AD-A079 707 WISCONSIN UNIV-MADISON MATHEMATICS RESEARCH CENTER F/G 12/1

ABSTRACT OPTIMALITY THEORY.(U)

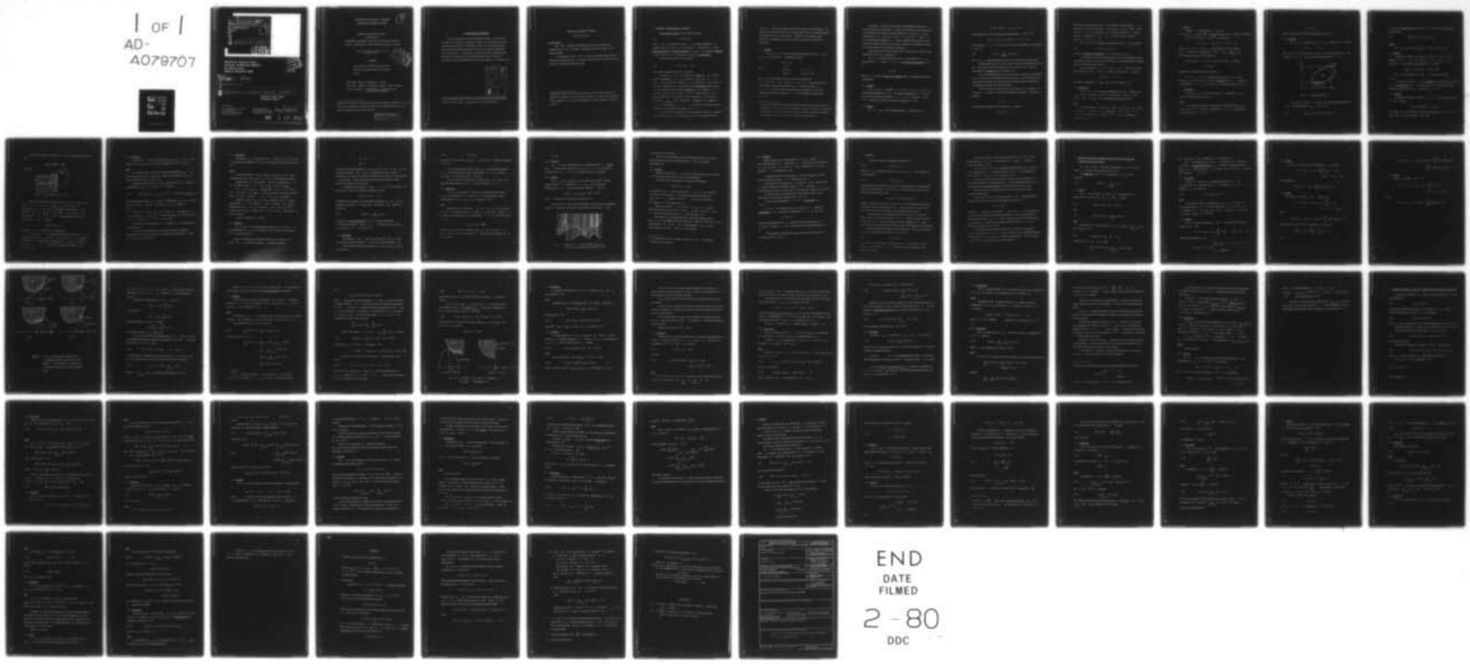
OCT 79 S DOLECKI DAAG29-75-C-0024

UNCLASSIFIED MRC-TSR-2006 NL

1 OF 1

AD-

A079707



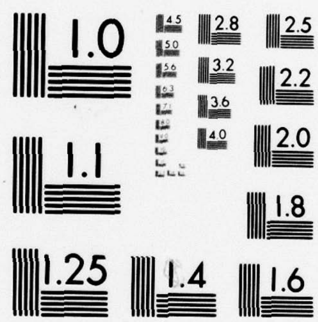
END

DATE

FILMED

2-80

DDC


1.0
4.5
5.0
5.6
6.3
2.8
3.2
3.6
4.0
2.5
2.2
2.0
1.1
1.8
1.25
1.4
1.6

9 MRC Technical Summary Report #2006

6 **ABSTRACT OPTIMALITY THEORY.**

Szymon Dolecki

10

3 B.S.

14 MRC-TSR-2006

LEVEL II

Mathematics Research Center
University of Wisconsin—Madison
610 Walnut Street
Madison, Wisconsin 53706

11 Oct[redacted]79

12 63

Received July 26, 1979

15 DAAG29-75-C-0024, AFOSR-79-0018

Sponsored by

U. S. Army Research Office
P.O. Box 12211
Research Triangle Park
North Carolina 27709

Air Force Office
of Scientific Research
Washington, D.C. 20332

221 200

Office of Naval Research
Arlington, Virginia 22217

UNIVERSITY OF WISCONSIN - MADISON

MATHEMATICS RESEARCH CENTER

3

ABSTRACT OPTIMALITY THEORY

Szymon Dolecki

Mathematics Department, University of Wisconsin, Madison

Institute of Mathematics, Polish Academy of Sciences, Warsaw

Technical Summary Report #2006
October 1979

ABSTRACT

We present the underlying ideas and basic notions of optimality theory in an abstract setting.

AMS (MOS) Subject Classification: 49 B27

Key words: Extremal, Lagrangian, Duality, Penalty functions.

Work Unit Number 5 (Mathematical Programming and Operations Research)

Sponsored in part by the United States Army under Contract No. DAAG29-75-C-0024, by the Air Force Office of Scientific Research Grant AFOSR-79-0018 and by the Office of Naval Research Contract 041-404.

SIGNIFICANCE AND EXPLANATION

This work presents an abstract optimality theory. The discussed concepts (e.g. extremal, Lagrangian, duality) are more general and much simpler than their familiar counterparts used in more concrete framework. They admit geometrical interpretations and help to clarify the underlying ideas of the theory. Consequently, the presented theory suggests further theoretical developments and possible applications (e.g. the use of Lagrangians of the form different from the usually used (3.7); Example 3.4.).

The responsibility for the wording and views expressed in this descriptive summary lies with MRC, and not with the author of this report.

ABSTRACT OPTIMALITY THEORY

Szymon Dolecki

Introduction

This work presents the underlying ideas and basic notions of optimality theory (e.g., extremal, Lagrangian, penalty, duality) in an abstract setting.

The mathematics of the work is simple, abounding with geometrical interpretations through which, it is hoped, the structure of the theory becomes apparent, motivating later developments.

Sponsored in part by the United States Army under Contract No. DAAG29-75-C-0024, by the Air Force Office of Scientific Research Grant AFOSR-79-0018 and by the Office of Naval Research Contract 041-404.

1. Solutions, representations, extremals

A minimization problem is represented as a triple

$$(f, A, (X, \tau)) \tag{1.1}$$

where (X, τ) is a topological space, f is a function defined on X valued in $\mathbb{R} \cup \{+\infty\}$ and A is a subset of X. We refer to f as the minimized function (or objective function); the set A is called the constraint of (1.1), and if $A = X$, the problem is said to be unconstrained.

It is common to write (1.1) in the form

$$f(x) \to \inf, \quad x \in A \; ; \quad \tau \tag{1.2}$$

and we shall sometimes follow this convention.

An element x_0 of A is said to be a solution of (1.1), if there exists a neighborhood Q of x_0 such that $f(x_0) \le f(x)$ for x in $A \cap Q$. The set of all the solutions of (1.1) is denoted by $\mathcal{R}(f, A, (X, \tau))$. Note that a solution of (1.1) corresponds to a local minimum (with respect to τ) in distinction from a (global) minimum (i.e. an element x_0 of A such that $f(x_0) \le f(x)$ for each x in A.) Observe that if $\tau_1 \le \tau_2$, then a solution of $(f, A, (X, \tau_1))$ is also a solution of $(f, A, (X, \tau_2))$. The number $\inf_{x \in A \cap Q} f(x)$ is called the Q-value of (1.1) (value, if $Q = X$). It may be finite, $+\infty$ or $-\infty$. We adopt the convention that the infimum of every function on the empty set is $+\infty$.

The reason for introducing the new terminology is two-fold. Primarily, a global minimum of (1.1) is a solution of (1.1) for the chaotic topology

$\tau = (\emptyset, X)$; thus the definition embraces both local and global minima. The theory we are going to present deals with solutions and it is immaterial whether they represent local or global minima. Secondly, the topology is made explicit in our definition, in order to avoid the confusion than can arise otherwise*.

1.1 <u>Example</u>

Let $f_0, \cdots, f_m$ be real-valued functions defined on a set X and let S be a subset of X. Consider the problem

$$
(1.3) \qquad \begin{array}{ll}
f_0(x) \to \inf & , \tau \\
x \in S & \\
f_i(x) = 0 & i = 1, \cdots, p \\
f_i(x) \le 0 & i = p+1, \cdots, m \;.
\end{array}
$$

Here f_0 is the objective function, while the constraint is $A = \{x : f_i(x) = 0 \;, \; i \le p, \; f_i(x) \le 0, \; i > p\} \cap S$. Frequently the set S is called the nonfunctional constraint of (1.3) and the functions f_i, $i = 1, \cdots, m$ are said to define the functional constraints at (1.3), more specifically equality constraint for $i \le p$ and inequality constraints for $p+1 \le i \le m$.

---

* A typical example is that of (local) strong and weak minima in the calculus of variations. The first refers to the topology of uniform convergence, the second to the Sobolev topology of uniform convergence together with first derivatives. The notions have nothing to do with strong and weak topologies.

Presumably, the above notions make a mathematician feel uneasy because of their lack of precision. Already the distinction between functional and nonfunctional constraints is quite vulnerable, as every set S may be represented by $\{x : \chi_S(x) \le 0\}$ (or by $\{x : \chi_S(x) = 0\}$), where χ_S is the characteristic function of S. Similar weaknesses can be seen in the distinction between equality and inequality constraints (an equality constraint can be represented as two inequality constraints).

What makes the discussed notions valid is that they do not refer to the constraint but to how it is represented.

Let Γ be a multifunction from a set Y into subsets of X, y_0 an element of Y. A couple (Γ, y_0) is a <u>constraint representation</u> of (1.1), if

$$A = \Gamma y_0 \ .$$

The set Y is called the <u>index set</u> (<u>parameter</u> set), its elements are indices (parameters).

1.2 <u>Example</u>

Let $Y = \{0, 1\}$. We define $\Gamma 1 = A$ and $\Gamma 0 = X \backslash A$. $(\Gamma, 1)$ is a constraint representation of (1.1). The inverse multifunction Γ^{-1} may be seen as the answer to the question: Is x in A? By setting $1 = \text{yes}$, $0 = \text{no}$ we obtain an answer in each case by looking at $\Gamma^{-1} x$.

1.3 <u>Example</u>

Let $f_1, \cdots, f_m$ be real-valued functions. Consider the set

$$A = \{x : f_i(x) \leq 0 , \quad i = 1, \cdots, m\} .$$

To represent this set we may choose the multifunction $\Gamma : \mathbb{R}^m \to 2^X$:

$$\Gamma(r_1, \cdots, r_m) = \{x : f_i(x) \leq r_i , \quad i = 1, \cdots, m\}$$

and an index $(0, \cdots, 0)$. But there are many other possibilities, for instance $\Delta : \mathbb{R} \to 2^X$:

(1.4)
$$\Delta r = \{x : \max_{1 \leq i \leq m} f_i(x) \leq r\}$$

and $r_0 = 0$. Another representation may be that of the previous example.

A minimization problem, as it is posed, includes a constraint representation. This is also true about problems given in the form (1.1), where the representation is simplest (that of Example 1.2).

The specific representation is the way in which a problem was presented to us, or a way of our comprehension of the problem. It may reflect the form under which a problem arises from the previous mathematical considerations or the mathematical model of a physical process.

We may loosely say that our access to the space X is indirect, via the space Y on which a constraint representation is defined. The question whether

$$x \in \Gamma y_0$$

amounts to the question (now in the space Y) whether

$$y_0 \in \Gamma^{-1} x .$$

This idea has a numerical aspect. If we consider the multifunction Γ of Example 1.3, then we observe that $\Gamma^{-1}x = (f_1(x), \cdots, f_m(x)) + \mathbb{R}_+^m$. Each element x of X is represented as an m-tuple of reals in Y. In order to see if x belongs to $\Gamma(r_1, \cdots, r_m)$ we need only check whether $f_i(x) \le r_i$, for $i = 1, \cdots, m$.

