On linear programmming duality and necessary and sufficient conditions in minimax theory.

J.B.G.Frenk*
P.Kas†
G.Kassay‡

March 31, 2004

Econometric Institute Report E.I.2004-14

Abstract

In this paper we discuss necessary and sufficient conditions for different minimax results to hold using only linear programming duality and the finite intersection property of compact sets. It turns out that these necessary and sufficient conditions have a clear interpretation within zero-sum game theory. In the last section we apply these results to derive necessary and sufficient conditions for strong duality for a general class of optimization problems.

keywords: Minimax theory, finite dimensional separation, game theory, Lagrangian and linear programming duality.

1 Introduction.

Let A and B be nonempty sets and $f : A \times B \to \mathbb{R}$ a given function. Since in this paper we consider Borel probability measures on A and B we assume without much loss of generality that A and B are topological spaces with Borel σ-algebras $\mathfrak{A}$ and $\mathfrak{B}$. A minimax result for the function f defined on $A \times B$ is a theorem which asserts that

$$\inf_{\mathbf{b}\in B} \sup_{\mathbf{a}\in A} f(\mathbf{a},\mathbf{b}) = \sup_{\mathbf{a}\in A} \inf_{\mathbf{b}\in B} f(\mathbf{a},\mathbf{b}).$$

*Corresponding author J.B.G.Frenk, Econometric Institute, Erasmus University, PO Box 1738, 3000DR Rotterdam, The Netherlands, E-mail: frenk@few.eur.nl

†P.Kas, Department of Mathematics, Eastern Mediterranean University, Mersin 10, Gazimağusa, TRNC, Turkey, E-mail: peter.kas@emu.edu.tr

‡G.Kassay, Department of Mathematics, Eastern Mediterranean University, Mersin 10, Gazimağusa, TRNC, Turkey, E-mail: gabor.kassay@emu.edu.tr

It is well known that the above equality has important implications in game theory and optimization. In general it is only possible to show that

$$\inf_{\mathbf{b}\in B}\sup_{\mathbf{a}\in A} f(\mathbf{a},\mathbf{b}) \geq \sup_{\mathbf{a}\in A}\inf_{\mathbf{b}\in B} f(\mathbf{a},\mathbf{b})$$

and due to the importance of this equality a lot of papers have appeared in the literature (for an extensive survey see [19] and for a more restrictive one see [7]) introducing sufficient conditions on the function f and the sets A and B for the reverse inequality to hold. To verify this a lot of different proof techniques have been used. Among the most important proof techniques are fixed point theorems, techniques from topology (connectedness) and versions of the Hahn-Banach theorem in finite and infinite dimensional topological vector spaces. The purpose of this paper is to derive for the above and some other related minimax results necessary and sufficient conditions on the function f and the sets A and B. At the same time we have tried to use elementary mathematics and keep the proofs as simple as possible. It turns out for the proof of these necessary and sufficient conditions that we only need either the separation result for finite dimensional disjoint convex sets (Hahn-Banach theorem in finite dimensional vector spaces) or the duality theorem of linear programming and some standard result on compact sets and lower semicontinuous functions. To introduce the other minimax results and their necessary and sufficient condition we first define the notion of a mixed strategy. For any set A let $\mathcal{P}_F(A)$ denote the convex set of all probability measures on A with finite support. If $\epsilon_{\mathbf{a}}$ represents the one-point probability measure concentrated on the point $\mathbf{a}\in A$, this means by definition that λ belongs to $\mathcal{P}_F(A)$ if and only if there exists some finite set $\{\mathbf{a}_1,...,\mathbf{a}_m\}\subseteq A$ and a vector $\mathbf{s}(\lambda) := (s_1(\lambda),...,s_m(\lambda))$ satisfying

$$\lambda = \sum_{i=1}^m s_i(\lambda)\epsilon_{\mathbf{a}_i}, \sum_{i=1}^m s_i(\lambda) = 1 \text{ and } s_i(\lambda) > 0, 1\leq i\leq m. \tag{1}$$

Within game theory (cf.[11]) the set $\mathcal{P}_F(A)$ is known as the set of mixed strategies available to a player having set A as its set of pure strategies. To clarify this name we observe that a player selecting the probability measure λ given by relation (1) will use the pure strategy $\mathbf{a}_i$ with probability $s_i(\lambda), 1\leq i\leq m$. A larger set of strategies is given by the convex set $\mathcal{P}(A)$ of Borel probability measures on A. To extend the minimax result involving the pure strategy sets A and B to a minimax result involving the strategy sets $\mathcal{P}(A)$ and $\mathcal{P}(B)$ we first extend the function f to the larger domain $\mathcal{P}_F(A)\times\mathcal{P}_F(B)$. Therefore introduce for any real valued function h defined on $A\times B$ the function $h_e:\mathcal{P}_F(A)\times\mathcal{P}_F(B)\to\mathbb{R}$ given by

$$h_e(\lambda,\mu) := \sum_{i=1}^m\sum_{j=1}^p s_i(\lambda)s_j(\mu)h(\mathbf{a}_i,\mathbf{b}_j) \tag{2}$$

with $\lambda \in \mathcal{P}_F(A)$ given by relation (1) and $\mu \in \mathcal{P}_F(B)$ given by

$$\mu = \sum_{j=1}^{p} s_j(\mu)\epsilon_{\mathbf{b}_j}, \sum_{j=1}^{p} s_j(\mu) = 1 \text{ and } s_j(\mu) > 0, 1 \leq j \leq p.$$

To extend the function h to the larger domain $\mathcal{P}_F(A) \times \mathcal{P}(B)$ we always assume that the function $h(\mathbf{a}, .) : B \to \mathbb{R}$ belongs for every $\mu \in \mathcal{P}(B)$ and $\mathbf{a} \in A$ to the set $\mathcal{L}^1_\mu(B)$ of Borel measurable functions on B (measurable with respect to the Borel σ-algebra $\mathfrak{B}$), which are Lebesgue absolutely integrable with respect to μ. The function $h_e : \mathcal{P}_F(A) \times \mathcal{P}(B) \to \mathbb{R}$ is now defined by

$$h_e(\lambda, \mu) := \sum_{i=1}^{m} s_i(\lambda) \int_B h(\mathbf{a}_i, \mathbf{b}) d\mu(\mathbf{b}) \tag{3}$$

with λ represented by relation (1). Finally, if we extend the function h to the largest domain $\mathcal{P}(A) \times \mathcal{P}(B)$, we assume that the function h belongs for every $\mu \in \mathcal{P}(B)$ and $\lambda \in \mathcal{P}(A)$ to the set $\mathcal{L}^1_{\lambda\otimes\mu}(A \times B)$ of Borel measurable functions on $A \times B$ (measurable with respect to the Borel product σ-algebra $\mathfrak{A}\otimes\mathfrak{B}$), which are Lebesgue absolutely integrable with respect to the Borel product measure $\lambda \otimes \mu$. In this case it is well known for any h belonging to $\mathcal{L}^1_{\lambda\otimes\mu}(A \times B)$ that the Fubini theorem holds (cf.[1],[15]) and so it follows that

$$\int_{A\times B} hd(\lambda \otimes \mu) = \int_A \int_B hd\mu d\lambda = \int_B \int_A hd\lambda d\mu. \tag{4}$$

The function $h_e : \mathcal{P}(A) \times \mathcal{P}(B) \to \mathbb{R}$ is now defined by

$$h_e(\lambda, \mu) := \int_{A\times B} hd(\lambda \otimes \mu) \tag{5}$$

and by relation (4) it follows that the function h_e is convex and concave in both arguments. Also for every $\mu \in \mathcal{P}(B)$ and $\lambda \in \mathcal{P}_F(A)$ the definition in relation (5) reduces to the definition in relation (3). The same holds for relation (5) and relation (2) in case $\mu \in \mathcal{P}_F(B)$ and $\lambda \in \mathcal{P}_F(A)$. Since the set A and B can be identified with the set of one point probability measures $(\epsilon_\mathbf{a})_{\mathbf{a}\in A}$ and $(\epsilon_\mathbf{b})_{\mathbf{b}\in B}$ it is obvious by relation (2) that the function h_e is indeed an extension of the function h. Consider now the following different minimax results given by

$$\inf_{\mu\in\mathcal{P}(B)} \sup_{\lambda\in\mathcal{P}(A)} f_e(\lambda, \mu) = \sup_{\lambda\in\mathcal{P}(A)} \inf_{\mu\in\mathcal{P}(B)} f_e(\lambda, \mu). \tag{6}$$

$$\inf_{\mu\in\mathcal{P}(B)} \sup_{\lambda\in\mathcal{P}_F(A)} f_e(\lambda, \mu) = \sup_{\lambda\in\mathcal{P}_F(A)} \inf_{\mu\in\mathcal{P}(B)} f_e(\lambda, \mu). \tag{7}$$

$$\inf_{\mu\in\mathcal{P}_F(B)}\sup_{\lambda\in\mathcal{P}_F(A)} f_e(\lambda,\mu) = \sup_{\lambda\in\mathcal{P}_F(A)}\inf_{\mu\in\mathcal{P}_F(B)} f_e(\lambda,\mu). \tag{8}$$

$$\inf_{\mathbf{b}\in B}\sup_{\lambda\in\mathcal{P}_F(A)} f_e(\lambda,\epsilon_{\mathbf{b}}) = \sup_{\lambda\in\mathcal{P}_F(A)}\inf_{\mathbf{b}\in B} f_e(\lambda,\epsilon_{\mathbf{b}}). \tag{9}$$

$$\inf_{\mathbf{b}\in B}\sup_{\mathbf{a}\in A} f(\mathbf{a},\mathbf{b}) = \sup_{\mathbf{a}\in A}\inf_{\mathbf{b}\in B} f(\mathbf{a},\mathbf{b}). \tag{10}$$

In the next section it will be verified that the minimax results considered in the above relations satisfy the following chain of strict inclusions

(10) $\Rightarrow$ (9) $\Rightarrow$ (8) $\Rightarrow$ (7) $\Rightarrow$ (6).

In this paper we derive in Section 2 for the minimax results mentioned in relations (7) up to (10) a necessary and sufficient condition on the function f and the sets A and B. In section 3 we apply the minimax results of Section 2 to derive results for the special case of Lagrangian duality in optimization.

2 On minimax results, inf-compactness and linear programming duality.

To derive a necessary and sufficient condition for the different minimax results we need the following well-known minimax theorem. For completeness an elementary proof of this result based on the separation theorem for finite dimensional convex sets is included. Before mentioning this minimax theorem we introduce the vector $\mathbf{e}^\top := (1, ..., 1)$ belonging to $\mathbb{R}^n$ and the $(n-1)$ dimensional unit simplex $\Delta_n \subseteq \mathbb{R}^n$ given by

$$\Delta_n := \{\alpha \in \mathbb{R}^n : \alpha^\top \mathbf{e} = 1, \alpha \geq \mathbf{0}\}.$$

Moreover, the set $\mathbb{R}^n_-$ denotes the non positive orthant $\{\mathbf{x} \in \mathbb{R}^n : \mathbf{x} \leq \mathbf{0}\}$ of $\mathbb{R}^n$.

