N71-19121

March, 1970

Report ESL-R-421
M.I.T. DSR Projects
72013 and 76265
NSF Grant GK-14152
NASA Grant ~~NGL-22-009(124)~~
NGL-22-009-124

ON THE DESIGN OF OPTIMAL AND SUBOPTIMAL FEEDBACK SYSTEMS

Keith M. Joseph

Electronic Systems Laboratory

MASSACHUSETTS INSTITUTE OF TECHNOLOGY, CAMBRIDGE, MASSACHUSETTS 02139

Department of Electrical Engineering

March, 1970 Report ESL-R-421

ON THE DESIGN OF OPTIMAL AND SUBOPTIMAL FEEDBACK SYSTEMS

by

Keith M. Joseph

This report consists of the unaltered thesis of Keith Millican Joseph, submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Electrical Engineering. This work has been supported in part by research grants awarded to the Massachusetts Institute of Technology, Electronic Systems Laboratory from the National Science Foundation, Grants GK-2645 and GK-14125 under MIT DSR Projects 71139 and 72013, and in part by the National Aeronautics and Space Administration, Grant NGL-22-009(124), under MIT DSR Project 76265.

Electronic Systems Laboratory
Department of Electrical Engineering
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

ON THE DESIGN OF OPTIMAL AND SUBOPTIMAL FEEDBACK SYSTEMS

by

KEITH MILLICAN JOSEPH

B.S., Massachusetts Institute of Technology
(1964)

M.S., Massachusetts Institute of Technology
(1966)

SUBMITTED IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
February, 1970

Signature of Author Keith M. Joseph
Department of Electrical Engineering, Feb. 27, 1970

Certified by Michael Athans
Thesis Supervisor

Accepted by ______________________
Chairman, Departmental Committee on Graduate Students

ON THE DESIGN OF OPTIMAL AND SUBOPTIMAL FEEDBACK SYSTEMS

by

KEITH MILLICAN JOSEPH

Submitted to the Department of Electrical Engineering on February 27, 1970 in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

ABSTRACT

The feasibility of taking practical engineering constraints into consideration when designing optimal nonlinear feedback control systems is investigated. The constraint of interest is a structural one which guarantees that any solution obtained subject to this constraint will be in a form for which there is a simple, direct, and easy means of implementation. A sequence of structural forms (interconnections of single-input-single-output function generators and ideal summers) from which the feedback control law will be constructed is specified while leaving the various synthesis functions free to be determined optimally. In this manner a structurally constrained optimal control problem (SCOCP) is formulated; the thesis establishes that such problems can be solved and that their solution is computationally feasible. A variety of optimal and suboptimal design procedures with explicit computational algorithms are developed for solving the SCOCP and examples illustrating their application are presented.

In addition to developing procedures for solving the SCOCP, the thesis contains two major theoretical contributions. First, an expression relating the suboptimality of a control law (i.e., the deviation between $U_s(\underline{x})$ and $U^*(\underline{x})$) to the suboptimality of its cost (i.e., the deviation between $J_s(\underline{x})$ and $J^*(\underline{x})$) is derived. This expression is used to establish a bound on the suboptimal cost in terms of only the bound on the deviation between the suboptimal and optimal control. Second, a set of scalar stability bounds is derived which specify the range over which the scalar value of the control can vary at each point in the state space and still produce a system which is asymptotically stable. These bounds are used to establish certain stability properties of optimal nonlinear systems.

Thesis Supervisor: Michael Athans

Title: Associate Professor of Electrical Engineering

ACKNOWLEDGMENT

I wish to express my gratitude to Professor Michael Athans for the inspiration, encouragement, and helpful criticism which he provided during the course of this research. I consider myself very fortunate to have had such an outstanding thesis supervisor. Special thanks are also extended to the other members of my thesis committee, Professors Jan Willems and Ian Rhodes, for their many comments and suggestions. I should also like to thank Professors George C. Newton, Jr. and Leonard Gould for their interest and involvement in my graduate education.

Many of my colleagues in the Decision and Control Sciences Group of the Electronic Systems Laboratory have made direct and indirect contributions to this research. In particular, I should like to thank Professor A. H. Levis, Professor W. S. Levine, Professor T. E. Fortmann, Dr. S. G. Greenberg, Dr. E. Tse, Mr. J. H. Davis, and Mr. R. Moore. I am especially grateful to Dr. Atif Debs for his many helpful suggestions during the course of this research.

Mrs. Brenda Dunteman of the ESL Publications Staff is deserving of special acknowledgment for her excellence in typing this thesis. I should also like to thank Mr. Arthur Giordani and Mr. Norman Darling of the ESL Drafting Department.

This research was carried out in the Decision and Control Sciences Group of the MIT Electronic Systems Laboratory with support extended by the National Science Foundation under grants GK-2645 and GK-14152 and by the National Aeronautics and Space Administration under grant NGL-22-009 (124).

CONTENTS

LIST OF ILLUSTRATIONS

CHAPTER I

INTRODUCTION

1.1 Introduction

During the past decade optimal control theory has emerged as a new and sophisticated approach to control system analysis and design. A prime motivation for this new approach has been its potential in establishing a standard reference or index of performance, as well as providing a method for computing the "best" or "optimal" control with respect to the specified index for a given dynamical system. The literature[3,11] contains a large number of interesting examples to which the theory has been successfully applied.

Despite the fact that theoretical techniques for solving many types of optimal control problems exist, the application of the theory to the solution of problems of engineering significance and to the design of engineering systems has been rather slow in gaining acceptance. While this may be partly attributed to the mathematical complexity of the problem formulation and the sophistication required to utilize effectively the various computational techniques, a more fundamental cause is that the optimal feedback control law for many problems invariably turns out to be a nonlinear multivariable function — a solution which is not amenable to direct or simple means of implementation. This is in contradistinction to conventional servo theory which has found wide acceptance and utility because of the simplicity with which it can be directly implemented.

From an engineering viewpoint, of course, theory is only as good as the extent to which it can be applied; and mathematical "solutions" which are impossible or impractical to implement or approximate are simply not "solutions" at all. Therefore, the purpose of this thesis will be to attempt to bridge the gap between "theory" and "practice" by imposing additional constraints upon the optimization problem which will guarantee that any solution so obtained will be in a form for which there are simple, direct, and practical means of implementation.

This thesis will concentrate exclusively on the design of optimal and suboptimal feedback control systems. There are several reasons for our interest in feedback rather than open-loop control. First, feedback controllers have certain inherent advantages over open-loop controllers. The sensitivity to plant variations is reduced by using a feedback controller instead of a nominally equivalent open-loop one. Another advantage of feedback controllers occurs whenever the system is subject to noise, inexact initial condition data, or other disturbances. In such cases the feedback controller will correct for the trajectory deviations which result while the open-loop controller will not. Second, feedback controllers may be considerably more desirable from an implementation viewpoint. The possibility of constructing a feedback controller with a small number of inexpensive analog components is considerably more attractive than an open-loop implementation requiring a digital computer with considerable data storage capacity. Third, the problem of designing optimal feedback controllers subject to structural constraints is of considerable theoretical interest; the majority of the previous research has concentrated on solving the open-loop optimization problem and relatively little attention has been directed toward implementation. Therefore, this thesis will concentrate exclusively on the design of optimal and suboptimal feedback control systems.

1.2 Historical Perspective

Despite the extensive research which has been conducted on optimal control systems in the past decade, the vast majority of the effort has been concentrated on solving the optimization problem and little concern has been directed toward the problem of implementation. Nevertheless, a few significant results have been obtained. Considering the general time-optimal problem, Neustadt[20] has developed a method of combining a rather small amount of precomputed information with "on-line" computation to reduce significantly the complexity of the "on-line" algorithm. While this is limited to time-optimal problems with computer implementation, it does provide a practical implementation scheme for a large class of such problems.

Melsa and Schultz [21] have developed a method for the closed-loop, approximately time-optimal control of a class of linear systems with total effort constraint. The method, based on a special class of solutions of the Hamilton-Jacobi equation called eigenvector scalar products, requires that the controller-computer solve only algebratic equations — not two-point boundary value problems. However, the method provides no measure of the degree of suboptimality of the implementation scheme.

ReKasius [9] has developed a suboptimal design procedure in which the control is restricted to be a finite polynomial in the state variables. This procedure was formulated as solving a sequence of suboptimal problems which converge to the true optimal problem and has the advantage that it allows one to observe the increasing complexity of the control law at each stage. Thus the procedure may be terminated at any desired level of control law complexity. However, the resulting control is not optimal; in fact, it is not even the best suboptimal approximation at the given level of complexity.

Durbeck[6] has discussed an approximation method based on assuming a solution $V(\underline{x})$ to the Hamilton-Jacobi equation having a known form with unknown coefficients. The form selected constrains the resulting control to be a finite polynomial in the state variables. The coefficients of $V(\underline{x})$ are determined by a minimization procedure. However, as Merriam[24] has noted, accurately approximating $V(\underline{x})$ does not necessarily result in an accurate approximation to the optimal control which depends on the partial derivatives of $V(\underline{x})$ with respect to the state variables. In fact, such a procedure may give rise to limit cycles and other instability problems as well as to a general degradation in the performance of the system constructed from these approximations.

Smith[10] has presented a heuristic method for designing easily implemented quasi-optimal, minimum-time controllers for high-order dynamic systems. The bang-bang controller is obtained by least-squares fitting points on the optimal switching surface with an easily implemented, linear-segment switching surface. This method is attractive because least-squares approximation is an analytic procedure which is readily applicable to high-order dynamic systems. However, choosing the coefficients of a quasi-optimal surface to optimize a least-squares fitting criterion is not exactly equivalent to minimizing the response time. As a result limit cycles and other undesirable phenomena may result, and considerable engineering judgement must be exercised in the application of this approach.

Saridis and ReKasius[22] consider the interesting and practical problem of a linear plant with an infinite-time, integral-type, quadratic performance index and state constraints. Their synthesis scheme is based on properties of the optimal constrained trajectories and utilizes a dual-mode controller to switch between the "constrained" and "unconstrained" portion of the state space. One major advantage of their method is that the control law can be synthesized from analog devices — operational amplifiers, diodes, and resistors.

Longmuir and Bohn[23] generalize Smith's[10] approach and consider synthesizing general suboptimal feedback control laws. A structure for the suboptimal control is assumed which is a linear combination of suitably chosen basis functions of the state variables. The coefficient multipliers for these functions are determined by the minimization of a mean-square error using data obtained from numerically computed optimal trajectories. However there is no procedure for selecting the appropriate or optimal basis functions; hence, the resulting control will in general not be the best suboptimal approximation at the given level of complexity.

Debs[44] considers the case of linear time-invariant systems in which the control law is restricted to be linear. The cost functional (containing terms which are quadratic in the control and quadratic and higher-than-quadratic in the state) is averaged over a quadratic surface in the state space. A method for determining the linear control law which minimizes this "averaged" cost is developed. However, this method is only applicable to linear systems.

While there have been some scattered results on the implementation of various types of optimal and suboptimal controls (particularly time-optimal ones), there are few approaches that have broad applicability. In addition, most existing methods possess serious defects — the generation of limit cycles by the prescribed control law, lack of estimation of suboptimality, inability to guarantee that the prescribed control is the best at the given level of complexity, etc. Therefore, the author feels that additional research on the practical implementation of optimal and suboptimal control laws is needed.

1.3 Outline of the Thesis

The purpose of this thesis is to attempt to make optimization theory more relevant to the design of engineering systems. This will be accomplished by incorporating structural constraints directly into the optimization problem which will guarantee that any solution obtained subject to these constrains will belong to a specific class of feedback structures for which there is a simple and direct means of implementation. In order to definitize the problem and obtain meaningful results, we shall restrict our attention to a specific class of practical, easily implemented, feedback structures — interconnections of single-input-single-output (SISO) function generators and ideal summers. This class of structures is appealing from both a theoretical and practical viewpoint. The recent mathematical results of Kolmogorov,[25-26] Lorentz,[37-39] and Sprecher[28-32] have established that such SISO configurations are capable of representing any continuous function of n variables. The synthesis functions required by their representation techniques are continuous; unfortunately however, they are extremely "wiggly" (in fact, nowhere differentiable) and hence difficult to implement. Nevertheless, the fact that such structures are capable of representing all continuous multivariable functions with continuous but non-differentiable synthesis functions provides a strong indication that such structures with smooth synthesis functions would be able to represent, or accurately approximate, a wide class of multivariable functions. These SISO structures are also appealing from an implementation viewpoint since they can be easily constructed by either analog (diode-resistor networks) or digital (data point storage with computer interpolation) means. Hence it is felt that such structures provide an appropriate compromise between theoretical completeness (the ability to represent any continuous multivariable function) and structural simplicity.

The thesis will establish that such structurally constrained optimization problems can be solved and that their solution is computationally feasible. Both optimal and suboptimal techniques will be developed and several examples illustrating the theory will be presented. In addition, several fundamental properties of optimal systems and their relation to suboptimal ones will be established which not only provide a rigorous justification for suboptimal design but, in addition, provide explicit bounds for evaluating the performance of suboptimal systems.

The organization of the thesis is as follows: Chapter II has a twofold purpose; first, to formulate a precise mathematical statement of the optimization and stability problems which will be considered in this thesis; second, to provide an analysis of the specific problems and peculiarities which result from the imposition of structural constraints upon optimization problems.

In Chapter III we explore the relationship between optimal and suboptimal systems. First, an expression relating the suboptimality of a control law (i.e., the deviation between $U_s(\underline{x})$ and $U^*(\underline{x})$) to the suboptimality of its cost (i.e., the deviation between $J_s(\underline{x})$ and $J^*(\underline{x})$) is derived. This leads directly to a useful procedure for suboptimal design. Second, we derive a set of scalar stability bounds which specify at each and every point $\underline{x}$ the range over which the scalar control can vary and still produce a system which is asymptotically stable. These bounds are then used to establish certain stability properties of optimal systems.

Chapter IV has a threefold purpose: First, to specify the set of feedback structures (interconnections of SISO function generators and ideal summers) which will form the constraint set of allowed implementations; second, to outline and discuss the general approach that will be taken in solving the optimization and stability problems; and third, to develop a mathematical technique, gradient projection, in a form which will be required by the algorithms developed in the later chapters for solving structurally constrained optimization problems.

Chapter V concentrates on developing computational algorithms for suboptimal design. A variety of procedures (each based on the results of Chapter III) are developed and explicit computational algorithms are presented.

In Chapter VI we develop a gradient procedure for solving structurally constrained optimization problems. A state space gradient function which relates the change in the cost to the change in the scalar value of the feedback control law at each point is defined and an analytic formula for its evaluation derived. This function is used to develop an explicit computational algorithm of the gradient type for solving structurally constrained optimization problems.

Chapter VII is concerned with the stability problem. An explicit computational algorithm is developed for solving the stability problem and an example is presented to illustrate the application of the algorithm and the properties of the stability bounds.

The concluding Chapter VIII contains a summary of the results obtained and recommendations for future research.

1.4 Contribution of this Research

The contribution of this research is twofold. First, the theoretical results of Chapter III relating optimal and suboptimal systems are quite significant. They provide a rigorous justification for suboptimal design, form a conceptual basis for constructing several practical design algorithms, and provide a partial characterization of the properties of optimal systems. Second, the algorithms developed in Chapters V and VI demonstrate the feasibility of incorporating practical engineering constraints directly into the optimization problem.

CHAPTER II

THE STRUCTURALLY CONSTRAINED CONTROL PROBLEM

2.1 Introduction

This chapter has a twofold purpose: First, to formulate a precise mathematical statement of the optimization and stability problems which will be considered; second, to provide an analysis of the specific problems and peculiarities which result from the imposition of structural constraints upon optimization problems.

Section 2.2 is concerned with formulating the structurally constrained control problem. The class of dynamical systems and cost functionals which will be considered is specified, the structural constraints are characterized in a very general manner, and the structurally constrained optimal control problem (SCOCP) is formally stated. Then a stability problem is formulated whose solution can be used to analyze the stability and sensitivity of feedback control laws — particularly those designed with the SCOCP formulation.

In Section 2.3 we shall examine the problems one encounters in attempting to modify the conventional optimization techniques (the Maximum Principle, Dynamic Programming, and the Hamilton-Jacobi equation) so that they will be applicable to structurally constrained optimization problems. Three of the four Necessary Conditions given by the Maximum Principle have meaningful structurally constrained equivalents. However the fourth — the differential statement of the Principle of Optimality — cannot be formulated in a computationally feasible manner. Similarly, a partial differential equation of the "Hamilton-Jacobi type" (but without incorporating the Principle of Optimality) will be developed for the structurally constrained case. Because of the apparent impossibility of formulating a computationally feasible statement of the Principle of Optimality for problems subject to structural constraints, none of the conventional optimization techniques can be modified to solve such problems. Nevertheless, several of the equations which are developed in this section will be of importance in the procedures developed in the later chapters for solving structurally constrained optimization problems.

2.2 Formulation of the Control Problem

We wish to consider plants defined over a state space X which are nonlinear, controllable, time-invariant, n^{th} order dynamical systems with a single input u(t) such that the state variables $x_1(t), x_2(t), \ldots, x_n(t)$ satisfy the equations

$$\dot{x}_i = f_i[x_1(t), \ldots, x_n(t)] + b_i u(t) \quad ; \quad i = 1, \ldots, n \tag{2.2.1}$$

or, in vector form,

$$\dot{\underline{x}}(t) = \underline{f}[\underline{x}(t)] + \underline{b}\, u(t) \quad ; \quad \forall \underline{x} \in X \tag{2.2.2}$$

where $\underline{f} \in D^{(2)}$ and $\underline{f}(\underline{0}) = 0$. In addition, we shall require that each state variable be measurable with its instantaneous value available for control purposes. We wish to minimize cost functionals of the form

$$J = \int_0^\infty \left[g[\underline{x}(t)] + \frac{1}{2} u^2(t) \right] dt \tag{2.2.3}$$

where $g \in D^{(2)}$ is a convex, bounded, positive definite function with $g(\underline{0}) = 0$. The origin is the target set. The above problem formulation will be called a Standard Optimization Problem (SOP). The SOP formulation is quite general and can be used to model many practical control problems.

The first problem which should be considered is that of existence: Does a unique optimal feedback control law exist for every SOP? If the system were linear, the answer would be affirmative; this has been established by Lee and Marcus [41] [Theorem 11, page 220] and Bridgland [1] [Theorem 1, page 268]. However, no existence theorem is known to the author for the nonlinear case under consideration. Therefore, we shall assume the existence of a solution for all SOP problems under consideration while cautioning the reader that this assumption may not be valid for every SOP formulation.

Since the system dynamics and cost functional are time-invariant and the terminal time is infinite, the optimal feedback control law $U^*(\underline{x})$ will be a time-invariant function of the state only. Unfortunately (from an engineering viewpoint) this function $U^*(\underline{x})$ will in general be a nonlinear multivariable function — a solution which has no obvious, direct, or simple means of implementation.

In a few ground-based applications [42] involving low order dynamical systems, a "brute-force" implementation scheme has been used in which the scalar value of $U^*(\underline{x})$ is computed on a grid of points spanning the state space and then stored in a multi-dimensional array in a computer data bank. Then the computer is used to interpolate between these data points to compute the correct value of the control as the state of the dynamical system moves through the state space. However, this approach requires an expensive computer system with large data storage facilities and, hence, is of limited utility. Other schemes (such as approximating $U^*(\underline{x})$ by a truncated multi-dimensional power series expansion)[6, 9, 10, 23] have been attempted, but all have met with a rather notable lack of success. In all of these procedures the approach has been to first solve the unconstrained optimization problem to determine $U^*(\underline{x})$; and then approximate $U^*(\underline{x})$ by a suboptimal control $U_s(\underline{x})$ which can be implemented by some feasible synthesis structure.

This procedure of separating the "control problem" and the "implementation problem" is a rather drastic suboptimal procedure; <u>minimizing the cost functional subject to an implementation constraint is not equivalent to selecting the structurally allowable control which is the best mathematical fit to $U^*(\underline{x})$</u>. In fact, the cost corresponding to a control selected by this separation procedure can be significantly greater than that of the optimal structurally constrained control law. Therefore, in this thesis we shall seek to incorporate the structural constraints directly into the optimization problem and hence obtain a solution which is guaranteed to be

1) representable by a prespecified synthesis structure
2) the best or optimal solution for this level of structural complexity

In the interest of generality, we shall refrain at this point from explicitely defining or characterizing the type of structural constraints or implementation structures which will be considered. This will be done in Chapter IV. For the present we shall rather abstractly characterize a structure and its associated structural constraint by a set S composed of all the feedback control laws which the structure is capable of generating. We shall use the symbol S to refer to both the structure itself and its set of control laws. Thus, the restriction that a control law $U_s(\underline{x})$ be capable of being generated by a structure S can be expressed as $U_s(\underline{x}) \epsilon S$.

For the SOP optimization problem formulated at the beginning of this section, the unconstrained optimal feedback control law $U^*(\underline{x})$ is a function only of the instantaneous state of the system. However, when a structural constraint is imposed upon the optimization problem, the optimal solution subject to this constraint will in general depend on the initial condition $\underline{x}_o$ and must be denoted as $U_s^*(\underline{x}, \underline{x}_o)$. This is not really a "feedback" control. In order to eliminate the dependence on the initial condition and convert the constrained optimal solution to the form $U_s^*(\underline{x})$, it is necessary to modify the optimization problem somewhat. The change that is made is to minimize the cost functional in an "average" sense. (A similiar idea has been used in references 14, 43, and 44.) If we view the initial state $\underline{x}_o$ as a random variable distributed over an initial condition subset $Q_o \subset X$ with a probability distribution $Q_o(\underline{x})$, then the expectation $\langle J \rangle$ of the cost is simply

$$\langle J \rangle = \int_{Q_o} J(\underline{x})\, Q_o(\underline{x})\, d\underline{x} \tag{2.2.4}$$

This average cost $\langle J \rangle$ is now independent of the initial condition $\underline{x}_o$ and hence the control $U_s^*(\underline{x})$ which minimizes $\langle J \rangle$ must also be independent of $\underline{x}_o$.

We can now formally state the structurally constrained optimal control problem which will be considered in this thesis.

Definition 2.1: Structurally Constrained Optimal Control Problem (SCOCP)

Given:

1) A nonlinear, time-invariant, controllable, n^{th} order, dynamical system of the form

$$\dot{\underline{x}} = \underline{f}(\underline{x}) + \underline{b}\, U_s(\underline{x}) \qquad \forall \underline{x} \in X$$

for which each state variable is directly measurable and where $\underline{f} \in D^{(2)}$ with $\underline{f}(\underline{0}) = 0$.

2) A target set which is the origin.

3) An initial condition probability distribution $Q_o(\underline{x})$ defined on an initial condition subset $Q_o \subset X$.

4) A cost functional $\langle J \rangle$ of the form

$$\langle J \rangle = \int_{Q_o} J_s(\underline{x})\, Q_o(\underline{x})\, d\underline{x}$$

where

$$J_s(\underline{x}) = \int_0^{\infty} \{g[\underline{x}(t)] + \tfrac{1}{2} U_s^2[\underline{x}(t)]\} dt$$

where $g \in D^{(2)}$ is a convex, bounded, positive definite function with $g(\underline{0}) = 0$.

5) A feedback structure with its constraint set S of allowed feedback control laws $U_s(\underline{x})$.

Determine: The optimal structurally constrained feedback control law $U_s^*(\underline{x}) \in S$ which minimizes $\langle J \rangle$.

The SCOCP which has just been defined will be solved in Chapter VI for the class of structural constraints defined in Chapter IV. Several suboptimal techniques for its approximate solution will be developed in Chapter V.

We shall now formulate a stability problem which will be considered in conjunction with the SCOCP. There are two motivations for this. First, it will provide a means of establishing asymptotic stability for control systems designed with suboptimal techniques. Second, it will provide a means of analyzing a given control system to determine its sensitivity and degree of stability. This is of considerable engineering significance since there is certain to be a deviation between the designed and implemented control law. Knowledge of how large a deviation is tolerable and its effect on system performance would be most useful. This same information would indicate the extent to which noise, sensor errors, and other disturbances could be expected to affect system performance.

Before proceeding to formulate the stability problem, we need to precisely define "stability" and "implementation set". By stability we shall mean "asymptotic stability over a set Ω " as used by LaSalle[46].

Definition 2.2: A dynamical system for which any trajectory $\underline{x}(t)$ originating in a closed bounded set $\Omega \subset X$ containing the origin remains in Ω for all t and approaches the origin as $t \to \infty$ possesses asymptotic stability over the set Ω.

Definition 2.3: An implementation set I is a rectangular closed subset of the state space defined by

$$I = \{\underline{x} : \hat{I}_i \leq x_i \leq I_i \quad ; \quad i = 1, \ldots, n\} \qquad (2.2.5)$$

over which a feedback control law is defined and implemented.

The definition of I as a rectangular region was motivated by physical considerations. The allowed dynamical range of each state variable or its sensor is usually a finite interval which is independent of the values of the other state variables; therefore, the region in the state space over which a system can operate and its state be measured is usually rectangular.

To simplify notation, we make the following definition.

Definition 2.4: The symbol $\underset{\ominus}{>}$ is defined as meaning "greater than or equal to" if $\|\underline{x}\| = 0$ and "strictly greater than" if $\|\underline{x}\| \neq 0$.

Now we can state the stability problem.

Definition 2.5: Stability Problem

Given the SOP dynamical system $\dot{\underline{x}} = \underline{f}(\underline{x}) + \underline{b}\, U_s(\underline{x})$ defined on a state space X and an initial condition set $Q \subset X$, determine an implementation set $I \subset X$ along with a set of upper and lower bounds, $T(\underline{x})$ and $B(\underline{x})$, defined for all $\underline{x} \in I$ such that all controls $U_s(\underline{x})$ satisfying

$$T(\underline{x}) \underset{\ominus}{>} U_s(\underline{x}) \underset{\ominus}{>} B(\underline{x}) \qquad \forall \underline{x} \in I$$

will generate asymptotically stable trajectories $\underline{x}(t) \in I$ for all initial conditions $\underline{x}_o \in Q$.

In Chapter III we will establish that such stability bounds exist and can be computed with a relatively minimal amount of computation from a knowledge of only $U^*(\underline{x})$. An explicit computational algorithm for solving this problem will be developed in Chapter VII and its applications discussed.

2.3 Analysis of Conventional Optimization Techniques

As is well known,[13] the Maximum Principle states the following four Necessary Conditions for the SOP problem formulation defined in Section 2.2.

NC1) $$H[\underline{x}^*(t), \underline{p}^*(t), u^*(t)] = g[\underline{x}^*(t)] + \frac{1}{2}[u^*(t)]^2 + \left\langle \underline{p}^*(t), \{\underline{f}[\underline{x}^*(t)] + \underline{b}\,u^*(t)\}\right\rangle = 0 \tag{2.3.1}$$

NC2) $$\dot{\underline{x}}^*(t) = \frac{\partial H}{\partial \underline{p}}[\underline{x}^*(t), \underline{p}^*(t), u^*(t)] = \underline{f}[\underline{x}^*(t)] + \underline{b}\,u^*(t) \quad ; \quad \underline{x}^*(\underline{0}) = \underline{x}_o \tag{2.3.2}$$

NC3) $$\dot{\underline{p}}^*(t) = -\frac{\partial H}{\partial \underline{x}}[\underline{x}^*(t), \underline{p}^*(t), u^*(t)] = -\frac{\partial g}{\partial \underline{x}}[\underline{x}^*(t)] - \left(\frac{\partial \underline{f}}{\partial \underline{x}}[\underline{x}^*(t)]\right)^T \underline{p}^*(t) \tag{2.3.3}$$

NC4) $$H[\underline{x}^*(t), \underline{p}^*(t), u^*(t)] = \operatorname*{Min}_{u} \{H[\underline{x}^*(t), \underline{p}^*(t), u(t)]\} \tag{2.3.4}$$

The approach we shall take in attempting to derive a set of structurally constrained equivalents to the above set of Necessary Conditions will be to first formulate a "Hamilton-Jacobi type" partial differential equation in the state space and then specialize it to a single trajectory. This partial differential equation will not initially incorporate the Principle of Optimality and hence will be valid for both optimal and suboptimal feedback control laws. Then the problem of incorporating the Principle of Optimality will be investigated.

Although in general the Hamilton-Jacobi formulation is not equivalent to the Maximum Principle because of its additional differentiability requirements, the two formulations are equivalent for the class of problems considered in this thesis where all allowed feedback control laws are restricted to be not only continuous but infinitely differentiable. We shall consider the SOP dynamical system

$$\dot{\underline{x}}(t) = \underline{f}[\underline{x}(t)] + \underline{b}\, U[\underline{x}(t)] \qquad \forall \underline{x} \in X \tag{2.3.5}$$

being driven by an arbitrary feedback control law $U(\underline{x}) \in D^{(2)}$ for which the resulting dynamical system is asymptotically stable (i.e., $\underline{x}(\infty) = \underline{0}$). The cost functional is

$$J = \int_0^{\infty} \{g[\underline{x}(t)] + \tfrac{1}{2} U^2[\underline{x}(t)]\}dt \qquad \forall \underline{x} \in X \tag{2.3.6}$$

and the target set is the origin.

