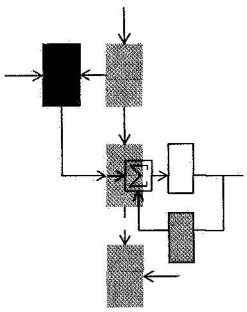
N69-21043
CR-10047

February, 1969

REPORT ESL-R-374
M.I.T. DSR Project 76265
NASA Grant NGL-22-009-124



OPTIMAL OUTPUT-FEEDBACK CONTROLLERS FOR LINEAR SYSTEMS

William S. Levine

Electronic Systems Laboratory

MASSACHUSETTS INSTITUTE OF TECHNOLOGY, CAMBRIDGE, MASSACHUSETTS 02139

Department of Electrical Engineering

February, 1969 Report ESL-R-374 Copy No.____

OPTIMAL OUTPUT-FEEDBACK CONTROLLERS FOR LINEAR SYSTEMS

by

William S. Levine

This report consists of the unaltered thesis of William S. Levine, submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at the Massachusetts Institute of Technology in January, 1969. This research was carried out at the M.I.T. Electronic Systems Laboratory with support extended by the National Aeronautic and Space Administration under Research Grant No. NGL-22-009(124), M.I.T. DSR Project No. 76265.

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

OPTIMAL OUTPUT-FEEDBACK CONTROLLERS FOR LINEAR SYSTEMS

by

WILLIAM SILVER LEVINE

Submitted to the Department of Electrical Engineering on January 22, 1969 in partial fulfillment of the requirements for the Degree of Doctor of Philosophy.

ABSTRACT

This research is concerned with the optimal control of linear systems with respect to a quadratic performance criterion. The optimization problem is formulated with the additional constraint that the control vector $\underline{u}(t)$ is a linear function of the output vector $\underline{y}(t)$ ($\underline{u}(t) = -\underline{F}(t)\underline{y}(t)$) rather than of the state vector $\underline{x}(t)$. The optimal feedback matrix $\underline{F}^*(t)$ is then chosen to minimize an "averaged" quadratic performance criterion.

The necessary conditions provided by the matrix minimum principle are used to determine the optimal feedback gain matrix $\underline{F}^*(t)$. This $\underline{F}^*(t)$ is then shown to satisfy the Hamilton-Jacobi equation thereby demonstrating that it is at least locally optimal. In addition, the existence of an optimal feedback gain matrix is proven.

A computer algorithm is developed to facilitate the calculation of $\underline{F}^*(t)$ for practical problems. This algorithm is programmed and used in the solution of several examples.

Finally, a time-invariant version of the above problem is formulated and solved. Again an algorithm for computing $\underline{F}^*$ (in this case, a constant matrix) is suggested. In addition, several examples are solved.

Thesis Supervisor: Michael Athans
Title: Associate Professor of Electrical Engineering

ACKNOWLEDGMENT

It is a great pleasure to be able to thank Prof. Michael Athans for his many contributions to the author's education. His encouragement and enthusiasm, his technical insight and suggestions have played a fundamental role in my graduate training. He has been an "optimal" thesis advisor.

It is also a pleasure to thank Professors Roger Brockett and Leonard A. Gould for their many constructive criticisms and comments while serving as readers for this thesis. Their suggestions have enriched the technical contents and improved the exposition of this research.

Many of the author's colleagues have also contributed to these results. In particular, Prof. J. C. Willems, Prof. A. H. Levis, Dr. D. L. Kleinman, T. Fortmann, S. G. Greenberg and J. H. Davis deserve special mention.

My wife Shirley has made essential, although entirely non-technical, contributions to this research.

I would also like to thank Carol Hewson of the ESL Publications Staff for her rapid and accurate typing of the final report and Arthur Giordani of the ESL Drafting Department for his help with the figures.

This research was carried out at the M. I. T. Electronic Systems Laboratory with support provided in part by the NASA Electronic Research Center under Grant NGL-22-009(124), and in part by the U.S. Department of Transportation under Contract C-85-65.

CONTENTS

CHAPTER I	INTRODUCTION page	1
CHAPTER II	THEORETICAL RESULTS - OPTIMAL OUTPUT FEEDBACK ON A FINITE INTERVAL	6
2. 1	Problem Formulation	6
2. 2	Statement of the Problem	9
2. 3	The Main Result	10
2. 4	Proof of the Main Result	11
2. 5	Existence and Uniqueness	17
CHAPTER III	COMPUTATION OF THE OPTIMAL FEED-BACK GAIN MATRIX	21
3. 1	Theoretical Algorithm	21
3. 2	The Computer Program	27
3. 3	Examples	29
CHAPTER IV	OPTIMAL TIME-INVARIANT OUTPUT-FEEDBACK PROBLEMS	46
4. 1	The Limiting Case, $T \rightarrow \infty$	46
4. 2	Reformulation of the Problem	49
4. 3	Solution Assuming $\underline{C} = \underline{I}$	50
4. 4	The Main Result	58
4. 5	Examples	66
CHAPTER V	CONCLUSIONS	71
APPENDIX A		73
APPENDIX B		78
REFERENCES		89

LIST OF FIGURES

1	"Optimal" Trajectories for Example 1 Plotted in the Phase Plane	page 31
2	Optimal Feedback Gain for Example 1a	32
3	Optimal Feedback Gain for Example 1b	32
4	Optimal Feedback Gain for Example 1c	32
5	"Optimal" Trajectories for Example 2 Plotted in the Phase Plane	36
6	Optimal Feedback Gain for Example 2	37
7	Optimal Feedback Gain for Example 2 with $T = 6$	37
8	The Ellipses on which the States are Distributed at $t = 0$, $t = 1$, $t = 1.6$ for the Optimal System of Example 2	38
9	"Optimal" Trajectories for Example 3 Plotted in the Phase Plane	42
10	Optimal Feedback Gain for Example 3	43

CHAPTER I

INTRODUCTION

The purpose of this thesis is to consider methods for the calculation of linear feedback controls for linear systems under the constraints that the control variables depend only on the outputs of the system and that the control be "optimal" in some well-defined sense. The approach that is taken is to create a precisely defined mathematical problem that corresponds to the rather vague physical problem above. This mathematical problem is then solved and its solutions interpreted physically. Before proceeding with this, the history and significance of the physical problem and some previous mathematical results are reviewed.

The problem of calculating linear feedback controls for linear systems has been one of the most widely studied problems in control theory for at least 35 years.[1,2] During these 35 years the theoretical techniques needed to design linear, time-invariant feedback controls for single-input, single-output, linear, and time-invariant systems has been very well developed. Furthermore, this theory has been used to design many systems that are in operation today. This same theory has also been applied with some success to multiple input, multiple-output, time-invariant, linear systems. However, the classical theory does not apply to time-varying linear systems. Furthermore, the classical theory cannot be applied to many multiple-input, multiple-output, time-invariant linear systems.

Meanwhile, beginning with Wiener's work on stationary time series and linear filtering and prediction problems,[3] interest has developed in the so-called "linear regulator problem."[4,5] Basically, the "linear regulator problem" is to find a control input to a linear system which minimizes the sum of the integral squared error and control energy. It happens that the solution of this problem is a linear feedback control. Thus, this "linear regulator problem" is closely related to the problem of calculating linear feedback controls for linear systems.

In the twenty years since its inception, the "linear regulator problem" has also been extensively studied. And, some remarkable theoretical and practical results have been obtained. In particular, the results obtained for this problem by R. E. Kalman[5,6,7,8] provide crucial background for this thesis. To briefly review Kalman's results, he begins with the linear system

$$\underline{\mathring{x}}(t) = \underline{A}(t)\underline{x}(t) + \underline{B}(t)\underline{u}(t) \tag{1.1}$$

and the performance criterion

$$J = \frac{1}{2}\underline{x}'(T)\underline{S}\,\underline{x}(T) + \frac{1}{2}\int_{t_0}^{T} [\underline{x}'(t)\underline{Q}(t)\underline{x}(t) + \underline{u}'(t)\underline{R}(t)\underline{u}(t)]\,dt \tag{1.2}$$

where $\underline{x}(t)$ is the state of the system and $\underline{u}(t)$ is the control. He then finds that the optimal control $\underline{u}^*(t) = -\underline{R}^{-1}(t)\underline{B}'(t)\underline{K}^*(t)\underline{x}^*(t)$ where $\underline{K}^*(t)$ is the solution of a matrix differential equation, the matrix Riccati equation. Note that this optimal control is a feedback control. Furthermore, if $T \to \infty$, $\underline{S} = 0$ and the system is time-invariant, completely controllable and observable, this feedback gain matrix is also time-invariant. This is a truly elegant result. If does have the practical drawback, however, that the feedback control depends on the <u>entire</u> state of the system. As a result, in practical applications it is necessary to augment the measurements of the state (the outputs) by either a Kalman filter[9] or some other state reconstructor.[10,*]

When one combines these two intimately related lines of research one sees an interesting gap. Classically, engineers have been quite successful using only output feedback, and in some cases, dynamic compensation. On the other hand, the "linear regulator problem" is not suited to the design of output feedback controls unless the output is equivalent to the state. Thus, there is a large class of practical problems for which the available theory could be improved. Specifically, the class of linear time-varying or time-invariant systems whose state vector has many more components than its output vector. The purpose of this thesis is to attempt to extend the available theory to cover as much of the above class of problems as possible.

There is a great deal of previous research that is applicable to the above class of problems. This research can be divided into three major groups:

1) Some of the early research on the "linear regulator problem" and on the optimization of the parameters in a system with fixed configuration is applicable to the above problem for time-invariant systems. The work of Newton, Gould and Kaiser[22] is an early example of this approach. Other examples are given and referenced by

---

*For the reader who wants an excellent treatise on Kalman's results augmented by some excellent research of his own on the same problem, the report by D. Kleinman[11] is highly recommended.

Willis.[23] The difficulty with these results has been that they are dependent on the initial conditions of the system. Thus, the results are not really feedback controls, nor, as it happens, do they apply to time-varying systems.

2) There has been some direct research on the relation between the approach listed in (1), the Kalman linear regulator and the Wiener linear regulator. Examples of this include Willis'[23] research and some of Kalman's[6] research. All of the results obtained however, apply only to time-invariant systems.

3) Several people have worked on the specific physical problem posed in this thesis.[24,25] In particular, Rekasius and Ferguson[24] recently published a paper dealing with the physical problem that is discussed herein. They take a completely different approach and obtain completely different results. Their results only apply to systems whose control is a scalar.

The results of this thesis are presented according to the following outline. In Chapter II, the mathematical problem is carefully formulated for linear, possibly time-varying, systems on a finite time interval $[t_0, T]$. Then, the necessary conditions which the solution to this problem must satisfy are derived and used to find the solution. However, this solution is not amenable to simple hand computation and so, in Chapter III, a computer algorithm is developed and programmed. This algorithm is used to solve for the optimal control in several examples. These examples are then analyzed at some length in an attempt to discover properties of optimal systems.

Unfortunately, the results of the first two chapters do not extend to the time-invariant case in precisely the same way as the

Kalman problem. As a result, in Chapter IV, appropriate modifications are made to obtain a time-invariant feedback solution. Necessary conditions, which lead to a set of algebraic equations, are derived and used to find the optimal control. Again, examples are worked and analyzed. Finally, the thesis is concluded with a brief summary of the results obtained and some suggestions for future research in Chapter V.

CHAPTER II

THEORETICAL RESULTS - OPTIMAL OUTPUT FEEDBACK ON A FINITE INTERVAL

As we stated in the introduction, we are interested in calculating linear output feedback controls that are "optimal" in some well-defined sense. In this chapter, we will begin by carefully formulating a precise optimization problem. This optimization problem, and a slight modification of this problem introduced in Chapter IV, will form the basic mathematical problem of this thesis.

Since there already exists a large body of theoretical knowledge about, and practical justification for, quadratic cost criteria applied to linear systems, we would like to use a quadratic type criterion. We show that we can use such a criterion and obtain meaningful results. In addition, our formulation includes the Kalman linear regulator[5] (state feedback) as the special case when the output vector is the state vector.

Once the problem has been formulated, we find its solution by application of the necessary conditions of the matrix minimum principle.[12] We next show that this same control satisfies the Hamilton-Jacobi equation. Finally, we prove that there exists a solution to the problem we have formulated and discuss its uniqueness.

2.1 Problem Formulation

Consider a linear system whose state vector $\underline{x}(t)$, control vector $\underline{u}(t)$ and output vector $\underline{y}(t)$ are related by

$$\underline{\dot{x}}(t) = \underline{A}(t)\underline{x}(t) + \underline{B}(t)\underline{u}(t) \tag{2.1.1}$$

$$\underline{y}(t) = \underline{C}(t)\underline{x}(t) \tag{2.1.2}$$

where:

$\underline{x}(t)$ is a real n-vector

$\underline{u}(t)$ is a real m-vector

$\underline{y}(t)$ is a real r-vector

Consider also the standard quadratic cost functional

$$J = \frac{1}{2}\underline{x}'(T)\underline{S}\,\underline{x}(T) + \frac{1}{2}\int_{t_0}^{T}\left[\underline{x}'(t)\underline{Q}(t)\underline{x}(t) + \underline{u}'(t)\underline{R}(t)\underline{u}(t)\right]dt \tag{2.1.3}$$

It is well known [5] that the optimal control can be generated by $\underline{u}(t) = -\underline{G}(t)\underline{x}(t)$ where the gain matrix $\underline{G}(t)$ can be evaluated through the solution of the Riccati equation.

Now suppose that one introduces the constraint that the control $\underline{u}(t)$ be generated via output linear feedback, i.e.

$$\underline{u}(t) = -\underline{F}(t)\underline{y}(t) \tag{2.1.4}$$

$$\text{or} \quad \underline{u}(t) = -\underline{F}(t)\underline{C}(t)\underline{x}(t) \tag{2.1.5}$$

where $\underline{F}(t)$, the <u>feedback gain matrix,</u> is to be determined. Under this constraint, the system equations (2.1.1 and 2.1.2) become

$$\underline{\dot{x}}(t) = \left[\underline{A}(t) - \underline{B}(t)\underline{F}(t)\underline{C}(t)\right]\underline{x}(t) \tag{2.1.6}$$

Thus, as expected, the choice of the gain matrix $\underline{F}(t)$ will govern the response of the closed-loop system. The closed-loop system response can be written as:

$$\underline{x}(t) = \underline{\Phi}(t, t_0)\underline{x}(t_0) \tag{2.1.7}$$

where $\underline{\Phi}(t, t_0)$ denotes the fundamental transition matrix for the system (2.1.6), defined by

$$\dot{\underline{\Phi}}(t, t_0) = [\underline{A}(t) - \underline{B}(t)\underline{F}(t)\underline{C}(t)]\underline{\Phi}(t, t_0) \ ; \ \underline{\Phi}(t_0, t_0) = \underline{I} \qquad (2.1.8)$$

If we substitute Eqs. (2.1.5) and (2.1.7) into the performance criterion (2.1.3) we deduce that, for any given initial state $\underline{x}(t_0)$ and any given feedback matrix $\underline{F}(t)$, the cost is given by

$$J = \frac{1}{2}\underline{x}'(t_0)\{\underline{\Phi}'(T, t_0)\underline{S}\,\underline{\Phi}(T, t_0) + \int_{t_0}^{T} \underline{\Phi}'(t, t_0)\,[\underline{Q}(t) + \underline{C}'(t)\underline{F}'(t)\underline{R}(t)\underline{F}(t)\underline{C}(t)]\,\underline{\Phi}(t, t_0)dt\}\,\underline{x}(t_0) \qquad (2.1.9)$$

At this point, Eqs. (2.1.6) and (2.1.9) form an optimization problem which, given an $\underline{x}(t_0)$, can be solved for an optimal $\underline{F}(t)$. Unfortunately, this optimal $\underline{F}(t)$ will in general depend on $\underline{x}(t_0)$. Thus, it would not really be a feedback control. In order to find an "optimal" value for $\underline{F}(t)$ that is independent of the initial state it is necessary to change the problem somewhat. The change that is made is to attempt to determine that $\underline{F}(t)$ which is optimal in an "average" sense (a similar idea was used in references 13 and 14). If we view the initial state $\underline{x}(t_0)$ as a random variable uniformly distributed over the surface of an n-dimensional unit sphere, then the expected value $\hat{J}$ of the cost (2.1.9) is simply:

$$\hat{J} = n\cdot[\mathcal{E}(J\,|\,\underline{x}(t_0) \text{ uniformly distributed on the surface of the unit sphere})] \qquad (2.1.10)$$

$$\hat{J} = \frac{1}{2}\,\mathrm{tr}[\underline{\Phi}'(T,t_0)\underline{S}\,\underline{\Phi}(T,t_0)]$$

$$+\frac{1}{2}\int_{t_0}^{T} \mathrm{tr}[\underline{\Phi}'(t,t_0)(\underline{Q}(t) + \underline{C}'(t)\underline{F}'(t)\underline{R}(t)\underline{F}(t)\underline{C}(t))\underline{\Phi}(t,t_0)]\,dt \tag{2.1.11}$$

The derivation of Eq. (2. 1. 11) from Eq. (2. 1. 10) can be found in reference 15.

This "average" cost $\hat{J}$ is now independent of the specific initial state $\underline{x}(t_0)$; it is still, of course, dependent on $\underline{F}(t)$. Thus, it is reasonable to seek a gain matrix $\underline{F}(t)$ which minimizes the average cost of Eq. (2. 1. 11) subject to the differential constraint of Eq. (2. 1. 8). It should be noted that the transition matrix $\underline{\Phi}(t,t_0)$ plays the role of the "state" and the matrix $\underline{F}(t)$ plays the role of the "control". Such problems can be readily attacked by the matrix minimum principle.[12]

2. 2 Statement of the Problem

Thus, we have formulated the following mathematical optimization problem:

Given the system described by the matrix differential equation

$$\dot{\underline{\Phi}}(t,t_0) = [\underline{A}(t) - \underline{B}(t)\underline{F}(t)\underline{C}(t)]\,\underline{\Phi}(t,t_0)\ ;\ \underline{\Phi}(t_0,t_0) = \underline{I} \tag{2.2.1}$$

and the performance functional

$$\hat{J} = \frac{1}{2}\,\mathrm{tr}[\underline{\Phi}'(T,t_0)\underline{S}\,\underline{\Phi}(T,t_0)]$$

$$+\frac{1}{2}\int_{t_0}^{T} \mathrm{tr}\{\underline{\Phi}'(t,t_0)[\underline{Q}(t) + \underline{C}'(t)\underline{F}'(t)\underline{R}(t)\underline{F}(t)\underline{C}(t)]\,\underline{\Phi}(t,t_0)\}\,dt \tag{2.2.2}$$

Find the matrix $\underline{F}^*(t)$ that minimizes $\hat{J}$ subject to the differential constraints imposed by the system (2.2.1), where:

$\underline{A}(t)$ is an nxn real matrix

$\underline{B}(t)$ is an nxm real matrix

$\underline{C}(t)$ is an rxn real matrix of full rank (rank r)

$\underline{\Phi}(t, t_0)$ is an nxn matrix

$\underline{S}$ and $\underline{Q}(t)$ are nxn symmetric positive semi-definite real matrices

$\underline{R}(t)$ is an mxm symmetric positive definite real matrix

$\underline{A}(t)$, $\underline{B}(t)$, $\underline{C}(t)$, $\underline{Q}(t)$ and $\underline{R}(t)$ are bounded and measurable

$\underline{F}(t)$ is the control for the given system and is composed of measurable, but otherwise unconstrained elements; it is an mxr real matrix

We remark that the smoothness conditions on $\underline{A}$, $\underline{B}$, $\underline{C}$, $\underline{Q}$ and $\underline{R}$ could be relaxed slightly.

2.3 The Main Result

The results of this chapter are summarized below. These results specify the properties of the optimal gain matrix $\underline{F}^*(t)$. We assume that $\underline{F}^*(t)$ exists.

The optimal gain matrix $\underline{F}^*(t)$, i.e., the one that minimizes the (average) cost subject to the constraints is given by

$$\underline{F}^*(t) = \underline{R}^{-1}(t)\underline{B}'(t)\underline{K}^*(t)\underline{\Phi}^*(t, t_0)\underline{\Phi}^{*'}(t, t_0)\underline{C}'(t)\underline{\psi}^{-1}(t) \qquad (2.3.1)$$

where:

(a) $\underline{\psi}(t) \triangleq \underline{C}(t)\underline{\Phi}^*(t, t_0)\underline{\Phi}^{*'}(t, t_0)\underline{C}'(t) > 0; \; \underline{\psi}(t) = \underline{\psi}'(t) \qquad (2.3.2)$

(b) $\underline{\Phi}^*(t, t_0)$ is the solution of:

$$\dot{\underline{\Phi}}^*(t, t_0) = [\underline{A}(t) - \underline{B}(t)\underline{F}^*(t)\underline{C}(t)]\,\underline{\Phi}^*(t, t_0)\,; \; \underline{\Phi}^*(t_0, t_0) = \underline{I} \qquad (2.3.3)$$

(c) $\underline{K}^*(t)$ is the solution of:

$$\dot{\underline{K}}^*(t) = -\underline{Q}(t) - \underline{C}'(t)\underline{F}^{*\prime}(t)\underline{R}(t)\underline{F}^*(t)\underline{C}(t) - [\underline{A}(t) - \underline{B}(t)\underline{F}^*(t)\underline{C}(t)]'\underline{K}^*(t)$$
$$-\underline{K}^*(t)[\underline{A}(t) - \underline{B}(t)\underline{F}^*(t)\underline{C}(t)] \quad (2.3.4)$$

with the boundary condition at the terminal time T:

$$\underline{K}^*(T) = \underline{S} \quad (2.3.5)$$

Remarks

The proof of these results proceeds as follows:

i) We shall show that $\underline{F}^*(t)$ satisfies the necessary conditions for optimality using the matrix minimum principle [12], in Theorem 2.1.

ii) We shall demonstrate that the Hamilton-Jacobi sufficiency conditions hold (provided that the solution exists) in Theorem 2.2.

2.4 Proof of the Main Result

Theorem 2.1 The matrices $\underline{F}^*(t)$, $\underline{K}^*(t)$, $\underline{\Phi}^*(t,t_0)$, $t \in [t_0, T]$, defined by Eqs. (2.3.1) to (2.3.5) satisfy the necessary conditions for optimality provided by the matrix minimum principle.

Proof: Let $\underline{P}(t,t_0)$ be an nxn "costate" matrix associated with $\underline{\Phi}(t,t_0)$. Then the (scalar) Hamiltonian function H for the optimization problem is given by

$$H = \frac{1}{2}\,\mathrm{tr}\{\underline{\Phi}'(t,t_0)[\underline{Q}(t) + \underline{C}'(t)\underline{F}'(t)\underline{R}(t)\underline{F}(t)\underline{C}(t)]\,\underline{\Phi}(t,t_0)\}$$
$$+\,\mathrm{tr}\,\{[\underline{A}(t) - \underline{B}(t)\underline{F}(t)\underline{C}(t)]\,\underline{\Phi}(t,t_0)\underline{P}'(t,t_0)\} \quad (2.4.1)$$

We note that the Hamiltonian is a quadratic function of $\underline{F}(t)$. Hence, a necessary condition that $\underline{F}^*(t)$ minimizes H is that the following gradient matrix vanishes:

1) $$\left.\frac{\partial \underline{H}}{\partial \underline{F}}\right|_* = \underline{R}(t)\underline{F}^*(t)\underline{C}(t)\underline{\Phi}^*(t,t_0)\underline{\Phi}^{*'}(t,t_0)\underline{C}'(t) - \underline{B}'(t)\underline{P}^*(t,t_0)\underline{\Phi}^{*'}(t,t_0)\underline{C}'(t) = 0 \tag{2.4.2}$$

Hence, $\underline{F}^*(t)$ is given by

$$\underline{F}^*(t) = \underline{R}^{-1}(t)\underline{B}'(t)\underline{P}^*(t,t_0)\underline{\Phi}^{*'}(t,t_0)\underline{C}'(t)\underline{\psi}^{-1}(t) \tag{2.4.3}$$

where:

$$\underline{\psi}(t) \triangleq \underline{C}(t)\underline{\Phi}^*(t,t_0)\underline{\Phi}^{*'}(t,t_0)\underline{C}'(t) \tag{2.4.4}$$

Note that $\underline{\psi}(t)$ is symmetric and at least positive semidefinite. Since $\underline{\Phi}(t,t_0)$ is a transition matrix and since $\underline{C}(t)$ is of rank r, then $\underline{\psi}^{-1}(t)$ exists and, so, $\underline{\psi}(t)$ is positive definite.

