Reprinted from:

DIFFERENTIAL GAMES AND RELATED TOPICS

SADDLE-POINTS AND CONVEX ANALYSIS *

R. Tyrrell ROCKAFELLAR **

The concept of a saddle-point in a minimax problem was introduced historically as a mathematical expression of an equilibrium between two opposing interests. The opposing interests were those of the players in a two-person game, which could have an economic interpretation. Recently saddle-points have been studied in the case of differential games arising from problems of optimal control. The analysis of such games is far more difficult, in that the correct definition of a strategy or pay-off involves many subtleties.

The abstract theory of saddle-points may be said to begin with the minimax theorem of von Neumann. A number of generalizations of von Neumann's theorem have been proved over the years, and the abstract question of the existence of saddle-points has been fairly well answered. However, research has not only been concerned with existence. Many results have been obtained lately which shed new light on the nature and properties of saddle-points and the possible methods of computing them. The purpose of these expository lectures is to describe some of these results, not as widely known as the standard existence theorems, in the hope that the results themselves, and the notions which they entail, may be useful in further developments.

The first three sections are concerned mainly with the presentation of the dual approach to minimax theory, which was initiated by Moreau [9] and the author [11] and subsequently pursued in [12], [15] and [18]. This is based on Fenchel's conjugacy correspondence for convex functions and its extension to concave-convex functions. Several of the theorems have not previously been stated in a form valid for infinite-dimensional spaces.

The fourth and final section discusses the idea of determining saddle-points by generalized methods of "steepest descent". Results concerning the Arrow-Hurwicz differential equation in nonlinear programming can be greatly broadened by means of the powerful, new theory of nonlinear, monotone operators.

* Supported in part by the Air Force Office of Scientific Research under grant no. AFOSR-1202-67B.

** Department of Mathematics, University of Washington, Seattle.

1. Minimax Problems and Concave-Convex Functions

In a minimax problem, one is given a function K on a product set $C \times D$, the values of K being real numbers or possibly $+\infty$ or $-\infty$, and one considers the quantities

$$(1.1) \qquad \sup_C \inf_D K = \sup_{u \in C} \,[\inf_{v \in D} K(u, v)]\ ,$$

$$\inf_D \sup_C K = \inf_{v \in D} \,[\sup_{u \in C} K(u, v)]\ .$$

In general these two quantities may not be equal, although it is always true, as is easily seen, that

$$(1.2) \qquad \sup_C \inf_D K \leqslant \inf_D \sup_C K\ .$$

If equality does hold, the common value is called the *saddle-value* in the minimax problem. For the saddle-value to exist, it is sufficient (but not necessary) that there exist a *saddle-point* of K with respect to $C \times D$, that is, a pair $(\bar{u}, \bar{v})$ such that

$$(1.3) \qquad \bar{u} \in C, \bar{v} \in D, \max_{u \in C} K(u, \bar{v}) = K(\bar{u}, \bar{v}) = \min_{v \in D} K(\bar{u}, v)\ .$$

Then the saddle-value is $K(\bar{u}, \bar{v})$.

One of the main goals of minimax theory is to establish conditions on K, C and D under which a saddle-point, or at least the saddle-value, exists. Results of this type are termed "minimax theorems". Another important goal is to elucidate the relationship between minimax problems and problems involving only minimization or maximization. Work in this direction is motivated by the Kuhn-Tucker theory of Lagrange multipliers and by duality considerations in mathematical programming. Still another goal is to find ways of computing saddle-points.

Convexity assumptions appear to be an essential ingredient in minimax theory. Most results require that C and D be convex subsets of (real) linear spaces U and V, respectively, and many require also that $K(u, v)$ be concave in $u \in C$ for fixed $v \in D$ and convex in $v \in D$ for fixed $u \in C$. This concave-convex case is the one we deal with here.

General minimax theorems can be proved using slightly weaker assumptions such as quasiconcavity-quasiconvexity in place of the concavity-convexity of K (see Sion [16]), but such results are limited in other respects by compactness assumptions on C or D, and they seem to have relatively little bearing on the second goal mentioned above. A remarkable equivalence is now known between minimax theory in the concave-convex case and the theory of duality in problems of minimizing a convex function over a convex set (see [12, 15]). The equivalence is based on Fenchel's notion of conjugate convex functions. No results of comparable scope have been found for minimax problems where K is only quasiconcave-quasiconvex, say, although there have been some investigations along this line (see Vogel [19]). This is an important reason for singling out the concave-convex case for special treatment.

Another reason is the fact that an arbitrary minimax problem can be "convexified" by various means. This idea is familiar, for example, in the theory of matrix games, where "pure strategies" are generalized to "mixed strategies" so that von Neumann's minimax theorem can be applied. Similarly, when C and D are infinite sets, "mixed strategies" can be defined measure-theoretically as probability distributions on C and D. There are other methods of convexification which take advantage of the linear structure already present when C and D are subsets of linear spaces: convex hull operations are applied to C, D, and in certain ways to K. In all these methods, the objective is to replace the given minimax problem by a concave-convex minimax problem in such a way that the saddle-points in the original problem (if any exist) can be identified with certain saddle-points in the new problem. We shall not discuss this further here, however.

Convexity assumptions are also fundamental for the dual approach to minimax theory which will be explained below.

