JOURNAL OF RESEARCH of the National Bureau of Standards — B. Mathematical Sciences
Vol. 74B, No. 2, April–June 1970

On Complementary Polar Conical Sets*

Christoph Witzgall**

(February 6, 1970)

Tucker has formulated the Duality Theorem of Linear Programming in terms of orthogonality properties of a pair of complementary orthogonal linear manifolds with respect to the positive orthant. This theorem is generalized by substituting complementary polar conical sets for complementary orthogonal linear manifolds, and the generalization is proved under simple stability assumptions. Equivalence to Fenchel's Duality Theorem for conjugate convex functions is established. There are strong parallelisms to work by Kretschmer.

Key words: Conjugate functions; convex cones; duality; linear programming; orthogonality; polarity.

1. Introduction

Two linear manifolds M and N in R^n are called *orthogonal*, if $(X_1-X_2)^T(Y_1-Y_2)=0$ holds whenever $X_1, X_2 \epsilon M$ and $Y_1, Y_2 \epsilon N$. They are called *complementary orthogonal*, if they are orthogonal, if their intersection is of dimension zero, and if their dimensions add up to n.

The following formulation of the duality theorem of Linear Programming has been given by Tucker [1]:[1]

THEOREM M: *Suppose* M *and* N *are complementary orthogonal linear manifolds both of which meet the nonnegative orthant* $\mathrm{R}^n_\oplus$. *Then there exist two nonnegative vectors* U *and* V *such that* $\mathrm{U}\epsilon\mathrm{M}$, $\mathrm{V}\epsilon\mathrm{N}$, *and* $\mathrm{U}^\mathrm{T}\mathrm{V}=0$ (fig. 1).

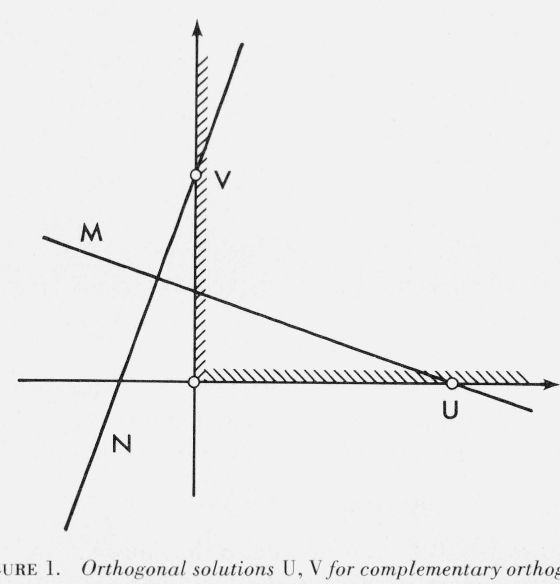

FIGURE 1. *Orthogonal solutions* U, V *for complementary orthogonal manifolds* M, N.

*An invited paper.

**Present address: Boeing Scientific Research Laboratories, Mathematics Research Laboratory, Seattle, Washington 98111.

[1] Figures in brackets indicate the literature references at the end of this paper.

382-233 O - 70 - 3

The purpose of this paper is to extend the above theorem to pairs of "complementary polar conical sets."

We say that two sets F and G are

complementary polar conical sets,

if $F=P+C$ and $G=Q+C^p$ for some points P, Q, and some cone C with $C=C^{pp}$. Here C^p denotes the (negative) *polar* of the cone C, i.e.,

$$C^p=\{Y \mid Y^TX \leq 0 \text{ for all } X\epsilon C\}.$$

We want to consider the existence of nonnegative vectors U and V such that $U\epsilon F$, $V\epsilon G$, and $U^TV=0$, where F and G are complementary polar conical sets. This statement generalizes *Theorem M*, since each pair of complementary orthogonal manifolds is also a pair of complementary polar conical sets (fig. 2). Carrying the generalization of *Theorem M* still further, we drop the nonnegativity hypothesis for U and V and require instead that

$$U\epsilon K,\ V\epsilon -K^p$$

for some given (closed) cone K (fig. 3).

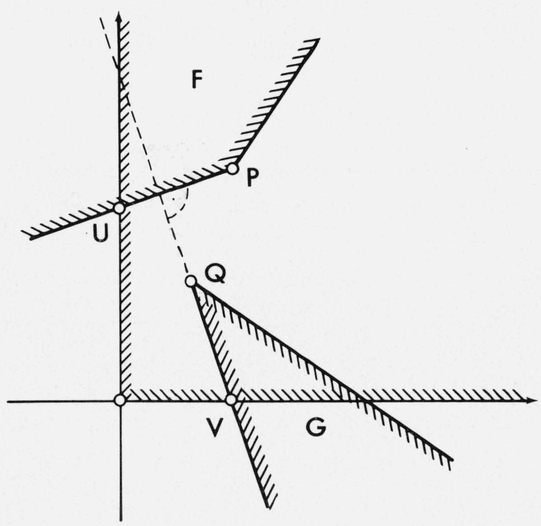

FIGURE 2. *Orthogonal solutions* U, V *for complementary polar conical sets* F, G *with respect to* $R^n_{\oplus}$.

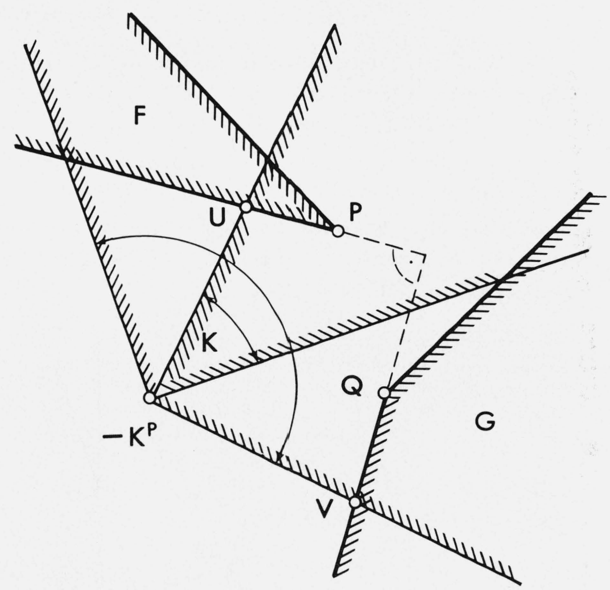

FIGURE 3. *Orthogonal solutions* U, V *for complementary polar conical sets* F, G *with respect to cones* K, $-K^p$.

