

AN EXTENSION OF FENCHEL'S DUALITY THEOREM TO SADDLE FUNCTIONS AND DUAL MINIMAX PROBLEMS

L. MCLINDEN

Fenchel's Duality Theorem (or more precisely, Rockafellar's extension of it) is extended here from the context of convex functions and dual convex extremum problems to that of saddle functions and dual minimax problems. The paper is written in the spirit of mathematical programming. Inequalities between optimal values are established, stable optimal solutions are characterized, strong duality theorems proved, and an existence criterion given. An associated Lagrangian saddle point problem is introduced and an extension of the Kuhn-Tucker Theorem derived. The proofs, which are necessarily different from the purely convex case, rely on recently developed pairs of dual operations on saddle functions, as well as on more widely known facts about conjugate saddle functions.

1. Introduction. In 1951, Fenchel [1] proved a fundamental and beautiful duality theorem linking the two convex extremum problems

$$\min_x \{f(x) - g(x)\}$$

and

$$\max_{x^*} \{g^*(x^*) - f^*(x^*)\},$$

where the function f is proper convex on R^n with conjugate f^* , and the function g is proper concave on R^n with conjugate g^* . The framework provided by this pair of problems allows one, by suitable choice of the functions f and g , to deduce duality results for many different constrained convex extremum problems. Since Fenchel's original result, various refinements and extensions of his theorem have been given by a number of authors. Important among these extensions is the one due to Rockafellar [6] (see also [8, §31]). Besides extending the setting from R^n to arbitrary paired locally convex Hausdorff topological vector spaces, it broadens Fenchel's model problems to include a continuous linear transformation A and its adjoint A^* , so that the problems become

$$\min_x \{f(x) - g(Ax)\}$$

and

$$\max_{y^*} \{g^*(y^*) - f^*(A^*y^*)\}.$$

This incorporation of a linear transformation permits much more direct treatment of many types of constraints.

In this paper we broaden Rockafellar's extension of Fenchel's Duality Theorem to the context of saddle point extremum problems and (extended-real-valued) saddle functions on finite-dimensional spaces. The two problems we consider in this regard are

$$(I) \quad \underset{x \in X \ y \in Y}{\text{maximin}} \{K(x, y) - L(A(x, y))\},$$

where

$$X \times Y = \{(x, y) \in \text{dom } K \mid A(x, y) \in \text{dom } L\},$$

and

$$(II) \quad \underset{z \in Z \ w \in W}{\text{minimax}} \{L^*(z, w) - K^*(A^*(z, w))\},$$

where

$$Z \times W = \{(z, w) \in \text{dom } L^* \mid A^*(z, w) \in \text{dom } K^*\}.$$

Here the function K is closed proper concave-convex on $R^m \times R^n$ with conjugate K^* (see [8] for definitions), and the function L is closed proper convex-concave on $R^p \times R^q$ with conjugate L^* . The transformation A has the special form $A = A_1 \times A_2$, where $A_1: R^m \rightarrow R^p$ and $A_2: R^n \rightarrow R^q$ are linear, and A^* is the adjoint of A . Since the saddle functions are allowed to be extended-real-valued, this framework can be made to handle problems which are of the general form of (I) except for the requirement that K and L be defined everywhere. This is done by the device of extending, as necessary, the domains of definition of K and L to the whole space by suitable assignment of the values $+\infty$ and $-\infty$. (See [8, pp. 349 and 398].)

The various results throughout the paper concerning (I) and (II) constitute the extension of Fenchel's Duality Theorem promised by the title. In a broader sense, the results of the paper can also be viewed as extending certain known duality results from the theory of convex programming to the context of "saddle programming". Indeed, the paper is written in this spirit, and accordingly one can think of (I) and (II) as a dual pair of mathematical programming problems, where (I), say, is the "primal" and (II) the "dual" problem.

We relate (I) and (II) also to a third saddle point problem, a "Lagrangian" problem of the form

$$(III) \quad \underset{(x, w)(y, z)}{\text{maximin}} \{K(x, y) + L^*(z, w) - \langle A(x, y), (z, w) \rangle\},$$

where the variables are constrained to satisfy

$$(x, y) \in \text{dom } K, \quad (z, w) \in \text{dom } L^*.$$

It is by studying (III) that we obtain many of our duality results linking (I) and (II). The technique of deriving duality results for pairs of extremum problems of *minimization* or *maximization* type via Lagrangian saddle point problems has been used by several authors. To our knowledge, however, its application to pairs of *saddle point* extremum problems is new.

Our first result states that the lower and upper saddle values in problems (I), (II) and (III) are related as follows:

$$\begin{array}{ccc} & \sup \inf (I) \leq \inf \sup (I) & \\ & \cong & \cong \\ \sup \inf (III) & & \inf \sup (III) . \\ & \cong & \cong \\ & \sup \inf (II) \leq \inf \sup (II) & \end{array}$$

Consequently, *if the saddle value in (III) exists ($\pm \infty$ allowed), then so do the saddle values in (I) and (II), and all three saddle values are equal.* This result can, of course, be combined with any condition which guarantees the existence of the saddle value in (III).

The above inequalities are a direct extension of the familiar and fundamental inequality “ $\inf \geq \sup$ ”, which holds for arbitrary dual pairs of convex programs, as well as for the usual nonlinear programming problem and its dual. This simple inequality has been extremely valuable as a stopping criterion in iterative methods of solution. Yet its extension to saddle point problems does not appear to be of similar utility. There is another relationship among the lower and upper saddle values of (I) and (II), however, which does seem to have potential as a stopping criterion. It is that *whenever (II) is strongly consistent,*

$$\sup \inf (II) \leq \sup \inf (I) \leq \inf \sup (I) \leq \inf \sup (II) ,$$

and dually, whenever (I) is strongly consistent,

$$\sup \inf (I) \leq \sup \inf (II) \leq \inf \sup (II) \leq \inf \sup (I) .$$

Problem (I) is defined to be strongly consistent if and only if there exists a pair $(x, y) \in \text{ri}(\text{dom } K)$ such that $A(x, y) \in \text{ri}(\text{dom } L)$, and similarly (II) is strongly consistent if and only if there exists a pair $(z, w) \in \text{ri}(\text{dom } L^*)$ such that $A^*(z, w) \in \text{ri}(\text{dom } K^*)$. (Here “ri” denotes the relative interior of a set.) Strong consistency will be a particularly useful hypothesis throughout and can be thought of as asserting the existence of a “strongly feasible” solution.

In addition to optimal solutions of (I) and (II), which are just the ordinary saddle points requested by the “maximin” and “minimax” notation, we study those optimal solutions which are “stable” in a

certain natural sense. We show, for instance, that *if (I) is strongly consistent, then every optimal solution of (I) is stable, and dually for (II).*

The celebrated Kuhn-Tucker Theorem [2] has an extension to the present context. Namely, *if (I) is strongly consistent, then for (x, y) to be an optimal solution of (I) it is necessary and sufficient that there exist a pair (z, w) such that (x, w, y, z) is a saddle point of the Lagrangian problem (III).* We actually prove somewhat more: *the hypothesis “(I) is strongly consistent” can be dropped, provided “optimal solution” is replaced by “stable optimal solution”.*

The saddle points of (III) are shown to be exactly those (x, w, y, z) which satisfy the subdifferential relations

$$A^*(z, w) \in \partial K(x, y), \quad A(x, y) \in \partial L^*(z, w).$$

We call this pair of relations the *Kuhn-Tucker conditions* corresponding to (I), (II), and (III). An analysis of the Kuhn-Tucker conditions shows, somewhat surprisingly, that the pairs (x, y) and (z, w) satisfying them are precisely the “stable” optimal solutions of (I) and (II). Our strongest duality result then follows easily: *The following four conditions are equivalent and imply that the optimal values in (I), (II) and (III) are equal:*

- (i) *there exists a stable optimal solution of (I);*
- (ii) *there exists a stable optimal solution of (II);*
- (iii) *there exists an optimal solution of (III);*
- (iv) *the Kuhn-Tucker conditions are satisfiable.*

We provide an existence theorem to go with this theorem. It is that *if (I) is strongly consistent and there exists a pair $(x, y) \in X \times Y$ such that the level sets*

$$\{x' \in X \mid K(x', y) - L(A(x', y)) \geq \alpha\}$$

and

$$\{y' \in Y \mid K(x, y') - L(A(x, y')) \leq \alpha\}$$

are bounded for every $\alpha \in R$, then the Kuhn-Tucker conditions are satisfiable, and moreover, the set of optimal solutions of (I) is bounded and (II) is strongly consistent. Naturally, a dual version can be stated in terms of (II).