A very useful technical device in the analysis of minimax problems in which C and D are subsets of linear spaces U and V is to represent the extrema as extrema over all of $U \times V$ by defining K in an appropriate manner outside of $C \times D$. One way to do this is to set

$$(1.4) \qquad K(u, v) = +\infty \text{ if } u \in C, v \notin D,$$

$$= -\infty \text{ if } u \notin C.$$

Equally good is to set

$$(1.5) \qquad K(u, v) = -\infty \text{ if } u \notin C, v \in D,$$

$$= +\infty \text{ if } v \notin D.$$

Observe that in both cases we have

$$(1.6)\qquad \sup_C \inf_D K = \sup_U \inf_V K,$$

$$\inf_D \sup_C K = \inf_V \sup_U K,$$

so that the saddle-value of the extended K with respect to $U \times V$ is the same as the saddle-value with respect to $C \times D$, if these exist. Moreover, $(\bar{u}, \bar{v})$ is a saddle-point with respect to $U \times V$ if and only if $(\bar{u}, \bar{v})$ belongs to $C \times D$ and is a saddle-point with respect to $C \times D$.

Thus in principle we need only consider minimax problems over all of $U \times V$, provided that we admit $+\infty$ and $-\infty$ as possible values of K. The fact that two (or more) definitions of K as in (1.4) and (1.5) can represent essentially the same problem, suggests however that in taking this approach we should regard problems as corresponding really to certain *equivalence classes* of functions, rather than individual functions. An equivalence relation for this purpose will be defined below.

The introduction of functions with possibly infinite values is technically convenient in some aspects of minimax theory, but a much more important and interesting justification will be given in terms of duality. The notion of equivalence of functions likewise has a dual motivation.

By a *concave-convex* function on $U \times V$, we shall always mean an (everywhere-defined) extended-real-valued function K with the property that $K(u, v)$ is concave as a function of u for every v and convex as a function of v for every u. (Here we speak of an extended-real-valued function f on V as *convex* if its epigraph

$$(1.7)\qquad \operatorname{epi} f = \{(v, \alpha) \mid v \in V, \alpha \in R^1\}$$

is convex as a subset of the linear space $V \times R^1$; see [15, §4] for a discussion of this definition and the way in which it generalizes the classical definition for finite functions. A function is concave if its negative is convex.) If K is obtained by (1.4) or (1.5) from a finite concave-convex function on $C \times D$, where C and D are convex sets, then K is concave-convex on $U \times V$, so that the device described above does not destroy concavity-convexity. We therefore limit attention henceforth to the study of saddle-values and saddle-points of concave-convex functions on all of $U \times V$.

The dual approach to minimax theory will now be explained. Let U^* and V^* denote (real) linear spaces paired with U and V, respectively. (Thus a certain bilinear form $\langle u, u^* \rangle$ is given on $U \times U^*$, and another such form is given on $V \times V^*$. For example, one could have $U = U^* = R^m$ and $V = V^* = R^n$ with $\langle \cdot, \cdot \rangle$ as the usual inner product.) Given a function K on $U \times V$, we define the functions $\overline{K}^*$ and $\underline{K}^*$ on $U^* \times V^*$ by

$$\overline{K}^*(u^*, v^*) = \inf_{u \in U} \sup_{v \in V} \{\langle u, u^* \rangle + \langle v, v^* \rangle - K(u, v)\}, \tag{1.8}$$

$$\underline{K}^*(u^*, v^*) = \sup_{v \in V} \inf_{u \in U} \{\langle u, u^* \rangle + \langle v, v^* \rangle - K(u, v)\}. \tag{1.9}$$

(If K is obtained by (1.4) or (1.5) from a function on a set $C \times D$, then the extrema in (1.8) and (1.9) are the same if taken instead over C and D, as already observed.) We then have $\underline{K}^* \leqslant \overline{K}^*$, and

$$\begin{aligned} \overline{K}^*(0, 0) &= -\sup\nolimits_U \inf\nolimits_V K, \\ \underline{K}^*(0, 0) &= -\inf\nolimits_V \sup\nolimits_U K. \end{aligned} \tag{1.10}$$

Thus the saddle-value of K exists if and only if

$$\overline{K}^*(0, 0) = \underline{K}^*(0, 0)\,. \tag{1.11}$$

The dual approach consists of deriving conditions for (1.11) through a general study of the properties of $\overline{K}^*$ and $\underline{K}^*$. Results not only about saddle-values, but also about saddle-points can be obtained in this manner. The success of the dual approach is tied to the following fact.

Theorem 1. *If K is concave-convex on $U \times V$, then $\overline{K}^*$ and $\underline{K}^*$ are concave-convex on $U^* \times V^*$.*

Of course, $\overline{K}^*$ and $\underline{K}^*$ may well have $+\infty$ and $-\infty$ as values. Thus minimax theory naturally leads to the consideration of extended-real-valued concave-convex functions, even if the minimax problems themselves are not expressed in terms of such functions. The functions $\overline{K}^*$ and $\underline{K}^*$ need not agree everywhere, but under mild regularity assumptions on K, as will be discussed below, they are "equivalent" in a sense which implies that they agree on a certain significant subset of $U^* \times V^*$. Minimax theorems correspond to conditions guaranteeing that $(0, 0)$ belong to this subset.

2. Continuity and Conjugate Equivalence Classes

Since minimax theory deals with the extremal values of various functions, it is reasonable to expect that conditions of continuity and compactness must ultimately be involved. Actually, the methods of convex analysis make it possible to go quite far before compactness is needed, and only a rather weak type of continuity is required. Such continuity can be "constructed" using the so-called closure operation for convex functions. This is an obvious virtue, but the consideration of weaker types of continuity than are usually encountered in analysis is not motivated only by "constructiveness" or the desire to obtain broader minimax theorems. It is also important to the dual approach: even if the function K in a given minimax problem has continuity properties which are simple and ordinary, the properties of the functions $\overline{K}^*$ and $\underline{K}^*$ may be more subtle, and we must be prepared to handle them. Although much of the discussion below concerns functions on $U \times V$, it should always be born in mind that one of our chief aims is to apply the results in a dual form to the study of $\overline{K}^*$ and $\underline{K}^*$ on $U^* \times V^*$.