Lemma 1 *If $C \subseteq \mathbb{R}^n$ is a convex set, then it follows that*

$$\inf_{\mathbf{x}\in C}\max_{\alpha\in\Delta_n} \alpha^\top\mathbf{x} = \max_{\alpha\in\Delta_n}\inf_{\mathbf{x}\in C} \alpha^\top\mathbf{x}.$$

Proof. It is obvious that

$$\inf_{\mathbf{x}\in C}\max_{\alpha\in\Delta_n} \alpha^\top\mathbf{x} \geq \max_{\alpha\in\Delta_n}\inf_{\mathbf{x}\in C} \alpha^\top\mathbf{x}. \tag{11}$$

To show that we actually have an equality in relation (11) we assume by contradiction that

$$\inf_{\mathbf{x}\in C}\max_{\alpha\in\Delta_n} \alpha^\top\mathbf{x} > \max_{\alpha\in\Delta_n}\inf_{\mathbf{x}\in C} \alpha^\top\mathbf{x} := \gamma. \tag{12}$$

Introduce now the mapping $H : C \to \mathbb{R}^n$ given by $H(\mathbf{x}) := \mathbf{x} - \beta\mathbf{e}$ with β satisfying

$$\inf_{\mathbf{x}\in C} \max_{\alpha\in\Delta_n} \alpha^\top \mathbf{x} > \beta > \gamma. \tag{13}$$

If we assume that $H(C) \cap \mathbb{R}^n_-$ is nonempty there exists some $\mathbf{x}_0 \in C$ satisfying $\mathbf{x}_0 - \beta\mathbf{e} \leq 0$. This implies $\max_{\alpha\in\Delta_n} \alpha^\top \mathbf{x}_0 \leq \beta$ and we obtain a contradiction with relation (13). Therefore $H(C)\cap\mathbb{R}^n_-$ is empty and since both sets are convex we may apply the separation result for finite dimensional disjoint convex sets (cf.[16]). Hence one can find some $\alpha_0 \in \Delta_n$ satisfying $\alpha_0^\top \mathbf{x} - \beta \geq 0$ for every $\mathbf{x} \in C$ and using also the definition of γ listed in relation (12) this implies that

$$\gamma \geq \inf_{\mathbf{x}\in C} \alpha_0^\top \mathbf{x} \geq \beta.$$

Hence we obtain a contradiction with relation (13) and the desired result is proved. □

Since it holds that $\max_{\alpha\in\Delta_n} \alpha^\top \mathbf{x} = \max\{x_1, ..., x_n\}$ for every $\mathbf{x} \in \mathbb{R}^n$ with x_i the ith component of the vector $\mathbf{x}$ an equivalent formulation of Lemma 1 is given by

$$\inf_{\mathbf{x}\in C} \max\{x_1, ..., x_n\} = \max_{\alpha\in\Delta_n} \inf_{\mathbf{x}\in C} \alpha^\top \mathbf{x} \tag{14}$$

for any convex set $C \subseteq \mathbb{R}^n$. Using Lemma 1 it is possible to give a short proof of Wald's minimax result. However, before discussing this result and its proof, let $\mathcal{F}(A_0)$ be the set of all finite subsets of the set $A_0 \subseteq A$ and denote by $|J|$ the cardinality of the set $J \in \mathcal{F}(A_0)$. Moreover, introduce on the set $\mathcal{P}(J), J \in \mathcal{F}(A)$ of all probability measures concentrated on J a topology τ_J with a neighborhood base of $\mu \in \mathcal{P}(J)$ given by the collection

$$N(\mu, \epsilon) := \{\lambda \in \mathcal{P}(J) : |s_i(\lambda) - s_i(\mu)| < \epsilon \text{ for every } 1 \leq i \leq |J|\},$$

$\epsilon > 0$. Since the set $\mathcal{P}(J)$ is isomorphic with $\Delta_{|J|}$ and $\Delta_{|J|} \subseteq \mathbb{R}^{|J|}$ is compact in the Euclidean topology we obtain that $\mathcal{P}(J)$ is compact in the topology τ_J. Moreover, by the definition of $\mathcal{P}(J)$ we obtain that

$$\mathcal{P}(J) = co(\{\epsilon_{\mathbf{a}}\}_{\mathbf{a}\in J}) \tag{15}$$

with $co(C)$ denoting the convex hull of a set C and this shows that $\mathcal{P}(J)$ is a convex compact set. Also it is easy to verify by the definition of $\mathcal{P}_F(A_0)$ with $A_0 \subseteq A$ that $\mathcal{P}_F(A_0)$ is a convex set and

$$\mathcal{P}_F(A_0) = co(\{\epsilon_{\mathbf{a}}\}_{\mathbf{a}\in A_0}) = \cup_{J\in\mathcal{F}(A_0)}\mathcal{P}(J). \tag{16}$$

An immediate consequence of Lemma 1 is the well-known Wald's minimax theorem. This result was already proved by Wald (cf.[23], [7]) by means of a more complicated approach.

Lemma 2 *For every set J belonging to $\mathcal{F}(A)$ it follows that*

$$\begin{aligned}\inf_{\mu\in\mathcal{P}_F(B)}\max{}_{\lambda\in\mathcal{P}(J)} f_e(\lambda,\mu) &= \inf_{\mu\in\mathcal{P}_F(B)}\max_{\mathbf{a}\in J} f_e(\epsilon_{\mathbf{a}},\mu)\\ &= \max_{\lambda\in\mathcal{P}(J)}\inf_{b\in B} f_e(\lambda,\epsilon_{\mathbf{b}})\\ &= \max_{\lambda\in\mathcal{P}(J)}\inf_{\mu\in\mathcal{P}_F(B)} f_e(\lambda,\mu).\end{aligned}$$

Proof. Let J belong to $\mathcal{F}(A)$ and introduce the mapping $L:\mathcal{P}_F(B)\to\mathbb{R}^{|J|}$ given by

$$L(\mu):=(f_e(\epsilon_{\mathbf{a}},\mu))_{\mathbf{a}\in J}.$$

Clearly the functions $\mu\to f_e(\epsilon_{\mathbf{a}},\mu),\mathbf{a}\in J$ are both convex and concave on $\mathcal{P}_F(B)$ and by the convexity of the set $\mathcal{P}_F(B)$ this implies that the range $L(\mathcal{P}_F(B))\subseteq\mathbb{R}^{|J|}$ is a convex set. Applying now relation (15) and Lemma 1 yields

$$\begin{aligned}\inf_{\mu\in\mathcal{P}_F(B)}\max_{\lambda\in\mathcal{P}(J)} f_e(\lambda,\mu) &= \inf_{\mathbf{x}\in L(\mathcal{P}_F(B))}\max_{\alpha\in\Delta_{|J|}}\alpha^\top\mathbf{x}\\ &= \max_{\alpha\in\Delta_{|J|}}\inf_{\mathbf{x}\in L(\mathcal{P}_F(B))}\alpha^\top\mathbf{x}\\ &= \max_{\lambda\in\mathcal{P}(J)}\inf_{\mu\in\mathcal{P}_F(B)} f_e(\lambda,\mu).\end{aligned}$$

Moreover, since the function $\lambda\to f_e(\lambda,\mu)$ is convex for every $\mu\in\mathcal{P}_F(B)$, it follows that

$$\max_{\lambda\in\mathcal{P}(J)} f_e(\lambda,\mu)=\max_{\mathbf{a}\in J} f_e(\epsilon_{\mathbf{a}},\mu) \tag{17}$$

for every $\mu\in\mathcal{P}_F(B)$,while by the concavity of the function $\mu\to f_e(\lambda,\mu)$ for every $\lambda\in\mathcal{P}(J)$ we obtain that

$$\inf_{\mu\in\mathcal{P}_F(B)} f_e(\lambda,\mu)=\inf_{\mathbf{b}\in B} f_e(\lambda,\epsilon_{\mathbf{b}}) \tag{18}$$

for every $\lambda\in\mathcal{P}(J)$. This completes the proof. □

For readers more familiar with the theory of linear programming an alternative proof of Wald's minimax theorem is also provided. Besides the strong duality theorem of linear programming we also need in this alternative proof a well known special case of a result on so-called inf-compact functions. Before mentioning this result we first introduce the following definition (cf.[2]).

Definition 3 *The function $k:B\to\mathbb{R}$ is called inf-compact if all its lower level sets $\{\mathbf{b}\in B:k(\mathbf{b})\leq r\},r\in\mathbb{R}$ are compact and it is called sup-compact if the function $-k$ is inf-compact.*

In case B is a Hausdorff topological space it can be shown (cf.[17]) that any compact set is closed and so an inf-compact function on a Hausdorff topological space is lower semicontinuous. For inf-compact functions the next result is well-known (cf.[2]).

Lemma 4 *If the functions $f(\mathbf{a},.) : B \to \mathbb{R}$ are lower semicontinuous for every $\mathbf{a} \in A$ and there exists some set $J_0 \in \mathcal{F}(A)$ such that the function $\max_{\mathbf{a}\in J_0} f(\mathbf{a},.)$ is inf-compact, then it follows that*

$$\sup_{J\in\mathcal{F}(A)} \inf_{\mathbf{b}\in B} \max_{\mathbf{a}\in J} f(\mathbf{a},\mathbf{b}) = \inf_{\mathbf{b}\in B} \sup_{\mathbf{a}\in A} f(\mathbf{a},\mathbf{b}). \quad (19)$$

Moreover, in both expressions the inf is attained and so we may replace inf by min in relation (19).

In the next section the well-known Slater condition in optimization theory is shown to be equivalent with the inf-compactness of the Lagrangian function and so Lemma 4 is useful in the next section. A symmetrical version of relation (19) is now given by

$$\inf_{I\in\mathcal{F}(B)} \sup_{\mathbf{a}\in A} \min_{\mathbf{b}\in I} f(\mathbf{a},\mathbf{b}) = \sup_{\mathbf{a}\in A} \inf_{\mathbf{b}\in B} f(\mathbf{a},\mathbf{b}) \quad (20)$$

and this holds if the functions $f(.,\mathbf{b}), \mathbf{b} \in B$ are upper semicontinuous on A and there exists some set $I_0 \in \mathcal{F}(B)$ such that the function $\min_{\mathbf{b}\in I_0} f(.,\mathbf{b})$ is sup-compact. In this case it follows that the sup is attained in both expressions and so we may replace $\sup$ by $\max$ in relation (20). Since in any compact space a closed subset of a compact set is compact (cf.[17]) it follows that the conditions of Lemma 4 are satisfied if the topological space B is compact and the functions $f(\mathbf{a},.)$, $\mathbf{a} \in A$ are lower semicontinuous. This well known special case of Lemma 4 will be used in the next proof.

Alternative proof of Wald's minimax result. By relation (16) it follows that

$$\inf_{\mu\in\mathcal{P}_F(B)} \max_{\mathbf{a}\in J} f_e(\epsilon_{\mathbf{a}},\mu) = \inf_{I\in\mathcal{F}(B)} \min_{\mu\in\mathcal{P}(I)} \max_{\mathbf{a}\in J} f_e(\epsilon_{\mathbf{a}},\mu).$$

Observe now for every $I \in \mathcal{F}(B)$ and $J \in \mathcal{F}(A)$ that the optimization problem

$$\min_{\mu\in\mathcal{P}(I)} \max_{\mathbf{a}\in J} f_e(\epsilon_{\mathbf{a}},\mu)$$

is a linear programming problem and applying the strong duality theorem for linear programming (cf.[4]) we obtain von Neumann's minimax result (cf.[21], [22]) given by

$$\min_{\mu\in\mathcal{P}(I)} \max_{\mathbf{a}\in J} f_e(\epsilon_{\mathbf{a}},\mu) = \max_{\lambda\in\mathcal{P}(J)} \min_{\mathbf{b}\in I} f_e(\lambda,\epsilon_{\mathbf{b}}).$$

Applying now the first equality in this proof yields

$$\inf_{\mu\in\mathcal{P}_F(B)} \max_{\mathbf{a}\in J} f_e(\epsilon_{\mathbf{a}},\mu) = \inf_{I\in\mathcal{F}(B)} \max_{\lambda\in\mathcal{P}(J)} \min_{\mathbf{b}\in I} f_e(\lambda,\epsilon_{\mathbf{b}}).$$

Moreover, by the compactness of the convex set $\mathcal{P}(J)$ (with respect to the topology τ_J) for any $J \in \mathcal{F}(A)$ and $\lambda \to f_e(\lambda,\epsilon_{\mathbf{b}})$ is continuous on

$\mathcal{P}(J)$ for every $\mathbf{b} \in B$, it follows by relation (20) replacing the set A by $\mathcal{P}(J)$ and the function $f(\mathbf{a}, \mathbf{b})$ by $f_e(\lambda, \epsilon_{\mathbf{b}})$ that

$$\inf_{I \in \mathcal{F}(B)} \max_{\lambda \in \mathcal{P}(J)} \min_{\mathbf{b} \in I} f_e(\lambda, \epsilon_{\mathbf{b}}) = \max_{\lambda \in \mathcal{P}(J)} \inf_{\mathbf{b} \in B} f_e(\lambda, \epsilon_{\mathbf{b}})$$

and so we obtain

$$\inf_{\mu \in \mathcal{P}_F(B)} \max_{\mathbf{a} \in J} f_e(\epsilon_{\mathbf{a}}, \mu) = \max_{\lambda \in \mathcal{P}(J)} \inf_{\mathbf{b} \in B} f_e(\lambda, \epsilon_{\mathbf{b}}).$$

Finally by relations (17) and (18) Wald's minimax result is verified.

