DUALITY AND STABILITY IN EXTREMUM PROBLEMS INVOLVING CONVEX FUNCTIONS

R. Tyrrell Rockafellar

In the theory of minimizing or maximizing functions subject to constraints, a given problem sometimes leads to a certain "dual" problem. The two problems are bound together like the strategy problems of the opposing players in a two-person game: neither can be solved without implicitly solving the other. The duality correspondence between linear programs is the best known example of this phenomenon. In the early 1950's Fenchel came up with a general theory of convex and concave functions on R^n which was capable of predicting and explaining the duality in many problems. This paper attempts a further development of Fenchel's theory, in both finite- and infinite-dimensional spaces. Fenchel's model problems are broadened by building a linear transformation into them. The stability of the extrema in these problems is investigated and shown to be a necessary and sufficient condition for the duality to manifest itself in full force. New light is thereby thrown on the "duality gaps" which are known to occur in some finite-dimensional convex programs and infinite-dimensional linear programs.

Let E and F be real vector spaces, finite- or infinite-dimensional, and let A be a linear transformation from E to F. Let f be a finite-valued convex function given on a nonempty convex set C in E. Let g be a finite-valued concave function given on a nonempty convex set D in F. We shall be concerned with the problem

(P)
$$\text{minimize } f(x) - g(Ax) ,$$
$$\text{subject to } x \in C \text{ and } Ax \in D .$$

This may be called a *convex program*, since it actually involves minimizing a certain convex function over the convex set in E consisting of the vectors satisfying the given constraints. If f and g are identically 0 on C and D, (P) reduces to the problem of finding a vector $x \in C$ such that $Ax \in D$.

The theory of conjugate functions devised by Fenchel [5] will enable us to construct a problem dual to (P). It is a concave program similar in form to (P), namely

(P*)
$$\text{maximize } g^*(y^*) - f^*(A^*y^*) ,$$
$$\text{subject to } y^* \in D^* \text{ and } A^*y^* \in C^* .$$

Here A^* is the linear transformation from F^* to E^* adjoint to A, where E^* and F^* are spaces paired with E and F. The elements f^*, C^*, g^*, D^*, are defined by the conjugate operation. (The pertinent facts about conjugate functions in paired spaces will be summarized briefly in § 2, as background for the precise formulation of (P) and (P*) in § 3.)

Practically the same construction leads from (P*) back to (P) again, so that (P) is in turn the dual of (P*). Therefore, for each theorem we prove which relates properties of (P) to those of (P*), there is a dual theorem in which the roles of the two problems are reversed.

Fenchel [6] initiated the finite-dimensional study of (P) and (P*) in the case where $E = F$ and A is the identity transformation. An account of Fenchel's elegant duality theory may also be found in Karlin's book [9, p. 218 ff.] (The original theory contained some minor errors which were reproduced by Karlin; these will be cited in § 7 along with counterexamples.) We have already devoted a paper [18] to an infinite-dimensional extension of this special case. In that paper it was also shown how the Lagrange multiplier theory of convex programs could be deduced by calculating the subdifferential of the function being minimized. Subdifferentials will again be useful here.

One advantage of the model problem (P), as opposed to the earlier one where $E = F$ and $A = I$, lies in its flexibility. For example, it is very easy to express a linear program as (P), and the dual problem (P*) then will be the familiar dual linear program (see § 3). This was not so in Fenchel's theory. A lengthy and complicated argument was needed there to derive the linear programming duality from the kind involving conjugate functions [6, p. 113 ff.]. Various other well known dual pairs of problems, such as quadratic programs, may also be viewed as instances of (P) and (P*). We shall not deal with such special cases below.

Our work here has been directed especially to finite-dimensional spaces. Many of the results appeared in that setting in the author's dissertation [16]. Certain theorems from [16] were announced in [17]. Generalization to infinite-dimensional spaces turns out to be virtually painless, however, and we have therefore broadened the statement of the theory to include it. We hope it may be possible to apply our results about infinite-dimensional convex programs to control theory or the calculus of variations, perhaps with A as a differential operator. Duality theorems for infinite-dimensional linear programs have already appeared in the work of Duffin [3], Kretchmer [10] and Fan [4].

The proofs of our main results are based on a device of perturbation explained in § 4. We investigate what happens to (P) if the set D and its function g are displaced from their original position by a

small amount $z \in F$. The infimum in the perturbed problem (P(z)) is a convex function of the displacement z, and its behavior at $z = 0$ can therefore be studied in the light of the theory of conjugate convex functions. This leads to a crucial notion of stability. Roughly speaking, we say (P) is "stably set" if the infimum changes only gradually when (P) is perturbed. In the unstable case, on the other hand, the infimum begins to drop off at an infinitely steep rate under the slightest displacement (or, what often amounts to the same thing, under the slightest relaxation of the constraints).

Stability criteria are developed in § 4. These mostly require that the constraints can be satisfied in some strong sense, e.g. that there exist some $x \in C$ such that Ax is an *interior* point of D.

Interest in stability has its own natural justification. It is surprising, though, that stability is also the condition on which the duality theorems in § 5 depend. We shall prove, for instance, that (P) is stably set and has a solution, if and only if (P*) is stably set and has a solution, and that the minimum in (P) then equals the maximum in (P*).

The stable case again is the one in which the solutions to (P) and (P*) can be characterized using subdifferentials. In § 8 we show that they are precisely the solutions to a certain system of "subdifferential equations" which we call the extremality conditions. For linear programs, the extremality conditions are the well known complementary slackness conditions. A minimax characterization of the solutions is given in § 9. It, too, is closely tied in with stability.

A weaker kind of duality between extrema in (P) and (P*) is brought to light in § 6. It resembles that first disclosed in the linear case by Duffin [3]. The study of this duality yields the two counterexamples given in § 7.

2. **Convex functions in paired spaces.** In this section, we shall review some facts and terminology which will be needed in the rest of the paper.

Let E and E^* be real vector spaces in duality with respect to a certain bilinear function $\langle \cdot, \cdot \rangle$. Assume that E and E^* have been assigned locally convex Hausdorff topologies compatible with this duality, so that the elements of each space can be identified with the continuous linear functionals on the other. We shall then speak of E and E^* as *topologically paired spaces*. The theory of such pairings may be found in [1, Chap. IV].

The reader who is interested primarily in the finite-dimensional case will not need a working knowledge of topological vector space theory in what follows. Instead, he can simply interpret E and E^* as R^n, with $\langle x, x^* \rangle$ as the ordinary inner product of two numerical

vectors x and x^*.

