TI 2002-093/4
Tinbergen Institute Discussion Paper

On Borel Probability Measures and Noncooperative Game Theory

J.B.G. Frenk[1]
G. Kassay[2]
V. Protassov[3]

[1] *Econometric Institute, Faculty of Economics, Erasmus University Rotterdam, and Tinbergen Institute,* [2] *Faculty of Mathematics and Computer Science, Babes Bolyai University, Cluj,* [3] *Dept. of Mechanics and Mathematics, Moscow State University, Moscow.*

Tinbergen Institute
The Tinbergen Institute is the institute for economic research of the Erasmus Universiteit Rotterdam, Universiteit van Amsterdam and Vrije Universiteit Amsterdam.

Tinbergen Institute Amsterdam
Keizersgracht 482
1017 EG Amsterdam
The Netherlands
Tel.: +31.(0)20.5513500
Fax: +31.(0)20.5513555

Tinbergen Institute Rotterdam
Burg. Oudlaan 50
3062 PA Rotterdam
The Netherlands
Tel.: +31.(0)10.4088900
Fax: +31.(0)10.4089031

On Borel Probability Measures and Noncooperative Game Theory.

J.B.G.Frenk
Econometric Institute, Erasmus University, Rotterdam

G.Kassay
Faculty of Mathematics and Computer Science, Babes Bolyai University,Cluj

V.Protassov
Dept. of Mechanics and Mathematics, Moscow State University,Moscow

September 11, 2002

Abstract

In this paper the well-known minimax theorems of Wald, Ville and Von Neumann are generalized under weaker topological conditions on the payoff function f and/or extended to the larger set of the Borel probability measures instead of the set of mixed strategies.

1 Introduction.

In this paper we will generalize the classical minimax theorems of von Neumann (cf.[18]), Ville (cf.[17]) and Wald (cf.[20]) in game theory under weaker topological conditions on the payoff function f. Also these results are extended to a larger class of strategies than the so-called class of mixed strategies (cf.[19]). Before presenting those results and the generalizations, we first need to introduce the following notations. Let A and B, unless stated otherwise, be nonempty Haussdorff spaces with Borel σ-algebras $\mathcal{A}$, respectively $\mathcal{B}$, and consider a payoff function $f : A \times B \to \mathbb{R}$. Denote now by $\mathcal{P}_F(A)$, respectively $\mathcal{P}_F(B)$, the set of all finite discrete Borel probability measures on $(A, \mathcal{A})$, respectively $(B, \mathcal{B})$. If ϵ_a represents the Borel probability measure concentrated on $a \in A$ then by definition λ belongs to $\mathcal{P}_F(A)$ if and only if there exists some finite set $\{a_1, ..., a_n\} \subseteq A$ and a finite set $\{\lambda_1, ..., \lambda_n\}$ of positive numbers satisfying $\sum_{i=1}^{n} \lambda_i = 1$ such that

$$\lambda = \sum_{i=1}^{n} \lambda_i \epsilon_{a_i}. \tag{1}$$

A similar observation applies to $\mathcal{P}_F(B)$, and so μ belongs to $\mathcal{P}_F(B)$ if and only if there exists some finite set $\{b_1, ..., b_m\} \subseteq B$ and a finite sequence $\{\mu_1, ..., \mu_m\}$ of positive numbers satisfying $\sum_{j=1}^{m} \mu_j = 1$ such that

$$\mu = \sum_{j=1}^{m} \mu_j \epsilon_{b_j}. \tag{2}$$

In noncooperative game theory (cf.[9], [19]) the sets $\mathcal{P}_F(A)$, respectively $\mathcal{P}_F(B)$ are called the set of mixed strategies of player 1, respectively player 2 and these strategies have a clear probabilistic interpretation. To measure the payoff for both players using mixed strategies we need to extend the payoff function $f : A \times B \to \mathbb{R}$ from the cartesian set of pure strategies to the cartesian set of mixed strategies. This extension $f_e : \mathcal{P}_F(A) \times \mathcal{P}_F(B) \to \mathbb{R}$ is defined by

$$f_e(\lambda, \mu) := \sum_{i=1}^{n} \sum_{j=1}^{m} \lambda_i \mu_j f(a_i, b_j) \tag{3}$$

with $\lambda \in \mathcal{P}_F(A), \mu \in \mathcal{P}_F(B)$ given by relations (1), respectively (2). Introducing the set $\mathcal{P}(A)$, respectively $\mathcal{P}(B)$ of all Borel probability measures on $(A, \mathcal{A})$, respectively $(B, \mathcal{B})$ it follows for A a finite set consisting of the elements $\{a_1, ..., a_n\}$ that $\mathcal{P}(A) = \mathcal{P}_F(A)$ and

$$\mathcal{P}_F(A) = \{\lambda : \lambda = \sum_{i=1}^{n} \lambda_i \epsilon_{a_i}, \sum_{i=1}^{n} \lambda_i = 1, \lambda_i \geq 0 \text{ for } 1 \leq i \leq n\}.$$

A similar observation also applies for B a finite set. In 1928 von Neumann (cf.[18]) published his famous minimax result for finite zero sum noncooperative games and this result in listed in the following theorem.

Theorem 1 *If A and B are finite sets, then it follows that*

$$\max_{\lambda \in \mathcal{P}(A)} \min_{\mu \in \mathcal{P}(B)} f_e(\lambda, \mu) = \min_{\mu \in \mathcal{P}(B)} \max_{\lambda \in \mathcal{P}(A)} f_e(\lambda, \mu).$$

The next minimax result due to Ville (cf.[17]) and published in 1938 is a generalization of the result of von Neumann and plays an important role in infinite zero sum noncooperative game theory (cf.[19]). In this theorem we need to assume that the pure strategy sets A and B are metric spaces.

Theorem 2 *If A and B are compact metric spaces and the function $f : A \times B \to \mathbb{R}$ is continuous, then*

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

Another generalization of von Neumann's result is due to Wald (cf.[20]) and published in 1945. This result plays a fundamental role in the theory of statistical decision functions (cf.[21]).

Theorem 3 *If B is a finite and A an arbitrary set, then it follows that*

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

Although the above results seem different it is possible to show that all these results and some other results proved more recently in the literature by sometimes different proofs can be easily deduced from each other and are equivalent to the well known separation result of a closed convex set and a point outside this set in a finite dimensional vector space (cf.[6]). This means that all these results are based on elementary mathematics. One of those equivalent results which plays an important role in this paper for the verification of the generalizations is given by a minimax result due to Kneser (cf.[11]) and proved in 1952. The proof of this result is very elementary, ingenious, and depends only on simple computations and the well known result (cf.[1]) that any upper semicontinuous function on a compact set attains its maximum (Weierstrass-Lebesgue lemma). Before mentioning this result we introduce for the function $f:A\times B\rightarrow\mathbb{R}$ the associated functions $f_a:B\rightarrow\mathbb{R}$ and $f_b:A\rightarrow\mathbb{R}$ given by $f_a(b)=f_b(a)=f(a,b)$.

Theorem 4 *If B is a nonempty convex, compact subset of a topological vector space, A is a nonempty convex subset of a vector space and the function $f:A\times B\rightarrow\mathbb{R}$ is affine in both variables and f_a is lower semicontinuous on B for every $a\in A$ then it follows that*

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

In this paper we will generalize the above results by weakening the topological conditions on the payoff function f and/or extending the set of mixed strategies. In particular, the Fubini-Tonelli theorem, the Riesz representation theorem, the separation theorem between disjoint convex set in normed linear spaces and the Banach-Alaoglu theorem play an important role in proving those generalizations. The first generalization under the strongest conditions is given by the following result.

