NOTES

Edited by **Ed Scheinerman**

A Simpler Proof of the Von Neumann Minimax Theorem

Hichem Ben-El-Mechaiekh and Robert W. Dimand

Abstract. This note provides an elementary and simpler proof of the Nikaidô-Sion version of the von Neumann minimax theorem accessible to undergraduate students. The key ingredient is an alternative for quasiconvex/concave functions based on the separation of closed convex sets in finite dimension, a result discussed in a first course in optimization or game theory.

1. INTRODUCTION. The minimax theorem, proving that a zero-sum two-person game must have a solution, was the starting point of the theory of strategic games as a distinct discipline. It is well known that John von Neumann [**15**] provided the first proof of the theorem, settling a problem raised by Emile Borel (see [**2, 8**] for detailed historical accounts).

Proofs of the minimax theorem based on the Brouwer fixed point theorem or the Knaster-Kuratowski-Mazurkiewicz (KKM) principle are elegant and short (see, e.g., [**2, 8**]) but cannot be considered elementary. Indeed, both fundamental results require substantial groundwork going beyond the typical North American undergraduate curriculum (e.g., such deep results as the nonretraction theorem of the unit ball onto its boundary in a Euclidean space or Sperner's lemma on the existence of complete labelings for a Euclidean simplex). A number of elementary proofs for the nonlinear case are worth mentioning. H. Brézis [**4**] and G. Garnir [**5**] provide elementary and simple proofs in the convex/concave case. Another elementary proof is given by I. Joó [**6**], who bases his argument on a lemma of F. Riesz on the nonempty intersection of a family of compact sets having the finite intersection property. A fourth, elementary but in our opinion not simple proof is due to J. Kindler and has appeared in the MONTHLY [**7**].

The aim of this note is to provide a simpler and very elementary proof of the Nikaidô-Sion version of the minimax theorem which is accessible to students in an undergraduate course in game theory. The proof is based on a result of Victor Klee [**9**] on convex covers of closed convex subsets of a Euclidean space. Klee's result derives in an easy way from the separation of convex sets in Euclidean spaces, a result often discussed in a first course on continuous optimization (which, in many curricula, is a prerequisite to the game theory course).

2. AN ALTERNATIVE FOR NONLINEAR SYSTEMS OF INEQUALITIES AND MINIMAX. Although von Neumann originally formulated the minimax theorem for linear forms, he quickly became well aware that convexity of level sets of the functionals involved was sufficient for the proof to hold true, thus providing an early definition of the concept of quasiconvexity before its time [**15**]. It was not until 1954 that H. Nikaidô [**11**] and later in 1958 M. Sion [**13**] formulated the minimax theorem for quasiconvex/concave and lower/upper semicontinuous functions.

http://dx.doi.org/10.4169/amer.math.monthly.118.07.636

Throughout this paper, all topological vector spaces are assumed to be vector spaces over the field $\mathbb{R}$ of real numbers equipped with a Hausdorff topology.

Definition 1. A real function $f : X \longrightarrow \mathbb{R}$ defined on a subset X of a topological vector space is:

(i) *quasiconvex* if $\forall\lambda \in \mathbb{R}$, the level set $\{x \in X : f(x) < \lambda\}$ is a convex subset of X;

(ii) *upper semicontinuous* (u.s.c.) if $\forall\lambda \in \mathbb{R}$, the level set $\{x \in X : f(x) < \lambda\}$ is an open subset of X.

A function f is *quasiconcave* if $-f$ is quasiconvex; it is *lower semicontinuous* (l.s.c.) if $-f$ is u.s.c. Note that f is quasiconvex on X if and only if $f(\mu x_1 + (1 - \mu)x_2) \leq \max\{f(x_1), f(x_2)\}$ for all $x_1, x_2 \in X$ and all $\mu \in [0, 1]$. Convex functions are clearly quasiconvex.