In order to prove that the $\underline{F}^*(t)$ given by Eq. (2.4.3) does indeed minimize the Hamiltonian we proceed as follows:

$$\text{Let } \underline{F}(t) \triangleq \underline{F}^*(t) + \Delta\underline{F}(t) \tag{2.4.5}$$

(with the arguments, t and t_0, suppressed for compactness)

$$\begin{aligned} H(\underline{F}) = {} & \tfrac{1}{2}\,\mathrm{tr}(\underline{\Phi}^{*'}\underline{Q}\,\underline{\Phi}^*) + \tfrac{1}{2}\,\mathrm{tr}[\underline{\Phi}^{*'}\underline{C}'(\underline{F}^*+\Delta\underline{F})'\underline{R}(\underline{F}^*+\Delta\underline{F})\underline{C}\,\underline{\Phi}^*] \\ & + \mathrm{tr}\,\{[\underline{A} - \underline{B}(\underline{F}^*+\Delta\underline{F})\underline{C}]\,\underline{\Phi}^*\underline{P}^{*'}\} \end{aligned} \tag{2.4.6}$$

$$\begin{aligned} \underline{H}(\underline{F}) = {} & \mathrm{tr}[\tfrac{1}{2}\,\underline{\Phi}^{*'}\underline{Q}\,\underline{\Phi}^* + \underline{A}\,\underline{\Phi}^*\underline{P}^{*'}] + \mathrm{tr}[-\underline{B}(\underline{F}^*+\Delta\underline{F})\underline{C}\,\underline{\Phi}^*\underline{P}^{*'}] \\ & + \mathrm{tr}[\tfrac{1}{2}\,\underline{C}\,\underline{\Phi}^*\underline{\Phi}^{*'}\underline{C}'\underline{F}^{*'}\underline{R}(\underline{F}^*+\Delta\underline{F})] + \mathrm{tr}[\tfrac{1}{2}\,\underline{C}\,\underline{\Phi}^*\underline{\Phi}^{*'}\underline{C}'\Delta\underline{F}'\underline{R}(\underline{F}^*+\Delta\underline{F})] \end{aligned} \tag{2.4.7}$$

$$\begin{aligned} H(\underline{F}) = {} & c_1 + \mathrm{tr}[-\underline{C}\,\underline{\Phi}^*\underline{P}^{*'}\underline{B}(\underline{F}^*+\Delta\underline{F})] + \mathrm{tr}[\tfrac{1}{2}\,\underline{C}\,\underline{\Phi}^*\underline{P}^{*'}\underline{B}(\underline{F}^*+\Delta\underline{F})] \\ & + \mathrm{tr}[\tfrac{1}{2}\,\underline{\psi}\Delta\underline{F}'\underline{R}(\underline{F}^*+\Delta\underline{F})] \end{aligned} \tag{2.4.8}$$

where c_1 is independent of $\underline{F}$

$$H(\underline{F}) = c_1 + tr[-\frac{1}{2}\underline{C}\,\underline{\Phi}^*\underline{P}^{*'}\underline{B}\,\underline{F}^*] + tr[-\frac{1}{2}\underline{C}\,\underline{\Phi}^*\underline{P}^{*'}\underline{B}\Delta\underline{F}]$$

$$+ tr[\frac{1}{2}\underline{\psi}\Delta\underline{F}'\underline{R}\,\underline{F}^*] + tr[\frac{1}{2}\underline{\psi}\Delta\underline{F}'\underline{R}\Delta\underline{F}] \qquad (2.4.9)$$

$$H(\underline{F}) = c_1 + tr[-\frac{1}{2}\underline{C}\,\underline{\Phi}^*\underline{P}^{*'}\underline{B}\,\underline{F}^*] + tr[\frac{1}{2}\underline{\psi}\Delta\underline{F}'\underline{R}\Delta\underline{F}]$$

$$+ tr[\frac{1}{2}\Delta\underline{F}'\underline{R}\,\underline{F}^*\underline{\psi}] + tr[-\frac{1}{2}\Delta\underline{F}'\underline{B}'\underline{P}^*\underline{\Phi}^{*'}\underline{C}'] \qquad (2.4.10)$$

$$H(\underline{F}) = c_2 + tr[\frac{1}{2}\underline{R}^{1/2}\Delta\underline{F}\underline{\psi}\Delta\underline{F}'\underline{R}^{1/2'}] \qquad (2.4.11)$$

where c_2 is a new constant independent of $\underline{F}$

But, $tr[\underline{R}^{1/2}\Delta\underline{F}\underline{\psi}\Delta\underline{F}'\underline{R}^{1/2'}] > 0$ for all $\Delta\underline{F} \neq 0$

because:

a) $\underline{R}^{1/2}\Delta\underline{F}\underline{\psi}\Delta\underline{F}'\underline{R}^{1/2'} = (\underline{R}^{1/2}\Delta\underline{F}\underline{\psi}^{1/2})(\underline{R}^{1/2}\Delta\underline{F}\underline{\psi}^{1/2})' \geq \underline{0}$

since $\underline{\psi} > 0$.

b) $tr[\underline{R}^{1/2}\Delta\underline{F}\underline{\psi}\Delta\underline{F}'\underline{R}^{1/2'}] = 0$ if and only if $\Delta\underline{F}'\underline{R}^{1/2'}\underline{v} = 0$

for all vectors $\underline{v}$. This is only possible if $\Delta\underline{F} = 0$.

Hence,

$$H(\underline{F}) > H(\underline{F}^*) \quad \text{for all} \quad \underline{F} \neq \underline{F}^* \qquad (2.4.12)$$

Thus, the $\underline{F}^*$ defined by (2.4.3) does indeed minimize the Hamiltonian.

2) Using the necessary conditions of the matrix minimum principle one deduces that the "costate" matrix $\underline{P}^*(t, t_0)$ satisfies the following matrix differential equation

$$\dot{\underline{P}}^*(t, t_0) = - \frac{\partial H}{\partial \underline{\Phi}(t, t_0)}\bigg|_* = - [\underline{Q}(t) + \underline{C}'(t)\underline{F}^{*\prime}(t)\underline{R}(t)\underline{F}^*(t)\underline{C}(t)] \underline{\Phi}^*(t, t_0)$$

$$- [\underline{A}(t) - \underline{B}(t)\underline{F}^*(t)\underline{C}(t)]'\underline{P}^*(t, t_0) \qquad (2.4.13)$$

and, of course, $\underline{\Phi}^*(t, t_0)$ satisfies the equation

$$\dot{\underline{\Phi}}^*(t, t_0) = [\underline{A}(t) - \underline{B}(t)\underline{F}^*(t)\underline{C}(t)] \underline{\Phi}^*(t, t_0) \; ; \; \underline{\Phi}^*(t_0, t_0) = \underline{I} \qquad (2.4.14)$$

Furthermore, at the terminal T, it is necessary that

$$\underline{P}^*(T, t_0) = \frac{\partial}{\partial \underline{\Phi}(T, t_0)} \operatorname{tr}[\underline{\Phi}'(T, t_0)\underline{S}\,\underline{\Phi}(T, t_0)]\bigg|_* = \underline{S}\,\underline{\Phi}^*(T, t_0) \qquad (2.4.15)$$

We claim that the solutions of Eqs. (2.4.13) and (2.4.14) are related by

$$\underline{P}^*(t, t_0) = \underline{K}^*(t)\underline{\Phi}^*(t, t_0) \qquad (2.4.16)$$

where $\underline{K}^*(t)$ is an nxn matrix to be determined. From Eq. (2.4.16)

$$\dot{\underline{P}}^*(t, t_0) = \dot{\underline{K}}^*(t)\underline{\Phi}^*(t, t_0) + \underline{K}^*(t)\dot{\underline{\Phi}}^*(t, t_0) \qquad (2.4.17)$$

Substituting Eqs. (2.4.13), (2.4.14), and (2.4.16) into Eq. (2.4.17) we obtain

$$-[\underline{Q}(t)+\underline{C}'(t)\underline{F}^{*\prime}(t)\underline{R}(t)\underline{F}^*(t)\underline{C}(t)]\underline{\Phi}^*(t, t_0)-[\underline{A}(t)-\underline{B}(t)\underline{F}^*(t)\underline{C}(t)]'\underline{K}^*(t)\underline{\Phi}^*(t, t_0)$$

$$= \dot{\underline{K}}^*(t)\underline{\Phi}^*(t, t_0)+\underline{K}^*(t)[\underline{A}(t)-\underline{B}(t)\underline{F}^*(t)\underline{C}(t)]\underline{\Phi}^*(t, t_0) \qquad (2.4.18)$$

which yields, since $\underline{\Phi}^*(t, t_0)$ is always non-singular

$$\dot{\underline{K}}^*(t) = -\underline{Q}(t)-\underline{C}'(t)\underline{F}^{*\prime}(t)\underline{R}(t)\underline{F}^*(t)\underline{C}(t)-\underline{K}^*(t)[\underline{A}(t)-\underline{B}(t)\underline{F}^*(t)\underline{C}(t)]$$

$$- [\underline{A}(t)-\underline{B}(t)\underline{F}^*(t)\underline{C}(t)]'\underline{K}^*(t) \qquad (2.4.19)$$

From Eqs. (2.4.15) and (2.4.16) we deduce that

$$\underline{K}^*(T) = \underline{S} \qquad (2.4.20)$$

Finally, from Eqs. (2.4.16) and (2.4.3) we have

$$\underline{F}^*(t) = \underline{R}^{-1}(t)\underline{B}'(t)\underline{K}^*(t)\underline{\Phi}^*(t,t_0)\underline{\Phi}^{*'}(t,t_0)\underline{C}'(t)\underline{\psi}^{-1}(t) \qquad (2.4.21)$$

This completes the proof that the matrices $\underline{F}^*(t)$, $\underline{\Phi}^*(t,t_0)$ and $\underline{K}^*(t)$, as stated in Section 2.3, satisfy the necessary conditions for optimality.

We remark that both $\underline{K}^*(t)$ and $\underline{\Phi}^*(t,t_0)$ satisfy matrix differential equations. It can be shown that Eqs. (2.4.14) and (2.4.19) satisfy (local) Lipschitz conditions; this implies that the solutions are (locally) unique.

We shall next prove that $\underline{\Phi}^*(t,t_0)$ and $\underline{K}^*(t)$ satisfy the Hamilton-Jacobi equation.

Theorem 2.2 $\underline{F}^*(t)$, $\underline{K}^*(t)$ and $\underline{\Phi}^*(t,t_0)$, defined by Eqs. (2.3.1)-(2.3.5) satisfy the sufficiency conditions that result from the application of the Hamilton-Jacobi theorem.

Proof: Define

$$\underline{V}^*(t) \triangleq \int_t^T \underline{\Phi}^{*'-1}(t,t_0)\underline{\Phi}^{*'}(\tau,t_0)[\underline{Q}(\tau) + \underline{C}'(\tau)\underline{F}^{*'}(\tau)\underline{R}(\tau)\underline{F}^*(\tau)\underline{C}(\tau)]$$
$$\underline{\Phi}^*(\tau,t_0)\underline{\Phi}^{*-1}(t,t_0)d\tau + \underline{\Phi}^{*'-1}(t,t_0)\underline{\Phi}^{*'}(T,t_0)\underline{S}\,\underline{\Phi}^*(T,t_0)\underline{\Phi}^{*-1}(t,t_0)$$
$$(2.4.22)$$

Differentiating with respect to time, we obtain:

$$\dot{\underline{V}}^*(t) = -\underline{Q}(t) - \underline{C}'(t)\underline{F}^{*'}(t)\underline{R}(t)\underline{F}^*(t)\underline{C}(t) - \underline{V}^*(t)[\underline{A}(t) - \underline{B}(t)\underline{F}^*(t)\underline{C}(t)]$$
$$- [\underline{A}(t) - \underline{B}(t)\underline{F}^*(t)\underline{C}(t)]'\underline{V}^*(t) \qquad (2.4.23)$$

with terminal condition $\underline{V}^*(T) = \underline{S}$.

Notice that the above equation is linear and thus a unique solution for $\underline{V}^*(t)$ exists. Notice also that $\underline{V}^*(t) = \underline{K}^*(t)$, provided $\underline{F}^*(t)$ exists, because they both satisfy the identical differential equation and boundary condition. We can evaluate the cost functional (2.2.2) to obtain

$$\hat{J}(\underline{F}^*, t) = \frac{1}{2}\, \mathrm{tr}[\underline{\Phi}^{*'}(t, t_0)\underline{V}^*(t)\underline{\Phi}^*(t, t_0)] \tag{2.4.24}$$

We can now compute the derivatives of $\hat{J}$:

$$\frac{\partial \hat{J}(\underline{F}^*, t, \underline{\Phi}^*)}{\partial t} = \frac{1}{2}\, \mathrm{tr}[\underline{\Phi}^{*'}(t, t_0)\dot{\underline{V}}^*(t)\underline{\Phi}^*(t, t_0)] \tag{2.4.25}$$

$$\frac{\partial \hat{J}(\underline{F}^*, t, \underline{\Phi}^*)}{\partial \underline{\Phi}^*(t, t_0)} = \underline{V}^*(t)\underline{\Phi}^*(t, t_0) = \underline{K}^*(t)\underline{\Phi}^*(t, t_0) = \underline{P}^*(t, t_0) \tag{2.4.26}$$

Let:

$$H(\underline{\Phi}^*(t, t_0), \underline{F}^*(t), t, \frac{\partial \hat{J}(\underline{F}^*, t, \underline{\Phi}^*)}{\partial \underline{\Phi}^*(t, t_0)}) \triangleq H^* \tag{2.4.27}$$

But:

$$H^* = \frac{1}{2}\, \mathrm{tr}[\underline{\Phi}^{*'}(t, t_0)\{\underline{Q}(t)+\underline{C}'(t)\underline{F}^{*'}(t)\underline{R}(t)\underline{F}^*(t)\underline{C}(t)\}\,\underline{\Phi}^*(t, t_0)]$$
$$+ \mathrm{tr}[\{\underline{A}(t)-\underline{B}(t)\underline{F}^*(t)\underline{C}(t)\}\,\underline{\Phi}^*(t, t_0)\underline{\Phi}^{*'}(t, t_0)\underline{V}^*(t)] \tag{2.4.28}$$

or:

$$H^* = \frac{1}{2}\, \mathrm{tr}\{\underline{\Phi}^{*'}(t, t_0)[\underline{Q}(t)+\underline{C}'(t)\underline{F}^{*'}(t)\underline{R}(t)\underline{F}^*(t)\underline{C}(t)]\underline{\Phi}^*(t, t_0)\}$$
$$+ \frac{1}{2}\, \mathrm{tr}\{\underline{\Phi}^{*'}(t, t_0)[\underline{A}(t)-\underline{B}(t)\underline{F}^*(t)\underline{C}(t)]'\underline{\Phi}^*(t, t_0)\} \tag{2.4.29}$$

or:

$$H^* = \frac{1}{2}\, \mathrm{tr}\{\underline{\Phi}^{*'}(t, t_0)[\underline{Q}+\underline{C}'\underline{F}^{*'}\underline{R}\,\underline{F}^*\underline{C}+\underline{V}^*(\underline{A}-\underline{B}\,\underline{F}^*\underline{C})$$
$$+ (\underline{A}-\underline{B}\,\underline{F}^*\underline{C})'\underline{V}^*]\underline{\Phi}^*(t, t_0)\} \tag{2.4.30}$$

Hence:

$$H^{*} = \frac{1}{2}\,\mathrm{tr}\{\underline{\Phi}^{*\prime}(t,t_0)[-\underline{\dot{V}}^{*}(t)]\,\underline{\Phi}^{*}(t,t_0)\} \tag{2.4.31}$$

This implies that

$$\frac{\partial \hat{J}(\underline{F}^{*},t,\underline{\Phi}^{*})}{\partial t} + H^{*} = 0 \tag{2.4.32}$$

In other words, the Hamilton-Jacobi equation is satisfied, as is the terminal condition.

Thus, we have verified all the assumptions of the Hamilton-Jacobi theorem [as given by Theorem 5.13, p. 360, of reference (4)].

To summarize the above results, we have formulated an optimal control problem that corresponds to optimizing the linear output feedback for a linear system. In addition, we have found a set of necessary conditions which must be satisfied by the optimal feedback control if it exists. Finally, we show that any solution of the necessary conditions is also a solution of the Hamilton-Jacobi equation. Thus, if we can find a control which satisfies the necessary conditions we know that it is at least a locally optimal control by the Hamilton-Jacobi theorem.

2.5 Existence and Uniqueness

In this section we will discuss the existence and uniqueness of solutions for the optimization problem defined in Section 2.2. We begin by stating and proving a theorem which implies, as we shall demonstrate, the existence of solutions.

Theorem 2.3 If we add the constraint that $|f_{ij}(t)| \leq M$ for all $i = 1,2,\ldots m$ and $j = 1,2,\ldots r$ to the assumptions given in the definition of the optimization problem in Section 2.2, then an optimal gain matrix $\underline{F}^{*}(t)$ exists.

Outline of Proof:

1) We prove that the set of reachable states at the terminal time T is closed and bounded (compact).

2) We prove that the performance criterion is defined and continuous on the set of reachable states.

3) Therefore, the minimum of the performance criterion is achieved.

Proof: Define $R(\underline{\Phi})$ to be the set of states $\underline{\Phi}(T, t_0)$ that can be reached by applying an admissible control $\underline{F}(t)$, $t \in [t_0, T]$, to the system described by Eq. (2.2.1) starting at $\underline{\Phi}(t_0, t_0) = \underline{I}$.

We claim that the set $R(\underline{\Phi})$ is closed. This can be deduced from any one of several published theorems on existence of solutions to optimal control problems.[20, 21] The crucial requirements in these theorems are:

a) existence and uniqueness of the solution to the differential equation (2.2.1) given a "control" $\underline{F}(t)$.

b) continuity of the right-hand side of the differential equation in $\underline{F}$ and $\underline{\Phi}$.

c) convexity of the set $\{(\underline{A}(t)-\underline{B}(t)\underline{F}(t)\underline{C}(t))\underline{\Phi}(t, t_0) \mid \underline{F}(t) \text{ allowable}\}$ for each $t, \underline{\Phi}(t, t_0)$.

All of these conditions are satisfied, as the reader can verify.

To show that $R(\underline{\Phi})$ is bounded we note that:

$$\frac{d}{dt} \| \underline{\Phi}(t, t_0) \| \leq \| \underline{A}(t) - \underline{B}(t)\underline{F}(t)\underline{C}(t) \| \cdot \| \underline{\Phi}(t, t_0) \| \tag{2.5.1}$$

thus;

$$\| \underline{\Phi}(T, t_0) \| \leq e^{\int_{t_0}^{T} \| \underline{A}(t) - \underline{B}(t)\underline{F}(t)\underline{C}(t) \| dt} \leq M_1 = \text{a constant} \tag{2.5.2}$$

Therefore, $R(\underline{\Phi})$ is compact.

To show that $\hat{J}$ is defined and continuous on $R(\underline{\Phi})$ it is sufficient to show that the integrand of $\hat{J}$ (Eq. 2.2.2) is Lipschitz in $\underline{\Phi}(t, t_0)$ ($\underline{\Phi}(t, t_0)$ bounded) for any $(t, \underline{F}(t))$. If we define the convenience,

$$\underline{M}(t) \triangleq \underline{Q}(t) + \underline{C}'(t)\underline{F}'(t)\underline{R}(t)\underline{F}(t)\underline{C}(t) \tag{2.5.3}$$

Then, by taking the norm of the difference of the integrands of $\hat{J}$ for two values of $\underline{\Phi}(t, t_0)$, we have:

$$|| \operatorname{tr}[\underline{\Phi}_1'(t, t_0)\underline{M}(t)\underline{\Phi}_1(t, t_0) - \underline{\Phi}_2'(t, t_0)\underline{M}(t)\underline{\Phi}_2(t, t_0)] || = a \tag{2.5.4}$$

$$a = || \operatorname{tr}\{ \underline{M}(t)[\underline{\Phi}_1(t, t_0) + \underline{\Phi}_2(t, t_0)] [\underline{\Phi}_1(t, t_0) - \underline{\Phi}_2(t, t_0)] \} || \tag{2.5.5}$$

$$a \leq n \cdot || \underline{M}(t)[\underline{\Phi}_1(t, t_0) + \underline{\Phi}_2(t, t_0)] || \circ || \underline{\Phi}_1(t, t_0) - \underline{\Phi}_2(t, t_0) || \tag{2.5.6}$$

$$a \leq \underline{K} \cdot || \underline{\Phi}_1(t, t_0) - \underline{\Phi}_2(t, t_0) || \tag{2.5.7}$$

Therefore, $\hat{J}$ is defined and continuous in $R(\underline{\Phi})$.

This completes the proof of Theorem 2.3 since a function that is defined and continuous on a compact set achieves its minimum on that set.

We have already shown, in the previous section, that if an optimal control exists it is characterized by Eq. (2.4.3) and that this $\underline{F}^*(t)$ is the unique H-minimal control. By taking the upper bound on $|f_{ij}(t)|$ large enough we can insure that $\underline{F}^*(t)$ is not on the boundary of the admissible $\underline{F}$'s. Thus, we have proved the existence of solutions to the problem defined in Section 2.2.

Finally, although we cannot offer a proof, we believe that the solutions to the optimization problem defined in Section 2.2 are not unique. The best intuitive evidence of this is contained in the proof of Lemma 3.1. If the solution to our optimization problem was unique, the method of proof used in the lemma would probably prove uniqueness.

CHAPTER III

COMPUTATION OF THE OPTIMAL FEEDBACK GAIN MATRIX

In order to implement the closed-loop system one must first be able to compute the feedback gain matrix $\underline{F}^*(t)$ for all $t \epsilon [t_0, T]$. It can be seen from the equations (given in the previous chapter) that specify $\underline{F}^*(t)$ that the computation of $\underline{F}^*(t)$ involves the solution of a non-linear two-point boundary value problem involving matrix differential equations. Very little prior work has been done in this area.

In this chapter, we begin by outlining an algorithm for computing $\underline{F}^*(t)$. Although we cannot prove that this algorithm converges to the optimal $\underline{F}^*(t)$ we can, and do, prove that the value of the cost functional decreases with each iteration. We next discuss the programming of this algorithm and, in fact, include a Fortran version of the program in Appendix B. Finally, we conclude the chapter with some examples that were calculated via the program.

3.1 Theoretical Algorithm

In this section, we outline a computational procedure which generates a sequence of matrices $\{\underline{K}_n(t)\}$, $\{\underline{F}_n(t)\}$, and $\{\underline{\Phi}_n(t, t_0)\}$ which hopefully converge to the optimal ones. We then prove that this algorithm has the property that the cost decreases at each iteration.

The algorithm for computing $\underline{K}_{n+1}(t)$, $\underline{F}_{n+1}(t)$ and $\underline{\Phi}_{n+1}(t, t_0)$ begins with a stored value for $\underline{F}_n(t)$, $t \epsilon [t_0, T]$. Knowing $\underline{F}_n(t)$, we compute $\underline{K}_{n+1}(t)$ by integrating the equation

$$\dot{\underline{K}}_{n+1}(t) = -\underline{K}_{n+1}(t)[\underline{A}(t)-\underline{B}(t)\underline{F}_n(t)\underline{C}(t)] - [\underline{A}(t)-\underline{B}(t)\underline{F}_n(t)\underline{C}(t)]'\underline{K}_{n+1}(t)$$
$$- \underline{Q}(t) - \underline{C}'(t)\underline{F}_n'(t)\underline{R}(t)\underline{F}_n(t)\underline{C}(t) \qquad (3.1.1)$$

backwards in time from the terminal condition $\underline{K}_{n+1}(T) = \underline{S}$. These values of $\underline{K}_{n+1}(t)$, $t \in [t_0, T]$ are stored. Then, they can be substituted in the equation for $\underline{F}_{n+1}(t)$ below:

$$\underline{F}_{n+1}(t) = \underline{R}^{-1}(t)\underline{B}'(t)\underline{K}_{n+1}(t)\underline{\Phi}_{n+1}(t,t_0)\underline{\Phi}'_{n+1}(t,t_0)\underline{C}'(t)\underline{\psi}^{-1}(t) \qquad (3.1.2)$$

where

$$\underline{\psi}(t) = \underline{C}(t)\underline{\Phi}_{n+1}(t,t_0)\underline{\Phi}'_{n+1}(t,t_0)\underline{C}'(t) \qquad (3.1.3)$$

Of course, since $\underline{\Phi}_{n+1}(t,t_0)$ is still unknown, we cannot actually compute $\underline{F}_{n+1}(t)$ yet. However, when Eqs. (3.1.2) and (3.1.3) are substituted into

$$\dot{\underline{\Phi}}_{n+1}(t,t_0) = [\underline{A}(t)-\underline{B}(t)\underline{F}_{n+1}(t)\underline{C}(t)]\underline{\Phi}_{n+1}(t,t_0) \;;\; \underline{\Phi}_{n+1}(t_0,t_0)=\underline{I} \quad (3.1.4)$$

it will be noted that Eq. (3.1.4) has only one unknown, $\underline{\Phi}_{n+1}(t,t_0)$. Thus, we have a non-linear ordinary matrix differential equation with a known initial condition which is integrated forwards in time for $\underline{\Phi}_{n+1}(t,t_0)$, $t \in [t_0, T]$. And, as each value of $\underline{\Phi}_{n+1}(t,t_0)$ is computed it is substituted into Eq. (3.1.2) thereby generating $\underline{F}_{n+1}(t)$, $t \in [t_0, T]$. These values are stored and used to begin the next iteration.

The iterations are begun with an initial guess for $\underline{F}_0(t)$. This initial guess does not determine whether the algorithm converges although it will affect the rate of convergence. We have obtained this initial guess by setting

$$\underline{K}_0(t) = \underline{K}(T) = \underline{S} \quad \text{and} \quad \underline{\Phi}_0(t,t_0) = \underline{\Phi}(t_0,t_0) = \underline{I} \qquad (3.1.5)$$

and substituting these values into Eq. (3.1.2). This results in an $\underline{F}_0(t)$ that matches the boundary conditions.

In the previous chapter, we proved that

$$\hat{J}(\underline{F}^*, t) = \frac{1}{2}\operatorname{tr}[\underline{\Phi}^{*\prime}(t, t_0)\underline{K}^*(t)\underline{\Phi}^*(t, t_0)] \qquad (3.1.6)$$

Thus,

$$\hat{J}(\underline{F}^*, t_0) = \frac{1}{2}\operatorname{tr}[\underline{K}^*(t_0)] \qquad (3.1.7)$$

A simple substitution shows that the cost, or performance, obtained by using the control $\underline{F}_n(t)$, $t \in [t_0, T]$, is given by

$$\hat{J}(\underline{F}_n(t), t_0) = \frac{1}{2}\operatorname{tr}[\underline{K}_{n+1}(t_0)] \qquad (3.1.8)$$

Thus, the lemma proven below guarantees that the value of the performance criterion decreases at each iteration.

