

DUALITY AND REGULARIZATION FOR INF-SUP PROBLEMS

Joachim Gwinner (Darmstadt)*

Werner Oettli (Mannheim)**

83/1988

*Technische Hochschule Darmstadt, Fachbereich Mathematik, Schlossgartenstr. 7,
D-6100 Darmstadt

**Universität Mannheim, Fakultät für Mathematik und Informatik, Schloss,
D-6800 Mannheim

DUALITY AND REGULARIZATION FOR INF-SUP PROBLEMS

JOACHIM GWINNER AND WERNER OETTLI

Introduction. Let X and Y be nonvoid sets and let $L : X \times Y \rightarrow \mathbf{R}$ be a real-valued function. This gives rise to a pair of optimization problems, namely the *primal* problem of finding

$$\alpha := \inf_{x \in X} \sup_{y \in Y} L(x, y) \quad (\text{primal value}),$$

and the *dual* problem of finding

$$\beta := \sup_{y \in Y} \inf_{x \in X} L(x, y) \quad (\text{dual value}).$$

The inequality $\alpha \geq \beta$ being always valid, one is interested in the equality $\alpha = \beta$ (*duality*). The classical von Neumann minimax theorem, as well as its infinite dimensional extension by Fan [3] and more recent refinements due to Fuchssteiner and König [4], which guarantee the duality, require among other conditions the compactness of at least one of the sets X or Y . However, the compactness assumption is too strong for many applications in optimization and related fields. Therefore we are led to consider *regularized values* $\alpha^* \leq \alpha$ and $\beta^* \geq \beta$ such that, under suitable assumptions, $\alpha^* = \beta^*$ (*regularized duality*) holds. In this paper we analyze this approach with three different regularization schemes and provide in each case sufficient conditions for regularized duality to hold:

1. We study the case where Y is a cone and $L(x, y) := f(x) + h(x, y)$, $h(x, \cdot)$ positively homogeneous. Then the primal value is given by

$$\alpha = \inf \{ f(x) \mid x \in X, (\forall y \in Y) h(x, y) \leq 0 \}.$$

We start from the representation $Y := \bigcup_{t \in T} Y(t)$, introduce

$$p(x, t) := \sup_{y \in Y(t)} h(x, y),$$

and consider

$$\alpha^* := \sup_{t \in T} \inf \{ f(x) \mid x \in X, p(x, t) \leq 1 \},$$

$$\beta^* := \sup_{t \in T} \inf_{x \in X} [f(x) + p(x, t)].$$

The determination of β^* may be understood as an abstract penalty method [2, Chapter 10], [5, Chapter 6] connected with the primal problem.

2. We embed L by a function $\varphi : X \times Y \times U \times Y \rightarrow \mathbf{R}$ such that $L(x, y) = \varphi(x, y, 0, 0)$, introduce $\psi(u, v) := \inf_{x \in X} \sup_{y \in Y} \varphi(x, y, u, v)$ ($u \in U, v \in Y$), and consider

$$\alpha^* := \sup_{v \in Y} \inf_{u \in U} \psi(u, v),$$

$$\beta^* := \beta.$$

In this way we extend the Fenchel-Rockafellar perturbation approach in optimization theory [8] to inf-sup problems.

3. We introduce

$$\begin{aligned} C &:= \{v \in \mathbb{R}^X \mid (\exists y \in Y) (\forall x \in X) L(x, y) \geq v(x)\}, \\ D &:= \{u \in \mathbb{R}^Y \mid (\exists x \in X) (\forall y \in Y) L(x, y) \leq u(y)\}, \end{aligned}$$

endow the linear function space \mathbb{R}^X with the product topology, and consider

$$\begin{aligned} \alpha^* &:= \inf_{u \in \text{conv } D} \sup_{y \in Y} u(y), \\ \beta^* &:= \sup_{v \in \text{cl conv } C} \inf_{x \in X} v(x), \end{aligned}$$

where *cl* and *conv* denote the topological closure and convex hull, respectively. Here we establish regularized duality without any further assumptions and, as an immediate consequence, obtain duality under suitable convexity and closedness assumptions. Thus we derive a version of von Neumann's minimax theorem that avoids compactness assumptions and extends an earlier result of Aubin [1, Theorem 2].