We assume in everything that follows that the spaces U and U^*, and similarly V and V^*, are furnished with locally convex Hausdorff topologies *compatible with the given pairings* [3, Chap. 4]. (Then the continuous linear functionals on U can be identified with the functionals of the form $u \to \langle u, u^* \rangle, u^* \in U^*$, while the continuous linear functionals on U^* can be identified with the functionals $u^* \to \langle u, u^* \rangle, u \in U$.) The reader who is not familiar with the theory of locally convex spaces can assume that $U = U^* = R^m$ and $V = V^* = R^n$ in the customary topologies.

A convex function f is said to be *closed* if f is lower semicontinuous and nowhere has the value $-\infty$, or if f is the constant function "$-\infty$". (Lower semicontinuity means that all level sets of the form $\{v \mid f(v) \leqslant \alpha\}$ are closed, or equivalently that the epigraph of f is a closed set.) If f is an arbitrary convex function on V, f need not be closed, but there exists a greatest closed convex function majorized by f (namely the pointwise supremum of the collection of all closed convex functions h such that $h(v) \leqslant f(v)$ for every v). This function is called the *closure* of f and denoted by cl f. It can be shown that one has

$$(2.1) \qquad (\mathrm{cl}\, f)(v) = \liminf_{w \to v} f(w) \quad \text{for all } v,$$

unless the "lim inf" is $-\infty$ for some v. In the latter event $(\mathrm{cl} f)(v) = -\infty$ for every v, while the "lim inf" is either $-\infty$ or $+\infty$, depending on the particular

v. Note that f and cl$f$ trivially have the same infimum over V, but that cl$f$ is "more likely" to attain this infimum, because it is lower semicontinuous.

For a concave function g, one obtains cl$g$ by closing the convex function $-g$ and multiplying the result by -1. Thus lower semicontinuity is replaced by upper semicontinuity for concave functions, and the special role of $-\infty$ is played by $+\infty$.

Concave-convex functions can be regularized by means of these closure operations. Given such a function K on $U \times V$, we denote by $\mathrm{cl}_1 K$ the function obtained by closing $K(u, v)$ as a concave function of u for each v. Similarly, $\mathrm{cl}_2 K$ is obtained by closing $K(u, v)$ as a convex function of v for each u. Although it is not obvious, it can be shown that $\mathrm{cl}_1 K$ and $\mathrm{cl}_2 K$ are again concave-convex, so that $\mathrm{cl}_2(\mathrm{cl}_1 K)$ and $\mathrm{cl}_1(\mathrm{cl}_2 K)$ are likewise well-defined and concave-convex. However, for interesting and significant reasons to be discussed in a moment, the operations cl_1 and cl_2 do not quite commute, so that in general

$$\mathrm{cl}_1 \mathrm{cl}_2 K \neq \mathrm{cl}_2 \mathrm{cl}_1 K . \tag{2.2}$$

This prevents us from having a single, natural closure operation for concave-convex functions, and we are brought instead to a concept of functions being "closed up to a certain equivalence".

Two concave-convex functions K and K' are defined to be *equivalent* if

$$\mathrm{cl}_1 K = \mathrm{cl}_1 K' \quad \text{and} \quad \mathrm{cl}_2 K = \mathrm{cl}_2 K' . \tag{2.3}$$

If $\mathrm{cl}_1 K$ and $\mathrm{cl}_2 K$ are equivalent to K, then K is said to be *closed*. Closedness is thus a continuity condition intimately related to, but slightly weaker than, the condition that $K(u, v)$ be upper semicontinuous in u and lower semicontinuous in v. It turns out that $\mathrm{cl}_1 \mathrm{cl}_2 K$ and $\mathrm{cl}_2 \mathrm{cl}_1 K$ are always closed and, except in pathological cases, equivalent. Furthermore, any saddle-point of K must also be a saddle-point of $\mathrm{cl}_1 \mathrm{cl}_2 K$ and $\mathrm{cl}_2 \mathrm{cl}_1 K$. In this sense, a minimax problem involving a concave-convex function can always be regularized to a problem in which the function is closed. (The definition of closedness given by Tynianski [18] is more restrictive and does not have this "constructibility" property.)

A highly illuminating example of equivalent, closed, concave-convex functions on $U \times V = R^1 \times R^1$ is obtained from the formula

$$K(u, v) = u^v \quad \text{for} \quad (u, v) \in C \times D = [0, 1] \times [0, 1] . \tag{2.4}$$

This defines a concave-convex function on the square $C \times D$, provided that $K(0, 0)$, which is ambiguous in the formula, is assigned a value in the interval $[0, 1]$. Either (1.4) or (1.5) can be used to extend K from $C \times D$ to a concave-convex function on the whole space. It is then found that

$$(2.5) \qquad \mathrm{cl}_1 \mathrm{cl}_2 K = \mathrm{cl}_1 K = \bar{K},$$

where $\bar{K}$ is the concave-convex function given by (2.4) and (1.5) with $\bar{K}(0, 0) = 1$. On the other hand,

$$(2.6) \qquad \mathrm{cl}_2 \mathrm{cl}_1 K = \mathrm{cl}_2 K = \underline{K},$$

where $\underline{K}$ is the concave-convex function given by (2.4) and (1.4) with $\underline{K}(0, 0) = 0$. The relations (2.5) and (2.6) imply that K, $\bar{K}$ and $\underline{K}$ are closed and equivalent to each other.