<u>Lemma 3.1</u>

Using the algorithm described above,

$$\operatorname{tr}[\underline{K}_{n+1}(t_0)] < \operatorname{tr}[\underline{K}_n(t_0)] \quad \text{for all } n \qquad (3.1.9)$$

<u>Proof</u>:

$$\frac{d}{dt}[\underline{K}_n(t)-\underline{K}_{n+1}(t)] = -[\underline{A}-\underline{B}\,\underline{F}_{n-1}\underline{C}]'\underline{K}_n-\underline{K}_n[\underline{A}-\underline{B}\,\underline{F}_{n-1}\underline{C}]-\underline{C}'\underline{F}_{n-1}'\underline{R}\,\underline{F}_{n-1}\underline{C}$$
$$+[\underline{A}-\underline{B}\,\underline{F}_n\underline{C}]'\underline{K}_{n+1}+\underline{K}_{n+1}[\underline{A}-\underline{B}\,\underline{F}_n\underline{C}]+\underline{C}'\underline{F}_n'\underline{R}\,\underline{F}_n\underline{C} \qquad (3.1.10)$$

$$\frac{d}{dt}[\underline{K}_n(t)-\underline{K}_{n+1}(t)] = -[\underline{A}-\underline{B}\,\underline{F}_n\underline{C}]'\underline{K}_n+[\underline{A}-\underline{B}\,\underline{F}_n\underline{C}]'\underline{K}_{n+1}-\underline{K}_n[\underline{A}-\underline{B}\,\underline{F}_n\underline{C}]$$
$$+\underline{K}_{n+1}[\underline{A}-\underline{B}\,\underline{F}_n\underline{C}]+(\underline{B}\,\underline{F}_{n-1}\underline{C})'\underline{K}_n-(\underline{B}\,\underline{F}_n\underline{C})'\underline{K}_n+\underline{K}_n(\underline{B}\,\underline{F}_{n-1}\underline{C})$$
$$-\underline{K}_n\underline{B}\,\underline{F}_n\underline{C}-\underline{C}'\underline{F}_{n-1}'\underline{R}\,\underline{F}_{n-1}\underline{C} + \underline{C}'\underline{F}_n'\underline{R}\,\underline{R}^{-1}\underline{R}\,\underline{F}_n\underline{C}+\underline{K}_n\underline{B}\,\underline{R}^{-1}\underline{B}\,\underline{K}_n$$
$$-\underline{K}_n\underline{B}\,\underline{R}^{-1}\underline{B}'\underline{K}_n \qquad (3.1.11)$$

$$\frac{d}{dt}[\underline{K}_n(t)-\underline{K}_{n+1}(t)] = -[\underline{A}-\underline{B}\,\underline{F}_n\underline{C}]'(\underline{K}_n-\underline{K}_{n+1})-(\underline{K}_n-\underline{K}_{n+1})[\underline{A}-\underline{B}\,\underline{F}_n\underline{C}]$$

$$-[\underline{C}'\underline{F}_{n-1}'\underline{R}-\underline{K}_n\underline{B}]\underline{R}^{-1}[\underline{R}\,\underline{F}_{n-1}\underline{C}-\underline{B}'\underline{K}_n]$$

$$+[\underline{C}'\underline{F}_n'\underline{R}-\underline{K}_n\underline{B}]\underline{R}^{-1}[\underline{R}\,\underline{F}_n\underline{C}-\underline{B}'\underline{K}_n] \qquad (3.1.12)$$

Integrating Eq. (3.1.12) and taking the trace, we obtain:

$$\mathrm{tr}[\underline{K}_{n+1}(t_0)]-\mathrm{tr}[\underline{K}_n(t_0)] = \int_{t_0}^{T} \mathrm{tr}\{\underline{\Phi}_n'(\tau,t_0)[\underline{C}'\underline{F}_{n-1}'\underline{R}-\underline{K}_n\underline{B})\underline{R}^{-1}(\underline{R}\,\underline{F}_{n-1}\underline{C}-\underline{B}'\underline{K}_n)$$

$$-(\underline{C}'\underline{F}_n'\underline{R}-\underline{K}_n\underline{B})\underline{R}^{-1}(\underline{R}\,\underline{F}_n\underline{C}-\underline{B}'\underline{K}_n)]\underline{\Phi}_n(\tau,t_0)\}d\tau \qquad (3.1.13)$$

We now wish to show that the integrand in Eq. (3.1.13) is positive for all τ. To do this, we first introduce three facts.

I) If $\underline{X}$ is a real symmetric $n\times n$ matrix, $\mathrm{tr}[\underline{X}] = \sum_{i=1}^{n} \underline{x}_i'\underline{X}\,\underline{x}_i$ where $\{\underline{x}_i\}$ is an arbitrary orthornormal basis for R^n. (3.1.14)

<u>Proof</u>: $\underline{x}_i = \underline{P}\underline{e}_i$ where $\underline{e}_i$ is an element of the natural basis and $\underline{P}$ is an orthogonal matrix ($\underline{P}\underline{P}' = \underline{I}$). Then $\mathrm{tr}[\underline{X}] = \mathrm{tr}[\underline{P}'\underline{X}\underline{P}] = \sum_{i=1}^{n} \underline{e}_i'\underline{P}'\underline{X}\,\underline{P}\underline{e}_i = \sum_{i=1}^{n} \underline{x}_i'\underline{X}\underline{x}_i$

II) $[\underline{C}(\tau)\underline{\Phi}_n(\tau,t_0)\underline{\Phi}_n'(\tau,t_0)\underline{C}'(\tau)]^{-1}\underline{C}(\tau)\underline{\Phi}_n(\tau,t_0) = [\underline{\Phi}_n'(\tau,t_0)\underline{C}'(\tau)]^{\dagger}$ where $\dagger$ denotes pseudo-inverse. (3.1.15)

<u>Proof</u>: See Appendix A, Corollary A.2

III) If $\underline{x}_o$ is an arbitrary real n-vector, $\underline{x}_0$ can be written as

$$\underline{x}_0 = \underline{x}_1 + \underline{x}_2 \text{ where } \begin{array}{l} \underline{x}_1 \in \mathcal{R}[\underline{\Phi}_n'(\tau,t_0)\underline{C}'(\tau)] \\ \underline{x}_2 \in \mathcal{R}[\underline{\Phi}_n'(\tau,t_0)\underline{C}'(\tau)]^{\perp} = \mathcal{N}[\underline{C}(\tau)\underline{\Phi}_n(\tau,t_0)] \end{array}$$

In order to evaluate the trace under the integral at time τ, we choose an orthonormal basis $\{\underline{x}_i\}$ such that

$$\begin{aligned} &\underline{x}_i \in \mathcal{R}\,[\underline{\Phi}_n'(\tau,t_0)\underline{C}'(\tau)] && \text{for } i \leq m \\ &\underline{x}_i \in \mathcal{R}\,[\underline{\Phi}_n'(\tau,t_0)\underline{C}'(\tau)]^{\perp} && \text{for } m < i \leq n \end{aligned} \tag{3.1.16}$$

Using facts I and III, we see that the integrand in Eq. (3.1.13)

$$\begin{aligned} \mathrm{tr}\{\underline{\Phi}_n'(\tau,t_0)[&(\underline{C}'\underline{F}_{n-1}'\underline{R}-\underline{K}_n\underline{B})\underline{R}^{-1}(\underline{R}\,\underline{F}_{n-1}\underline{C}-\underline{B}'\underline{K}_n) \\ &-(\underline{C}'\underline{F}_n'\underline{R}-\underline{K}_n\underline{B})\underline{R}^{-1}(\underline{R}\underline{F}_n\underline{C}-\underline{B}'\underline{K}_n)]\underline{\Phi}_n(\tau,t_0)\} \triangleq \mathrm{tr}\{\underline{I}_n\} \end{aligned} \tag{3.1.17}$$

$$\mathrm{tr}[\underline{I}_n] = \sum_{i=1}^{n} \underline{x}_i'\underline{I}_n\underline{x}_i \tag{3.1.18}$$

where $\underline{x}_i$ are elements of the special basis (3.1.16).

We see that for $\underline{x}_i \in \mathcal{N}\,[\underline{C}(\tau)\underline{\Phi}_n(\tau,t_0)]$ or equivalently, $\underline{x}_i$ such that $m < i \leq n$:

$$\underline{x}_i'\underline{I}_n\underline{x}_i = \underline{x}_i'\underline{\Phi}_n'(\tau,t_0)[\underline{K}_n\underline{B}\,\underline{R}^{-1}\underline{B}'\underline{K}_n-\underline{K}_n\underline{B}\,\underline{R}^{-1}\underline{B}'\underline{K}_n]\underline{\Phi}_n(\tau,t_0)\underline{x}_i = 0 \tag{3.1.19}$$

For all the other $\underline{x}_i$, $\underline{x}_i \in \mathcal{R}\,[\underline{\Phi}_n'(\tau,t_0)\underline{C}'(\tau)]$, we make use of fact II to show,

$$\begin{aligned} \underline{F}_n(\tau)\underline{C}(\tau)\underline{\Phi}_n(\tau,t_0) = {}& \underline{R}^{-1}(\tau)\underline{B}'(\tau)\underline{K}_n(\tau)\underline{\Phi}_n(\tau,t_0)\underline{\Phi}_n'(\tau,t_0)\underline{C}'(\tau) \\ &[\underline{C}(\tau)\underline{\Phi}_n(\tau,t_0)\underline{\Phi}_n'(\tau,t_0)\underline{C}'(\tau)]^{-1}\underline{C}(\tau)\underline{\Phi}_n(\tau,t_0) \end{aligned} \tag{3.1.20}$$

or,

$$\begin{aligned} \underline{F}_n(\tau)\underline{C}(\tau)\underline{\Phi}_n(\tau,t_0) = {}& \underline{R}^{-1}(\tau)\underline{B}'(\tau)\underline{K}_n(\tau)\underline{\Phi}_n(\tau,t_0) \\ &[\underline{\Phi}_n'(\tau,t_0)\underline{C}'(\tau)][\underline{\Phi}_n'(\tau,t_0)\underline{C}'(\tau)]^{\dagger} \end{aligned} \tag{3.1.21}$$

Using Theorem A2, of Appendix A

$$\underline{F}_n(\tau)\underline{C}(\tau)\underline{\Phi}_n(\tau, t_0)\underline{x}_i = \underline{R}^{-1}(\tau)\underline{B}'(\tau)\underline{K}_n(\tau)\underline{\Phi}_n(\tau, t_0)\underline{x}_i \quad \text{for } \underline{x}_i \in \mathcal{R}\,[\underline{\Phi}_n'\underline{C}'] \tag{3.1.22}$$

Thus, for the remaining $\underline{x}_i$, $i \leq m$,

$$\underline{x}_i'\underline{I}_n\underline{x}_i = \underline{x}_i'\underline{\Phi}_n'(\tau, t_0)[\underline{C}'\underline{F}_{n-1}'\underline{R}-\underline{K}_n\underline{B})\underline{R}^{-1}(\underline{R}\underline{F}_{n-1}\underline{C}-\underline{B}'\underline{K}_n)]\underline{\Phi}_n(\tau, t_0)\underline{x}_i \geq 0 \tag{3.1.23}$$

Hence, the integrand in Eq. (3.1.13) is positive and the lemma is proven.

This lemma does not guarantee convergence of the proposed algorithm. For example, the following sequence of matrices satisfies the condition of Eq. (3.1.9) and does not converge.

$$\underline{K}_n(t_0) = \begin{bmatrix} 1+(-1)^n & 0 \\ 0 & [1+\frac{1}{n^2} - (-1)^n] \end{bmatrix}; \quad n = 1, 2, \ldots \tag{3.1.24}$$

$$\mathrm{tr}[\underline{K}_n(t_0)] = 2 + \frac{1}{n^2} \rightarrow 2 \quad \text{as} \quad n \rightarrow \infty \tag{3.1.25}$$

However, the algorithm and the lemma are still useful in solving many problems. First, the algorithm is basically an approximation in policy space type of algorithm. This suggests that convergence, if it occurs, is likely to be fairly rapid and this suggestion is borne out by our experience (5 - 10 iterations were generally sufficient to solve the examples given in Section 3.3). Secondly, the non-convergence of the algorithm would seem to correspond to rather an odd behavior of the system. In particular, one likely cause of non-convergence suggests itself. That is a system with two distinct feedback gain matrices $\underline{F}(t)$, both of which give identical and minimal values of $\hat{J}$, but, both of which give distinct values of the performance matrix $\underline{K}(t_0)$. This possibility suggests, as we mentioned earlier, that more than one solution may exist for our optimization problem.

We remark that one could use a gradient algorithm to generate a solution to the optimization problem proposed in Section 2.2. Such an algorithm would be guaranteed to always reduce the cost at each iteration. However, it would tend to converge fairly slowly.

Finally, in the examples reported herein, we did not encounter any of the convergence problems mentioned above.

3.2 The Computer Program

The Fortran listing and an explanation of the use of the computer program are included in this thesis as Appendix B. The discussion in this section is intended to clarify the purposes and limitations of this program. It is hoped that the reader, armed with this discussion, can decide whether this program is adequate for his purposes. And, if it is not, he can make whatever revisions are necessary with a minimum of effort. With this in mind, we discuss our choice of integration routines and the accuracy of the program and suggest some possible improvements.

The program was essentially determined by three critical choices:

1) How would the necessary storage of $\underline{F}_n(t)$ and $\underline{K}_n(t)$ be accomplished?
2) What integration routine should be used to solve
   a) Equation (3.1.1) for $\underline{K}_{n+1}(t)$?
   b) Equation (3.1.4) for $\underline{\Phi}_{n+1}(t, t_0)$?

Since the program was intended to provide theoretical insight to further our understanding of the general problem rather than to solve specific control problems, the answers to the above questions were primarily dictated by the desire for an easy to write program. Thus, no great effort was expended to generate a particularly efficient (fast) or an extremely accurate program.

The first choice was to use only core storage since using tapes or disks involves a much greater programming effort. At the M.I.T. Computation Center the user has about 70,000 words available in core storage. This number determined the maximum dimensions of the problems we could solve to be approximately $2\, n^2\, N < 5 \times 10^4$ (where n is the dimension of the state-vector and N is the number of time steps).

The choice of an integration routine for Eq. (3.1.1), the equation for $\underline{K}_{n+1}(t)$, was simplified by the fact that it is a linear equation. As a result, a fourth order Runge-Kutta integration routine was easy to write and was, in fact, written. Thus, determination of $\underline{K}_{n+1}(t)$ is quite accurate. On the other hand, the choice of an integration routine for Eq. (3.1.4), the equation for $\underline{\Phi}_{n+1}(t, t_0)$, was fairly difficult. The equation is non-linear and involves the calculation of the inverse of a matrix at each step. As a result, a Runge-Kutta routine would have been complicated to write and comparatively time-consuming to run. As a result, Euler's Method was programmed as a first attempt at integrating Eq. (3.1.4). This routine performed well enough for our immediate purposes and so we have not yet replaced it by a better integration routine.

Having made these choices, and written the program, the question of accuracy arises. It was impossible to obtain a good theoretical estimate of the accuracy of the program. The accuracy was studied experimentally by solving the same problem using different step sizes in the integrations. These experiments suggest two conclusions:

1) The accuracy of the computation depends on $\alpha = \max_{t \epsilon [t_0, T]} || \underline{K}_n^*(t) ||$ and on the ratio of the "time-constant" of $\underline{\Phi}^*(t, t_0)$ to the step size. We found that the accuracy was of the order of magnitude of the step size provided the "time-constant" of $\underline{\Phi}^*(t, t_0) > 100 \times$ (step size) and $\alpha < 10^3$.

2) Reducing the step size, i.e. - improving the accuracy of the program, sometimes reduced the number of iterations required for convergence.

We believe that three improvements in the program would probably be useful. First, the routine for inverting the matrix $\underline{\psi}(t)$ could be more efficient and more accurate. Second, replacing the present integration routine for Eq. (3.1.4) by a Predictor-Corrector scheme would improve the accuracy and, possibly, the speed of the program. Finally, it would be useful to have more flexibility in the choice of a value of $\underline{F}_0(t)$. In particular, for some problems it is necessary to choose a time-varying initial matrix of control gains. All of these improvements will be made in the near future.

3.3 Examples

The computer program described in Appendix B was used to calculate the optimal linear output feedback control for several examples. These examples are discussed below because they provide additional information about the practicality of these theoretical results, about the accuracy and speed of the computer program and about linear output feedback control systems in general.

Example 1 :

The system is :

$$\dot{\underline{\Phi}}(t, 0) = [\underline{A} - \underline{B} f \underline{C}]\underline{\Phi}(t, 0) \quad ; \quad \underline{\Phi}(0, 0) = \underline{I} \tag{3.3.1}$$

and, the performance criterion is :

$$J = \frac{1}{2} \int_0^T tr[\underline{\Phi}'(t, 0)(\underline{Q} + f^2 \underline{C}'\underline{C})\underline{\Phi}(t, 0)]dt \tag{3.3.2}$$

The parameters are :

$$\underline{A} = \begin{bmatrix} 0 & 1 \\ -2 & -3 \end{bmatrix} \quad \underline{B} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad \underline{C} = \begin{bmatrix} 1 & 0 \end{bmatrix} \quad \underline{Q} = \begin{bmatrix} 10 & 0 \\ 0 & 10 \end{bmatrix} \tag{3.3.3}$$

The problem was solved for three values of T :

$$\text{a)} \quad T = 10 \tag{3.3.4}$$

$$\text{b)} \quad T = 8 \tag{3.3.5}$$

$$\text{c)} \quad T = 6 \tag{3.3.6}$$

The open-loop system is both controllable and observable and has the transfer function

$$\frac{y(s)}{u(s)} = \frac{1}{s^2 + 3s + 2} \tag{3.3.7}$$

Thus, the open-loop system has poles at s = -2, s = -1 and is therefore stable.

The solutions to the above problems are plotted in Figs. 1 through 4. Figure 1 is a plot of the optimal trajectory from each of

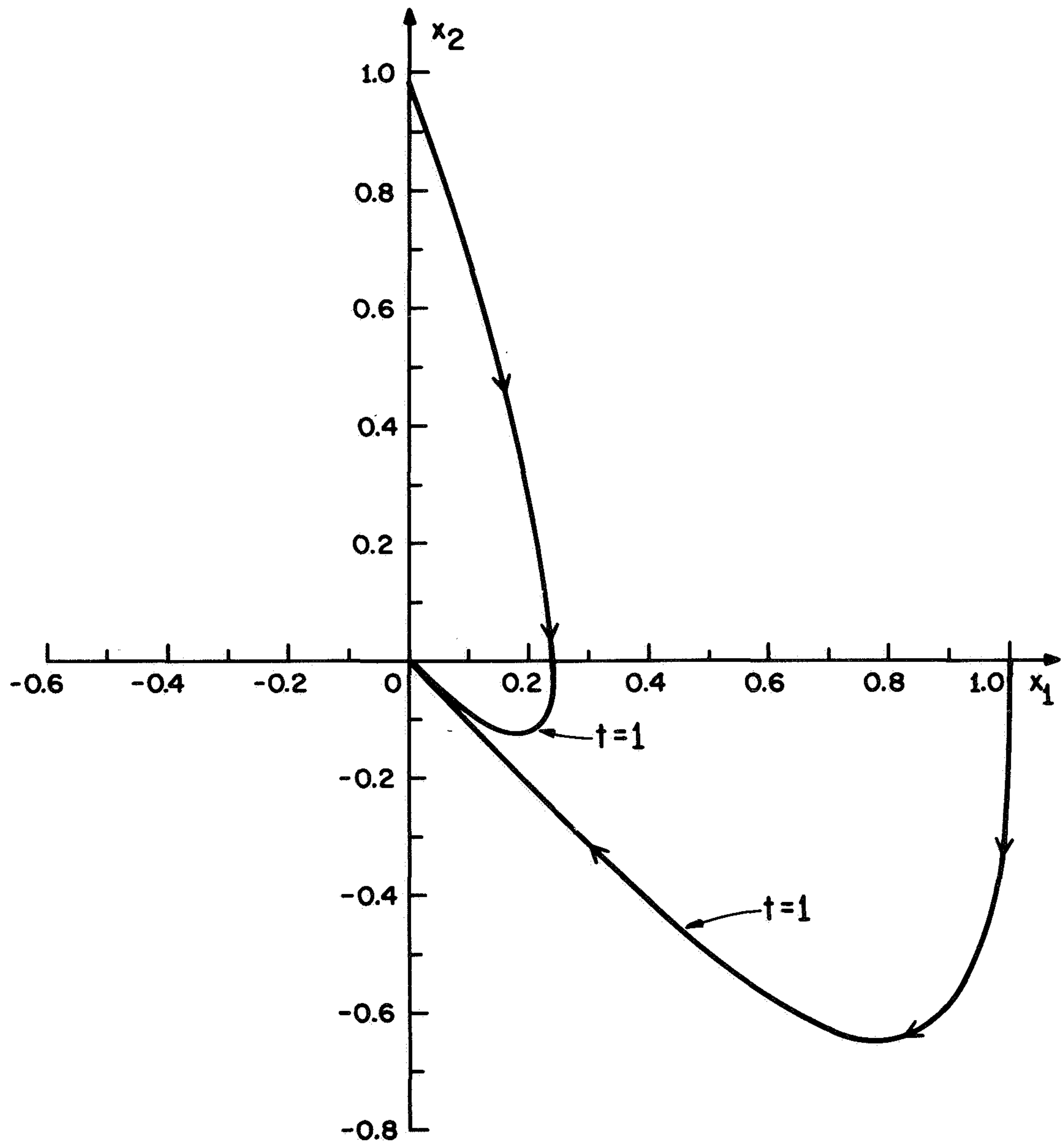


Fig. 1 "Optimal" Trajectories for Example 1 Plotted in the Phase Plane

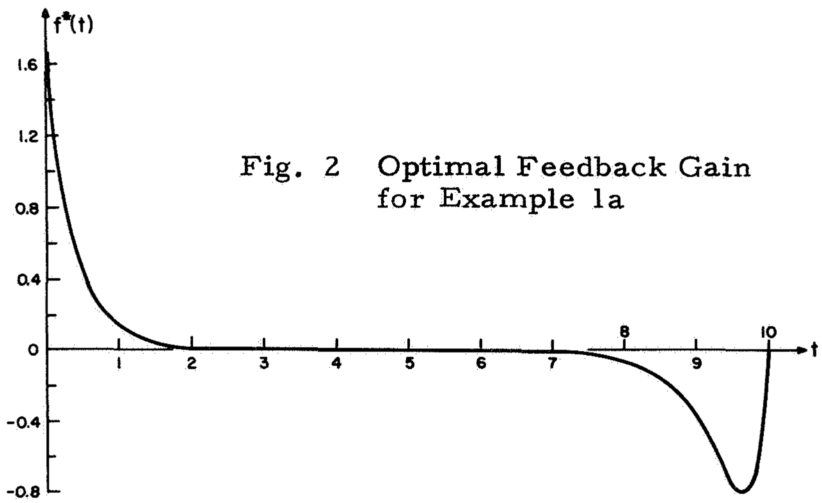


Fig. 2 Optimal Feedback Gain for Example 1a

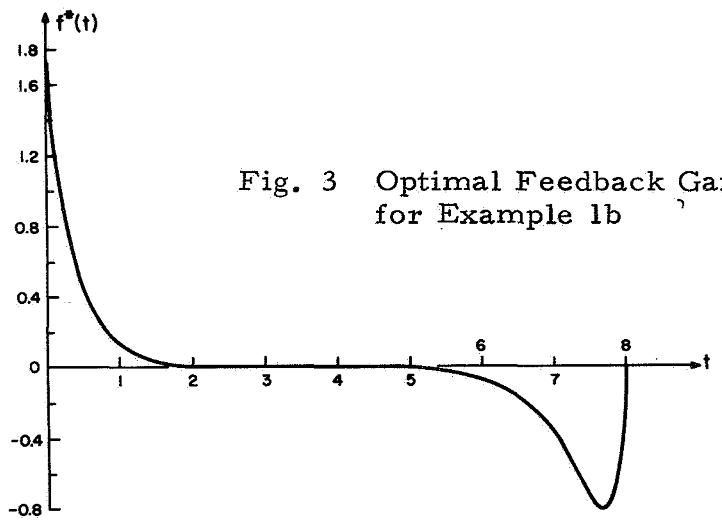


Fig. 3 Optimal Feedback Gain for Example 1b

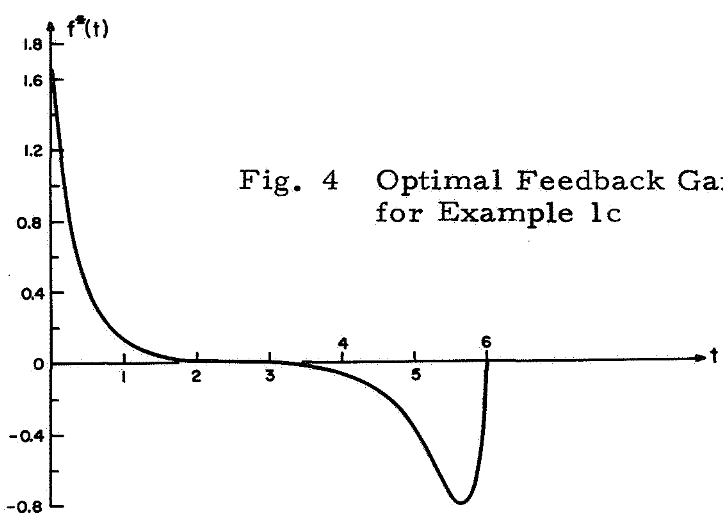


Fig. 4 Optimal Feedback Gain for Example 1c

two independent initial conditions plotted in the phase plane. Thus, in effect, it is a plot of the optimal transition matrix $\underline{\Phi}^{*}(t, 0)$. Figures 2, 3, and 4 are plots of the optimal feedback gain $f^{*}(t)$ versus time for each of the three intervals for which $f^{*}(t)$ was computed.

