AD/A-002 923

STRONG FENCHEL DUALITY

A. Ben-Tal

Texas University

Prepared for:

Office of Naval Research

October 1974

DISTRIBUTED BY:

NTIS

National Technical Information Service
U. S. DEPARTMENT OF COMMERCE

Unclassified
Security Classification

AD/A-002923

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author) Center for Cybernetic Studies The University of Texas	2a. REPORT SECURITY CLASSIFICATION Unclassified 2b. GROUP

3. REPORT TITLE

Strong Fenchel Duality

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)

5. AUTHOR(S) (First name, middle initial, last name)

A. Ben-Tal

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
	29	8

8a. CONTRACT OR GRANT NO. N00014-67-A-0126-0008; 0009 b. PROJECT NO NR 047-021 c. d.	9a. ORIGINATOR'S REPORT NUMBER(S) Center for Cybernetic Studies Research Report CCS 200 9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)

10. DISTRIBUTION STATEMENT

This document has been approved for public release and sale; its distribution is unlimited.

11. SUPPLEMENTARY NOTES Reproduced from best available copy.	12. SPONSORING MILITARY ACTIVITY Office of Naval Research (Code 434) Washington, D.C.

13. ABSTRACT

Fenchel's Duality Theorem concerns the problem of minimizing the difference of a convex function f and a concave function g. The duality resides in the connection between the above primal problem and the dual problem of minimizing the difference of the concave conjugate g* and the convex conjugate f*. In general a duality gap may exist between the two problems unless some regularity condition is imposed. Here a family of different duals is suggested for which a duality gap does not exist.

Unclassified

Security Classification

14 KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Duality						
Convex Programming						
Duality Gaps						

DD FORM 1 NOV 65 1473 (BACK)

S/N 0102-014-6800

ia

Unclassified

Security Classification

A-31409

Research Report CCS 200

STRONG FENCHEL DUALITY

by

A. Ben-Tal

October 1974

This research was partly supported by Project No. NR 047-021, ONR Contracts N00014-67-A-0126-0008 and N00014-67-A-0126-0009 with the Center for Cybernetic Studies, The University of Texas. Reproduction in whole or in part is permitted for any purpose of the United States Government.

CENTER FOR CYBERNETIC STUDIES

A. Charnes, Director
Business-Economics Building, 512
The University of Texas
Austin, Texas 78712
(512) 471-1821

ABSTRACT

Fenchel's Duality Theorem concerns the problem of minimizing the difference of a convex function f and a concave function g. The duality resides in the connection between the above primal problem and the dual problem of minimizing the difference of the concave conjugate g* and the convex conjugate f*. In general a duality gap may exist between the two problems unless some regularity condition is imposed. Here a family of different duals is suggested for which a duality gap does not exist.

1. Introduction

Fenchel's Duality Theorem concerns the problem of minimizing $f-g$ where f and g are convex and concave functions, respectively. The duality resides in the connection between minimizing $f-g$ and maximizing g^*-f^*, where g^* and f^* are the conjugates of g and f, respectively. More precisely,

$$(+) \qquad \inf(f - g) = \max(g^* - f^*)$$

provided the following regularity condition holds:

(*) The relative interiors of domain f and domain g possess a point in common.

For treatment of Fenchel's Duality in finite dimensions, see e.g. [2],[3], [6],[7], and [8], and in infinite dimensions, e.g. [1],[4] and [5].

If f and g are restricted to certain subfamilies of convex and concave functions, then (+) holds even without (*) being valid. Such subfamilies are the polyhedral convex and concave functions [6], or, more generally, the stable functions [7, chapter 5].

In this paper we are interested in finding duals, other than the Fenchel dual: (Sup $(g^* - f^*)$) for which a relation similar to (+) holds for <u>every</u> pair of convex and concave functions, whether (*) holds or not. Such duals, called "strong Fenchels' duals", are constructed in section 3.

In section 4, the results of section 3 are applied to Rockafellar's extension of Fenchel's duality [6], and to the well-known formulas for computing the conjugate function and the subdifferential of the sum of convex functions. A special duality result for a certain strong Fenchel dual is derived in section 5.

The terminology used in this paper is that of Rockafellar's book [6]. We list below some notations used in the sequel, for definitions and further details consult [6, Part I].

Let S be a nonempty convex subset of R^n, and let f and h be convex functions: $R^n \to R$. We denote by

ri S -- the relative interior of S

rbd S -- the relative boundary of S

aff S -- the affine hull of S

dim S -- the dimension of S

$\delta(\cdot|x)$ -- the indicator function of S

$f \,\square\, h$ -- the infimal convolution of f and h, i.e

$$(f \,\square\, h)(x) \overset{\Delta}{=} \inf_y (f(y) + g(x - y))$$

2. Fenchel's Duality

Let Φ be the set of all quadruples (f, g, A_1, A_2) such that

$$(1)\quad \begin{cases} A_1 \text{ and } A_2 \text{ are convex subsets of } R^n \\ f\colon R^n \to R \text{ is a proper convex function with } \operatorname{dom} f = A_1 \\ g\colon R^n \to R \text{ is a proper concave function with } \operatorname{dom} g = A_2 \\ A \overset{\Delta}{=} A_1 \cap A_2 \neq \emptyset \end{cases}$$