For the infinite time case under consideration, the cost will depend only on the initial state $\underline{x}_o$ and, of course, on $U(\underline{x})$. Let $J(\underline{x})$ denote the value of the cost at each point in the state space which results from the control $U(\underline{x})$. By this we mean that $J(\underline{x})$ is a scalar function of $\underline{x}$ whose numerical value at each point $\underline{x}$ is equal to the value of the cost functional defined by Equation (2.3.6) evaluated with $\underline{x}(t)$ being the solution of the differential equation defined by Equation (2.3.5) for which the point $\underline{x}$ is the initial condition. If $J(\underline{x})$ is a continuously differentiable function, we may deduce that along any system trajectory $\underline{x}(t)$

$$\frac{d}{dt}\left\{ J[\underline{x}(t)]\right\} = \left\langle \frac{\partial J[\underline{x}(t)]}{\partial \underline{x}}, \dot{\underline{x}}(t)\right\rangle = \left\langle \frac{\partial J[\underline{x}(t)]}{\partial \underline{x}}, \underline{f}[\underline{x}(t)] + \underline{b}\, U[\underline{x}(t)]\right\rangle \tag{2.3.7}$$

But from the cost functional it follows that

$$\frac{d}{dt}\left\{J[\underline{x}(t)]\right\} = -g[\underline{x}(t)] - \frac{1}{2}U^2[\underline{x}(t)] \qquad (2.3.8)$$

Combining these two equations gives

$$g(\underline{x}) + \frac{1}{2}U^2(\underline{x}) + \left\langle \frac{\partial J(\underline{x})}{\partial \underline{x}}, \underline{f}(\underline{x}) + \underline{b}\,U(\underline{x}) \right\rangle = 0 \quad \forall \underline{x} \in X \qquad (2.3.9)$$

The above equation must be valid at all points in the state space since some system trajectory must go through each point $\underline{x} \in X$. This partial differential equation is of the "Hamilton-Jacobi type" (without the Principle of Optimality) and will be of fundamental importance in Chapter III.

If we now select any point $\underline{x}_o \in X$ as an initial condition, define $\underline{x}(t)$ to be the solution of Equation (2.3.5) with $\underline{x}(0) = \underline{x}_o$, and define $\underline{p}(t)$ and $u(t)$ as

$$\underline{p}(t) = \frac{\partial J[\underline{x}(t)]}{\partial \underline{x}}, \quad u(t) = U[\underline{x}(t)] \qquad (2.3.10)$$

Equation (2.3.9) becomes

$$g[\underline{x}(t)] + \frac{1}{2}u^2(t) + \left\langle \underline{p}(t), \{\underline{f}[\underline{x}(t)] + \underline{b}\,u(t)\} \right\rangle = 0 \qquad (2.3.11)$$

Noting that this has precisely the same form as NC1, we define it to equal the constrained Hamiltonian $H_s(\underline{x}(t), \underline{p}(t), u(t))$ and state the first structurally constrained necessary condition:

SCNC1) $$H_s(\underline{x}(t), \underline{p}(t), u(t)) = g[\underline{x}(t)] + \frac{1}{2}u^2(t) + \left\langle \underline{p}(t), \{\underline{f}[\underline{x}(t)] + \underline{b}\,u(t)\} \right\rangle = 0 \qquad (2.3.12)$$

We can immediately mimic NC2 and state

SCNC2) $$\dot{\underline{x}}(t) = \frac{\partial H_s}{\partial \underline{p}}[\underline{x}(t), \underline{p}(t), u(t)] = \underline{f}[\underline{x}(t)] + \underline{b}\,u(t) \;;\; \underline{x}(0) = \underline{x}_o \qquad (2.3.13)$$

In SCNC3 given below we encounter the first deviation in form between the constrained and unconstrained necessary conditions. It can be derived most easily by differentiating Equation (2.3.9) with respect to $\underline{x}$ to obtain

SCNC3)

$$\dot{\underline{p}}(t) = -\left(\frac{\partial H_s}{\partial \underline{x}}[\underline{x}(t), \underline{p}(t), u(t)]\right) - \left(\frac{\partial H_s}{\partial u}[\underline{x}(t), \underline{p}(t), u(t)]\right)\left(\frac{\partial U[\underline{x}(t)]}{\partial \underline{x}}\right) \tag{2.3.14}$$

$$\dot{\underline{p}}(t) = -\left(\frac{\partial g}{\partial \underline{x}}[\underline{x}(t)]\right) - \left(\frac{\partial f}{\partial \underline{x}}[\underline{x}(t)]\right)^T \underline{p}(t) - \left(u(t) + \left\langle \underline{p}(t), \underline{b} \right\rangle\right)\left(\frac{\partial U[\underline{x}(t)]}{\partial \underline{x}}\right) \tag{2.3.15}$$

Equation (2.3.14) has been previously derived for the finite time case by Jacobson[45] [Equation 3.43, page 59]. This equation differs from that of NC3 in that an additional term

$$\left(u(t) + \left\langle \underline{p}(t), \underline{b} \right\rangle\right)\left(\frac{\partial U[\underline{x}(t)]}{\partial \underline{x}}\right) \tag{2.3.16}$$

is added to the right hand side. Note that in the special case in which $U(\underline{x}) = U^*(\underline{x})$, this added term is identically zero and SCNC3 and NC3 are identical. As Jacobson points out, the significance of this equation is that it provides a practical computational technique for propogating a <u>suboptimal</u> costate variable $\underline{p}(t)$ along a <u>suboptimal</u> trajectory generated by a <u>suboptimal feedback</u> control law. This will be required in Chapter VI where a gradient technique is developed for solving the constrained optimization problem.

The fourth necessary condition of the Maximum Principle is a differential statement of the Principle of Optimality. However, <u>when structural constraints are imposed upon an optimization problem, the</u>

Principle of Optimality cannot be formulated in a computationally feasible form. This type of constraint imposes a restrictive relationship between the values which the control may assume at any point and those it may assume at every other point. As a result if the control is specified at any point, this specification will restrict the values which it may assume at all other points; if the control is changed at any point, corresponding changes must occur at all other points. Therefore, the value of the optimal constrained cost at any point $\underline{x}$ will depend not only on $\underline{x}$ (and, of course, on the values of the control along the optimal trajectory from the point $\underline{x}$ to the target set), but also on the values which the control assumes at all points in the state space. It is precisely this dependence of the optimal cost at any point on the control at every point which prohibits the formulation of a meaningful statement of the Principle of Optimality for structurally constrained problems. Thus, there is no meaningful SCNC4, and we must reluctantly conclude that it is impossible to formulate a computationally feasible set of structurally constrained Necessary Conditions.

The negation of any meaningful statement of the Principle of Optimality by the imposition of structural constraints poses a rather challenging problem: How does one solve an optimization problem without using the Principle of Optimality? All of the major techniques (Maximum Principle, Hamilton-Jacobi Equation, and Dynamic Programming) are based upon it. Furthermore, the Principle of Optimality is so basic to these techniques and structural constraints seem to so thoroughly prohibit its application in a computationally feasible manner that the possibility of modifying these techniques sufficiently to overcome the difficulties seems remote. Therefore it becomes evident that alternative techniques must be developed.

While the preceeding analysis did not provide a method for solving the structurally constrained optimization problem, it did reveal the basic difficulties which arise from the imposition of structural constraints. These difficulties indicate some of the changes which must be made in formulating an effective solution to such problems. Specifically,

1) The optimality criterion must be global – not local. The conventional principle of optimality is a local condition valid at each and every point in the state space; at any given point it specifies a relationship between the values of certain functions at that point and is totally independent of their value elsewhere in the state space. However, the imposition of a structural constraint establishes a relationship between the values of the control at all points in the state space. This relationship prohibits the formulation of a statement of the Principle of Optimality in the conventional (local) sense. Hence, because of this global relationship on the values of the control, any meaningful optimality criterion must be global, not local.

2) The optimization problem must be formulated and solved in the state space – not in the time domain. The structural constraints are expressed in terms of functions defined over the state space and cannot be converted into the time domain

as constraints along a particular trajectory. A conversion to the time domain would become even more impossible for the problem of multiple trajectories resulting from initial condition probability distributions.

In this section we have examined the difficulties which result from the imposition of structural constraints on the feedback control law. Due to the impossibility of formulating a meaningful statement of the Principle of Optimality, none of the conventional methods (Maximum Principle, Hamilton-Jacobi Equation, or Dynamic Programming) can be used to solve such problems. Therefore, alternative techniques must be developed. Several such techniques — both optimal and suboptimal — will be developed in the following chapters. Several of the equations developed in this section, particularly Equation (2.3.9) and Equation (2.3.15), will be of great importance to these techniques.

CHAPTER III

PROPERTIES OF OPTIMAL AND SUBOPTIMAL SYSTEMS

3.1 Introduction

The purpose of this chapter is to investigate the relationship between optimal and suboptimal SOP systems. The major results are presented in Section 3.2 and their implications discussed.

Sections 3.3 and 3.4 investigate the relationship between the suboptimality of a control law (i.e., the deviation between $U_s(\underline{x})$ and $U^*(\underline{x})$) and the corresponding suboptimality of its cost (i.e., the deviation between $J_s(\underline{x})$ and $J^*(\underline{x})$). A simple mathematical expression relating the two is derived and used to establish bounds on the suboptimal cost. These bounds show that the fractional increase in the suboptimal cost is less than the square of the fractional deviation in the control. For example, a suboptimal control which is everywhere within 10% of the optimal will produce a suboptimal cost which is everywhere less that 1.25% above the optimal cost. This is a very important result; it may be viewed as a rigorous justification for using suboptimal design techniques — particularly those based on mathematically fitting $U_s(\underline{x})$ to $U^*(\underline{x})$.

Sections 3.5 through 3.8 investigate the stability properties of optimal and suboptimal systems. In Section 3.5 the existence of scalar stability bounds of the form

$$B(\underline{x}) \lesseqgtr U_s(\underline{x}) \lesseqgtr T(\underline{x})$$

is established. These bounds specify at each and every point in the state space a range over which the scalar control $U_s(\underline{x})$ can vary and still produce a system which is asymptotically stable. Section 3.6 discusses the various properties of these bounds. Section 3.7 establishes that all optimal SOP systems have certain stability properties in common. These are examined in Section 3.8; we demonstrate that in the special case of linear systems with quadratic criteria these stability properties are equivalent to those which can be deduced from Kalman's "solution to the inverse problem of optimal control".

3.2 Statement of Major Results

The major results of this chapter are contained in three basic theorems. Each will be stated and its implications discussed. The first, Theorem 3.1, relates suboptimal controls and suboptimal costs.

Theorem 3.1: Let $U^*(\underline{x})$ denote the optimal feedback control law and $J^*(\underline{x}) \in D^{(2)}$ the corresponding optimal cost of a SOP with dynamics $\dot{\underline{x}} = \underline{f}(\underline{x}) + \underline{b}\, u(t)$. Let $U_s(\underline{x})$ denote a feedback control law for the SOP system and $J_s(\underline{x}) \in D^{(2)}$ its corresponding cost. Then if $U_s(\underline{x})$ satisfies the inequality

$$|U^*(\underline{x}) - U_s(\underline{x})| \leq \gamma\, |U^*(\underline{x})| \quad ; \quad 0 \leq \gamma < \tfrac{1}{2} \tag{3.2.1}$$

or equivalently

$$(1-\gamma) \leq \left[\frac{U_s(\underline{x})}{U^*(\underline{x})}\right] \leq (1+\gamma) \quad ; \quad 0 \leq \gamma < \tfrac{1}{2} \tag{3.2.2}$$

at all points $\underline{x}$ along the suboptimal trajectory originating from any initial condition $\underline{x}_o$, then the suboptimal cost $J_s(\underline{x}_o)$ is bounded by

$$J_s(\underline{x}_o) \leq \left[1 + \left(\frac{\gamma^2}{1-2\gamma}\right)\right] J^*(\underline{x}_o) \quad ; \quad 0 \leq \gamma < \tfrac{1}{2} \tag{3.2.3}$$

Corollary 3.1: Given the same system, cost functional, and suboptimal control $U_s(\underline{x})$ of Theorem 3.1. Let $\Omega \subset X$ be a bounded closed set with the property that every solution of

$$\dot{\underline{x}} = \underline{f}(\underline{x}) + \underline{b}\, U_s(\underline{x}) \qquad \forall \underline{x} \in X \tag{3.2.4}$$

originating in Ω remains for all future time in Ω. If $U_s(\underline{x})$ satisfies the inequality of Equation (3.2.1) $\forall \underline{x} \in \Omega$, then the bound of Equation (3.2.3) holds $\forall \underline{x}_o \in \Omega$.

Theorem 3.1 and its associated corollary are very significant results; they may be viewed as a rigorous justification for using suboptimal controls and suboptimal design procedures. They establish that any control law $U_s(\underline{x})$ which is mathematically close to $U^*(\underline{x})$ will have a cost $J_s(\underline{x})$ which is close to $J^*(\underline{x})$; furthermore, they give an explicit bound on the suboptimal cost which depends only on the bound on the deviation between the optimal and suboptimal controls. For example, if a suboptimal control is everywhere within 10% of the optimal, its cost must be less than 1.25% above optimal. In addition, Equation (3.2.3) directly displays the functional relationship between the magnitude of the control law deviation and the corresponding cost increment. When deviations from optimality are small (i.e., γ near 0 such that $1 - 2\gamma \approx 1$), Equation (3.2.3) becomes

$$J_s(\underline{x}_o) - J^*(\underline{x}_o) \leq [\gamma^2]\, J^*(\underline{x}_o) \tag{3.2.5}$$

This states that small deviations from optimality will only induce second order increments in the cost — the fractional increase in the suboptimal cost is less than the square of the fractional deviation in the control.

Furthermore, since the integrand of the cost functional is positive definite and the term

$$\frac{\gamma^2}{1 - 2\gamma} \tag{3.2.6}$$

is a finite positive constant $\forall \gamma \in [0, \frac{1}{2})$, one can conclude that <u>any suboptimal control which satisfies (3.2.1) will produce a system which is asymptotically stable about the origin</u>.

Theorem 3.2 states a very simple but powerful result for establishing the stability of suboptimal control laws.

<u>Theorem 3.2</u>: Let $U^*(\underline{x})$ be the optimal feedback control law and $J^*(\underline{x}) \in D^{(2)}$ the optimal cost of any SOP with dynamics

$$\underline{\dot{x}} = \underline{f}(\underline{x}) + \underline{b}\, u(t)$$

for which $J^*(\underline{x}) \to \infty$ as $\|\underline{x}\| \to \infty$. Then the system

$$\underline{\dot{x}} = \underline{f}(\underline{x}) + \underline{b}\, U_s(\underline{x})$$

is asymptotically stable in the large if

i) $U_s(\underline{0}) = 0$

ii) $\mathrm{Sgn}[U_s(\underline{x})] = \mathrm{Sgn}[U^*(\underline{x})] \qquad \forall \{\underline{x}: U^*(\underline{x}) \neq 0\}$

iii) $|U_s(\underline{x})| \geq \frac{1}{2} |U^*(\underline{x})| \qquad \forall \{\underline{x}: U^*(\underline{x}) \neq 0\}$

One of the most attractive features of Theorem 3.2 is the simplicity with which it can be stated (i.e., "asymptotic stability if $U_s(\underline{0}) = 0$ and $U_s(\underline{x})$ has everywhere the same polarity and at least half the magnitude of $U^*(\underline{x})$"). This theorem clearly demonstrates the wide range over which the scalar value of a suboptimal control can vary and still produce an asymptotically stable system. In particular, it states that overdriving a system in the correct direction (i.e., having the same polarity but a much larger magnitude than U^*) will never endanger stability.

The major importance of Theorem 3.2, however, is not for establishing the stability of a particular system; stronger versions of this theorem (in which the bounds depend explicitely on $g(\underline{x})$) will be developed in Section 3.5 for this application. Rather, the importance of Theorem 3.2 is that since its bounds are a function of $U^*(\underline{x})$ only, they are valid for the entire class of all optimal SOP systems. Hence, this theorem may be used to partially characterize the properties of optimal SOP systems. This application is best illustrated by stating Theorem 3.3 which is a direct application of Theorem 3.2.

In preparation for stating Theorem 3.3, we make the following definition.

Definition 3.1: Let S_1 denote the SOP dynamical system defined by

$$\dot{\underline{x}} = \underline{f}(\underline{x}) + \underline{b}\, u(t)$$

$$y = -U^*(\underline{x})$$

where $U^*(\underline{x})$ is the optimal feedback control law for the above system for any SOP cost functional for which the optimal cost $J^*(\underline{x}) \in D^{(2)} \rightarrow \infty$ as $\|x\| \rightarrow \infty$.

Consider the system S_2 formed from S_1 as shown in Figure 3.1.

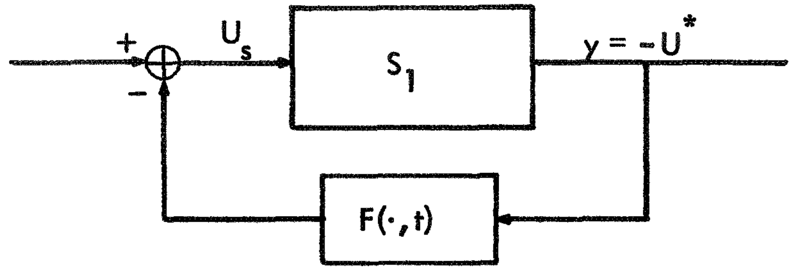

Figure 3.1 The System S_2

Note that if the nonlinear time-varying operator $F(\cdot, t)$ is a linear gain of unity, the system S_2 is merely the implementation of the optimal system S_1. We can now state Theorem 3.3.

Theorem 3.3: The system S_2 is asymptotically stable in the large if

1) $F(0, t) = 0 \qquad \forall t \in [0, \infty)$

2) $\frac{F(\sigma, t)}{\sigma} \geq \frac{1}{2} \qquad \forall t \in [0, \infty), \quad \sigma \neq 0$

If the conditions of Theorem 3.3 were both necessary and sufficient, they could be used to characterize optimal SOP systems and thus solve or at least state an equivalent criterion for the "inverse problem of optimal control". Unfortunately, although the conditions of the theorem are necessary, they are not sufficient. Nevertheless, the stability characterization contained in Theorem 3.3 is quite strong; in Section 3.8 we demonstrate that in the special case of linear systems with quadratic criteria the stability results of Theorem 3.3 are identical to those which can be deduced from Kalman's "solution to the inverse problem of optimal control". Thus, Theorem 3.3 can only serve to partially characterize the properties of optimal SOP systems; nevertheless, since this is an area in which very little is known, the theorem is of some significance.

3.3 Equivalence of Two Optimization Problems

The purpose of this section is to state, prove, and interpret a theorem which establishes the equivalence of two optimization problems. The proof of Theorem 3.1 and the suboptimal design procedures developed in Chapter V will be based on results derived from this theorem.

Theorem 3.4: Let $U^*(\underline{x})$ denote the optimal feedback control law and $J^*(\underline{x}) \in D^{(2)}$ the corresponding optimal cost of any SOP with dynamics

$$\dot{\underline{x}} = \underline{f}(\underline{x}) + \underline{b}\, U(\underline{x}) \qquad \forall \underline{x} \in X \tag{3.3.1}$$

and cost functional

$$J = \int_0^\infty [g(\underline{x}) + \tfrac{1}{2} U^2(\underline{x})]\, dt \tag{3.3.2}$$

Let S denote a set of feedback control laws $U_s(\underline{x})$ with corresponding costs $J_s(\underline{x}) \in D^{(2)}$ for which the dynamical system of Equation (3.3.1) is asymptotically stable. Consider the following two structurally constrained optimization problems.

Problem 1: Given the dynamical system of Equation (3.3.1), the cost functional of Equation (3.3.2), and an initial condition $\underline{x}_o \in X$. Determine the optimal control $U^*_{s_1}(\underline{x}) \in S$ which minimizes $J_s(\underline{x}_o)$.

Problem 2: Given the dynamical system of Equation (3.3.1), the cost functional

$$J_2 = \frac{1}{2} \int_0^{\infty} \left[U^*(\underline{x}) - U_s(\underline{x}) \right]^2 dt \qquad (3.3.3)$$

and the initial condition $\underline{x}_o \in X$. Determine the optimal control $U^*_{s_2}(\underline{x}) \in S$ which minimizes $J_2(\underline{x}_o)$.

These two optimization problems are equivalent in that the solution to the first, $U^*_{s_1}(\underline{x})$, is identical to the solution of the second, $U^*_{s_2}(\underline{x})$. For <u>any</u> control law $U_{s_1}(\underline{x}) \equiv U_{s_2}(\underline{x}) \in S$, the corresponding costs are related by

$$J_s(\underline{x}) \equiv J^*(\underline{x}) + J_2(\underline{x}) \qquad \forall\, \underline{x} \in X \qquad (3.3.4)$$

<u>Proof</u>: The proof proceeds as follows: First, a "Hamilton-Jacobi type" partial differential equation for the cost is formulated for both the unconstrained optimization problem and Problem 1. These two equations are combined to generate a third equation — a partial differential equation which is identical to that for the cost of Problem 2. By utilizing the specific properties of these equations, we complete the proof.

It was established in Chapter II, Equation (2.3.9), that the optimal control and cost of the SOP optimization problem must satisfy the partial differential equation

$$\left\langle \frac{\partial J^*(\underline{x})}{\partial \underline{x}}, [\underline{f}(\underline{x}) + \underline{b}\, U^*(\underline{x})] \right\rangle = -g(\underline{x}) - \frac{1}{2} U^{*2}(\underline{x}) \qquad \forall x \in X \qquad (3.3.5)$$

with the boundary condition

$$J^*(\underline{0}) = 0 \qquad (3.3.6)$$

where

$$U^*(\underline{x}) = -\left\langle \frac{\partial J^*(\underline{x})}{\partial \underline{x}}, \underline{b} \right\rangle \quad \forall \underline{x} \in X \tag{3.3.7}$$

Now we consider Problem 1. It was likewise established that the cost $J_s(\underline{x})$ corresponding to any feedback control law $U_{s_1}(\underline{x})$ for which the dynamical system of Equation (3.3.1) is asymptotically stable must satisfy the partial differential equation

$$\left\langle \frac{\partial J_s(\underline{x})}{\partial \underline{x}}, [\underline{f}(\underline{x}) + \underline{b}\, U_{s_1}(\underline{x})] \right\rangle = -g(\underline{x}) - \frac{1}{2} U_{s_1}^2(\underline{x}) \quad \forall \underline{x} \in X \tag{3.3.8}$$

with the boundary condition

$$J_s(\underline{0}) = 0 \tag{3.3.9}$$

The functions $J^*(\underline{x})$, $J_s(\underline{x})$, $U^*(\underline{x})$, and $U_{s_1}(\underline{x})$ are scalar functions of only the state $\underline{x}$ and are defined for all $\underline{x} \in X$. We now define two new functions, $U_o(\underline{x})$ and $J_o(\underline{x})$, as

$$U_{s_1}(\underline{x}) = U^*(\underline{x}) + U_o(\underline{x}) \qquad \forall \underline{x} \in X \tag{3.3.10}$$

$$J_s(\underline{x}) = J^*(\underline{x}) + J_o(\underline{x}) \qquad \forall \underline{x} \in X \tag{3.3.11}$$

Inserting these expressions for $U_{s_1}(\underline{x})$ and $J_s(\underline{x})$ into Equation (3.3.8) and subtracting Equation (3.3.5) gives

$$\left\langle \frac{\partial J_o(\underline{x})}{\partial \underline{x}}, \left[\underline{f}(\underline{x}) + \underline{b}\, \{U^*(\underline{x}) + U_o(\underline{x})\}\right] \right\rangle + U_o(\underline{x}) \left\langle \frac{\partial J^*(\underline{x})}{\partial \underline{x}}, \underline{b} \right\rangle$$

$$= -U^*(\underline{x})\, U_o(\underline{x}) - \frac{1}{2} U_o^2(\underline{x}) \qquad \forall \underline{x} \in X \tag{3.3.12}$$

Using Equation (3.3.7) we observe that the second term in the above expression is equal to $-U^*(\underline{x})\,U_o(\underline{x})$. Cancelling it with the corresponding term on the right hand side gives

$$\left\langle \frac{\partial J_o(\underline{x})}{\partial \underline{x}},\ \left[\underline{f}(\underline{x}) + \underline{b}\,\{U^*(\underline{x}) + U_o(\underline{x})\}\right] \right\rangle = -\frac{1}{2}\,U_o^2(\underline{x}) \quad \forall \underline{x} \in X \tag{3.3.13}$$

Using the definition given in Equation (3.3.10) we substitute

$$U_{s_1}(\underline{x}) = U^*(\underline{x}) + U_o(\underline{x}) \tag{3.3.14}$$

into the left hand side of Equation (3.3.13) and

$$U_o(\underline{x}) = U_{s_1}(\underline{x}) - U^*(\underline{x}) \tag{3.3.15}$$

into the right hand side to obtain

$$\left\langle \frac{\partial J_o(\underline{x})}{\partial \underline{x}},\ \left[\underline{f}(\underline{x}) + \underline{b}\,U_{s_1}(\underline{x})\right] \right\rangle = -\frac{1}{2}\left[U^*(\underline{x}) - U_{s_1}(\underline{x})\right]^2 \quad \forall \underline{x} \in X \tag{3.3.16}$$

It follows from Equations (3.3.6), (3.3.9), and (3.3.11) that the boundary condition for this partial differential equation is

$$J_o(\underline{0}) = 0 \tag{3.3.17}$$

Since

$$J_s(\underline{x}) = J^*(\underline{x}) + J_o(\underline{x}) \qquad \forall \underline{x} \in X \tag{3.3.18}$$

it follows that

$$\underset{U_{s_1} \in S}{\mathrm{Min}} \left[J_s(\underline{x}_o)\right] = \underset{U_{s_1} \in S}{\mathrm{Min}} \left[J^*(\underline{x}_o) + J_o(\underline{x}_o)\right] = J^*(\underline{x}_o) + \underset{U_{s_1} \in S}{\mathrm{Min}} \left[J_o(\underline{x}_o)\right]. \tag{3.3.19}$$

Thus the control law $U^*_{s_1}(\underline{x}) \in S$ which minimizes $J_s(\underline{x}_o)$ must also be the one which minimizes $J_o(\underline{x}_o)$.

Now we consider Problem 2. Equation (2.3.9) of Chapter II establishes that the partial differential equation for the cost of the dynamical system

$$\dot{\underline{x}} = \underline{f}(\underline{x}) + \underline{b}\, U_{s_2}(\underline{x}) \tag{3.3.20}$$

and cost functional

$$J_2 = \frac{1}{2} \int_0^{\infty} \left[U^*(\underline{x}) - U_{s_2}(\underline{x}) \right]^2 dt \tag{3.3.21}$$

is

$$\left\langle \frac{\partial J_2(\underline{x})}{\partial \underline{x}} , \left[\underline{f}(\underline{x}) + \underline{b}\, U_{s_2}(\underline{x}) \right] \right\rangle = -\frac{1}{2} \left[U^*(\underline{x}) - U_{s_2}(\underline{x}) \right]^2 \tag{3.3.22}$$

with the boundary condition

$$J_2(\underline{0}) = 0 \tag{3.3.23}$$

The above partial differential equation is identical to that of Equation (3.3.16) with $J_o(\underline{x})$ and $U_{s_1}(\underline{x})$ replaced by $J_2(\underline{x})$ and $U_{s_2}(\underline{x})$ respectively. Thus, if

$$U_{s_1}(\underline{x}) \equiv U_{s_2}(\underline{x}) \qquad \forall \underline{x} \in X \tag{3.3.24}$$

then

$$J_o(\underline{x}) \equiv J_2(\underline{x}) \qquad \forall \underline{x} \in X \tag{3.3.25}$$

Therefore, since $U_{s_1}(\underline{x})$ and $U_{s_2}(\underline{x})$ are restricted to belong to the same constraint set S, the control law $U^*_{s_2}(\underline{x})$ which minimizes $J_2(\underline{x}_o)$ must also be the one which minimizes $J_o(\underline{x}_o)$. Since we have previously established that this is $U^*_{s_1}(\underline{x})$, we can conclude

$$U^*_{s_1}(\underline{x}) \equiv U^*_{s_2}(\underline{x}) \qquad \forall \underline{x} \in X \tag{3.3.26}$$

Furthermore, by inserting Equation (3.3.25) into Equation (3.3.11) we can conclude that for any control law $U_{s_1}(\underline{x}) \equiv U_{s_2}(\underline{x}) \in S$, the costs corresponding to the two problems are related by

$$J_s(\underline{x}) \equiv J^*(\underline{x}) + J_2(\underline{x}) \qquad \forall \underline{x} \in X \tag{3.3.27}$$

In particular, for the optimal structurally constrained control laws $U^*_{s_1}(\underline{x}) \equiv U^*_{s_2}(\underline{x})$ and the initial condition $\underline{x}_o$

$$J^*_s(\underline{x}_o) \equiv J^*(\underline{x}_o) + J^*_2(\underline{x}_o) \tag{3.3.28}$$

Thus the theorem is proven.