The connection between the minimax theorem and the solvability of systems of linear inequalities and the crucial role played by convexity were first outlined by Jean André Ville in 1938, when he published the first elementary proof of the minimax theorem in an appendix to lecture notes of Emile Borel's Sorbonne course on the application of probability theory to games of chance [**14**] (Ville was a student of Borel).

Ville's fundamental result for the proof of the minimax theorem is in fact—as von Neumann and Morgenstern said in their celebrated *Theory of Games and Economic Behavior* [**16**]—an alternative for the solvability of linear systems of inequalities. This linear alternative can be extended to an alternative for nonlinear systems of inequalities: the key ingredient, here, for the Nikaidô-Sion versions of the von Neumann theorem.

Theorem 2 (An Alternative for Nonlinear Systems of Inequalities). *Let X and Y be two convex subsets of topological vector spaces, with Y compact, and let $\tilde{f}$, f, g, $\tilde{g} : X \times Y \longrightarrow \mathbb{R}$ be four functions satisfying:*

(i) $\tilde{f}(x, y) \leq f(x, y) \leq g(x, y) \leq \tilde{g}(x, y)$ *for all* $(x, y) \in X \times Y$*;*

(ii) $y \mapsto \tilde{f}(x, y)$ *is lower semicontinuous and quasiconvex on* Y*, for each fixed* $x \in X$*;*

(iii) $x \mapsto f(x, y)$ *is quasiconcave on* X*, for each fixed* $y \in Y$*;*

(iv) $y \mapsto g(x, y)$ *is quasiconvex on* Y*, for each fixed* $x \in X$*; and*

(iv) $x \mapsto \tilde{g}(x, y)$ *is upper semicontinuous and quasiconcave on* X*, for each fixed* $y \in Y$*.*

Then for any $\lambda \in \mathbb{R}$*, the following alternative holds:*

(A) *there exists* $\bar{x} \in X$ *such that* $\tilde{g}(\bar{x}, y) \geq \lambda$*, for all* $y \in Y$*; or*

(B) *there exists* $\bar{y} \in Y$ *such that* $\tilde{f}(x, \bar{y}) \leq \lambda$*, for all* $x \in X$*.*

Ville's result [**14**] corresponds to the case $\lambda = 0$, $X = \mathbb{R}^n_+$ (the positive cone in $\mathbb{R}^n$), $Y = \Delta$ (the standard n-simplex, clearly a closed, bounded, and hence compact set), and $\tilde{f}(x, y) = f(x, y) = g(x, y) = \tilde{g}(x, y) = -\sum_{j=1}^{p} y_j \varphi_j(x)$, a bilinear form in the variables $(x, y = (y_j)) \in \mathbb{R}^n_+ \times \Delta$ ($\varphi_j(x)$ being linear forms).

Corollary 3. *Under the same hypotheses as in Theorem 2, we have:*

$$\tilde{\alpha} = \sup_X \inf_Y \tilde{g}(x, y) \geq \min_Y \sup_X \tilde{f}(x, y) = \tilde{\beta}.$$

Proof. Assume that $\tilde{\alpha} < \tilde{\beta}$. Let λ be an arbitrary but fixed real number strictly between $\tilde{\alpha}$ and $\tilde{\beta}$. By Theorem 2, either there exists $\bar{y} \in Y$ such that $\tilde{f}(x, \bar{y}) \leq \lambda$ for all $x \in X$, and thus $\tilde{\beta} \leq \lambda < \tilde{\beta}$ which is impossible, or there exists $\bar{x} \in X$ such that $\tilde{g}(\bar{x}, y) \geq \lambda$ for all $y \in Y$, and thus $\tilde{\alpha} \geq \lambda > \tilde{\alpha}$, which is absurd. Hence $\tilde{\alpha} \geq \tilde{\beta}$. ∎

Note that the inequality in Corollary 3 is in fact equivalent to the alternative for nonlinear systems of inequalities. Indeed, if $\tilde{\alpha} \geq \tilde{\beta}$ and conclusion (A) in Theorem 2 fails, then $\exists \bar{y} \in Y$ such that $\sup_X \tilde{f}(x, \bar{y}) \leq \lambda$ and (B) is thus satisfied. Note also that the $\min_Y$ in Corollary 3 is justified by the fact that the supremum of a family of lower semicontinuous functions is also lower semicontinuous, and that a lower semicontinuous function on a compact domain achieves its minimum.