Consider the primal problem

$$(\mathcal{P}) \qquad \inf_{x \in A} (f - g)$$

Let f^* denote the (convex) conjugate of f, and g^* the (concave) conjugate of g, i.e.

$$f^*(x^*) \triangleq \sup_{x \in A_1} (\langle x^*, x \rangle - f(x))$$

$$g^*(x^*) \triangleq \inf_{x \in A_2} (\langle x^*, x \rangle - g(x))$$

Denote also $A_1^* \triangleq \text{dom } f^*$, $A_2^* \triangleq \text{dom } g^*$ and finally $A^* \triangleq A_1^* \cap A_2^*$.

The problem

$$(\mathcal{D}) \qquad \sup_{x^* \in A^*} (g^* - f^*)$$

is called the Fenchel Dual of $(\mathcal{P})$.

The following classical result relates $(\mathcal{P})$ and $(\mathcal{D})$.

Fenchel's Duality Theorem (e.g. [6, Theorem 31.1])

Let $(f, g, A_1, A_2) \in \Phi$. If

(2) $$\text{ri } A_1 \cap \text{ri } A_2 \neq \emptyset$$

then

(3) $$\inf_{A} (f - g) = \max_{A^*} (g^* - f^*).$$

If f and g are closed and

(4) $$\text{ri } A_1^* \cap \text{ri } A_2^* \neq \emptyset$$

then

(5) $$\min_{A} (f - g) = \max_{A^*} (g^* - f^*).$$

□

There are well-known examples, where neither (2) nor (4) holds, and in which $\inf\ (P) > \sup\ (D)$, i.e. there is a <u>duality gap</u>.

One such example is the following.

<u>Example 1</u> [7, p. 181-183]

Let

$$A_1 = \{(x, y) \in R^2 : x = 0,\ y \geq 0\}$$

$$A_2 = \{(x, y) \in R^2 : x \geq 0,\ y \geq 0\}$$

$$f(x,y) = \begin{cases} 0 & (x,y) \in A_1 \\ \infty & \text{otherwise} \end{cases}$$

$$g(x,y) = \begin{cases} 1 & (x,y) \in A_2 \text{ and } xy \geq 1 \\ \sqrt{xy} & (x,y) \in A_2 \text{ and } xy < 1 \\ -\infty & \text{otherwise} \end{cases}$$

Then

$$f^*(x^*,y^*) = \begin{cases} 0 & y^* \leq 0 \\ \infty & \text{otherwise} \end{cases}$$

$$g^*(x^*,y^*) = \begin{cases} -1 & x^* \geq 0,\ y^* = 0 \\ -\infty & \text{otherwise} \end{cases}$$

Therefore

$$\inf(f - g) = 0 > -1 = \sup(g^* - f^*).$$

<u>3. Strong Fenchel Duality</u>

For any subsets B_1, B_2, of R^n such that

(6) $$B_i \subset A_i, \qquad i = 1, 2$$

Let us denote the following:

$$f^*_{B_1}(x^*) \stackrel{\Delta}{=} \sup_{x\in B_1}(\langle x^*, x\rangle - f(x))$$

$$g^*_{B_1}(x^*) \stackrel{\Delta}{=} \inf_{x\in B_2}(\langle x^*, x\rangle - g(x))$$

$$B^*_1 \stackrel{\Delta}{=} \mathrm{dom}\ f^*_{B_1}, \quad B^*_2 \stackrel{\Delta}{=} \mathrm{dom}\ g^*_{B_2}, \quad B^* \stackrel{\Delta}{=} B^*_1 \cap B^*_2.$$

Also let $(\mathcal{D}; B_1, B_2)$ denote the following problem

$$(\mathcal{D}; B_1, B_2) \qquad \sup_{B^*}(g^*_{B_2} - f^*_{B_1}).$$

Whenever $B_1 = A_1$, and $B_2 = A_2$ they are omitted from the above notation, thus $f^*_{A_1} = f^*$, $g^*_{A_2} = g^*$, $(\mathcal{D}; A_1, A_2) = (\mathcal{D})$. A pair of convex subsets (B_1, B_2) is called <u>admissible</u> if it satisfies (6) and

(7) $$B_1 \cap B_2 = A$$

An admissible pair is called <u>strongly admissible</u> if in addition to (6) and (7) it satisfies

(8) $$\mathrm{ri}\ B_1 \cap \mathrm{ri}\ B_2 \neq \emptyset.$$

The following result is an elementary observation suggesting the possibility of constructing duals $(\mathcal{D}; B_1, B_2)$ without duality gaps.