Theorem 3.4 is a rather intriguing result with a number of interesting implications. Consider the problem of driving the dynamical system from some initial point $\underline{x}_o$ to the origin with a suboptimal control $U_s(\underline{x}) \in S$. <u>If the optimal control</u> $U^*(\underline{x})$ <u>is known at all points along the optimal trajectory</u> $\underline{x}^*(t)$, $J^*(\underline{x}_o)$ can be directly evaluated; similarly, if $U_s(\underline{x})$ is known along the suboptimal trajectory, $J_s(\underline{x}_o)$ may be directly evaluated. Then $J_2(\underline{x}_o)$, the increase in cost due to suboptimality, can be evaluated by

$$J_2(\underline{x}_o) = J_s(\underline{x}_o) - J^*(\underline{x}_o) \qquad \forall \underline{x}_o \in X \tag{3.3.29}$$

However, <u>Theorem 3.4 states that to evaluate</u> $J_2(\underline{x}_o)$ <u>you do not need to</u>

<u>know $U^*(\underline{x})$ along the optimal trajectory — in fact, you do not even need to know the optimal trajectory!</u> <u>Instead, $U^*(\underline{x})$ and $U_s(\underline{x})$ need only be known along the suboptimal trajectory</u>. Then $J_2(\underline{x}_o)$ may be evaluated from

$$J_2(\underline{x}_o) = \frac{1}{2} \int_0^{\infty} \left[U^*[\underline{x}(t)] - U_s[\underline{x}(t)] \right]^2 dt \tag{3.3.30}$$

where $\underline{x}(t)$ is the solution of

$$\dot{\underline{x}}(t) = \underline{f}[\underline{x}(t)] + \underline{b}\, U_s[\underline{x}(t)] \tag{3.3.31}$$

with $\underline{x}(0) = \underline{x}_o$; the $\underline{x}(t)$ used in Equation (3.3.30) is the suboptimal trajectory. Note that if $U^*(\underline{x}) \equiv U_s(\underline{x})$, $J_2(\underline{x}_o) = 0$ as it must.

3.4 <u>Proof of Theorem 3.1</u>

The proof of Theorem 3.1 is presented in this section. Before proceeding to construct the proof, we must first introduce a definition and prove a lemma.

As in Section 3.3 let

$$J^*(\underline{x}) = \text{Optimal Cost}$$

$$J_s(\underline{x}) = \text{Suboptimal Cost}$$

$$J_2(\underline{x}) = J_s(\underline{x}) - J^*(\underline{x})$$

We shall usually refer to $J_2(\underline{x})$ as the excess suboptimal cost. By definition

$$J_s = \int_0^{\infty} \left[g(\underline{x}) + \frac{1}{2} U_s^2(\underline{x}) \right] dt \tag{3.4.1}$$

and Theorem 3.4 establishes that if $J_s(\underline{x}) \in D^{(2)}$ and $J^*(\underline{x}) \in D^{(2)}$ then

$$J_2 = \frac{1}{2} \int_0^\infty \left[U^*(\underline{x}) - U_s(\underline{x}) \right]^2 dt \qquad (3.4.2)$$

where the $\underline{x}(t)$ used to evaluate both of the above integrals is that of the suboptimal trajectory generated by $U_s(\underline{x})$. Since

$$J^*(\underline{x}) = J_s(\underline{x}) - J_2(\underline{x}) \qquad \forall \underline{x} \in X \qquad (3.4.3)$$

it follows that

$$J^* = \int_0^\infty \left\{ g(\underline{x}) + \frac{1}{2} U_s^2(\underline{x}) - \frac{1}{2} \left[U^*(\underline{x}) - U_s(\underline{x}) \right]^2 \right\} dt \qquad (3.4.4)$$

where the above integral is likewise evaluated along the suboptimal trajectory.

We shall use the previous equations to define a rather interesting and useful scalar expression which provides a measure of the efficiency of the suboptimal control law at each point in the state space.

Definition 3.2: Let $E[\underline{x}, U_s(\underline{x})]$ denote the ratio of that portion of the suboptimal cost going into excess suboptimality, δJ_2, to that going into optimality, δJ^*, at each point in the state space. The analytic expression for $E[\underline{x}, U_s(\underline{x})]$

$$E[\underline{x}, U_s(\underline{x})] = \frac{\frac{1}{2}\left[U^*(\underline{x}) - U_s(\underline{x})\right]^2}{g(\underline{x}) + \frac{1}{2} U_s^2(\underline{x}) - \frac{1}{2}\left[U^*(\underline{x}) - U_s(\underline{x})\right]^2} \qquad (3.4.5)$$

is merely the ratio of the integrands of Equations (3.4.2) and (3.4.4).

Note that $E[\underline{x}, U_s(\underline{x})]$ will be zero at all points $\underline{x}$ where $U_s(\underline{x}) = U^*(\underline{x})$. The scalar value of $E[\underline{x}, U_s(\underline{x})]$ increases monotonically with the deviation between the scalar values of $U_s(\underline{x})$ and $U^*(\underline{x})$ over the range of values of $U_s(\underline{x})$ for which the denominator remains positive. The condition that the denominator be positive can be restated as

$$U_s(\underline{x})\, U^*(\underline{x}) > \frac{1}{2} U^{*2}(\underline{x}) - g(\underline{x}) \tag{3.4.6}$$

In Section 3.5 we will establish that all feedback control laws $U_s(\underline{x})$ which satisfy this inequality will produce asymptotically stable systems. The properties of such controls and their relation to $U^*(\underline{x})$ will be examined in Sections 3.6 through 3.8. There we shall demonstrate that the inequality is a very weak restriction and that virtually any $U_s(\underline{x})$ which may be said to approximate $U^*(\underline{x})$ will satisfy it. Since the numerator of $E[\underline{x}, U_s(\underline{x})]$ is positive semi-definite, the condition

$$E[\underline{x}, U_s(\underline{x})] \geq 0 \tag{3.4.7}$$

is equivalent to Equation (3.4.6). Hence any control $U_s(\underline{x})$ satisfying Equation (3.4.7) will produce an asymptotically stable system.

Now we state and prove Lemma 3.1.

Lemma 3.1: Let $U^*(\underline{x})$ denote the optimal feedback control law and $J^*(\underline{x}) \in D^{(2)}$ the corresponding optimal cost of a SOP. Let $U_s(\underline{x})$ denote a feedback control law for the SOP system and $J_s(\underline{x}) \in D^{(2)}$ its corresponding cost. If $U_s(\underline{x})$ is such that for some positive ϵ

$$0 \leq E[\underline{x}, U_s(\underline{x})] \leq \epsilon \tag{3.4.8}$$

at all points $\underline{x}$ along the suboptimal trajectory originating from any initial condition $\underline{x}_o$, then the suboptimal cost $J_s(\underline{x}_o)$ is bounded by

$$J_s(\underline{x}_o) \leq [1 + \epsilon]\ J^*(\underline{x}_o) \tag{3.4.9}$$

Proof: Since $E[\underline{x}, U_s(\underline{x})] \geq 0$, the denominator of $E[\underline{x}, U_s(\underline{x})]$ must be positive. Therefore,

$$\frac{1}{2}[U^*(\underline{x}) - U_s(\underline{x})]^2 \leq \epsilon \left\{ g(\underline{x}) + \frac{1}{2} U_s^2(\underline{x}) - \frac{1}{2}[U^*(\underline{x}) - U_s(\underline{x})]^2 \right\} \tag{3.4.10}$$

at all points $\underline{x}$ along the suboptimal trajectory. Since

$$J_2 = \frac{1}{2} \int_0^\infty [U^*(\underline{x}) - U_s(\underline{x})]^2\, dt \tag{3.4.11}$$

it follows that

$$J_2 \leq \epsilon \int_0^\infty \left\{ g(\underline{x}) + \frac{1}{2} U_s^2(\underline{x}) - \frac{1}{2}[U^*(\underline{x}) - U_s(\underline{x})]^2 \right\} dt \tag{3.4.12}$$

and thus from Equation (3.4.4) we obtain

$$J_2(\underline{x}_o) \leq \epsilon\ J^*(\underline{x}_o) \tag{3.4.13}$$

Therefore,

$$J_s(\underline{x}_o) \leq [1 + \epsilon]\ J^*(\underline{x}_o) \tag{3.4.14}$$

and the lemma is proven.

Now we are able to prove Theorem 3.1.

Proof of Theorem 3.1: If the suboptimal control $U_s(\underline{x})$ satisfies the inequality

$$|U^*(\underline{x}) - U_s(\underline{x})| \leq \gamma\ |U^*(\underline{x})| \quad ; \quad 0 \leq \gamma < \frac{1}{2} \tag{3.4.15}$$

or equivalently

$$(1-\gamma) \leq \left[\frac{U_s(\underline{x})}{U^*(\underline{x})}\right] \leq (1+\gamma) \quad ; \quad 0 \leq \gamma < \frac{1}{2} \tag{3.4.16}$$

at all points $\underline{x}$ along the suboptimal trajectory, it directly follows that the inequalities

$$\frac{1}{2}[U_s(\underline{x})]^2 \geq \frac{1}{2}(1-\gamma)^2[U^*(\underline{x})]^2 \tag{3.4.17}$$

and

$$\frac{1}{2}[U^*(\underline{x}) - U_s(\underline{x})]^2 \leq \frac{1}{2}\gamma^2[U^*(\underline{x})]^2 \tag{3.4.18}$$

or

$$-\frac{1}{2}[U^*(\underline{x}) - U_s(\underline{x})]^2 \geq -\frac{1}{2}\gamma^2[U^*(\underline{x})]^2 \tag{3.4.19}$$

are valid for all points $\underline{x}$ along the suboptimal trajectory. Adding Equations (3.4.17) and (3.4.19) gives

$$\frac{1}{2}U_s^2(\underline{x}) - \frac{1}{2}[U^*(\underline{x}) - U_s(\underline{x})]^2 \geq \frac{1}{2}(1-2\gamma)[U^*(\underline{x})]^2 \geq 0 \tag{3.4.20}$$

Since $g(\underline{x})$ is positive definite, it follows that

$$g(\underline{x}) + \frac{1}{2}U_s^2(\underline{x}) - \frac{1}{2}[U^*(\underline{x}) - U_s(\underline{x})]^2 \geq \frac{1}{2}(1-2\gamma)[U^*(\underline{x})]^2 \geq 0 \tag{3.4.21}$$

By using the inequalities in Equations (3.4.18) and (3.4.21), one can directly establish that

$$E[\underline{x}, U_s(\underline{x})] = \frac{\frac{1}{2}[U^*(\underline{x}) - U_s(\underline{x})]^2}{g(\underline{x}) + \frac{1}{2}U_s^2(\underline{x}) - \frac{1}{2}[U^*(\underline{x}) - U_s(\underline{x})]^2} \leq \frac{\gamma^2}{1-2\gamma} \tag{3.4.22}$$

for all points $\underline{x}$ along the suboptimal trajectory. Then, identifying ϵ as

$$\epsilon = \frac{\gamma^2}{1-2\gamma} \tag{3.4.23}$$

and using Lemma 3.1, it follows that

$$J_s(\underline{x}_o) \leq \left[1 + \left(\frac{\gamma^2}{1-2\gamma}\right)\right] J^*(\underline{x}_o) \quad ; \quad 0 \leq \gamma < \frac{1}{2} \tag{3.4.24}$$

and hence Theorem 3.1 is proven. Corollary 3.1 follows directly from Theorem 3.1 and the definition of Ω.

3.5 Derivation of Stability Bounds

In this section we shall establish the existence of a pair of scalar stability bounds, $T(\underline{x})$ and $B(\underline{x})$, for dynamical systems of the SOP type and prove the following two theorems.

Theorem 3.5: Let $U^*(\underline{x})$ be the optimal feedback control law and $J^*(\underline{x}) \in D^{(2)}$ the associated optimal cost of a SOP with dynamics $\dot{\underline{x}} = \underline{f}(\underline{x}) + \underline{b}\,u(t)$. Let $\Omega \subset X$ be a closed bounded set defined by $\Omega = \{\underline{x}\colon J^*(\underline{x}) \leq \gamma\}$, and let $T(\underline{x})$ and $B(\underline{x})$ be the set of stability bounds corresponding to $U^*(\underline{x})$. Then if

$$T(\underline{x}) \underset{\ominus}{\geq} U_s(\underline{x}) \underset{\ominus}{\geq} B(\underline{x}) \qquad \forall\, \underline{x} \in \Omega$$

the system $\dot{\underline{x}} = \underline{f}(\underline{x}) + \underline{b}\,U_s(\underline{x})$ is asymptotically stable over the set Ω.

Theorem 3.6: Let $U^*(\underline{x})$ be the optimal feedback control law and $J^*(\underline{x}) \in D^{(2)}$ the associated optimal cost of a SOP with dynamics $\dot{\underline{x}} = \underline{f}(\underline{x}) + \underline{b}\,u(t)$. Let $\hat{\Omega} \subset X$ be a closed bounded set with the property that every solution of $\dot{\underline{x}} = \underline{f}(\underline{x}) + \underline{b}\,U_s(\underline{x})$ originating in $\hat{\Omega}$ remains for all future time in $\hat{\Omega}$. Let $T(\underline{x})$ and $B(\underline{x})$ be the set of stability bounds corresponding to $U^*(\underline{x})$. Then if

$$T(\underline{x}) \underset{\ominus}{\geq} U_s(\underline{x}) \underset{\ominus}{\geq} B(\underline{x}) \qquad \forall\, \underline{x} \in \hat{\Omega}$$

the system $\dot{\underline{x}} = \underline{f}(\underline{x}) + \underline{b}\, U_s(\underline{x})$ is asymptotically stable over the set $\hat{\Omega}$.

These theorems will be used in Chapter VII to solve the stability problem formulated in Chapter II.

The proof of these two theorems is based on a Lyapunov argument and utilizes the following two theorems due to LaSalle[46].

Theorem 3.7: Consider the dynamical system $\dot{\underline{x}} = \underline{F}(\underline{x})$. Let $\Omega \subset X$ be a bounded closed set defined by $V(\underline{x}) \leq \gamma$, where $V(\underline{x})$ is a scalar function with continuous first partial derivatives for all $\underline{x} \in \Omega$ having the property that $V(\underline{x}) > 0$ and $\dot{V}(\underline{x}) < 0$ for all $\underline{x} \in \Omega$ except the origin where $V(\underline{0}) = 0$ and $\dot{V}(\underline{0}) = 0$. Then the system is asymptotically stable over the set Ω.

Theorem 3.8: Let $\hat{\Omega} \subset X$ be a bounded closed set with the property that every solution of $\dot{\underline{x}} = \underline{F}(\underline{x})$ starting in $\hat{\Omega}$ remains for all future time in $\hat{\Omega}$. Suppose there is a scalar function $V(\underline{x})$ with continuous first partial derivatives for all $\underline{x} \in \hat{\Omega}$ with the property that $V(\underline{x}) > 0$ and $\dot{V}(\underline{x}) < 0$ for all $\underline{x} \in \hat{\Omega}$ except the origin where $V(\underline{0}) = 0$ and $\dot{V}(\underline{0}) = 0$. Then the system is asymptotically stable over the set $\hat{\Omega}$.

We now proceed to develop an analytical expression for $T(\underline{x})$ and $B(\underline{x})$. The optimal cost $J^*(\underline{x})$ will be used as the Lyapunov function $V(\underline{x})$. This choice guarantees that $V(\underline{x}) > 0$ at all points $\underline{x}$ except the origin where $V(\underline{0}) = 0$. Theorems 3.7 and 3.8 both state that the dynamical system

$$\underline{\dot{x}} = \underline{f}(\underline{x}) + \underline{b}\, U_s(\underline{x}) \tag{3.5.1}$$

will be asymptotically stable over the specified domain if $\dot{V}(\underline{x}) < 0$ at all points $\underline{x}$ except the origin where $\dot{V}(\underline{0}) = 0$.

For the Lyapunov function $V(\underline{x}) = J^*(\underline{x})$, the expression for $\dot{V}(\underline{x})$ becomes

$$\dot{V}(\underline{x}) = \frac{d}{dt}[J^*(\underline{x})] = \left\langle \underline{\dot{x}}[\underline{x}, U_s(\underline{x})], \frac{\partial J^*(\underline{x})}{\partial \underline{x}} \right\rangle \tag{3.5.2}$$

Since in general

$$\underline{P}^*(\underline{x}) \triangleq \frac{\partial J^*(\underline{x})}{\partial \underline{x}}$$

is not zero at $\underline{x} = \underline{0}$, in order to insure that $\dot{V}(\underline{0}) = 0$ we must require

$$\underline{\dot{x}}[\underline{0}, U_s(\underline{0})] = \underline{0} = \underline{f}(\underline{0}) + \underline{b}\, U_s(\underline{0}) \tag{3.5.3}$$

but since

$$\underline{f}(\underline{0}) = \underline{0} \tag{3.5.4}$$

this condition becomes

$$U_s(\underline{0}) = 0 \tag{3.5.5}$$

Now for $\|\underline{x}\| > 0$ we must require

$$\dot{V}(\underline{x}) = \left\langle \underline{\dot{x}}[\underline{x}, U_s(\underline{x})], \underline{P}^*(\underline{x}) \right\rangle < 0 \tag{3.5.6}$$

and since

$$\underline{\dot{x}}[\underline{x}, U_s(\underline{x})] = \underline{f}(\underline{x}) + \underline{b}\, U_s(\underline{x}) \tag{3.5.7}$$

this condition becomes

$$\left\langle \underline{f}(\underline{x}), \underline{P}^*(\underline{x}) \right\rangle + U_s(\underline{x}) \left\langle \underline{b}, \underline{P}^*(\underline{x}) \right\rangle < 0 \tag{3.5.8}$$

or

$$U_s(\underline{x}) \left\langle \underline{b}, \underline{P}^*(\underline{x}) \right\rangle < -\left\langle \underline{f}(\underline{x}), \underline{P}^*(\underline{x}) \right\rangle \tag{3.5.9}$$

For the unconstrained optimization problem

$$H^*(\underline{x}, \underline{P}^*, U^*) = g(\underline{x}) + \frac{1}{2} U^{*2}(\underline{x}) + \left\langle \underline{P}^*(\underline{x}), \underline{f}(\underline{x}) \right\rangle + U^*(\underline{x}) \left\langle \underline{b}, \underline{P}^*(\underline{x}) \right\rangle = 0 \tag{3.5.10}$$

and from

$$\frac{\partial H^*}{\partial U^*} = 0 = U^*(\underline{x}) + \left\langle \underline{b}, \underline{P}^*(\underline{x}) \right\rangle \tag{3.5.11}$$

we find

$$U^*(\underline{x}) = -\left\langle \underline{b}, \underline{P}^*(\underline{x}) \right\rangle \tag{3.5.12}$$

Inserting this into Equation (3.5.9) gives

$$- U_s(\underline{x})\, U^*(\underline{x}) < -\left\langle \underline{f}(\underline{x}), \underline{P}^*(\underline{x}) \right\rangle \tag{3 5.13}$$

or

$$U_s(\underline{x})\, U^*(\underline{x}) > \left\langle \underline{f}(\underline{x}), \underline{P}^*(\underline{x}) \right\rangle \tag{3.5.14}$$

Observe that the direction of the inequality sign depends on whether $U^*(\underline{x})$ is positive or negative as shown by expressing Equation (3.5.14) as

$$\text{For } \{\underline{x}:\ U^*(\underline{x}) > 0\} \qquad U_s(\underline{x}) > + \frac{\left\langle \underline{f}(\underline{x}), \underline{P}^*(\underline{x}) \right\rangle}{|U^*(\underline{x})|} \tag{3.5.15}$$

$$\text{For } \{\underline{x}:\ U^*(\underline{x}) < 0\} \qquad U_s(\underline{x}) < - \frac{\left\langle \underline{f}(\underline{x}), \underline{P}^*(\underline{x}) \right\rangle}{|U^*(\underline{x})|} \tag{3.5.16}$$

This provides bounds on $U_s(\underline{x})$ except for the ambiguity occurring when $U^*(\underline{x}) = 0$, $\|\underline{x}\| \neq 0$. Recall that when $\|\underline{x}\| = 0$, $U_s(\underline{0}) = 0$. Equation (3.5.14) gives

$$0 > \left\langle \underline{f}(\underline{x}), \underline{P}^*(\underline{x}) \right\rangle \tag{3.5.17}$$

From Equation (3.5.10) we find that for $U^*(\underline{x}) = 0$

$$\left\langle \underline{f}(\underline{x}), \underline{P}^*(\underline{x}) \right\rangle = -g(\underline{x}) \tag{3.5.18}$$

and hence the above condition becomes

$$0 > -g(\underline{x}) \tag{3.5.19}$$

and since $g(\underline{x}) > 0$ for $\|\underline{x}\| \neq 0$, the above condition is satisfied; and being independent of $U_s(\underline{x})$, it is satisfied for any and all values of $U_s(\underline{x})$. Hence we conclude

$$\text{For } \{\underline{x}\colon U^*(\underline{x}) = 0\} \qquad \left\{ \begin{array}{ll} U_s(\underline{x}) = 0 & \|\underline{x}\| = 0 \\ & \\ -\infty < U_s(\underline{x}) < +\infty & \|\underline{x}\| \neq 0 \end{array} \right\} \tag{3.5.20}$$

The bounds which have just been established explicitly require the numerical value of the costate variable

$$\underline{P}^*(\underline{x}) = \frac{\partial J^*(\underline{x})}{\partial \underline{x}} \tag{3.5.21}$$

for all $\underline{x} \in X$. Since this might be difficult to compute and store, the bounds will be expressed in a format requiring only knowledge of $U^*(\underline{x})$. Inserting Equation (3.5.12) into Equation (3.5.10) gives

$$H^*(\underline{x}, \underline{P}^*, U^*) = g(\underline{x}) - \frac{1}{2} U^{*2}(\underline{x}) + \left\langle \underline{f}(\underline{x}), \underline{P}^*(\underline{x}) \right\rangle = 0 \tag{3.5.22}$$

from which we deduce

$$\left\langle \underline{f}(\underline{x}), \underline{P}^*(\underline{x}) \right\rangle = \frac{1}{2} U^{*2}(\underline{x}) - g(\underline{x}) \tag{3.5.23}$$

Using this substitution to eliminate the undesirable terms, the bounds become

$$
\begin{array}{ll}
\text{For} \quad \{\underline{x}:\ U^*(\underline{x}) > 0\} & U_s(\underline{x}) > \frac{1}{2}\,|U^*(\underline{x})| - \dfrac{g(\underline{x})}{|U^*(\underline{x})|} \\[2ex]
\text{For} \quad \{\underline{x}:\ U^*(\underline{x}) < 0\} & U_s(\underline{x}) < -\frac{1}{2}\,|U^*(\underline{x})| + \dfrac{g(\underline{x})}{|U^*(\underline{x})|} \\[2ex]
\text{For} \quad \{\underline{x}:\ U^*(\underline{x}) = 0\} & \left\{\begin{array}{ll} U_s(\underline{x}) = 0 & ; \quad \|\underline{x}\| = 0 \\ -\infty < U_s(\underline{x}) < +\infty & ; \quad \|\underline{x}\| \neq 0 \end{array}\right\}
\end{array}
\tag{3.5.24}
$$

From the structure of the above inequality constraints it is clear that one can define two bounding functions — an upper bound $T(\underline{x})$ and a lower bound $B(\underline{x})$ — such that the above conditions are equivalent to

$$
T(\underline{x}) \ \underset{\ominus}{\geq}\, U_s(\underline{x}) \ \underset{\ominus}{\geq}\, B(\underline{x}) \tag{3.5.25}
$$

These functions $T(\underline{x})$ and $B(\underline{x})$ are defined by Equations (3.5.26) and (3.5.27).

$T(\underline{x})$ Upper Bound

$$
\begin{array}{ll}
\text{For} \quad \{\underline{x}:\ U^*(\underline{x}) > 0\} & T(\underline{x}) = +\infty \\[2ex]
\text{For} \quad \{\underline{x}:\ U^*(\underline{x}) < 0\} & T(\underline{x}) = -\frac{1}{2}\,|U^*(\underline{x})| + \dfrac{g(\underline{x})}{|U^*(\underline{x})|} \\[2ex]
\text{For} \quad \{\underline{x}:\ U^*(\underline{x}) = 0\} & \left\{\begin{array}{ll} T(\underline{x}) = 0 & ; \quad \|\underline{x}\| = 0 \\ T(\underline{x}) = +\infty & ; \quad \|\underline{x}\| \neq 0 \end{array}\right\}
\end{array}
\tag{3.5.26}
$$

B($\underline{x}$) Lower Bound

$$
\begin{array}{ll}
\text{For } \{\underline{x}: U^*(\underline{x}) > 0\} & B(\underline{x}) = \frac{1}{2}\,|U^*(\underline{x})| - \dfrac{g(\underline{x})}{|U^*(\underline{x})|} \\[2ex]
\text{For } \{\underline{x}: U^*(\underline{x}) < 0\} & B(\underline{x}) = -\infty \\[2ex]
\text{For } \{\underline{x}: U^*(\underline{x}) = 0\} & \left\{\begin{array}{ll} B(\underline{x}) = 0 \quad ; & \|\underline{x}\| = 0 \\[2ex] B(\underline{x}) = -\infty \quad ; & \|\underline{x}\| \neq 0 \end{array}\right\}
\end{array}
\tag{3.5.27}
$$

Proof of Theorems 3.5 and 3.6: Since the choice $V(\underline{x}) = J^*(\underline{x})$ guarantees that $V(\underline{x}) > 0$ at all points $\underline{x}$ except the origin where $V(\underline{0}) = 0$ and since satisfaction of the stability bounds

$$T(\underline{x}) \;\underset{\ominus}{\geq}\; U_s(\underline{x}) \;\underset{\ominus}{\geq}\; B(\underline{x})$$

guarantees that over the specified domain $\dot{V}(\underline{x}) < 0$ at all points $\underline{x}$ except the origin where $\dot{V}(\underline{0}) = 0$, any system satisfying the conditions of Theorems 3.5 and 3.6 must satisfy the requirements of Theorems 3.7 and 3.8 and thus be asymptotically stable. Therefore, Theorems 3.5 and 3.6 are proven.

An example is presented in Section 7.3 which illustrates the application of these theorems to establish the stability of a typical control system.

3.6 Properties of the Stability Bounds

There are several important observations which can be made about these stability bounds.

1) Both $T(\underline{x})$ and $B(\underline{x})$ can be evaluated with a small amount of computation requiring only the numerical value of the optimal control $U^*(\underline{x})$.

2) The bounds do not even require $U_s(\underline{x})$ and $U^*(\underline{x})$ to have the same polarity at all points in the state space. (See the example in Section 7.3).

3) If $U^*(\underline{x})$ is continuous (as it almost certainly will be with the smoothness conditions of the SOP), then except for the origin, $T(\underline{x})$ is continuous for all $\{\underline{x}: T(\underline{x}) < \infty\}$ and $B(\underline{x})$ is continuous for all $\{\underline{x}: B(\underline{x}) > -\infty\}$.

4) For each point $\underline{x} \in X$ except the origin, the bounds are at most one-sided; either $T(\underline{x}) = +\infty$ or $B(\underline{x}) = -\infty$ or both $T(\underline{x}) = +\infty$ and $B(\underline{x}) = -\infty$. For $\{\underline{x}: U^*(\underline{x}) > 0\}$, $T(\underline{x}) = +\infty$ and $U_s(\underline{x})$ can be as positive as desired; for $\{\underline{x}: U^*(\underline{x}) < 0\}$, $B(\underline{x}) = -\infty$ and $U_s(\underline{x})$ can be as negative as desired; for $\{\underline{x}: U^*(\underline{x}) = 0\}$, $T(\underline{x}) = +\infty$ and $B(\underline{x}) = -\infty$ and $U_s(\underline{x})$ can have any finite value without endangering stability. Hence <u>one can conclude that overdriving the system in the correct direction</u> (i.e., <u>having the correct polarity for the control</u>) <u>will never endanger stability</u>; in fact, it would tend to enhance stability by moving the value of the control numerically further from the point of instability.

The phenomena of one-sided stability bounds is due exclusively to the control being linearly added to the system. The fact that both bounds are infinite when $U^*(\underline{x}) = 0$ results directly from the vector $\underline{b}$ being normal to the vector $\frac{\partial J^*(\underline{x})}{\partial \underline{x}}$ whenever $U^*(\underline{x}) = 0$; hence, $\frac{d}{dt}[J^*(\underline{x})] = -g(\underline{x})$ and is independent of the value of the control.

3.7 Proof of Theorems 3.2 and 3.3

In this section we shall prove Theorems 3.2 and 3.3. The proofs utilize the stability bounds developed in Section 3.5 and are based on the following theorem due to LaSalle.[46]

Theorem 3.9: Consider the dynamical system $\dot{\underline{x}} = \underline{F}(\underline{x})$. Let $V(\underline{x})$ be a scalar function with continuous first partial derivatives for all $\underline{x}$. Then if

i) $V(\underline{x}) > 0 \qquad \forall \underline{x} \neq \underline{0}$

ii) $\dot{V}(\underline{x}) < 0 \qquad \forall \underline{x} \neq \underline{0}$

$\dot{V}(\underline{0}) = 0$

iii) $V(\underline{x}) \to \infty$ as $\|\underline{x}\| \to \infty$

the dynamical system is ASIL (asymptotically stable in the large).

For convenience, both theorems will be restated; then their proofs given.

