70-5097

HEINS, Willy, 1932-
CONJUGATE FUNCTIONS, DUALITY THEORY AND OPTIMAL CONTROL.

Case Western Reserve University, Ph.D., 1969
Mathematics

University Microfilms, Inc., Ann Arbor, Michigan

CONJUGATE FUNCTIONS, DUALITY THEORY AND OPTIMAL CONTROL

by

WILLY HEINS

Submitted in partial fulfillment of the requirements for the Degree of Doctor of Philosophy

School of Engineering
CASE WESTERN RESERVE UNIVERSITY

January 1969

CASE WESTERN RESERVE UNIVERSITY

GRADUATE STUDIES

We hereby approve the thesis of

Willy Heins

candidate for the Doctor of Philosophy

degree.

Signed: A. Ritter
(Chairman)

J. G. Windeknecht

R. S. Varga

P. O. Fredrickson

Irving Leffman

Date December 27, 1968

CONJUGATE FUNCTIONS, DUALITY THEORY, AND OPTIMAL CONTROL

by

Willy Heins

ABSTRACT

Optimal control problems with linear equations of evolution, convex cost functions and constraints on controllers and responses are considered. A Caratheodory-Hamilton-Jacobi approach to such problems leads to Pontryagin's maximal principle in a generalized form and the construction of the dual problem. A very general and complete duality theory is presented for a wide range of optimal control problems, resulting in existence theorems for optimality and a generalized maximal principle.

Duality theory applied to interconnected optimal control problems, with convex cost functions of separable type, gives rise to a decomposition of such interconnected problems into subproblems. Optimality of the overall problem is expressed in terms of optimality of the subproblems.

ACKNOWLEDGEMENTS

I wish to express my sincere appreciation to Dr. Sanjoy K. Mitter for his encouragement, insight and guidance throughout the course of this thesis.

The Systems Research Center provided a stimulating environment and I am grateful to staff and students for their assistance, meaningful criticism and interesting discussions.

Appreciation is also extended to Miss Linda Cottrell for her competent typing of the manuscript in shortest possible time.

TABLE OF CONTENTS

Page

CHAPTER I

INTRODUCTION

This research is concerned with the study of classes of optimal control problems which enjoy certain convexity properties. In all the cases considered, the equations of evolution are linear and the performance functionals are convex while the controllers as well as the responses are constrained to lie in convex sets.

The techniques used are based on results obtained in modern convex analysis concerning the properties of real valued convex functions and their conjugates, defined on linear topological vector spaces. In the literature the concept of convex conjugate functions has resulted, among other things, in a very satisfactory duality theory for convex extremum problems initiated by Fenchel [1] and generalized by Rockafellar [8]. The very generality of Rockafellar's problem setting in topological vector spaces, allows one to formulate a unified duality theory for a wide variety of optimal control problems described by ordinary, partial or functional differential equations, as has been carried out here. Duality theory as known in mathematical programming was applied to control problems by Pearson [2], [3], using the classical Legendre transformation. The concept of a conjugate function may be thought of as a generalization of the Legendre transformation.

Further, the work of Pshenichnyi [4] and Demyanov, Rubinov [5] on convex extremum problems has applications in optimal control and has a direct relationship to Rockafellar's duality theory although the approaches taken by the authors are different. The former are less complete but in their work certain numerical aspects are taken into consideration.

Novel results have been obtained for the classes of problems considered.

Existence theorems for optimal controllers are given under conditions weaker than that in the literature. Pontryagin's maximal principle is formulated in a generalized form and is shown to be valid under the much weaker condition of subdifferentiability rather than the usual differentiability requirements. An intimate and natural relationship between the maximal principle and the dual problem has been established in the sense that the role of the first in control theory as a necessary and sufficient condition for optimality, is similar to that of the second in mathematical programming inasmuch the primal problem has a solution if and only if the dual has one.

Furthermore, results have been derived for decomposing large interconnected control systems into smaller subsystems.

Existence theorems for optimality of the subproblems and a characterization of an optimal controller for the overall problem in terms of optimal controllers to the subproblems are given.

The theory underlying the results is selfcontained in the sense that it develops along some main lines directly originating in the basic properties of conjugate functions.

The material presented in this thesis is divided into eight chapters.

Chapter II contains a summary of the properties of convex functions.

In Chapter III a Caratheodory-Hamilton-Jacobi theory is given for linear processes with convex costs and special attention is paid to the convexity of the problem and the technique of handling the constraints. The ideas outlined in Chapter III are particularly useful in Chapter IV to prove the generalized maximal principle. From there we proceed to construct another optimal control problem with properties expected to be present in a dual problem as known in the theory of mathematical programming.

Having shown in Chapter IV how the maximal principle leads in a natural way to duality, we present in Chapter V a general duality theory due to Rockafellar in terms of a general optimal control problem and apply this duality theory to control processes with ordinary, partial and functional differential equations. This is done in Chapters VI, VII and VIII respectively where in each case the maximal principle is obtained from the duality theory.

In Chapter IX we apply the duality theory of Chapter V to interconnected control processes to arrive at a technique for decomposing the process into subprocesses.

CHAPTER II

CONVEX FUNCTIONS ON TOPOLOGICAL VECTOR SPACES

Some of the properties of convex functions which are used in the sequel are summarized here. For a more complete treatment, see [6], [7], [8], and the references given there.

Let E and E* be real vector spaces in duality with respect to a certain real bilinear function $\langle \cdot , \cdot \rangle$. We shall assume that E and E* have been assigned locally convex Hansdorff topologies compatible with this duality, so that elements of each space can be identified with continuous linear functionals on the other. E and E* will then be referred to as topologically paired spaces.

Properties of Convex Functions

Definition 2.1

An infinite valued convex function f on E is an everywhere defined function with range in $[-\infty,+\infty]$ whose (upper) epigraph

$$\text{epi}(f)= \{(x,\mu)|x \in E,\ \mu \in R,\ \mu \geq f(x)\}$$

is a convex set in $E \oplus R$.

If f does not assume both $-\infty$ and $+\infty$ as values this definition of a convex function is equivalent to

$$\forall x_1 \in E, \forall x_2 \in E: f(\lambda x_1+(1-\lambda)x_2)\leq\lambda f(x_1)+(1-\lambda)f(x_2),$$

$$0 \leq \lambda \leq 1.$$

Definition 2.2

The set,

$$\mathrm{dom}(f) = \{x \in E \mid f(x) < +\infty\}$$

is the effective domain of f.

Note that dom(f) is the projection of the epigraph epi f of f on E.

Definition 2.3

A convex function f on E is said to be proper if $f(x) > -\infty$ for all $x \in E$ and $f(x) < +\infty$ for at least one $x \in E$.

If f is a proper convex function, then dom(f) is non-empty and f is finite there. On the other hand given a finite valued convex function f on a non-empty convex set C in E, we obtain a proper convex function f_o with effective domain C by

$$f_o(x) = \begin{cases} f(x), & x \in C \\ +\infty, & x \notin C. \end{cases}$$

Definition 2.4

The indicator function ψ_C of a non-empty convex set C in E is defined as

$$\psi_C(x) = \begin{cases} 0, & x \in C \\ +\infty, & x \notin C \end{cases}$$

Thus, the indicator function of a non-empty convex set is a proper convex function.

Definition 2.5

A convex function f on E is lower semi-continuous (l.s.c.) if for each $\mu \in R$, the convex level set

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

is a closed set in E.

Lower semi-continuity of convex functions is a constructive property. Given any convex function f on E, one constructs a l.s.c. convex function $\bar{f}$ on E by taking

$$\forall x \in E: \bar{f}(x) = \liminf_{z \to x} f(z)$$

Conjugate Convex Functions

Definition 2.6

Let f be a proper convex function on E. Its conjugate function f* on E* (with respect to the given bilinear function <·,·>) is defined by

$$\forall x^* \in E^*: f^*(x^*) = \sup_{x \in E} \{\langle x, x^* \rangle - f(x)\}$$

The function f* is a l.s.c. convex function but not necessarily proper. However, if f is a l.s.c. proper convex function, then f^* is also l.s.c. proper convex and

$$(f^*)^* = f.$$

Thus, a one-to-one correspondence between the l.s.c. proper convex functions on E and those on E^* is defined by the formulas

$$f^*(x^*) = \sup_{x \in E}\{\langle x,x^*\rangle - f(x)\}$$
$$f(x) = \sup_{x^* \in E^*}\{\langle x,x^*\rangle - f^*(x^*)\} \qquad (2.1)$$

Functions f and f^* satisfying (2.1) are said to be conjugate to each other.

Definition 2.7

An element $x^* \in E^*$ is said to be a subgradient of the convex function f at the point x if

$$f(y) \geq f(x) + \langle y-x,x^*\rangle\ , \quad \forall\, y \in E.$$

The set of all subgradients at x, denoted by $\partial f(x)$ is a weak* closed convex set in E^* which might be empty. If $\partial f(x)$ is non-empty, the convex function f is said to be subdifferentiable at x. If f is differentiable in the sense of Frechet, $\partial f(x)$ consists of a single point, namely the gradient $\nabla f(x)$ of f at x. Further, if f(x) is finite the one-sided directional derivative

$$f'(x;z) = \lim_{\lambda \downarrow 0} \frac{f(x+\lambda z)-f(x)}{\lambda}$$

exists, although it might be infinite and it is a positively homogeneous convex function of z. Then,

$$x^* \in \partial f(x) \leftrightarrow f'(x;z) \geq \langle z,x^* \rangle \ , \ \forall z \in E$$

and such an x* exists if and only if f'(x;z) is bounded below in z in some neighborhood of 0. A useful fact regarding l.s.c. proper convex functions on a Banach space is the following.

Theorem 2.8

Let E be a Banach space and f a l.s.c. proper convex function on E with effective domain C. Assume C has non-empty relative interior ri(C). Then f is continuous on ri(C).

A function g on E is said to be concave if -g is convex. The theory of concave functions therefore parallels that of convex functions with certain natural changes. In particular,

$$g^*(y^*) = \inf_{y \in E} \{\langle y,y^* \rangle - g(y)\}$$

$$g(y) = \inf_{y^* \in E^*} \{\langle y,y^* \rangle - g^*(y^*)\}$$

define a one-to-one correspondence between the upper semi-continuous (u.s.c.) proper concave functions on E and those on E*.

The following property of conjugate functions which follows easily from the definitions will be frequently used and is stated as a theorem.

Theorem 2.9

If f is a l.s.c. proper convex function on E and g is an u.s.c. proper concave function on E then,

$$x^* \in \partial f(x) \leftrightarrow x \in \partial f^*(x^*) \leftrightarrow f(x)+f^*(x^*) = \langle x,x^* \rangle$$

$$x^* \in \partial g(x) \leftrightarrow x \in \partial g^*(x^*) \leftrightarrow g(x)+g^*(x^*) = \langle x,x^* \rangle$$

Integrals with Convex Integrands

We give some results on integrals of convex functions of the type

$$\int_T f(t,x(t))dt \quad , \quad \int_T f^*(t,x^*(t))dt$$

where f(t,x) is a proper convex and l.s.c. function of x for each t and thus in general not continuous in x and f*(t,x*) is the conjugate of f(t,x) for each t. Moreover one should like to know what the relationship is between these integrals, regarded as functionals on the spaces to which the curves x(t) and x*(t) belong. The following facts were proved in [9] in case f is defined on R^{1+n} but the results can be readily extended to separable Hilbert spaces as encountered in optimal control theory. In fact, the extension will follow if one formulates Lemmas 1 and 2 of [9] for a separable Hilbert space and furthermore recalls that the results of [11], used in [9], were proved for a Hilbert space rather than for R^n.

The following two Lemmas correspond to Lemma 1 and 2 of our reference [9].

Definition 2.10

Let T be a measure space with σ-finite measure dt and L be a real vector space of measurable functions u from T to a separable Hilbert space H and consider,

$$I_f(u) = \int_T f(t,u(t))dt$$

where f is a function from T × H into $(-\infty,+\infty]$. Then f is called a normal convex integrand if it satisfies the conditions

(i) f(t,x) is a l.s.c. proper convex function on H for each fixed t

(ii) there is a countable collection U of measurable functions u from T to H such that

(a) for each $u \in U$, f(t,u(t)) is measurable in t

(b) for each t, $U_t \cap \text{dom}(f(t,x))$ is dense in dom(f(t,x)) where,

$$U_t = \{u(t) | u \in U\}.$$

Lemma 2.11

Suppose f(t,x) = F(x) for all t, where F is a l.s.c. proper convex function on H. Then f is a normal convex integrand.

Proof: By the separability of H, there exists a countable dense subset D of the non-empty convex effective domain of F.

Let U consist of the constant functions on T with values in D.

Then U satisfies conditions (a), (b) in Definition 2.10 and since F is l.s.c. and proper, F is a normal integrand.

Q.E.D.

Lemma 2.12

Let the function f(t,x) on T × H have values in $(-\infty,+\infty]$ such that f(t,x) is measurable in t for each fixed x and for each t, f(t,x) is a l.s.c. proper convex function in x with interior points in its effective domain. Then f is a normal convex integrand.

Proof: Let D be a countable dense subset of H and let U be the set of constant functions on T with values in D.

Then U satisfies condition (a) of Definition 2.10. Further D has a dense intersection with the interior of dom(f(t,x)) and therefore with dom f(t,x) because dom(f(t,x)) as a convex set with non-empty interior has no isolated points.

Q.E.D.

Definition 2.13

Let T be a measure space with a σ-finite measure dt, H be a separable Hilbert space and L a real vector space of measurable functions from T to H. Then, L is said to be decomposable if it satisfies the following conditions

(i). L contains every bounded measurable function from T to H which vanishes outside a set of finite measure.

(ii) if $u \in L$ and C is a set of finite measure in T, then L contains $\chi_C u$ where χ_C is the characteristic function of C.

In other words, if L is decomposable one can alter functions in L arbitrarily in a bounded fashion on every set of finite measure.

Namely, subtract χ_C u from u and add any bounded measurable function vanishing outside C. If L* is topologically paired to L with respect to $<\cdot,\cdot>$ such that $<u(t),u^*(t)>_H$ is summable in t for every $u \in L$, $u^* \in L^*$, then condition (i) of Definition 2.13 also implies that the functions in L* are summable on sets of finite measure. An important class of function spaces which are decomposable in this sense are the $L_p^n(T)$ spaces.

Finally we give the following theorem which relates the integrals of conjugate normal convex integrands as conjugate functionals.

Theorem 2.14

Let L and L* be topologically paired by means of the summable inner product on H, that is

$$<u,u^*> = \int_T <u(t),u^*(t)>_H dt \quad \forall u \in L, \forall u^* \in L^*$$

and suppose L, L* are decomposable. Let f be a normal convex integrand such that f(t,u(t)) is summable in t for at least one $u \in L$ and f*(t,u*(t)) is summable in t for at least one $u^* \in L^*$. Then the functionals I_f on L and I_{f^*} on L* where

$$I_f = \int_T f(t,u(t))dt \qquad I_{f^*} = \int_T f^*(t,u^*(t))dt$$

are proper convex functions conjugate to each other.

Proof: See [9]

CHAPTER III

CARATHEODORY-HAMILTON-JACOBI THEORY AND OPTIMAL CONTROL IN CONVEX CONTEXT

3.1 Introduction

In this chapter we will outline a Caratheodory-Hamilton-Jacobi (C-H-J) theory for optimal control processes which enjoy certain convexity properties. Modern convex analysis leads to immediate generalizations of some classical concepts and sets up a unified approach to such problems in the sense that motivations and results are stated in terms of some basic properties of convex functions.

Given a differential process S),

$$S) \qquad \dot{x} = f(t,x,u), \quad x(0) = x_0$$

with cost function

$$C(u)= \int_0^T L(t,\phi(u)(t),u(t))dt$$

where the response vector $\phi(u)(t) \in E^n$ and the control vector $u(t) \in E^m$.

Suppose, the response of S) must be steered from x_0 into some target set G in E^n, on [0,T], using controllers u(t) constrained to lie in a given subset Ω of E^m such that C(u) is minimized. The aim of the C-H-J theory is to characterize an optimal controller in terms of the solutions to the Hamilton-Jacobi partial differential equation

$$\frac{\partial S(x,t)}{\partial t} + \max_{z\varepsilon\Omega} \{<f(t,x,z), \nabla S(x,t)> - L(t,x,z)\} = 0$$

$$S(x,T) = 0 , \qquad \forall\ x \ \varepsilon\ G$$

For various ways leading to the Hamilton-Jacobi equation in optimal control with the non-classical Hamiltonian,

$$H(t,x,\eta) = \max_{z\varepsilon\Omega} \{<f(t,x,z),\eta> - L(t,x,z)\}$$

See [12], [15], [16].

Here, we are interested in solutions S of the Hamilton-Jacobi equation which have the property that S(x,t) is a convex function on E^n for fixed t. This leads to a class of problems where $S)$ is a linear differential equation and the integrand L in the cost has convex properties.

3.2 Conjugate Functions and C-H-J Theory

Following the terminology in [17] we have

Definition 3.1

A control process in E^n, $\dot{x} = f(t,x,u)$ with control constraint set $\Omega \subseteq E^m$ and cost function

$$C(u) = \int_0^T h(t,\phi(u)(t),u(t))dt$$

and Hamiltonian $H(t,x,\eta,z) = \langle f(t,x,z),\eta\rangle - h(t,x,z)$ has a feedback control $u^o(t,x,\eta)$ if

$$\max_{z\varepsilon\Omega} H(t,x,\eta,z) = H(t,x,\eta,u^o(t,x,\eta)).$$

Furthermore, if η is specified as a function of t and x then $\overline{u}(t,x)= u^o(t,x,\eta(t,x))$ is called a control law.

Given a sufficiently smooth control law, say $\overline{u}(t,x) \;\varepsilon\; C^1$ on E^{1+n}, a response $\phi(\overline{u})$ satisfying

$$\dot{x} = f(t,x,\overline{u}(t,x)) \quad , \quad \phi(\overline{u})(0) = x_o$$

and a controller $\overline{u}(t) = \overline{u}(t,\phi(\overline{u})(t))$ are uniquely determined.

Consider the optimal control problem in E^n,

$$\text{(P)} \qquad L) \quad \dot{x} = A(t)x + B(t)u \;, \quad x(0) = x_o$$

A(t) and B(t) are continuous matrices on [0,T].

The responses $\phi(u)$ of $L)$ are elements of $L_p[E^n;0,T]$, $p \geq 1$, and are constrained to lie in the convex subset X of $L_p[E^n;0,T]$ defined as

$$X=\{\phi \;\varepsilon\; L_p[E^n;0,T] \,|\, \phi(0)=x_o, \phi(t) \;\varepsilon\; G \subseteq E^n \text{ on } [0,T], G \text{ is convex}\}$$

The controllers u are in $L_r[E^m;0,T]$, $r > 1$, and the class of admissible controllers U is the convex subset of $L_r[E^m;0,T]$ defined as

$$U = \{u \in L_r[E^m;0,T]\ u(t) \in \Omega \subseteq E^m \text{ on } [0,T], \Omega \text{ is convex, response } \phi(u) \in X\}$$

$$C(u) = \int_0^T \{H(t,u(t)) + K(t,\phi(u)(t))\}dt$$

and satisfies the following hypotheses

(i) $H(t,z)$ is continuous on E^{1+m} and for each t, $H(t,z)$ is a convex function on E^m.

(ii) $H(t,u(t))$ is summable in t for each $u \in L_r[E^m;0,T]$.

(iii) $K(t,x)$ is continuous on E^{1+n} and for each t, $K(t,x)$ is a convex function on E^n.

(iv) $K(t,\phi(t))$ is summable in t for each $\phi \in L_p[E^n;0,T]$.

Before we consider the Hamilton-Jacobi equation of this optimal control problem we first prove a lemma.

Lemma 3.2

Given the optimal control problem (P), for $x_0 \in E^n$, $t_0 \in [0,T]$, let $U(x_0,t_0)$ be the class of controllers defined by

$$U(x_0,t_0)=\{u \in L_r[E^m;0,T] | u(t) \in \Omega \text{ on } [0,T], \text{ response } \phi(u,x_0)$$

such that $\phi(u,x_0)(t_0) = x_0,\ \phi(u,x_0)(T) \in G\}$

Assume that for each $x_o \varepsilon E^n$, $t_o \varepsilon [0,T]$, there exists a $u^o \varepsilon U(x_o,t_o)$ such that the cost function

$$C(u;x_o,t_o) = \int_{t_o}^{T} \{H(t,u(t)) + K(t,\phi(u,x_o)(t))\}dt$$

is minimized for u^o with respect to $U(x_o,t_o)$ and that the minimal cost is finite.

Thus,

$$C(u^o;x_o,t_o) = \min_{u \varepsilon U(x_o,t_o)} C(u;x_o,t_o)$$

exists and is finite. Write $V(x_o,t_o)$ rather than $C(u^o;x_o,t_o)$. Then $V(x_o,t_o)$ is a finite convex function on E^n for each $t_o \varepsilon [0,T]$.

Proof: If $t_o = T$, the conclusion is obvious.

Fix some $t_o \varepsilon [0,T)$. To simplify our notation we write $C(u;x_o)$ rather than $C(u;x_o,t_o)$.

Because the differential equation L) is linear we have for any x_o^1, x_o^2 in E^n and any two controllers $u_1 \varepsilon U(x_o^1,t_o)$, $u_2 \varepsilon U(x_o^2,t_o)$

$$\phi(\lambda u_1+(1-\lambda)u_2, \lambda x_o^1+(1-\lambda)x_o^2)(t) = \lambda\phi(u_1,x_o^1)(t)+(1-\lambda)\phi(u_2,x_o^2)(t)$$

$$0 \leq \lambda \leq 1, \quad t \varepsilon [t_o,T].$$

Compute,

$$C(\lambda u_1+(1-\lambda)u_2;\lambda x_o^1+(1-\lambda)x_o^2) =$$

$$=\int_{t_o}^{T} H(t,\lambda u_1(t)+(1-\lambda)u_2(t))dt+\int_{t_o}^{T} K(t,\phi(\lambda u_1+(1-\lambda)u_2,\lambda x_o^1+(1-\lambda)x_o^2)(t))dt$$

$$\leq\lambda\int_{t_o}^{T} \{H(t,u_1(t)) + K(t,\phi(u_1,x_o^1)(t))\}dt +$$

$$+(1-\lambda)\int_{t_o}^{T} \{H(t,u_2(t)) + K(t,\phi(u_2,x_o^2)(t))\}dt$$

Or,

$$C(\lambda u_1+(1-\lambda)u_2;\lambda x_o^1+(1-\lambda)x_o^2)\leq\lambda C(u_1;x_o^1)+(1-\lambda)C(u_2;x_o^2)$$

By hypothesis there exist controllers $u_1^o \in U(x_o^1,t_o),u_2^o \in U(x_o^2,t_o)$ such that

$$\min_{u\in U(x_o^1,t_o)} C(u;x_o^1) = C(u_1^o;x_o^1) \equiv V(x_o^1,t_o)$$

$$\min_{u\in U(x_o^2,t_o)} C(u;x_o^2) = C(u_2^o;x_o^2) \equiv V(x_o^2,t_o)$$

Thus,

$$C(\lambda u_1^o+(1-\lambda)u_2^o;\ x_o^1+(1-\lambda)x_o^2) \leq \lambda V(x_o^1,t_o)+(1-\lambda)V(x_o^2,t_o)$$

Furthermore there exists a controller $u_3^o \in U(\lambda x_o^1+(1-\lambda)x_o^2,t_o)$ such that

$$\min_{u\varepsilon U(\lambda x_o^1+(1-\lambda)x_o^2,t_o)} C(u;\lambda x_o^1+(1-\lambda)x_o^2) = C(u_3^o;\lambda x_o^1+(1-\lambda)x_o^2)$$

$$\equiv V(\lambda x_o^1+(1-\lambda)x_o^2,\ t_o)$$

Hence,

$$V(\lambda x_o^1+(1-\lambda)x_o^2) \leq \lambda V(x_o^1,t_o)+(1-\lambda)V(x_o^2,t_o)$$

Thus $V(x_o,t_o)$ is a convex function on E^n for each $t_o \varepsilon [0,T]$ and by hypothesis it is finite.

Q.E.D.

The following theorem shows that the Hamilton-Jacobi equation of the optimal control problem (P) has a concave solution.

Theorem 3.3

Given the optimal control problem (P). Assume in addition that

(i) the hypotheses of Lemma 3.2 hold

(ii) the gradient ∇V of V is continuous on $E^n \times (0,T)$

Then, $S(x,t) = -V(x,t)$ satisfies the Hamilton-Jacobi equation

$$\frac{\partial S}{\partial t} + \max_{z\varepsilon\Omega} \{\langle Ax+Bz,\nabla S\rangle - H(t,z)-K(t,x)\} = 0$$

$$S(x,T) = 0, \qquad \forall\ x \ \varepsilon\ G$$

Proof: We use the dynamic programming formalism to prove the theorem. Regard the initial state $x_o \varepsilon E^n$ and the initial time $t_0 \in [0,T]$ as variables.

For some $x_0 \in E^n$ and some $t_0 \in [0,T)$ choose $\delta > 0$ small enough such that $t_0 + \delta \leq T$.

According the notation in Lemma 3.2,

$$\min_{u \in U(x_0,t_0)} C(u;x_0,t_0) = C(u^0;x_0,t_0) \equiv V(x_0,t_0)$$

Such an u^0 exists and $V(x_0,t_0)$ is finite by hypothesis (i).

Let U_1 be the class of controllers u, $u(t) \in \Omega$ with responses $\phi(u,x_0)$ such that $\phi(u,x_0)(t_0)= x_0$ and $\phi(u,x_0)(t_0+\delta)=\phi(u^0,x^0)(t_0+\delta)$. Then,

$$V(x_0,t_0)=\min_{u \in U_1}[\int_{t_0}^{t_0+\delta} \{H(t,u(t))+K(t,\phi(u,x_0)(t))\}dt + V(\phi(u^0,x_0)(t_0+\delta),t_0+\delta)] \tag{3.1}$$

Write,

$$V(\phi(u^0,x_0)(t_0+\delta),t_0+\delta)=V(\phi(u^0,x_0)(t_0+\delta),t_0)+ +\delta V_t(\phi(u^0,x_0)(t_0+\delta),t_0)+o_1(\delta) \tag{3.2}$$

where $o_1(\delta)/\delta \to 0$ with $\delta \to 0$.

Furthermore,

$$V(\phi(u^0,x_0)(t_0+\delta),t_0)=V(\phi(u^0,x_0)(t_0),t_0)+ +\langle\phi(u^0,x_0)(t_0+\delta)-\phi(u^0,x_0)(t_0),\nabla V(\phi(u^0,x_0)(t_0+\theta\delta),t_0)\rangle,$$

$$0 \leq \theta \leq 1$$

$$V(\phi(u^o,x_o)(t_o+\delta),t_o) = V(\phi(u^o,x_o)(t_o),t_o) + \tag{3.3}$$

$$+\delta<\dot{\phi}(u^o,x_o)(t_o),\ \nabla V(\phi(u^o,x_o)(t_o+\theta\delta),t_o)> + o_2(\delta)$$

where $o_2(\delta)/\delta \to 0$ with $\delta \to 0$.

Substitution of (3.2) and (3.3) into (3.1), using the fact that

$$V(x_o,t_o) = V(\phi(u^o,x_o)(t_o),t_o)$$

we obtain,

$$0 = \min_{u\varepsilon U_1} [\int_{t_o}^{t_o+\delta} \{H(t,u(t))+K(t,\phi(u,x_o)(t))\}dt +$$

$$+ V_t(\phi(u^o,x_o)(t_o+\delta),t_o)+\delta<\dot{\phi}(u^o,x_o)(t_o),\nabla V(\phi(u^o,x_o)(t_o+\theta\delta),t_o)>$$

$$+ o_1(\delta) + o_2(\delta)$$

Divide by δ and let $\delta \to 0$,

$$0 = V_t(x_o,t_o)+ \min_{u(t_o)\varepsilon\Omega} [H(t_o,u(t_o))+K(t_o,x_o)+<A(t_o)x_o$$

$$+ B(t_o)u(t_o),\ \nabla V(x_o,t_o)>]$$

This equation holds for every $x_o \varepsilon E^n$ and every $t_o \varepsilon [0,T)$. Drop the subscripts and substitute $S(x,t) = -V(x,t)$. Then

$$\frac{\partial S}{\partial t} + \max_{z\varepsilon\Omega}\{<Ax+Bz,\nabla S> - H(t,z)-K(t,x)\} = 0$$

and $S(x,T) = 0, \ \forall x \ \varepsilon \ G$ is obvious.

