October, 1979 LIDS-TH-949

STOCHASTIC OPTIMIZATION FOR DISCRETE-TIME SYSTEMS

by

GREGORY S. LAUER

This report is based on the unaltered thesis of Gregory S. Lauer, submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at the Massachusetts Institute of Technology in September, 1979. The research was conducted at the M.I.T. Laboratory for Information and Decision Systems and was supported under National Science Foundation Grant No. NSF/ENG-77-19971.

Laboratory for Information and Decision Systems
Massachusetts Institute of Technology
Cambridge, MA 02139

STOCHASTIC OPTIMIZATION FOR DISCRETE-TIME SYSTEMS

by

GREGORY S. LAUER

B.S., University of Notre Dame

(1975)

M.S., University of Notre Dame

(1976)

SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE
DEGREE OF

DOCTOR OF PHILOSOPHY

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September, 1979

Signature of Author..
Department of Electrical Engineering,
September 24, 1979

Certified by..
Thesis Supervisor

Accepted by..
Chairman, Departmental Committee on Graduate Students

STOCHASTIC OPTIMIZATION FOR DISCRETE-TIME SYSTEMS

BY

GREGORY S. LAUER

Submitted to the Department of Electrical Engineering

on September 24, 1979 in partial fulfillment of the requirements

for the Doctor of Philosophy

ABSTRACT

This thesis considers the problem of the stochastic control of discrete-time systems for which there exist explicit constraints on the state variables and on the control laws. The presence of the control constraints implies that stochastic dynamic programming cannot be applied in a straightforward manner. Necessary conditions for the optimality of a control policy are thus obtained by deriving maximum principles for a finite time horizon equivalent deterministic problem.

The state variable constraints are incorporated into the maximum principle derivation by formulating the equivalent deterministic problem as a nonlinear programming problem in Banach spaces. The solution of this problem requires the extension of results known from nonlinear programming in finite dimensional spaces.

The extension of this problem formulation to infinite time horizon problems almost always precludes the existence of a stationary feasible control law. A new problem formulation is thus introduced for which an optimal stationary control policy exists. Necessary conditions for the optimality of a control policy are derived in a manner similar to that in the finite time horizon problem.

THESIS SUPERVISOR: Nils R. Sandell, Jr.

TITLE: Associate Professor

ACKNOWLEDGEMENTS

I want to thank my advisor Professor Nils Sandell for his constant guidance and patience without which this thesis would never have been completed. I want to thank my readers, Professors Dimitri Bertsekas and Michael Athans and Dr. David Castañon for their insightful comments and guidance. They have helped place my work in perspective and have helped keep me on track.

I also want to thank all the graduate students who have helped make my stay at M.I.T. enjoyable and instructive, in particular Khaled Yared and David Rossi. Lastly I want to thank my wife Janet for her love and patience which have helped make this thesis possible.

This research was conducted at the M.I.T. Laboratory for Information and Decision Systems and was supported under National Science Foundation Grant No. NSF/ENG-77-19971.

To my parents

TABLE OF CONTENTS

CHAPTER I

Introduction and Overview

Section I.0. Introduction

The basic problem that we shall consider in this thesis can be described as follows. We are given a fixed dynamical system and are allowed to make noise corrupted, partial observations of its state. Our task is to design a control strategy which chooses inputs to the system based on the observations previously made so as to force the system to behave in a desired manner.

Typically this problem is attacked by developing a mathematical model of the dynamical system. Unpredictable disturbances are modeled as random inputs to the dynamical system with known probability distributions. A cost is associated with each sequence of state and control values. Since there are random inputs into the dynamical system, for each control strategy the cost is a random variable. Our task reduces to minimizing the expected value of the cost over the set of all acceptable control strategies.

While this approach, known as stochastic optimal control, has been successful in producing analytical solutions to certain classes of problems (most notably LQG problems) its fundamental merit lies in that it provides a structured approach to the solution of the basic problem. In the many cases where no analytical solutions are available this approach allows one to gain insight into what good control strategies are like and it allows one to evaluate competing suboptimal control strategies.

There are, however, several problems associated with this approach. One is that it is quite difficult to model complex systems in a way in which the mathematical model leads to a tractable problem. Often this problem can be alleviated by allowing side constraints to be placed on the state of the mathematical model. In the problem formulation of Chapter III we shall allow explicit state constraints.

Another problem that occurs is that the "optimal control strategy" is undesirable because some physical constraint could not be incorporated into the mathematical formulation. For example, it may be that the control strategy is to be implemented on a small real time computer and thus the optimal control strategy must be simple enough to be implementable. Typically this problem is handled by specifying a priori a set of acceptable control strategies. The problem formulation of Chapter III will also allow one to place restrictions on the set of acceptable control strategies.

Also we note that it is often difficult to assign compatible costs to all the criteria with respect to which one wishes to optimize. For example, it may not be clear how to incorporate into a single cost function economic efficiency and environmental pollution. This problem also occurs when there are several competing criteria (which may be "compatible") for which the weighting is determined in a nonmathematical manner (e.g., by voting or committee discussion). To handle this type of problem we will allow vector valued objectives. The optimization (considered in section 3.9) now serves as a preprocessor which eliminates inferior solutions (that is, solutions for which there exists a strategy which produces a lower cost in all criteria).

Finally we note that will assume that the mathematical model has been formulated in discrete-time. We do this since it allows the control strategy to be implemented directly on a digital computer.

Section 1.1. Related Literature

The approach we shall take in attacking the problem just formulated is to reduce the basic (stochastic) problem to an equivalent deterministic problem by considering the evolution of the probability measure of the state variables. Necessary conditions for optimality will be developed by deriving a maximum principle for this equivalent deterministic problem. We are thus interested in two bodies of literature: that on the maximum principle for discrete-time deterministic systems and that on nonlinear mathematical programming in infinite dimensional spaces. The second body of literature is of interest because the maximum principle will be derived via a reformulation of the dynamical problem as a static mathematical programming problem in which the variables (probability measures) may be infinite dimensional.

A maximum principle for discrete-time systems which are described by the sum of term linear in the state and a term nonlinear in the control was first derived by Rozonoér [Roz 1]. He was also the first to point out the difficulty in directly applying the maximum principle of Pontryagin et. al. [Pon 1] to discrete-time systems nonlinear in the control. Halkin [Hal 1] seems to have been the first to give a rigorous and correct derivation of the discrete-time maximum principle for systems nonlinear in the state. The assumption under which this derivation was valid was that of the convexity of certain reachable sets. This condition was later weakened to that of directional convexity by Holtzman [Hol 1]. The concept of directional convexity was extended so as to be applicable to problems with a vector valued objective by DaCuna and Polak [DaC 1].

The problem of deriving a maximum principle for nonlinear discrete systems with phase (state) constraints was considered by several people. The most general formulation seems to be that of Cannon, Cullum and Polak [Ca 1] and [Ca 2].

A related approach which does not make the convexity assumptions but which makes certain differentiability assumptions was considered in Jordan and Polak [Jo 1]. This approach is subsumed by the results in [Ca 1].

None of the above considered the case in which the space was infinite dimensional. In Chapter III we extend some of the results of [Ca 1] to arbitrary Banach spaces. Since the problem formulation of [Ca 1] is that of a static mathematical programming problem we now consider infinite dimensional mathematical programming.

The Lagrange multiplier rule was extended to arbitrary spaces by Lynsternik and Sobolev [Ly 1]; the multiplier rule of Fritz John for inequality constraints was considered in arbitrary spaces by Hurwicz [Hur 1]. Since the requirements for separating convex cones are stricter in infinite dimensional spaces the multiplier rule for equality and inequality constraints of Mangasarian and Fromovitz [Man 2] was not immediately extended to Banach spaces.

Ioffe and Tihomirov [Io 1] consider the case in which the inequality constraints are defined in a finite dimensional space and the equality constraint is defined in an infinite dimensional space. The case in which the inequalities are defined in an infinite dimensional space and the equalities are defined in a finite dimensional space is considered by Halkin [Hal 3] and Dubovitskii and Mityutin [Dub 2]. Other related cases are considered in [Var 1], [Nag 1] and [Las 1].

Girsanov [Gir 1] extends [Dub 2] to handle infinite dimensional equality constraints but makes a restrictive assumption on set constraints. The problem with infinite dimensional equality and inequality constraints has been considered by Bazaraa and Goode [Baz 1], Neustadt [Neu 2], and many others (e.g., [Io 2], [Kur 1], [Tu 1], [Zo 1] and [Hal 2]). The work of Neustadt [Neu 1] subsumes almost all of these results and many of those on the maximum principle. He does not consider arbitrary linear equalities but the restrictions placed are weak enough that most interesting problems satisfy them.

We do note, however, that the assumptions made by Neustadt in [Neu 1] on certain set constraints are stronger than those of [Ca 1].

Section 1.2. Summary of Thesis

In Chapter II we extend the results of [Ca 1] to arbitrary Banach spaces. Here we concern ourselves only with nonlinear mathematical programming results of [Ca 1]; they will be applied to our basic problem in Chapter III. While it is possible to weaken the assumptions of [Neu 1] and obtain slightly more general results than we do in this chapter it requires considerably more sophisticated proofs. Thus for clarity we follow the format of [Ca 1].

In this chapter we also introduce the notion of K-linear independence, which is a natural extension of positive linear independence. The results of [Ca 1], derived under an assumption of linear independence of the active inequality gradients, hold under an assumption of positive linear independence. Thus our extension yields conditions weaker than those in [Ca 1].

In Chapter III, section one we introduce our problem formulation and derive an equivalent deterministic problem based on [Wit 2] and [San 1]. In section two we show that the problem formulation is quite general by reducing a variety of problems to that form. In section three we consider the relationship of a maximum principle for an unconstrained (in the state) system to the dynamic programming algorithm. In section four we compare our formulation with that of previous work. In sections five and six we extend maximum principle of [Ca 1] and apply them to our problem formulation. Extensions of these results are considered in section eight and the vector valued objective function is considered in section nine. Some concluding remarks are made in section ten.

Chapter IV, section one contains the extension of the problem to the infinite time horizon. An average cost per stage problem formulation is introduced in section two and we extend results for finite and countable state spaces to this formulation. In section three we introduce a new problem formulation, the steady state constrained problem formulation. A maximum principle is derived for this problem formulation. In section four we make some concluding remarks. Chapter V summarizes the thesis and contains some direction for future research.

Section 1.3. Contributions of the thesis

The following are felt to be the most significant contributions of this thesis.

1. Extension of the results of Canon, Cullum and Polak to Banach spaces.
2. Introduction of the notion of positive linear independence to Banach spaces.
3. Extension of the results of Witsenhausen and Sandell to problems with explicit state constraints.
4. Extension of the results for the average cost per step problem formulation from countable state spaces to Euclidean state space.

5. Introduction of the steady state constrained problem formulation.
6. Derivation of maximum principles for the steady state constrained problem formulation.

CHAPTER II

NONLINEAR PROGRAMMING

Section 2.0. Introduction

In this chapter we consider necessary conditions for the optimality of a solution to certain types of mathematical programming problems. The terminology, notation and basic results that we shall be using are contained in Appendix A.

We consider problems of the form:

$$\min f(x) \tag{P}$$

subject to

$$R(x) = 0 \quad G(x) \in K \quad x \in \Omega$$

where $R: X \to Y$, $G: X \to Z$ K, $f: X \to \mathbb{R}$,

$\Omega \in X$, and X, Y and Z are Banach spaces.

Necessary conditions that any $\hat{x}$ solving (P) must satisfy can be formulated in many ways, however we are looking for conditions of the following form:

$$\lambda < \nabla f(\hat{x}), \delta x> + <\underline{\phi}, \nabla R(\hat{x})\delta x> + <\underline{\psi}, \nabla G(\hat{x})\delta x> \leq 0$$
$$\forall \delta x \in P(\Omega). \tag{N}$$

Here $\nabla h(\hat{x})$ indicates the Fréchet derivative of $h(x)$ at $\hat{x}$, $\lambda \in \mathbb{R}$, $\underline{\phi} \in Y^*$, $\underline{\psi} \in Z^*$ where Y^* and Z^* are the dual spaces to Y and Z (see Appendix A), and $P(\Omega)$ is some transformation of the set Ω.

If the spaces Y and Z are finite dimensional then a triple $(\lambda, \underline{\phi}, \underline{\psi})$ satisfying (N) exist under very weak assumptions on the set Ω. If the spaces are infinite dimensional then in general additional assumptions are required. In this chapter several different sets of such assumptions will be considered. In spite of the differences between assumptions, the proofs of the theorems all consist of the following steps:

(i) convex approximations to various sets in (P) are constructed,

(ii) it is shown that if these sets have a nonempty intersection then $\hat{x}$ is not optimal, and

(iii) it is shown that if these sets have an empty intersection then they can be separated, which implies conditions of the form (N).

Two different types of assumptions will be made in this chapter. The first type is concerned with the existence of convex approximations and their separability. These assumptions are made primarily in Section 2.2 and are of similar nature to those made in [Ku 1], [Gi 1] and [Neu 1] for example.

The second type of assumption is concerned with the existence of hyperplanes of a particular type, ones for which, in the context of (N), $\lambda \neq 0$. The concept of K-linear independence, introduced in Section 2.3, is an assumption of this type. This concept is an extension of positive linear independence [Man 1] to arbitrary Banach spaces and leads to weaker conditions than the usual notion of linear independence.

While the existence of hyperplanes for which $\lambda \neq 0$ is important we shall not dwell long on this issue. The reason for this is twofold.

(i) it is frequently easier to attempt to find a set of multipliers with $\lambda \neq 0$ than to check that the assumptions are satisfied, and

(ii) most assumptions are too strong, that is, they imply every triple $(\lambda,\underline{\phi},\underline{\psi})$ satisfying (N) has $\lambda \neq 0$ when only one such triple is required.

In Sections 2.1 through 2.3 we consider problem (P) under various assumptions on Ω and $R(x)$. While the results are in some cases only slight extensions of existing results the proofs are new and simpler than existing ones. Consequently, the extension of the results in Sections 2.1 through 2.3 to nonscalar objective functions in Section 2.4 is very straightforward.

The results in Sections 2.1 through 2.3 are extensions of the results in [Ca 1] to Banach spaces. Section 2.4 extends the results of [Dac 1] from Euclidean space to arbitrary Banach spaces and the results of [Ku 1] and [Zo 1] from scalar optimization to vector valued optimization.

Section 2.1. Mathematical Programming

In this section we consider first order necessary conditions for the optimality of $\hat{x}$ in the following problem:

$$\min f(x) \tag{2.1.1}$$

subject to

$$r(x) = 0 \qquad x \in \Omega \subset X$$

where $f\colon\ X \to \mathbb{R}$, $r\colon\ X \to \mathbb{R}^n$, X is a Banach space and $f(x)$ and $r(x)$ are continuously Fréchet differentiable on X with derivatives $\nabla f(x)$ and $\nabla r(x)$.

The following notation will be useful:

(i) $\text{co}\{\delta x_1,\ldots,\delta x_k\}$ is the convex hull of the set $\{\delta x_1,\ldots,\delta x_k\}$ (see Appendix A),

(ii) $A - x = \{z \mid z = y-x,\ y \in A\}$, and

(iii) $o(x)$ denotes a function such that

$$\lim_{\|x\|\to 0} \|o(x)/\|x\|\| = 0 .$$

Recall that necessary conditions are to be derived by showing that the optimality of $\hat{x}$ implies the separation of certain convex sets.

Those convex sets, based on definitions in [Ca 1] and [Dac 1], will now be introduced.

A <u>quasilinear conical approximation</u> to the set Ω at $\hat{x}$, denoted $C(\hat{x},\Omega)$, is defined as any convex cone such that for any finite collection $\{\delta x_1,\ldots, \delta x_k\}$ of linearly independent vectors in $C(\hat{x},\Omega)$ there exists an $\epsilon^* > 0$ and a continuously Fréchet differentiable map ζ from $co\{\epsilon\delta x_1,\ldots,\epsilon\delta x_k\}$, for $0 < \epsilon \leq \epsilon^*$, into $\Omega - \hat{x}$ of the form

$$\zeta(\delta x) = \delta x + o(\delta x) \quad \forall \delta x \in co\{\epsilon\delta x_1,\ldots,\epsilon\delta x_k\}, \qquad (2.1.2)$$

$$0 < \epsilon \leq \epsilon^*$$

with $\zeta'(0) = I$.

A <u>conical approximation</u> $C(\hat{x},\Omega)$ to Ω at $\hat{x}$ is a quasilinear conical approximation for which $\zeta(\delta x) = \delta x$. Both of these sets can be considered as "first order" approximations to Ω at $\hat{x}$ since if $\delta x \in C(\hat{x},\Omega)$ then $\hat{x} + \epsilon\delta x$ is in Ω or "almost" in Ω for ϵ small enough.

We shall also need the following sets:

(i) $K(\hat{x}) = F(\hat{x})C(\hat{x},\Omega)$ where $F(x) = \begin{bmatrix} f(x) \\ r(x) \end{bmatrix}$ (2.1.3)

(ii) $K_e(\hat{x}) = \nabla r(\hat{x})C(\hat{x},\Omega)$ and (2.1.4)

(iii) $D = \{z \in \mathbb{R}^{n+1} | z = \beta(-1,0,\ldots,0),\ \beta > 0\}$. (2.1.5)

Their interpretation will be considered in the proof of the following theorem.

<u>Theorem 2.1.1</u> Let $\hat{x}$ be an optimal solution to problem (2.1.1). For any $C(\hat{x},\Omega)$ which is a quasilinear conical approximation to Ω at $\hat{x}$, there exists a nonzero vector $(\lambda,\psi_1,\ldots,\psi_n) \in \mathbb{R}^{n+1}$ with $\lambda \leq 0$ such that

$$\lambda < \nabla f(\hat{x}),\delta x> + <\underline{\psi},\nabla r(\hat{x})\delta x> \leq 0 \quad \forall \delta x \in \overline{C(\hat{x},\Omega)}. \qquad (2.1.6)$$

<u>Proof:</u> We outline the proof here. For more details see Appendix

B.

The strategy in this proof is to show that if D and $K(\hat{x})$ are not separable then $\hat{x}$ is not optimal and if they are separable then equation (2.1.6) follows. If $D \cap K(\hat{x}) \neq \phi$ then there exists a $\delta x \in C(\hat{x},\Omega)$ such that $\nabla r(\hat{x})\delta x = 0$ and $\nabla f(\hat{x})\delta x < 0$. Thus if ϵ is small enough $\hat{x} + \epsilon\delta x$ is almost in Ω, $r(\hat{x}) + \epsilon\nabla r(\hat{x})\delta x = 0$, and $f(\hat{x}) + \epsilon\nabla f(\hat{x})\delta x < f(\hat{x})$. Thus to "first order", $D \cap K(\hat{x}) \neq \phi$ contradicts the optimality of $\hat{x}$.

The actual contradiction is produced by showing that if D and $K(\hat{x})$ are not separable then

(i) $D \cap K(\hat{x}) \neq \phi$ and

(ii) $K_e(\hat{x}) = \mathbb{R}^n$.

The second fact is used to generate a "second order correction" $o(\delta x)$ such that for $x = \hat{x} + \epsilon\delta x + o(\epsilon\delta x)$, $r(x) = 0$ and $x \in \Omega$ for ϵ small enough. We then use the fact that the correction is of second order to show that, for ϵ small enough, $f(x) < f(\hat{x})$. This contradiction of the optimality of $\hat{x}$ implies D and $K(\hat{x})$ are separable.

The separability of D and $K(\hat{x})$ implies that there exists a $\underline{\phi} \in \mathbb{R}^{n+1}$, not zero, such that

$$\langle\underline{\phi},y\rangle \geq 0 \geq \langle\underline{\phi},z\rangle \quad \forall y \in D, \forall z \in K(\hat{x}) \tag{2.1.7}$$

which then implies, by equation (2.1.3), that $\lambda \leq 0$ and

$$\lambda \langle \nabla f(\hat{x}),\delta x\rangle + \langle\underline{\psi},\nabla r(\hat{x})\delta x\rangle \leq 0 \quad \forall \delta x \in C(\hat{x},\Omega), \tag{2.1.8}$$

where $\underline{\phi} = (\lambda,\underline{\psi})$. The continuity of $\nabla f(\hat{x})$ and $\nabla r(\hat{x})$ then imply equation (2.1.6).

In Chapter III we shall need to be able to consider nonlinear programming problems with operator equality constraints where the operator is a mapping between infinite dimensional Banach spaces. Thus we now extend Theorem 2.1.1 to problems which include linear operator equalities. Consider the following problem:

$$\min f(x) \tag{2.1.9}$$

subject to

$$r(x) = 0 \qquad Tx = 0 \qquad x \in \Omega \subset X$$

where f, r and X are as before (problem (2.1.1)) and T is a linear map from X to Y, a Banach space.

Assumption 2.1.1 $\mathcal{R}(\nabla T(\hat{x}))$ is closed.

Theorem 2.1.2 If $\hat{x}$ is an optimal solution to problem (2.1.40), $C(\hat{x},\Omega)$ is a conical approximation to Ω at $\hat{x}$ and Assumption 2.1.1 holds, then there exists a $\lambda \in \mathbb{R}$, $\lambda \leq 0$, and $\underline{\psi} \in \mathbb{R}^n$ and a $y^* \in Y^*$, not all zero, such that

$$\lambda\langle\nabla f(\hat{x}),\delta x\rangle + \langle\underline{\psi},\nabla r(\hat{x})\delta x\rangle + \langle y^*,T\delta x\rangle \leq 0 \quad \forall \delta x \in \overline{C(\hat{x},\Omega)} \tag{2.1.10}$$

Proof: We will reduce problem (2.1.9) to one in the same form as problem (2.1.1) and apply Theorem 2.1.1 to deduce equation (2.1.10). We outline the proof here, for the full proof see Appendix B.

An equivalent problem for that given in (2.1.9) is

$$\min f(x) \tag{2.1.11}$$

subject to

$$r(x) = 0 \qquad x \in \Omega \cap \mathcal{N}(\nabla T(\hat{x})),$$

where $\mathcal{N}(\nabla T(\hat{x}))$ is the null space of $\nabla T(\hat{x})$ (see Appendix A). Since T is linear $\mathcal{N}(\nabla T(\hat{x})) = \mathcal{N}(T)$. We then show that if $C(\hat{x},\Omega)$ is a conical approximation to Ω, $C(\hat{x},\Omega) \cap \mathcal{N}(T)$ is a conical approximation to $\Omega \cap \mathcal{N}(T)$. Theorem 2.1.1 then implies the existence of a $\lambda \leq 0$ and $\underline{\phi}$, not both zero such that

$$\lambda\langle\nabla f(\hat{x}),\delta x\rangle + \langle\underline{\phi},\nabla r(\hat{x})\delta x\rangle \leq 0 \quad \forall \delta x \in C(\hat{x},\Omega) \cap \mathcal{N}(T) \tag{2.1.12}$$

Let P_N be the projection of X onto $N(T)$. We are able to argue that (2.1.12) implies that there exist λ', ϕ' such that

$$\lambda'\langle\nabla f(\hat{x}),P_N\delta x\rangle + \langle\phi',\nabla r(\hat{x})P_N\delta x\rangle \quad \forall\delta x \in C(\hat{x},\Omega) \leq 0 \tag{2.1.13}$$

Equation (2.1.13) can be rewritten

$$\begin{aligned}&\lambda'\langle\nabla f(\hat{x}),\delta x\rangle + \langle\underline{\phi}',\nabla r(\hat{x})\delta x\rangle \\ &- \lambda'\langle\nabla f(\hat{x}),P_M\delta x\rangle - \langle\underline{\phi}',\nabla r(\hat{x})P_M\delta x\rangle \quad \forall\delta x \in C(\hat{x},\Omega) \leq 0\end{aligned} \tag{2.1.14}$$

where $P_N = I - P_M$. The assumption that $R(T)$ is closed leads to the fact that P_M is continuous.

Using the fact that $P_M x = 0$ for all $x \in N(T)$ we can show that

$$\lambda' P_M^*\nabla f(\hat{x}) + P_M^*\nabla r^*(\hat{x})\underline{\phi}' \in [N(T)]^{\perp} = R(T^*). \tag{2.1.15}$$

This implies that equation (2.1.14) can be rewritten

$$\lambda'\langle\nabla f(\hat{x}),\delta x\rangle + \langle\underline{\phi}',\nabla r(\hat{x})\delta x\rangle + \langle y^*,T\delta x\rangle \leq 0 \qquad \forall\delta x \in \overline{C(\hat{x},\Omega)} \tag{2.1.16}$$

which is the desired result.

We shall discuss the assumption that $R(T)$ is closed and why it is reasonable in Section 2.5. If, however, $R(T)$ is not closed we still have the following.

<u>Corollary 2.1.1</u> If $R(T)$ is not closed then the results of Theorem 2.1.2 hold but y^* need not be continuous.

<u>Proof:</u> The same as that of Theorem 2.1.2 except that P_N and P_M are not continuous so that y^* need only be a linear functional on Y (it need not be continuous).

It is sometimes useful to have conditions for $\lambda < 0$. As discussed in Section 2.0 we shall not consider this issue in detail, however, we have:

<u>Corollary 2.1.2</u> If $\langle\nabla f(\hat{x}),\delta x\rangle \geq 0$, $\forall\delta x \in C$ where

$$C = \{\delta x \mid T\delta x = 0,\ \nabla r(\hat{x})\delta x = 0,\ \forall\delta x \in \overline{C(\hat{x},\Omega)}\}$$

and all the assumptions of Theorem 2.1.2 hold then there exist $\lambda < 0$, $\psi \in \mathbb{R}^n$ and $y^* \in Y^*$ (the space of continuous linear functionals on Y) such that equation (2.1.10) holds.

Proof: This assumption clearly implies that

$$D \not\subset \overline{K(\hat{x})} \qquad (2.1.17)$$

where $K(\hat{x}) = \nabla F(\hat{x})(\overline{C(\hat{x},\Omega)} \cap N(T))$. Thus D and $K(\hat{x})$ can be strictly separated so that λ can be chosen strictly less than zero.

The following extensions of Theorem 2.1.2 will prove useful in Chapter III. They allow one to apply the results of Theorem 2.1.2 to problems for which it is not easy to find a conical approximation $C(\hat{x},\Omega)$.

Corollary 2.1.3 Suppose that x is a feasible point for problem (2.1.9), that is, $x \in \Omega$, $r(x) = 0$, and $Tx = 0$. If

(i) $\quad D \cap K(x) \neq \phi$ and $\qquad$ (2.1.18)

(ii) $\quad K_e(x) = \mathbb{R}^n$ $\qquad$ (2.1.19)

then there exists an $\tilde{x} \in \Omega$ such that $r(\tilde{x}) = 0$, $T\tilde{x} = 0$ and $f(\tilde{x}) < f(x)$.

Proof: This is proved in the course of proving theorems 2.1.1 and 2.1.2. The only difference is that we have not assumed x is optimal so this result does not contradict any assumptions.

Corollary 2.1.4 Let Ω' be any set such that if $x' \in \Omega'$ then there exists an $\bar{x} \in \Omega$ such that

(i) $\quad r(x) = r(x')$, $Tx = Tx'$, and $\qquad$ (2.1.20)

(ii) $\quad f(x) \leq f(x')$. $\qquad$ (2.1.21)

If $\hat{x} \in \Omega'$ is a solution of problem (2.1.1) and $C(\hat{x},\Omega')$ is a conical approximation to the set Ω' at $\hat{x}$ then the results of Theorem 2.1.2 hold for all $\delta x \in \overline{C(\hat{x},\Omega')}$.

Proof: Assume that

(i) $D \cap K(\hat{x}) \neq \phi$, and (2.1.22)

(ii) $K_e(\hat{x}) = \mathbb{R}^n$ (2.1.23)

where $K(\hat{x})$ and $K_e(\hat{x})$ are defined using $C(\hat{x},\Omega') \cap N$. Then since $\hat{x}$ is feasible, Corollary 2.1.3 implies that there exists an $x' \varepsilon \Omega$ such that $r(x') = 0$, $Tx' = 0$ and $f(x') < f(\hat{x})$. But by assumption this implies that there exists an $x^* \varepsilon \Omega$ such that $r(x^*) = r(x') = 0$ and $f(x^*) \leq f(x')$, so that

(i) $f(x^*) < f(x)$ (2.1.24)

(ii) $r(x^*) = 0$ (2.1.25)

(iii) $x^* \varepsilon \Omega$. (2.1.26)

This contradicts the optimality of $\hat{x}$, thus the assumptions (2.1.22) and (2.1.23) are not true and the result of the theorem follows from the proof of Theorems 2.1.1 and 2.1.2.

Corollary 2.1.5 If problem (2.1.9) is modified so that $T(x) = \tilde{T}x + t = 0$, that is so $T(x)$ is affine, then Theorem 2.1.2 holds for this new problem with equation (2.1.10) replaced by

$$\lambda\langle\nabla f(\hat{x}),\delta x\rangle + \langle\underline{\psi},\nabla r(\hat{x})\delta x\rangle + \langle y^*,\tilde{T}\delta x\rangle \leq 0 \quad \forall \delta x \ \varepsilon \ \overline{C(\hat{x},\Omega)}. \quad (2.1.27)$$

Proof: follows from the proof of Theorem 2.1.2 and the fact that $N(\tilde{T})$ is a conical approximation for the set

$$N = \{x | T(x) = 0\} . \quad (2.1.28)$$

Section 2.2. Extensions

In this section we will consider an extension of Theorem 2.1.2 to a problem with infinite dimensional nonlinear equality constraints. Consider the following problem:

$$\min f(x)$$

subject to (2.2.1)

$$T(x) = 0$$

$$x \in \Omega$$

where $f\colon X \to \mathbb{R}$, $T\colon X \to Y$, X and Y Banach spaces and f and T Fréchet differentiable.

By analogy with Theorem 2.1.2 and finite dimensional problems one might imagine that if $\hat{x}$ solves (2.2.1) then there would exist a $\lambda \leq 0$, $y^* \in Y^*$ <u>not both zero</u> such that

$$\lambda\langle\nabla f(\hat{x}),\delta x\rangle + \langle y^*, \nabla T(\hat{x})\delta x\rangle \leq 0 \quad \forall \delta x \in \overline{C(\hat{x},\Omega)} \tag{2.2.2}$$

where $C(\hat{x},\Omega)$ is a conical approximation to Ω at $\hat{x}$. As the following example shows however, equation (2.2.2) does not always hold true. Consider the problem in $X = \ell_2$,

$$\min f(\underline{x}) = \tfrac{1}{4}(x_1+1)^4 + \sum_{i=1}^{\infty} x_i^2 \qquad f\colon X \to \mathbb{R}$$

subject to

$$T(\underline{x}) = \begin{bmatrix} x_2 \\ \vdots \\ x_n \\ \vdots \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ \vdots \end{bmatrix} \qquad T\colon X \to X \tag{2.2.3}$$

$$\underline{x} \in \Omega = \{\underline{x} \mid x_1 \leq -n^2|x_n|,\ n > 1\}.$$

Clearly the minimum occurs at $\underline{x} = \underline{0}$ since $T(\underline{x}) = \underline{0}$ and $\underline{x} \in \Omega$ imply that $x_1 \leq 0$ and $x_n = 0, n = 2,\ldots$ but $f(\underline{x})$ is an increasing function in $|x_1|$.

To obtain equation (2.2.2) we must separate the sets D and $K(\hat{x})$ where

$$D = \{\underline{x} \mid x_1 < 0,\ x_n = 0,\ n = 2,\ldots\} \tag{2.2.4}$$

$$K(\hat{x}) = \nabla F(\hat{x}) C(\hat{x},\Omega), \tag{2.2.5}$$

$$F(\hat{x}) = \begin{bmatrix} f(\hat{x}) \\ T(\hat{x}) \end{bmatrix}. \tag{2.2.6}$$

Note that $\mathcal{R}(\nabla T(\hat{x})) = X$.

Now Ω is a closed convex cone so, since $\hat{x} = 0$, $C(\hat{x},\Omega) = \Omega$. Note that $D \subset K(\hat{x}) = \Omega$.

If (λ, y^*) is to separate D and $K(\hat{x})$ (which one might hope is possible since int $K(\hat{x}) = \phi$) then it must be that

$$<(\lambda,y^*),d> = 0 \leq <(\lambda,y^*),z> \ \forall d \varepsilon D \ \forall z \ \varepsilon \ K(\hat{x}) \ . \tag{2.2.7}$$

However, (2.2.7) implies that $\lambda = 0$ and so y^* must separate 0 and $\nabla T(\hat{x}) C(\hat{x},\Omega)$. But $\nabla T(\hat{x}) C(\hat{x},\Omega)$ is dense in X and thus has no support points. The conclusion is that D and $K(\hat{x})$ cannot be separated.

The basic problem is that, while closed convex subsets have support points dense in their boundary, it is possible to construct problems in which the ray of decreasing cost does not intersect any support points of the feasible set even though the ray is in the boundary of the set. Clearly additional assumptions must be made concerning the functions and sets in problem (2.2.1). Initially we shall assume Ω is a convex set with nonempty interior. Later, in corollaries 2.2.1 and 2.2.2 we shall weaken this assumption.

Consider the following problem:

$$\min f(x) \tag{2.2.8}$$

subject to

$$r(x) = 0 \qquad T(x) = 0 \qquad x \in \Omega$$

where $f\colon X \to \mathbb{R}$, $T\colon X \to Y$, $r\colon X \to \mathbb{R}^n$, X and Y are Banach spaces, f, r, and T are continuously Fréchet differentiable on X with derivatives $\nabla f(x)$, $\nabla r(x)$ and $\nabla T(\hat{x})$, and $\Omega \subset X$ is a set such that $\text{int}\,\Omega \neq \phi$ and $\overline{\Omega} \neq X$.

Define [Ca 1] the <u>radial cone</u>, $RC(\hat{x},\Omega)$, to the set Ω at $\hat{x}$ by

$$RC(\hat{x},\Omega) = \{\delta x \mid \exists \alpha > 0 \ni x + \alpha\delta x \in \Omega\} . \tag{2.2.9}$$

Note that, since Ω is convex $RC(\hat{x},\Omega)$ is a convex cone.

<u>Theorem 2.2.1</u> If $\hat{x}$ is a solution to problem (2.2.8), $\mathcal{R}(\nabla T(\hat{x}))$ is closed, Ω is convex, $\text{int}\,\Omega \neq \phi$, and $RC(\hat{x},\Omega)$ is a radial cone to the set Ω at $\hat{x}$ then there exists a $\lambda \leq 0$, $\underline{\phi} \in \mathbb{R}^n$ and a $y^* \in Y^*$, not all zero such that

$$\lambda\nabla\langle f(\hat{x}),\delta x\rangle + \langle\underline{\phi},\nabla r(\hat{x})\delta x\rangle + \langle y^*,\nabla T(\hat{x})\delta x\rangle \leq 0 \qquad \forall \delta x \in \overline{RC(\hat{x},\Omega)} \tag{2.2.10}$$

<u>Proof</u>: We will outline the proof here, for more details see Appendix B.

The proof of this theorem is very similar to that of Theorem 2.1.2. The only significant difference is that we use Luisternik's theorem to show that $\nabla T(\hat{x})$ is a quasilinear conical approximation to the set $A = \{x \mid T(x) = 0\}$. We then rewrite problem (2.2.8) by using $C(\hat{x},\Omega) \cap N(\nabla T(\hat{x}))$ as a quasilinear conical approximation to $\Omega \cap A$.

One of the reasons for our assumption that $\text{inf}\,\Omega \neq \phi$ is now clear - it guarantees that $\hat{x} + \zeta(\epsilon\delta x) \in \Omega$ for ϵ small enough, where $\zeta(\cdot)$ is the function used in the definition of a quasilinear approximation.

The rest of the proof mimics that of Theorem 2.1.2 except that

$\zeta(\epsilon\delta x)$ is used in place of $\epsilon\delta x$. Since $\zeta(\epsilon\delta x) = \delta x + o(\epsilon\delta x)$ this introduces no difficulty.

The assumption that Ω is a convex set and int $\Omega \neq \phi$ is quite restrictive and is not satisfied by some important sets (ones defined by inequality constraints, for example). The following two corollaries weaken the assumption on Ω under which Theorem 2.2.1 was derived.

Corollary 2.2.1 The results of Theorem 2.2.1 hold without the assumption that Ω is convex if a conical approximation $C(\hat{x},\Omega)$ is used such that

$$C(\hat{x},\Omega) = RC(\hat{x},\omega) \tag{2.2.11}$$

where ω is a convex set in Ω such that int $\omega \neq \phi$.

Corollary 2.2.2 The results of Theorem 2.2.1 hold without the assumption that Ω is convex if a conical approximation $C(\hat{x},\Omega)$ is used with the property that for each $\delta x \in C(\hat{x},\Omega)$ there exists an $\epsilon^* > 0$ and a $\delta > 0$ such that for all δz in a sphere of radius δ

$$\hat{x} + \epsilon(\delta x + \delta z) \in \Omega \quad \epsilon \in [0,\epsilon^*]. \tag{2.2.12}$$

Proof: This property is used in equations (B.59) through (B.61) in place of the convexity of Ω in the proof of Theorem 2.2.1.

Section 2.3. Inequality Constraints

In this section conical approximations will be given for sets defined by inequality constraints. These conical approximations will then be used to introduce inequality constraints into the mathematical programming problem.

Let g_i: $X \to Y_i$, $i = 1,\ldots,k$ be continuously Fréchet differentiable on X with derivatives $\nabla g_i(x)$. Let K_i be a closed convex cone in Y_i such that int $K_i \neq \phi$, $i = 1,\ldots,k$. Define the set Ω by

$$\Omega = \{x | g_i(x) \in K_i,\ i = 1,\ldots,k\} \tag{2.3.1}$$

Let $I(\hat{x}) = \{i | g_i(\hat{x}) \in K_i,\ g_i(\hat{x}) \notin \text{int } K_i,\ i \in (1,\ldots,k)\}$.

The <u>internal cone</u> [Ca 1] to Ω at $\hat{x}$ is defined by

$$IC(\hat{x},\Omega) = \{\delta x | \exists \beta > 0 \ni g_i(\hat{x}) + \beta \nabla g_i(\hat{x}) \delta x_i \in \text{int } K_i,\ \forall i \in I(\hat{x})\} \cup \{0\}. \tag{2.3.2}$$

The internal cone can be considered as a "first order" approximation to Ω at $\hat{x}$ since, for $i \in I(\hat{x})$ and $\delta x \in IC(\hat{x},\Omega)$, $g_i(\hat{x}) + \nabla g_i(\hat{x}) \delta x \in K_i$.

<u>Theorem 2.3.1</u> $IC(\hat{x},\Omega)$ is a conical approximation to the set $\Omega = \{x | g_i(x) \in K_i,\ i = 1,\ldots,k\}$ at the point $\hat{x}$.

<u>Proof</u>: $IC(\hat{x},\Omega)$ is a convex cone so there remains to show only that, for every finite set of independent vectors $\{\delta x_\ell\} \in IC(\hat{x},\Omega)$, there exists an $\epsilon > 0$ such that

$$\Sigma_\epsilon = co\{\hat{x},\ \hat{x}+\epsilon\delta x_1, \ldots, \hat{x}+\epsilon\delta x_\ell\} \subset \Omega. \tag{2.3.3}$$

Each vector $x \in \Omega$ can be written

$$x = \hat{x} + \epsilon\lambda \sum_{i=1}^{\ell} \alpha_i \delta x_i \tag{2.3.4}$$

where $\alpha_i \geq 0$, $i = 1,\ldots,\ell$, $\sum_{i=1}^{\ell} \alpha_i = 1$ and $\lambda \in [0,1]$.

Let β_{ij} be such that $g_j(\hat{x}) + \beta_{ij} \nabla g_j(\hat{x}) \delta x_i \in \text{int } K_j$ and define $\beta = \min_{\substack{i \in (0,\ldots,\ell) \\ j \in I(\hat{x})}} \beta_{ij}$. Let $\delta > 0$ be such that the open sphere of radius δ centered at $g_j(\hat{x}) + \beta \nabla g_j(\hat{x}) \delta x_i$, which we denote

$$S(g_j(\hat{x}) + \beta \nabla g_j(\hat{x}) \delta x_i, \delta), \tag{2.3.5}$$

is contained in K_j for all $i \in (0,\ldots,\ell)$ and for all $j \in I(\hat{x})$. Because K_j is a convex cone and $g_j(\hat{x}) \in K_j$

$$S(g_j(x) + \lambda\beta \sum_{i=1}^{\ell} \alpha_i \nabla g_j(\hat{x}) \delta x_i,\ \delta) \in K_j \tag{2.3.6}$$

$$\forall i \in (0,\ldots,\ell),\ \forall j \in I(\hat{x}),\ \forall \lambda \in [0,1],\ \forall \alpha_i \geq 0\ \sum_{i=1}^{\ell} \alpha_i = 1.$$

Since $g_j(\hat{x})$ is continuously Fréchet differentiable

$$g_j(\hat{x}+\lambda\beta \sum_{i=1}^{\ell} \alpha_i \delta_i) = g_j(x) + \lambda\beta \sum_{i=1}^{\ell} \alpha_i \nabla g_j(\hat{x})\delta x_i + o(\lambda\beta\nabla g_j(\hat{x})\delta x_i). \quad (2.3.7)$$

The set $co\{\hat{x},\ \hat{x}+\epsilon\delta x_1,\ldots,\hat{x}+\epsilon\delta x_k\}$ is compact, so there exists an $\epsilon' > 0$ such that

$$\|o(\epsilon\nabla g_j(x)x)\| \leq \epsilon\,\delta/2 \quad x \in \Sigma_{\epsilon\beta},\ \forall \epsilon \in [0,\epsilon'],\ \forall j \in I(\hat{x}). \quad (2.3.8)$$

Let $\epsilon'' = \min(\epsilon',1)\beta$ so that

$$g_j(x) \in K_j \quad \forall j \in I(\hat{x}) \quad \forall x \in \Sigma_{\epsilon''} \quad (2.3.9)$$

If $j \notin I(\hat{x})$ then $g_j(\hat{x}) \in \text{int } K_j$. Since $co\{\hat{x},\ \hat{x}+\epsilon\delta x_1,\ldots,\hat{x}+\epsilon\delta x_k\}$ is compact and since $g_j(x)$ is continuous, there exists an $\tilde{\epsilon} > 0$ such that

$$g_j(\hat{x}+\lambda\delta x) \in K_j \quad \forall j \notin I(\hat{x})\ \forall \lambda \in [0,\tilde{\epsilon}] \quad (2.3.10)$$

$$\forall \delta x \in co\{\delta x_1,\ldots,\delta x_k\}.$$

If $\epsilon^* = \min(\epsilon'',\tilde{\epsilon})$ then $\Sigma_{\epsilon^*} \subset \Omega$ and $IC(\hat{x},\Omega)$ is a conical approximation.