Theorem 5 *Let $f:A\times B\rightarrow\mathbb{R}$ be either bounded from above or below and measurable with respect to the Borel product σ-algebra $\mathcal{A}\otimes\mathcal{B}$. If A and B are compact Hausdorff spaces and f_a is lower semicontinuous for every $a\in A$ and f_b is upper semicontinous for every $b\in B$, then it follows that*

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

Since every finite set is clearly compact this result is a generalization of the minimax result of Von Neumann. Also, since the topological conditions on the function f are weaker, it is a generalization of the minimax theorem of Ville. To give an interpretation within game theory we observe that Theorem 5 shows that player 1, respectively player 2, using the mixed strategy sets $\mathcal{P}_F(A)$, respectively $\mathcal{P}_F(B)$ with A and B compact Hausdorff spaces can achieve under some topological properties on the payoff function f an ϵ-equilibrium for any $\epsilon > 0$. Moreover, if the value of the game is positive (this can be assumed without loss of generality by scaling the payoff function$)$ this value is equal to the optimal objective value of the primal problem

$$\begin{array}{ll} \sup & \|\mu\|_{tv} \\ & \int_B f_a d\mu \leq 1 \quad a \in A \\ & \mu \in \mathcal{M}_F(B) \end{array}$$

to be solved by player 2. In this optimization problem $\mathcal{M}_F(B)$ denotes the set of all finite discrete Borel measures μ on $(B, \mathcal{B})$ with (finite) total variation norm $\|\mu\|_{tv}$. The same value can be determind by player 1 solving the dual problem

$$\begin{array}{ll} \inf & \|\lambda\|_{tv} \\ & \int_A f_b d\mu \geq 1 \quad b \in B \\ & \lambda \in \mathcal{M}_F(A) \end{array}$$

and so Theorem 5 generalizes the duality theorem of linear programming. Both the optimal objective value of the primal and dual problem are the same but the above problems might not have an optimal solution within the sets $\mathcal{M}_F(B)$, respectively $\mathcal{M}_F(A)$. It can be shown that the optimal solution for both players exists in the larger set of Borel measures with a finite total variation norm and by scaling these solutions we obtain the optimal strategies belonging to the set of Borel probability measures on $(A, \mathcal{A})$, respectively $(B, \mathcal{B})$. In the next result the topological conditions on the function f are weaker than the conditions presented in Theorem 5. Under these conditions the extension of the payoff function to the domain $\mathcal{P}(A) \times \mathcal{P}(\mathcal{B})$ is well defined and given by $f_e(\lambda, \mu) := \int_{A \times B} f d(\lambda \times \mu)$ with $\lambda \times \mu$ the Borel probability product measure on $(A \times B, \mathcal{A} \otimes \mathcal{B})$.

Theorem 6 *Let $f : A \times B \to \mathbb{R}$ be either bounded from above or below and measurable with respect to the Borel product σ-algebra $\mathcal{A} \otimes \mathcal{B}$. If the set B is a compact Hausdorff space, A an arbitrary set and f_a is lower semicontinuous for every $a \in A$ then it follows that*

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

This result can be seen as a generalization of the minimax result of Ville. Again it shows that the two players can achieve an ϵ-equilibrium for every $\epsilon > 0$, when the strategy sets are given by $\mathcal{P}(B)$ and $\mathcal{P}(A)$. As before, one can easily construct the associated primal and dual optimization problems for determining the value of the game and so Theorem 6 also generalizes the duality theorem of linear programming. Finally we list the minimax result valid under the weakest topological conditions.

Theorem 7 *If B is a compact Haussdorf space, A is an arbitrary set, and the function f_a is lower semicontinuous for every $a \in A$, then it follows that*

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

Again this is a generalization of the minimax result of Wald and von Neumann and as before it has a clear interpretation in game theory. Considering now these generalizations one might wonder whether the same equality holds under weaker assumptions. The main assumptions in these generalizations are a compactness assumption on the set of pure strategies, a topological and a boundedness assumption on the function f. It turns out that these assumptions are critical and to show this we list some counterexamples in the last section.

2 On the Riesz Representation Theorem and Lower Semicontinuous Functions.

In this section we will gather results needed from functional analysis for the proof of the minimax result. Let B be a compact Hausdorff space and introduce the normed linear space $(C(B), \|.\|_\infty)$ of all continuous real valued functions h on B equipped with the supnorm

$$\|h\|_\infty := \sup_{x\in B} |h(x)| < \infty.$$

The set of all continuous linear functionals on $C(B)$ is given by the dual linear space $C(B)^*$ and this linear space has dual norm

$$\|x^*\|_d := \sup_{\|h\|_\infty\neq 0} \frac{|< x^*, h >|}{\|h\|_\infty}.$$

Also, let $(\mathcal{M}(B), \|.\|_{tv})$ denote the normed linear space of all finite signed Borel measures on the measurable space $(B, \mathcal{B})$ with $\mathcal{B}$ the Borel σ-algebra on B and $\|.\|_{tv}$ the total variation norm and consider for every $\mu \in \mathcal{M}(B)$ the continuous linear functional $I_\mu : C(B) \to \mathbb{R}$ defined by

$$< I_\mu, h >:= \int_B h d\mu.$$

Clearly I_μ belongs to the dual space $C(B)^*$ and by the so-called Riesz representation theorem it follows that all elements of the dual space $C(B)^*$ actually have this representation (cf.[4]).

Theorem 8 *For B a compact Hausdorff space the mapping $\mu \to I_\mu$ is a linear mapping of the space $(\mathcal{M}(B), \|.\|_{tv})$ onto the space $(C(K)^*, \|.\|_d)$ satisfying*

$$\|\mu\|_{tv} = \|I_\mu\|_d.$$

By the Riesz representation theorem we obtain that the unit ball $U := \{I_\mu \in C(B)^* : \|I_\mu\|_d \leq 1\}$ can be identified with the set $\{\mu \in \mathcal{M}(B) : \|\mu\|_{tv} \leq 1\}$ and since by the Banach-Alaoglu theorem (cf.[3]) the set U is weak*compact, the set $\{\mu \in \mathcal{M}(B) : \|\mu\|_{tv} \leq 1\}$ must also be weak*compact. It is shown in [13] that the Banach-Alaoglu theorem is an easy consequence of Tychonoff's theorem on the cartesian product of compact sets. Introducing now the set $\mathcal{P}(B)$ of Borel probability measures on $(B, \mathcal{B})$ we obtain

$$\mathcal{P}(B) = \{\mu \in \mathcal{M}(B) : \mu \text{ a positive measure and } \|\mu\|_{tv} = 1\}.$$

We will now show that this set $\mathcal{P}(B) \subseteq \{\mu \in \mathcal{M}(B) : \|\mu\|_{tv} \leq 1\}$ is closed in the weak*topology and hence weak*compact. It is well known (cf.[3]) by the definition of the weak*topology and Theorem 8 that a net $\{\mu_i, i \in I\} \subseteq \mathcal{M}(B)$ converges in the weak*topology to the finite signed Borel measure μ if and only if $\int_B h d\mu_i \to \int_B h d\mu$ for every h belonging to $C(B)$. This implies that any net of Borel probability measures converging in the weak*topology converges to a Borel probability measure and so it follows that $\mathcal{P}(B)$ is a weak*closed and hence a weak*compact subset of $\{\mu \in \mathcal{M}(B) : \|\mu\|_{tv} \leq 1\}$. Actually one can prove using the so-called theorem of approximation (cf.[2]) that the convex set $\mathcal{P}(B)$ of Borel probability measures is given by the weak*closure of all finite convex combinations of the so-called one point Borel probability measures ϵ_b concentrated on $b, b \in B$. This means $cl(\mathcal{P}_F(B)) = \mathcal{P}(B)$ with $\mathcal{P}_F(B)$ the already introduced set of all finite discrete probability measures on $(B, \mathcal{B})$. Summarizing we have the following result.

Theorem 9 *If the set B is a compact Hausdorff space then the set $\mathcal{P}(B)$ is weak*compact. Moreover, if $\mathcal{P}_F(B) \subseteq \mathcal{P}(B)$ denotes the set of all finite discrete probability measures on B,then it follows that $cl(\mathcal{P}_F(B)) = \mathcal{P}(B)$ with the closure taken in the weak*topology.*

The following result is a simple consequence of the Riesz representation theorem and the Hahn Banach theorem on normed linear spaces.

Lemma 10 *Let B be a compact Hausdorff space. For an arbitrary convex set $G \subseteq C(B)$ the following properties are equivalent:*

1. *For every $h \in G$ it holds that $\min_{x \in B} h(x) \leq 0$.*
2. *There exists a Borel probability measure μ on $(B, \mathcal{B})$ such that $\int_B h d\mu \leq 0$ for every $h \in G$.*

Proof. We first observe for every $h \in C(B)$ and B compact that the minimum of h over B is attained. To show $2 \Rightarrow 1$ assume by contradiction that there exists some $h \in G$ satisfying $\min_{x \in B} h(x) > 0$. This implies for every Borel probability measure μ on $(B, \mathcal{B})$ that $\int_B h d\mu \geq \min_{x \in B} h(x) > 0$ and so we obtain a contradiction. To verify $1 \Rightarrow 2$, it is clear that the convex set G does not intersect the convex cone $K^+ := \{h \in C(B) : \min_{x \in B} h(x) > 0\}$. Since the set K^+ is open in the normed linear space $(C(B), \|.\|_\infty)$ we may apply the separation theorem in normed linear spaces between two disjoint convex sets of which one set is open (cf.[15]), and so there exist some $x_0^* \in C(B)^*$ satisfying

$$\sup_{h \in G} < x_0^*, h > \leq \inf_{h \in K^+} < x_0^*, h > . \tag{4}$$

To show that x_0^* is a positive continuous linear functional we assume by contradiction that there exists some $h_0 \in K^+$ satisfying $< x_0^*, h_0 >< 0$. This implies using $th_0 \in K^+$ for every $t > 0$ that $\inf_{h \in K^+} < x_0^*, h >= -\infty$ and so by relation (4) we obtain $\sup_{h \in G} < x_0^*, h >= -\infty$. This is a contradiction and so it follows for every $h \in K^+$ that

$$\sup_{h \in G} < x_0^*, h > \leq \inf_{h \in K^+} < x_0^*, h >= 0 \tag{5}$$

or equivalently x_0^* is a positive continuous linear functional. By the Riesz representation theorem there exists some finite signed Borel measure μ satisfying

$$< x_0^*, h >= \int_B h d\mu$$

for every $h \in C(B)$ and since x_0^* is a positive continuous linear functional it must follow that μ is a finite Borel measure. By scaling we may assume without loss of generality that $\mu(B) = 1$ and so by relation (5) the desired result follows. □

To extend the above result to a larger class of functions, recall that the class of lower semicontinuous real valued functions on the compact Hausdorff space B is given by the next definition.