Most of Fenchel's results in [5] and [6] about convex functions have been generalized to infinite-dimensional spaces by Brøndsted [2] and Moreau [12]. We refer the reader to these papers for the details and the geometric motivation which are missing from our brief outline.

By an (infinite-valued) *convex function* f on E, we shall mean an everywhere-defined f, with values in the extended real interval $[-\infty, +\infty]$, whose (*upper*) *epigraph*

$$\{(x, \mu) \mid x \in E, \mu \in R, \mu \geqq f(x)\}$$

is a convex set in $E \oplus R$. If f does not assume both $+\infty$ and $-\infty$ as values, this convexity condition can be expressed as

$$f(\lambda x_1 + (1 - \lambda)x_2) \leqq \lambda f(x_1) + (1 - \lambda)f(x_2)$$

when $x_1 \in E$, $x_2 \in E$, $0 < \lambda < 1$. The set $\{x \mid f(x) < +\infty\}$, which is the projection of the epigraph of f on to E, is convex when f is convex. We call it the *effective domain* of f.

A convex function f on E is said to be *proper* if $f(x) > -\infty$ for all x, and $f(x) < +\infty$ for at least one x. Then the effective domain of f is nonempty, and f is finite there. Conversely, given a finite-valued convex function f on a nonempty convex set C in E, one can set $f(x) = +\infty$ for all $x \notin C$. In this way one gets a proper convex function on E having C as its effective domain.

A convex function f on E is *lower semi-continuous* (l.s.c.) if, for each real μ, the convex level set

$$\{x \in E \mid f(x) \leqq \mu\}$$

is closed. Since our duality theory will be directly applicable only to lower semi-continuous convex functions, it is important to realize that this is a constructive property. Given any convex function f on E, we can construct a l.s.c. convex function $\bar{f}$ on E, called the *l.s.c. hull* of f, by taking

$$(2.1) \qquad \bar{f}(x) = \liminf_{z \to x} f(z) \quad \text{for each} \quad x .$$

The epigraph of $\bar{f}$ is just the closure of the epigraph of f. For a l.s.c. convex function f which is not proper, the epigraph is a closed convex "vertical cylinder," so that f can not have any values other than $+\infty$ and $-\infty$.

A one-to-one correspondence between the l.s.c. proper convex functions f on E and f^* on E^* is defined by the formulas

$$(2.2) \qquad \begin{aligned} f^*(x^*) &= \sup_{x} \{\langle x, x^* \rangle - f(x)\} , \\ f(x) &= \sup_{x^*} \{\langle x, x^* \rangle - f^*(x^*)\} . \end{aligned}$$

Functions f and f^* satisfying (2.2) are said to be *conjugate* to each other.

A vector $x^* \in E^*$ is said to be a *sub-gradient* of the convex function f at the point $x \in E$ if

$$(2.3) \qquad f(x + z) \geq f(x) + \langle z, x^* \rangle \quad \text{for all} \quad z \,.$$

The set of all such sub-gradients at x is a closed convex (perhaps empty) set in E^* denoted by $\partial f(x)$. The (multiple-valued, or set-valued) mapping $\partial f \colon x \rightarrow \partial f(x)$ is called the *sub-differential* of f. If f is finite and differentiable at x in the ordinary finite-dimensional sense, $\partial f(x)$ consists of a single element, namely the ordinary gradient $\nabla f(x)$. More generally, if $f(x)$ is finite the one-sided directional derivative

$$(2.4) \qquad f'(x; z) = \lim_{\lambda \downarrow 0} [f(x + \lambda z) - f(x)]/\lambda$$

exists for every z, and it is a (positively homogeneous) convex function of z. Then $x^* \in \partial f(x)$ if and only if $f'(x; z) \geq \langle z, x^* \rangle$ for all z, and such an x^* exists if and only if $f'(x; z)$ is bounded below in z on some neighborhood of 0. The theory of directional derivatives and sub-differentials of convex functions has recently undergone considerable development; see [13], [14], [18], [19], and the papers cited there.

A function g is said to be concave if $-g$ is convex. The theory of concave functions thus parallels the theory just outlined, with only the obvious and natural changes. In particular, the formulas

$$(2.5) \qquad \begin{aligned} g^*(y^*) &= \inf_y \{ \langle y, y^* \rangle - g(y) \} \,, \\ g(y) &= \inf_{y^*} \{ \langle y, y^* \rangle - g^*(y^*) \} \,, \end{aligned}$$

define a one-to-one conjugate correspondence among upper semi-continuous proper concave functions.

3. **The dual programs.** Let E and E^* be topologically paired real vector spaces, and likewise F and F^*. Let A be a continuous linear transformation from E to F, and let A^* be its adjoint. Thus A^* is the continuous linear transformation such that

$$\langle Ax, y^* \rangle = \langle x, A^* y^* \rangle \quad \text{for all} \quad x \in E \quad \text{and} \quad y^* \in F^* \,.$$

(When $E = E^* = R^n$ and $F = F^* = R^m$, one can of course identify A with a certain $m \times n$ matrix and A^* with the $n \times m$ transpose matrix.) Furthermore, let f and f^* be lower semi-continuous proper convex functions on E and E^* conjugate to each other by formulas (2.2). Let g and g^* be upper semi-continuous proper concave functions on F and F^* conjugate to each other by (2.5). Let C be the convex set consisting of the points where f is finite (i.e. the effective domain of f), and define C^*, D and D^* similarly for f^*, g and g^*. *This notation will be*

in effect for the rest of the paper. It is summarized in Fig. 1 for convenient reference.
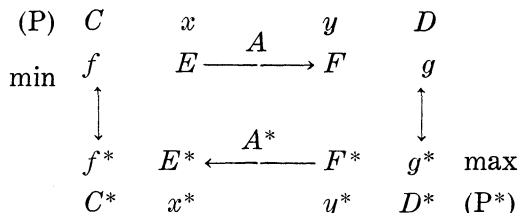
$$\text{(P)} \quad C \qquad x \qquad \quad y \qquad D$$

$$\min \quad f \qquad E \xrightarrow{\;A\;} F \qquad g$$

$$f^* \quad E^* \xleftarrow{\;A^*\;} F^* \quad g^* \quad \max$$

$$C^* \quad x^* \qquad \qquad y^* \quad D^* \quad \text{(P*)}$$

FIGURE 1.

The two extremum problems which we associate with this choice of elements are

(P) minimize $f(x) - g(Ax)$ over $x \in E$,

(P*) maximize $g^*(y^*) - f^*(A^*y^*)$ over $y^* \in F^*$.