Q.E.D.

One of the important practical effects of allowing convex functions to have infinite values is that every constrained problem can be formulated as an unconstrained problem. The technique is based on redefining convex functions in such a way that they have constraint sets as their effective domains. We will do this for the optimal control problem above and consider also in detail the consequences.

Define the functions h and k as follows. For each t, $h(t,z) = H(t,z)$ if $z \ \varepsilon \ \Omega$, $h(t,z) = +\infty$ if $z \notin \Omega$ and $k(t,x) = K(t,x)$ if $x \ \varepsilon \ G$, $k(t,x) = +\infty$ if $x \notin G$.

Then, for each t, $h(t,z)$ is a proper convex function on E^m with effective domain Ω and $k(t,x)$ is a proper convex function on E^n with effective domain G.

For the integral

$$\int_0^T \{h(t,u(t))+k(t,\phi(u)(t))\}dt$$

to make sense for all $u \ \varepsilon \ L_r[E^m;0,T]$ and all $\phi(u) \ \varepsilon \ L_p[E^n;0,T]$ we require, among other things, that for each t, $h(t,z)$ and $k(t,x)$ are l.s.c. But since lower semi-continuity of convex functions is a constructive property, h and k can be made so, if necessary, by lowering their values on the boundaries of their effective domains.

In addition to hypotheses (i),- (iv) in (P) assume that

(v) $h^*(t,u^*(t))$ is summable for at least one $u^* \in L_s[E^m;0,T]$ $\frac{1}{r}+\frac{1}{s}=1$.

(vi) $k^*(t,\phi^*(t))$ is summable for at least one $\phi^* \in L_q[E^n;0,T]$, $\frac{1}{p}+\frac{1}{q}=1$

(vii) Ω , G have non-empty interiors

Note that all the assumptions (i), - (vii) hold in most practical cases.

Then, h(t,z) is measurable in t for each fixed z and for each t, h(t,z) is a l.s.c. convex function with interior points in its effective domain Ω. Thus, by Lemma 2.12, h(t,z) is a normal convex integrand. Furthermore, $h(t,u(t)) = H(t,u(t))$ for every $u \in U$. From these facts and assumption (v) above it follows from Theorem 2.14 that

$$I_h(u) = \int_0^T h(t,u(t))dt$$

is a well defined proper convex functional on $L_r[E^m;0,T]$ with effective domain U. Similarly we have that,

$$I_k(\phi) = \int_0^T k(t,\phi(t))dt$$

is a well defined proper convex functional on $L_p[E^n;0,T]$ with effective domain X. Therefore,

$$C_1(u) = \int_0^T \{h(t,u(t))+k(t,\phi(u)(t))\}dt \qquad (3.4)$$

is defined and it has a finite value if and only if $u \in U$. Thus if we replace the cost function in (P) by (3.4) then we obtain an unconstrained problem.

The Hamiltonian of the optimal control problem (P) with cost function defined in (3.4) is

$$H(\eta,x,t) = \sup_{z\in\Omega}\{\langle Ax+Bz,\eta\rangle -h(t,z)-k(t,x)\} \qquad (3.5)$$

and is defined for $\eta \in E^n$, $x \in G$, $t \in [0,T]$ although it may have infinite values.

Following the terminology in [17], the sup operation in (3.5) is referred to as the maximal principle.

Note that (3.5) can be written as

$$H(\eta,x,t)=\langle Ax,\eta\rangle-k(t,x)+\sup_{z\in E^m}\{\langle z,B^*\eta\rangle-h(t,z)\} =$$

$$= \langle Ax,\eta\rangle-k(t,x) + h^*(t,B^*\eta)$$

where B* is the transposed of B.

Thus, the maximal principle here is nothing else as the sup operation in the definition of a conjugate function.

Another point we should like to mention is that the concept of a conjugate function can be viewed as a generalized Legendre transform, see [10]. Where in the classical calculus of variations the Hamiltonian is obtained as the Legendre transformation of the Lagrangian , with respect to the relevant variable, in our case the Hamiltonian is defined in terms of the generalized Legendre transform of the Lagrangian h(t,z) + k(t,x), with respect to u, namely the conjugate function h* of h. Therefore one could take this point of view in defining the Hamilton-Jacobi equation

$$\frac{\partial S}{\partial t} + \sup_{z \in \Omega}\{\langle Ax+Bz, \nabla S\rangle - h(t,z) - k(t,x)\} = 0 \qquad (3.6)$$

$$S(x,T) = 0, \quad \forall\, x \in G$$

in analogy with the classical case.

The Hamilton-Jacobi equation (3.6) is defined on G x [0,T] where h(t,z) = H(t,z) on int(G) x [0,T] and k(t,x) = K(t,x) on int(Ω) x [0,T]. Therefore, S(x,t) = -V(x,t) is a solution to (3.6) except perhaps on boundary points, a price to be paid for the lower semi-continuity of h and k.

Remark

We considered the constraint sets Ω and G time independent. However, in case Ω(t) and G(t) are given convex sets at each instant of time, all the results above carry over without complications.

The following Theorem gives a necessary and sufficient condition for (P) to have a feedback controller, in which case the Hamiltonian (3.5) is finite valued.

Theorem 3.4

The optimal control problem (P) with cost function (3.4) has a feedback controller $u^o(\eta,t)$ defined by

$$\max_{z\varepsilon\Omega} \{<Bz,\eta>-h(t,z)\} = <Bu^o,\eta>-h(t,u^o)$$

if and only if the conjugate function $h^*(t,z^*)$ of $h(t,z)$ is subdifferentiable at $B^*\eta$.

Furthermore, if for some specified $\eta(x,t)$ the control law $\bar{u}(x,t) = u^o(\eta(x,t),t)$ determines the response $\phi(\bar{u})$ of L) such that $\phi(\bar{u})(t) \varepsilon G$, $\forall t \varepsilon [0,T]$, then $\phi(\bar{u})$, whether an optimal response or not, satisfies

$$\dot{x}(t) \varepsilon \partial_\eta H(\eta(t),x(t),t) \text{ where } \eta(t) = \eta(x(t),t)$$

and the Hamiltonian H is defined as,

$$\forall t \varepsilon [0,T], \forall x \varepsilon G: H(\eta,x,t)=\max_{z\varepsilon\Omega}\{<Ax+Bz,\eta>-h(t,z)-k(t,x)\}$$

$$= <Ax,\eta>-k(t,x) + h^*(t,B^*\eta).$$

Proof: Assume (P) has a feedback controller $u^o(\eta,t)$. Then

$$h^*(t,B^*\eta) = \sup_{z\varepsilon E^m}\{<z,B^*\eta>-h(t,z)\}=$$

$$=<u^o,B^*\eta>-h(t,u^o)$$

$$h(t,u^o)+h^*(t,B^*\eta) = <u^o,B^*\eta> \rightarrow u^o(\eta,t) \ \varepsilon\ \partial h^*(t,B^*\eta)$$

That is, h^* is subdifferentiable at $B^*\eta$. On the other hand if h^* is subdifferentiable at $B^*\eta$ then there is a $u^o\ \varepsilon\ \partial h^*(t,B^*\eta) \rightarrow$ $\rightarrow B^*\eta\ \varepsilon\ \partial h(t,u^o) \rightarrow h^*(t,B^*\eta) = <u^o,B^*\eta>-h(t,u^o)$. Or,

$$\sup_{z\varepsilon E^m}\{<z,B^*\eta>-h(t,z)\}= <u^o,B^*\eta>-h(t,u^o),$$

that is u^o is a feedback controller.

Let the control law $\overline{u}(x,t) = u^o(\eta(x,t),t)$ determine the response $\phi(\overline{u})$ of L) satisfying

$$\dot{x} = A(t)x + B(t)\overline{u}(x,t),\ \phi(\overline{u})(0) = x_0$$

such that $\phi(\overline{u})(t)\ \varepsilon\ G,\ \ \forall\ t\ \varepsilon\ [0,T]$.

Then, the controller $\overline{u}(t) = u^o(\overline{\eta}(t),t)$, $\overline{\eta}(t) = \eta(\phi(\overline{u}(t),t)$ is determined and satisfies

$$\overline{u}(t)\ \varepsilon\ \partial h^*(t,B^*(t)\ \overline{\eta}(t))$$

Thus,

$$h^*(t,z^*) \geq h^*(t,B^*(t)\bar{\eta}(t)) + \langle z^* - B^*(t)\bar{\eta}(t), \bar{u}(t)\rangle, \forall\, z^* \in E^m$$

$$\rightarrow h^*(t,B^*(t)x^*) \geq h^*(t,B^*(t)\bar{\eta}(t)) + \langle B^*(t)x^* - B^*(t)\bar{\eta}(t), \bar{u}(t)\rangle, \forall\, x^* \in E^n$$

$$-\langle B(t)\bar{u}(t), x^*\rangle + h^*(t,B^*(t)x^*) \geq -\langle B(t)\bar{u}(t), \bar{\eta}(t)\rangle + h^*(t,B^*(t)\bar{\eta}(t))$$

$$\rightarrow \langle A(t)\phi(\bar{u})(t), x^*\rangle - \langle \dot{\phi}(\bar{u})(t), x^*\rangle + h^*(t,B^*(t)x^*)$$

$$\geq \langle A(t)\phi(\bar{u}(t), \bar{\eta}(t)\rangle - \langle \dot{\phi}(\bar{u})(t), \bar{\eta}(t)\rangle + h^*(t,B^*(t)\bar{\eta}(t))$$

$$\rightarrow \langle A(t)\phi(\bar{u})(t), x^*\rangle - k(t,\phi(\bar{u})(t)) + h^*(t,B^*(t)x^*)$$

$$\geq \langle A(t)\phi(\bar{u})(t), \bar{\eta}(t)\rangle - k(t,\phi(\bar{u})(t)) + h^*(t,B^*(t)\bar{\eta}(t)) + \langle x^* - \bar{\eta}(t), \dot{\phi}(\bar{u})(t)\rangle$$

That is,

$$H(x^*,\phi(\bar{u})(t),t) \geq H(\bar{\eta}(t),\phi(\bar{u})(t),t) + \langle x^* - \bar{\eta}(t), \dot{\phi}(\bar{u})(t)\rangle, \forall\, x^* \in E^n$$

$$\rightarrow \dot{\phi}(\bar{u})(t) \in \partial_\eta H(\bar{\eta}(t),\phi(\bar{u})(t),t)$$

since H is a convex function in η.

Q.E.D.

CHAPTER IV

MAXIMAL PRINCIPLE AND DUALITY THEORY FOR LINEAR DIFFERENTIAL CONTROL PROCESSES

4.1 Introduction

We propose to study linear optimal control problems with convex cost functions.

In Section 4.2 such a problem (P) is defined and the maximal principle of Pontryagin type is proved in generalized form. In Section 4.3 we construct the adjoint process (P^1) of (P) using the maximal principle and the notions of the C-H-J theory in Chapter III. The adjoint process (P^1) is defined in terms of dual space variables and conjugate functions and has duality relationship to (P).

Upper bounds on the minimum cost are given in terms of solutions to linear partial differential equations associated with the Hamilton-Jacobi equation of problem (P). This is carried out in Section 4.4.

4.2 Linear Differential Control Processes with Convex Cost Functions

Consider the optimal control problem (P) in E^n,

$$(P) \qquad L) \quad \dot{x} = A(t)x + B(t)u, \qquad x(0) = x_0$$

The responses $\phi(u)$ of $L)$ are in $L_p[E^n;0,T]$, $p > 1$, and are constrained to lie in the convex subset X

$X = \{\phi \in L_p[E^n;0,T] \mid \phi(0)=x_o, \phi(t) \in G \subseteq E^n$ on $[0,T]$, G is open convex$\}$

The controllers u are in $L_r[E^m;0,T]$, $r > 1$. The class of admissible controllers U is defined as

$U = \{u \in L_r[E^m;0,T] \mid u(t) \in \Omega \subseteq E^m$ on $[0,T]$, Ω is convex response $\phi(u) \in X\}$

The cost function is

$$C(u) = \int_0^T \{h(t,u(t)) + k(t,x(t))\}dt$$

where

(i) $h(t,z)$ is continuous in t for each fixed z in E^m and for each t, $h(t,z)$ is a l.s.c. proper convex function in z with effective domain Ω.

(ii) $h(t,u(t))$ is summable in t for all $u \in U$ and $h^*(t,u^*(t))$ is summable in t for at least one u^* in $L_s[E^m;0,T]$, $\frac{1}{r} + \frac{1}{s} = 1$

(iii) $k(t,x)$ is continuous in t for each fixed x in E^n and for each t, $k(t,x)$ is a l.s.c. proper convex function in x with effective domain G.

(iv) $k(t,\phi(t))$ is summable in t for all $\phi \in X$ and $k^*(t,\phi^*(t))$ is summable for at least one ϕ^* in $L_q[E^n;0,T]$, $\frac{1}{p} + \frac{1}{q} = 1$

(v) Ω , has non-empty interior.

Furthermore, we assume that for every $x_0 \in G$, $t_0 \in [0,T]$ the function $V(x_0,t_0)$,

$$V(x_0,t_0) = \min_{u \in U} C(u;x_0,t_0) \text{ exists and is finite}$$

and that $V(x_0,t_0)$ is in C^1 on $G \times (0,T)$.

Remark

The hypotheses (i) - (v) guarantee that the integrals

$$\int_0^T h(t,u(t))dt, \int_0^T k(t,\phi(t))dt$$

$$\int_0^T h^*(t,u^*(t))dt, \int_0^T k^*(t,\phi^*(t))dt$$

are well defined for every u, ϕ, u^*, ϕ^*, see Theorem 2.14. Integrals of conjugate functions will be used in the next section.

We have the following theorem which is of Pontryagin type in generalized form.

Theorem 4.1 (Generalized Pontryagin Maximal Principle)

Given the optimal control problem (P). A controller $\bar{u}$ with response $\phi(\bar{u})$ is an optimal controller if and only if there exists an n-vector $\bar{\eta}(t)$ satisfying the adjoint differential equation of L)

$$L^*) \qquad \dot{x}^* = -A^*(t)x^* + \bar{y}^* \quad , \qquad \bar{\eta}(T) = 0$$

for some $\dot{\overline{y}}^*(t) \in \partial k(t,\phi(\overline{u})(t))$ and such that the maximal principle

$$\max_{z\varepsilon\Omega}\{<B(t)z,\overline{\eta}(t)>-h(t,z)\}=<B(t)\overline{u}(t),\overline{\eta}(t)>-h(t,\overline{u}(t))$$

holds a.e. on $[0,T]$.

Proof: For some $t \varepsilon [0,T]$, $x \varepsilon G$ and some controller u with response $\phi(u)$ such that $\phi(u)(t) = x$, denote the corresponding cost on $[t,T]$ by $C(u;x,t,T)$.

For any $x \varepsilon G$ there exists an optimal controller $\overline{u}$ on $[0,T]$ with response $\phi(\overline{u})$ such that $\phi(\overline{u})(0) = x$.

We have on $[t,T]$, $0 \leq t < T$, that

$$C(\overline{u};\phi(\overline{u})(t),t,T)=\int_t^{t+\Delta}\{h(\tau,\overline{u}(\tau))+k(\tau,\phi(\overline{u})(\tau))\}d\dot{t}+C(\overline{u};\phi(\overline{u})(t+\Delta),t+\Delta,T)$$

where $C(\overline{u};\phi(\overline{u})(t+\Delta),t+\Delta,T)$ is the minimal cost on $[t+\Delta,T]$ according the Principle of Optimality.

Consider controllers u on $[t,T]$ with responses $\phi(u)$ such that $\phi(u)(t) = \phi(\overline{u})(t)$ where $u(\tau)$ is arbitrary on $[t,t+\Delta]$ but $\phi(u)(\tau)$ lies in G on $[t,t+\Delta]$ and $C(u;\phi(u)(t+\Delta),t+\Delta,T)$ is minimal on $[t+\Delta,t]$.

For such controllers u, the cost $C(u;\phi(\overline{u})(t),t,T)$ on $[t,T]$ is

$$C(u;\phi(\overline{u})(t),t,T)=\int_t^{t+\Delta} h(\tau,u(\tau))+k(\tau,\phi(u)(\tau))\}d\tau$$
$$+C(u;\phi(u)(t+\Delta),t+\Delta,T).$$

where

$$C(u;\phi(u)(t+\Delta),t+\Delta,T)=V(\phi(u)(t+\Delta),t+\Delta) \text{ is minimal.}$$

We have,

$$C(\bar{u};\phi(\bar{u})(t),t,T) \leq C(u;\phi(\bar{u})(t),t,T)$$

Or,

$$\int_t^{t+\Delta} \{h(\tau,\bar{u}(\tau))+k(\tau,\phi(\bar{u})(\tau))\}d\tau+V(\phi(\bar{u})(t+\Delta),t+\Delta) \tag{4.1}$$

$$\leq\int_t^{t+\Delta} \{h(\tau,u(\tau))+k(\tau,\phi(u)(\tau))\}d\tau+V(\phi(u)(t+\Delta),t+\Delta)$$

By hypothesis, $V(x,t)$ is in C^1 on $G \times (0,T)$ and by Lemma 3.2, $V(x,t)$ is a convex function on G for each t.

To simplify our notation let

$$\eta(\tau) = - \nabla V(\phi(u)(\tau),\tau),$$

and

$$\bar{\eta}(\tau) = - \nabla V(\phi(\bar{u})(\tau),\tau)$$

Then, by convexity,

$$V(\phi(\bar{u})(t+\Delta),t+\Delta)\geq V(\phi(u)(t+\Delta),t+\Delta)+\langle\phi(\bar{u})(t+\Delta)-\phi(u)(t+\Delta),-\eta(t+\Delta)\rangle \tag{4.2}$$

Substitution of (4.2) into (4.1) gives

$$\int_t^{t+\Delta}\{h(\tau,\overline{u}(\tau))+k(\tau,\phi(\overline{u})(\tau))\}d\tau+\langle\phi(\overline{u})(t+\Delta)-\phi(u)(t+\Delta),-\eta(t+\Delta)\rangle$$

$$\leq \int_t^{t+\Delta}\{h(\tau,u(\tau))+k(\tau,\phi(u)(\tau))\}d\tau$$

With,

$$\phi(\overline{u})(t+\Delta) = \phi(\overline{u})(t)+\Delta\dot{\phi}(\overline{u})(t) + o_1(\Delta)$$

$$\phi(u)(t+\Delta) = \phi(\overline{u})(t) + \Delta\dot{\phi}(u)(t) + o_2(\Delta)$$

where

$o_1(\Delta)/\Delta$ and $o_2(\Delta)/\Delta$ go to zero with Δ

$$\int_t^{t+\Delta}\{h(\tau,\overline{u}(\tau))+k(\tau,\phi(\overline{u})(\tau))\}d\tau+\Delta\langle\dot{\phi}(\overline{u})(t)-\phi(u)(t),-\eta(t+\Delta)\rangle+o_1(\Delta)+o_2(\Delta)$$

$$\leq \int_t^{t+\Delta}\{h(\tau,u(\tau))+k(\tau,\phi(u)(\tau))\}d\tau$$

$$\int_t^{t+\Delta}\{h(\tau,\overline{u}(\tau))+k(\tau,\phi(\overline{u})(\tau))\}d\tau+\Delta\langle B(u)\overline{u}(t)-B(t)u(t),-\eta(t+\Delta)\rangle+o_1(\Delta)+o_2(\Delta)$$

$$\leq \int_t^{t+\Delta}\{h(\tau,u(\tau))+k(\tau,\phi(u)(\tau))\}d\tau$$

$$\frac{1}{\Delta}\int_t^{t+\Delta}\{h(\tau,\overline{u}(\tau))+k(\tau,\phi(\overline{u})(\tau))\}d\tau+\langle u(t)-\overline{u}(t),B^*(t)\eta(t+\Delta)\rangle+(o_1(\Delta)+o_2(\Delta))/\Delta$$

$$\leq \frac{1}{\Delta}\int_t^{t+\Delta}\{h(t,u(\tau))+k(\tau,x(\tau))\}dt$$

Thus for $\Delta \to 0$,

$$h(t,\overline{u}(t))+k(t,\phi(\overline{u})(t))+\langle u(t)-\overline{u}(t),B^*(t)\overline{\eta}(t)\rangle$$

$$\leq h(t,u(t))+k(t,\phi(\overline{u})(t)) \text{ almost everywhere.}$$

Or,

$$h(t,\overline{u}(t))+\langle u(t)-\overline{u}(t),B^*(t)\overline{\eta}(t)\rangle\leq h(t,u(t)),\quad \forall\, u(t)\in\Omega,\ \text{a.e.}$$

Thus,

$$B^*(t)\overline{\eta}(t)\in\partial h(t,\overline{u}(t)),\qquad \text{a.e.} \tag{4.3}$$

and

$$\max_{z\in\Omega}\{\langle B(t)z,\overline{\eta}(t)\rangle-h(t,z)\}=\langle B(t)\overline{u}(t),\overline{\eta}(t)\rangle-h(t,\overline{u}(t)),\quad \text{a.e.}$$

We now show that $\overline{\eta}(t) = -\nabla V(\phi(\overline{u})(t),t)$ satisfies the adjoint differential equation

$$\dot{x}^*= -A^*(t)x^*+\overline{y}^*,\ \overline{y}^*(t)\in\partial k(t,\phi(\overline{u})(t)),\ \forall\, t\in[0,T)$$

For $x\in G$, $t\in[0,T)$, let u be the optimal controller on $[t,T]$ with response $\phi(u)$ such that $\phi(u)(t) = x$.

$\overline{y}^*(\tau) = \dot{x}^* + A^*(\tau)x^* \in \partial k(\tau,\phi(\overline{u})(\tau))$ implies

$$k(\tau,\phi(u)(\tau))\geq k(\tau,\phi(\overline{u})(\tau))+\langle\phi(u)(\tau)-\phi(\overline{u})(\tau),A^*(\tau)x^*(\tau)+\dot{x}^*(\tau)\rangle$$

and (4.3) implies

$$h(\tau,u(t)) \geq h(\tau,\overline{u}(\tau))+\langle u(\tau)-\overline{u}(\tau),\ B^*(\tau)\overline{\eta}(\tau)\rangle$$

So that
$$\int_t^{t+\Delta}\{h(\tau,u(\tau))+k(\tau,\phi(u)(\tau))\}d\tau\geq\int_t^{t+\Delta}\{h(\tau,\overline{u}(\tau))+k(\tau,\phi(\overline{u})(\tau))\}d\tau$$
$$+\int_t^{t+\Delta}\langle\phi(u)(\tau)-\phi(\overline{u}(\tau),\dot{x}^*(\tau)\rangle d\tau+\int_t^{t+\Delta}\langle B(\tau)u(\tau)-B(\tau)\overline{u}(\tau),\overline{\eta}(\tau)\rangle dt$$
$$+\int_t^{t+\Delta}\langle A(\tau)\phi(u)(\tau)-A(\tau)\phi(\overline{u})(\tau),x^*(\tau)\rangle d\tau$$

$$\int_t^{t+\Delta}\{h(\tau,u(\tau))+k(\tau,\phi(u)(\tau))\}d\tau \geq \int_t^{t+\Delta}\{h(\tau,\overline{u}(\tau))+k(\tau,\phi(\overline{u})(\tau))\}d\tau$$

$$+\int_t^{t+\Delta}<\phi(u)(\tau)-\phi(\overline{u})(\tau),\dot{x}^*(\tau)>d\tau+\int_t^{t+\Delta}<\dot{\phi}(u)(\tau)-\dot{\phi}(\overline{u})(\tau),x^*(\tau)>d\tau$$

$$+\int_t^{t+\Delta}\{<B(\tau)u(\tau)-B(\tau)\overline{u}(\tau),\overline{\eta}(\tau)-x^*(\tau)>d\tau$$

$$\int_t^{t+\Delta}\{h(\tau,u(\tau))+k(\tau,\phi(u)(\tau))\}d\tau \geq \int_t^{t+\Delta}\{h(\tau,\overline{u}(\tau))+k(\tau,\phi(\overline{u})(\tau))\}d\tau$$

$$+\int_t^{t+\Delta}\frac{d}{d\tau}<\phi(u)(\tau)-\phi(\overline{u})(\tau),x^*(\tau)>+\int_t^{t+\Delta}<B(\tau)u(\tau)-B(\tau)\overline{u}(\tau),\overline{\eta}(\tau)-x^*(\tau)>d\tau$$

Adding $V(\phi(u)(t+\Delta),t+\Delta)$ to both sides of the above inequality and using

$$V(\phi(u)(t+\Delta),t+\Delta)\geq V(\phi(\overline{u})(t+\Delta),t+\Delta)+<\phi(u)(t+\Delta)-\phi(\overline{u})(t+\Delta),-\overline{\eta}(t+\Delta)>$$

$$V(x,t)=V(\phi(u)(t+\Delta),t+\Delta)+\int_t^{t+\Delta}\{h(\tau,u(\tau))+k(\tau,\phi(u)(\tau))\}d\tau$$

$$V(\phi(\overline{u})(t),t)=V(\phi(\overline{u})(t+\Delta),t+\Delta)+\int_t^{t+\Delta}\{h(\tau,\overline{u}(\tau))+k(\tau,\phi(\overline{u})(\tau))\}d\tau$$

we find

$$V(x,t)\geq V(\phi(\overline{u})(t),t)+\int_t^{t+\Delta}\frac{d}{d\tau}<\phi(u)(\tau)-\phi(\overline{u})(\tau),x^*(\tau)>+<\phi(u)(t+\Delta)-\phi(\overline{u})(t+\Delta),$$

$$-\overline{\eta}(t+\Delta)>+\int_t^{t+\Delta}<B(\tau)u(\tau)-B(\tau)\overline{u}(\tau),\overline{\eta}(\tau)-x^*(\tau)>d\tau$$

$$V(x,t) \geq V(\phi(\overline{u})(t),t) + \langle x + \Delta\dot{\phi}(u)(t) - \phi(\overline{u})(t) - \Delta\dot{\phi}(\overline{u})(t), x^*(t) + \Delta\dot{x}^*(t) - \overline{\eta}(t)$$

$$- \dot{\overline{\eta}}(t)\rangle - \langle x - \phi(\overline{u})(t), x^*(t)\rangle + \Delta\langle B(t)u(t) - B(t)\overline{u}(t), \overline{\eta}(t) - x^*(t)\rangle + o(\Delta)$$

where $o(\Delta)/\Delta \to 0$ if $\Delta \to 0$

$$V(x,t) \geq V(\phi(\overline{u})(t),t) + \langle x - \phi(\overline{u})(t), -\overline{\eta}(t)\rangle + \Delta\langle \dot{\phi}(u)(t) - \dot{\phi}(\overline{u})(t), x^*(t) - \overline{\eta}(t)\rangle$$

$$+ \langle x - \phi(\overline{u})(t), \dot{x}^*(t) - \dot{\overline{\eta}}(t)\rangle + \Delta\langle B(t)u(t) - B(t)\overline{u}(t), \eta(t) - x^*(t)\rangle + o(\Delta)$$

$$V(x,t) \geq V(\phi(u)(t),t) + \langle x - \phi(\overline{u})(t), -\overline{\eta}(t)\rangle + \Delta\langle x - \phi(\overline{u})(t), \dot{x}^*(t) - \dot{\overline{\eta}}(t)\rangle$$

$$+ \Delta\langle A(t)x - A(t)\phi(\overline{u})(t), x^*(t) - \overline{\eta}(t)\rangle + o(\Delta)$$

$$V(x,t) \geq V(\phi(\overline{u})(t),t) + \langle x - \phi(\overline{u})(t), -\overline{\eta}(t)\rangle$$

$$+ \Delta\langle x - \phi(\overline{u})(t), \dot{x}^* + A^*(t)x^*(t) - \dot{\overline{\eta}}(t) - A^*(t)\overline{\eta}(t)\rangle + o(\Delta), \forall\, x \in G$$

But $-\overline{\eta}(t) = \nabla V(\phi(\overline{u})(t),t)$, so that we must have

$$0 = \Delta\langle x - \phi(\overline{u})(t), \dot{x}^*(t) + A^*(t)x^*(t) - \dot{\overline{\eta}}(t) - A^*(t)\overline{\eta}(t)\rangle + o(\Delta), \forall\, x \in E^n \tag{4.4}$$

since the last inequality actually holds for all x. Let $\Delta \to 0$ in (4.4) and obtain

$$0 = \langle x - \phi(\overline{u})(t), \dot{x}^*(t) + A^*(t)x^*(t) - \dot{\overline{\eta}}(t) - A^*(t)\overline{\eta}(t)\rangle, \forall\, x \in E^n$$

Hence, $\dot{\bar{\eta}} + A^*(t)\bar{\eta} = \dot{x}^* + A^*(t)x^* = \bar{y}^*$, $t \in [0,T)$

Furthermore, $\bar{\eta}(T) = 0$ since $V(x,T) = 0$, $x \in E^n$

On the other hand, assume that the controller $\bar{u}$ with response $\phi(\bar{u})$, $\phi(\bar{u})(0) = x_0$, is determined by

$$\max_{z\varepsilon\Omega}\{<B(t)z,\bar{\eta}(t)>-h(t,z)\}=<B(t)\bar{u}(t),\bar{\eta}(t)>-h(t,\bar{u}(t)), \quad \text{a.e.}$$

where $\bar{\eta}$ satisfies $\dot{x}^* = -A^*(t)x^* + \bar{y}^*$, $\bar{\eta}(T) = 0$ and $\bar{y}^*(t) \in \partial k(t,\phi(\bar{u})(t))$

These conditions imply $\bar{u}(t) \in \partial h^*(t,B^*(t)\bar{\eta}(t))$ and $k(t,x)$ is subdifferentiable at $\phi(\bar{u})(t) \rightarrow \phi(\bar{u})(t) \in G$, because $\partial k(t,x)$ is empty if $x \notin G$. Thus $\bar{u}(t)$ is an admissible controller. To show that $\bar{u}(t)$ is optimal, let $u(t)$ be any admissible controller with response $\phi(u),\phi(u)(0) = x_0$. Then

$$h(t,u(t))\geq h(t,\bar{u}(t))+<u(t)-\bar{u}(t),B^*(t)\bar{\eta}(t)>$$

$$k(t,\phi(u)(t))\geq k(t,\phi(\bar{u})(t))+<\phi(u)(t)-\phi(\bar{u})(t),\dot{\bar{\eta}}(t)+A^*(t)\bar{\eta}(t)>$$

Thus,

$$\int_0^T\{h(t,u(t))+k(t,\phi(u)(t))\}dt\geq\int_0^T\{h(t,\bar{u}(t))+k(t,\phi(\bar{u})(t))\}dt$$

$$+\int_0^T\{<\dot{\phi}(u)(t)-\dot{\phi}(\bar{u})(t),\bar{\eta}(t)>+<\phi(u)(t)-\phi(\bar{u})(t),\dot{\bar{\eta}}(t)>\}dt$$

$$\int_0^T \{h(t,u(t))+k(t,\phi(u)(t))\}dt \geq \int_0^T \{h(t,\bar{u}(t))+k(t,\phi(\bar{u})(t))\}dt$$

$$+\langle\phi(u)(T)-\phi(\bar{u})(T),0\rangle-\langle x_0-x_0,\bar{\eta}(0)\rangle$$

Thus,

$$\int_0^T \{h(t,u(t))+k(t,\phi(u)(t))\}dt \geq \int_0^T \{h(t,\bar{u}(t))+k(t,\phi(\bar{u})(t))\}dt$$

Hence, $\bar{u}$ is an optimal controller.