Our approach throughout is the one initiated by Rockafellar [5]. That is, we treat the saddle functions involved from the point of view of equivalence classes. This takes into account the phenomenon that, in general, many slightly different saddle functions give rise to the same saddle point problem. In particular, if \tilde{K} (respectively \tilde{L} , \tilde{K}^* and \tilde{L}^*) denotes any saddle function equivalent to K (respectively

L , K^* and L^*) in the sense of Rockafellar, one can form from (I), (II) and (III) a new trio of problems, call them (\tilde{I}) , (\tilde{II}) and (\tilde{III}) , by replacing the original saddle functions with the corresponding equivalent ones. One would then like to know conditions which guarantee that the new problems (\tilde{I}) , (\tilde{II}) and (\tilde{III}) are actually the same as (I), (II) and (III), respectively, in the sense of having the same lower and upper saddle values and saddle points. We show that, *with no qualification whatever (except of course that K and L be closed and proper)*, (\tilde{III}) is the same as (III), while for each of (I) and (II) the corresponding condition of strong consistency suffices.

The results in this paper are all new, inasmuch as they pertain to saddle point problems rather than purely convex problems. The papers of Rockafellar [7] and Tynjanskii [10] can be viewed as pointing toward the present results, though. In particular, [7] contains some results concerning (I) and (II) for the case when A is the identity transformation and L is the "convex-concave indicator function" of the nonnegative orthants. In [10] there are some results essentially about (I) and (II) in the case when A is again the identity but K and L are significantly less general than arbitrary closed proper saddle functions. (The saddle functions treated by Tynjanskii must essentially be the unique elements of their equivalence classes.) Neither [7] nor [10] considers the Lagrangian problem (III). As an application of the present results, we have developed elsewhere a decomposition principle for additively separable minimax problems subject to affine coupling constraints [4].

The results in this paper generalize a body of known duality results nearly intact from the convex function case to the saddle function case. Unfortunately, though, it is seldom true that the proofs also generalize directly. On the contrary, we have found it is usually necessary to devise different and far more complicated arguments. This is due to the essential way in which saddle point extremum problems differ from purely convex extremum problems. One of these differences is that one can typically represent a convex problem by a single "regularized" convex function, whereas (as noted above) a saddle point problem in general permits representation by many different saddle functions, each of which is "regularized". This means that in order to have a fully satisfying theory for saddle point problems, one must deal ultimately with whole equivalence classes of functions. Another difference between the two types of extremum problems is the obvious one, that of one versus two functional arguments. In performing various calculations and manipulations with the functions of one argument, there is usually no ambiguity over whether to take the supremum or the infimum of a particular func-

tion: one simply takes the “inf” or “sup” according as the function is convex or concave. With the two arguments involved in a saddle function, however, one is continually faced with the ambiguity of having to decide between taking the “sup inf” or the “inf sup”. These are two major difficulties not present in the purely convex case and with which we must deal. They will be amply illustrated in the proofs.

For our proofs, we draw heavily on the papers of Rockafellar [5] (actually, the revised presentation found in [8]) and McLinden [3]. In [5] (or [8]) one can find the fundamental facts about equivalence classes, closures, conjugates and subdifferentials of extended-real-valued saddle functions on finite-dimensional spaces. (For the infinite-dimensional case, see [9].) In [3] are located the basic facts about the operations of addition and composition, together with the corresponding dual operations, which are needed both for forming, and for taking the conjugates of, saddle functions of the type $K - LA$.

The remaining sections of the paper are as follows: Preliminaries (§ 2), Invariance Under Equivalence (§ 3), Duality for Optimal Values (§ 4), Optimal Solutions and the Kuhn-Tucker Conditions (§ 5), Stability (§ 6), Duality for Stable Optimal Solutions (§ 7), and An Existence Criterion (§ 8).

2. Preliminaries. In this section we fix the notation of the paper and make some definitions concerning problems (I), (II) and (III). We also introduce for convenience later on three everywhere-defined saddle functions which serve to represent (I), (II) and (III).

Generally, we follow the same procedure as in [3] regarding definitions, conventions and special notations and abbreviations. Thus, terms not defined here are to be understood as in Rockafellar [8], and the conventions regarding arithmetic computations involving $\pm\infty$ and the empty set \emptyset are those set forth in [8, p. 24]. We mention in particular that $\sup \emptyset = -\infty$ and $\inf \emptyset = +\infty$. Results from [8] are cited by a special abbreviation: the number of the result being cited is enclosed in parentheses. For example, Theorem 37.5 of [8] is cited simply as (37.5). The square bracket notation used in [3] to denote equivalence classes of saddle functions is also employed here. Lower and upper bars are used to denote convex and concave closures, respectively, of saddle functions.

Sometimes we need to write the product linear transformation A in terms of its “factors”. For this, we define linear transformations $A_1: R^m \rightarrow R^p$ and $A_2: R^n \rightarrow R^q$ and write $A(x, y) = (A_1x, A_2y)$. Then, of course, $A^*(z, w) = (A_1^*z, A_2^*w)$. Whenever possible, we omit cumbersome parentheses by writing $L(A(x, y)) = LA(x, y)$ and $K^*(A^*(z, w)) = K^*A^*(z, w)$. Also, we write $\text{dom } K = C_K \times D_K$, $\text{dom } K^* = C_{K^*} \times D_{K^*}$,

$\text{dom } L = C_L \times D_L$ and $\text{dom } L^* = C_{L^*} \times D_{L^*}$.

A pair (x, y) is an *optimal solution* of (I) if and only if (x, y) is a saddle point of $K - LA$ with respect to $X \times Y$, i.e., if and only if $(x, y) \in X \times Y$ and

$$K(x', y) - LA(x', y) \leq K(x, y) - LA(x, y) \leq K(x, y') - LA(x, y')$$

for all $(x', y') \in X \times Y$. The quantity $\sup_{x \in X} \inf_{y \in Y} \{K(x, y) - LA(x, y)\}$ will usually be denoted by $\sup_x \inf_y (K - LA)$ or $\sup \inf$ (I), and similarly for the "inf sup" quantity. With these abbreviations, *the optimal value in (I) exists* if and only if

$$\sup \inf (I) = \inf \sup (I) \in R ,$$

in which case this common value is called the *optimal value* in (I). Trivially, if (x, y) is an optimal solution of (I), then the optimal value in (I) exists and equals $K(x, y) - LA(x, y)$. Problem (I) is *consistent* if and only if $X \times Y$ is nonempty. The pairs $(x, y) \in X \times Y$ and the *feasible solutions* of (I). From the conventions $\sup \emptyset = -\infty$ and $\inf \emptyset = +\infty$, it is clear that (I) is consistent whenever the optimal value in (I) exists. Problem (I) is *strongly consistent* if and only if there exists a pair $(x, y) \in \text{ri}(\text{dom } K)$ such that $A(x, y) \in \text{ri}(\text{dom } L)$.