Note that, unless it is identically $+\infty$, the minimand in (P) is a l.s.c. proper convex function on E. It is finite at x if and only if

$$(3.1) \qquad\qquad x \in C \quad \text{and} \quad Ax \in D \ .$$

Now $+\infty$ is the worst possible candidate for a minimum, so alternatively we can view the minimization in (P) as taking place subject to (3.1) instead of over all of E, as in the problem (P) described in the introduction. We therefore call conditions (3.1) the *implicit constraints* of (P) and say that (P) is *consistent* if they are satisfied by at least one x. Similarly, the conditions

$$(3.2) \qquad\qquad y^* \in D^* \quad \text{and} \quad A^*y^* \in C^*$$

are called the implicit constraints of (P*), and (P*) is said to be consistent if they can be satisfied.

Let us now demonstrate how the above scheme fits in with the duality between linear programs. Assume that E and F have been partially ordered in the usual way, i.e. by naming a closed convex cone in each space to serve as nonnegative orthant. Give E^* and F^* the dual orderings (so that $x^* \geqq 0$ if and only if $\langle x, x^* \rangle \geqq 0$ for all $x \geqq 0$, etc.) In particular, one could take $E = E^* = R^n$ and $F = F^* = R^m$ with the standard coordinatewise partial ordering. Fix any $b \in F$ and $b^* \in E^*$. The problems

(lin P) minimize $\langle x, b^* \rangle$ in x subject to $x \geqq 0$ and $Ax \geqq b$,

(lin P*) maximize $\langle b, y^* \rangle$ in y^* subject to $y^* \geqq 0$ and $A^*y^* \leqq b^*$,

are the dual linear programs in this situation. To reformulate (lin P) in the model form (P), one need only set

$$f(x) = \langle x, b^* \rangle \text{ if } x \geqq 0, \ f(x) = +\infty \text{ otherwise },$$
$$g(y) = 0 \text{ if } y \geqq b, \ g(y) = -\infty \text{ otherwise }.$$

In this elementary case, the conjugate functions may be calculated at once from (2.2) and (2.5) as

$$f^*(x^*) = 0 \text{ if } x^* \leqq b^*, f^*(x^*) = +\infty \text{ otherwise },$$
$$g^*(y^*) = \langle b, y^* \rangle \text{ if } y^* \geqq 0, g^*(y^*) = -\infty \text{ otherwise }.$$

substitution of these into (P*) yields (lin P*) as desired. The implicit constraints in (P) and (P*) correspond to the explicit ones in (lin P) and (lin P*).

The duality between (P) and (P*) is not precisely symmetric. But it is clear that (P*) is equivalent to

(P′)   minimize $f'(y^*) - g'(A^*y^*)$ over $y^* \in F^*$,

where $f'(y^*) = -g^*(y^*)$ and $g'(x^*) = -f^*(x^*)$. The conjugates f' and g' are given by $f'^*(y) = -g(-y)$ and $g'^*(x) = -f(-x)$. The dual which our theory assigns to (P′), namely

(P′*)   maximize $g'^*(x) - f'^*(Ax)$ over $x \in E$,

is therefore in turn just the negative of (P). This fact allows us immediately to dualize to (P*) any results proved for (P).

It is convenient to denote the infimum in (P) and the supremum in (P*) by inf (P) and sup (P*). Notice that (P) is inconsistent if and only if inf (P) = $+\infty$, and that (P*) is inconsistent if and only if sup (P*) = $-\infty$. Hence, information about whether the implicit constraints (3.1) and (3.2) can be satisfied will appear in our theorems in the guise of some statement about inf (P) and sup (P*). This is one of the many technical advantages which result from allowing infinite-valued functions.

A vector x will be called a *solution* to (P) if x satisfies the implicit constraints (3.1) and the infimum is achieved at x. (Thus we do not speak of solutions when (P) is inconsistent, even though the minimum $+\infty$ is trivially achieved everywhere in that case.) The solutions to (P), when they exist, evidently form a closed convex set in E. Likewise, y^* is a solution to (P*) if the supremum in (P*) is achieved at y^* and is not $-\infty$. We follow the convention of writing min (P) and max (P*), instead of inf (P) and sup (P*), in order to indicate that an extremum is achieved.

A basic fact about the relationship between (P) and (P*) can be proved right away.

LEMMA 1. $\inf (P) \geqq \sup (P^*)$.

Proof. Let $x \in E$ and $y^* \in F^*$. The inequality

(3.3) $f(x) - g(Ax) \geqq g^*(y^*) - f^*(A^*y^*)$

holds trivially if the left side is $+\infty$, or if the right side is $-\infty$.
We may therefore assume x and y^* satisfy the implicit constraints of
(P) and (P*). All four functional values in (3.3) are then finite, so
that (3.3) follows from the inequalities

(3.4)
$$f(x) + f^*(A^*y^*) \geqq \langle x, A^*y^* \rangle ,$$
$$g(Ax) + g^*(y^*) \leqq \langle Ax, y^* \rangle ,$$

which are immediate from the definition of the conjugate correspondence.

COROLLARY. *If* (P) *and* (P*) *are both consistent, then* $\inf (P)$
and $\sup (P^*)$ *are both finite.*

4, Perturbation and stability. Some results about the stability
of (P) and (P*) will now be established. These will be very important
in the subsequent study of duality. For simplicity, we shall only deal
with (P). The theorems and definitions below are to be dualized to
(P*) in the natural way.

For each $z \in F$, we consider the perturbed problem obtained by
translating the graph of g by the amount z, namely

(P(z))   minimize $f(x) - g_z(Ax)$,   where $g_z(y) = g(y - z)$.

Obviously, (P(0)) = (P).

LEMMA 2. *The function h defined by $h(z) = \inf (P(z))$ is a convex
function on F.*

Proof. To prove that h is convex, we must prove that its epigraph
is a convex set in $F \oplus R$. It is enough actually to show that, if

(4.1) $h(z_1) < \mu_1 < \infty$, $h(z_2) < \mu_2 < \infty$, $0 < \lambda < 1$,

then

(4.2) $h(\lambda z_1 + (1 - \lambda)z_2) < \lambda\mu_1 + (1 - \lambda)\mu_2$.