Now let K be an arbitrary concave-convex function on $U \times V$. Since the infimum of a convex function is the same as the infimum of its closure, we have

$$(2.7) \qquad \inf_{v \in V} K(u, v) = \inf_{v \in V} \mathrm{cl}_2 K(u, v) \quad \text{for every } u \in U.$$

Similarly

$$(2.8) \qquad \sup_{u \in U} K(u, v) = \sup_{u \in U} \mathrm{cl}_1 K(u, v) \quad \text{for every } v \in V.$$

The expressions on the left in (2.7) and (2.8) suffice to determine the saddle-value and saddle-points of K, and hence these things depend only on $\mathrm{cl}_1 K$ and $\mathrm{cl}_2 K$. The definition of equivalence therefore yields:

Theorem 2. *Equivalent concave-convex functions have the same saddle-value and saddle-points (if any).*

Because of this fact and the regularization results described above, we adopt the view that the fundamental objects of study in (concave-convex) minimax theory are the *equivalence classes* of (extended-real-valued, everywhere-defined) *closed* concave-convex functions. Each such equivalence class corresponds to a single minimax problem.

We have already seen how the minimax problem for a function K on $U \times V$ leads to the study of the functions $\bar{K}^*$ and $\underline{K}^*$ on $U^* \times V^*$ defined by (1.8) and (1.9). In a parallel manner, the minimax problem for a function K^* on $U^* \times V^*$ leads to the study of the functions $\bar{K}$ and $\underline{K}$ on $U \times V$ defined by

$$\bar{K}(u,v) = \inf_{u^* \in U^*} \sup_{v^* \in V^*} \{\langle u, u^*\rangle + \langle v, v^*\rangle - K^*(u^*, v^*)\}, \tag{2.9}$$

$$\underline{K}(u,v) = \sup_{v^* \in V^*} \inf_{u^* \in U^*} \{\langle u, u^*\rangle + \langle v, v^*\rangle - K^*(u^*, v^*)\}. \tag{2.10}$$

The next theorem reveals a surprising connection between these situations.

Theorem 3 (*Duality*). *Let K be an arbitrary, closed, concave-convex function on $U \times V$, and let $\bar{K}^*$ and $\underline{K}^*$ be the concave-convex functions defined by* (1.8) *and* (1.9). *Then $\bar{K}^*$ and $\underline{K}^*$ are closed and equivalent, and they depend only on the equivalence class containing K. Moreover, if K^* is any concave-convex function on $U^* \times V^*$ equivalent to $\bar{K}^*$ and $\underline{K}^*$, then the concave-convex functions $\bar{K}$ and $\underline{K}$ defined by* (2.9) *and* (2.10) *are equivalent to K.*

According to Theorem 3, there is a one-to-one correspondence between equivalence classes of closed concave-convex functions on $U \times V$ and equivalence classes of closed concave-convex functions on $U^* \times V^*$. Corresponding equivalence classes are said to be *conjugate* to each other.

The main consequence of Theorem 3 for minimax theory is that it enables us, in taking the dual approach, to reduce the study of saddle-values to the study of the extent to which two closed, concave-convex functions in the same equivalence class must agree with each other. Results on this question are discussed in the next section. These results can be used to compare $\bar{K}^*$ and $\underline{K}^*$ not only with each other, but also with further elements K^* in the same equivalence class. Such other functions K^* arise, for example, through alternate maximization and minimization in different arguments:

$$K^*(u^*, v^*) = \inf_{u_1 \in U_1} \sup_{v_1 \in V_1} \inf_{u_2 \in U_2} \sup_{v_2 \in V_2} \{\langle u_1, u_1^*\rangle + \langle u_2, u_2^*\rangle + \langle v_1, v_1^*\rangle + \langle v_2, v_2^*\rangle - K(u,v)\} \tag{2.11}$$

where $u = (u_1, u_2) \in U_1 \times U_2 = U$, and so forth.

The proof of Theorem 3 is actually quite elementary – it is given for finite-dimensional spaces in [12, 15], and the extension to locally convex spaces is routine. The proof rests on the properties of Fenchel's conjugacy correspondence for convex functions and the following fact.

Theorem 4. *The formulas*

$$K(u, v) = \sup_{v^* \in V^*} \{\langle v, v^* \rangle - F(u, v^*)\}, \tag{2.12}$$

$$F(u, v^*) = \sup_{v \in V} \{\langle v, v^* \rangle - K(u, v)\}, \tag{2.13}$$

define a one-to-one correspondence between the closed, convex functions F on $U \times V^*$ and the closed concave-convex functions K on $U \times V$ satisfying

$$\mathrm{cl}_2 \mathrm{cl}_1 K = K. \tag{2.14}$$

Each equivalence class of closed, concave-convex functions on $U \times V$ contains exactly one K satisfying (2.14).

This result may be interpreted as saying that closed concave-convex functions are the "partial conjugates" of closed convex functions. If the analogous supremum were also taken over U in (2.12), one would get the convex function G on $U^* \times V$ conjugate to F:

$$G(u^*, v) = \sup_{u \in U} \sup_{v^* \in V^*} \{\langle u, u^* \rangle + \langle v, v^* \rangle - F(u, v^*)\} \tag{2.15}$$

The partial conjugacy relationship between concave-convex functions and convex functions is at the heart of the duality theory for convex programming problems which is expounded in [12] and [15].

3. Minimax Theorems and Subgradients

Theorems about the existence of saddle-values can be deduced in the dual approach from the following structure theorem. Here $\mathrm{rad}\, C$ denotes the set of all $u \in C$ such that every ray emanating from u contains points of C besides u.

Theorem 5. *Let K be a closed, concave-convex function on $U \times V$ which is not identically $+\infty$ or identically $-\infty$, and let*

$$C = \{u \in U \mid \mathrm{cl}_2 K(u, v) > -\infty \quad \text{for all } v \in V\}, \tag{3.1}$$

$$D = \{v \in V \mid \mathrm{cl}_1 K(u, v) < +\infty \quad \text{for all } u \in U\}. \tag{3.2}$$

Then C and D are nonempty, convex sets (not necessarily closed), and they depend only on the equivalence class containing K. Furthermore, K is finite on $C \times D$, $+\infty$ on $C \times [V \setminus \mathrm{cl} D]$ and $-\infty$ on $[U \setminus \mathrm{cl}\, C] \times D$.