Let (Γ, y_0) be a constraint representation of (1.1). The quadruple

(1.5) $$(f, \Gamma, y_0, (X, \tau))$$

is called a <u>problem representation</u> (of (1.1) relative to (Γ, y_0)).

As we shall see, some representations are more convenient than others. In some circumstances one may wish to change a given representation to another more convenient under some aspect.

Let a multifunction $\Gamma: Y \to 2^X$ and y_0 in Y define a constraint representation of (1.1). Let $\mathcal{B}$ be a family of subsets of Y such that

$$\Gamma^{-1}x \in \mathcal{B} \quad , \quad x \in X .$$

Consider a multifunction E from a set Z into (subsets) of Y.

1.4 <u>Proposition</u>.

Let (Γ, y_0) be a constraint representation of (1.1). If there exists z_0 in $E^{-1}y_0$ such that $z_0 \notin E^{-1}B$ for every set B from $\mathcal{B}$ with $y_0 \notin B$, then $(\Gamma E, z_0)$ is a constraint representation of (1.1).

<u>Proof</u>

We need proof that $\Gamma E z_0 = \Gamma y_0$. Since $y_0 \in E z_0$, $\Gamma y_0 \subset \Gamma E z_0$. Assume that x is not in Γy_0, equivalently, y_0 is not in $\Gamma^{-1}x$, thus by our assumptions z_0 is not in $E^{-1}\Gamma^{-1}x$ and, equivalently, x is not in $\Gamma E z_0$.

1.5 Example

Consider Γ from Example 1.3. The sets $\Gamma^{-1}x = (f_1(x), \cdots, f_m(x)) + \mathbb{R}_+^m$ belong to the class $\mathcal{B} = \{y + \mathbb{R}_+^m ; y \in Y (= \mathbb{R}^m)\}$.

Let $Z = \mathbb{R}$ and let $E : \mathbb{R} \to 2^{\mathbb{R}^m}$ be defined by

$$Er = \{(r_1, \cdots, r_m) : r_i \le r, \quad i = 1, \cdots, m\}.$$

Then $(\Delta, 0)$, $\Delta = \Gamma E$ (1.4), is a representation of the constraint $A = \Gamma 0$. Indeed, the multifunction $E^{-1}(r_1, \cdots, r_m) = \max_{1 \le i \le m} r_i + \mathbb{R}_+$ has the property required in Proposition 1.4.

Observe that if $\kappa : Z \to Y$ is one-to-one, then E given by

$$Ez = \{\kappa(z)\}$$

satisfies the assumptions of the proposition.

Let (X, τ), (Z, π) be topological spaces and let $\Omega : Z \to 2^X$. A point (z_0, x_0) from $\mathcal{G}(\Omega)$ is called a singular point of Ω, if there is a neighborhood Q of x_0 such that z_0 is a boundary point of $\Omega^{-1}Q$.

1.6 Proposition

A point (z_0, x_0) is a singular point of Ω, if and only if Ω is not lower semicontinuous at (z_0, x_0).

Proof

It is enough to reformulate the definition (keeping in mind that z_0 is always in $\Omega^{-1}Q$): there is a neighborhood Q of x_0 such that for each neighborhood W of y_0

$$W \not\subset \Omega^{-1} Q \ .$$

This is just contrary to lower semicontinuity at (z_0, x_0) .

1.7 Corollary

Let τ_1, τ_2 be topologies of X, π_1, π_2 topologies of Z such that

$$\tau_1 \leq \tau_2 \qquad \pi_1 \geq \pi_2 \ .$$

Let $\Omega : Z \to 2^X$ and assume that (z_0, x_0) is a singular point of Ω with respect to (π_1, τ_1), then it is singular with respect to (π_2, τ_2)

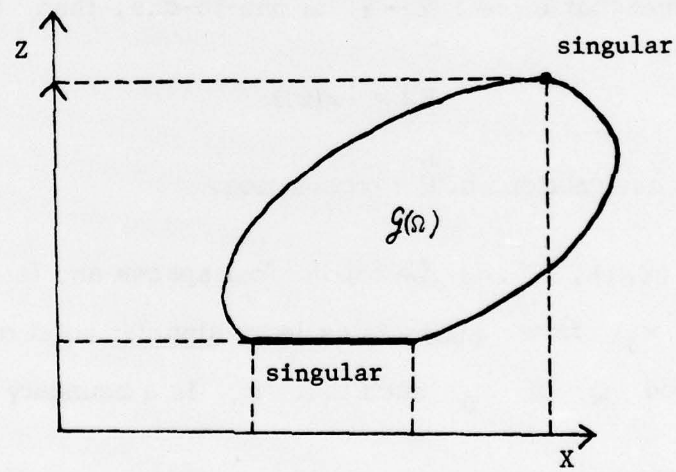


Fig. 1.1. Singular points of Ω

Let (Γ, y_0) (where $\Gamma : Y \to 2^X$) be a constraint representation of (1.1) . The multifunction $\Omega_{f,\Gamma} : Y \times \mathbb{R} \to 2^X$:

(1.6) $$\Omega_{f,\Gamma}(y, r) = \Gamma y \cap \{x : f(x) \leq r\}$$

is called the associated multifunction of (1.5). The inverse multifunction $\Omega_{f,\Gamma}^{-1}$ satisfies

$$(1.7) \qquad \Omega_{f,\Gamma}^{-1} x = \Gamma^{-1} x \times \{r : f(x) \le r\} .$$

Remark

If $r_1 \le r_2$ and $(y, r_1) \in \Omega_{f,\Gamma}^{-1} x$, then $(y, r_2) \in \Omega_{f,\Gamma}^{-1} x$.

1.8 Example

Consider the problem $(f_m, \Gamma 0, (X, \tau))$ where $\Gamma : \mathbb{R}^{m-1} \to 2^X$ is defined as $\Gamma(r_1, \cdots, r_{m-1}) = \{x : f_i(x) \le r_i, \ i = 1, \cdots, m-1\}$ (see Example 1.3). The associated multifunction Ω satisfies

$$\Omega_{f_m \Gamma}^{-1} x = (f_1(x), f_2(x), \cdots, f_m(x)) + \mathbb{R}_+^m .$$

Let σ be a topology of Y and ν be the natural topology of $\mathbb{R}$.

We say that an element x_0 is an extremal (of the problem representation (1.5)) with respect to σ, if $(y_0, f(x_0); x_0)$ is a singular point of the associated multifunction $\Omega_{f,\Gamma}$ with respect to $(\sigma \times \nu, \tau)$.

1.9 Example

Consider the problem of Example 1.9. Let σ be the natural topology of $\mathbb{R}^{m-1}$. Denote

$$\underset{\sim}{f}(x) = (f_1(x), f_2(x), \cdots, f_m(x)) .$$

An element x_0 is an extremal, if there is a neighborhood Q of x_0 such that $\underset{\sim}{f}(x_0)$ is on the boundary of $\underset{\sim}{f}(Q) + \mathbb{R}_+^m$.

Note that a sufficient condition for x_0 to be an extremal is that the sets

$$\underset{\sim}{f}(x_0) - \text{Int}\,\mathbb{R}_+^m\,, \quad \underset{\sim}{f}(Q)$$

be disjoint.

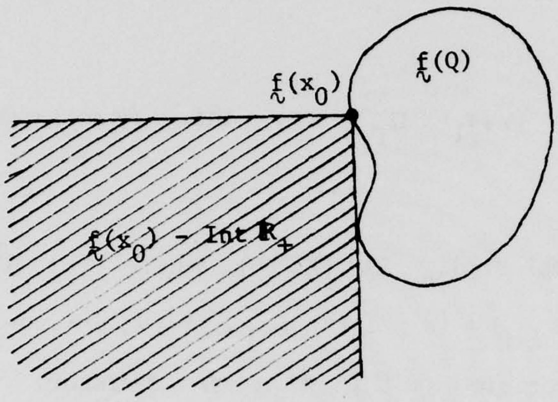


Fig. 1.2. Extremality of x_0 $(m = 2)$.

Indeed, the condition means that for each $\underset{\sim}{\xi} = (\xi_1, \cdots, \xi_m)$ such that $\xi_i > 0$, $i = 1,2,\cdots,m$, $\underset{\sim}{f}(x_0) - \underset{\sim}{\xi}$ is not in $\underset{\sim}{f}(Q)$, or $\underset{\sim}{f}(x_0) \not\in \underset{\sim}{f}(Q) + \underset{\sim}{\xi}$. If x_0 were not an extremal, then for some $\varepsilon > 0$ and each $\|\underset{\sim}{\eta}\| < \varepsilon$, $\underset{\sim}{f}(x_0) - \underset{\sim}{\eta} \in \underset{\sim}{f}(Q) + \mathbb{R}_+^m$. Let $\underset{\sim}{\eta} = (\eta_1, \cdots, \eta_m)$ be such that $\eta_i > 0$, $i = 1,2,\cdots,m$. Consequently there is a $\underset{\sim}{\xi}$ in $\mathbb{R}_+^m$ and an x in Q such that

$$\underset{\sim}{f}(x_0) = \underset{\sim}{f}(x) + \underset{\sim}{\xi} + \underset{\sim}{\eta} \;.$$

By setting $\underset{\sim}{\xi} = \underset{\sim}{\xi} + \underset{\sim}{\eta}$ we arrive at a contradiction.

Let (Γ, y_1) be a constraint representation of (1.1), where Γ is a multifunction from a topological space (Y_1, σ_1) into (X, τ) . Consider a multifunction E acting from a topological space (Y_2, σ_2) such that $(\Gamma E, y_2)$, y_2 in Y_2 , is another constraint representation of (1.1) (see Prop. 1.4).

1.10 Proposition

Suppose that E is lower semicontinuous at (y_2, y_1). If x_0 is an extremal of $(f, \Gamma E, y_2, (X, \tau))$, then it is an extremal of $(f, \Gamma, y_1, (X, \tau))$.

Proof

It is enough to show that if the associated multifunction $\Omega_{f,\Gamma}$ is lower semicontinuous at $(y_1, f(x_0), x_0)$, then the multifunction $\Omega_{f,\Gamma E}$ is lower semicontinuous at $(y_2, f(x_0), x_0)$.

Assume the former and let Q be a neighborhood of x_0. There are a neighborhood W_1 of y_1 and an $\varepsilon > 0$ such that

$$\Omega^{-1}_{f,\Gamma} Q = \{(y, r) : \exists_{x \in Q}\ y \in \Gamma^{-1} x,\ r \geq f(x)\} \supset W_1 \times (f(x_0) - \varepsilon, \infty) . \tag{1.8}$$

By the lower semicontinuity of E there is a neighborhood W_2 of y_2 such that $E^{-1} W_1 \supset W_2$. The set $W_2 \times (f(x_0) - \varepsilon, \infty)$ is a subset of $\Omega^{-1}_{f,\Gamma E} Q$.

Indeed, let z be in W_2 and $r > f(x_0) - \varepsilon$. Then there is a y in W_1 such that z is in $E^{-1} y$. On the other hand, by (1.8) there is an x in Q such that y is in $\Gamma^{-1} x$ and $r = f(x)$.

Observe that the multifunction E of Example 1.5 is everywhere lower semicontinuous.

The sense of Proposition 1.10 is that the change of a constraint representation with the aid of a lower semicontinuous multifunction E does not increase the set of extremals.

1.11 Proposition

An element x_0 is a solution of (1.1), if and only if it is an extremal of a problem representation (1.5) of (1.1) with respect to the discrete topology of Y.

Proof

Note that the family $\{\{y_0\} \times (f(x_0) - \varepsilon, f(x_0) + \varepsilon), \varepsilon > 0\}$ forms a neighborhood basis of $(y_0, f(x_0))$ in $Y \times \mathbb{R}$ in the considered topology.

Suppose that x_0 is an extremal: there is a neighborhood Q of x_0 such that for each $\varepsilon > 0$, $\{y_0\} \times (f(x_0) - \varepsilon, f(x_0) + \varepsilon) \not\subset \Omega_{f\Gamma}^{-1} Q$.

Let x be in $\Gamma y_0 \cap Q$. Consequently y_0 is in $\Gamma^{-1} x$ and, by (1.7), $(y_0, f(x))$ is in $\Omega_{f\Gamma}^{-1} Q$. By what we have just said and by Remark $f(x_0) - \varepsilon < f(x)$ for each $\varepsilon > 0$, thus x_0 is a solution.