<u>Proposition 1</u>

Let $(f, g, A_1, A_2) \in \Phi$ and let (B_1, B_2) be an admissible pair. Then

(9) $$\inf_A(f - g) \geq \sup_{B^*}(g^*_{B_2} - f^*_{B_1}) \geq \sup_{A^*}(g^* - f^*)$$

<u>Proof.</u>

From the definitions of $f^*_{B_1}$ and $g^*_{B_2}$ we derive

$$f^*_{B_1}(x^*) \geq \langle x^*, x\rangle - f(x) \qquad x \in B_1,\ x^* \in B^*_1$$

$$g^*_{B_2}(x^*) \leq \langle x^*, x\rangle - g(x) \qquad x \in B_2,\ x^*_2 \in B^*$$

Hence for every $x \in B_1 \cap B_2 = A$ and $x^* \in B^*$,

$$g^*_{B_2}(x^*) + g(x) \leq \langle x^*, x \rangle \leq f^*_{B_1}(x^*) + f(x)$$

implying

$$f(x) - g(x) \geq g^*_{B_2}(x^*) - f^*_{B_1}(x^*) \qquad x \in A,\quad x^* \in B^*$$

proving that first inequality in (9). To prove the second inequality note that

$$B_1 \subset A \Rightarrow \begin{cases} f^*_{B_1} \leq f^* \\ B^*_1 \supset A^*_1 \end{cases} \qquad B_2 \subset A_2 \Rightarrow \begin{cases} g^*_{B_1} \geq g^* \\ B^*_2 \supset B^*_1 \end{cases}$$

showing that

$$g^*_{B_2} - f^*_{B_1} \geq g^* - f^*$$

and

$$B^* \supset A^*$$

from which it follows that

$$\sup_{B^*}(g^*_{B_2} - f^*_{B_1}) \geq \sup_{A^*}(g^* - f^*) \qquad \square$$

A dual problem $(\mathcal{D}; B_1, B_2)$ is called a <u>strong Fenchel dual</u> if

$$\text{(10)} \qquad \inf(\mathcal{P}) = \max(\mathcal{D}; B_1, B_2)$$

for <u>every</u> $(f, g, A_1, A_2) \in \Phi$. This property is closely related to the strong admissibility of (B_1, B_2), as expressed in

<u>Proposition 2</u>

For every strongly admissible pair (B_1, B_2), the problem $(\mathcal{D}; B_1, B_2)$ is a strong Fenchel dual.

<u>Proof</u>

The admissibility of (B_1, B_2) implies

$$\text{(11)} \qquad \inf_{A}(f - g) = \inf_{B_1 \cap B_2}(\hat{f} - \hat{g})$$

where

$$\hat{f} \triangleq f + \delta(\,|B_1)$$

$$\hat{g} \triangleq g - \delta(\,|B_2)$$

clearly

$$(\hat{f}, \hat{g}, B_1, B_2) \in \Phi$$

moreover, by the strong admissibility of (B_1, B_2) it follows from **Fenchel's Duality Theorem** that

$$\text{(12)} \qquad \inf_{B_1 \cap B_2}(\hat{f} - \hat{g}) = \max(\hat{g}^* - \hat{f}^*)$$

but

$$\hat{f}^* = f^*_{B_1}, \quad \hat{g}^* = g^*_{B_2}$$

hence (11) and (12) implies

$$\inf_{A}(f - g) = \max_{B^*}(g^*_{B_2} - f^*_{B_1})$$

$\square$

The existence of a strongly admissible pair, i.e. the nonemptiness of

(13) $$S \stackrel{\Delta}{=} \{\text{all strongly admissible pairs}\}$$

is illustrated by the following simple example.

Example 2

Consider the pair

$$B_1 = A, \quad B_2 = A$$

then

$$B_1 = A \subset A_1, \; B_2 = A \subset A_2, \; B_1 \cap B_2 = A \cap A = A$$

$$\text{ri } B_1 \cap \text{ri } B_2 = \text{ri } A \cap \text{ri } A = \text{ri } A \neq \emptyset$$

hence (A, A) is strongly admissible. The fact ri $A \neq \emptyset$ indeed holds (in finite dimension spaces) for any nonempty convex set A. Note that Example 2 together with Proposition 2, produce our first strong duality relation

$$\inf_{A}(f - g) = \max(g^*_A - f^*_A).$$

4. Characterization of strong admissibility

Lemma 1

For any nonempty convex sets $S, T \subset R^n$

(14) $$T \cap \operatorname{ri} S = \emptyset$$

if, and only if

(15) $$S \cap T \subset \operatorname{rbd} S$$

consequently

(16) $$\operatorname{ri} S \cap \operatorname{ri} T = \emptyset$$

if, and only if

(17) $$[S \cap T \subset \operatorname{rbd} S] \vee [S \cap T \subset \operatorname{rbd} T]$$

<u>Proof</u>

First note that the equivalence (16)$\Leftrightarrow$(17) follows from the equivalence (14)$\Leftrightarrow$ (15) since

$$[\operatorname{ri} S \cap \operatorname{ri} T = \emptyset] \Leftrightarrow [T \cap \operatorname{ri} S = \emptyset] \vee [S \cap \operatorname{ri} T = \emptyset]$$

Indeed the implication ($\Leftarrow$) is trivial, and the implication ($\Rightarrow$) follows from the fact that the condition $\operatorname{ri} S \cap \operatorname{ri} T = \emptyset$, is necessary and sufficient for proper separation of S and T (see [6, Theorem 11.3]). Now, if $S \cap T = \emptyset$, the equivalence (14) $\Leftrightarrow$(15) is trivial.