In Wald's minimax result we do not assume anything except that the function f is finite valued. If we additionally assume that the functions $f(\mathbf{a}, .), \mathbf{a} \in A$ belong to $\mathcal{L}^1_\mu(B)$ for every $\mu \in \mathcal{P}(B)$, then the following result holds.

Lemma 5 *If the functions $f(\mathbf{a}, .), \mathbf{a} \in A$ belong to $\mathcal{L}^1_\mu(B)$ for every $\mu \in \mathcal{P}(B)$, then one may replace in Lemma 2 everywhere the set $\mathcal{P}_F(B)$ by $\mathcal{P}(B)$ without changing any values.*

Proof. Since the function $f(\mathbf{a}, .), \mathbf{a} \in A$ belong to $\cap_{\mu \in \mathcal{P}(B)} \mathcal{L}^1_\mu(B)$ we obtain for every $\lambda \in \mathcal{P}_F(A)$ and $\mu \in \mathcal{P}(B)$ that

$$f_e(\lambda, \mu) = \int_B f_e(\lambda, \epsilon_{\mathbf{b}}) d\mu(\mathbf{b}) \geq \inf_{\mathbf{b} \in B} f_e(\lambda, \epsilon_{\mathbf{b}})$$

and so using $(\epsilon_{\mathbf{b}})_{\mathbf{b} \in B} \subseteq \mathcal{P}_F(B) \subseteq \mathcal{P}(B)$ it follows that

$$\inf_{\mu \in \mathcal{P}(B)} f_e(\lambda, \mu) = \inf_{\mu \in \mathcal{P}_F(B)} f_e(\lambda, \mu) = \inf_{\mathbf{b} \in B} f_e(\lambda, \epsilon_{\mathbf{b}}) \quad (21)$$

for every $\lambda \in \mathcal{P}_F(A)$. Moreover, by relation (21) we obtain

$$\begin{aligned} \inf_{\mu \in \mathcal{P}_F(B)} \max_{\lambda \in \mathcal{P}(J)} f_e(\lambda, \mu) &\geq \inf_{\mu \in \mathcal{P}(B)} \max_{\lambda \in \mathcal{P}(J)} f_e(\lambda, \mu) \\ &\geq \max_{\lambda \in \mathcal{P}(J)} \inf_{\mu \in \mathcal{P}(B)} f_e(\lambda, \mu) \\ &= \max_{\lambda \in \mathcal{P}(J)} \inf_{\mathbf{b} \in B} f_e(\lambda, \epsilon_{\mathbf{b}}) \\ &= \max_{\lambda \in \mathcal{P}(J)} \inf_{\mu \in \mathcal{P}_F(B)} f_e(\lambda, \mu). \end{aligned}$$

By Lemma 2 and $\max_{\lambda \in \mathcal{P}(J)} f_e(\lambda, \mu) = \max_{\mathbf{a} \in J} f_e(\epsilon_{\mathbf{a}}, \mu)$ for every $\mu \in \mathcal{P}(B)$ the desired result follows. □

Although mentioned in Lemma 5 we list for further reference the useful observation that for any f satisfying the conditions of Lemma 5 it holds that

$$\inf_{\mu \in \mathcal{P}_F(B)} \max_{\mathbf{a} \in J} f_e(\epsilon_{\mathbf{a}}, \mu) = \inf_{\mu \in \mathcal{P}(B)} \max_{\mathbf{a} \in J} f_e(\epsilon_{\mathbf{a}}, \mu) \quad (22)$$

for any $J \in \mathcal{F}(A)$. Applying relation (16) we obtain the following useful implication of Wald's minimax result and its related version given by Lemma 5.

Lemma 6 *For any function $f : A \times B \to \mathbb{R}$ it follows that*

$$\sup_{J \in \mathcal{F}(A)} \inf_{\mu \in \mathcal{P}_F(B)} \max_{\mathbf{a} \in J} f_e(\epsilon_{\mathbf{a}}, \mu) = \sup_{\lambda \in \mathcal{P}_F(A)} \inf_{\mu \in \mathcal{P}_F(B)} f_e(\lambda, \mu).$$

Moreover, if the functions $f(\mathbf{a}, .), \mathbf{a} \in A$ belong to $\mathcal{L}^1_\mu(B)$ for every $\mu \in \mathcal{P}(B)$, then we may replace in the above equality without changing any values the set $\mathcal{P}_F(B)$ by $\mathcal{P}(B)$.

Proof. The first, respectively second part of this lemma is an immediate consequence of Lemma 2, respectively Lemma 5 and relation (16). □

To derive a natural necessary and sufficient condition for the equality in relation (7) we introduce the following class of functions.

Definition 7 *The function $f : A \times B \to \mathbb{R}$ belongs to the set $\mathcal{A}$ if*

$$\sup_{J \in \mathcal{F}(A)} \inf_{\mu \in \mathcal{P}(B)} \max_{\mathbf{a} \in J} f_e(\epsilon_{\mathbf{a}}, \mu) = \inf_{\mu \in \mathcal{P}(B)} \sup_{\mathbf{a} \in A} f_e(\epsilon_{\mathbf{a}}, \mu)$$

and the above expressions are well defined.

If the function f satisfies the conditions of Lemma 5 it follows by relation (22) that f belongs to $\mathcal{A}$ if and only if

$$\sup_{J \in \mathcal{F}(A)} \inf_{\mu \in \mathcal{P}_F(B)} \max_{\mathbf{a} \in J} f_e(\epsilon_{\mathbf{a}}, \mu) = \inf_{\mu \in \mathcal{P}(B)} \sup_{\mathbf{a} \in A} f_e(\epsilon_{\mathbf{a}}, \mu).$$

A game theoretic interpretation of the payoff function f belonging to the set $\mathcal{A}$ is given by the observation that for player 1 using strategy set $\mathcal{P}(B)$ and the minimax approach it does not make any difference whether his opponent given by player 2 selects a pure strategy from the set A or first considers all finite subsets of A and then selects from one of these finite subsets his pure strategy. It is now easy to show the following result.

Theorem 8 *If the functions $f(\mathbf{a}, .), \mathbf{a} \in A$ belong to $\mathcal{L}^1_\mu(B)$ for every $\mu \in \mathcal{P}(B)$, then it follows that relation (7), given by*

$$\inf_{\mu \in \mathcal{P}(B)} \sup_{\lambda \in \mathcal{P}_F(A)} f_e(\lambda, \mu) = \sup_{\lambda \in \mathcal{P}_F(A)} \inf_{\mu \in \mathcal{P}(B)} f_e(\lambda, \mu)$$

holds if and only if the function f belongs to the set $\mathcal{A}$.

Proof. Since the equality in relation (7) is the same as

$$\inf_{\mu \in \mathcal{P}(B)} \sup_{\mathbf{a} \in A} f_e(\epsilon_{\mathbf{a}}, \mu) = \sup_{\lambda \in \mathcal{P}_F(A)} \inf_{\mu \in \mathcal{P}(B)} f_e(\lambda, \mu).$$

the result follows immediately by the second part of Lemma 6. □

In the next lemma we list for the minimax result in relation (7) some sufficient topological conditions on f and the set B. To verify this result we need some standard results from the theory of Radon measures.

Lemma 9 *If the functions $f(\mathbf{a}, .), \mathbf{a} \in A$ are lower semicontinuous and belong to $\mathcal{L}^1_\mu(B)$ for every $\mu \in \mathcal{P}(B)$ and the set B is a compact Hausdorff space, then it follows that the minimax result in relation (7) holds.*

Proof. By Theorem 8 we need to check that the function f belongs to the set $\mathcal{A}$. To verify this we first observe using the Riesz representation theorem (cf.[17]) that the normed linear space of all finite signed Borel measures $(\mathcal{M}(B), \|.\|_{tv})$ with $\|.\|_{tv}$ denoting the total variation norm is isomorf with the dual space (equipped with the operator norm) of the set of all continuous real valued functions on the compact Hausdorf space B. This implies by the Banach Alaoglu theorem that the unit ball $\mathcal{S} := \{\mu \in \mathcal{M}(B) : \|\mu\|_{tv} \leq 1\}$ is compact in the weak* topology and since the convex set $\mathcal{P}(B) \subseteq \mathcal{S}$ is a closed subset of S (in the weak* topology) we obtain that $\mathcal{P}(B)$ is compact in the weak*topology. Moreover, since the functions $f(\mathbf{a}, .), \mathbf{a} \in A$, are lower semicontinuous it can be shown (cf.[6], [3]) that the function

$$\mu \to f_e(\epsilon_a, \mu) = \int_B f(\mathbf{a}, \mathbf{b}) d\mu(\mathbf{b})$$

is lower semicontinuous (in the weak*topology) for every $\mathbf{a} \in A$. Hence the conditions of Lemma 4 with B replaced by $\mathcal{P}(B)$ and f by $f_e(\epsilon_a, \mu)$ are satisfied and so it follows that

$$\sup_{J \in \mathcal{F}(A)} \inf_{\mu \in \mathcal{P}(B)} \max_{\mathbf{a} \in J} f_e(\epsilon_{\mathbf{a}}, \mu) = \inf_{\mu \in \mathcal{P}(B)} \sup_{\mathbf{a} \in A} f_e(\epsilon_{\mathbf{a}}, \mu)$$

or the function f belongs to $\mathcal{A}$. $\square$

To derive a natural necessary and sufficient condition for the equality in relation (8) we introduce the following class of functions.