Suppose now that $x_0 \in \Gamma y_0$ is not an extremal, that is, for each neighborhood Q of x_0 there is $r < f(x_0)$ such that $(y_0, r) \in \Omega_{f\Gamma}^{-1} Q$. Therefore there is an x in Q such that $r \geq f(x)$ and $y_0 \in \Gamma^{-1} x$. In other words $f(x) \leq r < f(x_0)$ and x is in $Q \cap \Gamma y_0$. The element x_0 is not a solution.

In view of Corollary 1.7 we have

1.12 Corollary

Let (Γ, y_0) be a constraint representation of (1.1) and let σ be any topology of Y. If x_0 is a solution of (1.1), then it is an extremal of (1.5) with respect to σ.

Here is an example of an extremal which is not a solution. Let $X = \mathbb{R}^2$ and τ be its natural topology. Consider the problem

$$x_2 \to \inf \ , \quad \tau$$
$$x_1^2 \le 0$$

and the constraint representation $(\{(x_1, x_2) : x_1^2 \le r\}, 0)$, $r \in \mathbb{R}$, $\mathbb{R}$ with its natural topology. Then $(0,0)$ is an extremal but not a solution. Conversely it is natural to ask when an extremal with respect to a given topology σ is a solution. Before providing an answer to this question we shall introduce the notion of value function.

A constraint representation (Γ, y_0), $\Gamma : Y \to 2^X$ of a problem (1.1) gives rise to a class of minimization problems, namely

$$(f, \Gamma y, (X, \tau)) , \quad y \in Y \ .$$

Sometimes these problems are called perturbed problems of (1.1). They are indexed by elements of Y (index set). The function $f\Gamma_Q$ defined on Y by the rule:

$$(1.9) \qquad f\Gamma_Q(y) = \inf_{x \in \Gamma y \cap Q} f(x)$$

is called the <u>Q-value function</u> of (1.5) (Q-primal functional, Q-perturbation function). When $Q = X$, it is called the value function and is denoted: $f\Gamma$. It is known [1]

1.13 <u>Proposition</u>

The epigraph of $f\Gamma_Q$ is equal to the closure of $\Omega_{f,\Gamma}^{-1} Q$ in the product of the discrete topology of Y and of the natural topology of $\mathbb{R}$.

The only elements that can belong to epi $f\Gamma_Q$ but not to $\Omega_{f,\Gamma}^{-1} Q$ are of the form

(1.10) $$(y, f\Gamma_Q(y))$$

An element (1.10) belongs to $\Omega^{-1}_{f,\Gamma} Q$, if and only if f attains its minimum in $Q \cap \Gamma y$.

It is an immediate observation, that x_0 is an extremal with respect to σ, if and only if there exists a neighborhood Q of x_0 such that $(y_0, f(x_0))$ is a boundary point of epi $f\Gamma_Q$.

We may now answer the question we have posed. Let (Γ, y_0) be a constraint representation of (1.1). We fix a topology σ of Y.

1.14 <u>Proposition</u>

Let x_0 be an extremal. If there is a neighborhood basis $\mathfrak{B}(x_0)$ of x_0 such that for each $Q \in \mathfrak{B}(x_0)$ the Q-value function is upper semicontinuous at y_0, then x_0 is a solution of (1.1).

<u>Proof</u>

Suppose that, on the contrary, x_0 is in Γy_0 but is not a solution of (1.1). It follows that for each $Q \in \mathfrak{B}(x_0)$, $\varepsilon(Q) = f(x_0) - f\Gamma_Q(y_0) > 0$. By upper semicontinuity there is a neighborhood W of y_0 such that for each y in W

$$f\Gamma_Q(y) < f(x_0) - \frac{\varepsilon(Q)}{2} .$$

Therefore $(y_0, f(x_0))$ is an interior point of epi $f\Gamma_Q$ for each Q in $\mathfrak{B}(x_0)$. This, in view of Proposition 1.6 contradicts the assumption that x_0 is an extremal.

We infer

1.15 Corollary

Let f be upper semicontinuous in a neighborhood of x_0. Assume that Γ is lower semicontinuous at (x, y_0) for x in Γy_0 and from a neighborhood of x_0. If x_0 is an extremal, then it is a solution.

1.16 Example

Let Γ be as in Example 1.9 and assume that $X = \mathbb{R}^n$ with its natural topology τ. Suppose that $f_1, \cdots, f_{m-1}$ are continuously differentiable at x_0. We shall show later (Corollary) that if

$$(f_1'(x_0), \cdots, f_{m-1}'(x_0))X + \mathbb{R}_+^{m-1} = \mathbb{R}^{m-1},$$

then Γ satisfies the assumptions of Corollary 1.15.

Proposition 1.13 offers precious information for analysis of minimization problems. It provides a geometrical interpretation (in the space $Y \times \mathbb{R}$)

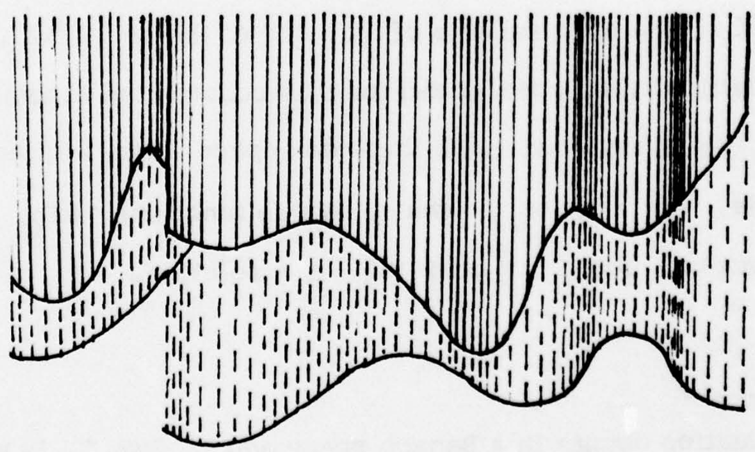


Fig. 1.3. Epigraph of a Q-value function, when its domain is a one-dimensional manifold in $\mathbb{R}^2$.

of extremality and minimality.

We are going to indicate how this interpretation helps to draw some interesting conclusions from general theorems. Take, for instance, Proposition 1.14.

1.17 Example

Suppose that the space Y of indices is given together with a topology ρ (for instance a Banach space with its strong topology).

If it happens that the domain of a constraint representation

$$\operatorname{dom} \Gamma = \{y : \Gamma y \neq \emptyset\}$$

is a boundary set in Y, then every element x (of Γy_0) is an extremal with respect to the topology ρ (this follows from the convention that $\inf_{x \in \Gamma y} f(x) = +\infty$, whenever $\Gamma y = \emptyset$)*. It is useful to introduce in Y a topology σ in the following way:

Induce the topology ρ on $Z = \operatorname{dom} \Gamma$ (a set $A \subset Z$ is open, whenever there exists an open set B for ρ such that $A = Z \cap B$). Let σ consist of all the open subsets of Z and of the set $Y \backslash Z$.

Proposition 1.14, specialized for thus constructed topology claims that, if x_0 is an extremal (with the index space $\operatorname{dom} \Gamma$), and if, for a basis of neighborhoods, $\{Q\}$ of x_0, the Q-value functions $f\Gamma_Q$ are upper semicontinuous on $\operatorname{dom} \Gamma$, then x_0 is a solution.

---

* Such a situation occurs in a Banach space when $\operatorname{dom} \Gamma$ is a manifold (of nonzero co-dimension).

1.18 Corollary

Suppose that there is a neighborhood Q of x_0 such that f is upper semicontinuous in Q. Assume that Γ is domain lower semicontinuous at (x, y_0) for each x in $Q \cap \Gamma y_0$. Then, an extremal of (1.5) with respect to σ, is a solution of (1.1).

A legitimate question at this moment is, why not choose as an index space the effective domain: $\Gamma^{-1} X$ rather than all of Y.

The motivation for staying with Y is two-fold. First, when knowing Γ it may be difficult to precisely describe $\Gamma^{-1} X$. Secondly, Y is usually chosen to enjoy a rich structure, which may not be shared by its subsets.

The former reason is linked with the controllability theory, where frequently it is very hard to determine what $\Gamma^{-1} X$ is.

We say that a multifunction $\Gamma : Y \rightarrow 2^X$ is controllable, if $\Gamma^{-1} X = Y$.

If (Y, σ) is a topological space, we say that Γ is (locally) controllable at y_0, if there is a neighborhood W of y_0 such that

$$\Gamma^{-1} X \supset W .$$

Notice that controllability at y_0 amounts to lower semicontinuity of Γ at (y_0, x) (for each x in X) with respect to the chaotic topology of X.

Controllability is lower semicontinuity with respect to the chaotic topology in both X and Y.

2. Interlude

Consider a standard minimization problem (1.1) :

(2.1) $$(f, A, (X, \tau)) .$$

Denote by τ_A the topology induced from (X, τ) on A and consider as well the problem

(2.2) $$(f, A, (A, \tau_A)) .$$

In the terminology of the previous section, the latter problem is unconstrained. One easily observes that (2.1) and (2.2) are equivalent in the sense that their sets of solutions are equal

$$\mathcal{R}(f, A, (X, \tau)) = \mathcal{R}(f, A, (A, \tau_A)) .$$

Formally, there is no reason to consider (2.1), as all information about what the problem consists of is given by (2.2). On the other hand, given (2.2) there are many ways of "embedding" it into problems of type (2.1) .

The motivation in formulating problems in the form (2.1) is that the space (X, τ) is usually nicer than (A, τ_A). "Nicer" in terms of a richer structure of (X, τ) (e.g. Banach space, analytic manifold, etc.) that enables us to use mathematical tools that have been developed for such structures.

The above thought inevitably entails another one: Why not try to replace (2.1) by an unconstrained problem on (X, τ), say

(2.3) $$(L, X, (X, \tau)) ,$$

and use precise tools available in the space (X, τ) to solve it? But be cautious! The function L in (2.3) must be nice too; otherwise it will resist precise tools.

We are concerned first with the replacement of (2.1) by (2.3) in the following sense: Let x_0 be a solution of (2.1). Find L so that x_0 is also a solution of (2.3).

Let $(\Gamma, y_0) : Y \to 2^X$ be a constraint representation of (2.1) and let σ be a topology of Y. Consider the scheme: Let x_0 be an extremal of (1.5). Find L so that x_0 is a solution of (2.3).

If we are able to reach the latter goal, then in view of Corollary 1.12, we shall also achieve our original objective. So the resulting problem ((2.3)) will constitute a <u>necessary condition</u> for solutions of the problem (2.1).

The role of the concept of extremal reveals on the occasion of the described replacement.

The inverse of the associated multifunction (1.7) completely characterizes (from the point of view of optimization) the points of X :

$$(y, r) \in \Omega^{-1}_{f,\Gamma} x ,$$

whenever x belongs to the constraint y and $f(x) \leq r$. As we pointed out in discussing the role of constraint representations, we face a minimization problem from the space $Y \times \mathbb{R}$ (Y-space of indices) through the multifunction $\Omega_{f,\Gamma}$. Therefore, our intervention (constructing the function L of (2.3)) may be done from the space $Y \times \mathbb{R}$.

The definition of extremal and Corollary 1.12 reflect an exceptional property of an element x_0 of X (being a solution) in an exceptional property of the image of x_0 : $(y_0, f(x_0))$ (being a boundary point). Now all depends on how we can take advantage of this new situation.

3. Relationship between constrained minimization and unconstrained minimization of the Lagrangian

Let X, Z be sets. Consider a multifunction $\Omega : Z \to 2^X$ and a family Ψ of (finite) real-valued functions on Z.

The Lagrangian of (Ω, Ψ) is the function $L : X \times \Psi \to \overline{\mathbb{R}}$ defined by

(3.1) $$L(x, \psi) = - \sup_{x \in \Omega^{-1} x} \psi(z) .$$

3.1 Lemma

Let Q be a subset of X and let x_0 be an element of Q. Suppose that z_0 from $\Omega^{-1} x_0$ and ψ_0 from Ψ satisfy

(3.2) $$\psi_0(z_0) \geq \psi_0(z) \quad \text{for} \quad z \in \Omega^{-1} Q .$$

Then

(3.3) $$-\psi(z_0) = L(x_0, \psi_0) = \min_{x \in Q} L(x, \psi_0) .$$

Proof

It follows from (3.2) that for each x in Q, $\psi_0(z_0) \geq \sup_{z \in \Omega^{-1} x} \psi_0(z)$, thus in view of (3.1)

(3.4) $$-\psi_0(z_0) \leq L(x, \psi_0) \quad \text{for} \quad x \in Q .$$

In particular for $x = x_0$ (3.4) yields

$$\psi_0(z_0) \geq -L(x_0, \psi_0) = \sup_{z \in \Omega^{-1} x_0} \psi_0(z) \geq \psi_0(z_0)$$

thus $L(x_0, \psi_0) = -\psi_0(z_0)$ plugged in (3.4) becomes (3.3).

