MATH. SCAND. 48 (1981), 189–204

A LAGRANGE MULTIPLIER THEOREM AND A SANDWICH THEOREM FOR CONVEX RELATIONS

J. M. BORWEIN

Abstract.

We formulate and prove various separation principles for convex relations taking values in an order complete vector space. These principles subsume the standard ones.

Introduction.

In this note we formulate and establish a set-valued version of the Lagrange multiplier theorem and the sandwich theorem for convex-operators ([7], [17], [26]). Certain preliminary definitions need to be made.

1.

Let Z, X and Y be real topological vector spaces. Let $S \subset Y$ be a convex cone which induces an *order-complete* ordering on Y by $y \geqq 0$ if $y \in S$. We assume throughout, for simplicity, that S is *pointed*: $S \cap -S = 0$. We will also assume that S is *normal* for the topology on Y. This is to say that there is a base of neighbourhoods V of zero in Y with

$$(V-S) \cap (S-V) \subset V. \tag{1.1}$$

A relation $H\colon X \to Y$ is *convex* if its *graph* in $X \times Y$

$$\operatorname{Gr} H = \{(x,y) \mid y \in H(x),\ x \in X\}, \tag{1.2}$$

is convex. Similarly H is a *convex process* if its graph is a convex cone in $X \times Y$ containing the origin. The *domain* of H, $D(H)$, is the set of points in X for which $H(x)$ is non-empty. The *inverse* relation H^{-1} is defined by

$$y \in H(x) \Leftrightarrow x \in H^{-1}(y) \tag{1.3}$$

and clearly H is convex (a convex process) exactly when H^{-1} is. The *range* of

Research Partially Supported by NSERC.
Received February 15, 1980.

H, $R(H)$, is $D(H^{-1})$. There is a one-to-one identification between relations $H: X \to Y$ and sets $\operatorname{Gr} H$ in $X \times Y$. In [7] the study of product space sets was used to produce algebraic separation theorems. Our results could be derived in this framework but we take for us the more natural course of looking at topologized relations. A relation $H: X \to Y$ is *lower semi-continuous* at x_0 in $D(H)$ if for every y in $H(x_0)$ and neighbourhood V of zero in Y one can find a neighbourhood U of zero in X with

$$x \in U + x_0 \Rightarrow H(x) \cap (V+y) \neq \emptyset . \tag{1.4}$$

The points at which H is lower semi-continuous will be denoted $\operatorname{LC}(H)$. A relation $H: X \to Y$ is *open* at y_0 in $R(H)$ if $H^{-1}: Y \to X$ is lower semi-continuous at y_0. This is equivalent to saying that for any x in $H^{-1}(y_0)$ and any neighbourhood V of zero in Y

$$y_0 \in \operatorname{int} H(V+x) . \tag{1.5}$$

For general relations it is possible for H to be lower semi-continuous at some y in $H(x_0)$ but not at every y. For convex relations this is not possible ([4] and Proposition 2.1 (a)). The reader is referred to [1], [12] for information on general relations, to [2], [4], [14], [20], [21] for information on convex relations and to [13], [18] for prerequisite material on ordered topological spaces.

We conclude this section by recalling the relationships between convex functions and relations. Let $f: D \subset X \to Y$. Then f is *B-convex* for some convex cone B exactly when its *epigraph*

$$\operatorname{Epi} f = \{(x,y) \mid y \in f(x)+B,\ x \in X\} , \tag{1.6}$$

is convex, and so exactly when the relation

$$H_f(x) = f(x)+B \tag{1.7}$$

is convex. Finally $D(H_f)=D$.

2. Properties of convex relations.

We begin by defining two ways of making new relations from old. Let H_1 and H_2: $H \to Y$. The negative *infimal convolution* $H_1 \boxminus H_2$ of H_1 and H_2 is defined by

$$\operatorname{Gr}(H_1 \boxminus H_2) = \operatorname{Gr} H_1 - \operatorname{Gr} H_2 . \tag{2.1}$$

The *homogenization* of $H: X \to Y$, cone H, is defined by

$$\operatorname{Gr}(\operatorname{cone} H) = \operatorname{cone}(\operatorname{Gr} H) . \tag{2.2}$$

It is clear from (2.1) and (2.2) that convolution and homogenization preserve convexity.

We now collect some information about convex relations.

PROPOSITION 2.1. *Let* $H_1, H_2\colon X \to Y$. *Let* $X = X_1 \times X_2$ *be the product of two topological vector spaces.*

(a) *If* $H_1 \subset H_2$ *with* H_1 *lower semi-continuous at* x_0 *and* H_2 *convex, then* H_2 *is lower semi-continuous at* x_0.
(b) *If* $H_1 \boxminus H_2$ *is convex, and* $H_1(\cdot, x_2)$ *is lower semi-continuous at* x_1 *and* $H_2(x_1, \cdot)$ *is lower semi-continuous at* x_2, *then* $H_1 \boxminus H_2$ *is lower-semi-continuous at zero.*
(c) *If* $f\colon X \to Y$ *is B-convex and B is normal, then f is continuous at x if and only if* H_f *is lower semi-continuous at x.*
(d) *If X is given the finest convex topology and H is convex* LC $(H) = \operatorname{core} D(H)$.