Theorem 3.2: Let $U^*(\underline{x})$ be the optimal feedback control law and $J^*(\underline{x}) \in D^{(2)}$ the optimal cost of any SOP with dynamics

$$\dot{\underline{x}} = \underline{f}(\underline{x}) + \underline{b}\, u(t)$$

for which $J^*(\underline{x}) \to \infty$ as $\|\underline{x}\| \to \infty$. Then the system

$$\underline{\dot{x}} = \underline{f}(\underline{x}) + \underline{b}\ U_s(\underline{x})$$

is asymptotically stable in the large if

i) $U_s(\underline{0}) = 0$

ii) $\mathrm{Sgn}\,[U_s(\underline{x})] = \mathrm{Sgn}\,[U^*(\underline{x})] \qquad \forall\{\underline{x}:\ U^*(\underline{x}) \neq 0\}$

iii) $|U_s(\underline{x})| \geq \frac{1}{2}|U^*(\underline{x})| \qquad \forall\{\underline{x}:\ U^*(\underline{x}) \neq 0\}$

Proof of Theorem 3.2: Let $V(\underline{x}) \equiv J^*(\underline{x})$. Then conditions i) and iii) of Theorem 3.9 are satisfied. Since $\underline{f}(\underline{0}) = 0$, requiring $U_s(\underline{0}) = 0$ insures that $\underline{\dot{x}}(\underline{0}) = 0$; hence, $\dot{V}(\underline{0}) = 0$. From the results of Section 3.5, Equation (3.5.24), it follows that $\dot{V}(\underline{x}) < 0$ for all $\{\underline{x}:\ \|\underline{x}\| \neq 0\}$ if

$$U_s(\underline{x}) > \frac{1}{2}|U^*(\underline{x})| - \frac{g(\underline{x})}{|U^*(\underline{x})|} \qquad \forall\{\underline{x}:\ U^*(\underline{x}) > 0\} \qquad (3.7.1)$$

$$U_s(\underline{x}) < -\frac{1}{2}|U^*(\underline{x})| + \frac{g(\underline{x})}{|U^*(\underline{x})|} \qquad \forall\{\underline{x}:\ U^*(\underline{x}) < 0\} \qquad (3.7.2)$$

Since $g(\underline{x})$ is positive definite, so is $\dfrac{g(\underline{x})}{|U^*(\underline{x})|}$. Hence, the conditions

$$U_s(\underline{x}) \geq \frac{1}{2}|U^*(\underline{x})| \qquad \forall\{\underline{x}:\ U^*(\underline{x}) > 0\} \qquad (3.7.3)$$

$$U_s(\underline{x}) \leq \frac{1}{2}|U^*(\underline{x})| \qquad \forall\{\underline{x}:\ U^*(\underline{x}) < 0\} \qquad (3.7.4)$$

are more restrictive than those of Equations (3.7.1) and (3.7.2); any $U_s(\underline{x})$ satisfying the latter set will guarantee that $\dot{V}(\underline{x}) < 0$ for all $\{\underline{x}:\ \|\underline{x}\| \neq 0\}$. Equations (3.7.3) and (3.7.4) are precisely equivalent to conditions ii) and iii) of Theorem 3.2. Thus, the three conditions of Theorem 3.2 guarantee that condition ii) of Theorem 3.9 is satisfied for

$V(\underline{x}) \equiv J^*(\underline{x})$, and the choice $V(\underline{x}) \equiv J^*(\underline{x})$ itself guarantees that conditions i) and iii) of Theorem 3.9 are satisfied. Therefore, the dynamical system must be ASIL; hence, Theorem 3.2 is proven.

The system S_2 was defined as being constructed from the optimal SOP system S_1 as shown in Figure 3.2

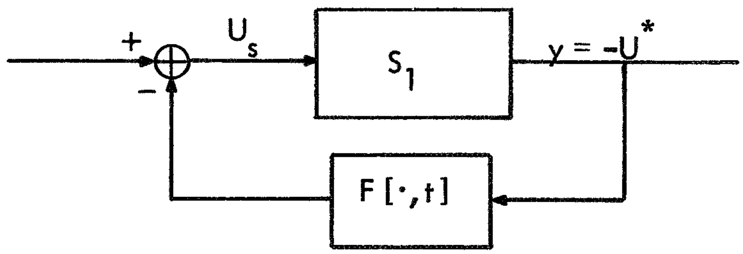

Figure 3.2 The System S_2

and Theorem 3.3 states

Theorem 3.3: The system S_2 is asymptotically stable in the large if

1) $F[0, t] = 0 \qquad \forall t \in [0, \infty)$

2) $\dfrac{F[\sigma, t]}{\sigma} \geq \dfrac{1}{2} \qquad \forall t \in [0, \infty), \; \sigma \neq 0$

Proof of Theorem 3.3: The proof of Theorem 3.3 follows directly from Theorem 3.2 and the definition of the system S_2. The feedback arrangement specifies that

$$U_s(\underline{x}) = - F[- U^*(\underline{x}), t] \tag{3.7.5}$$

For $\{\underline{x}: U^*(\underline{x}) = 0\}$: $F[0, t] = 0$; hence $U_s(\underline{x}) = 0$

For $\{\underline{x}: U^*(\underline{x}) > 0\}$: $F[-U^*(\underline{x}), t] \leq -\frac{1}{2}U^*(\underline{x})$; hence, $U_s(\underline{x}) \geq \frac{1}{2}U^*(\underline{x})$

For $\{\underline{x}: U^*(\underline{x}) < 0\}$: $F[-U^*(\underline{x}), t] \geq \frac{1}{2}U^*(\underline{x})$; hence, $U_s(\underline{x}) \leq \frac{1}{2}U^*(\underline{x})$

Since $U^*(\underline{0}) = 0$, the system S_2 subject to the restrictions of Theorem 3.3 satisfies all three conditions of Theorem 3.2. Hence, the system S_2 is ASIL and the theorem is proven.

The two main theorems of this section, Theorems 3.2 and 3.3, may be viewed from two perspectives. First, they may be reguarded as stability criteria which, although somewhat weaker than the criteria of Theorems 3.5 and 3.6, are nevertheless much simpler and easier to use. Theorems 3.2 and 3.3 clearly demonstrate the wide range over which the suboptimal control may vary and still produce a stable system. They establish the stability of a large class of suboptimal systems — particulary systems in which the suboptimal control is designed to approximate the optimal control.

From the second perspective, Theorems 3.2 and 3.3 may be viewed as stating a very important property of optimal SOP systems. Since these theorems apply to all optimal SOP systems but do not explicitly depend on either the system dynamics or cost functional, they are stated in a form which can be used to partially characterize the class of optimal SOP systems. This application will be considered in the following section.

3.8 Characterization of Optimal SOP Systems

In this section we seek to determine the extent to which Theorem 3.3 can be used to characterize optimal SOP systems. Ideally, we would like to establish that this theorem forms a necessary and sufficient condition for characterizing optimal SOP systems and thus solves or at least states an equivalent criterion for the "inverse problem of optimal control." Unfortunately, although the conditions of the theorem are necessary, they are not sufficient. Nevertheless, the stability characterization contained in Theorem 3.3 is quite strong; we shall demonstrate that in the special case of linear systems with quadratic criteria these stability properties are equivalent to those which can be deduced from Kalman's "solution to the inverse problem of optimal control".

First, we shall demonstrate that Theorems 3.2 and 3.3 are the strongest possible statements (of their type) which can be made; thus, they are the best necessary conditions for characterizing optimal SOP systems which can be obtained from the theory which has been developed. Conditions i) and ii) of Theorem 3.2 are clearly necessary and can't be changed; however, one might think that condition iii) could be strengthened by changing it to

$$|U_s(\underline{x})| \geq \gamma\,|U^*(\underline{x})| \qquad \forall \{\underline{x}:\ U^*(\underline{x}) \neq 0\}$$

where γ is some positive constant, $0 < \gamma < \frac{1}{2}$. The following counter-example will demonstrate that this is impossible and that the value of $\frac{1}{2}$ used in Theorems 3.2 and 3.3 is the least possible. Consider the scalar system

$$\dot{x}(t) = x(t) + u(t)$$

and cost functional

$$J = \frac{1}{2} \int_0^\infty [\alpha x^2 + u^2(t)]\, dt \quad ; \quad \alpha > 0 \tag{3.8.1}$$

for which the optimal feedback control is

$$U^*(\underline{x}) = -[1 + \sqrt{1+\alpha}\,]\, x \tag{3.8.2}$$

For very small values of α, $U^*(\underline{x})$ may be approximated as

$$U^*(\underline{x}) \cong -[2 + \frac{1}{2}\alpha]\, x \tag{3.8.3}$$

Now if we let

$$U_s(\underline{x}) = \gamma\, U^*(\underline{x}) = -\gamma\, [2 + \frac{1}{2}\alpha] x \tag{3.8.4}$$

the system dynamics become

$$\dot{x} = [1 - 2\gamma - \frac{\gamma}{2}\alpha] x \tag{3.8.5}$$

from which it is clear that the system will only be stable if

$$1 - 2\gamma - \frac{\gamma}{2}\alpha < 0 \tag{3.8.6}$$

In the limit of α approaching zero, this will require

$$\gamma \geq \frac{1}{2} \tag{3.8.7}$$

and, hence, condition (iii) cannot be improved. Thus, Theorems 3.2 and 3.3 cannot be improved.

Now that we have established that Theorem 3.3 is the strongest statement which can be made about the stability properties of optimal

SOP systems with the theory which has been developed, it is natural to ask: "How good is it? Does this theorem provide a complete characterization (i.e., both a necessary and sufficient condition) for optimal SOP systems, or is it only a necessary condition which provides a partial characterization of their properties?" Unfortunately, there are no results known to the author which specify the properties of optimal nonlinear systems of the SOP type; hence, a direct answer to these questions is not possible. However, Kalman[40] has completely characterized the class of optimal linear time-invariant systems with quadratic criteria. Therefore, we shall restrict our attention to this special case and compare the results of Theorem 3.3 with Kalman's "solution to the inverse problem of optimal control" and the stability properties which can be deduced from it.

We shall consider controllable, linear, time-invariant systems

$$\dot{\underline{x}} = \underline{A}\,\underline{x} + \underline{b}\,u(t) \tag{3.8.8}$$

with the cost functional

$$J = \frac{1}{2}\int_0^{\infty}\left[\left\langle \underline{x}(t), \underline{Q}\,\underline{x}(t)\right\rangle + u^2(t)\right] dt \tag{3.8.9}$$

where $\underline{Q}$ is a positive definite matrix. It is well known[13] that for this problem an optimal feedback control $U^*(\underline{x}) = -(\underline{k}^*)^T \underline{x}$ exists, is unique, and satisfies the two conditions required by Theorem 3.3 — $U^*(\underline{0}) = 0$ and $J^*(\underline{x}) \to \infty$ as $\|\underline{x}\| \to \infty$. Hence the results of Theorem 3.3 apply to every optimal linear system.

Let us consider the same system driven by any stable, observable feedback control law

$$U(\underline{x}) = -(\underline{k})^T \underline{x} \tag{3.8.10}$$

and define an output

$$y = -U(\underline{x}) = +(\underline{k})^T \underline{x} \tag{3.8.11}$$

and define the transfer function between the input and output as $S(j\omega)$

$$S(j\omega) = \frac{y(j\omega)}{u(j\omega)} = \underline{k}^T(\underline{I}j\omega - \underline{A})^{-1}\underline{b} \tag{3.8.12}$$

Then Kalman[40] states the following theorem.

Theorem 3.10: Given a linear, time-invariant, controllable plant with a stable, completely observable control law $U(\underline{x}) = -(\underline{k})^T \underline{x}$. Then $U(\underline{x}) = -(\underline{k})^T \underline{x}$ is an optimal control law corresponding to a cost functional of the type defined in Equation (3.8.9) if and only if

$$|1 + S(j\omega)|^2 > 1 \tag{3.8.13}$$

The above theorem is the classical "solution to the inverse problem of optimal control for linear systems."

Theorem 3.10 states that there is a portion of the complex plane into which the Nyquist plot of the transfer function $S(j\omega)$ of any optimal linear system may neither enter nor encircle. This excluded region, a circle of unit radius centered at -1, is shown in Figure 3.3. The restriction that the control law must be stable guarantees that $S(j\omega)$ can't encircle this region, and Equation (3.8.13) explicitly states that $S(j\omega)$ can't enter it. In addition, Theorem 3.10 states that every transfer function whose Nyquist plot neither encircles nor enters this region must be the transfer function of some optimal linear system. Thus, this simple frequency domain criterion provides an explicit necessary and sufficient condition for characterizing optimal linear systems.

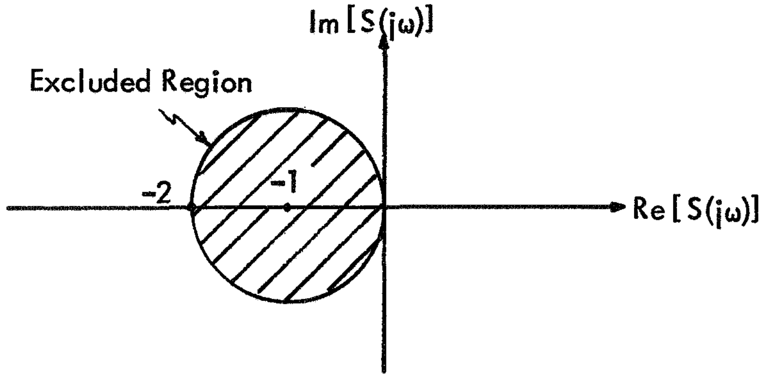

Figure 3.3 Excluded Nyquist Region

The stability properties of optimal linear systems can be determined from Theorem 3.10. Consider any transfer function $S(j\omega)$ which neither enters nor encircles the excluded region. This function $S(j\omega)$ must be the input-output transfer function of some optimal linear system and, hence, for some optimal SOP system S_1 (as defined in Definition 3.1). If the system S_2 is constructed from this system S_1 by adding the nonlinear time-varying operator $F(\cdot, t)$ as indicated in Figure 3.2, the Circle Theorem of I.W. Sandberg[50] states that S_2 is ASIL if

$$
\begin{aligned}
F[0,t] &= 0 && \forall t \in [0,\infty) \\
\frac{F[\sigma,t]}{\sigma} &\geq \frac{1}{2} && \forall t \in [0,\infty)\,,\ \sigma \neq 0
\end{aligned}
\tag{3.8.14}
$$

This is precisely the same result as stated by Theorem 3.3. Although Theorem 3.3 states nothing new about optimal <u>linear</u> systems, it does provide the strongest possible stability characterization which can be made. Hence, one might reasonably expect that the stability characterization for optimal nonlinear systems is also quite good.

The above results also demonstrate that Theorem 3.3 is not a sufficient condition for either the linear or general case. As is well known[50], the Circle Theorem is only a sufficient condition for stability— not a necessary and sufficient condition. Non-optimal linear systems exist which satisfy the conditions of Equation (3.8.14).

Therefore, we must conclude that Theorem 3.3 can only serve to partially characterize the properties of optimal SOP systems. Nevertheless, the stability characterization is quite good. Since very little is known about the properties of optimal nonlinear systems, this theorem is of some significance.

CHAPTER IV

THE FEEDBACK IMPLEMENTATION STRUCTURES

4.1 Introduction

We have thus far considered structural constraints of a completely general nature. For the SCOCP problem formulation to be meaningful, the set of feedback structures which will be used and its associated constraint set S must be specified. This will be done in this chapter.

In Section 4.2 we define and characterize the S(M, N, K) class of feedback structures which will form the constraint set of allowed implementations. These structures — interconnections of single-input-single-output (SISO) function generators and ideal summers — are simple and practical to implement, yet are capable of representing a wide class of control laws.

In Section 4.3 we outline and discuss the general approach that will be taken in solving the SCOCP problem formulation for the S(M, N, K) class of structural constraints.

In Section 4.4 we develop a modified version of the conventional gradient projection technique which allows the structural constraints of the S(M, N, K) structures to be incorporated directly into the optimization problem. This technique will be required by the design procedures developed in Chapters V and VI.

4.2 The S(M, N, K) Structures

We shall now define the class or set of feedback structures which will form the constraint set of allowable implementations. The structures in this set should be an appropriate compromise between structural simplicity and theoretical completeness — the ability to represent all optimal feedback control laws. Our choice of the particular set of structures which will be considered in this thesis was motivated by the recent mathematical results of Kolmogorov,[25-26] Lorentz,[37-39] and Sprecher.[28-32] They have established that any continuous function of n variables, $f(x_1, \ldots, x_n)$, can be represented as a superposition of functions of one variable as follows:

$$f(x_1, \ldots, x_n) = G\left[\sum_{i=1}^{n} \psi_i(x_i)\right] \tag{4.2.1}$$

where the functions $G(\cdot)$ and the $\psi_i(\cdot)$ are continuous. This rather remarkable result indicates that any continuous optimal feedback control law can be constructed by an appropriate interconnection of ideal summers and single-input-single-output (SISO) function generators. Since SISO function generators are easy and practical to construct, one would initially expect that this synthesis procedure would provide the solution to the implementation problem. Unfortunately, however, the synthesis functions required by the above representation technique are extremely "wiggly" (in fact, nowhere differentiable). The question of the applicability of the above representation scheme to the synthesis of optimal and suboptimal feedback control laws has been examined in detail by the author et al,[47] and the conclusion is that due to the poor analytic properties, this implementation procedure is useless as formulated. However, in the Kolmogorov-Lorentz-Sprecher formulation the functions $\psi_i(\cdot)$ were restricted to be topological transformations (i.e., the same set of functions $\psi_i(\cdot)$ is used for all $f(\underline{x})$) and only the function $G(\cdot)$ was allowed to depend on $f(\underline{x})$. It is this restriction which is primarily responsible for the poor analytic properties of the synthesis functions. If the restriction that the $\psi_i(\cdot)$ form a topological mapping is relaxed and these functions are allowed to depend on the particular function $f(\underline{x})$ which is to be represented, a substantial amount of additional flexibility will result and the new synthesis functions should be much smoother and hence much easier to approximate and implement. One can quickly verify that a large class of functions can be represented in this manner with smooth functions. In particular, sums, products, and exponentiations can be so generated. Thus it is felt that <u>since a particularly simple SISO structure is capable of representing all continuous multivariable functions under the rather severe requirement that the $\psi_i(\cdot)$ be a priori prespecified topological mappings, more general SISO structures utilizing only smooth synthesis functions all of which can depend on $f(\underline{x})$ should be able to represent or accurately approximate a wide class of optimal feedback control laws</u>. Hence, based upon the above considerations, general SISO structures with smooth synthesis functions will be selected as the members of our constraint set of allowable implementations.

Since SISO structures are to be used to implement the feedback control laws, one should at the outset define and classify all of the possible SISO structural forms. In addition, it would be particularly appealing if one could order them into <u>a sequence of monotonically increasing complexity</u>, for this would provide a direct and systematic means of comparing structural complexity with control law performance. Unfortunately, the generality and diversity of possible SISO structures are so great that this would not be feasible; however, a particular sequence of structural forms which recursively exploits the concept of "function-of-functions" can be defined. This concept was utilized by Kolmogorov, Lorentz, and Sprecher in their representations and the author feels that this type of structural arrangement can generate a wide variety of multivariable functions with a minimum number of SISO function generators. This sequence of structural forms will be denoted by S_1 and is defined below:

<u>Definition 4.1</u>: S_1 is the sequence of structural forms composed of members $S(M, N, K)$ which are defined by the following relation

$$U_s(\underline{x}) = \sum_{j=1}^{N} g_j \left[\sum_{i_1=1}^{n} \psi_{i_1}^{1} \left[\sum_{i_2=1}^{n} \psi_{i_1, i_2}^{2} \left[\cdots \left[\psi_{i_1, \ldots, i_M}^{M}(x_{i_M}) \right] \cdots \right] \right] \right] \tag{4.2.2}$$

for all points $\underline{x} \in I \subset X$. The definition of the implementation set I was given in Chapter II, Equation (2.2.5); n is the dimension of the state space. Each of the synthesis functions is restricted to be representable by a K^{th} order polynomial. To eliminate needless redundancy, only the function $g_1(\cdot)$ will possess a zeroth order coefficient. Thus, a typical synthesis function $\psi(s)$, $\psi \neq g_1$, will have the representation

$$\psi(s) = \sum_{k=1}^{K} A_k (s)^k \tag{4.2.3}$$

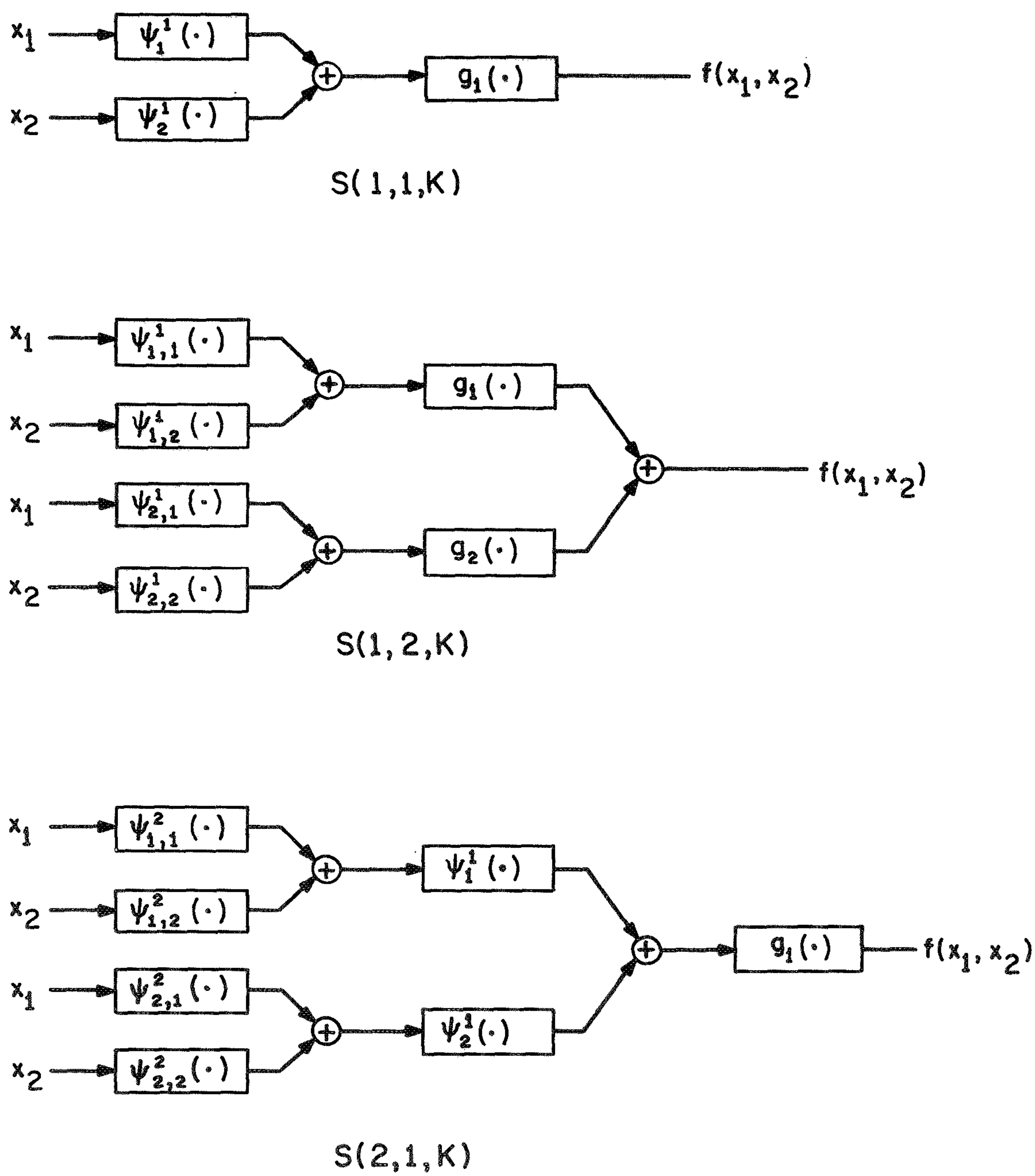

Figure 4.1 Block Diagram of Structural Forms for $n = 2$

and the function $g_1(s)$ will have the representation

$$g_1(s) = \sum_{k=0}^{K} B_k(s)^k \qquad (4.2.4)$$

The purpose of restricting the synthesis functions to be K^{th} order polynomials is to guarantee that they are "smooth". The structures S(1, 1, K), S(1, 2, K), and S(2, 1, K) are presented in block diagram form in Figure 4.1. As can be seen from the diagrams, these structures are in the form of trees which branch out and become sucessively more complex as M and N increase. N is the number of major branches of the tree structure [the number of $g(\cdot)$ functions] and M is the number of summing junctions encountered along any path from the inputs to the $g(\cdot)$ function. The total number A of SISO functions contained within the structure S(M, N, K) is given by the formula

$$A = \left[\frac{N}{n-1}\right]\left[n^{(M+1)} - 1\right] \qquad (4.2.5)$$

The index A serves as a means of ordering the various structures with respect to complexity. Alternatively, if one is considering a digital implementation, an index D equal to the total number of numerical constants (i.e., coefficients of the polynomials representing the synthesis functions) which must be stored is given by

$$D = \left[\frac{KN}{n-1}\right]\left[n^{(M+1)} - 1\right] + 1 \qquad (4.2.6)$$

Hence, with these indices A and D one can directly evaluate structural complexity and, presumably, the cost of implementation directly in terms of dollars, size, weight, etc.

4.3 The S(M, N, K) Design Procedure

In this section we wish to outline and discuss the general approach we shall take in solving the SCOCP problem formulation for the S(M, N, K) class of structural constraints. In formulating a design procedure which will be "meaningful" for physical problems, one should realize that any mathematical model is an abstraction of physical reality. In addition to merely "solving" a problem formulation for an "optimal" control law, one should critically evaluate both the model and its "optimal" solution to ascertain its validity and utility in solving the real physical problem. This can be accomplished by using both optimization and stability theory in the design process. Optimization theory should be used with a somewhat idealized model of the physical situation to determine an "optimal" solution with respect to a reasonable but somewhat arbitrary performance index. Then stability theory should be used to analyze this "optimal" solution with respect to the true physical situation, determine its applicability and limitations, and modify it as required. In this fashion one should be able to design practical control systems relevant to physical problems. We shall adopt this general approach in constructing our design procedure.

The SCOCP optimization problem for the S(M, N, K) class of structural constraints which we would like to be able to solve is:

> "Given a fixed number A of function generators determine both the optimal structural form $S(M^*, N^*, \cdot)$ composed of A function generators and its constituent functions which minimize the cost $\langle J \rangle$."

Unfortunately, the author has made no progress on solving this problem. The difficulty with this formulation is that it requires the optimization theory to determine not only the optimal synthesis functions but the optimal structural form as well. This is an extremely challenging problem which seems unlikely to be solved with currently available optimization techniques. Therefore, we must modify the problem somewhat. The problem which actually will be considered is:

> "Given a specific structure $S(M, N, K)$ composed of A function generators, determine the A optimal synthesis functions which minimize $\langle J \rangle$."

This formulation of the problem will be solved in Chapter VI and several suboptimal techniques for its approximate solution will be developed in Chapter V. The author envisions that one would initially start with the simpler structures [$S(1, 1, K)$, $S(1, 2, K)$, and $S(2, 1, K)$], solve the above problem for each, and systematically progress toward the more complex ones until a suitable control law and synthesis structure are found.

After obtaining what appears to be a suitable control law, the stability problem should be solved to analyze this solution. It is particularly important that this be done for structurally constrained problems which are designed by suboptimal procedures; control laws so designed may not be asymptotically stable over all regions of interest in the state space. In fact, it is even possible that the constraint set might not contain any asymptotically stable control laws. The solution of the stability problem provides a means of establishing the asymptotic stability of any proposed control law. More important, perhaps, the stability bounds provide a means of measuring the degree or extent of this stability. Specifically,

1) They indicate the general shape or type of feedback control laws which are stable and provide qualitative insight into the control problem.

2) They provide a means of specifying the required size of the implementation set I for a given dynamical system, control law, and initial condition set Q.

3) They indicate the appropriate direction (i.e., increasing or decreasing the scalar value of the control) which would be the safest direction to deviate from the optimal value. Since the stability bounds are one-sided, deviations in one direction will never threaten stability.

4) They provide both qualitative and quantitative information about the system stability. Extremely wide bounds indicate a very stable system and imply that state variable measurement errors, noise, or other disturbances would not pose serious problems. Narrow bounds would imply the opposite.

This type of information should allow one to determine the applicability and limitations of any given feedback control law.

Thus, by utilizing optimization and stability theory in the manner indicated, one should be able to design practical control systems relevant to physical problems.

4.4 Gradient Projection on S(M, N, K)

In this section we shall develop a modified version of the conventional gradient projection technique which allows the structural constraints of the S(M, N, K) structures to be incorporated directly into the optimization problem. Because of the apparent impossibility of utilizing any of the conventional optimization techniques for structurally constrained optimization problems, the procedures developed for their solution must be somewhat indirect — projected gradient techniques. These techniques first compute the unconstrained gradient function for a given control law. Then a gradient projection algorithm is used to project the unconstrained gradient function onto the synthesis structure such that all modifications of the given control law in the direction of the projected gradient will be representable by the given synthesis structure. By using this combination of conventional optimization theory in conjunction with gradient projection, constrained optimization problems can be solved in a relatively straightforward and computationally feasible manner.

The essential difference between the gradient projection problem we must consider and that of conventional gradient projection as formulated by Rosen[48-49] is the type of constraints employed. Rosen considers inequality constraints whereas we must consider a generalized type of equality constraint and insure that all control law variations remain on this constraint surface and hence representable by the S(M, N, K) structure.