Definition 11 *The function $\phi : B \to \mathbb{R}$ is called lower semicontinuous if for every $r \in \mathbb{R}$ the lower level set $L_\phi(r) := \{x \in B : \phi(x) \leq r\}$ is closed.*

Clearly a lower semicontinuous function is a Borel measurable function. In the next result we relate the class of lower semicontinuous functions to the class of continuous functions. Although this result is known, we list a short proof for completeness.

Theorem 12 *The following properties of a function $\phi : B \to \mathbb{R}$ on the compact Hausdorff space B are equivalent:*

1. *The function ϕ is lower semicontinuous.*
2. *$\phi(x) = \sup_{h \in H_\phi} h(x)$ with $H_\phi := \{h \in C(B) : h \leq \phi\}$ nonempty.*

Proof. To show $1 \Rightarrow 2$ we first observe using ϕ is lower semicontinuous and B compact that by the Weierstrass-Lebesgue lemma the function ϕ attains its minimum over B. Hence without loss of generality we may assume that $\phi \geq 0$. Clearly $\phi(x) \geq \sup_{h \in H_\phi} h(x)$ for every x and assume now by contradiction that $\phi(x_0) > r > \sup_{h \in H_\phi} h(x_0)$ for some $x_0 \in X$. Using now the lower semicontinuity of the function ϕ, there exists some open neighborhood U of x_0 satisfying $\phi(x) > r$ for every $x \in U$. Also, since B is a compact Hausdorff space, the set B is normal (cf.[4]) and the sets $\{x_0\}$ and $B \backslash U$ are closed and disjoint. Hence Urysohn's lemma holds and so one can find some $h \in C(B)$ satisfying $0 \leq h \leq 1$, $h(x_0) = 1$ and $h(x) = 0$ for every $x \in B \backslash U$. Taking now $h_r := rh$ it is easy to verify that $h_r \in C(B)$, $h_r \leq \phi$ and $h_r(x_0) = r$ and we obtain a contradiction. The implication $2 \Rightarrow 1$ is obvious and so we omit its proof. □

By the above result we see that a function is lower semicontinuous on a compact set B if and only if it can be pointwise approximated from below by continuous functions on B. Actually the set of lower semicontinuous functions is obtained from the normed linear space $(C(B), \|.\|_\infty)$ by addition of an extra operation: taking the supremum of an arbitrary set of functions. It is now easy to see that the set of lower semicontinuous functions is the smallest class of functions on B which contains $C(B)$ and is closed with respect to taking a supremum of an arbitrary set of functions belonging to this class. An immediate consequence of Theorem 12 is given by the next result.

Lemma 13 *Let B be a compact Hausdorff space. A real valued function ϕ on B is lower semicontinuous if and only if H_ϕ is nonempty and for every Borel* *probability measure μ on $(B, \mathcal{B})$ it holds that*

$$\sup_{h\in H_\phi} \int_B h d\mu = \int_B \phi d\mu.$$

Moreover, for ϕ lower semicontinuous it follows that the mapping $L : \mathcal{M}(B) \to (-\infty, \infty]$ given by $L(\mu) = \int_B \phi d\mu$ is lower semicontinuous in the weak* topology on $\mathcal{M}(B)$.

Proof. By Theorem 12 it follows that $\phi = \sup_{h\in H_\phi} h$ and this implies by the definition of the integral $\int \phi d\mu$ (cf.[2]) that $\sup_{h\in H_\phi} \int h d\mu = \int \phi d\mu$. To show the reverse implication we observe that ϵ_x is a Borel probability measure for every $x \in B$ and this implies by our assumption that

$$\phi(x) = \int_B \phi d\epsilon_x = \sup_{h\in H_\phi} \int_B h d\epsilon_x = \sup_{h\in H_\phi} h(x)$$

showing the desired result. To prove the last part we observe by the definition of the weak*topology that for every $h \in C(B)$ the mapping $\mu \to \int_B h d\mu$ is continuous in the weak*topology and using now $\int_B \phi d\mu = \sup_{h\in H_\phi} \int_B h d\mu$ the desired result follows. $\square$

We now list an extension of Lemma 10 to the class of lower semicontinuous functions on the compact Hausdorff space B.

Lemma 14 *Let B be a compact Hausdorff space. For an arbitrary convex set G of lower semicontinuous functions on B the following properties are equivalent:*

1. *For every $\phi \in G$ it holds that $\min_{x\in B} \phi(x) \leq 0$.*
2. *There exists a Borel probability measure μ on $(B, \mathcal{B})$ such that $\int_B \phi d\mu \leq 0$ for every $\phi \in G$.*

Proof. Again by the Weierstrass-Lebesgue lemma the function ϕ attains its minimum over B. As in Lemma 10 one can easily show $2 \Rightarrow 1$ and so we only verify $2 \Rightarrow 1$. Considering the set $G_0 := \cup_{\phi\in G} H_\phi \subseteq C(B)$ it follows by the convexity of the set G that also G_0 is convex. Also by our assumption we obtain that $\min_{x\in B} h(x) \leq 0$ for every $h \in G_0$. Hence we may apply Lemma 10 and so there exists some Borel probability measure μ on $(B, \mathcal{B})$ satisfying $\int h d\mu \leq 0$ for every $h \in G_0$. This implies for every $\phi \in G$ using $H_\phi \subseteq G_0$ that $\int_B \phi d\mu = \sup_{h\in H_\phi} \int_B h d\mu \leq 0$ and the proof is completed. $\square$

In the next example we construct a convex set G containing at least one Borel measurable and not lower semicontinuous function and for this set G

we show $\inf_{x\in B}\phi(x) \leq 0$ for every $\phi \in G$ and $\sup_{\phi\in G}\int_B \phi d\mu > 0$ for every Borel probability measure μ. This means that Lemma 14 does not hold if the convex set G contains at least one Borel measurable function which is not lower semicontinuous.

Example 15 *Let $\phi_0 : B \to \mathbb{R}$ be an arbitrary Borel measurable function bounded from below but not lower semicontinuous. For such a function the set H_{ϕ_0} is nonempty and so by Lemma 13 there exists some Borel probability measure μ_0 on $(B, \mathcal{B})$ satisfying*

$$\sup_{h\in H_{\phi_0}} \int_B h d\mu_0 < \int_B \phi_0 d\mu_0.$$

Without loss of generality (add a constant to the function ϕ_0) we may assume that

$$0 = \sup_{h\in H_{\phi_0}} \int_B h d\mu_0 < \int_B \phi_0 d\mu_0. \tag{6}$$

Introduce now the nonempty convex cone $G_0 := \{h \in C(B) : \int_B h d\mu_0 \leq 0\}$ and consider the convex set

$$G := \{\alpha\phi_0 + h : h \in G_0, 0 \leq \alpha \leq 1\}.$$

For this convex set G we will now verify that $\inf_{x\in B}\phi(x) \leq 0$ for every $\phi \in G$ and $\sup_{\phi\in G}\int_B \phi d\mu > 0$ for every Borel probability measure μ. To show that $\inf_{x\in B}\phi(x) \leq 0$ for every $\phi \in G$ we assume by contradiction that there exists some $0 \leq \alpha_0 \leq 1$ and $h_0 \in G_0$ satisfying

$$\beta := \inf_{x\in B}(\alpha_0\phi_0(x) + h_0(x)) > 0 \tag{7}$$

If $\alpha_0 = 0$ then by relation (7) it follows that $\int_B h_0 d\mu_0 \geq \beta > 0$ and this contradicts $h_0 \in G_0$. Therefore $\alpha_0 > 0$ and again by relation (7) we obtain $\phi_0(x) \geq \alpha_0^{-1}(\beta - h_0(x))$ for every $x \in B$. Since $h_0 \in C(B)$ this implies that the function $x \to \alpha_0^{-1}(\beta - h_0(x))$ belongs to H_{ϕ_0} and by relation (6) it must follow that