There are several interesting aspects to these plots. Probably the most striking feature is that all three plots of $f^{*}(t)$ are <u>identical</u> over the first three seconds and over the last three seconds. In fact, the time interval is divisible into three definite parts. Part one is an initial transient lasting about three seconds. Part two is a "steady state" value (which is approximately 0) that is held until part three begins. Part three is a terminal transient which lasts for slightly more than three seconds.

In this example, an explanation for the initial transient and the "steady state" value is suggested by Fig. 1. Notice, in Fig. 1, that all initial states are driven approximately to the line $x_1 = -x_2$ during the initial transient period of $f^{*}(t)$. This line has the following properties:

1) It is an eigenvector of the open-loop system for the eigenvalue $\lambda = -1$. That is, with $f = 0$, the state will decay to zero along the line $x_1 = -x_2$.

2) For the given system, if the initial condition is known to lie on the line $x_1 = -x_2$, and if we compute the time-invariant f which minimizes

$$J' = \frac{1}{2}\int_0^{\infty} [10\, x_1^2(t) + 10\, x_2^2(t) + f^2 x_1^2(t)]\, dt \qquad (3.3.8)$$

for that known initial condition,then the optimal $f = 0$. In fact, the Kalman optimal control for states on the line $x_1 = -x_2$ is zero.

Thus, the first two portions of the time-variation of the gain $f^*(t)$ seem to be explained by:

a) During the initial transient, the initial state, which is uniformly distributed in probability on the surface of the unit sphere, is driven onto the line $x_1 = -x_2$. This "identifies" the state.

b) During the "steady state" interval, the optimal feedback control for the, by now, known state is used.

Unfortunately, we have not yet found as satisfying a physical interpretation of the terminal transient of the feedback gain $f^*(t)$. However, we note that the "average" value of $f^*(t)$, because of the terminal transient, is approximately equal to zero, the "steady-state" value.

Finally, we compare the value of the performance criterion (3.3.2) for three alternative feedback control laws:

1) If $f(t) = 0$, then $\hat{J} = 15$

2) If $f(t) = f^*(t)$, as shown in Fig. 2, then $\hat{J} = 14.1$

3) If we use the Kalman optimal control, i.e. - let $\underline{C} = \underline{I}$, then $\hat{J} = 7.4$.

From these figures, we see that the best position feedback control is about 100% worse than the Kalman optimal control. On the other hand, the time variation in $f^*(t)$ improves the performance of the system by about 6% over the control $f(t) = 0$.

Example 2:

Example 2 should be studied in conjunction with example 3. In both examples, the system and the performance criterion are identical except for the choice of the output matrix $\underline{C}$. Thus, the two examples can be viewed as a study of the comparative value of position feedback versus velocity feedback in a second order servomechanism.

The system equation is identical to Eq. (3.3.1) and the performance criterion is identical to Eq. (3.3.2). The parameters are:

$$\underline{A} = \begin{bmatrix} 0 & 1 \\ -10 & -2 \end{bmatrix} \quad \underline{B} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad \underline{C} = \begin{bmatrix} 1 & 0 \end{bmatrix} \quad \underline{Q} = \begin{bmatrix} 10 & 0 \\ 0 & 10 \end{bmatrix} \tag{3.3.9}$$

$$T = 10 \tag{3.3.10}$$

The open-loop system is controllable and observable and has transfer function:

$$\frac{y(s)}{u(s)} = \frac{1}{s^2 + 2s + 10} \tag{3.3.11}$$

Thus, it is a stable system with poles at $s = -1 \pm j\sqrt{3}$.

The solution to this problem is plotted in Figs. 5 and 6. Again, Fig. 5 is effectively a phase plane plot of the optimal transition matrix $\underline{\Phi}^*(t, 0)$ while Fig. 6 is a plot of $f^*(t)$ versus time. Figure 7 is also devoted to this example except that in Fig. 7, $f^*(t)$ is determined for $T = 6$.

Again, as in the previous example, the graph of $f^*(t)$ (Fig. 6) has an evident initial transient, "steady state" value, and final transient. In Fig. 7, the "steady state" value does not appear because the two transient intervals overlap slightly. It is interesting to note that the initial transient in this example is almost twice as long as the initial transient in the previous example. Furthermore, the final transient exists for only half as long as in the previous example. However, the longest time constant of the open-loop system is identical in both

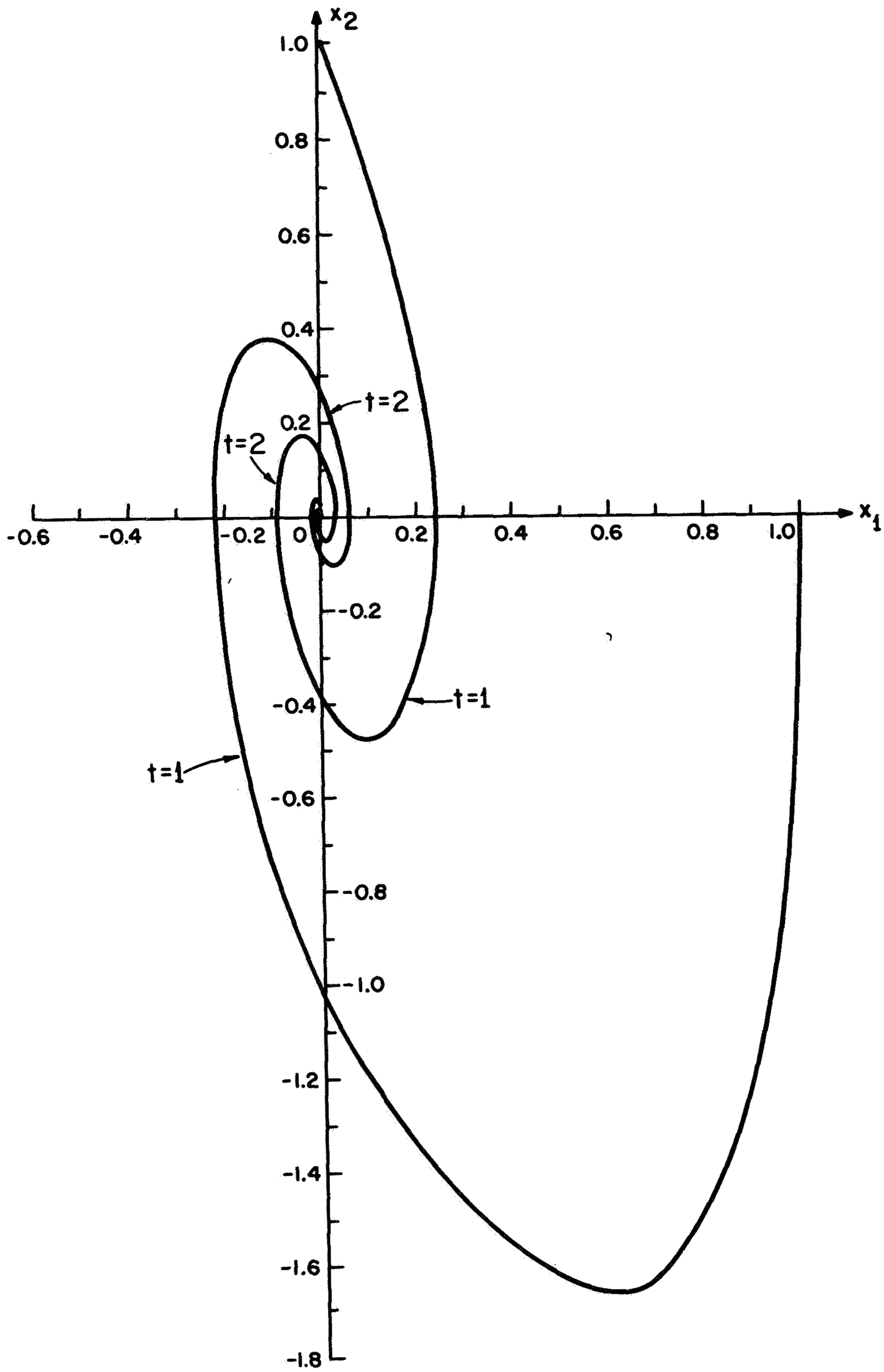


Fig. 5 "Optimal" Trajectories for Example 2
Plotted in the Phase Plane

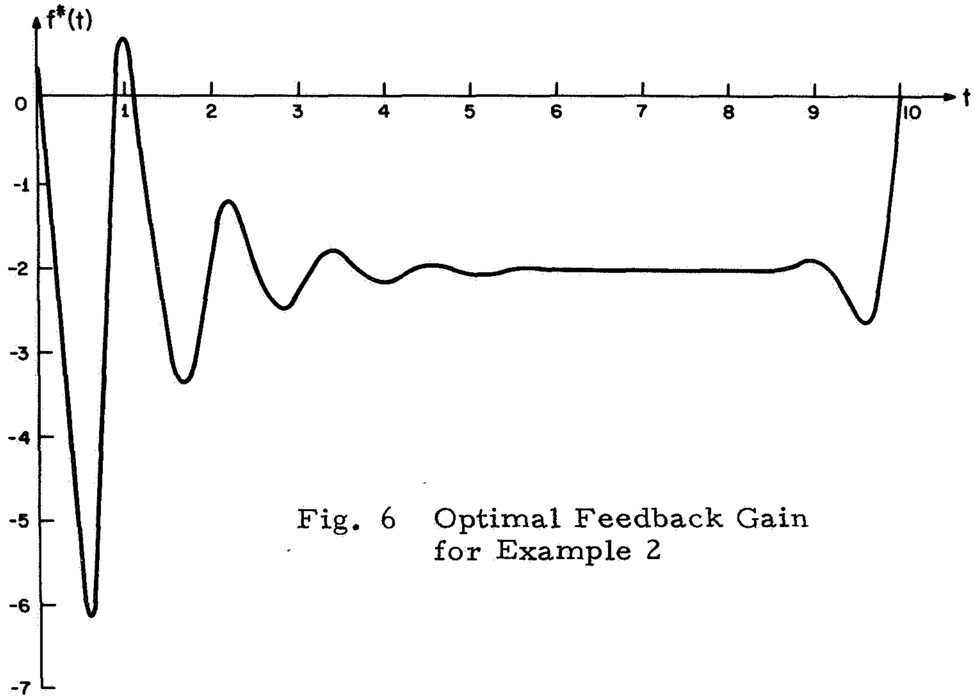


Fig. 6 Optimal Feedback Gain for Example 2

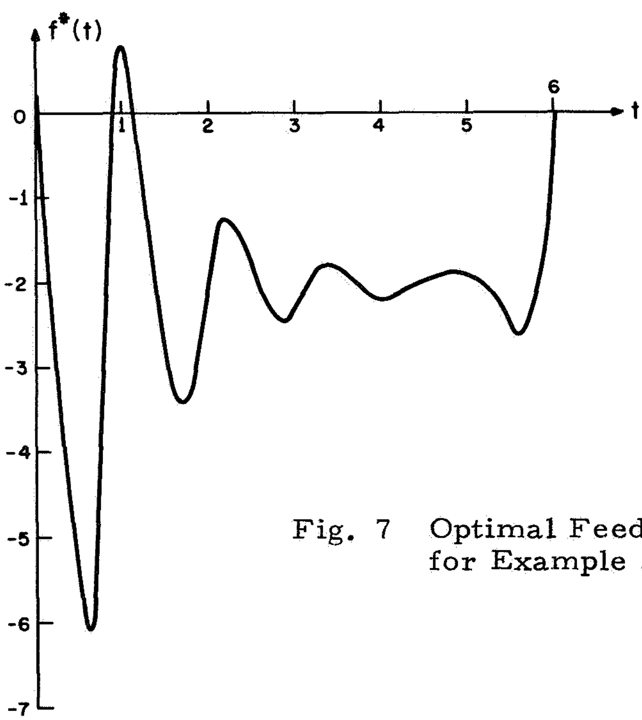


Fig. 7 Optimal Feedback Gain for Example 2 with T = 6

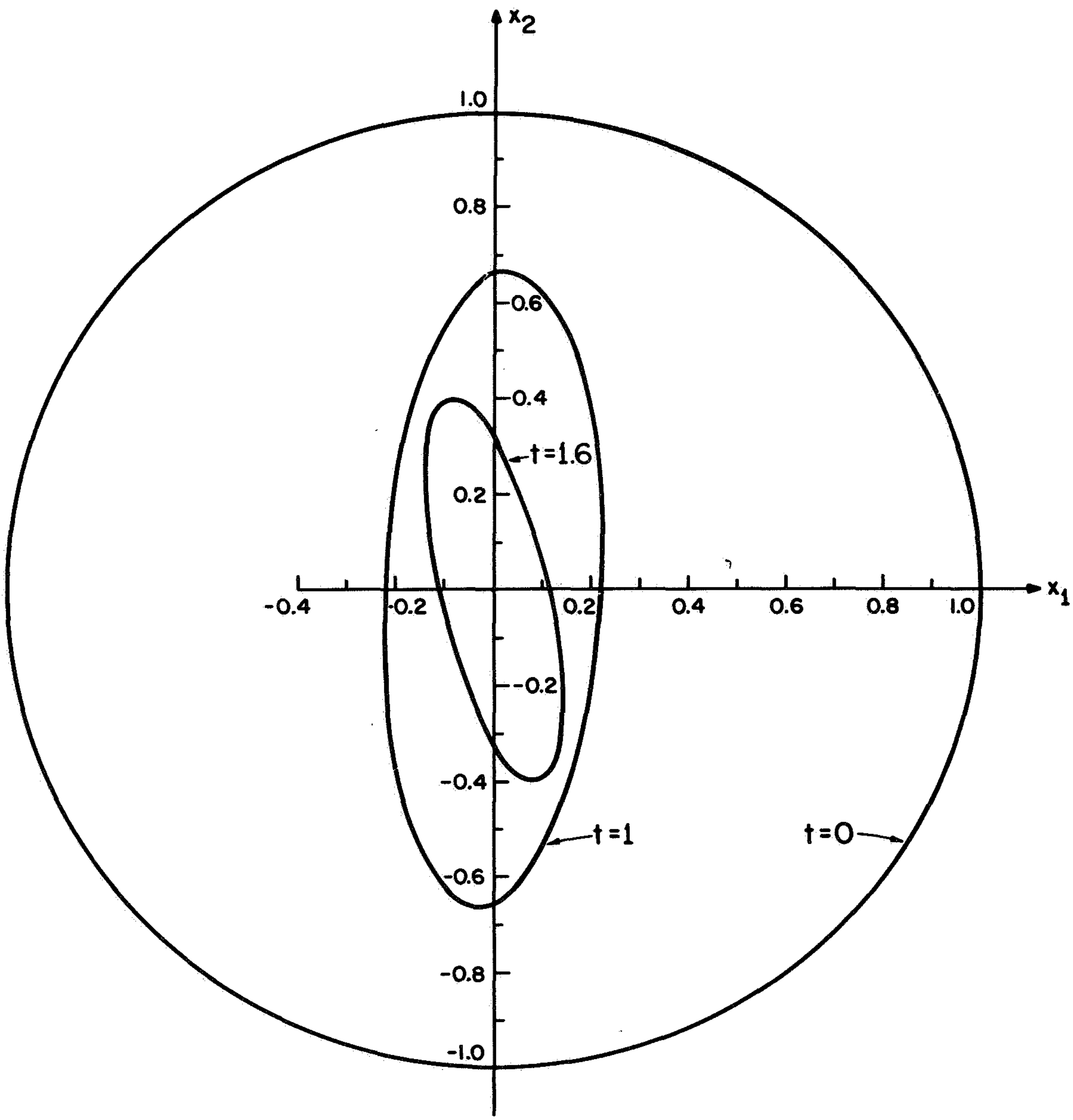


Fig. 8 The Ellipses on which the States are Distributed at $t=0$, $t=1$, $t=1.6$ for the Optimal System of Example 2

examples. Thus, one reasonable conjecture, that the length of the transient periods in $f^*(t)$ is determined by the open-loop time constants of the system, is probably false.

If we study the initial transient of $f^*(t)$ a bit more carefully, we note that it could be approximately described by a decaying exponential multiplied by a sinusoid. It does not seem too far-fetched to suggest that the period of that sinusoid is half the period of the natural oscillations of the open-loop system.

This suggests that a careful attempt to correlate the oscillations in the initial transient of $f^*(t)$ with the values of the optimal transition matrix $\underline{\Phi}^*(t, t_0)$ might be rewarding. First, we remark that the choice of the gain $f^*(t)$ must be based on two pieces of information:

1) The value of $x_1(t)$ (the position) is measured at each instant of time.
2) Although $x_2(t)$ (the velocity) is not measured and is not directly computable, some information about $x_2(t)$ can be obtained from knowledge of $x_1(t)$ and the following fact. The initial state of the system is known to have been uniformly distributed in probability on the surface of the unit sphere. This probability distribution is propagated by the system until, at each instant of time, the state of the system will be distributed in probability on the surface of an ellipse in the phase plane.

A glance at Fig. 1 shows that, in the previous example, after about $t = 1.5$ this ellipse is approximately a -45° line. Thus, knowledge of $x_1(t)$ implies precise knowledge of $x_2(t)$. In the present example, the ellipse on which the state is distributed is less obvious from Fig. 5. As a result, we have plotted these ellipses, for several values of the time, in Fig. 8.

Although it is not plotted in Fig. 8, our calculations show that the ellipse on which the states are distributed at $t = 2.2$ is identical, except in size, to the ellipse which is shown at $t = 1$. In other words, the orientation and the ratio of the length of the major-axis to the length of the minor-axis are the same for both ellipses. Similarly, the ellipse in Fig. 8 that corresponds to $t = 1.6$ is identical, except in size, to those at $t = .6$ and at $t = 2.8$. Thus, the period with which the ellipses repeat corresponds to the period of the transient oscillation in $f^*(t)$. The above are experimental conclusions. They are buttressed by the theoretical fact that, for a second-order, linear, time-invariant system with poles of the form $s = a \pm j\omega$ $(\omega \neq 0)$, the period with which the ellipses repeat is half the period of the natural oscillations of the system.

This correlation between the two periods suggests that we study the relation between the evolution of the ellipses and $f^*(t)$ even more closely. Unfortunately, this will require a great deal of additional computer programming. For the moment, we content ourselves with the following conjecture. The optimal control, $f^*(t)$ attempts to perform two operations simultaneously. The first is to improve its information by shaping the ellipse and thereby increasing the correlation between the measured variable and the unknown variable. The second, and probably most important, is to drive the state in a direction that minimizes the cost.

So far, our analysis has been concentrated on the initial transient. In Chapter IV we develop the tools needed to examine the "steady state" more closely.

We will discuss the performance $\hat{J}$ obtained for this example at the conclusion of the third example.

Example 3:

This example is identical to the previous example with the single exception that:

$$\underline{C} = [0 \quad 1] \tag{3.3.12}$$

Otherwise, the parameters are identical to those of Eqs. (3.3.9) and (3.3.10). Thus, the difference is that we now feed back the velocity rather than the position.

The open-loop system is again both controllable and observable and has transfer function:

$$\frac{y(s)}{u(s)} = \frac{s+2}{s^2 + 2s + 10} \tag{3.3.13}$$

It, therefore, has poles at $s = -1 \pm j\sqrt{3}$ and a zero at $s = -2$ and is stable.

The optimal transition matrix for this problem is plotted in Fig. 9 and the optimal feedback gain in Fig. 10. These graphs have two striking features:

1) The frequency of the oscillations in $f^*(t)$ is twice the frequency of the oscillations of $\underline{\Phi}^*(t, t_0)$.

2) The initial transient in $f^*(t)$, if it is a transient, lasts for nearly the entire time interval.

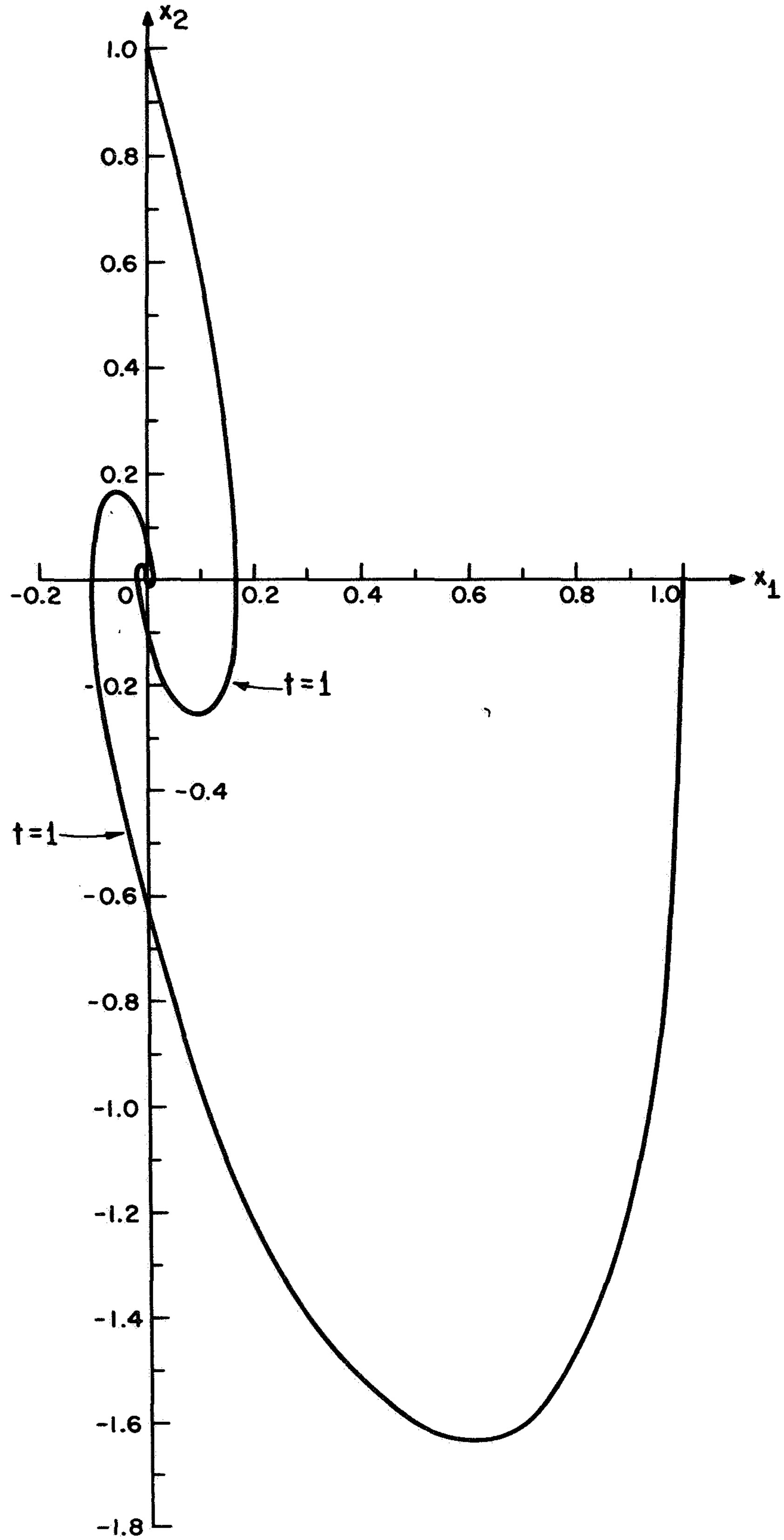


Fig. 9 "Optimal" Trajectories for Example 3 Plotted in the Phase Plane

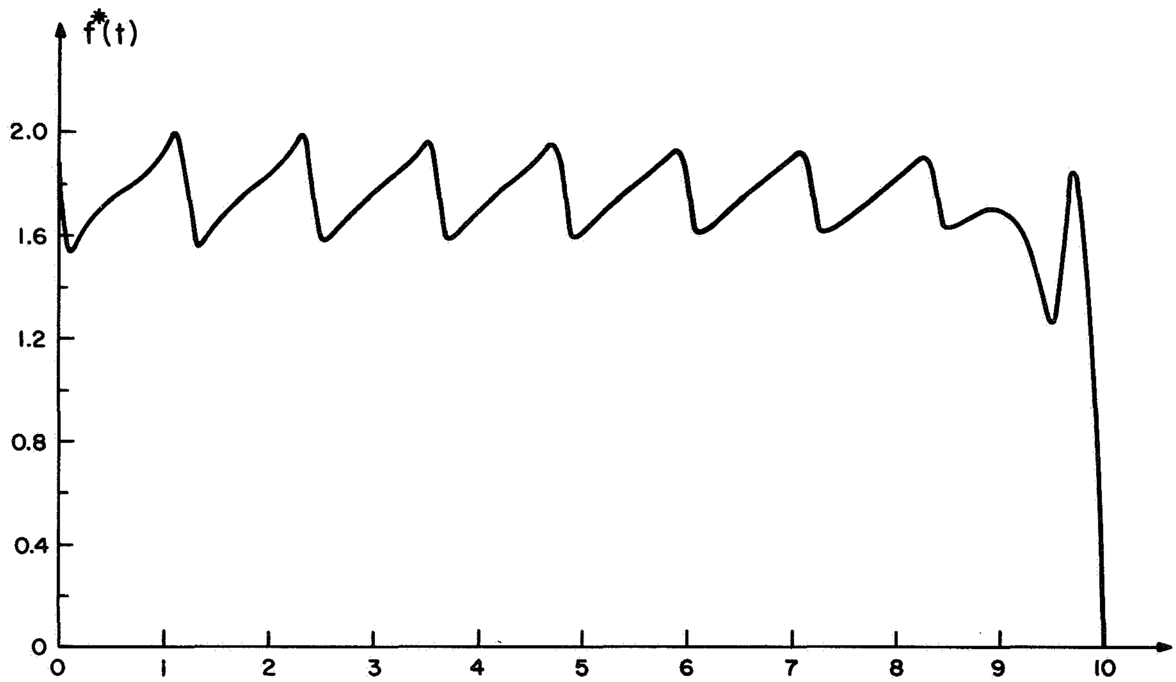


Fig. 10 Optimal Feedback Gain for Example 3

We believe that the first of these features is explained, as it was for the previous example, by the repetition frequency of the ellipses on which the state is distributed. The second feature must be due to the only change between this example and its predecessor, the change in the $\underline{C}$ matrix. Why the change from position to velocity feedback should produce this particular change in $f^*(t)$ is not yet understood.