In the case of manifolds, the vectors U and V are usually unique, whereas they are typically *not* unique in all other cases. Therefore we require in addition that the vectors $U-P$ and $V-Q$ be orthogonal—a condition satisfied in the case of manifolds. We say that U and V form a pair of

orthogonal solutions

of the complementary polar sets F and G with respect to the cone K.

It would appear natural to generalize even further by adding points P', Q' to K and $-K^p$, respectively. Orthogonal solutions would then be defined as satisfying

$$U\epsilon(P+C)\cap(P'+K),\ V\epsilon(Q+C^p)\cap(Q'-K^p)$$

$$(U-P)^T(V-Q)=0=(U-P')^T(V-Q').$$

However, the linear transformations $\bar{U}:=U-P'$, $\bar{V}:=V-Q'$, $\bar{P}:=P-P'$, $\bar{Q}:=Q-Q'$ reduce this case immediately to

$$\bar{U}\epsilon(\bar{P}+C)\cap K,\ \bar{V}\epsilon(\bar{Q}+C^p)\cap(-K^p),$$

$$(\bar{U}-\bar{P})^T(\bar{V}-\bar{Q})=0=\bar{U}^T\bar{V}.$$

Therefore it is no restriction of generality to assume $P'=Q'=0$, and this will be done throughout this paper.

2. Propositions

The first extension of *Theorem M* concerns the special class of *polyhedral* conical sets, that is, conical sets that are intersections of finitely many closed halfspaces. If $F=P+C$ is such a conical set, then so is $G=Q+C^p$, and $C=C^{pp}$. Here R may be any ordered field.

THEOREM P: *If* F *and* G *are complementary polar polyhedral conical sets both of which meet polyhedral cones* K *and* $-$K^p *respectively, then* F, G *possess orthogonal solutions* UϵK, V$\epsilon-$K^p.

If R is the field of real numbers, then the relation $C=C^{pp}$ characterizes closed cones. The following example shows that general complementary polar conical sets need not admit orthogonal solutions, even if both meet the nonnegative orthant $K=-K^p=R^n_\oplus$. Consider

$$C=\left\{\begin{pmatrix}x\\y\\z\end{pmatrix}\,\middle|\,-y+z\geqslant\sqrt{x^2+y^2+z^2}\right\}.$$

This is a circular cone with an opening angle of 90°. Clearly $C^p=-C$. Choose $P=Q=(1,0,0)^T$. Then (fig. 4)

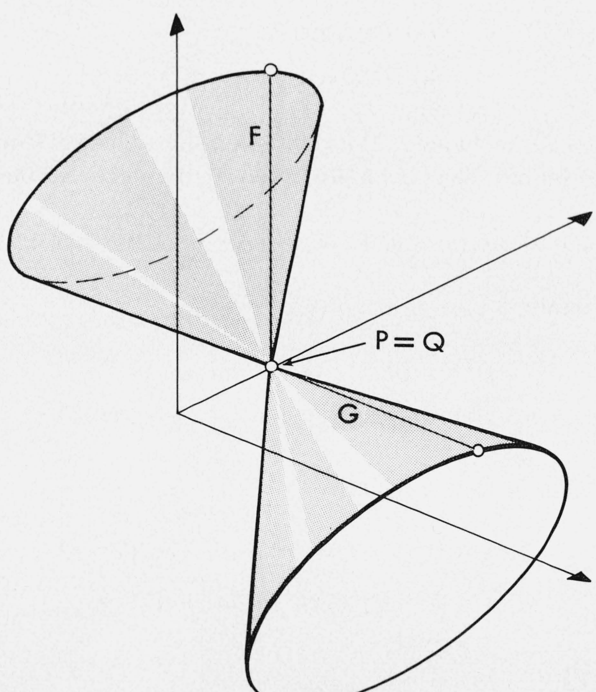

FIGURE 4. *Example for nonexistence of orthogonal solutions: Two circular cones with angle of 90° at apex, meeting the positive orthant in two rays* F *and* G.

$$F \cap R_{\oplus}^{n} = \{P + \theta(0, 0, 1)^T \mid \theta \geqslant 0\}$$

$$G \cap R_{\oplus}^{n} = \{Q + \theta(0, 1, 0)^T \mid \theta \geqslant 0\}$$

and these sets have no orthogonal points in common.

Thus we need additional conditions in order to extend *Theorem M* to general conical sets in the real space R^n.

A point D of a closed cone C is called

stable in C

if the conical hull $\mathscr{C}(C \cup (-D))$ is again closed.

If D lies in the relative interior of C, which we shall denote by C^I, then D is stable in C. Indeed,

$$D = \sum_{i=1}^{k} A_i u_i$$

where $A_i \epsilon C$, $u_i > 0$, and $k = \dim \mathscr{L}(C)$, where $\mathscr{L}(C)$ denotes the linear hull of C. Thus $A_1, \ldots,$ A_k, $-D$ span positively the linear hull of C. If therefore $D \epsilon C^I$ then $\mathscr{C}(C \cup (-D)) = \mathscr{L}(C)$, and linear subspaces of R^n are closed.

If D belongs to a polyhedral cone C, then $\mathscr{C}(C \cup (-D))$ is again a polyhedral cone and therefore closed. Thus every point of a polyhedral cone C is stable in C.

Further we need the notion of a

direction of infinity.

Vector D is a direction of infinity of F with respect to Q, and vector E a direction of infinity of G with respect to P, F

$$\begin{aligned} &\text{(i) } D \epsilon K,\ E \epsilon -K^p \\ &\text{(ii) } D \epsilon C,\ E \epsilon C^p \\ &\text{(iii) } D^T Q = 0,\ P^T E = 0, \end{aligned} \tag{1}$$

respectively. Zero vectors are included in this definition, since this will simplify statements later on. The name "directions of infinity" has been prompted by the following fact:

If U, V *are orthogonal solutions of* F, G, *then so are* $\mathrm{U} + \theta \mathrm{D}$, $\mathrm{V} + \tau \mathrm{E}$ *for all* $\theta, \tau \geqslant 0$. (2)

Indeed, (1.ii) and (1.iii) imply in view of polarity:

$$\begin{aligned} \mathrm{D^T V} &= \mathrm{D^T (V - Q)} \leqslant 0 \textit{ for all } \mathrm{V \epsilon G} \\ \mathrm{U^T E} &= \mathrm{(U - P)^T E} \leqslant 0 \textit{ for all } \mathrm{U \epsilon F}. \end{aligned} \tag{3}$$

From (1.i) one obtains again in view of polarity:

$$\begin{aligned} \mathrm{D^T V} &= \mathrm{D^T (V - Q)} = 0 \textit{ for all } \mathrm{V \epsilon G} \cap -\mathrm{K^p} \\ \mathrm{U^T E} &= \mathrm{(U - P)^T E} = 0 \textit{ for all } \mathrm{U \epsilon F} \cap \mathrm{K}. \end{aligned} \tag{4}$$

Finally, since $\mathrm{E + Q \epsilon G} \cap -\mathrm{K^p}$,

$$\mathrm{D^T E} = 0. \tag{5}$$

Proposition (2) now follows from (4) and (5): $(U + \theta D)^T (V + \tau E) = U^T V = 0$, $(U + \theta D - P)^T (V + \tau E - Q) = (U - P)^T (V - Q) = 0$. $U + \theta D \epsilon F \cap K$ and $V + \tau E \epsilon G \cap -K^p$ is plain, as $D \epsilon C \cap K$ and $E \epsilon C^p \cap -K^p$.