$$\int_B \alpha_0^{-1}(\beta - h_0) d\mu_0 \leq 0. \tag{8}$$

Also, since $h_0 \in G_0, \alpha_0 > 0$ and $\beta > 0$ we obtain

$$\int_B \alpha_0^{-1}(\beta - h_0) d\mu_0 \geq \alpha_0^{-1}\beta > 0.$$

and this contradicts relation (8). Therefore it must hold that $\inf_{x\in B}\phi(x)\leq 0$ *for every* $\phi\in G$ *and we have verified the first property of the convex set* G*. To show* $\sup_{\phi\in G}\int_B \phi d\mu > 0$ *for every Borel probability measure* μ *we first observe that for every Borel probability measure* $\mu\neq\mu_0$ *there exists some* $h\in C(B)$ *satisfying* $\int_B hd\mu\neq\int_B hd\mu_0$*. Without loss of generality we may assume* $\int_B hd\mu > \int_B hd\mu_0$ *(take* $-h$ *instead of* h*) and adding a constant to the function* h *one can find a function* $h_0\in C(B)$ *satisfying*

$$\int_B h_0 d\mu_0 \leq 0 \text{ and } \int_B h_0 d\mu > 0. \tag{9}$$

This shows $h_0\in G_0$ *and since* G_0 *is a convex cone, also* $\alpha h_0\in G_0$ *for every* $\alpha>0$*. This implies using relation (9) that*

$$\sup_{h\in G_0}\int_B hd\mu=\infty,$$

and since $G_0\subseteq G$ *it follows for every Borel probability measure* $\mu\neq\mu_0$ *that*

$$\sup_{\phi\in G}\int \phi d\mu \geq \sup_{h\in G_0}\int hd\mu=\infty.$$

Also for $\mu=\mu_0$ *we obtain by relation (6) that*

$$\sup_{\phi\in G}\int_B \phi d\mu_0 \geq \int_B \phi_0 d\mu_0 > 0$$

and so one may conclude that in this example there does not exist any Borel probability measure μ *satisfying* $\sup_{\phi\in G}\int_B\phi d\mu\leq 0$*. This means that Lemma 13 does not hold for the considered convex set* G *and our counterexample is completed.*

This concludes our discussion of consequences of the Riesz representation theorem. In the next section we will consider the application of the above results within game theory.

3 On Minimax Theorems.

Using the results of the previous section, we will prove in this section some of the generalizations of the minimax results of Von Neumann, Wald and Ville. As already observed the minimax results of Von Neumann, Wald and Ville are equivalent to the separation result in finite dimensional vector spaces between a closed convex set and a point outside this set. For our generalizations the stronger mathematical tools of the previous section are needed. For the key result given by Theorem 17 we give two different proofs. One proof uses Theorem 9 (a combination of the Banach-Alaoglu theorem and the Riesz representation theorem) and the last part of Lemma 13, based on Urysohn's lemma, to verify that the conditions of Kneser's minimax result hold and this yields the result. The other proof uses Lemma 14 based on the Riesz representation theorem, Urysohn's lemma and the separation result between disjoint convex set in normed linear spaces. We will start with the proof based on Lemma 14. Let $f : A \times B \to \mathbb{R}$ be given and consider the functions $f_a : B \to \mathbb{R}$ and $f_b : A \to \mathbb{R}$ given by

$$f_a(b) = f_b(a) := f(a, b). \tag{10}$$

If B is a compact Hausdorff space and f_a is lower semicontinuous on B for every $a \in A$ then by Theorem 12 and the definition of the integral we obtain for every finite signed Borel measure μ on $(B, \mathcal{B})$ that

$$\int_B f_a d\mu := \sup\{\int_B h d\mu : h \in H_{f_a}\} \leq \infty$$

and this implies that the integral $\int_B f_a d\mu$ is well defined for every $a \in A$ and μ a finite signed Borel measure. Hence it is possible to prove the following consequence of Lemma 14.

Lemma 16 *Let c_0 be a finite constant and B a compact Hausdorff space. If for every $a \in A$ the function f_a is lower semicontinuous then it follows that* $\sup_{\lambda \in \mathcal{P}_F(A)} \min_{b \in B} f_e(\lambda, \epsilon_b) \leq c_0$ *if and only if* $\sup_{a \in A} f_e(\epsilon_a, \mu) \leq c_0$ *for some* $\mu \in \mathcal{P}(B)$.

Proof. Replacing the function f by $f - c_0 1_{A \times B}$ with $1_{A \times B}(a, b) := 1$ for every $(a, b) \in A \times B$ we may assume without loss of generality that $c_0 = 0$. Introducing now for every $\lambda \in \mathcal{P}_F(A)$ the function $\phi_\lambda : B \to \mathbb{R}$ given by

$$\phi_\lambda(b) := f_e(\lambda, \epsilon_b) = \int_A f_b d\lambda$$

it follows that the set $G := \{\phi_\lambda : \lambda \in \mathcal{P}_F(A)\}$ is convex. Due to f_a is lower semicontinuous for every $a \in A$ the set G is also a subset of the set of lower semicontinuous functions on B and ϕ_λ attains its minimum over B. Since by definition

$$\sup_{\lambda \in \mathcal{P}_F(A)} \min_{b \in B} f_e(\lambda, \epsilon_b) \leq 0 \Leftrightarrow \min_{b \in B} \phi_\lambda(b) \leq 0 \text{ for every } \phi_\lambda \in G$$

and

$$\sup_{a\in A} f_e(\epsilon_a,\mu) \le 0 \Leftrightarrow \int_B \phi_{\epsilon_a} d\mu = \int_B f_a d\mu \le 0 \text{ for every } a \in A$$

we obtain the desired result by Lemma 14. □

An immediate consequence of Lemma 16 is given by the following result. This result will play a key role in this paper.

Theorem 17 *Let A be an arbitrary set and B a compact Hausdorff space. If the function f_a is lower semicontinuous for every $a \in A$ then it follows that*

$$\sup_{\lambda\in\mathcal{P}_F(A)} \min_{b\in B} f_e(\lambda,\epsilon_b) = \inf_{\mu\in\mathcal{P}(B)} \sup_{a\in A} f_e(\epsilon_a,\mu).$$

and there exists some $\mu \in \mathcal{P}(B)$ attaining the above infimum.

Proof. The result follows immediately by applying Lemma 16. □

The following remarks are immediate consequences of Theorem 17.

Remark 18 *In this remark we observe the following:*

1. *If B is a compact Hausdorff space and the function f_a is upper semicontinuous for every $a \in A$ instead of lower semicontinuous then we replace in Theorem 17 the function f by $-f$ and this yields the equality*

$$\inf_{\lambda\in\mathcal{P}_F(A)} \max_{b\in B} f_e(\lambda,\epsilon_b) = \max_{\mu\in\mathcal{P}(B)} \inf_{a\in A} f_e(\epsilon_a,\mu) \quad (11)$$

   Reversing the roles of the sets A and B we obtain by relation (11) for A a compact Hausdorff space and the function f_b is upper semicontinuous for every $b \in B$ that

$$\inf_{\mu\in\mathcal{P}_F(B)} \max_{a\in A} f_e(\epsilon_a,\mu) = \max_{\lambda\in\mathcal{P}(A)} \inf_{b\in B} f_e(\lambda,\epsilon_b) \quad (12)$$

2. *To compute the optimal Borel probability measure μ on $(B,\mathcal{B})$ satisfying the equality in Theorem 17 we need to solve the optimization problem*

$$\begin{array}{ll} \min & z \\ & \int_B f_a d\mu \le z \quad a \in A \\ & \mu \in \mathcal{P}(B) \end{array}$$

   By scaling the function f_a, the optimal solution of the above problem does not change and so we may assume that the optimal objective value *of this "generalized linear programming" problem is positive. Replacing now every feasible (μ, z) with $\mu \in \mathcal{P}(B)$ and $z > 0$ by the finite Borel measure $\overline{\mu} = z^{-1}\mu$ and using $\|\overline{\mu}\|_{tv} = \overline{\mu}(B) = z^{-1}$ we need to solve the primal optimization problem*

$$\begin{array}{ll} \max & \|\mu\|_{tv} \\ & \int_B f_a d\mu \leq 1 \quad a \in A \\ & \mu \in \mathcal{M}(B). \end{array} \tag{P}$$

It is easy to show that the result in Theorem 17 is actually a minimax result. Since $\epsilon_b \in \mathcal{P}(B)$ for every $b \in B$ and $\int_B \phi d\mu \geq \inf_{b\in B} \phi(b)$, for ϕ lower semicontinuous and $\mu \in \mathcal{P}(B)$ with B a compact Hausdorff space we obtain

$$\inf_{\mu\in\mathcal{P}(B)} \int_B \phi d\mu = \inf_{b\in B} \phi(b). \tag{13}$$