One expects velocity feedback to perform better than position feedback for this system. This is borne out by the following data for the system (3.3.1), criterion (3.3.2) and parameters

$$\underline{A} = \begin{bmatrix} 0 & 1 \\ -10 & -2 \end{bmatrix} \qquad \underline{B} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \qquad \underline{Q} = \begin{bmatrix} 10 & 0 \\ 0 & 10 \end{bmatrix} \tag{3.3.14}$$

These parameters correspond to examples 2 and 3.

We compare four alternative feedback controls.

1) Let $f_1(t) = 0$. The performance of the control can always be written as $\hat{J} = \mathrm{tr}[\underline{K}_f]$. In this case,

$$\underline{K}_{f1} = \begin{bmatrix} 28.5 & .5 \\ .5 & 2.75 \end{bmatrix} \tag{3.3.15}$$

2) Let $\underline{C} = [1 \quad 0]$. This yields pure position feedback and corresponds to example 2. Then, $f_2(t) = f^*(t)$ as shown in Fig. 6. The performance is given by $\mathrm{tr}[\underline{K}_{f2}]$ where:

$$\underline{K}_{f2} = \begin{bmatrix} 21.8 & .39 \\ .39 & 3.95 \end{bmatrix} \tag{3.3.16}$$

3) Let $\underline{C} = [0 \quad 1]$. This yields pure velocity feedback and corresponds to example 3. Then $f_3(t) = f^*(t)$ as shown in Fig. 10. And,

$$\underline{K}_{f3} = \begin{bmatrix} 20.6 & .5 \\ .5 & 1.88 \end{bmatrix} \qquad (3.3.17)$$

4) Let $\underline{C} = \underline{I}$. Let the feedback gain matrix be the Kalman optimal $\underline{F}^*(t)$. Then,

$$\underline{K}^* = \begin{bmatrix} 20.6 & .49 \\ .49 & 1.87 \end{bmatrix} \qquad (3.3.18)$$

Studying these performances leads to two observations:

a) The optimal velocity feedback control, $f_3(t)$, performs essentially as well as the best possible feedback control, the Kalman optimal control. Both of these controls are, in terms of $\hat{J}$, approximately 30% better than $f(t) = 0$.

b) The optimal position feedback control, $f_2(t)$, is about 16% better than $f(t) = 0$ in terms of $\hat{J}$. However, for some initial conditions (e.g. $\underline{x}_0' = [0 \quad 1]$) $f_2(t)$ is actually worse than $f(t) = 0$.

The second observation is explained by the fact that $\hat{J}$ is an "average" performance measure and $f_2(t)$ minimizes the "average" performance. The first observation partly justifies this research by demonstrating that excellent control laws are possible using only output feedback.

CHAPTER IV

OPTIMAL TIME-INVARIANT OUTPUT-FEEDBACK PROBLEMS

For many practical purposes one wants the matrix of feedback gains to be constant. In addition, the discovery of the conditions under which the optimal feedback gains are time-invariant is one of the important theoretical questions in optimal control theory. If one reasons by analogy to the standard linear regulator problem, one might conjecture that the optimal feedback matrix $\underline{F}^*(t)$ found in Chapter II is time-invariant under the added hypothesis that T, the terminal time, tends to ∞ and that $\underline{A}$, $\underline{B}$, $\underline{C}$, $\underline{Q}$ and $\underline{R}$ are constant. In the first section of this chapter we offer evidence that suggests this conjecture is false, at least in general.

After this, we assume that $\underline{F}^*$ is constant and derive the steady state optimal regulator solution by assuming that $\underline{C} = \underline{I}$. We then drop the hypothesis that $\underline{C}$ is invertible and derive the equivalent result to that of Chapter II, for $\underline{F}^*$ time-invariant. We conclude the chapter with several examples.

4.1 The Limiting Case, $T \rightarrow \infty$

The first problem we wish to discuss is the problem of Chapter II formulated as the exact analog of the Kalman linear regulator on the semi-infinite interval. Thus,

$$\dot{\underline{\Phi}}(t, t_0) = [\underline{A} - \underline{B}\,\underline{F}(t)\underline{C}]\underline{\Phi}(t, t_0) \quad ; \quad \underline{\Phi}(t_0, t_0) = \underline{I} \qquad (4.1.1)$$

$$\hat{J} = \frac{1}{2}\int_{t_0}^{\infty} \mathrm{tr}[\underline{\Phi}'(t,t_0)\{\underline{Q}+\underline{C}'\underline{F}'(t)\underline{R}\,\underline{F}(t)\underline{C}\}\underline{\Phi}(t,t_0)dt \qquad (4.1.2)$$

where $\underline{A}$, $\underline{B}$, $\underline{C}$, $\underline{Q}$ and $\underline{R}$ are constant real matrices of appropriate dimensions and properties (see Section 2.2).

The problem is to find a measurable $\underline{F}^*(t)$ which minimizes the performance criterion (4.1.2) subject to the constraint imposed by the system equation (4.1.1), assuming such an $\underline{F}^*(t)$ exists. In addition, one would like to find conditions on $\underline{A}$, $\underline{B}$, $\underline{C}$, $\underline{Q}$ and $\underline{R}$ which will guarantee the existence of an optimal $\underline{F}^*(t)$.

We were unable to solve this problem. We can, however, give some indication of the difficulties involved in its solution by discussing some of our attempts to solve it. One approach that we tried was to attempt to find $\underline{F}^*(t)$ as the limit, as $T \to \infty$, of the solutions of finite time problems. In particular, those finite time problems which were solved in Chapter II. It is well known that this approach works quite nicely in the case of full state feedback ($\underline{C}^{-1}$ exists). Unfortunately, when $\underline{C}$ is not invertible the solution of the finite time problems involves a two point boundary value problem. We found the technical difficulties in extending these two point boundary value problems to the semi-infinite interval insurmountable.

Another approach that was attempted is basically a version of the inverse problem of the calculus of variations. That is, given the problem described by Eqs. (4.1.1) and (4.1.2), assume an $\underline{F}^*(t)$ and try to find conditions on $\underline{A}$, $\underline{B}$, $\underline{C}$, $\underline{Q}$ and $\underline{R}$ which will guarantee that $\underline{F}^*(t)$ is optimal. Specifically, this approach was taken assuming

$\underline{F}^*(t)$ = constant and when that failed, assuming $\underline{F}^*(t)$ was periodic. No useful results were obtained for $\underline{F}^*(t)$ assumed periodic.

In connection with the possibility that $\underline{F}^*(t)$ might be time-invariant, the following result is useful.

We can prove that there exist cases where the optimal control $\underline{F}^*(t)$ for the problem described by Eqs. (4.1.1) and (4.1.2) must be time-varying by citing the following counter-example (due to Brockett and Lee, [16]). Given the time-invariant linear system:

$$\dot{\underline{\Phi}}(t, 0) = \left(\begin{bmatrix} 0 & 1 \\ -2 & 2 \end{bmatrix} - \begin{bmatrix} 0 \\ 1 \end{bmatrix} f(t) \begin{bmatrix} -4 & 3 \end{bmatrix} \right) \underline{\Phi}(t, 0) \quad ; \quad \underline{\Phi}(0, 0) = \underline{I} \qquad (4.1.3)$$

or, equivalently, in terms of the transforms :

$$\frac{y(s)}{u(s)} = \frac{3s - 4}{s^2 - 2s + 2} \qquad (4.1.4)$$

a) There exists no constant real f which stabilizes this system.

b) There exists a stabilizing f(t) given by:

$$f(t) = \begin{cases} 0 & \text{for } \; 0 \leq t - nT < T_1 \\ 1 & \text{for } \; T_1 \leq t - nT < T \end{cases} \qquad n = 0, 1, 2, \ldots \qquad (4.1.5)$$

$$T_1 = \tan^{-1} 3 \qquad T = \tan^{-1} 3 + \pi$$

Therefore, for the system (4.1.3) and any reasonable performance measure of the form of Eq. (4.1.2), the optimal output feedback control cannot be time-invariant.

The above example leads one to the belief that the conditions under which $\underline{F}^*(t) \rightarrow \underline{F}^*$ as $T \rightarrow \infty$ are very complex, especially since the system (4.1.3) is both controllable and observable. Rather than belabor a problem we have not been able to solve or bore the reader with our own intuition, we reformulate the problem in the following section and demand that $\underline{F}^*$ be constant. As we shall see, there is a real possibility that this approach will lead back to answers to the problem in this section.

4.2 Reformulation of the Problem

We will consider the following optimization problem:

Given the time-invariant linear system:

$$\dot{\underline{\Phi}}(t,0) = [\underline{A} - \underline{B}\,\underline{F}\,\underline{C}]\underline{\Phi}(t,0) \quad ; \quad \underline{\Phi}(0,0) = \underline{I} \tag{4.2.1}$$

It is well known that:

$$\underline{\Phi}(t,0) = e^{[\underline{A} - \underline{B}\,\underline{F}\,\underline{C}]t} \tag{4.2.2}$$

Given also the performance criterion:

$$\hat{J} = \frac{1}{2}\int_0^{\infty} \mathrm{tr}[\underline{\Phi}'(t,0)(\underline{Q} + \underline{C}'\underline{F}'\underline{R}\,\underline{F}\,\underline{C})\underline{\Phi}(t,0)]\,dt \tag{4.2.3}$$

Find that $\underline{F}^*$ which minimizes the performance criterion (4.2.3) subject to the constraint imposed by the system (4.2.1).

For the sake of completeness,

- $\underline{A}$ is an nxn real constant matrix
- $\underline{B}$ is an nxm real constant matrix
- $\underline{C}$ is an rxn real constant matrix of rank r
- $\underline{Q}$ is an nxn symmetric positive semi-definite real constant matrix
- $\underline{R}$ is an mxm symmetric positive definite real constant matrix
- $\underline{F}$ is an mxr real constant matrix

4.3 Solution Assuming $\underline{C} = \underline{I}$

In the case that $\underline{C} = \underline{I}$ we have the Kalman time-invariant linear regulator problem. The solution to this problem is well known and we shall, in fact, simply re-derive the conditions which $\underline{F}^*$, the optimal control, must satisfy. The derivation will proceed formally at first and then we will state and prove a theorem which guarantees the validity of all the prior assumptions. A similar derivation was given by Luenberger [17]. We remark that we develop the tools in this section which we will use in the following section where $\underline{C}$ is not invertible.

We begin by using Eq. (4.2.2) to rewrite the performance criterion, Eq. (4.2.3), as

$$\hat{J}(\underline{F}) = \frac{1}{2}\,\mathrm{tr}\int_0^\infty e^{[\underline{A}-\underline{B}\,\underline{F}]'t}(\underline{Q}+\underline{F}'\underline{R}\,\underline{F})e^{[\underline{A}-\underline{B}\,\underline{F}]t}dt \qquad (4.3.1)$$

It should be noted that $J(\underline{F})$ in Eq. (4.3.1) is a real function of $m \times r$ variables (the f_{ij}). A necessary condition for $\underline{F}^*$ to minimize such a function is that $\left.\frac{d\hat{J}}{d\underline{F}}\right|_{\underline{F}^*} = 0$. We shall simply calculate and evaluate the necessary derivative.

A key lemma in this calculation is the following, due to Kleinman [15].

<u>Lemma</u> - Let $f(\underline{X})$ be a trace function. Then if we can write

$$f(\underline{X}+\epsilon\,\Delta\underline{X}) - f(\underline{X}) = \epsilon\,\mathrm{tr}[\underline{M}(\underline{X})\Delta\underline{X}] \qquad (4.3.2)$$

as $\epsilon \to 0$, where $\underline{M}(\underline{X})$ is an $n \times r$ matrix, $\underline{X}$ is an $r \times n$ matrix, we have

$$\frac{\partial f(\underline{X})}{\partial \underline{X}} = \underline{M}'(\underline{X}) \qquad (4.3.3)$$

For completeness, a trace function is defined by:

Definition - $f(\cdot)$ is a trace function of the matrix $\underline{X}$ if $f(\underline{X})$ is of the form

$$f(\underline{X}) = \mathrm{tr}[\underline{F}(\underline{X})] \tag{4.3.4}$$

where $\underline{F}(\cdot)$ is a continuously differentiable mapping from the space of $r \times n$ matrices into the space of $n \times n$ matrices.

Example -

$$\text{Let} \quad \underline{F}(\underline{X}) = e^{(\underline{A}+\underline{B}\,\underline{X})t} \tag{4.3.5}$$

then,

$$\underline{F}(\underline{X}+\epsilon\,\underline{\Delta X}) = e^{(\underline{A}+\underline{B}\,\underline{X}+\epsilon\,\underline{B}\,\underline{\Delta X})t} \tag{4.3.6}$$

But, from p. 171, reference [18] we have that (4.3.6) is, to first order in ϵ,

$$F(\underline{X}+\epsilon\,\underline{\Delta X}) = e^{(\underline{A}+\underline{B}\,\underline{X})t} + \epsilon \int_0^t e^{(\underline{A}+\underline{B}\,\underline{X})(t-\sigma)}\,\underline{B}\,\underline{\Delta X}\, e^{(\underline{A}+\underline{B}\,\underline{X})\sigma}\, d\sigma \tag{4.3.7}$$

Hence,

$$\mathcal{L}(\underline{\Delta X}) = \int_0^t e^{(\underline{A}+\underline{B}\,\underline{X})(t-\sigma)}\,\underline{B}\,\underline{\Delta X}\; e^{(\underline{A}+\underline{B}\,\underline{X})\sigma}\, d\sigma \tag{4.3.8}$$

and so, since the trace operation commutes with integration, we obtain

$$\mathrm{tr}[\mathcal{L}(\underline{\Delta X})] = \mathrm{tr}\left[\int_0^t e^{(\underline{A}+\underline{B}\,\underline{X})\sigma}\, e^{(\underline{A}+\underline{B}\,\underline{X})(t-\sigma)}\, d\sigma \cdot \underline{B}\,\underline{\Delta X}\right] \tag{4.3.9}$$

$$= \mathrm{tr}\left[e^{(\underline{A}+\underline{B}\,\underline{X})t}\,\underline{B}\cdot(\underline{\Delta X})\right] \tag{4.3.10}$$

Therefore,

$$\frac{\partial}{\partial \underline{X}} \operatorname{tr}\left[e^{(\underline{A}+\underline{B}\,\underline{X})t}\right] = \underline{B}' e^{[\underline{A}+\underline{B}\,\underline{X}]' t} \tag{4.3.11}$$

Now, with the lemma and the example for guidance, we proceed with the derivation. We begin by defining the convenience

$$\underline{A}^{o} \triangleq \underline{A} - \underline{B}\,\underline{F} \tag{4.3.12}$$

Then,

$$\hat{J}(\underline{F}+\epsilon\underline{\Delta F}) = \frac{1}{2}\operatorname{tr}\left[\int_0^{\infty} e^{[\underline{A}^{o}-\epsilon\underline{B}\,\underline{\Delta F}]' t}(\underline{Q}+[\underline{F}'+\epsilon\underline{\Delta F}']\underline{R}[\underline{F}+\epsilon\underline{\Delta F}])\, e^{[\underline{A}^{o}+\epsilon\underline{B}\,\underline{\Delta F}]t} dt\right] \tag{4.3.13}$$

Using Eq. (4.3.7) from the example, we obtain the following equation, accurate to first order in ϵ,

$$\begin{aligned}
\hat{J}(\underline{F}+\underline{\Delta F}) = \frac{1}{2}\operatorname{tr}\int_0^{\infty} \Bigg\{ & e^{\underline{A}^{o\prime} t}(\underline{Q}+\underline{F}'\underline{R}\,\underline{F})e^{\underline{A}^{o} t} + \epsilon e^{\underline{A}^{o\prime} t}(\underline{\Delta F}'\underline{R}\,\underline{F})e^{\underline{A}^{o} t} + \epsilon e^{\underline{A}^{o\prime} t}(\underline{F}'\underline{R}\,\underline{\Delta F})e^{\underline{A}^{o} t} \\
& - \epsilon\left(\int_0^t e^{\underline{A}^{o\prime}(t-\sigma)}\underline{\Delta F}'\underline{B}' e^{\underline{A}^{o\prime}\sigma} d\sigma\right)(\underline{Q}+\underline{F}'\underline{R}\,\underline{F})e^{\underline{A}^{o} t} \\
& - \epsilon\, e^{\underline{A}^{o\prime} t}(\underline{Q}+\underline{F}'\underline{R}\,\underline{F})\left(\int_0^t e^{\underline{A}^{o}(t-\sigma)}\underline{B}\,\underline{\Delta F}\, e^{\underline{A}^{o}\sigma} d\sigma\right)\Bigg\} dt
\end{aligned} \tag{4.3.14}$$

Therefore,

$$\begin{aligned}
\hat{J}(\underline{F}+\epsilon\underline{\Delta F}) - \hat{J}(\underline{F}) = \frac{1}{2}\int_0^{\infty} \epsilon \operatorname{tr}\Bigg[& 2e^{\underline{A}^{o} t} e^{\underline{A}^{o\prime} t}\underline{F}'\underline{R}\,\underline{\Delta F} - \int_0^t e^{\underline{A}^{o}\sigma} e^{\underline{A}^{o\prime} t}(\underline{Q}+\underline{F}'\underline{R}\,\underline{F})e^{\underline{A}^{o}(t-\sigma)}\underline{B}\,\underline{\Delta F} d\sigma \\
& - \int_0^t e^{\underline{A}^{o}(t-\sigma)} e^{\underline{A}^{o\prime} t}(\underline{Q}+\underline{F}'\underline{R}\,\underline{F})e^{\underline{A}^{o}\sigma}\underline{B}\,\underline{\Delta F}\, d\sigma \Bigg] dt
\end{aligned} \tag{4.3.15}$$

Thus, using the lemma,

$$\frac{\partial \hat{J}}{\partial \underline{F}} = \int_0^\infty \underline{R}\underline{F} e^{\underline{A}^o t} e^{\underline{A}^{o'} t} dt - \frac{1}{2}\int_0^\infty \int_0^t \underline{B}' e^{\underline{A}^{o'}(t-\sigma)} (\underline{Q}+\underline{F}'\underline{R}\,\underline{F}) e^{\underline{A}^o t} e^{\underline{A}^{o'}\sigma} d\sigma\, dt$$

$$- \frac{1}{2}\int_0^\infty \int_0^t \underline{B}' e^{\underline{A}^{o'}\sigma} (\underline{Q}+\underline{F}'\underline{R}\,\underline{F}) e^{\underline{A}^o t} e^{\underline{A}^{o'}(t-\sigma)} d\sigma\, dt \qquad (4.3.16)$$

This is, of course, an answer. However, some manipulation is necessary before it can be used. Fortunately, this manipulation is possible. We begin by defining

$$\underline{\Gamma} \stackrel{\Delta}{=} \int_0^\infty \int_0^t \underline{B}' e^{\underline{A}^{o'}\sigma} (\underline{Q}+\underline{F}'\underline{R}\,\underline{F}) e^{\underline{A}^o t} e^{\underline{A}^{o'}(t-\sigma)} d\sigma\, dt \qquad (4.3.17)$$

Let $\sigma = t-\sigma_1$

Then, $d\sigma = -d\sigma_1$ (t is constant)

Note that at $\sigma = 0$, $\sigma_1 = t$

at $\sigma = t$, $\sigma_1 = 0$

With the above substitution,

$$\underline{\Gamma} = -\int_0^\infty \int_t^0 \underline{B}' e^{\underline{A}^{o'}(t-\sigma_1)} (\underline{Q}+\underline{F}'\underline{R}\,\underline{F}) e^{\underline{A}^o t} e^{\underline{A}^{o'}\sigma_1} d\sigma_1 dt \qquad (4.3.18)$$

Substituting Eq. (4.3.18) into Eq. (4.3.16) we obtain,

$$\frac{\partial \hat{J}}{\partial \underline{F}} = \int_0^\infty \underline{R}\underline{F} e^{\underline{A}^o t} e^{\underline{A}^{o'} t} dt - \int_0^\infty \int_0^t \underline{B}' e^{\underline{A}^{o'}(t-\sigma)} (\underline{Q}+\underline{F}'\underline{R}\,\underline{F}) e^{\underline{A}^o t} e^{\underline{A}^{o'}\sigma} d\sigma\, dt$$

(4.3.19)

Assuming that the required integrals exist, the next step is to interchange the order of integration. We begin this by defining :

$$\underline{\chi} = \int_0^{\infty} \int_0^{t} \underline{B}' e^{\underline{A}^{o'}(t-\sigma)} (\underline{Q}+\underline{F}'\underline{R}\,\underline{F}) e^{\underline{A}^{o}t} e^{\underline{A}^{o'}\sigma} \, d\sigma\, dt \tag{4.3.20}$$

Interchanging the order of integration:

$$\underline{\chi} = \int_0^{\infty} \int_{\sigma}^{\infty} \underline{B}' e^{\underline{A}^{o'}(t-\sigma)} (\underline{Q}+\underline{F}'\underline{R}\,\underline{F}) e^{\underline{A}^{o}t} e^{\underline{A}^{o'}\sigma} \, dt\, d\sigma \tag{4.3.21}$$

Let $\tau = t-\sigma$

Then, $d\tau = dt$ (σ is constant)

Note that at $t = \sigma$, $\tau = 0$

at $t = \infty$, $\tau = \infty$

With the above substitutions,

$$\underline{\chi} = \int_0^{\infty} \int_0^{\infty} \underline{B}' e^{\underline{A}^{o'}\tau} (\underline{Q}+\underline{F}'\underline{R}\,\underline{F}) e^{\underline{A}^{o}\tau} e^{\underline{A}^{o}\sigma} e^{\underline{A}^{o'}\sigma} \, d\tau\, d\sigma \tag{4.3.22}$$

or,

$$\underline{\chi} = \int_0^{\infty} \underline{B}' e^{\underline{A}^{o'}\tau} (\underline{Q}+\underline{F}'\underline{R}\,\underline{F}) e^{\underline{A}^{o}\tau} \, d\tau \int_0^{\infty} e^{\underline{A}^{o}\sigma} e^{\underline{A}^{o'}\sigma} \, d\sigma \tag{4.3.23}$$

Substituting Eq. (4.3.23) into Eq. (4.3.19) we obtain:

$$\frac{\partial \hat{J}}{\partial \underline{F}} = \underline{R}\,\underline{F} \int_0^{\infty} e^{\underline{A}^{o}\sigma} e^{\underline{A}^{o'}\sigma} \, d\sigma - \underline{B}' \int_0^{\infty} e^{\underline{A}^{o'}\tau} (\underline{Q}+\underline{F}'\underline{R}\,\underline{F}) e^{\underline{A}^{o}\tau} \, d\tau \int_0^{\infty} e^{\underline{A}^{o}\sigma} e^{\underline{A}^{o'}\sigma} \, d\sigma \tag{4.3.24}$$

Setting $\left.\frac{\partial \hat{J}}{\partial \underline{F}}\right|_{\underline{F}^*} = 0$, we obtain

$$\underline{F}^* = \underline{R}^{-1}\underline{B}' \int_0^{\infty} e^{[\underline{A}-\underline{B}\,\underline{F}^*]'t} (\underline{Q}+\underline{F}^{*'}\underline{R}\,\underline{F}^*) e^{[\underline{A}-\underline{B}\,\underline{F}^*]t} \, dt \tag{4.3.25}$$

Define:

$$\underline{K}^* = \int_0^\infty e^{[\underline{A}-\underline{B}\,\underline{F}^*]'t}(\underline{Q}+\underline{F}^{*\prime}\underline{R}\,\underline{F}^*)e^{[\underline{A}-\underline{B}\,\underline{F}^*]t}dt \qquad (4.3.26)$$

Equations (4.3.25) and (4.3.26) are fairly close to the solution of the problem, assuming that the required integral exists. The following theorem guarantees existence and uniqueness under the assumptions that are well known to be necessary.

Theorem 4.1 Given the linear time-invariant system (4.2.1) and the performance criterion (4.2.3) and

a) the matrix $[\underline{B}, \underline{A}\,\underline{B}, \underline{A}^2\underline{B}, \ldots, \underline{A}^{n-1}\underline{B}]$ is of rank n (controllability[7])

b) the matrix $[\underline{H}', \underline{A}'\underline{H}', \ldots, (\underline{A}')^{n-1}\underline{H}']$ is of rank n, where $\underline{Q} = \underline{H}'\underline{H}$; (observability[7])

then the constant matrix $\underline{F}^*$ which minimizes the cost functional (4.2.3) is given by

$$\underline{F}^* = \underline{R}^{-1}\underline{B}'\underline{K}^* \qquad (4.3.27)$$

and $\underline{K}^*$ is the unique positive definite solution of either Eqs. (4.3.26) and (4.3.25) or, equivalently, of Eq. (4.3.28) below:

$$\underline{0} = \underline{K}^*\underline{A} + \underline{A}'\underline{K}^* + \underline{Q} - \underline{K}^*\underline{B}\,\underline{R}^{-1}\underline{B}'\underline{K}^* \qquad (4.3.28)$$

Proof: Since the system (4.2.1) is controllable, there exists a positive definite symmetric nxn matrix $\underline{K}_0$ such that

$$\underline{F}_0 = \underline{R}^{-1}\underline{B}'\underline{K}_0 \qquad (4.3.29)$$

stabilizes the system (4.2.1). Thus, by Theorem 4, p. 231, of reference [18], the equation

$$\underline{0} = \underline{K}_1[\underline{A}-\underline{B}\,\underline{F}_0\underline{C}] + [\underline{A}-\underline{B}\,\underline{F}_0\underline{C}]'\underline{K}_1 + \underline{Q} + \underline{K}_0\underline{B}\,\underline{R}^{-1}\underline{B}\,\underline{K}_0 \tag{4.3.29}$$

possesses the unique solution:

$$\underline{K}_1 = \int_0^{\infty} e^{[\underline{A}-\underline{B}\,\underline{F}_0]'t}[\underline{Q}+\underline{K}_0\underline{B}\,\underline{R}^{-1}\underline{B}'\underline{K}_0]e^{[\underline{A}-\underline{B}\,\underline{F}_0]t}dt \tag{4.3.30}$$

Furthermore, $\underline{K}_1$ is positive definite by the assumption of observability.