If U and V are Banach spaces (the given "compatible" topologies being the norm topologies), then all the concave-convex functions equivalent to K agree on $[\mathrm{rad}\, C] \times V$ and $U \times [\mathrm{rad} D]$, are $+\infty$ on $[\mathrm{rad}\, C] \times [V \setminus D]$, and are $-\infty$ on $[U \setminus C] \times [\mathrm{rad} D]$.

Theorem 5 has previously been stated only for $U \times V = R^m \times R^n$ [11, 15], but essentially the same arguments work in the infinite-dimensional case. One uses the fact that, if f is a lower semicontinuous, convex function on a Banach space, then f is continuous on $\mathrm{rad} S$ whenever f does not take on $+\infty$ on S, and f has a finite lower bound on every bounded set where f does not take on $-\infty$ (see [8] and [13]); analogously for concave functions. (This is needed in proving that the convex functions $(\mathrm{cl}_1 K)(u, \cdot)$ on V are continuous on $\mathrm{rad} D$, so that $\mathrm{cl}_1 K$ and $\mathrm{cl}_2 \mathrm{cl}_1 K$ agree on $U \times [\mathrm{rad} D]$. Of course $\mathrm{cl}_1 K \geqslant \mathrm{cl}_2 K$, while the closedness of K implies that $\mathrm{cl}_2 \mathrm{cl}_1 K = \mathrm{cl}_2 K$ and $\mathrm{cl}_1 \mathrm{cl}_2 K = \mathrm{cl}_1 K$. From the latter relation we have

$$(\mathrm{cl}_1 K)(u, v) = \lim_{\epsilon \downarrow 0} \sup_{\|u' - u\| \leqslant \epsilon} \mathrm{cl}_2 K(u', v), \quad \forall v \in D. \tag{3.3}$$

Since for every $u' \in U$ and $v \in D$ the inequality

$$(\mathrm{cl}_2 K)(u', v) \leqslant (\mathrm{cl}_1 K)(u', v) < +\infty \tag{3.4}$$

holds, and $(\mathrm{cl}_1 K)(\cdot, v)$ is an upper semicontinuous, concave function on U, the supremum in (3.3) must be $< +\infty$ by the fact mentioned above. On the other hand, the supremum is lower semicontinuous and convex as a function of v, because this true of each of the functions $(\mathrm{cl}_2 K)(u', \cdot)$. Therefore the supremum in (3.3) is for any $\epsilon > 0$ a continuous function of $v \in \mathrm{rad}\, D$. This function majorizes the convex function $(\mathrm{cl}_1 K)(u, \cdot)$, and hence we may conclude that the latter is also continuous on $\mathrm{rad} D$.)

The convex set $C \times D$ defined by (3.1) and (3.2) is called the *effective domain* of K. On this set, K is finite. The properties in Theorem 5 imply that the relations (1.6) hold, so that the minimax problem for K on $U \times V$ is equivalent to the minimax problem for the restriction of K to $C \times D$.

To get results about this problem, we apply Theorem 5 to the equivalence class of closed, concave-convex functions which is *conjugate* to the equivalence class containing K. All the functions in this class, including $\overline{K}^*$ and $\underline{K}^*$, have the same effective domain $C^* \times D^*$, and it is not hard to see that this is given in terms of K by

$$(3.5) \qquad C^* = \{u^* \in U^* \mid \exists v \in D, \inf_{u \in C} [\langle u, u^* \rangle - K(u, v)] > -\infty\},$$

$$(3.6) \qquad D^* = \{v^* \in V^* \mid \exists u \in C, \sup_{v \in D} [\langle v, v^* \rangle - K(u, v)] < +\infty\}.$$

From Theorem 5 we have

$$(3.7) \qquad \overline{K}^*(u^*, v^*) = \underline{K}^*(u^*, v^*)$$

if U^* and V^* are Banach spaces (the given topologies being the norm topologies) and either $u^* \in \operatorname{rad} C^*$ or $v^* \in \operatorname{rad} D^*$.

In order to have a theorem asserting the existence of the saddle-value of K, we need only apply this fact in the case of $(u^*, v^*) = (0. 0)$.

Theorem 6. *Let K be a closed, concave-convex function on $U \times V$ which is not identically $+\infty$ or identically $-\infty$, and let C, D, C^* and D^* be the convex sets defined by* (3.1), (3.2), (3.5) *and* (3.6). *Suppose that U^* and V^* are Banach spaces* (*the given "compatible" topologies being the norm topologies*). *If either $0 \in \operatorname{rad} C^*$ or $0 \in \operatorname{rad} D^*$, then the saddle-value of K exists*, i.e. *one has*

$$(3.8) \qquad \sup_U \inf_V K = \inf_V \sup_U K = \sup_C \inf_D K = \inf_D \sup_C K .$$

The condition $0 \in \operatorname{rad} D^*$ in Theorem 6 is satisfied if there exists an element $u \in C$ such that the level sets

$$(3.9) \qquad \{v \in D \mid K(u, v) \leqslant \alpha\}, \quad \alpha \text{ real},$$

are all bounded. This follows from the fact that the level sets of a convex function are bounded if and only if the effective domain of the conjugate function is "radial" at 0. (For results on the duality between boundedness properties of the level sets of a convex function and continuity properties of the conjugate function at 0, see [8], [10], [13] and [2, Theorem 2].) A minimax theorem based on this condition has been proved by Moreau [9], moreover in a form valid for spaces more general than Banach spaces. (Another exposition of Moreau's result has been given by Ioffe and Tikhomirov [5].) Weaker conditions implying in the finite-dimensional case that $0 \in \operatorname{rad} D^*$ may be found in [11].