Given (4.1), we can always find real numbers $\mu_{11}, \mu_{12}, \mu_{21}, \mu_{22}$ and vectors
x_1 and x_2 in E such that

$$\mu_1 = \mu_{11} - \mu_{12} \quad \text{and} \quad \mu_2 = \mu_{21} - \mu_{22}$$
$$f(x_1) < \mu_{11} \quad \text{and} \quad g(Ax_1 - z_1) > \mu_{12} \ ,$$
$$f(x_2) < \mu_{21} \quad \text{and} \quad g(Ax_2 - z_2) > \mu_{22} \ .$$

Since f is a proper convex function,

$$f(\lambda x_1 + (1 - \lambda)x_2) < \lambda\mu_{11} + (1 - \lambda)\mu_{21} \ .$$

Likewise, by the concavity of g and the linearity of A,

$$g(A(\lambda x_1 + (1 - \lambda)x_2) - (\lambda z_1 + (1 - \lambda)z_2)) > \lambda\mu_{12} + (1 - \lambda)\mu_{22} \ .$$

Therefore, for $x = \lambda x_1 + (1 - \lambda)x_2$,

$$h(\lambda z_1 + (1 - \lambda)z_2) \leqq f(x) - g(Ax - (\lambda z_1 + (1 - \lambda)z_2))$$
$$< (\lambda\mu_{11} + (1 - \lambda)\mu_{21}) - (\lambda\mu_{12} + (1 - \lambda)\mu_{22}) = \lambda\mu_1 + (1 - \lambda)\mu_2 \ ,$$

which is (4.2). Thus h is convex as claimed.

We can now define what we mean by a stably set problem. This is best done by first describing instability.

Suppose that $\inf(P(0)) = \inf(P)$ is finite. For each $z \in F$, the directional derivative

(4.3) $$\lim_{\varepsilon \downarrow 0} [\inf(P(\varepsilon z)) - \inf(P)]/\varepsilon$$

exists by the convexity in Lemma 2. We shall say (P) is *unstably set*, if in every neighborhood of 0 one can choose vectors z for which this directional derivative is a negative number of arbitrarily large magnitude. (In the finite-dimensional case, it actually follows from the convexity of the directional derivative function that this happens if and only if the rate of change (4.3) is $-\infty$ in some direction.) We shall say that (P) is *stably set*, on the other hand, if it is consistent but not unstably set.

THEOREM 1. *Suppose there exists at least one $x \in E$, such that f is finite at x, and g is finite and continuous at Ax. Then* (P) *is stably set, and* $\inf(P(z))$ *is a continuous function of z in some neighborhood of $z = 0$.*

Proof. Let $h(z) = \inf(P(z))$ as in Lemma 2. Evidently the effective domain of h is

(4.4) $$\{z \mid h(z) < +\infty\} = \{Ax - y \mid f(x) < +\infty, g(y) > -\infty\}$$
$$= A(C) - D \ .$$

The hypothesis implies that $A(C)$ intersects the interior of D, and

hence that

(4.5) $0 \in A(C) - \text{int } D \subseteq \text{int } (A(C) - D)$.

Now it is a general fact [2] that a convex function is continuous at an interior point of its effective domain if and only if its epigraph has a nonempty interior. Applying this fact to the concave function g in view of our hypothesis, we see that

$$G = \{(y, \mu) \mid y \in F, g(y) \geqq \mu > -\infty\}$$

has a nonempty interior. But, for any (x, λ) in the epigraph of f, the set

$$\{(Ax - y, \lambda - \mu) \mid \mu \leqq g(y)\} = (Ax, \lambda) - G$$

is contained in the epigraph of h. The epigraph of h therefore has a nonempty interior, too. Hence h is continuous throughout the interior of its effective domain (4.4), which contains the origin by (4.5). We must still show this implies (P) is stably set. Stability would be automatic if $h(0) = -\infty$. We know on the other hand from (4.5) that $h(0) \neq +\infty$. In the remaining case, where $h(0)$ is finite, (P) is stably set if and only if the directional derivative function $h'(0; z)$ is bounded below in z in some neighborhood of $z = 0$. The conclusion we want will follow from two known elementary facts about the directional derivatives of a convex function:

(4.6)
$$h(z) \geqq h(0) + h'(0; z) \quad \text{for all} \quad z ,$$
$$h'(0; z) \geqq -h'(0; -z) \quad \text{for all} \quad z .$$

The first inequality implies, since h is continuous at 0, that $h'(0; z)$ has a finite upper bound in z on some neighborhood of 0. The second inequality translates the upper bound into a lower bound. This proves the theorem.

The stability condition in Theorem 1 implies that (P) is *super-consistent*, in the sense that there exists some x satisfying

(4.7) $x \in C \quad \text{and} \quad Ax \in \text{int } D$.

Conversely, super-consistency implies the condition in Theorem 1 in those situations where g is necessarily continuous on the interior of D. That would be true in finite-dimensional spaces, where a finite convex (or concave) function on an open set is always continuous, as is well known. More generally, it has been shown [20] that, in *tonnelé* spaces [1], a l.s.c. convex function is continuous at every interior point of its effective domain. The class of *tonnelé* spaces includes, besides

finite-dimensional spaces, all Banach spaces and all reflexive spaces. Thus we may state:

COROLLARY. *If* (P) *is super-consistent and* F *is a tonnelé space, then* (P) *is stably set, and* inf (P(z)) *is continuous in* z *in some neighborhood of* $z = 0$.

There is another useful consistency condition which guarantees stability in finite-dimensional spaces. We say that (P) is *strongly consistent* if there exists some $x \in E$ satisfying

(4.8) $x \in \text{ri } C \quad \text{and} \quad Ax \in \text{ri } D$.

Here "ri" denotes the relative interior of a finite-dimensional convex set, which is its interior with respect to the smallest affine manifold (translate of a subspace) containing it.

THEOREM 2. *If* E *and* F *are finite-dimensional and* (P) *is strongly consistent, then* (P) *is stably set.*

Proof. Due to finite-dimensionality, the convex function h, where $h(z) = $ inf (P(z)), is automatically continuous when restricted to the relative interior of its effective domain, which is given by (4.4). Thus if

(4.9) $0 \in \text{ri } (A(C) - D)$

we can show, by practically the same directional derivative argument used in the proof of Theorem 1, that (P) is stably set. Strong consistency, on the other hand, means that

(4.10) $0 \in A(\text{ri } C) - \text{ri } D$.

The problem is to establish that (4.10) implies (4.9). This is just a matter of the calculus of relative interiors. The following general rules, which lead to the equivalence of (4.9) and (4.10), are actually valid:

(4.11)
$$\text{ri } (A(C)) = A(\text{ri } C) ,$$
$$\text{ri } (D_1 + D_2) = \text{ri } D_1 + \text{ri } D_2$$

(where D_1 and D_2 are convex sets, $D_1 + D_2$ is their vector sum, etc.). The proof of these rules would be time-consuming, since we would be obliged to develop the elementary theory of relative interiors in some detail. Rather than proceed in this straightforward manner, we shall merely point out that Theorem 2 also follows from Theorem 3 below via an extension of Fenchel's theorem given by the author elsewhere

[17, Corollary 1]. This chain of proof, while extremely roundabout, is at least a space-saving expedient for us here.