First we wish to consider the formulation of a conventional gradient technique for determining the optimal unconstrained feedback control law $U^*(\underline{x})$ for a given optimization problem. All gradient techniques utilize a recursive algorithm in which a gradient function is computed for an existing control law and then form the new improved law by adding this gradient function times an appropriately selected scalar constant to the old law. Thus in solving for an optimal feedback control law the incrementing equation will be of the form

$$U_{\ell+1}(\underline{x}) = U_\ell(\underline{x}) + \alpha_\ell \hat{G}_\ell(\underline{x}) \qquad \forall \underline{x} \in X \ ; \ \ell = 0, 1, 2, \ldots \tag{4.4.1}$$

where $\hat{G}_\ell(\underline{x})$ is the gradient function for the control law $U_\ell(\underline{x})$ and α_ℓ is a scalar multiplier. This gradient function $\hat{G}_\ell(\underline{x})$ is a scalar function of $\underline{x}$ which indicates the rate of change of the cost $\langle J(U_\ell) \rangle$ with respect to a change in the scalar value of the control $U_\ell(\underline{x})$ at the point $\underline{x}$ with $U_\ell(\underline{x})$ held constant at all other points. Mathematically this could be stated as

$$\hat{G}_\ell(\underline{x}) = \frac{\partial \langle J(U_\ell) \rangle}{\partial U_\ell(\underline{x})} \qquad \forall x \in X \tag{4.4.2}$$

It directly follows from the above definition that the incremental change in the cost, $d\langle J(U_\ell) \rangle$, which results from an incremental change in the control, $dU_\ell(\underline{x})$, is given by

$$d\langle J(U_\ell) \rangle = \int_X \hat{G}_\ell(\underline{x})\, dU_\ell(\underline{x})\, d\underline{x} \tag{4.4.3}$$

For purposes of computation the state space X will be quantized and represented by a finite set of points $\{\underline{x}_i, i = 1, \ldots, r\}$. The set of integer subscripts $\{1, \ldots, r\}$ will be denoted by R. For this quantization the incrementing equation corresponding to Equation (4.4.1) becomes

$$U_{\ell+1}(\underline{x}_i) = U_\ell(\underline{x}_i) + \alpha_\ell G_\ell(\underline{x}_i) \qquad \forall i \in R \tag{4.4.4}$$

where the gradient function $G_\ell(\underline{x}_i)$ is defined as

$$G_\ell(\underline{x}_i) = \frac{\partial \langle J(U_\ell) \rangle}{\partial U_\ell(\underline{x}_i)} \qquad \forall i \in R \tag{4.4.5}$$

The set of quantization points $\{\underline{x}_i, i \in R\}$ are not arbitrary; they are selected such that the incremental change in the cost, $d\langle J(U_\ell)\rangle$, which results from an incremental change in the control, $dU_\ell(\underline{x})$, is given by

$$d\langle J(U_\ell)\rangle = \sum_{i=1}^{r} G_\ell(\underline{x}_i)\, dU_\ell(\underline{x}_i) \tag{4.4.6}$$

The existence of such gradient functions $G_\ell(\underline{x}_i)$ and corresponding quantization sets $\{\underline{x}_i, i \in R\}$ will be established in Chapter VI for problems of the SCOCP type in which there are no structural constraints. In the case of the suboptimal procedures of Chapter V, the cost is merely a weighted least-square-error criterion of the form

$$\langle J(U_\ell)\rangle = \sum_{i=1}^{r} W(\underline{x}_i)[U^*(\underline{x}_i) - U_\ell(\underline{x}_i)]^2 \tag{4.4.7}$$

for which the gradient function $G_\ell(\underline{x}_i)$ can be directly computed as

$$G_\ell(\underline{x}_i) = -2\, W(\underline{x}_i)\ [U^*(\underline{x}_i) - U_\ell(\underline{x}_i)] \qquad \forall i \in R \tag{4.4.8}$$

In all of these cases the gradient $G(\underline{x}_i)$ is the true, unconstrained gradient of the corresponding cost functional and the gradient algorithm can be expected to converge to the optimal solution which minimizes this cost functional.

If we now wish to impose structural constraints upon the problem and demand that the feedback controls $U_\ell(\underline{x})$ be restricted to belong to a particular structure $S(M, N, K)$, it is clear that the previously outlined gradient procedure must be modified. The required modification may be loosely stated as that of "projecting the gradient function $G_\ell(\underline{x}_i)$ onto the synthesis structure $S(M, N, K)$ such that all modifications of $U_\ell(\underline{x})$ in the direction of this projected gradient remain representable by $S(M, N, K)$."

The gradient function $G_\ell(\underline{x})$ indicates how the control $U_\ell(\underline{x})$ should be changed at each point $\underline{x}$; now we wish to determine how the control $U_\ell(\underline{x}) \in S(M, N, K)$ can be changed and still generate a $U_{\ell+1}(\underline{x}) \in S(M, N, K)$. Recall that each structure $S(M, N, K)$ has D coefficients which specify the control law. We shall assume that these are ordered and labeled $c_1, c_2, c_3, \ldots, c_D$. It will be convenient to think of this set of coefficients as specifing a D-dimensional Euclidean vector space C, and any specific set of values as specifing a point $\underline{c} \in C$. Since the scalar output of any structure $S(M, N, K)$ will depend upon both the state $\underline{x}$ and the coefficient vector $\underline{c}$, we shall denote the output of each structure as

$$U_s[\underline{x}, \underline{c}] \tag{4.4.9}$$

Since the coefficient vector $\underline{c}$ uniquely specifies the feedback control law, it is clear that any and all changes in the control law must be produced by changing $\underline{c}$. Hence the most general possible variation of $U_s[\underline{x}, \underline{c}] \in S(M, N, K)$ by a recursive, gradient-type algorithm would be one in which the recursion formula is of the form

$$U_s[\underline{x}_i, \underline{c}_{\ell+1}] = U_s[\underline{x}_i, \underline{c}_\ell + \alpha_\ell \underline{\delta c}_\ell] \quad \forall i \in R \tag{4.4.10}$$

where $\underline{\delta c}_\ell$ is a D component vector denoting the relative increment or change in each of the components of $\underline{c}$. Note that with this formulation all control laws $U_s[\underline{x}, \underline{c}_{\ell+1}]$ generated by any value of α_ℓ will be representable by S(M, N, K) — a fact which distinguishes this approach from conventional gradient projection.

We are now able to formulate the gradient projection problem. Note that the function $G(\underline{x}_i)$ has two basic properties

1) $G(\underline{x}_i)$ indicates the direction of steepest ascent (or descent) of the cost — i.e., $G(\underline{x}_i)$ indicates both the direction (positive or negative) and the relative amount that the control should be changed at each point $\underline{x}_i$ to most rapidly change $\langle J(U_\ell)\rangle$.

2) The magnitude of $G(\underline{x}_i)$ indicates the rate of change of the cost per unit change in the scalar value of the control at the point $\underline{x}_i$.

From the preceeding discussion it is clear that the gradient projection must convert the function $G(\underline{x}_i)$, $\underline{x}_i \in X$, into a vector $\underline{\delta c} \in C$. Since we wish this vector $\underline{\delta c}$ to be the projected equivalent of $G(\underline{x}_i)$, it is clear that $\underline{\delta c}$ should have the following two properties

1) At the point $\underline{c}_\ell \in C$, the vector $\underline{\delta c}_\ell$ must point in the direction of steepest ascent of the cost $\langle J(U_\ell)\rangle$.

2) The Euclidean length of $\underline{\delta c}_\ell$, $\|\underline{\delta c}_\ell\|$, must be equal to the rate of change of the cost $\langle J(U_\ell)\rangle$ per unit distance along the direction of the vector $\underline{\delta c}_\ell$ at the point $\underline{c}_\ell$.

These two conditions will uniquely specify $\underline{\delta c}_\ell$. The second is quite useful in that it directly relates changes in the coefficient space C to changes in the value of the cost. Note particularly that when a minimum is reached $\|\underline{\delta c}_\ell\| = 0$.

The process outlined above is equivalent to expressing the cost $\langle J(U_\ell)\rangle$ explicitely as a function of $\underline{c}$ and then directly evaluating $\underline{\delta c}$ by taking the gradient of $\langle J(\underline{c})\rangle$ as

$$\underline{\delta c} = \nabla[\langle J(\underline{c})\rangle] \tag{4.4.11}$$

Of course, the latter approach would be preferable; however expressing the cost as $\langle J(\underline{c})\rangle$ is completely impossible for the complex structures S(M, N, K). Hence we adopt the somewhat roundabout but computationally feasible process of first computing $G(\underline{x})$ and then projecting it on S(M, N, K) to obtain $\underline{\delta c}$.

We now precisely define the gradient projection problem.

<u>Definition 4.2</u>: <u>Gradient Projection on S(M, N, K)</u>

Given: A feedback control law $U_s[\underline{x}, \underline{c}_\ell] \in S(M, N, K)$ specified by a D component vector $\underline{c}_\ell$ and a scalar gradient function $G_\ell(\underline{x}_i)$ defined on the set $\{\underline{x}_i \,;\, i \in R\}$ for which the change in the cost, $d\langle J(U_\ell)\rangle$, which results from a change in the control law, $dU_\ell(\underline{x})$, is described by

$$d\langle J(U_\ell)\rangle = \sum_{i=1}^{r} G_\ell(\underline{x}_i)\, dU_\ell(\underline{x}_i) \tag{4.4.12}$$

Problem: Determine the vector $\underline{\delta c}_\ell$ which at the point $\underline{c}_\ell \in C$ points in the direction of the steepest ascent of the cost $\langle J(U_\ell(\underline{c}_\ell))\rangle$

and whose length, $\|\underline{\delta c}_\ell\|$, is equal to the rate of change of the cost per unit distance in this direction.

Before proceeding to solve the projection problem, we need to derive several preliminary results. First we need to express the change in the scalar value of the control at each point $\underline{x}_i$ in terms of the components of $\underline{\delta c}$. If we define $\delta U_s[\underline{x}_i, \underline{c}, \underline{\delta c}]$ as

$$\delta U_s[\underline{x}_i, \underline{c}, \underline{\delta c}] = U_s[\underline{x}_i, \underline{c} + \alpha \underline{\delta c}] - U_s[\underline{x}_i, \underline{c}] \tag{4.4.13}$$

then from the usual limiting process it follows that

$$\frac{dU_s[\underline{x}_i, \underline{c}, \underline{\delta c}]}{d\alpha} = \lim_{\alpha \to 0} \left[\frac{U_s[\underline{x}_i, \underline{c} + \alpha \underline{\delta c}] - U_s[\underline{x}_i, \underline{c}]}{\alpha} \right] \tag{4.4.14}$$

and that

$$\frac{dU_s[\underline{x}_i, \underline{c}, \underline{\delta c}]}{d\alpha} = \sum_{j=1}^{D} \left[\frac{\partial U_s[\underline{x}_i, \underline{c}]}{\partial c_j} \right] \delta c_j \tag{4.4.15}$$

Now if we define

$$D_j[\underline{x}_i, \underline{c}] = \frac{\partial U_s[\underline{x}_i, \underline{c}]}{\partial c_j} \tag{4.4.16}$$

we obtain

$$\frac{dU_s[\underline{x}_i, \underline{c}, \underline{\delta c}]}{d\alpha} = \sum_{j=1}^{D} D_j[\underline{x}_i, \underline{c}] \, \delta c_j \tag{4.4.17}$$

In formulating the above expression we have assumed the existence of the functions $D_j[\underline{x}_i, \underline{c}]$. This will now be established and explicit formulas for their computation developed.

From the definition of the S(M, N, K) structure, Equation (4.2.2), it is clear that the design of the structure was based on the "function-of-functions" concept. The basic building block of the structure is shown in Figure 4.2.

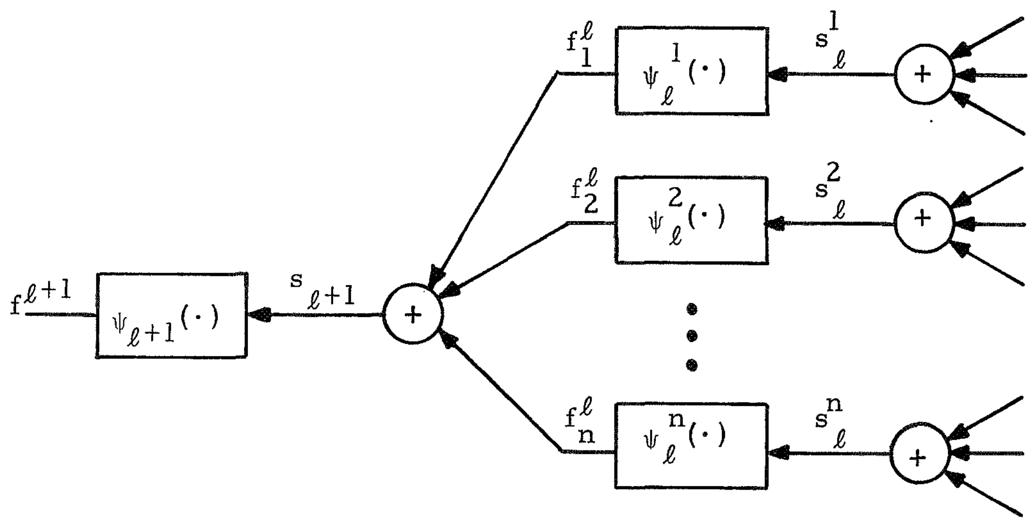

Figure 4.2 Building Block of S(M, N, K) Structure

For this type of structural formation the basic input-output relation of each function generator is

$$f^{\ell+1} = \psi_{\ell+1}[s_{\ell+1}] = \psi_{\ell+1}\left[\sum_{i=1}^{n} f_i^\ell\right] \tag{4.4.18}$$

Now we wish to consider some specific point $\underline{x}$ and some specific set of coefficients $\underline{c}$. For this $\underline{x}$ and $\underline{c}$, all the functions f^ℓ and s_ℓ will have a specific numerical value which we shall denote as $\overline{f^\ell}$ and $\overline{s_\ell}$. The differential relationship between the change $df^{\ell+1}$ in the scalar value of the output corresponding to a change df_i^ℓ in the scalar value of one of the inputs is given by

$$df^{\ell+1} = \left[\left(\frac{d\psi_{\ell+1}(s_{\ell+1})}{ds_{\ell+1}} \right)_{\overline{s_{\ell+1}}} \right] df_i^{\ell} \tag{4.4.19}$$

Now consider the general S(M, N, K) structure and any one of its adjustable coefficients, c_i. This coefficient c_i will belong to some synthesis function which we shall label ψ_o. In tracing the path from this function to the output, sequentially label each synthesis function encountered ψ_ℓ and its output f_ℓ ; $\ell = 1, \ldots, z$. Note that due to the tree structure of S(M, N, K), a change df_o in the scalar value of the output f_o of the function ψ_o will only induce a change in the inputs and outputs of those functions encountered along the path from ψ_o to the output. Thus repeated application of Equation (4.4.19) will give

$$d[U_s(\underline{x}, \underline{c})] = \left(\frac{d\psi_z(s_z)}{ds_z} \right)_{\overline{s_z}} \left(\frac{d\psi_{z-1}(s_{z-1})}{ds_{z-1}} \right)_{\overline{s_{z-1}}} \cdots \left(\frac{d\psi_1(s_1)}{ds_1} \right)_{\overline{s_1}} df_o \tag{4.4.20}$$

Now each of the synthesis functions ψ_ℓ is represented by a power series of the form

$$\psi_\ell(s_\ell) = \sum_{k=1}^{K} A_k^\ell (s_\ell)^k \tag{4.4.21}$$

Hence

$$\left(\frac{d\psi_\ell}{ds_\ell} \right)_{\overline{s_\ell}} = \sum_{k=1}^{K} k A_k^\ell (\overline{s_\ell})^{k-1} \tag{4.4.22}$$

The change df_o in the output of ψ_o will be produced by the change in the coefficient c_j which is one of the coefficients A_k^o in the power series representation for ψ_o. We shall denote it as $A^o_{k^*}$. From Equation (4.4.21) it directly follows that

$$df_o = (\overline{s_o})^{k^*} dc_j \tag{4.4.23}$$

Inserting Equations (4.4.22) and (4.4.23) into Equation (4.4.20) and recalling the definition of $D_j[\underline{x}_i, \underline{c}]$ gives

$$D_j[\underline{x}_i, \underline{c}] = \left[\sum_{k=1}^{K} k A^z_k (\overline{s_z})^k \right] \cdots \left[\sum_{k=1}^{K} k A^1_k (\overline{s_1})^k \right] (\overline{s_o})^{k^*} \tag{4.4.24}$$

Each of the coefficients A^ℓ_k ($\ell = 0, 1, \ldots, z$) in the above expression is some component of $\underline{c}$, and each of the terms $\overline{s_\ell}$ ($\ell = 0, 1, \ldots, z$) is a known function of $\underline{x}$ and $\underline{c}$ — each $\overline{s_\ell}$ is merely the input to one of the function generators in the control law $U_s[\underline{x}, \underline{c}]$. Thus the functions $D_j[\underline{x}_i, \underline{c}]$ can be directly computed for any control law $U_s[\underline{x}_i, \underline{c}] \in S(M, N, K)$. Note that for all finite $\underline{x}_i$ and $\underline{c}$, $D_j[\underline{x}_i, \underline{c}]$ will be finite. We summarize the preceeding results in the following theorem.

<u>Theorem 4.1</u>: For all structures $S(M, N, K)$ and any $U_s[\underline{x}, \underline{c}] \in S(M,N,K)$, the functions $D_j[\underline{x}, \underline{c}]$ defined as

$$D_j[\underline{x}, \underline{c}] = \frac{\partial U_s[\underline{x}, \underline{c}]}{\partial c_j} \tag{4.4.25}$$

exist, are finite, and can be directly computed with the formula

$$D_j[\underline{x}_i, \underline{c}] = \left[\sum_{k=1}^{K} k A^z_k (\overline{s_z})^k \right] \cdots \left[\sum_{k=1}^{K} k A^1_k (\overline{s_1})^k \right] (\overline{s_o})^{k^*} \tag{4.4.26}$$

where the terms involved have been previously defined.

Now we proceed to solve the projection problem. Equation (4.4.6) states that the change in the cost $d\langle J(U_\ell)\rangle$ which results from a change in the control $dU_\ell(\underline{x})$ is given by

$$d\langle J(U_\ell)\rangle = \sum_{i=1}^{r} G(\underline{x}_i)\, dU_\ell(\underline{x}_i) \tag{4.4.27}$$

Now we introduce the restriction that all variations in the control $dU_\ell(\underline{x})$ be such that $U_{\ell+1}(\underline{x}) \in S(M, N, K)$ for all values of α_ℓ. This requires

$$dU_\ell(\underline{x}_i) = dU_s[\underline{x}, \underline{c}, \underline{\delta c}] = \left[\sum_{j=1}^{D} D_j[\underline{x}_i, \underline{c}]\, \delta c_j\right] d\alpha_\ell \tag{4.4.28}$$

Therefore

$$d\langle J(U_\ell)\rangle = \sum_{i=1}^{r} G(\underline{x}_i) \left[\sum_{j=1}^{D} D_j[\underline{x}_i, \underline{c}]\, \delta c_j\right] d\alpha_\ell \tag{4.4.29}$$

The above equation directly relates the total change in the cost to the change in the coefficients of the S(M, N, K) structure. The change $d\langle J(U)\rangle$ in the cost, Equation (4.4.29), corresponds to a distance in C of $d\alpha \|\underline{\delta c}\|$, where the norm symbol will denote the usual Euclidean norm. Thus the first condition of the gradient projection problem — the requirement that $\underline{\delta c}$ point in the direction of steepest ascent — requires that $\underline{\delta c}$ be the vector which maximizes

$$\frac{d\langle J(U)\rangle}{d\alpha \|\underline{\delta c}\|} = \frac{1}{\|\underline{\delta c}\|} \left[\sum_{i=1}^{r} G(\underline{x}_i) \sum_{j=1}^{D} D_j[\underline{x}_i, \underline{c}]\, \delta c_j\right] \tag{4.4.30}$$

If we interchange summation signs in the above equation and define

$$F_j = \sum_{i=1}^{r} G(\underline{x}_i)\, D_j[\underline{x}_i, \underline{c}] \tag{4.4.31}$$

Equation (4.4.30) becomes

$$\frac{d\left\langle J(U_\ell)\right\rangle}{da\|\underline{\delta c}\|} = \frac{\sum_{j=1}^{D} F_j \delta c_j}{\|\underline{\delta c}\|} \tag{4.4.32}$$

Due to the linearity of $\underline{\delta c}$ in both the numerator and denominator of the above equation, the first condition will only specify $\underline{\delta c}$ to within a scalar multiple. The second condition will specify the magnitude of $\underline{\delta c}$. It requires that

$$\frac{d\left\langle J(U_\ell)\right\rangle}{da\|\underline{\delta c}\|} = \|\underline{\delta c}\| \tag{4.4.33}$$

Using Equation (4.4.32) the second condition becomes

$$\|\underline{\delta c}\|^2 = \sum_{j=1}^{D} F_j \delta c_j \tag{4.4.34}$$

Thus the two equations specifing $\underline{\delta c}$ are

$$\operatorname*{Max}_{\underline{\delta c} \in C} \left[\frac{\sum_{j=1}^{D} F_j \delta c_j}{\|\underline{\delta c}\|} \right] \tag{4.4.35}$$

subject to

$$\|\underline{\delta c}\|^2 = \sum_{j=1}^{D} F_j \delta c_j \tag{4.4.36}$$

The solution to these equations is rather easy and can be most easily established by defining θ as the angle between the vectors $\underline{F}$ and $\underline{\delta c}$ in the usual manner

$$\mathrm{Cos}(\theta) = \frac{\sum_{j=1}^{D} F_j \delta c_j}{\|\underline{F}\| \cdot \|\underline{\delta c}\|} \tag{4.4.37}$$

Then the first condition becomes

$$\mathrm{Max}\left[\frac{\|\underline{F}\| \cdot \|\underline{\delta c}\| \cdot \mathrm{Cos}(\theta)}{\|\underline{\delta c}\|}\right] = \mathrm{Max}\left[\|\underline{F}\| \cdot \mathrm{Cos}(\theta)\right] \tag{4.4.38}$$

with the obvious solution, $\mathrm{Cos}(\theta) = 1$. Hence we can conclude that $\underline{\delta c}$ and $\underline{F}$ are parallel and that

$$\underline{\delta c} = B\,\underline{F} \tag{4.4.39}$$

where B is an unknown scalar multiplier. The second condition requires

$$\|\underline{\delta c}\|^2 = \|\underline{\delta c}\| \cdot \|\underline{F}\| \cdot \mathrm{Cos}(\theta) = \|\underline{\delta c}\| \cdot \|\underline{F}\| \tag{4.4.40}$$

from which we conclude that $B = 1$ and hence

$$\underline{\delta c} = \underline{F} \tag{4.4.41}$$

The preceeding results are summarized in the following theorem.

Theorem 4.2: The solution of the problem of Gradient Projection on $S(M, N, K)$ is the vector $\underline{\delta c}$ which may be computed by the following formula

$$\delta c_j = \sum_{i=1}^{r} G(\underline{x}_i)\, D_j\,[\underline{x}_i, \underline{c}] \tag{4.4.42}$$

In this section we have developed a technique "gradient projection" which, when used in conjunction with the unconstrained gradient function $G(\underline{x})$, specifies the direction of steepest descent in the coefficient space

C for any control $U_s(\underline{x}, \underline{c}) \in S(M, N, K)$. In the following two chapters methods will be developed for computing $G(\underline{x})$ and explicit algorithms will be developed for solving the structurally constrained optimization problem. An example illustrating the application of these formulas in the solution of an optimization problem is presented in Section 6.5.

CHAPTER V

SUBOPTIMAL DESIGN TECHNIQUES

5.1 Introduction

The purpose of this chapter is to develop suboptimal design procedures for solving the SCOCP. All of the design techniques are based on appropriately "fitting" the suboptimal control $U_s(\underline{x}) \in S(M, N, K)$ to the unconstrained optimal feedback control law $U^*(\underline{x})$ so as to minimize some suboptimal design criterion. Hence, <u>all of the computational algorithms will require the numerical values of</u> $U^*(\underline{x})$ <u>as a function of the state</u> (i.e., <u>a dynamic programming table of the values of</u> $U^*(\underline{x})$).

Three suboptimal design criteria are proposed in Section 5.2 and the advantages of each are discussed. The first suboptimal design procedure (which can be used with any of the three criteria) is formulated in Section 5.3 and an explicit computational algorithm for its application is developed. The characteristics of this algorithm and the effectiveness of the three design criteria are discussed in Section 5.4.

A second suboptimal design procedure is formulated in Section 5.5. Although somewhat more complex and computationally demanding than the first, it is applicable to a wider variety of problems and produces a better suboptimal design. Its characteristics are discussed in Section 5.6.

5.2 Criteria For Suboptimal Design

The results of Chapter III provide a basis for developing a number of suboptimal design procedures for solving the SCOCP. In this section we shall consider three possible suboptimal design criteria; the following section will develop an explicit computational algorithm which can be used with any of the three. Then in Section 5.4 we will discuss the type of results one can expect from each of the criteria.

The following definition will be required.

Definition 5.1: Let $\Omega_o \subset X$ be the closed bounded set of minimal size for which the solution of $\dot{\underline{x}} = \underline{f}(\underline{x}) + \underline{b}\, U^*(\underline{x})$ for every initial condition $\underline{x}_o \in Q_o$ remains in Ω_o for all time.

The set Ω_o is merely that region of the state space spanned by all optimal trajectories originating in Q_o ; it can usually be determined by generating a few trajectories starting from critical points on the boundary of Q_o and utilizing the fact that trajectories can never intersect.

The results of Chapter III indicate that any of the following procedures should provide a good suboptimal design technique:

Criterion 1: Find the $U_s(\underline{x}) \in S(M, N, K)$ which minimizes $\langle \bar{J}_1 \rangle$

$$\langle \bar{J}_1 \rangle = \int_{\Omega_o} [U^*(\underline{x}) - U_s(\underline{x})]^2 \, d\underline{x}$$

The utilization of this least-square-error mathematical fitting criterion could be justified on the basis that such a criterion will generate a suboptimal control which tends to be close to the optimal everywhere; we have previously established in Theorem 3.1 that all such controls will generate costs which are likewise close to the optimal cost.

Criterion 2: Find the $U_s(\underline{x}) \in S(M, N, K)$ which minimizes $\langle \bar{J}_2 \rangle$

$$\langle \bar{J}_2 \rangle = \int_{\Omega_o} E^2[\underline{x}, U_s(\underline{x})]\, d\underline{x}$$

where

$$E[\underline{x}, U_s(\underline{x})] = \frac{\frac{1}{2}[U^*(\underline{x}) - U_s(\underline{x})]^2}{g(\underline{x}) + \frac{1}{2} U_s^2(\underline{x}) - \frac{1}{2}[U^*(\underline{x}) - U_s(\underline{x})]^2}$$

as defined in Section 3.4.

This criterion can be justified by observing that if

$$\epsilon = \underset{\underline{x} \in \Omega_o}{\text{Max}} \left[E[\underline{x}, U_s(\underline{x})] \right]$$

then from the results of Section 3.4 it follows that

$$J_s(\underline{x}) - J^*(\underline{x}) \leq \epsilon J^*(\underline{x}) \qquad \forall \underline{x} \in \Omega_o$$

The above criterion will guarantee that $E[\underline{x}, U_s(\underline{x})]$ never exceeds its average value by a large amount; hence, ϵ will tend to be as small as possible and the suboptimal cost should be close to the optimal everywhere.

Criterion 3: Find the $U_s(\underline{x}) \in S(M, N, K)$ which minimizes $\langle \bar{J}_3 \rangle$

$$\langle \bar{J}_3 \rangle = \int_{\Omega_o} \left[\| \underline{f}(\underline{x}) + \underline{b}\, U_s(\underline{x}) \| \right]^{-1} \left[U^*(\underline{x}) - U_s(\underline{x}) \right]^2 d\underline{x}$$

This criterion can be justified as follows: Theorem 3.4 establishes that minimizing the excess suboptimality $J_2(\underline{x}) = J_s(\underline{x}) - J^*(\underline{x})$ is equivalent to minimizing $J_s(\underline{x})$. This theorem further establishes that the cost functional for $J_2(\underline{x})$ is

$$J_2 = \frac{1}{2} \int_0^{\infty} [U^*(\underline{x}) - U_s(\underline{x})]^2 \, dt \tag{5.2.1}$$

where the integral is to be evaluated along the suboptimal trajectory. This expression can be transformed from an integral over time into an integral over the arc length of the suboptimal trajectory by using the following theorem of Rudin[52] [Theorem 6.35, page 125].

Theorem 5.1: Let $\underline{x}(\tau)$ be a continuous mapping of an interval $[0, t]$ into R^k. If $\dot{\underline{x}}(\tau)$ is continuous on $[0, t]$, then $\underline{x}(\tau)$ is rectifiable and has length $\ell(t)$

$$\ell(t) = \int_0^t \|\dot{\underline{x}}(\tau)\| \, d\tau \tag{5.2.2}$$

Since this integral is a Riemann integral, it is valid for $t = \infty$ if the integral converges. Since we are only considering dynamical systems which are asymptotically stable with finite cost, it is reasonable to assume that all trajectories generated by such systems will have a finite length. We shall make this assumption and use the above integral over the infinite time integral $[0, \infty)$.