<u>Theorem 2.3.2</u> If $IC(\hat{x},\Omega) \neq \{0\}$, $K_j - g_j(\hat{x})$ is contained in the range of $\nabla g_j(\hat{x})$ $(R(\nabla g_j(\hat{x}))$ and $R(\nabla g_j(\hat{x}))$ is closed, $j \in I(\hat{x})$, then the closure of $IC(\hat{x},\Omega)$ is given by

$$\overline{IC}(\hat{x},\Omega) = \{\delta x \mid \nabla g_j(\hat{x})\delta x \in \tilde{K}_j(\hat{x}),\ j \in I(\hat{x})\} \quad (2.3.11)$$

where $\tilde{K}_j = \overline{\{x \mid x = \alpha(y-g_j(\hat{x})),\ y \in K_j,\ \alpha > 0\}}$.

<u>Proof:</u> Let $A_j = \{\delta x \mid \nabla g_j(\hat{x})\delta x \in \tilde{K}_j(\hat{x})\}$ and let N_j denote the null-space of $\nabla g_j(\hat{x})$. Let X be written as the direct sum of N_j and a Banach

space M_j. (Such a space exists by [Ho 1, Theorem 1C]). The set A_j can be written

$$A_j = N_j + B_j^{-1}(\tilde{K}_j(\hat{x})) \tag{2.3.12}$$

where B_j is the bijection between M_j and $Z_j = \mathcal{R}(\nabla g_j(\hat{x}))$ defined by

$$B_j x = \nabla g_j(\hat{x})x \quad \forall x \in M_j. \tag{2.3.13}$$

For any surjective linear operator $T_j \in B(M_j, Z_j)$ and subset $U \subset X$, the set $T_j U$ is closed if and only if $U + N(T_j)$ (where $N(T_j)$ is the null space of T_j) is closed [Ho 1, Lemma 17H].

Thus A_j is closed since $\nabla g_j(\hat{x})(B_j^{-1}\tilde{K}_j(\hat{x})) = \tilde{K}_j(\hat{x})$ which is closed. Also

$$A(\hat{x},\Omega) = \bigcap_{j \in I(\hat{x})} A_j = \{\delta x \mid \nabla g_j(\hat{x})\delta x \in \tilde{K}_j(\hat{x}),\ j \in I(\hat{x})\} \tag{2.3.14}$$

is closed since the intersection of a finite number of closed sets is closed.

Clearly $\overline{IC}(\hat{x},\Omega) \subset A(\hat{x},\Omega)$ and, since $A(\hat{x},\Omega)$ is closed, $\overline{IC}(\hat{x},\Omega) A(\hat{x},\Omega)$. Since $IC(\hat{x},\Omega) \neq \{0\}$, there is a nonzero vector $\delta x^* \in IC(\hat{x},\Omega)$. Let δx be any vector in $A(\hat{x},\Omega)$. Since for any convex cone K, $x \in \text{int } K$, $y \in K$ implies $x+y \in \text{int } K$,

$$\delta x_j = \frac{1}{j}\delta x^* + \delta x \in IC(\hat{x},\Omega) . \tag{2.3.15}$$

But δx_j converges to δx so that $A(\hat{x},\Omega) \subset \overline{IC}(\hat{x},\Omega)$ thus

$$A(\hat{x},\Omega) = \overline{IC}(\hat{x},\Omega), \tag{2.3.16}$$

and equation (2.3.11) follows.

<u>Corollary 2.3.1</u> If Y_j is finite dimensional for all $j = 1,\ldots, k$, then Theorem 2.3.2 holds without the assumption that

$$\tilde{K}_j(\hat{x}) \subset R(\nabla g_j(\hat{x})) = \overline{R(\nabla g_j(\hat{x}))}. \tag{2.3.17}$$

Proof: The set $\nabla g_j(\hat{x})(B_j^{-1}\tilde{K}_j)$ is closed since $\tilde{K}_j$ is closed, convex and in a finite dimensional space and $\nabla g_j(\hat{x})B_j^{-1}$ is a continuous mapping. [Ho 1, Lemma 17I].

Theorem 2.3.3 The internal cone $IC(\hat{x},\Omega)$ defined in equation (2.3.2) is such that, if $\delta x \in IC(\hat{x},\Omega)$ then there exists an $\epsilon^* > 0$ and a $\delta^* > 0$ such that

$$\hat{x} + \epsilon(\delta x+\delta z) \in \Omega \quad \forall \epsilon \in [0,\epsilon^*] \quad \forall \delta z \in S(0,\delta^*). \tag{2.3.18}$$

Proof : Since $g_j(\hat{x})$ is Fréchet differentiable

$$g_j(\hat{x}+\alpha\delta x) = g_j(\hat{x}) + \alpha\nabla g_j(\hat{x})\delta x + o_j(\alpha\delta x). \tag{2.3.19}$$

Let β be such that

$$g_j(\hat{x}) + \beta\nabla g_j(\hat{x})\delta x \in \text{int } K_j, \quad \forall j \in I(\hat{x}) \tag{2.3.20}$$

and let α be defined by

$$\alpha = \max_{j \in I(x)} \| \nabla g_j(\hat{x}) \| .$$

Choose $\gamma > 0$ as the radius of a sphere such that

$$g_j(\hat{x}) + \beta\nabla g_j(\hat{x})\delta x + \delta y \in K_j \quad \forall j \in I(\hat{x}) \quad \forall y \in S(0,\gamma) \tag{2.3.21}$$

and let $\epsilon(\theta)$ be such that

$$\begin{aligned} \|o_j(\lambda\beta\nabla g_j(\hat{x})(\delta x+\delta y))\| &\le \lambda\theta \| \beta\nabla g_j(\hat{x})(\delta x+\delta y)\| \\ &\forall \lambda \in \| \lambda\beta\nabla g_j(\hat{x})(\delta x+\delta y)\| \in [0,\epsilon(\theta)] \\ &\forall j \in I(\hat{x}) \end{aligned} \tag{2.3.22}$$

For $j \in I(\hat{x})$ we can write

$$g_j(\hat{x} +\lambda\beta(\delta x+\delta z)) = g_j(\hat{x}) + \lambda\beta\nabla g_j(\hat{x})\delta x + \lambda\beta\nabla g_j(\hat{x})\delta z + o(\lambda\beta(\delta x+\delta z)). \tag{2.3.23}$$

and, since K_j is a convex cone,

$$g_j(\hat{x} + \lambda\beta(\delta x+\delta z)) \in K_j \tag{2.3.24}$$

if

$$\| \lambda\beta\nabla g_j(\hat{x})\delta z + o(\lambda\beta(\delta x+\delta z))\| \leq \lambda\gamma \ . \tag{2.3.25}$$

If $\delta z \in S(0,\delta)$, equation (2.3.25) becomes

$$\lambda\beta\alpha\delta + \lambda\beta\alpha\theta(\|\delta x\| + \delta) \leq \lambda\gamma \tag{2.3.26}$$

if $\lambda \ni \lambda\alpha\beta(\delta+\|\delta x\|) \leq \epsilon(\theta)$.

Let $\theta' = \gamma/(2\alpha\beta\|\delta x\|)$, $\delta' = \gamma/(2\alpha\beta(1+\theta))$ and $\lambda' = \epsilon(\theta)/\alpha\beta(\delta+\|\delta x\|)$ so that (2.3.18) holds for all $j \in I(\hat{x})$ with $\epsilon^* = \lambda'\beta$ and $\delta^* = \delta'$.

If $j \in I(\hat{x})$ then $g_j(\hat{x}) \in \text{int } K_j$. Let γ be such that

$$g_j(\hat{x}) + \delta y \in K_k \quad \forall j \notin I(\hat{x}) \quad \forall \delta y \in S(0,\gamma) \ . \tag{2.3.27}$$

Define

$$\alpha = \max_{j \notin I(\hat{x})} \|\nabla g_j(\hat{x})\| \tag{2.3.28}$$

and $\epsilon(\theta)$ so that

$$\|o_j(\lambda\nabla g_j(\hat{x})(\delta x+\delta z))\| \leq \lambda\theta\|\delta x+\delta z\| \tag{2.3.29}$$
$$\forall \lambda \ni \lambda\|\delta x+\delta z\| \in [0,\epsilon(\theta)]$$
$$\forall j \notin I(\hat{x}) \ .$$

Clearly,

$$g_j(\hat{x} + \lambda(\delta x+\delta z)) \in K_j \qquad \forall j \notin I(\hat{x}) \tag{2.3.30}$$

if

$$\|\lambda\nabla g_j(\hat{x})(\delta x+\delta z)+o(\lambda\nabla g_j(\hat{x})(\delta x+\delta z))\| \leq \gamma \ . \tag{2.3.31}$$

But equation (2.3.31) is satisfied if

$$\lambda\alpha(\|\delta x\|+\delta) + \alpha\lambda\theta(\|\delta x\|+\delta) \leq \gamma \tag{2.3.32}$$

and $\delta z \in S(0,\delta)$ and λ satisfies (2.2.29).

Let $\delta' = 1$, $\theta' = 1$ and $\lambda' = \gamma/(2\alpha(\|\delta x\| + 1))$ so that (2.3.18) is satisfied with $\delta^* = \delta'$, $\epsilon^* = \lambda'$, for all $j \notin I(\hat{x})$. Clearly by taking the minimum of the two just derived δ^* and ϵ^* equation (2.3.18) is satisfied for all $j = 1,\ldots,k$.

We can now extend Theorems 2.1.1 and 2.1.2 to the case where Ω includes inequality constraints. The following lemmas will be required. Recall (see Appendix A) that if K is a cone the dual cone K^* is the set of continuous linear functionals positive on K.

<u>Lemma 2.3.1</u> If K_1 and K_2 are closed convex cones then

$$(K_1 \cap K_2)^* = K_1^* + K_2^* \ . \tag{2.3.33}$$

<u>Proof:</u> (following Girsanov [Gi 1])

Note that $K^* = (co(K))^*$ since K^* consists of <u>linear</u> functionals positive on K. Let $Q = (K_1^* + K_2^*)$, then

$$Q = co(K_1^* \cup K_2^*) \tag{2.3.24}$$

Thus

$$Q^* = (co(K_1^* \cup K_2^*))^*$$

$$= (K_1^* \cup K_2^*)^* \tag{2.3.35}$$

But clearly $(K_1^* \cup K_2^*)^* = K_k^{**} \cap K_2^{**}$ so that the theorem is proven if $K^{**} = K$.

It is obvious that $K \subset K^{**}$, Assume that there exists an $x \in K^{**}$ such that $x \notin K$. Since K is a closed convex cone x can be strictly separated from K, that is,

$$\exists\, x^* \in X^* \ni \langle x^*,z\rangle \geq 0 \quad \forall z \in K$$
$$\langle x^*,x\rangle < 0. \tag{2.3.36}$$

Clearly $x^* \in K^*$, but if $\langle x,x^*\rangle < 0$ for some $x^* \in K^*$ then $x \notin K^{**}$. The contradiction implies that $K = K^{**}$ and thus that

$$(K_1 \cap K_2)^* = K_1^* + K_2^* \tag{2.3.27}$$

as desired.

The following generalization of linear independence will be useful. A set of linear operators $\{A_i\}$, where

$$A_i \in B(X,Y_i), \quad i = 1,\dots,k, \tag{2.3.38}$$

and X and Y_i, $i = 1,\dots,k$ are Banach spaces is said to be <u>linearly independent</u> if

$$\langle y^*,Ax\rangle = \langle y_1^*,A_1x\rangle + \dots + \langle y_k^*,A_kx\rangle = 0 \quad \forall x \in X \tag{2.3.39}$$

implies that $y^* = (y_1^*,\dots,y_k^*) = 0$.

<u>Lemma 2.3.2</u> The linear operator $A \in B(X,Y_1 x \dots xY_k)$ defined by $Ax = (A_1x,\dots,A_kx)$ is surjective if and only if the set $\{A_i\}$ is linearly independent and $\mathcal{R}(A_i)$ is closed for $i = 1,\dots,k$.

<u>Proof:</u> If A is surjective then clearly $\mathcal{R}(A) = Y$ and $\mathcal{R}(A_i) = Y_i$. Thus if $\langle y^*,Ax\rangle = 0 \ \forall x \in X$ then

$$\langle y^*,y\rangle = 0 \quad \forall y \in Y \tag{2.3.40}$$

which implies that $y^* = (y_1^*,\dots,y_k^*) = 0$.

If $\{A_i\}$ is linearly independent then

$$\langle y^*,Ax\rangle = \langle A^*y^*,x\rangle = 0 \quad \forall x \in X \tag{2.3.41}$$

implies that $y^* = 0$. But (2.3.41) implies $A^*y^* = 0$, so clearly $y^* \in \mathcal{N}(A^*)$ implies $y^* = 0$. That is,

$$\mathcal{N}(A^*) = \{0\}. \tag{2.3.42}$$

Recall (Appendix A) that for a set $U \subset X$, $[U]^\perp$ is defined as the set $\{x^* \in X^* \mid \langle x^*,u\rangle = 0, \forall u \in U\}$.

Thus [Theorem A. 5.1] $\mathcal{N}(A^*) = [\mathcal{R}(A)]^\perp = 0$. Define, for $K \subset Y^*$, ${}^\perp[K] = [K]^\perp \cap Y$. Clearly ${}^\perp[0] = Y$, and by [Lu 1, Theorem 5.7.1], if M is a closed subspace in X then ${}^\perp[M^\perp] = M$. Thus $\mathcal{R}(A) = Y$ and A is surjective.

Let $\{A_i\}$ be a set of linear operators, $A_i \in B(X,Y_i)$, where X, Y_i, $i = 1,\dots,k$ are Banach spaces and let $K_i \subset Y_i$, $i = 1,\dots,k$ be closed convex cones with nonempty interiors. The set $\{A_i\}$ is K-linearly independent if, for any $y^* = (y_1^*,\dots,y_k^*) \in K^* = K_1^* x \dots x K_k^*$.

$$\langle y^*,Ax\rangle = 0 \quad \forall x \in X \tag{2.3.43}$$

implies that $y^* = 0$.

Theorem 2.3.4 Let $A_i \in B(X,Y_i)$ and $\mathcal{R}(A_i)$ be closed where X,Y_i, $i = 1,\dots,k$ are Banach spaces. Let $K_i \subset Y_i$ be closed convex cones with nonempty interiors. Define

$$D = \{x | A_i x \in K_i,\ x \in X,\ i = 1,\dots,k\} \tag{2.3.44}$$

Then $D \neq \{0\}$ if and only if $\{A_i\}$ are K-linearly independent.

Proof: If $D = \{0\}$ then $\mathcal{R}(A) \cap K = \{0\}$ and $\mathcal{R}(A) \cap \text{int}\{K\} = \phi$. Since $\mathcal{R}(A)$ and $\text{int}\{K\}$ are convex sets they are separable, and since $\mathcal{R}(A)$ is a linear variety the separating functional is zero on $\mathcal{R}(A)$. That is, there exists a $0 \neq y^* \in Y^*$ such that

$$\langle y^*,y'\rangle = 0 \leq \langle y^*,y''\rangle \quad \forall y' \in \mathcal{R}(A) \quad \forall y'' \in K. \tag{2.3.45}$$

But clearly $y^* \in K^*$ and so $\{A_i\}$ are K-linearly dependent.

If $\{A_i\}$ are K-linearly independent then $y^* \in K^*$ and

$$\langle y^*,Ax\rangle = 0 \quad \forall x \in X \tag{2.3.46}$$

implies that $y^* = 0$. But (2.3.46) is equivalent to $A^*y^* = 0$, so K-linear independence implies that

$$\mathcal{N}(A^*) \cap K^* = \{0\}\ . \tag{2.3.47}$$

Since $\mathcal{N}(A^*) = \mathcal{R}(A)$ this implies

$$\mathcal{R}(A) \cap K^* = \{0\}\ . \tag{2.3.48}$$

Because K^* is the set of continuous linear functionals positive on K

$$K^* = (\overline{K})^*\ . \tag{2.3.49}$$

However int $K \neq \phi$ so $\overline{(\text{int } K)} = \overline{K} = K$ since K is closed. Note also that $\mathcal{R}(A)^{\perp} = \mathcal{R}(A)^*$ so that

$$\mathcal{R}(A)^* \cap (\text{int } K)^* = \{0\}. \tag{2.3.50}$$

Now if $\mathcal{R}(A) \cap \text{int } K = \phi$ then $\mathcal{R}(A)$ and int K are separable. That is, there exists a $0 \neq y^* \in Y^*$ such that

$$\langle y^*,y' \rangle = 0 \leq \langle y^*,y'' \rangle \ \forall y' \in \mathcal{R}(A), \ \forall y'' \in \text{int } K. \tag{2.3.51}$$

Clearly $y^* \in (\text{int } K)^*$ and $y^* \in \mathcal{R}(A)^*$ so that

$$\mathcal{R}(A) \cap \text{int } K = \phi \Rightarrow \exists\, 0 \neq y^* \ni y^* \in \mathcal{R}(A)^* \cap (\text{int } K)^*. \tag{2.3.52}$$

However equation (2.3.50) implies that if $y^* \in \mathcal{R}(A)^* \cap (\text{int } K)^*$ then $y^* = 0$ so (2.3.52) implies that

$$\mathcal{R}(A) \cap \text{int } K \neq \phi. \tag{2.3.53}$$

Since $0 \notin \text{int } K$ this implies that there exists a $y \neq 0$ such that $y \in \mathcal{R}(A) \cap K$, but then $D \neq \{0\}$.

<u>Theorem 2.3.5</u> If $\{A_i\}$ are K-linearly independent then

$$B^* = A^*K^*$$

where

$$B = \{x | Ax \in K\}.$$

<u>Proof:</u> From equation (2.3.48)

$$\mathcal{R}(A)^{\perp} \cap K^* = \{0\} \tag{2.3.54}$$

and clearly $\mathcal{R}(A) = \mathcal{R}(A)^*$ so Lemma 2.3.1 implies

$$\mathcal{R}(A)^{**} + K^{**} = \{0\}^* \tag{2.3.55}$$

but $\mathcal{R}(A)$ and K are both closed and convex so

$$\mathcal{R}(A) + K = Y. \tag{2.3.56}$$

An application of a theorem by Kurcyusz [Ku 1, Theorem 2.1] to equation (2.3.50) yields the desired result.

Consider the following problem:

$$\min f(x) \tag{2.3.57}$$

subject to

$$r(x) = 0 \qquad x \in \Omega$$

$$T(x) = 0 \quad G_i(x) \in K_i \qquad i = 1,\ldots,k$$

where $r: X \to \mathbb{R}^n$, $f: X \to \mathbb{R}$, $G_i: X \to Y_i$, $i = 1,\ldots,k$ are continuously Fréchet differentiable functions on X, $T \in B(X,Z)$, X, Y_i, $i = 1,\ldots,k$, Z are Banach spaces, K_i, $i = 1,\ldots,k$ are closed convex cones with non-empty interiors and Ω is an arbitrary set for which there exists a conical approximation $C(\hat{x},\Omega)$. Let $\tilde{K}_i = c\ell\{x \mid \exists \beta > 0 \ni x = \beta(y-G_i(\hat{x})),$ $y \in K_i\}$.

<u>Theorem 2.3.6</u> If $\hat{x}$ is a solution to problem (2.3.57) and $R(T)$, $R(\nabla G_i(\hat{x}))$, $i = 1,\ldots,k$ are closed then there exist $\lambda \leq 0$, $\underline{\phi} \in \mathbb{R}^n$, $z^* \in Z^*$, $y_i^* \in -\tilde{K}_i^*$, $i = 1,\ldots,k$, not all zero such that

$$\lambda\langle\nabla f(\hat{x}),\delta x\rangle + \langle\underline{\phi},\nabla r(\hat{x})\delta x\rangle + \langle z^*,T\delta x\rangle + \sum_{i=1}^{k}\langle y_i^*,\nabla G_i(\hat{x})\delta x\rangle \leq 0$$

$$\forall \delta x \in \overline{C(\hat{x},\Omega)} \tag{2.3.58}$$

<u>Proof:</u> Let $I(\hat{x}) = \{i \mid G_i(\hat{x}) \notin \text{int } K_i,\ i \in \{1,\ldots,k\}\}$. If $\{\nabla G_i(\hat{x})\}$, $i \in I(\hat{x})$ are $\tilde{K}$-linearly dependent then (2.3.58) can be satisfied trivially. Assume there that $\{\nabla G_i(\hat{x})\}$, $i \in I(\hat{x})$ are $\tilde{K}$-linearly independent.

Let $\Omega' = \Omega \cap \omega$ where $\omega = \{x \mid G_i(x) \in K_i,\ i = 1,\ldots,k\}$. Clearly $C(\hat{x},\Omega) \cap IC(\hat{x},\omega)$ is a conical approximation to Ω'. An application of Theorem 2.1.2 to problem (2.3.57) with Ω' replacing Ω and the inequality constraints implies that there exists a $\lambda \leq 0$, $\underline{\phi} \in \mathbb{R}^n$ and $z^* \in Z^*$ not all zero such that

$$\lambda\langle\nabla f(\hat{x}),\delta x\rangle + \langle\underline{\phi},\nabla r(\hat{x})\delta x\rangle + \langle z^*,T\delta x\rangle \leq 0 \tag{2.3.59}$$

$$\forall \delta x \in \overline{C(\hat{x},\Omega) \cap IC(\hat{x},\omega)}$$

But equation (2.3.59) implies that

$$\lambda\nabla f(\hat{x}) + \nabla r^*(\hat{x})\underline{\phi} + T^*z^* \in -[\overline{C(\hat{x},\Omega)} \cap \overline{IC(\hat{x},\omega)}]^* , \tag{2.3.60}$$

since int $[IC(\hat{x},\omega)] \neq \phi$ and $C(\hat{x},\Omega) \cap IC(\hat{x},\omega) \neq \phi$.

Lemma 2.3.1 implies that

$$-(\lambda\nabla f(\hat{x}) + r(\hat{x})\underline{\phi} + T^*z^*) \in \overline{C(\hat{x},\Omega)^*} + \overline{IC(\hat{x},\omega)^*} \tag{2.3.61}$$

so there exists a $k^* \in \overline{IC(\hat{x},\omega)^*}$ such that

$$\lambda\langle\nabla f(\hat{x}),\delta x\rangle + \langle\underline{\phi},\nabla r(\hat{x})\delta x\rangle + \langle z^*,T\delta x\rangle - \langle k^*,\delta x\rangle \leq 0 \tag{2.3.62}$$

$$\forall \delta x \in \overline{C(\hat{x},\Omega)}.$$

Theorem 2.3.5 implies that there exist $y_i^* \in \tilde{K}_i^*$, $i \in I(\hat{x})$ such that $k^* = \sum_{i \in I(\hat{x})} \nabla G_i^*(\hat{x})y_i^*$. Setting $y_i^* = 0$ for all $i \notin I(\hat{x})$ yields

$$\lambda\langle\nabla f(\hat{x}),\delta x\rangle + \langle\underline{\phi},\nabla r(\hat{x})\delta x\rangle + \langle z^*,T\delta x\rangle - \sum_{i=1}^{k}\langle y_i^*,\nabla G_i(\hat{x})\delta x\rangle \leq 0$$

$$\forall \delta x \in \overline{C(\hat{x},\Omega)} \tag{2.3.63}$$

Corollary 2.3.2 If $\{\nabla G_i(x)\}$ $i \in I(\hat{x})$ are $\tilde{K}$-linearly independent then there exist $\lambda \leq 0$, $\underline{\phi} \in \mathbb{R}^n$ and $z^* \in Z^*$ not all zero and $y_i^* \in \tilde{K}_i^*$ $i = 1,\dots,k$ not all zero, such that (2.3.58) holds.

Theorem 2.1.3 can also be extended in a similar way. Consider the following problem:

$$\min f(x)$$

subject to

$$r(x) = 0 \qquad x \in \Omega \tag{2.3.64}$$

$$T(x) = 0 \qquad G_i(x) \in K_i, \qquad i = 1,\dots,k$$

where all assumptions are as in (2.3.57) except $T: X \to Z$ is a continuously Fréchet differentiable mapping and $\mathcal{R}(\nabla T(\hat{x}))$ is closed. Also assume that either

(i) Ω is a convex set and int $\Omega \neq \phi$.

(ii) a $C(\hat{x},\Omega)$ exists as described in Corollary 2.2.1, or

(iii) a $C(\hat{x},\Omega)$ exists as described in Corollary 2.2.2.

Theorem 2.3.7 If $\hat{x}$ solves problem (2.3.64) and $R(\nabla G_i(\hat{x}))$, $i = 1,\ldots,k$ are closed then there exist $\lambda \leq 0$, $\underline{\phi} \in \mathbb{R}^n$, $z^* \in Z^*$ and $y_i^* \in K_i^*$, not all zero such that

$$\lambda\langle\nabla f(\hat{x}),\delta x\rangle + \langle\underline{\phi},\nabla r(\hat{x})\delta x\rangle + \langle z^*,\nabla T(\hat{x})\delta x\rangle - \sum_{i=1}^{k}\langle y_i^*,\nabla G_i(\hat{x})\delta x\rangle \leq 0$$

$$\forall \delta x \in \overline{C(\hat{x},\Omega)}. \qquad (2.3.65)$$

Proof: As in the proof of Theorem 2.3.6.

Corollary 2.3.3 If $\langle\nabla f(\hat{x}),\delta x\rangle \leq 0 \;\; \forall \delta x \in C$

where $C = \{\delta x \,|\, \nabla r(\hat{x})\delta x = 0,\ \nabla T(\hat{x})\delta x = 0,\ \delta x \in \overline{C(\hat{x},\Omega)} \cap \overline{IC(\hat{x},\omega)}\}$ and $\{\nabla G_i(\hat{x})\}$ are $\tilde{K}$-linearly independent, then there exist $\lambda < 0$, $\underline{\phi}$, z^* and y_i^* satisfying (2.3.65).

Proof: follows from Lemma 2.3.1, Theorem 2.3.5 and Corollary 2.1.2.

Section 2.4. Vector Valued Criteria

In this section we consider the extension of the results in the preceeding three sections to the problem in which a nonscalar valued criterion of optimality exists. Clearly a total ordering of solutions is not in general possible and thus minimization of the cost is not usually possible. The problem of minimizing a scalar function will thus be replaced with the problem of finding a K-noninferior solution when the criterion is not scalar valued. A vector $\hat{y} \in Y \supset K$ will be said to K-noninferior over Ω if $\hat{y} \in \Omega$ and for all $y \in \Omega$

$$\hat{y} - y \in K \Rightarrow \hat{y} = y. \qquad (2.4.1)$$

Of course $\hat{y}$ is not generally unique.

In what follows K will be assumed to be a closed convex cone with nonempty interior such that K is not the whole space Y. This implies

$$0 \notin \text{int } K. \qquad (2.4.2)$$

This notion of K-noninferiority reduces to Pareto optimality if Y is finite dimensional and K is the positive orthant.

Consider the following problem (analogous to problem (2.1.1)):

$$\begin{array}{lll} & \text{opt } f(x) & \\ \text{subject to} & r(x) = 0 \qquad x \in \Omega & . \end{array} \qquad (2.4.3)$$

where $f: X \to Y \supset K$, $r: X \to \mathbb{R}^n$ are continuously Fréchet differentiable mappings, X and Y are Banach spaces. Assume that a quasilinear conical approximation to Ω exists.

Theorem 2.4.1 If $\hat{x}$ is a K-noninferior solution to problem (2.4.3) and $C(\hat{x},\Omega)$ is a quasilinear conical approximation to Ω at $\hat{x}$ then there exists a vector $\underline{\phi} \in \mathbb{R}^n$ and a $y^* \in -K^*$, not both zero, such that

$$\langle y^*, \nabla f(\hat{x})\delta x\rangle + \langle \underline{\phi}, \nabla r(\hat{x})\delta x\rangle \leq 0 \quad \forall x \in \overline{C(\hat{x},\Omega)}. \qquad (2.4.4)$$

Proof: Similar to the proof of Theorem 2.1.1, see Appendix B for details.

The following theorems follow immediately from Theorem 2.4.1 since they do not depend on the cost function.

Assumption 2.4.1 The conical approximation $C(\hat{x},\Omega)$ has at least one of the following properties:

(i) $C(\hat{x},\Omega) = RC(\hat{x},\Omega)$ where Ω is convex and int $\Omega \neq \phi$.

(ii) $C(\hat{x},\Omega) = RC(\hat{x},\omega)$ where $\omega \subset \Omega$ is a convex set and int $\omega \neq \phi$.

(iii) for every $\delta x \in C(\hat{x},\Omega)$ there exists an $\epsilon^* > 0$ and a $\delta > 0$

such that

$$\hat{x} + \epsilon(\delta x+\delta z) \in \Omega \quad \forall \epsilon \in [0,\epsilon^*] \quad \forall \delta z \in S(0,\delta) \ . \tag{2.4.5}$$

Consider the following problem

$$\text{opt } f(x)$$

subject to

$$r(x) = 0 \qquad T(x) = 0$$
$$x \in \Omega \subset X \tag{2.4.6}$$

where $r: X \to \mathbb{R}^n$, $T: X \to Z$, $f: X \to Y \supset K$ are Fréchet differentiable mappings, K a closed convex cone such that int $K \neq \phi$ and $K \neq Y$, and $\mathcal{R}(\nabla T(\hat{x}))$ is closed.

<u>Theorem 2.4.2</u> If $\hat{x}$ is K-noninferior for problem (2.4.6), $T \in B(X,Z)$ and a conical approximation $C(\hat{x},\Omega)$ to Ω at $\hat{x}$ exists then there exists a $y^* \in Y^*$, $\underline{\phi} \in \mathbb{R}^n$, $z^* \in Z^*$ not all zero such that $-y^* \in K^*$ and

$$\langle y^*,\nabla f(\hat{x})\delta x\rangle + \langle\underline{\phi},\nabla r(\hat{x})\delta x\rangle + \langle z^*,T\delta x\rangle \leq 0 \qquad \forall \delta x \in \overline{C(\hat{x},\Omega)} . \tag{2.4.7}$$

<u>Theorem 2.4.3</u> If $\hat{x}$ is K-noninferior for problem (2.4.6) and a conical approximation $C(\hat{x},\Omega)$ to Ω at $\hat{x}$ exists which satisfies Assumption (2.4.1) then there exists a $y^* \in Y^*$, $\underline{\phi} \in \mathbb{R}^n$, $z^* \in Z^*$ not all zero such that $-y^* \in K^*$ and

$$\langle y^*,\nabla f(\hat{x})\delta x\rangle + \langle\underline{\phi},\nabla r(\hat{x})\delta x\rangle + \langle z^*,\nabla T(\hat{x})\delta x\rangle \leq 0 \qquad \forall \delta x \in \overline{C(\hat{x},\Omega)} . \tag{2.4.8}$$

<u>Corollary 2.4.1</u> Let Ω' be any set such that if $x' \in \Omega'$ then there exists an $x \in \Omega$ such that

(i) $r(x) = 0$, $Tx = 0$, and (2.4.9)

(ii) $f(x') - f(x) \in \text{int } K$. (2.4.10)

If $\hat{x} \in \Omega'$ is a solution to problem (2.4.6) with $T \in B(X,Z)$ and $C(\hat{x},\Omega')$ is a conical approximation to the set Ω' at $\hat{x}$ then the results of Theorem 2.4.2 hold for all $\delta x \in \overline{C(\hat{x},\Omega')}$.

Proof: see Corollary 2.1.4 and Theorem 2.4.1.

Section 2.5. Concluding Comments

In the next chapter it will be seen that the requirement that $\text{int}\Omega \neq \phi$ is unacceptable in Theorem 2.1.1. Thus a derivation like Dubovitskii and Milyutin's [Dub 1] in which is Ω assumed to have a conical approximation satisfying Assumption 2.4.1 (iii) is not possible.

On the other hand, the requirement that Assumption 2.4.1 hold in theorems 2.1.2 and 2.4.3 is not unreasonable. Basically it will be satisfied if a system is not overparameterized. It is possible to derive theorems for sets Ω such that $\text{int}\,\Omega \neq \phi$, which allow conical approximations, but restrictions must be placed on $T(x)$. However, there are restrictions which allow all the results of successive chapters to hold. (For example one could extend the theorems of Neustadt [Neu 1] from the sets he called finitely open in themselves to sets that have conical approximations.) These theorems are, however, much more difficult to derive and, for the cases we will consider, do not lead to significantly greater generality.

Another direction in which the theorems of this chapter could be generalized is to consider constraint and/or objective spaces in which the positive orthant K has an empty interior. However, the formulation in the next chapters is such that all constraints and/or objectives can naturally be formulated in either finite dimensional Euclidean space or infinite dimensional spaces in which $\text{int}\ K \neq \phi$.

A third possible generalization of these results is to weaken the assumption of Frechet differentiability. A significant amount of effort has gone on in this area ([Ps 1], [Baz 1], [Gi 1], etc.), however it should be noted that no weakening of the differentiability assumption on the equality constraints has been developed. The reason for this is that quasilinear conical approximation to the equality constraints in the nondifferentiable case do not generally exist except under very severe assumptions. Consider the case where $r(x)$ has a Gâteaux differential in all directions at $\hat{x}$. Clearly then for any $\delta x \in X$

$$g(\hat{x}+\alpha\delta x) = g(\hat{x}) + \nabla g(\hat{x};\alpha\,\delta x) + o(\alpha\,\delta x) \tag{2.5.1}$$

where $\nabla g(\hat{x};\delta x)$ denotes the Gâteaux differential at $\hat{x}$ in the direction δx. If $\nabla g(\hat{x};\delta x)$ is "almost" linear, then it can be shown that there exists a mapping $\phi(\delta x)$ such that

$$g(\hat{x}+\alpha\,\delta x + \phi(\alpha\,\delta x)) = 0 \quad \forall \delta x \in C \quad \alpha \geq 0 \tag{2.5.2}$$

where $C = \{x | \nabla g(\hat{x};\delta x) = 0\}$ and

$$\lim_{||x||\to 0} ||\phi(x)||/||x|| = 0. \tag{2.5.3}$$

(See [Io 1, Theorem 0.2.4] for exact details). The assumption that C is convex is, however, very severe since, if $\nabla g(\hat{x};\delta x)$ is linear, C is usually a subspace, thus if $\nabla g(\hat{x};\delta x)$ is nonlinear C is likely to be a hypersurface, which is rarely convex.

While it is possible to consider nondifferentiable inequality constraints this is not a significant generalization as far as the applications in the following chapters are concerned. Unfortunately the assumption that the equality constraint is differentiable is restrictive since that will correspond to assuming that, in an optimal control problem, the system dynamics are differentiable in the control.

Fortunately this problem can be circumvented by an ingenious application of Theorem 2.1.2, however, other severe assumptions on the convexity of certain sets must be made.

Finally, consider the assumption made in many of the theorems: That $\mathcal{R}(\nabla T(\hat{x}))$ be closed. This assumption is related to the difference between absolute and approximate reachability [Ku 1] and has a particularly nice interpretation if T(x) represents the linear dynamics of some system.

Example 2.4.1

Let $T(\underline{z}) = x - Au$ where A is the linear dynamics of some system, x is a terminal state and u a control input. If there is a terminal state constraint then the requirement that $\overline{\mathcal{R}(A)} = Y$ (range of A is dense in Y) is equivalent to the statement that there exists a control u such that x is arbitrarily close to any desired state x'. Superficially this seems to be a desirable property.

Consider the following choice for $A \in B(\ell_2, \ell_2)$

$$Au = (u_1, \frac{1}{2}u_2, \frac{1}{3}u_3 \ldots).$$

The operator A is one-to-one but $\mathcal{R}(A) \neq \ell_2$ since $y = (1, \frac{1}{2}, \frac{1}{3}, \ldots) \notin \mathcal{R}(A)$ even though $y \in \ell_2$ (it cannot be in $\mathcal{R}(A)$ since that would require that $(1,1,1,\ldots) \in \ell_2$). $\overline{\mathcal{R}(A)} = \ell_2$ since every finitely nonzero vector is in $\mathcal{R}(A)$.

If the terminal state constraint is to have x within ϵ of $(1, \frac{1}{2}, \frac{1}{3}, \ldots)$ then this is clearly possible but as $\epsilon \to 0$ $\|u\| \to \infty$.

In fact it can be shown that for any sequence $\{u_n\}$ such that $Au_n \to x \notin \mathcal{R}(A)$ $\|u\| \to +\infty$. Thus in some sense, the problem has been misformulated if $\mathcal{R}(A)$ is not closed since points which are "close" in the state space can correspond to points which are not at all close in the control space. Some work has been done on choosing the topology of the range space so that $\mathcal{R}(A)$ is closed, that is, choosing a topology so that a problem is well formulated. See [Ku 1] for more details.

CHAPTER III

MAXIMUM PRINCIPLES

Section 3.0. Introduction

In this chapter necessary conditions for the optimality of solutions to certain stochastic dynamic optimization problems will be developed. In the spirit of Chapter I the problem formulation will be general enough to include problems of decentralized control and problems with quite general types of constraints.

In Section 3.1 we will formulate the problem and derive an equivalent formulation. In Section 3.2 several examples will be given to illustrate the generality of the problem formulation. Section 3.3 will contain a comparison of this problem with others. In Section 3.4 a straightforward extension of the minimum principle in [San 1] is given and shown to be strongly related to dynamic programming algorithms. Sections 3.5 through 3.8 will contain maximum principles based on the results of Chapter II, as well as some examples. The proofs of Theorems 3.5.1 and 3.6.1 are extension to Banach spaces of the proofs in [Ca 1, Chapter IV]. Section 3.9 contains an extension of these results to the vector valued objective case. Theorem 3.9.2 is an extension of [DaC 1] to Banach spaces. Section 3.10 contains some concluding remarks.

Section 3.1. Problem Formulation

One of the problems faced in formulating a mathematical model for optimal control is choosing a cost function which penalizes appropriately all the undesirable types of system behavior. For example, minimizing a weighted sum of squares of state and control values may produce a control strategy that generates an unacceptably large transient in the control values. Likewise, minimizing the maximum control value may produce a control strategy that wastes energy. While it is sometimes possible to incorporate constraints on control and state values into the model of the system this can often produce an unacceptably complex model.

An alternative is to allow state and/or control constraints to be specified during the problem formulation. It is useful to define two different types of constraints. A structural constraint is one which specifies a class of acceptable parameters (i.e. linear or decentralized control laws). An operational constraint is one which specifies acceptable behavior for the dynamical system (i.e. upper and lower limits on control and state variables). This division of constraints is not always clear cut (for example, consider the class of control laws for which $\|u\| \leq 1$) but it does indicate the types of constraints that need to be considered.

In a deterministic problem the introduction of operational constraints leads to either a feasible or an infeasible problem. That is, either there exists a control strategy which satisfies the constraints or there doesn't. In the stochastic problem the situation is not so clear cut. The feasibility of the problem may depend on the values taken on by the random variables and thus a particular control strategy may either meet or violate some of the constraints.

Clearly the optimal control strategy must always satisfy the constraints. If the constraints are deterministic then this implies that for every sample path of the stochastic process the control must produce a feasible solution. We will consider deterministic constraints which are a function of the control law (rather than control values) only.

When considering state space constraints we shall allow a broader formulation.

We shall allow constraints which require that the expected value of a function of the state lie in some region or which require that the probability of having some function of the state lie in a certain region be greater than or equal to a given threshold. Such constraints will be called soft constraints. Note that these constraints can be "hardened" by setting the threshold to one. In this case the optimal controller must produce a feasible solution for almost all sample paths of the stochastic process.

Soft constraints thus can be seen to "almost" subsume deterministic constraints and to be capable of handling a wide variety of conditions on state variables. Soft constraints may also have some interest in their own right in problems where it is desired to prohibit certain types of dangerous operation during normal (high probability) conditions and to allow that type of operation during emergency (low probability) conditions.

The structural constraints can often be incorporated into the system dynamics or the parameter set over which one is optimizing. This will become clear in Section 3.2 where several examples will be considered.