We complete the proof of the theorem, in essence, by the following lemma.

<u>Lemma 4.1</u>

$$\text{Let} \quad \underline{F}_n = \underline{R}^{-1}\underline{B}'\underline{K}_n \tag{4.3.31}$$

and let $\underline{K}_{n+1}$ be the unique positive definite solution matrix of

$$\underline{0} = \underline{K}_{n+1}[\underline{A}-\underline{B}\,\underline{F}_n\underline{C}] + [\underline{A}-\underline{B}\,\underline{F}_n\underline{C}]'\underline{K}_{n+1} + \underline{Q} + \underline{K}_n\underline{B}\,\underline{R}^{-1}\underline{B}'\underline{K}_n \tag{4.3.32}$$

then,

a) $\underline{K}_{n+1}$ exists provided that $[\underline{A}-\underline{B}\,\underline{F}_0]$ is a stable matrix

b) $\underline{K}_{n+1} < \underline{K}_n$

<u>Proof</u>:

a) The proof is by induction. Assume that the matrix $[\underline{A}-\underline{B}\,\underline{F}_{n-1}]$ is stable. Then $\underline{K}_n$ exists, is unique and positive definite.

Define the Liapunov function

$$v_n(t) = \underline{x}'(t)\,\underline{K}_n\,\underline{x}(t) > \underline{0} \quad \text{for all } t \tag{4.3.33}$$

$$\dot{v}_n(t) = \underline{x}'(t)\,[(\underline{A}-\underline{B}\,\underline{F}_n)'\underline{K}_n + \underline{K}_n(\underline{A}-\underline{B}\,\underline{F}_n)]\,\underline{x}(t) \tag{4.3.34}$$

Let

$$\underline{V}_n \triangleq [\underline{A}-\underline{B}\,\underline{F}_n]'\underline{K}_n + \underline{K}_n[\underline{A}-\underline{B}\,\underline{F}_n] \qquad (4.3.35)$$

$$\underline{V}_n=[\underline{A}-\underline{B}\,\underline{F}_n]'\underline{K}_n+\underline{K}_n[\underline{A}-\underline{B}\,\underline{F}_n]+\underline{F}'_{n-1}\underline{B}'\underline{K}_n-\underline{F}'_{n-1}\underline{B}'\underline{K}_n+\underline{K}_n\underline{B}\,\underline{F}_{n-1}-\underline{K}_n\underline{B}\,\underline{F}_{n-1} \qquad (4.3.36)$$

$$\underline{V}_n=[\underline{A}-\underline{B}\,\underline{F}_{n-1}]'\underline{K}_n+\underline{K}_n[\underline{A}-\underline{B}\,\underline{F}_{n-1}]+\underline{F}'_{n-1}\underline{B}'\underline{K}_n+\underline{K}_n\underline{B}\,\underline{F}_{n-1}-2\underline{K}_n\underline{B}\,\underline{R}^{-1}\underline{B}'\underline{K}_n \qquad (4.3.37)$$

Using Eq. (4.3.32) and Eq. (4.3.31)

$$\underline{V}_n=-\underline{Q}-\underline{K}_{n-1}\underline{B}\,\underline{R}^{-1}\underline{B}'\underline{K}_{n-1}+\underline{K}_{n-1}\underline{B}\,\underline{R}^{-1}\underline{B}'\underline{K}_n+\underline{K}_n\underline{B}\,\underline{R}^{-1}\underline{B}'\underline{K}_{n-1}-2\underline{K}_n\underline{B}\,\underline{R}^{-1}\underline{B}'\underline{K}_n \qquad (4.3.38)$$

$$\underline{V}_n = -\underline{Q} - \underline{K}_n\underline{B}\,\underline{R}^{-1}\underline{B}'\underline{K}_n - (\underline{K}_{n-1}-\underline{K}_n)\underline{B}\,\underline{R}^{-1}\underline{B}'(\underline{K}_{n-1}-\underline{K}_n) \qquad (4.3.39)$$

Therefore $\underline{V}_n < \underline{0}$ and $\dot{v}_n(t) < 0$ for all t $\qquad$ (4.3.40)

Therefore, $[\underline{A}-\underline{B}\,\underline{F}_{n-1}]$ is stable implies $[\underline{A}-\underline{B}\,\underline{F}_n]$ is stable so that $\underline{K}_{n+1}$ exists and (a) is proven by induction from $\underline{F}_0$.

b) The proof that $\underline{K}_{n+1} < \underline{K}_n$ is obtained as follows:

$$\underline{0} = \underline{K}_{n+1}[\underline{A}-\underline{B}\,\underline{R}^{-1}\underline{B}'\underline{K}_n]+[\underline{A}-\underline{B}\,\underline{R}^{-1}\underline{B}'\underline{K}_n]'\underline{K}_{n+1}+\underline{K}_n\underline{B}\,\underline{R}^{-1}\underline{B}'\underline{K}_n$$
$$-\underline{K}_n[\underline{A}-\underline{B}\,\underline{R}^{-1}\underline{B}'\underline{K}_{n-1}]-[\underline{A}-\underline{B}\,\underline{R}^{-1}\underline{B}'\underline{K}_{n-1}]'\underline{K}_n-\underline{K}_{n-1}\underline{B}\,\underline{R}^{-1}\underline{B}'\underline{K}_{n-1} \qquad (4.3.41)$$

or,

$$\underline{0} = [\underline{K}_{n+1}-\underline{K}_n][\underline{A}-\underline{B}\,\underline{R}^{-1}\underline{B}'\underline{K}_n]+[\underline{A}-\underline{B}\,\underline{R}^{-1}\underline{B}'\underline{K}_n]'[\underline{K}_{n+1}-\underline{K}_n]$$
$$-[\underline{K}_n-\underline{K}_{n-1}]\underline{B}\,\underline{R}^{-1}\underline{B}'[\underline{K}_n-\underline{K}_{n-1}] \qquad (4.3.42)$$

This equation has a unique negative definite solution for $[\underline{K}_{n+1} - \underline{K}_n]$. Thus, $\underline{K}_{n+1} < \underline{K}_n$.

By the lemma just proven, the sequence of matrices $\{\underline{K}_n\}$ is a monotone decreasing sequence of positive definite matrices. Such a sequence must converge to a positive definite limit $\underline{K}^*$. This $\underline{K}^*$ must be a solution of Eq. (4.3.28) or, equivalently, of Eqs. (4.3.25) and (4.3.26). Uniqueness follows from a straightforward algebraic manipulation which may be found on p. 77, reference [11].

This completes the proof of the theorem. And, the theorem guarantees the existence of the integrals (4.3.20) and (4.3.21), thereby completing the derivation of the solution to the Kalman linear regulator problem.

4.4 The Main Result

In this section we relax the assumption that $\underline{C} = \underline{I}$ to the assumption that $\underline{C}$ is a real $r \times n$ constant matrix of rank r $(r \leq n)$. Thus, the results we obtain will apply exactly to the problem stated in Section 4.2. The results we actually obtain are somewhat more complicated than those of the previous section. However, the derivations and the structure of the solutions are quite similar. We remark that the existence of a constant $\underline{F}$ which stabilizes the system (4.2.1) is assumed throughout this section. If such an $\underline{F}$ does not exist, then $\hat{J}$ is infinite for all allowable controls and our problem is meaningless.

We begin by using Eq. (4.2.2) to rewrite the performance criterion, Eq. (4.2.3), as

$$\hat{J}(\underline{F}) = \frac{1}{2}\operatorname{tr}\left[\int_0^\infty e^{[\underline{A}-\underline{B}\,\underline{F}\,\underline{C}]'t}(\underline{Q}+\underline{C}'\underline{F}'\underline{R}\,\underline{F}\,\underline{C})e^{[\underline{A}-\underline{B}\,\underline{F}\,\underline{C}]t}dt\right] \tag{4.4.1}$$

Guided by our experience in the previous section, we will again calculate $\frac{\partial \hat{J}}{\partial \underline{F}}$ by application of the lemma and example of the previous section. Again, we define the convenience,

$$\underline{A}^{o} \overset{\Delta}{=} [\underline{A}-\underline{B}\,\underline{F}\,\underline{C}] \tag{4.4.2}$$

Then,

$$\hat{J}(\underline{F}+\epsilon\underline{\Delta F})=\frac{1}{2}\mathrm{tr}\left\{\int_{0}^{\infty} e^{[\underline{A}^{o}-\epsilon\underline{B}\,\underline{\Delta F}\,\underline{C}]'t}[\underline{Q}+\underline{C}'(\underline{F}'+\epsilon\underline{\Delta F}')\underline{R}(\underline{F}+\epsilon\underline{\Delta F})\underline{C}]\,e^{[\underline{A}^{o}-\epsilon\underline{B}\,\underline{\Delta F}\underline{C}]t}\,dt\right\} \tag{4.4.3}$$

Using the example, it is easy to show that, to first order in ϵ,

$$e^{[\underline{A}^{o}-\epsilon\underline{B}\,\underline{\Delta F}\underline{C}]t}=e^{\underline{A}^{o}t}-\epsilon\int_{0}^{t} e^{\underline{A}^{o}(t-\sigma)}\,\underline{B}\,\underline{\Delta F}\,\underline{C}\,e^{\underline{A}^{o}\sigma}\,d\sigma \tag{4.4.4}$$

Applying Eq. (4.4.4) to Eq. (4.4.3) we obtain, accurate to first order in ϵ,

$$\begin{aligned}\hat{J}(\underline{F}+\epsilon\underline{\Delta F})=\frac{1}{2}\mathrm{tr}\int_{0}^{\infty}\Bigg\{ & e^{\underline{A}^{o'}t}(\underline{Q}+\underline{C}'\underline{F}'\underline{R}\,\underline{F}\,\underline{C})e^{\underline{A}^{c}t}+2\epsilon\, e^{\underline{A}^{o'}t}(\underline{C}'\underline{F}'\underline{R}\,\underline{\Delta F}\,\underline{C})e^{\underline{A}^{o}t} \\ & -\epsilon\left(\int_{0}^{t} e^{\underline{A}^{o'}(t-\sigma)}\,\underline{C}'\underline{\Delta F}'\underline{B}'e^{\underline{A}^{o'}\sigma}\,d\sigma\right)(\underline{Q}+\underline{C}'\underline{F}'\underline{R}\,\underline{F}\,\underline{C})e^{\underline{A}^{o}t} \\ & -\epsilon\, e^{\underline{A}^{o'}t}(\underline{Q}+\underline{C}'\underline{F}'\underline{R}\,\underline{F}\,\underline{C})\left(\int_{0}^{t} e^{\underline{A}^{o}(t-\sigma)}\,\underline{B}\,\underline{\Delta F}\underline{C}\,e^{\underline{A}^{o}\sigma}\,d\sigma\right)\Bigg\}\,dt\end{aligned} \tag{4.4.5}$$

Then, to first order in ϵ,

$$\hat{J}(\underline{F}+\underline{\Delta F})-\hat{J}(\underline{F})=\frac{1}{2}\int_0^\infty \epsilon \;\mathrm{tr}\Big[2\underline{C}\,e^{\underline{A}^o t}\,e^{\underline{A}^{o'} t}\,\underline{C}'\underline{F}'\underline{R}\,\underline{\Delta F}$$

$$-\int_0^t \underline{C}\,e^{\underline{A}^o\sigma}\,e^{\underline{A}^{o'}t}\,(\underline{Q}+\underline{C}'\underline{F}'\underline{R}\,\underline{F}\,\underline{C})e^{\underline{A}^o(t-\sigma)}\,\underline{B}\,d\sigma\,\underline{\Delta F}$$

$$-\int_0^t \underline{C}\,e^{\underline{A}^o(t-\sigma)}\,e^{\underline{A}^{o'}t}\,(\underline{Q}+\underline{C}'\underline{F}'\underline{R}\,\underline{F}\,\underline{C})e^{\underline{A}^o\sigma}\,\underline{B}\,d\sigma\,\underline{\Delta F}\Big]dt \qquad (4.4.6)$$

Using Kleinman's lemma,

$$\frac{\partial\hat{J}}{\partial\underline{F}}=\int_0^\infty \underline{R}\,\underline{F}\,\underline{C}e^{\underline{A}^o t}\,e^{\underline{A}^{o'}t}\,\underline{C}'dt-\frac{1}{2}\int_0^\infty\int_0^t \underline{B}'e^{\underline{A}^{o'}(t-\sigma)}(\underline{Q}+\underline{C}'\underline{F}'\underline{R}\,\underline{F}\,\underline{C})e^{\underline{A}^o t}\,e^{\underline{A}^{o'}\sigma}\,\underline{C}'d\sigma\,dt$$

$$-\frac{1}{2}\int_0^\infty\int_0^t \underline{B}'e^{\underline{A}^{o'}\sigma}(\underline{Q}+\underline{C}'\underline{F}'\underline{R}\,\underline{F}\,\underline{C})e^{\underline{A}^o t}\,e^{\underline{A}^{o'}(t-\sigma)}\,\underline{C}'d\sigma\,dt \qquad (4.4.7)$$

This is nicely parallel to Eq. (4.3.16) and it is obvious that the identical substitutions will produce similar results. Thus,

$$\text{Let } \underline{\Gamma} \triangleq \int_0^\infty\int_0^t \underline{B}'e^{\underline{A}^{o'}\sigma}(\underline{Q}+\underline{C}'\underline{F}'\underline{R}\,\underline{F}\,\underline{C})e^{\underline{A}^o t}\,e^{\underline{A}^{o'}(t-\sigma)}\,\underline{C}'d\sigma\,dt \qquad (4.4.8)$$

Let $\sigma_1 = t-\sigma$

Then, $d\sigma = -d\sigma_1$ (t constant)

Note that at $\sigma = 0$, $\sigma_1 = t$

at $\sigma = t$, $\sigma_1 = 0$

With the above substitutions,

$$\underline{\Gamma} = \int_0^\infty \int_0^t \underline{B}' e^{\underline{A}^{o'}(t-\sigma_1)} (\underline{Q}+\underline{C}'\underline{F}'\underline{R}\,\underline{F}\,\underline{C}) e^{\underline{A}^{o}t} e^{\underline{A}^{o'}\sigma_1} d\sigma_1\, dt \qquad (4.4.9)$$

Therefore,

$$\frac{\partial \hat{J}}{\partial \underline{F}} = \int_0^\infty \underline{R}\,\underline{F}\,\underline{C}\, e^{\underline{A}^{o}t} e^{\underline{A}^{o'}t} \underline{C}'dt - \int_0^\infty \int_0^t \underline{B}' e^{\underline{A}^{o'}(t-\sigma)} (\underline{Q}+\underline{C}'\underline{F}'\underline{R}\,\underline{F}\,\underline{C}) e^{\underline{A}^{o}t} e^{\underline{A}^{o'}\sigma} \underline{C}'d\sigma\, dt$$

$$(4.4.10)$$

Let

$$\underline{\chi} \triangleq \int_0^\infty \int_0^t \underline{B}' e^{\underline{A}^{o'}(t-\sigma)} (\underline{Q}+\underline{C}'\underline{F}'\underline{R}\,\underline{F}\,\underline{C}) e^{\underline{A}^{o}t} e^{\underline{A}^{o'}\sigma} \underline{C}'d\sigma\, dt \qquad (4.4.11)$$

Interchanging the order of integration, assuming the required integrals exist,

$$\underline{\chi} = \int_0^\infty \int_\sigma^\infty \underline{B}' e^{\underline{A}^{o'}(t-\sigma)} (\underline{Q}+\underline{C}'\underline{F}'\underline{R}\,\underline{F}\,\underline{C}) e^{\underline{A}^{o}t} e^{\underline{A}^{o'}\sigma} \underline{C}'dt\, d\sigma \qquad (4.4.12)$$

Let $\tau = t-\sigma$

Then, $d\tau = dt$

Note that at $t = \sigma$, $\tau = 0$

at $t = \infty$, $\tau = \infty$

With the above substitutions,

$$\underline{\chi} = \int_0^\infty \underline{B}' e^{\underline{A}^{o'}\tau} (\underline{Q}+\underline{C}'\underline{F}'\underline{R}\,\underline{F}\,\underline{C}) e^{\underline{A}^{o}\tau} d\tau \int_0^\infty e^{\underline{A}^{o}\sigma} e^{\underline{A}^{o'}\sigma} \underline{C}'d\sigma \qquad (4.4.13)$$

Substituting Eq. (4.4.13) into Eq. (4.4.10) and setting $\left.\frac{\partial \hat{J}}{\partial \underline{F}}\right|_{\underline{F}^*} = 0$,

we obtain,

$$\underline{F}^* = \underline{R}^{-1}\underline{B}'\int_0^\infty e^{\underline{A}^{*'}\tau}(\underline{Q}+\underline{C}'\underline{F}^{*'}\underline{R}\underline{F}^*\underline{C})e^{\underline{A}^*\tau}d\tau \int_0^\infty e^{\underline{A}^*\sigma}e^{\underline{A}^{*'}\sigma}d\sigma \cdot \underline{C}'\left[\int_0^\infty \underline{C}e^{\underline{A}^*t}e^{\underline{A}^{*'}t}\underline{C}'dt\right] \tag{4.4.14}$$

where

$$\underline{A}^* = [\underline{A}-\underline{B}\underline{F}^*\underline{C}] \tag{4.4.15}$$

For some applications the form of Eq. (4.4.14) may be the most useful. However, we can obtain another form that is quite interesting by defining:

$$\underline{K}^* \triangleq \int_0^\infty e^{\underline{A}^{*'}\tau}[\underline{Q}+\underline{C}'\underline{F}^{*'}\underline{R}\underline{F}^*\underline{C}]e^{\underline{A}^*\tau}d\tau \tag{4.4.16}$$

$$\underline{L}^* \triangleq \int_0^\infty e^{\underline{A}^*\sigma}e^{\underline{A}^{*'}\sigma}d\sigma \tag{4.4.17}$$

Assuming that a $\underline{K}^*$, $\underline{L}^*$ and $\underline{F}^*$ exist such that $\underline{A}^*$, as defined in Eq. (4.4.15), is stable; assuming $\underline{K}^*$ and $\underline{L}^*$ are solutions of Eqs. (4.4.14), (4.4.16) and (4.4.17); then $\underline{K}^*$, $\underline{L}^*$ and $\underline{F}^*$ are also solutions of the following algebraic equations:

$$\underline{0} = \underline{K}^*[\underline{A}-\underline{B}\underline{F}^*\underline{C}] + [\underline{A}-\underline{B}\underline{F}^*\underline{C}]'\underline{K}^* + \underline{Q} + \underline{C}'\underline{F}^{*'}\underline{R}\underline{F}^*\underline{C} \tag{4.4.18}$$

$$\underline{0} = \underline{L}^*[\underline{A}-\underline{B}\underline{F}^*\underline{C}]' + [\underline{A}-\underline{B}\underline{F}^*\underline{C}]\underline{L}^* + \underline{I} \tag{4.4.19}$$

$$\underline{F}^* = \underline{R}^{-1}\underline{B}'\underline{K}^*\underline{L}^*\underline{C}'[\underline{C}\underline{L}^*\underline{C}']^{-1} \tag{4.4.20}$$

Note that Eq. (4.4.20) can be used to eliminate $\underline{F}^*$ from the other two equations. Furthermore, the existence of $\underline{C}^{-1}$ reduces the above equations to the single equation of the previous section, Eq. (4.3.28).

We remark that it is entirely possible that the above algebraic equations have solutions that are not also solutions of the integral equations, Eqs. (4.4.14), (4.4.16) and (4.4.17). Furthermore, these are only necessary conditions for a solution. These two caveats will be clarified somewhat by the following lemma which also provides an algorithm for computing $\underline{F}^*$.

Lemma 4.2

Let $$\underline{F}_{n-1} = \underline{R}^{-1}\underline{B}'\underline{K}_{n-1}\underline{L}_{n-1}\underline{C}'[\underline{C}\,\underline{L}_{n-1}\underline{C}']^{-1} \qquad (4.4.21)$$

where $\underline{K}_n$ is the solution of:

$$\underline{0} = \underline{K}_n[\underline{A}-\underline{B}\,\underline{F}_{n-1}\underline{C}] + [\underline{A}-\underline{B}\,\underline{F}_{n-1}\underline{C}]'\underline{K}_n + \underline{Q} + \underline{C}'\underline{F}'_{n-1}\underline{R}\,\underline{F}_{n-1}\underline{C} \qquad (4.4.22)$$

and $\underline{L}_{n-1}$ is the solution of:

$$\underline{0} = \underline{L}_{n-1}[\underline{A}-\underline{B}\,\underline{F}_{n-1}\underline{C}]' + [\underline{A}-\underline{B}\,\underline{F}_{n-1}\underline{C}]\underline{L}_{n-1} + \underline{I} \qquad (4.4.23)$$

a) Then, assuming $\underline{Q} > \underline{0}$ and $[\underline{A}-\underline{B}\,\underline{F}_{n-1}\underline{C}]$ stable, a unique and positive definite $\underline{K}_n$ exists.

b) Furthermore, assuming there exists a positive definite $\underline{L}_{n-1}$ which satisfies Eq. (4.4.23), then

$$\mathrm{tr}[\underline{K}_n] \leq \mathrm{tr}[\underline{K}_{n-1}] \qquad (4.4.24)$$

Proof

a) Existence and uniqueness of $\underline{K}_n$ is a direct consequence of Theorem 4, p. 231, of reference [18]. Positive definiteness was established for an identical equation, Eqs. (4.3.29) and (4.3.30), in the previous section.

b) Let $\underline{K}_n - \underline{K}_{n-1} \overset{\Delta}{=} \delta\underline{K}_n$ (4.4.25)

We next attempt to compute $\delta\underline{K}_n$.

$$0 = \underline{K}_n[\underline{A}-\underline{B}\,\underline{F}_{n-1}\underline{C}]+[\underline{A}-\underline{B}\,\underline{F}_{n-1}\underline{C}]'\underline{K}_n-\underline{C}'\underline{F}'_{n-1}\underline{R}\,\underline{F}_{n-1}\underline{C}+\underline{C}'\underline{F}'_{n-2}\underline{R}\,\underline{F}_{n-2}\underline{C}$$
$$-\underline{K}_{n-1}[\underline{A}-\underline{B}\,\underline{F}_{n-2}\underline{C}]-[\underline{A}-\underline{B}\,\underline{F}_{n-2}\underline{C}]'\underline{K}_{n-1} \qquad (4.4.26)$$

$$0 = \delta\underline{K}_n[\underline{A}-\underline{B}\,\underline{F}_{n-1}\underline{C}]+[\underline{A}-\underline{B}\,\underline{F}_{n-1}\underline{C}]'\delta\underline{K}_n-\underline{C}'\underline{F}'_{n-1}\underline{R}\,\underline{F}_{n-1}\underline{C}+\underline{C}'\underline{F}_{n-2}\underline{R}\,\underline{F}_{n-2}\underline{C}$$
$$+\underline{K}_{n-1}\underline{B}\,\underline{F}_{n-2}\underline{C}-\underline{K}_{n-1}\underline{B}\,\underline{F}_{n-1}\underline{C}+\underline{C}'\underline{F}'_{n-2}\underline{B}'\underline{K}_{n-1}-\underline{C}'\underline{F}'_{n-1}\underline{B}'\underline{K}_{n-1}$$
$$(4.4.27)$$

Adding and subtracting $\underline{K}_{n-1}\underline{B}\,\underline{R}^{-1}\underline{B}'\underline{K}_{n-1}$ and forming perfect squares in exactly the same manner as in the proof of Lemma 3.1, we obtain

$$0=\delta\underline{K}_n[\underline{A}-\underline{B}\underline{F}_{n-1}\underline{C}]+[\underline{A}-\underline{B}\,\underline{F}_{n-1}\underline{C}]'\delta\underline{K}_n+[\underline{C}'\underline{F}'_{n-2}-\underline{K}_{n-2}\underline{B}\,\underline{R}^{-1}]\underline{R}[\underline{F}_{n-2}\underline{C}-\underline{R}^{-1}\underline{B}'\underline{K}_{n-2}]$$
$$-[\underline{C}'\underline{F}'_{n-1}-\underline{K}_{n-2}\underline{B}\,\underline{R}^{-1}]\underline{R}[\underline{F}_{n-1}\underline{C}-\underline{R}^{-1}\underline{B}'\underline{K}_{n-2}] \qquad (4.4.28)$$

Define

$$\underline{I}_n \overset{\Delta}{=} [\underline{C}'\underline{F}'_{n-2}-\underline{K}_{n-2}\underline{B}\,\underline{R}^{-1}]\underline{R}[\underline{F}_{n-2}\underline{C}-\underline{R}^{-1}\underline{B}'\underline{K}_{n-2}]$$
$$-[\underline{C}'\underline{F}'_{n-1}-\underline{K}_{n-2}\underline{B}\,\underline{R}^{-1}]\underline{R}[\underline{F}_{n-1}\underline{C}-\underline{R}^{-1}\underline{B}'\underline{K}_{n-2}] \qquad (4.4.29)$$

Then,

$$\delta\underline{K}_n = \int_0^{\infty} e^{[\underline{A}-\underline{B}\,\underline{F}_{n-1}\underline{C}]'t}\,\underline{I}_n\,e^{[\underline{A}-\underline{B}\,\underline{F}_{n-1}\underline{C}]t}dt \qquad (4.4.30)$$

$$\mathrm{tr}[\delta \underline{K}_n] = \int_0^\infty \mathrm{tr}\left[e^{[\underline{A}-\underline{B}\,\underline{F}_{n-1}\underline{C}]'t}\, \underline{I}_n\, e^{[\underline{A}-\underline{B}\,\underline{F}_{n-1}\underline{C}]t} \right] dt \qquad (4.4.31)$$

$$= \mathrm{tr}\left[\underline{I}_n \int_0^\infty e^{[\underline{A}-\underline{B}\,\underline{F}_{n-1}\underline{C}]t}\, e^{[\underline{A}-\underline{B}\,\underline{F}_{n-1}\underline{C}]'t} dt \right] \qquad (4.4.32)$$

$$\mathrm{tr}[\delta \underline{K}_n] = \mathrm{tr}[\underline{I}_n \underline{L}_{n-1}] \qquad (4.4.33)$$

Since $\underline{L}_{n-1}$ is assumed to be positive definite, it can be factored uniquely into

$$\underline{\Phi}_{n-1}\underline{\Phi}_{n-1}' = \underline{L}_{n-1} \qquad (4.4.34)$$

Therefore,

$$\mathrm{tr}[\delta \underline{K}_n] = \mathrm{tr}[\underline{\Phi}_{n-1}' \underline{I}_n \underline{\Phi}_{n-1}] \qquad (4.4.35)$$

Equation (4.4.35) is identical in form to Eq. (3.1.17). Thus, application of the proof in Chapter III which follows Eq. (3.1.17) demonstrates that

$$\mathrm{tr}[\delta \underline{K}_n] < 0 \qquad (4.4.36)$$

which completes the proof.