We point out that this third scheme opens the access to a broad duality theory which we will present elsewhere.

Regularization (1). Let us introduce the following hypothesis

(H1): X is a convex subset (in a linear space); Y is a closed convex cone (in a linear topological space); $L(x, y) := f(x) + h(x, y)$, where $f(\cdot)$ and $(\forall y \in Y) h(\cdot, y)$ are convex functions; moreover $(\forall x \in X) h(x, \cdot)$ is concave, positively homogeneous, and upper semicontinuous.

By (H1),

$$\sup_{y \in Y} L(x, y) = \begin{cases} f(x) & \text{if } (\forall y \in Y) h(x, y) \leq 0, \\ +\infty & \text{otherwise.} \end{cases}$$

Thus, following the convention that $\inf \emptyset = +\infty$, our primal problem takes the following form

$$(1.1) \quad \alpha = \inf\{f(x) \mid x \in X, (\forall y \in Y) h(x, y) \leq 0\}.$$

The cone Y being not compact, we need the following hypothesis

(H2): We are given a family of nonvoid convex, compact sets $Y(t)$ ($t \in T$) such that $Y = \bigcup_{t \in T} Y(t)$.

Of course we could choose the family of one-element subsets of Y , but our subsequent results will be the more meaningful, the "smaller" - compared to Y - the index set T can be chosen.

Now by (H1) and by the compactness of $Y(t)$, we can apply the Fan-von Neumann minimax theorem [3, Theorem 2] to L in $X \times Y(t)$ and obtain

$$(1.2) \quad \sup_{y \in Y(t)} \inf_{x \in X} L(x, y) = \inf_{x \in X} \sup_{y \in Y(t)} L(x, y).$$

Using the real-valued function

$$p(x, t) := \max_{y \in Y(t)} h(x, y) \quad (x \in X, t \in T),$$

which for all $t \in T$ is convex with respect to x , we may rewrite (1.1) equivalently as follows

$$(1.3) \quad \alpha = \inf\{f(x) \mid x \in X, (\forall t \in T) p(x, t) \leq 0\}.$$

The regularized dual value is defined as

$$\beta^* := \sup_{t \in T} \inf_{x \in X} [f(x) + p(x, t)].$$

We obtain

$$(1.4) \quad \beta = \beta^*.$$

Indeed, using (1.2) we have

$$\begin{aligned} \beta &= \sup_Y \inf_X L(x, y) \\ &= \sup_T \sup_{Y(t)} \inf_X L(x, y) \\ &= \sup_T \inf_X \sup_{Y(t)} [f(x) + h(x, y)] \\ &= \sup_T \inf_X [f(x) + p(x, t)] = \beta^*. \end{aligned}$$

Remark. The function $f(x) + p(x, t)$ has the structure of a penalty function associated to the problem (1.1), and therefore the calculation of β^* may be considered as an abstract penalty procedure [2, Chapter 10]. The penalty function $f(x) + p(x, t)$ is "exact" in the following sense: If there exists $\hat{y} \in Y$ such that $\beta = \inf_X L(x, \hat{y})$, then there exists $\hat{t} \in T$ such that $\beta^* = \inf_X [f(x) + p(x, \hat{t})]$.