It is clear that various other stability criteria could be based on special assumptions about the nature of f and g. We shall not attempt to develop them in this paper. Strong consistency is unlikely to be of any help in infinite-dimensional spaces, because relative interiors are so badly behaved there; formulas (4.11) fail almost completely, even in Hilbert spaces.

5. Strong duality theorems. We shall now prove our strongest results about the relationship between the dual programs (P) and (P*).

THEOREM 3.   *If* (P) *is stably set, then*

$$\inf (P) = \max (P^*) .$$

The latter also implies conversely that (P) *is stably set. Dually,*

$$\min (P) = \sup (P^*)$$

if and only if (P*) *is stably set.*

Proof. Let $h(z) = \inf (P(z))$ as in Lemma 2. We shall show first that (P) is stably set if and only if the convex function h is sub-differentiable at 0, i.e. $\partial h(0)$ is not empty. It is a trivial consequence of the definition (2.3) of "sub-gradient," that $\partial h(0)$ is empty when $h(0) = +\infty$, while $\partial h(0) = F^*$ when $h(0) = -\infty$. Now $h(0) = \inf (P)$, so (P) is, by definition, stably set when $h(0) = -\infty$ and not stably set when $h(0) = +\infty$. The issue is thus reduced to the case where $\inf (P)$ is finite. Then (P) is stably set if and only if the directional derivative function $h'(0; z)$ is bounded below in some neighborhood of $z = 0$. But that is precisely the known condition for subdifferentiability which we cited in § 2.

Next we shall demonstrate that $y^* \in \partial h(0)$ if and only if

(5.1) $\inf (P) \leqq g^*(y^*) - f^*(A^*y^*) .$

The existence of a y^* satisfying (5.1) is, of course, equivalent to $\inf (P) = \max (P^*)$ by Lemma 1. By definition, $y^* \in \partial h(0)$ means

$$h(z) \geqq h(0) + \langle z, y^* \rangle \quad \text{for all} \quad z \in F .$$

This is equivalent to

$$h(0) \leqq \inf_z \{h(z) - \langle z, y^* \rangle\}$$

(5.2)
$$= \inf_z \inf_x \{f(x) - g(Ax - z) - \langle z, y^* \rangle\}$$
$$= \inf_x \inf_z \{\langle Ax - z, y^* \rangle - g(Ax - z) - \langle Ax, y^* \rangle + f(x)\} .$$

Now $h(0) = \inf (P)$ is finite, so we must have

$$-\infty < \inf_z \{\langle Ax - z, y^* \rangle - g(Ax - z)\} = g^*(y^*) < +\infty$$

for any $x \in C$, and hence trivially for every x. Thus (5.2) implies

$$\inf (P) \leqq \inf_x \{g^*(y^*) - \langle Ax, y^* \rangle + f(x)\}$$
$$= g^*(y^*) - \sup_x \{\langle x, A^*y^* \rangle - f(x)\} ,$$

which is the same as (5.1). This proves the first half of the theorem. The other half, involving the stability of (P*), follows now by duality.

THEOREM 4. *If* (P) *and* (P*) *are both stably set, then both have solutions and* $+\infty > \min (P) = \max (P^*) > -\infty$.

Proof. Theorem 3 implies that $\min (P) = \max (P^*)$ when (P) and (P*) are both stably set. But the minimand in (P) never has the value $-\infty$, so that an infimum of $-\infty$ could not be attained as our use of "min" is meant to indicate. Therefore, $\min (P) > -\infty$, and dually $\max (P) < +\infty$.

THEOREM 5. (P) *is stably set and has a solution, if and only if* (P*) *is stably set and has a solution.*

Proof. By Theorem 3, the condition that (P) be stably set and have a solution is equivalent to having $\min (P) = \max (P^*)$. Dually, this is equivalent to the condition that (P*) be stably set and have a solution.

6. **Weak duality theorems.** So far, we have given conditions guaranteeing that $\inf (P) = \sup (P^*)$. The theorem below explains the exact way in which $\inf (P)$ and $\sup (P^*)$ can fail to be equal. It does this by expressing $\sup (P^*)$ in terms of the situation in (P) itself.

THEOREM 6. *The formula*

$$\sup (P^*) = \liminf_{z \to 0} [\inf (P(z))]$$

is valid, except in the trivial case where the left side is $-\infty$ *and*

the right side $+\infty$.

Proof. Let $h(z) = \inf (P(z))$ as in Lemma 2, and let $\bar{h}$ be the l.s.c. hull of h, i.e.,

$$\bar{h}(y) = \liminf_{z \to y} h(z) .$$

Then $\bar{h}$ is a l.s.c. convex function on F. We want to prove that

(6.1) $\sup (P^*) = \bar{h}(0)$,

except when $\sup (P^*) = -\infty$ and $\bar{h}(0) = +\infty$. We consider first the case where $\bar{h}$ is proper. Then $\bar{h}$ has a conjugate $\bar{h}^*$ given by

$$\bar{h}^*(y^*) = \sup_y \{\langle y, y^* \rangle - \bar{h}(y)\} \quad \text{for each} \quad y^* \in F^* ,$$

and, since $\bar{h}$ is in turn the conjugate of $\bar{h}^*$, we have

(6.2) $\bar{h}(0) = \sup_{y^*} \{\langle 0, y^* \rangle - \bar{h}^*(y^*)\}$.

Observe that

$$-\bar{h}^*(y^*) = \inf_y \{\liminf_{z \to y} h(z) - \langle y, y^* \rangle\} = \inf_z \{h(z) - \langle z, y^* \rangle\} .$$

We have already calculated the latter infimum in the proof of Theorem 3, where it turned out to be $g^*(y^*) - f^*(A^*y^*)$. The conjugate function $\bar{h}^*$ is therefore just the negative of the maximand in (P*). Hence (6.1) follows from (6.2). Now we consider the other case, where the l.s.c. convex function $\bar{h}$ is not proper. Then $\bar{h}$ can have no values other than $+\infty$ and $-\infty$. But $\bar{h}$ can not be identically $+\infty$, since h is not (the effective domain of h being the nonempty set in (4.4)). Therefore $\bar{h}$ assumes the value $-\infty$ somewhere, so that

$$- \sup_y \{\langle y, y^* \rangle - \bar{h}(y)\} = -\infty \quad \text{for all} \quad y^* .$$

Our other calculation of this supremum above, where it was $-\bar{h}^*(y^*)$, is also still valid. Therefore the maximand in (P*) is identically $-\infty$. Thus, when $\bar{h}$ is improper, the left side of (6.1) is $-\infty$, while the right side might be either $-\infty$ or $+\infty$. The formula is therefore true to the extent claimed.