We are now able to formulate the general

THEOREM S: *Suppose* $F = P + C$ *and* $G = Q + C^p$ *are a pair of complementary polar sets. If* F *and* G *both meet the closed cones* K *and* $-K^p$ *respectively, and if all* [2] *directions of infinity of* F *are stable in* C *and* K, *while all directions of infinity of* G *are stable in* C^p *and* $-K^p$, *then* F, G *possess orthogonal solutions* $U \epsilon K$, $V \epsilon -K^p$.

Let $\mathscr{M}(S)$ denote the affine hull of a set $S \subseteq R^n$ and let S^I stand for the interior of S with respect to the relative topology in $\mathscr{M}(S)$. We call S^I the *relative interior* of S. Then we have as a corollary to *Theorem S*:

THEOREM I: *If* F *and* G *are complementary polar conical sets in the real space* R^n, *and if* $F^I \cap K^I \neq \emptyset$, $G^I \cap (-K^p)^I \neq \emptyset$, *then* F *and* G *possess orthogonal solutions* $U \epsilon K$, $V \epsilon -K^p$.

This follows immediately from the following

LEMMA: *If* $G^I \cap (-K^p)^I \neq \emptyset$, *then* $-D \epsilon C \cap K$ *holds for every direction of infinity* D *of* $F = P + C$.

PROOF: By (3), the plane $H = \{X \mid D^T X = 0\}$ is a supporting plane of G. Since $G^I \cap (-K^p)^I \neq \emptyset$ by hypothesis and $H \supseteq G \cap (-K^p)$ by (4), we have $H \cap G^I \neq \emptyset$ and $H \cap (-K^p)^I \neq \emptyset$. Any supporting plane of a convex set S which meets the relative interior S^I must contain S. Hence $H \supseteq G$ and $H \supseteq -K^p$. In view of (1.iii) we have $H \supseteq G - Q = C^p$. This is equivalent to $-D \epsilon C$. $-D \epsilon K$ follows from $H = -H \supseteq K^p$. —[3]

Theorem P for the real space R^n also follows from *Theorem S*. Indeed, if C and K are polyhedral cones, they have a finite number of generators, as have C^p and $-K^p$. But then so have $\mathscr{C}(C \cup \{-D\})$, $\mathscr{C}(K \cup \{-D\})$, $\mathscr{C}(C^p \cup \{-E\})$, $\mathscr{C}(-K^p \cup \{-E\})$, which are therefore again polyhedral cones and therefore closed, no matter what D and E are selected.

3. Uniqueness

Let us briefly examine the uniqueness of orthogonal solutions U, V, however only in the case $K = -K^p = R^n_{\oplus}$. If there exists an index i such that $u_i = v_i = 0$, we say that U and V have a *common zero*. In the case of manifolds, there is a known result:

THEOREM U: *Orthogonal solutions with a common zero exist if and only if the orthogonal solutions are not unique.*

The "only if" part of *Theorem U* is commonly deduced from the following statement:

THEOREM V: *If complementary orthogonal manifolds* M, N *possess orthogonal solutions at all, then they possess orthogonal solutions without a common zero.*

The analogous statements for complementary polar conical sets do not hold. Consider for example the following polyhedral conical sets:

$$F := \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \middle| \, x + y \geqslant 3,\ x - y + z \leqslant 1,\ -x + y + z \leqslant 1 \right\}$$

$$G := \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \middle| \, x + y \leqslant 3,\ -x + y + 2z \geqslant 2,\ x - y + 2z \geqslant 2 \right\}.$$

The above two sets are complementary polar, with $P = Q = (\frac{3}{2}, \frac{3}{2}, 1)^T$. The orthogonal solutions are

$$V = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix},\ U = \lambda \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} + \mu \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix},\ \lambda, \mu \geqslant 0,\ \lambda + \mu = 1.$$

---

[2] It will become apparent from the proof of Theorem S, that it suffices to require that there exists at least one stable direction in the relative interior of the cone of directions of infinity.

[3] The horizontal bar at the end of a paragraph marks termination of a proof.

They are not unique in spite of the absence of common zeros. Thus the "if" part of *Theorem U* is not valid for complementary polar conical sets.

As a counterexample for *Theorem V* consider the conical sets:

$$F=C:=\left\{\begin{pmatrix}x\\y\\z\end{pmatrix}\middle|\,-y+z\geqslant\sqrt{x^2+y^2+z^2}\right\},\ G=C^p:=-C.\ K=-K^p=R^n_\oplus.$$

Here $F\cap R^n_\oplus$ and $G\cap R^n_\oplus$ are the nonnegative parts of the z axis and the y axis, respectively. Hence $x=0$ for all $U\epsilon F\cap R^n_\oplus$ and $V\epsilon G\cap R^n_\oplus$. Note that in this example the existence of orthogonal solutions cannot be inferred from *Theorem S*.

4. Proofs

We proceed to prove *Theorem P*. The proof of *Theorem S* will then be based on *Theorem P*.