Q.E.D.

4.3 The Adjoint Problem

The maximal principle in the form of Theorem 4.1 suggests a process in dual space variables associated with (P) and the adjoint differential equation as equation of evolution. Because the maximal principle constitutes a necessary and sufficient condition for an optimal controller of (P), we are led to the question whether we can find such a process with the role of the maximal principle. This is the idea of duality in extremum problems.

Again, the maximal principle suggests

$$C^*(y^*) = \ell^*(-x^*(T))+\int_0^T \{k^*(t,y^*(t))+h^*(t,B^*(t)x^*(t))\}dt$$

as a cost related to the equation of evolution

$$\dot{x}^* = -A^*(t)x^* + y^*$$

To be more precise, we define the process (P') in E^n.

$$(P') \quad L^*) \quad \dot{x}^* = -A^*(t)x^* + y^*, \quad x^*(T) = 0$$

Given an initial state $x_0 \in G$ of L) we define $\{x_0^*\} = \nabla S(x_0,0)$ as a target set.

Further, the responses $\phi^*(y^*)$ of L^*) are constrained to lie in the (convex) set Σ_t^* on $(0,T)$, where Σ_t^* is the effective domain of the l.s.c. proper convex function $\bar{h}^*(t,x^*) = h^*(t,B^*(t)x^*)$. The class of admissible controllers Y^* consists of all measurable n-vectors y^* on $[0,T]$ such that $y^*(t)$ lies in the (convex) set G_t^*, where G_t^* is the effective domain of the l.s.c. proper convex function $k^*(t,x^*)$ and the corresponding response $\phi^*(y^*)$ satisfies $\phi^*(y^*)(T) = 0$, $\phi(y^*)(0) = x_0^*$ and $\phi^*(y^*)(t) \in \Sigma_t^*$, $\forall\, t \in (0,T)$.

The cost function which must be minimized is

$$C^*(y^*) = \int_0^T \{k^*(t,y^*(t))+h^*(t,B^*(t)\phi^*(y^*)(t))\}dt$$

The problem (P') defined above, will be called the adjoint problem of (P).

Note that all the constraint sets in (P') are defined as effective domains of proper convex functions so that they are non-empty.

It is not surprising that the conjugate function S^* of $S = -V$ turns out to be the minimal cost of (P') provided an optimal controller exists. To show this, the C-H-J theory is of no help because we do not know whether S^* is continuously differentiable on E^{n+1}. Proceed as follows. In problem (P) for every $x \in G$, every $t \in [0,T]$, an optimal controller exists by hypothesis

This property is the key to prove a similar result as Theorem 4.1 for (P'), which in turn will lead to a result which expresses the duality relationship between (P) and (P').

Lemma 4.2

Given the optimal control problem (P). Assume for $x_o \in E^n$, $t_o \in [0,T]$ the corresponding optimal controller $\bar{u}$ is such that $\bar{u}(t) \in \text{int}(\Omega)$. Then the adjoint problem (P') has an optimal controller $\bar{y}^*$ and the minimal cost is given by $C^*(\bar{y}^*) = -S^*(x^*(t_o),t_o)$, where S^* is the conjugate function of $S = -V$.

Proof: For some $x_o \in G$, $t_o \in [0,T]$ let $\bar{u}$ be the optimal controller with response $\phi(\bar{u})$.

Then it follows from Theorem 4.1 that there exists an $\bar{y}^* \in \partial k(t,\phi(\bar{u})(t))$ with response $\phi^*(\bar{y}^*)$ such that $\bar{u}(t) \in \partial h^*(t,B^*(t)\phi^*(\bar{y}^*)(t))$ and $\phi^*(\bar{y}^*)(T) = 0$.

Furthermore, $\bar{y}^*(t) \in \partial k(t,\phi(\bar{u})(t)) \rightarrow \phi(\bar{u})(t) \in \partial k^*(t,\bar{y}^*(t)) \rightarrow$ $\rightarrow \bar{y}^*(t) \in G_t^* \equiv \text{dom}(k^*(t,x^*))$. Also $\bar{u}(t) \in \partial h^*(t,B^*(t)\phi^*(\bar{y}^*)(t)) \rightarrow$ $\rightarrow B^*(t)\phi^*(\bar{y}^*)(t) \in \Sigma_t^* \equiv \text{dom}(h^*(t,B^*(t)x^*)$. We must also have that $\phi^*(\bar{y}^*)(t_o) = \nabla S(x_o,t_o)$. Hence $\bar{y}^*$ is an admissible controller of (P').

We show that $C^*(\bar{y}^*) = -S^*(\phi^*(\bar{y}^*)(t_o),t_o)$. Because G is an open set and $\bar{u}(t) \in \text{int}(\Omega)$, S satisfies the Hamilton-Jacobi equation of (P) on G x [0,T] since Theorem 3.3 applies in this case.

Therefore,

$$\int_{t_o}^{T}\{ \frac{\partial S}{\partial t}(\phi(\overline{u})(t),t)+<\dot{\phi}(\overline{u})(t),\phi^*(\overline{y}^*)(t)>\}dt=\int_{t_o}^{T}\{h(t,\overline{u}(t)+k(t,\phi(\overline{u})(t))\}dt$$

$$=\int_{t_o}^{T}\{<\overline{u}(t),B^*(t)\phi^*(\dot{\overline{y}}^*)(t)>-h^*(t,B^*(t)\phi^*(\overline{y}^*)(t))+<\phi(\overline{u})(t),\overline{y}^*(t)>+ \\ -k^*(t,\overline{y}^*(t))\}dt$$

$$= -\int_{t_o}^{T}\{h^*(t,B^*(t)\phi^*(\overline{y}^*)(t))+k^*(t,\overline{y}^*(t))\}dt+<\phi(\overline{u})(T),\phi^*(\overline{y}^*)(T)> + \\ -<x_o,\phi^*(\overline{y}^*)(t_o)>$$

$$= -\int_{t_o}^{T}\{h^*(t,B^*(t)\phi^*(\overline{y}^*)(t))+k^*(t,\overline{y}^*(t))\}dt+S(\phi(\overline{u})(T),T)+S^*(\phi^*(\overline{y}^*)(T),T) \\ -S(x_o,t_o)-S^*(\phi^*(\overline{y}^*)(t_o),t_o).$$

So that

$$0= -\int_{t_o}^{T}\{h^*(t,B^*(t)\phi^*(\overline{y}^*)(t))+k^*(t,\overline{y}^*(t))\}dt+S^*(\phi^*(\overline{y}^*)(T),T) \\ -S^*(\phi^*(\overline{y}^*)(t_o),t_o)$$

But $S^*(\phi^*(\overline{y}^*)(T),T) = 0$

Hence $C^*(\overline{y}^*) = -S^*(\phi^*(\overline{y}^*)(t_o))$.

It remains to show that $\overline{y}^*$ is optimal. Let y^* be any admissible controller with response $\phi^*(y^*)$. Then $y^*(t) \varepsilon G_t^*$, $\phi^*(y^*)(t) \varepsilon \Sigma_t^*$ and $\phi^*(y^*)(T) = 0$.

Now,

$$-S(x_o,t_o)=C(\bar{u})=\int_{t_o}^{T}\{h(t,\bar{u}(t))+k(t,\phi(\bar{u})(t))\}dt$$

$$=\int_{t_o}^{T}\{\sup_{u^*\varepsilon E^m}[<\bar{u}(t),u^*>-h^*(t,u^*)]\}dt$$

$$+\int_{t_o}^{T}\{\sup_{z^*\varepsilon E^n}[<\phi(\bar{u})(t),z^*>-k^*(t,z^*)]\}dt$$

$$\geq\int_{t_o}^{T}\sup_{v^*\varepsilon E^n}[<\bar{u}(t),B^*(t)v^*>-h^*(t,B^*(t)v^*)]\}dt$$

$$+\int_{t_o}^{T}\{\sup_{z^*\varepsilon E^n}[<\phi(\bar{u})(t),z^*>-k^*(t,z^*)]\}dt$$

$$-S(x_o,t_o)=C(\bar{u})\geq\int_{t_o}^{T}\{<B^*(t)\phi^*(y^*)(t),\bar{u}(t)>-h^*(t,B^*(t)\phi^*(y^*)(t))\}dt$$

$$+\int_{t_o}^{T}\{<\phi(\bar{u})(t),y^*(t)>-k^*(t,y^*(t))\}dt$$

$$=\int_{t_o}^{T}\{<\phi^*(y^*)(t),\dot{\phi}(\bar{u})(t)>+<\dot{\phi}^*(y^*)(t),\phi(\bar{u})(t)>\}dt$$

$$-\int_{t_o}^{T}\{k^*(t,y^*(t))+h^*(t,B^*(t)\phi^*(y^*)(t))\}dt$$

Thus,

$$-S(x_o,t_o)\geq -<x_o,\phi^*(y^*)(t_o)> - C^*(y^*)$$

Or,

$$S^*(\phi^*(y^*)(t_0),t_0) \geq -C^*(y^*)$$

Hence,

$$C^*(\overline{y^*}) = -S^*(\phi^*(y^*)(t_0),t_0) \leq C^*(y^*)$$

Q.E.D.

The lemma is used to prove Pontryagin's generalized maximal principle for problem (P').

Theorem 4.3

Consider the adjoint problem (P') of (P) for given initial state $x(0) = x_0$ in G. Assume the optimal controller $\overline{u}$ of (P) satisfies $\overline{u}(t) \in \text{int}(\Omega)$. Then, a controller $\overline{y}^*$ with response $\phi^*(\overline{y}^*)$ is optimal if and only if there exists an n-vector $\overline{x}$ such that $\overline{x}$ satisfies the differential equation

$$\dot{x} = A(t)x + B(t)\overline{u}(t), \quad \overline{x}(0) = x_0$$

$$\overline{u}(t) \in \partial h^*(t, B^*(t)\phi^*(\overline{y}^*)(t))$$

and

$$\max_{z^* \in G_t^*}\{\langle z^*,\overline{x}(t)\rangle - k^*(t,z^*)\} = \langle\overline{y}^*(t),\overline{x}(t)\rangle - k^*(t,\overline{y}^*(t))$$

a.e. on [0,T].

Proof: Similar to the proof of Theorem 4.1 using Lemma 4.2.

Q.E.D.

We can now give the following result of duality type

Theorem 4.4

Given the optimal control problem (P) and its adjoint problem (P'). Assume in addition that $h^*(t,z^*)$ and $k^*(t,x^*)$ are continuous in t for each fixed z^* and x^*. Then, for every $x_0 \in G$, $t_0 \in [0,T]$, there exists a corresponding optimal controller $\bar{u}(t) \in \text{int}(\Omega)$, an minimal cost $V(x_0,t_0)$ if and only if for every $x_0^* \in \Sigma_0^*$, $t_0 \in [0,T]$ the adjoint problem (P') has an optimal controller $\bar{y}^*$, $\bar{y}^*(t) \in \text{int}(G_t^*)$ and minimal cost $V^*(x_0^*,t_0)$. Moreover,

$$V(x_0,t_0) + V^*(x_0^*,t_0) = - \langle x_0, x^*_0 \rangle$$

Proof: Follows from Theorem 4.1 and Theorem 4.3 and the fact that the negative minimal cost satisfies the Hamilton-Jacobi equation.

Q.E.D.

Remark

In Theorem 4.4, V and V* are not conjugate functions.

Corollary 4.5

$$V(x_0,t_0) = -V^*(x_0^*,t_0) \leftrightarrow \langle x_0, x_0^* \rangle = 0.$$

In which case the adjoint process (P') is called the dual process of (P).

Proof: Immediate from Theorem 4.4.

$$-V^*(x_0^*,t_0) = \max_{y^*} - \int_0^T \{k^*(t,y^*(t)) + h^*(t,B^*(t)x^*(t))\}dt$$

justifies the name dual.

Q.E.D.

4.4 Quasilinearization of the Hamilton-Jacobi Equation of the Control Problem (P)

The Hamilton-Jacobi equation of (P),

$$\forall t \in [0,T],\ \forall x \in G: \frac{\partial S}{\partial t} + \max_{z\in\Omega}\{<Ax+Bz,\nabla S> - h(t,z) - k(t,x)\} = 0 \qquad (4.5)$$

$$S(x,T) = 0, \quad \forall\ x \in G$$

is of first order but nonlinear, because of the max-operation.

We take Bellman's approach, [13], to the quasilinearization of certain classes of nonlinear ordinary or partial differential equations to find a representation for the concave solution S to (4.5) along any admissible response $\phi(u)(t)$ of L) in terms of solutions to a linear partial differential equation, associated with (4.5). We will arrive at our result in a rather straight forward way using Theorem 4.2 and exploiting the convexity of the functions involved.

Drop the max-operation in (4.5) and obtain the linear partial differential equation for W(x,t;u) for some admissible u

$$\forall t \in [0,T],\quad x \in G: \frac{\partial W}{\partial t} + <Ax+Bu,\nabla W> - h(t,u) - k(t,x) = 0 \qquad (4.6)$$

$$W(x,T;u) = 0, \quad \forall\ x \in G$$

Let W(x,t;u) be a solution to (4.6) and suppose that W(x,t;u) is in C^1 on G x (0,T) for every admissible controller u of (P). For every admissible controller u, (4.5) can be written as

$$\frac{\partial S}{\partial t} + \langle Ax+Bu(t),\nabla S\rangle - h(t,u(t)) - k(t,x) = p(u(t)) \tag{4.7}$$

$$p(u(t)) \leq 0, \quad \forall\, t \in [0,T]\ ,\ u \text{ admissible}$$

Let $Z(x,t) = S(x,t) - W(x,t;u)$, then it follows from (4.6) and (4.7) that Z satisfies

$$\forall t \in [0,T], \forall\, x \in G: \frac{\partial Z}{\partial t} + \langle Ax+Bu(t),\nabla Z\rangle = p(u(t)) \leq 0$$

$$Z(x,T) = 0, \quad \forall\, x \in G$$

and Z is in C^1 on G x (0,T) since S and W are.

Let u be an admissible controller of (P) with response $\phi(u)$. Then $\phi(u)(t) \in G$ and the last expression for Z shows that

$$\frac{dZ}{dt}(\phi(u)(t),t) = \frac{\partial Z}{\partial t}(\phi(u)(t),t) + \langle \dot{\phi}(u)(t),\nabla Z(\phi(u)(t),t)\rangle = p(u(t)) \leq 0$$

Thus, $Z(\phi(u)(t),t) \geq 0, \quad \forall\, t \in [0,T]$, or

$$S(\phi(u)(t),t) \geq W(\phi(u)(t),t;u), \forall\, t \in [0,T] \tag{4.8}$$

We can go one step further and show that for an optimal controller $\overline{u}$ of (P) with response $\phi(\overline{u})$ we actually have

$$S(\phi(\overline{u})(t),t) = W(\phi(\overline{u})(t),t;\overline{u}) \tag{4.9}$$

Theorem 4.2 implies that $\overline{\eta}(t) = \nabla S(\phi(\overline{u})(t),t)$ is such that $\overline{u}(t) \in \partial h^*(t,B^*(t)\overline{\eta}(t))$ and $\dot{\overline{\eta}}(t) = -A^*(t)\overline{\eta}(t)+\overline{y}^*(t)$, $\overline{y}^*(t) \in \partial k(t,\phi(\overline{u})(t))$

Now,

$$\overline{u}(t) \in \partial h^*(t,B^*(t)\overline{\eta}(t)) \rightarrow h(t,\overline{u}(t))+h^*(t,B^*(t)\overline{\eta}(t))$$

$$=\langle\overline{u}(t),B^*(t)\overline{\eta}(t)\rangle$$

Thus,

$$\frac{\partial W}{\partial t}(\phi(\overline{u})(t),t;\overline{u})+\langle\dot{\phi}(\overline{u})(t),\nabla W(\phi(\overline{u})(t),t;\overline{u})\rangle=h(t,\overline{u}(t))+k(t,\phi(\overline{u})(t)) \qquad (4.10)$$

Furthermore,

$$\frac{\partial S}{\partial t}+\langle Ax,\nabla S\rangle-k(t,x)+\max_{z\in\Omega}\{\langle z,B^*\nabla S\rangle-h(t,z)\}= 0$$

$$\frac{\partial S}{\partial t}+ \langle Ax,\nabla S\rangle-k(t,x) + h^*(t,B^*\nabla S) = 0$$

implies

$$\frac{\partial S}{\partial t}(\phi(\overline{u})(t),t)+\langle A(t)\phi(\overline{u})(t),\overline{\eta}(t)\rangle-k(t,\phi(\overline{u})(t))+h^*(t,B^*(t)\overline{\eta}(t))=0$$

$$\frac{\partial S}{\partial t}(\phi(\overline{u})(t),t)+\langle A(t)\phi(\overline{u})(t),\overline{\eta}(t)\rangle-k(t,\phi(\overline{u})(t))+\langle\overline{u}(t),B^*(t)\overline{\eta}(t)\rangle -h(t,\overline{u}(t)) = 0$$

$$\frac{\partial S}{\partial t}(\phi(\overline{u})(t),t)+\langle\dot{\phi}(\overline{u})(t),\overline{\eta}(t)\rangle=h(t,\overline{u}(t))+k(t,\phi(\overline{u})(t)) \qquad (4.11)$$

From (4.10) and (4.11) it follows

$$\frac{dW}{dt}(\phi(\bar{u})(t),t;\bar{u}) = \frac{dS}{dt}(\phi(\bar{u})(t),t)$$

Thus, $W(\phi(\bar{u})(t),t;\bar{u}) = S(\phi(\bar{u})(t),t),\ \forall\ t \in [0,T].$

Therefore,

$$S(\phi(\bar{u})(t),t)=\max_{u\in U} W(\phi(u)(t),t;u)=W(\phi(\bar{u})(t),t;\bar{u})$$

where U is the class of admissible controllers for (P) and $\bar{u}$ an optimal controller.

Write $V(x,t) = -S(x,t)$ and $\tilde{W}(x,t;u) = -W(x,t;u)$ so it follows that $\tilde{W}$ satisfies

$$\forall t \in [0,T], \forall\ x \in G: \frac{\partial \tilde{W}}{\partial t} + \langle Ax+Bu, \nabla\tilde{W}\rangle + h(t,u) + k(t,x) = 0 \tag{4.12}$$

$$\tilde{W}(x,T;u) = 0, \quad \forall\ x \in G$$

and

$$V(\phi(u)(t),t) \leq \tilde{W}(\phi(u)(t),t;u), \quad \forall\ u \in U$$

$$V(\phi(\bar{u})(t),t) = \tilde{W}(\phi(\bar{u})(t),t;\bar{u}),$$

where $\bar{u}$ is an optimal controller. Thus the solutions $\tilde{W}(x,t;u)$ to (4.12) constitute upper bounds for the minimum cost $V(x_o,t_o)$.

Finally we would like to remark that these results are of interest from a computational viewpoint. It is possible to construct a monotonic increasing sequence of approximating solutions, converging quadratically to the solution S of the Hamilton-Jacobi equation of the process (P), see [18].

CHAPTER V

DUALITY THEORY FOR LINEAR OPTIMAL CONTROL PROBLEMS

5.1 Introduction

Many optimal control problems can be formulated as convex programming problems in some topological vector space and studied in the light of duality theory for convex programming problems. To motivate our discussion, let us consider the fixed endpoint-minimal effort control problem in its simplest form.

Given the linear process in E^n,

$$L) \qquad \dot{x} = A(t)x + B(t)u , \qquad x(0) = 0$$

with convex target set D in E^n and square summable controllers u in $L_2[E^m;0,T]$ with ranges in E^m and convex constraint set Ω in E^m. The task is to steer the response of L) from 0 into D on [0,T], such that the cost

$$C(u) = ||x(T) - \bar{x}|| + \int_0^T ||u(t)||^2 dt$$

is minimal, where $\bar{x}$ in D is the desired final state.

Define the functions Λ, f, g as

$$\Lambda u = x(T) = \int_0^T \Phi(T,t)\, B(t)\, u(t)dt, \quad \forall\, u \in L_2[E^m;0,T]$$

$$f(u) = \int_0^T ||u(t)||^2 dt \quad , \quad \forall\, u \in L_2[E^m;0,T]$$

$$g(x) = -||x - \bar{x}|| \quad , \quad \forall\, x \in E^n$$

and observe that Λ is a linear bounded transformation, f is a convex function and g is a concave function. If C is the class of admissible controllers, that is the class of all $u \in L_2[E^m;0,T]$ such that u(t) lies in Ω and the corresponding response steers 0 into D, then the optimal control problem in terms of Λ, f and g is

$$\text{(P)} \qquad \text{minimize } C(u) = f(u) - g(\Lambda u) \text{ subject to } u \in C \text{ and } \Lambda u \in D$$

The control problem as defined above is well studied by now, but things become immediately much more complicated, if, for instance we require the responses x to be constrained on the whole interval [0,T].

Further, very little is known in case L) is defined in some infinite dimensional Banach space rather than in E^n cases which arise as natural problem settings for many interesting control processes.

The fact that a wide variety of optimal control problems can be modelled in the form (P), for which an extensive duality theory is available, motivated us to undertake our studies. Section 5.2 contains a duality theory for convex extremum problems of type (P) in a terminology adapted to applications in control. All the results given there were obtained by Rockafellar, see [8]. The more interesting theorems for us are not directly obtainable from the basic concepts and are best presented in their full context. The proofs of the theorems are included for completeness. However, due to our formulation in Banach spaces rather than in topological vector spaces of the original theory, the proofs can sometimes be simplified.

5.2 Optimal Control Problems in Duality

As indicated in the introduction we consider

$$(P) \qquad \text{minimize } C(u) = f(u) - g(\Lambda u) \text{ subject to } u \in C,$$
$$\Lambda u \in D$$

as the model of an abstract, optimal control process (P) with linear equations of evolution L). To keep our notation and terminology transparent we let,

i) u is identified with the controller and an element of a Banach space E

ii) $x = \Lambda u$ is identified with the response of L) corresponding to u and an element of a Banach space F, where Λ is a linear and bounded map defined by L)

iii) C is the constraint set for u and is a convex subset of E

iv) D is the constraint set for the response $x = \Lambda u$ and is a convex subset of F

v) $C(u) = f(u) - g(\Lambda u)$ is the cost function

In more complicated situations, which we shall encounter the terms controller and response are to be taken in a broader sense, which however can be done without confusion in each particular case.

In all what follows we may assume f is a l.s.c. proper convex function with effective domain C and g is a u.s.c. proper concave function with effective domain D. Because if f has not already C as its effective domain we can always replace it by f_0 where,

$$f_0(u) = f(u) \quad , \quad u \in C$$
$$f_0(u) = +\infty \quad , \quad u \notin C$$

and if necessary lower the values of f_0 on the boundary of C so that it becomes l.s.c. Similarly the concave function g can be assumed to be u.s.c. with effective domain D. Therefore the minimand in (P) is a proper convex function or identically $+\infty$.

To define the dual problem (P*) of (P) let E* be a real Banach space, topologically paired with E with respect to a bilinear real valued function $\langle \cdot , \cdot \rangle$ on $E \times E^*$. That is, the elements of each space can be identified with continuous linear functionals on the other by means of $\langle \cdot , \cdot \rangle$. Further let F* be a real Banach space topologically paired to F with respect to $\langle \cdot , \cdot \rangle$. In most of the cases E*, F* are the dual spaces of E, F but there are interesting exceptions.

The dual (P*) of (P) is defined as

$$\text{(P*)} \quad \text{maximize } C^*(x^*) = g^*(x^*) - f^*(\Lambda^* x^*)$$
$$\text{subject to } x^* \in D^*, \Lambda^* x^* \in C^*$$

where f* is the conjugate of f with effective domain C*, g* is the conjugate of g with effective domain D* and Λ^* is the adjoint transformation of Λ. Of course, f^*, g^*, Λ^* are defined with respect to $<\cdot,\cdot>$. Thus the maximand in (P*) is a proper concave function or identically $-\infty$.