Of course, there is also a formula dual to the one in Theorem 6, expressing $\inf (P)$ in terms of the situation in (P*).

The "lim inf" in Theorem 6 can also be described as the lowest μ such that

$$f(x_i) - g(y_i) \to \mu$$

for some nets (or sequences) $\{x_i\}$ and $\{y_i\}$ satisfying

$$(6.3) \qquad x_i \in C , \qquad y_i \in D , \qquad Ax_i - y_i = z_i \to 0 .$$

The problem of determining this μ is an "asymptotic" twin of (P). Let us define (P) to be *normal* if it is equivalent to its twin, i.e. if

$$(6.4) \qquad \inf (\mathrm{P}) = \lim_{z \to 0} \inf \left[\inf \mathrm{P}(z) \right] .$$

(This means that the perturbation function h is lower semi-continuous at 0.) Thus (P) is normal if and only if the lowest μ described above can be reached with $Ax_i - y_i = 0$ for all i, instead of merely $Ax_i - y_i \to 0$, whenever sequences of type (6.3) exist at all. One can define normality of (P*) in a dual manner.

An obvious way to guarantee normality would be through various compactness conditions designed to prevent bad asymptotic behavior. We shall not go into these here. Actually, such conditions would essentially be dual to continuity conditions, like the one in Theorem 1, on the conjugate functions. This is shown by the general results in [15] and [20].

The geometric picture makes it clear that abnormal programs must really be very peculiar. Nonetheless they do exist—two examples are given in the next section.

On the brighter side of things, the notion of normality allows us to state, as a corollary to Theorem 6 and its dual, another theorem like those in §5.

THEOREM 7. (P) *is normal and* $\inf (\mathrm{P})$ *is finite, if and only if* (P*) *is normal and* $\sup (\mathrm{P}^*)$ *is finite. In that case*

$$\inf (\mathrm{P}) = \sup (\mathrm{P}^*) .$$

Conversely, the latter implies both (P) *and* (P*) *are normal.*

7. **Counterexamples.** In general, $\inf (\mathrm{P})$ and $\sup (\mathrm{P}^*)$ have to satisfy

$$(7.1) \qquad + \infty \geqq \inf (\mathrm{P}) \geqq \sup (\mathrm{P}^*) \geqq - \infty ,$$

as we saw in Lemma 1. We shall now show, however, that all the relationships compatible with this basic inequality can occur. The four cases where $\inf (\mathrm{P})$ and $\sup (\mathrm{P}^*)$ are finite and equal, or both $+ \infty$, or both $- \infty$, or oppositely infinite, can be disposed of immediately. All four are already well known from the special case where (P) and (P*) are finite-dimensional dual linear programs. The two cases where

only one extremum is finite, and the case where both are finite, but unequal, will be covered by counterexamples below. The first two are dual to each other, so they only require one construction.

Examples are already known [10] of infinite-dimensional linear programs for which (7.1) holds with strict inequalities. This is impossible for finite-dimensional linear programs. We shall see in Example 1, however, that it is possible in the finite-dimensional non-linear case even in the simpler context considered by Fenchel, where $E = F$ and A is the identity. This contradicts a theorem of Fenchel [6, p. 106] which is also in Karlin's book [9, p. 229]. The error was first discovered by J. Stoer. It arises in the proof of a preliminary theorem [6, p. 95], [9, p. 222], which states a formula for the conjugate of the sum of two convex functions. The given formula might not be valid outside the relative interior of the set $\Gamma_1 + \Gamma_2$ in question. (Fenchel has pointed out to the author that approximately the same error also occurs in a neighboring theorem [6, p. 97], [9, p. 223], where it was noticed by A. Brøndsted.)

Notice that, in both of the following examples, all the extrema are attained. Both programs (P) are abnormal and unstably set, according to Theorems 3 and 7. In both cases we assume that

$$E = E^* = F = F^* = R^2 \quad \text{and} \quad A = I .$$

EXAMPLE 1. ($+\infty > \min(P) > \max(P^*) > -\infty$). Let $f(\xi_1, \xi_2) = 0$ if $\xi_1 = 0$ but $= +\infty$ otherwise. Let $g(\xi_1, \xi_2) = \min\{1, (\xi_1\xi_2)^{1/2}\}$ if $\xi_1 \geq 0$ and $\xi_2 \geq 0$, with $g(\xi_1, \xi_2) = -\infty$ otherwise. For any $z = (\eta_1, \eta_2)$, (P(z)) is the problem

$$\text{minimize } f(\xi_1, \xi_2) - g(\xi_1 - \eta_1, \xi_2 - \eta_2) \text{ over } \xi_1 \text{ and } \xi_2 ,$$

which is the same as

$$\text{minimize } f(\xi_1 + \eta_1, \xi_2 + \eta_2) - g(\xi_1, \xi_2) \text{ over } \xi_1 \text{ and } \xi_2 .$$

The minimand is $+\infty$ unless $-\eta_1 = \xi_1 \geq 0$ and $\xi_2 \geq 0$, and in the latter case it is

$$\max\{-1, -(-\eta_1\xi_2)^{1/2}\} .$$

Therefore

$$\inf(P(\eta_1, \eta_2)) = \begin{cases} -1 & \text{if } \eta_1 < 0 , \\ 0 & \text{if } \eta_1 = 0 , \\ +\infty & \text{if } \eta_1 > 0 . \end{cases}$$

It follows from Theorem 6 that $\sup(P^*) = -1$, although $\inf(P) = 0$. Obviously $\inf(P)$ is attained at $\xi_1 = 0$ and $\xi_2 = 0$. On the other hand,

$$f^*(0, 0) = \sup \{-f(\xi_1, \xi_2) \mid (\xi_1, \xi_2) \in R^2\} = 0 \ ,$$
$$g^*(0, 0) = \inf \{-g(\xi_1, \xi_2) \mid (\xi_1, \xi_2) \in R^2\} = -1 \ .$$

The maximand in (P*), which is just $g^* - f^*$, hence attains its supremum at $\xi_1^* = 0$ and $\xi_2^* = 0$. We are therefore justified in writing max (P*) and min (P) in place of sup (P*) and inf (P).