For (II), we make the obvious, analogous definitions and observations. We just let L^* , K^* and A^* play the roles of K , L and A , respectively, and also, interchange the roles of "sup" and "inf" to take account of the fact that (II) is a convex-concave problem, whereas (I) is concave-convex.

Analogous definitions apply also to the optimal solutions and optimal value in (III). That is, the optimal solutions in (III) are the saddle points of the concave-convex function

$$(x, w, y, z) \longrightarrow K(x, y) + L^*(z, w) - \langle A(x, y), (z, w) \rangle ,$$

$$(x, w, y, z) \in C \times D ,$$

where $C = C_K \times D_{L^*}$ and $D = D_K \times C_{L^*}$, and the optimal value in (III) exists if and only if the saddle value of this function exists finitely.

Continuing in the spirit of mathematical programming, we call (I) the *primal problem*, (II) the *dual problem* and, for reasons which will become more apparent later, (III) the *Lagrangian problem*.

It is convenient sometimes to be able to deal with (I), (II) or (III) by means of a single everywhere-defined saddle function. For this purpose we define

$$H(x, y) = \begin{cases} K(x, y) - LA(x, y) & \text{if } x \in X \text{ and } y \in Y \\ +\infty & \text{if } x \in X \text{ and } y \notin Y \\ -\infty & \text{if } x \notin X, \end{cases}$$

$$J(z, w) = \begin{cases} L^*(z, w) - K^*A^*(z, w) & \text{if } z \in Z \text{ and } w \in W \\ +\infty & \text{if } z \notin Z \text{ and } w \in W \\ -\infty & \text{if } w \notin W, \end{cases}$$

and the *Lagrangian* function

$$M(x, w, y, z) = \begin{cases} K(x, y) + L^*(z, w) - \langle A(x, y), (z, w) \rangle & \text{if } (x, w) \in C \text{ and } (y, z) \in D \\ +\infty & \text{if } (x, w) \in C \text{ and } (y, z) \notin D \\ -\infty & \text{if } (x, w) \notin C. \end{cases}$$

It is easily checked, for instance, that H is concave-convex and its lower and upper saddle values, as well as its saddle points, are the same whether taken with respect to $X \times Y$ or to the whole space. Thus, H can be used to represent (I). Similarly, J represents (II) and M represents (III).

3. Invariance under equivalence. In this section we consider the question, "How much (if any) generality is lost in treating the saddle point problems (I), (II) and (III) by focusing our attention on fixed saddle functions K, L, K^* and L^* ?"

Suppose $\tilde{K}, \tilde{L}, \tilde{K}^*$ and \tilde{L}^* are any saddle functions equivalent to K, L, K^* and L^* , respectively. Let $(\tilde{I}), (\tilde{II})$ and (\tilde{III}) denote the trio of saddle point problems defined by replacing the given saddle functions in (I), (II) and (III) with the corresponding equivalent ones. Similarly, let \tilde{H}, \tilde{J} and \tilde{M} denote the associated everywhere-defined saddle functions (cf. § 2).

It is easy to see that the notions of "consistent", "strongly consistent", and "feasible solution" defined in § 2 are invariant under passage from (I) to (\tilde{I}) and from (II) to (\tilde{II}) . This follows immediately from the fact that equivalent saddle functions have the same effective domain.

However, concerning the lower and upper saddle values and the saddle points connected with our trio of problems, the situation is not so clear. The first two lemmas help to remedy this. According to the first, there is "usually" no generality lost in considering (I) rather than (\tilde{I}) , and similarly for (II) and (\tilde{II}) .

LEMMA 1. *If (I) is strongly consistent, then H and \tilde{H} are equivalent closed proper saddle functions, and thus (\tilde{I}) is the same as (I). Dually, if (II) is strongly consistent, then J and \tilde{J} are equivalent,*

closed proper saddle functions, and thus $(\tilde{\text{II}})$ is the same as (II).

Proof. Only the assertion about (I) will be proved, as the other is similar. By the blanket hypotheses, K and L themselves are closed and proper. If (I) is strongly consistent, then

$$A \operatorname{ri}(\operatorname{dom} K) \cap \operatorname{ri}(\operatorname{dom} L) \neq \emptyset .$$

By Theorems 1 and 4 of [3], these facts imply that the equivalence class $[K - LA]$ is well-defined (in the sense of the operations of addition and composition introduced in [3]) and moreover that $[K - LA]$ is closed and proper and contains the functions H and \tilde{H} . Then since H and \tilde{H} are equivalent, closed and proper, it follows from (36.4) and (36.3) that (I) and $(\tilde{\text{I}})$ are the same saddle point problems.

The next lemma shows that with regard to (III) and $(\tilde{\text{III}})$ we need no additional hypotheses whatsoever.

LEMMA 2. *The saddle functions M and \tilde{M} are equivalent, closed and proper, and thus $(\tilde{\text{III}})$ is the same as (III).*

Proof. Since K and L^* are each closed and proper, it follows from Theorem 7(i) of [3] that the function

$$(x, w, y, z) \longrightarrow \begin{cases} K(x, y) + L^*(z, w) & \text{if } (x, w) \in C \text{ and } (y, z) \in D \\ +\infty & \text{if } (x, w) \in C \text{ and } (y, z) \notin D \\ -\infty & \text{if } (x, w) \notin C \end{cases}$$

belongs to a closed proper equivalence class of (separable) saddle functions. Since the saddle function

$$(x, w, y, z) \longrightarrow -\langle A(x, y), (z, w) \rangle$$

is finite everywhere, (34.2.4) implies it is closed and proper (and in fact the unique element of its equivalence class). By Theorem 1 of [3], we can add together the two equivalence classes represented by these two concave-convex functions of (x, w, y, z) , obtaining another well-defined (in the sense of [3]) closed proper equivalence class. It is easy to check that both M and \tilde{M} belong to the resulting class. The last conclusion follows by (36.4) and (36.3), exactly as in Lemma 1.

In view of Lemmas 1 and 2, we can apply to (I), (II) and (III) all of the theory which has been developed for closed proper saddle functions, and in particular all of the various existence criteria and characterization results for saddle values and saddle points. To do this, for (III) we need no extra hypothesis, while for (I) and (II) the corresponding mild hypothesis of strong consistency suffices.

4. **Duality for optimal values.** First we prove that the lower and upper saddle values connected with (III) bracket those connected with (I) and (II).