The regularized primal value is defined as

$$\alpha^* := \sup_{t \in T} \inf \{f(x) \mid x \in X, p(x, t) \leq 1\}.$$

Since

$$\begin{aligned} \alpha &= \inf \{f(x) \mid x \in X, (\forall t \in T) p(x, t) \leq 0\} \\ &\geq \inf \{f(x) \mid x \in X, p(x, t) \leq 1\} \end{aligned}$$

holds for all $t \in T$, we obtain

$$(1.5) \quad \alpha \geq \alpha^*.$$

Finally to establish a relation between α^* and β^* we require a further – in this section our last – hypothesis

(H3): For every $t_1, t_2 \in T$ and every $\mu_1, \mu_2 \in \mathbb{R}_+$ there exists $t_3 \in T$ such that

$$\mu_1 p(\cdot, t_1) + \mu_2 p(\cdot, t_2) \leq p(\cdot, t_3).$$

THEOREM 1. *Let (H1), (H2), and (H3) be satisfied. If $\beta^* > -\infty$ or $\alpha^* < +\infty$, then $\alpha^* = \beta^*$ holds.*

Proof. Let us first show $\beta^* \leq \alpha^*$. To this end, fix $\varepsilon > 0$ and $t_1 \in T$. By (H3) there exists $t_2 \in T$ such that $p(\cdot, t_1) \leq \varepsilon p(\cdot, t_2)$. Hence

$$\begin{aligned} \inf_{x \in X} [f(x) + p(x, t_1)] &\leq \inf \{f(x) + p(x, t_1) \mid x \in X, p(x, t_2) \leq 1\} \\ &\leq \inf \{f(x) + \varepsilon p(x, t_2) \mid x \in X, p(x, t_2) \leq 1\} \\ &\leq \inf \{f(x) + \varepsilon \mid x \in X, p(x, t_2) \leq 1\} \\ &\leq \alpha^* + \varepsilon. \end{aligned}$$

Therefore $\beta^* \leq \alpha^* + \varepsilon$ for arbitrary $\varepsilon > 0$, proving our claim $\beta^* \leq \alpha^*$.

Now let us establish $\alpha^* \leq \beta^*$. Here we first consider the case that for every $t \in T$ there exists some $x \in X$ such that $p(x, t) \leq 1$. This entails that for every $t \in T$ there exists some $x \in X$ that satisfies $p(x, t) < 1$. Indeed, if for some $t_1 \in T$ and for all $x \in X$, $p(x, t_1) \geq 1$, then by (H3), there exists $t_2 \in T$ such that $p(\cdot, t_2) \geq 2p(\cdot, t_1) \geq 2$ contradicting our assumption. Hence, for all $t \in T$, the convex program

$$\gamma(t) := \inf\{f(x) \mid x \in X, p(x, t) \leq 1\}$$

satisfies the Slater condition, and consequently [2, p. 72] for all $t_1 \in T$ with $\gamma(t_1) > -\infty$ the Kuhn-Tucker condition holds: There exists $\hat{\mu} \in \mathbf{R}_+$ such that

$$\begin{aligned} \gamma(t_1) &\leq \inf_{x \in X} [f(x) + \hat{\mu}(p(x, t_1) - 1)] \\ &\leq \inf_{x \in X} [f(x) + \hat{\mu}p(x, t_1)] \\ &\leq \inf_{x \in X} [f(x) + p(x, t_2)] =: \delta(t_2), \end{aligned}$$

where $t_2 \in T$ is chosen according to (H3). Hence we conclude – what is clear if $\gamma(t_1) = -\infty$ – that for every $t_1 \in T$ there exists $t_2 \in T$ such that $\gamma(t_1) \leq \delta(t_2)$. Thus we arrive at $\alpha^* \leq \beta^*$ in the case considered. Now it remains to prove $\alpha^* \leq \beta^*$ in the case that $\gamma(t_1) = +\infty$ for some $t_1 \in T$. Then $\alpha^* = +\infty$, and by assumption, $\beta^* > -\infty$. So we have $p(x, t_1) > 1$ for every $x \in X$, and for some $t_2 \in T$,

$$f(\cdot) + p(\cdot, t_2) \geq \delta(t_2) > -\infty.$$

Because of (H3), for every $\kappa \in \mathbf{R}_+$, we find some $t_3 \in T$ such that

$$\begin{aligned} f(\cdot) + p(\cdot, t_3) &\geq f(\cdot) + p(\cdot, t_2) + \kappa p(\cdot, t_1) \\ &\geq \delta(t_2) + \kappa. \end{aligned}$$

