline{c}) \leq \overline{c}(\overline{x}) + \varepsilon \implies \overline{x} \in \partial_\varepsilon f^*(\overline{c}),$$

and if $f(\overline{x}) = f^{**}(\overline{x})$, *then*

$$f(\overline{x}) + f^*(\overline{c}) \leq \overline{c}(\overline{x}) + \varepsilon \iff \overline{x} \in \partial_\varepsilon f^*(\overline{c});$$

(vi) *for every* $\overline{x} \in X$ *and* $\varepsilon \geq 0$, *there holds* $\partial_\varepsilon f(\overline{x}) \subseteq \partial_\varepsilon f^{**}(\overline{x})$, *and if* $f(\overline{x}) = f^{**}(\overline{x})$, *then* $\partial_\varepsilon f(\overline{x}) = \partial_\varepsilon f^{**}(\overline{x})$;

(vii) *for every* $\overline{x} \in X$, *if* $\partial_\varepsilon f(\overline{x}) \neq \emptyset$ *for all* $\varepsilon > 0$, *then* $f(\overline{x}) = f^{**}(\overline{x})$;

(viii) $f^{***} = f^*$.

Proof.

(i) is obvious.

(ii) Let $x \in X$. Then $c(x) - f^*(c) \leq f(x)$ for all $c \in \mathcal{C}$, thus $f^{**}(x) \leq f(x)$.

(iii) Let $x \in X$ and $s := \sup\{\vartheta(x) \mid \vartheta \in \Theta,\ \vartheta \leq f\}$.
First we prove that $f^{**}(x) \leq s$, i.e., $c(x) - f^*(c) \leq s$ for every $c \in \mathcal{C}$. This is true if $f^*(c) = +\infty$. If $f^*(c) \in \mathbb{R}$, let $\vartheta := c - f^*(c)$. Then $\vartheta \in \Theta$ and, from the definition of f^*, $\vartheta = c - f^*(c) \leq f$. Thus $c(x) - f^*(c) \leq s$. If $f^*(c) = -\infty$, then $f \equiv +\infty$ and therefore $s = +\infty$.
Next we prove that $f^{**}(x) \geq s$. From the definition of s, for every real $\alpha < s$ there exists $\vartheta \leq f$ such that $\alpha < \vartheta(x)$ and $\vartheta = c + r$ with $c \in \mathcal{C}, r \in \mathbb{R}$. Then $c - f \leq -r$, thus $f^*(c) \leq -r$, and therefore $\alpha < c(x) + r \leq c(x) - f^*(c) \leq f^{**}(x)$.

(iv) Let $c \in \mathcal{C}$. From the definition, $f^*(c) \geq c(x) - f(x)$ for every $x \in X$.

(v) $\overline{c} \in \partial_\varepsilon f(\overline{x})$ is equivalent with $\overline{c}(x) - f(x) \leq \overline{c}(\overline{x}) + \varepsilon - f(\overline{x})$ for every $x \in X$, i.e., $f^*(\overline{c}) \leq \overline{c}(\overline{x}) + \varepsilon - f(\overline{x})$.
If $f(\overline{x}) \leq \overline{c}(\overline{x}) + \varepsilon - f^*(\overline{c})$ holds, then $f^*(c) - c(\overline{x}) \geq c(\overline{x}) - f(\overline{x}) - c(\overline{x}) = -f(\overline{x}) \geq f^*(\overline{c}) - \overline{c}(\overline{x}) - \varepsilon$ for every $c \in \mathcal{C}$, thus $\overline{x} \in \partial_\varepsilon f^*(\overline{c})$.
If $f(\overline{x}) = f^{**}(\overline{x})$ and $\overline{x} \in \partial_\varepsilon f^*(\overline{c})$, then $c(\overline{x}) - f^*(c) \leq \overline{c}(\overline{x}) + \varepsilon - f^*(\overline{c})$ for every $c \in \mathcal{C}$, hence $f(\overline{x}) = f^{**}(\overline{x}) \leq \overline{c}(\overline{x}) + \varepsilon - f^*(\overline{c})$.