Actually, since ϕ is lower semicontinuous, B a compact Hausdorff space and $\epsilon_b \in \mathcal{P}(B)$ for every $b \in B$, we obtain by the above equality that

$$\min_{\mu\in\mathcal{P}(B)} \int_B \phi d\mu = \min_{b\in B} \phi(b).$$

This implies with ϕ replaced by $b \to f_e(\lambda, \epsilon_b)$, $\lambda \in \mathcal{P}_F(A)$ that under the same conditions as in Theorem 17

$$\min_{\mu\in\mathcal{P}(B)} f_e(\lambda, \mu) = \min_{b\in B} f_e(\lambda, \epsilon_b)$$

for every $\lambda \in \mathcal{P}_F(A)$. Also it is easy to verify that

$$\sup_{a\in A} f_e(\epsilon_a, \mu) = \sup_{\lambda\in\mathcal{P}_F(A)} f_e(\lambda, \mu) \tag{14}$$

for every $\mu \in \mathcal{P}(B)$ and so the result in Theorem 17 is the same as the minimax result

$$\sup_{\lambda\in\mathcal{P}_F(A)} \min_{\mu\in\mathcal{P}(B)} f_e(\lambda, \mu) = \min_{\mu\in\mathcal{P}(B)} \sup_{\lambda\in\mathcal{P}_F(A)} f_e(\lambda, \mu). \tag{15}$$

The equality in relation (15) shows that players using the mixed strategy sets $\mathcal{P}_F(A)$ and $\mathcal{P}(B)$ can achieve for B a compact Hausdorff space and f satisfying some additional topological conditions an ϵ-equilibrium for every $\epsilon > 0$. Also the player using strategy set $\mathcal{P}(B)$ can achieve the value of the game. Clearly this result is a generalization of the minimax result of Wald. An alternative proof of relation (15) and hence of Theorem 17 is given by an application of Kneser's minimax result in combination with the Riesz representation theorem and the weak*compactness of $\mathcal{P}(B)$.

Proof. By Theorem 9 the set $\mathcal{P}(B)$ of Borel probability measures on $(B, \mathcal{B})$ is weak*compact. Also by Lemma 13 the mapping $\mu \to f_e(\epsilon_a, \mu)$ is lower semicontinuous in the weak*topology on $\mathcal{M}(B)$ for every $a \in A$ and this implies that the mapping $\mu \to f_e(\lambda, \mu)$ is also lower semicontinuous in the weak*topology for every $\lambda \in \mathcal{P}_F(A)$. Since the function $(\lambda, \mu) \to f_e(\lambda, \mu)$ is affine in both variables on $\mathcal{P}_F(A) \times \mathcal{P}(B)$ and $\mathcal{P}_F(A)$ is clearly convex the conditions of Kneser's minimax result hold and this shows the result. $\square$

Assuming for the moment that the integral $f_e(\lambda, \epsilon_b) = \int_A f_b d\lambda$ is well defined for every $\lambda \in \mathcal{P}(A)$ and $b \in B$ it follows that

$$\sup_{\lambda \in \mathcal{P}(A)} \inf_{b \in B} f_e(\lambda, \epsilon_b) \geq \sup_{\lambda \in \mathcal{P}_F(A)} \inf_{b \in B} f_e(\lambda, \epsilon_b). \tag{16}$$

Imposing the same conditions as in Theorem 17 this implies

$$\sup_{\lambda \in \mathcal{P}(A)} \inf_{b \in B} f_e(\lambda, \epsilon_b) \geq \min_{\mu \in \mathcal{P}(B)} \sup_{a \in A} f_e(\epsilon_a, \mu). \tag{17}$$

We are now interested under which conditions an equality occurs in relation (17). By relation (16) such an equality is stronger as the one verified in Theorem 17 and so it seems reasonable to impose, besides the conditions of Theorem 17, some additional condition on f. This additional condition is given by the assumption that the integral

$$f_e(\lambda, \mu) = \int_{A \times B} f d(\lambda \times \mu)$$

is well defined with $\lambda \times \mu$ denoting the Borel probability product measure of $\lambda \in \mathcal{P}(A)$ and $\mu \in \mathcal{P}(B)$. By the Fubini-Tonelli theorem (cf.[4]) it is well-known that this integral indeed exists and satisfies

$$\int_{A \times B} f d(\lambda \times \mu) = \int_A (\int_B f_a d\mu) d\lambda = \int_B (\int_A f_b d\lambda) d\mu \tag{18}$$

if the function $f : A \times B \to \mathbb{R}$ is measurable with respect to the Borel product σ-algebra $\mathcal{A} \otimes \mathcal{B}$ and is either bounded from below or above.

Theorem 19 *Let $f : A \times B \to \mathbb{R}$ be measurable with respect to $\mathcal{A} \otimes \mathcal{B}$ and either bounded from above or below. If either B is a compact Hausdorff space and f_a is lower semicontinuous for every $a \in A$ or A is a compact Hausdorff space and f_b is upper semicontinous for every $b \in B$ then it follows that*

$$\sup_{\lambda \in \mathcal{P}(A)} \inf_{b \in B} f_e(\lambda, \epsilon_b) = \inf_{\mu \in \mathcal{P}(B)} \sup_{a \in A} f_e(\epsilon_a, \mu).$$

Moreover, if B is a compact Hausdorff space and f_a lower semicontinuous then there exists some $\mu \in \mathcal{P}(B)$ attaining the above infimum, while for A a compact Hausdorff space and f_b upper semicontinuous, there exists some $\lambda \in \mathcal{P}(A)$ attaining the above supremum.

Proof. We first assume that B is a compact Hausdorff space and f_a is upper semicontinuous for every $a \in A$. Since by the first assumption the Fubini-Tonelli theorem holds the integral $f_e(\lambda, \epsilon_b)$ is well defined for every $\lambda \in \mathcal{P}(A)$ and so by relation (17) we only need to show that

$$\sup_{\lambda \in \mathcal{P}(A)} \inf_{b \in B} f_e(\lambda, \epsilon_b) \leq \inf_{\mu \in \mathcal{P}(B)} \sup_{a \in A} f_e(\epsilon_a, \mu).$$

Since μ and λ are probability measures we obtain by the Fubini-Tonelli theorem that

$$\inf_{b \in B} \int_A f_b d\lambda \leq \int_B (\int_A f_b d\lambda) d\mu = \int_A (\int_B f_a d\mu) d\lambda \leq \sup_{a \in A} \int_B f_a d\mu$$

and this implies

$$\sup_{\lambda \in \mathcal{P}(A)} \inf_{b \in B} f_e(\lambda, \epsilon_b) \leq \inf_{\mu \in \mathcal{P}(B)} \sup_{a \in A} f_e(\epsilon_a, \mu)$$

showing the desired result. To prove the result for A a compact Hausdorff space and f_b upper semicontinuous for every $b \in B$ we apply the first part with f replaced by $-f$ and the roles of A and B reversed. This implies

$$\begin{aligned} \sup_{\lambda \in \mathcal{P}(A)} \inf_{b \in B} f_e(\lambda, \epsilon_b) &= -\inf_{\lambda \in \mathcal{P}(A)} \sup_{b \in B} -f_e(\lambda, \epsilon_b) \\ &= -\sup_{\mu \in \mathcal{P}(B)} \inf_{a \in A} -f_e(\epsilon_a, \mu) \\ &= \inf_{\mu \in \mathcal{P}(B)} \sup_{a \in A} f_e(\epsilon_a, \mu) \end{aligned}$$

and so the second part is verified. To show the last part we observe by Lemma 13 that $\mu \to \sup_{a \in A} \int_B f_a d\mu$ is lower semicontinuous in the weak*topology and by the Weierstrass-Lebesgue lemma the infimum is attained. A similar proof applies for B compact and f_b is upper semicontinuous. □

The following remarks are immediate consequences of Theorem 19.

Remark 20 *In this remark we observe the following:*

1. *Let $f : A \times B \to \mathbb{R}$ be measurable with respect to $\mathcal{A} \otimes \mathcal{B}$ and either bounded from above or below. If B is a compact Hausdorff space and the function f_a is upper semicontinuous for every $a \in A$ instead of lower semicontinuous then we replace in Theorem 19 the function f by $-f$ and this yields the equality*

$$\inf_{\lambda \in \mathcal{P}(A)} \sup_{b \in B} f_e(\lambda, \epsilon_b) = \max_{\mu \in \mathcal{P}(B)} \inf_{a \in A} f_e(\epsilon_a, \mu). \tag{19}$$

A similar observation applies for A a compact Hausdorff space and f_b lower semicontinuous for every $b \in B$ yielding the equality

$$\min_{\lambda \in \mathcal{P}(A)} \sup_{b \in B} f_e(\lambda, \epsilon_b) = \sup_{\mu \in \mathcal{P}(B)} \inf_{a \in A} f_e(\epsilon_a, \mu) \tag{20}$$

2. *In Remark 18 we observed for B a compact Hausdorff space and f_a lower semicontinuous for every $a \in A$ that finding the Borel probability measure μ on $(B, \mathcal{B})$ attaining $\inf_{\mu \in \mathcal{P}(B)} \sup_{a \in A} f_e(\epsilon_a, \mu)$ boils down to solving the "generalized linear programming" problem*

$$\begin{array}{ll} \max & \|\mu\|_{tv} \\ & \int_B f_a d\mu \leq 1 \quad a \in A \\ & \mu \in \mathcal{M}(B). \end{array}$$