Consider the following dynamical equation:

$$x_{t+1} = f_t(x_t, w_t, u_t), \quad t = 0,\ldots,t-1 \qquad (3.1.1)$$

where x_t and x_{t+1} denote the system state at times t and t+1, where w_t is the disturbance at time t and u_t is the control at time t. Assume that $x_t \in X_t = R^{n_t}$, $w_t \in W_t = R^{m_t}$ and that the control

$u_t \varepsilon U_t = R^{\ell_t}$. Let Γ_t be an arbitrary set of Borel measurable functions from X_t to U_t. Assume also that f_t is jointly measurable in all of its arguments. For any $\gamma_t \varepsilon \Gamma_t$ equation (3.1.1) can be written

$$x_{t+1} = f_t(x_t, w_t, \gamma_t(x_t)) \qquad t = 0,\ldots,T-1, \tag{3.1.2}$$

and will also be written

$$x_{t+1} = f_t(x_t, w_t; \gamma_t) \qquad t = 0,\ldots,T-1. \tag{3.1.3}$$

A cost is associated with each sequence of $\{x_t\}$, $t = 0,\ldots T$, $\{u_t\}$, $t = 0,\ldots,T-1$ by

$$J_T = \sum_{t=0} h_t(x_t, u_t);\ h_T(x_T, u_T) \triangleq h_T(x_T) \tag{3.1.4}$$

where h_t: $X_t \times U_t \varepsilon \mathbb{R}$ is jointly Borel measurable in both arguments. For any choice of $\gamma_t \varepsilon \Gamma_t$, $t = 0,\ldots,T-1$, equation (3.1.4) can be written

$$J_T^\gamma = \sum_{t=0} h_t(x_t, \gamma_t(x_t)) \tag{3.1.5}$$

and will also be written

$$J_T^\gamma = \sum_{t=0} h_t(x_t; \gamma_t) \tag{3.1.6}$$

We assume that x_o, $w_o,\ldots,w_{T-1}$ are independent random variables with associated probability measures $\pi_o, \mu_o, \ldots, \mu_{T-1}$. The probability space $(\Omega, \beta(\Omega), P)$ is defined by

$$P(C) = \pi_o(A)\mu_o(B_o)\ldots\ \mu_{T-1}(B_{T-1}) \tag{3.1.7}$$

where $A \subset X_o$, $B_i \subset W_i$, $i = 0,\ldots T-1$, $C = AxB_o x\ldots xB_{T-1} \subset \Omega$, $\Omega = X_o \times W_o \times\ldots\times W_{T-1}$ and $\beta(\Omega)$ is the set of all Borel sets of Ω.

Let the <u>state transition stochastic kernal</u> [Bert 1] be defined by

$$p_t(B|x_t,u_t) = \mu_t(\{x|f_t(x_t,w,u_t) \varepsilon B\}). \quad (3.1.8)$$

As usual, if $\gamma_t \varepsilon \Gamma_t$ is chosen (3.1.8) may be written

$$p_t(B|x_t;\gamma_t) \quad (3.1.9)$$

Note that p_t is a Borel measurable function for fixed B and is a probability measure for fixed x_t,u_t (or γ_t,γ_t) [Bert 1, Proposition 26.1]. Furthermore, for any <u>policy</u> $\gamma = (\gamma_o,\ldots,\gamma_{T-1})$ in $\Gamma = \Gamma_o x \ldots \Gamma_{T-1}$ and any π_o there is a unique probability measure $\pi(\gamma,\pi_o)$ induced on $X = X_o x \ldots x X_T$ via equation (3.1.3) [Bert 1, Proposition 7.45]. That is,

$$\int h \, d\pi(\gamma,\pi_o) = \int_{X_o} \ldots \int_{X_T} h(x_o,\ldots,x_T) p_{T-1}(dx_T|x_{T-1};\gamma_T-1)\ldots p_o(dx_1|x_o;\gamma_o)\pi_o(dx_o) \quad (3.1.10)$$

for any h: $X \to \mathbb{R}^+$ ($\mathbb{R}^+ = \mathbb{R} \cup \{+\infty\}$) which is Borel measurable and for which $-\int h^- d\Pi(\gamma,\pi_o) < \infty$ where $h^-(x) = \min(h(x),0)$. It is now clear that the expected value of J_T^γ is well defined for each γ if $-\int J_T^{\gamma-} d\pi(\gamma,\pi_o) < +\infty$ and is given by

$$E\{J_T^\gamma\} = \int_I J_T^\gamma \, d\pi(\gamma,\pi_o) \quad (3.1.11)$$

Suppose that $q_t(x_t)$ is a Borel measurable function from X_T to $\mathbb{R}^{r_t}$ and that $\pi_t(\gamma,\pi_o)$ denotes the marginal distribution of $\pi(\gamma,\pi_o)$ on X_t. Then the expected value of $q_t(x_t)$ is defined and is given by

$$E\{q_t(x_t)\} = \int_{X_t} q_t(x_t) d\pi_t(\gamma,\pi_o) \quad (3.1.12)$$

The basic problem (BP) to be considered in this thesis can now be given:

$$\min_{\gamma \varepsilon \Gamma} E\{J_T^\gamma\} \quad (3.1.13)$$

subject to

$$x_{t+1} = f_t(x_t, w_t; \gamma_t) \qquad t = 0,\ldots,T-1 \tag{3.1.14}$$

$$E\{g_t(x_t)\} \in K_t \qquad t = 1,\ldots,T \tag{3.1.15}$$

$$E\{r_t(x_t)\} = 0 \qquad t = 1,\ldots,T \tag{3.1.16}$$

where equation (3.1.14) is defined by equation (3.1.3), g_t and r_t are defined as is q_t above, where r_t is equal to k_t and c_t respectively, K_t is the positive orthant in $\mathbb{R}^{k_t}$ and J_T^γ is defined by equations (3.1.4) through (3.1.6).

We have also have from [Bert 1] that

$$\pi_{t+1}(\gamma,\pi_o)(S) = \int_S \int_{X_t} p_t(dx_{t+1}|x_t;\gamma_t)d\pi_t(\gamma,\pi_o) \tag{3.1.17}$$

for any Borel subset in X_{t+1} and that

$$E\{J_T^\gamma\} = \sum_{t=0}^{T} \int_{X_t} h_t(x_t;\gamma_t)d\pi_t(\gamma,\pi_o). \tag{3.1.18}$$

If γ and π_o are given we shall sometimes write (3.1.17) and (3.1.18) as

$$\pi_{t+1}(x_{t+1}) = \int p_t(x_{t+1}|x_t;\gamma_t)\pi_t(dx_t) \tag{3.1.19}$$

and

$$E\{J_T\} = \sum_{t=0}^{T} \int h_t(x_t;\gamma_t)\pi_t(dx_t). \tag{3.1.20}$$

Note that π_t can be considered as an element in the Banach space Π_t of signed measures on $(X_t,\beta(X_t))$ with the total variation norm. Equation (3.1.19) thus defines a linear operator from Π_t into Π_{t+1} which will be denoted $P_t(\gamma_t)$. Equation (3.1.19) becomes

$$\pi_{t+1} = P_t(\gamma_t)\pi_t. \tag{3.1.21}$$

We will assume that $P_t(\gamma_t) \in B(\Pi_t,\Pi_{t+1})$ for all $\gamma_t \in \Gamma_t$. Equation (3.1.20) maps Π_t into $\mathbb{R}$ and thus defines a functional on Π_t. We also assume that linear equation (3.1.20) can be represented using elements from the dual space Π_t^* (recall from Appendix A that Π_t^* is the space of continuous linear functionals on Π_t). Equation (3.1.20) thus defines T+1 functionals on Π_t which will be denoted $h_t^*(\gamma_t)$. Equation (3.1.20) can be written

$$E\{J_T\} = \sum_{t=0}^{T} \langle h_t^*(\gamma_t),\pi_t\rangle \ . \tag{3.1.22}$$

Recall (from Appendix A) that $\langle x^*,x\rangle$ denotes the evaluation of the functional x^* at the point x; it should not be confused with the inner product notation sometimes used in Hilbert spaces. The star in $h_t^*(\gamma_t)$ indicates that it is an element of Π_t^*, which in this case is an integral functional with a kernal given by $h_t(x_t;\gamma_t)$.

Under similar assumptions on g_t and r_t (3.1.15) and (3.1.16) become

$$E\{g_t(x_t)\} = \int g_t(x_t)\pi_t(dx_t) \in K_t \tag{3.1.23}$$

and

$$E\{r_t(x_t)\} = \int r_t(x_t)\pi_t(dx_t) = 0 \tag{3.1.24}$$

or

$$G_t\pi_t \in K_t \tag{3.1.25}$$

and

$$R_t\pi_t = 0. \tag{3.1.26}$$

Note that the basic problem (BP) has a corresponding equivalent deterministic problem (EDP) given by:

$$\min_{\gamma\in\Gamma} \sum_{t=0}^{T} \langle h_t^*(\gamma_t),\pi_t\rangle \tag{3.1.27}$$

subject to

$$\pi_{t+1} = P_t(\gamma_t)\pi_t \qquad t = 0,\ldots,T-1 \tag{3.1.28}$$

$$G_t\pi_t \in K_t \qquad t = 1,\ldots,T \tag{3.1.29}$$

$$R_t\pi_t = 0 \qquad t = 1,\ldots,T \tag{3.1.30}$$

where, as before, $h_T(\gamma_T) \triangleq h_T$.

Note that, while π_t contains measures which are not probability measures, we do not need to constrain π_t to be in the set of probability measures. This occurs because $P_t(\gamma_t)$ is such that, if π_t is a probability measure, then π_{t+1} is a probability measure. Since π_o is given and is a probability measure it follows that π_t, $t = 1,\ldots,T$ are also probability measures.

Note that the EDP is a deterministic control problem if the state is taken to be π_t and the control is γ_t. The EDP is thus <u>linear</u> in the state and nonlinear in the control.

Now assume that q_t: $X_t \to \mathbb{R}$ is a Borel measurable function and a constraint is given by

$$\pi_t(\gamma,\pi_o)(\{x_t | q_t(x_t) \leq d\}) \geq \alpha \tag{3.1.31}$$

where $\alpha \in [0, 1]$, then the constraint is equivalent to

$$E\{I(q_t,d)\} \geq \alpha \tag{3.1.32}$$

where

$$I(q_t,d)(x) = \begin{cases} 1 & q_t(x_t) \leq d \\ 0 & q_t(x_t) > 0 \end{cases} \quad . \tag{3.1.31}$$

Note that $I(q_t,d)$ is Borel measurable so that (3.1.32) is well defined. Thus constraints of the form (3.1.32) can be reduced to ones of the form (3.1.15) by noting that equation (3.1.32) is equivalent to

$$E\{I(q_t,d) - \alpha\} \geq 0 \qquad (3.1.32)$$

and thus can be written as

$$E\{I(q_t,d) - \alpha\} \in K \qquad (3.1.33)$$

where $K = \{r \in \mathbb{R} | r \geq 0\}$.

While constraints of this type are not differentiable with respect to x_t this will be unimportant in what follows. This occurs because we shall consider the EDP, which is differentiable with respect to the "state" π_t.

Finally, note that the EDP is nonlinear mathematical programming problem in a Banach space, thus the theory developed in Chapter II is applicable.

Section 3.2. Generality of the Problem Formulation

The purpose of this section is to illustrate the generality of the basic problem formulation. This will be done by reducing several more common problems to that form. Throughout this section we shall assume that the measurability assumptions of Section 3.1 are satisfied by all system functions (i.e. f_t, g_t, h_t, r_t, etc.).

Example 3.2.1 Full State Feedback (Noiseless Observations) The problem here is to

$$\underset{u_t \in U_t}{\text{minimize}} \; E\{\sum_{t=0}^{T} h_t(x_t,u_t)\} \qquad (3.2.1)$$

subject to

$$x_{t+1} = f_t(x_t,w_t,u_t) \qquad t = 0,\ldots,T-1 \qquad (3.2.2)$$

by choosing a function γ_t: $X_t \rightarrow U_T$. Assume that $x_o,w_o,\ldots,w_{T-1}$ are independent random variables.

This problem is trivially identified with the BP if we restrict γ_t to be Borel measurable. Since Borel measurable mappings include every practical feedback function this is not a severe restriction. This restriction does, however, preclude the straightforward application of dynamic programming algorithms. The reason is that Borel measurability of the cost-to-go is not necessarily retained under the operations in an iteration of the dynamic programming algorithm. For a detailed discussion of this and other measure theoretic problems see [Bert 1], [Str 1], [Min 1], etc.

<u>Example 3.2.2</u> <u>Noisy Observations and Perfect Memory</u> The problem is to

$$\underset{u_t \in U_t}{\text{minimize}}\ E\ \{\sum_{t=0}^{T} h_t(x_t,u_t)\} \tag{3.2.3}$$

subject to

$$x_{t+1}=f_t(x_t,w_t,u_t) \tag{3.2.4}$$

by choosing functions γ_t: $Y_t \to U_T$ where $y_t \in Y_t$ is defined by

$$y_t = (z_t,\ldots,z_o,u_{t-1},\ldots,u_o) \tag{3.2.5}$$

and

$$z_t = g_t(x_t,v_t)\ . \tag{3.2.6}$$

Assume that $x_o,w_o,\ldots,w_{T-1},v_o,\ldots,v_{T-1}$ are independent random variables.

Define the augmented state variable $\tilde{x}_t$ by

$$\tilde{x}_t = (x_t,y_t) \tag{3.2.7}$$

a new system equation by

$$\tilde{x}_{t+1} = \tilde{f}_t(\tilde{x}_t,\tilde{w}_t;\gamma_t) = (f_t(x_t,w_t;\gamma_t),g_{t+1}(f_t(x_t,w_t;\gamma_t),v_{t+1})) \tag{3.2.8}$$

where $w_t = (w_t, v_{t+1})$. Let the cost be determined by

$$\tilde{h}_t(\tilde{x}_t, u_t) = h_t(x_t, u_t) \quad . \tag{3.2.9}$$

This problem can now be rewritten

$$\min_{u_t \in U_t} E\{\sum_{t=0}^{T} \tilde{h}_t(\tilde{x}_t, u_t)\} \tag{3.2.10}$$

subject to

$$\tilde{x}_{t+1} = \tilde{f}_t(\tilde{x}_t, \tilde{w}_t, u_t) \tag{3.2.11}$$

by choosing functions γ_t: $X_t \times Y_t \to U_t$ such that $\gamma_t((x_t', y_t)) = \gamma_t((x_t'', y_t))$ for all x_t', x_t'' in X_t.

Clearly this last restriction defines Γ_t and (3.2.10) and (3.2.11) are already in the form of the BP. Note again that we require γ_t to be Borel measurable.

Example 3.2.3 Linear Decentralized Control with Soft Constraints

The problem is to

$$\underset{u_t^1,\, u_t^2}{\text{minimize}}\ E\{\sum_{t=0}^{T} (x_t'Q_tx_t + u_t^{1'}R_t^1u_t^1 + u_t^{2'}R_t^2u_t^2)\} \tag{3.2.12}$$

$$(R_T^1 = R_T^2 = 0)$$

subject to

$$x_{t+1} = A_tx_t + B_t^1u_t^1 + B_t^2u_t^2 + \zeta_t \tag{3.2.13}$$

and

$$E\{x_t'D_tx_t\} \leq dt \tag{3.2.14}$$

by choosing linear functions γ_t^1: $Y_t^1 \to U_t^1$, γ_t^2: $Y_t^2 \to U_t^2$

where $Y_t^1 \in y_t^1$ and $y_t^2 \in Y_t^2$ are given by:

$$y_t^1 = C_t^1x_t + \theta_t^1 \tag{3.2.15}$$

$$y_2^t = C_t^2 x_t + \theta_t^2 \qquad (2.3.16)$$

Assume that $x_o, \zeta_o, \ldots, \zeta_{T-1}, \theta_o^1, \ldots, \theta_{T-1}^1, \theta_o^2, \ldots, \theta_{T-1}^2$ are independent random variables.

Since γ_t^1 and γ_t^2 are linear functions they can be characterized by matrices G_t^1 and G_t^2, thus equation (3.2.13) can be rewritten

$$x_{t+1} = A_t x_t + B_t^1 G_t^1 C_t^1 x_t + B_t^2 G_t^2 C_t^2 x_t + B_t^1 G_t^1 \theta_t^1 + B_t^2 G_t^2 \theta_t^2 + \zeta_t \ . \qquad (3.2.17)$$

Thus the structural constraint that γ_t^1 and γ_t^2 be linear has been incorporated into the system dynamics. Equation (3.2.14) is obviously equivalent to

$$E\{d - x_t' D_t x_t\} \geq 0 \ . \qquad (3.2.18)$$

But equations (3.2.12), (3.2.17) and (3.2.18) are obviously in the form of the BP with $\Gamma_t = \{(G_t^1, G_t^2)\}$.

<u>Example 3.2.4</u> <u>Estimation</u> Here we are given a system and allowed to make noisy measurements

$$x_{t+1} = f_t(x_t, w_t) \qquad t = 0, \ldots, T-1 \qquad (3.2.19)$$

$$y_t = g_t(x_n, v_t) \ . \qquad (3.2.20)$$

The problem is to develop an estimate $\hat{x}_t$ of the state based on the measurements $(y_o, \ldots, y_t)$ so as to

$$\text{minimize } E\{\sum_{t=0}^{T} ||x_t - \hat{x}_t||^2\} \ . \qquad (3.2.21)$$

Since an optimal estimator is usually quite difficult to determine a suboptimal (but hopefully good) estimator is often used. A common one is

$$\hat{x}_{t+1} = f_t(\hat{x}_t, \overline{w}_t) + H_t(\hat{y}_t - g_t(\hat{x}_t, \overline{v}_t)) \quad t = 0, \ldots, T-1 \qquad (3.2.22)$$

where $\bar{w}_t = E\{w_t\}$, $\bar{v}_t = E\{v_t\}$ and $\{H_t\}$ is a sequence of matrices which map X_t into Y_t. The optimization problem is to choose $\{H_t\}$ so as to perform the minimization in (3.2.21). We assume that $v_o,\ldots v_{T-1}$, $w_o,\ldots,w_{T-1}$ are independent random variables.

This problem is easily put into the form of the BP by augmenting the state x_t with $\hat{x}_t$ and letting $\Gamma_t = \{H_t\}$.

Example 3.2.5 Communications

Consider a communication problem in which it is desired to send a message (x_o) over a noisy channel. We require that the encoding and decoding be recursive, that is, the transmission at time t, denoted z_t, depends only on x_o and z_{t-1}, and the estimate of x_o at time t, denoted $\hat{x}_t$ depends only on $\hat{x}_{t-1}$ and $y_{t-1} = z_{t-1} + v_{t-1}$, where $v_o,\ldots,v_T$ are the independent random variables representing the channel noise. We wish to minimize $||\hat{x}_T - x_o||^2$ subject to the constraint that $\sum_{t=0}^{T-1} ||z_t||^2 \leq 1$. Assume $\hat{x}_o$ is given.

This problem can be put in the BP form as follows: choose q_t and f_t, $t = 0,\ldots,T-1$ so as to

$$\text{minimize } E\{(x_o - \hat{x}_T)'I(x_o - x_T)\} \tag{3.2.23}$$

subject to

$$\tilde{x}_{t+1} = \begin{bmatrix} x_o \\ z_{t+1} \\ y_{t+1} \\ \hat{x}_{t+1} \\ u_{t+1} \end{bmatrix} = \begin{bmatrix} x_o \\ q_t(x_o, z_t) \\ z_t + v_t \\ f_t(\hat{x}_t, y_t) \\ u_t - z_t' I z_t \end{bmatrix} = \tilde{f}_t(\tilde{x}_t, \tilde{w}; \gamma_t) \tag{3.2.24}$$

where

$$I(u,d) = \begin{cases} 1 & u \geq d \\ 0 & u \leq d \end{cases} \quad (3.2.27)$$

$$\tilde{I}(u,d) = \begin{cases} 1 & u = d \\ 0 & u \neq d \end{cases} \quad . \quad (3.2.28)$$

Note that this example illustrates two features of this problem formulation.

(i) it is possible to include non-additive cost function by augumenting the state dynamics, and

(ii) it is possible to handle "global" constraints (that is, constraints which affect state and/or control variables at several time instances) by augmenting the state dynamics.

Of course augmenting the state dynamics increases the complexity but it allows one to handle much more sophisticated problems.

Example 3.2.6 Control with Information Constraints In decentralized control problems it is interesting to consider situations in which each controller is allowed to act not only on local information but also on some delayed information from other controllers and/or observers [Wit 1]. The problem here is

$$\underset{u_t^i \in U_t^i}{\text{minimize}} \; E\{\sum_{t=0}^{T} h_t(x_t,u_t^1,\ldots,u_t^n)\} \quad (3.2.29)$$

subject to

$$x_{t+1} = f_t(x_t,u_t^1,\ldots,u_t^N,w_t) \qquad t = 0,\ldots,T-1 \quad (3.2.30)$$

by selecting functions g_t^i, $t = 0,\ldots,T-1$, $i = 1,\ldots,N$ which map the information set I_t^i into control values u_t^i. The information set consists of observations y_τ^k and controls u_τ^j available to controller i at time t from any of the K observers and N controllers.

This is obviously in the BP formulation Γ_t is identified with the set of Borel measurable maps from I_t^i into U_t^i.

Section 3.3 Comparison with Other Formulations

The basic problem formulation is based on the "Standard Form" of Witsenhausen [Wit 2] since the unconditional probability measure is used as the "state". There are two basic differences:

(i) We allow explicit state space constraints (the operational constraints of Section 3.1), and

(ii) We use the state evolution equation introduced in the finite-state, finite-memory problem of Sandell [San 1].

The first difference precludes the use of the minimum principles developed in [Wit 2] and [San 1]. This requires us to make certain differentiability and/or convexity assumptions that cannot hold in the discrete-state formulation of [San 1]. The second difference allows us (as it did Sandell) to extend the results derived for finite time horizon problems to infinite time horizon problems. This will be considered in Chapter IV.

The extension of this formulation to a vector valued cost criterion can be considered as a problem with several teams, each of which have their own cost criterion. However, we assume that the teams are cooperating and thus the game formulation of Castañon [Cas 1] does not arise.

Section 3.4 Relationship to Dynamic Programming

In this section we convert the minimum principle in [Wit 2] (as formulated in [San 1]) to a maximum principle applicable to our basic problem formulation. The extension is very straightforward and involves primarily a change of notation. We then show that this maximum principle is strongly related to the dynamic programming algorithm. To do all of this requires that we restrict our attention to the following unconstrained version of the EDP introduced in Section 3.0.

$$\min_{\gamma\varepsilon\Gamma} \sum_{t=0}^{T} \langle h_t^*(\gamma_t), \pi_t \rangle \tag{3.4.1}$$

subject to

$$\pi_{t+1} = P_t(\gamma_t)\pi_t \qquad t = 0,\ldots,T-1 \tag{3.4.2}$$

$$\pi_o \text{ given.}$$

Define the forced adjoint (or costate) equation by

$$\phi_t^* = P_t^*(\gamma_t)\phi_{t+1}^* - h_t^*(\gamma_t); \quad \phi_T \equiv -h_T \tag{3.4.3}$$

Note that this is a functional difference equation since $\phi_t^* \varepsilon \Pi_t^*$.

<u>Lemma 3.3.1</u> Let $\gamma = (\gamma_o,\ldots,\gamma_{T-1})$ be given and let the corresponding states and costates be given by equations (3.4.2) and (3.4.3) where π_o is also assumed given. Then

$$\langle \phi_t^*, \pi_t \rangle = - \sum_{\tau=t} \langle h_\tau^*(\gamma_\tau), \pi_\tau \rangle \tag{3.4.4}$$

<u>Proof</u>: By backward induction. Equation (3.4.4) obviously holds for $t = T$, so assume it holds for an arbitrary $t < T$. Then

$$\langle \phi_{t-1}^*, \pi_{t-1} \rangle = \langle P_{t-1}(\gamma_{t-1})\phi_t - h_{t-1}^*(\gamma_{t-1}), \pi_{t-1} \rangle \tag{3.4.5}$$

$$= \langle \phi_t^*, P_{t-1}(\gamma_{t-1})\pi_{t-1} \rangle - \langle h_{t-1}^*(\gamma_{t-1}), \pi_{t-1} \rangle$$

$$= \langle \phi_t^*, \pi_t \rangle - \langle h_t^*, (\gamma_{t-1}), \pi_{t-1} \rangle$$

$$= - \sum_{\tau=t} \langle h_\tau^*(\gamma_\tau), \pi_\tau \rangle \ ,$$

so that (3.4.4) also holds for t-1. Thus (3.4.4) is valid for $t = 0,\dots,T$.

Theorem 3.4.1 If $\hat{\gamma} = (\hat{\gamma}_o,\dots,\hat{\gamma}_{T-1})$ is optimal in the problem defined by equations (3.4.1) and (3.4.2) for a given π_o, and the corresponding states and costates (generated by equations (3.4.2) and (3.4.3)) are denoted $\hat{\pi}_t$ and $\hat{\phi}^*_t$ then

$$\langle P^*_t(\hat{\gamma}_t)\hat{\phi}^*_{t+1} - h^*_{t+1}(\hat{\gamma}_{t+1}),\hat{\pi}_t\rangle \geq \langle P^*_t(\gamma_t)\hat{\phi}^*_{t+1} - h^*_{t+1}(\gamma_t),\hat{\pi}_t\rangle \qquad (3.4.6)$$
$$\forall \gamma_t \in \Gamma_t$$

Proof: Let $\gamma' = (\hat{\gamma}_o,\dots,\hat{\gamma}_{t-1},\gamma_t,\hat{\gamma}_{t+1},\dots,\hat{\gamma}_{T-1})$ and let π_t' and $\phi^*_t{}'$ denote the corresponding states and costates. Note that

$$\hat{\pi}_\tau = \pi_\tau \qquad \tau = 0,\dots,t \qquad (3.4.7)$$

since π_t depends only on $\gamma_o,\dots,\gamma_{t-1}$ and note that

$$\phi^*_\tau = \phi^*_\tau{}' \qquad \tau = t+1,\dots,T \qquad (3.4.8)$$

since ϕ^*_t depends only on $\gamma_{t+1},\dots,\gamma_T$. Let

$$J_T = -\langle\hat{\phi}^*_o,\hat{\pi}_o\rangle\ , \qquad J_T' = -\langle\phi^*_o{}',\pi_o'\rangle. \qquad (3.4.9)$$

From Lemma 3.4.1 and the optimality of $\hat{\gamma}$ one has

$$\hat{J}_T \leq J_T' \ . \qquad (3.4.10)$$

By the definition of γ' and equation (3.4.10)

$$\sum_{\tau=t}\langle h^*_\tau(\hat{\gamma}_\tau),\hat{\pi}_\tau\rangle \geq \sum_{=t}\langle h^*_\tau(\gamma_\tau'),\pi_\tau\rangle \qquad (3.4.11)$$

which by Lemma 3.4.1 is just

$$\langle\hat{\phi}^*_t,\hat{\pi}_t\rangle \geq \langle\phi^*_t{}',\pi_t'\rangle. \qquad (3.4.12)$$

Now equations (3.4.3), (3.4.7) and (3.4.8) imply

$$\langle\hat{\phi}^*_t,\hat{\pi}_t\rangle \geq \langle P^*_t(\gamma_t)\hat{\phi}^*_{t+1} - h^*_{t+1}(\gamma_{t+1}),\hat{\pi}_t \quad \forall \gamma_t \in \Gamma_t\ . \qquad (3.4.13)$$

Clearly (3.4.13) implies (3.4.6).

While we are interested in problems which are not readily handled by standard dynamic programming techniques it is interesting to note that Theorem 3.4.1 leads to Bellman's equation for two interesting problems. Before showing that this is the case however, we need to make two points. The first is that Theorem 3.4.1 is only a necessary condition of optimality while dynamic programming yields necessary and sufficient condition for optimality. Thus our derivation of Bellman's equation yields only a necessary condition of optimality.

The second point is that our derivation is a formal one. For example, we will, to facilitate comparison with standard results, assume that probability densities exist. Furthermore we will manipulate equations without providing justification of a measure theoretic nature. Note, however, that such justifications are needed (and provided in [Bert 1]) in any derivation of the dynamic programming algorithm. Thus, with a great deal of effort, it is possible to provide mathematically rigorous justification (under very mild assumptions) for all the steps in the following derivations.

Example 3.4.1 Consider the following problem: choose a sequence of functions γ_t: $X_t \to U_T$ so as to

$$\text{minimize } E\{\sum_{t=0} h_t(x_t,u_t)\} \tag{3.4.14}$$

subject to

$$x_{t+1} = f_t(x_t,w_t,u_t), \qquad t = 0,\ldots,T-1 \tag{3.4.15}$$

where $u_t = \gamma_t(x_t)$, a probability measure π_o for the state x_o is given, and $w_o,\ldots,w_{T-1}$ are independent random variables. An equivalent deterministic problem is given by

$$\underset{\gamma_t \in \Gamma_t}{\text{minimize}} \sum_{t=0}^{T} \langle h_t^*(\gamma_t), \pi_t \rangle \tag{3.4.16}$$

subject to

$$\pi_{t+1} = P_t(\gamma_t)\pi_t \qquad t = 0,\ldots,T-1, \tag{3.4.17}$$

where π_o is given.

It will be shown in Section 3.7 that since $h_T(x_T)$ is Borel measurable and since

$$q(x_t) = \int \bar{p}_t(x_{t+1}|x_t;\gamma_t)\pi_{t+1}(dx_{t+1}) \tag{3.4.18}$$

is Borel measurable the costate variables defined by (3.4.3) are such that

$$\langle \phi_t^*, \pi_t \rangle = \int \phi_t(x_t)\pi_t(dx_t) \tag{3.4.19}$$

where $\phi_t(x_t)$ is a Borel measurable function. Thus if probability densities $p_t(x_t)$, corresponding to the probability measures π_t, exist then Theorem 3.4.1 becomes

$$\int [\int \hat{\phi}_{t+1}(x_{t+1})p_t(x_{t+1}|x_t;\hat{\gamma}_t)dx_{t+1} - h_t(x_t;\hat{\gamma}_t)]\hat{p}_t(x_t)dx_t \tag{3.4.20}$$

$$= \max_{\gamma_t \in \Gamma_t} \{\int [\int \hat{\phi}_{t+1}(x_{t+1})p_t(x_{t+1}|x_t;\gamma_t)dx_{t+1} - h_t(x_t;\gamma_t)]\hat{p}_t(x_t)dx_t\}$$

$$= \int \max_{u_t \in U_t} [\int \hat{\phi}_{t+1}(x_{t+1})p_t(x_{t+1}|x_t,u_t)dx_{t+1} - h_t(x_t,u_t)]\hat{p}_t(x_t)dx_t$$

but by equation (3.4.3) and the definition of $\hat{u}_t$ and $\hat{\gamma}_t$

$$\hat{\phi}_t(x_t) = \min_{u_t \in U_t} \{\int -\hat{\phi}_{t+1}(x_{t+1}p_t(x_{t+1}|x_t,u_t)dx_{t+1} + h_t(x_t,u_t)\} \tag{3.4.21}$$

which is exactly Bellman's equation for stochastic dynamic programming in the case of noiseless observations [Bert 2] with $-\hat{\phi}_{t+1}(x_{t+1})$ the expected cost-to-go from x_{t+1}.

Example 3.4.2 Consider the following problem: choose a sequence of functions γ_t: $X_t \times Y_o \times \ldots \times Y_t \to U_t$ such that

$$\gamma_t(x_t',y_o,\ldots,y_t) = \gamma_t(x_t'',y_o,\ldots,y_t) \qquad \forall x_t',x_t'' \in X_t \tag{3.4.22}$$

so as to

$$\text{minimize } E\{\sum_{t=0}^{T} h_t(x_t,u_t = \gamma_t(x_t,y^t))\} \tag{3.4.23}$$

subject to

$$x_{t+1} = f_t(x_t,w_t,u_t) \qquad t = 0,\ldots,T-1 \tag{3.4.24}$$

where $u_t = \gamma_t(x_t,y^t)$ and

$$y^t = (y_o,\ldots,y_t) \tag{3.4.25}$$

and

$$y_t = g_t(x_t,v_t) \tag{3.4.26}$$

and $w_o,\ldots,w_{T-1},v_o,v_{T-1}$ are independent random variables.

An equivalent deterministic problem is:

$$\underset{\gamma_t \in \Gamma_t}{\text{minimize}} \sum_{t=0}^{T} \langle h_t^*(\gamma_t),\pi_t\rangle \tag{3.4.27}$$

subject to

$$\pi_{t+1} = P_t(\gamma_t)\pi_t \tag{3.4.28}$$

where π_o is given and equation (3.4.28) is a recursive equation for probability measures corresponding to the augmented states (x_t,y^t).

Assuming that probability densities exist, Theorem 3.4.1 becomes, for this example,

$$\int\int[\int\int\hat{\phi}_{t+1}(x_{t+1},y^{t+1})p_t(x_{t+1},y^{t+1}|x_t,y^t;\hat{\gamma}_t)dx_{t+1}dy^{t+1}-h_t(x_t;\hat{\gamma}_t)]$$
$$\hat{p}_t(x_t,y^t)dx_t dy^t =$$

$$\max_{\gamma_t \in \Gamma_t} \int [\iint \hat{\phi}_{t+1}(x_{t+1},y^{t+1}) p_t(x_{t+1},y^{t+1}|x_t,y^t;\gamma_t) dx_{t+1} dy^{t+1} - h_t(x_t;\gamma_t)] \cdot$$

$$p_t(x_t,y^t) dx_t dy^t =$$

$$\max_{u_t \in U_t} [\int \{\iint \hat{\phi}_{t+1}(x_{t+1},y^{t+1}) p_t(x_{t+1},y^{t+1}|x_t,y^t,u_t) dx_{t+1} dy^{t+1} - h_t(x_t,u_t)\}$$

$$\cdot \hat{p}_t(x_t|y^t) dx_t] \hat{p}_t(y^t) dy^t \tag{3.4.29}$$

Now

$$\int p_t(x_{t+1},y^{t+1}|x_t,y^t,u_t) \hat{p}_t(x_t|y^t) dx_t = p_t(x_{t+1} y^{t+1}|y^t,u_t) \tag{3.4.30}$$

and

$$\int h_t(x_t,u_t) \hat{p}_t(x_t|y^t) dx_t = E\{h_t(x_t,u_t)|y^t,u_t\} \tag{3.4.31}$$

and also note that

$$p_t(x_{t+1},y^{t+1}|y^t,u_t) \equiv p_t(x_{t+1}|y^{t+1},y^t,u_t) p_t(y^{t+1}|y^t,u_t)$$

$$= p_t(x_{t+1}|y^{t+1},u_t) p_t(y^{t+1}|y^t,u_t) \tag{3.4.32}$$

so that (3.4.29) becomes

$$\int \max_{u_t \in U_t} [E\{\int \hat{\phi}_{t+1}(x_{t+1},y^{t+1}) p_t(x_{t+1}|y^{t+1},u_t) dx_{t+1} - h_t(x_t,u_t)|y^t,u_t\}]$$

$$\hat{p}_t(y^t) dy^t . \tag{3.4.33}$$

Now let

$$\hat{J}_t(p_t(x_t|y^t)) = -\int \hat{\phi}_{t+1}(x_t,y^t) \hat{p}_t(x_t|y^t) dx_t \tag{3.4.34}$$

which is clearly the expected cost-to-go given that the observations have implied a conditional density $p_t(x_t|y^t)$. Thus equation (3.4.3) implies that

$$\hat{J}_t(p_t(x_t|y^t)) = \min_{u_t \in U_t} E\{J_{t+1}(p_t(x_t|y^t,u_t)) + h_t(x_t,u_t)|y^t,u_t\} \tag{3.4.35}$$

which is Bellman's equation for stochastic dynamic programming with imperfect state information [Bert 2].

Theorem 3.4.1 thus subsumes the necessary conditions of dynamic programming when the problem is such that dynamic programming is applicable. In cases where the problem has structural constraints dynamic programming is not readily applied; we shall see in the next sections that Theorem 3.4.1 and its generalization are still applicable.

Section 3.5 A Maximum Principle

In this section we shall derive a maximum principle for the basic problem. As might be expected from the development in Chapter II some assumptions must be made about the nature of the sets and functions in the BP.

Consider the following problem:

$$\underset{\gamma_t \in \Gamma_t}{\text{minimize}}\ E\{\sum_{t=0}^{T} h_t(x_t;\gamma_t)\};\ h_T(x_T) \triangleq h_T(x_T;\gamma_T) \quad (3.5.1)$$

subject to

$$x_{t+1} = f_t(x_t,w_t;\gamma_t) \qquad t = 0,\ldots,T-1 \quad (3.5.2)$$

and

$$E\{g_t(x_t)\} \in K_t \qquad t = 1,\ldots,T \quad (3.5.3)$$

$$E\{r_t(x_t)\} = 0 \qquad t = 1,\ldots,T \quad (3.5.4)$$

Recall that $x_t \in X_t = \mathbb{R}^{n_t}$, $w_t \in W_t = \mathbb{R}^{m_t}$, γ_t: $X_t \to U_t = \mathbb{R}^{\ell_t}$, g_t: $X_t \to \mathbb{R}^{k_t}$, K_t is the closed positive orthant in $\mathbb{R}^{k_t}$, r_t: $X_t \to \mathbb{R}^{e_t}$, h_t: $X_t \times \Gamma_t \to R^+$, f_t: $X_t \times W_t \times \Gamma_t \to X_{t+1}$ and Γ_t is a set of Borel measurable mappings from $\mathbb{R}^{n_t}$ into $\mathbb{R}^{\ell_t}$. All the measurability assumptions of Section 3.1 are assumed to hold.

Recall that an equivalent deterministic problem corresponding to equations (3.5.1) through (3.5.4) exists and is given by:

$$\underset{\gamma \in \Gamma}{\text{minimize}} \sum_{t=0} \langle h_t^*(\gamma_t), \pi_t \rangle \tag{3.5.5}$$

subject to

$$\pi_{t+1} = P_t(\gamma_t)\pi_t \qquad t = 0,\ldots,T-1 \tag{3.5.6}$$

$$G_t \pi_t \in K_t \qquad t = 1,\ldots,T \tag{3.5.7}$$

$$R_t \pi_t = 0 \qquad t = 1,\ldots,T \tag{3.5.8}$$

The following generalization of convexity [DaC 1] will be needed in what follows. Note that it is a weaker notion than convexity in that it is implied by but does not imply convexity.

Let P be a convex cone in X. A subset S of X is <u>P-directionally convex</u> if, for every x_1 and x_2 in S and $\lambda \in [0,1]$, there exists an $x(\lambda) \in P$ such that

$$\lambda x_1 + (1-\lambda)x_2 + x(\lambda) \in S \tag{3.5.9}$$

Let M_t denote the set of all measures in π_t which are probability measures. Let M* be the polar cone for M (that is $M^* = \{m^* \in \Pi^* | \langle m^*, m\rangle \geq 0, \forall m \in M\}$). Note that M* is a convex cone and M is a convex set.

Define, for t = 0,...,T-1, the set

$$F_t(\Gamma_t) = \{(\pi^*, P) \in \Pi^* \times B(\Pi_t, \Pi_{t+1}) \,|\, \langle \pi^*, P\rangle = (h_t^*(\gamma_t), P_t(\gamma_t))\, \gamma_t \in \Gamma_t\} \tag{3.5.10}$$

and let $F_T = (h_t^*, 0)$.

<u>Assumption 3.5.1</u> The set $F_t(\Gamma_t)$ is $(-M_t^*, \underline{0})$-directionally convex, $t = 0,\ldots,T-1$.

This implies that

(i) $P_t(\Gamma_t)$ is a convex set, and

(ii) if, for any $\gamma_t', \gamma_t'' \in \Gamma_t$, $\gamma_t(\lambda)$ is such that

$$P_t(\gamma_t(\lambda)) = \lambda P_t(\gamma_t') + (1-\lambda)P_t(\gamma_t'')$$

then (3.5.11)

$$\lambda h_t^*(\gamma_t') + (1-\lambda)h_t^*(\gamma_t'') - h_t^*(\gamma_t(\lambda)) \in M^*.$$

Define the set $F_t(\Gamma_t)\pi_t$ by (3.5.12)

$$F_t(\Gamma_t)\pi_t = \{(r,\pi_{t+1}) \in \mathbb{R} \times \Pi_{t+1} \mid r = \langle\pi_t^*\pi_t\rangle, \pi_{t+1} = P\pi_t, (\pi_t^*P) \in F_t(\Gamma_t)\}$$

<u>Lemma 3.5.1</u> If $F_t(\Gamma_t)$ is $(-M_t^*,\underline{0})$-directionally convex then $F_t(\Gamma_t)\pi_t$ is $(-1,\underline{0})$-directionally convex for all $\pi_t \in M_t$.

<u>Proof:</u> Let $z_t \in F_t(\Gamma_t)$ be denoted $(\pi^*,P)_t$. By assumption for any $z_t', z_t'' \in R_t(\Gamma_t)$ and $\lambda \in [0,1]$ there exists $z_t(\lambda)$ such that

$$\lambda(\pi^*,P)_t' + (1-\lambda)(\pi^*,P)_t'' + (\pi^*(\lambda),\underline{0})_t \in F_t(\Gamma_t) \tag{3.5.13}$$

where $\pi^*(\lambda) \in -M_t^*$. Thus, since $\pi_t \in M_t$,

$$\lambda(\pi^*,P)_t'\pi_t + (1-\lambda)(\pi^*,P)_t''\pi_t + (\pi^*(\lambda),\underline{0})_t\pi_t \in F_t(\Gamma_t)\pi_t \tag{3.5.14}$$

implies that, for some $\beta_t \geq 0$

$$\lambda(\pi^*,P)_t'\pi_t + (1-\lambda)(\pi^*,P)_t''\pi_t + \beta_t(-1,\underline{0}) \in F_t(\Gamma_t)\pi_t \tag{3.5.15}$$

that is, $F_t(\Gamma_t)\pi_t$ is $(-1,\underline{0})$-directionally convex for all $\pi_t \in M_t$.