PROOF OF THEOREM P: Let $U\epsilon K$, $V\epsilon-K^p$ be vectors in F, G respectively. By the definition of the polar $(U-P)^T(V-Q)\leqslant 0$ as well as $U^TV\geqslant 0$. Thus

$$U^TV-(U-P)^T(V-Q)=U^TQ+P^TV-P^TQ\geqslant 0 \tag{6}$$

for $U\epsilon F\cap K$ and $V\epsilon G\cap(-K^p)$, and

$$U^TQ+P^TV-P^TQ=0$$

is necessary and sufficient for $U^TV=0$ and $(U-P)^T(V-Q)=0$ to hold simultaneously. The existence of a pair of orthogonal solutions U and V is therefore equivalent to the following linear program having an optimal value of zero:

$$\textit{Minimize } \mathrm{U^TQ+P^TV-P^TQ} \textit{ over } \mathrm{U, V} \textit{ subject to} \tag{7}$$
$$\mathrm{U\epsilon P+C,\ U\epsilon K}$$
$$\mathrm{V\epsilon Q+C^p,\ V\epsilon-K^p.}$$

Program (7) is *separable:* The pair U, V is an optimal solution of (7) if and only if U and V are optimal solutions of the following two programs respectively:

$$\textit{Minimize } \mathrm{U^TQ} \textit{ over } \mathrm{U} \textit{ subject to} \tag{8}$$
$$\mathrm{U\epsilon P+C,\ U\epsilon K,}$$

and

$$\textit{Minimize } \mathrm{P^TV-P^TQ} \textit{ over } \mathrm{V} \textit{ subject to}$$
$$\mathrm{V\epsilon Q+C^p,\ V\epsilon-K^p.} \tag{9}$$

By hypothesis, C and K are polyhedral cones

$$C=\{S\,|\,A^TS\leqslant 0\},\ K=\{BZ\,|\,Z\geqslant 0\}$$

$$C^p=\{AT\,|\,T\geqslant 0\},\ -K^p=\{W\,|\,B^TW\geqslant 0\}.$$

One has $U=BZ=P+S$ with $Z\geqslant 0$, $A^TS\leqslant 0$, and $V=Q+AT$ with $T\geqslant 0$, $B^TV\geqslant 0$. In terms of

Z and T the programs (8) and (9) thus reduce to

$$\textit{Minimize } \mathrm{Z}^T \mathrm{B}^T \mathrm{Q} \textit{ with } \mathrm{Z} \textit{ subject to}$$
$$-\mathrm{A}^T\mathrm{BZ} \geqslant -\mathrm{A}^T\mathrm{P} \tag{10}$$
$$\mathrm{Z} \geqslant 0.$$

$$\textit{Maximize } -\mathrm{P}^T\mathrm{AT} \textit{ with } \mathrm{T} \textit{ subject to}$$
$$-\mathrm{B}^T\mathrm{AT} \leqslant \mathrm{B}^T\mathrm{Q} \tag{11}$$
$$\mathrm{T} \geqslant 0.$$

The programs (10) and (11) are clearly *duals* of each other. The duality theorem of linear programming then gives $Z^TB^TQ=-P^TAT$ or

$$U^TQ+P^TV-P^TQ=0,$$

as $U^T=Z^TB^T$ and $AT=V-Q$, for optimal solutions Z and T of programs (10) and (11), respectively. The existence of optimal solutions follows from the hypothesis, that F and G meet K and $-K^p$, respectively, and that therefore feasible solutions to both programs exist.—

PROOF OF THEOREM S: Let D_F be the set of all directions of infinity of $F=P+C$, and similarly let D_G be the set of all directions of infinity of $G=Q+C^p$. Plainly, D_F and D_G are closed convex cones. Moreover, by the stability hypothesis of *Theorem S*,

the four cones (12)

$$\mathscr{C}(C\cup -D_F)=C-D_F=C+L_F,\ L_F:=\mathscr{L}(D_F)=D_F-D_F,$$
$$\mathscr{C}(C^p\cup -D_G)=C^p-D_G=C^p+L_G,\ L_G:=\mathscr{L}(D_G)=D_G-D_G,$$
$$\mathscr{C}(K\cup -D_F)=K-D_F=K+L_F,$$
$$\mathscr{C}(-K^p\cup -D_G)=-K^p-D_G=-K^p+L_G$$

are closed.

To see, for instance that $C+L_F$ is closed, let D_0 be any point in the relative interior D_F^I of D_F. Then $\mathscr{C}(D_F\cup\{-D_0\})=L_F$ and therefore $C+L_F=\mathscr{C}(C\cup\{-D_0\})$. The latter cone is closed since D_0 is stable by hypothesis.—

Next we show that without loss of generality, we may assume that

$$C\cap K\supseteq L_F \textit{ and } C^p\cap(-K^p)\supseteq L_G. \tag{13}$$

To prove this we shall use on several occasions the following simple lemma whose verification is left to the reader:

Let L *be a linear manifold in* R^n, *and let* A *and* B *be any sets in* R^n. *Then* $\mathrm{A}\subseteq\mathrm{L}$ *implies*

$$\mathrm{A}+(\mathrm{B}\cap\mathrm{L})=(\mathrm{A}+\mathrm{B})\cap\mathrm{L}. \tag{14}$$

We return to proving that (13) constitutes no restriction of generality. Suppose the pairs of cones C, C^p and $K, -K^p$ do not satisfy (13). Then we consider the following modified pairs

$$\hat{C}:=(C+L_F)\cap L_G^\perp,\ \hat{C}^p:=(C^p+L_G)\cap L_F^\perp.$$
$$\hat{K}:=(K+L_F)\cap L_G^\perp,\ -\hat{K}^p:=(-K^p+L_G)\cap L_F^\perp.$$

The (closed) cones $\hat{C}$ and $\hat{C}^p$ are indeed polars of each other, as are $\hat{K}$ and $\hat{K}^p$, since by (5), we have

$$L_F\subseteq L_G^\perp \textit{ and } L_G\subseteq L_F^\perp, \tag{15}$$

and (14) then gives

$$\hat{C} = (C \cap L_G^\perp) + L_F,$$

$$\hat{K} = (K \cap L_G^\perp) + L_F.$$

Polarizing these equations yields in view of (12)

$$\hat{C}^p = \overline{(C^p + L_G)} \cap L_F^\perp = (C^p + L_G) \cap L_F^\perp$$

$$\hat{K}^p = \overline{(K^p + L_G)} \cap L_F^\perp = (K^p + L_G) \cap L_F^\perp,$$

and finally

$$-\hat{K}^p = (-K^p - L_G) \cap -(L_F^\perp) = (-K^p + L_G) \cap L_F^\perp.$$

Next we verify that $\hat{C}$, $\hat{C}^p$ and $\hat{K}$, $-\hat{K}^p$ indeed satisfy (13). To this end we denote by $\hat{C}_F$, $\hat{C}_G$ the cones of directions of infinity of $\hat{F} := P + \hat{C}$, $\hat{G} := Q + \hat{C}^p$ with respect to $\hat{K}$, $-\hat{K}^p$. Let then