(vi) Let $\overline{c} \in \partial_\varepsilon f(\overline{x})$. From (ii) and (v), $f^{**}(\overline{x}) \leq f(\overline{x}) \leq \overline{c}(\overline{x}) + \varepsilon - f^*(\overline{c})$ holds. Thus $f^{**}(x) - \overline{c}(x) \geq \overline{c}(x) - f^*(\overline{c}) - \overline{c}(x) = -f^*(\overline{c}) \geq f^{**}(\overline{x}) - \overline{c}(\overline{x}) - \varepsilon$ for every $x \in X$, and therefore $\overline{c} \in \partial_\varepsilon f^{**}(\overline{x})$.
If $f(\overline{x}) = f^{**}(\overline{x})$ and $\overline{c} \in \partial_\varepsilon f^{**}(\overline{x})$, then from (ii) follows $\overline{c} \in \partial_\varepsilon f(\overline{x})$.

(vii) Let $\varepsilon > 0$ and choose $\overline{c} \in \partial_\varepsilon f(\overline{x})$. From (v), $f(\overline{x}) \leq \overline{c}(\overline{x}) - f^*(\overline{c}) + \varepsilon \leq f^{**}(\overline{x}) + \varepsilon$. Since ε was arbitrary, $f(\overline{x}) \leq f^{**}(\overline{x})$ follows.

(viii) From (ii) and (i) we obtain $f^{***} \geq f^*$. Let $c \in \mathcal{C}$. Then $c(x) - f^{**}(x) \leq f^*(c)$ for all $x \in X$, thus $f^{***}(c) \leq f^*(c)$. □

3. The basic hypothesis

To proceed further, let X and $\mathcal{C}$ be as before. Our only structural tool will be the following hypothesis (A), which we require to be satisfied for a given function $f: X \to \overline{\mathbb{R}}$:

(A) *For every* $\overline{x} \in X$ *and every* $\alpha \in \mathbb{R}$ *with* $\alpha < f(\overline{x})$ *there exists* $c \in \mathcal{C}$ *such that* $f(x) - c(x) \geq \alpha - c(\overline{x})$ *for all* $x \in X$.

We note some simple consequences: If $f \not\equiv -\infty$, then hypothesis (A) implies in particular that $f(x) > -\infty$ for all $x \in X$ and $f^* \not\equiv +\infty$. Indeed, if we choose $\alpha < f(\overline{x})$, then from (A) there exists $c \in \mathcal{C}$ such that $c(x) - f(x) \leq c(\overline{x}) - \alpha$ for all $x \in X$, thus $f^*(c) \leq c(\overline{x}) - \alpha < +\infty$. Moreover, if $f(\overline{x}) \in \mathbb{R}$, then (A) implies $\partial_\varepsilon f(\overline{x}) \neq \emptyset$ for all $\varepsilon > 0$; simply choose $\alpha := f(\overline{x}) - \varepsilon$ in (A). However, (A) does not imply $\partial_0 f(\overline{x}) \neq \emptyset$.

PROPOSITION 2. *For any* $f: X \to \overline{\mathbb{R}}$ *the validity of hypothesis* (A) *is equivalent to* $f = f^{**}$.

Proof. Let (A) hold for f. From Proposition 1(ii) we know that $f^{**} \leq f$. To prove $f(\overline{x}) \leq f^{**}(\overline{x})$ for arbitrary $\overline{x} \in X$, let $\alpha \in \mathbb{R}$, $\alpha < f(\overline{x})$. Then there exists $c \in \mathcal{C}$ such that

$$c(\overline{x}) - \alpha \geq c(x) - f(x) \quad \forall x \in X,$$

which implies $c(\overline{x}) - \alpha \geq f^*(c)$, and therefore $f^{**}(\overline{x}) \geq c(\overline{x}) - f^*(c) \geq \alpha$.