EXAMPLE 2. $(+\infty > \min (P) > \max (P^*) = -\infty.)$ Again   let $f(\xi_1, \xi_2) = 0$ if $\xi_1 = 0$, but $= +\infty$ otherwise. Let $g(\xi_1, \xi_2) = (\xi_1\xi_2)^{1/2}$ if $\xi_1 \geqq 0$ and $\xi_2 \geqq 0$, but $= -\infty$ otherwise. Arguing as in Example 1 we get

$$\inf (P(\eta_1, \eta_2)) = \begin{cases} -\infty & \text{if} \quad \eta_1 < 0 \ , \\ 0 & \text{if} \quad \eta_1 = 0 \ , \\ +\infty & \text{if} \quad \eta_1 > 0 \ . \end{cases}$$

Therefore sup (P*) $= -\infty$ by Theorem 6. The maximand in (P*) is then identically $-\infty$, so the supremum is trivially attained. On the other hand,

$$\inf (P) = \inf (P(0, 0)) = 0 \ ,$$

and this is attained at $\xi_1 = 0$ and $\xi_2 = 0$.

8. **Extremality conditions.** Solutions to convex programs have often been characterized as the solutions to certain systems of equations and inequalities. This is analogous to the situation in the calculus, where one minimizes a function by solving the equations which arise when the gradient is set equal to zero. The theory of subdifferentials is the bridge connecting these two situations, as we have already tried to demonstrate in [18]. We shall now explain how the same idea can be put to work here.

Consider first the finite-dimensional case where the functions f and g are everywhere finite and differentiable. Then $f(x) - g(Ax)$ is minimal if and only if

(8.1) $0 = \nabla(f - g \circ A)(x) = \nabla f(x) - A^*(\nabla g(Ax)) \ .$

This leads us to ask whether the same condition, but with ordinary gradients replaced by sub-gradients, i.e. the condition that

(8.2) $0 \in \partial(f - g \circ A)(x) = \partial g(x) - A^*(\partial g(Ax)) \ ,$

might play an equally substantial role in the general theory. As a matter of fact, it is trivially true that $f - g \circ A$ achieves a finite minimum precisely at those points x where 0 is a subgradient.

Therefore, the real question in (8.2) lies in whether the subdifferential of $f - g \circ A$ can be reduced to those of f and g in the manner indicated. We shall see in a moment that this is very closely related to the question of the stability of (P).

It is known that the definitions of "sub-gradient" and "conjugate" imply the equivalence of the three conditions

$$(8.3) \qquad x^* \in \partial f(x) \,, \qquad x \in \partial f^*(x^*) \,, \qquad f(x) + f^*(x^*) \leqq \langle x, x^* \rangle \,.$$

Likewise, in the concave case

$$(8.4) \qquad y^* \in \partial g(y) \,, \qquad y \in \partial g^*(y^*) \,, \qquad g(y) + g^*(y^*) \geqq \langle y, y^* \rangle \,,$$

are equivalent. Therefore 0 belongs to the set at the right of (8.2) if and only if

$$(8.5) \qquad Ax \in \partial g^*(y^*) \quad \text{and} \quad A^*y^* \in \partial f(x)$$

for some y^*. We call relations (8.5) the *extremality conditions* for (P) and (P*).

The extremality conditions can be expressed in many ways using the equivalences in (8.3) and (8.4). For instance, we can put them in the form: $x \in \partial f^*(x^*)$ and $y^* \in \partial g(y)$, where $x^* = A^*y^*$ and $y = Ax$. This says that the extremality conditions can be satisfied if and only if it is possible to complete a circuit via the four mappings indicated in Fig. 2.

FIGURE 2.
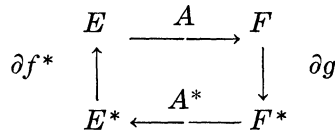
THEOREM 8. (P) *and* (P*) *are stably set if and only if the extremality conditions can be satisfied. In fact, in that case, $\bar{x}$ is a solution to* (P) *and $\bar{y}^*$ is a solution to* (P*), *if and only if $\bar{x}$ and $\bar{y}^*$ satisfy the extremality conditions.*

Proof. Suppose first that $\bar{x}$ and $\bar{y}^*$ satisfy the extremality conditions. By means of the equivalence in (8.3) and (8.4), we can express these conditions in the alternate form:

$$(8.7) \qquad \begin{aligned} f(\bar{x}) + f^*(A^*\bar{y}^*) &\leqq \langle \bar{x}, A^*\bar{y}^* \rangle \,, \\ g(A\bar{x}) + g^*(\bar{y}^*) &\geqq \langle A\bar{x}, \bar{y}^* \rangle \,. \end{aligned}$$

These imply that

(8.8) $$f(\bar{x}) - g(A\bar{x}) \leqq g^*(\bar{y}^*) - f^*(A^*\bar{g}^*) \, .$$

Hence $\bar{x}$ is a solution to (P) and $\bar{y}^*$ is a solution to (P*) by Lemma 1. Furthermore, we have $\min(\text{P}) = \max(\text{P*})$, so that (P) and (P*) are stably set by Theorem 3. Conversely, suppose (P) and (P*) are both stably set. By Theorem 4 both (P) and (P*) have solutions, and for any two such solutions $\bar{x}$ and $\bar{y}^*$ (8.8) is true. Now

$$f(x) + f^*(x^*) \geqq \langle x, x^* \rangle$$

is true all x and x^* by definition of the conjugate correspondence in (2.2), and similarly $g(y) + g^*(y^*) \leqq \langle y, y^* \rangle$. Therefore (8.8) implies

$$\langle \bar{x}, A^*\bar{y}^* \rangle \leqq f(\bar{x}) + f^*(A^*\bar{y}^*) \leqq g(A\bar{x}) + g^*(\bar{y}^*) \leqq \langle A\bar{x}, \bar{y}^* \rangle \, .$$

This in turn yields the extremality conditions in the form of (8.7).

COROLLARY. *Suppose* (P) *is stably set. Then $\bar{x}$ is a solution to* (P) *if and only if there exists a $\bar{y}^*$ such that $\bar{x}$ and $\bar{y}^*$ satisfy the extremality conditions.*

Proof. This is immediate from Theorem 5 and the present theorem.