Let τ be a topology of X and π a topology of Z. We say that Ω has the Ψ-<u>separation property at</u> x_0, if there is a neighborhood basis $\mathcal{B}(x_0)$ of x_0 such that for each Q in $\mathcal{B}(x_0)$ the set $\Omega^{-1}Q$ is Ψ-separated with respect to π (see [1]).

3.2 <u>Proposition</u>

Assume Ω to have the Ψ-separation property at x_0. Let $(z_0, x_0) \in \mathcal{G}(\Omega)$ be a singular point of Ω. Then there is a ψ_0 in Ψ such that x_0 is a solution of

$$(3.5) \qquad (L(\cdot, \psi_0), X, (X, \tau)) \quad .$$

<u>Proof</u>

By singularity there is a neighborhood Q_0 of x_0 such that z_0 is a boundary point of $\Omega^{-1}Q$ for each neighborhood $Q \subset Q_0$ of x_0. By Ψ-separation, for Q in $\mathcal{B}(x_0)$ and $Q \subset Q_0$, there is a ψ_0 such that (3.2) holds. Hence (3.3) is valid.

3.3 <u>Example</u>

Let Ω be that of Example 1.9. Let Ψ be the class of all linear forms on $\mathbb{R}^m$. Then

$$\Psi = \{\psi : \psi(r_1, \cdots, r_m) = \sum_{i=1}^{m} (-\lambda_i) r_i , \quad \lambda_i \in \mathbb{R}, \quad i = 1, \cdots, m\} \quad .$$

The Lagrangian for this case is

$$L(x, \lambda_1, \lambda_2, \cdots, \lambda_m) = \begin{cases} \sum_{i=1}^{m} \lambda_i f_i(x), & \text{if } \lambda_i \geq 0, \quad i = 1, \cdots, m \\ -\infty , & \text{otherwise} \quad . \end{cases}$$

3.4 Example

Consider Ω from the previous example but Ψ of the form

$$\Psi = \{\psi : \psi(r_1, \cdots, r_m) = -k \sum_{i=1}^{m} (r_i - s_i)^2 , \quad k \geq 0 ,$$

$$s_i \in \mathbb{R} , \quad i = 1, 2, \cdots, m\} .$$

The Lagrangian is

$$L(x, k, s_1, \cdots, s_m) = k \sum_{i=1}^{m} \max^2 (f_i(x) - s_i, 0)$$

3.5 Example

The same Ω as before. Now Ψ is the class

$$\Psi = \{\psi : \psi(r_1, \cdots, r_m) = -\lambda r_m - k \sum_{i=1}^{m-1} (r_i - s_i)^2$$

$$\lambda \geq 0 , \quad k \geq 0 , \quad s_i \in \mathbb{R}\} .$$

Then

$$L(x, k, s_1, \cdots, s_{m-1}, \lambda) = \lambda f_m(x) + k \sum_{i=1}^{m-1} \max^2 (f_i(x) - s_i, 0) .$$

This Lagrangian may be written in another way. By denoting

$$k \sum_{i=1}^{m-1} (r_i - s_i)^2 = k \sum_{i=1}^{m-1} r_i^2 + \sum_{i=1}^{m-1} \lambda_i r_i + c ,$$

we get

$$L(x,k,\lambda_1,\lambda_2,\cdots,\lambda_m,c) = \lambda_m f_m(x) + k \sum_{i=1}^{m-1} \max^2\left(-\frac{\lambda_i}{2k}, f_i(x)\right)$$

$$+ \sum_{i=1}^{n-1} \lambda_i \max\left(-\frac{\lambda_i}{2k}, f_i(x)\right) + c \quad .$$

3.6 Example

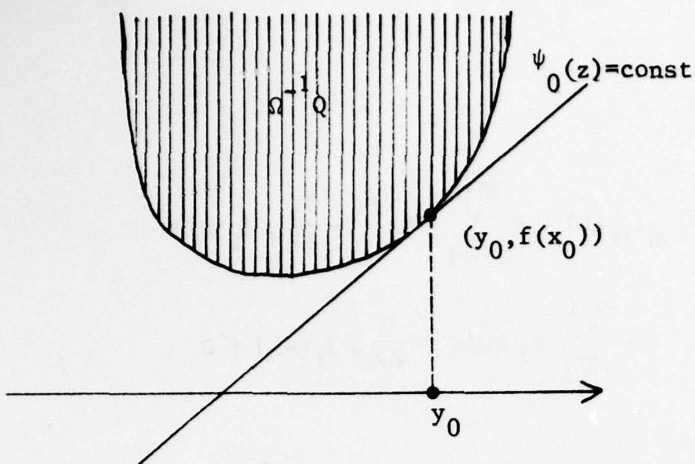
Again Ω as before. Let Ψ be

$$\Psi = \{\psi : \psi(r_1,\cdots,r_m) = -\lambda r_m - k \sum_{i=1}^{m-1} | r_i - s_i | ,$$

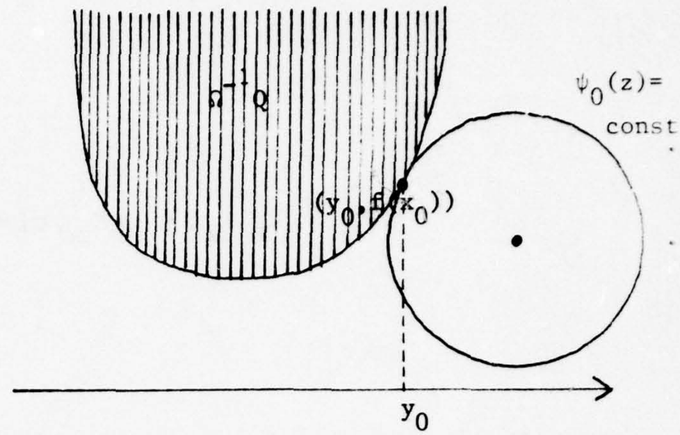
$$\lambda \geq 0 , \quad k \geq 0 , \quad s_i \in \mathbb{R}\} \quad .$$

Then

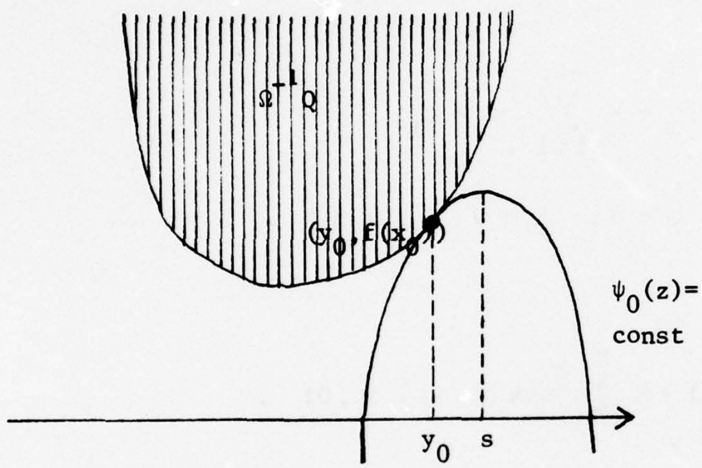
$$L(x,k,s_1,\cdots,s_m) = \lambda f_m(x) + k \sum_{i=1}^{m} \max(f_i(x) - s_i, 0) \quad .$$



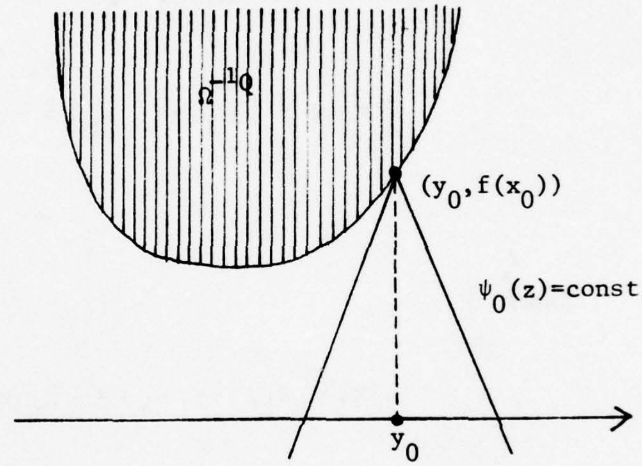
Ex.3.3



Ex.3.4.



Ex.3.5.



Ex.3.6.

Fig. 3.1. If Ω has convex graph, then it has the Ψ-separation property with respect to all the classes Ψ from Examples 3.3 through 3.6. It is enough to take bases of convex neighborhoods.

Let $\Omega_i : Y_i \to 2^X$, $i = 1, 2, \cdots, m$ and let Ψ_i be families of real-valued functions on Y_i for $i = 1, \cdots, m$. Denote by L_i the Lagrangian of (Ω_i, Ψ_i).

Consider the multifunction $\Omega : Y_1 \times \cdots \times Y_m \to 2^X$

$$\Omega(y_1, \cdots, y_m) = \bigcap_{i=1}^{m} \Omega_i y_i$$

and the family Ψ

$$\psi(y_1, \cdots, y_m) = \sum_{i=1}^{m} \psi_i(y_i) .$$

Then the Lagrangian of (Ω, Ψ) satisfies

$$L(x, \sum_{i=1}^{m} \psi_i) = \sum_{i=1}^{m} L(x, \psi_i) .$$

For instance Example 3.3 is of this type.

We are now concerned with Lagrangians of (Ω, Ψ), where Ω is the associated multifunction of a problem representation (1.5) and Ψ is a family (of functions on $Y \times \mathbb{R}$) of type

$$\Psi = \{\psi : \psi(y, r) = -\lambda r + \varphi(y) ; \quad \lambda \geq 0 , \quad \varphi \in \Phi\} , \tag{3.6}$$

Φ being a class of (finite) real-valued functions on Y. In view of (1.7) (3.6) the Lagrangian is of the type just discussed and may be written

$$L(x, \varphi, \lambda) = \lambda f(x) - \sup_{y \in \Gamma^{-1}x} \varphi(y) . \tag{3.7}$$

The term $- \sup_{y \in \Gamma^{-1}x} \varphi(y)$ is of course the Lagrangian of (Γ, Φ).

Example 3.3, 3.5 and 3.6 are of the considered type. As a consequence of Proposition 3.2 we have the following necessary condition of optimality.

3.7 Corollary

Suppose that the associated multifunction of (1.5) has the Ψ-separation property at x_0. If x_0 is a solution of (1.1), then there is a ψ_0 in (3.6) such that x_0 is a solution of (3.5).

Proof

By Corollary 1.12, x_0 is an extremal, that is $(y_0, f(x_0), x_0)$ is a singular point of Ω. By Proposition 3.2 x_0 is a solution of (3.5) for some ψ_0 in Ψ.

For Lagrangians of type (3.7) we have that

$$(3.8)\qquad \inf_{x \in Q} L(x, \varphi, \lambda) = \inf_{y \in Y} (\lambda f \Gamma_Q(y) - \varphi(y)) .$$

Indeed, on recalling (1.9),

$$\begin{aligned}\inf_{x \in Q} L(x, \varphi, \lambda) &= \inf_{x \in Q} (\lambda f(x) + \inf_{y \in \Gamma^{-1}x} (-\varphi(y))) \\ &= \inf_{x \in Q} \inf_{y \in \Gamma^{-1}x} (\lambda f(x) - \varphi(y)) \\ &= \inf_{y \in Y} (\lambda \inf_{x \in Q \cap \Gamma y} f(x) - \varphi(y)) \\ &= \inf_{y \in Y} (\lambda f \Gamma_Q(y) - \varphi(y)) .\end{aligned}$$

3.8 Example

Let Ω be that of Example 1.9. We assume that X is a normed space and that the functions $f_1, \cdots, f_m$ are convex. In this case we have that

$$\mathcal{R}(f_m, \Gamma 0, (X, \tau)) = \mathcal{R}(f_m, \Gamma 0, (X, \tau_0))$$

where τ stands for the natural topology of X while $\tau_0 = (\emptyset, X)$ denotes the chaotic topology. Let Ψ stand for the class of all nonzero linear forms on $\mathbb{R}^m$. Ω has the Ψ-separation property with respect to both τ and τ_0. As a consequence of Corollary 3.4 we have that if x_0 is a solution of $(f_m, \Gamma 0, (X, \tau_0))$, then there are numbers $\lambda_1 \geq 0, \cdots, \lambda_m \geq 0$ not simultaneously equal to zero such that

$$\sum_{i=1}^{m} \lambda_i f_i(x_0) = \min_{x \in X} \sum_{i=1}^{m} \lambda_i f_i(x) .$$

Indeed, there exists a ψ_0 $(\psi_0(r_1, \cdots, r_m) = \sum_{i=1}^{m} (-\lambda_i) r_i)$ such that

(3.9) $$\psi_0(f_1(x_0), \cdots, f_m(x_0)) \geq \psi_0(z), \quad \text{for} \quad z \in \Omega^{-1} Q .$$

We infer that $\lambda_i \geq 0$ for all i by setting in (3.9)

$$z = (f_1(x_0), \cdots, f_m(x_0)) + \xi \quad \text{in (3.9) for all } \xi \text{ from } \mathbb{R}_+^m .$$

From now on we shall assume that classes Φ that occur in (3.6) satisfy

(3.10) $$\mu\varphi + c \in \Phi , \quad \text{if} \quad \varphi \in \Phi , \quad \mu \geq 0 , \quad c \in \mathbb{R} .$$

All the classes of Examples 3.3 through 3.6 have the property that $\mu\varphi$ is in Φ provided that $\mu \geq 0$ and φ is in Φ. They acquire the property (3.10) if we complete them by setting

$$(3.11) \qquad \tilde{\Phi} = \{\varphi + c : \varphi \in \Phi, \quad c \in \mathbb{R}\} .$$

Augmenting the class Φ in this manner does not alter the Ψ-separation property.