Let $v_T = \langle h_T,\pi_T\rangle$ and let $v_t \in F_t(\Gamma_t)\pi_t$ be denoted (v_t^o,v_t^1) where $v_t^o \in \mathbb{R}$ and $v_t^1 \in \Pi_{t+1}$. Let $z \in Z$ be defined by

$$z = (\pi_t,\ldots,\pi_T,v_o,\ldots,v_T) \tag{3.5.16}$$

where

$$Z = \Pi_1 \times\ldots\times \Pi_T \times \mathbb{R} \times \Pi_1 \times\ldots\times \mathbb{R} \times \Pi_T \times \mathbb{R}. \tag{3.5.17}$$

Define the affine map T from Z into $\Pi_1 \times\ldots\times \Pi_T$ by

$$Tz = \begin{bmatrix} \pi_1 - v_o^1 \\ \vdots \\ \pi_T - v_{T-1}^1 \end{bmatrix}, \tag{3.5.18}$$

R: $Z \to \mathbb{R}^{e_1} \times \ldots \times \mathbb{R}^{e_T}$ by

$$Rz = [R_1\pi_1, \ldots, R_T\pi_T] \tag{3.5.19}$$

and let S_t be defined by

$$S_t = \{\pi_t | G_t\pi_t \in K_t\}, \qquad t = 1, \ldots, T \tag{3.5.20}$$

It is now possible to rewrite problem (3.5.5) (hereafter a problem is referred to by the equation number of the cost function) as:

$$\underset{z \in \Omega}{\text{minimize}}\ f(z) = \sum_{t=0}^{T} v_t^o \tag{3.5.21}$$

subject to

$$Tz = 0, \qquad Rz = 0 \tag{3.5.22}$$

where

$$\Omega = \{z | \pi_t \in S_t,\ t = 1, \ldots, T,\ v_t \in F_t(\Gamma_t)\pi_t,\ t = 0, \ldots, T\}. \tag{3.5.23}$$

Note that $\mathcal{R}(T) = \Pi_1 \times \ldots \times \Pi_T$, that is, $\mathcal{R}(T)$ is closed. Thus Problem (3.5.21) is now in a form to which the theorems of Chapter II can be applied. Since a conical approximation to Ω is not easily found we shall (following [Ca 1]) introduce another set Ω' for which a conical approximation can be determined. It will be shown that this set satisfies the assumptions of Corollary 2.1.4 and thus the results of Theorem 2.1.2 hold with Ω replaced Ω'.

Let Ω' be defined by

$$\Omega' = \{z | \pi_t \in S_t,\ t = 1, \ldots, T,\ v_t \in co(F_t(\Gamma_t)\pi_t), t = 0, \ldots, T\}. \tag{3.5.24}$$

If $z = (\pi_1, \ldots, \pi_T, v_o, \ldots, v_T) \in \Omega'$ then, since $F_t(\Gamma_t)\pi_t$ is $(-1,\underline{0})$-directionally convex, there exists $\tilde{v}_t \in F_t(\Gamma_t)\pi_t$ such that $\tilde{v}_t^1 = v_t^1$ and $\tilde{v}_t^o \leq v_t^o$, $t = 0, \ldots, T$. Thus if $\tilde{z} = (\pi_1, \ldots, \pi_T, \tilde{v}_o, \ldots, \tilde{v}_T)$ then $T\tilde{z} = Tz$, $R\tilde{z} = Rz$ and $f(\tilde{z}) \leq f(z)$. This implies that Ω' satisfies the

the hypothesis of Corollary 2.1.4, with respect to the set Ω and T and R as defined in equations (3.5.18) and (3.5.19).

Assume that $\hat{z}$ is an optimal solution to problem (3.5.21). We now show that a conical approximation to Ω' exists at $\hat{z}$.

Lemma 3.5.2 Let $C(\hat{z},\Omega')$ be defined by

$$C(\hat{z},\Omega') = \{\delta z = (\delta\pi_1,\ldots,\delta\pi_T,\delta v_o,\ldots,\delta v_T) \mid \delta\pi_t \varepsilon IC(\hat{\pi}_t,S_t),\ t = 1,\ldots,T,$$
$$\delta v_t - F_t(\hat{\gamma})\delta\pi_t \ \varepsilon\ RC(\hat{v}_t, co(F_t(\Gamma_t)\hat{\pi}_t)),\ t = 0,\ldots,T \quad (3.5.25)$$

then $C(\hat{z},\Omega')$ is a conical approximation to Ω' at $\hat{z}$ where Ω' is defined in (3.5.24). ($IC(\cdot,\cdot)$ is defined in equation (2.3.2) and $RC(\cdot,\cdot)$ is defined in equation (2.2.9)).

Proof: Recall that IC $(\hat{\pi}_t,S_t)$ is a conical approximation to S_t at $\hat{\pi}_t$ and that $RC(\hat{v}_t,co(F_t(\Gamma_t)\hat{\pi}_t))$ is a conical approximation to $co(F_t(\Gamma_t)\pi_t)$ since it is convex. It is easy to verify that $C(z,\Omega')$ is a convex cone, thus we need only show that for any finite set of independent vectors $\{\delta z_1,\ldots,\delta z_\ell\}$ there exists an $\epsilon > 0$ such that

$$\hat{z} + \delta z \in \Omega' \ \forall \delta z \ \varepsilon\ co(\delta z_1,\ldots,\delta z_\ell). \quad (3.5.26)$$

Let $\delta z_j = (\delta\pi_{1j},\ldots,\delta\pi_T,\delta v_{oj},\ldots,\delta v_{Tj})$. Since $IC(\cdot,\cdot)$ and $RC(\cdot,\cdot)$ are conical approximations there exists an $\epsilon^* > 0$ such that, for $t = 1,\ldots,T, \forall \epsilon \in [0,\epsilon^*]$

$$\hat{\pi}_t + \delta\pi_t \in S_t \ \forall \delta\pi_t \ \varepsilon\ co(\delta\pi_{t1},\ldots,\delta\pi_{t\ell}) \quad (3.5.27)$$

and, for $t = 0,\ldots,T$, $\forall\ \epsilon\ \varepsilon\ [0,\epsilon^*]$

$$\hat{v}_t + (\delta v_t - F_t(\hat{\gamma}_t)\delta\pi_t)\ \varepsilon\ co(F_t(\Gamma_t))\hat{\pi}_t \quad (3.5.28)$$

$$\forall (\delta v_t - F_t(\hat{\gamma}_t)\delta\pi_t)\ \varepsilon\ co(\delta v_{t1} - F_t(\hat{\gamma}_t)\delta\pi_{t1},\ldots,\ v_{t\ell} - F_t(\hat{\gamma}_t)\delta\pi_{t\ell}).$$

That is, for all $\alpha_i \geq 0$, $i = 1,\ldots,\ell$ such that $\sum_{i=1}^{\ell} \alpha_i \leq 1$,

$$\hat{\pi}_t + \sum_{i=1}^{\ell} \alpha_i \delta\pi_{ti}\ \varepsilon\ S_t,\quad t = 1,\ldots,T \quad (3.5.29)$$

$$\hat{v}_t + * \sum_{i=1}^{\ell} \alpha_i (\delta v_{ti} - F_t(\hat{\gamma}_t)\delta\pi_{ti}) \; \varepsilon \; co \; (F_t(\Gamma_t))\hat{\pi}_t, \; t = 0,\ldots,T, \tag{3.5.29}$$

or

$$\hat{\pi}_t + *\delta\pi_t(\underline{\alpha}) \; \varepsilon \; S_t, \qquad t = 1,\ldots,T \tag{3.5.30}$$

$$\hat{v}_t + *\delta v_t(\underline{\alpha}) \; \varepsilon \; co(F_t(\Gamma_t))\hat{\pi}_t + *F_t(\hat{\gamma}_t)\delta\pi_t(\underline{\alpha})$$

$$\varepsilon \; co(F_t(\Gamma_t))(\hat{\pi}_t + *\delta\pi_t(\underline{\alpha})), \quad t = 0,\ldots,T \tag{3.5.31}$$

where $\underline{\alpha} = (\alpha_1,\ldots,\alpha_\ell)$.

But then ϵ^* is such that $z = \hat{z} + \delta z \; \varepsilon \; \Omega'$ for all $\delta z \; \varepsilon \; co(\delta z_1,\ldots,\delta z_\ell)$ since $z = \hat{z} + \delta z$ is such that the corresponding $\pi_t = \hat{\pi}_t + \delta\pi_t$ and $v_t = \hat{v}_t + \delta v_t$ satisfy $\pi_t \; \varepsilon \; S_t$, $t = 1,\ldots,T$ and $v_t \; \varepsilon \; co(F_t(\Gamma_t))\pi_t$, $t = 0,\ldots,T$, thus $C(\hat{z},\Omega')$ is a conical approximation to the set Ω' at $\hat{z}$.

All the requirements of Corollary 2.1.4 are now satisfied, thus we can apply Theorem 2.1.2 using Ω'.

Let $\tilde{K}_t^*$ be the dual cone to $\tilde{K}_t$ where $\tilde{K}_t = \overline{\{y \mid \exists \; \beta > 0 \ni (y - G_t\hat{\pi}_t) \; \varepsilon \; K_t\}}$

<u>Theorem 3.5.1</u> <u>(Maximum Principle)</u>

If $\hat{\gamma} = (\hat{\gamma}_o,\ldots,\hat{\gamma}_{T-1})$ is optimal in problem (3.5.5), Assumption 3.5.1 holds and $(\hat{\pi}_o,\ldots,\hat{\pi}_T)$ is the corresponding "state" trajectory, then there exist costates $(\hat{\phi}_o^*,\ldots,\hat{\phi}_T^*)$, $\hat{\phi}_t^* \; \varepsilon \; \Pi_t^*$, vectors $k_t^* \; \varepsilon \; \tilde{K}_t^*$, a scalar $\lambda \leq 0$, and vectors $\psi_t^* \; \varepsilon \; [\mathbb{R}^{e_t}]^*$, not all zero, such that:

(i) $\hat{\phi}_t^* = P_t^*(\hat{\gamma}_t)\hat{\phi}_{t+1}^* + \lambda h_t^*(\hat{\gamma}_t) + G_t^* k_t^* + R_t^*\psi_t^*, \; t=T-1,\ldots,0,$ (3.5.32)

(ii) $\hat{\phi}_T = G_t^* k_T^* + R_T^*\psi_T^* + \lambda h_T$ (3.5.33)

(iii) $\langle k_t^*, G_t\hat{\pi}_t \rangle = 0, \; t = 1,\ldots,T$ (3.5.34)

(iv) $H_t(\hat{\pi}_t,\hat{\gamma}_t,\hat{\phi}_{t+1}^*) \geq H_t(\hat{\pi}_t,\gamma_t,\hat{\phi}_{t+1}^*) \;\; \forall \gamma_t \; \varepsilon \; \Gamma_t, \; t = 0,\ldots,T-1$ (3.5.35)

where

$$H_t(\hat{\pi}_t,\gamma_t,\hat{\phi}^*_{t+1}) = \lambda\langle h^*_t(\gamma_t),\hat{\pi}_t\rangle + \langle\hat{\phi}^*_{t+1},P_t(\gamma_t)\hat{\pi}_t\rangle. \quad (3.5.36)$$

Proof: For the full proof see Appendix C. We outline that proof here.

(i) if R_t is not surjective or if G_t is not positively linearly independent in the rows corresponding to $G^i_t\hat{\pi}_t = 0$ then the Theorem follows trivially. Under the opposite assumption (3.5.34) holds.

(ii) Corollary 2.1.4 implies the existence of $\lambda \leq 0$, $\hat{\phi}^*_t \in \Pi^*_t$ and $\psi^*_t \in [\mathbb{R}^{e_t}]^*$ not all zero such that

$$\lambda \sum_{t=0}^{T} \delta v^o_t + \sum_{t=0}^{T-1} \langle -\hat{\phi}^*_{t+1},\delta\pi_{t+1}-\delta v^1_t\rangle + \sum_{t=0}^{T}\langle\psi^*_t,R_t\delta\pi_t\rangle \leq 0 \quad \forall\,\delta z \in \overline{C(\hat{z},\Omega')} \quad (3.5.37)$$

(iii) if $\delta z = (0,\ldots,0,\delta v_t,0,\ldots,) \in \overline{C(\hat{z},\Omega')}$, $t = 0,\ldots,T-1$ then

$$\lambda v^o_t + \langle\hat{\phi}^*_{t+1},\delta v^1_t\rangle \leq 0 \quad \forall\,\delta v_t \in RC(\hat{v}_t,co(F_t(\Gamma_t))\hat{\pi}_t) \quad (3.5.38)$$

this, and properties of radial cones implies (3.5.35).

(iv) if $\delta z = (0,\ldots,0,\delta\pi_T,0,\ldots,0,\delta v_T)$ where $\delta v_T = h_T\delta\pi_T$, then (3.5.37) plus Theorem 2.3.5 implies (3.5.33) and

(v) if $\delta z = (0,\ldots,\delta\pi_t,0,\ldots,0,\delta v_t,0,\ldots,0)$, $t = 0,\ldots,T-1$ where

$$F_t(\hat{\gamma}_t)\delta\pi_t = \delta v_t \quad (3.5.39)$$

then (3.5.37) plus Theorem 2.3.5 implies (3.5.32).

We note the following about Theorem 3.5.1.

1) We have assumed that K_t is the positive orthant in $\mathbb{R}^{k_t}$, thus K_t is the set of vectors the components of which are each greater than or equal to zero. This assumption is used in the derivation of equation (3.5.34)

2) No differentiability assumptions have been made about the original problem (problem 3.5.1). In fact, the functions f_t, h_t, r_t

and g_t need not even be continuous with respect to x_t. Of course, severe restrictions are placed on f_t and h_t by the convexity assumption (Assumption 3.5.1).

3) Assumption 3.5.1 corresponds to the directional convexity assumptions of [Ca 1] which are usually placed on the reachable sets. Since the equations of the EDP are linear in π_t, convexity of the reachable set corresponding to any one π_t implies the convexity of the reachable set for all π_t's. Thus Assumption 3.5.1 is in terms of γ_t only.

4) Note that Ω and Ω' do not contain any interior points since M_t does not. This is why it was important in Chapter II to consider sets which contain no interior points (that is, which do not satisfy the conditions of Corollary 2.1.1 or 2.2.2).

The following corollaries consider various sets of assumptions under which the results of Theorem 3.5.1 can be strengthened or extended.

Let G_t: $\Pi_t \to \mathbb{R}^{k_t}$ be decomposed into maps G_t^i: $\Pi_t \to \mathbb{R}$, $i = 1,\dots,k_t$ and let $\{G_t^i\}^I$ be the set of maps such that $G_t^i \in \{G_t^i\}^I$ implies $G_t^i\hat{\pi}_t = 0$. Let $K_t^i = \{r \in \mathbb{R} | r \geq 0\}$. Note that for this decomposition of K_t, K_t-linear independence is equivalent to positive linear independence.

<u>Corollary 3.5.1</u> If, in addition to the assumptions of Theorem 3.5.1, the following hold:

(i) R_t is surjective, $t = 1,\dots,T$, and

(ii) $\{G_t^i\}^I$ are positively linearly independent, then the results of Theorem 3.5.1 hold and $(\phi_o^*,\dots,\phi_T^*),\lambda,(\psi_1^*,\dots,\psi_T^*)$ are not all zero.

<u>Proof</u>: Follows immediately from the proof of Theorem 3.5.1.

<u>Corollary 3.5.2</u> If $R_t \equiv 0$ and $G_t \equiv 0$, $t = 1,\dots,T$ then Theorem 3.5.1 holds with $\lambda = -1$.

Proof: If not, $\hat{\phi}_T^* \equiv 0$ and thus, by equation (3.5.32) $\hat{\phi}_t^* \equiv 0$, $t = 0,\ldots,T-1$. But this contradicts the existence of a nonzero multiplier. Thus $\lambda < 0$ and clearly all $\hat{\phi}_t^*$ can be normalized so that $\lambda = -1$.

If K_t is not the positive orthant then Theorem 3.5.1 and Corollary 3.5.1 do not hold. The reason is that an active constraint need not have the form $G_t\hat{\pi}_t = 0$, rather an active constraint is one for which $G_t\ddot{\hat{\pi}}_t \in \partial K_t$.

Corollary 3.5.3 If equation (3.5.34) is replaced by $\langle k_t^*, G_t\hat{\pi}_t\rangle \geq 0$, $t = 1,\ldots,T$, $\mathcal{R}(G_t)$ is closed and $\tilde{K}_t \subset \mathcal{R}(G_t)$, then Theorem 3.5.1 holds for any set of closed convex cones $\{K_t\}$ such that int $K_t \neq \phi$.

Proof: Follows from the proof of Theorem 3.5.1 and Theorem 2.3.2.

Corollary 3.5.4 If

(i) R_t is surjective, $t = 1,\ldots,T$, and

(ii) either G_t is K_t-linearly independent or $G_t\hat{\pi}_t \in$ int K_t, $t = 1,\ldots,T$, then Corollary 3.5.3 holds and $(\hat{\phi}_o^*,\ldots,\hat{\phi}_T^*),\lambda,$ $(\psi_1^*,\ldots,\psi_2^*)$ are not all zero.

Proof: Follows from the proof of Theorem 3.5.1.

The assumption that G_t has its range in a finite dimensional space is inessential. Theorem 2.1.2 which was used in the proof of Theorem 3.5.1 is valid if G_t has its range in any Banach space. Of course the function $g_t(x_t)$ corresponding to G_t must be measurable, thus the range of g_t must be in a measurable space. But Borel sets always exist in a Banach space (they are generated by the spheres with rational radii), thus it is meaningful to require that $g_t(x_t)$ be Borel measurable.

Corollary 3.5.5 Theorem 3.5.1 and Corollaries 3.5.1 through 3.5.4 hold if g_t: $X_t \to Y_t$ is a Borel measurable function and Y_t is a Banach space.

Proof: As above.

Corollary 3.5.5 is interesting because ℓ_∞ is a Banach space and its positive orthant has a nonempty interior. Thus it is possible to consider a problem with a countably infinite number of inequality constraints.

Section 3.6 Another Optimality Condition

In Section 3.5 we had to make Assumption 3.5.1 to derive Theorem 3.5.1. It is, of course, a very restrictive assumption and it is the purpose of this section to consider a weaker result that can be derived under different conditions.

Recall that the EDP is:

$$\underset{\gamma \in \Gamma}{\text{minimize}} \sum_{t=0}^{T} \langle h_t^*, (\delta_t), \pi_t \rangle \tag{3.6.1}$$

subject to

$$\pi_{t+1} = P_t(\gamma_t)\pi_t \qquad t = 0,\ldots,T-1 \tag{3.6.2}$$

$$G_t\pi_t \in K_t \qquad t = 1,\ldots,T \tag{3.6.3}$$

$$R_t\pi_t = 0 \qquad t = 1,\ldots,T \tag{3.6.4}$$

where all the sets and functions are as in problem (3.5.5).

We make the following assumptions.

Assumption 3.6.1 The functionals $h_t^*(\gamma_t)$, $t = 0,\ldots,T-1$ and the linear operators $P_t(\gamma_t)$, $t = 0,\ldots,T-1$ are continuously Fréchet differentiable on Γ_t, $t = 0,\ldots,T-1$.

Note that this is a condition on the EDP. Later in this chapter we will consider this assumption in terms of the BP.

Assumption 3.6.2 The sets Γ_t, $t = 0,\ldots,T-1$ each lie in a Banach space G_t, $t = 0,\ldots,T-1$ and each set satisfies at least one of the following conditions:

(i) Γ_t contains a convex set A_t such that int $A_t \neq \phi$, or

(ii) Γ_t contains a set B_t such that if $\hat{\gamma}_t + \delta\gamma_t \in B_t$ then there exists an $\epsilon_t^* > 0$ and an $\alpha_t^* > 0$ such that

$$\hat{\gamma}_t + \epsilon(\delta\gamma_t + \delta\zeta_t) \in B_t, \ \forall \epsilon \in [0, \epsilon_t^*], \quad \forall \delta\zeta_t \in S(0, \alpha_t^*) \ . \tag{3.6.5}$$

Let us consider for a moment the assumption that $\Gamma_t \subset G_t$ where G_t is a Banach space. We shall see that this is always true if we slightly modity the set Γ_t.

Let $\eta_t(\gamma_t, \pi_t)$ be a probability measure on U_t defined by

$$\eta_t(\gamma_t, \pi_t)(B) = \pi_t(\{x | \gamma_t(x) \in B\}) \tag{3.6.6}$$

where B is any Borel set in X_t and $\gamma_t \in \Gamma_t$, where Γ_t is an arbitrary set in the space $M(X_t, U_t)$ of Borel measurable mappings from X_t into U_t. Let $N(\pi_t) = \{\zeta_t \in M(X_t, U_t) | \pi_t(\{x| \, \|\zeta_t(x)\| > \alpha\}) = 0, \forall \alpha \geq 0\}$. The set $N(\pi_t)$ induces equivalence classes in $M(X_t, U_T)$ if γ_1 and γ_2 in $M(X_t, U_t)$ are said to be equivalent whenever $\gamma_1 - \gamma_2 \in N(\pi_t)$, this will be denoted $\gamma_1 \equiv \gamma_2$.

If $\gamma_1 \equiv \gamma_2$ then [Du 1, Definition 3.2.13]

$$\pi_t(\{x | \gamma_1(x) \in B\}) = \pi_t(\{x | \gamma_2(x) \in B\}) \tag{3.6.8}$$

for any Borel set B. Thus

$$\eta_t(\gamma_1, \pi_t) = \eta_t(\gamma_2, \pi_t) \quad \forall \gamma_1 \equiv \gamma_2 \in \Gamma_t \ .$$

The mapping (I, γ_t), where I is the identity map in X, thus induces a unique probability measure on $X_t \times U_T$ for all equivalent γ_t. Then $f_t(x_t, w_t, u_t)$ induces a unique probability measure π_{t+1} on X_{t+1} for all equivalent γ_t.

By [Du 1, Theorem 3.6.5] $M(X_t, U_t)/N(\pi_t)$ is a Banach space if the norm is defined by

$$||f|| = \inf_{\alpha>0} \arctan \{\alpha + TV(\{x \,|\, |f(x)| > \alpha\}, \pi_t)\} \tag{3.6.7}$$

where $TV(E,\pi)$ is the total variation [Du 1, page 97] of the set E with respect to measure π. Note that the principle value of the arctan is used in equation (3.6.7).

Thus $\Gamma_t/N(\pi_t)$ lies in a Banach space. Note that restricting attention to $\gamma_t \in \Gamma_t/N(\pi_t)$ does not change the cost associated with problem (3.6.1) so this substitution is innocuous. Thus, without loss of generality, we may always assume that Γ_t lies in a Banach space G_t.

Note, however, that it is not always advantageous to take G_t as the space just defined. If, for example, Γ_t is a subset of continuous linear maps then int $\Gamma_t = \phi$ if G_t is taken to be $M(X_t,U_t)/N_t(\pi_t)$. However Γ_t may have interior points when considered as a set in the closed subspace $B(X_t,U_t)$ of $M(X_t,U_t)/N(\pi_t)$.

For notational convenience we shall continue to write Γ_t even when it refers to an equivalence class of functions.

The purpose of Assumption 3.6.2 is to allow us to construct conical approximations of a type that will allow us to apply Theorem 2.2.1. Clearly if part (i) holds then a conical approximation exists which satisfies Corollary 2.2.1 (where, if Γ_t is convex we take $A_t = \Gamma_t$). If part (ii) holds then there exists a conical approximation satisfying Corollary 2.2.2. For the remainder of this section we shall assume that any conical approximation to Γ_t at $\hat{\gamma}_t$ is constructed so as to satisfy Corollary 2.2.1 or Corollary 2.2.2.

To transform problem (3.6.1) into a form to which Theorem 2.2.1 can be applied we need only define

$$z = (\pi_1,\ldots,\pi_T,\gamma_o,\ldots,\gamma_{T-1}) \tag{3.6.9}$$

$$Z = \Pi_1 \times \ldots \times \Pi_T \times \Gamma_o \times \ldots \times \Gamma_{T-1} \tag{3.6.10}$$

$$R(z) = [R_1\pi_1, \ldots, R_T\pi_T] \tag{3.6.11}$$

$$T(z) = \begin{bmatrix} \pi_1 - P_o(\gamma_o)\pi_o \\ \vdots \\ \pi_T - P_{T-1}(\gamma_{T-1})\pi_{T-1} \end{bmatrix} \tag{3.6.12}$$

$$S_t = \{\pi_t | G_t\pi_t \in K_t\} \qquad t = 1,\ldots,T \tag{3.6.13}$$

$$f(z) = \sum_{t=0} \langle h_t^*(\gamma_t), \pi_t \rangle \tag{3.6.14}$$

and

$$\Omega = S_1 \times \ldots \times S_T \times \Gamma_o \times \ldots \times \Gamma_{T-1} \tag{3.6.15}$$

Problem (3.6.1) can now be written

$$\underset{z \in \Omega}{\text{minimize}}\ f(z) \tag{3.6.16}$$

subject to

$$T(z) = 0 \quad Rz = 0 \tag{3.6.17}$$

Note that $C(\hat{z},\Omega) = IC(\hat{\pi}_1,S_1) \times \ldots \times IC(\hat{\pi}_T,S_T) \times C(\hat{\gamma}_o,\Gamma_o) \times \ldots \times C(\hat{\gamma}_{T-1},\Gamma_{T-1})$ is a conical approximation to Ω at $\hat{z}$.

Recall that $\mathcal{R}(\nabla T(\hat{x}))$ must be closed to be able to apply Theorem 2.2.1 to problem (3.6.16). But $\nabla T(\hat{z})$ is given by

$$\begin{bmatrix} I & 0 & & 0 & -\nabla_\gamma P_o(\hat{\gamma}_o)\pi_o & 0 & & 0 \\ -P_1(\hat{\gamma}_1) & I & 0 \ldots & 0 & 0 & -\nabla_\gamma P_1(\hat{\gamma}_1)\pi_1 & 0 \ldots & 0 \\ & & P_{T-1}(\hat{\gamma}_{T-1}) & I & 0 & & & -\nabla_\gamma P_{T-1}(\hat{\gamma}_{T-1})\pi_{T-1} \end{bmatrix} \tag{3.6.18}$$

It is now clear that $\mathcal{R}(\nabla T(\hat{z})) = \Pi_1 \times \ldots \times \Pi_T$ since for any $\tilde{\pi} = (\tilde{\pi}_1,\ldots,\tilde{\pi}_T) \in \Pi_1 \times \ldots \times \Pi_T$ one can choose

$$\tilde{z} = (\tilde{\pi}_1, \tilde{\pi}_2 + P_1(\hat{\gamma}_1)\tilde{\pi}_1, \ldots, \tilde{\pi}_T + P_{T-1}(\hat{\gamma}_{T-1})\tilde{\pi}_{T-1}, 0, \ldots, 0) \tag{3.6.19}$$

so that $\tilde{\pi} = T(\tilde{z})\tilde{z}$ and thus $\mathcal{R}(\nabla T(\hat{z}))$ is closed.

The application of Theorem 2.2.1 (and Corollaries 2.2.1 and 2.2.2) to problem 3.6.16) thus yields the following theorem:

Theorem 3.6.1 (Quasi Maximum Principle)

If $\hat{\gamma} = (\hat{\gamma}_o, \ldots, \hat{\gamma}_{T-1})$ is optimal in problem (3.6.1), Assumptions 3.6.1 and 3.6.2 hold and $(\hat{\pi}_o, \ldots, \hat{\pi}_T)$ is the corresponding "state" trajectory, then there exist costates $(\hat{\phi}^*_o, \ldots, \hat{\phi}^*_T)$, $\hat{\phi}^*_t \in \Pi^*_t$, vectors $k^*_t \in \tilde{K}^*_t$, $\psi^*_t \in [\mathbb{R}^{e_t}]^*$, $t = 1, \ldots, T$ and a scalar $\lambda \leq 0$, not all zero and conical approximations, $C(\hat{\gamma}_t, \Gamma_t)$, to the set Γ_t at $\hat{\gamma}_t$, $t = 0, \ldots, T-1$ such that

(i) $\hat{\phi}^*_t = P^*_t(\hat{\gamma}_t)\hat{\phi}_{t+1} + \lambda h^*_t(\hat{\gamma}_t) + G^*_t k^*_t + R^*_t \psi^*_t$, $t = T-1, \ldots, 0$, (3.6.20)

(ii) $\hat{\phi}^*_T = G^*_T k^*_T + R^*_T \psi^*_T + \lambda h^*_T$ (3.6.21)

(iii) $\langle k^*_t, G_t \hat{\pi}_t \rangle = 0$, $t = 1, \ldots, T$ (3.6.22)

(iv) $\nabla_{\gamma_t} H_t(\hat{\pi}_t, \hat{\gamma}_t, \hat{\phi}^*_{t+1})\delta\gamma_t \leq 0 \quad \forall \delta\gamma_t \in \overline{C(\hat{\gamma}_t, \Gamma_t)}$, $t=0, \ldots, T-1$ (3.6.23)

where $H_t(\hat{\pi}_t, \hat{\gamma}_t, \hat{\phi}^*_{t+1})$ is given in equation (3.5.36).

Proof: For the full proof see Appendix C. We outline that proof here.

(i) if R_t is not surjective or if G_t is not positively linearly independent in the rows corresponding to $G^i_t \hat{\pi}_t = 0$ then the theorem follows trivially. Under the opposite assumption (3.6.22) holds.

(ii) Theorem 2.2.1 implies the existence of a $\lambda \leq 0$ and $\phi^*_t \in \Pi^*_t$ and $\psi^*_t \in [\mathbb{R}^{e_t}]^*$ not all zero such that

$$\lambda\left(\sum_{t=1}^{T-1}\langle h_t^*(\hat{\gamma}_t),\delta\pi_t\rangle + \sum_{t=0}^{T-1}\langle\nabla_\gamma h_t^*(\hat{\gamma}_t)\delta\gamma_t,\hat{\pi}_t\rangle + \langle h_T^*\delta\pi_T\rangle\right) +$$

$$\sum_{t=0}^{T}\langle\phi_{t+1}^*,\delta\pi_{t+1} - P_t(\hat{\gamma}_t)\delta\pi_t - \nabla_\gamma P_t(\hat{\gamma}_t)\delta\gamma_t\hat{\pi}_t\rangle +$$

$$\sum_{t=1}\langle\psi_t^*,R_t\delta\pi_t\rangle \leq 0 \quad \forall\delta z_t \in \overline{C(\hat{z},\Omega)} \tag{3.6.24}$$

(iii) if $\delta z = (0,\ldots,\delta\gamma_t,0,\ldots,0) \in C(\hat{z},\Omega)$, $t = 0,\ldots,T-1$ then equation (3.6.23) follows,

(iv) if $\delta z = (0,\ldots,0,\delta\pi_t,0,\ldots,0) \in C(\hat{z},\Omega)$, $t = 0,\ldots,T-1$ then Theorem 2.3.5 implies equation (3.6.20), and

(v) if $\delta z = (0,\ldots,0,\delta\pi_T,\ldots,0) \in C(\hat{z},\Omega)$ then Theorem 2.3.5 implies equation (3.6.21).

We note the following about Theorem 3.6.1:

1) All the Corollaries of Theorem 3.5.1 also hold for Theorem 3.6.1 since they all deal with conditions under which the Farkas-Minkowski lemma (Theorem 2.3.5) hold.

2) If Γ_t is a Banach space, then equation (3.6.23) clearly holds with equality for all $\delta\gamma_t$.

3) Note that Assumption 3.6.2 requires that Γ_t have interior points. We have seen that if Γ_t is the class of linear feedback laws then this holds trivially. In many cases Γ_t will be a parameter set for a class of feedback laws (e.g. matrices in linear laws, coefficients in polynomical laws, etc.). In this Γ_t will have interior points if the class of feedback laws is not "overparameterized", and thus it should not be too difficult a requirement to meet.

4) As was noted in section 2.5, Assumption 3.6.2 can be weakened to a requirement that a conical approximation exists. However, the proofs required to demonstrate this are significantly more complex than those of Chapter II. We do not feel that the concomitant loss of

clarity is justified in this case.

5) The requirement of Fréchet differentiability in Assumption 3.6.1 is on the functionals and operators of the EDP. While this is still much weaker than Assumption 3.5.1, it is not easily checked.

We shall now consider a class of problems for which the requirement of Fréchet differentiability can be stated on the BP. Recall that the BP is:

$$\underset{\gamma \in \Gamma}{\text{minimize}} \ E\{\sum_{t=0}^{T} h_t(x_t;\gamma_t)\}; h_T(x_T;\gamma_T) \triangleq h_T(x_T) \qquad (3.6.25)$$

subject to

$$x_{t+1} = f_t(x_t,w_t;\gamma_t) \qquad t = 0,\ldots,T-1 \qquad (3.6.26)$$

and

$$E\{g_t(x_t)\} \in K_t \qquad t = 1,\ldots,T \qquad (3.6.27)$$

$$E\{r_t(x_t)\} = 0 \qquad t = 1,\ldots,T. \qquad (3.6.28)$$

Also, μ_t is the probability measure associated with the random variable w_t so that the state transition stochastic kernel is given by

$$p_t(B|x_t;\gamma_t) = \mu_t(\{w|f_t(x_t w_t;\gamma_t) \in B\}) \qquad (3.6.29)$$

for all Borel sets B in X_{t+1}.

If $f_t(x_t,w_t;\gamma_t)$ has the following form:

$$x_{t+1} = f_t(x_t;\gamma_t) + w_t \qquad t = 1,\ldots,T-1 \qquad (3.6.30)$$

then (3.6.29) can be written

$$p_t(B|x_t;\gamma_t) = \int_B \mu_t(dx_{t+1} - f_t(x_t;\gamma_t)). \qquad (3.6.31)$$

If μ_t has a corresponding density p_t^{ω} then (3.6.31) becomes

$$p_t(B|x_t;\gamma_t) = \int_B p_t^{\omega}(x_{t+1} - f_t(x_t;\gamma_t))dx_{t+1} \qquad (3.6.32)$$

and we write (3.6.26) as

$$p_{t+1}(x_{t+1}) = \int_{X_t} p_t(x_{t+1}|x_t;\gamma_t)p_t(x_t)dx_t \qquad (3.6.33)$$

or

$$p_{t+1}(x_{t+1}) = \int_{X_t} p_t^{\omega}(x_{t+1} - f_t(x_t;\gamma_t))p_t(x_t)dx_t \,. \qquad (3.6.34)$$

Thus $\nabla\gamma_t P_t(\gamma_t)\pi_t$ becomes

$$\frac{d}{d\gamma_t}\int_{X_t} p_t^{\omega}(x_{t+1} - f_t(x_t;\gamma_t))p_t(x_t)dx_t \qquad (3.6.35)$$

which, by [Rud 1, Theorem 9.42] can be written as

$$\int_{X_t} \frac{d}{d\gamma_t} p_t^{\omega}(x_{t+1} - f_t(x_t;\gamma_t))p_t(x_t)dx_t \qquad (3.6.36)$$

if $p_t^{\omega}(x_{t+1} - f_t(x_t;\gamma_t))$ is a continuous function of γ_t. This is certainly the case if $f_t(x_t;\gamma_t)$ is Fréchet differentiable with respect to γ_t and if $p_t^{\omega}(x)$ is differentiable. Then (3.6.36) becomes

$$-\int_{X_t} (\frac{d}{dx} p_t^{\omega})(x_{t+1} - f_t(x_t;\gamma_t))\frac{d}{d\gamma_t} f_t(x_t;\gamma_t)p_t(x_t)dx_t \qquad (3.6.37)$$

by the Fréchet derivative chain rule (Theorem A.7.1).

Thus if μ_t is twice differentiable, $f_t(x_t,w_t;\gamma_t)$ has the form specified in equation (3.6.30) and if $f_t(x_t;\gamma_t)$ is Fréchet differentiable with respect to γ_t, then $P_t(\gamma_t)$ is Fréchet differentiable with respect to γ_t. If $h_t(x_t;\gamma_t)$ is also Fréchet differentiable with respect to γ_t then Assumption 3.6.1 is satisfied.

Note that differentiability of $f_t(x_t;\gamma_t)$ and $h_t(x_t;\gamma_t)$ with respect to x_t is <u>not</u> required, in fact continuity is not required. Also note that Fréchet differentiability with respect to γ_t implies nothing about the continuity or differentiability of $\gamma_t(x)$.

By the definition of Frechet differentiability (section A.7) we have that

$$\frac{d}{d\gamma_t} f_t(x_t;\gamma_t) \triangleq \frac{d}{d\gamma_t} f_t(x_t,u_t) \tag{3.6.38}$$

$$= \frac{d}{d} f(x_{t1}\gamma_t(x_t) + \epsilon\delta\gamma_t(x_t))\Big|_{\epsilon=0} \tag{3.6.39}$$

$$= \frac{d}{du} f(x_t,\gamma_t(x_t))\delta\gamma_t(x_t), \tag{3.6.40}$$

thus we need only require that $h_t(x_t,u_t)$ and $f_t(x_t,w_t,u_t)$ be differentiable with respect to u_t.

Thus we have:

<u>Theorem 3.6.2</u> If in problem (3.6.25) the functions $h_t(x_t,u_t)$ and $f_t(x_t,w_t,u_t)$ (defined by equation (3.6.30)) are differentiable with respect to u_t, the measures μ_t are twice differentiable and Assumption 3.6.2 holds then the results of Theorem 3.6.1 follow.

<u>Proof</u>: Already given.

<u>Corollary 3.6.1</u> If Γ_t is a class of linear feedback laws for which Assumption 3.6.2 holds then the LQG problem satisfies the assumptions of Theorem 3.6.1.

<u>Proof</u>: Since the cost is quadratic in $u_t h_t$ is differentiable, since the dynamics are linear in u_t, f_t is differentiable and since μ_t is Gaussian it is twice differentiable. Theorem 3.6.2 then implies that Corollary 3.6.1 holds.

Section 3.7. Extension and Alternate Derivation

In this section we consider slightly stronger versions of Theorems 3.5.1 and 3.6.1 as well as an alternative derivation of these theorems.

One of the drawbacks of Theorems 3.5.1 and 3.6.1 is that, in general, a satisfactory representation of the continuous linear functionals on Π is not available [Du 1, Section 4.15]. However, due to the special structure of those problems it is possible to give an effective representation of ϕ_t^* if h_t, g_t and r_t are bounded functions.

Note that we are distinguishing between ϕ_t^*, which is a functional on Π_t and $\phi_t(x_t)$ which is a Borel measurable function on $\mathbb{R}^{n_t}$. For example

$$\langle h_T^*, \pi_T \rangle = \int h_T(x_T)\pi_T(dx_T), \tag{3.7.1}$$

thus h_T^* is defined by an integral functional on Π_T with a kernel $h_T(x_T)$ which is Borel measurable. What we are doing is showing that ϕ_t^* has a similar representation and that Theorems 3.5.1 and 3.6.1 can be written in terms of these kernels.

Recall from Theorems 3.5.1 and 3.6.1 (specifically equations (3.5.33) and (3.6.21)) that

$$\phi_T^* = \lambda h_T^* + G_T^* k_T^* + R_T^* \psi_T^* \tag{3.7.2}$$

where $h_T^* \in \Pi_T^*$, G_T: $\Pi_T \to \mathbb{R}^{\ell_T}$, R_T: $\Pi_T \to \mathbb{R}^{e_T}$, $k_T^* \in [\mathbb{R}^{\ell_T}]^*$ and $\psi_T^* \in [\mathbb{R}^{e_T}]^*$. Also,

$$\langle G_T^* k_T^*, \pi_T \rangle = \langle k_T^*, G_T \pi_T \rangle \tag{3.7.3}$$

$$= \sum_{i=1}^{\ell_T} k_t^i \int g_T^i(x_T)\pi_T(dx_T) \tag{3.7.4}$$

$$= \left(\sum_{i=1}^{\ell_T} k_T^i g_T^i(x_T)\right)\pi_T(dx_T) \quad \dagger \tag{3.7.5}$$

and similarly

$$\langle R_T^*, \psi_T^*, \pi_T \rangle = \int \left(\sum_{i=1}^{e_T} \psi_T^i r_T^i(x_T)\pi_T(dx_T)\right) . \tag{3.7.6}$$

Clearly

$$\langle \phi_T^*, \pi_T \rangle = \int \phi_T(x_T)\pi_T(dx_T) \tag{3.7.7}$$

---

† This is valid since $k_T < +\infty$. If $k_T = +\infty$ (as in Corollary 3.5.5) then we need to make further assupmtions about $g_T^i(x_T)$. For example, if $g_T^i(x_T) \geq 0$ for all i then Lebesque's monotone convergence Theorem [Rud 1] implies (3.7.5). We shall see however that other problems arise when $k_t = +\infty$.

where ϕ_T is Borel measurable and bounded since g, h and r are Borel measurable and bounded.

Also recall

$$\langle P^*_{T-1}(\gamma_{T-1})\phi^*_T, \pi_{T-1}\rangle = \langle \phi^*_T, P_{T-1}(\gamma_{T-1})\pi_{T-1}\rangle \tag{3.7.8}$$

$$= \int \phi_T(x_T)\,[\int p_t(dx_T|x_{T-1};\gamma_{T-1})\pi_{T-1}(dx_{T-1})] \tag{3.7.9}$$

$$= \int\, [\int \phi_T(x_T)p_{T-1}(dx_T|x_{T-1};\gamma_{T-1})]\pi_{T-1}(dx_{T-1}) \tag{3.7.10}$$

By [Bert 1, Proposition 7.29].

$$\int\, \phi_T(x_T)p_{T-1}(dx_T|x_{T-1};\gamma_T) \tag{3.7.11}$$

is a bounded Borel measurable function of x_{T-1}. An argument similar to that leading to equation (3.7.7) implies

$$\langle \phi^*_{T-1}, \pi_{T-1}\rangle = \int \phi_{T-1}(x_{T-1})\pi_{T-1}(dx_{T-1}) \quad . \tag{3.7.12}$$

A trivial induction argument then yields the fact that ϕ^*_t has a representation as an integral functional with a Borel measurable kernel.