Definition 10 *The function $f : A \times B \to \mathbb{R}$ belongs to the set $\mathcal{B}$ if*

$$\sup_{J \in \mathcal{F}(A)} \inf_{\mu \in \mathcal{P}_F(B)} \max_{\mathbf{a} \in J} f_e(\epsilon_{\mathbf{a}}, \mu) = \inf_{\mu \in \mathcal{P}_F(B)} \sup_{\mathbf{a} \in A} f_e(\epsilon_{\mathbf{a}}, \mu).$$

A game theoretic interpretation of the payoff function f belonging to the set $\mathcal{B}$ is given by the observation that for player 1 using the mixed strategy set $\mathcal{P}_F(B)$ and the minimax approach it does not make any difference whether his opponent given by player 2 selects a pure strategy from the set A or first considers all finite subsets of A and then selects from one of these finite subsets his pure strategy. If we know additionally that the set B is a compact Hausdorff space and the functions $f(\mathbf{a}, .), \mathbf{a} \in A$ are lower semicontinuous and belong to $\mathcal{L}^1_\mu(B)$ for every $\mu \in \mathcal{P}(B)$, then the definition of the set $\mathcal{B}$ can be simplified. If this holds we know by relation (22) and Lemma 9 that

$$\sup_{J \in \mathcal{F}(A)} \inf_{\mu \in \mathcal{P}_F(B)} \max_{\mathbf{a} \in J} f_e(\epsilon_{\mathbf{a}}, \mu) = \inf_{\mu \in \mathcal{P}(B)} \sup_{\mathbf{a} \in A} f_e(\epsilon_{\mathbf{a}}, \mu) \tag{23}$$

and so under the above conditions we obtain that

$$f \in \mathcal{B} \Longleftrightarrow \inf_{\mu \in \mathcal{P}(B)} \max_{\mathbf{a} \in A} f_e(\epsilon_{\mathbf{a}}, \mu) = \inf_{\mu \in \mathcal{P}_F(B)} \sup_{\mathbf{a} \in A} f_e(\epsilon_{\mathbf{a}}, \mu).$$

Observe in this case the game theoretic interpretation of the set $\mathcal{B}$ becomes easier and is given by the observation that player 1 using the strategy set $\mathcal{P}(B)$ can restrict himself to the strategy set $\mathcal{P}_F(B)$. One can now show the following result.

Theorem 11 *It follows that relation (8), given by*

$$\inf_{\mu \in \mathcal{P}_F(B)} \sup_{\lambda \in \mathcal{P}_F(A)} f_e(\lambda, \mu) = \sup_{\lambda \in \mathcal{P}_F(A)} \inf_{\mu \in \mathcal{P}_F(B)} f_e(\lambda, \mu)$$

holds if and only if the function f belongs to $\mathcal{B}$.

Proof. Apply a similar proof as in Theorem 8 and use the first part of Lemma 6. □

The minimax result listed in relation (8) is of importance in the theory of zero-sum games. It states that both players should use the set of mixed strategies to achieve the (maybe not attainable) value of a zero-sum game. If the function f is continuous on $A \times B$ and the sets A and B are compact sets in a metric space Ville (cf.[20], [7]) showed that relation (8) holds. Applying the result that any continuous function on a compact set in a metric space is uniformly continuous (cf.[13]) it is easy to verify that the function f belongs to the set $\mathcal{B}$ and so Ville's minimax result follows from Theorem 11. To derive a necessary and sufficient condition for the equality in relation (9) we introduce the following class of functions.

Definition 12 *The function $f : A \times B \to \mathbb{R}$ belongs to the set $\mathcal{C}$ if*

$$\sup_{J \in \mathcal{F}(A)} \inf_{\mu \in \mathcal{P}_F(B)} \max_{\mathbf{a} \in J} f_e(\epsilon_{\mathbf{a}}, \mu) = \inf_{\mathbf{b} \in B} \sup_{\mathbf{a} \in A} f(\mathbf{a}, \mathbf{b}).$$

A game theoretic interpretation of the payoff function f belonging to the set $\mathcal{C}$ is given by the observation that for player 1 using the mixed strategy set $\mathcal{P}_F(B)$ and the minimax approach it does not make any difference whether his opponent given by player 2 selects a pure strategy from the set A or first considers all finite subsets of A and then selects from one of these finite subsets his pure strategy. Moreover, the payoff function for player 1 is such that his mixed strategy set is always dominated by his pure strategy set. This means that player 1 can restrict himself to the set of pure strategies instead of using the set of mixed strategies. By relation (23) we obtain for B a compact Hausdorff space and the functions $f(\mathbf{a}, .), \mathbf{a} \in A$ are lower semicontinuous and belong to $\mathcal{L}^1_\mu(B)$ for every $\mu \in \mathcal{P}(B)$ that

$$f \in \mathcal{C} \Longleftrightarrow \inf_{\mu \in \mathcal{P}(B)} \sup_{\mathbf{a} \in A} f_e(\epsilon_{\mathbf{a}}, \mu) = \inf_{\mathbf{b} \in B} \sup_{\mathbf{a} \in A} f(\mathbf{a}, \mathbf{b}).$$

Again in this case the game theoretic interpretation of the set $\mathcal{C}$ becomes easier and is given by the observation that player 1 using the strategy set $\mathcal{P}(B)$ can restrict himself to the pure strategy set B. One can now show the following result. Observe a sufficient condition for the listed minimax result was discussed in [12].

Theorem 13 *It follows that relation (9), given by*

$$\inf_{\mathbf{b}\in B}\sup_{\lambda\in\mathcal{P}_F(A)} f_e(\lambda,\epsilon_{\mathbf{b}}) = \sup_{\lambda\in\mathcal{P}_F(A)}\inf_{\mathbf{b}\in B} f_e(\lambda,\epsilon_{\mathbf{b}}).$$

holds if and only if the function f belongs to $\mathcal{C}$.

Proof. The equality in relation (9) is the same as

$$\inf_{\mathbf{b}\in B}\sup_{\mathbf{a}\in A} f(\mathbf{a},\mathbf{b}) = \sup_{\lambda\in\mathcal{P}_F(A)}\inf_{\mu\in\mathcal{P}_F(B)} f_e(\lambda,\mu)$$

Applying now the first part of Lemma 6 yields the desired result. □

Finally we derive a necessary and sufficient condition for a minimax result involving the pure strategy sets A and B.

Definition 14 *The function $f : A\times B\to\mathbb{R}$ belongs to the set $\mathcal{D}$ if*

$$\sup_{\lambda\in\mathcal{P}_F(A)}\inf_{\mathbf{b}\in B} f_e(\lambda,\epsilon_{\mathbf{b}}) = \sup_{\mathbf{a}\in A}\inf_{\mathbf{b}\in B} f(\mathbf{a},\mathbf{b}).$$

A game theoretic interpretation of the payoff function f belonging to the set $\mathcal{C}$ is given by the observation that for player 2 using the mixed strategy set $\mathcal{P}_F(A)$ and the minimax approach his mixed strategy set is always dominated by his pure strategy set. This means that player 2 can restrict himself to the set of pure strategies instead of using the set of mixed strategies. One can now show the most useful minimax result.

Theorem 15 *It follows that relation (10), given by*

$$\inf_{\mathbf{b}\in B}\sup_{\mathbf{a}\in A} f(\mathbf{a},\mathbf{b}) = \sup_{\mathbf{a}\in A}\inf_{\mathbf{b}\in B} f(\mathbf{a},\mathbf{b}).$$

holds if and only if the function f belongs to the set $\mathcal{C}\cap\mathcal{D}$.

Proof. If the function f belongs to the set $\mathcal{C}\cap\mathcal{D}$ then by Theorem 13 we obtain that

$$\begin{aligned}\inf_{\mathbf{b}\in B}\sup_{\mathbf{a}\in A} f_e(\mathbf{a},\mathbf{b}) &= \inf_{\mathbf{b}\in B}\sup_{\lambda\in\mathcal{P}_F(A)} f_e(\lambda,\epsilon_{\mathbf{b}})\\ &= \sup_{\lambda\in\mathcal{P}_F(A)}\inf_{\mathbf{b}\in B} f_e(\lambda,\epsilon_{\mathbf{b}})\\ &= \sup_{\mathbf{a}\in A}\inf_{\mathbf{b}\in B} f(\mathbf{a},\mathbf{b}).\end{aligned}$$

To show the reverse implication consider an arbitrary λ belonging to $\mathcal{P}_F(A)$. By relation (16) there exists some $J_0\in\mathcal{F}(A)$ such that $\lambda\in\mathcal{P}(J_0)$ and this implies

$$\begin{aligned}\inf_{\mathbf{b}\in B} f_e(\lambda,\epsilon_{\mathbf{b}}) &\leq \inf_{\mathbf{b}\in B}\sup_{\mathbf{a}\in J_0} f(\mathbf{a},\mathbf{b})\\ &\leq \sup_{J\in\mathcal{F}(A)}\inf_{\mathbf{b}\in B}\sup_{\mathbf{a}\in J} f(\mathbf{a},\mathbf{b}).\end{aligned}$$

Applying the minimax equality yields

$$\begin{aligned}\sup_{\lambda\in\mathcal{P}_F(A)}\inf_{\mathbf{b}\in B} f_e(\lambda,\epsilon_{\mathbf{b}}) &\le \sup_{J\in\mathcal{F}(A)}\inf_{\mathbf{b}\in B}\sup_{\mathbf{a}\in J} f(\mathbf{a},\mathbf{b})\\ &\le \inf_{\mathbf{b}\in B}\sup_{\mathbf{a}\in A} f(\mathbf{a},\mathbf{b})\\ &= \sup_{\mathbf{a}\in A}\inf_{\mathbf{b}\in B} f(\mathbf{a},\mathbf{b}).\end{aligned}$$

Since the reverse inequality trivially holds we obtain

$$\sup_{\lambda\in\mathcal{P}_F(A)}\inf_{\mathbf{b}\in B} f_e(\lambda,\epsilon_{\mathbf{b}}) = \sup_{\mathbf{a}\in A}\inf_{\mathbf{b}\in B} f(\mathbf{a},\mathbf{b}) \tag{24}$$

or the function f belongs to $\mathcal{D}$. Again by the minimax equality and (24) we obtain

$$\sup_{\lambda\in\mathcal{P}_F(A)}\inf_{\mathbf{b}\in B} f_e(\lambda,\epsilon_{\mathbf{b}}) = \inf_{\mathbf{b}\in B}\sup_{\mathbf{a}\in A} f(\mathbf{a},\mathbf{b})$$

and this shows by Theorem 13 that the function f belongs to $\mathcal{C}$. □

This concludes our discussion of the necessary and sufficient conditions for the different minimax results. We will now investigate in the next subsection in more detail these different function classes and show how they are related.

2.1 On the relations between the different minimax results.

In this subsection we investigate in more detail the relations between the different minimax results given by relations (6) up to (10). Introducing the notation L_i and R_i for the left and right-hand side of relation (i) we obviously obtain that

$$L_{10} = L_9 \ge L_8 \ge L_7 = L_6 \ge R_6 \ge R_7 = R_8 = R_9 \ge R_{10}. \tag{25}$$

This implies that

$$(10) \Rightarrow (9) \Rightarrow (8) \Rightarrow (7) \Rightarrow (6).$$

Below we show by means of some counterexamples that none of the arrows in relation (25) can be reversed. In the first counterexample we show an instance for which (9) holds and (10) does not hold.

Example 16 *Let $A = [0,1] \subset \mathbb{R}$, $B = \{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\} \subset \mathbb{R}$ and introduce the function $f : A \times B \to \mathbb{R}$ given by*

$$f(\mathbf{a},\mathbf{b}) = \begin{cases} \mathbf{a}^2 & \textit{if } \mathbf{b} = \mathbf{b}_1 \\ (\mathbf{a}-1)^2 & \textit{if } \mathbf{b} = \mathbf{b}_2 \\ 2^{-1} & \textit{if } \mathbf{b} = \mathbf{b}_3 \end{cases}.$$

For this bifunction we have

$$L_{10} := \min_{\mathbf{b}\in B}\sup_{\mathbf{a}\in A} f(\mathbf{a},\mathbf{b}) = 1/2,$$

while

$$R_{10} := \sup_{\mathbf{a}\in A} \min_{\mathbf{b}\in B} f(\mathbf{a},\mathbf{b}) = 1/4$$

and so (10) does not hold. Since $L_{10} = L_9 = 2^{-1}$ and it is obvious to check that $R_9 = 2^{-1}$, we obtain that (9) holds.

In the next counterexample we show an instance for which (8) holds and (9) does not hold.