From the viewpoint of mathematical programming techniques, (P) and (P*) can be seen as generalized dual models for linear programs. On the other hand, the relationship between (P), (P*) can be explained directly as a generalization of a conjugate function as follows. Let f be a l.s.c. proper convex function with effective domain C. Then, for some fixed $\bar{u} \in C$,

$$f(\bar{u}) = \sup_{u^* \in E^*} \{<\bar{u},u^*> - f^*(u^*)\} =$$

$$= \sup_{u^* \in C^*} \{<\bar{u},u^*> - f^*(u^*)\}$$

where f^* is the unique conjugate of f with effective domain C^*.

Let $g^*(u^*) = <\bar{u},u^*>$, then g^* is an u.s.c. proper concave function with unique conjugate g where,

$$g(u) = \inf_{u^* \in E^*} \{<u,u^*> - g^*(u^*)\} = \inf_{u^* \in E^*} <u-\bar{u},u^*>$$

$$g(u) = \begin{cases} 0 , & u = \bar{u} \\ -\infty, & u \neq \bar{u} \end{cases}$$

Therefore, $f(\bar{u}) = \inf_{u \in C}\{f(u) - g(u)\} = \sup_{u^* \in C^*}\{g^*(u^*) - f^*(u^*)\}$

A generalization of the above is obtained if g is any u.s.c. proper concave function with effective domain D,

$$\inf_{u \varepsilon C} f(u) - g(Iu)$$

$$\sup_{u^* \varepsilon C^*} g^*(Iu^*) - f^*(u^*)$$

where I is the identity operator. Replacing I by any linear bounded operator Λ gives rise to

$$(P) \quad \inf f(u) - g(\Lambda u) \quad , u \varepsilon C, \; \Lambda u \varepsilon D.$$

$$(P^*) \quad \sup g^*(x^*) - f^*(\Lambda^* x^*), x^* \varepsilon D^*, \Lambda^* x^* \varepsilon C^*$$

Note that the minimization in (P) can be carried out over all of E (unconstrained problem) because of the fact that C and D are the effective domains of f and g respectively. For the same reason the maximization in (P*) can be taken over all of F*.

We adopt the following terminology.

A controller $\bar{u}$ is a solution to (P) if and only if $\bar{u} \varepsilon C$, $\Lambda \bar{u} \varepsilon D$ and the infimum of C(u) is finite and attained at $\bar{u}$. A solution $\bar{u}$ to (P) will be called an optimal controller.

The response constraint set D is said to be reachable if and only if there exists a $u \varepsilon C$ such that $\Lambda u \varepsilon D$.

Similarly $\bar{x}^*$ is a solution to (P*) if and only if $\bar{x}^* \varepsilon D^*$,

$\Lambda^*\bar{x}^* \in C^*$ and supremum $C^*(x^*)$ is finite and attained at $\bar{x}^*$.

Further, in case inf $C(u)$ or sup $C^*(x^*)$ is attained at some point we shall write min $C(u)$ or max $C^*(x^*)$.

Before presenting the duality theory for (P) and (P*) we summarize:

Primal: (P) minimize $C(u) = f(u) - g(\Lambda u)$ subject to $u \in C$

$\Lambda u \in D$

where,

f is l.s.c. convex with effective domain C, g is u.s.c. concave with effective domain D and Λ is a linear bounded transformation from a Banach space E to a Banach space F.

Dual: (P*) maximize $C^*(x^*) = g^*(x^*) - f^*(\Lambda^* x^*)$ subject to $x^* \in D^*$,

$\Lambda^* x^* \in C^*$

where,

f^* is the conjugate function of f with effective domain C^*, g^* is the conjugate function of g with effective domain D^* and Λ^* is the adjoint transformation of Λ from the Banach space F^* to the Banach space E^*.

Lemma 5.1

$$\inf C(u) \geq \sup C^*(x^*)$$

Proof: Let $u \in E$ and $x^* \in F^*$. The inequality $f(u) - g(\Lambda u) \geq g^*(x^*) - f^*(\Lambda^* x^*)$ holds trivially if the right hand side is $-\infty$ or if the left hand side is $+\infty$.

Assume that both sides are finite. Then necessarily $u \in C$,

$\Lambda u \in D$, $x^* \in D^*$ and $\Lambda^* x^* \in C^*$. From the definition of a conjugate function, it follows that

$$f(u) + f^*(\Lambda^* x^*) \geq \langle u, \Lambda^* x^* \rangle$$

$$g(\Lambda u) + g^*(x^*) \leq \langle \Lambda u, x^* \rangle = \langle u, \Lambda^* x^* \rangle$$

Thus, $f(u) - g(\Lambda u) \geq g^*(x^*) - f^*(\Lambda^* x^*)$ for arbitrary u, x^*.
Hence,

$$\inf C(u) \geq \sup C^*(x^*)$$

Q.E.D.

In duality theory the concept of stability is very important, where a stably set process (P) is defined as follows.

Consider the perturbed process (P(z)) for some $z \in F$ where

$$(P(z)) \quad \text{minimize } C(u,z) = f(u) - g(\Lambda u - z),\ u \in C,\ \Lambda u \in D.$$

Definition 5.2

If $\inf C(u,0) = \inf C(u)$ is finite then the process (P) is said to be stably set if in some neighborhood N of the origin in F,

$$\lim_{\varepsilon \downarrow 0} \frac{\inf C(u,\varepsilon z) - \inf C(u)}{\varepsilon} > -\infty,\ \forall z \in N.$$

We shall adopt the convention that in case $\inf C(u)$ is $-\infty(+\infty)$ we say that (P) is stably set (unstably set).

We must verify that the limit of the definition is well defined for every z. This will follow from the lemma below where it is proved that the function p defined as $p(z) = \inf_u C(u,z)$ is a convex function on F, so that its one-sided directional derivative at 0,

$$p'(0;z) = \lim_{\varepsilon \downarrow 0} \frac{\inf C(u,\varepsilon z) - \inf C(u)}{\varepsilon}$$

exists for all z although it may be infinite.

Lemma 5.3

The function p defined by $p(z) = \inf_u C(u,z)$ is a convex function on F.

Proof: To prove that p is convex, we show that its epigraph, epi(p) is a convex set in $F \oplus R$. Let (z_1,a_1) and (z_2,a_2) be two points in epi(p), that is

$$p(z_1) \leq a_1 < +\infty \ , \ p(z_2) \leq a_2 < +\infty$$

Thus, $\inf f(u) - g(\Lambda u - z_1) \leq a_1 \rightarrow \exists \ u_1$ such that

$$f(u_1) - g(\Lambda u_1 - z_1) \leq a_1$$

Let $a_1 - \{f(u_1) - g(\Lambda u_1 - z_1)\} = \delta_1 \geq 0$ and

$$a_{11} = f(u_1) + \frac{\delta}{2} , \qquad a_{12} = g(\Lambda u_1 - z_1) - \frac{\delta}{2}$$

Then, $\quad f(u_1) \leq a_{11}, \; g(\Lambda u_1 - z_1) \geq a_{12}, \;\; a_{11} - a_{12} = a_1$

Similarly we can find numbers a_{21}, a_{22} and a $u_2 \varepsilon E$ such that

$$f(u_2) \leq a_{21}, \; g(\Lambda u_2 - z_2) \geq a_{22}, \;\; a_{21} - a_{22} = a_2$$

Since f is a proper convex function,

$$f(\lambda u_1 + (1-\lambda)u_2) \leq \lambda a_{11} + (1-\lambda)a_{21}, \; 0 \leq \lambda \leq 1$$

and by the concavity of g and the linearity of Λ,

$$g(\Lambda(\lambda u_1 + (1-\lambda)u_2) - (\lambda z_1 + (1-\lambda)z_2)) \geq \lambda a_{12} + (1-\lambda)a_{22}.$$

Thus for $u = \lambda u_1 + (1-\lambda)u_2$,

$$\begin{aligned} p(\lambda z_1 + (1-\lambda)z_2) &\leq f(u) - g(\Lambda u - (\lambda z_1 + (1-\lambda)z_2)) \\ &\leq (\lambda a_{11} + (1-\lambda)a_{21}) - (\lambda a_{12} + (1-\lambda)a_{22}) \\ &= \lambda a_1 + (1-\lambda)a_2 \end{aligned}$$

This shows that epi(p) is a convex set.

Q.E.D.

We have the following sufficient condition for the stability of (P) in which case $\inf_u C(u,z)$ is a smooth function in some neighborhood of the origin.

Theorem 5.4

If for the linear process L), D is reachable for some $u \varepsilon C$ such that $\Lambda u \varepsilon$ int D then (P) is stably set and $\inf_u C(u,z)$ is a continuous function in some neighborhood of the origin.

Proof: Let $p(z) = \inf_u C(u,z)$, $z \varepsilon F$. The effective domain of p is,

$$\text{dom}\,(p) = \{z | p(z) < +\infty\}$$

$$= \{\Lambda u-\omega | f(u) < +\infty,\ g(\omega) > -\infty\} = \Lambda(C)-D.$$

Since F is a Banach space we know that p is continuous on int dom(p). Therefore what we want to show is 0ε int dom(p). By hypothesis $\Lambda u \varepsilon$ int D for some $u \varepsilon C$ what implies

$$0 \varepsilon \Lambda(C) - \text{int}\, D \subseteq \text{int}(\Lambda(C)-D) = \text{int dom}(p)$$

It remains to show that this fact implies stability. Stability would be by definition if $p(0) = \inf C(u) = -\infty$ and since 0ε int dom(p) we have that $p(0) \neq +\infty$. Thus we may assume that p(0) is finite. Then (P) is stably set if and only if the directional derivative function $p'(0;z)$ is bounded below in z in some neighborhood of z = 0. But for the convex function p it is known that

$$p(z) \geq p(0) + p'(0;z) \quad , \quad \forall z$$

$$p'(0;z) \geq -p'(0;-z) \quad , \quad \forall z$$

Since p is continuous at $z = 0$, the first inequality implies $p'(0;z)$ is bounded above in z in some neighborhood of $z = 0$. The conclusion we want follows from the second inequality.

Q.E.D.

<u>Example</u>

(P) minimize $f(u) - g(\Lambda u)$, $u \varepsilon C$, $\Lambda u \varepsilon D$.

If $0 \varepsilon C$ and $0 \varepsilon \text{int}(D)$ then (P) is stably set.

<u>Example</u>

(P) minimize $f(u) - g(\Lambda u)$, $u \varepsilon C$, g is finite everywhere (no constraints on the responses). If there is a $u \varepsilon C$, then (P) is stably set.

The following theorems give the relationship between (P) and its dual (P*).

<u>Theorem 5.5</u>

The process (P) is stably set if and only if

$$\inf C(u) = \max C^*(x^*)$$

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

$$\min C(u) = \sup C^*(x^*)$$

<u>Proof:</u> Let $p(z) = \inf_u C(u,z)$ as before.

We first show that (P) is stably set if and only if p is subdifferentiable at 0, that is $\partial p(0)$ is non-empty. Assume (P) is stably set then $p(0) = -\infty$ or $p(0)$ is finite. If $p(0) = -\infty$ then

$\partial p(0) = F^*$ and if $p(0)$ is finite, $p'(0;z)$ is defined for all z and bounded below in z in some neighborhood of 0, since (P) is stably set. But this implies that p is subdifferentiable at 0. On the other hand if p is subdifferentiable at 0 then $p(0) = -\infty$ or $p(0)$ is finite. If $p(0) = -\infty$, (P) is stably set by definition and in case $p(0)$ is finite then $p'(0;z)$ exists and is bounded below in z in some neighborhood of 0, what implies that (P) is stably set.

Thus, (P) is stably set $\leftrightarrow$ $\partial p(0)$ is non-empty.

If $p(0) = -\infty$, then by Lemma 5.1

$$p(0) = \inf C(u) = -\infty \geq \sup C^*(x^*) \qquad \text{so that}$$

$$\inf C(u) = C^*(x^*), \forall x^*, \text{ or } \inf C(u) = \max C^*(x^*).$$

It remains to show that for $p(0)$ finite,

$$x^* \in \partial p(0) \leftrightarrow \inf C(u) = \max C^*(x^*)$$

In view of Lemma 5.1 it suffices to show that

$$x^* \in \partial p(0) \leftrightarrow \inf C(u) = C^*(x^*) \qquad (5.1)$$

By definition,

$$x^* \in \partial p(0) \leftrightarrow p(z) \geq p(0) + \langle z,x^* \rangle \ , \ \forall z$$

Or,

$$p(0) \leq \inf_{z} \{p(z) - \langle z,x^* \rangle\}$$

$$= \inf_{z} \inf_{u} \{f(u) - g(\Lambda u-z) - \langle z,x^* \rangle\}$$

$$p(0) \leq \inf_{u} \inf_{z} \{\langle \Lambda u-z,x^* \rangle - g(\Lambda u-z) - \langle \Lambda u,x^* \rangle + f(u)\} \quad (5.2)$$

Since $p(0) = \inf_{u} C(u)$ is finite

$$-\infty < \inf_{z} \{\langle \Lambda u-z,x^* \rangle - g(\Lambda u-z)\} = g^*(x^*) < +\infty ,$$

for $u \varepsilon C$ and thus for all $u \varepsilon E$.

From (5.2),

$$\inf_{u} C(u) \leq \inf\{g^*(x^*) - \langle \Lambda u,x^* \rangle + f(u)\}$$

$$= g^*(x^*) - \sup_{u} \{\langle u,\Lambda^* x^* \rangle - f(u)\}$$

$$= g^*(x^*) - f^*(\Lambda^* x^*) = C^*(x^*)$$

which is (5.1).

This proves the first part of the theorem. The second part follows dually.

Q.E.D.

<u>Theorem 5.6</u>

If (P) and (P*) are both stably set then both have solutions and

$$-\infty < \min C(u) = \max C^*(x^*) < +\infty$$

<u>Proof</u>: Theorem 5.5 implies min C(u) = max C*(x*) if (P) and (P*) are both stably set. But the minimand in (P) never has the value $-\infty$, so that an infimum of $-\infty$ cannot be attained at some u, as the use of "min" is meant to indicate.

Thus min C(u) > $-\infty$ and dually max C*(x*) < $+\infty$.

Q.E.D.

<u>Theorem 5.7</u>

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

<u>Proof</u>: By Theorem 5.6, (P) is stably set and has a solution if and only if min C(u) = max C*(x*) if and only if (P*) is stably set and has a solution.

Q.E.D.

Optimal controllers, that is solutions to (P), can be characterized by certain subdifferentiability conditions on f and g in analogy with the ordinary case in the calculus when f and g are differentiable. To get a better insight into the nature of the next theorem note that the convex function $f - g \circ \Lambda$ attains a finite minimum in precisely those points $\bar{u}$ where 0 is a subgradient. Thus optimal controllers $\bar{u}$ satisfy

$$0 \in \partial(f - g \circ \Lambda)(\bar{u})$$

We will prove in a moment that,

$$0 \in \partial f(u) - \Lambda^*(\partial g(\Lambda u)) \leftrightarrow \Lambda u \in \partial g^*(x^*), \Lambda^* x^* \in \partial f(u) \quad (5.3)$$

for some $x^* \in F^*$.

Provided,

$$0 \in \partial(f - g \circ \Lambda)(u) \leftrightarrow 0 \in \partial f(u) - \Lambda^*(\partial g(\Lambda u)) \quad (5.4)$$

holds, the right hand side of (5.3) constitutes a convenient form for characterizing optimal controllers. Indeed, (5.4) is true under the stability condition, see the next theorem.

To verify (5.3) assume that $0 \in \partial f(u) - \Lambda^*(\partial g(\Lambda u)$ for some $u \in E$. Then there exists a $u^* \in \partial f(u)$ and $x^* \in \partial g(\Lambda u)$ such that $u^* = \Lambda^* x^*$, and $x^* \in \partial g(\Lambda u) \rightarrow \Lambda u \in \partial g^*(x^*)$. On the other hand if u, x^* satisfy the right hand side of (5.3) then $x^* \in \partial g(\Lambda u) \rightarrow \Lambda^* x^* \in \Lambda^*(\partial g(\Lambda u))$ so that $0 \in \partial f(u) - \Lambda^*(\partial g(\Lambda u))$. This proves (5.3).

<u>Theorem 5.8</u>

The process (P) and its dual (P*) are stably set, with solutions $\bar{u}$, $\bar{x}^*$ respectively, if and only if $\bar{u}$, $\bar{x}^*$ satisfy $\Lambda\bar{u} \in \partial g^*(\bar{x}^*)$ and $\Lambda^*\bar{x}^* \in \partial f(\bar{u})$.

<u>Proof</u>: Suppose (P), (P*) are stably set. By Theorem 5.6, (P) and (P*) have both solutions, $\bar{u}$, $\bar{x}^*$ and

$$f(\bar{u}) - g(\Lambda\bar{u}) = g^*(\bar{x}^*) - f^*(\Lambda^*\bar{x}^*) \tag{5.5}$$

Further, $f(\bar{u}) + f^*(\Lambda^*\bar{x}^*) \geq \langle\bar{u},\Lambda^*\bar{x}^*\rangle$

$$g(\Lambda\bar{u}) + g^*(\bar{x}^*) \leq \langle\Lambda\bar{u},\bar{x}^*\rangle = \langle\bar{u},\Lambda^*\bar{x}^*\rangle .$$

So that in view of (5.5),

$$\langle\bar{u},\Lambda^*\bar{x}^*\rangle = f(\bar{u}) + f^*(\Lambda^*\bar{x}^*) = g(\Lambda\bar{u})+g^*(\bar{x}^*)=\langle\Lambda\bar{u},\bar{x}^*\rangle$$

what implies $\Lambda\bar{u} \in \partial g^*(\bar{x}^*)$, $\Lambda^*\bar{x}^* \in \partial f(\bar{u})$.

On the other hand suppose $\bar{u}$, $\bar{x}^*$ be such that

$$\Lambda\bar{u} \in \partial g^*(\bar{x}) \quad \text{and} \quad \Lambda^*\bar{x}^* \in \partial f(\bar{u})$$

Then,

$$f(\bar{u}) + f^*(\Lambda^*\bar{x}^*) = \langle\bar{u},\Lambda^*\bar{x}^*\rangle$$

$$g(\Lambda\bar{u}) + g^*(\bar{x}^*) = \langle\Lambda\bar{u},\bar{x}^*\rangle = \langle\bar{u},\Lambda^*\bar{x}^*\rangle$$

Thus, $\bar{u}$ is a solution to (P) and $\bar{x}^*$ is a solution to (P*) by Lemma 5.1. Furthermore, min $C(u)$ = max $C^*(x^*)$ so that (P) and (P*) are both stably set by Theorem 5.5.

Q.E.D.

<u>Corollary 5.8</u>

If the process (P) is stably set, then $\bar{u}$ is a solution to (P) if and only if there exists an $\overline{x}^*$ such that $\Lambda\overline{u} \in \partial g^*(\overline{x}^*)$ and $\Lambda^*\overline{x}^* \in \partial f(\overline{u})$.

<u>Proof</u>: Immediate from Theorems 5.7 and 5.8.

Q.E.D.

CHAPTER VI

LINEAR DIFFERENTIAL CONTROL PROCESSES

6.1 Introduction

The duality theory of Chapter V is applied to a class of linear differential processes with convex cost functions. In Section 6.2 the parameters f, g, Λ, f*, g*, Λ* of the model (P) are determined explicitly. An existence and uniqueness theorem for optimal controllers is given under weaker conditions than appears in the literature. After introduction of the adjoint differential equation of the process, we arrive at Pontryagin's maximal principle in generalized form and in so doing reveal the existence of an intimate relationship between duality and maximal principle in optimal control. Section 6.3 deals with control processes where the initial state is allowed to vary over some convex set. The results obtained follow without difficulties from those in Section 6.2.

6.2 Linear Optimal Control with Convex Cost and Constrained Responses

Consider the linear control process in E^n,

$$(P)\quad L)\quad \dot{x} = A(t)x + B(u)\, u, \qquad x(0) = 0$$

where A(t), B(t) are continuous matrices on [0,T].

The responses $\phi(u)$ of L) are in $L_p[E^n;0,T]$, $p > 1$, and are required to lie in the convex subset X of $L_p[E^n;0,T]$ where,

$$X = \{\phi \in L_p[E^n;0,T] | \phi(0)=0, \phi(t) \in G_t \subseteq E^n \text{ a.e. on } (0,T], G_t \text{ is convex}\}$$

The controllers u are in $L_r[E^m;0,T]$, $r > 1$, and the class of admissible controllers U is a convex subset of $L_r[E^m;0,T]$ where,

$$U = \{u \in L_r[E^m;0,T] | u(t) \in \Omega \subseteq E^m, \text{a.e. on } [0,T], \Omega \text{ is convex, response } \phi(u) \in X\}$$

The cost function is

$$C(u) = \ell(\phi(u)(T)) + \int_0^T \{h(t,u(t)) + k(t,\phi(u)(t))\}dt$$

Hypothesis:

(i) $\ell(y)$ is a l.s.c. proper convex function on E^n with effective domain G_T

(ii) $h(t,z)$ is measurable in t for each fixed $z \in E^m$ and for each t, $h(t,z)$ is a l.s.c. convex function in z with effective domain Ω

(iii) $h(t,u(t))$ is summable in t for all $u \in U$ and $h^*(t,u^*(t))$ is summable in t for at least one u^* in $L_s[E^m;0,T]$, $\frac{1}{r} + \frac{1}{s} = 1$

(iv) $k(t,y)$ is measurable in t for each fixed $y \in E^n$ and for each t, $k(t,y)$ is a l.s.c. convex function in $\tilde{y}$ with effective domain G_t

(v) $k(t,\phi(t))$ is summable in t for all $\phi \in X$ and $k^*(t,\phi^*(t))$ is summable in t for at least one ϕ^* in $L_q[E^n;0,T]$, $\frac{1}{p}+\frac{1}{q}=1$

(vi) Ω, G_t have non-empty interiors, $\forall\, t \in (0,T]$

To formulate this optimal control problem in the form of the model in Chapter V we define,

$$f(u) = \int_0^T h(t,u(t))dt \quad , \quad \forall\, u \in L_r[E^m;0,T]$$

By hypotheses (ii), (vi) h is a normal convex integrand and by hypothesis (iii) f is a proper convex function with effective domain U. Furthermore f and f* are conjugate functions where,

$$f^*(u^*) = \int_0^T h^*(t,u^*(t))dt, \; \forall\, u^* \in L_s[E^m;0,T] \qquad (6.1)$$

The function f* is proper by hypothesis (iii) and it is automatically l.s.c. as the conjugate of the proper convex function f, which in turn implies that f is also l.s.c. Similarly define g, g* where

$$g((\phi,y))=-\ell(y)-\int_0^T k(t,\phi(t))dt, \forall(\phi,y) \in L_p[E^n;0,T]\oplus E^n$$

$$g^*((\phi^*,y^*))=-\ell^*(-y^*)-\int_0^T k^*(t,-\phi^*(t))dt, \forall(\phi^*,y^*)\in L_q[E^n;0,T]\oplus E^n \qquad (6.2)$$

are u.s.c. proper concave functions as it follows from hypotheses (iv), (v), (vi). The effective domain of g is evidently $X \oplus G_T$.

If $\phi(u)$ is the response of $L)$ for some u then $\phi(u)$ is defined by

$$\phi(u)(t)=\int_0^t \Phi(t,\tau)B(\tau)u(\tau)d \quad , \quad t \in [0,T] \qquad (6.3)$$

where Φ is the fundamental matrix of $L)$. Define the linear bounded transformation Λ from $L_r[E^m;0,T]$ to $L_p[E^n;0,T] \oplus E^n$ by

$$u \overset{\Lambda}{\rightarrow} (\phi(u),\phi(u)(T)) \text{ and } \phi(u) \text{ defined by } (6.3)$$

In terms of f, g, Λ, the optimal control problem (P) is

(P) minimize $f(u) - g(\Lambda u)$ subject to $u \in U$
and $\Lambda u \in X \oplus G_T$

which is precisely the form of the model in Chapter V. The dual problem (P*) of (P) is

(P*) maximize $g^*((\phi^*,y^*)) - f^*(\Lambda^*(\phi^*,y^*))$ subject to
$(\phi^*,y^*) \in X^* \oplus G_T^*$ and $\Lambda^*((\phi^*,y^*)) \in U^*$

where $X^* \oplus G_T^*$, U^* are the effective domains of g^*, f^*, given by (6.2), (6.1), respectively.

It remains to calculate the linear bounded adjoint transformation Λ^* of Λ.

We have,

$$\forall\ (\phi^*,y^*) \in L_q[E^n;0,T]\oplus E^n : <\Lambda^*(\phi^*,y^*),u> = <(\phi^*,y^*),\Lambda u>,$$

$$\forall\ u \in L_r[E^m;0,T].$$

Thus,

$$<\Lambda u,(\phi^*,y^*)> = \int_0^T <\int_0^t \Phi(t,\tau)B(\tau)u(\tau)dt,\phi^*(t)>_{E^n} dt + <\int_0^T \Phi(T,\tau)B(\tau)u(\tau),y^*>dt_{E^n}$$

$$= \int_0^T\int_0^t <\Phi(t,\tau)B(\tau)u(\tau),\phi^*(t)>_{E^n} d\tau dt + \int_0^T <\Phi(T,\tau)B(\tau)u(\tau),y^*>_{E^n} dt =$$

$$= \int_0^T\int_\tau^T <\Phi(t,\tau)B(\tau)u(\tau),\phi^*(t)>_{E^n} dt d\tau + \int_0^T <\Phi(T,\tau)B(\tau)u(\tau),y^*>_{E^n} dt =$$

$$= \int_0^T <u(\tau),\int_\tau^T B^*(\tau)\Phi^*(t,\tau)\phi^*(t)dt>_{E^n} d\tau + \int_0^T <u(\tau),B^*(\tau)\Phi^*(T,\tau)y^*>_{E^n} dt$$

Hence, $\Lambda^*(\phi^*,y^*)$ is given by

$$\Lambda^*(\phi^*,y^*)(\tau) = B^*(\tau)\{\int_\tau^T \Phi^*(t,\tau)\phi^*(t)dt + \Phi^*(T,\tau)y^*\},\ \tau \in [0,T] \qquad (6.4)$$

where B^*, Φ^* are the transposed matrices of B and Φ respectively.

Thus the duality theory in Chapter V applies to the present case with f, f^*, g, g^*, Λ, Λ^* as defined above and $C = U$, $D = X \oplus G_T$ and $C^* = U^*$, $D^* = X^* \oplus G_T^*$.

We shall now show how the duality theory gives us an existence theorem for optimal control under weaker conditions than those reported in the literature.

Theorem 6.1

Given the optimal control problem (P). Assume in addition that (i) Ω is bounded, (ii) for each $t \varepsilon [0,T]$, $h(t,z) \geq H(t)$ for all $z \varepsilon E^m$ where $H(t)$ is summable on $[0,T]$, (iii) for the process L), $X \oplus G_T$ is reachable. Then there exists a controller $\overline{u} \varepsilon U$ such that $\min_{u \varepsilon U} C(u) = C(\overline{u})$. Moreover if f is strictly convex, then $\overline{u}$ is unique.

Proof: For arbitrary $u^* \varepsilon L_s[E^m;0,T]$,

$$f^*(u^*)=\int_0^T h^*(t,u^*(t))dt=\int_0^T [\sup_{z\varepsilon E^m}\{<z,u^*(t)>-h(t,z)\}]\ dt =$$

$$=\int_0^T [\sup_{z\varepsilon\Omega}\{<z,u^*(t)>-h(t,z)\}]dt\leq M\int_0^T ||u^*(t)||dt-\int_0^T H(t)dt < +\infty$$

for some positive constant M.

That is, the effective domain of f^* is all of $L_s[E^m;0,T]$. The function g^* is proper, thus there exists a $(\overline{\emptyset}^*,\overline{y}^*)$ at which g^* is finite and $\Lambda^*(\overline{\emptyset}^*,\overline{y}^*)$ is evidently in the interior of dom(f^*). From the dual version of Theorem 5.4 it follows that (P^*) is stably set.