Let $x_0 \in \Gamma y_0$ where Γ is that of (1.5). We say that a problem representation (1.5) is <u>Φ-normal at x_0</u>, if there are a neighborhood Q of x_0 and an element φ_0 of Φ such that

$$(3.12) \qquad -f(x_0) + \varphi_0(y_0) \geq -r + \varphi_0(y) \quad \text{for each} \quad (y,r) \in \Omega^{-1}Q .$$

In other words (3.12) means that one may separate $z_0 = (y_0, f(x_0))$ from $\Omega^{-1}Q$ by a ψ_0 of type

$$\psi_0(y,r) = -r + \varphi_0(y) .$$

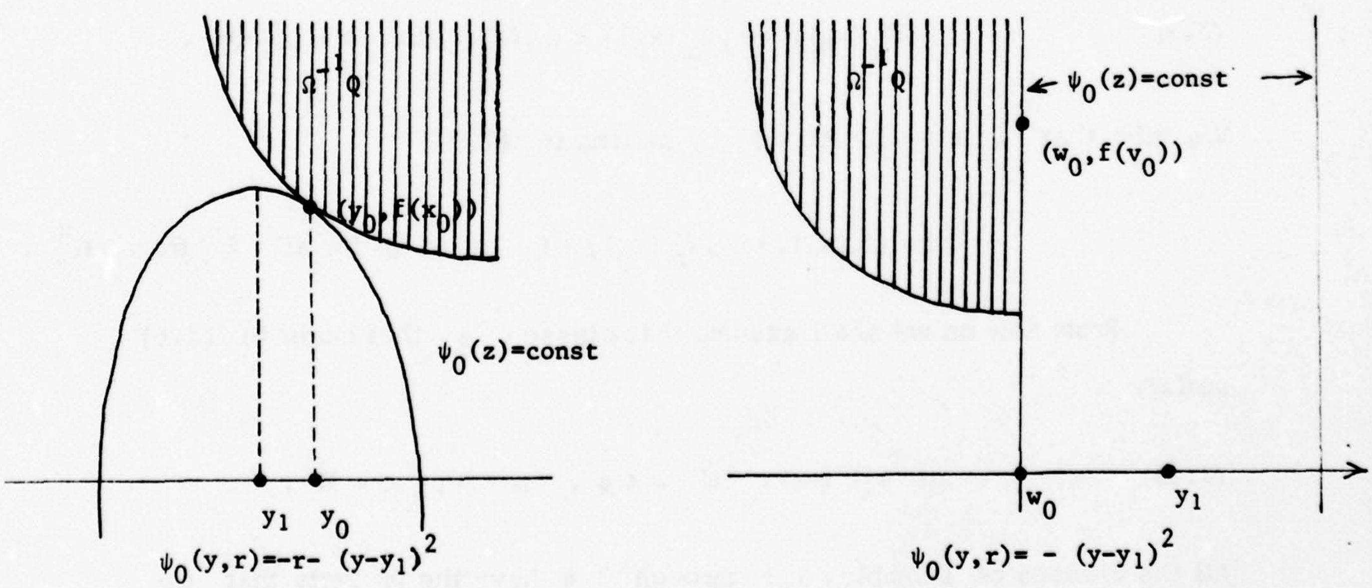


Fig. 3.2. Φ-normal at x_0 and not Φ-normal at v_0.

The class Ψ of Example 3.5.

3.9 Proposition

If a problem representation of (1.1) is Φ -normal at x_0, then x_0 is a solution of (1.1).

Proof

By definition (3.2) (in its special form 3.12) holds. By Lemma 3.1

$$f(x_0) - \varphi_0(y_0) = \min_{x \in Q} L(x, \varphi_0, 1)$$

and in view of (3.8)

$$f(x_0) - \varphi_0(y_0) = \inf_{y \in Y} (f\Gamma_Q(y) - \varphi_0(y)) \tag{3.13}$$

from which $f(x_0) \leq f\Gamma_Q(y_0)$, hence x_0 is a solution of (1.1).

3.10 Corollary

If a representation of (1.1) is Φ -normal at x_0, then the Q -value function is Φ -subdifferentiable at y_0. Moreover, if φ_0 is a subgradient of $f\Gamma_Q$ at y_0, then the corresponding ψ_0 satisfies

$$0 = \psi_0(y_0, f(x_0)) \geq \psi_0(z) \quad \text{for} \quad z \in \Omega^{-1}Q\ . \tag{3.14}$$

Proof

Indeed, by putting $f(x_0) = f\Gamma_Q(y_0)$ in (3.13) we get

$$f\Gamma_Q(y_0) - \varphi_0(y_0) \leq f\Gamma_Q(y) - \varphi_0(y) \qquad y \in Y\ . \tag{3.15}$$

Hence $\hat{\varphi}$ $(\hat{\varphi}(y) = \varphi_0(y) - f\Gamma_Q(y_0) + \varphi_0(y_0))$ is a subgradient of $f\Gamma_Q$ at

So far we have related unconstrained minimization of Lagrangians to weak separation of boundary points of $\Omega_{f\Gamma}^{-1} Q$ with the aid of functions from Ψ and obtained necessary optimality conditions.

We may derive analogous conditions without appealing to the existence of solutions.

In the case when solutions do not exist the mentioned results are no longer necessary conditions for optimality but still provide useful information about the structure of the problem.

The discussed method is based on the separation of boundary points of $\text{epi}\, f\Gamma_Q$ rather than those of $\Omega_{f\Gamma}^{-1} Q$. Proposition 1.13 reflects the advantage of this approach.

Analogously to Corollary 3.7 we have

3.11 <u>Proposition</u>

Suppose that the associated multifunction of a problem representation (1.5) has the Ψ-separation property at x_0. If $f\Gamma_Q(y_0) > -\infty$, then there are a neighborhood Q of x_0 and a ψ_0 of the form

$$\psi_0(y, r) = -\lambda_0 r + \varphi_0(y) \tag{3.16}$$

such that

$$\lambda_0 f\Gamma_Q(y_0) - \varphi_0(y_0) = \inf_{x \in Q} L(x, \lambda_0, \varphi_0) \tag{3.17}$$

$$= \inf_{y \in Y} (\lambda_0 f\Gamma_Q(y) - \varphi_0(y)) .$$

<u>Proof</u>

Each function of type (3.16) is continuous in the product of the discrete and the natural topologies, thus $\sup_{\Omega_{f\Gamma}^{-1} Q} \psi_0 = \sup_{\text{epi}\, f\Gamma_Q} \psi_0$. Let

$z_0 = (y_0, f\Gamma_Q(y_0))$. By Ψ-separation there is a ψ_0 of type (3.16) so that (in view of (3.1)) (3.4) holds, which in virtue of (3.8) amounts to (3.17).

The counterpart of normality is strong duality which consists in separating of boundary points of epi $f\Gamma_Q$ by elements of the class

$$\Phi \times 1 = \{\psi : \psi(y,r) = -r + \varphi(y) , \quad \varphi \in \Phi\} .$$

We say that <u>strong Φ-duality holds at x_0</u> (for a problem representation (1.5)) if there is a neighborhood Q of x_0 such that the point $(y_0, f\Gamma_Q(y_0))$ may be weakly separated from epi $f\Gamma_Q$ by a function from $\Phi \times 1$.

Certainly Φ-normality at x_0 implies strong Φ-duality at x_0 .

3.12 <u>Proposition</u>

Strong Φ-duality holds at x_0, if and only if there is a neighborhood Q such that the Q-value function is Φ-subdifferentiable at y_0 .

A function φ_0 is a subgradient of $f\Gamma_Q$ at y_0, whenever $\varphi_0 \times 1$ separates $(y_0, f\Gamma_Q(y_0))$ from epi $f\Gamma_Q$.

<u>Proof</u>

Strong Φ-duality at x_0 means the existence of φ_0 and a neighborhood Q of x_0 such that

(3.18) $$-f\Gamma_Q(y_0) + \varphi_0(y_0) \geq -r + \varphi_0(y) \quad \text{for each} \quad (y,r) \in \text{epi } f\Gamma_Q .$$

(3.18) is equivalent to

(3.19) $$f\Gamma_Q(y_0) - \varphi_0(y_0) \leq f\Gamma_Q(y) - \varphi_0(y) , \qquad y \in Y$$

which signifies that φ_0 is a subgradient of $f\Gamma_Q$ at y_0 .

In view of (3.8) formula (3.19) is equivalent to

(3.20) $$f\Gamma_Q(y_0) - \varphi_0(y_0) = \inf_{x \in Q} L(x, \varphi_0, y)) = \inf_{x \in Q} (f(x) - \sup_{y \in \Gamma^{-1}x} \varphi_0(y))$$

Frequently we use Lagrangians referred to a point (which differ from those we have been studying so far by a function depending merely on Ψ). The most important is <u>Lagrangian of Kurcyusz</u> (of a problem representation (1.5) with respect to Φ).

(3.21) $$L(x, \varphi, y_0) = f(x) - \sup_{y \in \Gamma^{-1}x} \varphi(y) + \varphi(y_0) \quad .$$

It is an immediate observation that (see (3.20)).

3.13 <u>Proposition</u>

An element φ_0 is a subgradient of $f\Gamma_Q$ at y_0, if and only if

(3.22) $$f\Gamma_Q(y_0) = \inf_{x \in Q} L(x, \varphi_0, y_0) \quad .$$

Formula (3.22) enables us to find the Q -value by computing the Q -value for a single unconstrained problem.

We know that Φ -subdifferentiability implies Φ -convexity (more generally for some classes of sets Ψ -separation implies Ψ -convexity).

We say that <u>weak Φ -duality holds at x_0</u> if there is a neighborhood Q of x_0 such that $(y_0, f\Gamma_Q(y_0))$ may be weakly almost separated from epi $f\Gamma_Q$ by a function from $\Phi \times 1$.

3.14 Proposition

Weak Φ-duality holds at x_0, if and only if there exists a neighborhood Q of x_0 such that the Q-value function is Φ-convex at y_0.

Proof

By definition weak Φ-duality holds at x_0, whenever for some neighborhood Q of x_0 and for each $\varepsilon > 0$ there is an element φ of Φ such that

$$f\Gamma_Q(y_0) - \varphi(y_0) - \varepsilon \le f\Gamma_Q(y) - \varphi(y), \qquad y \in Y$$

which, by the definition, amounts to Φ-convexity of $f\Gamma_Q$ at x_0.

3.15 Proposition

Weak Φ-duality holds at y_0, if and only if there is a neighborhood Q of x_0 such that

$$f\Gamma_Q(y_0) = \sup_{\varphi \in \Phi} \inf_{x \in Q} L(x, \varphi, y_0) \tag{3.23}$$

where L is the Lagrangian of Kurcyusz (3.21).