Similarly, the condition $0 \in \operatorname{rad} C^*$ in Theorem 6 is satisfied if there exists an element $v \in D$ such that the level sets

$$(3.10) \qquad \{u \in C \mid K(u, v) \geqslant \alpha\}, \quad \alpha \text{ real},$$

are all bounded. In particular, therefore, we have:

Corollary. *Let K be a closed concave-convex function on $U \times V$, and let C and D be defined by* (3.1) *and* (3.2). *Suppose that U^* and V^* are Banach spaces (the given "compatible" topologies being the norm topologies). If either C or D is bounded, then* (3.8) *holds.*

The boundedness of C implies actually that $\operatorname{rad} C^*$ is all of U^* in Theorem 6, while the boundedness of D implies that $\operatorname{rad} D^*$ is all of V^*.

The closedness requirement on K is implied by (but generally not equivalent to) the condition that C and D be closed, and that K be upper semicontinuous in u and lower semicontinuous in v on $C \times D$. Thus, for example, the Corollary asserts the existence of the saddle-value of the K defined on $R^1 \times R^1$ by (2.4) and (1.4) with $K(0, 0)$ an arbitrary number in $[0, 1]$, even though this function has a bad discontinuity at $(0, 0)$. Incidentally, if one takes K to be, not the function just described, but a member of the *conjugate* class, then one has by duality an example of concave-convex function K on $R^1 \times R^1$ (actually finite and continuous everywhere) such that $\overline{K}^*(0, 0) = 1$ and $\underline{K}^*(0, 0) = 0$ (with $C^* \times D^* = [0, 1] \times [0, 1]$), and consequently

$$-1 = \sup_u \inf_v K(u, v) < \inf_u \sup_u K(u, v) = 0 .$$

If both of the conditions $0 \in \operatorname{rad} C^*$ and $0 \in \operatorname{rad} D^*$ hold in Theorem 6, then the saddle-value of K is necessarily finite. We shall see below that a much stronger conclusion can be drawn in this case.

Up till now we have concentrated on the existence of saddle-values, but the dual approach also leads easily to results about saddle-points. Just as the study of saddle-values is connected with the analysis of continuity properties of concave-convex functions, the study of saddle-points is connected with differentiability properties.

Let K be a concave-convex function on $U \times V$. An element (u^*, v^*) of $U^* \times V^*$ is called a *subgradient* of K at the point (u, v), if the concave-convex function

$$(3.11) \qquad K - \langle \cdot, u^* \rangle - \langle \cdot, v^* \rangle$$

has a saddle-point on $U \times V$ at (u, v). It can be shown, as an easy extension of similar results for convex functions (see [2] and [15, §35]), that if K happens to be differentiable at (u, v) in the usual sense of Gâteaux (or Fréchet), then K has a unique subgradient at (u, v), namely the usual gradient $\nabla K(u, v) \in U^* \times V^*$. In general, however, K may have no subgradients, at a given point. The set of all subgradients (u^*, v^*) at (u, v) is at all events a closed convex subset of $U^* \times V^*$ which we denote by $\partial K(u, v)$. The multifunction

$$(3.12) \qquad \partial K \colon (u, v) \to \partial K(u, v)$$

from $U \times V$ to $U^* \times V^*$ is called the *subdifferential* of K. The subdifferential of a concave-convex function on $U^* \times V^*$ is defined similarly as a multifunction from $U^* \times V^*$ to $U \times V$.

By definition, K has a saddle-point at $(\bar{u}, \bar{v})$ if and only if

$$(3.13) \qquad (0, 0) \in \partial K(\bar{u}, \bar{v}) .$$

The next theorem allows us to put this condition in a dual form. The first assertion of the theorem is immediate from Theorem 2, and the rest can be deduced from Theorem 4.

Theorem 7. *Let K be a closed, concave-convex function on $U \times V$. The subdifferential ∂K then depends only on the equivalence class containing K. Furthermore, the subdifferential corresponding to the conjugate equivalence class of closed, concave-convex functions is the inverse of ∂K, in the sense that for any member K^* of this class one has*

$$(3.14) \qquad (u, v) \in \partial K^*(u^*, v^*) \Leftrightarrow (u^*, v^*) \in \partial K(u, v) .$$

Combining (3.14) with the saddle-point condition (3.13) we see that, if K is closed, the saddle-points of K are precisely the elements of $\partial K^*(0, 0)$, where K^* is any member of the conjugate equivalence class. *Thus K has a saddle-point if and only if K^* has a subgradient at* $(0, 0)$.

Conditions for the existence of subgradients are fortunately very easy to derive from theorems about supporting hyperplanes to convex sets. Employing the well-known results in the case of convex functions, together with the fact mentioned above that a lower semicontinuous, convex function on a Banach space is continuous on any open set where it does not take on $+\infty$, we get:

Theorem 8. *Let K be a closed, concave-convex function on $U \times V$ which is not identically $+\infty$ or identically $-\infty$, and let C and D be the convex sets defined by* (3.1) *and* (3.2). *If U and V are Banach spaces (the given "compatible" topologies being the norm topologies), then $\partial K(u, v)$ is a nonempty, bounded set for every $(u, v) \in \operatorname{rad}(C \times D)$, whereas $\partial K(u, v)$ is empty for every $(u, v) \notin C \times D$.*

A minimax theorem is now obtained by applying this result to the conjugate class and then invoking the conditions in Theorem 6

Corollary (*Existence of Saddle-points*). *Let K be a closed, concave-convex function on $U \times V$. If U^* and V^* are Banach spaces (the given "compatible" topologies being the norm topologies) and both of the conditions $0 \in \operatorname{rad} C^*$ and $0 \in \operatorname{rad} D^*$ are satisfied (where C^* and D^* are given by* (3.5) *and* (3.6)), *then K has at least one saddle-point, and the set of all such saddle-points is closed, convex and bounded.*

Here, of course, one can substitute for the conditions $0 \in \operatorname{rad} C^*$ and $0 \in \operatorname{rad} D^*$ the stronger conditions discussed following Theorem 6.