From Theorem 5.5 it follows that there exists a $\overline{u} \varepsilon L_r[E^m;0,T]$ such that the inf $C(u)$ is attained at $\overline{u}$, which we write as,

$$\min_{u \in L_r[E^m;0,T]} C(u) = C(\bar{u}) \tag{6.5}$$

It remains to show that $\bar{u}$ is admissible, that is $\bar{u} \in U$.

$X \oplus G_T$ is reachable by hypothesis which implies that there is an admissible controller $u \in U$ with response $\phi(u)$ such that $\phi(u)(0) = 0$, $\phi(u)(t) \in G_t$ a.e. on $[0,T]$. In other words $u \in \mathrm{dom}(f)$ and $(\phi(u),\phi(u)(T)) \in X \oplus G_T = \mathrm{dom}(g)$. Thus, for this controller u, $f(u) - g(\Lambda u)$ is finite and therefore,

$$\min_{u} C(u) < +\infty \tag{6.6}$$

Furthermore, $g^*(\bar{\phi}^*,\bar{y}^*) - f^*(\Lambda(\bar{\phi}^*,\bar{y}^*))$ is finite so that

$$\sup_{(\phi^*,y^*)} C^*(\phi^*,y^*) \equiv \sup_{(\phi^*,y^*)} g^*(\phi^*,y^*)-f^*(\Lambda^*(\phi^*,y^*)) > -\infty \tag{6.7}$$

From Lemma 5.1, (6.6) and (6.7) it follows

$$+\infty > \min_{u} C(u) \geq \sup_{(\phi^*,y^*)} C^*(\phi^*,y^*) > -\infty \tag{6.8}$$

Hence, from (6.5) and (6.8)

$$\min_{u} C(u) = C(\bar{u}) \quad \text{is finite}$$

and thus necessarily $\bar{u} \in U$, that is $\bar{u}$ is admissible.

The uniqueness of $\bar{u}$ in case f is strictly convex follows trivially from a contradiction argument.

Q.E.D.

Now we will demonstrate the relationship between duality and maximal principle for the optimal control problem (P).

Define the Hamiltonian $H(\eta,x,u)$ as,

$$H(\eta,x,z) = \langle Ax + Bz,\eta\rangle - h(t,z) - k(t,x)$$

Assuming that k is differentiable with respect to x, the adjoint response $\eta(t)$ according a controller u with response $\phi(u)$ is the solution to the Hamiltonian differential system.

$$\dot{\phi}(u)(t) = \frac{\partial H}{\partial \eta}(\eta(t),\phi(u)(t),u(t)) = A(t)\phi(u)(t) + B(t)u(t)$$

$$\dot{\eta}(t) = -\frac{\partial H}{\partial x}(\eta(t),\phi(u)(t),u(t)) = -A^*(t)\eta(t) - \frac{\partial k}{\partial x}(t,\phi(u)(t))$$

The expression between the brackets in equation (6.4) satisfies the above equations if $\phi^*(t) = \frac{\partial k}{\partial x}(t,\phi(u)(t))$.

Indeed, equation (6.4) which appears in the dual problem (P*) is the connection between duality and Hamiltonian approach of the problem on hand.

In case we only require k to be subdifferentiable with respect to x rather than differentiable the situation is somewhat more complicated and we must write

$$\dot{\eta}(t) = -A^*(t)\eta(t) + x^*(t,\phi(u)(t)), x^*(t,\phi(u)(t)) \in \partial k(t,\phi(u)(t))$$

rather than $\dot{\eta}(t) = -A^*(t)\eta(t) - \frac{\partial k}{\partial x}(t,\phi(u)(t))$.

Thus we arrive at the following.

Theorem 6.2 (Generalized Pontryagin Maximal Principle)

Given the optimal control problem (P). Assume there exists an admissible controller $u \in U$ with response $\phi(u)$ of $L)$ such that $(\phi(u),\phi(u)(T))$ is in the interior of the constraint set $X \oplus G_T$. Then, a controller $\bar{u}$ with response $\phi(\bar{u})$ of $L)$ is an optimal controller if and only if there exists a vector $\bar{\eta} \in L_q[E^n;0,T]$ satisfying

$$\dot{\eta} = -A^*(t)\eta + x^*(t,\phi(\bar{u})(t)), -\bar{\eta}(T) \in \partial\, \ell(\phi(\bar{u})(T))$$

where $x^*(t,\phi(\bar{u}(t)) \in \partial k(t,\phi(\bar{u})(t))$ almost everywhere on $[0,T]$ and such that the maximal principle

$$\langle B(t)\bar{u}(t),\bar{\eta}(t)\rangle - h(t,\bar{u}(t)) = \max_{z \in \Omega} [\langle B(t)z,\bar{\eta}(t)\rangle - h(t,z)]$$

holds almost everywhere on $[0,T]$.

Proof: By hypothesis there is a $u \in U$ with response $\phi(u)$ such that $(\phi(u),\phi(u)(T)) \in \text{int}(X \oplus G_T)$. It follows from Theorem 5.4 that (P) is stably set.

Assume, $\bar{u}$ is an optimal controller, in other words $\bar{u}$ is a solution of (P). From Theorem 5.7, we must have that (P*) is stably set and has a solution $(\overline{\phi}^*,\overline{y}^*)$. Moreover, $\overline{u}$ and $(\overline{\phi}^*,\overline{y}^*)$ satisfy

$$\Lambda\overline{u} \in \partial g^*((\overline{\phi}^*,\overline{y}^*)),\ \Lambda^*((\overline{\phi}^*,\overline{y}^*)) \in \partial f(\overline{u})$$

by Theorem 5.8.

$$\Lambda\overline{u}=(\phi(\overline{u}),\phi(\overline{u})(T)) \in \partial g^*((\overline{\phi}^*,\overline{y}^*)) \rightarrow (\overline{\phi}^*,\overline{y}) \in \partial g((\phi(\overline{u}),\phi(\overline{u})(T)))$$

that is,

$$-\ell(y)-\int_0^T k(t,\phi(t))dt \leq -\ell(\phi(\overline{u})(T))-\int_0^T k(t,\phi(\overline{u}(t))dt+\int_0^T <\phi(t)-\phi(\overline{u}(t),$$

$$\overline{\phi}^*(t)>dt +<y-\phi(\overline{u})(T),\overline{y}^*>,\ \forall(\phi,y) \in L_p[E^n;0,T] \oplus E^n.$$

Because ϕ is independent of y it follows that

$$\ell(y) \geq \ell(\phi(\overline{u})(T)) + <y-\phi(\overline{u})(T),-y^*>,\ \forall\ y \in E^n \tag{6.9}$$

$$\int_0^T k(t,\phi(t))dt \geq \int_0^T k(t,\phi(\overline{u})(t))dt+\int_0^T <\phi(t)-\phi(\overline{u})(t),-\overline{\phi}^*(t)>dt,$$

$$\forall\ \phi \in L_p[E^n;0,T] \tag{6.10}$$

Suppose, there is a $\phi \in L_p[E^n;0,T]$ such that for some interval Δ of finite length, $\Delta \subseteq [0,T]$,

$$k(t,\phi(t))-<\phi(t),-\overline{\phi}^*(t)><k(t,\phi(\overline{u})(t))-<\phi(\overline{u})(t),-\overline{\phi}^*(t)>$$

holds on Δ.

Let $\tilde{\phi} \in L_p[E^n;0,T]$ be defined by

$$\tilde{\phi}(t) = \begin{cases} \phi(t), & t \in \Delta \\ \phi(\overline{u})(t), & t \notin \Delta \end{cases}$$

Then, $\int_0^T \{k(t,\phi(t)) - \langle\tilde{\phi}(t), -\overline{\phi}^*(t)\rangle\}dt <$

$$< \int_0^T \{k(t,\phi(\overline{u})(t)) - \langle\phi(\overline{u})(t), -\overline{\phi}^*(t)\rangle\}dt$$

which contradicts (6.10).

Therefore,

$$k(t,\phi(t)) \geq k(t,\phi(\overline{u})(t)) + \langle\phi(t)-\phi(\overline{u})(t), -\overline{\phi}^*(t)\rangle$$

almost everywhere on [0,T] for each $\phi \in L_p[E^n;0,T]$. Thus actually,

$$k(t,y) \geq k(t,\phi(\overline{u})(t)) + \langle y-\phi(\overline{u})(t), -\overline{\phi}^*(t)\rangle$$

almost everywhere for each $y \in E^n$. That is,

$$-\overline{\phi}^*(t) \in \partial k(t,\phi(\overline{u})(t)) \tag{6.11}$$

almost everywhere. Furthermore, $\Lambda^*(\overline{\phi}^*,\overline{y}^*) \in \partial f(\overline{u})$ means,

$$\int_0^T h(t,u(t))dt \geq \int_0^T h(t,\overline{u}(t))dt + \int_0^T \langle u(t)-\overline{u}(t), \Lambda^*(\overline{\phi}^*,\overline{y}^*)(t)\rangle dt$$

for all $u \in L_r[E^m;0,T]$, which implies as above that point wise

$$h(t,z) \geq h(t,\bar{u}(t)) + \langle z-\bar{u}(t), \Lambda^*(\bar{\phi}^*,\bar{y})(t)\rangle$$

almost everywhere for each $z \in E^m$. That is,

$$\Lambda^*(\bar{\phi}^*,\bar{y}^*)(t) \in \partial h(t,\bar{u}(t) \qquad (6.12)$$

almost everywhere. But we know from equation (6.4) that

$$\Lambda^*(\bar{\phi}^*,\bar{y}^*)(t)=B^*(t)\{\int_t^T \Phi^*(\tau,t)\bar{\phi}^*(\tau)d\tau+\Phi^*(T,t)\bar{y}^*\}= B^*(t)\bar{\eta}(t)$$

where $\bar{\eta}(t) = \Phi^*(T,t)\bar{y}^* + \int_T^t \Phi^*(\tau,t)(-\bar{\phi}^*(\tau))dt.$

Thus $\bar{\eta} \in L_q[E^n;0,T]$ such that,

$-\bar{\eta}(T)= -\bar{y}^* \in \partial\ell(\phi(\bar{u})(T))$ by (6.9) and $\bar{\eta}(t)$ satisfies

$$\dot{\eta} = -A^*(t)\eta - \bar{\phi}^*(t), \; -\bar{\eta}(T) \in \partial\ell(\phi(\bar{u})(T))$$

where $-\bar{\phi}^*(t) \in \partial k(t,\phi(\bar{u})(t))$ a.e. according (6.11).

Moreover, $\Lambda^*(\bar{\phi}^*,\bar{y})(t) = B^*(t)\bar{\eta}(t) \in \partial h(t,\bar{u}(t))$ a.e. by (6.12) and thus,

$$\langle B(t)\bar{u}(t),\bar{\eta}(t)\rangle - h(t,\bar{u}(t)) = \max_{z \in \Omega} \langle B(t)z,\bar{\eta}(t)\rangle - h(t,z)$$

almost everywhere.

This proves one part of the theorem.

To prove the second part, assume there is a vector $\overline{\eta} \in L_q[E^n;0,T]$ satisfying

$$\dot{\eta} = -A^*(t)\eta + \overline{\phi}^* \qquad -\overline{\eta}(T) \in \partial\ell(\phi(\overline{u})(T))$$

$$\overline{\phi}^*(t) \in \partial k(t,\phi(\overline{u})(t))$$

where $\phi(\overline{u})$ is the response of the controller $\overline{u}$ determined by the maximal principle,

$$\langle B(t)\overline{u}(t),\overline{\eta}(t)\rangle - h(t,\overline{u}(t)) = \max_{z\in\Omega}\langle B(t)z,\overline{\eta}(t)\rangle - h(t,z) \tag{6.13}$$

almost everywhere.

Since Ω is the effective domain of $h(t,z)$ we can write (6.13) as

$$\langle B(t)\overline{u}(t),\overline{\eta}(t)\rangle - h(t,\overline{u}(t)) = \sup_{z\in E^m}\langle z,B^*(t)\overline{\eta}(t)\rangle - h(t,z)$$

$$= h^*(t,B^*(t)\overline{\eta}(t)$$

almost everywhere, from which it readily follows that

$$B^*(t)\overline{\eta}(t) \in \partial h(t,\overline{u}(t)) \qquad \text{almost everywhere.}$$

But $\Lambda^*((-\overline{\phi}^*,\overline{\eta}(T))(t) = B^*(t)\overline{\eta}(t)$, so that

$$\Lambda^*((-\overline{\phi}^*,\overline{\eta}(T)) \in \partial f(\overline{u}) \tag{6.14}$$

Furthermore, $\overline{\phi}^*(t) \in \partial k(t,\phi(\overline{u})(t))$ and $-\overline{\eta}(T) \in \partial \ell(\phi(\overline{u})(T))$ imply $\phi(\overline{u})(t) \in \partial k^*(t,\overline{\phi}^*(t))$ and $\phi(\overline{u})(T) \in \partial \ell^*(-\overline{\eta}(T))$ respectively and we must have that

$$\Lambda\overline{u} = (\phi(\overline{u}),\phi(\overline{u})(T)) \in \partial g^*((-\overline{\phi}^*,\overline{\eta}(T)) \tag{6.15}$$

By Theorem 5.8 it follows that $\overline{u}$ is a solution to (P), that is $\overline{u}$ is an optimal controller.

Q.E.D.

Remark

In the above we considered the initial state of (P) to be zero but this is not essential. The results in the next section include the situation where the arbitrary initial state belongs to an arbitrary convex set.

Example

We consider an explicit condition for which an important class of control problems is stably set.

Given the linear control process in E^n,

$$\dot{x} = A(t)x + B(t)u \quad , \quad x(0) = x_0$$

The controllers u belong to some Banach space U and are constrained to lie in the closed, convex subset C of U,

$$C = \{u \in U| \quad ||u||_U \leq \rho, \ \rho > 0\}$$

The target set D is the closed convex subset of E^n,

$$D = \{x \in E^n \mid \ ||x||_{E^n} \leq \varepsilon , \quad \varepsilon > 0\}$$

Necessary and sufficient conditions for a controller $u \in C$ such that the corresponding response $\phi(u)$ has its endpoint in int (D) might be of interest in itself and are first derived.

Write,

$$\phi(u)(T) = \Phi(T,0)x_0 + \int_0^T \Phi(T,t)B(t)u(t)dt$$

$$= \Phi(T,0)x_0 + Lu$$

Then,

$$\phi(u)(T) \in \text{int}(D) \leftrightarrow <\Phi(T,0)x_0,x^*>-\rho||L^*x^*||<\varepsilon||x^*||, \forall x^* \in E^n$$

Suppose $\phi(u)(T) \in \text{int}(D)$. Then there is a $\overline{u} \in C$ such that

$$||\Phi(T,0)x_0 + L\overline{u}||<\varepsilon \rightarrow \inf_{u\varepsilon C}||\Phi(T,0)x_0 + Lu|| < \varepsilon$$

Or,

$$\inf_{u\varepsilon C} \sup_{||x^*||\leq 1} <\Phi(T,0)x_0 + Lu, x^*> < \varepsilon$$

$$\sup_{||x^*||\leq 1} \inf_{u\varepsilon C} <\Phi(T,0)x_0 + Lu, x^*> < \varepsilon$$

$$\sup_{||x^*||\leq 1} \{ <\Phi(T,0)\ x_0,x^*> - \rho||L^*x^*||\} < \varepsilon$$

Or,

$$<\Phi(T,0)x_0,x^*> - \rho||L^*x^*||< \varepsilon \, || \, x^*||, \forall \, x^* \varepsilon E^n$$

On the other hand let F be the attainable set, that is

$$F = \{x \, \varepsilon \, E^n | \, x = \Phi(T,0)x_0 + Lu, \quad u \, \varepsilon \, C\}$$

and assume,

$$(*) \quad <\Phi(T,0)x_0,x^*> - \rho \, ||L^*x^*|| < \varepsilon \, ||x^*||, \quad \forall \, x^* \varepsilon E^n$$

If F $\cap$ int(D) is empty, then F and int(D) can be separated by a hyperplane, since F is a closed and bounded convex set. Thus, there is an $\overline{x}^* \varepsilon E^n$ and a real number α such that

$$<x,x^*> \geq \alpha \quad , \qquad \forall \, x \, \varepsilon \, F$$

$$<x,x^*> \leq \alpha \quad , \qquad \forall \, x \, \varepsilon \, \text{int}(D)$$

Thus,

$$\sup_{x \varepsilon \text{int}(D)} <x,\overline{x}^*> \leq \alpha \rightarrow \sup_{x \varepsilon D} <x,\overline{x}^*> \leq \alpha$$

Or,

$$\varepsilon \, ||\overline{x}^*|| \leq \alpha$$

Further,

$$<x,\overline{x}^*> \geq \alpha \quad , \forall \, x \, \varepsilon \, F \rightarrow \inf_{u \varepsilon C} <\Phi(T,0)x_0 + Lu,\overline{x}^*> \geq \alpha$$

Or,

$$<\Phi(T,0)x_0,\overline{x}^*> - \rho||L^*\overline{x}^*|| \geq \alpha \geq \varepsilon\ ||\overline{x}^*||$$

which contradicts (*).

In particular, the control process is stably set if (*) holds, according Theorem 5.4.

Example

In this example we present a case where the primal problem is defined in a function space, but where its dual is defined in E^n.

Given the control process in E^n,

$$\dot{x} = A(t)x + B(t)u,\ x(0)=c,\ u \in L_2[E^m;0,T]$$

Constraint set $U = \{u \in L_2[E^m;0,T] | \int_0^T ||u(t)||^2_{E^m} dt \leq \rho,\ \rho > 0\}$

The task is to steer the response $\phi(u)$ on $[0,T]$ from c such that $||\phi(u)(T)||^2_{E^n}$ is minimized.

The attainable set F is defined as

$$F = \{\phi(u)(T) | \phi(u)(T)=\Phi(T,0)c+\int_0^T \Phi(T,t)B(t)u(t)dt\}$$

and F is closed and bounded.

Define the functions f, g as follows

$$f: L_2[E^m;0,T]\oplus E^n \to E^1 \ni\ f((u,x)) = \begin{cases} 0, & u \in U,\ x = c \\ +\infty, & \text{otherwise} \end{cases}$$

$$g: E^n \to E^1 \ni g(x) = -||x||^2 , \quad \forall x \in E^n.$$

Further let Λ be the linear bounded map defined by

$$\Lambda: L_2[E^m;0,T] \oplus E^n \to E^n \ni \Lambda(u,x) = \Phi(T,0)x + \int_0^T \Phi(T,t)B(t)u(t)dt$$

The adjoint Λ^* of Λ is defined by

$$\Lambda^* : E^n \to L_2[E^m;0,T] \oplus E^n \ni \Lambda^* x^* = (u^*, y^*) \quad \text{where}$$

$$u^*(t) = B^*(t)\, \Phi^*(T,t)\, x^* \quad , \quad t \in [0,T] \quad .$$

$$y^* = \Phi^*(T,0)\, x^*$$

The control process can be written as

(P) minimize $f((u,x)) - g(\Lambda(u,x)), u \in L_2[E^m;0,T], x \in E^n$

and its dual is

(P*) maximize $g^*(x^*) - f^*(\Lambda^* x^*) , \quad x^* \in E^n$

Note that (P*) involves a maximization in a finite dimensional space where (P) is a minimization problem in an infinite dimensional space.

Existence of an optimal controller $\overline{u} \in U$

(P) $\inf_{u,x} f((u,x)) - g(\Lambda(u,x)) = \inf_{u\varepsilon U} ||\Lambda(u,c)||^2$

$$= \inf_{\phi(u)(T)\varepsilon F} ||\phi(u)(T)||^2,$$

and since F is compact the infimum is attained at some $\phi(\overline{u})(T) \varepsilon F$ corresponding to a controller $\overline{u} \varepsilon U$ which is thus optimal.

(P) is stably set

For any $u \varepsilon U$, $f((u,c)) = 0$ and $\Lambda(u,c)$ belongs to the interior of the effective domain of g. Thus (P) is stably set by Theorem 5.4.

But the fact that (P) is stably set and has a solution implies that (P*) has a solution, say $\overline{x}^*$, and is stably set by Theorem 5.7.

Thus we have,

Characterization of the optimal controller $\overline{u}$ by the maximal principle as a necessary and sufficient condition.

The solutions $(\overline{u},c)$ of (P) and $\overline{x}^*$ of (P*) satisfy

$$\Lambda(\overline{u},c) \;\varepsilon\; \partial g^*(\overline{x}^*), \quad \Lambda^*\overline{x}^* \;\varepsilon\; \partial f((\overline{u},c))$$

by Theorem 5.8.

But, $\Lambda^*\overline{x}^* \;\varepsilon\; \partial f((\overline{u},c))$ is by definition,

$$f((u,x)) \geq f((\overline{u},c)) + \int_0^T \langle u(t)-\overline{u}(t), \overline{u}^*(t)\rangle dt + \langle x, \overline{y}^*\rangle$$

in which $\Lambda^*\bar{x}^* = (\bar{u}^*, \bar{y}^*)$, $\bar{u}^*(t) = B^*(t)\,\Phi^*(T,t)\bar{x}^*$ and $\bar{y}^* = \Phi^*(T,0)\bar{x}^*$.

Since $f((u,x))$ is zero for $u \in U$, $x = c$ and $+\infty$ otherwise the above inequality yields

$$0 \geq \int_0^T \langle u(t)-\bar{u}(t), B^*(t)\Phi^*(T,t)\bar{x}^*\rangle dt, \quad \forall\, u \in U$$

from which it follows that,

$$(*)\quad \forall\, u \in U: \langle u(t)-\bar{u}(t), B^*(t)\Phi^*(T,t)\bar{x}^*\rangle \leq 0, \text{ a.e. on } [0,T].$$

Define $\bar{\eta}(t)$ by

$$\bar{\eta}(t) \quad = \quad \Phi^*(T,t)\bar{x}^*$$

and write (*) as

$$\max_{z \in E^m} \langle B(t)z, \bar{\eta}(t)\rangle = \langle B(t)\bar{u}(t), \bar{\eta}(t)\rangle \quad \text{a.e.}$$

and $\bar{\eta}(t)$ satisfies the adjoint differential equation

$$\dot{\eta} \quad = \quad -A^*(t)\eta$$

with $\bar{\eta}(T)$ determined as follows

$$\phi(\bar{u})(T) = \Lambda(\bar{u},c) \in \partial g^*(\bar{x}^*) \rightarrow \bar{x}^* = \bar{\eta}(T) \in \partial g(\phi(\bar{u})(T))$$

and because g is actually differentiable everywhere,

$$\overline{\eta}(T) = -2\phi(\overline{u})(T).$$

<u>Example</u> (Reachability)

Given the linear control process in E^n,

$$L)\quad \dot{x} = A(t)x + B(t)u, x(0)=0,\ u \varepsilon L_r[E^m;0,T],\ r > 1$$

Class of admissible controllers U is the closed convex set defined by

$$U = \{u \varepsilon L_r[E^m;0,T] | u(t) \varepsilon \Omega \subseteq E^m \text{ a.e. on } [0,T], \Omega \text{ is closed convex}\}$$

Assume the responses $\phi(u)$ of $L)$ in $L_p[E^n;0,T]$, $p > 1$, and let X be the closed convex set defined by,

$$X = \{\phi \varepsilon L_p[E^m;0,T] | \phi(0)=0, \phi(t) \varepsilon G_t \subseteq E^n \text{ a.e. on } (0,T],\ G_t \text{ is closed, convex for } t \varepsilon (0,T]\}$$

Give necessary and sufficient conditions for a controller $\overline{u} \varepsilon U$ such that the response $\phi(\overline{u}) \varepsilon X$.

This problem of reachability can be formulated as a convex minimization problem.

Let f be the indicator function of the convex set U,

$$f(u) = \begin{cases} 0, & u \in U \\ +\infty, & u \notin U \end{cases} \qquad \text{and thus} \qquad f^*(u^*) = \sup_{u \in U} \langle u, u^* \rangle$$

Let g be the negative indicator function of the convex set X,

$$g(\phi) = \begin{cases} 0, & \phi \in X \\ -\infty, & \phi \notin X \end{cases} \qquad \text{and thus} \qquad g^*(\phi^*) = \inf_{\phi \in X} \langle \phi, \phi^* \rangle$$

Further define the map Λ as,

$$\Lambda : L_r[E^m;0,T] \to L_p[E^n;0,T] \ni \Lambda u(t) = \int_0^t \Phi(t,\tau)B(\tau)u(\tau)d\tau \ ,$$

$$t \in [0,T], \quad \forall\, u \in L_r[E^m;0,T]$$

Then the adjoint Λ^* of Λ is

$$\Lambda^* : L_q[E^n;0,T] \to L_s[E^m;0,T], \ \frac{1}{p} + \frac{1}{q} = 1 \quad , \quad \frac{1}{r} + \frac{1}{s} = 1$$

such that

$$\Lambda^*\phi^*(t) = B^*(t)\int_t^T \Phi^*(\tau,t)\phi^*(\tau)d\tau, \ t \in [0,T], \ \forall\, \phi^* \in L_q[E^n;0,T]$$

Or,

$$\Lambda^*\phi^*(t) = B^*(t)\eta(t), \ \eta(t) = \int_t^T \Phi^*(\tau,t)\phi^*(\tau)d\tau.$$

The reachability problem can be written as,

$$(P) \quad \text{minimize } f(u) - g(\Lambda u), \quad u \in L_r[E^m;0,T]$$

and its dual is

$$(P^*) \quad \text{maximize } g^*(\phi^*) - f^*(\Lambda^*\phi^*), \quad \phi^* \in L_q[E^n;0,T]$$

Note that since U and X are closed, f is l.s.c. and g is u.s.c. so that the general results apply for (P) and (P*).

Sufficient condition for reachability.

Suppose for some $\bar{u} \in U$, there exists an $\bar{\phi}^*$ such that,

$$\Lambda\bar{u} \in \partial g^*(\bar{\phi}^*) \ , \quad \Lambda^*\bar{\phi}^* \in \partial f(\bar{u})$$

Then, $\bar{u}$ is a solution to (P) and $\bar{\phi}^*$ is a solution to (P*), see Theorem 6.2.

$$\Lambda^*\bar{\phi}^* \in \partial f(\bar{u}) \rightarrow f(u) \geq f(\bar{u}) + \langle u-\bar{u},\Lambda^*\bar{\phi}^*\rangle, \quad \forall\, u \in L_r[E^m;0,T]$$

$$\rightarrow 0 \geq \int_0^T \langle u(t)-\bar{u}(t),\ \Lambda^*\bar{\phi}^*(t) \rangle\, dt, \qquad \forall\, u \in U$$

or $$0 \geq \int_0^T \langle u(t)-\bar{u}(t),\ B^*(t)\bar{\eta}(t) \rangle\, dt; \qquad \forall\, u \in U$$

where $$\bar{\eta}(t) = \int_t^T \Phi^*(\tau,t)\ \bar{\phi}^*(\tau)d\tau$$

From the last inequality it follows that,

$$\langle B(t)\bar{u}(t),\bar{\eta}(t)\rangle \geq \langle B(t)u,\bar{\eta}(t)\rangle, \quad u \in \Omega, \text{ a.e. on } [0,T]$$

Or,

$$\langle B(t)\bar{u}(t),\bar{\eta}(t)\rangle = \max_{z\varepsilon\Omega} \langle B(t)z,\bar{\eta}(t)\rangle \text{ a.e. on } [0,T]$$

and $\bar{\eta}(t)$ satisfies,

$$\dot{\eta} = -A^*(t)\eta + \bar{\phi}^* \quad , \quad \bar{\eta}(T) = 0$$

where $\bar{\phi}^*$ is determined by

$$\Lambda\bar{u} \; \varepsilon \; \partial g^*(\bar{\phi}^*) \rightarrow \bar{\phi}^* \; \varepsilon \; \partial g(\Lambda\bar{u})$$

$$g(\phi) \leq g(\Lambda\bar{u}) + \langle \phi-\Lambda\bar{u},\bar{\phi}^*\rangle \quad , \quad \forall\, \phi \; \varepsilon \; L_p[E^n;0,T]$$

Or, $$0 \leq \int_0^T \langle \phi(t)-\phi(\bar{u})(t), \bar{\phi}^* \rangle \, dt, \quad \forall\, \phi \; \varepsilon \; X$$

implying $0 \leq \langle \phi(t)-\phi(\bar{u})(t),\bar{\phi}^*(t)\rangle \quad , \quad \forall\, \phi(t) \; \varepsilon \; G_t$, a.e. on $[0,T]$

which can be written as

$$\langle \phi(\bar{u})(t),\bar{\phi}^*(t)\rangle = \min_{x\varepsilon G_t} \langle x,\bar{\phi}^*(t)\rangle, \text{ a.e. on } [0,T]$$

If the problem (P) is stably set, the above conditions are also necessary by Theorem 6.2.