If one can find a stabilizing initial guess for the feedback gain matrix then the above lemma can be used in a computer algorithm that is essentially similar to the one used in Chapter III. Again, convergence is not guaranteed but is likely for well-behaved systems. It is conjectured that the solutions of the necessary conditions are <u>not</u> unique. Furthermore, it is conjectured that convergence will not occur unless the initial guess is "close enough" to optimality. This conjecture is based on two facts:

1) The algorithm is basically Newton's method and this type of behavior is characteristic of Newton's method.

2) The Lyapunov argument of the previous section, when it is applied to this problem, shows that stability of $[\underline{A}-\underline{B}\,\underline{F}_n\underline{C}]$ and existence of $\underline{F}_{n+1}$ does not guarantee stability of $[\underline{A}-\underline{B}\,\underline{F}_{n+1}\underline{C}]$ unless $||\underline{F}_{n+1}-\underline{F}_n||$ is "small enough".

4.5 Examples

We have worked two examples that are of some theoretical interest. They are included below.

Example 1:

a) The system is described by

$$\ddot{x} + f\dot{x} + x = 0 \tag{4.5.1}$$

where f is the feedback gain and x is a scalar function of time. This system is identical to:

$$\dot{\underline{\Phi}}(t,0) = [\underline{A}-\underline{B}f\underline{C}]\underline{\Phi}(t,0) \qquad \underline{\Phi}(0,0) = \underline{I} \tag{4.5.2}$$

$$\text{with} \quad \underline{A} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \quad \underline{B} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad \underline{C} = [0 \quad 1]$$

The system is controllable and observable. The performance criterion is given by

$$\hat{J} = \frac{1}{2}\int_0^{\infty} \operatorname{tr}\left\{\underline{\Phi}'(t,0)\,[\underline{Q}+\underline{C}'\underline{f}^2\underline{C}]\,\underline{\Phi}(t,0)\right\} dt \tag{4.5.3}$$

$$\text{with} \quad \underline{Q} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$$

The solution, as the reader can verify by substituting into Eqs. (4.4.18)-(4.4.20) is:

$$f^* = \sqrt{2/3} \tag{4.5.4}$$

minimizes the performance criterion (4.5.3) constrained by Eq. (4.5.2).

b) It can be shown, by direct substitution, that

$$\hat{J} = \frac{1}{2}\int_0^\infty (x^2+f^2\dot{x}^2)dt\Bigg|_{\substack{x(0)=0\\ \dot{x}(0)=1}} + \frac{1}{2}\int_0^\infty (x^2+f^2\dot{x}^2)dt\Bigg|_{\substack{x(0)=1\\ \dot{x}(0)=0}} \tag{4.5.5}$$

The f which minimizes Eq. (4.5.5) subject to the constraint imposed by Eq. (4.5.1) can be computed by a procedure suggested by Brockett. We include the calculations to demonstrate the method.

Multiplying Eq. (4.5.1) by $\dot{x}$, we obtain

$$\ddot{x}\dot{x} + f\dot{x}^2 + x\dot{x} = 0 \tag{4.5.6}$$

Integrating from 0 to ∞, we obtain

$$\int_0^\infty (\dot{x}\ddot{x} + f\dot{x}^2 + x\dot{x})dt = 0 \tag{4.5.7}$$

Therefore,

$$\int_0^\infty \dot{x}^2 dt = -\frac{1}{f}\left[(x^2+\dot{x}^2)\Big|_0^\infty\right] \tag{4.5.8}$$

Multiplying Eq. (4.5.1) by $(\dot{x}+fx)$, we obtain,

$$(\dot{x} + f x)(\ddot{x} + f\dot{x}) + (\dot{x} + f x)x = 0 \qquad (4.5.9)$$

Integrating from 0 to ∞, we see that

$$\int_0^\infty (\dot{x} + f x)(\ddot{x} + f\dot{x})dt + \int_0^\infty x\dot{x}\,dt + \int_0^\infty f x^2 dt = 0 \qquad (4.5.10)$$

Thus,

$$\int_0^\infty x^2 dt = -\frac{1}{f}\left[\dot{x} + f x)^2 + x^2\right]_0^\infty \qquad (4.5.11)$$

Therefore,

$$\frac{1}{2}\int_0^\infty (x^2 + f^2\dot{x}^2)dt = -\frac{1}{2f}\left[\dot{x}+f x)^2+x^2+f^2x^2+f^2\dot{x}^2\right]_0^\infty \qquad (4.5.12)$$

We assume (as is the case) that the minimizing f produces a stable system so that $x(\infty) = \dot{x}(\infty) = 0$.

Thus, if $x(0) = 1$, $\dot{x}(0) = 0$ then

$$\frac{1}{2}\int_0^\infty (x^2 + f^2\dot{x}^2)dt = \frac{1}{2f}(1 + f^2) \qquad (4.5.13)$$

And, if $x(0) = 0$, $\dot{x}(0) = 1$ then

$$\frac{1}{2}\int_0^\infty (x^2 + f^2\dot{x}^2)dt = \frac{1}{2f}(1 + 2f^2) \qquad (4.5.14)$$

Substituting these results into Eq. (4.5.15), we obtain a specific function for $\hat{J}$ that is,

$$\hat{J}(f) = \frac{2 + 3f^2}{2f} \tag{4.5.15}$$

Differentiating this expression, setting the derivative equal to zero and recognizing that we must choose that f for which the resulting system is stable gives

$$f^* = \sqrt{2/3} \tag{4.5.16}$$

There is a third technique which could be used to solve this example. One could compute the Laplace transforms of x and $\dot{x}$ for the two sets of initial data. Then, Parseval's theorem can be used to obtain an expression for $\hat{J}$ in terms of an integral of these Laplace transforms. The integral tables in Appendix F of Newton, Gould and Kaiser[22] can be used to evaluate this integral directly. One thus arrives at Eq. (4.5.15) and proceeds from there.

Example 2:

This example is identical to Example 4 of Chapter III except that $T = \infty$ and $\underline{F}^*$ is constant by hypothesis. The parameters are:

$$\underline{A} = \begin{bmatrix} 0 & 1 \\ -10 & -2 \end{bmatrix} \qquad \underline{B} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \qquad \underline{C} = [0 \quad 1] \qquad \underline{Q} = 10\underline{I}, \qquad \underline{R} = 1$$

The solution, obtained by substituting into Eqs. (4.4.18)-(4.4.20) and solving is:

$$f^* = 1.7 \tag{4.5.17}$$

Two reasonable conjectures about the relation between the constant $\underline{F}^*$ of this chapter and the time-varying $\underline{F}^*(t)$ of Chapter III, when they are calculated for identical systems and for cost criteria whose only difference is in whether T is infinite or not, are:

1) $\underline{F}^*$ = the "average" value of $\underline{F}^*(t)$ computed for T large.

2) $\underline{F}^*$ = the "steady-state" value of $\underline{F}^*(t)$ computed for T large. By "steady-state" we mean a constant value of $\underline{F}^*(t)$ maintained for a time interval between the two terminal transients, if such a constant value exists.

It should be noted that the above example and Example 4 in Chapter III support either hypothesis although the first conjecture is supported more strongly.

CHAPTER V

CONCLUSIONS

In the previous chapters we have studied two very closely related output feedback problems. For the first of these problems, the linear output feedback control of a linear system with respect to a quadratic criterion for a finite interval, we found conditions which the optimal control must satisfy. In addition, we derived and programmed a computer algorithm which can be used to compute this optimal control. The second problem, treated in Chapter IV, is identical to the first except that the system is assumed time-invariant as well as linear, the interval is semi-infinite and we demand that the feedback matrix be time-invariant. Necessary conditions that the solution to this problem must satisfy are found. In addition, a number of examples of both types are included.

We believe that these results are quite interesting, both theoretically and practically. From a practical viewpoint, one can use these results for two purposes:

1) To design linear feedback controls, especially when the state vector has many more components than the output vector.

2) To study the cost-effectiveness of changing the measurements in a linear system. In other words, one can solve the problems discussed in this thesis for several different candidates for $\underline{C}$, compare the cost of buying each $\underline{C}$ with the performance obtained by it, and choose the best one.

Both of these applications have been illustrated in the examples included in the previous chapters.

From a theoretical viewpoint, we believe these results represent a contribution to quadratic optimization problems for linear systems. In addition, these results will help span the gap between classical control theory and modern control theory.

We believe that there are many potentially useful extensions of this research. For example, in the classical design of feedback controls it is well known that dynamical compensation is often useful. Thus, it would probably be useful to extend our results so that they might be used to calculate "optimal" compensators. Another interesting question is how does additive noise in the output vector $\underline{y}(t)$ affect our results. We have briefly studied still another possible extension of these results in Chapter IV. That is the inverse problem; When is a linear system optimal with respect to the performance criteria used in this thesis?

APPENDIX A

ON THE PSEUDO-INVERSE OF A MATRIX[19]

The purpose of this appendix is to develop those properties of the pseudo-inverse of a matrix that are relevant to our research. Since our concern is with matrices we restrict ourselves to the consideration of linear transformations (matrices) mapping a finite dimensional vector space into a finite dimensional vector space. All of these vector spaces are defined on the complex field $\mathcal{C}$ although all our results are equally true for vector spaces on the real field. We have closely followed reference 19, Zadeh and Desoer, in this appendix.

With the above comments in mind, we make the following definitions :

Let $\mathcal{X}$ be an m-dimensional linear vector space defined on $\mathcal{C}$

$\mathcal{Y}$ be an n-dimensional linear vector space defined on $\mathcal{C}$

$\underline{A}$ be an arbitrary $n \times m$ matrix of complex numbers

<u>Definition A.1</u> - The <u>range of a matrix</u> $\underline{A}$ is the set $\mathcal{R}(A)$ defined by:

$$\mathcal{R}(\underline{A}) = \{\underline{y} \in \mathcal{Y} \mid \underline{y} = \underline{A}\,\underline{x} \quad \text{for some} \quad \underline{x} \in \mathcal{X}\} \tag{A.1}$$

<u>Definition A.2</u> - The <u>null space of a matrix</u> $\underline{A}$ is the set $\mathcal{N}(A)$ defined by:

$$\mathcal{N}(\underline{A}) = \{\underline{x} \in \mathcal{X} \mid \underline{A}\,\underline{x} = 0\} \tag{A.2}$$

That is, $\mathcal{N}(\underline{A})$ is the set of all vectors of $\mathcal{X}$ that $\underline{A}$ maps into the zero vector of $\mathcal{Y}$.

Definition A. 3 - A subspace $\mathcal{B}$ of a finite dimensional vector space $\mathcal{X}$ is a set of vectors of $\mathcal{C}^n$ such that if $\underline{x}$ and $\underline{y}$ are in $\mathcal{B}$, then for all complex numbers α and β, $\alpha\underline{x} + \beta\underline{y} \in \mathcal{B}$.

Definition A. 4 - Let $\mathcal{M}$ and $\mathcal{N}$ be two subspaces of a vector space $\mathcal{X}$. $\mathcal{X}$ is said to be the direct sum of $\mathcal{M}$ and $\mathcal{N}$, written $\mathcal{M} \oplus \mathcal{N} = \mathcal{X}$, if any $\underline{x} \in \mathcal{X}$ may be written in one and only one way as $\underline{x} = \underline{y} + \underline{z}$ where $\underline{y} \in \mathcal{M}$ and $\underline{z} \in \mathcal{N}$.

Definition A. 5 - Let $\underline{A}$ be an $n \times m$ matrix. The adjoint matrix $\underline{A}'$ of $\underline{A}$ is a matrix such that

$$< \underline{A}\underline{x}, \underline{y} > = < \underline{x}, \underline{A}'\underline{y} > \quad \text{for all} \quad \underline{x} \in \mathcal{C}^n, \ \underline{y} \in \mathcal{C}^m \tag{A. 3}$$

Definition A. 6 - Let $\mathcal{M}$ be a subspace of $\mathcal{C}^n$. The orthogonal complement of $\mathcal{M}$, denoted by $\mathcal{M}^{\perp}$, is the set of all vectors of $\mathcal{C}^n$ that are orthogonal to all vectors of $\mathcal{M}$.

Theorem A. 1 - Let $\underline{A}$ be a matrix in $\mathcal{C}^n$; then

$$\text{I)} \quad \mathcal{C}^n = \mathcal{R}(\underline{A}) \oplus \mathcal{N}(\underline{A}') \tag{A. 4}$$

II) $\mathcal{N}(\underline{A}')$ is the orthogonal complement of $\mathcal{R}(\underline{A})$ (A. 5)

$$\text{i. e.} \quad \mathcal{N}(\underline{A}') = \mathcal{R}(\underline{A})^{\perp} \tag{A. 6}$$

Proof:

Let $\underline{y} \in \mathcal{R}(\underline{A})^{\perp}$ (see Definition A. 6)

Since $\underline{A}(\underline{A}'\underline{y}) \in \mathcal{R}(\underline{A})$ and since $\underline{y} \in \mathcal{R}(\underline{A})^{\perp}$, we have

$0 = < \underline{y}, \underline{A}(\underline{A}'\underline{y}) > = < \underline{A}'\underline{y}, \underline{A}'\underline{y} > = || \underline{A}'\underline{y} ||^2 = 0$

Therefore, $\underline{A}'\underline{y} = 0$ and $\underline{y} \in \mathcal{N}(\underline{A}')$

Thus, we have proved that $\underline{y} \in \mathcal{R}(\underline{A})^{\perp} \Rightarrow \underline{y} \in \mathcal{N}(\underline{A}')$ (A. 7)

Let $\underline{z} \in \mathcal{N}(\underline{A}')$

then, for all $\underline{x}$, $0 = < \underline{A}'\underline{z}, \underline{x} > = < \underline{z}, \underline{A}\underline{x} >$

that is, $\underline{z}$ is orthogonal to all vectors in $\mathcal{R}(\underline{A})$.

Thus, we have proved $\underline{z} \in \mathcal{N}(\underline{A}') \Rightarrow \underline{z} \in \mathcal{R}(\underline{A})^{\perp}$ (A.8)

Therefore, (II) is proven and $\mathcal{N}(\underline{A}') = \mathcal{R}(\underline{A})^{\perp}$

(I) follows from the fact that $\mathcal{C}^{n} = \mathcal{R}(\underline{A}) \oplus \mathcal{R}(\underline{A})^{\perp}$

Definition A.7 - Let $\underline{A}$ be an $n \times m$ matrix mapping $\mathcal{C}^{m} \rightarrow \mathcal{C}^{n}$

The pseudo-inverse of $\underline{A}$ is denoted by $\underline{A}^{\dagger}$ and satisfies the conditions:

(I) $\underline{A}^{\dagger}\underline{A}\,\underline{x} = \underline{x}$ for all $\underline{x} \in \mathcal{N}(\underline{A})^{\perp} = \mathcal{R}(\underline{A}')$ (A.9)

(II) $\underline{A}^{\dagger}\underline{z} = \underline{0}$ for all $\underline{z} \in \mathcal{R}(\underline{A})^{\perp} = \mathcal{N}(\underline{A}')$ (A.10)

(III) $\underline{A}^{\dagger}(\underline{y}+\underline{z}) = \underline{A}^{\dagger}\underline{y} + \underline{A}^{\dagger}\underline{z}$ for all $\underline{y} \in \mathcal{R}(\underline{A})$, $\underline{z} \in \mathcal{R}(\underline{A})^{\perp}$ (A.11)

Corollary A.1 -

$$\mathcal{R}(\underline{A}^{\dagger}) = \mathcal{R}(\underline{A}') = \mathcal{N}(\underline{A})^{\perp} \quad (A.12)$$

$$\mathcal{N}(\underline{A}^{\dagger}) = \mathcal{N}(\underline{A}') = \mathcal{R}(\underline{A})^{\perp} \quad (A.13)$$

Theorem A.2 -

(I) $\underline{A}^{\dagger}\underline{A}$ is the orthogonal projection of $\mathcal{C}^{n}$ onto $\mathcal{R}(\underline{A}') = \mathcal{N}(\underline{A})^{\perp}$ (A.14)

(II) $(\underline{A}^{\dagger})^{\dagger} = \underline{A}$ (A.15)

(III) $\underline{A}^{\dagger}\underline{A}\,\underline{A}^{\dagger} = \underline{A}^{\dagger}$ (A.16)

(IV) $\underline{A}\,\underline{A}^{\dagger}\underline{A} = \underline{A}$ (A.17)

(V) $\underline{A}\,\underline{A}^{\dagger}$ is the orthogonal projection of $\mathcal{C}^{n}$ onto $\mathcal{R}(\underline{A}) = \mathcal{N}(\underline{A}')^{\perp}$ (A.18)

Proof:

of (I):

Let $\underline{x}$ be an arbitrary vector in $\mathcal{C}^{n}$

Consider the orthogonal decomposition $\underline{x} = \underline{x}_1 + \underline{x}_2$ where

$$\underline{x}_1 \in \mathcal{N}(\underline{A})^{\perp}, \quad \underline{x}_2 \in \mathcal{N}(\underline{A}) \quad (A.19)$$

Then, $\underline{A}^{\dagger}\underline{A}\,\underline{x} = \underline{A}^{\dagger}\underline{A}\,\underline{x}_1$ by the definition of $\underline{x}_2$ above (A.20)

But $\underline{A}^{\dagger}\underline{A}\,\underline{x}_1 = \underline{x}_1$ from Definition A.7, Part I (Eq. A.9) (A.21)

This proves (I)

of (II):

$$\left.\begin{aligned} \mathcal{R}([\underline{A}^{\dagger}]^{\dagger}) &= \mathcal{N}(\underline{A}^{\dagger})^{\perp} = \mathcal{R}(\underline{A}) && \text{(A.22)} \\ \mathcal{N}([\underline{A}^{\dagger}]^{\dagger}) &= \mathcal{R}(\underline{A}^{\dagger})^{\perp} = \mathcal{N}(\underline{A}) && \text{(A.23)} \end{aligned}\right\}\text{ by Corollary A.1}$$

Next, we verify that $\underline{A}$ satisfies the conditions I-III of Definition A.7 for $(\underline{A}^{\dagger})^{\dagger}$

a) Let $\underline{x} \in \mathcal{N}(\underline{A}^{\dagger})^{\perp} = \mathcal{R}(\underline{A})$ by (A.22) (A.24)

then $\underline{x} = \underline{A}\,\underline{y}$ for $\underline{y} \in \mathcal{N}(\underline{A})^{\perp}$ (A.25)

Therefore $\underline{A}\,\underline{A}^{\dagger}\underline{x} = \underline{A}\,\underline{A}^{\dagger}\underline{A}\,\underline{y} = \underline{A}\,\underline{y}$ from Eq. A.9 (A.26)

But $\underline{A}\,\underline{y} = \underline{x}$ and condition I is verified

b) Let $\underline{z} \in \mathcal{R}(\underline{A}^{\dagger})^{\perp} = \mathcal{N}(\underline{A})$

then $\underline{A}\,\underline{z} = \underline{0}$ by the line above, thus verifying Condition II (A.27)

c) Condition III is trivially satisfied by $\underline{A}$.

of (III):

Let $\underline{y} \in \mathcal{C}^n$ and $\underline{y} = \underline{y}_1 + \underline{y}_2$ with $\underline{y}_1 \in \mathcal{R}(\underline{A})$, $\underline{y}_2 \in \mathcal{R}(\underline{A})^{\perp}$ (A.28)

By Definition A.7, $\underline{A}^{\dagger}\underline{y} = \underline{A}^{\dagger}\underline{y}_1 \in \mathcal{N}(\underline{A})^{\perp}$ (A.29)

Therefore $\underline{A}^{\dagger}\underline{A}(\underline{A}^{\dagger}\underline{y}) = \underline{A}^{\dagger}\underline{A}\,\underline{A}^{\dagger}\underline{y}_1 = \underline{A}^{\dagger}\underline{y}_1 = \underline{A}^{\dagger}\underline{y}$ (A.30)

Thereby proving III

of (IV):

(II) and (III) imply (IV) trivially

of (V):

Let $\underline{y} \in \mathcal{C}^n$ and $\underline{y} = \underline{y}_1 + \underline{y}_2$ with $\underline{y}_1 \in \mathcal{R}(\underline{A})$, $\underline{y}_2 \in \mathcal{R}(\underline{A})^{\perp}$

Then $\underline{A}\,\underline{A}^{\dagger}\underline{y} = \underline{A}\,\underline{A}^{\dagger}\underline{y}_1 = \underline{y}_1$ (see Eqs. A.28-A.30) (A.31)

Thus, for any $\underline{y} \in \mathcal{C}^n$, $\underline{A}\,\underline{A}^{\dagger}\underline{y} = \underline{y}_1 \in \mathcal{R}(\underline{A})$ (A.32)

Theorem A.3 - $(\underline{A}')^{\dagger} = (\underline{A}^{\dagger})'$ (A.33)

Proof: (see Zadeh and Desoer)

Theorem A.4 - Let $\underline{S}$ be the hermitian positive semi-definite matrix defined by

$$\underline{S} = \underline{A}'\underline{A} \tag{A.34}$$

Then,

$$\underline{A}^{\dagger} = \underline{S}^{\dagger}\underline{A}' \tag{A.35}$$

Proof: (see Zadeh and Desoer)

Corollary A.2 - Let $\underline{A}$ be an $n \times m$ matrix, $n \geq m$, of full rank (rank m). Then,

$$\underline{A}^{\dagger} = (\underline{A}'\underline{A})^{-1}\underline{A}' \tag{A.36}$$

Proof: By Theorem A.4,

$$\underline{A}^{\dagger} = (\underline{A}'\underline{A})^{\dagger}\underline{A}' \tag{A.37}$$

But $(\underline{A}'\underline{A})$ is a non-singular [actually positive definite] $m \times m$ symmetric matrix. And, the pseudo-inverse of an invertible matrix is equal to the inverse of the matrix.

APPENDIX B

The computer program used to compute the solutions for the examples in Chapter III is listed on the following pages. The programming language used is the M. I. T. version of Fortran IV for the IBM System/360 Operating System and the IBM System/360 Model 44 Programming System.