Conversely, if $f^{**} = f$, then (A) is satisfied. In fact, let $\alpha < f(\overline{x}) = f^{**}(\overline{x})$. Then there exists $c \in \mathcal{C}$ such that $\alpha \leq c(\overline{x}) - f^*(c) \leq c(\overline{x}) - (c(x) - f(x))$ for all $x \in X$. □

Once it is granted that $f = f^{**}$, the biconjugate is not needed any more for our purposes. We collect the relevant results as follows.

THEOREM 1. *Let* $f: X \to \overline{\mathbb{R}}$ *satisfy* (A), *and let*

$$f^*(c) := \sup_{x \in X}(c(x) - f(x)) \quad \forall c \in \mathcal{C}.$$

Then for every $x \in X$, $c \in \mathcal{C}$, $\varepsilon \geq 0$,

$$f(x) + f^*(c) \geq c(x),$$
$$c \in \partial_\varepsilon f(x) \iff f(x) + f^*(c) \leq c(x) + \varepsilon \iff x \in \partial_\varepsilon f^*(c),$$
$$f(x) = \sup_{c \in \mathcal{C}}(c(x) - f^*(c)).$$

Proof. This follows from Proposition 1(iv), (v), and Proposition 2. □

4. Examples

Let us quote two examples for hypothesis (A).

1. Convex conjugation (Fenchel, Moreau, Rockafellar)

Let X be a locally convex topological vector space. Let X^* denote the topological dual of X, i.e., the space of all continuous linear functionals defined on X. We denote the value of $\xi \in X^*$ at $x \in X$ by $\langle \xi, x \rangle$ instead of $\xi(x)$. Let $\mathcal{C} := X^*$. In this setting we have

PROPOSITION 3. *Let* $f: X \to \mathbb{R} \cup \{+\infty\}$ *be proper (i.e.,* $\not\equiv +\infty$*). Then* f *is convex and lower semicontinuous if, and only if, condition* (A) *holds.*

Proof. If (A) holds, then from Propositions 2 and 1(iii) it follows that $f(x) = \sup\{\vartheta(x) \mid \vartheta \in \Theta,\ \vartheta \leq f\}$. Thus f, being the supremum of a family of continuous affine functions, is lower semicontinuous and convex.

Conversely, let f be proper, convex, and lower semicontinuous. Let $\overline{x} \in X$ and $\alpha \in \mathbb{R}$, $\alpha < f(\overline{x})$. Then $(\overline{x}, \alpha) \notin \operatorname{epi} f := \{(x, r) \in X \times \mathbb{R} \mid f(x) \leq r\}$. Then (see [5, pp. 14/15]) there exists a nonvertical hyperplane which separates $(\overline{x}, \alpha)$ from the closed, convex set $\operatorname{epi} f$. So there exists $\xi \in X^*$ such that $\langle \xi, x \rangle - f(x) \leq \langle \xi, \overline{x} \rangle - \alpha$ for all $x \in X$. Hence (A) holds. □

f^* becomes in this case the Fenchel conjugate of f. If f is proper, convex, and lower semicontinuous, then f^* is proper, convex, and weak* lower semicontinuous. Compare [3, Chapter I]. Moreover, if $f(\overline{x}) \in \mathbb{R}$, then $\partial_\varepsilon f(\overline{x}) \neq \emptyset$ for all $\varepsilon > 0$. But $\partial_0 f(\overline{x})$ may be empty.

2. Nonconvex conjugation (Flores-Bazán [8])

Let X be a metric space. Let $\mathcal{C}$ denote the set of all real-valued continuous functions defined on X. In this setting we have

PROPOSITION 4. *Let* $f: X \to \mathbb{R} \cup \{+\infty\}$. *Then* f *is lower semicontinuous if, and only if, condition* (A) *holds.*

Proof. If (A) is satisfied, then from Propositions 2 and 1(iii) it follows that $f(x) = \sup\{\vartheta(x) \mid \vartheta \in \Theta,\ \vartheta \leq f\}$. Hence f, as a supremum of continuous functions, is lower semicontinuous.