Differentiating Equation (5.2.2) with respect to time gives

$$\frac{d\ell(t)}{dt} = \|\dot{\underline{x}}(t)\| \qquad \text{or} \qquad dt = \frac{d\ell(t)}{\|\dot{\underline{x}}(t)\|} \tag{5.2.3}$$

For the dynamical system

$$\dot{\underline{x}} = \underline{f}(\underline{x}) + \underline{b}\, U_s(\underline{x}) \tag{5.2.4}$$

this becomes

$$dt = \left[\| \underline{f}[\underline{x}(\ell)] + \underline{b}\, U_s[\underline{x}(\ell)] \| \right]^{-1} d\ell \tag{5.2.5}$$

where $\underline{x}(\ell)$ is to be interpreted as

$$\underline{x}(\ell) \equiv \underline{x}[t(\ell)]$$

Observing that $\ell(0) = 0$ and defining $\ell(\infty) = L$, we now convert the integration over time into an integration over the arc length of the suboptimal trajectory as

$$J_2 = \frac{1}{2} \int_0^L \left[\| \underline{f}[\underline{x}(\ell)] + \underline{b}\, U_s[\underline{x}(\ell)] \| \right]^{-1} \left[U^*[\underline{x}(\ell)] - U_s[\underline{x}(\ell)] \right]^2 d\ell \tag{5.2.6}$$

As Equation (5.2.6) indicates, the integrand in this cost functional specifies at each point $\underline{x}$ the amount of incremental excess suboptimality (i.e., the difference between the increment added to the suboptimal cost and that added to the optimal cost) for any trajectory passing through this point $\underline{x}$. Since minimizing the excess suboptimality is equivalent to minimizing the suboptimal cost, the process of minimizing the average incremental suboptimality over Ω_o (i.e., Criterion 3) should be an excellent design procedure.

5.3 The SDP1 Algorithm

In this section we shall formulate the first suboptimal design problem (SDP1) and develop an explicit computational algorithm for its solution. The SDP1 formulation is designed for use with any of the three criteria discussed in the previous section. As stated, none of these criteria are in a computationally feasible form because of the impossibility of analytically performing the indicated integration. Instead, we shall do the integration numerically by approximating the integral by a summation over a set of r equally spaced grid points which span Ω_o. As in Section 4.4, the set of integer subscripts will be denoted by R and the set of grid points by $\{\underline{x}_i, i \epsilon R\}$. Since the integrand of all three criteria should be smooth, replacing the integration by this summation should introduce only small errors for any reasonable quantization level.

For each of the quantized criteria, the unconstrained scalar gradient function $G(\underline{x}_i)$ defined in Section 4.4 can be directly computed at each grid point $\underline{x}_i$. The analytic expression for $G(\underline{x}_i)$ in each case is

Criterion 1: $$G(\underline{x}_i) = -2\,[U^*(\underline{x}_i) - U_s(\underline{x}_i)]$$

Criterion 2: $$G(\underline{x}_i) = -\frac{[U^*(\underline{x}_i) - U_s(\underline{x}_i)]^3\,[g(\underline{x}_i) + \frac{1}{2}U_s(\underline{x}_i)\,U^*(\underline{x}_i)]}{[g(\underline{x}_i) + U_s(\underline{x}_i)\,U^*(\underline{x}_i) - \frac{1}{2}(U^*(\underline{x}_i))^2]^3}$$

Criterion 3: $$G(\underline{x}_i) = -\frac{[U^*(\underline{x}_i) - U_s(\underline{x}_i)]}{\|\dot{\underline{x}}(\underline{x}_i)\|}\left[1 + \left(\frac{[U^*(\underline{x}_i) - U_s(\underline{x}_i)]}{2\,\|\dot{\underline{x}}(\underline{x}_i)\|}\right)\left(\frac{d\,\|\dot{\underline{x}}(\underline{x}_i)\|}{dU_s(\underline{x}_i)}\right)\right]$$

where

$$\|\dot{\underline{x}}(\underline{x}_i)\| = \|\underline{f}(\underline{x}_i) + \underline{b}\,U_s(\underline{x}_i)\|$$

Thus, for any of the three criteria and any given control law $U_s(\underline{x}) \epsilon$ S(M, N, K), the gradient projection algorithm can be used to compute the vector $\underline{\delta c}$ specifying the direction of steepest ascent in the coefficient space C for the cost of the particular criterion used.

We shall assume that one of the three criteria has been selected for use. Using the notation of Section 4.4, let the quantized cost of the selected criterion be denoted as $\left\langle \overline{J}(\underline{c}) \right\rangle$ whose unconstrained gradient function is $G(\underline{x}_i)$. We are now able to formally state the first suboptimal design problem.

Definition 5.2: Suboptimal Design Problem 1 (SDP1)

Given a structurally constrained optimal control problem and its unconstrained optimal feedback control law $U^*(\underline{x})$. Determine the optimal constrained solution $U_s^*(\underline{x}) \in S(M, N, K)$ specified by the D component vector $\underline{c}^*$ which minimizes the cost $\left\langle \overline{J}(\underline{c}) \right\rangle$.

The following algorithm based on a recursive gradient procedure can be used to solve the above problem.

Algorithm for the Solution of SDP1

1) Determine the set Ω_o and select an appropriate set of grid points which span it.

2) Set the iteration index ℓ to 0 and start with some initial control law $U[\underline{x}, \underline{c}_o] \in S(M, N, K)$.

3) Compute $G_\ell(\underline{x}_i)$ $\forall i \in R$.

4) Compute $\underline{\delta c}_\ell$ with the gradient projection algorithm.

5) Determine the optimal value of a_ℓ, a_ℓ^*, for which the cost $\left\langle \overline{J}(\underline{c}_\ell - a_\ell \underline{\delta c}_\ell) \right\rangle$ will be a minimum.

6) Set $\underline{c}_{\ell+1} = \underline{c}_\ell - a_\ell^* \underline{\delta c}_\ell$ and then increment the index ℓ by 1.

7) Evaluate some preselected termination criterion and then either return to step 3 or stop as indicated.

5.4 Evaluation of the SDP1 Algorithm

In evaluating a suboptimal design procedure, the two considerations of prime importance are its computational requirements and the suboptimality of its solutions. These will be discussed in the following paragraphs. However, there is a third major consideration for the SDP1 algorithm. This algorithm requires that the numerical values of the optimal unconstrained feedback control law $U^*(\underline{x})$ be known as a function of the state. This data could be obtained by solving the unconstrained optimization problem by dynamic programming. However, dynamic programming solutions generally require considerable computational time. Thus, if $U^*(\underline{x})$ is unknown, the computational requirements for computing $U^*(\underline{x})$ must be considered in addition to those for the SDP1 algorithm.

From a computational standpoint the SDP1 algorithm is extremely attractive. It does not require the generation of any system trajectories or the solution of any differential equations. The only computation which is required is that of evaluating scalar functions at the grid points. Hence, the required computational time should be quite small for each iteration. The author would estimate that with a modern computer (i.e., system 360/65, etc.) the iteration time would be from 1 to 10 seconds for a four state variable system and a state space quantization of 10,000 grid points. Although gradient procedures typically exhibit slow convergence near the optimum, this rapid iteration rate should more than compensate and allow one to solve three and four state variable problems in less than five minutes of computation time.

The SDP1 algorithm has been tested with all three criteria on several two dimensional problems in which the system and structurally constrained suboptimal control were linear, the unconstrained optimal control was nonlinear, and the initial condition probability distribution $Q_o(\underline{x})$ was uniform over a hypersphere centered at the origin. Although each of the criteria produced reasonably good results (within 6% of the optimal constrained solution), the third criterion was somewhat better than the second and both were considerably better than the first. In all of the cases tested, the third criterion produced control laws whose cost was within 2% of the optimal constrained cost. Based on these results, the author concludes that the third criterion is the best choice for formulating a suboptimal design procedure. However, it should be noted that the other criteria have certain advantages; the second criterion will produce the best guaranteed bound on performance and the computational formulas of the first criterion are somewhat simpler than those of the other two. Hence, they may be preferred in certain applications.

An important restriction on the class of problems for which the SDP1 algorithm is applicable should be noted. The basic assumption of the SDP1 formulation is that minimizing the various criteria at each point $\underline{x} \in \Omega_o$ with uniform weighting is a good suboptimal design procedure. For cases in which $Q_o(\underline{x})$ is a uniform distribution over a hypersphere, the resulting trajectories tend to be uniformly distributed throughout Ω_o and this assumption appears to be valid. However, if $Q_o(\underline{x})$ is not uniform and if the resulting trajectories tend to concentrate in certain portions of Ω_o, the utilization of a uniform weighting is a highly suboptimal procedure. Hence, the SDP1 algorithm should only be used for problem formulations in which $Q_o(\underline{x})$ will generate trajectories which tend to be uniformly distributed throughout Ω_o.

In conclusion, for problems in which $U^*(\underline{x})$ is known or can be computed and for which $Q_o(\underline{x})$ generates trajectories which are uniformly distributed throughout Ω_o, the SDP1 algorithm seems to provide an effective method for solving structurally constrained optimization problems.

5.5 The SDP2 Algorithm

In this section we shall formulate a second suboptimal design problem SDP2 and develop an explicit computational algorithm for its solution. Its major advantage over the SDP1 formulation is that it is applicable to problems with completely general initial condition probability distributions $Q_o(\underline{x})$.

The basic assumption of the SDP1 formulation is that minimizing a given criterion at each point $\underline{x} \in \Omega_o$ with uniform weighting is a good suboptimal design procedure. For cases in which $Q_o(\underline{x})$ is a uniform distribution over a hypersphere, the resulting trajectories tend to be uniformly distributed throughout Ω_o and this assumption appears to be valid. However, if $Q_o(\underline{x})$ is not uniform and if the resulting trajectories tend to concentrate in certain portions of Ω_o, the utilization of uniform weighting is a highly suboptimal procedure. Instead, a weighting function

<u>should be introduced which correctly compensates for the trajectory density at each point</u>. We shall incorporate this modification into the formulation of the second suboptimal design problem. Actually, a weighting function itself will not be used; instead, <u>the set of grid points will be distributed throughout Ω_o proportional to the trajectory density, thus achieving the same result</u>.

The cost functional of the SCOCP is

$$\langle J \rangle = \int_{Q_o} J_s(\underline{x}_o)\, Q_o(\underline{x}_o)\, d\underline{x}_o \tag{5.5.1}$$

Using the definition of the excess suboptimality $J_2(\underline{x}_o)$

$$J_2(\underline{x}_o) = J_s(\underline{x}_o) - J^*(\underline{x}_o) \tag{5.5.2}$$

we can express $\langle J \rangle$ as

$$\langle J \rangle = \int_{Q_o} J^*(\underline{x}_o)\, Q_o(\underline{x}_o)\, d\underline{x}_o + \int_{Q_o} J_2(\underline{x}_o)\, Q_o(\underline{x}_o)\, d\underline{x}_o \tag{5.5.3}$$

Since the first integral is independent of the control law $U_s(\underline{x})$, minimizing $\langle J \rangle$ is equivalent to minimizing the second integral which we shall denote as $\langle J_2 \rangle$.

$$\langle J_2 \rangle = \int_{Q_o} J_2(\underline{x}_o)\, Q_o(\underline{x}_o)\, d\underline{x}_o \tag{5.5.4}$$

Since minimizing the suboptimal cost $\langle J \rangle$ by minimizing $\langle J_2 \rangle$ is substantially easier than attempting to minimize $\langle J \rangle$ directly, the SDP2 problem will be formulated as that of minimizing $\langle J_2 \rangle$.

If we were considering linear systems and linear feedback control laws, a transition matrix could be used to express the propagation of the initial condition probability distribution. However, for the nonlinear systems and nonlinear control laws under consideration, no computationally feasible equations exist for propagating the initial condition probability distribution. Thus an alternative method must be developed for accomplishing this. The procedure we shall use is to quantize $Q_o(\underline{x})$ (i.e., represent $Q_o(\underline{x})$ by a finite set of impulses distributed throughout Q_o) as defined below.

Definition 5.3: The distribution $\hat{Q}_o(\underline{x})$ will be called the quantized initial condition probability distribution of $Q_o(\underline{x})$ if

$$\hat{Q}_o(\underline{x}) = \sum_{k=1}^{\alpha} Q_k \delta\left[\|(\underline{x} - \underline{x}_k)\|\right] \qquad (5.5.5)$$

where α is some positive integer, each $\underline{x}_k \in Q_o$, each $Q_k \geq 0$, $\sum_{i=1}^{\alpha} Q_k = 1$,

and the points $\underline{x}_k$ and weights Q_k are selected so that

$$\int_{Q_o} J_2(\underline{x}) Q_o(\underline{x}) d\underline{x} \approx \int_{Q_o} J_2(\underline{x}) \hat{Q}_o(\underline{x}) d\underline{x} = \sum_{k=1}^{\alpha} Q_k J_2(\underline{x}_k) \qquad (5.5.6)$$

Since all of the $U_s(\underline{x}) \in S(M, N, K)$ are infinitely differentiable, their corresponding costs $J_2(\underline{x})$ should be extremely smooth functions; hence, the error introduced by replacing $Q_o(\underline{x})$ with $\hat{Q}_o(\underline{x})$ should be small for any reasonable quantization.

With the replacement of $Q_o(\underline{x})$ by $\hat{Q}(\underline{x})$, the expression for $\langle J_2 \rangle$ becomes

$$\langle J_2 \rangle = \sum_{k=1}^{a} Q_k \, J(\underline{x}_k) \tag{5.5.7}$$

and thus $\langle J_2 \rangle$ is merely the weighted sum of the costs for a finite set of initial conditions. Utilizing Equation (5.2.6) we can express $\langle J_2 \rangle$ as

$$\langle J_2 \rangle = \frac{1}{2} \sum_{k=1}^{a} Q_k \left[\int_0^{L(\underline{x}_k)} \left[\| \underline{f}[\underline{x}(\ell)] + \underline{b}\, U_s[\underline{x}(\ell)] \| \right]^{-1} \left[U^*[\underline{x}(\ell)] - U_s[\underline{x}(\ell)] \right]^2 d\ell \right] \tag{5.5.8}$$

thus converting $\langle J_2 \rangle$ into the sum of a finite set of integrals along trajectories in the state space.

By utilizing the Newton-Cotes integration formulas to approximate the trajectory integrals, the cost functional $\langle J_2 \rangle$ can be converted into the summation of a scalar function over a finite set of grid points. These formulas replace the integral by a weighted sum of the values of the integrand at equally spaced points along the integration range as indicated

$$\int_0^{L} F(\ell)\, d\ell \approx \frac{L}{\beta} \sum_{j=0}^{\beta} W_j(\beta)\, F(\ell_j) \tag{5.5.9}$$

where

$$\ell_j = \frac{L}{\beta} j \tag{5.5.10}$$

and $W_j(\beta)$ is a prespecified weighting function dependent only on the index j and β, the number of intervals between the $\beta + 1$ data points.

These integration formulas are extremely good — particularly if the integrand is smooth. For example, the value of the integral of e^x over $[0, 1]$ can be computed by the 6 data point Newton-Cotes formula with an error of less than 10^{-9}. For the class of problems under consideration in which all of the suboptimal controls are infinitely differentiable, the integrand in Equation (5.5.8) should be smooth; therefore, the Newton-Cotes integration formulas should be very effective. The Newton-Cotes integration technique is described in considerable detail in references 53 and 54 where the values of the weighting function $W_j(\beta)$ are tabulated for all values of β up to 20.

If the Newton-Cotes integration formula with $\beta + 1$ equally spaced data points is used for each of the trajectories, $\langle J_2 \rangle$ can be expressed as

$$\langle J_2 \rangle = \frac{1}{2} \sum_{k=1}^{\alpha} Q_k \frac{L(\underline{x}_k)}{\beta} \sum_{j=0}^{\beta} W_j(\beta) \left[\| \underline{f}(\underline{x}_{k,j}) + \underline{b}\, U_s(\underline{x}_{k,j}) \| \right]^{-1} \left[U^*[\underline{x}_{k,j}] - U_s[\underline{x}_{k,j}] \right]^2 \tag{5.5.11}$$

where $L(\underline{x}_k)$ is the length of the k^{th} trajectory starting from the initial condition $\underline{x}_k$ and $\underline{x}_{k,j}$ is the j^{th} data point along this trajectory. Equation (5.5.11) can be placed in a more convenient form if a single index i is used to label each of the grid points. Each value of i will have a unique, one-to-one correspondence with each index pair (k, j) as specified by the formula

$$i = \alpha(k-1) + j + 1 \tag{5.5.12}$$

We define

$$r = \alpha(\beta + 1) \tag{5.5.13}$$

$$\underline{x}_i = \underline{x}_{k,j} \tag{5.5.14}$$

$$\hat{W}_i = \frac{Q_k\, L(\underline{x}_k)\, W_j(\beta)}{\beta} \tag{5.5.15}$$

The cost $\langle J_2 \rangle$ expressed in terms of the new notation is

$$\langle J_2 \rangle = \frac{1}{2} \sum_{i=1}^{r} \left[\hat{W}_i\right] \left[\left\| \underline{f}(\underline{x}_i) + \underline{b}\, U_s(\underline{x}_i) \right\|\right]^{-1} \left[U^*(\underline{x}_i) - U_s(\underline{x}_i)\right]^2 \tag{5.5.16}$$

Thus by using the initial condition probability distribution $Q_o(\underline{x})$ and the system dynamics to specify a particular set of grid points, we have been able to express $\langle J_2 \rangle$ as merely a summation of a scalar function over a set of grid points.

Although the cost functional $\langle J_2 \rangle$ is similiar in form to $\langle \bar{J}_3 \rangle$, the cost functional used in SDP1 with Criterion 3, $\langle J_2 \rangle$ differs in two important respects. First, $\langle \bar{J}_3 \rangle$ is not the cost functional of the SCOCP; it is merely the cost functional of a suboptimal design problem which was selected because its optimal solution is expected to be close to that of the SCOCP. In contradistinction, $\langle J_2 \rangle$ represents the excess cost of the SCOCP which results from suboptimality (i.e., $\langle J_2 \rangle = \langle J_s \rangle - \langle J^* \rangle$); thus, except for errors introduced by the quantizations employed, minimizing $\langle J_2 \rangle$ is precisely equivalent to minimizing the cost functional of the SCOCP. Second, whereas the grid points $\underline{x}_i$ in $\langle \bar{J}_3 \rangle$ are fixed, specified a priori, and completely independent of the suboptimal control, the grid points $\underline{x}_i$ used in $\langle J_2 \rangle$ depend on the particular control used and will change as the control is changed.

If the unconstrained scalar gradient function $G(\underline{x}_i)$ for $\langle J_2 \rangle$ defined in Section 4.4 by Equation (4.4.5) as

$$G(\underline{x}_i) = \frac{\partial \left\langle J_2[U_s(\underline{x})] \right\rangle}{\partial U_s(\underline{x}_i)} \tag{5.5.17}$$

could be computed at each grid point $\underline{x}_i$ for $\langle J_2 \rangle$, one could formulate a simple gradient algorithm for solving the SCOCP. Due to the similarity in form between $\langle \overline{J}_3 \rangle$ and $\langle J_2 \rangle$, one might suggest that the correct expression for $G(\underline{x}_i)$ should be

$$G(\underline{x}_i) = - \frac{[W_i][U^*(\underline{x}_i) - U_s(\underline{x}_i)]}{\|\dot{\underline{x}}(\underline{x}_i)\|} \left[1 + \left(\frac{[U^*(\underline{x}_i) - U_s(\underline{x}_i)]}{2\|\dot{\underline{x}}(\underline{x}_i)\|} \right) \left(\frac{d\|\dot{\underline{x}}(\underline{x}_i)\|}{dU_s(\underline{x}_i)} \right) \right] \tag{5.5.18}$$

where

$$\|\dot{\underline{x}}(\underline{x}_i)\| = \|\underline{f}(\underline{x}_i) + \underline{b}\, U_s(\underline{x}_i)\| \tag{5.5.19}$$

This expression would be correct if the grid points were fixed as in the case of $\langle \overline{J}_3 \rangle$ in SDP1 and did not depend upon the control $U_s(\underline{x})$. <u>However since the grid points of</u> $\langle J_2 \rangle$ <u>do depend on</u> $U_s(\underline{x})$, <u>Equation (5.5.18) is not the correct expression for</u> $G(\underline{x}_i)$. In Chapter VI we shall develop a method for computing the correct gradient function for $\langle J_2 \rangle$ and then formulate an algorithm for solving the SCOCP. However, the procedure which will be developed there for computing $G(\underline{x}_i)$ is considerably more complex than that of computing the function defined by Equation (5.5.18).

Experience has shown that although the function defined by Equation (5.5.18) is not the correct gradient function for $\langle J_2 \rangle$, this function is a good approximation to it. As the second suboptimal design problem we shall propose utilizing a gradient algorithm in which one attempts to minimize $\langle J_2 \rangle$ by using the expression of Equation (5.5.18) for the gradient of $\langle J_2 \rangle$. <u>Since this expression is not the correct gradient for</u>

$\langle J_2 \rangle$, the algorithm will not converge to the optimal solution of the SCOCP and hence is a suboptimal procedure. Nevertheless, in the cases tested this algorithm has produced control laws for which the corresponding costs are less than .5% above the optimal constrained cost. Thus, this algorithm seems to provide a computationally feasible procedure for approximately solving the general SCOCP at an acceptable level of suboptimality. We are now able to formally state the second suboptimal design problem.

Definition 5.4: Suboptimal Design Problem 2 (SDP2)

Given a structurally constrained optimal control problem and its unconstrained optimal feedback control law $U^*(\underline{x})$. Determine a suboptimal control law $\hat{U}_s(\underline{x}) \in S(M, N, K)$ specified by a D component vector $\hat{\underline{c}}$ which approximately minimizes $\langle J_2 \rangle$ by using a recursive gradient algorithm in which the function defined by Equation (5.5.18) is used as the gradient function of $\langle J_2 \rangle$.

The following algorithm can be used to solve the SDP2.

Algorithm for the Solution of SDP2

1) Select an appropriate set of quantized initial conditions which approximate $Q_o(\underline{x})$ and select β, the number of intervals used in the Newton-Cotes integration formula.

2) Set the iteration index ℓ to 0 and start with some initial control law $U_s[\underline{x}, \underline{c}_o] \in S(M, N, K)$.

3) Using $U_s[\underline{x}, \underline{c}_\ell]$, generate a trajectory for each of the quantized initial conditions and determine the set of grid points $\{\underline{x}_i,\ i \in R\}$.

4) Use Equation (5.5.18) to compute the approximate gradient function at each grid point.

5) Compute $\underline{\delta c}_\ell$ with the gradient projection algorithm.

6) Determine the optimal value of α_ℓ, α_ℓ^*, for which the cost $\left\langle J_2(\underline{c}_\ell - \alpha_\ell \underline{\delta c}_\ell)\right\rangle$ will be a minimum.

7) Set $\underline{c}_{\ell+1} = \underline{c}_\ell - \alpha_\ell^* \underline{\delta c}_\ell$ and then increment the index ℓ by 1.

8) Evaluate some preselected termination criterion and then either return to step 3 or stop as indicated.

5.6 Evaluation of the SDP2 Algorithm

In comparing the SDP2 and SDP1 procedures, one concludes that the SDP2 formulation is more accurate, applicable to more general problems, but has more extensive computational requirements. Like the SDP1 algorithm, it requires the numerical values of the optimal unconstrained feedback control law $U^*(\underline{x})$ as a function of the state.

The SDP2 formulation is more general than that of the SDP1 in that it may be used with completely general initial condition probability distributions instead of only those which generate trajectories which are uniformly distributed throughout Ω_0.

The author has tested the SDP2 algorithm on several linear two-dimensional problems which could be solved analytically. In the cases tested, the SDP2 algorithm produced control laws whose costs varied from .12% to .5% above the optimal constrained cost. This is somewhat better than the 1% to 2% above optimal of the SDP1 algorithm.

From a computational standpoint, the SDP2 algorithm is considerably more complex. A new set of trajectories and grid points must be generated on each iteration. Furthermore, the value of $U^*(\underline{x})$ must be computed at each of these grid points – presumably by interpolation from the dynamic programming table. Both of these operations will require considerable computational time – probably increasing the iteration time by a factor of 5 to 10.

A possible compromise between the SDP1 and SDP2 algorithms should be mentioned. The set of trajectories and grid points could be computed on only the first or second iteration and held fixed for all future iterations. This might give a sufficiently accurate approximation to the correct weighting for the given $Q_o(\underline{x})$ to produce a good suboptimal design, yet retain the computational advantages of the SDP1 algorithm. The applicability of this procedure remains a topic for future research.

CHAPTER VI

SOLUTION OF THE OPTIMIZATION PROBLEM

6.1 Introduction

The purpose of this chapter is to develop an optimal design procedure for solving the SCOCP. Unlike the suboptimal techniques of the preceeding chapter, this design procedure does not require the numerical values of $U^*(\underline{x})$ or any other precomputed information. An expression for the Fréchet derivative of the cost with respect to a variation in the feedback control law is developed in Section 6.2. This expression is utilized in Section 6.3 to formulate an explicit computational algorithm for solving the SCOCP. The characteristics and computational requirements of this algorithm are discussed in Section 6.4. An example illustrating the procedure is presented in Section 6.5.

6.2 Derivation of the Feedback Fréchet Derivative

An expression for the Fréchet derivative of the cost with respect to a variation in the feedback control law is derived in this section for the case in which the initial condition is a single point $\underline{x}_o$; the following section extends the results to the more general case of an arbitrary initial condition probability distribution $Q_o(\underline{x})$.

The SCOCP problem formulation considers dynamical systems of the form

$$\dot{\underline{x}} = \underline{F}[\underline{x}, U_s(\underline{x})] = \underline{f}(\underline{x}) + \underline{b}\, U_s(\underline{x}) \qquad (6.2.1)$$

with cost functionals

$$J = \int_0^\infty L[\underline{x}, U_s(\underline{x})]\, dt = \int_0^\infty [g(\underline{x}) + \frac{1}{2} U_s^2(\underline{x})]\, dt \qquad (6.2.2)$$

We shall use the notation $\underline{F}_x$ and $\underline{L}_x$ to denote the total (not partial) derivative of the functions $\underline{F}[\underline{x}, U_s(\underline{x})]$ and $L[\underline{x}, U_s(\underline{x})]$ with respect to $\underline{x}$.

$$\underline{F}_x = \frac{\partial \underline{f}(\underline{x})}{\partial \underline{x}} + \underline{b} \left(\frac{\partial U_s(\underline{x})}{\partial \underline{x}} \right)^T \qquad (6.2.3)$$

$$\underline{L}_x = \frac{\partial g(\underline{x})}{\partial \underline{x}} + U_s(\underline{x}) \left(\frac{\partial U_s(\underline{x})}{\partial \underline{x}} \right) \qquad (6.2.4)$$

Note that $\underline{F}_x$ is a matrix while $\underline{L}_x$ is a vector. Let $U(\underline{x})$ be any feedback control law for which the dynamical system of Equation (6.2.1) is asymptotically stable, $J(\underline{x})$ its cost, and $\underline{x}(t)$ its trajectory from the initial condition $\underline{x}_o$. Let $U_1(\underline{x}) = U(\underline{x}) + \delta U(\underline{x})$, where $\delta U(\underline{x})$ is any infinitesimal variation for which $\delta U(\underline{0}) = 0$. Let $J_1(\underline{x})$ be the cost corresponding to $U_1(\underline{x})$ and $\underline{x}_1(t)$ its trajectory from the initial condition $\underline{x}_o$. We define $\underline{h}(t)$ as

$$\underline{h}(t) = \underline{x}_1(t) - \underline{x}(t) \qquad \forall\, t \in [0, \infty) \qquad (6.2.5)$$

Then the incremental change in the cost δJ defined as

$$\delta J = J_1(\underline{x}_o) - J(\underline{x}_o) \qquad (6.2.6)$$

can be expressed to first order as

$$\delta J = \int_0^\infty \left[\left\langle \underline{L}_x, \underline{h}(t) \right\rangle + U[\underline{x}(t)]\, \delta U[\underline{x}(t)] \right] dt \qquad (6.2.7)$$

where $\underline{L}_x$ is evaluated along the nominal trajectory $\underline{x}(t)$. The second term in the above integrand is already in the form of a Fréchet derivative (i.e., a function of $\underline{x}(t)$ times $\delta U[\underline{x}(t)]$); the only remaining task is to convert the first term into this form.