THEOREM 1.

$$\begin{array}{ccc} & \sup \inf (\text{I}) \leq \inf \sup (\text{I}) & \\ & \cong & \cong \\ \sup \inf (\text{III}) & & \inf \sup (\text{III}) . \\ & \cong & \cong \\ & \sup \inf (\text{II}) \leq \inf \sup (\text{II}) & \end{array}$$

Proof. By (36.1), we have

$$\begin{aligned} \sup \inf (\text{III}) &\leq \sup_{C_K} \inf_{D_K} \sup_{D_{L^*}} \inf_{C_{L^*}} M \\ &= \sup_{C_K} \inf_{D_K} (K - \varphi) , \end{aligned}$$

where

$$\varphi(x, y) = \inf_{D_{L^*}} \sup_{C_{L^*}} (\langle \cdot, A_1 x \rangle + \langle \cdot, A_2 y \rangle - L^*) .$$

But (37.1.1) and (36.3) imply that $\varphi = \bar{L}A$. Hence,

$$\begin{aligned} \sup \inf (\text{III}) &\leq \sup_{C_K} \inf_{D_K} (K - \bar{L}A) \\ &= \sup_x \inf_{D_K} (K - \bar{L}A) \\ &\leq \sup_x \inf_y (K - \bar{L}A) \\ &\leq \sup_x \inf_y (K - LA) \\ &= \sup \inf (\text{I}) . \end{aligned}$$

The first equality here follows from the fact that \bar{L} is convex-closed, and hence $\bar{L}(u, \cdot)$ is constantly $+\infty$ whenever $u \notin \text{dom}_1 L$, while the last inequality follows from the fact that $L \leq \bar{L}$. (See (34.2).) On the other hand, by (36.1) we also have

$$\begin{aligned} \sup \inf (\text{III}) &\leq \sup_{D_{L^*}} \inf_{C_{L^*}} \sup_{C_K} \inf_{D_K} M \\ &= \sup_{D_{L^*}} \inf_{C_{L^*}} (L^* - \psi) , \end{aligned}$$

where

$$\psi(z, w) = \inf_{C_K} \sup_{D_K} (\langle \cdot, A_1^* z \rangle + \langle \cdot, A_2^* w \rangle - K) .$$

Observing that $\psi = \bar{K}^* A^*$ and continuing similarly to the above calculation, we obtain

$$\sup \inf \text{ (III)} \leq \sup \inf \text{ (II)} .$$

The proof of the remaining two nontrivial inequalities is entirely analogous.

COROLLARY. *If the optimal value in (III) exists, then so do the optimal values in (I) and (II), and all three values are equal.*

According to this corollary, any condition which implies that M has a finite saddle value also yields a duality theorem linking the optimal values in (I) and (II). Since M is closed and proper (Lemma 2), such conditions are well-known.

The next theorem gives rather surprising "boxing in" inequalities between the lower and upper saddle values connected with (I) and with (II).

THEOREM 2. *If (I) is strongly consistent, then*

$$\sup \inf \text{ (I)} \leq \sup \inf \text{ (II)} \leq \inf \sup \text{ (II)} \leq \inf \sup \text{ (I)} .$$

Dually, if (II) is strongly consistent, then

$$\sup \inf \text{ (II)} \leq \sup \inf \text{ (I)} \leq \inf \sup \text{ (I)} \leq \inf \sup \text{ (II)} .$$

Proof. Define two auxiliary functions as follows:

$$P(x, u, y, v) = \begin{cases} K(x, y) - L(u, v) & \text{if } (x, u) \in C_P \text{ and } (y, v) \in D_P \\ +\infty & \text{if } (x, u) \in C_P \text{ and } (y, v) \notin D_P \\ -\infty & \text{if } (x, u) \notin C_P , \end{cases}$$

where $C_P = C_K \times C_L$ and $D_P = D_K \times D_L$, and

$$Q(x, u, y, v) = \begin{cases} 0 & \text{if } (x, u) \in C_Q \text{ and } (y, v) \in D_Q \\ +\infty & \text{if } (x, u) \in C_Q \text{ and } (y, v) \notin D_Q \\ -\infty & \text{if } (x, u) \notin C_Q , \end{cases}$$

where $C_Q = \{(x, u) \mid u = A_1x\}$ and $D_Q = \{(y, v) \mid v = A_2y\}$. By Theorem 7(i) of [3], P is a (separable) closed proper concave-convex function with $\text{dom } P = C_P \times D_P$, and by (34.2.4) Q is a closed proper concave-convex function with $\text{dom } Q = C_Q \times D_Q$. Assume (I) is strongly consistent, i.e., $\text{Ari}(\text{dom } K) \cap \text{ri}(\text{dom } L) \neq \emptyset$. Since $\text{dom } Q$ is an affine set and hence relatively open, the strong consistency hypothesis is equivalent to the condition $\text{ri}(\text{dom } P) \cap \text{ri}(\text{dom } Q) \neq \emptyset$. Hence Theorem 1 of [3] implies that the equivalence class $[P + Q]$ is well-defined, closed and proper with $\text{dom } (P + Q) = \text{dom } P \cap \text{dom } Q = C \times D$. Therefore, (36.3) implies that

$$\begin{aligned} \sup \inf (P + Q) &= \sup_{(x,u) \in C} \inf_{(y,v) \in D} \{K(x, y) - L(u, v)\} \\ &= \sup_{(x,u) \in C} \inf_{y \in Y} \{K(x, y) - L(u, A_2 y)\} \\ &= \sup_{x \in X} \inf_{y \in Y} \{K(x, y) - L(A_1 x, A_2 y)\} \\ &= \sup \inf (I). \end{aligned}$$

On the other hand, Theorem 2 of [3] implies that

$$\begin{aligned} \sup \inf (P + Q) &= -(\overline{P + Q})^*(0, 0) \\ &= -(\overline{P^* \square Q^*})(0, 0), \end{aligned}$$

where \square denotes the operation of extremal convolution developed in [3] (and, as usual, an upper bar signifies the greatest element of an equivalence class). Now it follows from Theorem 7(iv) of [3] that one conjugate of P is the function

$$(s, z, t, w) \longrightarrow \begin{cases} K^*(s, t) - L^*(-z, -w) & \text{if } (s, z) \in C_{P^*} \text{ and } (t, w) \in D_{P^*} \\ +\infty & \text{if } (s, z) \in C_{P^*} \text{ and } (t, w) \notin D_{P^*} \\ -\infty & \text{if } (s, z) \notin C_{P^*}, \end{cases}$$

where $C_{P^*} = C_{K^*} \times (-C_{L^*})$ and $D_{P^*} = D_{K^*} \times (-D_{L^*})$, and it is an easy exercise to compute that one conjugate of Q is the function

$$(s, z, t, w) \longrightarrow \begin{cases} 0 & \text{if } (s, z) \in C_{Q^*} \text{ and } (t, w) \in D_{Q^*} \\ +\infty & \text{if } (s, z) \in C_{Q^*} \text{ and } (t, w) \notin D_{Q^*} \\ -\infty & \text{if } (s, z) \notin C_{Q^*}, \end{cases}$$

where $C_{Q^*} = \{(s, z) \mid s = -A_1^* z\}$ and $D_{Q^*} = \{(t, w) \mid t = -A_2^* w\}$. Using these two functions together with the fact that $[P^* \square Q^*]$ is well-defined, we obtain

$$\begin{aligned} -(\overline{P^* \square Q^*})(0, 0) &\leq -\inf_{(t,w) \in D_{Q^*}} \sup_{(s,z) \in C_{Q^*}} \{K^*(s, t) - L^*(-z, -w)\} \\ &= \sup_{(t,w) \in D_{Q^*}} \inf_{z \in Z} \{L^*(z, -w) - K^*(A_1^* z, t)\} \\ &= \sup_{w \in W} \inf_{z \in Z} \{L^*(z, w) - K^*(A_1^* z, A_2^* w)\} \\ &= \sup \inf (II). \end{aligned}$$

Combining facts, we obtain

$$\sup \inf (I) = -(\overline{P^* \square Q^*})(0, 0) \leq \sup \inf (II).$$

Similarly,

$$\inf \sup (I) = -(\underline{P^* \square Q^*})(0, 0) \geq \inf \sup (II).$$

COROLLARY. *If (I) is strongly consistent and the optimal value in (I) exists, then (II) is consistent and the optimal value in (II)*

exists and equals the optimal value in (I). Dually, if (II) is strongly consistent and the optimal value in (II) exists, then (I) is consistent and the optimal value in (I) exists and equals the optimal value in (II).

5. **Optimal solutions and the Kuhn-Tucker conditions.** In this section we show there is a close connection between the optimal solutions of (I) and (II) and the pairs (x, y) and (z, w) which satisfy the subdifferential relations

$$A^*(z, w) \in \partial K(x, y), \quad A(x, y) \in \partial L^*(z, w).$$

This pair of relations will be called the *Kuhn-Tucker conditions*. The key to their connection with the optimal solutions of (I) and (II) lies in the following result.