Conversely, let f be lower semicontinuous. Let $\overline{x} \in X$ and $\alpha \in \mathbb{R}$ such that $\alpha \leq f(\overline{x})$. Define a function g as $g(x) := -\infty$ for $x \neq \overline{x}$ and $g(\overline{x}) := \alpha$. Then g is upper semicontinuous and $g \leq f$. From Hahn's result, quoted in the introduction, we obtain $c \in \mathcal{C}$ such that $f(x) - c(x) \geq 0 \geq \alpha - c(\overline{x})$ for all $x \in X$. Hence (A) is satisfied. □

In this situation, if f is lower semicontinuous, then $\partial_0 f(\overline{x}) \neq \emptyset$ for all $\overline{x} \in X$ with $f(\overline{x}) \in \mathbb{R}$.

5. Global optimization

Here, using hypothesis (A), we obtain a condition for global optimality in certain nonconvex minimization problems. Compare also [10, Theorem 4.4, 12, 17].

THEOREM 2. *Let* $h\colon X \to \mathbb{R} \cup \{+\infty\}$ *satisfy hypothesis* (A), *let* $g\colon X \to \overline{\mathbb{R}}$ *be arbitrary, and let* $\psi = g - h$.* *Let* $x_0 \in X$ *be such that* $h(x_0) \in \mathbb{R}$. *Then* x_0 *minimizes* ψ *on* X *if, and only if,*

$$\partial_\varepsilon h(x_0) \subseteq \partial_\varepsilon g(x_0) \quad \forall\, \varepsilon \geq 0. \tag{1}$$

Proof. If x_0 minimizes ψ, then $g(x) - h(x) \geq g(x_0) - h(x_0)$ for all $x \in X$. Let $\varepsilon \geq 0$ and $c \in \partial_\varepsilon h(x_0)$. Then $-c(x) \geq -h(x) + h(x_0) - c(x_0) - \varepsilon$ for all $x \in X$, thus

$$g(x) - c(x) \geq g(x) - h(x) + h(x_0) - c(x_0) - \varepsilon \geq g(x_0) - c(x_0) - \varepsilon$$

for all $x \in X$, so $c \in \partial_\varepsilon g(x_0)$.

Now assume that (1) holds true. Let $\overline{x} \in X$ be arbitrary and $\alpha \in \mathbb{R}$ such that $\alpha < h(\overline{x})$. From hypothesis (A) there exists $c \in \mathcal{C}$ such that

$$h(x) - c(x) \geq \alpha - c(\overline{x}) \quad \forall\, x \in X. \tag{2}$$

Let $\varepsilon := h(x_0) - c(x_0) - \alpha + c(\overline{x})$. Choosing $x = x_0$ in (2) it follows that $\varepsilon \geq 0$. Using ε we can rewrite (2) as

$$h(x) - c(x) \geq h(x_0) - c(x_0) - \varepsilon \quad \forall\, x \in X,$$

i.e., $c \in \partial_\varepsilon h(x_0)$. From (1), $c \in \partial_\varepsilon g(x_0)$, thus

$$g(\overline{x}) - c(\overline{x}) \geq g(x_0) - c(x_0) - \varepsilon = g(x_0) - h(x_0) + \alpha - c(\overline{x}),$$

and therefore $g(\overline{x}) - \alpha \geq g(x_0) - h(x_0)$. Since $\alpha < h(\overline{x})$ was arbitrary, we obtain $g(\overline{x}) - h(\overline{x}) \geq g(x_0) - h(x_0)$. Hence, x_0 minimizes ψ on X. □

6. Duality I

We replace the set X by the product $X \times Y$ of two nonempty sets X, Y. We assume that nonempty sets $\mathcal{C}_X, \mathcal{C}_Y$ of real-valued functions on X, Y, respectively, are given, and we define $\mathcal{C}$ as the set of all $c = (\xi, \eta)\colon X \times Y \to \mathbb{R}$, $c(x, y) = \xi(x) + \eta(y)$, where $\xi \in \mathcal{C}_X$, $\eta \in \mathcal{C}_Y$.