Proof

In view of (3.8) and on recalling the definition of the Fenchel transform [2]

$$\begin{aligned}\inf_{x \in Q} L(x, \varphi, y_0) &= \inf_{y \in Y} (f\Gamma_Q(y) - \varphi(y) + \varphi(y_0)) \\ &= -f\Gamma_Q^*(\varphi) + \varphi(y_0) .\end{aligned}$$

Therefore

$$\sup_{\varphi \in \Phi} \inf_{x \in Q} L(x, \varphi, y_0) = f\Gamma_Q^{**}(y_0) .$$

By the Moreau-Fenchel theorem ([2]), $f\Gamma_Q^{**} = f\Gamma_Q^{\Phi}$; $f\Gamma_Q$ is Φ -convex at y_0, whenever $f\Gamma_Q(y_0) = f\Gamma_Q^{\Phi}(y_0)$ thus in view of Proposition 3.14 the proof is complete.

Formula (3.23) enables us to compute the Q -value of a problem with a priori chosen precision by computing the Q -value for a single unconstrained problem.

It is the Ψ -separation property of associated multifunctions that makes possible the replacement of minimization problems by unconstrained minimization of the Lagrangian.

The resulting optimality conditions (see e.g. Corollary 3.7) carry the more information the smaller is the class Ψ, what is due to the existential qualifier "there is a ψ_0 in Ψ". Also, classes Ψ that are used in practice are chosen to make easy the computation of (3.1).* Consequently, good classes Ψ are those small and special.

Therefore, in such cases, the Ψ -separation is an exceptional property of multifunction. It may be expected that in the overwhelming majority of cases associated multifunctions will miss such a property.

As we shall see, sometimes a change of representation or a change of the class Ψ will be sufficient to overcome the impasse.

---

*Note that all the considered concrete associated multifunction of the type

$$\Omega^{-1}x = \begin{cases} \mathfrak{F}(x) + C, & x \in S \\ \emptyset, & \text{otherwise} \end{cases}$$

where $\mathfrak{F}$ is a mapping into Z and C is a closed convex one.

But most commonly we shall need substitute the associated multifunction by another multifunction enjoying the Ψ-separation property and preserving extremality.

The space Z is considered with two topologies π_1 and π_2. Let $\Omega : Z \to 2^X$. A multifunction $\Omega'_{(t_0, x_0)} : Z \to 2^X$ <u>is close to Ω at</u> (z_0, x_0) in the sense (π_1, π_2) if the lower semicontinuity of $\Omega'_{(z_0, x_0)}$ at (z_0, x_0) with respect to π_2 implies the lower semicontinuity of Ω at (z_0, x_0) with respect to π_1.

3.16 Proposition

Let $\Omega_{f\Gamma}$ be the associated multifunction of (1.5) and let $\Omega'_{(y_0, f(x_0), x_0)}$ have the Ψ-separation property at x_0 (with respect to π_2) and be close to $\Omega_{f\Gamma}$ at $(y_0, f(x_0), x_0)$ in the sense of (π_1, π_2), where $\pi_1 = \sigma_1 \times \nu$ (product of the topology σ_1 of Y and the natural topology ν of $\mathbf{R}$).

If x_0 is a solution of (1.5), then there is a ψ_0 in Ψ such that x_0 is a solution of (3.5) where L is the Lagrangian of $(\Omega'_{(y_0, f(x_0), x_0)}, \Psi)$.

Proof

Left to the reader.

3.17 Example

Let $f_1, \cdots, f_m$ be continuous functions on a Banach space X and differentiable at x_0. Let $\Omega : \mathbf{R}^m \to 2^X$ be defined by

$$\Omega(r_1, \cdots, r_m) = \{x : f_i(x) \le r_i, \quad i = 1, \cdots, m\} .$$

Let Ψ be the class of all linear forms on $\mathbf{R}^m$. The multifunction

$$\Omega'_{(f_1(x_0), \cdots, f_m(x_0), x_0)}(r_1, \cdots, r_m) = \{x : f_i(x_0) + f_i'(x_0)(x - x_0) \le r_i , \quad i = 1, \cdots, m\}$$

has the Ψ-separation property. It is known [1], that $\Omega'(f_2(x_0), \cdots, f_m(x_0), x_0)$ is close to Ω at $(f_2(t_0), \cdots, f_m(x_0), x_0)$ in (π, π), π the natural topology of $\mathbf{R}^m$.

We wish to emphasize that the multifunction $\Omega'_{(z_0, x_0)}$ was chosen for a given (z_0, x_0) from the graph of Ω. In many applications there is a routine procedure of finding close multifunctions (see Example 3.17).

4. Sufficient optimality conditions; improved necessary optimality conditions

A necessary condition for x_0 to be a solution of a problem (1.1) with a representation

$$(4.1) \qquad (f, \Gamma, y_0, (X, \tau))$$

is under some additional hypotheses, the existence of a ψ_0 in a class Ψ of real-valued functions such that x_0 is a solution of

$$(4.2) \qquad (L(\cdot, \psi_0), X, (X, \tau)),$$

where L is the Lagrangian of the associated multifunction and of the class Ψ.

Our next objective is to compare sets of solutions of (4.1) and (4.2).

In general the sets of all the solutions of (4.1) and (4.2)

$$(4.3) \qquad \Re(f, \Gamma y_0, (X, \tau)), \quad \Re(L(\cdot, \psi_0), X, (X, \tau))$$

are very weakly related.

We start by relating some subsets of (4.3)**, namely the sets

$$(4.4) \qquad \Re(f, \Gamma y_0 \cap Q, (Q, \tau_0)), \quad \Re(L(\cdot, \psi_0), Q, (Q, \tau_0))$$

where Q is an open set and τ_0 is the chaotic topology in Q.

Let $L(x, \varphi, \lambda)$ be the Lagrangian (3.7).

** See Exercise 9.

4.1 Proposition

Suppose that there is a neighborhood Q of x_0 and φ_0 in Φ such that φ_0 is a subgradient of $f\Gamma_Q$ at y_0. Then

$$\mathcal{R}(f, \Gamma y_0 \cap Q, (Q, \tau_0)) \subset \mathcal{R}(L(\cdot, \varphi_0, 1), Q, (Q, \tau_0)) \quad . \tag{4.5}$$

Proof

Let $\hat{x}$ belong to $\mathcal{R}(f, \Gamma y_0, (Q, \tau_0))$; thus $\hat{x}$ is in $\Gamma y_0 \cap Q$ and $f(\hat{x}) = f\Gamma_Q(y_0)$. Consequently for each function φ in Φ

$$f(\hat{x}) - \varphi(y_0) \geq f(\hat{x}) - \sup_{y \in \Gamma^{-1}\hat{x}} \varphi(y) = L(\hat{x}, \varphi, 1) \quad .$$

On the other hand, by (3.8),

$$f(\hat{x}) - \varphi_0(y_0) = f\Gamma_Q(y_0) - \varphi_0(y_0) = \inf_{x \in Q} L(x, \varphi_0, 1) \quad .$$

Hence, $L(\hat{x}, \varphi_0, 1) = \inf_{x \in Q} L(x, \varphi_0, 1)$.

Since the Lagrangian of Kurcyusz (3.21) differs from $L(x, \varphi_0, 1)$ by $\varphi_0(y_0)$, it follows from (3.22) that if $\hat{x}$ is a solution of $(f, \Gamma y_0 \cap Q, (Q, \tau_0))$ and φ_0 is a subgradient of $f\Gamma_Q$, then

$$f(\hat{x}) = f\Gamma_Q(y_0) = L(\hat{x}, \varphi_0, y_0) \quad . \tag{4.5}$$

4.2 Proposition

If φ_0 is a strict subgradient of $f\Gamma_Q$ at y_0 and if all $\Gamma^{-1}x$ are singletons, then

$$\mathcal{R}(f, \Gamma y_0 \cap Q, (Q, \tau_0)) = \mathcal{R}(L(\cdot, \varphi_0, 1), Q, (Q, \tau_0)) \quad . \tag{4.6}$$

Proof

In view of Proposition 4.1 it is enough to prove inclusion $\supset$. In our case the Lagrangian becomes

$$L(x, \varphi, 1) = f(x) - \varphi(\Gamma^{-1}x)$$

where by $(\Gamma^{-1}x)$ we denote the (only) element of $\Gamma^{-1}x$. Let $\hat{x}$ belong to the right hand side of (4.6). If $\hat{x}$ is in Γy_0, then (in view of (3.8))

$$f(\hat{x}) - \varphi_0(y_0) = \inf_{x \in Q} L(x, \varphi_0, 1) \leq f\Gamma_Q(y) - \varphi_0(y), \qquad y \in Y$$

thus $f(\hat{x}) \leq f\Gamma_Q(y_0)$ and $\hat{x}$ is in $\mathcal{R}(f, \Gamma y_0 \cap Q, (Q, \tau_0))$. If $\hat{x}$ were not in $\Gamma y_0 \cap Q$ then (denoting $\{\hat{y}\} = \Gamma^{-1}\hat{x}$)

$$L(\hat{x}, \varphi_0, 1) = f(\hat{x}) - \varphi_0(\hat{y}) \geq f\Gamma_Q(\hat{y}) - \varphi_0(\hat{y}) > f\Gamma_Q(y_0) - \varphi_0(y_0) .$$

On the other hand, by (3.8),

$$\inf_{x \in Q} L(x, \varphi_0, 1) \leq f\Gamma_Q(y_0) - \varphi_0(y_0),$$

leading to a contradiction.

4.3 Proposition

Suppose that all the sets $\Gamma^{-1}x$ are closed. If φ_0 is a decisive subgradient of $f\Gamma_Q$ at y_0 and if for each x in $\Gamma y_0 \cap Q$

$$\varphi_0(y_0) = \max_{y \in \Gamma^{-1}x} \varphi_0(y), \tag{4.7}$$

then (4.6) holds.

Proof

Let $\hat{x}$ belong to $\mathcal{R}(L(\cdot, \varphi_0, 1), Q, (Q, \tau_0))$.

If $\hat{x}$ is in Γy_0, then by (4.7) the thesis is proved.

Suppose that $\hat{x}$ is not in Γy_0. The set $\Gamma^{-1}\hat{x}$ is closed not containing y_0 hence by decisive subdifferentiability

$$(4.8)\qquad f\Gamma_Q(y_0) - \varphi_0(y_0) < \inf_{y \in \Gamma^{-1}\hat{x}} (f\Gamma_Q(y) - \varphi_0(y)) .$$

In virtue of (3.8)

$$L(\hat{x}, \varphi_0, 1) = f(\hat{x}) - \sup_{y \in \Gamma^{-1}\hat{x}} \varphi_0(y) \le \inf_{y \in \Gamma^{-1}\hat{x}} (f\Gamma_Q(y) - \varphi_0(y))$$

$$(4.9)\qquad \le \inf_{y \in \Gamma^{-1}\hat{x}} (f(\hat{x}) - \varphi_0(y))$$

$$= f(\hat{x}) - \sup_{y \in \Gamma^{-1}\hat{x}} \varphi_0(y) = L(\hat{x}, \varphi_0, 1) .$$

Again in view of (3.8) and by (4.8), (4.9)

$$\inf_{x \in Q} L(x, \varphi_0, 1) \le f\Gamma_Q(y_0) - \varphi_0(y_0) < L(\hat{x}, \varphi_0, 1)$$

4.4 <u>Example</u>

Let Φ be the following class of functions defined on a normed space Y

$$\Phi_1 = \{-k \| \cdot - y \| + c , \quad k \ge 0 , \quad y \in Y , \quad c \in \mathbb{R}\} .$$

Recall that this class has the property that if a function g is Φ-subdifferentiable at y_0, then there are k_0 and c_0 such that

$$\varphi_0(y) = -k_0 \| y - y_0 \| + c_0$$

is a decisive subgradient of g at y_0. For this φ_0, (4.7) is always satisfied.

Note that for all multifunctions Γ from the previous examples, the sets $\Gamma^{-1}x$ are closed.

In order to improve the necessary optimality conditions (like Corollary 3.7) we impose further requirements on the ψ_0 from Ψ used for separating z_0 from $\Omega^{-1}Q$.

Namely, we require that ψ_0 separates essentially.

This avoids trivial separation but does not exclude the situation where the Lagrangian is constant on Q.

4.5 <u>Example</u>

Let f be a linear continuous form on a Banach space X. Consider the following problem representation

$$(f, \{x : -f(x) \leq y\}, 0, (X, \tau_0)) .$$

The associated multifunction satisfies $\Omega^{-1}x = (-f(x), f(x)) + \mathbb{R}^2_+$, each point satisfying the constraint is an extremal, and its image may be separated by the function $\psi_0(r_1, r_2) = -r_1 - r_2$ from $\Omega^{-1}X$. We have that for each boundary point z_0 of $\Omega^{-1}X$

$$\psi_0(z_0) = \sup_{z \in \Omega^{-1}X} \psi_0(x) > \inf_{z \in \Omega^{-1}X} \psi_0(x) = -\infty$$

but the corresponding Lagrangian is null.