PROOF. (a) Let $y_2 \in H_2(x_0)$ and let W be a neighbourhood of zero in Y. Let $y_1 \in H_1(x_0)$ and pick a neighbourhood V of zero with $tV + t(y_1 - y_2) \subset W$ for some $0 < t < 1$. Let $U + x_0 \subset H_1^{-1}(V + y_1)$. Then for $tu + x_0$ in $tU + x_0$ one has, by convexity,

$$(2.3) \qquad (1-t)y_2 + tH_2(u + x_0) \subset H_2(tu + x_0)$$

and since $H_2(u + x_0)$ meets $V + y_1$,

$$(2.4) \qquad H_2(tu + x_0) \cap ((1-t)y_2 + t(V + y_1)) \neq \varnothing .$$

Thus $H_2(tu + x_0)$ meets $W + y_2$. Since W and y_2 are arbitrary H_2 is lower semi-continuous at x_0.

(b) Let $y_1 \in H_1(x_1, x_2)$, $y_2 \in H_2(x_1, x_2)$. Let V be a neighbourhood in Y. Pick U_1 and U_2 neighbourhoods of zero in X_1 and X_2 with

$$(2.5) \quad H_1(u_1 + x_1, x_2) \cap (V + y_1) \neq \varnothing; \qquad H_2(x_1, u_2 + x_2) \cap (V + y_2) \neq \varnothing$$

for u_1 in U_1, u_2 in U_2. Thus

$$(2.6) \quad [H_1(u_1 + x_1, x_2) - H_2(x_1, u_2 + x_2)] \cap [(V - V) + (y_1 - y_2)] \neq \varnothing .$$

Since $H_1(u + x_1, x_2) - H_2(x_1, u_2 + x_2)$ lies in $(H_1 \boxminus H_2)\,(u_1, u_2)$ and $V - V$ may be made arbitrarily small, (2.6) shows that $H_1 \boxminus H_2$ is lower semi-continuous at $(0, y_1 - y_2)$. Now part (a) applied to $H_1 \boxminus H_2$ and K defined by

$$(2.7) \qquad K(x) = \begin{cases} H_1 \boxminus H_2(x) & x \neq 0 \\ y_1 - y_2 & x = 0 , \end{cases}$$

shows that $H_1 \boxminus H_2$ is lower semi-continuous at zero.

(c) It is clear that if f is continuous at x_0 then H_f is lower semi-continuous as the sum of lower semi-continuous maps $\{f\}$ and B. If B is normal the converse holds as follows. Assume H_f is lower semi-continuous at x_0. Let V be given and select a symmetric neighbourhood U in X with

$$f(x)+B \cap (f(x_0)+V) \neq \emptyset \tag{2.8}$$

whenever x lies in $U+x_0$. Thus when $x-x_0 \in U$

$$f(x)-f(x_0) \subset V-B\,. \tag{2.9}$$

Since f is convex

$$f(x_0)-f(x_0+(x_0-x)) \subset f(x)-f(x_0)-B\,, \tag{2.10}$$

and as U is symmetric

$$f(x)-f(x_0) \subset (V-B) \cap (B-V) \tag{2.11}$$

whenever $x-x_0$ lies in U. Since B is normal, f is continuous.

(d) Let V be a neighbourhood in Y. Let $x_0 \in \operatorname{core} D(H)$ and $y_0 \in H(x_0)$. Let $x \in D(H)$. Then $x \in H^{-1}(y)$ and for small $t>0$

$$tx+(1-t)x_0 \subset H^{-1}(ty+(1-t)y_0) \subset H^{-1}(y_0+V) \tag{2.12}$$

since V is absorbing and H^{-1} is convex. For any x in X $\alpha x+(1-\alpha)x_0$ lies in $D(H)$ for small positive α and (2.12) shows that

$$(t\alpha)x+(1-t\alpha)x_0 \in H^{-1}(y_0+V)\,. \tag{2.13}$$

Thus $H^{-1}(y_0+V)-x_0$ is a convex absorbing set in X and $x_0 \in \mathrm{LC}\,(H)$.

Similar arguments show that whenever H is convex and LC$\,(H)$ is not empty LC$\,(H)=\operatorname{core} D(H)$.

PROPOSITION 2.2. *Let H: $X_1 \times X_2 \to Z$ and F: $X_1 \times X_2 \to Y$. Suppose that $0 \in H(x_1,x_2)$ and that $H(\cdot,x_2)$ is open at zero. Suppose that $F(\cdot,x_2)$ is lower semi-continuous at x_1. Then FH^{-1}: $Z \to Y$ defined by*

$$FH^{-1}(z) = \{F(x_1,x_2) \mid z \in H(x_1,x_2)\} \tag{2.14}$$

is lower semi-continuous at zero, if FH^{-1} is convex.