Let $F\colon X \times Y \to \overline{\mathbb{R}}$ satisfy (A). Then

$$F^*(\xi, \eta) = \sup_{x,y}(\xi(x) + \eta(y) - F(x, y)), \tag{3a}$$

* Here $\infty - \infty := \infty$.

$$F(x, y) = \sup_{\xi,\eta}(\xi(x) + \eta(y) - F^*(\xi, \eta)). \tag{3b}$$

We define

$$f(x) := F(x, y_0) - \xi_0(x), \quad f^\#(\eta) := F^*(\xi_0, \eta) - \eta(y_0),$$

with fixed elements $\xi_0 \in \mathcal{C}_X$, $y_0 \in Y$. We consider the pair of problems

(P) $\min\{f(x) \mid x \in X\}$,
(D) $\min\{f^\#(\eta) \mid \eta \in \mathcal{C}_Y\}$.

The Lagrangian for problem (P) is defined as

$$L(x; \eta) := \inf_y (F(x, y) - \xi_0(x) + \eta(y_0) - \eta(y)),$$

and the Lagrangian for problem (D) is defined as

$$L^\#(\eta; x) := \inf_\xi (F^*(\xi, \eta) - \eta(y_0) + \xi_0(x) - \xi(x)).$$

Note that, from (3),

$$f^\#(\eta) = \sup_x(-L(x; \eta)) \quad \text{and} \quad f(x) = \sup_\eta(-L^\#(\eta; x)) .$$

THEOREM 3. *For every $x \in X$, $\eta \in \mathcal{C}_Y$ there holds*

$$f(x) + f^\#(\eta) \geq 0. \tag{4}$$

For every $\overline{x} \in X$, $\overline{\eta} \in \mathcal{C}_Y$, $\varepsilon \geq 0$ there holds

$$\begin{aligned}(\xi_0, \overline{\eta}) \in \partial_\varepsilon F(\overline{x}, y_0) &\Longleftrightarrow f(\overline{x}) + f^\#(\overline{\eta}) \leq \varepsilon \\ &\Longleftrightarrow (\overline{x}, y_0) \in \partial_\varepsilon F^*(\xi_0, \overline{\eta}),\end{aligned}$$

$$f(\overline{x}) + f^\#(\overline{\eta}) \leq \varepsilon \Longrightarrow \begin{cases} f(\overline{x}) - \varepsilon \leq f(x) & \forall\, x \in X, \\ f^\#(\overline{\eta}) - \varepsilon \leq f^\#(\eta) & \forall\, \eta \in \mathcal{C}_Y. \end{cases}$$

Proof. Rewriting $f(x) + f^\#(\eta) \geq 0$ as $F(x, y_0) + F^*(\xi_0, \eta) \geq \xi_0(x) + \eta(y_0)$, and $f(\overline{x}) + f^\#(\overline{\eta}) \leq \varepsilon$ as $F(\overline{x}, y_0) + F^*(\xi_0, \overline{\eta}) \leq \xi_0(\overline{x}) + \overline{\eta}(y_0) + \varepsilon$, we see that (4) and the equivalences are immediate from Theorem 1.

If $f(\overline{x}) + f^\#(\overline{\eta}) \leq \varepsilon$ holds, then from (4) it follows that $f(\overline{x}) - \varepsilon \leq -f^\#(\overline{\eta}) \leq f(x)$ for every x, and $f^\#(\overline{\eta}) - \varepsilon \leq -f(\overline{x}) \leq f^\#(\eta)$ for every η. □