THEOREM 3. *In order that $(x, y) \in R^m \times R^n$ be an optimal solution of (I), it is sufficient that*

$$(0, 0) \in \partial K(x, y) - A^* \partial L(A(x, y));$$

this condition is also necessary when (I) is strongly consistent. Dually, in order that $(z, w) \in R^p \times R^q$ be an optimal solution of (II), it is sufficient that

$$(0, 0) \in \partial L^*(z, w) - A \partial K^*(A^*(z, w));$$

this condition is also necessary when (II) is strongly consistent.

Proof. Assume $(0, 0) \in \partial K(x, y) - A^* \partial L(A(x, y))$. Thus, there exists a pair (z, w) such that

$$A^*(z, w) \in \partial K(x, y), \quad (z, w) \in \partial L(A(x, y)).$$

Then (37.4) implies

$$(x, y) \in \text{dom } K, \quad A(x, y) \in \text{dom } L,$$

and that

$$\begin{aligned} K(x', y) - \langle x', A_1^* z \rangle - \langle y, A_2^* w \rangle \\ \leq K(x, y) - \langle x, A_1^* z \rangle - \langle y, A_2^* w \rangle \\ \leq K(x, y') - \langle x, A_1^* z \rangle - \langle y', A_2^* w \rangle, \quad \forall (x', y'), \end{aligned}$$

and

$$\begin{aligned} \langle u, z \rangle + \langle A_2 y, w \rangle - L(u, A_2 y) \\ \leq \langle A_1 x, z \rangle + \langle A_2 y, w \rangle - L(A_1 x, A_2 y) \\ \leq \langle A_1 x, z \rangle + \langle v, w \rangle - L(A_1 x, v), \quad \forall (u, v). \end{aligned}$$

Since $K(x, \cdot)$ and $-L(A_1x, \cdot)$ are never $-\infty$ and $K(\cdot, y)$ and $-L(\cdot, A_2y)$ are never $+\infty$, we can “add” the two systems of inequalities to obtain

$$\begin{aligned} K(x', y) - L(u, A_2y) + \langle u - A_1x', z \rangle \\ \leq K(x, y) - L(A_1x, A_2y) \\ \leq K(x, y') - L(A_1x, v) + \langle v - A_2y', w \rangle \end{aligned}$$

for all (x', y') and (u, v) . Taking $(u, v) = A(x', y')$ in this, together with the fact that $(x, y) \in X \times Y$, we see that (x, y) is an optimal solution of (I). This establishes the sufficiency. On the other hand, let (x, y) be any optimal solution of (I) and assume (I) is strongly consistent. Then $(0, 0) \in \partial H(x, y)$. But by strong consistency we have

$$\begin{aligned} \partial H(x, y) &= \partial(K - LA)(x, y) \\ &= \partial K(x, y) - \partial(LA)(x, y) \\ &= \partial K(x, y) - A^* \partial L(A(x, y)) \end{aligned}$$

by Lemma 1 and the subdifferential formulas in Theorems 1 and 4 of [3]. The proof of the dual assertion is analogous.

Of course, when K and L are actually differentiable, the subdifferentials in the above “sufficient and usually necessary” conditions (cf. [8, p. 333]) reduce to the ordinary gradients.

We remark that, as far as “sufficiency” is concerned, the conditions in the theorem can be improved to handle the case when the saddle functions are not closed. Indeed, suppose K and L are merely proper, that $\tilde{K} \in [K]$ and $\tilde{L} \in [L]$, that

$$(0, 0) \in \partial \tilde{K}(x, y) - A^* \partial \tilde{L}(A(x, y)),$$

and that $\tilde{K}(x, y)$ and $\tilde{L}A(x, y)$ are finite. Then (x, y) is an optimal solution of (I). This can be shown by the reader as an exercise, using (37.4.1), (37.4) and (23.3).

We can draw several corollaries from Theorem 3.

COROLLARY 1. *Assume $(x, y) \in R^m \times R^n$ and $(z, w) \in R^p \times R^q$ satisfy the Kuhn-Tucker conditions, i.e.,*

$$A^*(z, w) \in \partial K(x, y) \quad \text{and} \quad A(x, y) \in \partial L^*(z, w).$$

Then (x, y) is an optimal solution of (I) and (z, w) is an optimal solution of (II).

Proof. By the theorem and (37.5).

We omit the dual version of the next corollary.

COROLLARY 2. *If (II) is strongly consistent and has an optimal solution, then (I) has an optimal solution and the optimal values in (I) and (II) are equal.*

Proof. By the theorem and the corollary to Theorem 2.

Stronger duality results are possible. For these, it will be helpful to introduce a notion of stability for the optimal solutions of (I) and (II).

6. **Stability.** A typical application of the present results concerning problems (I), (II) and (III) entails formulating some given constrained saddle point problem as (I), say, in such a way that the term $-LA$ represents (possibly some subset of) the explicit constraints of the problem. It is natural to ask about the sensitivity of the optimal solutions with respect to various perturbations of the explicit constraints. In this section we formulate a notion of stability to deal with this for the case in which the perturbations correspond to translations of L . As will be seen in the next section, such stability is intimately related to the strongest possible duality between optimal solutions of (I) and (II).

For each $(x, y) \in X \times Y$, define auxiliary functions g_y on R^p and f_x on R^q by

$$g_y(u) = \sup_{x_u} \{K(\cdot, y) - LA(\cdot, y)\} \quad \text{and} \quad f_x(v) = \inf_{y_v} \{K(x, \cdot) - LA(x, \cdot)\},$$

where

$$X_u = \{x' \in C_K \mid A_1 x' \in C_L - u\} \quad \text{and} \quad Y_v = \{y' \in D_K \mid A_2 y' \in D_L - v\}.$$

We shall say that an optimal solution (x, y) of (I) is *stable* if and only if $\lim_{\lambda \downarrow 0} 1/\lambda [g_y(\lambda u) - g_y(0)] < +\infty, \forall u$, and $\lim_{\lambda \downarrow 0} 1/\lambda [f_x(\lambda v) - f_x(0)] > -\infty, \forall v$. It follows from Lemma 3 (below) and (23.1) that the limits just written do in fact exist, $+\infty$ and $-\infty$ being allowed as limits. Stability of optimal solutions of (II) is defined analogously, using the functions h_w on R^m and k_z on R^n defined for each $(z, w) \in Z \times W$ by

$$h_w(s) = \inf_{z_s} \{L^*(\cdot, w) - K^*A^*(\cdot, w)\}$$

and

$$k_z(t) = \sup_{w_t} \{L^*(z, \cdot) - K^*A^*(z, \cdot)\},$$

where

$$Z_s = \{z' \in C_{L^*} \mid A_1^* z' \in C_{K^*} - s\} \quad \text{and} \quad W_t = \{w' \in D_{L^*} \mid A_2^* w' \in D_{K^*} - t\}.$$

It will be shown in the next section that this notion of stability is "invariant under equivalence" in the sense used in §3.

Stability can be described heuristically using the terminology of game theory. Let (I) represent a two-person zero-sum game on $X \times Y$ with payoff function $K - LA$, and suppose Player 1's objective is to pick a strategy $x \in X$ achieving $\max_x \inf_y \{K - LA\}$ while Player 2's objective is to pick a strategy $y \in Y$ achieving $\min_y \sup_x \{K - LA\}$. Then for this game, an optimal strategy pair (x, y) is stable if and only if (i) no slight perturbation of Player 1's strategy space from X to X_u yields an infinite marginal *increase* in the amount Player 1 receives from Player 2 and (ii) no slight perturbation of Player 2's strategy space from Y and Y_v yields an infinite marginal *decrease* in the amount Player 2 pays Player 1.