Thus,

if $0 \; \varepsilon \; U$, $0 \; \varepsilon \; \text{int}(X)$ then (P) is stably set and there is a non-zero controller $\bar{u} \; \varepsilon \; U$

with response $\phi(\bar{u}) \; \varepsilon \; X$ if and only if there exists a vector $\bar{\eta} \; \varepsilon \; L_q[E^n;0,T]$ such that

$$\langle B(t)\bar{u}(t)\rangle = \max_{z\varepsilon\Omega} \langle B(t)z,\bar{\eta}(t)\rangle, \quad \text{a.e.}$$

where $\bar{\eta}(t)$ satisfies

$$\dot{\eta} = -A^*\eta + \bar{\phi}^*, \qquad \bar{\eta}(T) = 0$$

with

$$\langle \phi(\bar{u})(t),\bar{\phi}^*(t)\rangle = \min_{x\varepsilon G_t} \langle x,\bar{\phi}^*(t)\rangle, \text{ a.e.}$$

6.3 Linear Optimal Control with Convex Cost, Constrained Responses and Variable Initial States

A larger class of problems than the one considered in the previous section is obtained if we allow the initial state of the process to vary over some convex set and seek to minimize the cost function with respect to all such allowable initial states and the given class of admissible controllers. The cost functions may have a term which depends explicitly on the initial state. Again, problems of this class can be suitably formulated as convex programs of the form considered in Chapter V. Since the process studied in this section is a direct extension of the one in Section 6.2, it suffices to limit a detailed treatment of the problem to those parts of the theory where substantial differences arise.

Consider the linear control process in E^n,

$$\text{(P)} \quad L) \quad \dot{x} = A(t)x+B(t)u, \ x(0)=x \ \varepsilon \ G_0$$

is a convex subset of E^n. The matrices A(t) and B(t) are continuous on [0,T]. The responses $\phi(u)$ of L) are in $L_p[E^n;0,T]$, $p > 1$, and are constrained to lie in the convex subset X of $L_p[E^n;0,T]$ where

$$X=\{\phi \in L_p[E^n;0,T] \mid \phi(0)\in G_0, \phi(t)\in G_t \subseteq E^n, \text{a.e. on } (0,T], G_t \text{ convex}\}$$

The controllers u are in $L_r[E^m;0,T]$, $r > 1$, and the class of admissible controllers U is the convex subset of $L_r[E^m;0,T]$ defined by,

$$U = \{u \in L_r[E^m;0,T] \; u(t) \in \Omega \subseteq E^m \text{ a.e. on } [0,T], \Omega \text{ is convex, response } \phi(u) \in X\}$$

The cost function is,

$$C(u,x)=\ell_0(x)+\ell_1(\phi(u)(T))+\int_0^T \{h(t,u(t))+k(t,\phi(u)(t))\}dt$$

Hypotheses:

(i) ℓ_0, ℓ_1 are l.s.c. proper convex functions with effective domain G_0 and G_T respectively.

(ii) h, k satisfy the same hypotheses as in the previous section

(iii) Ω, G_t have non-empty interiors, $\forall\, t \in (0,T]$

To formulate the above control process in the model form of Chapter V, we define the following functions.

$$f : L_r[E^m;0,T] \oplus E^n \to E^1 \qquad \text{such that}$$

$$f((u,y))=\ell_0(y)+\int_0^T h(t,u(t))dt, \quad (u,y) \in L_r[E^m;0,T]\oplus E^n \qquad (6.15)$$

$$g : L_p[E^n;0,T] \oplus E^n \to E^1 \qquad \text{such that}$$

$$g((\phi,y))=-\ell_1(y)-\int_0^T k(t,\phi(t))dt, \quad (\phi,y) \in L_p[E^n;0,T]\oplus E^n \quad (6.16)$$

$$\Lambda : L_r[E^m;0,T] \oplus E^n \to L_p[E^n;0,T] \oplus E^n \quad \text{such that}$$

$\Lambda((u,y))=(\phi(u),\phi(u)(T))$ where $\phi(u)$ is defined by

$$\phi(u)(t)=\Phi(t,0)y+\int_0^t \Phi(t,\tau)B(\tau)u(\tau)d\tau, \ t \in [0,T] \qquad (6.17)$$

and Φ is the fundamental matrix of L). Then Λ is a linear, bounded transformation.

In the same fashion as in Section 6.2, it follows that f is a l.s.c. proper convex function with effective domain $U \oplus G_0$ and conjugate f^*,

$$f^*((u^*,y^*))=\ell_0^*(y^*)+\int_0^T h(t,u^*(t))dt, \ \forall(u^*,y^*) \in L_s[E^m;0,T]\oplus E^n$$

with some non-empty effective domain $U^* \oplus G_0^*$, in $L_s[E^m;0,T] \oplus E^n$.

g is an u.s.c. proper concave function with effective domain $X \oplus G_T$ and conjugate g^*,

$$g^*((\phi^*,y^*))=-\ell_1^*(-y^*)-\int_0^T k^*(t,-\phi^*(t))dt,\ \forall\ (\phi^*,y^*)\in L_q[E^n;0,T]\oplus E^n$$

with some non-empty effective domain $X^* \oplus G_T^*$ in $L_q[E^n;0,T] \oplus E^n$.

In terms of f, g and Λ, the process (P) can be written as

$$\text{(P)}\quad \text{minimize } C(u,x) = f((u,x))-g(\Lambda(u,x)) \text{ subject to}$$
$$(u,x) \in U \oplus G_0,\quad \Lambda(U,x) \in X \oplus G_T$$

and the dual process (P*) is accordingly given by

$$\text{(P*)}\quad \text{maximize } C^*(\phi^*,y^*)=g^*((\phi^*,y^*))-f^*(\Lambda^*(\phi^*,y^*))$$
$$\text{subject to } (\phi^*,y^*) \in X^* \oplus G_T^* \text{ and}$$
$$\Lambda^*(\phi^*,y^*) \in U^* \oplus G_0^*$$

A computation shows that Λ^* is determined by

$$\forall\ (\phi^*,y^*)\in L_q[E^n;0,T]\oplus E^n:\Lambda^*(\phi^*,y^*)=(u^*,x^*)\in L_s[E^m;0,T] \oplus E^n$$

where
$$u^*(t)=B^*(t)\{\int_t^T \Phi^*(\tau,t)\phi^*(\tau)d\tau + \Phi^*(T,t)y^*\}$$

$$x^* = \int_0^T \Phi^*(t,0)\phi^*(t)dt + \Phi^*(T,0)y^* \tag{6.20}$$

Thus the duality theory of Chapter V applies to the present situation with f, g, Λ defined by (6.15), (6.16), (6.17), $C = U \oplus G_0$, $D = X \oplus G_T$ and f^*, g^*, Λ^* defined by (6.18), (6.19), (6.20) and

$C^* = U^* \oplus G_0^*$, $D^* = X^* \oplus G_T^*$.

The existence theorem for an optimal pair $(\overline{u},\overline{x})$ is

Theorem 6.3

Given the optimal control problem (P). Assume in addition that (i) Ω is bounded and G_0 is compact, (ii) for each $t \in [0,T]$, $h(t,z) \geq H(t)$ for all $z \in E^m$, where $H(t)$ is summable on $[0,T]$, (iii) for L) $X \oplus G_T$ is reachable. Then there exists a pair $(\overline{u},\overline{x}) \in U \oplus G_0$ such that $\min_{(u,x)} C(u,x) = C(\overline{u},\overline{x})$. Moreover, if f is strictly convex, $(\overline{u},\overline{x})$ is a unique optimal pair.

Proof: Similar to Theorem 6.1. Q.E.D.

As in Section 6.2, the defining equations (6.20) of $\Lambda^*(\phi^*,y^*)$ connect the duality and Hamiltonian approaches resulting in a generalized maximal principle of Pontryagin type. Thus,

Theorem 6.4

Given the optimal control process (P) with variable initial state. Assume there is an admissible controller $u \in U$ with response $\phi(u)$ of L) such that $(\phi(u),\phi(u)(T))$ lies in the interior of $X \oplus G_T$. Then, a pair $(\overline{u},\overline{x})$ is optimal with respect to the set of admissible controllers U and the set of allowable initial states G_0 if and only if there exists a vector $\overline{\eta} \in L_q[E^n;0,T]$ satisfying

$$\dot{\eta} = -A^*(t)\eta + x^*(t,\phi(\overline{u})(t)), \quad -\overline{\eta}(T) \in \partial\ell_1(\phi(\overline{u})(T))$$

$$\overline{\eta}(0) \in -\partial\,\ell_0(\overline{x})$$

where $x^*(t,\phi(\overline{u})(t)) \in \partial k(t,\phi(\overline{u})(t))$ almost everywhere and such that the maximal principle

$$<B(t)\overline{u}(t),\overline{\eta}(t)>-h(t,\overline{u}(t))=\max_{z\in\Omega}\{<B(t)z,\overline{\eta}(t)>-h(t,z)\}$$

holds almost everywhere.

Proof: The proof parallels that of Theorem 6.2, except that here the condition $\Lambda^*(\overline{\phi}^*,\overline{y}^*) = (\overline{u}^*,\overline{x}^*) \in \partial f((\overline{u},\overline{x}))$ holds if and only if $\overline{u}^*(t) \in \partial h(t,\overline{u}(t)$ and $\overline{x}^* \in \partial \ell_0(\overline{x})$ where $\overline{u}^*$, $\overline{x}^*$ are defined by (6.20) and

$$\overline{x}^* = \Phi^*(T,0)\overline{y}^* + \int_0^T \Phi^*(t,0)\overline{\phi}^*(t)dt = \overline{\eta}(0)$$

Q.E.D.

Remark

To verify that Theorem 6.4 reduces to Theorem 6.2 for fixed initial state $x = x_0$, we show that in this case the condition $\overline{\eta}(0)\in\partial\ell_0(x_0)$ is automatically satisfied.

Although for fixed initial state the term ℓ_0 in the cost function is a constant and thus could be left out we prefer to take it in consideration and define ℓ_0 as,

$$\ell_0(x) = \begin{cases} 0\ , & x = x_0 \\ +\infty, & x \neq x_0. \end{cases}$$

The advantage is to have a problem formulation which precisely fits the theoretical setting of Theorem 6.4, since it is the particular case of (P) with $G_o = \{x_o\}$.

But

$$\bar{\eta}(0) \in \partial\ell_o(x_o) \leftrightarrow \ell_o(x) \geq \ell_o(x_o) + \langle x - x_o, \bar{\eta}(0)\rangle, \ \forall x$$

and it follows from the definition of ℓ_o that $\partial\ell(x_o)$ is all of E^n so that indeed $\bar{\eta}(0) \in \partial\ell_o(x_o)$ is always satisfied.

CHAPTER VII

CONTROL PROCESSES WITH PARTIAL DIFFERENTIAL EQUATIONS

7.1 Introduction

A class of linear hyperbolic or parabolic distributed processes can be formulated as an abstract Cauchy problem (initial value problem) in some function space and the solution is obtained as a semi-group of linear bounded operators on the function space to itself. The duality theory in Chapter V applies to such problems, if defined in a proper way. In Appendix I of this thesis a summary is given of the theory of semi-groups in connection with partial differential equations.

7.2 Linear Distributed Processes with Convex Cost Functions

Given the linear distributed process in E^n, $t \geq 0$, with time independent coefficients

$$\frac{\partial y}{\partial t}(x,t) = A(x)\frac{\partial^2 y}{\partial x^2}(x,t)+B(x)\frac{\partial y}{\partial x}(x,t)+C(x)y(x,t)+D(x)u(x,t)$$

$$y(x,0) = y_0(x)$$

where the matrices $A(x),\ldots,D(x)$ are $n \times n$ and x ranges over some subset X of E^n. The distributed controllers $u(x,t)$ are given on the boundary $\partial(X)$ of X for every $t \geq 0$.

We formulate the above process as an abstract Cauchy problem in a way described in Appendix I and write

$$\dot{y}(t) = Ay(t) + Bu(t),\ y(0) = y_0 \in D(A) \qquad (7.1)$$

Assume,

the solution $\phi(u)(t)$ belongs to a Banach space B_1 for each t, u(t) belongs to a Banach space B_2 for each t.

$A \equiv A(x) \frac{\partial^2}{\partial x^2} + B(x) \frac{\partial}{\partial x} + C(x)$ is a closed operator with dense domain D(A) and non-empty resolvent set.

B is a linear bounded transformation from B_2 into D(A).

For (7.1) to be a well posed problem some additional conditions on A, B and the control functions u(t) are needed.

In the usual way, the uniqueness of solutions $\phi(u)$ to (7.1) follows from the uniqueness of solutions ϕ to the homogeneous equation

$$\dot{y}(t) = Ay(t) \quad , \quad y(0) = y_0 \in D(A) \qquad (7.2)$$

According theorems A_1 and A_2 of the Appendix, a necessary and sufficient condition for (7.2) to have a unique solution for each $y_0 \in D(A)$ and $t > 0$ is that A be the infinitesimal generator of a semi-group S(t) of linear bounded transformations on B_1 to itself which is strongly continuous at the origin. Moreover, the solution is then given by

$$\phi(t) = S(t)\, y_0$$

and satisfies

$$\lim_{t \to o^+} ||\phi(t) - y_o|| = 0 \quad .$$

Here S(t) is required to be strongly continuous for $t \geq 0$. Necessary and sufficient conditions for a closed operator A with dense domain and non-empty resolvent set to be the infinitesimal generator of such a strongly continuous semi-group are given in [20].

In analogy with the finite dimensional case, we want (7.1) to have solutions of the form

$$\phi(u)(t) = S(t)y_o + \int_0^t S(t-s)Bu(s)ds \tag{7.3}$$

Indeed, see [23], it has been shown that (7.3) is the unique solution to (7.1) for each $y_o \in D(A)$, $t > 0$, under the conditions that

(i) A is the infinitesimal generator of a strongly continuous semi-group S(t).

(ii) B is linear and bounded.

(iii) u(t) is strongly measurable and Bochner integrable on every finite interval in $(0,\infty)$.

(iv) $Bu(t) \in D(A)$ for almost every $t > 0$

(v) $\int_{t_1}^{t_2} ||ABu(t)||dt < \infty$ on every finite interval $[t_1,t_2]$ in $(0,\infty)$.

The duality theory of Chapter V applies to distributed control processes defined by (7.1) and (7.3) with a wide variety of cost functions.

For example, the fixed endpoint problem on [0,T] with cost function

$$\ell(\phi(u)(T)) = ||\phi(u)(T)-\overline{y}||$$

where $\overline{y}$ is the given desired final state. This problem and the time optimal control problem were treated in [23]. Here we propose to study linear distributed processes with more general convex cost functions of the type

$$C(u) = \ell(\phi(u)(T))+\int_0^T \{h(t,u(t))+k(t,\phi(u)(t))\}dt$$

where ℓ, h, k are convex functions.

The class of problems includes as important special cases

(i) state regulator problem with ℓ a given convex function,

$$h(u(t))=||u(t)||, k(\phi(u)(t))=||\phi(u)(t)||$$

(ii) fixed endpoint problem with $\ell(\phi(u)(T)) =||\phi(u)(T)-\overline{y}||$ for given $\overline{y}$ and $h \equiv 0$, $k \equiv 0$.

To have a suitable problem setting, we consider the following function spaces.

If X is a, not necessarily bounded, subset of E^n, H_1 will be a separable Hilbert space of square Bochner summable functions on X,

$$\int_X ||y(x)||^2_{E^n}\, dx < \infty \quad , \qquad \forall\, y \in H_1$$

H_2 is a separable Hilbert space of square Bochner summable functions on the boundary $\partial(X)$ of X,

$$\int_{\partial(X)} ||z(x)||^2_{E^n}\, dx < \infty \quad , \qquad \forall\, z \in H_2.$$

Further, $L_p[H_1;0,T]$, $p > 1$, is the space of strongly measurable functions ϕ on $0 \leq t \leq T$ with range in H_1 such that

$$\int_0^T ||\phi(t)||^p_{H_1}\, dt < \infty$$

and $L_r[H_2;0,T]$, $r > 1$, is the space of strongly measurable functions u on $0 \leq t \leq T$ with range in H_2 such that

$$\int_0^T ||u(t)||^r_{H_2}\, dt < \infty$$

Note in particular that $u \in L_r[H_2;0,T]$, $r > 1$, implies u is Bochner summable on [0,T].

We can now give a precise definition of a distributed linear control process with convex cost function.

Consider the linear distributed control process in the separable Hilbert space H_1

$$\text{(ACP)} \qquad LP) \quad \dot{y}(t) = Ay(t)+Bu(t),\ y(0) = 0$$

where A is the infinitesimal generator of the strongly continuous semi-group S(t) and B is a bounded linear transformation from H_2 into the domain D(A) of A.

The responses $\phi(u)$ of $LP)$ are elements of $L_p[H_1;0,T]$, $p > 1$, and are required to be in the convex set Y where,

$$Y = \{\phi \in L_p[H_1;0,T] | \phi(0)=0, \phi(t) \in G_t \subseteq H_1 \text{ a.e. on } [0,T], \; G_t \text{ is convex}\}$$

The class of admissible controllers U is the convex subset of $L_r[H_2;0,T]$ defined by

$$U = \{u \in L_r[H_2;0,T] | u(t) \in \Omega \subseteq H_2 \text{ a.e. on } [0,T],\ \Omega \text{ is convex response } \phi(u) \in Y\}$$

The cost function is

$$C(u) = \ell(\phi(u)(T)) + \int_0^T \{h(t,u(t)) + k(t,\phi(u)(t))\}dt$$

and satisfies

(i) $\ell(y)$ is a l.s.c. proper convex function on H_1 with effective domain G_T.

(ii) $h(t,z)$ is a normal convex integrand (see Chapter II) with effective domain Ω such that $h(t,u(t))$ is summable in t for all $u \in U$ and $h^*(t,u^*(t))$ is summable in t for at least one $u^* \in L_s[H_2;0,T]$, $\frac{1}{r} + \frac{1}{s} = 1$.

(iii) $k(t,y)$ is a normal convex integrand with effective domain G_t such that $k(t,\phi(t))$ is summable in t for all $\phi \in Y$ and $k^*(t,\phi^*(t))$ is summable in t for at least one $\phi^* \in L_q[H_1;0,T]$, $\frac{1}{p} + \frac{1}{q} = 1$

In problem (ACP) defined above, LP) has a unique solution, as a Bochner integral

$$\phi(u)(t) = \int_0^t S(t - s)\, Bu(s)\, ds, \qquad t \in [0,T].$$

To formulate (ACP) as a convex program of the form in Chapter V, proceed as follows.

Let f be the function defined by

$$f(u) = \int_0^T h(t,u(t))dt, \quad \forall u \in L_r[H_2;0,T] \qquad (7.4)$$

In view of hypothesis (ii) f is a proper convex function on $L_r[H_2;0,T]$ with effective domain U. Further, f and the function f* on $L_s[H_2;0,T]$ where

$$f^*(u^*) = \int_0^T h^*(t,u^*(t))dt, \quad \forall u^* \in L_s[H_2;0,T] \qquad (7.5)$$

are conjugate to each other. The function f^* is a proper convex function by hypothesis (ii) and is automatically l.s.c. as the conjugate of the proper convex function f, which in turn implies that f is also l.s.c. Similarly, define the function g on $L_p[H_1;0,T] \oplus H_1$ where

$$g((\phi,y)) = -\ell(y) - \int_0^T k(t,\phi(t))dt, \forall(\phi,y) \in L_p[H_1;0,T] \oplus H_1 \qquad (7.6)$$

It follows by hypothesis (iii) that g is an u.s.c. proper concave function with effective domain $Y \oplus G_T$ and has the u.s.c. proper concave conjugate g^* on $L_q[H_1;0,T] \oplus H_1$ where

$$g^*((\phi^*,y^*)) = -\ell^*(-y^*) - \int_0^T k^*(t,-\phi^*(t))dt, \forall(\phi^*,y^*) \in L_q[H_1;0,T] \oplus H_1 \qquad (7.7)$$

In particular, f^* and g^* have non-empty effective domains which we shall denote by U^* and $Y^* \oplus G_T^*$ respectively.

Finally, define the transformation Λ by

$$\Lambda: L_r[H_2;0,T] \to L_p[H_1:0,T] \oplus H_1 \qquad \text{such that}$$

$$\forall\, u \in L_r[H_2;0,T] : \Lambda u = (\phi(u),\phi(u)(T)) \qquad \text{where}$$

$$\phi(u)(t) = \int_0^t S(t-s)\, Bu(s)ds, \quad \forall\, t \in [0,T] \qquad (7.8)$$

Then Λ is a linear bounded transformation. The problem (ACP) can be written as the convex program (P)

$$\text{(P)} \quad \text{minimize } f(u) - g(\Lambda u), \text{ subject to } u \in U \text{ and } \Lambda u \in Y \oplus G_T$$

and accordingly the dual program (P*) of (P) is

$$\text{(P*)} \quad \text{maximize } g^*((\phi^*,y^*)) - f^*(\Lambda^*(\phi^*,y^*)) \text{ subject to } (\phi^*,y^*) \in Y^* \oplus G_T^* \text{ and } \Lambda^*(\phi^*,y^*) \in U^*.$$

where Λ^*, the adjoint transformation of Λ, is defined by

$$\Lambda^*(\phi^*,g^*)(t) = B^*\{\int_t^T S^*(s-t)\phi^*(s)ds + S^*(T-t)y^*\}, \tag{7.9}$$

$$\forall (\phi^*,y^*) \in L_q[H_1;0,T] \oplus H_1$$

in which $S^*(t)$ is the adjoint semi-group of $S(t)$ and B^* the adjoint transformation of B.

Now that we have shown how (P) and (P*) satisfy the conditions of the general models in Chapter V, it suffices to remark that all the duality theorems make sense for the control problem (ACP).

<u>Theorem 7.1</u>

Given the optimal control problem (ACP). Assume in addition that i) Ω is bounded, ii) for each $t \in [0,T]$, $h(t,z) \geq H(t)$ for all

$z \in H_2$ where $H(t)$ is summable on $[0,T]$, iii) for the process LP), $Y \oplus G_T$ is reachable.

Then there exists a controller $\bar{u} \in U$ such that $\min_{u \in U} C(u) = C(\bar{u})$. Moreover if f is strictly convex then $\bar{u}$ is unique.

Proof: We prove that f* is finite on all of $L_s[H_2;0,T]$. The rest of the proof is identical to that of Theorem 6.1.

For arbitrary $u^* \in L_s[H_2;0,T]$

$$f^*(u^*) = \int_0^T h^*(t,u^*(t))dt = \int_0^T [\sup_{z \in H_2}\{<z,u^*(t)>-h(t,z)\}]dt$$

$$= \int_0^T [\sup_{z \in \Omega}\{<z,u^*(t)>-h(t,z)\}]dt$$

$$\leq M \int_0^T ||u^*(t)||dt - \int_0^T H(t)dt < +\infty$$

for some positive constant M.

Furthermore $f^*(u^*) > -\infty$ for every u* because f* is a proper convex function.

Q.E.D.

Because A is the infinitesimal generator of the strongly continuous semi-group S(t), there is a real number w_0 (see Appendix I) such that the resolvent $R(\lambda;A)$ of A can be expressed as

$$R(\lambda;A)y = \int_0^\infty e^{-\lambda t} S(t)y \, dt, \quad \lambda > w_0$$

which implies

$$\lambda R(\lambda;A)y \to y \quad \text{if} \quad \lambda \to \infty$$

This and the fact that H_1 is reflexive guarantee, [20], that the infinitesimal generator A^* of $S^*(t)$ is the adjoint transformation of A.

Furthermore, if the semi-group $S(t)$ is compact, $S^*(t)$ has the same properties of $S(t)$ and the abstract Cauchy problem

$$\dot{\eta}(t) = - A^*\eta(t), \quad \eta(T) = y^*$$

is well posed with unique solution $\eta(t) = S^*(T-t)y^*$, $0 \leq t \leq T$, see [23].

Equation (7.9) can be written as

$$\Lambda^*(\phi^*,y^*)(t) = B^*\eta(t)$$

where

$$\eta(t) = S^*(T-t)y^* + \int_t^T S^*(T-t)\phi^*(t)dt \qquad (7.10)$$

satisfies

$$\dot{\eta}(t) = - A^*\eta(t) + \phi^*(t), \quad \eta(T) = y^*$$

The above observations were made to prove

<u>Theorem 7 .2</u> (Generalized Pontryagin Maximal Principle)

Given the distributed control problem (ACP). Assume in addition that the semi-group $S(t)$ is compact and that there exists an admissible controller $u \in U$ with response $\phi(u)$ of LP) such that $(\phi(u),\phi(u)(T))$ lies in the interior of the constraint set $Y \oplus G_T$. Then, a controller $\bar{u}$ with response $\phi(\bar{u})$ of LP) is an optimal controller if and only if there exists a vector $\bar{\eta}$ in $L_q[H_1;0,T]$ satisfying

$$\dot{\eta}(t) = - A^*\eta(t) + y^*(t,\phi(\bar{u})(t)),-\bar{\eta}(T) \in \partial\ell(\phi(\bar{u})(T))$$

with $y^*(t,\phi(\bar{u})(t)) \in \partial k(t,\phi(\bar{u})(t))$ a.e. and such that

$$<B\bar{u}(t),\bar{\eta}(t)>-h(t,\bar{u}(t))=\max_{z\in\Omega}\{<Bz,\bar{\eta}(t)>-h(t,z)\} \text{ a.e.}$$

Proof: Similar to that of Theorem 6.2.

Q.E.D.

Remark: For examples of calculation of the infinitesimal generator of semi-groups which are solutions to partial differential equations, see [19], [21].

CHAPTER VIII

CONTROL PROCESSES WITH FUNCTIONAL DIFFERENTIAL EQUATIONS

8.1 Introduction

Optimal control processes whose equations of evolution are linear functional differential equations with constant coefficients are again properly defined in a function space. Following the techniques outlined in [24], we let the process evolve in a Banach space and obtain the responses as linear bounded transformations (semi-groups) of the control inputs. In Section 8.2 the preliminary mathematical results are given and in Section 8.3 an optimal control problem is defined and it is shown how the general duality theory can be applied to this problem. As in the previous chapter the solutions to the functional differential equation define a semi-group with infinitesinial generator A. In this case, however, the Banach space fails to be reflexive and it becomes a complicated matter when the adjoint semi-group can be identified with the solutions of the equation determined by A*. A priori introduction of the adjoint functional differential equation is the way out but it leads to a duality theory in Banach spaces which are not each others dual.

8.2 Linear Functional Differential Equations in Banach Spaces

In this section we follow J.K. Hale in [24] where the proofs of the theorems can be found.