$\hat{D} \epsilon \hat{C}_F \subseteq \hat{C} \cap \hat{K}$. As $C + L_F = C - C_F$, since $L_F = C_F - C_F$ and $C_F \subseteq C$, we have $\hat{D} = D_1 - D_2$ where $D_1 \epsilon C$ and $D_2 \epsilon C_F$. Similarly, $\hat{D} = D_3 - D_4$ where $D_3 \epsilon K$ and $D_4 \epsilon C_F$. We claim $D_1 + D_4 \epsilon C_F$. Indeed, $(D_1 + D_4)^T Q = (\hat{D} + D_2 + D_4)^T Q = \hat{D}^T Q + D_2^T Q + D_4^T Q = 0$; $D_1 + D_4 \epsilon C$ as $D_1 \epsilon C$, $D_4 \epsilon C_F \subseteq C$; $D_1 + D_4 = D_2 + D_3 \epsilon K$ as $D_2 \epsilon C_F \subseteq K$, $D_3 \epsilon K$. Clearly, $D_2 + D_4 \epsilon C_F$. Thus $\hat{D} = (D_1 + D_4) - (D_2 + D_4) \epsilon C_F - C_F = L_F$. This proves $\hat{C}_F \subseteq L_F$. Now $L_F \subseteq \hat{C} \cap \hat{K}$ by the definitions of $\hat{C}$ and $\hat{K}$, and in view of $L_F \subseteq L_G^\perp$ (15). Moreover, $\hat{D}^T Q = 0$ for every $\hat{D} \epsilon L_F = C_F - C_F$ by (1.iii). Thus, and by an analogous argument for $\hat{G}$,

$$\hat{C}_F = L_F \textit{ and } \hat{C}_G = L_G. \tag{16}$$

This clearly implies (13) for $\hat{C}$ and $\hat{K}$. Moreover it shows that $\mathscr{C}(\hat{C} \cup \{-\hat{D}\}) = \hat{C}$, $\mathscr{C}(\hat{C}^p \cup \{-\hat{E}\}) = \hat{C}^p$ for all $\hat{D} \epsilon \hat{C}_F$ and $\hat{E} \epsilon \hat{C}_G$. In other words, all directions in infinity of $\hat{F}$, $\hat{G}$ are stable (the reader should recall at this point, that the construction of F, G did require that the original sets F, G had only stable directions of infinity).

By (4),

$$F \cap K \subseteq L_G^\perp \textit{ and } G \cap (-K^p) \subseteq L_F^\perp. \tag{17}$$

Since $P \epsilon L_G$ by (1.iii), application of (14) yields

$$\hat{F} = P + \hat{C} = P + (C + L_F) \cap L_G^\perp = (P + C + L_F) \cap L_G^\perp.$$

Hence,

$$\hat{F} \cap \hat{K} = (P + C + L_F) \cap (K + L_F) \cap L_G^\perp \supseteq F \cap K \cap L_G^\perp$$

By (17), and an analogous argument for G,

$$\hat{F} \cap \hat{K} \supseteq F \cap K \textit{ and } \hat{G} \cap (-\hat{K}^p) \supseteq G \cap (-K^p). \tag{18}$$

This shows that if F, G meet K, $-K^p$, respectively, then $\hat{F}$, $\hat{G}$ meet $\hat{K}$, $-\hat{K}^p$.

Altogether, we have seen that if F, G, K satisfy the hypotheses of *Theorem S*, then so do $\hat{F}$, $\hat{G}$, $\hat{K}$. It remains to be shown that if $\hat{F}$, $\hat{G}$ possess orthogonal solutions $\hat{U} \epsilon \hat{K}$, $\hat{V} \epsilon -\hat{K}^p$, then orthogonal solutions $U\epsilon K$, $V\epsilon-K^p$ of F, G can be derived from them. As $\hat{U}\epsilon\hat{F}=P+C-C_F=F-C_F$, *there exist* $U_1\epsilon F$, $D_1\epsilon C_F$ such that $\hat{U}=U_1-D_1$. As $\hat{U}\epsilon\hat{K}$, there exist $U_2\epsilon K$ and $D_2\epsilon C_F$ such that $\hat{U}=U_2-D_2$. Then $U:=U_1+D_2=U_2+D_1$ belongs to $F\cap K$. Similarly, there exist $V_1\epsilon G$, $V_2\epsilon-K^p$ and E_1, $E_2\epsilon C_G$ such that $V:=V_1+E_2=V_2+E_1\epsilon G\cap(-K^p)$. U and V satisfy the orthogonality conditions in view of (4) and (5). Hence they are orthogonal solutions of F, G with respect to K. If therefore *Theorem S* holds for cones satisfying (13), then it holds in general.

Consider now a pair of complementary polar conical sets $F=P+C$ and $G=Q+C^p$ with the cones C and K satisfying (13). Moreover, we suppose that $F\cap K\neq\emptyset$ and $G\cap -K^p\neq\emptyset$. Then we approximate C and K by increasing sequences of polyhedral cones $C^{(k)}$ and $K^{(k)}$,

$$\begin{aligned} C^{(1)}\subseteq C^{(2)}\subseteq\ .\ .\ .\subseteq C&=\overline{\cup\, C^{(k)}} \\ K^{(1)}\subseteq K^{(2)}\subseteq\ .\ .\ .\ \subseteq K&=\overline{\cup\, K^{(k)}}, \end{aligned}\tag{19}$$

such that for $k=1, 2, \ldots$ and $F^{(k)}:=P+C^{(k)}$,

$$F^{(k)}\cap K^{(k)}\neq\emptyset, \tag{20}$$

and

$$L_F\subseteq C^{(k)}\subseteq L_G^{\perp},\ L_F\subseteq K^{(k)}\subseteq L_G^{\perp}. \tag{21}$$

We construct such a sequence of cones $C^{(k)}$ from a sequence $\{X^{(k)}\}_{k=1,2,\ldots}$ of points everywhere dense in C. Since $F\cap K\neq\emptyset$, we may assume that $X^{(1)}+P\epsilon F\cup K$ and put

$$C^{(k)}:=\mathscr{C}\{X^{(1)},\ldots,X^{(k)}\}+L_F$$

Both $\mathscr{C}\{X^{(1)},\ldots,X^{(k)}\}$ and L_F are polyhedral cones, and the sum of two polyhedral cones is again polyhedral. Sequence $K^{(k)}$ is formed similarly. $C^{(k)}\cap K^{(k)}\subseteq C\cap K\subseteq L_G^{\perp}$ by (4). Thus (21) is satisfied.