The extremality conditions tie together situations which otherwise would seem unrelated. They become a system of partial derivative equations when f and g^* are differentiable in the usual way. At the other end of the spectrum, consider the case where (P) and (P*) are the dual linear programs (lin P) and (lin P*) described in § 3. The extremality conditions, when we express them in form (8.7) and substitute the particular functions in question, then reduce to

$$\bar{x} \geqq 0 \, , \quad b^* - A^*\bar{y}^* \geqq 0 \, , \quad \langle \bar{x}, b^* - A^*\bar{y}^* \rangle = 0 \, ,$$
$$A\bar{x} - b \geqq 0 \, , \quad \bar{y}^* \geqq 0 \, , \quad \langle A\bar{x} - b, \bar{y}^* \rangle = 0 \, .$$

These are the complementary slackness conditions for linear programs.

9. Minimax characterization. The finite-valued function K on $C \times D^*$ defined by

$$K(x, y^*) = f(x) + g^*(y^*) - \langle Ax, y^* \rangle$$

will be called the *Kuhn-Tucker function* of (P) and (P*). It is, of course, convex in x and concave in y^*. A pair $(\bar{x}, \bar{y}^*)$ is said to be a *saddle-point* of K if $\bar{x} \in C, \bar{y}^* \in D^*$, and

(9.1) $$K(\bar{x}, \bar{y}^*) = \min_{x \in C} K(x, \bar{y}^*) = \max_{y^* \in D^*} K(\bar{x}, y^*) \, .$$

Then the real number $K(\bar{x}, \bar{y}^*)$ in (9.1) is called the *minimax* of K.

THEOREM 9. *The Kuhn-Tucker function K has a saddle-point if and only if* (P) *and* (P*) *are stably set. In the latter case,* $(\bar{x}, \bar{y}^*)$ *is a saddle-point of K if and only if $\bar{x}$ is a solution to* (P) *and $\bar{y}^*$ is a solution to* (P*). *Then*

$$(9.2) \qquad \text{minimax } K = \min(\text{P}) = \max(\text{P*}) .$$

Proof. Observe first that, for any $\bar{x} \in C$ and $\bar{y}^* \in D^*$,

$$\inf_{x \in C} K(x, \bar{y}^*) = g^*(\bar{y}^*) - \sup_{x \in C} \{\langle x, A^* \bar{y}^* \rangle - f(x)\} = g^*(\bar{y}^*) - f^*(A^* \bar{y}^*) .$$

$$\sup_{y^* \in D^*} K(\bar{x}, y^*) = f(\bar{x}) - \inf_{y^* \in D^*} \{\langle A\bar{x}, y^* \rangle - g^*(y^*)\} = f(\bar{x}) - g(A\bar{x}) .$$

Therefore $(\bar{x}, \bar{y}^*)$ is a saddle-point of K if and only if

$$(9.3) \qquad f(\bar{x}) - g(A\bar{x}) = g^*(\bar{y}^*) - f^*(A^* \bar{y}^*) ,$$

in which case the latter is also the minimax of K. Of course, by Lemma 1, (9.3) is equivalent to $\min(\text{P}) = \max(\text{P*})$ being attained at $\bar{x}$ and $\bar{y}^*$. The desired conclusion therefore follows from Theorem 3.

COROLLARY. *Suppose* (P) *is stably set. Then $\bar{x}$ is a solution to* (P) *if and only if there exists a $\bar{y}^*$ such that $(\bar{x}, \bar{y}^*)$ is a saddle-point of the Kuhn-Tucker function K.*

Proof. Obvious from Theorem 5 and the theorem above.

REFERENCES

1. N. Bourbaki, *Espaces Vectoriels Topologiques*, (Hermann & Cie., Paris, 1955).

2. A. Brøndsted, *Conjugate convex functions in topological vector spaces*, Mat. Fys. Medd. Dansk. Vid. Selsk. **34** (1964).

3. R. J. Duffin, *Infinite programs*, in [11], 157-170.

4. K. Fan, *A generalization of the Alaoglu-Bourbaki theorem and its applications*, Math. Zeitschr. **88** (1965), 48-60.

5. W. Fenchel, *On conjugate convex functions*, Canad. J. Math. **1** (1949), 73-77.

6. ———, *Convex Cones, Sets and Functions*, lecture notes (mimeograph, Princeton 1953).

7. D. Gale, H. W. Kuhn and A. W. Tucker, *Linear programming and the theory of games*, in *Activity Analysis of Production and Allocation* (T. C. Koopmans, ed.), Wiley, New York, 1951.

8. A. J. Goldman and A. W. Tucker, *Theory of linear programming*, in [11], 53-98.

9. S. Karlin, *Mathematical Methods and Theory in Games, Programming and Economics*, Vol. I, Addison Wesley, Reading, Mass., 1960.

10. K. S. Kretchmer, *Programmes in paired spaces*, Canad. J. Math. **13** (1961), 221-238.

11. H. W. Kuhn and A. W. Tucker (eds.), *Linear Inequalities and Related Systems*, Ann. of Math. Study **38** (Princeton, 1956).

12. J.-J. Moreau, *Fonctions convexes en dualité* (multigraph, Séminaires de Mathematique, Faculté des Sciences, Université de Montpellier, 1962)

13. ————, *Etude locale d'une fonctionelle convexe* (multigraph, Séminaires de Mathematique, Faculté des Sciences, Université de Montpellier, 1963).

14. ————, *Fonctionelles sous-differentiables*, C. R. Acad. Sci. **257** (1963), 4117–4119.

15. ————, *Sur la polaire d'une fonctionelle semi-continue supérieurment*, C. R. Acad. Sci. **258** (1964).

16. R. T. Rockafellar, *Convex Functions and Dual Extremum Problems* (Dissertation, multilith, Harvard, 1963).

17. ————, *Duality theorems for convex functions*, Bull. Amer. Math. Soc. **70** (1964), 189–192.

18. ————, *Extension of Fenchel's duality theorem for convex functions*, Duke Math. J. **33** (1966), 81–90.

19. ————, *Characterization of the subdifferentials of convex functions*, Pacific J. Math. **17** (1966), 497–510.

20. ————, *Level sets and continuity of conjugate convex functions*, Trans. Amer. Math. Soc. **123** (1966), 46–63.

21. ————, *Minimax theorems and conjugate saddle-functions*, Math. Scand. **14** (1964), 151–173.

22. ————, *Helly's theorem and minima of convex functions*, Duke Math. J. **32** (1965), 381–398.

23. ————, *Conjugates and Legendre transforms of convex functions*, Canad, J. Math. (Jan., 1967).

24. ————, *A monotone convex analog of linear algebra*, to appear in the proceedings of the 1965 Copenhagen Colloquium on Convexity.

25. P. Wolfe, *A duality theorem for non-linear programming*, Q. Appl. Math. **19** (1961), 239–244.

UNIVERSITY OF WASHINGTON