Even if the Lagrangian is constant and thus if the corresponding problem (3.5) is trivial, it may provide us with useful information. Take, as an example Proposition 3.16 : The Lagrangian we consider is that relative to a multifunction close to the associated one at the examined point. In this case (3.5) becomes the condition that the examined point should satisfy.

Essential separation may be performed, in some cases, with the aid of regular functions. A function ψ_0 is called regular, if for each open set Q the set $\psi_0(Q)$ is open.

4.6 Proposition

Suppose that a set A has a nonempty interior. If for an element z_0 of A and a regular function ψ_0 we have that

$$\psi_0(z_0) = \sup_{z \in A} \psi_0(z) ,$$

then z_0 is on the boundary of A and the separation is essential:

$$\psi_0(z_0) > \inf_{z \in A} \psi_0(z) .$$

Proof

Left to the reader.

This proposition applies in the cases where $\Omega^{-1}Q$ has a nonempty interior. In Example 1.17 we discussed the case in which $\Gamma^{-1}X$ is a boundary set in Y (hence $\Omega^{-1}Q$ must not have interior points). In this case one may require that the separating function is regular when restricted to $\Gamma^{-1}X \times \mathbb{R}$.

Note that a class Ψ of type (3.2) is composed only of regular functions, if we remove the subclass $\{\psi : \psi(y,r) = \varphi(y), \ \varphi \text{ is not regular on } Y\}$.

Let $\Gamma_i : Y_i \to 2^X$, $i = 1,2,\cdots,m$ be multifunctions. Define the multifunction $\Gamma : Y_1 \times Y_2 \times \cdots \times Y_m \to 2^X$ by

(4.10)
$$\Gamma(y_1, y_2, \cdots, y_m) = \bigcap_{i=1}^{m} \Gamma_i y_i \ .$$

Consider now a constraint representation $(\Gamma, \hat{y})$ of a minimization problem (1.1), where Γ is of type (4.10).

The first representation of Example 1.3, and many other representations we shall deal with, are of the discussed type.

The pairs $(\Gamma_1, \hat{y}_1), \cdots, (\Gamma_m, \hat{y}_m)$ are called <u>components</u> of the representation $(\Gamma, \hat{y})$.

We say that a component $(\Gamma_j, \hat{y}_j)$ <u>is active at x_0</u> $(x_0 \in \Gamma\hat{y})$, if x_0 is a boundary point of $\Gamma_j \hat{y}_j$.

A class Ψ is of type (3.6) where

$$\Phi = \{\sum_{i=1}^{m} \varphi_i \ ; \ \varphi_i \in \Phi_i\}$$

and for every i, Φ_i is a family of real-valued functions on Y_i satisfying (3.10).

4.7 <u>Proposition</u>

Suppose that there is a neighborhood Q of x_0 such that $(\hat{y}, f(x_0))$ is essentially separated from $\Omega_{f\,\Gamma}^{-1} Q$ by an element ψ of the form

(4.11)
$$\psi(y_1, \cdots, y_m, r) = -\lambda r + \sum_{i=1}^{m} \varphi_i(y_i) \ .$$

If $(\Gamma_j, \hat{y}_j)$ is not active at x_0, then there is a neighborhood $\hat{Q}$ of x_0 such that the element $\hat{\psi}$

(4.12)
$$\hat{\psi}(y_1, \cdots, y_m, r) = -\lambda r + \sum_{\substack{i=1 \\ i \neq j}}^{m} \varphi_i(y_i)$$

separates $(\hat{y}, f(x_0))$ essentially from $\Omega_{f\Gamma}^{-1}\hat{Q}$.

Proof

Since $(\Gamma_j, \hat{y}_j)$ is not active at x_0, there is a neighborhood $\hat{Q} \subset Q$ at x_0 such that

$$\hat{y}_j \in \Gamma_j^{-1}x\ , \quad \text{for every} \quad x \in \hat{Q}\ .$$

By our assumption, for each x in $\hat{Q}$

$$-\lambda f(x_0) + \sum_{i=1}^{m} \varphi_i(\hat{y}_i) \geq \sup_{(y_1, \ldots, y_m, r) \in \Omega_{f\Gamma}^{-1}x} (-\lambda r + \sum_{i=1}^{m} \varphi_i(y_i))$$

$$= -\lambda f(x) + \sum_{i=1}^{m} \sup_{y_i \in \Gamma^{-1}x} \varphi_i(y_i)$$

$$\geq -\lambda f(x) + \sum_{\substack{i=1 \\ i \neq j}}^{m} \sup_{y_i \in \Gamma^{-1}x} \varphi_i(y_i) + \varphi_j(\hat{y}_j)\ .$$

The proof is complete.

The meaning of Proposition 4.7 is that by ignoring nonactive components of constraint representation we may but improve the necessary optimality conditions.

5. Duality

Necessary conditions for the optimality of x_0 discussed in Section 3 assure the existence of a function ψ_0 such that x_0 is a solution of (3.5).

It turns out that ψ_0 is itself a solution of a certain minimization problem, called the dual problem.

We restrict these considerations to the chaotic topology τ_0 (to *global* solutions) and thus we abbreviate the notation e.g. $(f, A, (X, \tau_0)) = (f, A)$.

We start by introducing dual multifunctions.

Let Γ be a multifunction from Y into subsets of X. Given a family Φ of (finite) real-valued functions on Y we define the Φ-dual multifunction Γ^{Φ} (acting from $\mathbb{R}^X$ into subsets of Φ):

$$(5.1) \qquad \Gamma^{\Phi} f = \{\varphi \in \Phi : \sup_{y \in \Gamma^{-1} x} \varphi(y) \le -f(x), \quad x \in X\} .$$

An equivalent representation is

$$(5.2) \qquad \Gamma^{\Phi} f = \{\varphi \in \Phi : \varphi(y) \le (-f)\Gamma(y), \quad y \in Y\} .$$

In other words, the set $\Gamma^{\Phi}(-f)$ *consists of all those elements of* Φ that are majorized by the value function of $(f, \Gamma y_0)$.

Indeed we have that the condition in (5.1) is equivalent to

$$0 \le \inf_{x \in X} (-f(x) + \inf_{y \in \Gamma^{-1} x} (-\varphi(y)))$$

$$= \inf_{x \in X} \inf_{y \in \Gamma^{-1} x} (-f(x) - \varphi(y))$$

$$= \inf_{y \in Y} (\inf_{x \in \Gamma y} (-f(x)) - \varphi(y))$$

$$= \inf_{y \in Y} ((-f)\Gamma(y) - \varphi(y))$$

which amounts to the condition in (5.2). Certainly,

$$(5.3)\qquad Y \xrightarrow{\ \Gamma\ } 2^X \qquad\qquad 2^\Phi \xleftarrow{\ \Gamma^\Phi\ } \mathbb{R}^X .$$

5.1 Example

Let X and Y be normed spaces and let Φ denote the set of all affine continuous forms on Y: $\Phi = \{\varphi = \varphi_0 + c : \varphi_0 \in Y^*, \quad c \in \mathbb{R}\}$. Let

$$\Gamma y = A^{-1}(y)$$

where A is a continuous linear operator. In view of (5.1) for any $f : X \to \mathbb{R}$

$$\Gamma^\Phi(-f) = \{\varphi_0 + c : A^*\varphi_0(x) + c \le f(x), \quad x \in X\} .$$

Notice that for any function f, $\Gamma^\Phi(-f) = \Gamma^\Phi(-\bar{co} f)$.

5.2 Example

Let C be a closed convex cone in Y. Let

$$\Gamma y = \{x : y \in Ax + C\}$$

and take Φ from Example 5.1. We have that

$$\sup_{y \in Ax + C} \varphi(y) = \begin{cases} \varphi(Ax), & \text{if } \varphi_0 \in C^* \\ +\infty, & \text{otherwise} \end{cases},$$

thus

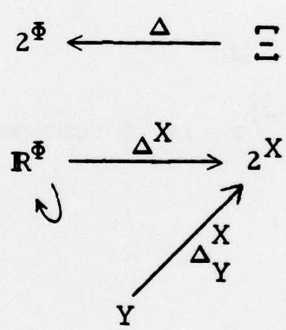
$$\Gamma^{\Phi} f = \{\varphi_0 + c : \ A^{*}\varphi_0 + c \le -f , \ \ \varphi_0 \in C^{*}\} .$$

Consider a multifunction Δ acting from a set Ξ of real-valued functions on X into subsets of Φ. Since the set X may be viewed as a subset of the real-valued functions on Ξ by the formula*

$$x(f) = f(x) , \qquad f \in \Xi \quad ,$$

we may consider the X-dual multifunction of Δ

(5.4)



By the definition

(5.5) $$\Delta^{X} g = \{x : \ \sup_{\xi \in \Delta^{-1}\varphi} x(\xi) \le -g(\varphi) , \ \ \varphi \in \Phi\} .$$

The restriction to Y (viewed as a subset of $\mathbf{R}^{\Phi}$) is denoted by Δ^{X}_{Y} .

---

* We write $X \hookrightarrow \mathbf{R}^{\Xi}$. Note, that two different elements x_1 , x_2 may define the same function on Ξ . This happens when $f(x_1) = f(x_2)$ for each $f \in \Xi$.

We are now in a position to give meaning to the X-dual of the Φ-dual of a multifunction and its restriction to Y. We define

$$\Gamma^{\Phi X} = (\Gamma^{\Phi})^X , \qquad \Gamma^{\Phi X}_Y = (\Gamma^{\Phi})^X_Y \tag{5.6}$$

where $\Xi = \mathbb{R}^X$.

5.3 <u>Proposition</u>

The X-dual of the Φ-dual of a multifunction Γ (restricted to Y) is equal to the Φ-hull of Γ:

$$\Gamma^{\Phi X}_Y = \mathrm{co}_{\Phi}\,\Gamma \quad . \tag{5.7}$$

Consequently, if for each x, $\Gamma^{-1}x$ is Φ-convex, then

$$\Gamma^{\Phi X}_Y = \Gamma \quad .$$

<u>Proof</u>

By definition, x is in $(\Gamma^{\Phi X}_Y)y$, whenever

$$\sup_{f \in (\Gamma^{\Phi})^{-1}\varphi} f(x) \le -\varphi(y) \quad \text{for each } \varphi \in \Phi \quad . \tag{5.8}$$

Now, f is in $(\Gamma^{\Phi})^{-1}\varphi$, whenever

$$-\sup_{y \in \Gamma^{-1}x} \varphi(y) \ge f(x) \quad . \tag{5.9}$$

The supremum of all these real-valued functions f that fulfill (5.9) is equal $-\sup_{y \in \Gamma^{-1}x} \varphi(y)$, hence condition (5.8) becomes

$$(5.10) \qquad \varphi(y) \le \sup_{z \in \Gamma^{-1}x} \varphi(z) \quad \text{for each} \quad \varphi \in \Phi .$$

By definition , (5.10) amounts to

$$y \in co_{\Phi} \Gamma^{-1} x$$

or equivalently $x \in co_{\Phi} \Gamma y$.

5.4 Proposition

Let Ξ be a subset of $\mathbb{R}^X$. If the Lagrangian $L(\cdot, \varphi)$ of (Γ, Φ) is Ξ -convex for each φ in Φ, then

$$(5.11) \qquad \Gamma^{\Phi X}_{\Xi Y} = co_{\Phi} \Gamma .$$

Proof

By definition x is in $(\Gamma^{\Phi X}_{\Xi Y})y$, whenever

$$\sup_{\xi \in (\Gamma^{\Phi}_{\Xi})^{-1}\varphi} \xi(x) \le -\varphi(y) \quad \text{for each} \quad \varphi .$$

A function ξ is in $(\Gamma^{\Phi}_{\Xi})^{-1}\varphi$, whenever

$$(5.12) \qquad L(x, \varphi) = - \sup_{y \in \Gamma^{-1}x} \varphi(y) \ge \xi(x) .$$

By the Ξ -convexity of the Lagrangian the supremum of those ξ from Ξ that fulfill (5.12) is equal to $-\sup_{y \in \Gamma^{-1}x} \varphi(y)$, thus (5.10) holds.

This proves the proposition.