First, we need to establish an identity. From Equations (6.2.1), (6.2.3), and (6.2.5) it follows that to first order

$$\dot{\underline{h}}(t) = \underline{F}_x \underline{h}(t) + \underline{b}\, \delta U[\underline{x}(t)] \qquad \forall\, t \in [0, \infty) \tag{6.2.8}$$

where $\underline{F}_x$ is evaluated along the nominal trajectory $\underline{x}(t)$. Integrating Equation (6.2.8) from 0 to t yields

$$\underline{h}(t) - \int_0^t \underline{F}_x \underline{h}(\tau)\, d\tau = + \underline{b} \int_0^t \delta U[\underline{x}(\tau)]\, d\tau \tag{6.2.9}$$

Since $U(\underline{x})$ is asymptotically stable and $\delta U_1(\underline{x})$ is an infinitesimal variation with $\delta U(\underline{0}) = 0$, $U_1(\underline{x})$ must also be asymptotically stable. Therefore, $\underline{x}(\infty) = \underline{x}_1(\infty) = \underline{0}$ and $\underline{h}(\infty) = \underline{0}$. Thus, it follows that

$$\int_0^\infty \underline{F}_x \underline{h}(\tau)\, d\tau = - \underline{b} \int_0^\infty \delta U[\underline{x}(\tau)]\, d\tau \tag{6.2.10}$$

In Section 2.3 we established that the equation describing the propagation of the costate variable $\underline{p}(t)$ defined as

$$\underline{p}(t) = \frac{\partial J[\underline{x}(t)]}{\partial \underline{x}} \tag{6.2.11}$$

along a trajectory $\underline{x}(t)$ resulting from a feedback control law $U(\underline{x})$ is

$$\dot{\underline{p}}(t) = -(\underline{F}_x)^T \underline{p}(t) - \underline{L}_x \tag{6.2.12}$$

Taking the inner product of both sides of this equation with $\underline{h}(t)$ and integrating from zero to infinity gives

$$\int_0^{\infty} \left\langle \underline{L}_x, \underline{h}(t) \right\rangle dt = - \int_0^{\infty} \left\langle \dot{\underline{p}}(t), \underline{h}(t) \right\rangle dt - \int_0^{\infty} \left\langle \underline{p}(t), \underline{F}_x \underline{h}(t) \right\rangle dt \tag{6.2.13}$$

Integrating the third integral by parts and regrouping terms gives

$$\int_0^{\infty} \left\langle \underline{L}_x, \underline{h}(t) \right\rangle dt = - \int_0^{\infty} \left\langle \dot{\underline{p}}(t), \left[\underline{h}(t) - \int_0^{t} \underline{F}_x \underline{h}(\tau) d\tau \right] \right\rangle dt - \left\langle \underline{p}(\infty), \int_0^{\infty} \underline{F}_x \underline{h}(\tau) d\tau \right\rangle \tag{6.2.14}$$

Using the identity established by Equation (6.2.9) to convert the second integral into the equivalent form indicated below and then integrating it by parts gives

$$- \int_0^{\infty} \left\langle \dot{\underline{p}}(t), \underline{b} \int_0^{t} \delta U[\underline{x}(\tau)]\, d\tau \right\rangle dt = + \int_0^{\infty} \left\langle \underline{p}(t), \underline{b} \right\rangle \delta U[\underline{x}(t)]\, dt$$

$$- \left\langle \underline{p}(\infty), \int_0^{\infty} \underline{b}\ \delta U[\underline{x}(\tau)]\, d\tau \right\rangle \tag{6.2.15}$$

Inserting this expression into Equation (6.2.14) and observing that the last term in Equation (6.2.14) is the negative of the last term in Equation (6.2.15) as indicated by Equation (6.2.10), one finally obtains

$$\int_0^{\infty} \left\langle \underline{L}_x, \underline{h}(t) \right\rangle dt = \int_0^{\infty} \left\langle \underline{p}(t), \underline{b} \right\rangle \delta U[\underline{x}(t)]\, dt \tag{6.2.16}$$

Inserting the above expression into Equation (6.2.7) gives the correct expression for the feedback Fréchet derivative

$$\delta J = \int_0^\infty \left[U[\underline{x}(t)] + \left\langle \underline{p}(t), \underline{b} \right\rangle \right] \delta U[\underline{x}(t)]\, dt \tag{6.2.17}$$

where $\underline{p}(t)$ satisfies the differential equation

$$\dot{\underline{p}}(t) = -\left(\frac{\partial g}{\partial \underline{x}}[\underline{x}(t)] \right) - \left(\frac{\partial \underline{f}}{\partial \underline{x}}[\underline{x}(t)] \right)^T \underline{p}(t) - \left[U[\underline{x}(t)] + \left\langle \underline{p}(t), \underline{b} \right\rangle \right] \left(\frac{\partial U}{\partial \underline{x}}[\underline{x}(t)] \right) \tag{6.2.18}$$

The gradient of the cost with respect to $U[\underline{x}(t)]$ is merely that portion of the integrand of the Fréchet derivative which multiplies $\delta U[\underline{x}(t)]$; thus, it is given by the following expression

$$U_s[\underline{x}(t)] + \left\langle \underline{b}, \underline{p}(t) \right\rangle \tag{6.2.19}$$

for all points along the trajectory $\underline{x}(t)$.

It should be noted that the derivation and expression for the feedback Fréchet derivative will reduce to that of the conventional open-loop Fréchet derivative if the state dependence of the control $U(\underline{x})$ is removed (i.e., if the partial derivative of the control with respect to $\underline{x}$ is set equal to zero) and the terms $U[\underline{x}(t)]$ and $\delta U[\underline{x}(t)]$ are interpreted as $u(t)$ and $\delta u(t)$ respectively. The Fréchet derivative, Equation (6.2.17), is formally the same in the two cases; however, the values of $\underline{p}(t)$ are different since the differential equations for the propagation of $\underline{p}(t)$ differ.

The preceeding results should be compared with those of Jacobson[45] who has developed both first and second order algorithms for computing the unconstrained optimal control for single trajectory problems from a feedback (Differential Dynamic Programming) viewpoint. Jacobson first formulates a second order algorithm and then obtains a first order algorithm by setting all second order terms to zero. The gradient function in the resulting first order algorithm is identical with that derived above except for the fact that the term $\frac{\partial U[\underline{x}(t)]}{\partial \underline{x}}$ appearing in the differential equation for $\underline{p}(t)$, Equation (6.2.18), is set equal to zero because his procedure for evaluating $\frac{\partial U[\underline{x}(t)]}{\partial \underline{x}}$ requires second order terms. Therefore, Jacobson's first order algorithm does not utilize a "correct" expression for the gradient function. Nevertheless, this approach is feasible for unconstrained optimization problems since this approximation converges to the correct expression for the gradient as the optimum is approached; hence, his first order algorithm can be expected to converge to the correct unconstrained optimal control. In contradistinction, Jacobson's first order algorithm can not be used for structurally constrained optimization problems in which the unconstrained optimal control is not a member of the constraint set since its gradient is incorrect everywhere except in the vicinity of the unconstrained optimum. In his second order algorithm, however, Jacobson uses Equation (6.2.18) with $\frac{\partial U[\underline{x}(t)]}{\partial \underline{x}}$ included for propagating $\underline{p}(t)$. Thus the correct expression for the feedback gradient is a hybrid of Jacobson's first and second order algorithms. Fortunately, in the case of structural constraints of the S(M, N, K) type (where the correct expression for the gradient must be used if convergence to the optimal constrained solution is desired), one can compute simple analytic expressions for $\frac{\partial U[\underline{x}]}{\partial \underline{x}}$ from the expression specifying the structural constraint; thus, no second order terms need to be computed or propagated in order to compute the correct expression for the feedback gradient.

6.3 Algorithm for the Optimal Solution of the SCOCP

An explicit computational algorithm for solving the SCOCP with a quantized initial condition probability distribution is developed in this section. The algorithm does not require any precomputed information and should converge to the optimal constrained solution.

First, we must convert the cost functional of the SCOCP

$$\langle J \rangle = \int_{Q_o} J(\underline{x})\, Q_o(\underline{x})\, d\underline{x} \tag{6.3.1}$$

into a form which is compatible with the gradient projection algorithm. As in Chapter V, we must use a quantized initial condition probability distribution $\hat{Q}_o(\underline{x})$ since no computationally feasible equations exist for propagating $\hat{Q}_o(\underline{x})$ with nonlinear dynamics. The distribution $\hat{Q}_o(\underline{x})$ was defined in Definition 5.3. With this restriction the cost functional can be expressed as

$$\langle J \rangle = \sum_{k=1}^{\alpha} Q_k\, J(\underline{x}_k) \tag{6.3.2}$$

If there is a variation $\delta U(\underline{x})$ in the feedback control law $U(\underline{x})$, it follows from the results of the previous section that the corresponding variation in the cost functional is given by

$$\delta \langle J \rangle = \sum_{k=1}^{\alpha} Q_k\, \delta J(\underline{x}_k) \tag{6.3.3}$$

where

$$\delta J(\underline{x}_k) = \int_0^\infty \left[U[\underline{x}_k(t)] + \left\langle \underline{p}_k(t), \underline{b} \right\rangle\right] \delta U[\underline{x}_k(t)]\, dt \qquad (6.3.4)$$

The function $\underline{x}_k(t)$ is the trajectory generated by $U(\underline{x})$ from the initial condition $\underline{x}_k$ and $\underline{p}_k(t)$ is its costate variable. Since we established in Chapter III that

$$\underline{p}(t) \equiv \frac{\partial J[\underline{x}(t)]}{\partial \underline{x}} \qquad (6.3.5)$$

Equation (6.3.4) can be expressed as

$$\delta J(\underline{x}_k) = \int_0^\infty \left[U[\underline{x}_k(t)] + \left\langle \underline{P}[\underline{x}_k(t)], \underline{b} \right\rangle\right] \delta U[\underline{x}_k(t)]\, dt \qquad (6.3.6)$$

where we define

$$\underline{P}(\underline{x}) \triangleq \frac{\partial J(\underline{x})}{\partial \underline{x}} \qquad (6.3.7)$$

The expression for $\delta J(\underline{x}_k)$ can be converted from an integral over time into an integral over the arc length of the trajectory as was done in Section 5.2. Using the notation defined there we can express $\delta J(\underline{x}_k)$ as

$$\delta J(\underline{x}_k) = \int_0^{L(\underline{x}_k)} \left\{ \left[\| \underline{f}[\underline{x}(\ell)] + \underline{b}\, U[\underline{x}(\ell)] \|\right]^{-1} \left[U[\underline{x}(\ell)] + \left\langle \underline{P}[\underline{x}(\ell)], \underline{b} \right\rangle\right] \delta U[\underline{x}(\ell)] \right\} d\ell \qquad (6.3.8)$$

To simplify notation, the function $\overline{G}(\underline{x})$ will be defined as

$$\overline{G}(\underline{x}) = \left[\| \underline{f}(\underline{x}) + \underline{b}\, U(\underline{x}) \|\right]^{-1} \left[U(\underline{x}) + \left\langle \underline{P}(\underline{x}), \underline{b} \right\rangle\right] \qquad (6.3.9)$$

Therefore the expression for $\delta \langle J \rangle$ becomes

$$\delta\langle J\rangle = \sum_{k=1}^{a} \Omega_k \int_0^{L(\underline{x}_k)} \overline{G}\,[\underline{x}_k(\ell)]\,\delta U\,[\underline{x}_k(\ell)]\,d\ell \qquad (6.3.10)$$

By utilizing the Newton-Cotes integration formulas to approximate the trajectory integrals, the variation in the cost functional $\delta\langle J\rangle$ can be converted into a summation of a scalar function over a finite set of grid points. This same technique was utilized in the SDP2 algorithm. These integration formulas replace the integral by a weighted sum of the values of the integrand at equally spaced points along the integration range as indicated

$$\int_0^L F(\ell)\,d\ell \approx \frac{L}{\beta} \sum_{j=0}^{\beta} W_j(\beta)\,F(\ell_j) \qquad (6.3.11)$$

where

$$\ell_j = \frac{L}{\beta}\,j \qquad (6.3.12)$$

and $W_j(\beta)$ is a prespecified weighting function dependent only on the index j and β, the number of intervals between the $\beta + 1$ data points.

If the Newton-Cotes integration formula with $\beta + 1$ equally spaced data points is used for each of the trajectories, $\delta\langle J\rangle$ can be expressed as

$$\delta\langle J\rangle = \sum_{k=1}^{a} \Omega_k \frac{L(\underline{x}_k)}{\beta} \sum_{j=0}^{\beta} W_j(\beta)\,\overline{G}(\underline{x}_{k,j})\,\delta U(\underline{x}_{k,j}) \qquad (6.3.13)$$

where $L(\underline{x}_k)$ is the length of the k^{th} trajectory starting from the initial condition $\underline{x}_k$ and $\underline{x}_{k,j}$ is the j^{th} data point along this trajectory. The above expression can be placed in a more convenient form if a single index i is used to label each of the grid points. Each value of i will have a unique, one-to-one correspondence with each index pair (k, j) as specified by the formula

$$i = \alpha(k-1) + j + 1 \tag{6.3.14}$$

We define

$$r = \alpha(\beta + 1) \tag{6.3.15}$$

$$\underline{x}_i = \underline{x}_{k,j} \tag{6.3.16}$$

$$G(\underline{x}_i) = \frac{Q_k \, L(\underline{x}_k) \, W_j(\beta)}{\beta} \, \overline{G}(\underline{x}_{k,j}) \tag{6.3.17}$$

Then $\delta\langle J\rangle$ may be expressed in the new notation as

$$\delta\langle J\rangle = \sum_{i=1}^{r} G(\underline{x}_i) \, \delta U(\underline{x}_i) \tag{6.3.18}$$

which is the required form for using the gradient projection algorithm.

The function $G(\underline{x}_i)$ will be called the state space gradient function since it directly indicates the rate of change of the cost with respect to a change in the scalar value of the control at each grid point. The preceeding notation may be condensed into the following formula for evaluating $G(\underline{x}_i)$ at each grid point.

$$G(\underline{x}_i) = \frac{Q_i L_i W_i}{\beta} \left[\|\underline{f}(\underline{x}_i) + \underline{b}\, U(\underline{x}_i)\|\right]^{-1} \left[U(\underline{x}_i) + \langle \underline{P}(\underline{x}_i), \underline{b}\rangle\right] \tag{6.3.19}$$

where

Q_i is the probability associated by $\hat{Q}_o(\underline{x})$ with the initial condition of the trajectory containing $\underline{x}_i$,

L_i is the length of the trajectory containing $\underline{x}_i$,

W_i is the value of the Newton-Cotes weighting function associated with $\underline{x}_i$, and

β is the number of intervals between the $\beta + 1$ grid points of each trajectory.

The quantities Q_i, W_i, and β are prespecified constants and do not have to be evaluated. L_i could be evaluated as each trajectory is generated to compute the grid points. Since $\underline{f}(\underline{x})$ and $U(\underline{x})$ are available in an analytic form, they can be directly evaluated at each grid point. The only term in the formula for $G(\underline{x}_i)$ which is difficult to evaluate is

$$\underline{P}(\underline{x}) = \frac{\partial J(\underline{x})}{\partial \underline{x}} \tag{6.3.20}$$

There are two possible methods of evaluating $\underline{P}(\underline{x}_i)$. First, $J(\underline{x})$ could be evaluated at each grid point as the trajectories were computed. Then a numerical differentiation formula could be used to estimate $\underline{P}(\underline{x}_i)$. However, such formulas have large error bounds and it is very doubtful if accurate results could be obtained. Therefore, the author does not recommend this approach.

A better procedure would be to use the differential equation for the propagation of $\underline{p}(t)$, Equation (6.2.18), to compute $P(\underline{x}_i)$ at each grid point. This approach is used in the conventional open-loop gradient technique and can be applied in this application as follows. One would select a terminal time T sufficiently large so that all trajectories would be very close to the origin at $t = T$. In the vicinity of the origin a linear quadratic approximation to the SCOCP dynamics and cost functional would be constructed. The cost $\hat{J}(\underline{x})$ for this approximation could be expressed as

$$\hat{J}(\underline{x}) = \frac{1}{2}\langle \underline{x}, \underline{M}\, \underline{x} \rangle \tag{6.3.21}$$

for which

$$\hat{\underline{P}}(\underline{x}) = \underline{M}\, \underline{x} \tag{6.3.22}$$

$\hat{\underline{P}}(\underline{x})$ should provide a reasonably good estimate for $\underline{P}(\underline{x})$ in the vicinity of the origin. Thus, it could be used to evaluate $\underline{p}(T)$ for each trajectory and by integrating Equation (6.2.18) backwards in time, the value of $\underline{P}(\underline{x}_i)$ could be computed at each grid point.

We shall use the latter procedure for evaluating $\underline{P}(\underline{x}_i)$ in the following algorithm which can be used to solve the SCOCP.

Algorithm for the Solution of the SCOCP

1) Select an appropriate set of quantized initial conditions which approximate $Q_o(\underline{x})$; select β, the number of intervals used in the Newton-Cotes integration formulas; and select the terminal time T.

2) Set the iteration index ℓ to 0 and start with some initial control law $U_s[\underline{x}, \underline{c}_o] \in S(M, N, K)$.

3) Using $U_s[\underline{x}, \underline{c}_\ell]$, generate a trajectory for each of the quantized initial conditions and determine the set of grid points $\{\underline{x}_i, i \in R\}$.

4) In the manner discussed, use Equation (6.2.18) to propagate $\underline{p}(t)$ along each trajectory to compute $\underline{P}(x_i)$ at each grid point.

5) Use Equation (6.3.19) to evaluate $G(\underline{x}_i)$ at each grid point.

6) Compute $\underline{\delta c}_\ell$ with the gradient projection algorithm.

7) Determine the optimal value of α_ℓ, α_ℓ^*, for which the cost $\left\langle J(\underline{c}_\ell - \alpha_\ell \underline{\delta c}_\ell)\right\rangle$ will be a minimum.

8) Set $\underline{c}_{\ell+1} = \underline{c}_\ell - \alpha_\ell^* \underline{\delta c}_\ell$ and increment the index ℓ by 1.

9) Evaluate some preselected termination criterion and then either return to step 3 or stop as indicated.

6.4 Characteristics of the Algorithm

The algorithm developed in the preceeding section has two major advantages over the suboptimal design procedures of Chapter V. First, it is an optimal design procedure; hence, it will produce a better system design than the suboptimal techniques. Second, it does not require the numerical values of $U^*(\underline{x})$ as a function of the state. Since computing $U^*(\underline{x})$ is a major task, this is a significant advantage.

The major disadvantage of the algorithm is its extensive computational requirements. Twice as many trajectories must be generated to evaluate the gradient function as in the SDP2 algorithm. Additional sets of trajectories must be generated in the process of evaluating α_ℓ^*. The author would estimate that each iteration would require from 3 to 5 times as much computational time as that for each iteration of the SDP2 algorithm. Nevertheless, it should be computationally feasible for problems having up to four state variables if the number of quantized initial conditions is limited to 50 points.

6.5 A Two-Dimensional Example

In this section a simple two-dimensional example is presented which illustrates the application of both the gradient projection algorithm and the optimal algorithm for solving the SCOCP. We shall consider the dynamical system

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= u \end{aligned} \tag{6.5.1}$$

and cost functional

$$J = \frac{1}{2} \int_0^\infty [x_1^2 + 2x_2^2 + u^2]\, dt \tag{6.5.2}$$

The optimal feedback control law for this problem is

$$U^*(\underline{x}) = -x_1 - 2x_2 \tag{6.5.3}$$

Now we impose the structural constraint that the control law must be of the form

$$U_s(\underline{x}) = -K[x_1 + x_2] \tag{6.5.4}$$

where K may be any positive constant. Clearly, $U^*(\underline{x})$ is not a member of this constraint set. For simplicity of presentation, we shall only consider the single trajectory problem for which the initial condition is the point $(\hat{x}_1, \hat{x}_2)$; the general case is merely a sum of such problems.

First, we shall determine the solution to this problem (i.e., the equation for K^*, the value of K which minimizes the cost, in terms of $\hat{x}_1$ and $\hat{x}_2$) by the simplest possible method; then, we will demonstrate that application of the algorithm of Section 6.3 will lead to the same result. The expression for the cost in terms of K, $\hat{x}_1$, and $\hat{x}_2$ can be determined in the following manner. Compute the Laplace transform of the time functions $x_1(t)$ and $x_2(t)$ which would result from driving the dynamical system of Equation (6.5.1) with the control law of Equation (6.5.4) from the initial condition $(\hat{x}_1, \hat{x}_2)$. Then, Parseval's theorem can be used to obtain an expression for the cost in terms of integrals of these Laplace transforms. The integral tables in Appendix F of Newton, Gould, and Kaiser [55] can be used to evaluate these integrals directly. The resulting expression for the cost is

$$J[\hat{x}_1, \hat{x}_2, K] = \frac{1}{4}\left\{[K+3+K^{-1}]\hat{x}_1^2 + [K+1+2K^{-1}+K^{-2}]\hat{x}_2^2 + [2K+2K^{-1}]\hat{x}_1\hat{x}_2\right\} \tag{6.5.5}$$

The equation for K^* is

$$\frac{\partial J[\hat{x}_1, \hat{x}_2, K]}{\partial K} = \frac{1}{4K^3}\left\{(\hat{x}_1+\hat{x}_2)^2 K^3 - [(\hat{x}_1+\hat{x}_2)^2+\hat{x}_2^2]K - [2\hat{x}_2^2]\right\} = 0 \tag{6.5.6}$$

It can be shown that the above equation has one and only one real solution which is always positive. The expressions for the costate variables $P_1(\underline{x}, K)$ and $P_2(\underline{x}, K)$ are

$$P_1(\underline{x}, K) = \frac{1}{2}[K+3+K^{-1}]x_1 + \frac{1}{2}[K+K^{-1}]x_2 \tag{6.5.7}$$

$$P_2(\underline{x}, K) = \frac{1}{2}[K + 1 + 2K^{-1} + K^{-2}] x_2 + \frac{1}{2}[K + K^{-1}] x_1 \tag{6.5.8}$$

By directly computing their time derivative along any trajectory of the dynamical system of Equation (6.5.1) driven by the control law of Equation (6.5.4), one can verify that the equation for the propagation of the costate variable, Equation (6.2.18), is correct for this example.

Now we wish to demonstrate that the algorithm of Section 6.3 will lead to the same result. Since this example is being solved analytically, the integrals can and will be evaluated precisely instead of approximated by the Newton-Cotes integration formulas. Thus, instead of using the approximate expression for $\delta\langle J\rangle$ of Equation (6.3.18)

$$\delta\langle J\rangle = \sum_{i=1}^{r} G(\underline{x}_i)\ \delta U(\underline{x}_i) \tag{6.5.9}$$

we shall use the exact expression of Equation (6.3.10)

$$\delta\langle J\rangle = \int_0^L \overline{G}[\underline{x}(\ell)]\ \delta U[\underline{x}(\ell)]\, d\ell \tag{6.5.10}$$

where

$$\overline{G}(\underline{x}) = \left[\|\dot{\underline{x}}\|\right]^{-1}\left[U(\underline{x}) + \langle \underline{P}(\underline{x}), \underline{b}\rangle\right] \tag{6.5.11}$$

and

$$\|\dot{\underline{x}}\| = \|\underline{f}(\underline{x}) + \underline{b}\ U(\underline{x})\| \tag{6.5.12}$$

Correspondingly, instead of using the quantized solution to the gradient projection algorithm of Equation (4.4.42)

$$\delta c_j = \sum_{i=1}^{r} G(\underline{x}_i)\, D_j[\underline{x}_i, \underline{c}] \tag{6.5.13}$$

we shall use the integral solution

$$\delta c_j = \int_0^L \overline{G}[\underline{x}(\ell)]\, D_j[\underline{x}(\ell), \underline{c}]\, d\ell \tag{6.5.14}$$

Suppose that we are given any control law

$$U_s(\underline{x}) = -K(x_1 + x_2) \quad ; \quad K > 0 \tag{6.5.15}$$

We can evaluate the gradient function $\overline{G}(\underline{x})$ for this control law from Equation (6.5.11). Since $\underline{b}^T = [0, 1]$, it follows from Equations (6.5.15) and (6.5.8) that

$$\overline{G}(\underline{x}) = \frac{1}{2}\left[\|\dot{\underline{x}}\|\right]^{-1}\left\{[-K+1+2K^{-1}+K^{-2}]\,x_2 + [-K+K^{-1}]\,x_1\right\} \tag{6.5.16}$$

Since there is only one coefficient K to vary, we identify $c_1 = K$ and $\delta c_1 = \delta K$ and compute $D_1[\underline{x}(\ell), \underline{c}]$.

$$D_1[\underline{x}(\ell), \underline{c}] = \frac{\partial U_s[\underline{x}(\ell), \underline{c}]}{\partial c_1} = \frac{d}{dK}\left\{-K[x_1(\ell)+x_2(\ell)]\right\} = -[x_1(\ell)+x_2(\ell)] \tag{6.5.17}$$

Inserting Equations (6.5.16) and (6.5.17) into (6.5.14) yields

$$\delta K = -\frac{1}{2}\int_0^L \left[\|\dot{\underline{x}}(\ell)\|\right]^{-1}\left\{[x_1(\ell)+x_2(\ell)]\,[(-K+K^{-1})\,x_1(\ell) + (-K+1+2K^{-1}+K^{-2})\,x_2(\ell)]\right\} d\ell \tag{6.5.18}$$

Converting from an integral over the arc length of the trajectory to an integral over time gives

$$\delta K = -\frac{1}{2} \int_0^{\infty} \Big\{ [-K + K^{-1}] x_1^2(t) + [-K + 1 + 2K^{-1} + K^{-2}] x_2^2(t)$$
$$+ [-2K + 1 + 3K^{-1} + K^{-2}] x_1(t) x_2(t) \Big\} dt \qquad (6.5.19)$$

If the above integrals are evaluated (by the Laplace transform technique described previously) along the trajectory generated by the dynamical system of Equation (6.5.1) driven by the control law of Equation (6.5.15) from the initial condition $(\hat{x}_1, \hat{x}_2)$, the result is

$$\delta K = \frac{1}{4K^3} \Big\{ (\hat{x}_1 + \hat{x}_2)^2 K^3 - [(\hat{x}_1 + \hat{x}_2)^2 + \hat{x}_2^2] K - [2\hat{x}_2^2] \Big\} \qquad (6.5.20)$$

Comparing this with Equation (6.5.6) we conclude

$$\delta K = \frac{\partial J[\hat{x}_1, \hat{x}_2, K]}{\partial K} = \nabla \langle J(K) \rangle \qquad (6.5.21)$$

as was the claim of the gradient projection algorithm. Since $-\delta K$ always points in the direction of steepest descent in the coefficient space and is zero only at the optimal solution K^*, convergence to this optimal solution is guaranteed for this problem.

The preceeding example was solved analytically to avoid numerical approximations. Although time domain methods were used to evaluate the integrals, the essential "logic" of the algorithms was preserved and their validity demonstrated.

CHAPTER VII

STABILITY ANALYSIS

7.1 Introduction

This chapter is concerned with formulating an explicit computational algorithm for solving the stability problem. The algorithm is developed in Section 7.2. A two-dimensional example is presented in Section 7.3 which illustrates the application of the algorithm and the properties of the stability bounds.

7.2 Solution of the Stability Problem

In this section we shall state an explicit computational algorithm for solving the stability problem formulated in Chapter II. This algorithm is a direct application of Theorem 3.5 of Chapter III.

Stability Problem Algorithm

1) Select a SOP with optimal feedback control $U^*(\underline{x})$ and optimal cost $J^*(\underline{x})$.

2) Determine $\gamma = \underset{\underline{x} \in Q}{\text{Max}} \; [J^*(\underline{x})]$

3) Define the set $\Omega = \{\underline{x} : \; J^*(\underline{x}) \leq \gamma\}$

4) Select the implementation set $I = \{\underline{x}: \; \hat{I}_i \leq x_i \leq I_i \; ; \; i = 1, \ldots, n\}$ where the 2n scalars $\hat{I}_i$ and I_i are defined as

$$\hat{I}_i \leq \underset{\underline{x} \in \Omega}{\text{Min}} \; [x_i]$$

$$I_i \geq \underset{\underline{x} \in \Omega}{\text{Max}} \; [x_i]$$

5) Compute the stability bounds $T(\underline{x})$ and $B(\underline{x})$ for all $\underline{x} \in \Omega$ from Equations (3.5.26) and (3.5.27).

6) Use Theorem 3.5 to conclude asymptotic stability for the system $\dot{\underline{x}} = \underline{f}(\underline{x}) + \underline{b}\, U_s(\underline{x})$ for all trajectories starting from initial conditions $\underline{x}_o \in Q$ if

$$T(\underline{x}) \underset{\ominus}{\geq} U_s(\underline{x}) \underset{\ominus}{\geq} B(\underline{x}) \qquad \forall \underline{x} \in \Omega$$

The stability algorithm requires the numerical values of both the optimal control $U^*(\underline{x})$ and the optimal cost $J^*(\underline{x})$ as a function of the state; the author assumes that this information would be available in a standard dynamic programming table. The values of $J^*(\underline{x})$ are only needed to specify I; the values of $U^*(\underline{x})$ are sufficient to compute the stability bounds. Since the process of specifying I and evaluating the stability bounds only requires the evaluation of a simple mathematical expression at each data point, the computational time required to perform the algorithm should be extremely small.