Thus suppose that

(18) $$S \cap T \neq \emptyset$$

Let (14) hold. Then

(19) $$(S \cap T) \cap \operatorname{ri} S = (S \cap \operatorname{ri} S) \cap T = T \cap \operatorname{ri} S = \emptyset$$

Since $(S \cap T) \subset S$ it follows from (19) that

$$S \cap T \subset S - \text{ri } S \subset \text{cl } S - \text{ri } S = \text{rbd} S.$$

Suppose now that (15) holds. Clearly rbd$S \cap$ ri $S = \emptyset$ hence, by (15), $(S \cap T) \cap$ ri $S = \emptyset$ and, by (19), $T \cap$ ri $S = \emptyset$.

□

Corollary 1.1

The set $\mathcal{S}$ of all strongly admissible pairs is given by

(20) $\mathcal{S} = \{\text{convex pairs } (B_1, B_2): A \subset B_i \subset A_i,\ A \not\subset \text{rbd } B_i,\ i = 1,2\}$

Proof from Lemma 1

$$\text{ri } B_1 \cap \text{ri } B_2 \neq \emptyset \Longleftrightarrow B_1 \cap B_2 \not\subset \text{rbd } B_i \qquad i = 1, 2$$

Now the fact $[A \subset B_i \subset A_i,\ i = 1, 2]$ is equivalent to $[B_1 \cap B_2 = A,\ B_i \subset A_i \quad i = 1, 2]$ and hence the result.

□

The following lemma will enable us to find an important subset of $\mathcal{S}$.

Lemma 2

For any non-empty convex sets $S, T \subset R^n$ and any convex subsets P, Q such that

(21) $$S \cap T \subset P \subset S \cap \text{aff}(S \cap T)$$

(22) $$S \cap T \subset Q \subset T \cap \text{aff}(S \cap T)$$

it follows that

(23) $\qquad \text{ri } P \cap \text{ri } Q \neq \emptyset.$

<u>Proof</u>

If (23) is false then

$$[Q \cap \text{ri } P = \emptyset] \vee [P \cap \text{ri } Q = \emptyset].$$

Thus, without loss of generality, suppose that

(24) $\qquad Q \cap \text{ri } P = \emptyset$

This is equivalant, by Lemma 1 to

$$P \cap Q \subset \text{rbd } P.$$

Since this means that $P \cap Q$ is a convex subset of the relative boundary of the convex set P it follows [Corollary 6.3.3] that

(25) $\qquad \dim(P \cap Q) < \dim P.$

On the other hand (21) and (22) imply

$$S \cap T \subset P \cap Q \subset (S \cap T) \cap \text{aff}(S \cap T)$$

Hence

(26) $\qquad S \cap T = P \cap Q$

Moreover

$$\dim P \leq \dim [S \cap \text{aff}(S \cap T)] \leq$$
$$\leq \dim \text{aff}(S \cap T) = \dim(S \cap T).$$

Hence by (26)

$$\dim(P) \leq \dim(P \cap Q)$$

ntradicting (25). Thus (24) is false. Similarly

$$P \cap \text{ri } Q \neq \emptyset$$

Proving (23). □

Restating Lemma 2, we obtain

Corollary 2.

The set

$$(27) \qquad \mathcal{A} \triangleq \{\text{convex pairs } (B_1,B_2):\ A \subset B_i \subset A_i \cap \text{aff } A,\ i=1,2\}$$

consists of strongly admissible pairs, i.e. $\mathcal{A} \subset \mathcal{S}$.

Remarks

1. A necessary condition for (B_1, B_2) to be strongly admissible is that

$$\text{ri } A \subset \text{ri } B_i, \quad i = 1, 2$$

This follows from the implication (see [6, Corollary 6.5.2]):

$$\left.\begin{array}{l} A \subset B_i \\ A \not\subset \text{rbd } B_i \end{array}\right\} \Rightarrow \text{ri } A \subset \text{ri } B_i$$

2. There are pairs (A_1, A_2) for which $\mathcal{A} = \mathcal{S}$, such as the pair (A_1, A_2) given in Example 1.

There are of course pairs (A_1, A_2) for which $\mathcal{A} \neq \mathcal{S}$. Consider the example A_1 = a circle in the plane, A_2 = a radius of the circle. Then $(B_1, B_2) = (A_1, A_2) \in \mathcal{S}$ but $(B_1, B_2) \notin \mathcal{A}$.

3. If

$$A_2 \cap \text{ri } A_1 \neq \emptyset$$

then the set

$$\mathcal{A}_1 \triangleq \{(B_1, B_2): A \subset B_1 \subset A_1, A \subset B_2 \subset A_2 \cap \text{aff } A\}$$

is contained in $\mathcal{S}$ (and, clearly, contains $\mathcal{A}$). This fact follows actually from the proof of Lemma 2.