The sequences of the polar cones $C^{(k)p}$ and $-K^{(k)p}$ are decreasing sequences which approximate C^p and $-K^p$, respectively,

$$\begin{aligned} C^{(1)p}\supseteq C^{(2)p}\supseteq\ .\ .\ .\ \supseteq C^p&=\cap\, C^{(k)p} \\ -K^{(1)p}\supseteq -K^{(2)p}\supseteq\ .\ .\ .\ \supseteq -K^p&=\cap -K^{(k)p}, \end{aligned}\tag{22}$$

and satisfy for $k=1, 2, \ldots$ and $G^{(k)}:=Q+C^{(k)p}$,

$$G^{(k)}\cap(-K^{(k)p})\neq\emptyset \tag{23}$$

and

$$L_G\subseteq C^{(k)p}\subseteq L_F^{\perp} \text{ and } L_G\subseteq -K^{(k)p}\subseteq L_F^{\perp}. \tag{24}$$

Indeed if $X\epsilon C$, then there exists a sequence $\{X^{(k)}\}_{k=1,2,\ldots}$ such that $X^{(k)}\epsilon C^{(k)}$ and $X^{(k)}\to X$. For every $Y\epsilon\cap C^{(k)p}$ and all k we have therefore $Y^TX^{(k)}\leqslant 0$. Hence $Y^TX\leqslant 0$ for all $X\epsilon C$, which proves $Y\epsilon C^p$. The inclusion $C^p\subseteq\cap C^{(k)p}$ is trivial. The same argument gives $-K^p=\cap -K^{(k)p}$. The properties (23) and (24) follow immediately from $G^{(k)}\supseteq G$, $-K^{(k)p}\supseteq -K^p$, and $L_F^{\perp}\subseteq C^{(k)}$, $-K^{(k)}\subseteq L_G^{\perp}$ from polarizing (21).

In what follows, we will replace sequences (19), and thereby sequences (22), repeatedly by suitable subsequences. The terms "sequences (19)" and "sequences (22)" will always refer to the *current*, and not to the original specimens. Similarly, $C^{(k)}$, $K^{(k)}$ will be the kth elements of the sequences (19) as they stand *after* modifications, and $U^{(k)}\epsilon K^{(k)}$, $V^{(k)}\epsilon-K^{(k)p}$ will always be orthogonal solutions of $F^{(k)}=P+C^{(k)}$, $G^{(k)}=Q+C^{(k)p}$. Such orthogonal solutions exist by *Theorem P*.

The orthogonal solutions $U^{(k)}$, $V^{(k)}$ are in general not uniquely determined. They can be modified by adding arbitrary directions $D \epsilon L_F$, $E \epsilon L_G$ respectively. Indeed $D \epsilon L_F$, $E \epsilon L_G$ are directions of infinity of $F^{(k)}$, $G^{(k)}$ with respect to $K^{(k)}$, and we have seen that adding such multiples leads to new orthogonal solutions. Now let $D_1, \ldots, D_s$ be a basis of L_F. Then there exist multipliers α_i such that $\hat{U}^{(k)} := U^{(k)} - \Sigma \alpha_i D_i \epsilon L_F$. It follows that without restriction of generality, we may assume

$$U^{(k)} \epsilon L_F^{\perp} \text{ and } V^{(k)} \epsilon L_G^{\perp}. \tag{25}$$

Now either sequences (19) can be replaced by subsequences such that the limits

$$U = \lim U^{(k)},\ V = \lim V^{(k)}$$

exist, or $||U^{(k)}|| \to \infty$ or $||V^{(k)}|| \to \infty$.

Assume that $||U^{(k)}|| \to \infty$. If sequences (19) are replaced by a suitable subsequence, then the sequence $\{U^{(k)}/||U^{(k)}||\}_{k-1,2,\ldots}$ converges towards a direction D, $||D|| = 1$. We have

$$D = \lim \frac{U^{(k)} - P}{||U^{(k)} - P||} = \lim \frac{U^{(k)}}{||U^{(k)}||}.$$

As C and K are closed by hypothesis,

$$D \epsilon K \cap C. \tag{26}$$

Select any point $\bar{U} \epsilon F^{(1)} \cap K$. Then $\bar{U} \epsilon F^{(k)} \cap K^{(k)}$, and since $U^{(k)}$ is an optimal solution of program (8) over $F^{(k)} \cap K^{(k)}$, we have

$$U^{(k)T} Q \leq \bar{U}^T Q \text{ for all } k.$$

In other words, the sequence $\{U^{(k)T} Q\}_{k=1,2,\ldots}$ is bounded above, whence

$$D^T Q = \lim \frac{U^{(k)T} Q}{||U^{(k)}||} \leq 0. \tag{27}$$

Select any point $\bar{V} \epsilon G \cap (-K^p)$. Then $D^T \bar{V} \geq 0$ as $D \epsilon K$ and $\bar{V} \epsilon -K^p$, and $D^T(\bar{V} - Q) \leq 0$ as $D \epsilon C$ and $\bar{V} - Q \epsilon C^p$. Hence $D^T Q \geq 0$. Together with (27), this gives

$$D^T Q = 0.$$

In other words, D is a direction of infinity of F, and

$$D \epsilon L_F$$

by (13). On the other hand, $D \epsilon L_F^{\perp}$ by (25). Thus $D = 0$, which contradicts $||D|| = 1$. Therefore $||U^{(k)}|| \to \infty$ cannot hold.

Assume then $||V^{(k)}|| \to \infty$. For suitable subsequences of sequences (23),

$$E = \lim \frac{V^{(k)} - Q}{||V^{(k)} - Q||} = \lim \frac{V^{(k)}}{||V^{(k)}||} \epsilon C^p \cap (-K^p) \cap L_G$$

and $||E|| = 1$. Select any point $\bar{V} \epsilon G \cap (-K^p)$. Then $\bar{V} \epsilon G^{(k)} \cap (-K^{(k)p})$ as sequences (22) are decreasing. Since $V^{(k)}$ is an optimal solution of program (9),

$$P^T V^{(k)} \leq P^T \bar{V} \text{ for all } k,$$

and therefore

$$P^T E = \lim \frac{P^T V^{(k)}}{||V^{(k)}||} \leqslant 0.$$

It then follows as before that $E \epsilon L_G$, contradicting $||E||=1$ and $E \epsilon L_G^{\perp}$.