The Nikaidô-Sion formulation [**11, 13**] of the minimax theorem of von Neumann [**15**, Theorem 3.4] follows immediately with $\tilde{f} = f = g = \tilde{g}$.

Theorem 4. *Let X and Y be convex subsets of topological vector spaces, with Y compact, and let f be a real function on $X \times Y$ such that:*

(i) *$x \mapsto f(x, y)$ is upper semicontinuous and quasiconcave on X for each fixed $y \in Y$; and*
(ii) *$y \mapsto f(x, y)$ is lower semicontinuous and quasiconvex on Y for each fixed $x \in X$.*

Then:

$$\sup_X \min_Y f(x, y) = \min_Y \sup_X f(x, y).$$

Proof. Let $\tilde{\alpha} = \alpha = \sup_X \min_Y f(x, y)$ and $\tilde{\beta} = \beta = \min_Y \sup_X f(x, y)$. Since the inequality $\alpha \leq \beta$ is always true, it follows from Corollary 3 that $\alpha = \beta$. ∎

3. A SIMPLER PROOF OF THE ALTERNATIVE. The basis for the proof of the alternative for nonlinear systems of inequalities is the following result of Victor Klee [**9**] (see also Claude Berge [**3**] for extensions).

Lemma 5. *Let C and $C_1, \ldots, C_n$ be closed convex sets in a Euclidean space satisfying:* (i) *$C \cap \bigcap_{i=1, i\neq j}^n C_i \neq \emptyset$ for $j = 1, 2, \ldots, n$; and* (ii) *$C \cap \bigcap_{i=1}^n C_i = \emptyset$. Then $C \not\subseteq \bigcup_{i=1}^n C_i$.*

Proof. (We reproduce Victor Klee's proof as, truly, one cannot do any better.) One may assume with no loss of generality that the sets C and C_i, $i = 1, \ldots, n$, are all compact (otherwise, one may replace C by the compact convex finite polytope $C' := \mathrm{Conv}\{y_j : j = 1, \ldots, n\}$, where the $y_j \in C \cap \bigcap_{i=1, i\neq j}^n C_i$ are provided by (i), and C_i by $C_i' := C_i \cap C'$). The proof is by induction on n.

If $n = 1$, (i) asserts that C is nonempty and (ii) that C and C_1 are disjoint. Thus, clearly $C \not\subseteq C_1$. Suppose that the thesis holds for $n = k - 1$ and consider the case $n = k$, i.e., assume that $\{C, \{C_i\}_{i=1}^k\}$ is a collection of compact convex sets such that for $j = 1, \ldots, k$, $C \cap \bigcap_{i=1, i\neq j}^k C_i \neq \emptyset$, and $(C \cap C_k) \cap \bigcap_{i=1}^{k-1} C_i = \emptyset$. The disjoint compact convex sets $(C \cap C_k)$ and $\bigcap_{i=1}^{k-1} C_i$ can be strictly separated by a hyperplane H (a short and elementary proof of the separation theorem in finite dimensions can be found in [**10**]). Putting $C' := H \cap C$ and $C_i' := H \cap C_i$, it follows that $C' \cap \bigcap_{i=1}^{k-1} C_i' = \emptyset$. Moreover, for a given arbitrary $j_0 \in \{1, \ldots, k - 1\}$, let $y_0 \in C \cap \bigcap_{i=1, i\neq j_0}^k C_i$, so that $y_0 \in C \cap C_k$, and let $y_k \in C \cap \bigcap_{i=1}^{k-1} C_i$ be arbitrary. Clearly, the points y_0 and y_k are strictly separated by H. The intersection $\bar{z}$ of the line segment $[y_0, y_k]$ with H belongs to C as well as to $\bigcap_{i=1, i \neq j_0}^{k-1} C_i$. The integer j_0 being arbitrary, hypotheses (i) and (ii) are verified for the collection $\{C', \{C'_i\}_{i=1}^{k-1}\}$. By the induction hypothesis, $C' = H \cap C \not\subseteq \bigcup_{i=1}^{k-1} C'_i = \bigcup_{i=1}^{k-1} C_i \cap H$. Since $(H \cap C) \cap C_k = \emptyset$, it follows that $H \cap C \not\subseteq \bigcup_{i=1}^{k} C_i \cap H$, and thus $C \not\subseteq \bigcup_{i=1}^{k} C_i$. ∎