This shows that $\beta^* = +\infty$, and we arrive at $\beta^* = \alpha^*$. ■

From the preceding theorem, (1.4), and (1.5) we derive the following

COROLLARY 1. *Let (H1), (H2), and (H3) be satisfied. Suppose that $\alpha < +\infty$ or $\beta > -\infty$. Then $\alpha \geq \alpha^* = \beta^* = \beta$ holds.*

The case $\alpha = +\infty$ and $\beta = -\infty$ cannot be excluded by our hypotheses. This case occurs already in linear programming, as soon as the primal and the dual program do not possess feasible solutions; see e.g. [2, p. 21].

Example. We suppose (H1) and that $Y = \bigcup_{t>0} t \cdot B$, where B is convex and compact with $0 \in B$.

We fix $Y(t) := t \cdot B$ ($t > 0$). Then (H2) and (H3) are satisfied. Moreover using

$$P(x) := \max_{y \in B} h(x, y),$$

we have

$$p(x, t) = t \cdot P(x).$$

Since $0 \in B$, the function P is nonnegative. By monotony we obtain

$$\begin{aligned} \beta^* &= \lim_{t \rightarrow \infty} \inf_{x \in X} [f(x) + t \cdot P(x)], \\ \alpha^* &= \lim_{t \rightarrow \infty} \inf \{f(x) \mid x \in X, P(x) \leq \frac{1}{t}\}, \end{aligned}$$

whereas from (1.3)

$$\alpha = \inf \{f(x) \mid x \in X, P(x) \leq 0\}.$$

In particular, if $h(x, y) := \langle H(x), y \rangle$, $H : X \rightarrow \mathbf{R}^m$, $Y := \mathbf{R}_+^m$, and $B = \{y \in \mathbf{R}_+^m \mid \sum_i y_i \leq 1\}$, then the primal problem (1.1) reads

$$(1.6) \quad \alpha = \inf \{f(x) \mid x \in X, H_1(x) \leq 0, \dots, H_m(x) \leq 0\},$$

and moreover,

$$f(x) + p(x, t) = f(x) + t \cdot \max \{0, H_1(x), \dots, H_m(x)\}.$$

Thus we arrive at a well-known exact penalty function for the convex program (1.6); see [5, Chapter 6], [6].

Regularization (2). Let, as in the introduction, $L : X \times Y \rightarrow \mathbf{R}$ and the optimal values α and β be given. Here we need a further set U and suppose that Y and U are contained in linear spaces E_1 and E_2 , respectively, such that $0_{E_1} \in Y$ and $0_{E_2} \in U$. We embed L by a function $\varphi : X \times Y \times U \times Y \rightarrow \mathbf{R}$ in the following way

$$L(x, y) = \varphi(x, y, 0, 0),$$

and introduce the subsequent hypotheses:

$$(H1) \quad (\forall x \in X, y \in Y, v \in Y) \quad \varphi(x, y, 0, v) \leq \varphi(x, y, 0, 0),$$

$$(H2) \quad (\forall x \in X, y \in Y, u \in U) \quad \varphi(x, y, 0, 0) \leq \varphi(x, y, u, y).$$

Note that both hypotheses are satisfied in the setting

$$\varphi(x, y, u, v) := L(x, y) - \langle u, y - v \rangle,$$

where $\langle \cdot, \cdot \rangle$ denotes the duality form for the spaces E_1 and E_2 in duality.

In addition, we define for $u \in U$, $v \in Y$,

$$\psi(u, v) := \inf_{x \in X} \sup_{y \in Y} \varphi(x, y, u, v).$$

Obviously $\alpha = \psi(0, 0)$, and the regularized primal value is defined as

$$\alpha^* := \sup_{v \in Y} \inf_{u \in U} \psi(u, v).$$

We claim that (H1) implies

$$(2.1) \quad \alpha \geq \alpha^*.$$

Indeed, for all $v \in Y$, by (H1)

$$\begin{aligned} \alpha &= \inf_{x \in X} \sup_{y \in Y} \varphi(x, y, 0, 0), \\ &\geq \inf_{x \in X} \sup_{y \in Y} \varphi(x, y, 0, v), \\ &\geq \inf_{u \in U} \inf_{x \in X} \sup_{y \in Y} \varphi(x, y, u, v) \\ &= \inf_{u \in U} \psi(u, v), \end{aligned}$$

which results in

$$\alpha \geq \sup_{v \in Y} \inf_{u \in U} \psi(u, v) = \alpha^*,$$

proving our claim.