Example 17 *Take $A = [0,1]$, $B = \{\mathbf{b}_1, \mathbf{b}_2\} \subset \mathbb{R}$ and introduce the function $f : A \times B \to \mathbb{R}$ given by*

$$f(\mathbf{a},\mathbf{b}) = \begin{cases} \mathbf{a}^2 & \text{if } \mathbf{b} = \mathbf{b}_1 \\ (\mathbf{a}-1)^2 & \text{if } \mathbf{b} = \mathbf{b}_2 \end{cases} .$$

Consider now the probability measure $\lambda^* \in \mathcal{P}_F(A)$ given by $\lambda^* = 2^{-1}\epsilon_{\mathbf{a}_1} + 2^{-1}\epsilon_{\mathbf{a}_2}$ with $\mathbf{a}_1 = 0$ and $\mathbf{a}_2 = 1$. It is easy to check that

$$\min_{\mathbf{b}\in B} f_e(\lambda^*, \epsilon_{\mathbf{b}}) = 2^{-1}$$

and so it follows that $R_9 \geq 2^{-1}$. Moreover, we observe by the definition of the sets A and B that

$$\begin{aligned} L_8 &= \inf_{\mu\in\mathcal{P}_F(B)} \sup_{\mathbf{a}\in A} f_e(\epsilon_{\mathbf{a}}, \mu) \\ &= \inf_{0\leq s_1(\mu)\leq 1} \sup_{\mathbf{a}\in A} \{s_1(\mu) f(\mathbf{a},\mathbf{b}_1) + (1 - s_1(\mu)) f(\mathbf{a},\mathbf{b}_2)\} \\ &= \inf_{0\leq s_1(\mu)\leq 1} \max\{s_1(\mu), 1 - s_1(\mu)\} = 2^{-1}. \end{aligned}$$

Since we already know that $L_8 \geq R_9 = R_8$ and $R_9 \geq 2^{-1}$ we obtain

$$L_8 = R_9 = R_8 = 2^{-1}.$$

It is now easy to check that $L_9 = 1$ and hence we have found an instance for which (8) holds and (9) does not hold.

In order to construct an instance for which (7) holds and (8) does not hold we first need to introduce the set C_0 of all (real valued) sequences converging to 0. It is well-known that the space C_0 endowed with the norm

$$\|\mathbf{a}\|_{C_0} = \sup_{k\in N} |a_k|$$

is a Banach space. Let $A = \{\mathbf{a} = (a_k) \in C_0 : a_1 = 0\}$, $B = [0,1] \subset \mathbb{R}$ and introduce the function $f : A \times B \to \mathbb{R}$ given by

$$f(\mathbf{a},\mathbf{b}) = f((a_k),\mathbf{b}) = \begin{cases} 1 \text{ if there exist some } k \in \mathbb{N} \text{ such that } \mathbf{b} = a_k \\ 0 \text{ otherwise .} \end{cases} \tag{26}$$

One can now show the following lemma.

Lemma 18 *The function f listed in relation (26) belongs to $\mathcal{L}^1_{\lambda\otimes\mu}$ for every $\lambda \in \mathcal{P}(A)$ and $\mu \in \mathcal{P}(B)$.*

Proof. Since the function f is bounded it is sufficient to verify that f is Borel measurable. Clearly the set $A \times B$ is closed in $C_0 \times R$ and so it is Borel-measurable. To show that the function f is Borel-measurable on $A \times B$ it is sufficient to check that the set

$$\begin{aligned} S &= \{(\mathbf{a}, \mathbf{b}) \in A \times B : f(\mathbf{a}, \mathbf{b}) = 1\} \\ &= \{(\mathbf{a}, \mathbf{b}) \in A \times B : \text{there exist some } k \in \mathbb{N} \text{ such that } \mathbf{b} = a_k\} \end{aligned}$$

is measurable. Its complementary set will then be measurable by the definition of a σ-algebra. To verify that S is Borel measurable we show that it is closed. Let $(\mathbf{a}^n, \mathbf{b}^n)$ be an arbitrary sequence in S converging to $(\mathbf{a}, \mathbf{b})$. We have to prove that there exists some $k \in \mathbb{N}$ such that $\mathbf{b} = a_k$. By our assumption $\mathbf{a}^n \to \mathbf{a}$ in C_0 and $\mathbf{b}^n \to \mathbf{b}$ in $\mathbb{R}$ and so

$$\lim_{n\uparrow\infty} ||\mathbf{a}^n - \mathbf{a}|| = \lim_{n\uparrow\infty} \sup_{k\in N} |a_k^n - a_k| = 0.$$

Since for each $n \in \mathbb{N}$ there exists some k such that $\mathbf{b}^n = a_k^n$ consider for each fixed $n \in \mathbb{N}$ the smallest index $k(n)$ satisfying $\mathbf{b}^n = a_{k(n)}^n$. Due to

$$|\mathbf{b}^n - a_{k(n)}| = |a_{k(n)}^n - a_{k(n)}| \leq ||\mathbf{a}^n - \mathbf{a}|| \to 0 \text{ if } n \to \infty$$

it follows that

$$|\mathbf{b} - a_{k(n)}| \leq |\mathbf{b} - \mathbf{b}^n| + |\mathbf{b}^n - a_{k(n)}| \to 0 \text{ if } n \to \infty. \qquad (27)$$

We now distinguish the following two cases: If the sequence $(k(n))_{n\in N}$ is bounded and so it takes only a finite number of distinct values there exists a constant subsequence $(k(n_i))_{i\in\mathbb{N}}$ with $n_1 < n_2 < n_3 < ...$ of the sequence $(k(n))_{n\in\mathbb{N}}$. This means that $k(n_i) = k_0$ for every $i \in \mathbb{N}$ and so $a_{k(n_i)} = a_{k_0}$ for every $i \in \mathbb{N}$. Hence by relation (27) we obtain that $a_{k_0} = \mathbf{b}$ and so the vector $(\mathbf{a}, \mathbf{b})$ belongs to S. If, on the other hand, the sequence $(k(n))_{n\in N}$ is unbounded, i.e. $\lim_{n\to\infty} k(n) = \infty$, there exists a strictly increasing subsequence $(k(n_i))_{i\in\mathbb{N}}$ of $(k(n))_{n\in\mathbb{N}}$, i.e.

$$k(n_1) < k(n_2) < k(n_3) < ...$$

Again by relation (27) and $\mathbf{a}$ belongs to the Banach space C_0 we obtain that $\lim_{i\to\infty} a_{k(n_i)} = \mathbf{b} = 0$. Since by the definition of set A we know that $a_1^n = 0$ for every $n \in N$, it follows that $a_1 = 0$ and so $(\mathbf{a}, \mathbf{b})$ belongs to S. This completes the proof of the lemma. $\square$

We will now list the counterexample for which (7) holds and (8) does not hold

Example 19 *Let $f : A \times B \to \mathbb{R}$ be the function defined in relation (26) and consider some $\lambda \in \mathcal{P}_F(A)$. Hence there exists a finite number of sequences $\mathbf{a}^i = (a_k^i)_{k\in\mathbb{N}}, 1 \leq i \leq m$, belonging to A and some vector $\mathbf{s}(\lambda) = (s_1(\lambda), ..., s_m(\lambda))$, $s_i(\lambda) > 0$ and $\sum_{i=1}^m s_i(\lambda) = 1$ such that*

$$\lambda = \sum_{i=1}^m s_i(\lambda)\epsilon_{\mathbf{a}^i}.$$

Since the set $[0,1]$ contains more than a countable number of elements one can now choose a number $\mathbf{b} \in [0,1]$ such that ***none*** *of the above sequences $\mathbf{a}^i, 1 \leq i \leq m$, contain this number. Using this number and the definition of f it can be easily seen that*

$$\inf_{\mathbf{b}\in[0,1]} f_e(\lambda, \epsilon_{\mathbf{b}}) = \inf_{\mathbf{b}\in[0,1]} \sum_{i=1}^{m} s_i(\lambda) f(\mathbf{a}^i, \mathbf{b}) = 0$$

and so $R_8 = 0$. On the other hand, consider some $\mu \in \mathcal{P}_F(B)$. By definition one can find some finite set $\{\mathbf{b}_1, ...\mathbf{b}_p\} \subseteq [0,1]$ and a vector $\mathbf{s}(\mu) = (s_1(\mu), ..., s_p(\mu))$, $s_j(\mu) > \mathbf{0}$ and $\sum_{j=1}^{p} s_j(\mu) = 1$ such that

$$\mu = \sum_{j=1}^{p} s_j(\mu)\epsilon_{\mathbf{b}_j}.$$

Introducing the element $\mathbf{a}_0 := (0, \mathbf{b}_1, ..., \mathbf{b}_p, 0, 0, ...) \in C_0$ it is obvious by the definition of f that

$$\begin{aligned}\sup_{\mathbf{a}\in A} f_e(\epsilon_{\mathbf{a}}, \mu) &= \sup_{\mathbf{a}\in A} \sum_{j=1}^{p} s_j(\mu) f(\mathbf{a}, \mathbf{b}_j) \\ &\geq \sum_{j=1}^{n} s_j(\mu) f(\mathbf{a}_0, \mathbf{b}_j) = 1.\end{aligned}$$

Since f is bounded by 1 this shows that

$$L_8 := \inf_{\mu\in\mathcal{P}_F(B)} \sup_{\mathbf{a}\in A} f_e(\epsilon_{\mathbf{a}}, \mu) = 1$$

and so we have verified that (8) does not hold. To see that (7) holds, observe that $R_7 = R_8 = 0$ and let μ_0 be the Lebesgue measure on $[0,1]$. Obviously $\mu_0 \in \mathcal{P}(B)$ and since for every $\mathbf{a} \in A$ the function $f(\mathbf{a}, .)$ takes the value 1 on a countable set and zero elsewhere and by Lemma 18 f belongs to $\mathcal{L}^1_{\lambda\otimes\mu}$ for every $\lambda \in \mathcal{P}(A)$ and $\mu \in \mathcal{P}(B)$, we obtain

$$\int_0^1 f(\mathbf{a}, \mathbf{b}) d\mu_0(\mathbf{b}) = 0$$

for every $\mathbf{a} \in A$. Hence it follows that $L_8 = 0$ and so (8) holds.

We now list an instance for which (6) holds and (7) does not hold.

Example 20 *Let $A := [0,1]$ and $B := \{(b_k)_{k\in\mathbb{N}} \in C_0 : b_1 = 0\}$ and introduce the function $f : A \times B \to \mathbb{R}$ given by*

$$f(\mathbf{a}, \mathbf{b}) = \begin{cases} 0 \text{ if there exist some } k \in \mathbb{N} \text{ such that } \mathbf{a} = b_k \\ 1 \text{ otherwise.} \end{cases}.$$

As in Example 19 one can verify for every $\lambda \in \mathcal{P}_F(A)$ that

$$\inf_{\mathbf{b}\in B} f_e(\lambda, \epsilon_{\mathbf{b}}) = 0$$

and so $R_7 = 0$. On the other hand, by Lemma 18 the function f is Borel measurable and if λ_0 is the Lebesgue measure on $[0, 1]$ we obtain as before that

$$\int_0^1 f(\mathbf{a}, \mathbf{b}) d\lambda_0(\mathbf{a}) = 1. \tag{28}$$

for every $\mathbf{b} \in B$. Also it is easy to verify by a similar argument as used in Example 19 that

$$\sup_{\mathbf{a}\in A} f_e(\epsilon_{\mathbf{a}}, \mu) = \sup_{\mathbf{a}\in[0,1]} \sum_{j=1}^{p} s_j(\mu) f(\mathbf{a}, \mathbf{b}^j) = 1 \tag{29}$$

for every $\mu \in \mathcal{P}_F(B)$. Using now relations (28) and (29) we obtain that

$$1 = L_6 \geq R_6 \geq 1$$

and so (6) holds. Moreover, since $R_7 = 0$ and $L_7 = L_6 = 1$ it follows that (7) does not hold.