The condition $(\xi_0, \overline{\eta}) \in \partial_\varepsilon F(\overline{x}, y_0)$, i.e.,

$$F(x, y) - \xi_0(x) - \overline{\eta}(y) \geq F(\overline{x}, y_0) - \xi_0(\overline{x}) - \overline{\eta}(y_0) - \varepsilon \quad \forall\, x, y,$$

is equivalent with

$$L(x;\overline{\eta}) \geq f(\overline{x}) - \varepsilon \quad \forall x.$$

The latter is the ε-Kuhn–Tucker condition for problem (P) at $\overline{x}$ with multiplier $\overline{\eta}$. Similarly, the condition $(\overline{x}, y_0) \in \partial_\varepsilon F^*(\xi_0, \overline{\eta})$ is equivalent with

$$L^\#(\eta;\overline{x}) \geq f^\#(\overline{\eta}) - \varepsilon \quad \forall \eta.$$

In the next theorem, hypothesis (A) is not needed for the function F, but for the perturbation function $h\colon Y \to \overline{\mathbb{R}}$ of problem (P), defined as

$$h(y) := \inf_{x\in X}(F(x, y) - \xi_0(x)).$$

We observe that

$$\begin{aligned} f^\#(\eta) &= \sup_{x,y}(\xi_0(x) + \eta(y) - F(x, y)) - \eta(y_0) \\ &= \sup_{y}(-h(y) + \eta(y)) - \eta(y_0) = h^*(\eta) - \eta(y_0). \end{aligned}$$

THEOREM 4. *Assume that the perturbation function $h\colon Y \to \overline{\mathbb{R}}$ together with $\mathcal{C}_Y$ fulfills hypothesis* (A). *Let $\overline{x} \in X$ be such that $f(\overline{x}) \in \mathbb{R}$ and $f(\overline{x}) \leq f(x)$ for all $x \in X$. Then for every $\varepsilon > 0$ there exists $\overline{\eta} \in \mathcal{C}_Y$ such that*

$$f(\overline{x}) + f^\#(\overline{\eta}) \leq \varepsilon.$$

Proof. Since $f(\overline{x}) = \inf_{x\in X} f(x) = \inf_{x\in X}(F(x, y_0) - \xi_0(x)) = h(y_0) \in \mathbb{R}$, it follows from hypothesis (A) for every $\varepsilon > 0$ that there exists $\overline{\eta} \in \mathcal{C}_Y$ such that

$$h(y) - \overline{\eta}(y) \geq f(\overline{x}) - \varepsilon - \overline{\eta}(y_0) \quad \forall y \in Y,$$

hence from the definition of h,

$$F(x, y) - \xi_0(x) + \overline{\eta}(y_0) - \overline{\eta}(y) \geq f(\overline{x}) - \varepsilon \quad \forall x \in X, y \in Y.$$

Taking the infimum over all x, y we obtain $-f^\#(\overline{\eta}) \geq f(\overline{x}) - \varepsilon$. □

EXAMPLE 1. *Let $\varphi\colon X \to \mathbb{R} \cup \{+\infty\}$ and $T\colon X \to Y$. Set*

$$F(x, y) := \begin{cases} \varphi(x), & \textit{if } T(x) = y, \\ +\infty & \textit{else.} \end{cases}$$

Then $f(x) = \varphi(x) - \xi_0(x)$ if $T(x) = y$, and $f(x) = +\infty$ otherwise. Hence problem (P) takes the form

$$(P') \qquad \min\{\varphi(x) - \xi_0(x) \mid x \in X,\ T(x) = y_0\},$$

and the Lagrangian for (P′) becomes

$$L(x;\eta) = \inf_y (F(x,y) - \xi_0(x) + \eta(y_0) - \eta(y))$$
$$= \varphi(x) - \xi_0(x) + \eta(y_0) - \eta(T(x)).$$

7. Duality II

Here the underlying space is the product set $X \times X$, where X is a real vector space. Let $\mathcal{C}_X$ be a nonempty set of real-valued functions on X such that $\xi(0) = 0$ for all $\xi \in \mathcal{C}_X$. We define $\mathcal{C}$ as the set of all $c = (\xi, \eta)\colon X \times X \to \mathbb{R}$ which are of the form $c(x,y) = \xi(x-y) + (\xi - \eta)(y)$ with $\xi, \eta \in \mathcal{C}_X$.