Theorem 3.5.1 and 3.6.1 can now be written in terms of these kernels. Since the equations for both theorems are similar we give only those corresponding to (3.5.32) through (3.5.36).

$$\phi_t(x_t) = \int \phi_{t+1}(x_{t+1})p_t(dx_{t+1}|x_t;\gamma_t) + \lambda h_t(x_t;\gamma_t) + \sum_{i=1}^{k_t} k^i_t g^i_t(x_t) + \sum_{j=1}^{e_t} \psi^i_t r^i_t(x_t) \qquad t = T-1,\ldots,0 \tag{3.7.13}$$

$$\phi_T(x_T) = \lambda h_T(x_T) + \sum_{i=1}^{k_t} k^i_T g^i_T(x_T) + \sum_{j=1}^{e_t} \psi^i_T r^i_T(x_T) \tag{3.7.14}$$

$$\sum_{i=1}^{k_t} k^i_t \int g^i_t(x_t)\pi_t(dx_t) = 0 \qquad\qquad t = 1,\ldots,T \tag{3.7.15}$$

$$H_t(\hat{\pi}_t,\hat{\gamma}_t,\hat{\phi}_{t+1}) \geq H_t(\hat{\pi}_t,\gamma_t,\hat{\phi}_{t+1}) \quad \forall \gamma_t \in \Gamma_t, t = 0,\ldots,T-1 \quad (3.7.16)$$

where

$$H_t(\hat{\pi}_t,\gamma_t,\hat{\phi}_{t+1}) = \lambda\int h_t(x_t;\gamma_t)\pi_t(dx_t) + \int[\int\phi_{t+1}(x_{t+1})p_t(dx_{t+1}|x_t;\gamma_t)]\pi_t(dx_t) \quad (3.7.17)$$

Note that if g_t: $X_t \times \Gamma_t \to Y_t$ where $Y_t = \ell_\infty$ then the positive orthant K_t is a convex cone with nonempty interior. Here we could consider decomposing $g_t(x_t)$ such that

$$g_t^i: \; X_t \times \Gamma_t \to \mathbb{R} . \qquad i = 1,\ldots,+\infty \qquad (3.7.18)$$

However, in general we cannot characterize ℓ_∞^* as a sequence and thus (3.7.4) does not necessarily follow from (3.7.5). Thus we cannot determine whether ϕ_t^* has a representation as an integral functional with a Borel measurable kernel by proceeding in a similar manner to that in equations (3.7.1) through (3.7.12). It is not clear to the author whether such a representation exists.

In theorem 3.4.1 the function $-\phi_t(x_t)$ has the interpretation of "the expected-cost-to-go" given that the state is x_t and the optimal control law is used over the remaining time steps. Such an interpretation is not available for Theorems 3.5.1 and 3.6.1. However $-\phi_t(x_t)$ is closely related to the cost-to-go function of dynamic programming and an interesting result is available.

Consider a BP for which Γ is a particular class of control laws (e.g. linear). Standard dynamic programming results are not applicable to this problem since one cannot minimize the expected cost-to-go as a function of x_t. Rather, one must choose a single control law which will be used for all values of x_t such that the expected cost-to-go is minimized. Clearly this minimization will depend on π_t, thus it is appropriate to apply dynamic programming to the EDP.

If we proceed as in [San 1, Chapter IV] the dynamic programming recursion is given by

$$V_T(\pi_T) = \langle h_T^*, \pi_T \rangle \tag{3.7.19}$$

$$V_t(\pi_t) = \min_{\gamma_t \in \Gamma_t} \{\langle h_t^*(\gamma_t), \pi_t \rangle + V_{t+1}(P_t(\gamma_t)\pi_t) \,\big|$$

$$G_{t+1}P_t(\gamma_t)\pi_t \in K_{t+1},\; R_{t+1}P_t(\gamma_t)\pi_t = 0\},\; t = 0,\ldots,T-1 \tag{3.7.20}$$

The dependence of the control law on π_t can then be eliminated by use of the forward equation

$$\pi_{t+1} = P_t(\hat{\gamma}_t(\pi_t))\pi_t \qquad t = 0,\ldots,T-1 . \tag{3.7.21}$$

Note that the dynamic programming iteration leads to a mapping $\hat{\gamma}_t(\pi_t)$. The $\hat{\gamma}_t$ in the theorems of preceeding sections corresponds to $\hat{\gamma}_t(\hat{\pi}_t)$. This embedding of a single problem in a family of problems is, of course, the distinguishing feature of dynamic programming.

Under conditions discussed in Chapter II and section 3.5 multipliers exist such that

$$\hat{\gamma}_t(\pi_t) = \operatorname{argmin}\{\lambda(\langle h_t(\gamma_t), \pi_t \rangle + V_{t+1}(P_t(\gamma_t)\pi_t)) +$$

$$k_{t+1}^* G_{t+1}P_t(\gamma_t)\pi_t + \psi_{t+1}^* R_{t+1}P_t(\gamma_t)\pi_t \,|\, \gamma_t \in \Gamma_t\}. \tag{3.7.22}$$

We can write $V_{t+1}(P_t(\gamma_t)\pi_t)$ as

$$\sum_{\tau=t+1}^{T} \langle h_\tau^*(\hat{\gamma}_\tau(\pi_\tau)), P_\tau(\hat{\gamma}_\tau(\pi_\tau)) \ldots P_{t+1}(\hat{\gamma}_{t+1}(\pi_{t+1}))P_t(\gamma_t)\pi_t \rangle \tag{3.7.23}$$

where the π_t's are determined by equation (3.7.21).

It is easy to verify that $\langle \hat{\phi}_{t+1}^*, P_t(\gamma_t)\pi_t \rangle$ is given by

$$\sum_{\tau=t+1} \{\lambda \langle h_\tau^*(\hat{\gamma}(\hat{\pi}_\tau)) + G_\tau^* k_\tau^* + R_\tau^* \psi_\tau^*,\; P_\tau(\hat{\gamma}_\tau(\hat{\pi}_\tau)) \ldots$$
$$P_{t+1}(\hat{\gamma}_{t+1}(\hat{\pi}_{t+1}))P_t(\gamma_t)\pi_t \rangle\}. \tag{3.7.24}$$

If $\gamma_t = \hat{\gamma}_t$ and $\pi_t = \hat{\pi}_t$ then, because $<k_\tau^*, G_\tau \pi_\tau> = 0$ and $<\psi^*, R_\tau \hat{\pi}_\tau> = 0$,

$$<\phi_{t+1}^*, P_t(\hat{\gamma}_t)\hat{\pi}_t> = \lambda' V_{t+1}(P_t(\hat{\gamma}_t)\hat{\pi}_t). \tag{3.7.25}$$

Clearly this result depends on the linearity of $G_t\pi_t$ and $R_t\pi_t$ in π_t and upon evaluating both functionals at $\hat{\pi}_{t+1}$.

Section 3.8. Examples

Let Π_t be the space of signed measures of bounded variations then quadratic functions induce continuous linear functionals on Π_t and we can consider the following LQG problems.

Example 3.8.1

Consider the following problem which is usually solved by use of the matrix minimum principle [Ath 1].

$$\underset{G_t^1, G_t^2}{\text{minimize}}\ E\{x_T' Q_T x_T + \sum_{\tau=0}^{T-1} x_\tau' Q_\tau x_\tau + u_\tau^{1'} R_\tau^1 u_\tau^1 + u_\tau^{2'} R_\tau^2 u_\tau^2\} \tag{3.8.1}$$

subject to

$$x_{t+1} = A_t x_t + B_t^1 u_t^1 + B_t^2 u_t^2 + \zeta_t \qquad t = 0,\ldots,T-1 \tag{3.8.2}$$

$$y_t^1 = C_t^1 x_t^1 \tag{3.8.3}$$

$$y_t^2 = C_t^2 x_t^2 \tag{3.8.4}$$

$$u_t^1 = -G_t^1 y_t^1 \tag{3.8.5}$$

$$u_t^2 = -G_t^2 y_t^2 \tag{3.8.6}$$

where (of course) $x_o, \zeta_o, \ldots, \zeta_{T-1}$ are uncorrelated zero mean Gaussian random variables with associated covariances of $\Sigma_o, \Xi_o, \ldots, \Xi_{T-1}$.

This can be rewritten as

$$\underset{G_t^1, G_t^2}{\text{minimize}}\ E\{x_T' Q_T x_T + \sum_{\tau=0}^{T-1} x_\tau' \bar{Q}_\tau x_\tau\} \tag{3.8.7}$$

subject to

$$x_{t+1} = \bar{A}_t x_t + \zeta_t \qquad t = 0,\ldots,T-1$$

where

$$\bar{A}_t = A_t - B_t^1 G_t^1 C_t^1 - B_t^2 G_t^2 C_t^2 \tag{3.8.9}$$

$$\bar{Q}_t = Q_t + C_t^{1'} G_t^{1'} R_t^1 G_t^1 C_t^1 + C_t^{2'} G_t^{2'} R_t^2 G_t^2 C_t^2. \tag{3.8.10}$$

Note that $\phi_t(x_t)$ as determined by equation (3.4.3) is a quadratic form as the following induction argument shows.

Clearly $\phi_T(x_T) = -x_T' Q_T x_T$ is quadratic. Assume that $\phi_{t+1}(x_{t+1})$ is quadratic, that is

$$\phi_{t+1}(x_{t+1}) = x_{t+1}' K_{t+1} x_{t+1} + k_{t+1} \tag{3.8.11}$$

then

$$\phi_t(x_t) = \int \phi_{t+1}(x_{t+1}) p_t(x_{t+1} | x_t; G_t^1, G_t^2) dx_{t+1} - x_t' \bar{Q}_t x_t \tag{3.8.12}$$

$$= E\{\phi_{t+1}(x_{t+1}) | x_t\} - x_t' \bar{Q}_t x_t \tag{3.8.13}$$

$$= E\{(\bar{A}_t x_t + \zeta_t)' K_{t+1} (\bar{A}_t x_t + \zeta_t)\} + k_{t+1}\} - x_t' \bar{Q}_t x_t \tag{3.8.14}$$

$$= x_t' [\bar{A}_t' K_{t+1} \bar{A}_t - \bar{Q}_t] x_t + tr\{K_{t+1} \Xi_t\} + k_{t+1} \tag{3.8.15}$$

$$= x_t' K_t x_t + k_t$$

where

$$K_t = \bar{A}_t' K_{t+1} \bar{A}_t - \bar{Q}_t \tag{3.8.17}$$

$$k_t = k_{t+1} + tr\{K_{t+1} \Xi_t\} \tag{3.8.18}$$

thus $\phi_t(x_t)$ is a quadratic form.

Now $H_t(\hat{\pi}_t, \hat{\gamma}_t, \hat{\phi}_{t+1})$ can be calculated as:

$$\int[\int \phi_{t+1}(x_{t+1})p(x_{t+1}|x_t;G_t^1,G_t^2)dx_{t+1} - x_t'\bar{Q}_t x_t]dx_t = \tag{3.8.19}$$

$$[x_t'(\bar{A}_t' K_{t+1}\bar{A}_t - \bar{Q}_t)x_t + k_t]\ p_t(x_t)dx_t = \tag{3.8.20}$$

$$\mathrm{tr}\ (\bar{A}_t' K_{t+1}\bar{A}_t - \bar{Q}_t)\ _t\ + k_t \tag{3.8.21}$$

where

$$\Sigma_{t+1} = A_t\Sigma_t A_t' + \Xi_t;\ \Sigma_o \text{ given .} \tag{3.8.22}$$

Let $\hat{K}_t$, $\hat{k}_t$ and $\hat{\Sigma}_t$ correspond to $\hat{\phi}_t$ and $\hat{\pi}_t$ so that equation (3.4.6) becomes

$$H_t(\hat{\pi}_t,(\hat{G}_t^1,\hat{G}_t^2),\hat{\phi}_{t+1}) = \max_{G_t^1,G_t^2}\ [\mathrm{tr}\{(\bar{A}_t^{-1}\hat{K}_{t+1}\bar{A}_t - \bar{Q}_t)\hat{\Sigma}_t\} + \hat{k}_t]. \tag{3.8.23}$$

We perform the maximization by setting the Fréchet derivative to zero. Recall that this derivative is given by (section A.7)

$$\frac{d}{d\epsilon} H_t(\hat{\pi}_t,(\hat{G}_t^1 + \epsilon\,\delta G_t^1,\ \hat{G}_t^2 + \epsilon\,\delta G_t^2),\ \hat{\phi}_{t+1})\Big|_{\epsilon=0} \tag{3.8.24}$$

Expanding (3.8.24) yields

$$\begin{aligned}&\mathrm{tr}\{[(A_t - B_t^1(\hat{G}_t^1 + \epsilon\,\delta G_t^1)C_t^1 - B_t^2(\hat{G}_t^2 + \epsilon\delta G_t^2)C_t^2)'\hat{K}_{t+1}\\ &\qquad (A_t - B_t^1(\hat{G}_t^1 + \epsilon\,\delta G_t^1)C_t^1 - B_t^2(\hat{G}_t^2 + \epsilon\delta G_t^2)C_t^2) -\\ &Q_t + C_t^{1'}(G_t^1 + \epsilon\,\delta G_t^1)'R_t^1(G_t^1 + \epsilon\,\delta G_t^1)C_t^1 + C_t^{2'}(G_t^2 + \epsilon\,\delta G_t^2)'R_t^2(G_t^2 + \epsilon\delta G_t^2)C_t^2]\hat{\Sigma}_t\}\\ &\qquad\qquad + k_t\end{aligned} \tag{3.8.25}$$

If we differentiate (3.8.25) with respect to ϵ, evaluate it at $\epsilon = 0$ and the set the result to zero then one has

$$\begin{aligned}\mathrm{tr}\{[&C_t^{1}{}'\delta G_t^{1}{}' R_t^1\hat{G}_t^1 C_t^1 + C_t^{1}{}'\hat{G}_t^{1}{}' R_t^1\delta G_t^1 C_t^1 + \\ &C_t^{2}{}'\delta G_t^{2}{}' R_t^2\hat{G}_t^2 C_t^2 + C_t^{2}{}'\hat{G}_t^{2}{}' R_t^2\delta G_t^2 C_t^2 + \\ &C_t^{1}{}'\delta G_t^{1}{}' B_t^1\hat{K}_{t+1}(A_t-B_t^1\hat{G}_t^1 C_t^1) + (A_t-B_t^1\hat{G}_t^1 C_t^1)'\hat{K}_{t+1}B_t^1\delta G_t^1 C_t^1 + \\ &C_t^{2}{}'\delta G_t^{2}{}' B_t^{2}{}'\hat{K}_{t+1}(A_t-B_t^2\hat{G}_t^2 C_t^2) + (A_t-B_t^2\hat{G}_t^2 C_t^2)'K_{t+1}B_t^2\delta G_t^2 C_t^2]\hat{\Sigma}_t\} = 0.\end{aligned} \tag{3.8.26}$$

Using the properties of the trace operator yields:

$$\begin{aligned}&2\ \mathrm{tr}\{\delta G_t^{1}{}'[R_t^1\hat{G}_t^1 C_t^1\hat{\Sigma}_t C_t^{1}{}' + B_t^{1}{}'\hat{K}_{t+1}(A_t-B_t^1\hat{G}_t^1 C_t^1)\hat{\Sigma}_t C_t^{1}{}']\} + \\ &2\ \mathrm{tr}\{\delta G_t^{2}{}'[R_t^2\hat{G}_t^2 C_t^2\hat{\Sigma}_t C_t^{2}{}' + B_t^{2}{}'\hat{K}_{t+1}(A_t-B_t^2\hat{G}_t^2 C_t^2)\hat{\Sigma}_t C_t^{2}{}']\} = 0\end{aligned} \tag{3.8.27}$$

which, since δG_t^1 and δG_t^2 are arbitrary, implies

$$R_t^1\hat{G}_t^1 C_t^1\hat{\Sigma}_t C_t^{1}{}' + B_t^{1}{}'\ \hat{K}_{t+1}(A_t-B_t^1\hat{G}_t^1 C_t^1)\hat{\Sigma}_t C_t^{1}{}' = 0 \tag{3.8.28}$$

and

$$R_t^2\hat{G}_t^2 C_t^2\hat{\Sigma}_t C_t^{2}{}' + B_t^{2}{}'\hat{K}_{t+1}(A_t-B_t^2\hat{G}_t^2 C_t^2)\hat{\Sigma}_t C_t^{2}{}' = 0 \tag{3.8.29}$$

If $C_t^1\hat{\Sigma}_t C_t^{1}{}'$, $C_t^2\hat{\Sigma}_t C_t^{2}{}'$, $R_t^1 - B_t^{1}{}'\hat{K}_{t+1}B_t^1$, and $R_t^2 - B_t^{2}{}'\hat{K}_{t+1}B_t^2$ are invertible, then

$$\hat{G}_t^1 = -(R_t^1-B_t^{1}{}'\hat{K}_{t+1}B_t^1)^{-1}B_t^{1}{}'\hat{K}_{t+1}A_t\hat{\Sigma}_t C_t^{1}{}'(C_t^1\hat{\Sigma}_t C_t^{1}{}')^{-1} \tag{3.8.30}$$

$$\hat{G}_t^2 = -(R_t^2-B_t^{2}{}'\hat{K}_{t+1}B_t^2)^{-1}B_t^{2}{}'\hat{K}_{t+1}A_t\hat{\Sigma}_t C_t^{2}{}'(C_t^2\hat{\Sigma}_t C_t^{2}{}')^{-1}\ . \tag{3.8.31}$$

Of course to obtain the values for $\hat{G}_t^1$ and $\hat{G}_t^2$ requires the solution of a nonlinear two point boundary value problem.

<u>Example 3.8.2</u> Consider the same problem as in Example 3.8.1 except the following additional constraints are introduced:

$$E\{d_t - x_t' D_t x_t\} \geq 0 \qquad t = 1,\ldots,T \tag{3.8.32}$$

It is easy to show that for this problem $\phi_t(x_t)$ also has a quadratic form. Clearly

$$\phi_T(x_T) = \lambda x_T' Q_T x_T + \mu_T(d_T - x_T' D_T x_T)$$
$$= x_T'(Q_T - \mu_T D_T) x_T + \mu_T d_T \qquad (3.8.33)$$

is quadratic (here μ_T corresponds to k_T^* in equation (3.6.21)). And if

$$\phi_{t+1}(x_{T=1}) = x_{t+1} K_{t+1} x_{t+1} + k_{t+1} \qquad (3.8.34)$$

then

$$\phi_t(x_t) = x_t'[\overline{A}_t' K_{t+1}\overline{A}_t + \lambda\overline{Q}_t]x_t + tr\{K_{t+1} \Xi_t\} + k_{t+1}$$
$$+ \mu_t(d_t - x_t' D_t x_t) \qquad (3.8.35)$$
$$= x_t'[\overline{A}_t' K_{t+1}\overline{A}_t + \lambda\overline{Q}_t - \mu D_t]x_t + tr\{K_{t+1} \Xi_t\}$$
$$+ k_{t+1} + \mu_t d_t \, ,$$

which is also quadratic. Equations (3.8.17) and (3.8.18) are thus replaced with

$$K_t = \overline{A}_t' K_{t+1}\overline{A}_t + \lambda\overline{Q}_t - \mu_t D_t \qquad (3.8.36)$$

$$k_t = k_{t+1} + tr\{K_{t+1} \Xi_t\} + \mu_t d_t \, . \qquad (3.8.37)$$

The equation for $H_t(\hat{\pi}_t, \gamma_t, \hat{\phi}^*_{t+1})$ is now given by:

$$H_t(\hat{\Sigma}_t, (G_t^1, G_t^2), (\hat{K}_{t+1}, \hat{k}_{t+1})) = tr\{(\overline{A}_t^{-1}\hat{K}_{t+1}\overline{A}_t + \lambda\overline{Q}_t - \mu D_t)\Sigma_t\} + k_t. \qquad (3.8.38)$$

Note that equation (3.6.23) holds with equality since $\Gamma_t = B(X_t, U_t)$ and thus leads to almost the same optimality condition as in the previous example since D_t is independent of G_t^1 and G_t^2. The only difference in fact is the presence of λ. Equations (3.8.30) and (3.8.31) become:

$$\hat{G}_t' = (\lambda R_t^1 + B_t^{1'}\hat{K}_{t+1}B_t^1)^{-1} B_t^{1'}\hat{K}_{t+1}A_t\hat{\Sigma}_t C_t^{1'}(C_t^1\hat{\Sigma}_t C_t^{1'})^{-1} \qquad (3.8.39)$$

$$\hat{G}_t^2 = (\lambda R_t^2 + B_t^{2'}\hat{K}_{t+1}B_t^2)^{-1}B_t^2\hat{K}_{t+1}A_t\hat{\Sigma}_t\, C_t^{2'}\,(C_t^2\Sigma_t C_t^2)^{-1}. \qquad (3.8.40)$$

Thus if $\lambda < 0$ the only difference between the solutions of these two problems in the equation for K_t. For the constrained version

$$K_t = \bar{A}_t^{-1}K_{t+1}\bar{A}_t - \bar{Q}_t - \mu_t D_t \qquad (3.8.41)$$

where

$$\mu_t = 0 \text{ if } E\{x_t' D_t x_t\} < d_t \text{ and}$$

$$\mu_t \geq 0 \text{ if } E\{x_t' D_t x_t\} = d_t \;.$$

<u>Example 3.8.3</u> Consider the same problem as in Example 3.8.1 except that we also require that $\hat{G}_t^1$ and $\hat{G}_t^2$ satisfy

$$\hat{G}_t^1 \in K_t^1 \qquad\qquad \hat{G}_t^2 \in K_t^2 \qquad (3.8.42)$$

where K_t^1 and K_t^2 are the cones of matrices with only non-negative entries. The associated conical approximation $C(\hat{G}_t,\Gamma_t)$ is $IC(\hat{G}_t^1, \text{int } K_t^1) \times IC(\hat{G}_t^2, \text{int } K_t^2)$, where $IC(\hat{G}_t^1, \text{int } K_t^1) = \{0\}$ if $\hat{G}_t^1 \notin \text{int } K_t^1$ and $IC(G_t^1, \text{int } K_t^1) = \text{int } K_t^1 \cup \{0\}$ if $\hat{G}_t^1 \in \text{int } K_t^1$ and similarly for $IC(\hat{G}_t^2, \text{int } K_t^2)$.

The derivation proceeds as in Example 3.8.1 up to equation (3.8.22). We now, however, apply equation (3.6.23) so that we have

$$\frac{d}{d\epsilon} H_t(\hat{\pi}_t, (\hat{G}_t^1 + \epsilon\delta G_t^1, \hat{G}_t^2 + \epsilon\delta G_t^2), \hat{\phi}_{t+1})\Big|_{\epsilon = 0} \leq 0 \qquad (3.8.43)$$

$$\forall\, (\delta G_t^1, \delta G_t^2) \in \overline{C(\hat{G}_t,\Gamma_t)} \;.$$

This implies that there exist elements k_t^{*1} and k_t^{*2} (not to be confused with k_t) in $(K_t^1)^*$ and $(K_t^2)^*$ such that

$$\mathrm{tr}\{\delta G_t^{1'}[-\lambda R_t^1\hat{G}_t^1C_t^1\hat{\Sigma}_tC_t^{1'} + B_t^{1'}\hat{K}_{t+1}(A_t-B_t^1\hat{G}_t^1C_t^1)\hat{\Sigma}_tC_t^{1'}] +$$

$$\mathrm{tr}\{\delta G_t^{2'}[-\lambda R_t^2\hat{G}_t^2C_t^2\hat{\Sigma}_tC_t^{2'} + B_t^{2'}\hat{K}_{t+1}(A_t-B_t^2\hat{G}_t^2C_t^2)\hat{\Sigma}_tC_t^{2'}] =$$

$$-(\langle k_t^*,\delta G_t^1\rangle + \langle k_t^*,\delta G_t^2\rangle). \tag{3.8.44}$$

Since G_t^1 and G_t^2 are matrices the space in which they exist is isomorphic to a finite Euclidean space, thus every linear functional on $G_t^{1,2}$ can be expressed as a weighted linear combination of the entries in $G_t^{1,2}$. Thus if δG_t^i is an m_t^i by n_t^i matrix, $i = 1,2$, $t = 0,\ldots,T$, and a_t^i is an n_t^i-vector, b_t^i is an n_t^i-vector and $\tilde{D}_t^i$ is an m_t^i by n_t^i matrix then

$$\begin{aligned}\langle k_t^{*i},\delta G_t^i\rangle &= a_t^{i'}\tilde{D}_t^i\delta G_t^i b_t^i \\ &= \mathrm{tr}\{a_t^{i'}\tilde{D}_t^i\delta G_t^i b_t^i\} \\ &= \mathrm{tr}\{\delta G_t^{i'}(b_t^i a_t^i\tilde{D}_t^i)'\} \triangleq \mathrm{tr}\{\delta G_t^{i'}\tilde{D}_t^i\}\end{aligned} \tag{3.8.45}$$

for some appropriate choice of a_t^i, b_t^i and $\tilde{D}_t^i$.

Thus (3.8.44) becomes

$$\begin{aligned}&\mathrm{tr}\{\delta G_t^{1'}[-\lambda R_t^1\hat{G}_t^1C_t^1\hat{\Sigma}_tC_t^{1'}+B_t^{1'}\hat{K}_{t+1}(A_t-B_t^1\hat{G}_t^1C_t^1)\hat{\Sigma}_tC_t^{1'}+D_t^1\} + \\ &\mathrm{tr}\{\delta G_t^{2'}[-\lambda R_t^2\hat{G}_t^2C_t^2\hat{\Sigma}_tC_t^{2'}+B_t^{2'}\hat{K}_{t+1}(A_t-B_t^2\hat{G}_t^2C_t^2)\hat{\Sigma}_tC_t^{2'}+D_t^2\} = 0\end{aligned} \tag{3.8.46}$$

which, under the invertibility assumptions of Example 3.8.1 implies, since δG_t^1 and δG_t^2 are arbitrary, that

$$\hat{G}_t^1 = (\lambda R_t^1+B_t^{1'}\hat{K}_{t+1}B_t^1)^{-1}(B_t^{1'}\hat{K}_{t+1}A_t\hat{\Sigma}_tC_t^{1'}+D_t^1)(C_t^1\hat{\Sigma}_tC_t^{1'})^{-1} \tag{3.8.47}$$

$$\hat{G}_t^2 = (\lambda R_t^2+B_t^{2'}\hat{K}_{t+1}B_t^2)^{-1}(B_t^{2'}\hat{K}_{t+1}A_t\hat{\Sigma}_tC_t^{2'}+D_t^2)(C_t^2\hat{\Sigma}_tC_t^{2'})^{-1} \tag{3.8.48}$$

where D_t^1 or D_t^2 are zero if $\hat{G}_t^1$ or $\hat{G}_t^2$ are in the interior of K_t^1 or K_t^2 respectively. If $\hat{G}_t^1$ or $\hat{G}_t^2$ are not in the interior of K_t^1 or K_t^2 then clearly D_t^1 or D_t^2 are in the set of rank one matrices with non-negative entries.

Section 3.9. Extension of Results to Vector Valued Criteria

The vector valued cost criterion can be a useful tool for formulating certain types of optimization problems [Pol 1]. For problems in which the various criteria do not have any units that allow a natural comparison (e.g. kilowatt-hours and pollution index) the use of a vector cost criterion provides an alternative to arbitrary amalgamation of the criteria. Of course a solution must eventually be chosen; the value of this approach is in problems where some selection scheme is available which is not easily formulated in a mathematical setting (e.g. voting or discussion among interested groups) [Zio 1].

The vector valued optimization procedure is thus a preprocessor which eliminates inferior solutions. While an interesting problem is how one can generate new noninferior points in response to a higher level decision process we shall consider the more basic problem of necessary conditions for noninferiority.

Recall that a point is noninferior with respect to a set of feasible points $\Omega \subset X$ and a cost cone $K \subset X$ if $\hat{x} \in \Omega$ and, for all $x \in \Omega$, $\hat{x} - x \in K$ implies that $\hat{x} = x$. We assume that the cone K is such that $K \neq X$.

An alternative definition is in terms of a set of continuous linear functionals K^*. A point $\hat{x}$ is noninferior if there does not exist an $x \in \Omega$ and a $k_1^* \in K^*$ such that $\langle k_1^*, x\rangle < \langle k_1^*, \hat{x}\rangle$ and, for all $k^* \in K^*$, $\langle k^*, x\rangle \leq \langle k^*, \hat{x}\rangle$. These two definitions are equivalent if K is the intersection of the dual cone of K^* ($K^{**} \in X^{**}$) and X.

In extending the results of Chapter III to vector criterion there are two alternatives. One is to reduce the dynamic vector optimization problem to a sequence of static vector optimization problems. The other is to reduce the dynamic vector optimization problem to a sequence of static scalar optimization problems. We shall use both approaches, the first in Theorem 3.9.1, and the second in Theorems 3.9.2 and 3.9.3. Consider the following problem:

$$\underset{\gamma \in \Gamma}{\text{optimize}} \; E\{\sum_{t=0}^{T} h_t(x_t;\gamma_t)\}; h_T(x_T;\gamma_T) \triangleq h_T(x_T) \tag{3.9.1}$$

subject to

$$x_{t+1} = f_t(x_t,\omega_t;\gamma_t) \qquad t = 0,\dots,T-1 \tag{3.9.2}$$

$$E\{g_t(x_t)\} \in K_t \qquad t = 1,\dots,T \tag{3.9.3}$$

$$E\{r_t(x_t)\} = 0 \qquad t = 1,\dots,T \tag{3.9.4}$$

where all functions and sets are as in section 3.1 except h_t: $X_t \times \Gamma_t \to \mathbb{R}^c$, $t = 0,\dots,T-1$, h_T: $X_T \to \mathbb{R}^c$ and optimality is defined with respect to K (the positive orthant of $\mathbb{R}^c$).

The equivalent deterministic problem is given by:

$$\underset{\gamma \in \Gamma}{\text{optimize}} \; \sum_{t=0}^{T} H_t(\gamma_t)\pi_t \tag{3.9.5}$$

subject to

$$\pi_{t+1} = P_t(\gamma_t) \; t \qquad t = 0,\dots,T-1 \tag{3.9.6}$$

$$G_t\pi_t \in K_t \qquad t = 1,\dots,T \tag{3.9.7}$$

$$R_t\pi_t = 0 \qquad t = 1,\dots,T \tag{3.9.8}$$

where $H_T(\gamma_T) \triangleq H_T$.

To derive the theorem corresponding to Theorem 3.4.1 we shall need to expand the notion of an adjoint operator. Let X, Y and Z be Banach spaces and let $T \varepsilon B(X,Y)$. Define an operator Φ mapping X into Z by the relation

$$\Phi x = B'Tx \tag{3.9.9}$$

where B' is a fixed element in B(Y,Z). Clearly Φ is a linear operator and since B' and T are bounded linear operators, so is Φ [Lu 1, Proposition 6.2.3]. But then Φ is continuous [Lu 1, Proposition 6.2.1] so $\Phi \varepsilon B(X,Z)$.

Now consider the relation $T^\dagger$ associating to each $B \varepsilon B(Y,Z)$ the corresponding $A_B \varepsilon B(X,Z)$, as just defined. We call $T^\dagger$ a <u>generalized</u> adjoint. Note that

$$A_B = T^\dagger B \tag{3.9.10}$$

where

$$A_B x = BTx \qquad x \varepsilon X, \tag{3.9.11}$$

so that

$$A_{\alpha B} x = \alpha BTx = \alpha A_B x \tag{3.9.12}$$

and

$$A_{B_1+B_2} = (B_1+B_2)Tx = A_{B_1} x + A_{B_2} x . \tag{3.9.13}$$

Thus $T^\dagger$ is a linear operator mapping elements of B(Y,Z) into elements in B(X,Z). Also note that

$$||T^\dagger Bx|| = ||BTx|| \le ||B|| \ ||T|| \ ||x|| \tag{3.9.14}$$

so that

$$||T^\dagger B|| \le ||T|| \ ||B|| \tag{3.9.15}$$

and

$$||T^\dagger|| \le ||T|| . \tag{3.9.16}$$

so

$$T^{\dagger} \in B(B(Y,Z),B(X,Z)) \qquad 3.9.17)$$

We can now state the equivalent of Lemma 3.4.1. Define the forced generalized adjoint equation (costate equation) by

$$\phi_t = P_t^{\dagger}(\gamma_t)\phi_{t+1} - H_t(\gamma_t); \phi_T = -H_T \qquad (3.9.18)$$

<u>Lemma 3.9.1</u> Let $\gamma = (\gamma_o, \ldots, \gamma_{T-1})$ be given and let the corresponding states and costates be given by equations (3.9.6) and (3.9.18) where π_o is assumed to be given. Then

$$\phi_t \pi_t = - \sum_{\tau=t}^{T} H_\tau(\gamma_\tau)\pi_\tau . \qquad (3.9.19)$$

<u>Proof</u>: By backward induction. Equation (3.9.19) obviously holds for $t = T$, so assume it holds for an arbitrary $t < T$. Then

$$\phi_{t-1}\pi_{t-1} = \langle P^{\dagger}_{t-1}(\gamma_{t-1})\phi_{t-1} - H_{t-1}(\gamma_{t-1}), \pi_{t-1}\rangle \qquad (3.9.20)$$

$$= \phi_t P_{t-1}(\gamma_{t-1})\pi_{t-1} - H_{t-1}(\gamma_{t-1})\pi_{t-1} \qquad (3.9.21)$$

$$= \phi_t \pi_t - H_{t-1}(\gamma_{t-1})\pi_{t-1} \qquad (3.9.22)$$

$$= \sum_{\tau=t}^{T} H_\tau(\gamma_\tau)\pi_\tau - H_{t-1}(\gamma_{t-1})\pi_{t-1} \qquad (3.9.23)$$

$$= \sum_{\tau=t-1}^{T} H_\tau(\gamma_\tau)\pi_\tau , \qquad (3.9.24)$$

so that equation (3.9.19) holds for $t = 0, \ldots, T$.

<u>Theorem 3.9.1</u> If $\hat{\gamma} = (\hat{\gamma}_o, \ldots, \hat{\gamma}_{T-1})$ is optimal in problem (3.9.5) for a given π_o and the corresponding states and costates are denoted $\hat{\pi}_t$ and $\hat{\phi}_t$ then, for all $\gamma_t \in \Gamma_t$, if

$$(P_t^{\dagger}(\hat{\gamma}_t)\hat{\phi}_{t+1} - H_{t+1}(\tilde{\gamma}_{t+1}))\hat{\pi}_t = (P_t^{\dagger}(\gamma_t)\hat{\phi}_{t+1} - H_{t+1}(\gamma_{t+1}))\hat{\pi}_t \in -K \quad (3.9.25)$$

then

$$(P_t^{\dagger}(\hat{\gamma}_t)\hat{\phi}_{t+1} - H_{t+1}(\hat{\gamma}_{t+1}))\hat{\pi}_t = (P_t^{\dagger}(\gamma_t)\hat{\phi}_{t+1} - H_{t+1}(\gamma_{t+1}))\hat{\pi}_t \qquad (3.9.26)$$

Proof: Let $\gamma' = (\hat{\gamma}_o,\ldots,\hat{\gamma}_{t-1},\gamma_t,\hat{\gamma}_{t+1},\ldots,\hat{\gamma}_{T-1})$ and let π_t' and ϕ_t' denote the corresponding states and costates. Note that $\hat{\pi}_\tau = \pi_\tau'$, $\tau = 0,\ldots,t$ and $\hat{\phi}_\tau = \phi_\tau'$, $\tau = t+1,\ldots,T$. Let $\hat{J}_T = -\hat{\phi}_o\hat{\pi}_o$ and $J_T' = -\phi_o'\pi_o'$. By the optimality of $\hat{\gamma}$ and equation (3.9.19) one has that, if

$$\hat{J}_T - J_T' \in K \tag{3.9.27}$$

then $\hat{J}_T = J_T'$. Thus if

$$-\sum_{\tau=t}^{T} H_\tau(\gamma_\tau)\hat{\pi}_\tau + \sum_{\tau=t}^{T} H_\tau(\hat{\gamma}_\tau)\hat{\pi}_\tau \in K \tag{3.9.28}$$

then

$$\sum_{\tau=t}^{T} H_\tau(\gamma_\tau)\hat{\pi}_\tau = \sum_{\tau=t}^{T} H_\tau(\hat{\gamma}_\tau)\hat{\pi}_\tau, \tag{3.9.29}$$

which, by equations (3.9.6), (3.9.19) implies equations (3.9.25) and (3.9.26).

Theorem 3.9.1 is useful if a method for solving static vector optimization problems is available (e.g. [Lin 1]). Otherwise it may be more useful to convert the static vector optimizations into static scalar optimizations. This can be done by deriving the analogues of Theorems 3.5.1 and 3.6.2 using Theorems 2.4.2 and 2.4.3. Since this derivation is almost identical to that of sections 3.5 and 3.6 we merely comment on the differences and state the analogous theorems without proof.

To derive the theorem corresponding to Theorem 3.5.1 we introduce the mapping

$$F_t(\gamma_t) = (H_t(\gamma_t), P_t(\gamma_t)) \tag{3.9.30}$$

$$F_T = H_T \,. \tag{3.9.31}$$

Let $B(M_t,K)$ be the set of linear operators in $B(\pi_t,\mathbb{R}^c)$ that map elements in M_t into elements in K. The appropriate assumption to make is

Assumption 3.9.1 The set $F_t(\Gamma_t)$ is $(-B(M_t,K),\underline{0})$-directionally convex, $t = 0,\ldots,T-1$.

This assumption implies that $F_t(\Gamma_t)\pi_t$ is $(-K,0)$-directionally convex for all $\pi_t \in M_t$. As before, the reason the assumption does not depend on π_t is the fact that the problem is linear in π_t.

Corresponding to Theorem 3.5.1 we have

Theorem 3.9.2 (Vector Maximum Principle) If $\hat{\gamma} = (\hat{\gamma}_0,\ldots,\hat{\gamma}_{T-1})$ is optimal in problem (3.9.6), Assumption 3.9.1 holds and $(\hat{\pi}_0,\ldots,\hat{\pi}_T)$ is the corresponding "state" trajectory then there exist costates $(\hat{\phi}_0^*,\ldots,\hat{\phi}_T^*)$, $\hat{\phi}_t^* \in \Pi_t^*$, vectors $k_t^* \in \tilde{K}_t^*$, vectors $\psi_t^* \in [\mathbb{R}^{e_t}]^*$, $t = 1,\ldots,T$, and an element λ^*, not all zero, such that, $-\lambda^* \in K^*$ and

$$\text{(i)} \quad \hat{\phi}_t^* = P_t^*(\hat{\gamma}_t)\hat{\phi}_{t+1}^* + H_t^*(\hat{\gamma}_t)\lambda^* + G_t^*k_t^* + R_t^*\psi_t^*, \quad t = T-1,\ldots,0, \qquad (3.9.32)$$

$$\text{(ii)} \quad \hat{\phi}_T^* = G_T^*k_T^* + R_T^*\psi_T^* + H_T^*\lambda^* \qquad (3.9.33)$$

$$\text{(iii)} \quad \langle k_t^*, G_t\hat{\pi}_t\rangle = 0, \quad t = 1,\ldots,T \qquad (3.9.34)$$

$$\text{(iv)} \quad H_t(\hat{\pi}_t,\hat{\gamma}_t,\hat{\phi}_{t+1}^*) \geq H_t(\hat{\pi}_t,\gamma_t,\hat{\phi}_{t+1}^*) \quad \forall \gamma_t \in \Gamma_t, \quad t = 0,\ldots,T-1 \qquad (3.9.35)$$

where

$$H_t(\hat{\pi}_t,\gamma_t,\hat{\phi}_{t+1}^*) = \langle\lambda^*, H_t(\gamma_t)\hat{\pi}_t\rangle + \langle\hat{\phi}_{t+1}^*, P_t(\gamma_t)\hat{\pi}_t\rangle. \qquad (3.9.36)$$

Proof: The same as that of Theorem 3.5.1 except that Corollary 2.4.1 is used in place of Corollary 2.1.4.

Since none of the comments and corollaries in section 3.5 depend on the cost function they also apply to Theorem 3.9.2. The remarks in section 3.7 on a representation of $\hat{\phi}_t^*$ also apply to Theorem 3.9.2 since $\lambda^*\hat{\phi}_t^*$ is the sum of bounded Borel measurable functions.

Corresponding to Assumption 3.6.1 we have

Assumption 3.9.2 The linear operators $H_t(\gamma_t)$, $t = 0,\ldots,T-1$ and $P_t(\gamma_t)$, $t = 0,\ldots,T-1$ are continuously Fréchet differentiable on Γ_t, $t = 0,\ldots,T-1$ respectively.

Corresponding to Theorem 3.6.1 we have

Theorem 3.9.3 (Vector QuasiMaximum Principle) If $\hat{\gamma} = (\hat{\gamma}_0,\ldots,\hat{\gamma}_{T-1})$ is optimal in problem (3.9.6), Assumptions 3.9.2 and 3.6.2 hold and $(\hat{\pi}_0,\ldots,\hat{\pi}_T)$ is the corresponding "state" trajectory, then there exist costates $(\hat{\phi}_0^*,\ldots,\hat{\phi}_T^*)$, $\hat{\phi}_t^* \in \Pi_t^*$, vectors $k_t^* \in \tilde{K}_t^*$, vectors $\psi_t^* \in [\mathbb{R}^{e_t}]^*$, $t = 1,\ldots,T$ an element $\lambda^* \in -K^*$, not all zero, and conical approximations, $C(\hat{\gamma}_t,\Gamma_t)$, to the sets Γ_t at $\hat{\gamma}_t$, $t = 0,\ldots,T-1$, such that

$$\text{(i)}\quad \hat{\phi}_t^* = P_t^*(\hat{\gamma}_t)\hat{\phi}_{t+1}^* + H_t^*(\hat{\gamma}_t)\lambda^* + G_t^*k_t^* + R_t^*\psi_t^*,\ t = T-1,\ldots,0 \tag{3.9.37}$$

$$\text{(ii)}\quad \hat{\phi}_T^* = G_T^*k_T^* + R_T^*\psi_T^* + H_T^*\lambda^* \tag{3.9.38}$$

$$\text{(iii)}\quad \langle k_t^*, G_t\hat{\pi}_t\rangle = 0,\ t = 0,\ldots,T \tag{3.9.39}$$

$$\text{(iv)}\quad \nabla_{\gamma_t}H_t(\hat{\pi}_t,\hat{\gamma}_t,\hat{\phi}_{t+1}^*)\delta\gamma_t \le 0 \qquad \forall\delta\gamma_t \in \overline{C(\hat{\gamma}_t,\Gamma_t)} \tag{3.9.40}$$

where $H_t(\hat{\pi}_t,\hat{\gamma}_t,\hat{\phi}_{t+1}^*)$ is given in equation (3.9.36).