5.5 Example

We shall consider the X-dual of the multifunction Γ^{Φ} from Example 5.1. Observe that the Lagrangian of (Γ, Φ)

$$L(x,\varphi) = -\varphi(Ax) = -A^*\varphi_0(x) - c$$

is Ξ-convex for each φ, if Ξ is the class of all continuous affine forms on X. Therefore, in view of Propositions 5.3 and 5.4,

$$\Gamma^{\Phi X}_{\Xi Y} = \Gamma^{\Phi X}_{Y} \, .$$

The set

$$(\Gamma^{\Phi}_{\Xi})^{-1}(\varphi_0 + c) = \{f \in \Xi : A^*\varphi_0(x) + c \le -f(x), \ \ x \in X\}$$
$$= \{-A^*\varphi_0 - c\} \, .$$

According to the definition, x is in $\Gamma^{\Phi X}_{\Xi Y}\, y$ whenever

$$-A^*\varphi_0(x) - c \le -\varphi_0(y) - c$$

for each φ_0 in Y^*. In other words, whenever $\varphi_0(Ax-y) = 0$, $\varphi_0 \in Y^*$ or $Ax = y$. This means that $\Gamma^{\Phi X}_{\Xi Y} = \Gamma$. This follows, of course from Proposition 5.3, since all $\Gamma^{-1}x = \{Ax\}$ are Φ convex.

Consider a problem representation

(5.13) $$(f, \Gamma, y_0)$$

where $\Gamma : Y \to 2^X$. The Φ-dual of (5.13) is defined as

(5.14) $$(-y_0, \Gamma^{\Phi}, -f)$$

where Γ^{Φ} is the Φ-dual multifunction of Γ. The <u>X-dual</u> of (5.14) is

$$(5.15) \qquad (f, \Gamma^{\Phi X}, y_0) .$$

Let $L_{\Phi} : \Phi \times X \times \mathbb{R}^X \to \overline{\mathbb{R}}$ be the Lagrangian of Kurcyusz of (5.14) and $L : X \times \Phi \times Y \to \overline{\mathbb{R}}$ be the Lagrangian of Kurcyusz of (5.13).

5.6 <u>Proposition</u>

The following equality holds

$$(5.16) \qquad -L_{\Phi}(\varphi, x, -f) = L(x, \varphi, y_0) .$$

<u>Proof</u>

By definition

$$L_{\Phi}(\varphi, x, -f) = -y_0(\varphi) - \sup_{g \in (\Gamma^{\Phi})^{-1}\varphi} x(g) - x(f) ,$$

thus in view of (3.21) we should merely show that

$$\inf_{g \in (\Gamma^{\Phi})^{-1}\varphi} (-g(x)) = \sup_{y \in \Gamma^{-1}x} \varphi(y) .$$

The set $(\Gamma^{\Phi})^{-1}\varphi$ is composed of all these real-valued functions g on X which satisfy $\sup_{y \in \Gamma^{-1}x} \varphi(y) \leq -g(x)$, thus (5.16) holds.

5.7 <u>Proposition</u>

If φ_0 is a subgradient of $f\Gamma$ at y_0, then it is a solution of (5.14).

Proof

If an element φ_0 is a subgradient of $f\Gamma$, then

(5.17) $$\varphi_0(y) \le f\Gamma(y)\,, \qquad y \in Y\,,$$

in other words, recalling (5.2), φ_0 is in $\Gamma^{\Phi}(-f)$, and if for all φ in $\Gamma^{\Phi}(-f)$,

(5.18) $$\varphi(y_0) \le \varphi_0(y_0)\ .$$

Thus φ_0 is a solution (5.14).

5.8 Proposition

If φ_0 is a solution of (5.14) and if $f\Gamma$ is Φ-convex at y_0, then φ_0 is a subgradient of $f\Gamma$ at y_0.

Proof

Let φ_0 be a solution of (5.14). In view of (5.18) $\sup_{\varphi \le f\Gamma} \varphi(y_0) = \varphi_0(y_0)$ and by Φ-convexity of $f\Gamma$ at y_0 $\varphi_0(y_0) = f\Gamma(y_0)$ which together with (5.17) completes the proof.

Certainly, the above propositions may be applied to establishing the relationship between solutions of (5.15) and the subgradients of (5.14). However, such results seem to be of little practical value, because of the difficulty in interpreting what X-subdifferentiability (X-convexity) of the value function of (5.14) means.

Instead, we present

5.9 Lemma

If $f\Gamma$ is Φ-convex at y_0 and x_0 is a solution of (5.13), then x_0 is a subgradient of the value function of (5.14) at $-f$.

Proof

The value function of (5.14) may be represented by

$$(5.19)\qquad (-y_0)\Gamma^{\Phi}(-g) = \inf_{\varphi \le g\Gamma} (-\varphi(y_0)) = -(g\Gamma)^{\Phi}(y_0) \,.$$

By Φ -convexity of $f\Gamma$ at y_0

$$(-y_0)\Gamma^{\Phi}(-f) = -f\Gamma(y_0) \,.$$

Therefore, taking into account that $f\Gamma(y_0) = f(x_0)$,

$$(-y_0)\Gamma^{\Phi}(-f) + f(x_0) = -f\Gamma(y_0) + f(x_0) = 0 \,.$$

On the other hand since $x_0 \in \Gamma y_0$, we have for each g that

$$(-y_0)\Gamma^{\Phi}(-g) + g(x_0) = -(g\Gamma)^{\Phi}(y_0) + g(x_0)$$

$$\ge -g\Gamma(y_0) + g(x_0) \ge 0 \,.$$

This signifies that $(-y_0)\Gamma^{\Phi}(\cdot) - x_0(\cdot)$ attains its minimum equal to zero at $-f$. The proof is complete.

5.10 Proposition

Assume that weak Φ -duality holds. If x_0 is a solution of (5.13) and φ_0 is a solution of (5.14), then (x_0, φ_0) is a saddle point of the Lagrangian of Kurcyusz (3.21):

$$(5.20)\qquad L(x_0, \varphi, y_0) \le L(x_0, \varphi_0, y_0) \le L(x, \varphi_0, y_0)$$

for each $x \in X$ and each $\varphi \in \Phi$.

Proof

By Proposition 5.8, φ_0 is a subgradient of $f\Gamma$ at y_0, thus by Proposition 3.13 the second inequality holds.

By Lemma 5.9, x_0 is a subgradient of the value function of (5.14) at $-f$. Hence by Proposition 3.13 (applied to (5.14)) and by (5.16) the first inequality holds.

Exercises

1. Parametric problems consist in minimizing over X

$$f(x, w_0)$$

provided that $z_0 \in \Delta^{-1}(x, w_0)$, where $f : X \times W \to \mathbb{R}$ and $\Delta : Z \to 2^{X \times W}$. Find a problem representation (1.5) for the general parametric problem.

2. (Rockafellar)

In particular, let $f : X \times W \to \mathbb{R} \cup \{+\infty\}$. Consider the problem

$$(f(\cdot, w_0), X, (X, \tau_0)) .$$

Determine a constraint representation (Γ, w_0), $\Gamma : W \to 2^{X \times W}$ that yields a problem representation of the type

$$(f, \Gamma, w_0, (X \times W, \tau_0 \times \tau_1)) .$$

Show that the Lagrangian of the associated multifunction and of the class $\Phi \times 1$ (see p. 31) is of the form

$$L(x, w, \varphi) = \inf_{z \in Y} (f(x, z) - \varphi(z)) .$$

3. Let R be a normed space, K a closed convex cone in R, f a mapping from a topological space (X, τ) into R, A a subset of X. A Pareto minimization problem (vector minimization) is

$$(f, (R, K), A, (X, \tau)) .$$

A solution of the problem is an element x_0 of A such that there is a neighborhood Q of x_0 such that for each x in $A \cap Q$, $f(x_0) - f(x) \notin K$. Our problem (1.1) is a special case of Pareto minimization.

If (Γ, y_0) is a constraint representation, then the associated multifunction Ω is defined by

$$\Omega(y,z) = \{x : z \in f(x) + K\} \cap \Gamma y .$$

Find a geometrical interpretation of singular points. What is the form of the Lagrangian of Ω and of a class

$$\Psi = \{\psi_1 + \psi_2 : \psi_1 : Z \to \mathbb{R}, \ \psi_2 : Y \to \mathbb{R}\} .$$

4. Consider a set $f_1, \cdots, f_m$ of real-valued functions on a topological space (X, τ). Let σ be the natural topology of $\mathbb{R}^m$. Then x_0 is an extremal for either all of or none of the following representations.

$$(f_j, \Gamma_j, (f_1(x_0), \cdots, f_{j-1}(x_0), f_{j+1}(x_0), \cdots, f_m(x_0)) \ (X, \tau)$$

where

$$\Gamma_j(r_1, \cdots, r_{j-1}, r_{j+1}, \cdots, r_m) = \{x : f_i(x) \le r_i , \quad i \ne j\} .$$

5. a) ([1]) Let A be a subset of a set X. A family P of functions $p : X \to \mathbb{R} \cup \{+\infty\}$ is called a penalty system for A, if

(i) $\lambda p \in P$ for each $\lambda > 0$ and $p \in P$

(ii) $p(x) \le 0$ for each $x \in A$ and $p \in P$

(iii) for each $x \notin A$ there is $p \in P$ such that $p(x) > 0$

(iv) for each $x \in A$ there is $p \in P$ such that $p(x) > -\infty$.

Let $f : X \to \mathbb{R} \cup \{+\infty\}$. Show that if P is a penalty system of A, then

$$\inf_{x \in X} \sup_{p \in P} (f(x) + p(x)) = \inf_{x \in A} f(x) .$$

b) (Dolecki-Kurcyusz [2]) Let Φ be a class of real-valued functions on Y satisfying (3.10), let $\Gamma : Y \to 2^X$.

Define

$$P(y_0) = \{- \sup_{y \in \Gamma^{-1}x} \varphi(y) + \varphi(y_0) , \; \varphi \in \Phi\} .$$

Show that if for each x the set $\Gamma^{-1}x$ is Φ-convex and $\bigcap_{y \in Y} \Gamma y = \emptyset$, then for each y_0 $P(y_0)$) is a penalty system for Γy_0.

6. [1] Let Φ be composed of regular functions and let Ψ be of type (3.6). Show that if $f\Gamma_Q$ is upper semicontinuous at y_0 and $(y_0, f\Gamma_Q(y_0))$ may be separated from epi $f\Gamma_Q$ by an element ψ of Ψ, then strong Φ-duality holds.

7. Using the definition, find $\Gamma_Y^{\Phi X}$ from Example 5.2.

8. Prove Proposition 3.16.

9. Show that for a problem representation (4.1)

$$\mathfrak{R}(f, \Gamma y_0, (X, \tau)) = \bigcup_{Q \in \mathfrak{B}} \Omega_{f,\Gamma}(y_0, f\Gamma_Q(y_0)) \cap Q ,$$

where $\mathfrak{B}$ is a basis for τ.

10. We say that a multifunction $\Gamma : Y \longrightarrow 2^X$ is Φ-observable at (y_0, x_0), whenever for each neighborhood Q of x_0 there exists a neighborhood W of y_0 such that

$$co_\Phi \Gamma^{-1} Q \supset W .$$

Show that if Γ, Φ are those of Example 5.1, then the Φ-observability of Γ (at (0,0)) amounts to the existence of k>0 such that

$$\|A^* \phi\| \geq k \|\phi\| , \qquad \phi \in Y^* .$$

Bibliography

[1] S. Dolecki, Abstract study of optimality conditions, J. Math. Anal. Appl., to appear.

[2] S. Dolecki, S. Kurcyusz, On Φ-convexity in extremal problems, SIAM J. Control Optim. 16 (1978), 277-300.

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER 2006	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) ABSTRACT OPTIMALITY THEORY		5. TYPE OF REPORT & PERIOD COVERED Summary Report - no specific reporting period
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Szymon Dolecki		8. CONTRACT OR GRANT NUMBER(s) AFOSR-79-0018 DAAG29-75-C-0024 041-404
9. PERFORMING ORGANIZATION NAME AND ADDRESS Mathematics Research Center, University of Wisconsin 610 Walnut Street Madison, Wisconsin 53706		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 5 - Mathematical Programming and Operations Research
11. CONTROLLING OFFICE NAME AND ADDRESS See Item 18 below		12. REPORT DATE October 1979
		13. NUMBER OF PAGES 58
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

U. S. Army Research Office
P.O. Box 12211
Research Triangle Park
North Carolina 27709

Air Force Office
of Scientific Research
Washington, D.C. 20332

Office of Naval Research
Arlington, Virginia 22217

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Extremal, Lagrangian, Duality, Penalty functions.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

We present the underlying ideas and basic notions of optimality theory in an abstract setting.