Proof of Theorem 2. Suppose that the alternative for nonlinear systems of inequalities does not hold, i.e., both (A) and (B) fail. This amounts to saying that the collection of open level sets $\{U_y := \{x \in X : \tilde{g}(x, y) < \lambda\} : y \in Y\}$ is a cover of X and the collection $\{V_x := \{y \in Y : \tilde{f}(x, y) > \lambda\} : x \in X\}$ is an open cover of Y (U_y and V_x are open due to the semicontinuity hypotheses).

Since Y is compact, $\{V_x : x \in X\}$ admits a finite subcover $\{V_{x_k} : k = 1, \dots, m\}$. The convex polytope $C := \mathrm{Conv}\{x_k : k = 1, \dots, m\}$ lies in a finite-dimensional subspace L of the underlying linear topological space containing X. The subspace L being homeomorphic to a Euclidean space (see, e.g., Rudin [**12**]), C is also compact. Thus, it can be covered by a finite subcollection $\{U_i = U_{y_i} \cap L : i = 1, \dots, n\}$. One can drop indices from $i = 1, \dots, n$ so as to make the cover $\{U_i\}$ minimal, in the sense that $C \subseteq \bigcup_{i=1}^{n} U_i$ but $C \not\subseteq \bigcup_{i=1, i \neq j}^{n} U_i$ for $j = 1, \dots, n$. For $i = 1, 2, \dots, n$, let $C_i := \{x \in L : \tilde{g}(x, y_i) \geq \lambda\} = L \setminus U_i$, a closed convex subset of L. The fact that C is covered by $\{U_i\}$ is precisely the emptiness of the intersection $C \cap \bigcap_{i=1}^{n} C_i$. The minimality of $\{U_i\}$ is nothing else than $C \cap \bigcap_{i=1, i \neq j}^{n} C_i \neq \emptyset$ for $j = 1, 2, \dots, n$. Lemma 5 implies the existence of $x_0 \in C$ with $x_0 \notin C_i$, and thus $g(x_0, y_i) \leq \tilde{g}(x_0, y_i) < \lambda$ for $i = 1, 2, \dots, n$. The quasiconvexity of $g(x_0, \cdot)$ implies the existence of $x_0 \in C$ such that

$$g(x_0, y) < \lambda, \forall y \in D := \mathrm{Conv}\{y_i : i = 1, \dots, n\}.$$

A similar argument (left to the reader) yields the existence of $y_0 \in D$ such that:

$$f(x, y_0) > \lambda, \quad \forall x \in C.$$

Thus,

$$\lambda < f(x_0, y_0) \leq g(x_0, y_0) < \lambda,$$

a contradiction. ∎

4. CONCLUDING REMARKS. A related alternative was provided in [**1**], based on the Brouwer fixed point theorem.

One can opt for a more geometric expression of the infsup inequality in Corollary 3 as an alternative between the existence of a coincidence between "assymmetric" relations representing level sets of numerical functions, and the existence of what game theorists call *maximal elements* for those relations. Such a coincidence result is instrumental in proving the existence of an equilibrium for generalized games involving more than one preference relation. We elect to include it here, not only for the simplicity of its proof, but also in order to come back to a geometric point of view which might be more appealing to some students.