PROOF. Pick y in $F(x_1,x_2) \subset FH^{-1}(0)$ and a neighbourhood V in Y. Pick a neighbourhood U_1 in X with

$$F(x,x_2) \cap (V+y) \neq \emptyset \tag{2.15}$$

if $x \in x_1 + U_1$. Pick a neighbourhood W in Z with

$$W \subset H(x_1 + U_1, x_2) . \tag{2.16}$$

Then if z lies in W, (x, x_2) lies in $H^{-1}(z)$ for some x in $x_1 + U_1$ and $F(x, x_2)$ meets $V + y$. Thus

$$FH^{-1}(z) \cap (V+y) \neq \emptyset , \tag{2.17}$$

and FH^{-1} is lower semi-continuous at one y and so at any y in $FH^{-1}(0)$ as in Proposition 2.1 (a).

REMARK. The best possible case, in some senses, which can occur in Proposition 2.2 is that X_1 be zero dimensional. Then the requirement on H is that $0 \in \operatorname{int} H(x_2)$ and the requirement on F is merely that $x_2 \in D(F)$. Here we have suppressed the dummy variable x_1. We will call H *strongly open* at 0 if $0 \in \operatorname{int} H(x_2)$. The primary example of such H is H_f given by (1.7) when *Slater's condition* holds:

$$f(x_2) \in -\operatorname{int} B, \quad x_2 \in X . \tag{2.18}$$

The other extreme case occurs when X_2 is zero dimensional. (See also [20].)

The following open mapping and closed graph theorem extends the usual linear one [23] and gives useful sufficient conditions for a relation to be open or lower semi-continuous.

PROPOSITION 2.3. *Let $H: X \to Y$ be a convex relation with a closed graph. Let X and Y be Fréchet spaces. Then H is lower semi-continuous at any point in* core $D(H)$ *and open at any point in* core $R(H)$.

PROOF. The Banach space case is given in [14], [20] albeit with slightly different terminology. The present theorem is a special case of one in [2]. Let us sketch a proof for Banach spaces due to Jameson [14] which holds for CS-*closed* relations. A relation is CS-closed if its graph is a CS-closed set. A set C is CS-closed if whenever $\lambda_n \geqq 0$, $\sum_{n=0}^{\infty} \lambda_n = 1$, $x_n \in c$ and $\sum_{n=1}^{\infty} \lambda_n x_n = x$ exists then x is in C. Clearly closed and open convex sets are CS-closed as are many others. A set is CS-compact if it is CS-closed and $\sum_{n=1}^{\infty} \lambda_n x_n$ always exists.

To prove the proposition it suffices by symmetry to consider the open case. Let us normalize things so that $0 \in H(0)$, $0 \in \operatorname{core} R(H)$. Let B be the unit ball in X. One verifies first that

$$0 \in \operatorname{core} H(B) .$$

Then one observes that B is CS-compact being convex, bounded and complete.

It is easily verified that the CS-closed image of a CS-compact set is CS-closed. Thus $H(B)$ is a CS-closed set. The Baire Category Theorem and the previous equation show that

$$0 \in \operatorname{int} \overline{H(B)} .$$

The final step is to show that CS-closed sets are *semi-closed* (have the same interior as their closure). This is very close to a lemma of Holmes on *ideally convex* sets [11].

The result remains true for CS-closed relations between complete metrizable linear spaces or for closed relations between a Frechet space and a barreled space.

For two relations $H_1, H_2: X \to Y$ we write $H_1(x) \geqq H_2(x)$ if $y_1 \geqq y_2$ whenever $y_1 \in H_1(x)$, $y_2 \in H_2(x)$.

PROPOSITION 2.4. *Let $H: X \to Y$ be a convex relation. Suppose that H is lower semi-continuous at zero. Then if*

$$H(0) \geqq 0 \tag{2.19}$$

there exists a continuous linear operator $T: X \to Y$ with

$$H(x) \geqq T(x), \qquad \forall x \in X . \tag{2.20}$$

PROOF. Let $K = \operatorname{cone} H$. Since $0 \in \mathrm{LC}(H)$ and K is a convex process, $D(K) = X$. Let

$$k(x) = \inf_s \{y \mid y \in K(x)\} . \tag{2.21}$$

Since K is a convex process

$$K(x) + K(-x) \subset K(0) \subset S ,$$

by (2.19), and any point in $-K(-x)$ is a lower bound for $K(x)$. Thus $k(x)$ exists since S is order complete. Since

$$\alpha K(x) + \beta K(y) \subset K(\alpha x + \beta y) + S \tag{2.22}$$

for any $\alpha, \beta > 0$ and x, y in X, standard properties of infima show

$$\alpha \inf K(x) + \beta \inf K(y) \in \inf K(\alpha x + \beta y) + S \tag{2.23}$$

and it follows that $k: X \to Y$ is sublinear. Moreover, $H(x) \subset k(x) + S$ so that it follows from Proposition 2.1(a) that H_k is lower semi-continuous at zero. Proposition 2.1(c) shows that k is continuous at zero and so is continuous throughout X. The Hahn–Banach extension theorem [6, p. 123] guarantees the existence of $T: X \to Y$ linearly with

$$T(x) \leqq k(x) \leqq H(x), \quad \forall x \in X, \tag{2.24}$$

Finally Proposition 2.1 (c) shows T is continuous since S is normal. Indeed all such T are equicontinuous.