For any two real numbers α, β, $\alpha \leq \beta$, $C[E^n;\alpha,\beta]$ is the Banach space of continuous functions on $[\alpha,\beta]$ with range in E^n and norm defined as

$$\forall\, \phi \in C[E^n;\alpha,\beta]:\ ||\phi|| = \sup_{\alpha \leq \theta \leq \beta} ||\phi(\theta)||_{E^n}$$

For any $\alpha \geq 0$, $K > 0$ and any continuous n-vector x on $[-\alpha,K]$, define for fixed $t \in [0,K]$ the function x_t on $[-\alpha,0]$ by

$$\forall\, t \in [0,K] : x_t(\theta) = x(t+\theta), \qquad \forall\, \theta \in [-\alpha,0]$$

Thus $x_t \in C[E^n;-\alpha,0]$ and coincides with the segment of the function x on $[t-\alpha,t]$.

Let $f(\phi)$ be an n-vector valued linear function defined on $C[E^n;-\alpha,0]$ and consider the linear functional differential equation with constant coefficients.

$$\dot{x} = f(x_t) \tag{8.1}$$

where $\dot{x}$ denotes the right hand derivative of x.

Definition 8.1

Let $t_o \geq 0$ be any real number and $\phi \in C[E^n;-\alpha,0], ||\phi|| \leq H$. The function $x(t_o,\phi)$ is said to be a solution to (8.1) with initial function ϕ at t_o if there is a number $K > 0$ such that

(i) for each $t_o \leq t \leq t_o + K$, $x_t(t_o,\phi)$ is defined, belongs to $C[E^n;-\alpha,0]$, $||x_t(t_o,\phi)|| \leq H$.

(ii). $x_{t_o}(t_o,\phi) = \phi$

(iii) $x(t_o,\phi)$ satisfies (8.1) for $t_o \leq t \leq t_o + K$.

If f satisfies a Lipschitz condition then, [14], equation (8.1) has a unique solution which depends continuously on ϕ for every $\phi \in C[E^n;-\alpha,0]$. We shall assume that f is Lipschitzian (well posed problems). Then f is continuous on $C[E^n;-\alpha,0]$ and has a representation as a Stieltjes integral with respect to some matrix m with elements of bounded variation on $[-\alpha,0]$,

$$f(\phi) = \int_{-\alpha}^{0} [dm(\theta)]\phi(\theta), \qquad \forall \phi \in C[E^n;-\alpha,0] \qquad (8.2)$$

This shows in particular that (8.1) includes all linear differential-difference equations with constant coefficients of the form

$$\dot{x}(t) = \sum_{k=o}^{n} A_k x(t-\tau_k), \qquad \tau_k \geq 0$$

Without loss of generality, take $t_o = 0$ and write $x(\phi)$ rather than $x(0,\phi)$ for the solution of (8.1) with initial function ϕ at zero.

For $\phi \in C[E^n;-\alpha,0]$, let $x(\phi)$ be the corresponding solution of (8.1) and define the transformation $S(t)$ on $C[E^n;-\alpha,0]$ to itself by

$$\forall t \geq 0: x_t(\phi) = S(t)\phi \,, \quad \forall \phi \in C[E^n;-\alpha,0] \qquad (8.3)$$

It is shown in [24], that S(t) is a semi-group of linear bounded transformations with S(0) = I and which is strongly continuous for all $t \geq 0$ and compact for all $t \geq \alpha$.

Thus, see Appendix I, the infinitesinial generator A of S(t) defined by

$$A\phi = \lim_{t \to 0^+} \frac{S(t)\phi - \phi}{t}$$

is a closed and linear operator with dense domain D(A).

From the definition (8.3) of S(t) it follows that $S(t)\phi$ is defined as

$$\begin{aligned} S(t)\phi(\theta) &= \phi(t+\theta), \quad t+\theta \leq 0 \\ S(t)\phi(\theta) &= \phi(0) + \int_0^{t+\theta} f(S(\tau)\phi)d\tau, \quad t+\theta > 0, -\alpha \leq \theta \leq 0 \end{aligned} \tag{8.4}$$

A calculation, using (8.4) shows that $A\phi$ is defined as

$$A\phi(\theta) = \begin{cases} \dfrac{d\phi(\theta^+)}{d\theta}, & -\alpha \leq \theta < 0 \\ \dfrac{d\phi(0^+)}{d\theta} = \int_{-\alpha}^{0} [dm(\theta)]\phi(\theta) = f(\phi) \end{cases} \tag{8.5}$$

Equation (8.1) can be written, in view of (8.2), as

$$\dot{x}(t) = \int_{-\alpha}^{0} [dm(\theta)]\, x(t+\theta), \qquad t \geq 0 \tag{8.6}$$

The adjoint equation of (8.6) is

$$\dot{y}(s) = - \int_{-\alpha}^{0} [dm^*(\theta)]\, y(s-\theta), \quad s \le 0 \tag{8.7}$$

Because if we let L, M be the operators defined by

$$Lx(t)=\dot{x}(t)- \int_{-\alpha}^{0}[dm(\theta)]x(t+\theta), My(s)=\dot{y}(s)+ \int_{-\alpha}^{0}[dm^*(\theta)]y(s-\theta)$$

then,

$$<y(t),Lx(t)>+<My(t),x(t)>= \frac{d}{dt}[<y(t),x(t)>-\int_{-\alpha}^{0}\int_{0}^{\theta}<y(t+\xi-\theta), [dm(\theta)]x(t+\xi)>d\xi.]$$

For $\phi \in C[E^n;-\alpha,0]$, $\psi \in C[E^n;0,\alpha]$ define the bilinear real valued function (ψ,ϕ) where,

$$(\psi,\phi)=<\psi(0),\phi(0)>- \int_{-\alpha}^{0}\int_{0}^{\theta} <\psi(\xi-\theta),[dm(\theta)]\phi(\xi)>d\xi. \tag{8.8}$$

If $\psi \in C[E^n;0,\alpha]$ then the adjoint equation (8.7) has a solution $y(\psi)$ with initial function ψ at zero and defined for all $s \le \alpha$. As above let

$$\forall\, s \le 0 : y_s(\psi)(\theta)=y(\psi)(s+\theta),\ 0 \le \theta \le \alpha$$

and define the transformation $S^*(s)$ where

$$\forall\ s \leq 0 : y_s(\psi) = S^*(s)\ \psi,\ \forall\ \psi\ \varepsilon\ C[E^n;0,\alpha] \qquad (8.9)$$

It follows that $S^*(s)$, $s \leq 0$, has all of the same properties as $S(t)$, $t \geq 0$.

The infinitesimal generator A_1^*, of $S^*(s)$ is defined by

$$A_1^*\ \psi = \lim_{s \to 0^-} \frac{S^*(s)\ \psi - \psi}{s}$$

A calculation shows that if A^* is the operator defined as $A^* = -A_1^*$ then $A^*\psi$ is defined by

$$A^*\psi(\theta) = \begin{cases} \dfrac{d\psi(\theta^-)}{d\theta}, & 0 \leq \theta \leq \alpha \\ -\dfrac{d\psi(0^-)}{d\theta} = \displaystyle\int_{-\alpha}^{0} [dm(\theta)]\psi(-\theta) & \end{cases} \qquad (8.10)$$

and $S^*(s)$, A^* satisfy the relationship,

$$\forall\psi\ \varepsilon\ D(A^*): \frac{dS^*(s)\psi}{ds} = -A^*S^*(s)\psi = -S^*(s)A^*\psi, s \leq 0$$

Finally it can be shown that

$$(\psi, A\phi) = (A^*\psi, \phi), \quad \forall\phi\ \varepsilon\ D(A) \qquad (8.11)$$

$$\forall\psi\ \varepsilon\ D(A^*).$$

where $(\cdot,\cdot)$ is defined in (8.8).

8.3 The Fixed Endpoint Problem

In what follows $L_p(C[E^n;-\alpha,0])$, $p > 1$, will denote the space of strongly measurable functions x on [0,T] with range in $C[E^n;-\alpha,0]$ such that

$$\int_0^T ||x(t)||^p_{C[E^n;-\alpha,0]} \, dt < +\infty$$

Similarly, $L_r(C[E^m;-\alpha,0])$, $r > 1$, is the space of strongly measurable functions u on [0,T] with range in $C[E^m;-\alpha,0]$.

Consider the following fixed endpoint problem.

Given the linear control process in $C[E^n;-\alpha,0]$

$$\text{(ACP)} \qquad LF) \; \dot{x}_t = Ax_t + Bu_t, \quad x_0 = 0$$

where A is the infinitesimal generator of the strongly continuous semi-group $S(t)$, $t \geq 0$, defined by (8.5) and B is a constant linear transformation from $C[E^m;-\alpha,0]$ into the domain $D(A)$ of A.

Note that differentiation with respect to t in $LF)$ is taken in the strong topology of $C[E^n;-\alpha,0]$.

The responses $x(u)$ of $LF)$ are elements in $L_p(C[E^n;-\alpha,0])$, $p > 1$. The target set G is a convex subset of $C[E^n;-\alpha,0]$ defined by

$$G = \{x \in C[E^n;-\alpha,0] \mid \; ||x-\bar{x}|| \leq \varepsilon, \; \varepsilon > 0, \bar{x} \text{ given}\} .$$

The class of admissible controllers U is a closed convex subset of $L_r(C[E^m;-\alpha,0])$ defined by

$$U=\{u \in L_r(C[E^m;-\alpha,0])|u_t \in \Omega\subseteq C[E^m;-\alpha,0] \text{ on } [0,T],$$
$$\Omega \text{ is closed convex, response endpoint } x_T(u) \in G\}$$

The cost function is

$$C(u) = ||x_T(u) - \bar{x}||$$

To show how the general duality theory applies to the above problem, define

$$f(u) = \begin{cases} 0, u \in U \\ +\infty, u \notin U \end{cases}$$

then f is a proper l.s.c. convex function on $L_r(C[E^m;-\alpha,0])$.

$$g(x) = \begin{cases} -||x-\bar{x}||, x \in G \\ -\infty, x \notin G \end{cases}$$

then g is an u.s.c. proper concave function on $C[E^n;-\alpha,0]$.

$$\Lambda : L_r(C[E^m;-\alpha,0]) \to C[E^n;-\alpha,0] \quad \text{such that}$$

$$\Lambda u = \int_0^T S(T-\tau)Bu_\tau d\tau, \ \forall\, u \in L_r(C[E^m;-\alpha,0])$$

Then Λ is a linear bounded transformation. The optimal control problem (ACP) above is in terms of f, g, Λ

(P) minimize $f(u) - g(\Lambda u)$, subject to all u in $L_r(C[E^m;-\alpha,0]$

The dual problem (P*) is

(P*) maximize $g^*(x^*)-f^*(\Lambda^*x^*)$, subject to all x^* in $C[E^n;0,\alpha]$

where f* and g* are the conjugate functions of f and g with respect to the bilinear function $(\cdot,\cdot)$ defined in (8.8).

The adjoint transformation Λ^* remains to be determined explicitly.

$$\Lambda^*: C[E^n;0,\alpha] \to L_s(C[E^m;0,\alpha], \quad \frac{1}{r} + \frac{1}{s} = 1$$

such that $\forall\, u \in L_r(C[E^m;-\alpha,0]),\ \forall\, \psi \in C[E^n;0,\alpha]$

$$\int_0^T (u_t,(\Lambda^*\psi)_t)dt = (\Lambda u\ ,\psi) =$$

$$=<\Lambda u\ (0),\psi(0)>- \int_{-\alpha}^{0}\int_0^{\theta} <\int_0^T S(T-t)Bu_t(\xi-\theta)dt,[dm(\theta)]\psi(\xi)>d\xi$$

$$=<\int_0^T S(T-t)Bu(t)dt,\psi(0)>-\int_0^T \int_{-\alpha}^{0}\int_0^{\theta}<S(T-t)Bu_t(\xi-\theta),[dm(\theta)]\psi(\xi)>d\xi dt$$

$$=\int_0^T \{<u_t(0),B^*S^*(t-T)\psi(0)> - \int_{-\alpha}^{0}\int_0^{\theta} <u_t(\xi-\theta),[dm(\theta)]B^*S^*(t-T)\psi(\xi)>d\xi\}dt$$

Thus $\Lambda^*\psi$ is defined by

$$(\Lambda^*\psi)_t \quad = \quad B^*S^*(t-T)\ \psi\ , T \geq t \geq 0 \qquad (8.12)$$

Theorem 8.1

Given the optimal control problem (ACP). If U is weakly compact and there exists an admissible controller $u \varepsilon U$ then there exists a controller $\bar{u} \varepsilon U$ which is optimal, that is,

$$\min_{u \varepsilon U} C(u) \quad = \quad C(\bar{u})$$

Proof: If U is weakly compact then,

$$f^*(u^*) = \sup_{u \varepsilon L_r(C[E^m;-\alpha,0])}\{(u,u^*)-f(u)\} = \sup_{u \varepsilon U} (u,u^*).$$

Since (u,u^*) is weakly continuous and U is weakly compact f^* is finite on all of $L_s(C[E^m;0,\alpha])$. The rest of the proof parallels that of Theorem 6.1.

Q.E.D.

If we define for some $\psi^o \ \varepsilon\ C[E^n;0,\alpha]$ the function η^o in $L_q(C[E^n;0,\alpha])$, $\frac{1}{p} + \frac{1}{q}$, where

$$\eta_t^o \quad = \quad S^*(t-T)\ \psi^o; \qquad T \geq t \geq 0$$

Then η^o is the unique solution of

$$\dot{\eta}_t = -A^*\eta_t \quad , \eta^o_T = \psi^o$$

where A* is defined in (8.10).

This fact, and the connection of η^o_t with $(\Lambda^*\psi^o)_t$ according to (8.12) leads to

Theorem 8.2 (Generalized Maximal Principle)

Given the functional differential process (ACP). Assume there is an admissible controller $u \varepsilon U$ with response $x(u)$ of LF) such that $x_T(u)$ lies in the interior of G.

Then, a controller $\bar{u}$ with response $x(\bar{u})$ is an optimal controller if and only if there exists a vector $\bar{\eta}$ in $L_q(C[E^n;0,\alpha])$ satisfying

$$\eta_t = - A^*\eta_t, \qquad -\bar{\eta}_T \ \varepsilon \ \partial g(x_T(\bar{u}))$$

almost everywhere and such that

$$<B\bar{u}_t,\bar{\eta}_t> = \max_{u \varepsilon \Omega} <Bu,\bar{\eta}_t>$$

almost everywhere.

Proof: Similar to Theorem 6.2.

Q.E.D.

CHAPTER IX

INTERCONNECTED CONTROL PROCESSES

9.1 Introduction

It frequently occurs that control processes of large dimension, that is with many inputs and many outputs, have the structure of mutually interconnected subprocesses of smaller dimension. Here, we consider such a linear interconnected optimal control process with cost function of separable type and show how the duality theory is of help in finding subprocesses which make up the dual of the interconnected process. In particular we obtain a decomposition theorem which relates the optimal controllers of the individual dual subprocesses to the optimal controller of the original overall process. For related work, see References [25] to [30].

9.2 Linear Interconnected Control Processes with Convex Cost Functions of Separable Type

Let E(n) denote the Euclidean vector space of dimension n, $L_p[E(n);0,T]$ the space of p-summable functions on [0,T] with range in E(n) and consider the linear interconnected control process in

$E(\sum_{i=1}^{N} n_i)$

$$\text{(P)} \qquad L) \quad \dot{x}_i = A_i(t)x_i + B_i(t)u_i + v_i$$

$$x_i(0) = 0 \qquad\qquad i = 1,\dots,N$$

$$T) \quad v_i(t) = \sum_{\substack{j=1 \\ \neq i}}^{N} F_{ij}(t)x_j(t)$$

where v_i, x_i are in $L_p[E(n_i);0,T]$, $p > 1$, u_i in $L_r[E(m_i);0,T]$, $r > 1$, and $A_i(t)$, $B_i(t)$, $F_{ij}(t)$ are continuous matrices of appropriate order, $i, j = 1,\ldots,N$.

The responses $\phi_i(u_1,\ldots,u_N)$ of L) are constrained to lie in the convex subsets X_i of $L_p[E(n_i);0,T]$ where

$$X_i=\{\phi_i \in L_p[E(n_i);0,T] | \phi_i(0)=0, \phi_i(t) \in G_i(t) \subseteq E(n_i)$$
$$\text{a.e. } G_i(t) \text{ is convex}\} \quad i = 1,\ldots,N$$

Let U_i be the convex subset in $L_r[E(m_i);0,T]$ defined by

$$U_i=\{u_i \in L_r[E(m_i);0,T] | u_i(t) \in \Omega_i \subseteq E(m_i) \text{ a.e. } \Omega_i \text{ is convex}\}$$
$$i = 1,\ldots,N$$

The class U of admissible controllers $u = (u_1,\ldots,u_N)$ is defined as

$$U=\{u=(u_1,\ldots,u_N) \in L_r[E(\sum_{i=1}^{N} m_i);0,T] | u_i \in U_i,$$
$$\text{responses } \phi_i(u_1,\ldots,u_N) \in X_i, i=1,\ldots,N\}$$

The cost function is

$$C(u_1,\ldots,u_N)=\sum_{i=1}^{N}\int_0^T\{h_i(t,u_i(t))+k_i(t,\phi_i(u_1,\ldots,u_N)(t))\}\, dt$$

and satisfies as before,

(i) $h_i(t,Z_i)$, $i = 1,\dots,N$, is a normal convex integrand (see Chapter 2) such that $h_i(t,u_i(t))$ is summable in t for all $u_i \in U_i$ and $h_i^*(t,u_i^*(t))$ is summable in t for at least one $u_i^* \in L_s[E(m_i);0,T]$, $\frac{1}{r} + \frac{1}{s} = 1$

(ii) $k_i(t,y_i)$, $i = 1,\dots,N$, is a normal convex integrand such that $k_i(t,\phi_i(t))$ is summable in t for all $\phi_i \in X_i$ and $k_i^*(t,\phi_i^*(t))$ is summable in t for at least one ϕ_i^* in $L_q[E(n_i);0,T]$, $\frac{1}{p} + \frac{1}{q} = 1$.

Extremum problems of this type were considered in [30] from a mathematical programming point of view and it was shown how a dual problem can be constructed via the Lagrangean function of the overall problem and further how a decomposition of the dual can be obtained. But, due to the strong assymmetry between primal and dual problem one may arrive at the conclusion that decomposition could be obtained via duality. This is not so, we feel, and may lead to confusion.

Here we take a different approach and apply the duality theory of Chapter V, after appropriate modeling of the process (P).

Set,

$$x(t) = \begin{pmatrix} x_1(t) \\ \vdots \\ x_N(t) \end{pmatrix}, \quad u(t) = \begin{pmatrix} u_1(t) \\ \vdots \\ u_N(t) \end{pmatrix} \quad \text{and write } L) \text{ as}$$

$$L) \quad x = A(t)x + B(t)\,u, \quad x(0) = 0 \qquad (9.1)$$

where,

$$A(t) = \begin{pmatrix} A_1(t) & F_{12}(t)- & - & - & -F_{1N}(t) \\ F_{21}(t) & A_2(t) & - & - & -F_{2N}(t) \\ \vdots & & & & \vdots \\ F_{N1}(t) & F_{N2}(t)- & - & - & -A_N(t) \end{pmatrix}, \quad B(t) = \begin{pmatrix} B_1(t) & & 0 \\ & \ddots & \\ 0 & & B_N(t) \end{pmatrix}$$

Define the functions f, g and their conjugates f*, g*.

It follows from hypothesis (i) in (P) that the functions f_i, $i = 1,\ldots,N$,

$$f_i(u_i) = \int_0^T h_i(t,u_i(t))dt \tag{9.2}$$

are l.s.c. proper convex functions on $L_r[E(m_i);0,T]$ with effective domains U_i and conjugate functions

$$f_i^*(u_i^*) = \int_0^T h_i^*(t,u_i^*(t))dt \tag{9.3}$$

on $L_s[E(m_i);0,T]$. In particular f_i^* has non-empty effective domain which we denote by U_i^*. Let f be defined by

$$f:L_r[E(m_1);0,T]\oplus\text{--}\oplus\ _r[E(m_N);0,T] \to E(1)$$

such that

$$f(u)=(f_1\oplus\text{--}\oplus f_N)(u_1,\ldots,u_N)=f_1(u_1)+\ldots+f_N(u_N),$$
$$\forall\ u \in L_r[\sum_{i=1}^{N}(m_i);0,T] \tag{9.4}$$

It is clear that f is a l.s.c. proper convex function with effective domain $U = U_1 \oplus \text{--} \oplus U_N$ and conjugate f^* where

$$f^*(u^*) = (f_1^* \oplus \text{---} \oplus f_N^*)(u^*), \qquad u^* \varepsilon\ L_s[\sum_{i=1}^{N}(m_i);0,T] \tag{9.5}$$

with effective domain $U^* = U_1^* \oplus \text{---} \oplus U_N^*$.

From hypothesis (ii) in (P) it follows that the functions g_i, $i = 1,\dots,N$,

$$g_i(\phi_i) \quad = \quad - \int_0^T k_i(t,\phi_i(t))\ dt \tag{9.6}$$

are u.s.c. proper concave functions on $L_p[E(n_i);0,T]$ with effective domains X_i and conjugate functions

$$g_i^*(\phi_i^*) \quad = \quad - \int_0^T k_i^*(t,-\phi^*(t))\ dt$$

In particular g_i^* has non-empty effective domain and we denote by X_i^*.

Let g be the function defined by

$$g:L_p[E(n_1);0,T] \oplus \text{--} \oplus L_p[E(n_N);0,T] \to E(1)$$

such that

$$g(\phi)=(g_1 \oplus \text{--} \oplus g_N)(\phi)=g_1(\phi_1)+\text{--}+g_N(\phi_N), \forall\, \phi\ \varepsilon\ L_p[E(\sum_{i=1}^{N} n_i);0,T] \tag{9.11}$$

then g is an u.s.c. proper concave function with effective domain $X = X_1 \oplus \text{---} \oplus X_N$ and conjugate g^* given by

$$g^*(\phi^*)=(g_1^*\oplus\text{---}\oplus g_N^*)(\phi^*), \quad \forall\ \phi^* \in L_q[\sum_{i=1}^{N}(n_i);0,T] \qquad (9.12)$$

with effective domain $X^* = X_1^*\oplus\text{---}\oplus X_N^*$. Finally define the map Λ by,

$$\Lambda : L_r[E(m_1);0,T]\oplus\text{--}\oplus L_r[E(m_N);0,T] \to L_p[E(n_1);0,T]\oplus\text{--}\oplus L_p[E(n_N);0,T]$$

such that

$$\Lambda u = \phi(u)\ , \quad \forall\ u,$$

and $\phi(u)$ is defined by

$$\phi(u)(t)= \int_0^t \Phi(t,\tau)B(\tau)u(\tau)dt, \quad t \in [0,T] \qquad (9.13)$$

where Φ is the fundamental matrix of (9.1). Thus Φ satisfies

$$\dot{\Phi} = A(t)\ \Phi\ , \qquad \Phi(0) = I$$

which is a system of $N \times N$ matrix differential equations

$$\begin{pmatrix} \dot{\phi}_{11} & \cdots & \dot{\phi}_{12} \\ \vdots & & \vdots \\ \dot{\phi}_{N1} & \cdots & \dot{\phi}_{NN} \end{pmatrix} = \begin{pmatrix} A_1(t) & F_{12}(t) \cdots & F_{N1}(t) \\ \vdots & & \vdots \\ F_{N1}(t) & F_{N2}(t) \cdots & A_N(t) \end{pmatrix} \begin{pmatrix} \phi_{11} & \cdots & \phi_{1N} \\ \vdots & & \vdots \\ \phi_{N1} & \cdots & \phi_{NN} \end{pmatrix}$$

$$\begin{aligned}\dot{\phi}_{11} &= A_1(t)\phi_{11}+F_{12}(t)\phi_{21}+\cdots+F_{N1}(t)\phi_{N1} \\ \dot{\phi}_{N1} &= F_{N1}(t)\phi_{1N}+F_{N2}(t)\phi_{2N}+\cdots+A_N(t)\phi_{NN}\end{aligned} \quad ,\phi_{ij}(o)=\begin{cases} I, & i=j \\ 0, & i\neq j\end{cases} \tag{9.14}$$

The interconnected control process (P) in terms of f, g, Λ defined by (9.4), (9.11) and (9.13) becomes

(P) minimize $f(u) - g(\Lambda u)$, subject to $u \in U$, $\Lambda u \in X$

and consequently the dual (P*) is

(P*) maximize $g^*(\phi^*) - f^*(\Lambda^*\phi^*)$, subject to $\phi^* \in X^*$, $\Lambda^*\phi^* \in U^*$

where f^*, g^* are defined by (9.5), (9.12) and Λ^* is the adjoint of Λ.

It follows from (9.14) that

$\forall\, \phi^* \in L_q[E(n_1);0,T]\oplus\cdots\oplus L_q[E(n_N);0,T]:\Lambda^*\phi^*\in L_s[E(m_1);0,T]\oplus\cdots\oplus L_s[E(m_N);0,T]$

such that

$$\Lambda^*\phi^*(t)=\begin{pmatrix} B_1^*(t) & & 0 \\ & \ddots & \\ 0 & & B^*(t)\end{pmatrix}\begin{pmatrix}\phi_{11}^*(T,t)\phi_{21}^*(T,t)\cdots\phi_{N1}^*(T,t) \\ \phi_{12}^*(T,t)\phi_{22}^*(T,t)\cdots\phi_{N2}^*(T,t) \\ \vdots \quad \ddots \quad \vdots \\ \phi_{1N}^*(T,t)\phi_{2N}^*(T,t)\cdots\phi_{NN}^*(T,t)\end{pmatrix}\begin{pmatrix}\phi_1^*(t) \\ \phi_2^*(t) \\ \vdots \\ \phi_N^*(t)\end{pmatrix}$$

Thus the general duality theory of Chapter V can be applied to the interconnected control problem (P).

We shall write (P) and (P*) in a form which emphasizes the sub-problem-structure of both (P) and (P*).

For $\Lambda u = (\phi_1,\ldots,\phi_N), u \in L_r[E(\sum_{i=1}^{N} m_i);0,T]$, define the map Λ_i,

$$\Lambda_i : L_r[E(\sum_{i=1}^{N} m_i);0,T] \to L_p[E(n_i);0,T] \quad \text{such that } \Lambda_i u = \phi_i, \quad i = 1,\ldots,N$$

and similarly for $\Lambda^*\phi^* = (u_1^*,\ldots,u_N^*)$, $\phi^* \in L_q[E(\sum_{i=1}^{N} n_i);0,T]$, define the map Λ_i^* by

$$\Lambda_i^* : L_q[E(\sum_{i=1}^{N}(n_i);0,T] \to L_s[E(m_i);0,T] \text{ such that } \Lambda_i^*\phi^* = u_i^*, \quad i = 1,\ldots,N.$$

Note that Λ_i^* is not the adjoint of Λ_i. From the above definitions it follows that

$$\langle \Lambda u,\phi^*\rangle = \sum_{i=1}^{N} \langle \Lambda_i u,\phi_i^*\rangle \ , \ \langle u,\Lambda^*\phi^*\rangle = \sum_{i=1}^{N} \langle u_i,\Lambda_i^*\phi^*\rangle$$

We can write (P) as,

$$(P) \quad \inf_{u\varepsilon U}\{f(u)-g(\Lambda u)\} = \inf_{u\varepsilon U}[f(u)- \inf_{\phi^*\varepsilon X^*}\{<\Lambda u,\phi^*>-g^*(\phi^*)\}] =$$

$$= \inf_{u\varepsilon U}[f(u) + \sup_{\phi^*\varepsilon X^*}\{g^*(\phi^*) - <\Lambda u,\phi^*>\}]$$

Or, in view of (9.12) and (9.15)

$$(P) \quad \inf_{u\varepsilon U}[f(u) + \sum_{i=1}^{N} \sup_{\phi_i^*\varepsilon X_i^*}\{g_i^*(\phi_i^*) -<\Lambda_i u,\phi_i^*>\}] \tag{9.16}$$

Similarly we can write (P*) in the form

$$(P^*) \quad \sup_{\phi^*\varepsilon X^*}[g^*(\phi^*)+ \sum_{i=1}^{N} \inf_{u_i\varepsilon U_i}\{f_i(u_i)- <u_i,\Lambda_i^*\phi^*>\}]. \tag{9.17}$$

For fixed u, (P) has N separate subproblems (P_i) of the form

$$(P_i) \quad \sup_{\phi_i^*\varepsilon X_i^*}\{g_i^*(\phi_i^*)- <\Lambda_i u,\phi_i^*>\}, \; i = 1,2,\ldots,N \tag{9.18}$$

where for fixed ϕ^*, the dual (P*) has N separated subproblems (P_i^*) of the form

$$(P_i^*) \quad \inf_{u_i\varepsilon U_i}\{f_i(u_i)- <u_i,\Lambda_i^*\phi^*>\}, \; i = 1,2,\ldots,N \tag{9.19}$$

Note that (P) as well as (P*) has the same subproblem-structure.