The preceeding algorithm was based on Theorem 3.5 and is the best result which can be obtained with the techniques developed in this chapter for establishing stability bounds. However for any particular $U_s(\underline{x})$, say $U_1(\underline{x})$, <u>the size of the implementation set I_1 as determined by this algorithm will usually be somewhat larger than necessary</u>. One would like to be able to determine the implementation set of <u>minimal size</u> $\hat{I}_1 \subset I_1$ for which the system $\dot{\underline{x}} = \underline{f}(\underline{x}) + \underline{b}\, U_1(\underline{x})$ would have asymptotic stability for all $\underline{x}_o \in Q$. Theorem 3.6 can be used to establish this if one can determine $\hat{\Omega}_1$, the region of the state space actually spanned by all trajectories of this system starting from each and every initial condition $\underline{x}_o \in Q$. A determination of $\hat{\Omega}_1$ can usually be accomplished by generating a few trajectories starting from critical initial conditions on the boundry of Q and utilizing the fact that trajectories can never intersect. However, developing an explicit, efficient procedure for determining $\hat{\Omega}_1$ depends strongly on the pecularities of the particular problem under consideration and will not be pursued further in this thesis.

7.3 A Two-Dimensional Example

In order to illustrate the application of the preceeding algorithm as well as provide an indication of the nature of the stability bounds, the following example is presented. Consider the dynamical system

$$\dot{x}_1 = x_2$$

$$\dot{x}_2 = U(\underline{x})$$

and the cost functional

$$J = \int_0^\infty \left[\frac{1}{2}x_1^2 + x_2^2 + \frac{1}{2}U^2(\underline{x})\right] dt$$

with

$$Q = \{\underline{x} : \ -4 \leq x_1 \leq 4, \ -4 \leq x_2 \leq 4\}$$

for which

$$U^*(\underline{x}) = -x_1 - 2x_2$$

and

$$J^*(\underline{x}) = x_1^2 + x_2^2 + x_1 x_2$$

It is readily established that

$$\gamma = \underset{\underline{x} \in Q}{\text{Max}} [J^*(\underline{x})] = 48$$

Hence

$$\Omega = \{\underline{x}: x_1^2 + x_2^2 + x_1 x_2 \leq 48\}$$

and from the plot of Q and Ω shown in Figure 7.1, it is evident that a reasonable choice for the implementation set I would be $I = \{\underline{x}: -8 \leq x_1 \leq 8; -8 \leq x_2 \leq 8\}$. The stability bounds can be readily computed and are given by the following expressions

$\{\underline{x}: \|\underline{x}\| = 0\}$	$T(\underline{x}) = 0$	$B(\underline{x}) = 0$
$\{\underline{x}: -x_1 - 2x_2 = 0; \|\underline{x}\| \neq 0\}$	$T(\underline{x}) = +\infty$	$B(\underline{x}) = -\infty$
$\{\underline{x}: -x_1 - 2x_2 > 0\}$	$T(\underline{x}) = +\infty$	$B(\underline{x}) = -\phi(\underline{x})$
$\{\underline{x}: -x_1 - 2x_2 < 0\}$	$T(\underline{x}) = -\phi(\underline{x})$	$B(\underline{x}) = -\infty$

where

$$\phi(\underline{x}) = \frac{2x_1 x_2 + x_2^2}{x_1 + 2x_2}$$

The implementation set I is again shown in Figure 7.2. The dashed line represents the locus $U^*(\underline{x}) = 0 = -x_1 - 2x_2$; $U^*(\underline{x})$ is negative at all points above and to the right of this line and positive at all points below and to the left. Four additional lines denoted A, B, C, and D are also indicated. For each of these lines the stability bounds $T(\underline{x})$ and $B(\underline{x})$ along with the optimal control $U^*(\underline{x})$ are plotted in Figure 7.3. Any control $U_s(\underline{x})$ which does not enter a shaded area will produce on asymptotically stable system. The plots in Figure 7.3 illustrate the important characteristics of the stability bounds (discussed in Section 3.6) and are typical of the results one can expect for a general system. The bounds are generally quite broad and establish asymptotic stability for a wide variety of suboptimal controls. In particular, any $U_s(\underline{x})$ which is uniformly "close" to $U^*(\underline{x})$ over the control space can be expected to be asymptotically stable, thus justifying design procedures based on fitting $U_s(\underline{x})$ to $U^*(\underline{x})$.

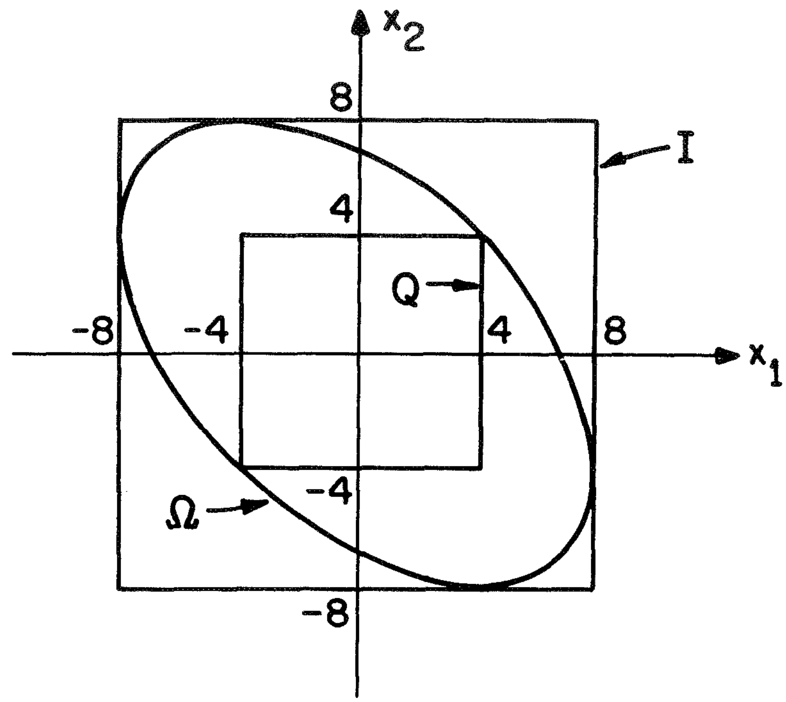

Figure 7.1 Plot of Q, Ω, and I

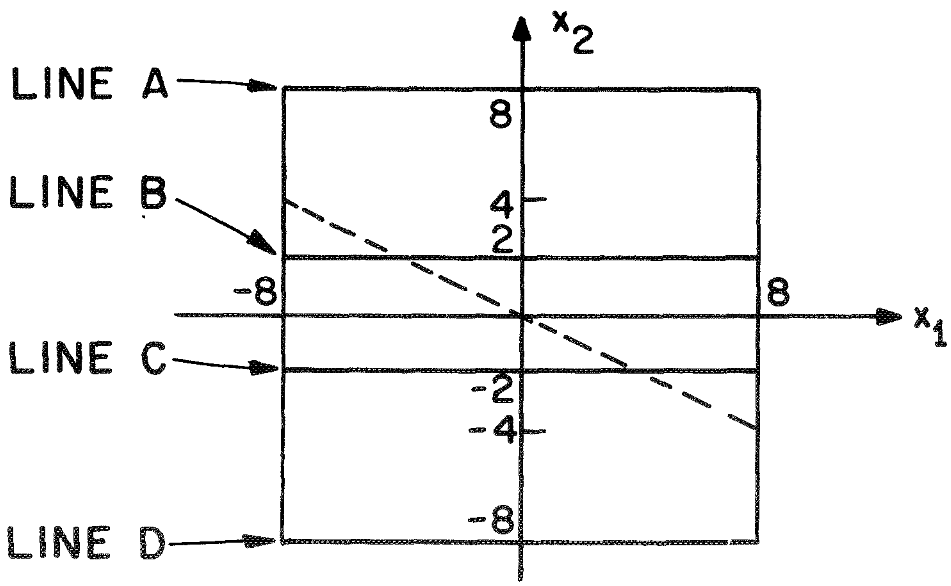

Figure 7.2 The Implementation Set I

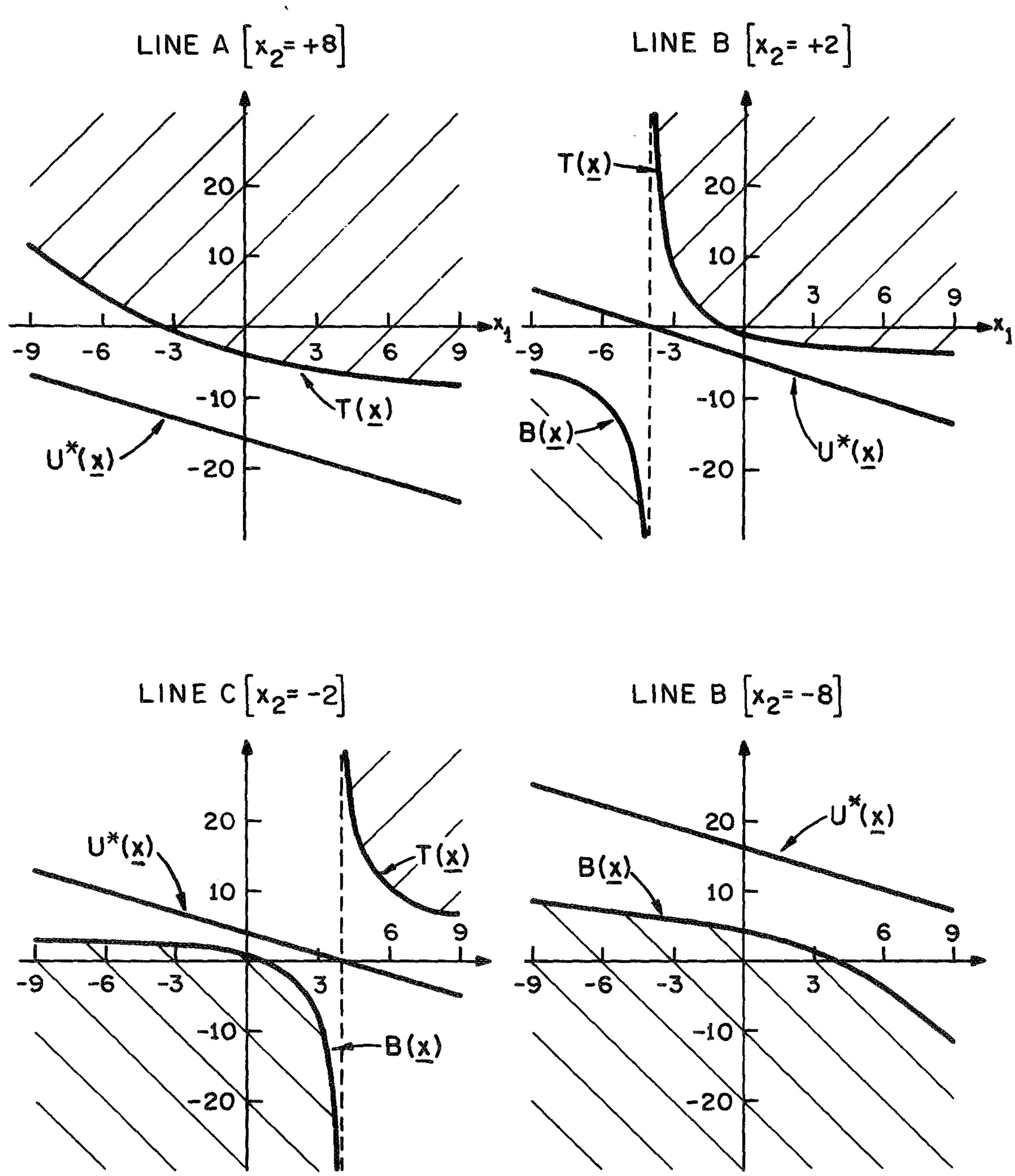

Figure 7.3 Stability Bounds

CHAPTER VIII

SUMMARY AND CONCLUSIONS

The design of practical, easily implemented, feedback controllers for nonlinear dynamical systems has been considered in this thesis. Structural constraints which restrict the feedback control law to belong to specific classes of feedback structures were introduced and imposed upon the optimization problem in order to guarantee that every control law designed subject to these constraints would be in a form for which there is a simple and direct means of implementation. The structurally constrained optimal control problem (SCOCP) was formulated and the specific peculiarities which result from the imposition of structural constraints upon optimization problems were analyzed and discussed.

The major theoretical contributions of the thesis were developed in Chapter III where we investigated certain properties of optimal and suboptimal systems. Two significant results were established which not only provide considerable insight into the properties of suboptimal control systems but, in addition, provide a direct basis for constructing a variety of suboptimal design procedures. First, a simple mathematical expression relating the suboptimality of a control law to the corresponding suboptimality of its cost was derived and used to establish a bound on the suboptimal cost as stated in Theorem 3.1. The importance of this bound becomes evident when specific numerical values are considered — for example, a suboptimal control which is everywhere within 10% of the optimal will produce a suboptimal cost which is everwhere less than

1.25% above the optimal cost. Second, we established the existence of scalar stability bounds which specify at each and every point in the state space a range over which the scalar suboptimal control can vary and still produce a system which is asymptotically stable. These stability bounds were used to prove Theorem 3.2 which guarantees "asymptotic stability if $U_s(\underline{0}) = 0$ and $U_s(\underline{x})$ has everywhere the same polarity and at least half the magnitude of $U^*(\underline{x})$". This theorem and the stability bounds from which it was derived clearly demonstrate the wide range over which the scalar value of a suboptimal control can vary and still produce an asymptotically stable system. The significance of these results is twofold: First, they provide a rigorous justification for using suboptimal controls and suboptimal design procedures; second, they provide an explicit bound for evaluating the extent or degree of this suboptimality. The final portion of this chapter was devoted to proving Theorem 3.3 (which establishes that all optimal SOP systems have certain stability properties in common) and exploring the implications of this theorem.

Algorithms for the optimal and suboptimal solution of the SCOCP were developed in Chapters V and VI. Two basic suboptimal design procedures were formulated; both require the numerical value of $U^*(\underline{x})$ as precomputed input data. The SDP1 algorithm requires only a relatively minor amount of computation and has produced control laws whose cost was found to be within 2% of the optimal constrained cost in all of the cases tested. However, it is restricted to problem formulations in which $Q_o(\underline{x})$ will generate trajectories which tend to be uniformly distributed throughout Ω_o. The SDP2 algorithm is valid for completely general initial condition probability distributions. It has produced control laws whose cost was found to be within .5% of the optimal constrained cost in all of the cases tested; however, the computational requirements of this algorithm are substantially greater (by a factor of 5 to 10) than those of the SDP1 algorithm. Unlike the suboptimal techniques, the optimal design procedure does not require the numerical values of $U^*(\underline{x})$ or any other precomputed information and should converge to the optimal constrained solution. However, its computational requirements are extensive — from 3 to 5 times those of the SDP2 algorithm.

The author has drawn two major conclusions from this research. First, the utilization of suboptimal control laws and suboptimal design procedures are definitely justified for optimal control problems of the SCOCP type. Second, the incorporation of structural constraints directly into the optimization problem is not only possible but computationally feasible; hence, it is possible to determine solutions to the optimization problem in a "meaningful" form which can be directly used to implement the control law.

There are numerous extentions and applications of the theory developed in this thesis which merit further investigation. They range from purely theoretical to entirely numerical studies. We shall enumerate these additional research problems while categorizing them for the benefit of future investigators.

THEORETICAL STUDIES

1) Vector Control Problem

The SCOCP problem formulation considered in this thesis was restricted to scalar controls. It can be generalized to the vector case by considering dynamical systems of the form

$$\dot{\underline{x}} = \underline{f}(\underline{x}) + \underline{B}\ \underline{U}(\underline{x})$$

where

$\underline{x}$ and $\underline{f}$ are n vectors

$\underline{U}$ is an m vector

$\underline{B}$ is an n x m matrix

with cost functionals

$$J = \int_0^{\infty} \left[g(\underline{x}) + \frac{1}{2} \| \underline{U}(\underline{x}) \|^2 \right] dt$$

One should be able to follow the approach of Chapter III to develop equivalent results for the vector case. The author would conjecture that Theorem 3.4 would remain valid if the cost functional J_2 were converted to

$$J = \int_0^{\infty} \left[\| \underline{U}^*(\underline{x}) - \underline{U}_s(\underline{x}) \| \right]^2 dt$$

and that Theorem 3.1 would also remain valid if the condition specifying γ were changed to

$$\| \underline{U}^*(\underline{x}) - \underline{U}_s(\underline{x}) \| \leq \gamma \| \underline{U}^*(\underline{x}) \|$$

Similiarly, he would also conjecture that stability bounds of the form

$$T(\underline{x}) \gtreqless \left\langle \underline{U}^*(\underline{x}), \underline{U}_s(\underline{x}) \right\rangle \gtreqless B(\underline{x})$$

could be developed for establishing the asymptotic stability of such systems. The validity of these conjectures is currently under investigation by the author.

In addition, one should be able to extend all of the optimal and sub-optimal design procedures to the vector case. Of course, a separate S(M, N, K) structure would be required for implementing each component of the control law. This remains a topic for future research.

2) Finite Time Problem

Suppose the infinite time SCOCP problem formulation were changed into that of a free end-point, finite time problem. It would be of considerable interest to determine which, if any, of the theorems of Chapter III could be converted into equivalent results which would be applicable to this finite time problem.

3) Inverse Problem for Optimal SOP Systems

In Section 3.8 we investigated the extent to which Theorem 3.3 could be used to characterize optimal SOP systems and concluded that it was a necessary but not sufficient condition. Nevertheless, the characterization was quite strong and came close to solving the "inverse problem of optimal control for SOP systems". Many unanswered questions remain. The author believes that further research is needed to determine whether extensions of the theory developed in this thesis can be used to either solve the "inverse problem" or deduce additional properties of optimal SOP systems.

4) Stability Theory

A problem which has attracted considerable interest in the literature of the last few years is that of determining the stability of feedback systems containing a single, time-varying, nonlinear element as shown in Figure 8.1.

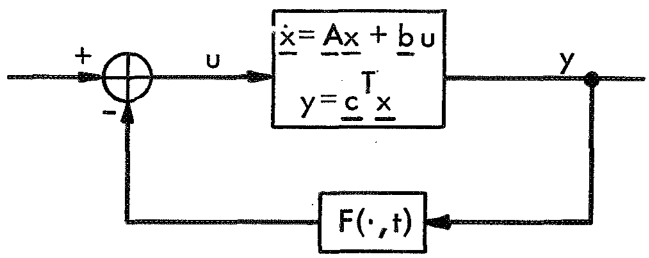

Figure 8.1 Nonlinear Feedback System

If the transfer function $\frac{y(s)}{u(s)}$ does not enter or encircle the circle whose diameter extends from the point $s = 0$ to $s = -\alpha$, then Kalman's solution to the inverse problem of optimal control states that $\gamma y = \gamma \underline{c}^T \underline{x}$ is the optimal feedback control law for some linear quadratic problem for all $\gamma \geq \frac{2}{\alpha}$. Using Theorem 3.3, we can prove that the above system will be ASIL for all $F[\cdot, t]$ satisfying

$$F[0, t] = 0 \qquad \forall t \in [0, \infty)$$

$$\frac{F[\sigma, t]}{\sigma} \geq \frac{1}{\alpha} \qquad \forall t \in [0, \infty),\ \sigma \neq 0$$

The above result is identical with the Circle Theorem of I.W. Sandberg.[50] However, if the stability bounds of Theorem 3.5 and 3.6 are used instead of using Theorem 3.3 (where the term

$$\frac{g(\underline{x})}{|U^*(\underline{x})|}$$

was set to zero), a stronger result than the Circle Theorem can be obtained. An investigation of the applicability of this procedure should prove an interesting topic for future research.

NUMERICAL STUDIES

1) Computerize Algorithms

Develop a computer program for the SDP1, SDP2, and optimal algorithms. Determine the best method for generating the system trajectories. Test the algorithms on a number of problems and construct a table giving an estimate of the computational time required by each algorithm as a function of the number of state variables and initial condition quantizations.

2) Analyze S(M, N, K) Structures

Use the computer programs to solve the same optimization problems for a variety of the S(M, N, K) structures. Try to determine which structures produce the best results with the least complexity. Try to correlate the various structures with the types of control laws or control problems for which they are best suited.

3) Hybrid Algorithm

Develop a computer program for the algorithm obtained by combining the SDP1 and SDP2 algorithms in the manner indicated in Section 5.6 on page 101. Compare its computational requirements and accuracy with those of the SDP1 and SDP2 algorithms.

The previous suggestions are by no means exhaustive of the possible extentions of this thesis; they are intended merely to indicate some of the more attractive possibilities.

BIBLIOGRAPHY

1. Bridgland, T.F., "On the Existence of Optimal Feedback Controls" J. SIAM on Control, Series A, Vol. 2, 1963, pp. 261-274.

2. Pontryagin, L.S., The Mathematical Theory of Optimal Processes, J. Wiley and Sons, N.Y., 1962.

3. Athans, M., "The Status of Optimal Control Theory and Applications for Deterministic Systems," IEEE Trans. on Automatic Control, Vol. AC-11, No. 3, July, 1966, pp. 580-596.

4. Bellman, R., Dynamic Programming, Princeton University Press, Princeton, N.J., 1957.

5. Bellman, R., and Dreyfus, S., Applied Dynamic Programming, Princeton University Press, Princeton, N.J., 1962.

6. Durbeck, R.C., "An Approximation Technique for Suboptimal Control", IEEE Trans. on Automatic Control, Vol. AC-10, No. 2, April, 1965, pp. 144-150.

7. Friedland, B., "A Technique of Quasi-Optimum Control", IEEE Trans. on Automatic Control, Vol. AC-11, No. 3, July, 1966.

8. Melsa, J.L., "Closed-Loop, Sub-Optimal Control Employing the Second Method of Liapunov", Report CSL-65-1, Control Systems Laboratory, University of Arizona, Tucson, Arizona, Jan., 1965.

9. Rekasius, Z.V., "Suboptimal Design of Intentionally Nonlinear Controllers", IEEE Trans. on Automatic Control, Vol. AC-9, No.4, October, 1964, pp. 380-386.

10. Smith, F.W., "Design of Quasi-Optimal Minimum-Time Controllers", IEEE Trans. on Automatic Control, Vol. AC-11, No. 1, January, 1966, pp. 71-77.

11. Paiewonsky, B., "Optimal Control: A Review of Theory and Practice", AIAA Journal, Vol. 3, No. 11, Nov., 1965, pp. 1985-2006.

12. Dreyfus, S.E., Dynamic Programming and the Calculus of Variations, Academic Press, N.Y., 1965.

13. Athans, M. and Falb, P.L., Optimal Control: An Introduction to the Theory and Its Applications, McGraw-Hill Book Co., N.Y., 1966.

14. Kleinman, D., "Suboptimal Design of Linear Regulator Systems Subject to Computer Storage Limitations", Electronic Systems Laboratory Report ESL-R-297(MIT), Feb., 1967.

15. Kopp, R.E., McGill, R., Moyer, H.G., and Pinkham, G., "Several Trajectory Optimization Techniques", in Computing Methods in Optimization Problems, A.V. Balakrishnan and L.W. Neustadt, Editors, Academic Press, N.Y., N.Y., 1964.

16. Kalman, R.E., and Koepke, R.W., "Optimal Synthesis of Linear Sampling Control Systems using Generalized Performance Indicees", Trans. ASME, Vol. 80, Nov., 1958, pp. 1820-1838.

17. Bellman, R., Glicksberg, I., and Gross, O., "On the 'bang-bang' control problem", Q. Appl. Math, Vol. 14, No. 1, 1956, pp. 11-18.

18. Larson, R.E., State Increment Dynamic Programming, New York: American Elsevier, 1967.

19. Larson, R.E., "A Survey of Dynamic Programming Procedures", IEEE Trans. on Automatic Control, Vol. AC-12, Dec., 1967, pp. 767-774.

20. Neustadt, L.W., "Synthesizing Time-Optimal Control Systems", Journal Math. Anal. Appl., Vol. 1, 1960, pp. 484-493.

21. Melsa, J.L., and Schultz, D.G., "A Closed-Loop Approximately Time-Optimal Control Method for Linear Systems", IEEE Trans. on Automatic Control, Vol. AC-12, No. 1, Feb., 1967, pp. 94-97.

22. Saridis, G.N., and Rekasius, Z.V., "Design of Approximately Optimal Feedback Controllers for Systems with Bounded States", IEEE Trans. on Automatic Control, Vol. AC-12, No. 4, August, 1967, pp. 373-379.

23. Longmuir, A.G., and Bohn, E.V., "The Synthesis of Suboptimal Feedback Control Laws", IEEE Trans. on Automatic Control, Vol. AC-12, No. 6, December, 1967, pp. 755-758.

24. Merriam, C.W., Optimization Theory and the Design of Feedback Control Systems, McGraw-Hill, New York, 1964, p. 238.

25. Kolmogorov, A.N., "On the Representation of Continuous Functions of Several Variables as the Superposition of Continuous Functions of One Variable and of Addition", Doklady, An USSR, Vol. 114, No. 5, 1957.

26. Kolmogorov, A.N., DAN USSR, Vol. 108, No. 2, 1956.

27. Arnol'd, V.I., DAN USSR, Vol. 114, No. 4, 1957.

28. Sprecher, D.A., "On the Structure of Continuous Functions of Several Variables", Ph.D. Thesis, Dept. of Mathematics, University of Maryland, 1963.

29. Sprecher, D.A., "On the Structure of Continuous Functions of Several Variables", Trans. American Math Society, Vol. 115, Issue 3, March, 1965, pp. 340-355.

30. Sprecher, D.A., "A Representation Theorem for Continuous Functions of Several Variables", Proc. of the American Math Society, Vol. 16, No. 2, April, 1965, pp. 200-203.

31. Sprecher, D.A., "On the Structure of Representations of Continuous Functions of Several Variables as Finite Sums of Functions of One Variable", Proc. of the American Math Society, Vol. 17, No. 1, February, 1966, pp. 98-105.

32. Sprecher, D.A., "On Admissibility in Representations of Functions of Several Variables as Finite Sums of Functions of One Variable", Article in Numerical Solution of Partial Differential Equations, Academic Press, 1966.

33. Vituskin, A.G., Theory of Transmission and Processing of Information, Permagon Press, New York, 1961.

34. Vituskin, A.G., "Proof of existence of analytic functions of several variables, not representable by linear superpositions of continuously differential functions of fewer variables", Dokl., Vol. 156, 1964.

35. Vituskin, A.G., "Some properties of linear superpositions of smooth functions", Dokl., Vol. 156, 1964.

36. Henkin, G.M., "Linear Superpositions of Continuously Differential Functions", Dokl., Vol. 157, 1964.

37. Lorentz, G.G., Approximation of Functions, Holt, Rinehart, and Winston, N.Y., N.Y., 1966.

38. Lorentz, G.G., "Metric Entropy, Widths, and Superpositions of Functions", American Math Monthly, No. 69, 1962.

39. Lorentz, G.G., "Metric Entropy and Approximation", Bulletin of the American Math Society, November, 1966, Vol. 72, No. 6, pp. 903-937.

40. Kalman, R.E., "When is a Linear Control System Optimal", Journal of Basic Engineering, March, 1964, pp. 51-60.

41. Lee and Markus, Foundations of Optimal Control Theory, John Wiley and Sons, New York, 1967.

42. Larson, R.E., "Optimum Guidance Laws for an ABM", Stanford Research Institute, SRI Project 5188, Jan., 1966.

43. Levine, W.S., "Optimal Output-Feedback Controllers for Linear Systems", Electronic Systems Laboratory Report ESL-R-374(MIT), Feb., 1969.

44. Debs, A.S., "Optimal and Suboptimal Feedback Control for Deterministic Systems", Electronic Systems Laboratory Report ESL-TM-375(MIT), Feb., 1969.

45. Jacobson, D.H., "Differential Dynamic Programming Methods for Determining Optimal Control of Non-linear Systems", Control Systems Research Report, Centre for Computing and Automation, Imperial College, University of London, 1966.

46. LaSalle, J.P., "The Extent of Asymptotic Stability", Proc. N.A.S., Vol. 46, 1960, pp. 363-365.

47. Athans, M.; Debs, A.S.; and Joseph, K.M.; "On Approximation of Functions, Topological Mappings, and Nonlinear Feedback Control Systems", Electronic Systems Laboratory Report ESL-R-389 (MIT), June, 1969.

48. Rosen, J.B., "The Gradient Projection Method for Nonlinear Programming. Part I. Linear Constraints", J.Soc. Indust. Appl. Math., Vol. 8, No. 1, 1960, pp. 181-217.

49. Rosen, J.B., "The Gradient Projection Method for Nonlinear Programming. Part II. Nonlinear Constraints", J. Soc. Indust. Appl. Math., Vol. 9, No. 4, 1961, pp. 514-532.

50. Sandberg, I.W., "A Frequency Domain Condition for the Stability of Systems Containing a Single Time-Varying Nonlinear Element". Bell Systems Technical Journal, Vol. 43, page 1601, 1964.

51. Zames, G., "On the Input-Output Stability of Time-Varying Nonlinear Feedback Systems - Part II", IEEE Trans. on Automatic Control, Vol. AC-11, No. 3, 1966, pp. 465-476.

52. Rudin, W., Principles of Mathematical Analysis, McGraw-Hill Book Company, New York, 1964.

53. Hildebrand, F.B., Introduction to Numerical Analysis, McGraw-Hill Book Company, New York, 1956.

54. Lowan, A.N., and H. Salzer; "Table of Coefficients in Numerical Integration Formulae", J. Math. and Phys., 24:1-21, 1943.

55. Newton, G.C., Gould, L.A., and Kaiser, J.F., Analytical Design of Linear Feedback Controls, John Wiley and Sons Inc., New York, 1961.

MASSACHUSETTS INSTITUTE OF TECHNOLOGY