We collect some useful facts concerning f_x, g_y and (I) in the following lemma. The analogous assertions concerning h_w, k_z and (II) are omitted.

LEMMA 3. *The function f_x is convex for each $x \in X$, and the function g_y is concave for each $y \in Y$. A pair (x, y) is an optimal solution of (I) if and only if $(x, y) \in X \times Y$ and $f_x(0) = g_y(0) \in R$, and in this event (x, y) is stable if and only if f_x and g_y are subdifferentiable at the origin.*

Proof. Let $x \in X$, and consider the function

$$\varphi(v, y) = K(x, y) - L(A_1x, v + A_2y) + \delta(y | Y_v),$$

where $\delta(\cdot | S)$ denotes the indicator function of a set S (i.e., $\delta(s | S)$ equals 0 when $s \in S$ and $+\infty$ otherwise). It is easy to see that φ is convex in (v, y) jointly, and hence the convexity of $f_x(v) = \inf_y \varphi(v, y)$ is immediate from (5.7). The concavity of g_y is proved similarly. The characterization of optimal solutions is immediate from the definitions and the fact that $K - LA$ is necessarily finite on $X \times Y = X_0 \times Y_0$. The characterization of stability is immediate from the first two assertions of the lemma together with (23.2) and (23.3).

For any $x \in X$, it can also be shown that $v \in \text{ri}(\text{dom } f_x)$ if and only if $(A_2 \text{ri } D_x) \cap (\text{ri } D_L - v) \neq \emptyset$, in which case $\partial f_x(v) \neq \emptyset$. Similarly for g_y . From these facts and Lemma 3 it follows that, if (I) is strongly consistent, *every* optimal solution of (I) is stable. This result appears below as Theorem 5, where a different proof is given.

7. Duality for stable optimal solutions. We are now ready to derive the following strong, "three-way" duality result.

THEOREM 4. *For $(x, y) \in R^m \times R^n$ and $(z, w) \in R^p \times R^q$, the following four conditions are equivalent:*

- (i) $(x, y) \in X \times Y, f_x(0) = g_y(0) \in R$ and $-(z, w) \in \partial g_y(0) \times \partial f_x(0)$;
- (ii) $(z, w) \in Z \times W, h_w(0) = k_z(0) \in R$ and $-(x, y) \in \partial h_w(0) \times \partial k_z(0)$;
- (iii) (x, w, y, z) is a saddle point of M ;
- (iv) $A^*(z, w) \in \partial K(x, y)$ and $A(x, y) \in \partial L^*(z, w)$.

Moreover, any of these four conditions implies that

$$\begin{aligned} K(x, y) - LA(x, y) &= L^*(z, w) - K^*A^*(z, w) \\ &= K(x, y) + L^*(z, w) - \langle A(x, y), (z, w) \rangle . \end{aligned}$$

Proof. We first prove that (i) is equivalent to (iv). Observe that in the presence of $f_x(0) = g_y(0) = \alpha$,

$$-(z, w) \in \partial g_y(0) \times \partial f_x(0)$$

is equivalent to

$$g_y(u) + \langle z, u \rangle \leq \alpha \leq f_x(v) + \langle w, v \rangle$$

for each (u, v) , which is equivalent to

$$\begin{aligned} K(x', y) - L(u + A_1x', A_2y) + \langle z, u \rangle \\ \leq \alpha \leq K(x, y') - L(A_1x, v + A_2y') + \langle w, v \rangle \end{aligned}$$

for each (u, v) and each $(x', y') \in X_u \times Y_v$, which in turn is equivalent to

$$\begin{aligned} K(x', y) - L(u', A_2y) + \langle z, u' - A_1x' \rangle \\ \leq \alpha \leq K(x, y') - L(A_1x, v') + \langle w, v' - A_2y' \rangle \end{aligned}$$

for each $(u', v') \in \text{dom } L$ and each $(x', y') \in \text{dom } K$. Since $(x, y) \in X \times Y$ and $f_x(0) = g_y(0) = \alpha$ implies $\alpha = K(x, y) - LA(x, y)$, it follows that (i) is equivalent to $(x, y) \in X \times Y$ and

$$\begin{aligned} K(x', y) - L(u', A_2y) + \langle z, u' - A_1x' \rangle \\ \leq K(x, y) - L(A_1x, A_2y) \\ \leq K(x, y') - L(A_1x, v') + \langle w, v' - A_2y' \rangle \end{aligned}$$

for each $(x', y') \in \text{dom } K$ and each $(u', v') \in \text{dom } L$. By (36.3) this is equivalent to the two conditions

$$(x, y) \text{ solves } \underset{C_K \ D_K}{\text{maximin}} \{K - \langle \cdot, A_1^*z \rangle - \langle \cdot, A_2^*w \rangle\}$$

and

$$A(x, y) \text{ solves } \underset{C_L \ D_L}{\text{minimax}} \{L - \langle \cdot, z \rangle - \langle \cdot, w \rangle\} .$$

By (36.3) again, together with (37.4), these conditions are equivalent to

$$A^*(z, w) \in \partial K(x, y) \quad \text{and} \quad (z, w) \in \partial L(A(x, y)) .$$

But by (37.5) these last conditions are the same as the Kuhn-Tucker conditions. The proof that (ii) is equivalent to (iv) is analogous. To show (iii) equivalent to (iv), note first that by Lemma 2, M is closed and proper with effective domain $C \times D$. Hence (iii) is equivalent to $(x, w, y, z) \in C \times D$ and

$$\begin{aligned} & K(x', y) + L^*(z, w') - \langle A_1 x', z \rangle - \langle A_2 y, w' \rangle \\ & \leq K(x, y) + L^*(z, w) - \langle A_1 x, z \rangle - \langle A_2 y, w \rangle \\ & \leq K(x, y') + L^*(z', w) - \langle A_1 x, z' \rangle - \langle A_2 y', w \rangle \end{aligned}$$

for each $(x', w') \in C$ and each $(y', z') \in D$. Using the facts that $K(x, y)$ and $L^*(z, w)$ are finite, $K(x, \cdot)$ and $L^*(\cdot, w)$ are never $-\infty$, and $K(\cdot, y)$ and $L^*(z, \cdot)$ are never $+\infty$, one can easily show that this is equivalent to the conditions

$$(x, y) \text{ solves } \underset{C_K \times D_K}{\text{maximin}} \{K - \langle \cdot, A_1^* z \rangle - \langle \cdot, A_2^* w \rangle\}$$

and

$$(z, w) \text{ solves } \underset{C_{L^*} \times D_{L^*}}{\text{minimax}} \{L^* - \langle \cdot, A_1 x \rangle - \langle \cdot, A_2 y \rangle\},$$

which by (36.3) and (37.4) are equivalent to (iv). The last assertion can be deduced from the Corollary to Theorem 1. However, we prefer to give another, more direct proof. Observe first that by (37.5) and (37.4), $A^*(z, w) \in \partial K(x, y)$ if and only if $A^*(z, w)$ is a saddle point of $\langle \cdot, x \rangle + \langle \cdot, y \rangle - K^*$. On the other hand, the definition of the conjugacy correspondence for saddle functions implies that, when the saddle value of $\langle \cdot, x \rangle + \langle \cdot, y \rangle - K^*$ exists, its value is $K^{**}(x, y)$, which equals $K(x, y)$ when K is closed. Therefore, the condition $A^*(z, w) \in \partial K(x, y)$ implies $(K(x, y)$ is finite and)

$$\langle A_1^* z, x \rangle + \langle A_2^* w, y \rangle - K^* A^*(z, w) = K(x, y).$$

Similarly, the condition $A(x, y) \in \partial L^*(z, w)$ implies $(L^*(z, w)$ is finite and)

$$\langle z, A_1 x \rangle + \langle w, A_2 y \rangle - LA(x, y) = L^*(z, w).$$

The last assertion of the theorem follows from these two equations.