Here we simply set

$$\beta^* := \beta = \sup_{y \in Y} \inf_{z \in X} \varphi(x, y, 0, 0).$$

To establish a relation between α^* and β^* we require the additional hypothesis

$$(H3) \quad (\forall x \in X, \bar{y} \in Y) \quad \inf_{u \in U} \sup_{y \in Y} \varphi(x, y, u, \bar{y}) \leq \varphi(x, \bar{y}, 0, 0),$$

which we shall discuss in detail in an example after presenting the results of this section.

THEOREM 2. *Let (H2) and (H3) be satisfied. Then there holds $\alpha^* = \beta^*$, i.e.,*

$$(2.2) \quad \sup_{\bar{y} \in Y} \inf_{u \in U} \psi(u, \bar{y}) = \sup_{\bar{y} \in Y} \inf_{z \in X} \varphi(x, \bar{y}, 0, 0).$$

Proof. Take $\bar{y} \in Y$ arbitrarily. Then by (H2),

$$\begin{aligned} \inf_{z \in X} \varphi(x, \bar{y}, 0, 0) &\leq \inf_{z \in X} \inf_{u \in U} \varphi(x, \bar{y}, u, \bar{y}) \\ &\leq \inf_{u \in U} \inf_{z \in X} \sup_{y \in Y} \varphi(x, y, u, \bar{y}) \\ &= \inf_{u \in U} \psi(u, \bar{y}). \end{aligned}$$

Conversely by (H3),

$$\begin{aligned} \inf_{u \in U} \psi(u, \bar{y}) &= \inf_{z \in X} \inf_{u \in U} \sup_{y \in Y} \varphi(x, y, u, \bar{y}) \\ &\leq \inf_{z \in X} \varphi(x, \bar{y}, 0, 0). \end{aligned}$$

Therefore equality holds in the estimates above leading to the desired relation (2.2). ■

The preceding theorem yields at once the following

COROLLARY 2. *Let (H2) and (H3) be satisfied. Then there holds $\alpha = \beta$ if, and only if, $\alpha = \alpha^*$.*

Examples. (i) Define

$$\varphi(x, y, u, v) := L(x, y) - G(u, y - v),$$

where $G : U \times (Y - Y) \rightarrow \mathbf{R}$ satisfies $G(0, \cdot) \equiv 0$, $G(\cdot, 0) \equiv 0$ and

$$(2.3) \quad (\forall x \in X, \bar{y} \in Y, \varepsilon > 0)(\exists u \in U)(\forall y \in Y) \quad L(x, y) - L(x, \bar{y}) \leq G(u, y - \bar{y}) + \varepsilon.$$

Then all our hypotheses are satisfied. This being obvious for (H1) and (H2), (H3) follows from (2.3). Indeed, (2.3) means that for any fixed $x \in X, \bar{y} \in Y$,

$$(\forall \varepsilon > 0)(\exists u \in U)(\forall y \in Y) \quad \varphi(x, y, u, \bar{y}) \leq \varphi(x, \bar{y}, 0, 0) + \varepsilon,$$

hence

$$(\forall \varepsilon > 0)(\exists u \in U) \sup_{y \in Y} \varphi(x, y, u, \bar{y}) \leq \varphi(x, \bar{y}, 0, 0) + \varepsilon,$$

hence

$$\inf_{u \in U} \sup_{y \in Y} \varphi(x, y, u, \bar{y}) \leq \varphi(x, \bar{y}, 0, 0);$$

thus (H3). – The condition (2.3) can be interpreted as follows: For any $x \in X$, $\bar{y} \in Y$, $\varepsilon > 0$ there exists some $u \in U$ such that the function $G(u, \cdot)$ is an ε -support from above to $L(x, \cdot)$ at \bar{y} . This is a natural assumption, when $L(x, \cdot)$ is concave and $G(u, \cdot)$ is linear.