4. Some related results.

Rockafellar's extension of Fenchel duality

Suppose that

$$(28)\quad \begin{cases} C_1 \text{ is a nonempty convex subset of } R^n \\ C_2 \text{ is a nonempty convex subset of } R^m \\ f: R^n \to R \text{ is a proper convex function, dom } f = C_1 \\ g: R^m \to R \text{ is a proper concave function, dom } g = C_2 \\ M: R^n \to R^m \text{ is a linear transformation with inverse } M^{-1} \end{cases}$$

where, for any $S \subset R^m$

$$M^{-1}S \triangleq \{x: Mx \in S\} .$$

Suppose further that

$$(29)\qquad C \triangleq C_1 \cap M^{-1}C_2 \neq \emptyset$$

and

$$(30)\qquad C_2 \subset \text{Range } M$$

Let the set of all (f, g, M, C_1, C_2) satisfying (28) - (30) be denoted by Ψ.

Consider the problem

$$\text{(31)} \qquad \inf(f - gM)$$

The infimum is taken effectively on the (nonempty, convex) set C. Note that there is no loss of generality in assuming (30), for if (f, g, M, A_1, A_2) satisfy (28), (29) but not (30), one can consider instead of (31) the equivalent problem

$$\text{(32)} \qquad \inf\{(f - \bar{g}M): \ x \in C_1 \cap M^{-1}\bar{C}_2\}$$

where

$$\bar{g} \triangleq g - \delta(\cdot \mid \text{range } M)$$

and

$$\bar{C}_2 \triangleq \text{dom } \hat{g}$$

Clearly then

$$(f, \bar{g}, M, C_1, \bar{C}_2) \in \Psi.$$

Rockafellar (see e.g. [6. Corollary 31.2.1] and [7]) proves that if $(f, g, M, C_1, C_2) \in \Psi$ and

$$\text{(33)} \qquad \text{ri } C_1 \cap M^{-1}(\text{ri } C_2) \neq \emptyset$$

Then

$$\text{(34)} \qquad \inf(f - gM) = \text{Max}(g^* - f^*M^*)$$

where M^* is the adjoint of M.

The results of the previous section can be used here to derive the following.

<u>Theorem 1</u>

Let $(f, g, M, C_1, C_2) \in \Psi$, and let $D_1 \subset R^n$, $D_2 \subset R^m$ be any convex subsets satisfying

(35) $\quad C \subset D_1 \subset C_1, \quad C \subset M^{-1}D_2 \subset M^{-1}C_2$

(36) $\quad C \not\subset \text{rbd } D_1$

(37) $\quad C \not\subset M^{-1}(\text{rbd } D_2)$

Then

(38) $\quad \inf(f - gM) = \max(g^*_{D_2} - f^*_{D_1} M^*)$.

In particular, (38) holds if

(39) $$C \subset D_1 \subset \text{aff}(C) \cap C_1$$
$$C \subset M^{-1}D_2 \subset \text{aff}(C) \cap M^{-1}C_2$$

<u>Proof</u>

First we collect some properties of M^{-1} needed below,

(40) $\quad S \subset T \Rightarrow M^{-1}S \subset M^{-1}T$

(41) $\quad M^{-1}(S \cap T) = M^{-1}S \cap M^{-1}T$

(42) $\quad M^{-1}(S \sim T) = M^{-1}S \sim M^{-1}T$

(43) $\quad \text{Range } M \supset S,\ S \neq \emptyset \Rightarrow M^{-1}S \neq \emptyset,$

Finally (see e.g. [6, Theorem 6.7])

(44) $$M^{-1}(\text{ri } S) \neq \emptyset \Rightarrow \begin{cases} \text{ri}(M^{-1}S) = M^{-1}(\text{ri } S), \\ \text{cl}(M^{-1}S) = M^{-1}(\text{cl } S). \end{cases}$$

Now, by (42), (37) is equivalent to

(45) $$C \not\subset M^{-1}(\text{cl } D_2) \sim M^{-1}(\text{ri } D_2),$$

Also

$$\emptyset \neq \text{ri } D_2 \subset D_2 \subset C_2 \subset \text{Range } M, \text{ by (35) and (30)}$$

Hence, by (43)

(46) $$M^{-1}(\text{ri } D_2) \neq \emptyset$$

and thus, by (44) - (46)

(47) $$C \not\subset \text{cl}(M^{-1}D_2) \sim \text{ri}(M^{-1}D_2) = \text{rbd}(M^{-1}D_2).$$

It follows that $S = D_1$ and $T = M^{-1}D_2$ are two subsets of R^n satisfying $S \cap T = C$ (by(35))and $S \cap T \not\subset \text{rbd } S$, $S \cap T \not\subset \text{rbd } T$ and hence, by Lemma 1, $\text{ri } S \cap \text{ri } T \neq \emptyset$, or in view of (46) and (44):