Our first corollary to this theorem follows by Lemma 3, the definitions, and the general fact that a saddle function evaluated at a saddle point gives the saddle value.

COROLLARY 1. *The following four conditions are equivalent and imply that the optimal values in (I), (II) and (III) are equal:*

- (i) *there exists a stable optimal solution of (I);*
- (ii) *there exists a stable optimal solution of (II);*

- (iii) *there exists an optimal solution of (III);*
- (iv) *the Kuhn-Tucker conditions are satisfiable.*

By Lemma 3 and its "dual" version, we obtain another corollary to Theorem 4.

COROLLARY 2. *A pair (x, y) is a stable optimal solution of (I) if and only if there exists a pair (z, w) which together with (x, y) satisfies the Kuhn-Tucker conditions. Dually, a pair (z, w) is a stable optimal solution of (II) if and only if there exists a pair (x, y) which together with (z, w) satisfies the Kuhn-Tucker conditions.*

According to this corollary, stable optimal solutions of (I) and (II) are expressible solely in terms of the Kuhn-Tucker conditions. Since by (37.4.1) and (37.1.1) the Kuhn-Tucker conditions are "invariant under equivalence" in the sense used in § 3, it follows that our notion of stability is also invariant under equivalence.

The preceding corollary and the next theorem taken together give a sharp extension of the Kuhn-Tucker Theorem to our problems (I), (II) and (III).

THEOREM 5. *If (I) is strongly consistent, then every optimal solution of (I) is stable. Dually, if (II) is strongly consistent, then every optimal solution of (II) is stable.*

Proof. Assume (I) is strongly consistent. Then by Theorem 3, (x, y) is an optimal solution of (I) if and only if

$$(0, 0) \in \partial K(x, y) - A^* \partial L(A(x, y)).$$

By (37.5) it follows that (x, y) is an optimal solution of (I) if and only if there exists a pair (z, w) which together with (x, y) satisfies the Kuhn-Tucker conditions. But by conditions (iv) and (i) of Theorem 4, together with Lemma 3, this implies (x, y) is stable. The proof of the dual assertion is analogous.

8. An existence criterion. Our object in this section is to establish a criterion for the existence of (stable) optimal solutions. This is done in the theorem below. Although considerably more general existence criteria can be formulated, the one given here hopefully strikes a reasonable balance between generality and ease of applicability. Also, the criterion here yields the useful fact of dual strong consistency as a byproduct. We state the result in terms of (I); naturally, a dual version can be phrased in terms of (II).

THEOREM 6. Assume (I) is strongly consistent and that there exists a pair $(x, y) \in X \times Y$ such that the sets

$$\{x' \in X \mid K(x', y) - LA(x', y) \geq \alpha\}$$

and

$$\{y' \in Y \mid K(x, y') - LA(x, y') \leq \alpha\}$$

are bounded for every $\alpha \in R$. Then the Kuhn-Tucker conditions are satisfiable. Moreover, the set of optimal solutions of (I) is bounded and (II) is strongly consistent.

For the proof of this theorem we shall use several lemmas. These involve recession functions, which will be indicated by the notation "rec". That is,

$$(\text{rec } f)(y) = \sup \{f(y' + y) - f(y') \mid y' \in \text{dom } f\}$$

when f is proper convex, and

$$(\text{rec } g)(x) = \inf \{g(x' + x) - g(x') \mid x' \in \text{dom } g\}$$

when g is proper concave.

The first lemma is a slightly different form of (37.2). Rockafellar chose to formulate this result for the particular choice $S = \text{ri}(\text{dom}_1 K)$ and $T = \text{ri}(\text{dom}_2 K)$, in which case (34.2) allows the functions \underline{K} and \bar{K} appearing in the formulas to be replaced simply by K .

LEMMA 4. Let K be any closed proper concave-convex function on $R^m \times R^n$. Then

$$\sup \{\langle y, y^* \rangle \mid y^* \in \text{dom}_2 K^*\} = \sup \{\text{rec } \underline{K}(x, \cdot)(y) \mid x \in S\}$$

for any set S (not necessarily convex) satisfying $\text{ri}(\text{dom}_1 K) \subset S \subset \text{dom}_1 K$. Similarly,

$$\inf \{\langle x, x^* \rangle \mid x^* \in \text{dom}_1 K^*\} = \inf \{\text{rec } \bar{K}(\cdot, y)(x) \mid y \in T\}$$

for any set T (not necessarily convex) satisfying $\text{ri}(\text{dom}_2 K) \subset T \subset \text{dom}_2 K$.

Proof. We prove only the first assertion. Write $\text{dom}_2 K^* = D^*$. From the formulas given for D^* and $\text{ri } D^*$ in the proof of (37.2), it is clear that

$$\text{ri } D^* \subset \bigcup_{x \in S} \text{dom } f(x, \cdot) \subset D^*,$$

where f is the convex function related to K by

$$f(x, y^*) = \sup_y \{ \langle y, y^* \rangle - K(x, y) \} .$$

Hence, much as in the proof of (37.2), we can write

$$\begin{aligned} \sup_{y^* \in D^*} \{ \langle y, y^* \rangle \} &= \sup_{x \in S} \{ \sup \{ \langle y, y^* \rangle \mid y^* \in \text{dom } f(x, \cdot) \} \} \\ &= \sup_{x \in S} \{ \text{rec } f(x, \cdot)^*(y) \} . \end{aligned}$$

Since $f(x, \cdot)^* = \text{cl } K(x, \cdot) = \underline{K}(x, \cdot)$, the proof is complete.

The next lemma gives a “primal-space” characterization of the important hypothesis of dual strong consistency, and the subsequent lemma gives a simpler sufficiency condition. Besides being useful for the proof of Theorem 6, these two lemmas should prove useful in applying the earlier results which involve the hypothesis of dual strong consistency.

LEMMA 5. *Problem (II) is strongly consistent if and only if (i) for every y ,*

$$\sup_{x \in C_K} \{ \text{rec } \underline{K}(x, \cdot)(y) \} \leq \inf_{u \in C_L} \{ \text{rec } \bar{L}(u, \cdot)(A_2 y) \}$$

implies

$$\sup_{x \in C_K} \{ \text{rec } \underline{K}(x, \cdot)(-y) \} \leq \inf_{u \in C_L} \{ \text{rec } \bar{L}(u, \cdot)(-A_2 y) \} ,$$

and (ii) for every x ,

$$\inf_{y \in D_K} \{ \text{rec } \bar{K}(\cdot, y)(x) \} \geq \sup_{v \in D_L} \{ \text{rec } \underline{L}(\cdot, v)(A_1 x) \}$$

implies

$$\inf_{y \in D_K} \{ \text{rec } \bar{K}(\cdot, y)(-x) \} \geq \sup_{v \in D_L} \{ \text{rec } \underline{L}(\cdot, v)(-A_1 x) \} .$$

Proof. We shall show that condition (i) is equivalent to $A_2^* \text{ri} D_{L^*} \cap \text{ri} D_{K^*} \neq \emptyset$. A similar argument shows (ii) equivalent to $A_1^* \text{ri} C_{L^*} \cap \text{ri} C_{K^*} \neq \emptyset$. First, observe that for a subspace S and a nonempty convex set T , the condition $S \cap \text{ri} T \neq \emptyset$ is equivalent to the condition

$$\forall s^*, s^* \in S^\perp \text{ and } \sup_{t \in T} \langle t, s^* \rangle \leq 0 \implies -s^* \in S^\perp \text{ and } \sup_{t \in T} \langle t, -s^* \rangle \leq 0 .$$