Given a subset A in a cartesian product of sets $X \times Y$, denote by A^c its complement, and by $A[x]$ and $A[y]$ the sections $\{y \in Y : (x, y) \in A\}$ and $\{x \in X : (x, y) \in A\}$ respectively; also, denote by A^{-1} the inverse relation $\{(y, x) \in Y \times X : (x, y) \in A\}$ and by 1_A the characteristic function of A: $1_A(x, y) := 1$ if $(x, y) \in A$, and 0 otherwise.

Let us now call a pair of relations $(\tilde{A}, A)$ in a cartesian product $X \times Y$ of subsets of topological vector spaces a *von Neumann pair* if

(i) $\tilde{A} \subseteq A$,
(ii) $A[y]$ is convex, for all $y \in Y$, and
(iii) $\tilde{A}[x]$ is open and $Y \setminus \tilde{A}[x]$ is convex for all $x \in X$.

Theorem 6. *Let $(\tilde{A}, A)$ and $(\tilde{B}, B)$ be two pairs of relations in the cartesian product $X \times Y$ of two nonempty convex subsets X and Y in topological vector spaces. Assume that Y is compact and that $(\tilde{A}, A)$ and $(\tilde{B}^{-1}, B^{-1})$ are von Neumann pairs.*

Then one of the following must hold:

(1) (Maximal Element) $\tilde{A}[\bar{y}] = \emptyset$ *for some* $\bar{y} \in Y$ *or* $\tilde{B}[\bar{x}] = \emptyset$ *for some* $\bar{x} \in X$.
(2) (Coincidence) $A \cap B \neq \emptyset$.

Proof. Assume possibility (1) fails, that is:

$$\tilde{A}[y] \neq \emptyset \text{ for all } y \in Y \text{ and } \tilde{B}[x] \neq \emptyset \text{ for all } x \in X.$$

We show that A and B must have a coincidence.

Define $\tilde{f}, f, g, \tilde{g} : X \times Y \longrightarrow \mathbb{R}$ as:

$$\tilde{f} := 1_{\tilde{A}},\ f := 1_A,\ g := 1_{B^c},\ \text{and } \tilde{g} := 1_{\tilde{B}^c},$$

and let $\lambda = 1/2$ in Theorem 2.

The reader can readily verify that:

- $\tilde{f} \leq f$ and $g \leq \tilde{g}$;
- the failure of possibility (1) amounts to the failure of the infsup inequality $\sup_X \inf_Y \tilde{g}(x, y) \geq \min_Y \sup_X \tilde{f}(x, y)$ in Corollary 3, and equivalently to the failure of the conclusion of Theorem 2 ($\tilde{A}[y] \neq \emptyset$ for all $y \in Y$ is the negation of alternative (B), while $\tilde{B}[x] \neq \emptyset$ for all $x \in X$ negates alternative (A));
- all hypotheses (ii) to (iv) of Theorem 2 hold true (for arbitrary $\lambda \in \mathbb{R}$, level sections of the numerical functions defined above are either $\emptyset$, all of X or Y, or the appropriate sections of the relations involved).

Naturally, the only remaining possibility is that the middle inequality in hypothesis (i) of Theorem 2 fails, that is:

$$\exists (x_0, y_0) \in X \times Y \text{ with } g(x_0, y_0) < f(x_0, y_0).$$

This can only happen when $g(x_0, y_0) = 0 < \lambda = \frac{1}{2} < 1 = f(x_0, y_0)$, i.e.,

$$(x_0, y_0) \in A \cap B. \qquad \blacksquare$$

ACKNOWLEDGMENTS. We wish to thank an anonymous referee for constructive remarks and the correction of a few errors, and for bringing to our attention references [**4, 5**], where other elementary proofs for the convex/concave cases of the minimax theorem are provided.