We may therefore assume that limits U, V exist. $U \epsilon F \cap K$ and $V \epsilon G \cap (-K^p)$, as C and K are closed. $U^{(k)T} V^{(k)} = (U^{(k)} - P)^T (V^{(k)} - Q) = 0$ for all k by orthogonality, and this carries over to U and V. The latter are therefore orthogonal solutions of F and G. This completes the proof of *Theorem S*.

5. Fenchel Duality

Fenchel [2, 3] considers the following pair of dual programs

$$\begin{aligned} &\textit{Minimize}\ f(X) - g(X) \\ &\textit{Maximize}\ g^c(Y) - f^c(Y). \end{aligned} \tag{28}$$

Here f is a convex function $f|R^n \rightarrow R \cup \{+\infty\}$ and g is a concave function $g|R^n \rightarrow R \cup \{-\infty\}$. It is convenient to consider f and g as defined everywhere on R^n and to admit infinite function values. The obvious interpretation is used, whenever infinite values occur in the convexity (concavity) conditions:

$$f(\lambda_1 X_1 + \lambda_2 X_2) \leqslant \lambda_1 f(X_1) + \lambda_2 f(X_2)$$

$$g(\lambda_1 X_1 + \lambda_2 X_2) \geqslant \lambda_1 g(X_1) + \lambda_2 g(X_2)$$

for all $\lambda_1, \lambda_2 \geqslant 0$ with $\lambda_1 + \lambda_2 = 1$. The following "domains of finiteness"

$$K(f) := \{X \epsilon R^n | f(X) < +\infty\},\ K(g) := \{X \epsilon R^n | g(X) > -\infty\}$$

are convex under these circumstances.

The function f^c is the "convex conjugate" of f, and the function g^c the "concave conjugate" of g. They are defined by

$$f^c(Y) := \sup_X\ (Y^T X - f(X)),\ g^c(Y) := \inf_X\ \{Y^T X - g(X)\}.$$

From this definition,

$$f^c(Y) \geqslant Y^T X - f(X),\ g^c(Y) \leqslant Y^T X - g(X)$$

and therefore

$$f(X) - g(X) \geqslant g^c(Y) - f^c(Y)$$

for arbitrary X, $Y \epsilon R^n$. In other words,

$$\inf_X\ (f(X) - g(X)) \geqslant \sup_Y\ (g^c(Y) - f^c(Y)).$$

The question examined by Daulity Theory is under which circumstances inf = sup, or stronger, min = max. The conditions

$$K(f) \cap K(g) \neq \emptyset \textit{ and } K(f^c) \cap K(g^c) \neq \emptyset \tag{29}$$

are obviously necessary for min = max to hold. Another, in general necessary, condition is that the functions

$$\text{f} \textit{ and } \text{g} \textit{ are "closed"}, \tag{30}$$

that is, lower semicontinuous convex and upper semicontinuous concave, respectively. In this case, $f^{cc}=f$ and $g^{cc}=g$.

Except for polyhedral functions, hypotheses (29) and (30) are not sufficient for min = max, not even for inf = sup to hold. So-called "duality gaps" may occur, as first communicated by Stoer in a letter to Karlin. The stronger hypotheses,

$$K(f)^I \cap K(g)^I \neq \emptyset \text{ and } K(f^c)^I \cap K(g^c)^I \neq \emptyset, \tag{31}$$

together with (30) ensure min = max in the general case. Rockafellar [5] and Stoer [6, 7] investigated "stability" hypotheses weaker than (31) but stronger than (29), for which min = max or inf = sup obtains.

This situation is analogous to the one examined in this paper: *Theorem P* holds for polyhedral cones but not in general. *Theorem I* holds in general but is too weak to yield *Theorem P* in the special case of polyhedral cones. *Theorem S* finally gives a general result subject to stability conditions. One is therefore lead to expect a relationship between Fenchel's Duality theorem and *Theorem S*. We proceed to show that they are in fact equivalent. More precisely, we prove:

THEOREM E: *For each triple* $F = P + C$, $G = Q + C^p$, K, *where* C, K *are arbitrary closed cones and* P, Q *arbitrary points, there exist closed functions* f *and* g, *convex and concave respectively, such that* F, G *have orthogonal solutions* $U \epsilon K$, $V \epsilon -K^p$ *if and only if min = max for the corresponding programs (28), and vice versa, for each pair of such closed functions* f, g *there exists a triple* F, G, K *such that again equivalence holds between the statement min = max and the existence of orthogonal solutions.*

PROOF: Given a triple $(F=P+C, \quad G=Q+C^p, K)$, define

$$f(X) := \begin{cases} Q^T X \text{ for } X \epsilon K \\ \infty \text{ otherwise} \end{cases}$$

$$g(X) := \begin{cases} 0 \text{ for } X \epsilon P+C \\ -\infty \text{ otherwise.} \end{cases}$$

Then

$$f^c(Y) = \sup_X \left(Y^T X - f(X)\right) = \sup_{X \epsilon K} (Y-Q)^T X = \begin{cases} 0 \text{ for } Y \epsilon Q + K^p \\ \infty \text{ otherwise.} \end{cases}$$

Putting $Z := X - P$, one has similarly

$$g^c(Y) = \inf_X \left(Y^T X - g(X)\right) = \inf_{X \epsilon P+C} Y^T X = Y^T P + \inf_{Z \epsilon C} Y^T Z,$$

and therefore

$$g^c(Y) = \begin{cases} Y^T P \text{ if } Y \epsilon -C^p \\ -\infty \text{ otherwise.} \end{cases}$$

Programs (28) then become

Minimize $Q^T X$ *with* X *subject to*

$$X \epsilon (P + C) \cap K$$

Maximize $Y^T P$ *with* Y *subject to*

$$Y \epsilon (Q + K^p) \cap (-C^p)$$

Substituting $U := X$, $V := Q - Y$, and transforming the maximization program into a minimization program yields programs (8) and (9), and it has been shown that the existence of orthogonal solutions is equivalent to min + min = 0 for these programs.