This can be proved using (11.3) and (11.1). Now take S and T to be the sets

$$S = \{ (w, t) \mid t = A_2^* w \} , \quad T = D_{L^*} \times D_{K^*} .$$

Trivially, $S^\perp = \{ (v, y) \mid v + A_2 y = 0 \}$ and

$$\sup_{(w, t) \in T} \langle (w, t), (v, y) \rangle = \sup_{w \in D_{L^*}} \langle w, v \rangle + \sup_{t \in D_{K^*}} \langle t, y \rangle .$$

Also, Lemma 4 implies that

$$\sup_{w \in D_{L^*}} \langle w, v \rangle = -\inf_{u \in C_L} \{\text{rec } \bar{L}(u, \cdot)(-v)\}$$

and

$$\sup_{t \in D_{K^*}} \langle t, y \rangle = \sup_{x \in C_K} \{\text{rec } \underline{K}(x, \cdot)(y)\} .$$

Combining these facts, we see that

$$(v, y) \in S^\perp \quad \text{and} \quad \sup_{(w, t) \in T} \langle (w, t), (v, y) \rangle \leq 0$$

occur if and only if $(v = -A_2y$ and)

$$\sup_{x \in C_K} \{\text{rec } \underline{K}(x, \cdot)(y)\} \leq \inf_{u \in C_L} \{\text{rec } \bar{L}(u, \cdot)(A_2y)\} .$$

Hence it is clear that (i) holds if and only if $A_2^* \text{ri}D_{L^*} \cap \text{ri}D_{K^*} \neq \emptyset$, and the proof is complete.

LEMMA 6. *If there exists a pair $(x, y) \in X \times Y$ such that the level sets*

$$\{x' \mid \bar{K}(x', y) - \underline{L}A(x', y) \geq \alpha\} \quad \text{and} \quad \{y' \mid \underline{K}(x, y') - \bar{L}A(x, y') \leq \alpha\}$$

are bounded for every $\alpha \in R$, then (II) is strongly consistent.

Proof. Assume such a pair (x, y) exists. Then by (34.3), (9.5) and (9.3), the function

$$f = \underline{K}(x, \cdot) - \bar{L}A(x, \cdot)$$

is closed proper convex and

$$\text{rec } f = \text{rec } \underline{K}(x, \cdot) - (\text{rec } \bar{L}(A_1x, \cdot))A_2 .$$

By hypothesis, the level sets of f are bounded, so that by (8.7), f has no recession vectors. Thus,

$$\forall y', \text{rec } \underline{K}(x, \cdot)(y') \leq \text{rec } \bar{L}(A_1x, \cdot)(A_2y') \implies y' = 0 .$$

Now let y' be any vector satisfying the hypothesis of condition (i) of Lemma 5. Since $x \in C_K$ and $A_1x \in C_L$, we conclude that

$$\text{rec } \underline{K}(x, \cdot)(y') \leq \text{rec } \bar{L}(A_1x, \cdot)(A_2y') ,$$

which by the fact just established implies $y' = 0$. Clearly $y' = 0$ satisfies the conclusion of condition (i) of Lemma 5. In a similar way, one can verify that condition (ii) of Lemma 5 is satisfied. Strong consistency of (II) now follows by Lemma 5.

Finally, we need the following simple result.

LEMMA 7. *Let f be a proper convex function on R^n . Then the following two conditions are equivalent:*

- (i) $\{x \mid f(x) \leq \alpha\}$ is bounded for every $\alpha \in R$;
- (ii) $\{x \mid \text{cl} f(x) \leq \alpha\}$ is bounded for every $\alpha \in R$.

These conditions imply

- (iii) $(\text{rec } f)(x) > 0$ for every $x \neq 0$;

the converse holds when f is closed.

Proof. By (7.6) and (8.7).

Proof of Theorem 6. By strong consistency, Theorem 3 and (37.5), the Kuhn-Tucker conditions are satisfiable if and only if H has a saddle point. By strong consistency and Lemma 1, H is closed and proper. Therefore, by (37.5.3) the condition $(0, 0) \in \text{int}(\text{dom } H^*)$ implies both that the Kuhn-Tucker conditions are satisfiable and (by (23.4)) that the set of optimal solutions of (I) is bounded. To establish the condition $(0, 0) \in \text{int}(\text{dom } H^*)$, it suffices by (13.3) and Lemma 4 to show that

$$\text{rec } \underline{H}(x, \cdot)(y') > 0 \text{ for every } y' \neq 0$$

and

$$\text{rec } \bar{H}(\cdot, y)(x') < 0 \text{ for every } x' \neq 0.$$

But these two conditions are immediate consequences of the boundedness hypothesis, Lemma 7, and the identities

$$\underline{H}(x, \cdot) = \text{cl}(K(x, \cdot) - LA(x, \cdot) + \delta(\cdot \mid Y))$$

and

$$\bar{H}(\cdot, y) = \text{cl}(K(\cdot, y) - LA(\cdot, y) - \delta(\cdot \mid X)).$$

It remains only to show that (II) is strongly consistent. Using the strong consistency of (I) once more, one can show routinely that

$$\underline{H}(x, \cdot) = \underline{K}(x, \cdot) - \bar{L}A(x, \cdot)$$

and

$$\bar{H}(\cdot, y) = \bar{K}(\cdot, y) - \underline{L}A(\cdot, y).$$

By these identities combined with the previous ones, the boundedness hypothesis, and Lemma 7, it follows that

$$\{x' \mid \bar{K}(x', y) - \underline{L}A(x', y) \geq \alpha\} \quad \text{and} \quad \{y' \mid \underline{K}(x, y') - \bar{L}A(x, y') \leq \alpha\}$$

are bounded for each $\alpha \in R$. Hence (II) is strongly consistent by Lemma 6.

REFERENCES

1. W. Fenchel, *Convex Cones, Sets and Functions*, mimeographed lecture notes, Princeton University, 1951.
2. H. W. Kuhn and A. W. Tucker, *Nonlinear programming*, in Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability, University of California Press, Berkeley, Calif., 1951, 481-492.
3. L. McLinden, *Dual operations on saddle functions*, Trans. Amer. Math. Soc., **179** (1973), 363-381.
4. ———, *A decomposition principle for minimax problems*, in Decomposition of Large-Scale Problems, D. M. Himmelblau, ed., North-Holland, Amsterdam, 1973, 427-435.
5. R. T. Rockafellar, *Minimax theorems and conjugate saddle-functions*, Math. Scand., **14** (1964), 151-173.
6. ———, *Duality and stability in extremum problems involving convex functions*, Pacific J. Math., **21** (1967), 167-187.
7. ———, *A general correspondence between dual minimax problems and convex programs*, Pacific J. Math., **25** (1968), 597-611.
8. ———, *Convex Analysis*, Princeton University Press, Princeton, N.J., 1970.
9. ———, *Saddle-points and convex analysis*, in Differential Games and Related Topics, H. W. Kuhn and G. P. Szegö, eds., American Elsevier, New York, 1971, 109-127.
10. N. T. Tynjanskii, *Conjugate concave-convex functions in linear topological spaces and their saddle-points*, Math. USSR-Sb., **7** (1969), 503-531.

Received September 29, 1972. Sponsored by the United States Army under Contract No. DA-31-124-ARO-D-462.

MATHEMATICS RESEARCH CENTER, UNIVERSITY OF WISCONSIN
Current address: University of Illinois, Urbana