(ii) Let more specifically, for any $u \in U$, $G(u, \cdot)$ be linear. Suppose (H3) is satisfied. Then

$$\begin{aligned} \psi(u, v) &= \inf_{x \in X} \sup_{y \in Y} [L(x, y) - G(u, y - v)] \\ &= \inf_{x \in X} \sup_{y \in Y} [L(x, y) - G(u, y)] + G(u, v) \\ &=: \sigma(u) + G(u, v), \end{aligned}$$

where σ is the *perturbation function* for the primal problem. This gives

$$\begin{aligned} \alpha &= \psi(0, 0) = \sigma(0), \\ \alpha^* &= \sup_{v \in Y} \inf_{u \in U} \psi(u, v) = \sup_{v \in Y} \inf_{u \in U} [\sigma(u) + G(u, v)]. \end{aligned}$$

Now suppose in addition that $\alpha = \sigma(0)$ is finite and, for any $u \in U \setminus \{0\}$, $\sup_{v \in Y} G(u, v) = +\infty$. Then we obtain

$$\sigma(0) = \inf_{u \in U} \sup_{v \in Y} [\sigma(u) + G(u, v)],$$

and our result $\alpha = \beta \Leftrightarrow \alpha = \alpha^*$ reads

$$\begin{aligned} \inf_{x \in X} \sup_{y \in Y} L(x, y) &= \sup_{y \in Y} \inf_{x \in X} L(x, y) \\ \Leftrightarrow \inf_{u \in U} \sup_{v \in Y} [\sigma(u) + G(u, v)] &= \sup_{v \in Y} \inf_{u \in U} [\sigma(u) + G(u, v)]. \end{aligned}$$

(iii) To arrive at the standard case of convex optimization, we choose

$$L(x, y) := f(x) + \langle y, H(x) \rangle, \quad H : X \rightarrow \mathbf{R}^m, \quad Y := \mathbf{R}_+^m, \quad G(u, v) := \langle v, u \rangle.$$

Then

$$\alpha = \inf_{x \in X} \sup_{y \in Y} L(x, y) = \inf \{f(x) \mid x \in X, H(x) \leq 0\},$$

and

$$\begin{aligned} \sigma(u) &= \inf_{x \in X} \sup_{y \in Y} [f(x) + \langle y, H(x) \rangle - \langle y, u \rangle] \\ &= \inf \{f(x) \mid x \in X, H(x) \leq u\} \end{aligned}$$

is the classical perturbation function.

Thus the existence of a saddle value $\alpha = \beta$ of the Lagrangian L is intimately connected to the existence of a saddle value of the “condensed” Lagrangian $L^*(u, v) := \sigma(u) + G(u, v)$. This is already known for the special inf-sup problem of the classical Lagrangian stemming from convex optimization (compare [8, Theorems 15,16], [7, p. 213]). Here we have recovered this connection in a more general framework.

Regularization (3). Let, as in the introduction, $L : X \times Y \rightarrow \mathbf{R}$ and the optimal values α and β be given. In this section we use the linear function spaces

$$\mathbf{R}^X = \{v \mid v : X \rightarrow \mathbf{R}\}, \quad \mathbf{R}^Y = \{u \mid u : Y \rightarrow \mathbf{R}\}.$$