Order completeness is only essential in Proposition 2.4 to ensure that k exists. If this can be established in some other way one may use results of Zowe [24] or the present author [3] to complete the proof. If a purely algebraic result is desired one can place the order topology for S on Y [18] and the finest locally convex topology on X. The same refers to all the results in the next section.

The following proposition is very similar to one of Robinson's in [20].

PROPOSITION 2.5. *Let $H_1: X \to Y_1$, $H_2: X \to Y_2$ be convex relations between topological vector spaces such that for some x_0 in $H_1^{-1}(0) \cap H_2^{-1}(0)$,*

(i) $0 \in \operatorname{int} H(x_0)$, *$H_2$ is open at* 0,
(ii) *H_1 is lower semi-continuous at x_0.*

Then H defined by

$$H(x) = (H_2(x), H_2(x)) \tag{2.25}$$

is open at 0.

PROOF. Let W_1 be a neighbourhood of zero in Y_1 with

$$W_1 - W_1 \subset H_1(x_0)\,. \tag{2.26}$$

Fix a neighbourhood U_0 of zero in X and pick a neighbourhood U_1 of zero in X with $U_1 + x_0 \subset H_1^{-1}(W_1)$ and a neighbourhood W_2 of zero in Y_2 with

$$W_2 \subset H_2(x_0 + \dot{U}_2); \quad U_2 = U_1 \cap U_0\,. \tag{2.27}$$

Let $(y_1, y_2) \in (W_1, W_2)$. Then $y_2 \in H_2(x_2)$, $x_2 - x_0 \in U_2 \subset U_1$. Thus $x_2 \in H_1^{-1}(y_0)$ for some y_0 in W_1. Also

$$(y_1 - y_0, 0) \in (W_1 - W_1, 0) \subset (H_1(x_0), H_2(x_0)) = H(x_0) \tag{2.28}$$

and

$$(y_0, y_2) \in (H_1(x_2), H_2(x_2)) = H(x_2)\,. \tag{2.29}$$

Thus

$$\tfrac{1}{2}(y_1, y_2) \in \tfrac{1}{2}H(x_0) + \tfrac{1}{2}H(x_2) \subset H(x_0 + \tfrac{1}{2}(x_2 - x_0)) \tag{2.30}$$

and

$$\tfrac{1}{2}(W_1, W_2) \subset H(x_0 + \tfrac{1}{2}U_2) \subset H(x_0 + \tfrac{1}{2}U_0)\,. \tag{2.31}$$

Since U_0 is arbitrary H is open at zero.

An important application of the proposition is that convex constraint sets of the form

$$0 \in g(x) + B, \qquad 0 \in H_2(x) \tag{2.32}$$

will give rise to an open relation (H_g, H_2) whenever H_2 is open at 0 and g satisfies Slater's condition (2.18) for some x in $H_2^{-1}(0)$ at which g is continuous. This is of use in applying Theorem 3.1.

In connection with Proposition 2.5 let us observe that the product of open relations is generally not open. Indeed even for continuous linear operators this can fail.

3. The Sandwich Theorem and The Lagrange Multiplier Theorem.

We now derive two equivalent consequences of Proposition 2.4 from which all other separation principles follow easily. In the next section we show some of the other derivations in outline as application of our methods.

Let $F\colon X = X_1 \times X_2 \to Y$, $H\colon X = X_1 \times X_2 \to Z$ be relations. We consider the program

$$\text{(P)} \qquad \mu = \inf_s \{F(x) \mid 0 \in H(x)\}\,. \tag{3.1}$$

THEOREM 3.1. (Lagrange Multipliers) *Let* (P) *be as above.* (a) *Suppose that the* *perturbation relation*

$$K = FH^{-1}\,, \tag{3.2}$$

is convex and lower semi-continuous at zero. There is then a continuous linear operator $T\colon Z \to Y$ with

$$F(x) + TH(x) \geqq \mu, \qquad \forall x \in X\,. \tag{3.3}$$

(b) *In particular, if F and H are convex and for some (x_1, x_2) in $H^{-1}(0)$,*

(i) *$H(\cdot, x_2)$ is open at zero,*
(ii) *$F(\cdot, x_2)$ is lower semi-continuous at x_1,*

then K is convex and lower semi-continuous at 0 *and* (*a*) *applies.*

PROOF. (b) Proposition 2.2 shows that K satisfies the hypotheses in (a).

(a) By Proposition 2.4 one may solve

$$K(z)-\mu \geqq T(z) \quad z \text{ in } Z . \tag{3.4}$$

Let z lie in $H(x)$. Then $F(x)$ lies in $K(z)=F(H^{-1}(z))$ and so for any z in $H(x)$ and y in $F(x)$

$$y-T(z) \geqq \mu \tag{3.5}$$

which is equivalent to (3.3).