Let $F\colon X \times X \to \overline{\mathbb{R}}$ satisfy (A) with respect to $\mathcal{C}$. Then

$$F^*(\xi,\eta) = \sup_{x,y}(\xi(x-y) + (\xi-\eta)(y) - F(x,y)), \tag{5a}$$

$$F(x,y) = \sup_{\xi,\eta}(\xi(x-y) + (\xi-\eta)(y) - F^*(\xi,\eta)). \tag{5b}$$

We define

$$f(x) := F(x,x), \qquad f^{\#}(\eta) := F^*(\eta,\eta).$$

We consider the pair of problems

(P) $\min\{f(x) \mid x \in X\}$,
(D) $\min\{f^{\#}(\eta) \mid \eta \in \mathcal{C}_X\}$.

As Lagrangian functions for problems (P), (D) we choose

$$L(x,y;\eta) := F(x,y) - \eta(x-y),$$
$$L^{\#}(\xi,\eta;x) := F^*(\xi,\eta) - (\xi-\eta)(x),$$

respectively. From (5) follows $f^{\#}(\eta) = \sup_{x,y}(-L(x,y;\eta))$ and $f(x) = \sup_{\xi,\eta}(-L^{\#}(\xi,\eta;x))$.

THEOREM 5. *For every $x \in X$, $\eta \in \mathcal{C}_X$, there holds*

$$f(x) + f^{\#}(\eta) \geq 0.$$

For every $\overline{x} \in X$, $\overline{\eta} \in \mathcal{C}_X$, $\varepsilon \geq 0$, there holds

$$(\overline{\eta},\overline{\eta}) \in \partial_\varepsilon F(\overline{x},\overline{x}) \iff f(\overline{x}) + f^{\#}(\overline{\eta}) \leq \varepsilon \iff (\overline{x},\overline{x}) \in \partial_\varepsilon F^*(\overline{\eta},\overline{\eta}),$$

$$f(\overline{x}) + f^{\#}(\overline{\eta}) \leq \varepsilon \implies \begin{cases} f(\overline{x}) - \varepsilon \leq f(x) & \forall\, x \in X, \\ f^{\#}(\overline{\eta}) - \varepsilon \leq f^{\#}(\eta) & \forall\, \eta \in \mathcal{C}_X. \end{cases}$$

Proof. We proceed as in the proof of Theorem 3. □

The condition $(\overline{\eta}, \overline{\eta}) \in \partial_\varepsilon F(\overline{x}, \overline{x})$, i.e.,

$$F(x, y) - \overline{\eta}(x - y) \geq F(\overline{x}, \overline{x}) - \overline{\eta}(0) - \varepsilon \quad \forall x, y,$$

is equivalent with

$$L(x, y; \overline{\eta}) \geq f(\overline{x}) - \varepsilon \quad \forall x, y.$$

Likewise the condition $(\overline{x}, \overline{x}) \in \partial_\varepsilon F^*(\overline{\eta}, \overline{\eta})$ is equivalent with

$$L^\#(\xi, \eta; \overline{x}) \geq f^\#(\overline{\eta}) - \varepsilon \quad \forall \xi, \eta.$$

To obtain an analog of Theorem 4 we have to assume that the perturbation function of (P), namely

$$h(z) := \inf\{F(x, y) \mid x - y = z\} \quad (z \in X),$$

satisfies hypothesis (A) with respect to $\mathcal{C}_X$. Let $\overline{x} \in X$ be such that $f(\overline{x}) \in \mathbb{R}$ and $f(\overline{x}) \leq f(x)$ for all $x \in X$. Then for every $\varepsilon > 0$ there exists $\overline{\eta} \in \mathcal{C}_X$ such that

$$f(\overline{x}) + f^\#(\overline{\eta}) \leq \varepsilon.$$

Indeed: $f(\overline{x}) = \inf_{x \in X} f(x) = \inf\{F(x, y) \mid x - y = 0\} = h(0) \in \mathbb{R}$. From hypothesis (A) for every $\varepsilon > 0$ there exists $\overline{\eta} \in \mathcal{C}_X$ such that

$$h(z) - \overline{\eta}(z) \geq h(0) - \varepsilon - \overline{\eta}(0) = f(\overline{x}) - \varepsilon \quad \forall z \in X,$$

hence from the definition of h,

$$F(x, y) - \overline{\eta}(x - y) \geq f(\overline{x}) - \varepsilon \quad \forall x, y \in X.$$