Proof: The same as that of Theorem 3.6.1 except that Theorem 2.4.2 is used in place of Theorem 2.2.1.

Again, all the comments in section 3.6 apply to Theorem 3.9.3 and $\hat{\phi}_t^*$ can be represented as an integral functional with a Borel measurable kernel.

Finally, note that the assumptions that H_t: $X_t \times \Gamma_t \to \mathbb{R}^c$ and K is the positive orthant in $\mathbb{R}^c$ are inessential. An arbitrary Banach space Y can replace R^c and any convex cone K such that $K \neq Y$ and int $K \neq \phi$ can be used and Theorems 3.9.1 through 3.9.3 will still be correct. However it need not be the case that ϕ_t^* will have a representation as an integral functional with a Borel measurable kernel. (See the comments in section 3.7 regarding the extension of K_t to arbitrary Banach spaces).

Section 3.10. Concluding Remarks

The convexity assumption of section 3.5 is clearly very restrictive. If it is possible to find a subset of Γ_t for which it holds then if $\hat{z}$ is in that subset a necessary condition for optimality can be formulated in terms of that subset. Of course it is not easy to determine an appropriate subset.

Theorem 3.6.1 seems much more useful than Theorem 3.5.1 since the differentiability assumption is more likely to be true than the convexity assumption. It is interesting to note that differentiability or continuity of any function with respect to x was never assumed. This is due to the form of the equivalent deterministic problem which has as its only variables π_t and u_t. This justifies the comment in section 2.5 that it was not necessary to consider extending the theorems of Chapter II to problems with nondifferentiable constraints.

We have not considered problems with an unspecified horizon. Unlike the continuous time problem there are not boundary conditions that imply when a final time can be optimal. Obviously this occurs since a discrete-time problem is not differentiable with respect to time.

CHAPTER IV

An Infinite Horizon Problem Formulation

Section 4.0. Introduction

In the previous chapter we have considered the necessary conditions for optimal control of a constrained dynamical system over a finite time horizon. In this chapter we consider the same problem except that the time horizon is infinite. The reason for considering this extension is that it often provides insight into the control of a system over a long (but infinite) time horizon. In particular, if the optimal infinite time horizon control policy is stationary it may be desirable, for reasons of simplicity, to use it as a finite time horizon control policy.

In section 4.1 we will consider the basic infinite time horizon problem and its equivalent deterministic problem. This has been considered in detail by Bertsekas [Bert 1, Chapter 9] so we will merely present the results that we will need. In section 4.1 necessary conditions for the optimal policy for the infinite time horizon problem will also be developed. Because of the explicit state constraints the optimal policy is rarely stationary. In section 4.2 we modify the problem by introducing a new cost criterion. We then extend results for finite or countably infinite state spaces to our continuous state formulation. Finally we develop an equivalent deterministic problem.

In section 4.3 the EDP is modified by introducing state constraints. Necessary conditions for an optimal stationary policy are then derived. Finally in section 4.4 we make some concluding remarks.

Section 4.1. An Infinite Horizon Problem

Consider the following problem:

$$\min_{\gamma \in \Gamma} E\{ \sum_{t=0}^{\infty} h_t(x_t;\gamma_t) \} \tag{4.1.1}$$

subject to

$$x_{t+1} = f_t(x_t,w_t;\gamma_t) \qquad t = 0,\ldots \tag{4.1.2}$$

$$E\{g_t(x_t)\} \in K_t \qquad t = 1,\ldots \tag{4.1.3}$$

$$E\{r_t(x_t)\} = 0 \qquad t = 1,\ldots \tag{4.1.4}$$

where h_t, f_t, g_t and r_t are defined as in section 3.1. To make the problem well-defined and to allow us to derive a meaningful equivalent deterministic problem we must place some restrictions on $h_t(x_t;\gamma_t)$.

Following Bertsekas [Bert 1] we consider the following three cases:

(i) $0 \le h_t(x_t;\gamma_t)$ for all $x_t \in X_t$ and $\gamma_t \in \Gamma_t$ (4.1.5)

(ii) $0 \ge h_t(x_t;\gamma_t)$ for all $x_t \in X_t$ and $\gamma_t \in \Gamma_t$ (4.1.6)

(iii) $h_t(x_t;\gamma_t) = \alpha^t \tilde{h}_t(x_t;\gamma_t)$ where $|\tilde{h}_t(x_t;\gamma_t)| \le b < \infty$ (4.1.7)
for all $x_t \in X_t$ and $\gamma_t \in \Gamma_t$ and $\alpha \in (0,1)$.

For these three cases it can easily be shown that $E\{ \sum_{t=0}^{\infty} h_t(x_t;\Gamma_t)\}$ is well-defined (that is, $\liminf_{T \to \infty} E\{ \sum_{t=0} h_t(x_t;\gamma_t)\}$ and $\limsup_{T \to \infty} E\{ \sum_{t=0} h_t(x_t;\gamma_t)\}$ are equal, though they may both be $\pm \infty$). It can also be shown that to each policy $\gamma = (\gamma_o,\gamma_1,\ldots)$ there corresponds a unique sequence of probability measures $\{\pi_t(\gamma,\pi_o)\}$ (as before π_o is assumed to be given) such that

$$E\{ \sum_{t=0}^{\infty} h_t(x_t;\gamma_t)\} = \sum_{t=0}^{\infty} \int h_t(x_t;\gamma_t)\pi_t(\gamma,\pi_o)(dx_t) \tag{4.1.8}$$

$$E\{g_t(x_t)\} = \int g_t(x_t)\pi_t(\gamma,\pi_o)(dx_t) \tag{4.1.9}$$

and

$$E\{r_t(x_t)\} = \int r_t(x_t)\pi_t(\gamma,\pi_o)(dx_t) \tag{4.1.10}$$

As before we shall usually suppress the arguments and write π_t for $\pi_t(\gamma,\pi_o)$. We make the same assumptions on $P_t(\gamma_t)$, $h_t^*(\gamma_t)$, G_t and R_t as before.

The equivalent deterministic problem can thus be written as:

$$\min_{\gamma\epsilon\Gamma} \sum_{t=0}^{\infty} \langle h_t^*(\gamma_t),\pi_t\rangle \tag{4.1.11}$$

subject to

$$\pi_{t+1} = P_t(\gamma_t)\pi_t \qquad t = 0,\ldots \tag{4.1.12}$$

$$G_t\pi_t \; \epsilon \; K_t \qquad t = 1,\ldots \tag{4.1.13}$$

$$R_t\pi_t = 0 \qquad t = 1,\ldots \tag{4.1.14}$$

The probability measures π_t, $t = 0,\ldots$ are again considered as elements of the Banach space Π_t of signed measures. Note that the product space $\Pi = \Pi_x \times \Pi_1 \times\ldots$ is also a Banach space if we take the norm on Π to be

$$||\pi|| = ||(\pi_o,\pi_1,\ldots)|| = \sup_{j\,\epsilon\,(0,\ldots)} ||\pi_j|| \tag{4.1.15}$$

where $||\pi_j||$ is the total variation norm on π_j.

The development of a maximum principle and a quasi maximum principle now proceeds as in sections 3.5 and 3.6. The only difference is that the equality constraint

$$Rz = [R_1\pi_1,R_2\pi_2,\ldots] = 0 \tag{4.1.16}$$

is included in the constraint $T(z) = 0$. We do this since $\mathcal{R}(R)$ is infinite dimensional. Clearly the same assumptions as in Chapter III lead to the same results (except for the terminal constraints, equations (3.5.34) and (3.6.21)).

Theorem 4.1.1 (Maximum Principle) If $\hat{\gamma} = (\hat{\gamma}_o, \hat{\gamma}_1, \ldots)$ is optimal in problem (4.1.11), Assumption 3.5.1 holds for $t = 0,1,\ldots$ and $(\hat{\pi}_o, \hat{\pi}_1, \ldots)$ is the corresponding "state" trajectory, then there exist costates $(\hat{\phi}_o^*, \hat{\phi}_1^*, \ldots)$, $\hat{\phi}_t^* \in \Pi_t^*$, $t = 0,1,\ldots$, vectors $k_t^* \in \tilde{K}_t^*$, $\psi_t^* \in [R^{e_t}]^*$, $t = 1,\ldots$ and a scalar $\lambda \leq 0$ not all zero, such that

$$\text{(i)} \quad \hat{\phi}_t^* = P_t^*(\hat{\gamma}_t)\hat{\phi}_{t+1}^* + \lambda h_t^*(\hat{\gamma}_t) + G_t^* k_t^* + R_t^* \psi_t^*, \; t = 0,1,\ldots \qquad (4.1.17)$$

$$\text{(ii)} \quad \langle k_t^*, G_t^* \hat{\pi}_t \rangle = 0, \; t = 1,\ldots \qquad (4.1.18)$$

$$\text{(iii)} \quad H_t(\hat{\pi}_t, \hat{\gamma}_t, \hat{\phi}_{t+1}^*) \geq H_t(\hat{\pi}_t, \gamma_t, \hat{\phi}_{t+1}^*) \;\; \forall \gamma_t \in \Gamma_t, \; t = 0,1,\ldots \qquad (4.1.19)$$

where

$$H_t(\hat{\pi}_t, \gamma_t, \hat{\phi}_{t+1}^*) = \lambda \langle h_t^*(\gamma_t), \hat{\pi}_t \rangle + \langle \hat{\phi}_{t+1}^*, P_t(\gamma_t)\hat{\pi}_t \rangle. \qquad (4.1.20)$$

Proof: see proof of Theorem 3.5.1.

Theorem 4.1.2 (Quasi Maximum Principle) If $\hat{\gamma} = (\hat{\gamma}_o, \hat{\gamma}_1, \ldots)$ is optimal in problem (4.1.11), Assumptions 3.6.1 and 3.6.2 hold for $t = 0,1,\ldots$, and $(\hat{\pi}_o, \hat{\pi}_1, \ldots)$ is the corresponding "state" trajectory, then there exist costates $(\hat{\phi}_o^*, \hat{\phi}_1^*, \ldots)$, $\hat{\phi}_t^* \in \Pi_t^*$, $t = 0,1,\ldots$, vectors $k_t^* \in \tilde{K}_t^*$, $\phi_t^* \in [R^{e_t}]^*$, $t = 1,\ldots$, and a scalar $\lambda \leq 0$, not all zero, and conical approximations, $C(\hat{\gamma}_t, \Gamma_t)$, to the set Γ_t at $\hat{\gamma}_t$, $t = 0,1,\ldots$, such that:

$$\text{(i)} \quad \hat{\phi}_t^* = P_t^*(\hat{\gamma}_t)\hat{\phi}_{t+1} + \lambda h_t^*(\hat{\gamma}_t) + G_t^* k_t^* + R_t^* \psi_t^*, \; t = 0,1,\ldots \qquad (4.1.21)$$

$$\text{(ii)} \quad \langle k_t^*, G_t^* \hat{\pi}_t \rangle = 0, \; t = 1,\ldots \qquad (4.1.22)$$

$$\text{(iii)} \quad \nabla\gamma_t H_t(\hat{\pi}_t, \hat{\gamma}_t, \hat{\phi}_{t+1}^*)\delta\gamma_t \leq 0 \;\; \forall \delta\gamma_t \in \overline{C(\hat{\gamma}_t, \Gamma_t)} \qquad (4.1.23)$$

where $H_t(\hat{\pi}_t, \hat{\gamma}_t, \hat{\phi}_{t+1}^*)$ is given by equation (4.1.20).

Unfortunately neither of these theorems is very useful for developing a stationary optimal policy. In fact it is clear that in general no feasible stationary policy exists even if all the mappings f_t, h_t, r_t and g_t are time invariant. The reason for this is clear: the state constraints may require a time varying control policy even if the cost-to-go does not.

Section 4.2. An EDP for an Unconstrainted Problem

In this section we consider necessary conditions for an optimal policy to be stationary. Since the formulation of section 4.1 does not usually allow a feasible stationary policy we will consider a different problem in this section. The unconstrained version of the problem is:

$$\min_{\gamma\varepsilon\Gamma} \lim_{T\to\infty} \frac{1}{T+1} E\{\sum_{t=0}^{T} h(x_t;y)\} \tag{4.2.1}$$

subject to

$$x_{t+1} = f(x_t,w_t;\gamma),\ t = 0,1,... \tag{4.2.2}$$

where h, f and Γ are time invarient, $(w_o,w_1,...)$ are independent identically distributed random variables and the policy is $(\gamma,\gamma,...)$.

We consider this problem formulation because it leads to an equivalent deterministic problem which is easily extended to problems with state constraints. Unfortunately this formulation has several technical problems associated with the existence of certain limits. Before considering these problems note that if f and h satisfy the measurability requirements of section 3.1 then there exists a stochastic transition kernal $p(x|y;\gamma)$ which maps probability measures on X into probability measures on X. For each $\gamma \varepsilon \Gamma$ and π_o there corresponds a sequence of probability measures $\{\pi_t\}$ such that

$$\lim_{T\to\infty} \frac{1}{T+1} E\{\sum_{t=0}^{T} h(x_t;\gamma)\} = \lim_{T\to\infty} \frac{1}{T+1} \sum_{t=0}^{T} \int h(x;\gamma)\pi_t(dx) \tag{4.2.3}$$

Again we can write (4.2.3) and (4.2.2) in functional operation by letting π_t be an element of Π, the space of signed measures on X, by assuming $h^*(\gamma)$ is an element of Π^* for all $\gamma \varepsilon \Gamma$ and by letting $P(\gamma)$ denote a linear operator in $B(X,X)$ such that

$$\int h(x;\gamma)\pi_t(dx) = \langle h^*(\gamma),\pi_t\rangle \tag{4.2.4}$$

$$\int p(x|y;\gamma)\pi_t(dy) = P(\gamma)\pi_t . \tag{4.2.5}$$

The problem can thus be written

$$\min_{\gamma\in\Gamma} \lim_{T\to\infty} \frac{1}{T+1} \sum_{t=0}^{T} \langle h^*(\gamma),\pi_t\rangle \tag{4.2.6}$$

subject to

$$\pi_{t+1} = P(\gamma)\pi_t \tag{4.2.7}$$

where π_o is given.

Before deriving another equivalent deterministic problem let us consider some properties of problem (4.2.6). These will be extensions of the results in [Kus 1], [Bert 2], [Ros 1] and [Ito 1]. The following definitions are required.

A finite invariant measure for $P(\gamma)$ is defined as any measure $\pi(\gamma) \varepsilon \Pi$ such that

$$\pi(\gamma) = P(\gamma)\pi(\gamma) \tag{4.2.8}$$

$$\int \pi(\gamma)(dx) < \infty \tag{4.2.9}$$

The n-step stochastic transition kernal $p^{(n)}(x|y;\gamma)$ is defined by

$$p^{(n)}(x|y;\gamma) = \int p(x|z;\gamma)p^{(n-1)}(dz|y;\gamma) \tag{4.2.10}$$

$$p^{(0)}(x|y;\gamma) = p(x|y;\gamma). \tag{4.2.11}$$

Let the corresponding linear operator be denoted $P^{(n)}(\gamma)$ and let

$$P^T(\gamma) = \frac{1}{T+1} \sum_{t=0}^{T} P^{(t)}(\gamma). \tag{4.2.12}$$

Recall that $\beta(X)$ is the set of Borel sets of X. Let m be any measure defined on $\beta(X)$ and let $L^P(X,\beta(X),m)$, $1 \le p<\infty$, be defined

[Yos 1, page 32] as the space of Borel measurable functions f such that

$$\int |f(x)|^p m(dx) < \infty. \tag{4.2.13}$$

Let

$$\| f\|_p = [\int |f(x)|^p m(dx)]^{1/p} \tag{4.2.14}$$

and consider the space $L_p(X,\beta(X),m)$ of equivalence classes in $L^p(X,\beta(X),m)$ defined by the equivalence relation

$$f_1 \equiv f_2 \text{ if } \int|f_1(x) - f_2(x)|^p m(dx) = 0 .$$

This space is a Banach space.

Let X_p^* denote the space of continuous linear functionals that can be represented by an integral functional with kernal in $L_p(X,\beta(X),m)$. Note that X_1^* is a subspace of $\mathbb{I}^*$.

<u>Assumption 4.2.1</u> An invariant probability measure $\pi(\gamma)$ exists for $P(\gamma)$ for each $\gamma \in \Gamma$.

By [Yos 1, Theorem 13.1.1] the adjoint operator for $P(\gamma)$, denoted $P(\overset{*}{\gamma})$, can be defined so that $P^*(\gamma) \in B(X_2^*,X_2^*)$, where $m = \pi(\gamma)$. Since $L^2(X,\beta(X),\pi(\gamma))$ is reflexive [Lu 1, Section 5.6], it is a locally sequentially weakly compact space [Yos 1, Theorem 5.2.1] and thus the corollary to [Yos 1, Theorem 8.3.1] can be applied. Thus for each $f^* \in X_2^*$ there exists a $\tilde{f}^* \in X_2^*$ such that

(i) $$\lim_{T\to\infty} \|P^T(\overset{*}{\gamma})f^* - f^*\|_2 = 0 , \tag{4.2.15}$$

if $P^{\infty}(\gamma)^*$ is defined by $P^{\infty}(\overset{*}{\gamma})f^* = f^*$ then $P^{\infty}(\overset{*}{\gamma}) \in B(X_2^*,X_2^*)$ and

(ii) $$P^{\infty *}(\gamma) = P^{\infty *}(\gamma) = P^*(\gamma)P^{\infty *}(\gamma) = P^{\infty *}(\gamma)P^*(\gamma) \tag{4.2.16}$$

(iii) $$R(P^{\infty *}(\gamma)) = N(I - P^*(\gamma)) \tag{4.2.17}$$

(iv) $N(P^{\infty *}(\gamma)) = \overline{R(I - P^*(\gamma))} = R(I - P^{\infty *}(\gamma))$, and

(v) $$X_2 = \overline{R(I - P^*(\gamma))} + N(I - P^*(\gamma)). \tag{4.2.19}$$

Lemma 4.2.1 If Assumption 4.2.1 holds and $\mathcal{R}(I - P^*(\gamma))$ is closed then the linear operator $I - P^*(\gamma) + P^{\infty *}(\gamma)$ is invertible.

Proof: By equations (4.2.17) and (4.2.19)

$$\mathcal{R}(I - P^*(\gamma) + P^{\infty}(\gamma) = X_2 \ . \tag{4.2.20}$$

By [Yos 1, Theorem 8.3.1], $\overline{\mathcal{R}(I - P^*(\gamma))} \cap \mathcal{N}(I - P^*(\gamma)) = \{0\}$ so $\mathcal{R}(I - P^*(\gamma)) \cap \mathcal{N}(I - P^*(\gamma)) = \{0\}$. Thus by equation (4.2.17) $\mathcal{R}(I - P^*(\gamma)) \cap \mathcal{N}(P^{\infty *}(\gamma) = \{0\}$, that is, if x^* is such that

$$(I - P^*(\gamma))x^* = -P^{\infty}(\gamma)^* x^* \tag{4.2.21}$$

then

$$(I - P^*(\gamma))x^* = P^{\infty *}(\gamma)x^* = 0 \ . \tag{4.2.22}$$

Equivalently $x^* \in \mathcal{N}(I - P^*(\gamma))$ and $x^* \in \mathcal{N}(P^{\infty}(\gamma) = \mathcal{R}(I - P^*(\gamma))$; Equation (4.2.19) now implies $x^* = 0$. This, however, implies $\mathcal{N}(I - P^*(\gamma) + P^{\infty}(\overset{*}{\gamma})) = \{0\}$ since

$$\mathcal{N}(I - P^*(\gamma) + P^{\infty *}(\gamma)) = \{x^* | (I - P^*(\gamma))x^* = -P^{\infty *}(\gamma)x^*\}. \tag{4.2.23}$$

Clearly $I - P^*(\gamma) + P^{\infty *}(\gamma)$ is one-to-one and thus is is invertible.

Lemma 4.2.2 If the assumptions of Lemma 4.2.1 hold and $h^*(\gamma) \in X_2^*$ then there exists a $\phi^*(\gamma) \in X_2^*$ such that

$$J^*(\gamma) + \phi^*(\gamma) = h^*(\gamma) + P^*(\gamma)\phi^*(\gamma) \tag{4.2.24}$$

where $J^*(\gamma)$ is defined by

$$J^*(\gamma) = P^{\infty *}(\gamma)h^*(\gamma) . \tag{4.2.25}$$

Proof: (following [Bert 1]).

From equation (4.2.15) $J^*(\gamma)$ exists. By Lemma 4.2.1 we can define

$$\phi^*(\gamma) = (I - P^*(\gamma) + P^{\infty *}(\gamma))^{-1}(I - P^{\infty *}(\gamma)h^*(\gamma) \tag{4.2.26}$$

which clearly exists in X_2^*. Now equation (4.2.26) implies, by equation (4.2.16), that

$$P^{\infty *}(\gamma)h^*(\gamma) = 0. \qquad (4.2.27)$$

Thus (4.2.25) and (4.2.26) imply

$$J^*(\gamma) + \phi^*(\gamma) = h^*(\gamma) + P^*(\gamma)\phi^*(\gamma) . \qquad (4.2.28)$$

These results can be extended to the class of functions in which we are really interested: $L^1(X,\beta(X),\pi(\gamma))$. By [Yos 1, Theorem 13.1.1] $P^*(\gamma)$ can be defined so that $P^*(\gamma) \varepsilon B(X_1^*,X_1^*)$, where $m = \pi(\gamma)$. Because X_2^* is X_1^*-dense in X_1^* (that is, for every $\epsilon > 0$ and for every $f^* \varepsilon X_1^*$, there exists an $\tilde{f}^* \varepsilon X_1^* \quad X_2^*$ such that $||f^* - \tilde{f}^*||_1 < \epsilon$) [Yos 1, Theorem 13.2.2] and since $P^*(\gamma)$ is such that

$$||P^*(\gamma)f^*||_p \leq ||f^*||_p \quad \forall f^* \varepsilon X_p^*, \; p = 1,2 \qquad (4.2.29)$$

[Yos 1, Theorem 13.1.1] we have

<u>Lemma 4.2.3</u> If Assumption 4.2.1 holds then for each $f^* \varepsilon X_1^*$ there exists an $\tilde{f}^* \varepsilon X_1$ such that

$$\lim_{T\to\infty} || P^{T*}(\gamma)f^* - \tilde{f}^*||_1 = 0, \qquad (4.2.30)$$

and if $P^{\infty *}(\gamma)$ is defined by $P^{\infty *}(\gamma)f^* = \tilde{f}^*$ then $P^{\infty *}(\gamma) \quad \varepsilon B(X_1^*,X_1^*)$. If $R(I - P^*(\gamma))$ is closed and if $h^*(\gamma) \varepsilon X_1^*$ for all $\gamma \varepsilon \Gamma$, then there exists a $\phi^*(\gamma) \varepsilon X_1^*$ such that

$$J^*(\gamma) + \phi^*(\gamma) = h^*(\gamma) + P^*(\gamma)\phi^*(\gamma), \qquad (4.2.31)$$

where

$$J^*(\gamma) = P^{\infty *}(\gamma)h^*(\gamma). \qquad (4.2.32)$$

<u>Proof</u>: If $f^* \varepsilon X_2^*$ then $f^* \varepsilon X_1^*$ since by Schwartz' inequality $||f^*||_1 \leq ||f^*||_2 \cdot [\int_X \pi(\gamma)(dx)]^{1/2} = ||f^*||_2$. Thus (4.2.30) holds if $f^* \varepsilon X_1^* \cap X_2^*$. Since X_2 is X_1-dense in X_1, for any $\epsilon > 0$ and for any $f^* \varepsilon X_1$ there exists an $f_\epsilon^* \varepsilon X_1 \cap X_2$ such that $||f_\epsilon^* - f^*||_1 < \epsilon$. But by (4.2.29)

$$||P^{T*}(\gamma)f_\epsilon^* - P^{T*}(\gamma)f^* \,||_1 \le ||f_\epsilon^* - f^*||_1 \le \epsilon \qquad (4.2.33)$$

and since (4.2.30) holds for f_ϵ^* it holds for f^*. The corollary to [Yos 1, Theorem 8.3.2] can now be applied to yield the analogues to (4.2.15) through (4.2.19). The results corresponding to Lemmas 4.2.1 and 4.2.2 clearly hold and imply equations (4.2.31) and (4.2.32).

These results can be extended to yield a pointwise convergence theorem (note that this is not now the case since $P^\infty(\gamma)^*$ is defined implicitly by equation (4.2.30)).

Theorem 4.2.1 If Assumption 4.2.1 holds, $\mathcal{R}(I - P^*(\gamma))$ is closed and $h(\gamma) \in L_1(X,\beta(X),\pi(\gamma))$ for all $\gamma \in \Gamma$, then there exists a $\phi(\gamma) \in L_1(X,\beta(\gamma),\pi(\gamma))$ such that

$$J(\gamma)(x) + \phi(\gamma)(x) = h(\gamma)(x) + \int\phi(y)p(dy|x;\gamma) \quad \pi(\gamma)\text{-a.e.} \qquad (4.2.34)$$

where

$$J(\gamma)(x) = \lim_{T\to\infty} \frac{1}{T+1} \sum_{t=0}^{T} \int h(y)p^{(t)}(dy|x;\gamma) \qquad (4.2.35)$$

and the limit in (4.2.33) exists and is finite $\pi(\gamma)$-a.e.

Proof: From Lemma 4.2.3

$$\lim_{T\to\infty} ||P^{T*}(\gamma)h^*(\gamma) + \phi^*(\gamma) - h^*(\gamma) - P^*(\gamma)\phi^*(\gamma)||_1 = 0 \qquad (4.2.36)$$

and by [Yos 1, Theorem 13.3.5] and the definition of X_p^* equation (4.2.33) follows.

Note that Theorem 4.2.1 extends similar results in [Bert 1] and [Kus 1] which were derived for finite and countable state spaces. Note that for a finite state space, Assumption 4.2.1 always holds and $\pi(\gamma)$-a.e. convergence reduces to the usual notion of pointwise convergence.

We shall not extend this analysis further. The reason is that to do so would require examining the conditions under which $J^*(\gamma)\pi(\gamma) = J^*(\gamma)\pi_o$. While this can be done in a satisfactory manner if π_o is absolutely continuous [Hal 5, page 124] with respect to $\pi(\gamma)$ (see Halmos [Hal 4] and Ito [Ito 1] for details) it is often the case that $\pi(\gamma)$ is not absolutely continuous with respect to π_o, especially if π_o is degenerate. Thus we shall make the following assumption.

Assumption 4.2.2 Let $\pi^T(\gamma) = \frac{1}{T+1} \sum_{t=0}^{T} \pi_t$ where the π_t are generated by equation (4.2.7). Assume that $\pi^T(\gamma)$ converges weakly (in the sense of [Bil 1]) to an invariant probability measure $\pi(\gamma)$, that is

$$\lim_{T\to\infty} \int h(x)\pi^T(\gamma)(dx) = \int h(x)\pi(\gamma)(dx) \tag{4.2.37}$$

for all bounded uniformly continuous functions $h(x)$.

Note that $\pi(\gamma)$ is independent of π_o thus $\pi(\gamma)$ is a unique invariant probability measure. This implies [Do 1, page 214] that only one ergodic class exists. Note that we have not assumed that π_t converges to $\pi(\gamma)$. The convergence assumption is a Cesaro convergence assumption on $\pi^t(\gamma)$, thus we allow cyclically moving sets [Do 1, page 211].

Lemma 4.2.4 If $\pi(\gamma)$ is the unique invariant probability measure for $P(\gamma)$ then $\pi(\gamma)$ is the unique solution to

$$\pi(\gamma) = P(\gamma)\pi(\gamma) \tag{4.2.38}$$

$$\int \pi(\gamma)(dx) = 1 \, . \tag{4.2.39}$$

Proof: If $\pi(\gamma)$ is not the unique solution then let $\tilde{\pi}(\gamma)$ be another solution. Clearly $\tilde{\pi}(\gamma)(x)$ is not greater than or equal to zero for all x since then $\tilde{\pi}(\gamma)$ would be an invariant probability measure, and (4.2.39) implies that $\tilde{\pi}(\gamma)(x)$ is not less than zero for all x. Thus the Jordan decomposition of $\tilde{\pi}(\gamma)$ [Hal 5, page 123] is nontrivial. That is, there exist sets Ω_+ and Ω_- in X such that $\Omega_+ \cap \Omega_- = \phi$, positive measures $\tilde{\pi}(\gamma)_+$ and $\tilde{\pi}(\gamma)_-$ such that

$$\tilde{\pi}(\gamma) = \tilde{\pi}(\gamma)_+ - \tilde{\pi}(\gamma)_- \tag{4.2.40}$$

$$\tilde{\pi}(\gamma)_+(A) = 0 \text{ if } A \cap \Omega_+ = \phi \tag{4.2.41}$$

$$\tilde{\pi}(\gamma)_-(A) = 0 \text{ if } A \cap \Omega_- = \phi. \tag{4.2.42}$$

Let $\int\tilde{\pi}(\gamma)_+(dx) = m_+$ and $\int\tilde{\pi}(\gamma)_-(dx) = m_-$. Since $\tilde{\pi}(\gamma)$ is totally finite so are $\tilde{\pi}(\gamma)_+$ and $\tilde{\pi}(\gamma)_-$ that is, m_+ and m_- are finite. Clearly $P(\gamma)$ is such that

$$\int_{\Omega_+} P(\gamma)\tilde{\pi}(\gamma)_+(dx) = \int_X P(\gamma)\tilde{\pi}(\gamma)_+(dx) = m_+ \tag{4.2.43}$$

$$\int_{\Omega_-} P(\gamma)\tilde{\pi}(\gamma)_-(dx) = \int_X P(\gamma)\tilde{\pi}(\gamma)_-(dx) = m_- \tag{4.2.44}$$

and, if we let

$$\hat{\pi}(\gamma)_+ = P(\gamma)\tilde{\pi}(\gamma)_+ \tag{4.2.45}$$

$$\hat{\pi}(\gamma)_- = P(\gamma)\tilde{\pi}(\gamma)_- \tag{4.2.46}$$

then $\hat{\pi}(\gamma)_+(x)$ and $\hat{\pi}(\gamma)_-(x)$ are greater than or equal to zero for all x and $\hat{\pi}(\gamma)_+ - \hat{\pi}(\gamma)_- = \tilde{\pi}(\gamma)_+ - \tilde{\pi}(\gamma)_-$.

Pick $A \subset \Omega_+$, then

$$\hat{\pi}(\gamma)_+(A) - \hat{\pi}(\gamma)_-(A) = \tilde{\pi}(\gamma)_+(A). \tag{4.2.47}$$

If $\hat{\pi}(\gamma)_+(A) < \tilde{\pi}(\gamma)_+(A)$ then (4.2.46) cannot hold since $\hat{\pi}(\gamma)_-(A) \geq 0$. If $\hat{\pi}(\gamma)_+(A) > \tilde{\pi}(\gamma)_+(A)$ then $\hat{\pi}(\gamma)_-(A) > 0$. But then $A \cap \Omega_- = \phi$ contradicts (4.2.43), thus $\hat{\pi}(\gamma)_+(A) = \tilde{\pi}(\gamma)_+(A)$ for all Borel sets $A \subset \Omega_+$. A similar argument shown the same holds for $\hat{\pi}(\gamma)_-$ and $\hat{\pi}(\gamma)_-$. But this implies the existence of two finite invariant positive measures for $P(\gamma)$

different from $\pi(\gamma)$. Thus $\tilde{\pi}(\gamma)_+$ and $\tilde{\pi}(\gamma)_-$ can be normalized so that they are invariant probability measures. This contradicts the uniqueness of $\pi(\gamma)$ and thus the lemma is proved.

This lemma is useful since it allows us to uniquely describe the invariant probability measure without using inequality constraints to ensure positivity.

<u>Assumption 4.2.3</u> For each $\gamma \varepsilon \Gamma$, $h(x;\gamma)$ is a uniformly continuous bounded function on X.

Problem 4.2.6 can now clearly be written as

$$\min_{\gamma\varepsilon\Gamma} \langle h^*(\gamma), \pi(\gamma)\rangle \tag{4.2.48}$$

subject to

$$\pi(\gamma) = P(\gamma)\pi(\gamma) \tag{4.2.49}$$

$$\int \pi(\gamma)(dx) = 1. \tag{4.2.50}$$

Section 4.3. A Constrained EDP

In this section we will introduce a variation on the state constraints of section 4.1. We shall require $\hat{\gamma} \varepsilon \Gamma$ be such that

$$\lim_{T\to\infty} \int r(x)\pi^T(\gamma)(dx) = 0 \tag{4.3.1}$$

and

$$\lim_{T\to\infty} \int g(x)\pi^T(\gamma)(dx) \; \varepsilon \; K \tag{4.3.2}$$

where K is a closed convex cone with nonempty interior. Constraints (4.3.1) and (4.3.2) can be considered as "asymptotic constraints", that is, they act to constrain the steady state behavior of the system.

<u>Assumption 4.3.1</u> The functions r: $X \to \mathbb{R}^e$ and g: $X \to \mathbb{R}^k$ are uniformly continuous and bounded.

The constrained equivalent deterministic problem can now be stated

$$\min_{\gamma\varepsilon\Gamma} \langle h^*(\gamma), \pi(\gamma)\rangle \quad (4.3.3)$$

subject to

$$\pi(\gamma) = P(\gamma)\pi(\gamma) \quad (4.3.4)$$

$$G\pi(\gamma) \ \varepsilon\ K \quad (4.3.5)$$

$$R\pi(\gamma) = 0 \quad (4.3.6)$$

$$\mathbb{1}^*\pi(\gamma) = 1, \quad (4.3.7)$$

where

$$\mathbb{1}^*\pi(\gamma) = \int \pi(\gamma)(dx). \quad (4.3.8)$$

The EDP is now formulated in a manner in which the theorems of Chapter II can be applied to yield results paralleling those in Chapter III. The following assumption is needed.

Assumption 4.3.2 $\mathcal{R}(I - P(\gamma))$ is closed for all $\gamma \ \varepsilon\ \Gamma$.

Since the null space of $I - P(\gamma)$ is not empty Assumption 4.3.2 implies that $\mathcal{R}(T)$ is closed where

$$T = \begin{bmatrix} \mathbb{1}^* \\ I - P(\gamma) \end{bmatrix} \quad (4.3.9)$$

Define $F(\Gamma)$ by

$$F(\Gamma) = \{(\pi^*, P)\ \varepsilon\ \Pi^* \times B(\Pi,\Pi) \mid (\pi^*, P) = (h^*(\gamma), P(\gamma)), \gamma\varepsilon\Gamma\} \quad (4.3.8)$$

Assumption 4.3.3 $F(\Gamma)$ is $(-M^*, \underline{0})$-directionally convex, where, as in section 3.5, M^* is the polar cone to M, the set of probability measures.

As usual $\tilde{K}$ is defined as

$$\tilde{K} = \{x \mid x = \beta(y - G\hat{\pi}(\hat{\gamma}), y\ \varepsilon\ K, \beta > 0\}. \quad (4.3.9)$$

Theorem 4.3.1 (Steady State Maximum Principle)

If Assumptions 4.2.2, 4.2.3, 4.3.1 and 4.3.3 hold and $\hat{\gamma}$ is optimal in problem (4.3.3) then there exist a $\hat{\phi}^* \in \Pi^*$, $k^* \in \tilde{K}^*$ and $\psi^* \in [\mathbb{R}^e]^*$ and scalars λ, and c, not all zero, such that $\lambda \leq 0$ and

$$\text{(i)} \quad \hat{\phi}^* = P^*(\hat{\gamma})\hat{\phi}^* + \lambda h^*(\hat{\gamma}) + G^*k^* + R^*\psi^* + c^* \tag{4.3.10}$$

$$\text{(ii)} \quad <k^*, G\pi(\hat{\gamma})> = 0 \tag{4.3.11}$$

$$\text{(iii)} \quad H(\pi(\hat{\gamma}), \hat{\gamma}, \hat{\phi}^*> = \max_{\gamma \in \Gamma} H(\pi(\hat{\gamma}), \gamma, \hat{\phi}^*) \tag{4.3.12}$$

where

$$H(\pi(\hat{\gamma}), \gamma, \hat{\phi}^* = <P^*(\gamma)\hat{\phi}^* + \lambda h^*(\gamma), \pi(\hat{\gamma})> \tag{4.3.13}$$

Proof: Straightforward modification of the proof of Theorem 3.5.1.

Corollary 4.3.1 In Theorem 4.3.1, $c = -<\lambda h(\hat{\gamma}), \pi(\hat{\gamma})>$.

Proof: From (4.3.10) we have

$$<\hat{\phi}^*, \pi(\hat{\gamma})> = <P^*(\hat{\gamma})\hat{\phi}^*, \pi(\hat{\gamma})> + \lambda<h(\hat{\gamma}), \pi(\hat{\gamma})> + <G^*k^*, \pi(\hat{\gamma})>$$
$$+ c<\underline{\pi}^*, \pi(\hat{\gamma})> + <R^*\psi^*, \pi(\hat{\gamma})>. \tag{4.3.14}$$

But (4.3.11) implies $<G^*k^*, \pi(\hat{\gamma})> = 0$ and the optimality of $\pi(\hat{\gamma})$ and the linearity of (4.3.6) implies $<R^*\psi^*, \pi(\hat{\gamma})> = 0$. Since $\pi(\hat{\gamma})$ is an invariant measure for $P(\hat{\gamma})$ we also have

$$<P^*(\hat{\gamma})\hat{\phi}^*, \pi(\hat{\gamma})> = <\hat{\phi}^*, P(\hat{\gamma})\pi(\hat{\gamma})> = <\hat{\phi}^*, \pi(\hat{\gamma})>. \tag{4.3.15}$$

Thus (4.3.14) becomes, since $<\mathbb{1}^*, \pi(\hat{\gamma})> = \int \pi(\hat{\gamma})(dx) = 1$,

$$0 = \lambda<h^*(\hat{\gamma}), \pi(\hat{\gamma})> + c \tag{4.3.16}$$

which implies the corollary.

The theorem corresponding to Theorem 3.6.1 is easily derived. The appropriate assumptions are:

Assumption 4.3.4 The set Γ lies in a Banach space G and satisfies one of the following conditions:

(i) Γ contains a convex set A such that int $A \neq \phi$, or

(ii) Γ contains a subset B such that if $\delta\gamma \in B$ then there exists an $\epsilon^* > 0$ and an $\alpha^* > 0$ such that:

$$\hat{\gamma} + \epsilon(\delta\gamma+\delta\zeta) \in \Gamma \quad \forall \epsilon \in [0,\epsilon^*], \forall \delta\zeta \in S(0,\alpha^*) \tag{4.3.17}$$

Assumption 4.3.5 The functional $h^*(\gamma)$ and the linear operator $P(\gamma)$ are continuously Fréchet differentiable on Γ.

Theorem 4.3.2 (Steady State Quasi Maximum Principle) If Assumptions 4.2.2, 4.2.3, 4.3.1, 4.3.2, 4.3.4 and 4.3.5 hold and $\hat{\gamma}$ is optimal in problem (4.3.3) then there exist a $\hat{\phi}^* \in \Pi^*$, $k^* \in \tilde{K}^*$ and $\psi^* \in [\mathbb{R}^e]^*$ and scalars λ and c not all zero and a conical approximation; $C(\hat{\gamma},\Gamma)$, to the set Γ at $\hat{\gamma}$, such that $\lambda \leq 0$ and

$$\text{(i)} \quad \hat{\phi}^* = P^*(\hat{\gamma})\hat{\phi}^* + \lambda h^*(\hat{\gamma}) + G^*k^* + R^*\psi^* + c^* \tag{4.3.18}$$

$$\text{(ii)} \quad \langle k^*, G\pi(\hat{\gamma})\rangle = 0 \tag{4.3.19}$$

$$\text{(iii)} \quad \nabla_\gamma H(\pi(\hat{\gamma}),\hat{\gamma},\hat{\phi}^*)\delta\gamma \leq 0 \quad \forall \delta\gamma \in \overline{C(\hat{\gamma},\Gamma)} \tag{4.3.20}$$

where

$$H(\pi(\hat{\gamma}),\hat{\gamma},\hat{\phi}^*) = \langle P^*(\hat{\gamma})\hat{\phi}^* + \lambda h^*(\hat{\gamma}), \pi(\hat{\gamma})\rangle. \tag{4.3.21}$$

Proof: Straightforward modification of the proof of Theorem 3.6.1.

Of course Corollary 4.3.1 also applies to Theorem 4.3.2 as do many of the corollaries of Chapter III. Since these are quite obvious we do not consider them here.

Section 4.4.

The introduction of the cost criterion which represents the average cost per stage in section 4.2 brought with it several difficulties but it allowed us to introduce the "asymptotic constraints". These constraints allowed us to derive necessary conditions for optimality of a stationary policy. This would not have been possible if the exponentially weighted cost $(h_t^*(\gamma) = \alpha^t h^*(\gamma))$ had been used since the cost cannot be described solely in terms of the invariant probability measure.

The value of "asymptotic constraints" does not lie in having many problems fit that formulation (since there aren't many); rather it would seem to lie in the insight that can be gained into steady state control laws.

Note that in this problem formulation we must assume that the range of $I - P(\gamma)$ is closed. In Chapter III we were able to show that the range was closed in the finite horizon problem. The reason for this is that we have lost one degree of freedom in transforming the "state" equation to steady state form. As we noted in Chapter II this does not seem to be an unreasonable assumption.

The boundedness assumption made on the set and constraint functions is not restrictive on physical grounds. Certainly infinite values would be unreasonable physically. For mathematical simplicity however unbounded costs are sometimes used (for example, in LQG problems) and these cannot be handled in the formulation of section 4.3. These problems can, however, often be modified so that these assumptions are valid or they can be handled by other methods (for example the LQG problem is handled by a wide variety of techniques [San 3], [Ath 1], etc.).