The above examples showed that none of the implications in relation (25) can be reversed. To conclude this section we give an example which shows that (6) can also fail.

Example 21 *Let $A = B := [0, \infty) \subset \mathbb{R}$ and consider the function $f : A \times B \to \mathbb{R}$ given by*

$$f(\mathbf{a}, \mathbf{b}) = \begin{cases} 1 \textit{ if } \mathbf{a} \geq \mathbf{b} \\ 0 \textit{ otherwise}. \end{cases}$$

For any $\lambda \in \mathcal{P}(A)$ it follows that

$$\int_0^\infty f(\mathbf{a}, \mathbf{b}) d\lambda(\mathbf{a}) = \lambda([\mathbf{b}, \infty)) = 1 - \lambda([0, \mathbf{b}))$$

for every $\mathbf{b} \geq 0$ and so we obtain that $R_6 = 0$. On the other hand, for any $\mu \in \mathcal{P}(B)$ we observe that

$$\int_0^\infty f(\mathbf{a}, \mathbf{b}) d\mu(\mathbf{b}) = \mu([0, \mathbf{a}))$$

for every $\mathbf{a} \geq 0$ and so it follows that $L_6 = 1$. Hence (6) does not hold.

In the next section we apply the minimax results derived in the previous sections to Lagrangian duality.

3 Application to Lagrangian duality.

Before applying the results of the first section to the Lagrangian dual problem we first need to introduce some well-known notions. Let Y be a normed linear space and $K \subseteq Y$ some closed convex cone. Introduce now on Y the partial ordering $\leq_K$ defined by

$$\mathbf{y}_1 \leq_K \mathbf{y}_2 \Longleftrightarrow \mathbf{y}_2 - \mathbf{y}_1 \in K.$$

If Y^* denotes the topological dual space of Y, let $K^* \subseteq Y^*$ be the so-called dual cone given by

$$K^* := \{\mathbf{y}^* \in Y^* :< \mathbf{y}^*, \mathbf{y} >\geq 0 \text{ for every } \mathbf{y} \in K\}$$

with $< \mathbf{y}^*, \mathbf{y} >:= \mathbf{y}^*(\mathbf{y})$. This means that

$$K^* = \{\mathbf{y}^* \in Y^* :< \mathbf{y}^*, \mathbf{y} >\geq 0 \text{ for every } \mathbf{y} \geq_K \mathbf{0}\}$$

and so the dual cone K^* denotes the space of all continuous positive linear functionals on Y. If X is some topological space and $h : X \to \mathbb{R}$ and $g : X \to Y$ are some given mappings, consider then for the nonempty feasible region

$$D := \{\mathbf{x} \in X : g(\mathbf{x}) \geq_K \mathbf{0}\}$$

the general primal optimization problem

$$v(\text{P}) := \sup\{h(\mathbf{x}) : \mathbf{x} \in D\}. \tag{P}$$

To derive the Lagrangian dual of the optimization problem (P) we first introduce the so-called Lagrangian function $\theta : K^* \to (-\infty, \infty]$ given by

$$\theta(\mathbf{y}^*) := \sup_{\mathbf{x} \in X}\{h(\mathbf{x})+ < \mathbf{y}^*, g(\mathbf{x}) >\}$$

Since it is easy to verify that

$$\theta(\mathbf{y}^*) \geq v(P) \tag{30}$$

for every $\mathbf{y}^* \in K^*$ and we like to approximate $v(P)$ by means of the Lagrangian function, it is natural to consider the so-called Lagrangian dual given by

$$v(\text{D}) := \inf_{\mathbf{y}^* \in K^*} \theta(\mathbf{y}^*). \tag{D}$$

By relation (30) it is clear that $v(\text{D}) \geq v(\text{P})$ and in the remainder of this section we are interested under which necessary and sufficient conditions we actually have an equality. Whether or not one has an equality (no duality gap) plays a cental role in the theory of optimization and so a lot of papers and books have discussed this topic. In this section we will also pursue this question and although most of the sufficient conditions are already known we like to stress that there are virtually no papers in the literature trying to derive necessary and sufficient conditions. Using now the minimax approach and imposing for noncompact sets X the well-known Slater type regularity condition, it is possible to give a necessary and sufficient condition for equality of the optimal objective value of the primal and dual problem. The same holds for compact sets X without this regularity condition. Moreover, we show that the Slater type condition is actually equivalent to the inf-compactness of the Lagrangian bifunction and hence this regularity condition is nothing else than a compactness type condition. To start with the analysis of the Lagrangian dual and its relation to the primal problem, we first give an alternative expression for $v(P)$.

Lemma 22 *If the function $f : X \times K^* \to \mathbb{R}$ is given by*

$$f(\mathbf{x}, \mathbf{y}^*) := h(\mathbf{x}) + < \mathbf{y}^*, g(\mathbf{x}) >, \tag{31}$$

then it follows that

$$v(P) = \sup_{\mathbf{x} \in \mathbf{X}} \inf_{\mathbf{y}^* \in K^*} f(\mathbf{x}, \mathbf{y}^*).$$

Proof. If the vector $\mathbf{x}$ belongs to the set D, then clearly $< \mathbf{y}^*, g(\mathbf{x}) > \geq 0$ for every $\mathbf{y}^*$ belonging to K^* and so we obtain

$$\inf_{\mathbf{y}^* \in K^*} f(\mathbf{x}, \mathbf{y}^*) = h(\mathbf{x}). \tag{32}$$

Moreover, since K is a closed convex cone, we may apply the bipolar theorem given by $K^{**} = K$ (cf.[9]) and so for $\mathbf{x}$ belonging to $X \backslash D$ the bipolar theorem implies $g(\mathbf{x}) \notin K^{**}$. Hence there exists some $\mathbf{y}_0^* \in K^*$ satisfying $< \mathbf{y}_0^*, g(\mathbf{x}) > < 0$ and since $\alpha \mathbf{y}_0^* \in K^*$ for every $\alpha > 0$ this implies that

$$\inf_{\mathbf{y}^* \in K^*} f(\mathbf{x}, \mathbf{y}^*) = -\infty. \tag{33}$$

Since the set D is nonempty we know $v(P) > -\infty$ and this implies by relations (32) and (33) that

$$v(P) = \sup_{\mathbf{x} \in D} \inf_{\mathbf{y}^* \in K^*} f(\mathbf{x}, \mathbf{y}^*) = \sup_{x \in X} \inf_{\mathbf{y}^* \in K^*} f(\mathbf{x}, \mathbf{y}^*)$$

showing the desired result. □

By Lemma 22 and the definition of the Lagrangian dual problem (D) it follows that there exists no duality gap if and only if the minimax result in relation (10) holds with A replaced by X and B by K^*. For the bifunction $f : X \times K^* \to \mathbb{R}$, listed in relation (31), one can now show the following result.

Lemma 23 *It follows for every $J \in \mathcal{F}(X)$ and the bifunction $f : X \times K^* \to \mathbb{R}$ given by relation (31) that*

$$\inf_{\mu \in \mathcal{P}_F(K^*)} \max_{x \in J} f_e(\epsilon_{\mathbf{x}}, \mu) = \inf_{\mathbf{y}^* \in K^*} \max_{\mathbf{x} \in J} f(\mathbf{x}, \mathbf{y}^*).$$

Proof. For every μ belonging to $\mathcal{P}_F(K^*)$ there exists by definition some finite set $\{\mathbf{y}_1^*, ..., \mathbf{y}_p^*\} \subseteq K^*$ and a vector $\mathbf{s}(\mu) = (s_1(\mu), ..., \mathbf{s}_p(\mu))$ such that

$$\mu = \sum_{j=1}^{p} s_j(\mu)\epsilon_{\mathbf{y}_j^*}, s_j(\mu) > 0, \sum_{j=1}^{p} s_j(\mu) = 1.$$

This yields for every J belonging to $\mathcal{F}(X)$ and f given by relation (31) that

$$\max_{\mathbf{x}\in J} f_e(\epsilon_{\mathbf{x}}, \mu) = \max_{\mathbf{x}\in J} f_e(\epsilon_{\mathbf{x}}, \sum_{j=1}^{p} s_j(\mu)\mathbf{y}_j^*). \tag{34}$$

Since the dual cone $K^* \subseteq Y^*$ is convex this implies that $\sum_{j=1}^{p} s_j(\mu)\mathbf{y}_j^* \in K^*$ and hence we obtain by relation (34) that

$$\max_{\mathbf{x}\in J} f_e(\epsilon_{\mathbf{x}}, \mu) \geq \inf_{\mathbf{y}^*\in K^*} \max_{\mathbf{x}\in J} f(\mathbf{x}, \mathbf{y}^*).$$

This shows

$$\inf_{\mu\in\mathcal{P}_F(K^*)} \max_{\mathbf{x}\in J} f_e(\epsilon_{\mathbf{x}}, \mu) \geq \inf_{\mathbf{y}^*\in K^*} \max_{\mathbf{x}\in J} f(\mathbf{x}, \mathbf{y}^*).$$

and since the reverse inequality trivially holds the desired equality follows. □

To show that under some additional assumption the function f, listed in relation (31), actually belongs to the set $\mathcal{C}$ it is by Lemma 23 sufficient and necessary to show that

$$\sup_{J\in\mathcal{F}(X)} \inf_{\mathbf{y}^*\in K^*} \max_{\mathbf{x}\in J} f(\mathbf{x}, \mathbf{y}^*) = \inf_{\mathbf{y}^*\in K^*} \sup_{\mathbf{x}\in X} f(\mathbf{x}, \mathbf{y}^*). \tag{35}$$

To verify this, we need to check the conditions of Lemma 4 and so we have to introduce a convenient topology on the set Y^*. As we shall see later the strong topology on Y^* generated by the operator norm

$$\|.\|_d := \sup_{\|\mathbf{y}\|=1} | < \mathbf{y}^*, \mathbf{y} > |$$

is not suitable. The proper topology to define on Y^* is now given by the weak*topology. Remember in the weak*topology on Y^* the neighborhood base of zero is given by sets of the form

$$\mathcal{N}(\mathbf{y}_1, ..., \mathbf{y}_k, \epsilon) := \{\mathbf{y}^* \in Y^* : | < \mathbf{y}^*, \mathbf{y}_i > | < \epsilon, 1 \leq i \leq k\} \tag{36}$$

with $\epsilon > 0$ and $\{\mathbf{y}_1, ..., \mathbf{y}_k\}$ some finite subset of Y. It is also well-known that the net $\{\mathbf{y}_i^*\}_{i\in I} \subseteq Y^*$ converges in the weak*topology to $\mathbf{y}^* \in Y^*$ (notation $\mathbf{y}_i^* \to^* \mathbf{y}^*$) if and only if

$$\lim_{i\in I} < \mathbf{y}_i^*, \mathbf{y} > = < \mathbf{y}^*, \mathbf{y} > \tag{37}$$

for every $\mathbf{y} \in Y$. Using the weak*topology it is obvious by relation (37) that the function $f(\mathbf{x}, .) : K^* \to \mathbb{R}$, given by

$$f(\mathbf{x}, \mathbf{y}^*) := h(\mathbf{x}) + < \mathbf{y}^*, g(\mathbf{x}) > \tag{38}$$

is continuous (in the weak*topology) for every $\mathbf{x} \in X$. Hence to apply Lemma 4 and show that f belongs to $\mathcal{C}$ we still need to verify the inf-compactness property. To check this property we introduce the following well-known regularity condition.