THEOREM 3.2. (Sandwich Theorem) *Let H_1 and H_2: $X \to Y$ be convex relations such that for some point (x_1,x_2) in $X_1 \times X_2 = X$*

(*i*) *$H_1(\cdot,x_2)$ is lower semi-continuous at x_1,*
(ii) *$H_2(x_1,\cdot)$ is lower semi-continuous at x_2.*
Suppose that

$$H_1(x) \geqq H_2(x), \quad \forall x \in X . \tag{3.6}$$

There then exists a continuous linear operator T: $X \to Y$ and a constant b in Y with

$$H_1(x) \geqq T(x)-b \geqq H_2(x), \quad \forall x \in X . \tag{3.7}$$

PROOF. Let $H=H_1 \boxminus H_2$. Proposition 2.1(b) shows that H is lower semi-continuous at 0. Since H is convex and $H(0) \geqq 0$, (by 3.6), we may apply Proposition 2.4 and find a continuous linear T with

$$T(x) \leqq H(x) \quad \forall x \in X . \tag{3.8}$$

If $x=x_1-x_2$, then $H(x)$ contains $H_1(x_1)-H_2(x_2)$ and so

$$T(x_1)-H_1(x_1) \leqq T(x_2)-H_2(x_2) \tag{3.10}$$

for any x_1,x_2 in X. Let b satisfy

$$\sup_x T(x_1)-H_1(x_1) \leqq b \leqq \inf_x T(x_2)-H(x_2) . \tag{3.11}$$

Since S is order complete, (3.10) shows that b exists. Then (3.11) yields (3.7).

Much as in Theorem 3.1 one need only require that $H_1 \boxminus H_2$ be convex and lower semi-continuous at zero.

Note that Theorem 3.2 contains Proposition 2.4 as the special case where $K_1 = K$ and $\mathrm{Gr}(K_2)=0$. Note also, that Theorem 3.2 can be derived from Theorem 3.1 by setting $F(x_1,x_2,x_3,x_4)=H_1(x_1,x_2)-H_2(x_3,x_4)$ and

(3.12) $$H(x_1, x_2, x_3, x_4) = (x_1 - x_3, x_2 - x_4) .$$

Under assumptions (i) and (ii) of Theorem 3.2, F is lower semi-continuous in x_1 and x_4 at a point where H is trivially open in x_1 and x_4 at zero.

The full perturbational duality theory [8], [22] can be viewed as a form of (P) by replacing

(3.13) $$\inf_{x}{}_s f(x, u) = h(u)$$

by

(3.14) $$\inf_s \{r \mid u \in H(x, r)\} ,$$

with

(3.15) $$u \in H(x, r) \Leftrightarrow f(x, u) \leqq r .$$

Then Proposition 2.2 gives conditions for h to be continuous at 0.

Whenever $B \subset Z$ is convex and $S \subset Y$ is a convex cone we let B^s denote all (*positive*) continuous linear operators T: $Z \to Y$ with $T(B) \subset S$.

EXAMPLE 3.3. Let f: $X \to Y$ be S-convex and let g: $X \to Z$ be B-convex. Consider

(3.16) $$(\mathrm{P}_1)\ \mu = \inf_s \{f(x) \mid g(x) \leqq 0\} = \inf_s \{H_f(x) \mid 0 \in H_g(x)\} .$$

Suppose that either

(3.17) (i) $f(\hat{x})$ is defined for some $\hat{x}$ with $g(\hat{x}) \in -\operatorname{int} B$,

or that X and Z are Fréchet with H_g closed and

(3.18) (ii) f is continuous at some $\hat{x}$ with $g(\hat{x}) \in -B$ and $0 \in \operatorname{core} g(X) + B$.

Then in both cases H_g is open at zero and Theorem 3.1 (b) applies. We derive that

(3.19) $$\mu = \max_s \{\inf_x (f(x) + Tg(x)) \mid T \in B^s\} ,$$

which is the vector "strong duality" result. The fact that any solution T to (3.3) lies in B^s relies on the fact that an order complete cone is linearly closed [6].

A special case of Example 3.3 is worth examining further as it plays a central role in vector Kuhn–Tucker Theory [4], [9], [16].

THEOREM 3.4. *Let K_1 and K_2 be convex cones in X and Z respectively. Let A: $X \to Z$ be linear and suppose the convex relation defined by*

(3.20)
$$H_A(x) = \begin{cases} A(x)-K_2 & \text{if } x \in K_1 \\ \varnothing & \text{if } x \notin K_1 \end{cases},$$

is open at zero. Then

(3.21)
$$(K_1 \cap A^{-1}(K_2))^s = K_2^s A + K_1^s .$$

PROOF. Let $T_0 \in (K_1 \cap A^{-1}(K_2))^s$. Then

(3.22) (P_2)
$$0 = \inf\{T_0(x) \mid 0 \in H_A(x)\} ,$$

and as T_0 is continuous and H_A is open at 0 we may apply Theorem 3.1 (or Example 3.3) and deduce that $T\colon Z \to Y$ exists with

(3.23)
$$T_0(x) + T(A(x) - K_2) \subset S \quad \text{if } x \in K_1 ,$$

Thus $T_0 + TA \in K_1^s$ and $-T \in K_2^s$. This shows that T_0 lies in $K_2^s A + K_1^s$.