In order to prove the converse direction, we transform the minimization program (28) into

$$\textit{Minimize } z_1 - z_2 \textit{ with } z_1, z_2, X \textit{ subject to} \tag{32}$$

$$z_1 \geqslant f(X),\ z_2 \leqslant g(X).$$

Introducing the cones

$$C_f := \mathscr{C}\left[\begin{pmatrix} X \\ z \\ 1 \end{pmatrix} \middle| z \geqslant f(X)\right]$$

$$C_g := \mathscr{C}\left[\begin{pmatrix} X \\ z \\ 1 \end{pmatrix} \middle| z \leqslant g(X)\right],$$

and the linear subspace

$$L := \left\{\begin{bmatrix} X \\ z_1 \\ 0 \\ X \\ z_2 \\ 0 \end{bmatrix} \middle| z_1, z_2 \epsilon R, X \epsilon R^n\right\},$$

we formulate program

$$\textit{Minimize } z_1 - z_2 \textit{ subject to} \tag{33}$$

$$\begin{bmatrix} X_1 \\ z_1 \\ t_1 \\ X_2 \\ z_2 \\ t_2 \end{bmatrix} \epsilon (\bar{C}_f \times \bar{C}_g) \cap \left[\begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} + L\right]. \tag{34}$$

We proceed to prove, that

$$\textit{programs } (32) \textit{ and } (33) \textit{ are equivalent.} \tag{35}$$

We have to show, that if (z_1, z_2, X) is a feasible solution of (32), then $(t_1 = 1, t_2 = 1, z_1, z_2, X_1 = X, X_2 = X)$ is a feasible solution of (33), and vice versa.

For an auxiliary argument, consider any closed convex set K in R^k, and the convex cone

$$H(K) := \mathscr{C}\left\{\begin{pmatrix} W \\ 1 \end{pmatrix} \middle| W \epsilon K\right\}$$

Then $H(K)$ is not necessarily closed. However, $\overline{H(K)}$ and $H(K)$ differ only in points $\begin{pmatrix} W \\ w \end{pmatrix}$ with $w = 0$. Indeed, let

$$\begin{pmatrix} W^{(n)} \\ w^{(n)} \end{pmatrix} \rightarrow \begin{pmatrix} W \\ w \end{pmatrix}$$

be a convergent sequence of nonzero elements of $H(K)$, i.e., $W^{(n)}/w^{(n)} \epsilon K$. If $w > 0$, then

$W^{(n)}/w^{(n)} \to W/w \epsilon K$ as K is closed. But then $\begin{pmatrix} W \\ w \end{pmatrix} \epsilon H(K)$. Thus only if $w=0$ may $\begin{pmatrix} W \\ w \end{pmatrix}$ be in $\overline{H(K)}$ but not in $H(K)$.

As the functions f and g are closed, so are the convex sets

$$\left\{\begin{pmatrix} X_1 \\ z_1 \end{pmatrix} | z_1 \geqslant f(X_1)\right\}, \left\{\begin{pmatrix} X_2 \\ z_2 \end{pmatrix} | z_2 \leqslant g(X_2)\right\}.$$

The above argument can therefore be employed to show that

$$(\bar{C}_f \times \bar{C}_g) \cap T = (C_f \times C_g) \cap T,$$

where T denotes the plane characterized by $t_1 = t_2 = 1$. In the expression (34) for the feasible region of (33), one can then replace $(\bar{C}_f \times \bar{C}_g)$ by $(\bar{C}_f \times \bar{C}_g) \cap T$, $(C_f \times C_g) \cap T$ and $(C_f \times C_g)$ successively, without changing the set. It is then plain that the points in (34) correspond to the feasible solutions of (32).

Program (33) *is of type* (8).

The product $\bar{C}_f \times \bar{C}_g$ plays the role of K, and the linear subspace L plays the role of C. Furthermore

$$Q = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \\ -1 \\ 0 \end{bmatrix}, \; P = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}.$$

We have to show, that the corresponding program (9),

$$\textit{Minimize } w_1 + w_2 \textit{ subject to} \tag{36}$$

$$\begin{bmatrix} Y_1 \\ s_1 \\ w_1 \\ Y_2 \\ s_2 \\ w_2 \end{bmatrix} \epsilon - (\bar{C}_f \times \bar{C}_g)^p \cap \left[\begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \\ -1 \\ 0 \end{bmatrix} + L^{\perp}\right]$$

is indeed equivalent to the maximization program (28). To this end, we note that

$$L^{\perp} = \left\{\begin{bmatrix} -Y \\ 0 \\ w_1 \\ Y \\ 0 \\ w_2 \end{bmatrix} \middle| w_1, w_2 \epsilon R, \; Y \epsilon R^n \right\}.$$

All feasible solutions of program (36) are therefore necessarily of the form

$$\begin{bmatrix} -Y \\ 1 \\ w_1 \\ Y \\ -1 \\ w_2 \end{bmatrix}.$$

These points must moreover belong to $-(\bar{C}_f \times \bar{C}_g)^p$. As $(\bar{C}_f \times \bar{C}_g)^p = (\bar{C}_f \times \bar{C}_g)^p = (C_f \times C_g)^p$, all feasible solutions are characterized by

$$-Y^T X + z_1 + w_1 \geq 0 \text{ and } Y^T X - z_2 + w_2 \geq 0 \textit{ whenever } z_1 \geq f(X) \textit{ and } z_2 \leq g(X),$$

in addition to being of the form (37). These conditions are clearly equivalent to

$$-Y^T X + f(X) + w_1 \geq 0 \textit{ for all } X \epsilon R^n$$

and

$$Y^T X - g(X) + w_2 \geq 0 \textit{ for all } X \epsilon R^n,$$

respectively. These in turn are equivalent to

$$w_1 \geq f^c(Y) \textit{ and } w_2 \geq -g^c(Y).$$

Hence program (36) becomes

$$\textit{Minimize } f^c(Y) - g^c(Y)$$

or

$$\textit{Maximize } g^c(Y) - f^c(Y),$$

which was to be shown.—

The main part of this work was carried out at the National Bureau of Standards in Washington, D.C., where the author benefited from the interest and encouragement of A. J. Goldman. The author acknowledges the thorough and detailed examination of the manuscript by the referee of this Journal.

6. References

[1] Tucker, A. W., Abstract Structure of the Simplex Method, The RAND Symposium on Linear Programming, The RAND Corporation, Report R–351, June 21, 1959, pp. 33–34.
[2] Fenchel, W., On conjugate convex functions, Canad. J. Math. **1**, 73–77 (1949).
[3] Fenchel, W., Convex cones, sets and functions, Lecture notes by D. W. Blackett, Princeton University 1953.
[4] Kretschmer, K. S., Programs in paired spaces, Canad. J. Math. **13**, 221–238 (1961).
[5] Rockafellar, R. T., Convex functions and dual extremum problems, Diss. Harvard Univ., Cambridge, Mass. (1963).
[6] Stoer, J., Duality in convex programming, Habilitationsschrift. Technische Hochschüle München (1964).
[7] Stoer, J., Duality in nonlinear programming and the minimax theorem. Numer. Math. **5**, 371–379 (1963).

(Paper 74B2–322)