Taking the infimum over all x, y we obtain $-f^\#(\overline{\eta}) \geq f(\overline{x}) - \varepsilon$. □

EXAMPLE 2. *Let X be a topological vector space and $\mathcal{C}_X := X^*$. Let $F(x, y) := h(x) + g(y)$. Then $F^*(\xi, \eta) = \sup_{x,y}(\xi(x) - \eta(y) - F(x, y)) = h^*(\xi) + g^*(-\eta)$, where h^*, g^* are the Fenchel conjugates of h, g. Thus $f(x) = h(x) + g(x)$, $f^\#(\eta) = h^*(\eta) + g^*(-\eta)$. One obtains the dual pair introduced by Fenchel. Compare [3], Chapter I, for further details.*

Acknowledgments

The authors are indebted to Professor Michel Théra (Limoges) for calling their attention to reference [16], and to one of the referees for additional bibliographical remarks.

References

1. Balder, E.J. (1977), An extension of duality-stability relations to nonconvex optimization problems, *SIAM Journal of Control and Optimization* 15: 329–343.

2. Borwein, J.M. and Théra, M. (1992), Sandwich theorems for semicontinuous operators, *Canadian Mathematical Bulletin* 35: 463–474.
3. Brézis, H. (1987), *Analyse fonctionnelle*, Masson, Paris.
4. Dolecki, S. and Kurcyusz, S. (1978), On Φ-convexity in extremal problems, *SIAM Journal of Control and Optimization* 16: 277–300.
5. Ekeland, I. and Temam, R. (1976), *Convex Analysis and Variational Problems*, North-Holland, Amsterdam.
6. Engelking, R. (1989), *General Topology, Revised and Completed Edition*, Heldermann, Berlin.
7. Fenchel, W. (1949), On conjugate convex functions, *Canadian Journal of Mathematics* 1: 73–77.
8. Flores-Bazán, F. (1995), On a notion of subdifferentiability for non-convex functions, *Optimization* 33: 1–8.
9. Hahn, H. (1917), Über halbstetige und unstetige Funktionen, *Sitzungsberichte Akad. Wiss. Wien Abt. IIa* 126: 91–110.
10. Hiriart-Urruty, J.-B. (1989), From convex optimization to nonconvex optimization. Necessary and sufficient conditions for global optimality, in: F.H. Clarke, V.F. Demyanov, F. Giannessi (eds.), *Nonsmooth Optimization and Related Topics*, pp. 219–239, Plenum, New York.
11. Moreau, J.J. (1970), Inf-convolution, sous-additivité, convexité des fonctions numériques, *Journal de Mathématiques Pures et Appliquées* 49: 109–154.
12. Oettli, W. (1995), Kolmogorov conditions for vectorial optimization problems, *OR Spektrum* 17: 227–229.
13. Pallaschke, D. and Rolewicz, S. (1997), *Foundations of Mathematical Optimization*, Kluwer Academic Publishers, Dordrecht/Boston/London.
14. Rockafellar, R.T. (1974), *Conjugate Duality and Optimization*, Society for Industrial and Applied Mathematics, Philadelphia.
15. Stoer, J. and Witzgall, C. (1970), *Convexity and Optimization in Finite Dimensions I*, Springer Verlag, Berlin.
16. Stromberg, K.R. (1981), *An Introduction to Classical Real Analysis*, Wadsworth International, Belmont.
17. Tuy, H. and Oettli, W. (1994), On necessary and sufficient conditions for global optimality, *Revista de Matemáticas Aplicadas* 15: 39–41.