The operation of the program, and of the various subroutines used, is explained by comment cards preceding each operation. The data cards needed to provide the program with the input data are explained in comment cards at the beginning of the listing on the next page.

```
//RICCATI  JOB  (M4219,3719,2,2000,750,SRI=0),'LEVINE',MSGLEVEL=1
//TEST EXEC FORC,PARM.C='EBCDIC,MAP,DECK'
//C.SYSIN  DD *
C      .............................................................
C      THIS IS THE MAIN PROGRAM FOR CALCULATION OF THE OPTIMAL OUTPUT
C      FEEDBACK GAINS
C
C      INPUT DATA
C          N...........N IS THE DIMENSION OF THE STATE VECTOR (PHI IS NXN)
C          M...........M IS THE DIMENSION OF THE CONTROL VECTOR (F IS MXLR)
C          LR..........LR IS THE DIMENSION OF THE OUTPUT VECTOR
C          INTMAX.......IS THE NUMBER OF TIME STEPS INTO WHICH THE INTERVAL
C                       IS DIVIDED. HENCE, THE TERMINAL TIME.
C          ISEE.........IS THE NUMBER OF TIME STEPS BETWEEN EACH PRINTOUT
C          MAXITS.......IS THE MAXIMUM NUMBER OF ITERATIONS WE WILL TRY
C          MORE.........=1 SIGNIFIES ADDITIONAL COMPUTATIONS ARE TO BE DONE
C
C          THE SECOND DATA CARD SPECIFIES  PRINTING FORMATS
C          THE SECOND,THIRD AND FOURTH FIELDS OF A TYPICAL SECOND CARD
C          FOLLOW
C            (' ',2E12.3)(' ',4E12.3)(' ', E12.3)
C          THE ABOVE CARD IS USEABLE     FOR N=4,M=2,LR=1.
C
C          EPSLO.........CONVERGENCE OCCURS IF DELTA COST<EPSLO
C          H.............IS THE STEP SIZE
C
C          A,B,C,Q,R,S,PHIO ARE READ  BY A READ NAMELIST
C      .............................................................
       DIMENSION PN(3),PM(3),PNN(3),PMLR(3)
       DIMENSION S(2,2),Q(2,2),R(2,2),PHIO(2,2),DUM(2,2),
     1 BT(2,2),RIB(2,2),PHI(2,2),PHIT(2,2),FEED(2,2),SQ(2,2),COST1(2,2),
     2   COST2(2,2)
       COMMON CK(2,2,10001),F(2,2,10001),A(2,2),B(2,2),C(2,2),H,
     1   INTMAX,N,M,LR
 100   READ (5,3001) N,M,LR,INTMAX,ISEE,MAXITS,MORE
       READ (5,3002) (PN(I),I=1,3),(PM(J),J=1,3),(PNN(K),K=1,3),
     1 (PMLR(L),L=1,3)
       READ (5,3003) EPSLO,H
C
C      WE WRITE THE INPUT DATA
       WRITE (6,4001) INTMAX
       WRITE (6,4002) EPSLO,H
       NAMELIST/ZAP/A,B,C,Q,R,S,PHIO
       READ (5,ZAP)
       WRITE (6,2001)
       WRITE (6,PN) ((A(I,J),J=1,N),I=1,N)
       WRITE (6,2002)
       WRITE (6,PM) ((B(I,J),J=1,M),I=1,N)
       WRITE (6,2003)
       WRITE (6,PN) ((C(I,J),J=1,N),I=1,LR)
       WRITE (6,2004)
       WRITE (6,PN) ((Q(I,J),J=1,N),I=1,N)
       WRITE (6,2005)
       WRITE (6,PM)  ((R(I,J),J=1,M),I=1,M)
       WRITE (6,2006)
       WRITE (6,PN)  ((S(I,J),J=1,N),I=1,N)
       WRITE (6,2007)
```

```
      WRITE (6,PN)  ((PHIO(I,J),J=1,N),I=1,N)
C
C     FIRST STEP OF COMPUTATIONS
C
      ITS=0
C
C     COMPUTATION OF R-INVERSE
      DO 101 J=1,M
      DO 101 I=1,M
  101 DUM(I,J)=R(I,J)
      CALL VECT (DUM,M)
C
C     COMPUTATION OF B'
      DO 102 J=1,M
      DO 102 I=1,N
  102 BT(J,I)=B(I,J)
C
C     R-INVERSE TIMES B-TRANSPOSE IS DEFINED AS RIB
      CALL MULT (DUM,BT,RIB,M,N,M)
C
C     WE KNOW AND STORE K(TERMINAL TIME)=S
      DO 103 J=1,N
      DO 103 I=1,N
      CK(I,J,INTMAX)=S(I,J)
      PHI(I,J)  =PHIO(I,J)
  103 PHIT(J,I) =PHIO(I,J)
C
C     PSOL IS USED TO COMPUTE F(0,T)
      CALL PSOL1 (PHI,RIB,INTMAX,FEED,ITS)
      DO 104 L=1,INTMAX
      DO 104 J=1,LR
      DO 104 I=1,M
  104 F(I,J,L)=FEED(I,J)
C
C     COMPUTATION OF THE COST IS SET UP
      CALL MULT (PHIO,PHIT,SQ,N,N,N)
      CALL MULT (S,SQ,COST1,N,N,N)
C
C     RSOL BEGINS THE ITERATIVE LOOP.  IT COMPUTES K(N+1,T) FROM F(N,T)
  105 CALL RSOL (Q,R)
C
C     THIS IS A CHECK ON THE COMPUTATUIONS
      WRITE (6,PN) (( CK(I,J,1),J=1,N),I=1,N)
      WRITE (6,PN) (( CK(I,J,21),J=1,N),I=1,N)
C
C     COMPUTATION OF NEW COST
      DO 106 J=1,N
      DO 106 I=1,N
      PHI(I,J)=PHIO(I,J)
      COST2(I,J)=0.0
      DO 106 K=1,N
  106 COST2(I,J)=COST2(I,J)+CK(I,K,1)*SQ(K,J)
C
C     THIS CAUSES THE ITERATION COUNTER TO INCREASE BY 1.
      ITS=ITS+1
C
C     CHECK FOR CONVERGENCE.  ONLY AFTER THE THIRD ITERATION
```

```
      IF (ITS-2) 108,108,116
  116 TR=0.0
      DO 107 I=1,N
  107 TR=TR+(COST1(I,I)-COST2(I,I))
      IF (TR-EPSLO) 110,110,108
C
C     IF CONVERGENCE, IDONE=2 AND WE PRINT DATA.  IF NOT, IDONE=1 AND PROCEED
  108 DO 109 I=1,N
  109 COST1(I,I)=COST2(I,I)
      IDONE=1
      IF (ITS-MAXITS) 113,110,110
  110 IDONE=2
      WRITE (6,4003) TR,ITS
      WRITE (6,4004) ISEE
      WRITE (6,2008)
      DO 111 L=1,INTMAX,ISEE
      WRITE (6,2009)
  111 WRITE (6,PNN) ((CK(I,J,L),J=1,N),I=1,N)
      WRITE (6,2010)
      DO 112 L=1,INTMAX,ISEE
      WRITE (6,2009)
  112 WRITE (6,PMLR)  ((F(I,J,L),J=1,LR),I=1,M)
      WRITE (6,2011)
C
C     PSOL IS CALLED TO COMPUTE F(N,T) AND PHI(N,T) GIVEN K(N,T)
  113 CALL PSOL (PHI,RIB,IDONE,ISEE,ITS,PNN)
C
C     WE HAVE ALREADY DETERMINED CONVERGENCE AND USED THIS TO SET IDONE
C     IF CONVERGENCE, CHECK FOR MORE DATA.  IF NOT, REITERATE.
      GO TO (105,114),IDONE
  114 IF (MORE) 115,115,100
 3001 FORMAT (7I6)
 3002 FORMAT (3A4,3A4,3A4,3A4)
 3003 FORMAT (2E11.4)
 2001 FORMAT ('0   THE A-MATRIX IS PRINTED BELOW')
 2002 FORMAT ('0   THE B-MATRIX IS PRINTED BELOW')
 2003 FORMAT ('0   THE C-MATRIX IS PRINTED BELOW')
 2004 FORMAT ('0   THE Q-MATRIX IS PRINTED BELOW')
 2005 FORMAT ('0   THE R-MATRIX IS PRINTED BELOW')
 2006 FORMAT ('0   THE S-MATRIX IS PRINTED BELOW')
 2007 FORMAT ('0   THE INITIAL CONDITION MATRIX IS PRINTED BELOW')
 2008 FORMAT ('0   THE K-MATRIX IS PRINTED BELOW')
 2009 FORMAT ('0')
 2010 FORMAT ('0   THE FEEDBACK MATRIX IS PRINTED BELOW')
 2011 FORMAT ('0   THE TRANSITION MATRIX IS PRINTED BELOW')
 4001 FORMAT ('0   THE INTERVAL IS DIVIDED INTO ',I4,' PARTS')
 4002 FORMAT ('0 CONVERGENCE OCCURS IF DELTA COST IS LESS THAN ',E11.4,
     1  ' THE STEP SIZE IS=',E11.4)
 4003 FORMAT ('0 DELTA COST IS=',E11.4,'THIS IS ITERATION',I4)
 4004 FORMAT ('0 THE OUTPUT MATRICES ARE PRINTED ONLY AT T=',I4,'*H')
  115 END
C     ..............................................................
C     SUBROUTINE RSOL
C
C     PURPOSE
C         TO COMPUTE K(N+1,T), GIVEN F(N+1,T)
C
```

```
C     COMMENT
C         INPUT DATA IS PARTLY TRANSFERRED THROUGH COMMON
C     ..................................................................
C
      SUBROUTINE RSOL (XQ,XR)
      DIMENSION XQ(2,2),XR(2,2),XDUM(2,2),FC(2,2),BFC(2,2),RFC(2,2),
     1  CFRFC(2,2),ABFC(2,2),XABFC(2,2),D(2,2,4)
      COMMON RK(2,2,10001),XF(2,2,10001),XA(2,2),XB(2,2),XC(2,2),H,
     1 INTMAX,N,M,LR
C
C     COMPUTATION BEGINS AT T=INTMAX, THE TERMINAL TIME
  100 NDELT=INTMAX
C
C     FIRST STEP OF RUNGE-KUTTA ROUTINE BEGINS
  101 DO 102 J=1,N
      DO 102 I=1,N
  102 XDUM(I,J)=RK(I,J,NDELT)
C
      L=0
C
      DO 103 J=1,N
      DO 103 I=1,M
      FC(I,J)=0.0
      DO 103 K=1,LR
  103 FC(I,J)=FC(I,J)+XF(I,K,NDELT)*XC(K,J)
C
      CALL MULT (XB,FC,BFC,N,N,M)
      CALL MULT (XR,FC,RFC,M,N,M)
C
      DO 105 J=1,N
      DO 105 I=1,N
      CFRFC(I,J)=0.0
      DO 104 K=1,M
  104 CFRFC(I,J)=CFRFC(I,J)+FC(K,I)*RFC(K,J)
  105 ABFC(I,J)=XA(I,J)-BFC(I,J)
C
  106 L=L+1
C
      CALL MULT (XDUM,ABFC,XABFC,N,N,N)
C
C     EVALUATION OF THE PARTIAL SLOPE IN RUNGE KUTTA ROUTINE
      DO 107 J=1,N
      DO 107 I=1,N
  107 D(I,J,L)=H*(XABFC(I,J)+XABFC(J,I)+XQ(I,J)+CFRFC(I,J))
C
C     LOGIC FOR ROUTING TO EACH PHASE OF ONE RUNGE KUTTA STEP
      GO TO (108,108,110,112),L
C
  108 DO 109 J=1,N
      DO 109 I=1,N
  109 XDUM(I,J)=.5*D(I,J,L)+RK(I,J,NDELT)
C
      GO TO 106
C
  110 DO 111 J=1,N
      DO 111 I=1,N
  111 XDUM(I,J)=D(I,J,L)+RK(I,J,NDELT)
```

```
C
      GO TO 106
C
C     CALCULATION OF K(N+1,T-1)
  112 DO 113 J=1,N
      DO 113 I=1,N
  113 RK(I,J,NDELT-1)=RK(I,J,NDELT)+(D(I,J,1)+2.*D(I,J,2)+2.*D(I,J,3)
     1   +D(I,J,4))/6.
C
C     TIME IS STEPPED BACKWARDS ONE STEP
      NDELT=NDELT-1
C
C     IF T IS NOT ZERO, WE BEGIN THE NEXT STEP.  IF T=0, WE RETURN TO
C     MAIN.
      IF (NDELT-1) 114,114,101
C
  114 RETURN
      END
C     ..............................................................
C     SUBROUTINE PSOL
C
C     PURPOSE
C         TO COMPUTE F(N,T) AND PHI(N,T) GIVEN K(N,T)
C
C     COMMENT
C         INPUT DATA IS PARTLY TRANSFERRED THROUGH COMMON
C     ..............................................................
C
      SUBROUTINE PSOL (YPHI,YRIB,IDONE,ISEE,NSTART,PNN)
      DIMENSION PNN(3),FC(2,2),YPHI(2,2),YRIB(2,2),YFEED(2,2),CPHI(2,2),
     1 PHIC(2,2),CPPC(2,2),DUM1(2,2),DUM2(2,2),DUM3(2,2),
     2 BFC(2,2),ABFC(2,2),ABFCY(2,2)
      COMMON PK(2,2,10001),YF(2,2,10001),YA(2,2),YB(2,2),YC(2,2),H,
     1 INTMAX,N,M,LR
C
C     THE VALUE OF ICE DETERMINES WHETHER A VALUE OF PHI IS PRINTED
      ICE=ISEE
C
C     COMPUTATION BEGINS AT T=0,OR NDELT=1
      NDELT=1
C
C     IDONE =2 MEANS THE ITERATIONS HAVE CONVERGED AND WE WANT TO
C     COMPUTE THE OPTIMAL PHI.  WE DO NOT RECOMPUTE F.  IF IDONE=1, WE
C     COMPUTE A NEW F AND A NEW PHI.
  100 GO TO (102,106),IDONE
C
C     THIS ENTRY IS USED TO COMPUTE THE INITIAL F.
      ENTRY PSOL1 (YPHI,YRIB,NDELT,YFEED,NSTART)
C
C     COMPUTATION OF C TIMES PHI
  102 CALL MULT (YC,YPHI,CPHI,LR,N,N)
C
C     COMPUTATION OF C TIMES PHI TRANSPOSE
      DO 103 J=1,N
      DO 103 I=1,LR
  103 PHIC(J,I)=CPHI(I,J)
C
```

```
C     COMPUTATION OF C*PHI*PHI'*C'
      CALL MULT (CPHI,PHIC,CPPC,LR,LR,N)
C
C     COMPUTATION OF NC*PHI*PHI'*C') INVERSE
      CALL VECT (CPPC,LR)
C
C     ETC. ETC.
      CALL MULT (PHIC,CPPC,DUM1,N,LR,LR)
      CALL MULT (YPHI,DUM1,DUM2,N,LR,N)
      DO 104 J=1,LR
      DO 104 I=1,N
      DUM3(I,J)=0.0
      DO 104 K=1,N
  104 DUM3(I,J)=DUM3(I,J)+PK(I,K,NDELT)*DUM2(K,J)
      CALL MULT (YRIB,DUM3,YFEED,M,LR,N)
C
C     YFEED IS NOW THE VALUE OF F AT THIS TIME AND THIS ITERATION
C     IF NSTART=0, THIS IS F(0,T) AND WE RETURN TO MAIN.  IF NSTART 0,
C     WE CONTINUE
      IF (NSTART) 114,114,112
C
C     THIS STORES THE NEW VALUE OF F IN THE PROPER PLACE
  112 DO 105 J=1,LR
      DO 105 I=1,M
  105 YF(I,J,NDELT)=YFEED(I,J)
C
C     KNOWING F, WE BEGIN COMPUTING THE VALUE OF PHI AT THE NEXT TIME
  106 DO 101 J=1,N
      DO 101 I=1,M
      FC(I,J)=0.0
      DO 101 K=1,LR
  101 FC(I,J)=FC(I,J)+YF(I,K,NDELT)*YC(K,J)
C
      CALL MULT (YB,FC,BFC,N,N,M)
C
      DO 107 J=1,N
      DO 107 I=1,N
  107 ABFC(I,J)=YA(I,J)-BFC(I,J)
C
C     IF IDONE=1, WE SKIP THE WRITING ROUTINE.  IF IDONE=2, WE WRITE
C     EVERY ICE VALUES OF PHI
      GO TO (210,109),IDONE
C
C     ROUTINE FOR WRITING PHI
  109 ICE=ICE+1
      IF (ICE-ISEE) 210,113,113
  113 ICE=0
      WRITE (6,1000)
      WRITE (6,PNN) ((YPHI(I,J),J=1,N),I=1,N)
C
C     CONTINUATION OF THE COMPUTATION OF NEXT PHI
  210 CALL MULT (ABFC,YPHI,ABFCY,N,N,N)
C
      DO 108 J=1,N
      DO 108 I=1,N
  108 YPHI(I,J)=YPHI(I,J)+H*ABFCY(I,J)
C
```

```
 1000 FORMAT ('0')
C
C       WE STEP THE TIME ONE STEP
  110 NDELT=NDELT+1
C
C       IF WE HAVE REACHED THE TERMINAL TIME, WE RETURN TO MAIN.
        IF (NDELT-INTMAX) 100,111,111
  111 IF (IDONE-2)114,115,115
  115 WRITE (6,1000)
      WRITE (6,PNN) ((YPHI(I,J),J=1,N),I=1,N)
  114 RETURN
C
C       DEBUG PACKET PRINTS USEFUL DATA
        DEBUG SUBTRACE,INIT(YRIB,IDONE)
        AT 112
        IF (NDELT-1) 300,300,301
  300 DISPLAY DUM3,DUM2,DUM1,CPPC,CPHI,FC
  301 CONTINUE
        END
C
C       ..............................................................
C       ..............................................................
C
C       SUBROUTINE MULT
C
C       PURPOSE
C          TO COMPUTE THE PRODUCT OF TWO MATRICES.
C          GAMMA(N X M) = ALPHA(N X L) * BETA(L X M)
C
C       USAGE
C          CALL MULT(ALPHA,BETA,GAMMA,N,M,L)
C
C       DESCRIPTION OF PARAMETERS
C          ALPHA- N X L  REAL MATRIX
C          BETA - L X M  REAL MATRIX
C          GAMMA-  N X M  REAL MATRIX
C          N     - NUMBER OF ROWS IN ALPHA
C          M     - NUMBER OF COLUMNS IN BETA
C          L     - NUMBER OF COLUMNS(ROWS) IN ALPHA(BETA)
C
C       ..............................................................
C
        SUBROUTINE MULT(ALPHA,BETA,GAMMA,N,M,L)
        DIMENSION ALPHA(2,2),BETA(2,2),GAMMA(2,2)
        DO 10 I=1,N
        DO 10 J=1,M
        GAMMA(I,J)=0.0
        DO 10 K=1,L
        GAMMA(I,J)=GAMMA(I,J)+ALPHA(I,K)*BETA(K,J)
   10 CONTINUE
        RETURN
        END
C       ..............................................................
C
C       SUBROUTINE VECT
C
C       PURPOSE
```

```
C          TO CONVERT A SQUARE MATRIX TO VECTOR MODE=O,
C          TO CALL THE MATRIX INVERSION SUBROUTINE AND
C          TO RECONVERT THE INVERTED VECTOR TO MATRIX FORM.
C
C       USAGE
C          CALL VECT(RMAT,M)
C
C       DESCRIPTION OF PARAMETERS
C          M    - THE DIMENSION OF THE SQUARE MATRIX
C          RMAT - THE MATRIX TO BE INVERTED AND ITS INVERSE
C
C       REMARKS
C          THE INVERSE IS STORED IN THE LOCATIONS OF THE INPUT MATRIX.
C
C       SUBROUTINES REQUIRED
C          INVERT
C
C       ..................................................................
C
        SUBROUTINE VECT(RMAT,M)
        DIMENSION RMAT(2,2),AMAT(4)
C       MATRIX TO VECTOR CONVERSION
        JNOT=0
    150 JNOT=JNOT+1
        IF(M.LT.JNOT)GO TO 180
        KONE=1+M*(JNOT-1)
        KEND=M*JNOT
        DO 170 K=KONE,KEND
        I=K-M*(JNOT-1)
        AMAT(K)=RMAT(I,JNOT)
    170 CONTINUE
        GO TO 150
    180 CONTINUE
C       MATRIX INVERSION
        CALL INVERT(AMAT,M,M)
C       VECTOR TO MATRIX CONVERSION
        KNOT=M*M
        DO 190 K=1,KNOT
        J=(K-1)/M+1
        I=K-M*(J-1)
        RMAT(I,J)=AMAT(K)
    190 CONTINUE
        RETURN
        END
C
C       ..................................................................
C       ..................................................................
C
C       SUBROUTINE INVERT
C
C       PURPOSE
C          TO INVERT A REAL SQUARE MATRIX
C
C       USAGE
C          CALL INVERT(A,NN,N)
C
C       DESCRIPTION OF PARAMETERS
```

```
C          A     - REAL SQUARE MATRIX TO BE INVERTED
C          NN    - ORDER OF MATRIX A
C          N     - MAXIMUM ORDER OF A. SET EQUAL TO NN.
C
C       METHOD
C          THE INVERSE OF A IS COMPUTED AND STORED IN A.
C
C       REMARKS
C          THIS SUBROUTINE IS A SLIGHTLY MODIFIED VERION OF THE
C          IBM SHARE NO. 1533 MATRIX INVERSION SUBROUTINE.
C
C       ..............................................................
C
        SUBROUTINE INVERT(A,NN,N)
        DIMENSION A(4),M(2),C(2)
        IF(NN.NE.1)GO TO 80
        A(1)=1./A(1)
        GO TO 300
     80 DO 90 I=1,NN
        M(I)=-I
     90 CONTINUE
        DO 140 I=1,NN
C       LOCATE LARGEST ELEMENT
        D=0.0
        DO 112 L=1,NN
        IF(M(L).GT.0)GO TO 112
        J=L
        DO 110 K=1,NN
        IF(M(K).GT.0)GO TO 108
        IF(ABS(D)-ABS(A(J)))105,105,108
    105 LD=L
        KD=K
        D=A(J)
    108 J=J+N
    110 CONTINUE
    112 CONTINUE
C       INTERCHANGE ROWS
        TEMP=-M(LD)
        M(LD)=M(KD)
        M(KD)=TEMP


        L=LD
        K=KD
        DO 114 J=1,NN
        C(J)=A(L)
        A(L)=A(K)
        A(K)=C(J)
        L=L+N
    114 K=K+N
C       DIVIDE COLUMN BY LARGEST ELEMENT
        NR=(KD-1)*N+1
        NH=NR+N-1
        DO 115 K=NR,NH
    115 A(K)=A(K)/D
C       REDUCE REMAINING ROWS AND COLUMNS
```

```
      L=1
      DO 135 J=1,NN
      IF(J.NE.KD)GO TO 130
      L=L+N
      GO TO 135
  130 DO 134 K=NR,NH
      A(L)=A(L)-C(J)*A(K)
  134 L=L+1
  135 CONTINUE
C     REDUCE ROW
      C(KD)=-1.0
      J=KD
      DO 140 K=1,NN
      A(J)=-C(K)/D
      J=J+N
  140 CONTINUE
C     INTERCHANGE COLUMNS
      DO 200 I=1,NN
      L=0
  150 L=L+1
      IF(M(L).NE.I)GO TO 150
      K=(L-1)*N+1
      J=(I-1)*N+1
      M(L)=M(I)
      M(I)=I
      DO 200 L=1,NN
      TEMP=A(K)
      A(K)=A(J)
      A(J)=TEMP
      J=J+1
  200 K=K+1
  300 CONTINUE
      RETURN
      END
C
C     ..............................................................
/*
```

REFERENCES

1. Nyquist, H., "Regeneration Theory," Bell System Tech. J., 11, pp. 126-147 (1932)

2. Hazen, H. L., "Theory of Servomechanisms," J. Franklin Inst., 218, pp. 543-580 (1934)

3. Wiener, N., The Extrapolation, Interpolation and Smoothing of Stationary Time Series, Technology Press, M.I.T., Cambridge, Mass., 1949

4. Athans, M. and Falb, P. F., Optimal Control: An Introduction to the Theory and its Applications, McGraw-Hill Book Co., New York, 1966

5. Kalman, R. E., "Contribution to the Theory of Optimal Control," Bol. Soc. Mat. Mexico, pp. 102-119 (1960)

6. Kalman, R. E., "When is a Linear System Optimal?," J. Basic Engineering, (ASME Trans.) Vol. 86, pp. 1-10 (1964)

7. Kalman, Ho, and Narendra, "Controllability of Linear Dynamical Systems," Contributions to Differential Equations, Vol. 1, 1962

8. Kalman, R. E., "Mathematical Description of Linear Dynamical Systems," J. SIAM on Control, Ser. A., Vol. 1, No. 2, pp. 152-192 (1963)

9. Kalman, R. E. and Bucy, R. S., "New Results in Linear Filtering and Prediction Theory," J. Basic Engineering, (ASME Trans.), Vol. 83, pp. 95-108, March, 1961

10. Luenberger, D. G., "Observers for Multivariable Systems," IEEE Trans. on Automatic Control, Vol. AC-11, pp. 190-197, April, 1966

11. Kleinman, D. L., "On the Linear Regulator Problem and the Matrix Riccati Equation," Electronic Systems Laboratory Report 271, Mass. Inst. of Tech., (1966)

12. Athans, M., "The Matrix Minimum Principle," Information and Control, Vol. 11, pp. 592-606, Nov.-Dec. (1967)

13. Kleinman, D. L. and Athans, M., "The Design of Suboptimal Linear Time-Varying Systems," IEEE Trans. on Automatic Control, Vol. AC-13, pp. 150-159, April (1968)

REFERENCES (Contd.)

14. Kleinman, D. L., Fortmann, T., and Athans, M., "On the Design of Linear Systems with Piecewise Constant Feedback Gains," Preprints of Ninth Joint Automatic Control Conference, pp. 698-710, June (1968)

15. Kleinman, D. L., "Suboptimal Design of Linear Regulator Systems Subject to Computer Storage Limitations," Electronic Systems Laboratory Report 297, Mass. Inst. of Tech., (1967)

16. Brockett, R. W. and Lee, H. B., "Frequency-Domain Instability Criteria for Time-Varying and Nonlinear Systems," Proceedings of the IEEE, Vol. 55, No. 5, pp. 604-619, May (1967)

17. Luenberger, D. G., "A New Derivation of the Quadratic Loss Control Equation," IEEE Transactions on Automatic Control, Vol. AC-10, No. 2, p. 202, April (1965)

18. Bellman, R., Introduction to Matrix Analysis, McGraw-Hill Book Co., New York (1960)

19. Zadeh, L. A. and Desoer, C. A., Linear Systems Theory, McGraw-Hill Book Co., New York (1963)

20. Roxin, E., "The Existence of Optimal Controls," Michigan Math. J., 9, pp. 109-119 (1962)

21. Filippov, A. F., "On Certain Questions in the Theory of Optimal Control," J. SIAM on Control, Ser. A, Vol. 1, No. 1, pp. 76-84 (1962)

22. Newton, G. C., Gould, L. A., and Kaiser, J. F., Analytical Design of Linear Feedback Controls, John Wiley and Sons Inc., New York, 1961

23. Willis, B. H., "On the Least-Squares Optimization of Constant Linear Regulator Systems," Ph. D. thesis, Dept. of Aeronautics and Astronautics, M.I.T., September, 1965

24. Ferguson, J. D. and Rekasius, Z. V., "Optimal Linear Systems with Incomplete State Measurements," Proceedings Sixth Annual Allerton Conference on Circuit and Systems Theory, pp. 671-679, October (1968)

25. Rekasius, Z. V., "Optimal Linear Regulators with Incomplete State Feedback," IEEE Transactions on Automatic Control, Vol. AC-12, pp. 296-299, June, 1967.