(3.21) will hold if either: X and Z are Fréchet, A is continuous and K_1 and K_2 are closed and

(3.24)
$$A(K_1) - K_2 = Z ,$$

or if

(3.25)
$$A(K_1) \cap K_2^0 \neq \varnothing .$$

Condition (3.24) extends, by much different arguments, a scalar result in [15]. Condition (3.25) extends the Krein–Rutman theorem [11], [18]. One can often directly show that H_A is open when (3.24) or (3.25) fail. At the expense of a little more notation one can formulate similar results when K_1 and K_2 are convex sets (containing zero) but not cones. Theorem 3.4 is a special subgradient result but the openness condition on H_A emerges more clearly in the Lagrange multiplier framework.

An example of Zowe's [25] can be interpreted as showing that in an non-lattice polyhedral ordering in $\mathbf{R}^3$ it is possible for k in Proposition 2.4 to fail to exist. Since order-completeness is essentially equivalent to the Hahn–Banach extension theorem holding [6], this is hardly surprising.

4. Alternate formulations.

We now sketch the derivation of several of the equivalent variants of the separation principle. We begin by defining the *conjugate* P^* of a relation $F\colon X \to Y$. If T lies in $B[X, Y]$, the continuous linear operators on X into Y, we set

$$F^*(T) = \sup\{T(x)-y \mid y \in F(x)\}\,. \tag{4.1}$$

Then if $F_1 \subset F_2$, $F_1^* \leqq F_2^*$. Also $H_f^* = f^*$ is the standard conjugate of f [25]. F^* is a convex operator from $B[X, Y]$ into Y when it is defined.

If $y_0 \leqq F(x_0)$ we define the *subgradient* of F at (x_0, y_0) by

$$\partial_{y_0}F(x_0) = \{T \in B[X, Y] \mid T(h) \leqq F(x_0+h)-y_0\}\,. \tag{4.2}$$

If $\{f\} = F$ is single-valued and $y_0 = f(x_0) - \varepsilon$ for some ε in S then

$$\partial_\varepsilon f(x_0) = \partial_{y_0}F(x_0) \tag{4.3}$$

where the left-hand side is the standard *ε-subgradient* of f at x_0 [5], [10], [22].

If y_0 actually lies in $F(x_0)$ then

$$F^*(T)+y_0 \geqq T(x_0)\,, \tag{4.4}$$

with equality exactly when

$$T \in \partial_{y_0}F(x_0)\,. \tag{4.5}$$

Indeed, for any $y_0 \leqq F(x_0)$

$$T \in \partial_{y_0}F(x_0) \Leftrightarrow F^*(T) \leqq T(x_0)-y_0\,. \tag{4.6}$$

THEOREM 4.1. (Subgradient Formula) *Let $F_1\colon X \to Y$ and $F_2\colon Z \to Y$ be convex relations. Let $A\colon X \to Z$ be linear. Suppose that*

$$\mathrm{LC}\,(F_2) \cap A(D(F_1)) \neq \emptyset\,. \tag{4.7}$$

a) *Then for any y in Y*

$$\partial_y(F_1+F_2A)(x_0) = \bigcup\,\{\partial_{y_1}F_1(x_0)+A^T\partial_{y_2}F_2(Ax_0) \mid y_1+y_2 \geqq y\}\,. \tag{4.8}$$

b) *Letting $F_1 = 0$, this becomes*

$$\partial_y(F_2A)(x_0) = A^T\partial_yF_2(Ax_0)\,. \tag{4.9}$$

c) *Letting $F_1 = \{f_1\}$, $F_2 = \{f_2\}$ and $y = f_1(x_0)+f_2(Ax_0)-\varepsilon$, this reduces to*

$$\partial_\varepsilon(f_1+f_2A)(x_0) = \bigcup\,\{\partial_{\varepsilon_1}f_1(x_0)+A^T\partial_{\varepsilon_2}f_2(Ax_0) \mid \varepsilon_1+\varepsilon_2 \leqq \varepsilon\}\,. \tag{4.10}$$

PROOF. (a) Let T_0 lie in the left hand side of (4.8). Set

(4.11)
$$F(x) = F_1(x_1)+F_2(x_2)-T_0(x_1-x_0) \quad \text{and} \quad H(x) = A(x_1)-x_2\,.$$

Apply Theorem 3.1 to derive the existence of T in $B[X, Y]$ such that for all x_1 in X_1 and x_2 in X_2, (4.11) $$F_1(x_1)+F_2(x_2)-T_0(x_1-x_0)+T(Ax_1-x_2) \geqq y\,.$$

Note that H is open in x_2 at any point for which F is lower semi-continuous in x_2.