4. Saddle-points and "Steepest Descent"

Recent years have seen the rapid development of the theory of a class of nonlinear operators closely related to the subdifferentials of convex and concave-convex functions. An important part of this work consists of generalizations of the method of "steepest descent" in its continuous form. Without going into much detail, we would like to point out the connection between the new results and the theory of saddle-points, particularly the Arrow-Hurwicz differential equation [1, p. 118].

A multifunction T from $U \times V$ to $U^* \times V^*$ is called a *monotone operator* if the inequality

$$(4.1) \qquad \langle u_0 - u_1, u_0^* - u_1^* \rangle + \langle v_0 - v_1, v_0^* - v_1^* \rangle \geqslant 0$$

holds whenever

$$(u_0^*, v_0^*) \in T(u_0, v_0) \quad \text{and} \quad (u_1^*, v_1^*) \in T(u_1, v_1) .$$

It is called a *maximal* monotone operator if it is a monotone operator whose graph

$$(4.2) \qquad G(T) = \{(u, v, u^*, v^*) \mid (u^*, v^*) \in T(u, v)\}$$

is not properly included in the graph $G(T')$ of any other monotone operator T' from $U \times V$ to $U^* \times V^*$. Every monotone operator is embedded in a maximal monotone operator, as can be shown using Zorn's Lemma. (Usually monotonicity is defined for mappings from a space X to a space X^*. We take $X = U \times V$ and $X^* = U^* \times V^*$ here because of our intention of treating saddle-points.)

This concept of monotonicity is decidedly not very intuitive, although a number of heuristic justifications can be given. The main reason for considering it is that it turns out, perhaps rather surprisingly, to be a concept arising in many situations and leading to deep mathematical results. The key to applying these results to minimax theory is the following fact proved in [14].

Theorem 9. *Let K be a closed, concave-convex function on $U \times V$ which is not identically $+\infty$ or identically $-\infty$. Let T be the multifunction from $U \times V$ to $U^* \times V^*$ defined by*

$$(4.3) \qquad (u^*, v^*) \in T(u, v) \Leftrightarrow (-u^*, v^*) \in \partial K(u, v) .$$

Then T is a monotone operator, and if U and V are reflexive Banach spaces T is maximal.

More general criteria for maximality are given in [14], but for present purposes we shall in fact only discuss the simple Euclidean case where $U = U^* = R^m$ and $V = V^* = R^n$.

Let K and T be as in Theorem 9, and let

$$(4.4)\qquad E = \{(u, v) \mid T(u, v) \neq \emptyset\} = \{(u, v) \mid \partial K(u, v) \neq \emptyset\}\,.$$

The set E is not necessarily convex, but according to Theorem 8 we have

$$(4.5)\qquad C \times D \supset E \supset \operatorname{int}(C \times D)\,,$$

where $C \times D$ is convex. For each (u, v), let $-S(u, v)$ denote the unique element of $T(u, v)$ nearest to $(0, 0)$. (Such an element exists, because $T(u, v)$ is a closed, convex set.) We want to consider the differential equation

$$(4.6)\qquad \begin{aligned} &(d/dt)^+(u(t), v(t)) = S(u(t), v(t)), \quad 0 \leqslant t < \infty\,, \\ &(u(0), v(0)) = (a, b) \in E\,, \end{aligned}$$

where $(d/dt)^+$ denotes the right derivative with respect to t. (The solutions to this equation are thus to be trajectories $(u(t), v(t))$ in E which are continuous and right-differentiable in t for $0 \leqslant t < \infty$.)

This is a rather unlikely looking differential equation, but it is a generalization of "steepest descent", and it includes the Arrow-Hurwicz equation as a special case. Suppose, for example, that K is of the form

$$(4.7)\qquad \begin{aligned} K(u, v) &= L(u, v) \quad \text{if } u \in C, v \in D\,, \\ &= +\infty \qquad \text{if } u \in C, v \notin D\,, \\ &= -\infty \qquad \text{if } u \notin C\,, \end{aligned}$$

where C and D are nonempty, closed, convex subsets of R^m and R^n, respectively, and L is a continuously differentiable function on $R^m \times R^n$ which is concave-convex relative to $C \times D$. (Note that K is indeed a closed, concave-convex function on $R^m \times R^n$ whose saddle-points are the saddle-points of L with respect to $C \times D$). It can be verified from the definitions that in this case one has $E = C \times D$, and for each $(u, v) \in C \times D$ the vector $S(u, v)$ is the projection of the vector

$$(4.8)\qquad (\nabla_u L(u, v), -\nabla_v L(u, v))$$

on the closed convex cone generated by the translate $[C \times D] - (u, v)$.

(Thus $S(u, v)$ is the nearest vector to (4.8) which gives a "feasible direction of motion" in $C \times D$ from (u, v). If $(u, v) \in \operatorname{int}(C \times D)$, the vectors $S(u, v)$ and (4.8) coincide.) "Steepest descent" in the classical sense is the case of (4.6) where L is actually independent of u and D is all of R^n. Similarly, "steepest ascent" is the case where L is independent of v and $C = R^m$. In general, (4.6) involves "steepest descent" in v and "steepest ascent" in u simultaneously. The Arrow-Hurwicz equation is the case where L is the Lagrangian function in an ordinary convex programming problem, and C and D are the nonnegative orthants.