CHAPTER V

Summary, Conclusions and Suggestions for Future Research

Section 5.0. Introduction

This chapter summarizes the results of this thesis and considers some of the conclusions that can be drawn from this research. Some suggestions of areas for future research are made.

Section 5.1. Summary

In Chapter I we consider some of the problems that can arise in large scale or nonclassical problems. These considerations lead naturally to the study of constrained systems which is the problem formulation that we use in the remainder of the thesis. A survey of related topics and results that will be used later is then given. An overview of the remaining chapters closes Chapter I.

Chapter II is concerned with nonlinear mathematical programming in Banach spaces. The results in this chapter are extensions and/or simplifications of existing results. We also extend these results to a problem formulation which includes a nonscalar objective function.

In Chapter III the mathematical formulation of the problem of Chapter I is introduced. An equivalent deterministic problem formulation is given and shown to be general enough to subsume several types of problems. The results of Chapter II are then used to prove a maximum and a quasimaximum principle. These theorems are extended to problems with nonscalar objective functions and are compared to the results of dynamic programming. Consideration is also given to different assumptions under which these theorems hold. These results are then applied to three different constrained LQG problems.

In Chapter IV the infinite time horizon problem is introduced and maximum and quasi maximum principles are derived. It is seen that the optimal control law for these problems is rarely stationary so a new problem formulation, known as the steady state constrained problem formulation, is introduced. Several results are derived for this problem which are extensions of results known for finite and countable state spaces. Finally, under an erogicity assumption, a maximum and a quasi maximum principle are derived for the steady state constrained problem formulation.

Section 5.2. Conclusions

The fundamental problem with the necessary conditions that we have derived in this thesis is the restrictive conditions that have been required for them to hold. In particular the directional convexity assumption required for the maximum principle is not easily satisfied (even for LQG problems!). The differentiability assumption required for the quasi maximum principle, while not as restrictive as directional convexity, is not satisfied for some interesting problems. These problems are, of course, common to all applications of maximum principles to discrete time systems.

Another drawback to the formulation of this thesis is that the results are necessary conditions for optimality and thus are not as strong as those obtained via dynamic programming. However this is often not a serious difference since the constrained optimization performed in each DP iteration is often done by solving necessary conditions for optimality.

The necessary conditions are satisfied for locally optimal solutions and thus clearly need not be sufficient. This can be seen in problems for which a signalling strategy ([San 1], [Ho 1]) is optimal: the necessary conditions are satisfied for any strategy which cannot be improved by a change made in one time instant. Since the signalling strategy requires cooperation between control strategies at various time instances there will clearly be many locally optimal solutions to the necessary conditions.

Often, however, we are uninterested in signalling strategies and the above mentioned problem does not arise. This particularly true in cases where a class of control strategies of interest are selected a priori.

The major benefit of the problem formulation of this thesis is that control and state constraints are handled in a natural and explicit manner. Thus one would expect that insight into the optimal control strategy could be obtained in a broader variety of problems than when dynamic programming is used. This is especially true when constraints on the complexity of the control law exist, since dynamic programming cannot be applied in an intuitive manner.

While we have not considered algorithms for using the necessary conditions to obtain a solution, the maximum principle formulation has some advantages over the dynamic programming formulation. This is particularly true if an initial feasible solution is known since successively better solutions can then be obtained by using the gradient of the Hamiltonian. Thus a recursive algorithm can be formulated using the maximum principle.

If dynamic programming can be applied to the basic problem then a one point boundary value problem results as opposed to the two point boundary value problem that results from using the results of this thesis. Thus our results are most useful when dynamic programming cannot be applied to the basic problem.

Section 5.3. Suggestions for Future Research

There are many different areas in which further work is required. We shall make some comments in three areas.

A major difficulty in the straightforward application of existing algorithms for the solution of necessary conditions to this problem is that it is posed in an infinite dimensional space. Two possible techniques for circumventing this problem are discretizing the state (the probability measures) and parameterizing the space of measures. Both techniques work by reducing an infinite dimensional problem to an approximate finite dimensional problem.

An alternative approach is to note that we do not need to evaluate the π_t's, we only need to know $\langle h_t^*(\gamma), \pi_t \rangle$, $G_t \pi_t$, etc. It is possible to use Monte Carlo simulations to evaluate these quantities for any given γ. An iterative procedure can then be developed to minimize the Hamiltonian. Two different approaches seem reasonable: (i) for the finite horizon problem, multiple simulations of the system over the whole time horizon could be used to evaluate the expected costs [Qua 1], (ii) for the infinite horizon problem, under stationarity and ergodicity assumptions, the cost could be evaluated by one long simulation.

Another interesting area concerns extending the class of constraints that are allowed in our problem formulation. In this thesis we have allowed constraints on γ_t and constraints on π_t; an interesting extension is to consider constraints on γ_t and π_t together (e.g. $E\{g(u_t)\} \in K_t$). Some results have been obtained in this area [Lea 1] but only under very restrictive assumptions and in the finite dimensional case.

A third area of interest is that of using the structure of the problem to obtain further results. In particular, we are interested in the following questions:

(i) if we consider probabalistic (mixed) control strategies does this allow us to prove directional convexity?

(ii) is there an analogue to the bang-bang principle?

(iii) is there a compromise between the assumption of directional convexity and differentiability that will allow the derivation of a maximum principle?

APPENDIX A

MATHEMATICAL PRELIMINARIES

A.1. Linear Topological Vector Spaces

A collection S of subsets of a set X forms a topology for X if:

(i) $\phi \in X,\ X \in S$ (A.1)

(ii) $X_1 \cap X_2 \in S$ if $X_1, X_2 \in S$ (A.2)

(iii) $\bigcup_{i=1}^{\infty} X_i \in S$ if $X_i \in S,\ i = 1,2,\ldots$. (A.3)

The pair (X,S) is called a topological space.

A collection F of subsets of X is said to be a basis of the topology S if

(i) $X_o \in F \Rightarrow X_o \in S$ (A.4)

(ii) $X_o \in S \Rightarrow\ X_i \in F\ \ \bigcup_{i=1}^{\infty} X_i = X_o$. (A.5)

Thus, if a basis F is given, the topology consists of all sets that are unions of sets in F.

The sets of S are open sets in X. Any open set that contains a point $x \in X$ is a neighborhood of x. If A is a set in X then $a \in A$ is an interior point if there exists a neighborhood of a contained in A.

A set B in X is closed if its complement $B\backslash X$ is open. The closure of B is the intersection of all the closed sets containing B.

A sequence $\{x_n\}$ of points in X converges to x if, for any neighborhood of x, there exists an N such that all x_n, $n \geq N$ are in the neighborhood.

We will be interested only in locally convex Hausdorff topological spaces, that is, topological spaces in which there exists a basis of the topology consisting of convex sets and in which for any two distinct points there exist nonintersecting neighborhoods of those points.

If X_1 and X_2 are topological spaces with topologies S_1 and S_2 then the topological product of X_1 and X_2 is denoted $X_1 \times X_2$. The topological space $X_1 \times X_2$ consists of all pairs (x_1,x_2) where $x_1 \varepsilon X_1$ and $x_2 \varepsilon X_2$. The basis for the topology of $X_1 \times X_2$ is defined as all sets of the form:

$$\{(x_1,x_2) | x_1 \varepsilon A_1, x_2 \varepsilon A_2, A_1 \varepsilon S_1, A_2 \varepsilon S_2\}. \quad (A.6)$$

It is known that the topological product of locally convex Hausdorff spaces is a locally convex Hausdorff space.

Now consider a <u>mapping</u> f: $X_1 \rightarrow X_2$ where (X_1,S_1) and (X_2,S_2) are topological spaces. The mapping f is <u>continuous at x_o</u> if, for any open set $A_2 \subset X_2$ with $f(x_o) \varepsilon A_2$, there exists a neighborhood A_1 of x_o such that $f(A_1) \subset A_2$. A mapping f is <u>continuous</u> if A is continuous at every point $x \varepsilon X_1$.

If X is a real linear space then (X,S) is a <u>real linear topological space</u> if S is such that $f(x,y) = x+y$: $X \times X \rightarrow X$ and $g(\lambda,x) = \lambda x: E^1 \times X \rightarrow X$ are continuous, where E^1 is the space of real numbers with topology defined by the basis of sets of the form $\{\lambda | \; |\lambda-\lambda_o| < r\}$.

If a linear space is <u>normed</u> then a topology can be defined via a basis consisting of all sets of the form

$$S(x_o,r) = \{x | \; \|x-x_o\| < r\} , \quad (A.7)$$

where $\|x\|$ is the norm on the space. This topology is the <u>natural</u> topology for a normed linear space. In a normed linear space with the natural topology a sequence $\{x_n\}$ converges if and only if

$$\lim_{n \rightarrow \infty} \|x_n-x\| \rightarrow 0 . \quad (A.8)$$

A sequence $\{x_n\}$ in a normed linear space† is a <u>Cauchy</u> sequence if,

† the natural topology is assumed unless otherwise specified.

for any $\epsilon > 0$, there exists a number N such that $||x_n - x_m|| < \epsilon$ for all $n,m \geq N$. Clearly if $\{x_n\}$ converges to x then $\{x_n\}$ is a Cauchy sequence, however, the converse is not generally true. A normed linear space X is complete if every Cauchy sequence converges to an element in X. A complete normed space is a Banach space.

Note that if B_1 and B_2 are Banach spaces with norms $||\cdot||_1$ and $||\cdot||_2$ then the produce topology induced by the metric $||(x_1,x_2)|| = ||x_1||_1 + ||x_2||_2$ makes $B_1 \times B_2$ a Banach space.

A.2 Properties of Normed Linear Spaces

The dimension of a linear space X is defined as equal to the maximum number of linearly independent elements in X. We shall see that the properties of finite and infinite dimensional spaces have some interesting differences, but first, some more definitions.

A set $A_o \subset X$ is compact if, for every sequence of open sets $\{A_i\}$ with the property $\bigcup_{i=1}^{\infty} A_i \supset A_o$, there exists a finite subsequence $A_{i_1}, \ldots, A_{i_n}$ such that $\bigcup_{j=1}^{n} A_{i_j} \supset A_o$.

A set $A_o \subset X$ is sequentially compact if every sequence of points in A_o, $\{a_i\}$ has a convergent sequence.

Theorem A.2.1 In a normed linear space a set is sequentially compact if and only if it is compact.

Proof: see Dunford and Schwartz [1958] pg. 22.

Let us define, for a normed linear space X, the closed unit sphere $\overline{S}(0,1) = \{x | x \in X, ||x|| \leq 1\}$.

Theorem A.2.2 The closed unit sphere $\overline{S}(0,1)$ is compact if and only if X is finite dimensional

Proof: see Dunford and Schwartz [1958] pg. 245.

Thus in an infinite dimensional space a closed and bounded set need not be compact. Compactness is, therefore, a much more restrictive requirement in infinite dimensional spaces than in finite dimensional ones. This is important because, for a large class of functionals, a maximum is attained on a compact set.

An upper semicontinuous functional is a mapping $f: X \to \mathbb{R}$ such that, for every $\epsilon > 0$ and x_o, there exists a $\delta(x_o,\epsilon) > 0$ such that $f(x) - f(x_o) < \epsilon$ for all x such that $||x-x_o|| < \delta(x_o,\epsilon)$.

Theorem A.2.3 An upper semicontinuous functional on a compact subset K of a normed linear space X achieves a maximum on K.

Proof: see Luenberger [1969] pg. 40.

A set $K \subset X$ is dense in X if the closure of K, $\bar{K}$, is equal to X.

Theorem A.2.4 There exist two complimentary dense sets in X if and only if X is infinite dimensional.

Proof: see [Kle 1] .

These sets will have to be excluded in a later section on separating hyperplanes.

A set A in a linear space X is convex if $x_1, x_2 \in A$ imply that $\lambda x_1 + (1-\lambda)x_2 \in A$ for all $\lambda \in [0,1]$.

The convex hull of a set A is the smallest convex set that contains A and it is denoted coA. If $A \subset X$ a linear topological space then the closed convex hull of A, $\overline{co}A$, is the set of points

$$\{x | x = \sum_{i=1}^{n} \lambda_i x_i,\ x_i \in X,\ \sum_{i=1}^{n} \lambda_i = 1,\ \lambda_i \geq 0,\ n \geq 1\} . \qquad (A.9)$$

A convex cone C is a set such that if $x \in C$ then $\lambda x \in C$ for all $\lambda > 0$ and if $x,y \in C$ then $x + y \in C$.

Convex cones will be used in two ways in this thesis. First, they will be used as approximations to certain sets that arise in optimization problems. This is important since we can only separate convex sets. Second, they will be used to induce a partial order on a space. This partial ordering generalizes the notion of inequality associated with scalars.

This partial ordering property will be used in two different ways. If K is the positive orthant in $\mathbb{R}^n$ then $g(x) \in K$ is equivalent to $g^i(x) \geq 0$ $i = 1,\ldots,n$. Thus if K is an arbitrary cone in a Banach space Y, $g(x) \in K$ is a generalization of the inequality constraint.

A cone K can also be used to induce a notion of optimality. If $\Omega \subset X$ is a set of feasible points in X and K is a cone in X then $\hat{x} \in \Omega$ is noninferior on Ω with respect to K if for all $x \in \Omega$, $\hat{x} - x \in K$ implies $\hat{x} = x$. If K is the positive orthant of $\mathbb{R}^n$ this just requires that there exist no feasible point $x \neq \hat{x}$ less than or equal to $\hat{x}$ in each component.

<u>Example A.2.1</u> A convex cone without any interior points, dense in the whole space.

Let $X = \ell_2$ and define C_+ as the set of finitely nonzero sequences for which the last nonzero element is positive, that is,

$$C_+ = \{x \mid x = (\alpha_1,\ldots,\alpha_n,0,0,\ldots),\alpha_n > 0,\ n \geq 1\}.$$

Furthermore, if $C_- = \{x \mid -x \in C_+\} \cup \{0\}$, then $C_+ \cup C_- = \ell_2$, and $\overline{C}_+ = \overline{C}_- = \ell_2$. Clearly C_+ and C_- are convex cones without interior points.

<u>Example A.2.2</u> A closed convex cone without any interior points, not contained in a proper subspace.

Let $X = \ell_2$ and define P_+ as the positive orthant, that is

$$P_+ = \{x | x = (\alpha_1, \ldots, \alpha_n, \ldots.), \ \alpha_n \geq 0, \ n = 1,2,\ldots\} \ .$$

The set P_+ is clearly closed. That it has no interior points can be verified by noting that for any $\epsilon > 0$ and $x' = (\alpha_1', \alpha_2', \ldots) \in P_+$ there is an $x \notin P_+$ such that $||x'-x|| < \epsilon$. This occurs because $\lim_{n\to\infty} |x_n| = 0$.

A.3. Conjugate Spaces

Let B denote a normed linear space and let $\mathbb{R}$ denote the real numbers with the Euclidean topology. A <u>functional</u> f is a mapping $f\colon \ B \to \mathbb{R}$. A functional is continuous at x_o if and only if, for every $\epsilon > 0$, there exists a $\delta > 0$ such that

$$|f(x) - f(x_o)| < \epsilon \quad x \quad ||x-x_o|| < \delta \ .$$

A linear functional is continuous everywhere if it is continuous at any point. It can also be shown that a linear functional is bounded if any only if it is continuous. A functional f is bounded if there exists a constant M such that

$$|f(x)| \leq M \, ||x|| \qquad \forall x. \tag{A.10}$$

Let B* denote the space of all continuous linear functionals defined on B. B* is conjugate (or dual) to B. The space B* is a linear space since the operations of addition of functionals and scalar multiplication can be defined:

$$(f_1 + f_2)(x) = f_1(x) + f_2(x) \tag{A.11}$$

$$(\lambda f_1)(x) = \lambda f_1(x) \ . \tag{A.12}$$

Henceforth we denote arbitrary functionals in B* by x*, y*, etc., and we denote the real number resulting from evaluating x* at x by <x*,x>.

A topology for B* can be defined in various ways. The <u>strong topology</u> of B* is generated by the basis consisting of all sets of the form

$$V(X,r,x_o^*) = \{x^* \mid \sup_{x \varepsilon X} |\langle x^*,x\rangle - \langle x_o^*,x\rangle| < r\}$$

where X is any bounded subset of B, that is, a set such that, for any neighborhood N of the origin, there exists a $\lambda < \infty$ such that $X \subset \lambda N$. If X is a normed linear space then this topology can also be generated as the natural topology corresponding to the norm

$$\| x^* \| = \sup_{\|x\| \leq 1} \langle x^*,x\rangle. \tag{A.13}$$

Another interesting topology for B* is generated by basis of the form

$$W(X,r,x_o^*) = \{x^* \mid |\langle x^*,x\rangle - \langle x_o^*,x\rangle| < r, \quad x \varepsilon X\}$$

where X is any subset of B which contains a finite number of elements $x_i \varepsilon B$. This topology is known as the <u>weak* topology</u>. It is interesting to note that the weak* topology is not induced by any metric unless B* is finite dimensional.

This can be seen by recalling from Theorem 2.2.2 that the closed unit sphere is compact in a normal space if and only if the space is finite dimensional and noting that

<u>Theorem A.3.1</u> The closed unit sphere in the conjugate space B* of the Banach space B is compact in the weak* topology.

<u>Proof:</u> see Dunford and Schwartz [1958] pg. 424.

The basis defined by $V(X,r,x_o^*)$ and $W(X,r,x_o^*)$ are identical in finite dimension spaces; thus the weak* and strong topologies are identical in finite dimensional spaces.

It can be shown that, if B is a normed linear space, then B* is complete, and thus a Banach space in its strong topology but B* is a Banach space in its weak* topology if and only if B* is finite dimensional. Note also if $B = B_1 \times B_2$ then $B^* = B_1^* \times B_2^*$.

Since we have defined two different topologies on B^* the terms open, closed, convergent, compact, etc. have two different meanings. If no topology is specified then the strong topology is assumed. Note that since every open set in the weak* topology is open in the strong topology, strong convergence implies weak* convergence and strong compactness implies weak* compactness, however, since a set is closed if its complement is open, sets that are weak* closed are strongly closed, but not necessarily conversely.

If K is a convex cone in a Banach space B, then we can define the <u>dual</u> (or <u>conjugate</u>) <u>cone</u> K^* in B^*:

$$K^* = \{x^* | x^* \in B^*, \langle x^*, x\rangle \geq 0 \ \forall x \in K\}. \tag{A.14}$$

It is easy to verify that K^* is a convex cone in B^* and that, if $\overline{K} \neq B$, K^* contains nonzero elements. In fact we can show that.

<u>Theorem A.3.2</u> If K is a convex cone, then the cone K^* is weak* closed.

<u>Proof:</u> Pshenichnyi [1971], pg. 34.

<u>Example A.3.1</u> A set closed in the strong topology but not in the weak* topology.

Let $S = \{x^* | \ \|x^*\| = 1\}$. S is clearly closed in the strong topology, however it can be shown (see Bourbaki [Bo 1]) that S is not closed in the weak* topology.

<u>A.4. Separation Theorems and Support Points</u>

One of the more important theorems from functional analysis (at least with respect to optimization problems) is the Hahn-Banach Theorem.

<u>Theorem A.4.1</u> Let K_1 and K_2 be convex sets in a real normed

linear vector space X such that K_1 has interior points and K_2 int $K_1 \neq \phi$. Then there is a closed hyperplane H separating K_1 and K_2; that is, there exists an $x^* \varepsilon X^*$, $x^* \neq 0$, such that

$$\sup_{x \varepsilon K_1} \langle x^*,x\rangle \leq \inf_{x \varepsilon K_2} \langle x^*,x\rangle \qquad \text{(A.15)}$$

Proof: see Dunford and Schwartz [1958] , pg. 412.

Corollary A.4.1 If X is a finite dimensional space then the requirement of an internal point can be removed. Two convex sets are separable if

(i) their union is contained in a subspace, or

(ii) the intersection of their relative interiors is empty.

Proof: see Canon, Cullum and Polak [1970] pg. 247.

The requirement of an interior point in infinite dimensional spaces can be eliminated if $K_1 \cup K_2 \subset M$, a subspace such that $\dim M < \dim X$.

If X is a locally convex linear topological space then we can use the following theorem to get stronger results.

Theorem A.4.2 If K_1 and K_2 are closed convex subsets in a locally convex linear topological space X and K_1 is compact, then $K_1 + K_2$ is closed and convex.

Proof: Dunford and Schwartz [1958], pg. 414.

Theorem A.4.3 If K_1 and K_2 are disjoint closed convex subsets in a locally convex linear topological space X and if K_1 is compact, then there exist constants c and ϵ, $\epsilon > 0$ and a continuous linear functional $x^* \varepsilon X^*$, $x^* \neq 0$, such that

$$\langle x^*,x\rangle \leq c-\epsilon < c \leq \langle x^*,y\rangle \quad \forall x \varepsilon K_2 \quad \forall y \varepsilon K_1 \qquad \text{(A.16)}$$

Proof: [Du 1, Theorem 5.2.7.10]. We can separate K_1 and K_2 if and only if we can separate K_1-K_2 and $\{0\}$. By theorem 2.4.2 K_1-K_2 is closed and convex and since $K_1 \quad K_2 = \phi$, $0 \notin K_1-K_2$. Thus there exists an open neighborhood N about 0 disjoint from K_1-K_2, so by the Hahn-Banach theorem we can separate N and K_1-K_2. Since $x^* \neq 0$ there exists a $y \in K_1-K_2$ such that $\langle x^*,y\rangle = 1$, but $\alpha y \in N$ for α small enough which implies that $\langle x^*,x\rangle \geq \epsilon$ for some $\epsilon > 0$ and all $x \in K_1-K_2$. The theorem follows by placing $c = \inf_{x \in K_1} \langle x^*,x\rangle$ and using the linearity of x^*.

Using theorems A.4.1 and A.4.3 we can now separate a point p from any set with an interior point or from any closed set if the point is not in the set. We can strengthen these results by insisting that X be a Banach space.

Let A be a closed convex subset in X. We say that x^* <u>supports</u> A at a point p, and that p is a <u>support point</u> of A if $\langle x^*,p\rangle = c$ and $\langle x^*,x\rangle \geq c$ for all $x \in A$.

<u>Theorem A.4.4</u> If A is a closed convex subset of a Banach space X, then the support points of A are dense in the boundary of A.

<u>Proof</u>: see Holmes [1975], pg. 166.

If A is a convex set in X and $\bar{A} \neq X$ then this theorem shows that the set of points $\{p | p \notin A,\ p \in \bar{A}$, and $\exists\ x^* = 0,\ c \ni \langle x^*,p\rangle = c$, $\langle x^*,x\rangle \geq c \ \ \forall x \in \bar{A}\}$ is dense in the boundary of A. That is, the set of points that can be separated from a convex set A is dense in the boundary of $\bar{A}$.

<u>Example A.4.1</u> A closed convex set with boundary points which are not support points.

Let $X = \ell_2$ and define S by

$$S = \{x | x = (\alpha_1, \ldots, \alpha_n, \ldots,), \ \alpha_i \ \varepsilon \ [0,1/i], \ i = 1, \ldots\}.$$

Since $S \subset P_+$ (see Example A.2.2) S has no interior points, thus all $x \ \varepsilon \ S$ are boundary points. The point $p = (\frac{1}{2}, \frac{1}{4}, \frac{1}{6} \ldots)$ is not a support point as the following argument demonstrates.

If $x^* \neq 0$ supports S at p then there exists a c such that $\langle x^*, p\rangle = c$ and $\langle x^*, x\rangle \geq c \ \forall x \ \varepsilon \ S$. Since $0 \ \varepsilon \ S$, $c \leq 0$. If $c < 0$ then $x' = (1, 1/2, 1/3, \ldots) \ \varepsilon \ S$ is such that $\langle x^*, x'\rangle = 2c < \langle x^*, p\rangle$, thus $c = 0$. If $\langle x^*, p\rangle = 0$ then $x^* = (\alpha_1^*, \alpha_2^*, \ldots)$ must be such that $\alpha_j^* < 0$ for some $j \geq 0$. But then $x = (0, \ldots, 0, \alpha_i, 0, \ldots) \ \varepsilon \ S$ where $\alpha_i = 1/i$ is such that $\langle x^*, x\rangle < 0$, thus no $x^* \ \varepsilon \ X^*$ separates p and S.

Note, however, that for any $\epsilon > 0$ there exists a $p' \ \varepsilon \ S$ such that $\| p-p' \| < \epsilon$ and p' is a support point of S. If $1/n < \epsilon$ then let $p' = p - (0, \ldots, 0, \alpha_n, 0, \ldots)$ where $\alpha_n = \frac{1}{n}$, and let $x^* = (0, \ldots, 0, \alpha_n^*, 0, \ldots)$ where $\alpha_n^* = 1$, so that $\langle x^*, p'\rangle = 0$ and $\langle x^*, x\rangle \geq 0 \quad x \ \varepsilon \ S$. Thus the set of support points is dense in S.

<u>Example A.4.2</u> A convex cone S with no interior points which cannot be separated from points not in S.

Let $X = \ell_2$ and define $S = C_+$ (see Example A.2.1). If x^* separates $p \notin S$ and S then by continuity it also separates p and $\bar{S}$. Since $\bar{S} = \ell_2$ no x^* separates $p \notin S$ and S.

A.5. Linear Operators and Adjoints

Consider two normed linear topological spaces X and Y and let the mapping $A: X \to Y$ be linear. It is easily seen that a linear operator is continuous at every point in X if it is continuous at any point.

The space of all linear operators from X into Y is denoted B(X,Y) and it is clear that B(X,Y) is a linear space. This space can be topologized by introducing the metric

$$\|A\| = \sup_{\|x\| \leq 1} \|Ax\| \ . \qquad \text{(A.17)}$$

As was the case with linear functionals it can be shown that a linear operator is bounded if and only if it is continuous. Finally, note that if Y is a Banach space then B(X,Y) is also a Banach space.

If A ε B(X,Y) then a functional ϕ^* on X may be defined by $\phi^*(x) = \langle y^*, Ax \rangle$ where $y^* \ \varepsilon \ Y^*$. It is easily seen that ϕ^* is linear and continuous, thus $\phi^* \ \varepsilon \ X^*$. Note that ϕ^* associates a particular x^* to each y^*; this mapping will be denoted A^*. It can be seen that A^* is linear and continuous in the weak* (and thus in the strong) topologies of X* and Y*. The adjoint operator $A^* \ \varepsilon \ B(Y^*, X^*)$ associated with A ε B(X,Y) is thus defined by the equation

$$\langle x, A^*y^* \rangle = \langle Ax, y^* \rangle \ . \qquad \text{(A.18)}$$

Let $\mathcal{R}(A)$ denote the range of A, that is,

$$\mathcal{R}(A) = \{y \mid y = Ax, \ x \ \varepsilon \ X\}, \qquad \text{(A.19)}$$

Let $\mathcal{N}(A)$ denote the null space of A, that is,

$$\mathcal{N}(A) = \{x \mid Ax = 0, \ x \ \varepsilon \ X\} \qquad \text{(A.20)}$$

and define, for a set S X, the orthogonal complement (annihilator) of S by

$$[S] = \{x^* \mid x^* \ \varepsilon \ X, \ \langle x^*, x \rangle = 0 \text{ for all } x \ \varepsilon \ S\}. \qquad \text{(A.21)}$$

Theorem A.5.1 Let X and Y be normed spaces and let A ε B(X,Y). Then $[\mathcal{R}(A)]^{\perp} = \mathcal{N}(A^*)$.

Proof: see Luenberger [1969], pg. 155.

Theorem A.5.2 Let X and Y be Banach spaces and let A ε B(X,Y). Let $\mathcal{R}(A)$ be closed. Then

$$R(A^*) = [N(A)]^{\perp}. \quad \text{(A.22)}$$

Proof: see Luenberger [1969], pg. 156.

Example A.5.1 An operator A such that $R(A)$ is not closed.

Let $A: \ell_2 \to \ell_2$ be defined by

$Ax = A(\alpha_1, \alpha_2, \ldots) = (\alpha_1, \frac{1}{2}\alpha_2, \frac{1}{3}\alpha_3 \ldots)$.

If x is finitely nonzero (see Example A.2.1) then $x \in \ell_2$, Ax is finitely nonzero and $Ax \in \ell_2$. Since the set of finitely nonzero sequences is dense in ℓ_2, $\overline{R(A)} = \ell_2$. But $R(A) \neq \ell_2$ since $(1, \frac{1}{2}, \frac{1}{3}, \ldots) \in R(A)$ would imply $(1, 1, \ldots) \in \ell_2$ which is not true.

A.6. Minkowski-Farkas Lemma

Let X be a linear topological space, Y a locally convex linear space and let $T \in B(X,Y)$. We let $P_Y \subset Y$ be a closed convex cone and define $y' \geq y''$ to mean $y' - y'' \in P_Y$. Denote the dual cone to P_Y by P_Y^* and define $y^{*\prime} \geq y^{*\prime\prime}$ to mean $y^{*\prime} - y^{*\prime\prime} \in P_Y^*$.

Define:

$$Z_T = \{x^* \in X^* | x^* = T^*y^*, \ y^* \geq 0\} \quad \text{(A.23)}$$

and

$$V_T = \{x^* \in X | \ x \in X, \ TX \geq 0 \Rightarrow \langle x^*, x\rangle \geq 0\} . \quad \text{(A.24)}$$

Theorem A.6.1 (Minkowski-Farkas Lemma) $Z_T = V_T$ if and only if Z_T is convex and closed in the weak* topology.

Proof: Hurwicz [Hu 1].

Recall that in finite dimensional Euclidean spaces Farkas' lemma states that [Ca 1]:

If $a_1, \ldots, a_k$ and b are a finite set of vectors in E^n, then

$$b^T x \geq 0 \quad \text{(A.25)}$$

for all $x \in E^n$ satisfying

$$a_i^T \geq 0 \qquad i = 1,\ldots,k \tag{A.26}$$

if and only if

$$b = \sum_{i=1}^{k} \mu_i a_i,$$

with $\mu_i \geq 0$ for $i = 1,\ldots,k$.

The Minkowski-Farkas Lemma can be seen as a generalization of the above if we note that (A.26) can be rewritten as $Ax \geq 0$ if we let the a_i define a mapping $A\colon\ E^n \to E^k$ and if we define $\geq$ by the positive orthant of E^k.

A.7. Fréchet Derivatives

The concept of derivative and differential can be extended to normed linear spaces. In the following let X and Y be normed linear spaces and let f be a (possibly nonlinear) transformation:

$$f\colon\ X \to Y.$$

The function f is <u>Fréchet differentiable</u> at x if for each $h \in X$ there exists $\delta f(x;h) \in Y$ which is linear and continuous with respect to h such that

$$\lim_{\|h\|\to 0} \|h\|^{-1}\|f(x+h) - f(x) - \delta f(x;h)\| = 0, \tag{A.27}$$

and $\delta f(x;h)$ is known as the <u>Fréchet differential of f</u> at x with increment h.

If f is Fréchet differentiable at x then it is continuous at x. The Fréchet differential is unique and is given by

$$f(x_j h) = \frac{d}{d\alpha} f(x+\alpha h)\Big|_{\alpha=0} \tag{A.28}$$

Since $\delta f(x;h)$ is linear with respect to h it has the form

$$\delta f(x_j h) = f'(x)h \tag{A.29}$$

where

$$f': \quad X \to B(X,Y). \tag{A.30}$$

The mapping f' is the Frechet derivative of f and the linear operator $f'(x) \varepsilon B(X,y)$ is the Frechet derivative of f at x. If f is a functional the f' is also called the gradient of f at x.

Let $X = X' \times X''$ be the Banach product of the two Banach spaces X' and X". If

$$f: \quad X = X' \times X'' \to Y \tag{A.31}$$

then $\delta_{x'} f(x_{o}; h)$ is the partial Fréchet differential of f with respect to x' at x_o and is defined by

$$\delta_{x'} f(x_o; h) = \delta f(x_o; (h, 0_{x''})) \tag{A.32}$$

where

$$h \,\varepsilon\, X', \; 0_{x''} \,\varepsilon\, X'' \text{ and } (h, 0_{x''}) \quad X = X' \times X''.$$

Frequently the Fréchet derivative of $f(\hat{x})$ at $\hat{x}$ will be denoted $\nabla f(\hat{x})$. Partial Fréchet derivatives will be denoted $\nabla_i f(\hat{x}_1, \ldots, \hat{x}_n)$, where i denotes the argument being differentiated.

Finally we have the following extension of the chain rule of differentiation.

Theorem A.7.1 Let S map an open set $D \subset X$ into an open set $E \subset Y$ and let P map E into a normed space Z. Let T = PS and suppose S is Fréchet differentiable at $x \,\varepsilon\, D$ and P is Fréchet differentiable at $y = S(x) \,\varepsilon\, E$. Then T is Frechet differentiable at x and $T'(x) = P'(S(x))S'(x)$.

Proof: see [Lu 1, page 176].

A.8. The Implicit Function Theorem

Theorem A.8.1 Let X and Y be Banach spaces and let $f: X \times Y \to X$ be continuous in an open neighborhood N of a point (x_o, y_o) for which $f(x_o, y_o) = 0$. If $\partial_x f$ exists in a neighborhood of (x_o, y_o), is continuous at (x_o, y_o) and if $\partial_x f(x_o, y_o)$ is nonsingular then there exist open neighborhoods $N_1 \subset X$ of x_o and $N_2 \subset Y$ of y_o such that, for each $x \in N_2$ the equation

$$f(x,y) = 0 \tag{A.33}$$

has a unique solution $\hat{x} = g(y)$ where $g: N_2 \subset Y \to X$ is continuous.

If $\partial_y f$ exists at (x_o, y_o) then g is Fréchet differentiable at y_o and

$$g'(y_o) = -[\partial_x f(x_o, y_o)]^{-1} \partial_y f(x_o, y_o). \tag{A.34}$$

Proof: see Ortega and Rheinbolt [1970] pp. 128-129.

APPENDIX B

The notation used in this and the following appendices as well as many theorems from linear and convex analysis are given in Appendix A.

Proof of Theorem 2.1.1: Recall the following three sets defined in Section 2.1:

$$K(\hat{x}) = F(\hat{x})C(\hat{x},\Omega) \text{ where } F(x) = \begin{matrix} f(x) \\ r(x) \end{matrix} \tag{B.1}$$

$$K_e(\hat{x}) = r(\hat{x})C(\hat{x},\Omega) \text{ and} \tag{B.2}$$

$$D = \{z \in \mathbb{R}^{n+1} | z = \beta(-1,0,\ldots,0),\ \beta > 0\} \tag{B.3}$$

Assume that D and $K(\hat{x})$ are not separable. We now show that this contradicts the optimality of $\hat{x}$. Note that D and $K(\hat{x})$ not separable implies that

(i) $$D \cap K(\hat{x}) \neq \phi \tag{B.4}$$

(ii) $$K_e(\hat{x}) = \mathbb{R}^n . \tag{B.5}$$

Equation (B.5) follows from Corollary A.4.1 which implies that if $0 \notin \text{int } K_e(\hat{x})$ then a nonzero vector $\underline{\phi}$ exists such that

$$\langle\underline{\phi},y\rangle \leq 0 \quad \forall y \in K_e(\hat{x}). \tag{B.6}$$

But then $(0,\underline{\phi})$ would separate D and $K(\hat{x})$. Thus $0 \in K_e(\hat{x})$ and, since $K_e(\hat{x})$ is a cone, $K_e(\hat{x}) = \mathbb{R}^n$.

Let $N = \mathcal{N}(\nabla r(\hat{x}))$ and decompose X into the direct sum N + M where M is isomorphic to X/N. Since $\nabla r(\hat{x})$ is surjective there exists a bijective map A: $M \leftrightarrow \mathbb{R}^n$ such that

$$Ax = \nabla r(\hat{x})x \qquad \forall x \in M. \tag{B.7}$$

Let Σ be a simplex in $\mathbb{R}^n$ with vertices $y_o,\ldots,y_n$ such that

(i) $$0 \in \text{int } \Sigma \text{ and} \tag{B.8}$$

(ii) $\nabla r(\hat{x})x' = 0$ (B.11)

Such an x' exists since, from (2.1.7), $D \cap K(\hat{x}) \neq \phi$.

Because Φ is a simplex in X (let its vertices be $x_o,\ldots,x_n$)

$$<\nabla f(\hat{x}),x> \leq \max_{i=0,\ldots,n} <\nabla f(\hat{x}),x_i> \quad \forall x \in \Phi. \tag{b.12}$$

Thus there exists a $\lambda > 0$ such that

$$<\nabla f(\hat{x}),x> < 0 \qquad \forall x \in \lambda x' + \Phi, \tag{B.13}$$

where $\lambda x' + \Phi \subset C(x,\Omega)$. (B.14)

Let $\bar{x} = \lambda x'$.

Define g(x,y) by

$$g(x,y) = r(\hat{x} + \zeta(A^{-1}y + x)) \tag{B.15}$$

and note that

(i) $g(0,0) = 0$ (B.16)

(ii) $\nabla_1 g(x,y) = \nabla r(\hat{x} + \zeta(A^{-1}y+x))\zeta'(A^{-1}y+x)$ (B.17)

(iii) $\nabla_2 g(x,y) = \nabla r(\hat{x} + \zeta(A^{-1}y+x))\zeta'(A^{-1}y+x)A^{-1}$ (B.18)

(iv) $\nabla_2 g(0,0) = I$. (B.19)

By Theorem A.8.1 equations (B.16-B-.19) imply that there exists a neighborhood U of $\{0\} \in X$ and a function $\psi(x)$ such that

(i) $\psi(0) = 0$ (B.20)

(ii) $r(\hat{x} + \zeta(A^{-1}\psi(x) + x)) = 0 \quad \forall x \in U$ (B.21)

(iii) $\psi'(0) = -\nabla r(\hat{x})$. (B.22)

Choose $x = \alpha\bar{x}$ and recall that $\bar{x} \in N$ so that

$$y(\lambda) = \psi(\lambda\bar{x}) = \psi(0) + \lambda\psi'(0)\bar{x} + o(\lambda\bar{x}) = o(\lambda\bar{x}), \tag{B.23}$$

where

$$\lim_{|\alpha|\to 0} ||o(\alpha x)||/|\alpha| = 0. \tag{B.24}$$

In light of equation (B.24) it is clear that there exists an $\epsilon > 0$ such that

$$||y(\lambda)|| \leq \lambda\rho \qquad \forall \lambda \in [0,\epsilon] \tag{B.25}$$

Thus

$$\lambda\bar{x} + A^{-1}y(\lambda) \in \lambda(\bar{x} + \Phi) \subset C(\hat{x},\Omega) \qquad \forall \lambda \in [0,\eta]. \tag{B.26}$$

Now by the definition of ζ there exists an $\eta \in (0,\epsilon]$ such that

$$\hat{x} + \zeta(\lambda\bar{x} + A^{-1}y(\lambda)) \in \Omega \quad \forall \lambda \in [0,\eta]. \tag{B.27}$$

Recall that f(x) is differentiable so that

$$f(\hat{x}+\zeta(\lambda\bar{x}+A^{-1}y(\lambda))) = f(\hat{x}) + <\nabla f(\hat{x}),\zeta(\lambda x+A^{-1}y(\lambda))> + o(\zeta(\lambda x+A^{-1}y(\lambda))) \tag{B.28}$$

which, using (2.1.2) and (B.23) can be written

$$f(\hat{x} + \zeta(\lambda\bar{x}+A^{-1}y(\lambda))) = f(\hat{x}) + <\nabla f(\hat{x}),\lambda\bar{x}> + o'(\lambda\bar{x}) . \tag{B.29}$$

Now since

$$<\nabla f(\hat{x}),\lambda\bar{x}> < 0 \tag{B.30}$$

there exists an $\eta^* \in (0,\eta]$ such that, letting $\delta x^* = \zeta(\eta^*\bar{x}+A^{-1}y(\eta^*))$,

(i) $f(\hat{x}+\delta x^*) < f(\hat{x})$ (B.31)

(ii) $r(\hat{x}+ \delta x^*) = 0$ (B.32)

(iii) $x + \delta x^* \in \Omega$ (B.33)

but this contradicts the optimality of $\hat{x}$ thus D and $K(\hat{x})$ are separable. That is, there exists a vector $\underline{\phi} \in \mathbb{R}^{n+1}$ such that $\underline{\phi} \neq 0$ and

$$<\underline{\phi},y> \geq <\underline{\phi},z> \quad \forall y \in D \qquad \forall z \in K(\hat{x}) . \tag{B.34}$$

Since $0 \in K(\hat{x})$ $<\underline{\phi},y> \geq 0$ $\forall y \in D$, which implies that $\phi_o \leq 0$. Now if there exists a $z' \in K(\hat{x})$ such that $<\underline{\phi},y> > 0$ then, for any y in D, there is an $\alpha > 0$ such that $\alpha z' \varepsilon K(\hat{x})$ and $<\alpha z',\underline{\phi}> > <y,\underline{\phi}>$, thus $<\underline{\phi},z> \leq 0$ for all $z \varepsilon K(\hat{x})$. Clearly continuity implies then that

$$<\underline{\phi},z> \leq 0 \qquad \forall z \varepsilon \overline{K(\hat{x})} \tag{B.35}$$

However, $\nabla F(\hat{x})\overline{C(\hat{x},\Omega)} \subset \overline{K(\hat{x})}$, so letting $\underline{\phi} = (\lambda,\underline{\psi})$ and using equation (2.1.3), one has

$$\lambda<\nabla f(\hat{x}),\delta x> + <\underline{\psi},\nabla r(\hat{x})\delta x> \leq 0 \quad \forall x \varepsilon \overline{C(\hat{x},\Omega)} \tag{B.36}$$

<u>Proof of Theorem 2.1.2</u>: Let $N = \mathcal{N}(T)$ and let $X = N + M$. Consider the set $C(\hat{x},\Omega) \cap N$. For any finite set of linearly independent vectors $\{\delta x_k\} \subset C(\hat{x},\Omega) \cap N$ clearly there exists an $\epsilon > 0$ such that

$$\hat{x} + \delta x \subset \Omega \quad \forall \delta x \varepsilon co\{\delta x_1,\ldots,\delta x_k\} \tag{B.37}$$

If $x \varepsilon \Omega \cap N$,

$$\hat{x} + \delta x \varepsilon \Omega \cap N \qquad x \varepsilon co\{\delta x_1,\ldots,\delta x_k\} \tag{B.38}$$

so that $C(\hat{x},\Omega) \cap N$ is a conical approximation to the set $\Omega \cap N$ at $\hat{x} \varepsilon \Omega \cap N$.