Before we can relate the solutions of the subproblems (P_i), (P_i^*) to those of the overall problems (P) and (P*) we have to make some preparations.

Introduce the function $K(u,\phi^*)$ where

$$K(u,\phi^*) = f(u) + g^*(\phi^*) - \langle \Lambda u,\phi^* \rangle \qquad (9.20)$$

Then, K is a l.s.c. proper convex function of u, with effective domain U and an u.s.c. proper concave of ϕ^* with effective domain X^*.

A point $(\bar{u},\bar{\phi}^*)$, $\bar{u} \in U$, $\bar{\phi}^* \in X^*$ is a saddle-point of $K(u,\phi)$ if

$$K(\bar{u},\bar{\phi}^*) = \min_{u \in U} K(u,\bar{\phi}^*) = \max_{\phi^* \in X^*} K(\bar{u},\phi^*)$$

Theorem 9.1

The function $K(u,\phi^*)$ defined in (9.20) has a saddle-point $(\bar{u},\bar{\phi}^*)$ if and only if (P) and (P*) are stably set in which case $\bar{u}$ is a solution (optimal controller) of (P) and $\bar{\phi}^*$ is a solution of (P*).

Proof: For any $\bar{u} \in U$, $\bar{\phi}^* \in X^*$,

$$\inf_{u \in U} K(u,\bar{\phi}^*) = g^*(\bar{\phi}^*) - f^*(\Lambda^*\bar{\phi}^*)$$

$$\sup_{\phi^* \in X^*} K(\bar{u},\phi^*) = f(\bar{u}) - g(\Lambda\bar{u})$$

Thus, $(\bar{u},\bar{\phi}^*)$ is a saddle-point of $K(u,\phi^*)$ if and only if

$$f(\bar{u}) - g(\Lambda\bar{u}) = g^*(\bar{\phi}^*) - f^*(\Lambda^*\bar{\phi}^*)$$

which, in view of Lemma 5.1, is equivalent to

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

are attained at $\bar{u}$ and $\bar{\phi}^*$.

The conclusion of the theorem follows from Theorem 5.5.

Q.E.D.

Corollary 9.2

Assume the process (P) is stably set. Then $\bar{u}$ is a solution of (P) if and only if there exists an $\bar{\phi}^*$ such that $(\bar{u},\bar{\phi}^*)$ is a saddle-point of $K(u,\phi^*)$.

Proof: Immediate from Theorem 9.1 and Theorem 5.7.

Q.E.D.

We can now prove the decomposition theorem for the interconnected control problem (P) and its dual (P*).

Theorem 9.3

Assume (P) and (P*) are both stably set. If for fixed $\bar{\phi}^*$, $\bar{u}_i$ solves the subproblem (9.19), $i = 1,\ldots,N$ of (P*) then $\bar{u} = (\bar{u}_1,\ldots,\bar{u}_N)$ is a solution to (P) if and only if $\bar{\phi}^*$ is a solution to (P*). Dually, if for fixed $\bar{u}$, $\bar{\phi}_i^*$ solves the subproblem (9.18), $i = 1,\ldots,N$, of (P) then $\bar{\phi}^* = (\bar{\phi}_1^*,\ldots,\bar{\phi}_N^*)$ is a solution of (P*) if and only if

$\overline{u}$ is a solution of (P).

Proof: Let for some $\overline{\phi}^* \in X^*$, $\overline{u}_i$ be a solution to (P_i^*), $i = 1,\ldots,N$. Assume $\overline{u} = (\overline{u}_1,\ldots,\overline{u}_N)$ is a solution to (P). Because $\overline{u}_i$ solves (P_i^*) it readily follows that

$$-f_i^*(\Lambda_i^*\overline{\phi}^*) = f_i(\overline{u}_i) - \langle\overline{u}_i,\Lambda_i^*\overline{\phi}^*\rangle \quad ,i=1,..,N \tag{9.21}$$

and thus

$$-f^*(\Lambda^*\overline{\phi}^*) = f(\overline{u}) - \langle\overline{u},\Lambda^*\overline{\phi}^*\rangle \tag{9.22}$$

Therefore,

$$K(\overline{u},\overline{\phi}^*) = g^*(\overline{\phi}^*) + f(\overline{u}) - \langle\Lambda\overline{u},\overline{\phi}^*\rangle = \tag{9.23}$$

$$= g^*(\overline{\phi}^*) - f^*(\Lambda^*\overline{\phi}^*) \leq \sup_{\phi^*\in X^*} K(\overline{u},\phi^*)$$

Since (P) is stably set and has a solution $\overline{u}$, it follows from Theorem 5.6 that (P*) has a solution and thus by Theorem 9.1 we must have,

$$\min\ (P) = \sup_{\phi^*\in X^*} K(\overline{u},\phi^*) \tag{9.24}$$

Hence, from (9.23) and (9.24)

$$g^*(\overline{\phi}^*) - f^*(\Lambda^*\overline{\phi}^*) \leq \min\ (P)$$

But by Lemma 5.1,

$$\min (P) \geq g^*(\bar{\phi}^*) - f^*(\Lambda^*\bar{\phi}^*)$$

so that

$$\min (P) = g^*(\bar{\phi}^*) - f^*(\Lambda^*\bar{\phi}^*)$$

what implies that $\bar{\phi}^*$ is a solution of (P^*).

On the other hand assume $\bar{\phi}^*$ is a solution of (P^*). Then,

$$\max (P^*) = g^*(\bar{\phi}^*) - f^*(\Lambda^*\bar{\phi}^*)$$

Or in view of (9.22), (9.23)

$$\max (P^*) = g^*(\bar{\phi}^*) + f(\bar{u}) - \langle \Lambda\bar{u},\bar{\phi}^* \rangle = K(\bar{u},\bar{\phi}^*)$$

But (P^*) is stably set and has a solution implies (P) has a solution, by Theorem 5.7, such that

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

That is,

$$\min_{u\varepsilon U} \max_{\phi^*\varepsilon X^*} K(u,\phi^*) = \max_{\phi^*\varepsilon X^*} \min_{u\varepsilon U} K(u,\phi^*)$$

and in view of (9.23),

$$\min_{u\varepsilon U} \max_{\phi^*\varepsilon X^*} K(u,\phi^*) = \max_{\phi^*\varepsilon X^*} \min_{u\varepsilon U} K(u,\phi^*) = K(\bar{u},\bar{\phi}^*)$$

Hence, $(\bar{u},\bar{\phi}^*)$ is a saddle-point of $K(u,\phi^*)$ implying that $\bar{u}$ is a solution of (P) by Theorem 9.1. This proves the first part of the theorem and the second part follows dually.

Q.E.D.

The following theorem guarantees that the subproblems (P_i) and (P_i^*) in (9.18) and (9.19) are stably set and do have solutions.

Theorem 9.4

Given the interconnected control problem (P) and its dual (P*). Assume that, for i = 1,...,N

(i) the effective domain Ω_i of $h_i(t,z_i)$ is bounded

(ii) $h_i(t,z_i) \geq H_i(t)$, $\forall z_i \in \Omega_i$ where H_i is a summable function on [0,T].

(iii) the effective domain $G_i^*(t)$ of $k_i^*(t,y_i^*)$ is bounded

(iv) $k_i^*(t,y_i^*) \geq K_i(t)$, $\forall y_i^* \in G_i^*(t)$ where K_i is a summable function on [0,T].

Then, the subproblems (P_i) in (9.18) and (P_i^*) in (9.19) are stably set and have solutions.

Proof:

$$(P_i) \quad \sup_{\phi_i^* \in X_i^*} \{g_i^*(\phi_i^*) - \langle \Lambda_i u, \phi_i^* \rangle\}, \text{ for some } u$$

has a dual process

$$\inf_{\phi_i} \{\ell_i(\phi_i) - g_i(\phi_i)\}$$

where

$$\ell_i(\phi_i) = \sup_{\phi_i^*}\{<\phi_i,\phi_i^*> - <\Lambda_i u,\phi_i^*>\} =$$

$$= \sup_{\phi_i^*} <\phi_i-\Lambda_i u,\phi_i^*> = \begin{cases} 0, & \phi_i = \Lambda_i u \\ +\infty, & \phi_i \neq \Lambda_i u \end{cases}$$

Further, according (9.6)

$$g_i(\phi_i) = - \int_0^T k(t,\phi_i(t))dt = - \int_0^T \{\inf_{x_i^* \; G_i^*} <\phi_i(t),\phi_i^*>-k^*(t,\phi_i^*)\}dt$$

$$= \int_0^T \{\sup_{\phi_i^* \varepsilon G_i^*} <-\phi_i(t),\phi_i^*> + k^*(t,\phi_i(t)\}dt$$

$$\geq \int_0^T \{\sup_{\phi_i^* \varepsilon G_i^*} <-\phi_i(t),\phi_i^*> + K_i(t)\}dt > - \infty$$

We also know that g_i is a proper concave function so that g_i is finite on all of $L_p[E(n_i);0,T]$. Thus the dual problem of (P_i) is stably set by Theorem 5.4. Clearly, $\Lambda_i u$ is also the solution to this dual problem from which it follows by Theorem 5.7 that (P_i) is stably set and has a solution.

$$(P_i^*) \quad \inf_{u_i \varepsilon U_i} \{f_i(u_i) - <u_i,\Lambda_i^*\phi^*>\}, \text{ for some } \phi^*,$$

has a dual,

$$\sup_{u_i^*} \{p_i^*(u_i^*) \; - \; f_i^*(u_i^*)\}$$

where

$$p_i^*(u_i^*) = \inf_{u_i}\{<u_i,u_i^*> - <u_i,\Lambda_i^*\phi^*>\}$$

$$= \inf_{u_i} <u_i,u_i^*-\Lambda_i^*\phi^*> = \begin{cases} 0, & u_i = \Lambda_i^*\phi^* \\ -\infty, & u_i \neq \Lambda_i^*\phi^* \end{cases}$$

From (9.2) we have

$$f_i^*(u_i^*) = \int_0^T h_i^*(t,u_i^*(t))dt = \int_0^T \{\sup_{z_i \varepsilon \Omega_i} <z_i,u_i^*(t)> - h_i(t,z_i)\}dt$$

$$\leq \int_0^T \{\sup_{z_i \varepsilon \Omega_i} <z_i,u_i^*(t)> - H_i(t)\}dt < +\infty$$

Because f_i^* is also proper it follows that f_i^* is finite on all of $L_s[E(m_i);0,T]$. It follows by Theorem 5.4 that the dual of (P_i^*) is stably set. Furthermore it has $\Lambda_i^*\phi^*$ as a solution so that (P_i^*) is stably set and has a solution by Theorem 5.7.

Q.E.D.

CHAPTER X

SUMMARY AND CONCLUSIONS

A study of linear optimal control problems with convex cost functions and convex constraints on controllers and responses has been carried out, based on some recent results in the theory of convex functions and their conjugates in topological vector spaces. The interesting aspects of modern convex analysis for applications in control theory are centered on the concepts of infinite valued functions and subgradients. Infinite valued convex functions make it possible to deal with constrained problems as if they were unconstrained by redefining the convex functions involved in such a way that they are finite on the constraint sets and infinite elsewhere. This leads to functions which are discontinuous on the boundaries of the constraint sets. The functions, however, are required to be lower-semi-continuous and this is a constructive property. A convex function which is finite valued on a non-empty convex set and infinite valued elsewhere can be made lower-semi-continuous if necessary by redefining its values on the boundary of this convex set. The concept of the subgradient of a convex function allows generalizations of results which only could be obtained under differentiability conditions. Theorem 4.1, the maximal principle of Pontryagin's type in generalized form is such a novel result.

The theory of convex functions in topological vector spaces is also the context of R.T. Rockafellar's elegant and general duality theory.

In this thesis, applications of this duality theory have been made to optimal control problems with ordinary, partial or functional differential equations, providing a unified theory for a wide range of problems.

The main results in control theory obtained from this duality theory are existence theorems for optimal controllers and Pontryagin's generalized maximal principle. The existence theorems in all the cases presented are novel because of the weaker conditions under which they apply (boundedness of the control constraint set rather than compactness).

To the author's knowledge, the maximal principle has not appeared in the literature in the form presented here, relying on subdifferentiability rather than on the stronger condition of differentiability. Duality theory also leads to a technique of decomposing large interconnected control problems with cost functions of separable type into smaller subproblems. Due to the symmetry of primal and dual problem, the subsystem structure of the primal is reflected in a similar structure of the dual. This is an advantage which is not obtained if the dual is defined in the calssical way by means of the Lagrangian function of the problem. Optimal controllers of the interconnected problem can be expressed in terms of optimal controllers of the subproblems. Existence theorems for optimality of the subproblems are given.

The results obtained in the thesis have engineering significance, specially for state constrained optimal control problems. Practical considerations often require that the response of the system is constrained to lie in some set after some time has elapsed. The C-H-J theory in Chapter III and its applications in Chapter IV have been given for the case that the constraint sets are time independent. The essential point of the theory is to deal with response constrained problems and it follows from the proofs of the theorems that extensions are readily made to those cases with time varying constraint sets. These include the situation mentioned above where the system is required to operate in such a way that its response lies in some convex set after some time has elapsed.

Of special practical interest is the bang-bang control problem because of the simplicity of its implementation in control systems. It is clear that the theory presented in the thesis does not allow a nonconvex set of admissible bang-bang controllers. However, theoretical investigation of certain problems shows that the bang-bang controller is optimal with respect to a wider class of admissible controllers. For instance consider the example of reachability on page 89 and specialize to the one dimensional case for simplicity. Let the class of admissible controllers U be given by

$$U = \{u \in L_r[E^1;0,T] \mid |u(t)| \leq 1 \text{ a.e. on } [0,T]\}$$

and let the response constraint set X be given by

$$X = \{\phi \in L_p[E^1;0,T] \mid \phi(0)=0,\ \phi(t) \in G_t \subseteq E^1 \text{ a.e. on } (0,T]\}$$

According to the maximal principle, a controller $\bar{u}$ has its response $\phi(\bar{u})$ in X (X is reachable) if there is an $\bar{\eta}$ in $L_q[E^1;0,T]$, such that

$$\bar{u}(t)\ B(t)\ \bar{\eta}(t) = \max_{|z| \leq 1} z\ B(t)\ \bar{\eta}(t) \quad \text{a.e. on } [0,T]$$

Or,

$$\bar{u}(t) = \text{sgn}[B(t)\bar{\eta}(t)], \qquad \text{a.e. on } [0,T]$$

and $\bar{\eta}$ satisfies

$$\dot{\eta} = -A(t)\eta + \bar{\phi}^* \quad , \qquad \bar{\eta}(T) = 0$$

where

$$\phi(\bar{u})(t)\bar{\phi}^*(t) = \min_{x \in G_t} x\ \bar{\phi}^*(t), \qquad \text{a.e. on } [0,T]$$

The results obtained in this thesis are believed to be of importance in the numerical solution of optimal control problems. In connection with extremum problems some numerical results on convex functions and subgradients have been obtained, see [4] and the references given there. Further investigation in this direction however, is needed as a next step in the applications of the results to practical problems.

APPENDIX I

SEMI-GROUPS AND SOLUTIONS OF LINEAR PARTIAL DIFFERENTIAL EQUATIONS

1. The Initial Value Problem

Initial value problems described by certain classes of linear partial differential equations give rise to linear bounded transformations with semi-group property. Intuitively this can be seen on hand of the following example. The ideas brought in here will be summarized more precisely later on.

Consider the linear partial differential equation, with time invariant coefficients

$$y_t(x,t) = a(x)y_{xx}(x,t)+b(x)y_x(x,t)+c(x)y(x,t), \quad t \geq 0 \qquad (A.1)$$
$$-\infty < x < +\infty$$

and suppose (A.1) together with a certain function class of initial values is well posed in the sense of Hadamard. Thus, to every initial value $f(x)$ in the function class there corresponds uniquely a solution $y(x,t)$ of (A.1) which, in some sense, tends to $f(x)$ when $t \to 0^+$. For a fixed t each $y(x,t)$ belongs to the same function class and may be considered as the image of $f(x)$ under a linear transformation $S(t)$, $y(x,t) = S(t)f(x)$. Then $y(x,t+s) = S(t+s)f(x)$ but if the solution y is unique, we may consider $y(x,s)$ as the initial value so that $y(x,t+s) = S(t)y(x,s)$. Therefore the linear transformation has to satisfy the semi-group property $S(t+s) = S(t)S(s)$ and $S(0) = I$, the identity transformation.

According this intuitive idea of a semi-group $\{S(t)|t \geq 0\}$ as a solution, one relates to the partial differential equation (A.1) a function space in which the solutions are required to lie and imposes the semi-group property $S(t+s) = S(t)S(s)$ together with the strong continuity condition

$$\lim_{h \to o^+} ||S(t+h)y - S(t)y|| = 0, \ \forall \ y$$

The connection between (A.1) and a particular semi-group is made up by the infinitesimal generator A of the semi-group where

$$Ay \equiv \lim_{h \to o^+} \frac{S(t+h)y - S(t)y}{h}$$

The domain D(A) of A consists of all y for which the limit exists. Thus if $y(x,t) = S(t)f(x)$ then

$$\frac{y(x,t+h)-y(x,t)}{h} = \frac{S(t+h)-S(t)}{h} f(x) = \frac{S(h)-I}{h} y(x,t)$$

For $y \varepsilon D(A)$ and $h \to o^+$ we find, at least formally, that

$$Ay = a(x)y_{xx} + b(x)y_x + c(x)y$$

These intuitive ideas lead one to consider a solution to the initial value problem for (A.1) as a linear transformation $y(x,t)=S(t)f(x)$ where S(t) is a semi-group generated by the differential operator

$A = a(x)\frac{\partial^2}{\partial x^2} + b(x)\frac{\partial}{\partial x} + c(x)$. The partial differential equation (A.1) is written as

$$\dot{y}(t) = Ay(t) \quad , \quad y(0) = f$$

and the initial value problem formulated as an abstract Cauchy problem in some function space in which the solutions y(x,t) are required to lie.

2. Summary of Semi-Group Theory

A family of linear bounded operators $\{S(t)|t > 0\}$ on a Banach space B to itself is called a semi-group if $S(t+s) = S(t)S(s)$ for $t, s > 0$.

We assume that S(t) is continuous in the strong operator topology for $\forall t > 0$. Then

$$w_o = \lim_{t\to\infty} \log||S(t)||/t < \infty$$

The semi-group is said to be of class C_o if

$$\lim_{t\to o^+} ||S(t)y-y|| = 0, \ \forall \ y \in B$$

which is written as $\lim_{t\to o^+} S(t)y = y, \ \forall \ y \in B$.

Then it follows that given $w > w_o$, there exists an $M > 0$ such that

$$||S(t)|| \leq M \exp(wt), \qquad t > 0$$

Further there is a set D dense in B such that for x in D,

$$\lim_{t \to o^+} \left|\left| \frac{S(t)x - x}{t} - y \right|\right| = 0$$

for some y in B. Of course y depends on the particular choice of x and writing y = Ax, x ε D,

$$\lim_{t \to o^+} \left|\left| \frac{S(t)x - x}{t} - Ax \right|\right| = 0$$

defines the linear operator A on D, the domain of A which we shall denote by D(A).

We write,

$$\lim_{t \to o^+} \frac{S(t)x - x}{t} \equiv Ax \quad , \quad \forall\, x \,\varepsilon\, D(A)$$

The linear operator A is called the infinitesimal generator of the semi-group S(t) and one proves that A is a closed operator, its resolvent $R(\lambda,A)$ exists and

$$R(\lambda,A)y = \int_0^\infty \exp(-\lambda t)S(t)y\,dt, \qquad \text{Re } \lambda > w_0$$
$$\forall\, y \,\varepsilon\, B$$

Further, $\forall\, y \in D(A)$, $\lim_{t \to 0^+} S(t)y = y$ and by definition $S(0) = I$.

Finally for some $y \in D(A)$ and $h > 0$ consider

$$\frac{1}{h}\{S(t+h)y-S(t)y\}= \frac{1}{h}\{S(h)-I\}S(t)y=S(t)[\frac{1}{h}\{S(h)y-y\}].$$

Because $y \in D(A)$ the right hand side converges to $S(t)Ay$ if $h \to 0^+$. Therefore the middle member converges so that

$$\forall\, y \in D(A) : S(t)y \in D(A) \qquad (A.2)$$

Hence the "right derivative" exists and satisfies

$$\forall\, y \in D(A): \frac{dS^+}{dt}(t)y=AS(t)y=S(t)Ay, \quad \forall\, t > 0$$

But for $0 < h < t$,

$$\frac{1}{h}\{S(t)y-S(t-H)y\} = S(t-h)\,\frac{S(h)y-y}{h}$$

and it follows that the "left derivative" exists and equals $S(t)Ay$. Thus,

$$\forall\, y \in D(A): \frac{dS}{dt}(t)y=AS(t)y=S(t)Ay, \quad \forall t > 0 \qquad (A.3)$$

Note that (A.2) makes it possible to continue the solution of a partial differential equation in time.

The derivative with respect to time as defined by (A.3) is taken in the strong topology of the Banach space B. Thus if A is identified with the differential operator equation (A.1), $\frac{\partial y}{\partial t}$ is a generalized partial derivative, namely, $\frac{\partial y}{\partial t} = \frac{d}{dt} S(t)f$, which is obtained by having A act on f according (A.3) and S(t) on D(A) provides a solution to the abstract Cauchy problem

$$\dot{y}(t) = Ay(t) \quad , \quad y(0) = f$$

3. <u>A General Abstract Cauchy Problem</u>

(ACP) Given a linear operator U with domain and range in a complex Banach space B and given an element y_0 in B, find a function $y(t) = y(t;y_0)$ such that,

(i) y(t) is strongly absolutely continuous and continuously differentiable in each finite subinterval of $[0,\infty)$

(ii) $y(t) \varepsilon D(U)$ and $U\{y(t)\} = \dot{y}(t)$, $t > 0$

(iii) $\lim_{t \to 0^+} y(t;y_0) = y_0$

The following two theorems are due to R.S. Phillips [22], and constitute necessary and sufficient conditions for (ACP) to have a unique solution.

<u>Theorem A.1</u>

If U is a closed operator with dense domain D(U) in B and U = A is the infinitesimal generator of class C_0, then (ACP) has a unique solution.

On the other hand,

Theorem A.2

Let U be a closed linear operator with dense domain D(U) in B and non-empty resolvent set. Suppose that for each $y_0 \in D(U)$ there is a unique solution to (ACP). Then U generates a semi-group S(t) of class C_0 such that

$$S(t)y_0 = y(t;y_0) \text{ for all } y_0 \in D(U)$$

REFERENCES

[1] Fenchel, W.: "Convex Cones, Sets, and Functions," Lecture Notes Princeton University, Spring 1951.

[2] Pearson, J.D.:"Reciprocity and Duality in Control Programming Problems", J.Math.Anal.Appl. 10, 1965,No.2, pp.388-468.

[3] Pearson, J.D.:"Dualtiy and a Decomposition Technique", J.SIAM, Series A, Control, 4, 1966, No. 1, pp.164-172.

[4] Pshenichnyi ,B.N.:"Linear Optimal Control Problems", J.SIAM, Series A,Control,4,1966,No.4,pp.577-593.

[5] Demyanov,V.F.,Rubinov,A.M.:"Minimization of Functionals in Normed Spaces", J.SIAM,Series A,Control,6,1968, No.1,pp.73-88.

[6] Brondsted,A.:"Conjugate Convex Functions in Topological Vector Spaces", Mat.Fys.Med.Dan.Vid.Selsk,34,1964,No.2, pp.1-27.

[7] Rockafellar,R.T.:"Convex Analysis", Lecture Notes, Princeton University, Spring 1966.

[8] Rockafellar,R.T.:"Duality and Stability in Extremum Problems Involving Convex Functions", Pac.J.Math.,21,1967, No.1,pp.167-187.

[9] Rockafellar,R.T.:"Integrals Which are Convex Functionals", Pac. J.Math.,24,1968,No.3,pp.525-539.

[10] Rockafellar,R.T.:"Conjugates and Legendre Transforms of Convex Functions", Canad.J.Math.,19, 1967, No. 1.

[11] Moreau,J.J.:"Proximite'et Dualite' dansune Espace Hilbertien", Bul.Soc.Math.France, 93,1965,pp.273-299.

[12] Bellman,R.:"Dynamic Programming and a New Formalism in the Calculus of Variations, I",Rivista Math.Univ.Parma, 6,1955,pp.193-213.

[13] Bellman,R.:"Functional Equations in the Theory of Dynamic Programming, V, Positivity and Quasilinearization", Proc.Natl.Acad.Sci., U.S.A.,41,1955,pp.743-746.

[14] Bellman,R., Cooke,R.:"Differential-Difference Equations", Acad. Press, New York, 1963.

[15] Halkin, H.:"The Principle of Optimal Evolution", Int.Symp. Nonl.Diff.Eqs. and Nonl.Mech., 1963, Eds. LaSalle, J.P., Lefschetz,S.

[16] Snow,D.R.:"Caratheodory-Hamilton-Jacobi Theory in Optimal Control," J.Math.Anal.Appl., 17,1967,pp.99-118.

[17] Lee, E.B.,Markus,L.:"Foundations of Optimal Control Theory", John Wiley, New York,1967.

[18] Kalaba,R.:"On Nonlinear Differential Equations, the Maximum Operation and Monotone Convergence", J.Math. and Mech.,8,1959,No.4,pp.519-574.

[19] Hille,E.:"Functional Analysis and Semi-Groups",Am.Math.Soc., 1948.

[20] Hille,E., Phillips,R.S.:"Functional Analysis and Semi-Groups", Am.Math.Soc., 1957.

[21] Phillips,R.S.:"Semi-Group Methods in the Theory of Partial Differential Equations", Mathematics for the Engineer, Second Series, pp.100-132, Ed. Beckenbach, E.F., McGraw-Hill, 1961.

[22] Phillips,R.S.:"A Note on the Abstract Cauchy Problem", Proc. Natl.Ac.Sci., U.S.A.,40,April 1954.

[23] Balakrishnan .,A.V.:"Optimal Control Problems in Banach Spaces", J.SIAM,Series A,Control,3,1965,No.1,pp.152-180.

[24] Hale,J.K.:"Linear Functional-Differential Equations with Constant Coefficients", Contr.Diff.Eqs,2,1963,pp.291-319.

[25] Mesarovic,M.D.:"A Conceptual Framework for the Studies of Multi-Level, Multi-Goal Systems", Report SRC 110-A-67-47, Case Western Reserve University.

[26] Mesarovic,M.D., Lefkowitz,I., Pearson,J.D.:"Advances in Multilevel Control", IFAC Symp.,1965, Tokyo, Japan.

[27] Lefkowitz,I:"Multilevel Approach Applied to Control Systems Design",Trans,A.S.M.E. ,Series D,88,1966,No.2,pp. 392-398.

[28] Takahara,Y.:"Multilevel Systems and Uncertainties", Report SRC 99-A-66-42, Case Western Reserve University.

[29] Takahara,Y.;"Multilevel Approach to Dynamic Optimization", Report SRC, 59-A-64-21, Case Western Reserve University.

[30] Lasdon, L.S.,:"Duality and Decomposition in Mathematical Programming", Trans.,IEEE, SSC-4,1968,No.2,pp.86-100.