Equivalently

(4.12) $$\inf F_1(x_1)-(T_0-TA)(x_1-x_0)+\inf F_2(x_2)-T(x_2-Ax_0) \geqq y\,.$$

Letting y_1 and y_2 be the respective infima in (4.12) we see that

(4.13) $$T_0-TA \in \partial_{y_1}F(x_0), \qquad TA \in A^T\partial_{y_2}F_2(Ax_0)\,.$$

The converse is straight forward.

(b) and (c) follow easily.

NOTE. The abuse of notation involved in writing A^T above is justified by the scalar case. In general the operator A^TS is just SA.

The scalar version of Theorem 4.1 (c) may essentially be found in [10]. The existence of subgradients for $y \leqq F(x_0)$ and x_0 in LC (F) is a special case of (4.8) in which $\operatorname{Gr} F_1=\{(x_0,0)\}$. This is in turn just a rephrasing of Proposition 2.4.

The Fenchel Theorem is an easy consequence of the last result, or of course of the Lagrange multiplier result itself.

THEOREM 4.2. (Fenchel Duality) *Suppose that F_1, F_2, A are as in Theorem* 4.1. *Then*

(4.14) $$\mu \;=\; \inf(F_1+F_2A)(x) \;=\; \max -F_1^*(A^TT)-F_2^*(-T)$$

whenever the left-hand side exists.

PROOF. By definition

(4.15) $$0 \in \partial_\mu(F_1+F_2A)(0)\,.$$

The previous result shows that for some $\mu_1+\mu_2 \geqq \mu$

(4.16) $$0 \;=\; A^TT_2+T_1$$

and

(4.17) $$T_1 \in \partial_{\mu_1}F_1(0), \qquad T_2 \in \partial_{\mu_2}F_2(0)\,.$$

Now (4.6) shows that

(4.18) $$F_1^*(T_1)+F_2^*(T_2) \;\leqq\; -\mu_1-\mu_2 \;\leqq\; -\mu\,.$$

In conjunction with (4.16), (4.18) shows that the maximum in (4.14) exceeds μ. The converse inequality follows from (4.2).

Note that by allowing subgradients with μ not in $F(x_0)$ we have established the Fenchel Theorem from the subgradient result, which usually needs attainment, rather than vice versá.

Theorem 4.2, may be extended to the case where A is an arbitrary convex relation. This allows one to recover simultaneously the formula (4.14) and the formula for the subgradient or conjugate of fG where G is an S-convex mapping and f is an S-increasing convex mapping. Indeed,

$$\inf H_f H_G(x) = fG(x) \quad \text{and} \quad (fG)^* = (H_f H_G)^*$$

and one proceeds as in (4.11).

The special case of Fenchel's original result with $A = I$ is indeed easily obtainable directly from the Sandwich Theorem. We may then assume (i) and (ii) of that theorem rather than (4.7).

We finish with a vector geometric duality theorem [19] and a result on affine minorants.

THEOREM 4.3. (Geometric Duality) *Let* $F\colon X \to Y$ *and* $H\colon X \to X$, *given by* $H(x) = x + K$, *satisfy Theorem* 3.1(b). *Let* K *be a convex cone. Then*

$$\inf\{F(x) \mid x \in K\} = \max\{-F^*(T) \mid T \in K^s\} \tag{4.19}$$

whenever the left-hand side exists.

PROOF: This follows directly from the Lagrange multiplier theorem and the definitions of K^s and F^*. Assuming that F is lower semi-continuous at some point of K, it is also a consequence of the Fenchel Theorem with $F = F_1$ and F_2 with $\operatorname{Gr} F_2 = (K, 0)$.

The Fenchel Theorem follows from (4.19) with $K = \operatorname{Gr} A$ and $F(x) = F_1(x_1) + F_2(x_2)$.

A particularly nice applicatoon of the Sandwich Theorem is the following form of Pshenichnii's characterization of constrained convex optimality without his assumption of infimal attainment.

THEOREM 4.4. *Let* $f\colon X \to Y \cup \{+\infty\}$ *be an* S*-convex function and* $C \subset X$ *a convex set such that either*

$$\text{(a)} \qquad \operatorname{int} C \cap \operatorname{dom} f \neq \varnothing, \tag{4.20}$$

or

$$\text{(b)} \qquad C \cap \operatorname{cont} f \neq \varnothing. \tag{4.21}$$

Then the following are equivalent:

(4.22) (1) $\mu \leqq \inf_s \{f(x) \mid x \in C\}$,

(2) *There exists* $T \in B[X, Y]$ *and* $p \in Y$ *such that*

$$T(x) - p \leqq f(x) - \mu \ , \tag{4.23}$$

and

$$p \leqq \inf\{T(x) \mid x \in C\} \ . \tag{4.24}$$

PROOF. Apply the Sandwich Theorem to $f - \mu$ and $-i(\cdot \mid C)$ (the *indicator* function of C).