(48) $$\text{ri } D_1 \cap M^{-1} \text{ ri } D_2 \neq \emptyset.$$

From (35) it follows that

$$\inf(f - gM) = \inf(\hat{f} - \hat{g}M)$$

where

$$\hat{f} \triangleq f + \delta(\,|D_1)$$
$$\hat{g} \triangleq g - \delta(\,|D_2).$$

Clearly

$$(\hat{f}, \hat{g}, M, D_1, D_2) \in \Psi$$

and hence (38) follows from the validity of the regularity conditions (48). Finally, (39) implies, by lemma 2, that $\mathrm{ri}\, D \cap \mathrm{ri}\, M^{-1}D_1 \neq \emptyset$ which, again, by (46) and (44), implies (48), proving the last assertion of the theorem. □

<u>The conjugate of the sum of convex functions</u>

<u>Theorem 2</u>

Let $(f, -h, A_1, A_2) \in \Phi$. Then, the infimum in $f^*_{B_1} \,\square\, h^*_{B_2}$ is attained and

(49) $$(f + h)^* = f^*_{B_1} \,\square\, h^*_{B_2}$$

for every $(B_1, B_2) \in S$ (see (20)) and, in particular, for every $(B_1, B_2) \in A$ (see (27))

<u>Proof:</u>

$$(f + h)^*(y^*) = \sup(\langle y^*, x\rangle - [f(x) + h(x)])$$
$$= -\inf(f(x) - [\langle y^*, x\rangle - h(x)])$$
$$= -\inf(f(x) - g(x))$$

where

$$g(x) \triangleq \langle y^*, x\rangle - h(x)$$

Now, $(f, g, A_1, A_2) \in \Phi$, and (B_1, B_2) are strongly admissible (Corollaries 1.1 and 2.1), hence by proposition 2

$$-\inf(f - g) = -\max(g^*_{B_2} - f^*_{B_1})$$

A simple calculation shows that

$$g^*_{B_2}(x^*) = -h^*_{B_2}(y^* - x^*)$$

so

$$(f + h)^* (y^*) = -\inf(f - g) = -\max(g^*_{B_2} - f^*_{B_1}) =$$
$$= -\max(-h^*_{B}(y^* - x^*) - f^*_{B_1}(x^*)) =$$
$$= \min(f^*_{B_1}(x^*) + h^*_{B_2}(y^* - x^*)) = (f^*_{B_1} \square h^*_{B_2})(y^*)$$

$\square$

Theorem 2 generalized [6, Theorem 16.4]. The "infimal convulution formula" (19) was first obtained by Fenchel [2]. See also [7].

The subdifferential of the sum of convex functions

Let f be a convex function, and S a subset of dom f. Consider for $x \in S$, the set $\partial_S f(x)$ off all $x^* \in R^n$ such that

$$f(z) > f(x) + \langle x^*, z - x\rangle, \quad \forall z \in S$$

We write $\partial f(x)$ for $\partial_{\text{dom } f} f(x)$, thus actually

$$\partial_S f(x) = \partial(f(x) + \delta(x|S)).$$

Theorem 3

Let $(f, -h, \Lambda_1, \Lambda_2) \in \Phi$. Then

(50) $$\partial(f + h) = \partial_{B_1} f(x) + \partial_{B_2} h(x)$$

for every $(B_1, B_2) \in S$ and, in particular, for every $(B_1, B_2) \in \Lambda$.

Proof

$$\partial(f + h) = \partial_{A_1 \cap A_2}(f + h) = \partial_{B_1 \cap B_2}(\hat{f} + \hat{h})$$

where

$$\hat{f} = f + \delta(\cdot|B_1)$$

$$\hat{h} = h + \delta(\cdot|B_2)$$

the last equality is justified by the fact that (B_1, B_2) are admissible. Now, since $(\hat{f}, \hat{h}, B_1, B_2) \in \Phi$ and (B_1, B_2) are strongly admissible, it follows that (see [6, Theorem 23.8]

$$\partial_{B_1 \cap B_2}(\hat{f} + \hat{h}) = \partial\hat{f} + \partial\hat{h}$$

but $\partial\hat{f} = \partial_{B_1} f$ $\quad \partial\hat{h} = \partial_{B_2} h$ and hence (50) follows. $\square$

5. A special result for the strong Fenchels' dual $(\mathcal{D}; A, A)$

It was shown in Example 2 that (A, A) is a strongly admissible pair, and hence

(51) $$\inf_A(f - g) = \max(g^*_A - f^*_A).$$

The following theorem adds to the validity of (50) an explicit connection between the optimal solutions of $(\mathcal{P})$ and $(\mathcal{D}; A, A)$. The proof does not rely on Fenchel's Duality Theorem, or its traditional proofs(e.g. [6], [7], and [4]) and in fact does not utilize separation arguments. This is significant in deriving generalizations of (51) for nonconvex functions.

Theorem 4

Let $(f, g, A_1, A_2) \varepsilon \Phi$ and suppose further that $f, g \varepsilon C^1$. Let $\bar{x} \varepsilon A$ be an optimal solution of (P). Then any x^* belonging to the interval

$$[\nabla f(\bar{x}), \nabla g \bar{x}]$$

is an optimal solution of $(D;\, A, A)$ and (51) is valid.