In case A is a compact Hausdorff space and f_b is upper semicontinuous for every $b \in B$ we obtain under the conditions of Theorem 19 that

$$\max_{\lambda \in \mathcal{P}(A)} \inf_{b \in B} f_e(\lambda, \epsilon_b) = \inf_{\mu \in \mathcal{P}(B)} \sup_{a \in A} f_e(\epsilon_a, \mu).$$

To compute the optimal $\lambda \in \mathcal{P}(A)$ attaining the above maximum one can show similarly that the dual "generalized linear programming" problem

$$\begin{array}{ll} \min & \|\lambda\|_{tv} \\ & \int_A f_b d\lambda \geq 1 \quad b \in B \\ & \lambda \in \mathcal{M}(A) \end{array}$$

needs to solved. Actually for the primal objective value

$$\inf_{\mu \in \mathcal{P}(B)} \sup_{a \in A} f_e(\epsilon_a, \mu)$$

positive, the above optimization problems with max replaced by sup and min by inf have the same optimal objective value and so Theorem 19 generalizes the duality principle in linear programming. In case both A and B are finite sets Theorem 19 reduces to Von Neumann's minimax result, which can be derived by the duality principle. As we saw, the generalized duality principle holds under the assumptions of Theorem 19 with only one of those problems having an optimal feasible solution. In case both problems have an optimal solution we need to assume by Theorem 19 that A and B are compact Hausdorff spaces and f_b is upper semicontinuous for every $b \in B$ and f_a is lower semicontinuous for every $a \in A$. In the next subsection we will show by means of counterexamples that this generalized duality principle will fail if the conditions of Theorem 19 are weakened.

Since in Theorem 19 the Fubini-Tonelli theorem holds we obtain in a similar way as in relation (13) that

$$\inf_{b \in B} f_e(\lambda, \epsilon_b) = \inf_{\mu \in \mathcal{P}(B)} f_e(\lambda, \mu)$$

for every $\lambda \in \mathcal{P}(A)$ and

$$\sup_{a\in A} f_e(\epsilon_a,\mu) = \sup_{\lambda\in\mathcal{P}(A)} f_e(\lambda,\mu)$$

for every $\mu \in \mathcal{P}(B)$. This shows that the result in Theorem 19 is the same as the minimax result

$$\sup_{\lambda\in\mathcal{P}(A)} \inf_{\mu\in\mathcal{P}(B)} f_e(\lambda,\mu) = \inf_{\mu\in\mathcal{P}(B)} \sup_{\lambda\in\mathcal{P}(A)} f_e(\lambda,\mu). \qquad (21)$$

Hence under the conditions of Theorem 19 any two players with the strategy sets $\mathcal{P}(A)$ and $\mathcal{P}(B)$ can achieve an ϵ-equilibrium for every $\epsilon > 0$. Moreover, in case A is compact the player with strategy set $\mathcal{P}(A)$ can achieve the value of the game, while for B compact the player with strategy set B can achieve this value. Finally, inspired by the second part of Remark 20 we list the following consequence of Theorem 19 and 17.

Theorem 21 *Let $f : A \times B \to \mathbb{R}$ be measurable with respect to $\mathcal{A} \otimes \mathcal{B}$ and either bounded from above or below. If B is a compact Hausdorff space and f_a is lower semicontinuous for every $a \in A$ and A is a compact Hausdorff space and f_b is upper semicontinous for every $b \in B$ then it follows that*

$$\sup_{\lambda\in\mathcal{P}_F(A)} \min_{b\in B} f_e(\lambda,\epsilon_b) = \inf_{\mu\in\mathcal{P}_F(B)} \max_{a\in A} f_e(\epsilon_a,\mu).$$

Proof. Since B is a compact Hausdorff space and f_a is lower semicontinuous for every $a \in A$ it follows by Theorem 17 that

$$\sup_{\lambda\in\mathcal{P}_F(A)} \min_{b\in B} f_e(\lambda,\epsilon_b) = \inf_{\mu\in\mathcal{P}(B)} \sup_{a\in A} f_e(\epsilon_a,\mu). \qquad (22)$$

Using now A is compact and f_b is upper semicontinous for every $b \in B$ we obtain by relation (12) that

$$\inf_{\mu\in\mathcal{P}_F(B)} \max_{a\in A} f_e(\epsilon_a,\mu) = \max_{\lambda\in\mathcal{P}(A)} \inf_{b\in B} f_e(\lambda,\epsilon_b) \qquad (23)$$

Applying now Theorem 19 to the last parts of relations (23) and (22) yields the desired result. □

Using relation (14) it is easy to verify that the result of Theorem 21 is the same as the minimax result

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

and this minimax result is clearly a generalization of the minimax result of Von Neumann. We will now show some easy consequences of Theorem17, thereby generalizing earlier results to be found in the minimax literature. Before mentioning those generalizations we introduce for convenience the following class of functions.

Definition 22 *The function $f : A \times B \to \mathbb{R}$ belongs to the class $\mathcal{C}$ if*

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

Moreover, the function $f : A \times B \to \mathbb{R}$ belongs to the class $\mathcal{D}$ if

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

We first start with an improvement of the main minimax result proved by Kassay and Kolumban (cf.[10]). Observe the usual topological conditions are imposed beforehand.

Lemma 23 *Let B be a compact Hausdorff space and A an arbitrary set. If the function f_a is lower semicontinuous for every $a \in A$ then it follows that $f \in \mathcal{C}$ if and only if*

$$\sup_{\lambda\in\mathcal{P}_F(A)} \min_{b\in B} f_e(\lambda, \epsilon_b) = \min_{b\in B} \sup_{\lambda\in\mathcal{P}_F(A)} f_e(\lambda, \epsilon_b).$$

Proof. To show $f \in \mathcal{C}$ implies the desired equality we observe by Theorem 17 and $f \in \mathcal{C}$ that

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

Since f_a is lower semicontinuous for every $a \in A$ it follows that $\sup_{a\in A} f_a$ is also lower semicontinous and by the Lebesgue-Weierstrass theorem and B compact this yields

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

Applying now relation (14) yields the desired equality. To show the reverse implication, it follows immediately by Theorem 17 and our assumption that f belongs to $\mathcal{C}$. □

A second easy consequence is given by a characterization for which functions f actually a minimax result for f holds. This result generalizes for different sets of generalized convex functions the well known minimax results of Ky-Fan (cf.[5]), König (cf.[12]), Neumann (cf.[14]) and Jeyakumar (cf.[8]). The class of functions considered by these authors are a proper subclass of the set $\mathcal{C} \cap \mathcal{D}$.

Lemma 24 *Let B be a compact Hausdorff space and A an arbitrary set. If the function f_a is lower semicontinuous for every $a \in A$ then it follows that $f \in \mathcal{C} \cap \mathcal{D}$ if and only if*

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

Proof. To show $f \in \mathcal{C} \cap \mathcal{D}$ implies the desired equality we apply Lemma 23 and use $f \in \mathcal{D}$. The reverse implication is obvious using Theorem 17. $\square$

Finally we list some counterexamples showing that the conditions mentioned in Theorems 17, 19 and 21 cannot be deleted from these theorems. In this paper we have used three types of conditions. These conditions are given by:

1. (Topological) The function f_a is lower semicontinuous or the function f_b is upper semicontinuous.
2. (Compactness) The set A or B is a compact Hausdorff space.
3. (Boundedness) The function f is either bounded from above or below.

In the first counterexample we show that Theorem 19 is not correct if only conditions 2 and 3 hold. Observe this is also a counterexample for Theorem 17. Actually in this counterexample both sets A and B are compact metric spaces (hence condition 2 is replaced by a stronger condition) and the function f is uniformly bounded from above and below (also stronger than condition 3). However, the function f_a is not lower semicontinuous for every $a \in A$ and f_b is not upper semicontinous for some $b \in B$. Clearly by Theorem 19, for both A and B compact metric spaces and f is bounded from below or above, the minimax result should hold if either f_a is lower semicontinuous or f_b is upper semicontinuous.