Definition 24 *A point $\mathbf{x}_0$ is called a Slater point of the feasible region $D := \{\mathbf{x} \in X : g(\mathbf{x}) \geq_K \mathbf{0}\}$ if $\mathbf{x}_0 \in X$ and $g(\mathbf{x}_0) \in int(K)$.*

As shown by the following result the existence of a Slater point $\mathbf{x}_0$ of the set D is the same as the inf-compactness (in the weak*topology) of the function $f_{\mathbf{x}_0}$.

Lemma 25 *The point $\mathbf{x}_0$ is a Slater point of the set D if and only if the function $f(\mathbf{x}_0, .) : K^* \to \mathbb{R}$, given by*

$$f(\mathbf{x}_0, \mathbf{y}^*) = h(\mathbf{x}_0) + < \mathbf{y}^*, g(\mathbf{x}_0) >,$$

is inf-compact (in the weak*topology).

Proof. We need to show for every $r \in \mathbb{R}$ that the set $L(r) := \{\mathbf{y}^* \in K^* : f_{\mathbf{x}_0}(\mathbf{y}^*) \leq r\}$ is compact in the weak*topology. Since $g(\mathbf{x}_0)$ belongs to $int(K)$ one can find some $\epsilon > 0$ such that

$$g(\mathbf{x}_0) + \mathcal{N}_1(\epsilon) \subseteq K \tag{39}$$

with $\mathcal{N}_1(\epsilon) := \{\mathbf{y} \in Y : \|\mathbf{y}\| \leq \epsilon\}$. Consider now some $\mathbf{y}^* \in K^*$. Since $\|\mathbf{y}^*\|_d := \sup_{\|\mathbf{y}\|=1} | < \mathbf{y}^*, \mathbf{y} > |$ there exists some $\mathbf{y}_0 \in Y$ satisfying

$$\|\mathbf{y}_0\| = 1 \text{ and} < \mathbf{y}^*, \mathbf{y}_0 > \geq \frac{1}{2}\|\mathbf{y}^*\|_d. \tag{40}$$

This implies by relation (39) and (40) that

$$< \mathbf{y}^*, g(\mathbf{x}_0) > = < \mathbf{y}^*, g(\mathbf{x}_0) - \epsilon\mathbf{y}_0 > + \epsilon < \mathbf{y}^*, \mathbf{y}_0 > \geq \epsilon\|\mathbf{y}^*\|_d \tag{41}$$

and so we obtain for every $\mathbf{y}^*$ belonging to $L(r)$ that

$$\epsilon\|\mathbf{y}^*\|_d \leq < \mathbf{y}^*, g(\mathbf{x}_0) > \leq r - h(\mathbf{x}_0).$$

Hence we have shown that

$$L(r) \subseteq \{\mathbf{y}^* \in K^* : \|\mathbf{y}^*\|_d \leq \epsilon^{-1}(r - h(\mathbf{x}_0))\} \tag{42}$$

and since by Alaoglu's theorem (cf.[8]) the last set in relation (42) is weak*compact and $L(r)$ is weak*closed we obtain that the set $L(r)$ is weak*compact. To show the reverse implication, let $f(\mathbf{x}_0, .)$ be inf-compact (with respect to the weak*topology) and take $r := h(\mathbf{x}_0) + \|g(\mathbf{x}_0)\|$. Observe now for every $\mathbf{y}^*$ belonging to K^* and satisfying $\|\mathbf{y}^*\|_d \leq 1$ that

$$f(\mathbf{x}_0, \mathbf{y}^*) \leq h(\mathbf{x}_0) + \|\mathbf{y}^*\|_d \|g(\mathbf{x}_0)\| \leq r$$

and so it follows that

$$\{\mathbf{y}^* \in K^* : \|\mathbf{y}^*\|_d \leq 1\} \subseteq L(r). \tag{43}$$

Assume now by contradiction that there exists some nonzero $\mathbf{y}_0^* \in L(r)$ satisfying $< \mathbf{y}_0^*, g(\mathbf{x}_0) > \leq 0$ and so by the definition of $L(r)$ we obtain that $\alpha \mathbf{y}_0^* \in L(r)$ for every $\alpha > 1$. Since $\mathbf{y}_0^* \neq \mathbf{0}$ there exists some $\mathbf{y}_0 \in Y$ such that $< \mathbf{y}_0^*, \mathbf{y}_0 > \neq 0$ and consider now for this $\mathbf{y}_0$ the open set $\mathcal{N}(\mathbf{y}_0, 1) \subseteq Y^*$ containing $\mathbf{0}^{*}$. Since the vector space Y^* equipped with the weak*topology is a topological vector space and by assumption the set $L(r)$ is weak*compact it follows by part b of Theorem 1.15 of [18] that the lower level set $L(r)$ is bounded. Since $\alpha \mathbf{y}_0^* \in L(r)$ for every $\alpha > 1$ we obtain using $\mathbf{y}_0^* \neq 0$ that $\|\alpha \mathbf{y}_0^*\|_d = \alpha \|\mathbf{y}_0^*\|_d \uparrow \infty$ if $\alpha \uparrow \infty$ and this contradicts the boundedness of $L(r)$. Hence for every nonzero $\mathbf{y}^* \in L(r)$ it follows that $< \mathbf{y}^*, g(\mathbf{x}_0) >> 0$ and by relation (43) we obtain that $< \mathbf{y}^*, g(\mathbf{x}_0) >> 0$ for every $\mathbf{y}^* \in K^* \backslash \{\mathbf{0}^*\}$. This shows (cf.[10]) that $g(\mathbf{x}_0)$ belongs to int(K) and so $\mathbf{x}_0$ is a Slater point of the set D. □

Using now Lemma 25 and Theorem 15 one can verify the following important result.

Theorem 26 *If the set D contains a Slater point, then it follows that $v(D) = v(P)$ if and only if the function f, given by relation (31), belongs to $\mathcal{D}$, i.e*

$$\sup_{\lambda \in \mathcal{P}_F(X)} \inf_{\mathbf{y}^* \in K^*} f_e(\lambda, \epsilon_{\mathbf{y}^*}) = \sup_{\mathbf{x} \in X} \inf_{\mathbf{y}^* \in K^*} f(\mathbf{x}, \mathbf{y}^*).$$

Moreover, the dual problem (D) has an optimal solution.

Proof. By Lemma 25 and using $f_{\mathbf{x}}(\mathbf{y}^*) := h(\mathbf{x}) + < \mathbf{y}^*, g(\mathbf{x}) >$ is continuous in the weak*topology we obtain that the conditions of Lemma 4 are satisfied and so relation (35) holds. Hence the function f, listed in relation (31), belongs to the set $\mathcal{C}$. Applying now Theorem 15 yields $v(\mathrm{D}) = v(\mathrm{P})$ if and only if $f \in \mathcal{D}$. Actually, by the inf-compactness of $f_{\mathbf{x}_o}$ with $\mathbf{x}_0$ the Slater point, it also holds by Lemma 4 that

$$\inf_{\mathbf{y}^* \in K^*} \sup_{\mathbf{x} \in X} f(\mathbf{x}, \mathbf{y}^*) = \min_{\mathbf{y}^* \in K} \sup_{\mathbf{x} \in X} f(\mathbf{x}, \mathbf{y}^*)$$

and this shows that the dual problem has an optimal solution. □

In the next example we will consider an important class of optimization problems for which the Lagrangian dual can be simplified.

Example 27 *Let X be a normed linear space with $L \subseteq X$ some closed linear subspace, $\mathbf{b} \in X$ and $K \subseteq X$ some closed convex cone and consider the conic convex programming problem given by*

$$v(CP) := \sup\{< \mathbf{x}_0^*, \mathbf{x} >: \mathbf{x} \in K \cap (L + \mathbf{b})\} \tag{CP}$$

with $\mathbf{x}_0^*$ *some element of the topological dual space* X^* *of* X *and* $K \cap (L + \mathbf{b})$ *nonempty (for the finite dimensional version of a conic convex programming problem, see [14]). Since* L *is a closed linear subspace and hence a closed convex cone, it follows that a conic convex optimization problem can be written as*

$$v(CP) = \sup\{< \mathbf{x}_0^*, \mathbf{x} >: \mathbf{x} \in D\} \text{ and } D := \{\mathbf{x} \in K : \mathbf{x} - \mathbf{b} \geq_L \mathbf{0}\}.$$

Since L *is a linear subspace it is easy to verify that*

$$L^* = \{\mathbf{x}^* \in X^* :< \mathbf{x}^*, \mathbf{x} > = 0 \text{ for every } \mathbf{x} \in L\}$$

and the space L^* *is mostly denoted in the literature by* $L^\perp$*. The Lagrangian function* $\theta : L^\perp \to (-\infty, \infty]$ *is now given by*

$$\begin{aligned} \theta(\mathbf{x}^*) &= \sup_{\mathbf{x} \in K}\{< \mathbf{x}_0^*, \mathbf{x} > + < \mathbf{x}^*, \mathbf{x} - \mathbf{b} >\} \\ &= - < \mathbf{x}^*, \mathbf{b} > + \sup_{\mathbf{x} \in K} < \mathbf{x}_0^* + \mathbf{x}^*, \mathbf{x} > . \end{aligned}$$

To analyse $\sup_{\mathbf{x} \in K} < \mathbf{x}_0^* + \mathbf{x}^*, \mathbf{x} >$ *we observe the following. If* $\mathbf{x}_0^* + \mathbf{x}^* \notin -K^*$ *there exists some* $\mathbf{x}_0 \in K$ *such that* $< \mathbf{x}_0^* + \mathbf{x}^*, \mathbf{x}_0 >> 0$ *and using* $\alpha \mathbf{x}_0 \in K$ *for every* $\alpha > 0$ *this implies that*

$$\sup_{\mathbf{x} \in K} < \mathbf{x}_0^* + \mathbf{x}^*, \mathbf{x} >= \infty.$$

Moreover, if $\mathbf{x}_0^* + \mathbf{x}^* \notin -K^*$ *it is obvious that* $\sup_{\mathbf{x} \in K} < \mathbf{x}_0^* + \mathbf{x}^*, \mathbf{x} >= 0$ *and so we obtain*

$$\sup_{\mathbf{x} \in K} < \mathbf{x}_0^* + \mathbf{x}^*, \mathbf{x} >= \begin{cases} 0 & \text{if } \mathbf{x}_0^* + \mathbf{x}^* \in -K^* \\ \infty & \text{otherwise} \end{cases} .$$

This shows

$$\theta(\mathbf{x}^*) = \begin{cases} - < \mathbf{x}^*, \mathbf{b} > & \text{if } \mathbf{x}_0^* + \mathbf{x}^* \in -K^* \\ \infty & \text{otherwise} \end{cases}$$

and we have shown that for the conic convex programming problem (CP) *the Lagrangian dual problem* (D) *has the form*

$$v(D) = \inf\{- < \mathbf{x}^*, \mathbf{b} >: \mathbf{x}_0^* + \mathbf{x}^* \in -K^*, \mathbf{x}^* \in L^\perp\}.$$

Since $L^\perp = -L^\perp$ *this reduces to*

$$v(D) = \inf\{< \mathbf{x}^*, \mathbf{b} >: \mathbf{x}^* \in L^\perp \cap (K^* + \mathbf{x}_0^*)\}.$$

Clearly the dual decision variables $\mathbf{x}^*$ *in the dual problem belong to the topological dual* X^* *of* X*. To simplify this dual problem we assume that the set* X *is a real Hilbert space. Since it is well-known that any continuous linear functional* $\mathbf{x}^*$ *on a real Hilbert space* X *can be written as*

$$< \mathbf{x}^*, \mathbf{x} > = (\mathbf{c}, \mathbf{x})$$

for some $\mathbf{c} \in X$ *with* $(.,.)$ *denoting the inner product on the real Hilbert space (cf.[13]) it follows that a conic convex programming problem on* X *has the form*

$$\sup\{(\mathbf{c}, \mathbf{x}) : \mathbf{x} \in K \cap (L + \mathbf{b})\} \qquad \text{(HCP)}$$

with $\mathbf{c} \in X$*. The associated Lagrangian dual is then given by*

$$\inf\{(\mathbf{b}, \mathbf{x}) : \mathbf{x} \in L^{\perp} \cap (K^* + \mathbf{c})\}.$$