Proof

Since f is convex on A_1, it satisfies the gradient inequality

$$f(x) \geq f(\bar{x}) + \langle x - \bar{x}, \nabla f(\bar{x})\rangle \qquad x \varepsilon A_1$$

and hence, in particular

(52) $$f^*_A(\nabla f(\bar{x})) = \langle \nabla f(\bar{x}), \bar{x}\rangle - f(\bar{x}) \geq \langle \nabla f(\bar{x}), x\rangle - f(x), \quad x \varepsilon A$$

and

(53) $$\langle \nabla f(\bar{x}), x - \bar{x}\rangle \leq f(x) - f(\bar{x}), \qquad x \varepsilon A.$$

A necessary condition for $\bar{x}$ to solve (P) is (see e.g. [4, Theorem 2, p. 175])

$$\langle \nabla f(\bar{x}) - \nabla g(\bar{x}), x - \bar{x}\rangle \geq 0 \qquad x \varepsilon A$$

or rearranging terms

(54) $$\langle \nabla g(\bar{x}), x - \bar{x}\rangle \leq \langle \nabla f(\bar{x}), x - \bar{x}\rangle \qquad x \varepsilon A$$

(53) with (54) imply

$$\langle \nabla g(\bar{x}), x - \bar{x}\rangle \leq f(x) - f(\bar{x})$$

or

$$\langle \nabla g(\bar{x}), \bar{x}\rangle - f(\bar{x}) \geq \langle \nabla g(\bar{x}), x\rangle - f(x) \qquad x \in A$$

i.e.

$$(55) \qquad f^*_A(\nabla g(\bar{x})) = \langle \nabla g(\bar{x}), \bar{x}\rangle - f(\bar{x}).$$

But, similar to (52),

$$(56) \qquad g^*_A(\nabla g(\bar{x})) = \langle \nabla g(\bar{x}), \bar{x}\rangle - g(\bar{x}).$$

Now, (55) and (56) show that

$$(57) \qquad f(\bar{x}) - g(\bar{x}) = g^*_A(\nabla g(\bar{x})) - f^*_A(\nabla g(\bar{x})).$$

Since (see Proposition 1)

$$f(\bar{x}) - g(\bar{x}) > g^*_A(x^*) - f^*_A(x^*) \qquad \text{for every } x^*$$

it follows from (56) that $x^* = \nabla g(\bar{x})$ is an optimal solution of $(\mathcal{D}, A, A)$, and that (51) is valid.

Similar to (57), it can be shown that

$$f(\bar{x}) - g(\bar{x}) = g^*_A(\nabla f(\bar{x})) - f^*_A(\nabla f(\bar{x}))$$

which proves that $x^* = \nabla f(\bar{x})$ is also an optimal solution of $(\mathcal{D}, A, A)$.

Finally, $(\mathcal{D}, A, A)$ being a concave program implies that its solution set is convex, and hence every $x^* \in [\nabla f(\bar{x}), \nabla g(\bar{x})]$ is an optimal solution. $\square$

Corollary 4.1

Dual program ($\mathcal{D}$; A, A) has a unique optimal solution only if primal problem ($\mathcal{P}$) has an optimal solution which is a critical point of its objective function.

Proof

Let $\bar{x}$ be an optimal solution of ($\mathcal{P}$). If ($\mathcal{D}$; A, A) has a unique maximizer, it follows from Theorem 4 that $\nabla f(\bar{x}) = \nabla g(\bar{x})$ i.e. $\nabla f(\bar{x}) - \nabla g(\bar{x}) = 0$, hence $\bar{x}$ is a critical point.

□

Theorem 1 is illustrated in the following

Example 3

Let f and g be, respectively, a strictly convex and a strictly concave functions: $R \to R$ such that $f - g$ is strictly monotone. Let $\text{dom } f \cap \text{dom } g = [a, b]$ $(a < b)$. Clearly then

$$\min_{[a,b]}(f - g) = f(a) - g(a)$$

and, furthermore, f' is strictly increasing, g' is strictly decreasing and $f' > g'$. Hence

(58) $$g'(b) < g'(a) < f'(a) < f'(b).$$

Let L_h denote the Legandre transform of h, i.e.

$$L_h(x^*) \triangleq \langle x^*, h'^{-1}(x^*)\rangle - h(h'^{-1}(x^*))$$

By the calculus then

$$
(59)\quad f^*_A(x^*) = \begin{cases} ax^* - f(a) & -\infty < x^* \le f'(a) \\ L_f(x^*) & f'(a) \le x^* < f'(b) \\ bx^* - f(b) & f'(b) \le x^* < \infty \end{cases}
$$

$$
(60)\quad g^*_A(x^*) = \begin{cases} bx^* - g(b) & -\infty < x^* \le g'(b) \\ L_g(x^*) & g'(b) \le x^* < g'(a) \\ ax^* - g(a) & g'(a) \le x^* < \infty \end{cases}
$$

Combining the information in (58) - (60) we derive the graphical representation of the dual objective function $g^*_A - f^*_A$ (see figure 1) from which the conclusions of Theorem 1 are evident.