REFERENCES

1. H. Ben-El-Mechaiekh, P. Deguire, and A. Granas, Points fixes et coincidences pour les fonctions multivoques I (applications de Ky Fan), *C. R. Acad. Sci. Paris* **295** (1982) 337–340.
2. H. Ben-El-Mechaiekh and R. Dimand, Von Neumann, Ville, and the Minimax Theorem, *Int. Game Theory Rev.* **12** (2010) in print. http://dx.doi.org/10.1142/S0219198910002556

3. C. Berge, Sur une propriété combinatoire des ensembles convexes, *C. R. Acad. Sci. Paris* **248** (1959) 2698–2699.
4. H. Brézis, *Opérateurs Maximaux Monotones et Semi-Groupes de Contractions dans les Espaces de Hilbert*, North-Holland Mathematics Studies, no. 5, North-Holland, Amsterdam, 1973.
5. H. G. Garnir, *Fonctions de Variables Réelles, tome I et tome II*, Vander, Louvain, 1970.
6. I. Joó, A simple proof for von Neumann's minimax theorem, *Acta Sci. Math.* **42** (1980) 91–94.
7. J. Kindler, A simple proof of Sion's minimax theorem, *Amer. Math. Monthly* **112** (2005) 356–358. http://dx.doi.org/10.2307/30037472
8. T. H. Kjeldsen, John von Neumann's conception of the minimax theorem: A journey through different mathematical contexts, *Arch. Hist. Exact Sci.* **56** (2001) 39–68. http://dx.doi.org/10.1007/s004070100041
9. V. L. Klee, Jr., On certain intersection properties of convex sets, *Canad. J. Math.* **3** (1951) 272–275. http://dx.doi.org/10.4153/CJM-1951-031-2
10. M. Magill and M. Quinzii, *Theory of Incomplete Markets*, MIT Press, Cambridge, MA, 1996.
11. H. Nikaidô, On von Neumann's minimax principle, *Pacific J. Math.* **4** (1954) 65–72.
12. W. Rudin, *Functional Analysis*, 2nd ed., McGraw-Hill, New York, 1991.
13. M. Sion, On general minimax theorems, *Pacific J. Math.* **8** (1958) 171–176.
14. J. A. Ville, Sur la théorie générale des jeux où intervient l'habileté des joueurs, in *Traité du Calcul des Probabilités et de ses Applications, Applications aux Jeux de Hasard*, vol. IV, 2, É. Borel, et al., eds., Gauthier-Villars, Paris, 1938, 105–113.
15. J. von Neumann, Sur la théorie des jeux, *C. R. Acad. Sci. Paris* **186** (1928) 1689–1691.
16. J. von Neumann and O. Morgenstern, *Theory of Games and Economic Behavior*, Princeton University Press, Princeton, NJ, 1944.

Department of Mathematics, Brock University, St. Catharines, Ontario, L2S 3A1, Canada
hmechaie@brocku.ca

Department of Economics, Brock University, St. Catharines, Ontario, L2S 3A1, Canada
dimand@brocku.ca

Probabilistically Proving that $\zeta(2) = \pi^2/6$

Luigi Pace

Abstract. We give a short proof of the identity $\zeta(2) = \pi^2/6$ using tools from elementary probability. Related identities, due to Euler, are also briefly discussed.

1. INTRODUCTION.

Let us consider the zeta function

$$\zeta(s) = \sum_{k=1}^{\infty} \frac{1}{k^s}$$

for real $s > 1$. The purpose of this note is to give a short, natural proof of the identity $\zeta(2) = \pi^2/6$ using tools from elementary probability. Inspiration comes from the simple proof given in [**2**]. Probability contributes in two ways. First, it adds motivation to the starting point of calculations, which is viewed as the solution of a general problem. Second, it puts the later steps into a smooth flow, exploiting a simple symmetry.

http://dx.doi.org/10.4169/amer.math.monthly.118.07.641