Despite the nonstandard character of the differential equation (4.6), a great deal can be said about its solutions. The general results of Kato [6, 7] and Browder [4, cf. Theorem 9.23] for maximal monotone operators can be applied, in view of Theorem 9. In this way one can deduce:

Theorem 10. *Let K and T be as in Theorem 9, where $U = U^* = R^m$ and $V = V^* = R^n$. Then for each $(a, b) \in E$ the differential equation* (4.6) *has a unique solution $(u(t), v(t))$, $0 \leqslant t < +\infty$.*

Furthermore, suppose that K has a saddle-point $(\bar{u}, \bar{v})$ with the property that, if $(u, v) \in E$ is such that the identity

$$(4.9) \qquad K((1-\lambda)\bar{u}+\lambda u, (1-\mu)\bar{v}+\mu v) = (1-\lambda)(1-\mu)K(\bar{u}, \bar{v}) + \lambda(1-\mu)K(u, \bar{v}) + (1-\lambda)\mu K(\bar{u}, v) + \lambda\mu K(u, v)$$

holds for $0 \leqslant \lambda \leqslant 1$ and $0 \leqslant \mu \leqslant 1$, then $u = \bar{u}$ and $v = \bar{v}$. Then $(\bar{u}, \bar{v})$ is the unique saddle-point of K, and for each $(a, b) \in E$ the solution $(u(t), v(t))$ to (4.6) *satisfies*

$$(4.10) \qquad \lim_{t \to +\infty} (u(t), v(t)) = (\bar{u}, \bar{v}) .$$

The condition in the second half of Theorem 10 (which is by no means the most general condition that can be given for convergence to a saddle-point) is satisfied in particular if the saddle-point $(\bar{u}, \bar{v})$ is such that the convex function $K(\bar{u}, \cdot)$ is "strictly convex at $\bar{v}$" (not affine along any line segment including $\bar{v}$), while the concave function $K(\cdot, \bar{v})$ is "strictly concave at $\bar{u}$". An example where the condition does not hold, and (4.10) fails even though the saddle-point $(\bar{u}, \bar{v})$ is unique, is provided by

$$K(u, v) = uv, \quad (u, v) \in R^2, \quad \text{with} \quad (\bar{u}, \bar{v}) = (0, 0) .$$

Results of this type were developed for the Arrwo-Hurwicz equation by Uzawa [1]. For a discussion of related discrete methods of determining saddle-points, see Tremoliers [17].

References

[1] K.J. Arrow, L. Hurwicz and H. Uzawa, Studies in Linear and Nonlinear Programming (Stanford Univ. Press, Palo Alto, Cal., 1958).

[2] E. Asplund and R.T. Rockafellar, Gradients of convex functions, Trans. Am. Math. Soc. 139 (1969) 443–467.

[3] N. Bourbaki, Espaces Vectoriels Topologiques (Hermann, Paris, 1955).

[4] F.E. Browder, Nonlinear operators and nonlinear equations of evolution in Banach spaces, in: Nonlinear Functional Analysis, Part 2, Proceedings of Symposia Pure Math. 18 (Am. Math. Soc., 1970).

[5] A.D. Ioffe and V.M. Tikhomirov, Duality of convex functions and extremum problems, Uspekhi Mat. Nauk 23 (1968) 4–116.

[6] T. Kato, Nonlinear semigroups and evolution equations, J. Math. Soc. Japan 19 (1967) 508–520.

[7] T. Kato, Accretive operators and nonlinear evolution equations in Banach spaces, in: Nonlinear Functional Analysis, Part 1, Proceedings of Symposia in Pure Math. 18 (Am. Math. Soc., 1970), 138–161.

[8] J.J. Moreau, Fonctionelles Convexes, lecture notes, Collège de France, 1967.

[9] J.J. Moreau, Théorèmes 'inf-sup', Compt. Rend. Acad. Sci. Paris 258 (1964) 2720–2722.

[10] J.J. Moreau, Sur la fonction polaire d'une fonction semi-continue supérieurement, Compt. Rend. Acad. Sci. Paris 258 (1964) 1128–1131.

[11] R.T. Rockafellar, Minimax theorems and conjugate saddle-functions, Math. Scand. 14 (1964) 151–173.

[12] R.T. Rockafellar, A general correspondence between dual minimax problems and convex programs, Pacific J. Math. 25 (1968) 597–564.

[13] R.T. Rockafellar, Level sets and continuity of conjugate convex functions, Trans. Am. Math. Soc. 123 (1966) 46–63.

[14] R.T. Rockafellar, Monotone operators associated with saddle-functions and minimax problems, in: Nonlinear Functional Analysis, Part 1, Proceedings of Symposia in Pure Math. 18 (Am. Math. Soc., 1970), 241–250.

[15] R.T. Rockafellar, Convex Analysis (Princeton Univ. Press, Princeton, N.J., 1970).

[16] M. Sion, On general minimax theorems, Pacific J. Math. 8 (1958) 171–176.

[17] R. Tremoliers, Algorithmes pour la resolution de problèmes posés sour la form 'min-max. $f(x, y)$'. Cas non lineaire, report no. Inf6909, Institut de Recherche d'Informatique et d'Automatique, Rocquencourt, May, 1969.

[18] N.T. Tynianski, Conjugate concave-convex functions in linear topological spaces and their saddle-points, Mat. Sbornik 78 (1969) 512–541.

[19] W. Vogel, Duale Optimierungsaufgaben und Sattelpunktsätze, Unternehmensforschung 13 (1969) 1–28.