We provide \mathbf{R}^X with the product topology π . With *conv* we denote the convex hull in \mathbf{R}^X or \mathbf{R}^Y , and with *cl* the π -closure in \mathbf{R}^X . For any $\mu \in \mathbf{R}$ we denote by $\underline{\mu} \in \mathbf{R}^X$ the function having constant value μ on X . We define

$$\begin{aligned} C &:= \{v \in \mathbf{R}^X \mid (\exists y \in Y)(\forall x \in X) L(x, y) \geq v(x)\}, \\ D &:= \{u \in \mathbf{R}^Y \mid (\exists x \in X)(\forall y \in Y) L(x, y) \leq u(y)\}, \end{aligned}$$

and consider

$$\begin{aligned} \alpha^* &:= \inf_{u \in \text{conv } D} \sup_{y \in Y} u(y), \\ \beta^* &:= \sup_{v \in \text{cl conv } C} \inf_{x \in X} v(x). \end{aligned}$$

THEOREM 3. There holds $\alpha \geq \alpha^* = \beta^* \geq \beta$.

Proof. (i) $\alpha \geq \alpha^*$, since

$$\begin{aligned} \alpha &= \inf_{x \in X} \sup_{y \in Y} L(x, y) = \inf_{u \in D} \sup_{y \in Y} u(y) \\ &\geq \inf_{u \in \text{conv } D} \sup_{y \in Y} u(y) = \alpha^*. \end{aligned}$$

Likewise $\beta^* \geq \beta$ is verified.

(ii) Let us show $\alpha^* \geq \beta^*$. Fix $u \in \text{conv } D$ arbitrarily, i.e.,

$$u = \sum_{i=1}^k \rho_i u_i, \quad \sum_{i=1}^k \rho_i = 1, \quad (\rho_1, \dots, \rho_k) \in \mathbf{R}_+^k \quad \text{for some } k \in \mathbf{N},$$

where $u_i \in D$, hence for some $x_i \in X$

$$(\forall y \in Y) \quad u_i(y) \geq L(x_i, y).$$

Fix $v \in \text{conv } C$ arbitrarily, i.e.,

$$v = \sum_{j=1}^l \tau_j v_j, \quad \sum_{j=1}^l \tau_j = 1, \quad (\tau_1, \dots, \tau_l) \in \mathbf{R}_+^l \quad \text{for some } l \in \mathbf{N},$$

where $v_j \in C$, hence for some $y_j \in Y$

$$(\forall x \in X) \quad L(x, y_j) \geq v_j(x).$$

Consequently,

$$\begin{aligned} \sum_{j=1}^l \tau_j u(y_j) &= \sum_{j=1}^l \tau_j \sum_{i=1}^k \rho_i u_i(y_j) \\ &\geq \sum_{j=1}^l \tau_j \sum_{i=1}^k \rho_i L(x_i, y_j) \end{aligned}$$

$$\begin{aligned}
&= \sum_{i=1}^k \rho_i \sum_{j=1}^l \tau_j L(x_i, y_j) \\
&\geq \sum_{i=1}^k \rho_i \sum_{j=1}^l \tau_j v_j(x_i) \\
&= \sum_{i=1}^k \rho_i v(x_i).
\end{aligned}$$

By

$$\sup_{y \in Y} u(y) \geq \sum_{j=1}^l \tau_j u(y_j)$$

it follows for the fixed $u \in \text{conv } D$ that

$$(\forall v \in \text{conv } C) \quad \sup_{y \in Y} u(y) \geq \sum_{i=1}^k \rho_i v(x_i).$$

Since the linear functional $v \mapsto \sum_{i=1}^k \rho_i v(x_i)$ is π -continuous, it follows

$$(\forall v \in \text{cl conv } C) \quad \sup_{y \in Y} u(y) \geq \sum_{i=1}^k \rho_i v(x_i).$$

Hence

$$(\forall u \in \text{conv } D, \forall v \in \text{cl conv } C) \quad \sup_{y \in Y} u(y) \geq \inf_{x \in X} v(x),$$

which results in

$$\alpha^* = \inf_{u \in \text{conv } D} \sup_{y \in Y} u(y) \geq \sup_{v \in \text{cl conv } C} \inf_{x \in X} v(x) = \beta^*.$$