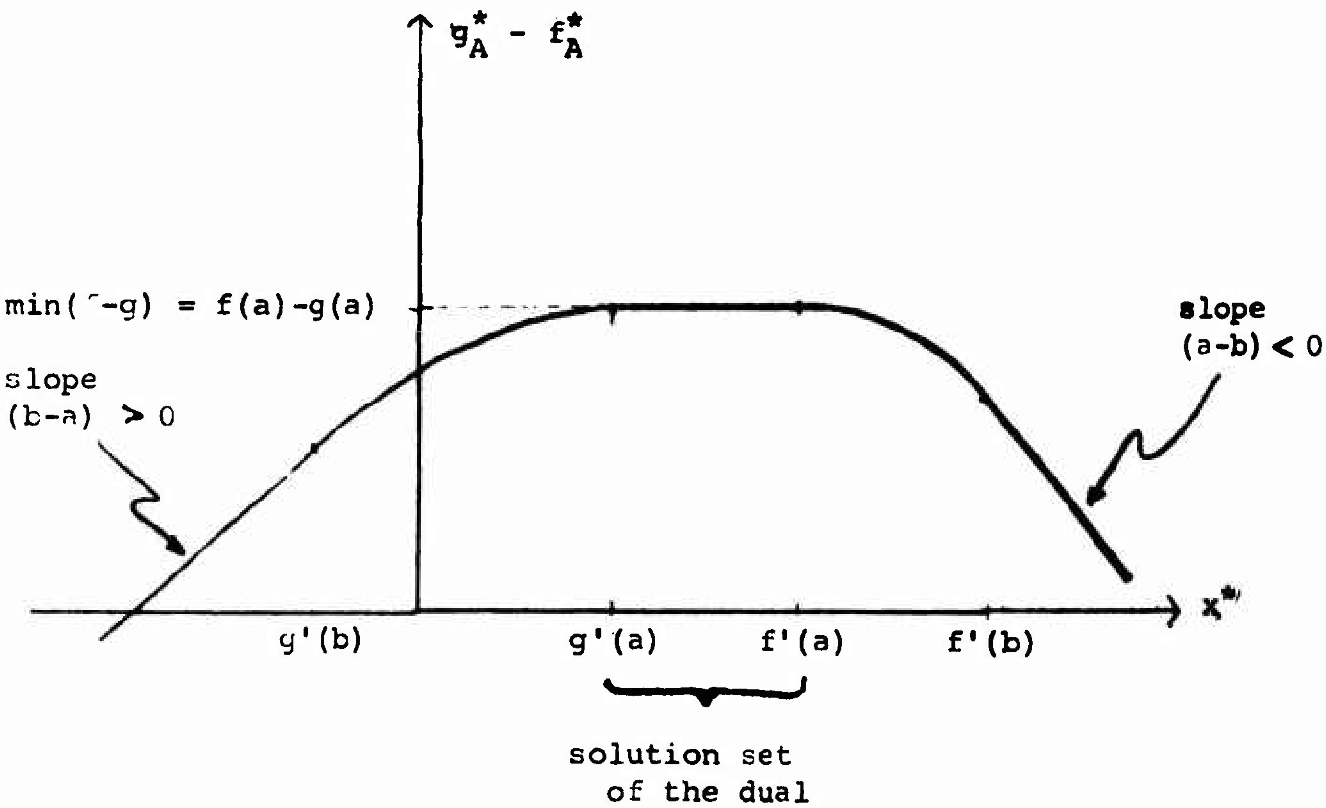

FIGURE 1

REFERENCES

(1) Dieter, V., "Dual Extremal Problems in locally convex linear spaces," Proceedings of the Colloquium on Convexity, Copenhagen, 1965, W. Fenchel, Editor.

(2) Fenchel, W., "Convex Cones, Sets, and Functions," Lecture Notes, Department of Mathematics, Princeton University, Princeton New Jersey, 1953.

(3) Karlin, S., Mathematical Methods and Theory in Games, Programming, and Economics, Vol. I, Addison-Wesley, Reading, Mass. 1959.

(4) Luenberger, D. G., Optimization By Vector Space Methods, John Wiley, New York, 1969.

(5) Moreau, J. J., "Convexity and Duality," Functional Analysis and Optimization, E. R. Caianello, Ed., Academic Press, New York, 1966, 145-169.

(6) Rockafellar, R. T., Convex Analysis, Princeton University Press, Princeton, New Jersey, 1969.

(7) Stoer, J. and Witzgall, C., Convexity and Optimization in Finite Dimensions I, Springer-Verlag, New York, Heidelberg, Berlin, 1970.

(8) Whinston, A., "Some Applications of the Conjugate Function Theory to Duality," Nonlinear Programming, J. Abadie (ed.) North Holland Publishing Company, Amsterdam, John Wiley & Sons, Inc., New York, 1967.