Example 25 *Let* $A = B = [0,1]$ *and introduce the function* $f : [0,1] \times [0,1] \to \mathbb{R}$ *given by*

$$f(a,b) := \begin{cases} 1 & \textit{for } 0 < a < b \\ 1 & \textit{for } b = 0 \\ 0 & \textit{otherwise} \end{cases}$$

This function is bounded from above and below. Also for every $0 < a \leq 1$ *we obtain* $\{b \in B : f_a(b) \leq 0\} = (0,a]$ *and* $\{b \in B : f_0(b) \leq 0\} = (0,1]$ *and so* f_a *is not lower semicontinuous for every* $a \in A$*. Similarly for every* $0 < b < 1$ *it follows that*

$$\{a \in A : f_b(a) \geq 1\} = (0,b) \tag{24}$$

and so f_b *is not upper semicontinuous for every* $0 < b < 1$*. Also by relation (24) we obtain for* $\lambda \in \mathcal{P}(A)$ *and* $0 < b < 1$ *that*

$$0 \leq \int_A f_b d\lambda = \lambda((0,b))$$

and this shows $\inf_{b\in B} f_e(\lambda, \epsilon_b) = 0$ *for every* $\lambda \in \mathcal{P}(A)$. *At the same time, for* $\mu \in \mathcal{P}(B)$ *and* $0 < a < 1$ *it follows that*

$$1 \geq \int_B f_a d\mu = \mu(\{0\} \cup (a, 1])$$

and this implies $\sup_{a\in A} f_e(\epsilon_a, \mu) = 1$. *Hence the conclusion of Theorems 17 and 19 do not hold.*

In the next more complicated counterexample we construct an example with A and B compact metric spaces, f bounded from above and below, f_b is continuous for every $b \in B$ and f_a upper semicontinuous for every $a \in A$ (not lower semicontinuous) and show that the conclusion of Theorem 21 does not hold.

Example 26 *Let* $A = B = [0, 1]$ *and consider for any* $n \in \mathbb{N}$ *a continuous mapping* $\phi_n : [0, 1] \to \Pi_{i=1}^n [0, 1]$ *given by*

$$\phi_n(t) := (\phi_{n1}(t), ..., \phi_{nn}(t))$$

onto the n-dimensional cube $\Pi_{i=1}^n [0, 1]$ *satisfying* $\phi_n(0) = \phi_n(1) = \mathbf{0}$. *To construct such a continuous surjective curve we use for* $n = 2$ *the so-called Peano space filling curve* ϕ_2 *(cf.[7]) and use induction on* n *and the composition of functions*

$$[0, 1] \stackrel{\phi_2}{\to} [0, 1] \times [0, 1] \stackrel{h}{\to} \Pi_{i=1}^n [0, 1]$$

with $h(s, t) := (p(s), \phi_{n-1}(t))$ *and*

$$p(s) = \begin{cases} s & \text{for } 0 \leq s \leq \frac{1}{2} \\ 1 - s & \text{for } \frac{1}{2} \leq s \leq 1 \end{cases} .$$

Introduce now the nonnegative function $f : [0, 1] \times [0, 1] \to [0, \infty)$ *given by*

$$f(a, b) := \begin{cases} 2\exp(1) & \text{for } b = 0 \\ 2\exp(1)\Pi_{k=1}^n p_{nk}(a, b)^{\frac{1}{n}} & \text{for } b \in (0, 1] \end{cases}$$

with $p_{nk} : [0, 1] \times (0, 1]$ *defined by*

$$p_{nk}(a, b) := |a - \phi_{nk}(2^n b - 1)| \text{ for } b \in (2^{-n}, 2^{-n+1}], n \in \mathbb{N}.$$

To determine an upperbound on the function f *we observe for every* $(a, b) \in [0, 1] \times (0, 1]$ *that* $p_{nk}(a, b) \leq 1$ *and this implies*

$$\sup_{(a,b)\in\{0,1]\times[0,1]} f(a, b) \leq 2\exp(1). \tag{25}$$

Hence we have shown that the function f is bounded from above and below. To list the topological properties of the function f it is obvious that the function f_b is continuous for every $0 \leq b \leq 1$. Also for every $a \in [0,1]$ and $b_0 \in (0,1]$ it is clear that $\lim_{b \to b_0} f_a(b) = f_a(b_0)$ and by relation (25) $\limsup_{b \downarrow 0} f_a(b) \leq 2\exp(1) = f_a(0)$. This shows that f_a is upper semicontinuous, and to prove that f_a is not continuous we consider for $a \in [0,1)$ the sequence $b_n = 2^{-n+1}, n \in \mathbb{N}$. Using $\phi_{nk}(0) = 0, 1 \leq k \leq n$ it follows that

$$\lim_{n \uparrow \infty} f_a(b_n) = 2a\exp(1) < f_a(0)$$

and so f_a is not lower semicontinuous for $0 \leq a < 1$. We will now verify for every $\lambda \in \mathcal{P}_F(A)$ that

$$\min_{b \in [0,1]} f_e(\lambda, \epsilon_b) = 0. \tag{26}$$

To show this, we observe for every $\lambda \in \mathcal{P}_F(A)$ that

$$f_e(\lambda, \epsilon_b) = \sum_{i=1}^{n} \lambda_i f(a_i, b) \tag{27}$$

for some finite set $\{a_1, ..., a_n\} \subseteq \Pi_{i=1}^n [0,1]$ and positive numbers $\lambda_i, 1 \leq 1 \leq n$ satisfying $\sum_{i=1}^n \lambda_i = 1$. Since the mapping $\phi_n : [0,1] \to \Pi_{i=1}^n [0,1]$ with $\phi_n(0) = \phi_n(1) = \mathbf{0}$ is surjective onto the hypercube, there exists some $0 < t_0 < 1$ with $\phi_n(t_0) = (a_1, ..., a_n)$, and this implies for every $1 \leq i \leq n$ that

$$\Pi_{k=1}^n |a_i - \phi_{nk}(2^n b_0 - 1)|^{\frac{1}{n}} = 0$$

with $2^{-n} < b_0 := (t_0 + 1)2^{-n} \leq 2^{-n+1}$. Hence it follows that $f(a_i, b_0) = 0$ for every $1 \leq i \leq n$, and so by relation (27) we obtain $f_e(\lambda, \epsilon_{b_0}) = 0$. Applying the nonnegativity of the function f the result in relation (27) now follows. To give a lower bound on the value

$$\sup_{a \in [0,1]} f_e(\epsilon_a, \mu) = \sup_{a \in [0,1]} \int_0^1 f_a d\mu$$

for every $\mu \in \mathcal{P}(B)$ we continue as follows. Since the function $x \to \ln(x)$ is concave on $(0, \infty)$ it follows by Jenssen's inequality (cf.[16]) that

$$\ln \int_0^1 \psi(t) dt \geq \int_0^1 \ln(\psi(t)) dt$$

for every positive continuous function ψ on $[0,1]$. This shows for every $2^{-n} < b \leq 2^{-n+1}$, $n \in \mathbb{N}$ and $b_k := \phi_{nk}(2^n b - 1)$ that

$$\int_0^1 f_b(a)da \geq 2\exp(\int_0^1 \ln(\Pi_{k=1}^n |a-b_k|^{\frac{1}{n}})da + 1) \tag{28}$$
$$= 2\exp(\frac{1}{n}\sum\nolimits_{k=1}^n \int_0^1 \ln(|a-b_k|)da + 1).$$

Since $0 \leq b_k \leq 1$ for every $1 \leq k \leq n$ and the function $p : [0,1] \to \mathbb{R}$ given by

$$p(x) = \int_0^1 \ln(|a-x|)da = \int_0^x \ln(a)da + \int_0^{1-x} \ln(a)da$$

achieves its minimum at $x = \frac{1}{2}$ (check by differentiation), we obtain by relation (28) that

$$\int_0^1 f_b(a)da \geq 2\exp(\int_0^1 \ln(|a - \frac{1}{2}|)da + 1) = 1.$$

Hence we have shown that $\int_0^1 f_b(a)da \geq 1$ for every $0 \leq b \leq 1$ and by the Fubini-Tonelli theorem this yields for every $\mu \in \mathcal{P}(B)$ that

$$\int_0^1 f_e(\epsilon_a, \mu)da = \int_0^1 (\int_0^1 f_a d\mu)da = \int_0^1 (\int_0^1 f_b(a)da)d\mu \geq 1.$$

This shows $\sup_{a\in[0,1]} f_e(\epsilon_a, \mu) \geq \int_0^1 f_e(\epsilon_a, \mu)da \geq 1$ for every $\mu \in \mathcal{P}(B)$ and so the conclusion of Theorem 21 does not hold.

We will now consider two examples which show that the compactness assumption given by condition 2 cannot be deleted in the above theorems. In the first counterexample of this kind we show that there exists open sets A and B and a uniformly bounded continuous function $f : A \times B \to [0,1]$ (much stronger than conditions 1 and 3!), for which the conclusion of Theorems 17 or 19 does not hold.

Example 27 *Let $A = B = (0,1)$ and introduce the function $f : A \times B \to [0,1]$ given by $f(a,b) := h(1 - \frac{b}{a})$ with $h : \mathbb{R} \to [0,1]$ defined by*

$$h(t) = \begin{cases} 1 & \text{for } t \leq 0 \\ 1-2t & \text{for } 0 < t \leq \frac{1}{2} \\ 0 & \text{for } t > \frac{1}{2} \end{cases} .$$

Since the function h is continuous on $\mathbb{R}$ it follows that also the function $f : A \times B \to \mathbb{R}$ is continuous on $A \times B$. Moreover, for every $0 < b < \frac{1}{2}$ and $\lambda \in \mathcal{P}(A)$ we obtain

$$0 \leq \int_A f_b d\lambda \leq \lambda((0, 2b])$$

and so $\inf_{b \in B} f_e(\lambda, \epsilon_b) = 0$ for every $\lambda \in \mathcal{P}(A)$. Also we obtain for every $0 < a < \frac{1}{2}$ and $\mu \in \mathcal{P}(B)$ that

$$\mu([a, 1)) \leq \int_B f_a d\mu \leq 1$$

and this shows $\sup_{a \in A} f_e(\epsilon_a, \mu) = 1$ for every $\mu \in \mathcal{P}(B)$. Hence the conclusions of Theorems 17 and 19 do not hold.