(iii) It remains to show $\alpha^* \leq \beta^*$. Assume this is not true. Then there exists some $\mu \in \mathbb{R}$ such that $\alpha^* > \mu > \beta^*$. This means that, for any $v \in \text{cl conv } C$, $\mu > \inf_{x \in X} v(x)$ holds. Therefore $\underline{\mu} \notin \text{cl conv } C$. The separation theorem yields a $\gamma \in \mathbb{R}$ and a π -continuous linear functional $\Lambda \neq 0$ on \mathbb{R}^X such that

$$(\forall v \in \text{cl conv } C) \quad \Lambda(v) \leq \gamma < \Lambda(\underline{\mu}).$$

Now Λ can be represented by

$$\Lambda(v) = \sum_{n=1}^m \lambda_n v(x_n)$$

with appropriate $m \in \mathbb{N}$, $\lambda_n \in \mathbb{R}$, $x_n \in X$ ($n = 1, \dots, m$). Because of $C - \mathbb{R}_+^X \subset C$, a standard argument shows $\lambda_n \geq 0$ ($n = 1, \dots, m$). Since $\Lambda \neq 0$, $\sum \lambda_n > 0$ holds. Therefore we can assume $\sum \lambda_n = 1$. In particular we obtain, since, for any $y \in Y$, $v = L(\cdot, y) \in C$,

$$(\forall y \in Y) \quad \sum_{n=1}^m \lambda_n L(x_n, y) \leq \gamma < \mu.$$

Since $L(x_n, \cdot) \in D$ ($n = 1, \dots, m$), it follows that $\bar{u} := \sum_{n=1}^m \lambda_n L(x_n, \cdot) \in \text{conv } D$. Thus we arrive at

$$(\forall y \in Y) \quad \bar{u}(y) \leq \gamma < \mu,$$

hence

$$\sup_{y \in Y} \bar{u}(y) < \mu,$$

hence

$$\alpha^* = \inf_{u \in \text{conv } D} \sup_{y \in Y} u(y) < \mu,$$

contradicting $\alpha^* > \mu > \beta^*$. ■

Actually part (iii) of the proof above shows somewhat more: If $\beta^* < +\infty$, then there exists $v^* \in \text{cl conv } C$ such that $\beta^* = \inf_{x \in X} v^*(x)$.

COROLLARY 3. *If D is convex and if C is convex and π -closed, then $\alpha = \beta$ holds.*

Concerning duality, this corollary includes an earlier result of Aubin [1, Theorem 2]. There more concrete conditions for convexity and π -closedness are given.

REFERENCES

- [1] J.-P. Aubin: Théorème du minimax pour une classe de fonctions. C.R. Acad. Sci. Paris Sér. A-B 274 (1972) A455-A458.
- [2] E. Blum, W. Oettli: *Mathematische Optimierung*, Springer, Berlin, 1975.
- [3] K. Fan: Minimax theorems. *Proc. Nat. Acad. Sci. USA* 39 (1953) 42-47.
- [4] B. Fuchssteiner, H. König: New versions of the Hahn-Banach theorem, pp. 255-266, in: *General Inequalities 2* (E.F. Beckenbach; ed.), Birkhäuser, Basel, 1980.
- [5] E.G. Gol'stein: *Theory of convex programming*, Translations of mathematical monographs, vol. 36, AMS, 1972.
- [6] S.P. Han, O.L. Mangasarian: Exact penalty functions. *Math. Programming* 17 (1979) 251-269.
- [7] W. Oettli: Optimality conditions for programming problems involving multivalued mappings, pp. 195-226, in: *Modern Applied Mathematics* (B. Korte; ed.), North-Holland, Amsterdam, 1982.
- [8] R.T. Rockafellar: *Conjugate duality and optimization*. Regional Conference in Applied Mathematics, SIAM, 1974.

Technische Hochschule Darmstadt
 Fachbereich Mathematik
 Schlossgartenstr. 7
 D-6100 Darmstadt

Universität Mannheim
 Fakultät für Mathematik und Informatik
 Schloss
 D-6800 Mannheim