For a Hilbert space X *the sets* $L^{\perp}$ *and* K^* *are given by*

$$L^{\perp} = \{\mathbf{x} \in X : (\mathbf{x}, \mathbf{c}) = 0 \textit{ for every } \mathbf{c} \in L\}$$

and

$$K^* = \{\mathbf{x} \in X : (\mathbf{x}, \mathbf{c}) \geq 0 \textit{ for every } \mathbf{c} \in K\}.$$

Hence in this case the dual is defined on the original space and a special instance of optimization problem (HCP) *is now given by a so-called positive semidefinite programming problem defined on the Hilbert space of all* $n \times n$ *symmetric real valued matrices equipped with the Frobenius norm*

$$\|A\|_F := \sqrt[2]{tr(AA^{\top})}$$

with $tr(AB) := \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} b_{ij}$ *(cf.[5]). In this case the set* K *is given by the set all symmetric positive semidefinite matrices and the dual cone* K^* *of this set is again the set of all symmetric positive semidefinite matrices (cf.[5]).*

In case we do not assume that there exists a Slater point one can still come up with a necessary and sufficient condition for the absense of a duality gap. As before (reverse the roles of X and K^*) we introduce the bifunction $f : K^* \times X \to \mathbb{R}$ given by

$$f(\mathbf{y}^*, \mathbf{x}) := h(\mathbf{x}) + < \mathbf{y}^*, g(\mathbf{x}) > . \qquad (44)$$

It is now easy to show the following result.

Lemma 28 *The function* $-f : K^* \times X \to \mathbb{R}$ *with* f *listed in relation (44), belongs to the set* $\mathcal{D}$*.*

Proof. By the definition of the set $\mathcal{D}$ we need to show that

$$\inf_{\lambda \in \mathcal{P}_F(K^*)} \sup_{\mathbf{x} \in X} f_e(\lambda, \epsilon_{\mathbf{x}}) = \inf_{\mathbf{y}^* \in K^*} \sup_{\mathbf{x} \in X} f(\mathbf{y}^*, \mathbf{x}).$$

Observe, if λ belongs to $\mathcal{P}_F(K^*)$, there exists by definition some finite set $\{\mathbf{y}_1^*, ..., \mathbf{y}_m^*\} \subseteq K^*$ and a vector $\mathbf{s}(\mu) = (s_1(\mu), ..., s_m(\mu))$ such that

$$\lambda = \sum_{i=1}^{m} s_i(\lambda) \epsilon_{\mathbf{y}_i^*}, s_i(\lambda) > 0, \sum_{i=1}^{m} s_i(\lambda) = 1.$$

Since K^* is a convex cone we obtain that $\sum_{i=1}^m s_i(\lambda)\mathbf{y}_i^*$ belongs to K^* and this implies with f given by relation (44) that

$$\sup_{\mathbf{x}\in X} f_e(\lambda, \epsilon_{\mathbf{x}}) = \sup_{\mathbf{x}\in X} f(\sum_{i=1}^m s_i(\lambda)\mathbf{y}_i^*, \mathbf{x}). \qquad (45)$$

Hence by relation (45) we obtain

$$\sup_{\mathbf{x}\in X} f_e(\lambda, \epsilon_{\mathbf{x}}) \geq \inf_{\mathbf{y}^*\in K^*} \sup_{\mathbf{x}\in X} f(\mathbf{y}^*, \mathbf{x})$$

and so

$$\inf_{\lambda\in\mathcal{P}_F(K^*)} \sup_{\mathbf{x}\in X} f_e(\lambda, \epsilon_{\mathbf{x}}) \geq \inf_{\mathbf{y}^*\in K^*} \sup_{\mathbf{x}\in X} f(\mathbf{y}^*, \mathbf{x}).$$

This shows the desired result. □

An immediate consequence of Theorem 15 and Lemma 28 is given by the following result. Observe in this result we do not assume the existence of a Slater point or the compactness of the set X.

Theorem 29 *It follows that $v(D) = v(P)$ if and only if the function f given by relation (44) satisfies*

$$\inf_{J\in\mathcal{F}(K^*)} \sup_{\mu\in\mathcal{P}_F(X)} \inf_{\mathbf{y}^*\in J} f_e(\epsilon_{\mathbf{y}^*}, \mu) = \sup_{\mathbf{x}\in X} \inf_{\mathbf{y}^*\in K^*} f(\mathbf{y}^*, \mathbf{x}).$$

Proof. The above equality means that $-f$ belongs to the set $\mathcal{C}$. By Lemma 22 we know that

$$v(P) = \sup_{\mathbf{x}\in \mathbf{X}} \inf_{\mathbf{y}^*\in K^*} f(\mathbf{y}^*, \mathbf{x})$$

and so the above result is a consequence of Theorem 15 and Lemma 28. □

Using Theorem 29 one can show the following important result.

Theorem 30 *If the set X is a compact Hausdorff space and the functions $f(\mathbf{y}^*, .), \mathbf{y}^* \in K^*$ with f listed in relation (44) are upper semicontinuous and belong to $\mathcal{L}^1_\mu(X)$ for every $\mu \in \mathcal{P}(X)$, then it follows that $v(P) = v(D)$ if and only if*

$$\sup_{\mu\in\mathcal{P}(X)} \inf_{\mathbf{y}^*\in K^*} f_e(\epsilon_{\mathbf{y}^*}, \mu) = \sup_{\mathbf{x}\in X} \inf_{\mathbf{y}^*\in K^*} f(\mathbf{y}^*, \mathbf{x}).$$

Proof. Since $f(\mathbf{y}^*, .)$ belongs to $\mathcal{L}^1_\mu(X)$ for every $\mu \in \mathcal{P}(X)$ we obtain by relation (22) that

$$\sup_{\mu\in\mathcal{P}_F(X)} \inf_{\mathbf{y}^*\in J} f_e(\epsilon_{\mathbf{y}^*}, \mu) = \sup_{\mu\in\mathcal{P}(X)} \inf_{\mathbf{y}^*\in J} f_e(\epsilon_{\mathbf{y}^*}, \mu) \qquad (46)$$

for every $J \in \mathcal{F}(X)$. Since X is a compact Hausdorff space it follows (cf.[3]) that $\mathcal{P}(X)$ is compact in the weak*topology. Moreover, due to the upper semicontinuity of the functions $f(\mathbf{y}^*, .)$, it can be shown (cf.[6], [3]) that the function $\mu \to f_e(\epsilon_{\mathbf{y}^*}, \mu)$ is upper semicontinuous in the weak*topology for every $\mathbf{y} \in K^*$. Applying now relation (20) with A replaced by $\mathcal{P}(X)$, B by K^* and the function f by $f_e(\mathbf{y}^*, \mu)$ it follows using also relation (46) that

$$\inf_{J \in \mathcal{F}(K^*)} \sup_{\mu \in \mathcal{P}_F(X)} \inf_{\mathbf{y}^* \in J} f_e(\epsilon_{\mathbf{y}^*}, \mu) = \sup_{\mu \in \mathcal{P}(X)} \inf_{\mathbf{y}^* \in K^*} f_e(\epsilon_{\mathbf{y}^*}, \mu).$$

This shows by Theorem 29 the desired result. □

This concludes the section on Lagrangian duality. Finally we like to observe that all the above results can be more easily proved for finite dimensional optimization problems, i.e $X \subseteq \mathbb{R}^n$ and $Y = \mathbb{R}^m$. In this case the set $Y^* = \mathbb{R}^m$ is finite dimensional and instead of the Banach Alaoglu theorem and the weak*topology on Y^* we use the ordinary Euclidean topology on $\mathbb{R}^m$ and the result that a set $C \subseteq \mathbb{R}^m$ is compact if and only if the set C is closed and bounded.

References

[1] Ash, R.B. *Real Analysis and Probability*. Academic Press, New York, 1972.

[2] Aubin, J.B. *Optima and Equilibra: An Introduction to Nonlinear Analysis*. Graduate Texts in Mathematics. Springer Verlag, Berlin, 1993.

[3] Choquet, G. *Lectures on Analysis 1: Integration and Topological Vector Spaces*. Mathematics Lecture Note Series. W.A.Benjamin, Reading, Massachusetts, 1976.

[4] Chvátal, V. *Linear Programming*. Freeman, New York, 1999.

[5] Faigle, U., Kern, W. and Stil, G. *Algorithmic Principles of Mathematical Programming*, volume 24 of *Kluwer texts in Mathematical Sciences*. Kluwer Academic Publishers, Dordrecht, 2002.

[6] Frenk, J.B.G, Kassay, G and Protassov, V. *On Borel probability measures and noncooperative game theory*. Technical Report ERS-2002-85-LIS, ERIM, Erasmus University, Rotterdam, 2002.

[7] Frenk, J.B.G, Kassay, G and Kolumbán, J. *Equivalent results in minimax theory*. Technical Report ERS-2002-08-LIS, ERIM, Erasmus University, Rotterdam, 2002, revised version to appear in European Journal of Operational Research.

[8] Friedman, A. *Foundations of Modern Analysis*. Dover Publications, New York, 1982.

[9] Holmes, R.B. *Geometric Functional Analysis and its Applications*. Springer Verlag, New York, 1975.

[10] Jahn, J. *Mathematical Vector Optimization in Partially Ordered Spaces*. Peter Lang, Frankfurt am Main, 1986.

[11] Jones, A.N. *Game Theory: Mathematical Models of Conflict*. Horwood Publishing, Chicester, 2000.

[12] Kassay, G. and Kolumbán, J. *On a generalized sup-inf problem*. Journal of Optimization Theory and Applications. 91:651–670, 1996.

[13] Kolmogorov, A.N. and Fomin, S.V. *Introductory Real Analysis*. Dover Books on Mathematics. Dover Publications, New York, 1975.

[14] Nesterov, Y and Nemirowskii, A. *Interior-point Polynomial Algorithms in Convex Programming*, volume 13 of *SIAM studies in Applied Mathematics*. SIAM, Philadelphia, 1994.

[15] Neveu, J. *Mathematical Foundations of the Calculus of Probability*. Holden Day, San Francisco, 1965.

[16] Rockafellar, R. *Convex Analysis*, volume 28 of *Princeton Mathematical Series*. Princeton University Press, Princeton, 1972.

[17] Royden, H.L. *Real Analysis*. Macmillan, New York, 1988.

[18] Rudin, W. *Functional Analysis*. International Series in Pure and Applied Mathematics. McGraw-Hill, New York, 1991.

[19] Simons, S. *Minimax theorems and their proofs*. In Du. D.Z. and P.M Pardalos, editors, *Minimax and Applications*, pages 1–23. Kluwer Academic Publishers, Dordrecht, 1995.

[20] Ville, J. *Sur la théorie générale des jeux au intervient l'habilité des jouers 2(5)*. In E.Borel et al., editors, *Traité du calcul des probabilités et de ses applications*, pages 105–113. Gauthier-Villars Cie, Paris, 1938.

[21] von Neumann, J. *Zur Theorie der Gesellschaftsspiele*. Math.Ann. 100:295–320, 1928.

[22] von Neumann, J. and Morgenstern, O. *Theory of Games and Economic Behavior*. Princeton University Press, Princeton, 1944.

[23] Wald, A. *Generalization of a theorem by von Neumann concerning zero-sum two-person games*. Annals of Mathematics, 46(2):281–286, 1945.