If f is real valued and C is given by $H^{-1}(0)$ where H is a *polyhedrally convex* relation, Theorem 4.4 and the Farkas lemma combine to show that

$$H^{-1}(0) \cap \operatorname{cont} f \neq \emptyset \tag{4.25}$$

is a constraint qualification for the associated convex program.

This in turn can be used to develop convex quadratic duality theory without assuming or first showing primal attainment. Indeed this comes out for nothing along the way.

5. Conclusion.

Many other separation and duality results can be formulated in this vein. The main contention of this author is that armed with Theorems 3.1 and 3.2 in their present state of generality it is generally very easy to establish any particular result that one wishes to. One advantage of Theorem 3.1 over an equivalent, perturbational development ([22]) is that our Lagrange multiplier theorem retains the intuition of its functional progenitor and has considerable suggestive value. Moreover, many manipulations are simplified by using topologized relations as the proofs of section three indicate. Indeed certain unifying notions such as openness of constraint sets really only make sense in this generality.

REFERENCES

1. C. Berge, *Espaces topologiques*, Dunod, Paris, 1959.
2. J. M. Borwein, *Multivalued convexity; a unified approach to equality and inequality constraints*, Math. Programming 13 (1977), 163–180.
3. J. M. Borwein, *Continuity and differentiability of convex operators, Dalhousie Research Report*, 1980. (*To appear in Proc. London Math. Soc.*)
4. J. M. Borwein, *Convex relations in analysis and optimization*, NATO Advanced Study Institute on Generalized Concavity in Optimization and Economics, North Holland Publ. co. (To appear).
5. A. Brøndsted, *On the subdifferential of the supremum of two convex functions*, Math. Scand. 31 (1972), 225–230.
6. M. M. Day, *Normed linear spaces*, 3rd edition (Ergebnisse der Mathematik und ihrer Grenzgebiete 21), Springer-Verlag, Berlin - Heidelberg - New York, 1973.
7. R. Deumlich, K. H. Elster, and R. Neise, *Recent results on the separation of convex sets*, Math. Operationsforsch. Statist. Ser. Optim. 9 (1978), 273–296.
8. I. Ekeland and R. Temam, *Analyse convexe et problèmes variationnels*, Dunod, Paris, 1973.
9. M. Guignard, *Generalized Kuhn–Tucker conditions for mathematical programming in a Banach space*, SIAM J. Control Optim., 7 (1969), 232–241.
10. J.-B. Hirrart-Urruty, *Lipschitz r-continuity of the approximate subdifferential of a convex function*, Math. Scand. 47 (1980), 123–134.
11. R. B. Holmes, *Geometric functional analysis and its applications* (Graduate Texts in Mathematics 24), SpringerVerlag, Berlin - Heidelberg - New York, 1975.
12. P. Huard (editor), *Point to set maps and mathematical programming*, Math. Programming Stud. 10, (1979).
13. G. J. Jameson, *Ordered linear spaces*, (Lecture Notes in Mathematics 141), Springer-Verlag, Berlin - Heidelberg - New York, 1970.
14. G. J. Jameson, *Convex series*, Proc. Cambridge Philos., Soc. 72 (1972), 37–47.
15. S. Kurcyusz, *On existence and nonexistence of Lagrange multipliers in Banach space*, J. Optim. Theory Appl. 20 (1976), 81–110.
16. S. Kurcyusz and J. Zowe, *Regularity and stability for the mathematical programming problem in Banach spaces*, Appl. Math. Optim. 5 (1979), 49–62.
17. C. Malivert, J. P. Penot, and M. Thera, *Un prolongement du théorème de Hahn–Banach*, C. R. Acad. Sci. Paris Sér. A 286 (1978), 165–168.
18. A. L. Peressini, *Ordered topological vector spaces*, Harper and Row, New York - London, 1967.
19. E. L. Peterson, *Geometric programming*, SIAM Rev. 18 (1976), 1–51.
20. S. Robinson, *Regularity and stability for convex multivalued functions*, Math. Optr. Res. 1 (1976), 130–143.
21. S. Robinson, *Normed convex processes*, Trans. Amer. Math. Soc. 174 (1972), 127–140.
22. R. T. Rockafellar, *Convex analysis* (Princeton Mathematical Series 28), Princeton University Press, Princeton, 1970.
23. W. Rudin, *Functional analysis*, McGraw–Hill, New York, Düsseldorf, London, 1973.
24. J. Zowe, *Linear maps majorized by a sublinear map*, Arch. Math. (Basel) 26 (1975), 637–645.
25. J. Zowe, *A duality theorem for a convex programming problem in an order complete vector lattice*, J. Math. Anal. Appl. 50 (1975), 273–287.
26. J. Zowe, *Sandwich theorems for convex operators with values in an ordered vector space*, J. Math. Appl. 66 (1978), 292–296.

CARNEGIE-MELLON MATHEMATICS DEPARTMENT
PITTSBURGH, PENNSYLVANIA 15213
U.S.A.