Since $T(x)$ is Fréchet differentiable $\nabla T(\hat{x}) = T \varepsilon B(X,Y)$ and thus [Lu, Prob. 6.8] $N = \mathcal{N}(T)$ is closed so that $M = X/N$ is a Banach space [Lu 1, Prop. 2.14.1]. Note that $\mathcal{R}(T) \equiv Z$ is closed so it is also a Banach space. The restriction of T to M, $\hat{T}$: $M \rightarrow Z$, is in $B(M,Z)$ and is bijective. The Banach Inverse Theorem [Lu 1, Theorem 6.4.1] then implies $\hat{T}^{-1} \varepsilon B(Z,M)$. Clearly $Z + X/Z = Y$ is closed, thus [Ho 1, Lemma 17H] $M = \hat{T}^{-1}(Z)$ is closed and [Ho 1, Theorem 17I] the projection operator P_M: $X \rightarrow M$ is continuous as is $P_N = I - P_M$: $X \rightarrow N$.

Note that problem (2.1.9) can be written

$$\min f(x)$$

subject to (B.39)

$$r(x) = 0 \qquad x \in \Omega \cap N.$$

An application of Theorem 2.1.1 leads to the existence of a $\underline{\phi} \in \mathbb{R}^{n+1}$ such that

$$\langle \underline{\phi}, F(\hat{x})\delta x \rangle \leq 0 \quad \forall \delta x \in C(\hat{x},\Omega \cap N). \tag{B.40}$$

Note that the conical approximation in equation (B.40) can be chosen to be

$$C(\hat{x},\Omega \cap N) = C(\hat{x},\Omega) \cap N. \tag{B.41}$$

Recall that in the proof of Theorem 2.1.1 $\underline{\phi}$ was chosen to separate the sets D and $K(\hat{x})$. In problem (B.39)

$$K(\hat{x}) = \nabla F(\hat{x})C(\hat{x},\Omega \cap N) = \nabla F(\hat{x})C(\hat{x},\Omega) \cap \nabla F(\hat{x})N. \tag{B.42}$$

If $\nabla F(\hat{x})N = \mathbb{R}^{n+1}$ then $K(\hat{x}) = \nabla F(\hat{x})C(\hat{x},\Omega)$ in which case

$$\lambda\langle \nabla f(\hat{x}),\delta x\rangle + \langle \underline{\phi},\nabla r(\hat{x})\delta x\rangle + \langle y^*,T\delta x\rangle \leq 0 \quad \forall x \in \overline{C(\hat{x},\Omega)} \tag{B.43}$$

where $y^* \in [\mathcal{R}(T)]^{\perp}$. If $\nabla F(\hat{x})N = \mathbb{R}^m$, $m < n+1$, then clearly $\underline{\phi} \neq 0$ can be chosen so that

$$\langle \underline{\phi},z\rangle = 0 \qquad \forall z \in \nabla F(\hat{x})N \tag{B.44}$$

and thus so that

$$\langle \underline{\phi},z\rangle = 0 \qquad \forall z \in \nabla F(\hat{x})P_N C(\hat{x},\Omega) \tag{B.45}$$

That is, $\underline{\phi}$ separates D and $K'(\hat{x})$, where

$$K'(\hat{x}) = \nabla F(\hat{x})P_N C(\hat{x},\Omega) \tag{B.46}$$

This implies that there exists a $\lambda \leq 0$ and a $\underline{\psi} \in \mathbb{R}^n$, not both zero, such that

$$\lambda\langle\nabla f(\hat{x}),\delta x\rangle + \langle\underline{\psi},\nabla r(\hat{x})\delta x\rangle \le 0 \quad \forall \delta x \in \overline{P_N C(\hat{x},\Omega)} \tag{B.47}$$

or

$$\lambda\langle\nabla f(\hat{x}),P_N\delta x\rangle + \langle\underline{\psi},\nabla r(\hat{x})P_N\delta x\rangle \le 0 \quad \forall \delta x \in \overline{C(\hat{x},\Omega)}. \tag{B.48}$$

Since $P_N = I - P_M$ equation (B.48) becomes

$$\begin{aligned} &\lambda\langle\nabla f(\hat{x}),\delta x\rangle + \langle\underline{\psi},\nabla r(\hat{x})\delta x\rangle \\ &- \lambda\langle\nabla f(\hat{x}),P_M\delta x\rangle - \langle\underline{\psi},\nabla r(\hat{x})P_M\delta x\rangle \le 0 \quad \forall \delta x \in \overline{C(\hat{x},\Omega)} \end{aligned} \tag{B.49}$$

But

$$\begin{aligned} &\lambda\langle\nabla f(\hat{x}),P_M\delta x\rangle + \langle\underline{\psi},\nabla r(\hat{x})P_M\delta x\rangle = \\ &\langle\lambda P_M{}^*\nabla f(\hat{x}) + P_M{}^*\nabla r^*(\hat{x})\underline{\psi},\delta x\rangle = 0 \quad \forall \delta x \in N, \end{aligned} \tag{B.50}$$

thus, since $P_M{}^*$ and $\nabla r^*(\hat{x})$ are continuous,

$$\lambda P_M{}^*\nabla f(\hat{x}) + P_M{}^*\nabla r^*(\hat{x})\psi \in [N]^{\perp} \subset X^*. \tag{B.51}$$

However $\mathcal{R}(T)$ is closed and thus, by Theorem A.5.2

$$[N]^{\perp} = \mathcal{R}(T^*) \tag{B.52}$$

so that there exists a $y^* \in Y^*$ such that

$$-T^*y^* = \lambda P_M{}^*\nabla f(\hat{x}) + P_M{}^*\nabla r^*(\hat{x})\underline{\psi}\ . \tag{B.53}$$

Equation (B.49) thus becomes the desired

$$\lambda\langle\nabla f(\hat{x}),\delta x\rangle + \langle\underline{\psi},\nabla r(\hat{x})\delta x\rangle + \langle y^*,T\delta x\rangle \le 0 \quad \forall \delta x \in \overline{C(\hat{x},\Omega)} \tag{B.54}$$

<u>Proof of Theorem 2.2.1</u>: If $\mathcal{R}(\nabla T(\hat{x})) \ne Y$ then clearly there exists a nonzero $y^* \in Y$ such that

$$\langle y^*,\ \nabla T(\hat{x})\delta x\rangle = 0 \quad \forall \delta x \in X \tag{B.55}$$

so that (2.2.10) can be trivially satisfied. Assume then that

$\mathcal{R}(\nabla T(\hat{x})) = Y$.

Let $N = \mathcal{N}(\nabla T(\hat{x}))$ and $X = N + M$. By Luisternik's Theorem [Lui 1, Theorem VI.46.1] there exists a function that maps N into A where $A = \{x | T(x) = 0\}$. In particular, there exists a mapping $\gamma(x)$ such that

$$T(\hat{x}+\gamma(\delta x)) = 0 \qquad \forall \delta x \in U \cap N \tag{B.56}$$

where U is some neighborhood of $\hat{x}$ and

$$\gamma(\delta x) = \delta x + o(\delta x) \tag{B.57}$$

where

$$\lim_{\|x\| \to 0} \|o(x)\| / \|x\| = 0 \ . \tag{B.58}$$

This implies that N is a quasilinear conical approximation to A with $\zeta(\delta x) = \gamma(\delta x)$.

Consider a finite set of linearly independent vectors $\{\delta x_1, \ldots, \delta x_k\}$ in $RC(\hat{x}, \Omega^{\circ}) \cap N$. Clearly $RC(\hat{x}, \Omega^{\circ})$ is a conical approximation to the set Ω° since $RC(\hat{x}, \Omega^{\circ})$ is convex. Thus there exists an $\epsilon > 0$ such that

$$\hat{x} + \delta x \in \Omega^{\circ} \qquad x \in co\{\epsilon\delta x_1, \ldots, \epsilon\delta x_k\} \ . \tag{B.59}$$

Since $x + \epsilon\delta x_j \in \Omega^{\circ}$, $j = 1, \ldots, k$ there exist $\alpha_j > 0$ $j = 1, \ldots, k$ such that

$$\hat{x} + \epsilon\delta x_j + \delta z \in \Omega^{\circ} \qquad \forall \delta z \in S(0, \alpha_j) \ j = 1, \ldots, k \tag{B.60}$$

Let $\alpha = \min_{j=1,\ldots,k} \alpha_j$. Since Ω° is convex

$$\hat{x} + \epsilon \sum_{j=1}^{k} \beta_j \delta x_j + \delta z \in \Omega^{\circ} \qquad \forall \delta z \in S(0, \alpha) \ . \tag{B.61}$$

where $\sum_{j=1}^{k} \beta_j = 1$, $\beta_j \geq 0$.

Now by the definition of $\zeta(\delta x) = \gamma(\delta x)$ in equations (B.57) and (B.58) there exists an $\epsilon^* > 0$ such that

$$\zeta(\lambda\delta x) \in S(\lambda\delta x, \lambda\alpha) \quad \forall\, \delta x = \sum_{j=1}^{k} \beta_j \beta x_j, \quad \sum_{j=1}^{k} \beta_j = 1, \quad \beta_j \geq 0 \tag{B.62}$$

But (B.62) implies that

$$\hat{x} + \zeta(\delta x) \in \Omega^\circ \quad \forall\, \delta x \in \text{co}\{\epsilon^*\delta x_1, \ldots, \epsilon^*\delta x_k\}, \tag{B.63}$$

and thus $RC(\hat{x},\Omega^\circ) \cap N$ is a quasilinear conical approximation to the set $\Omega \cap A$.

Arguing as in the proof of Theorem 2.1.2 yields the existence of a $\lambda \leq 0$ and a $\underline{\phi} \in \mathbb{R}^n$ not both zero and a nonzero $y^* \in Y^*$ such that

$$\lambda\langle\nabla f(\hat{x}),\delta x\rangle + \langle\underline{\phi},\nabla r(\hat{x})\delta x\rangle + \langle y^*,\nabla T(\hat{x})\delta x\rangle \leq 0 \quad \forall\, \delta x \in \overline{RC(\hat{x},\Omega^\circ)} \tag{B.64}$$

Note that

$$\overline{RC(\hat{x},\Omega^\circ)} = \{\delta x \mid \exists\; \delta x_1(\epsilon) \in RC(\hat{x},\Omega^\circ),\ \epsilon > 0 \quad \|\delta x - \delta x_1(\epsilon)\| \leq \epsilon\} \tag{B.65}$$

$$\overline{RC(\hat{x},\overline{\Omega})} = \{\delta x \mid \exists\; \delta x_2(\epsilon) \in RC(\hat{x},\overline{\Omega}),\ \epsilon > 0 \quad \|\delta x - \delta x_2(\epsilon)\| \leq \epsilon\} \tag{B.66}$$

$$\overline{RC(\hat{x},\overline{\Omega})} = \{\delta x \mid \exists\, \overline{x}(\epsilon) \in \Omega^\circ,\ \epsilon > 0,\ \ni \|\overline{x}(\epsilon) - (x+\delta x)\| \leq \epsilon\}. \tag{B.67}$$

Choose $\delta x \in \overline{RC(\hat{x},\Omega)}$ and note that if

$$\delta x_3(\eta) = \overline{x}(\eta) - \hat{x} \tag{B.68}$$

where $\overline{x}(\eta) \in \Omega^\circ$ is defined in (B.67) then

$$\delta x_3(\eta) \in RC(\hat{x},\Omega^\circ)\ . \tag{B.69}$$

Corresponding to δx is a $\delta x_2(\phi)$ such that

$$\delta x_2(\phi) \in RC(\hat{x},\overline{\Omega}) \tag{B.70}$$

and

$$\|\delta x-\delta x_2(\phi)\| \leq \phi, \tag{B.71}$$

so that there exists a $\delta x_3(\eta,\phi) \in RC(\hat{x},\Omega^\circ)$ such that

$$\|\delta x_3(\eta,\phi)-\delta x_2(\phi)\| \leq \eta . \tag{B.72}$$

Letting $\phi = \eta = \epsilon/2$ one has that there exists a $\delta x_3(\epsilon/2,\epsilon/2) \in RC(\hat{x},\Omega)$ such that

$$\|\delta x-\delta x_3(\epsilon/2,\epsilon/2)\| \leq \|\delta x-\delta x_2(\epsilon/2)\|+\|\delta x_2(\epsilon/2)-\delta x_3(\epsilon/2,\epsilon/2)\| \leq \epsilon \tag{B.73}$$

so that $\delta x \in \overline{RC(\hat{x},\Omega^\circ)}$.

Clearly $\overline{RC(\hat{x},\Omega^\circ)} \subset \overline{RC(\hat{x},\bar{\Omega})}$ so that (B.73) implies $\overline{RC(\hat{x},\Omega^\circ)} = \overline{RC(\hat{x},\Omega)} = \overline{RC(\hat{x},\bar{\Omega})}$ and thus

$$\lambda\langle\nabla f(\hat{x}),\delta x\rangle + \langle\underline{\phi}\nabla r(\hat{x})\delta x\rangle + \langle y^*,\nabla T(\hat{x})\delta x\rangle \leq 0 \qquad \forall\, \delta x \in \overline{RC(\hat{x},\Omega)}. \tag{B.74}$$

<u>Proof of Theorem 2.4.1</u>: Define the following three sets:

$$K(\hat{x}) = \nabla F(\hat{x})C(\hat{x},\Omega) \text{ where } \quad F(x) = \begin{matrix} f(x) \\ r(x) \end{matrix} \tag{B.75}$$

$$K_e(\hat{x}) = \nabla r(\hat{x})C(\hat{x},\Omega) \text{ and} \tag{B.76}$$

$$D = \{z \in Y \times \mathbb{R}^n \mid z = (-k,\underline{0}),\ k \in \text{int } K\}. \tag{B.77}$$

As before we show that if D and $K(\hat{x})$ are not separable then $\hat{x}$ is not optimal. Note that if D and $K(\hat{x})$ are not separable then

(i) $D \cap K(\hat{x}) \neq \phi$ (B.78)

(ii) $K_e(\hat{x}) = \mathbb{R}^n$ (B.79)

Let $N = N(\nabla r(\hat{x}))$, $X = N + M$ and A be a bijective map $A\colon\ M \to \mathbb{R}^n$

equal to $\nabla r(\hat{x})$ on M. Let Σ be a simplex in $\mathbb{R}^n$ with vertices $v_o,\ldots,v_n$ such that

$$\text{(i)}\quad 0 \in S(0,\rho) \subset \text{int}\ \Sigma \text{ for some } \rho > 0 \tag{B.80}$$

$$\text{(ii)}\quad \Phi = A^{-1}\Sigma \subset C(\hat{x},\Omega). \tag{B.81}$$

Since, from (B.78), $D \cap K(\hat{x}) \neq \phi$, there exists a $x' \in C(\hat{x},\Omega)$ such that:

$$\text{(i)}\quad -\nabla f(\hat{x})x' \in \text{int } K \tag{B.82}$$

$$\text{(ii)}\quad \nabla r(\hat{x})x' = 0. \tag{B.83}$$

Let the vertices of Φ be denoted $x_o,\ldots,x_n$ and note that

$$\|\nabla f(\hat{x})x\| \leq \max_{i=0,\ldots,n} \|\nabla f(\hat{x})x_i\| \qquad \forall x \in \Phi. \tag{B.84}$$

Thus there exists a $\lambda > 0$ such that

$$-\nabla f(\hat{x})x \in \text{int } K \qquad \forall x \in \lambda x' + \Phi \tag{B.85}$$

where $$\lambda x' + \Phi \subset C(\hat{x},\Omega) . \tag{B.86}$$

Let $\bar{x} = \lambda x'$ and define $g(x,y)$ by

$$g(x,y) = r(\hat{x}) + \zeta(A^{-1}y + x) \tag{B.87}$$

and note:

$$\text{(i)}\quad g(0,0) = 0 \tag{B.88}$$

$$\text{(ii)}\quad \nabla_1 g(x,y) = \nabla r(\hat{x}+\zeta(A^{-1}y+x))\zeta'(A^{-1}y+x) \tag{B.89}$$

$$\text{(iii)}\quad \nabla_2 g(x,y) = \nabla r(\hat{x}+\zeta(A^{-1}y+x))\zeta'(A^{-1}y+x)A^{-1} \tag{B.90}$$

$$\text{(iv)}\quad \nabla_2 g(0,0) = I \tag{B.91}$$

Theorem A.8.1 again implies that there exists a neighborhood U of $\{0\} \in X$ and a function $\psi(x)$ such that

$$\text{(i)}\quad \psi(0) = 0 \tag{B.92}$$

$$\text{(ii)}\quad r(\hat{x}+\zeta(A^{-1}\psi(x)+x)) = 0 \qquad \forall x \in U \tag{B.93}$$

$$\text{(iii)}\quad \psi'(0) = -\nabla r(\hat{x}). \tag{B.94}$$

Let $x = \alpha\bar{x} \in N$ so that

$$y(\lambda) = \psi(\lambda\bar{x}) = o(\lambda\bar{x}) \tag{B.95}$$

where

$$\lim_{|\alpha|\to 0} \|o(\alpha x)\|/|\alpha| = 0 \ . \tag{B.96}$$

Let $\epsilon > 0$ be such that

$$\|y(\lambda)\| \leq \lambda\rho \qquad \forall \lambda \in [0,\epsilon] \tag{B.97}$$

Clearly

$$\lambda\bar{x} + A^{-1}y(\lambda) \in \lambda(\bar{x}+\Phi) \subset C(\hat{x},\Omega) \qquad \forall\lambda \in [0,\epsilon] \tag{B.98}$$

thus there exists an $\eta \in (0,\epsilon]$ such that

$$\hat{x} + \zeta(\lambda\bar{x}+A^{-1}y(\lambda)) \in \Omega \qquad \forall\lambda \in [0,\eta] \tag{B.99}$$

Recall that

$$\begin{aligned} f(\hat{x} + \zeta(\bar{x}+A^{-1}y(\lambda))) &= f(\hat{x}) + \nabla f(\hat{x})\zeta(\lambda\bar{x}+A^{-1}y(\lambda)) + \\ &\quad o(\zeta(\lambda\bar{x} + A^{-1}y(\lambda))) \\ &= f(\hat{x}) + \lambda\nabla f(\hat{x})\bar{x} + o'(\lambda\bar{x}) \end{aligned} \tag{B.100}$$

Since $-\nabla f(\hat{x})\bar{x} \in \text{int } K$ there exists a $\delta > 0$ such that $\nabla f(\hat{x})\bar{x} + \delta y \in \text{int } K \ \forall \delta y \in S(0,\delta)$. Let $\eta^* \in [0,\eta]$ be such that

$$\|o'(\lambda\bar{x})\| \leq \lambda\delta \qquad \forall \lambda \in [0,\eta^*] \tag{B.101}$$

Thus $x^* = \zeta(\eta^*\bar{x} + A^{-1}y(\eta^*))$ is such that:

(i) $f(\hat{x}) - f(\hat{x} + \delta x^*) \in \text{int } K$ (B.102)

(ii) $r(\hat{x} + \delta x^*) = 0$ (B.103)

(iii) $\hat{x} + \delta x^* \in \Omega$

but since $0 \notin \text{int } K, f(\hat{x}) \neq f(\hat{x} + \delta x^*)$ which contradicts the K-non-inferiority of $\hat{x}$, thus D and $K(\hat{x})$ can be separated. That is, there exists a $z^* \in [Y \times \mathbb{R}^n]^*$ such that $z^* \neq 0$ and

$$\langle z^*,z'\rangle \geq \langle z^*,z''\rangle \quad \forall z' \in D \quad \forall z'' \in K(\hat{x}) \tag{B.105}$$

Since $0 \in K(\hat{x})$

$$<z^*,z'> \geq 0 \quad \forall z' \in D \tag{B.106}$$

but the definition of D (equation B.77)) implies then that $z^* = (y^*,\underline{\phi})$ with $-y^* \in K^*$ and $\underline{\phi} \in \mathbb{R}^n$.

The definition of $\nabla F(\hat{x})$ and continuity then imply that

$$<y^*,\nabla f(\hat{x})\delta x> + <\underline{\phi},\nabla r(\hat{x})\delta x> \leq 0 \quad \forall \delta x \in \overline{C(\hat{x},\Omega)} \; . \tag{B.107}$$

APPENDIX C

Proof of Theorem 3.5.1:

(i) If R_t is not surjective then there exists a nonzero ψ_t such that $R_t^* \psi_t^* = 0$ and Theorem 3.5.1 can be satisfied trivially. Thus assume R_t is surjective for $t = 1,\ldots,T$. If $\{G_t^i\}$ (as defined for Corollary 3.5.1) are positively linearly dependent then a nonzero $k_t^* \in \tilde{K}_t^*$ exists such that $G_t^* k_t^* = 0$ and thus Theorem 3.5.1 can be satisfied trivially. Thus assume $\{G_t^i\}^I$ is positively linearly independent for $t = 1,\ldots,T$. Note that $\langle k_t^*, G_t \hat{\pi}_t \rangle = 0$, $t = 1,\ldots,T$ since we can choose $k_t^i = 0$ if $(G_t \hat{\pi}_t)^i \neq 0$, thus equation (3.5.34) holds.

(ii) Theorem 2.1.2 and Corollary 2.1.4 imply that there exists a $\lambda \leq 0$, $\underline{\psi} \in R^{e_0+\ldots+e_T}$, $-\hat{\phi}^* \in (\Pi_1 \times \ldots \times \Pi_T)^*$ not all zero such that

$$\lambda \sum_{t=0}^{T} \delta v_t^o + \sum_{t=0}^{T-1} \langle -\hat{\phi}_{t+1}^*, \delta\pi_{t+1} - \delta v_t^1 \rangle + \sum_{t=0}^{T} \langle +\psi_t, R_t \delta\pi_t \rangle \leq 0$$

$$\delta z \in C(\hat{z}, \Omega') \qquad (C.1)$$

where we have made use of the fact that

$\hat{\phi}^* = (\hat{\phi}_1^*,\ldots,\hat{\phi}_T^*)$, $\hat{\phi}_t^* \in \Pi_t^*$, $t = 1,\ldots,T$ and

$\psi = (\psi_1,\ldots,\psi_T)$, $\psi_t \in [\mathbb{R}^{e_t}]$, $t = 1,\ldots,T$.

(iii) Consider equation (C.1). Let $\delta z = (0,\ldots,0,\delta v_t,0,\ldots,0)$ be in $C(\hat{z},\Omega')$. Then (C.1) implies that

$$\lambda \delta v_t^o + \langle \hat{\phi}_{t+1}^*, \delta v_t^1 \rangle \leq 0 \qquad \delta v_t \in RC(\hat{v}_t, co(F_t(\Gamma_t)\ \hat{\pi}_t)) \qquad (C.2)$$

Note that this can be written

$$\lambda (v_t^o - \hat{v}_t^o) + \langle \hat{\phi}_{t+1}^*, v_t^1 - \hat{v}_t^1 \rangle \leq 0 \qquad v_t \in \hat{v}_t + RC(\hat{v}_t, co(F_t(\Gamma_t)\hat{\pi}_t)) \quad (C.3)$$

and by continuity

$$\lambda(v_t^o - \hat{v}_t^o) + \langle \hat{\phi}^*_{t+1}, v_t^1 - \hat{v}_t^1 \rangle \leq 0 \qquad v_t \; \varepsilon \; \overline{(\hat{v}_t + RC(\hat{v}_t, co(F_t(\Gamma_t)\hat{\pi}_t)))}. \quad (C.4)$$

It is readily verified that, for a convex set U,

$$\overline{\hat{u} + RC(\hat{u}, U)} \quad U \qquad (C.5)$$

so

$$\overline{\hat{v}_t + RC(\hat{v}_t, co(F_t(\Gamma_t)\hat{\pi}_t))} \quad co\;(F_t(\Gamma_t)\hat{\pi}_t) \quad F_t(\Gamma_t)\hat{\pi}_t. \qquad (C.6)$$

Thus

$$\lambda(\hat{v}_t^o - v_t^o) + \langle \hat{\phi}^*_{t+1}, \hat{v}_t^1 - v_t^1 \rangle \geq 0 \qquad v_t \; \varepsilon \; F_t(\Gamma_t)\hat{\pi}_t \qquad (C.7)$$

so

$$\lambda(\langle h_t(\hat{\gamma}_t), \hat{\pi}_t \rangle - \langle h_t(\gamma_t), \hat{\pi}_t \rangle) + \langle \hat{\phi}^*_{t+1}, P(\hat{\gamma}_t)\hat{\pi}_t - P_t(\gamma_t)\hat{\pi}_t \rangle \geq 0 \qquad (C.8)$$

$$\gamma_t \; \varepsilon \; \Gamma_t$$

which obviously implies (3.5.35).

(iv) Now let $\delta z = (0,\ldots,0,\delta\pi_T,0,\ldots,0,\delta v_T,0,\ldots,0)$ where

$$\delta\pi_T \; \varepsilon \; \overline{IC(\hat{\pi}_T, S_T')}, \qquad (C.9)$$

and $\delta v_T = h_T^* \delta\pi_T$. Equation (C.1) becomes

$$\lambda \langle h_T^*, \delta\gamma_T \rangle + \langle -\hat{\phi}_T^*, \delta\pi_T \rangle + \langle \psi_T^*, R_T \delta\pi_T \rangle \leq 0 \qquad \delta\pi_T \; \varepsilon \; \overline{IC(\hat{\pi}_T, S_T)} \qquad (C.10)$$

thus

$$\hat{\phi}_T^* - \lambda h_T^* - R_T^* \psi_T^* \; \varepsilon \; [\overline{IC(\hat{\pi}_T, S_T)}]^*. \qquad (C.11)$$

By assumption $\{G_T^i\}^I$ is positively linearly independent so that Theorem 2.3.5 implies that there exists a $k_T^* \; \varepsilon \; \tilde{K}_T^*$ such that

$$\hat{\phi}_T^* - \lambda h_T^* - R_T^* \psi_T^* = G_T^* k_T^* \qquad (C.12)$$

which is just equation (3.5.33).

(v) Let $\delta z = (0,\ldots,0,\delta\pi_t,0,\ldots,0,\delta v_t,0,\ldots,0)$ where

$$\delta v_t = F_t(\hat{\gamma}_t)\delta\pi_t \tag{C.13}$$

and $\delta\pi_t \in \overline{IC(\hat{\pi}_t,S_t)}$, $t = 1,\ldots,T-1$.

Equation (C.1) becomes

$$\lambda v_t^o - \langle\phi_t^*,\delta\pi_t\rangle + \langle\phi_{t+1}^*,\delta v_t^1\rangle + \langle\psi_t^*,R_t\delta\pi_t\rangle \leq 0 \qquad \delta\pi_t \in \overline{IC(\hat{\pi}_t,S_t)} \tag{C.14}$$

which can be written

$$-\lambda\langle h_t^*(\hat{\gamma}_t),\delta\pi_t\rangle + \langle\phi_t^*,\delta\pi_T\rangle - \langle\phi_{t+1}^*,P_t(\hat{\gamma}_t)\delta\pi_t\rangle - \langle\psi_t,R_t\delta\pi_t\rangle \geq 0 \qquad \delta\pi_t \in \overline{IC(\hat{\pi}_t,S_t)} \tag{C.15}$$

Theorem 2.3.5 then implies

$$\phi_t^* - P_t^*(\hat{\gamma}_t)\phi_{t+1}^* - \lambda h_t^*(\hat{\gamma}_t) - R_t^*\psi_t^* \in G_t^*\tilde{K}_t^* \tag{C.16}$$

which is equivalent to equation (3.5.32), for $t = 1,\ldots,T-1$. Equation (3.5.32) is then used to define ϕ_o^*

Proof of Theorem 3.6.1:

(i) If R_t is not surjective then there exists a nonzero ψ_t^* such that $R_t^*\psi_t^* = 0$ and Theorem 3.5.1 can be satisfied trivially. Thus assume R_t is surjective for $t = 1,\ldots,T$. If $\{G_t^i\}^I$ (as defined for Corollary 3.5.1) are positively linearly dependent then a nonzero $k_t^* \in \tilde{K}_t^*$ exists such that $G_t^*k_t^* = 0$ and thus Theorem 3.5.1 can be satisfied trivially. Thus assume $\{G_t^i\}^I$ is positively linearly independent for $t = 1,\ldots,T$. Note that $\langle k_t^*,G_t\hat{\pi}_t\rangle = 0$, $t = 1,\ldots,T$ since we can choose $\kappa_t^i = 0$ if $(G_t\hat{\pi}_t)^i \neq 0$, thus equation (3.5.34) holds.

(ii) Theorem 2.2.1 implies that there exist a $\lambda \le 0$, $\psi \in [R^{e_+ \ldots + e_+}]^*$, $-\hat{\phi}^* \in [\Pi_1 \times \ldots \times \Pi_T]^*$ not all zero such that

$$\lambda\left(\sum_{t=1}^{T} \langle h_t^*(\hat{\gamma}_t), \delta\pi_t\rangle + \sum_{t=0}^{T-1} \langle \nabla_\gamma h_t^*(\hat{\gamma}_t)\delta\gamma_t, \hat{\pi}_t \rangle\right) +$$

$$\sum_{t=0}^{T-1} \langle -\hat{\phi}^*_{t+1}, \delta\pi_{t+1} - P_t(\hat{\gamma}_t)\delta\pi_t - \nabla_\gamma P_t(\hat{\gamma}_t)\delta\gamma_t \hat{\pi}_t \rangle +$$

$$\sum_{t=0}^{T} \langle \psi_t^*, R_t \delta\pi_t \rangle \le 0 \qquad \delta z_t \in \overline{C(\hat{z}, \Omega)} \tag{C.17}$$

(iii) Consider equation (C.17). Let $\delta z = (0, \ldots, 0, \delta\gamma_t, 0, \ldots, 0)$ where $t = 0, \ldots, T-1$, be in $C(\hat{z}, \Omega)$. Then

$$\lambda \langle \nabla_\gamma h_t^*(\hat{\gamma}_t)\delta\gamma_t, \hat{\pi}_t \rangle + \langle \hat{\phi}^*_{t+1}, \nabla_\gamma P_t(\hat{\gamma}_t)\delta\gamma_t \hat{\pi}_t \rangle \le 0 \tag{C.18}$$

$$\delta\gamma_t \in \overline{C(\hat{\gamma}_t, \Gamma_t)}.$$

which is just equation (3.6.23).

(iv) Let $\delta z = (0, \ldots, \delta\pi_t, 0, \ldots, 0)$, where $t = 0, \ldots, T-1$, be in $C(\hat{z}, \Omega)$. Then equation (C.17) becomes

$$\lambda \langle h_t^*(\hat{\gamma}_t), \delta\pi_t \rangle + \langle \hat{\phi}^*_{t+1}, P_t(\hat{\gamma}_t)\delta\pi_t \rangle - \langle \hat{\phi}^*_t, \delta\pi_t \rangle + \langle \psi_t^*, R_t \delta\pi_t \rangle \le 0$$

$$\delta\pi_t \in \overline{IC(\hat{\pi}_t, S_t)} \tag{C.19}$$

Theorem 2.3.5 now implies that there exists a $k^* \in \tilde{K}^*$ such that

$$\phi_t^* - P_t^*(\hat{\gamma}_t)\phi^*_{t+1} + \lambda h_t^*(\hat{\gamma}_t) + G_t^* k_t^* + R_t^* \psi_t^*, \quad t = T-1, \ldots, 0 \tag{C.20}$$

which is just equation (3.6.20)

(v) If $\delta z = (0, \ldots, \delta\pi_T, 0, \ldots, 0)$ the same argument implies equation (3.6.21).

BIBLIOGRAPHY

[Ath 1] M. Athans, "The role and use of the stochastic linear-quadratic-Gaussian problem in control system design", IEEE Trans. A.C., vol. AC-16, no. 6, 1971.

[Ath 2] M. Athans, "The matrix minimum principle", Inf. and Contr., vol. 11, 1968.

[Baz 1] M. Bazaraa, C. Shetty, J. Goode and M. Nashed, "Nonlinear programming without differentiability in Banach spaces: necessary and sufficient constraint qualification", J. Appl. Func. Anal., vol. 11, 1977.

[Baz 2] M. Bazaraa and J. Goode, "Necessary optimality criteria in mathematical programming in normed linear spaces", J. Opt. Th. Appl., vol. 11, no. 3, 1973.

[Bert 1] D. Bertsekas and S. Shreve, Stochastic Optimal Control: The Discrete Time Case, Academic Press, N.Y., 1979.

[Bert 2] D. Bertsekas, Stochastic Optimal Control, Academic Press, N.Y., 1976.

[Bil 1] P. Billingsley, Convergence of Probability Measures, John Wiley and Sons, N.Y., 1968.

[Bo 1] N. Bourbaki, Espaces Vectoriels Topologique (Eléments de Mathematique, XVIII, Première Partie; Livre V), Hermann et Cie, Paris, 1953.

[Bou 1] D. Bourgin, "Linear topological spaces," Amer. Jour. Math., LXV, 1943.

[Ca 1] M. Canon, C. Cullum, Jr. and E. Polak, Theory of Optimal Control and Mathematical Programming, McGraw-Hill, N.Y., 1970.

[Ca 2] M. Canon, C. Cullum, Jr. and E. Polak, "Constrained minimization problems in finite dimensional spaces," SIAM J. Contr., vol. 4, 1966.

[Cas 1] D. Castañon, "Equilibria in stochastic dynamic games of Stakleberg type", Ph.D. dissertation, Electrical Engineering, M.I.T.

[Cra 1] B. Craven, "Nonlinear programming in locally convex spaces", J. Opt. Th. Appl., vol. 10, no. 4, 1972.

[DaC 1] N. DaCunha and E. Polak, "Constrained minimization under vector-valued criteria in linear topological spaces", in Mathematical Theory of Control, A. Balakrishnan and L. Neustedt, Eds., Academic Press, N.Y., 1967.

[Di 1] J. Dieudonné, "Sur le théorème de Hahn-Banach", Bull. Soc. Math. France, vo. 50, 1922.

[Do 1] J. Doob, Stochastic Processes, John Wiley and Sons, N.Y., 1953.

[Du 1] N. Dunford and J. Schwartz, Linear Operators Part I: General Theory, Interscience Publishers, N.Y., 1958.

[Dub 1] A. Dubovitskii and A. Milyutin, "Extremum problems with constraints", Soviet Math. Dokl., vol. 4, 1963.

[Dub 2] A. Dubovitskii and A. Milyutin, "Extremum problems in the presence of restrictions", USSR Comput. Math. and Math. Phys., vol. 5, no. 3, 1965.

[Gir 1] I. Girsanov, Lectures on Mathematical Theory of Extremum Problems, Springer-Verlag, N.Y., 1972.

[Hal 1] H. Halkin, "A maximum principle of the Pontryagin type for systems described by nonlinear difference equations", SIAM J. Contr., vol. 4, 1966.

[Hal 2] H. Halkin, "A satisfactory treatment of equality and operator constraints in the Dubovitskii-Milyutin optimization formalism", J. Opt. Th. Appl., vol. 6, 1970.

[Hal 3] H. Halkin, "Nonlinear nonconvex programming in an infinite dimensional space", in Mathematical Theory of Control, A. Balakrishnan and L. Neustadt, Eds., Academic Press, N.Y., 1967.

[Hal 4] P. Halmos, Measure Theory, Springer-Verlag, N.Y., 1970.

[Hal 5] P. Halmos, Lectures on Ergodic Theory, The Mathematical Society of Japan, 1956.

[Hol 1] J. Holtzman, "Convexity and the maximum principle for discrete systems", IEEE Trans. A.C., vol. AC-11, no. 1, 1966.

[Hin 1] K. Hinderer, Foundations of Non-stationary Dynamic Programming with Discrete Time Parameter, Springer-Verlag, N.Y., 1970.

[Ho 1] R. Holmes, Geometric Functional Analysis and Its Applications, Springer-Verlag, N.Y., 1975.

[Ho 2] Y. Ho and K. Chu, "Team decision theory and information structures in optimal control problems - part I", IEEE Trans. A.C., vol. AC-17, no. 1, 1972.

[Hur 1] L. Hurwicz, "Programming in linear spaces", in Studies in Linear and Nonlinear Programming by K. Arrow, L. Hurwicz and H. Uzawa, Stanford Univ. Press, Stanford, 1958.

[Io 1] A. Ioffe and V. Tihomirov, Theory of Extremal Problems, North-Holland Publishing Company, N.Y., 1979.

[Io 2] A. Ioffe, "Necessary and sufficient conditions for a local minimum, 1: a reduction theorem and first order conditions", SIAM J. Contr., vol. 17, no. 2, 1979.

[Ito 1] Y. Ito, "Necessary and sufficient conditions for an invariant measure", in Ergodic Theory, F. Wright, Ed., Academic Press, N.Y., 1963.

[Jo 1] B. Jordan and E. Polak, "Theory of a class of discrete optimal control systems", J. Elec. Contr., vol. 17, 1964.

[Kle 1] V. Klee, Jr., "Convex Sets in Linear Spaces", Duke Math. J., XVIII, 1951.

[Kur 1] S. Kurcyusz, "On the existence and nonexistence of Lagrange multipliers in Banach spaces", J. Opt. Th. Appl., vol. 20, no. 1, 1976.

[Kus 1] H. Kushner, Introduction to Stochastic Optimal Control, McGraw-Hill, N.Y., 1970.

[Las 1] I. Lasiecka, "Generalization of the Dubovitskii-Milyutin optimality conditions", J. Opt. Th. Appl., vol. 24, no. 3, 1978.

[Lea 1] J. Leake and J. Ortega, "A discrete-time maximum principle for systems with state constrained controls", SIAM J. Contr., vol., 15, no. 3., 1977.

[Lu 1] D. Luenberger, Optimization by Vector Space Methods, John Wiley and Sons, N.Y., 1969.

[Ly 1] L. Lyusternik and V. Sobolev, Elements of Functional Analysis, Frederick Ungar, 1961.

[Man 1] O. Mangasarian, Nonlinear Programming, McGraw-Hill, N.Y., 1969.

[Man 2] O. Mangarsarian and S. Fromovitz, "The Fritz John necessary optimality conditions in the presence of equality and inequality constraints", J. Math. Anal. Appl., vol. 17, 1967.

[Mau 1] H. Maurer and J. Zowe, "First and second-order necessary and sufficient optimality conditions for infinite diemensional programming problems", Appl. Math. Optim., vol. 5, 1979.

[Nag 1] Y. Nagahisa, and Y. Sakawa, "Nonlinear programming in Banach spaces", J. Opt. Th. Appl., vol. 4, no. 3, 1969.

[Neu 1] L. Neustadt, Optimization, Princeton University Press, Princeton, 1976.

[Nue 2] L. Neustadt, "An abstract variational theory with applications to a broad class of optimization problems I: General theory", SIAM J. Contr., vol. 4, 1966.

[Or 1] J. Ortega and W. Rheinboldt, Iterative Solution of Nonlinear Equations in Several Variables, Academic Press, 1970.

[Pol 1] E. Polak and A. Payne "On multicriterion optimization", in Directions in Large-Scale Systems, Y. Ho and S. Mitter, Eds., Plenum, N.Y., 1976.

[Ps 1] B. Pshenichnyi, Necessary Conditions for an Extremum, Marcel Dekker, N.Y., 1971.

[Qua 1] J. Quadrat, "Methodes de simulation en programmation dynamique stochastique", Docteur-Ingenieur These, Mathematique, Sciences de Paris VI.

[Ros 1] S. Ross, "Arbitrary state Markovian decision models", Annals Math. Stat., vol. 39, 1968.

[Roz 1] L. Rozonoér, "The maximum principle of L.S. Pontryagin in optimal systems theory III", Automat. Remote Contr. vol. 20, 1959.

[Rud 1] W. Rudin, Principles of Mathematical Analysis McGraw-Hill, N.Y., 1976.

[San 1] N. Sandell, Jr., "Control of finite-state, finite-memory stochastic systems", Ph.D. dissertation, Eelctrical Engineering, M.I.T.

[San 2] N. Sandell, Jr., P. Varaiya, M. Athans and M. Safonov, "Survey of decentralized control methods for large scale systems", IEEE Trans. A.C., vol. AC-23, no. 2, 1978.

[San 3] N. Sandell, Jr. and M. Athans, "Solution of some nonclassical LQG sotchastic decision problems", IEEE Trans. A.C., vol. AC-19, no. 2, 1974.

[Str 1] C. Striebel, Optimal Control of Discrete Time Stochastic Systems, Springer-Verlag, N.Y., 1975.

[Tu 1] H. Tuy, "On necessary conditions for optimality", in Colloquia Mathematica Societatis János Bolyai 12. Progress in Operations Research, Eger, 1974.

[Var 1] P. Varaiya, "Nonlinear Programming in Banach space", SIAM J. Appl. Math., vol. 15, no. 2, 1967.

[Wit 1] H. Witsenhausen, "Separation of estimation and control for discrete-time systems", Proc. IEEE, vol. 59, no. 9, 1971.

[Wit 2] H. Witsenhausen, "A standard form for sequential stochastic control", Math. Sys. Th., vol. 7, 1973.

[Yos 1] K. Yosida, Functional Analysis, Springer-Verlag, N.Y., 1971.

[Zio 1] S. Zionts and J. Wallenius, "An interactive programming method for solving the multiple criteria problem", Management Sci., vol. 22, 1976.

[Zo 1] J. Zowe and S. Kureyusz, "Regularity and stability for the mathematical programming problem in Banach spaces", J. Appl. Func. Anal., vol. 10, 1976.