In the second counterexample related to the compactness assumption we construct a continuous function f on $A \times B$, bounded from above and below, with A compact and B not, and show that the conclusion of Theorem 21 does not hold.

Example 28 *Let $A = [0, 1]$ and $B = (0, 1]$ and consider the function f defined in Example 26. It is easy to see that this function f is continuous on $A \times B$. As in Example 26 we obtain $\inf_{0 < b \leq 1} f_e(\lambda, \epsilon_b) = 0$ for every $\lambda \in \mathcal{P}_F(A)$ and $\sup_{a \in [0,1]} f_e(\epsilon_a, \mu) \geq 1$ for every $\mu \in \mathcal{P}(B)$. This shows that the conclusion of Theorem 21 does not hold.*

Finally in the last counterexample we show that the boundedness condition in Theorem 19 cannot be omitted. Actually in this counter example we construct compact sets A and B together with a function f neither bounded from above or below satisfying f_a and f_b are continuous for every $a \in A$ and $b \in B$ and show that the Fubini-Tonelli theorem does not hold. This implies that also the conclusion of Theorem 19 does not hold.

Example 29 *Let $A = B = [0, 1]$. We will now construct a function f satisfying f_a and f_b continuous for every $a \in A$ and $b \in B$, which is not bounded from above or below. To carry out the construction of this function consider a continuously differentiable function $\theta : \mathbb{R} \to \mathbb{R}$ satisfying $\theta(t) < 0$ for $0 < t < \frac{1}{2}$ and $\theta(t) = 0$ otherwise. Clearly such a function exists and introduce now the function $h : \mathbb{R} \to \mathbb{R}$ given by*

$$h(t) = \begin{cases} \frac{\theta(t) - \theta(t - \frac{1}{2})}{t} & \text{for } t \neq 0 \\ 0 & \text{for } t = 0. \end{cases} \tag{29}$$

Since θ is continuously differentiable with $\theta'(t) = 0$ for every $t < 0$ we obtain by relation (29)

$$\lim_{t\downarrow 0} h(t) = \lim_{t\downarrow 0} \frac{\theta(t)}{t} = \theta'(0) = \lim_{t\uparrow 0} \theta'(t) = 0 = h(0)$$

and $\lim_{t\uparrow 0} h(t) = 0 = h(0)$. This shows that the function h is continuous at 0 and since this function is clearly continuous at $t \neq 0$ the function h is continuous in $\mathbb{R}$ and satisfies by relation (29) supp$(h) \subseteq [0,1]$. If $\frac{1}{2} \leq b \leq 1$, it follows by relation (29) that

$$\int_b^1 th(t)dt = \int_b^1 \theta(t) - \theta(t - \frac{1}{2})dt = -\int_{b-\frac{1}{2}}^{\frac{1}{2}} \theta(t)dt \geq 0$$

and for $0 \leq b < \frac{1}{2}$

$$\int_b^1 th(t)dt = \int_b^1 \theta(t) - \theta(t - \frac{1}{2})dt = -\int_0^b \theta(t)dt \geq 0.$$

Also for this function h we obtain

$$\int_0^1 h(t)dt = \int_0^{\frac{1}{2}} \theta(t)(t^{-1} - (t + \frac{1}{2})^{-1})dt < 0$$

and by scaling we may assume that we have constructed a continuous function h with supp$(h) \subseteq [0,1]$ satisfying

$$\int_0^1 h(t)dt = -1 \text{ and } \int_b^1 th(t)dt \geq 0 \text{ for every } b \in [0,1]. \tag{30}$$

Introduce now the function $g : A \times B \rightarrow \mathbb{R}$ given by

$$g(a,b) = \begin{cases} a^{-3}h(\frac{b}{a}) & \text{for } 0 < a \leq 1, 0 \leq b \leq 1 \\ 0 & \text{for } a = 0, 0 \leq b \leq 1. \end{cases}$$

The function $g_a : [0,1] \rightarrow \mathbb{R}$ given by $g_a(b) := g(a,b)$ is continuous for every $0 \leq b \leq 1$ and the function $g_b : [0,1] \rightarrow \mathbb{R}$ given by $g_b(a) := g(a,b)$ is also continuous for every $0 \leq b \leq 1$. Also by relation (30) it follows that

$$\int_0^1 g(a,b)db = -a^{-2} \leq -1$$

for every $0 < a < 1$, $\int_0^1 g(a,b)db = 0$ for $a = 0$ and

$$\int_0^1 g(a,b)da = b^{-2} \int_b^1 th(t)dt \geq 0$$

for every $0 \leq b \leq 1$. *Now we finally introduce the function* $f : A \times B \to \mathbb{R}$ *given by*

$$f(a,b) = g(a,b) + g(1-a,b) - g(b,a) - g(1-b,a). \tag{31}$$

By the previous observations it follows that the functions f_a *and* f_b *are continuous. Also one can check by relation (31) that* $\int_0^1 f(a,b)db \leq -1$ *for every* $0 \leq a \leq 1$ *and* $\int_0^1 f(b,a)da \geq 1$ *for every* $0 \leq b \leq 1$. *This shows*

$$\sup_{\lambda \in \mathcal{P}(A)} \inf_{b \in B} f_e(\lambda, \epsilon_b) \geq \inf_{b \in B} \int_0^1 f(a,b)da \geq 1$$

and

$$\inf_{\mu \in \mathcal{P}(B)} \sup_{a \in A} f_e(\epsilon_a, \mu) \leq \sup_{a \in A} \int_0^1 f(a,b)db \leq -1$$

and so the concludion of Theorem 19 does not hold.

This concludes our discussion of the generalizations of the minimax results of Wald, Ville and Von Neumann. An important issue related to the above results would be to derive computational procedures for finding good approximations of the optimal strategies within the set of Borel probability measures. This might be a topic of future research.

References

[1] Aubin, J.B. *Optima and Equilibra: An Introduction to Nonlinear Analysis*, volume 140 of *Graduate Texts in Mathematics*. Springer Verlag, Berlin, 1993.

[2] Choquet, G. *Lectures on Analysis Volume 1: Integration and Topological Vector Spaces*. W.A.Benjamin, London, 1969.

[3] Conway, J.B. *A course in Functional Analysis*. Graduate texts in Mathematics. Springer Verlag, New York, 1990.

[4] Dudley, R.M. *Real Analysis and Probability*. Wadsworth and Brooks/Cole, Pacific Grove, 1989.

[5] Fan, K. Minimax theorems. *Proc.Nat.Acad.Sci.U.S.A.*, 39:42–47, 1953.

[6] Frenk, J.B.G., Kassay., G. and J.Kolumban. Equivalent Results in Minimax theory. To appear in European Journal of Operational Research.

[7] Hocking, J.G. and G.S.Young. *Topology*. Addison-Wesley series in Mathematics. Addison-Wesley, 1961.

[8] Jeyakumar, V. A generalization of a minimax theorem of Fan via a theorem of the alternative. *Journal of Optimization Theory and Applications*, 48(3):525–533, 1986.

[9] Jones, A.J. *Game Theory*. Horwood Publishers, Cichester, 2000.

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

[11] Kneser, H. Sur un theoreme fondamental de la theorie des jeux. *Comptes Rendus Acad.Sci.Paris*, 234:2418–2420, 1952.

[12] König, H. Über das von Neumannsche Minimax-Theorem. *Arch.Math.*, 19:482–487, 1968.

[13] Lassonde, M. Hahn Banach theorems for convex functions. In B.Ricceri and S.Simons, editors, *Minimax Theory and Applications*, pages 135–145. Kluwer Academic Publishers, Dordrecht, 1996.

[14] Neumann, M. Bemerkungen zum von Neumannschen minimaxtheorem. *Arch.Math.*, 29:96–105, 1977.

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

[16] van Tiel, J. *Convex Analysis: An Introductory Text*. Wiley, New York, 1988.

[17] 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.

[18] von Neumann, J. Zur theorie der gessellschaftsspiele. *Math.Ann*, 100:295–320, 1928.

[19] Vorob'ev, N.N. *Game Theory:Lectures for Economists and Systems Scientists*. Springer Verlag, New York, 1977.

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

[21] Wald, A. *Statistical Decision Functions, 2nd Edition*. Chelsea, New York, 1971.