INFORMATION TO USERS

This material was produced from a microfilm copy of the original document. While the most advanced technological means to photograph and reproduce this document have been used, the quality is heavily dependent upon the quality of the original submitted.

The following explanation of techniques is provided to help you understand markings or patterns which may appear on this reproduction.

1. The sign or "target" for pages apparently lacking from the document photographed is "Missing Page(s)". If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting thru an image and duplicating adjacent pages to insure you complete continuity.

2. When an image on the film is obliterated with a large round black mark, it is an indication that the photographer suspected that the copy may have moved during exposure and thus cause a blurred image. You will find a good image of the page in the adjacent frame.

3. When a map, drawing or chart, etc., was part of the material being photographed the photographer followed a definite method in "sectioning" the material. It is customary to begin photoing at the upper left hand corner of a large sheet and to continue photoing from left to right in equal sections with a small overlap. If necessary, sectioning is continued again — beginning below the first row and continuing on until complete.

4. The majority of users indicate that the textual content is of greatest value, however, a somewhat higher quality reproduction could be made from "photographs" if essential to the understanding of the dissertation. Silver prints of "photographs" may be ordered at additional charge by writing the Order Department, giving the catalog number, title, author and specific pages you wish reproduced.

5. PLEASE NOTE: Some pages may have indistinct print. Filmed as received.

75-24,622

BURBANK, Harry Marvin, Jr., 1940-
    OPTIMAL CONTROL AND IDENTIFICATION OF
    STOCHASTIC SYSTEMS USING DIFFERENTIAL GAME
    THEORY.

    New Jersey Institute of Technology, D.Eng.Sc.,
    1975
    Engineering, electronics and electrical

OPTIMAL CONTROL AND IDENTIFICATION OF STOCHASTIC SYSTEMS

USING DIFFERENTIAL GAME THEORY

BY

HARRY MARVIN BURBANK,JR.

A THESIS

PRESENTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE

OF

DOCTOR OF ENGINEERING SCIENCE IN ELECTRICAL ENGINEERING

AT

NEW JERSEY INSTITUTE OF TECHNOLOGY

Newark, New Jersey
1975

APPROVAL OF DISSERTATION

OPTIMAL CONTROL AND IDENTIFICATION OF STOCHASTIC SYSTEMS

USING DIFFERENTIAL GAME THEORY

BY

HARRY MARVIN BURBANK, JR.

FOR

DEPARTMENT OF ELECTRICAL ENGINEERING

NEW JERSEY INSTITUTE OF TECHNOLOGY

BY

FACULTY COMMITTEE

APPROVED: _____ CHAIRMAN

_____

_____

_____

_____

NEWARK, NEW JERSEY

JUNE, 1975

PREFACE

It was the original intent of this work to provide a theoretical basis to study several problems associated with electrocardiagrams. Preliminary reading and discussions sparked the interest in this area. It became clear that non-stationary statistics were involved, leading to the formulations given in the dissertation. As the theoretical study progressed, the applicability of the work to the area of electroencephalograms was also evident. In each step , with every hurdle encountered, the intricacies involved lent a stronger and stronger intuitive notion that a fruitful path was being followed. Hopefully now, the joining of theory and application can be pursued with an approach that will bear much fruit.

The completion of this dissertation would not have been possible without the patience, understanding and endurance of family and some good friends. To them is given a simple, sincere thank you. Also, grateful thanks are given to Margaret, Sister Clarissa, Bobbie, Karen, Regina, Diana, Rosemary, Karen, Ruth, and Gary who all helped put it together.

ACKNOWLEDGEMENTS

ABSTRACT

This dissertation deals with linear systems subjected to stochastic disturbances. The class of stochastic processes considered is the class of second order stochastic processes characterized by having finite continuous covariance. The properties of the covariance provide means to formulate optimization problems without the difficulties present when the covariance is not finite or continuous.

The first aspect studied was several classes of optimal control problems. The effects of the stochastic processes were approximated by the effects of its first two moments. This procedure resulted in allowing optimal system controls to be found whatever the first two moments of the stochastic input were, or "worst case" optimal controls were found. Differential game theory was used to solve the "worst case" problem.

Then, a model reference adaptive control system was employed to permit simultaneous parameter identification and control to be obtained in an on-line environment. The parameter identification was accomplished using gradient or steepest descent techniques. The control inputs were updated as the parameters were changed yielding sub-optimal control of the physical system. In addition, minimum error covariance estimation of linear systems with second order stochastic disturbances was developed.

TABLE OF CONTENTS

LIST OF FIGURES

Figure	Title
1.1	General linear problem
1.2	General linear problem without measurement
6.3.2.1	Model reference adaptive system
7.1.1.1	Analog computer diagram of example
7.1.2.1	Plot of z_1
7.1.2.2	Plot of λ_1
7.1.2.3	Plot of u_1
7.1.2.4	Plot of u_2
7.1.2.5	Plot of K
7.1.2.6	Plot of z_2
7.1.2.7	Plot of z_3
7.1.2.8	Plot of λ_2
7.1.2.9	Plot of λ_3
7.1.2.10	Plot of u_3
7.1.3.1	Identification of a
7.1.3.2	Identification of b
7.1.3.3	Trajectory of z_1 from model reference simulation
7.1.3.4	Trajectory of z_1 from optimal solution
7.1.3.5	Trajectory of u_1 from model reference simulation
7.1.3.6	Trajectory of u_1 from optimal solution
7.1.3.7	Plot of noise v used in model reference simulation
7.1.3.8	Regions of parameter space
7.1.3.9	Identification of a for other noise inputs
7.1.3.10	Identification of b for other noise inputs

LIST OF THEOREMS

Theorem

CHAPTER 1

INTRODUCTION

In this dissertation a class of optimal control and
identification problems is studied. The approach developed
permits treatment of some physical situations not yet fully
explored and allows a different approach to physical situa-
tions previously dealt with. The overall problem is to
optimally control a plant in the presence of noise distur-
bances and sumultaneously identify parameters of the plant.
The general linear problem is pictured below.
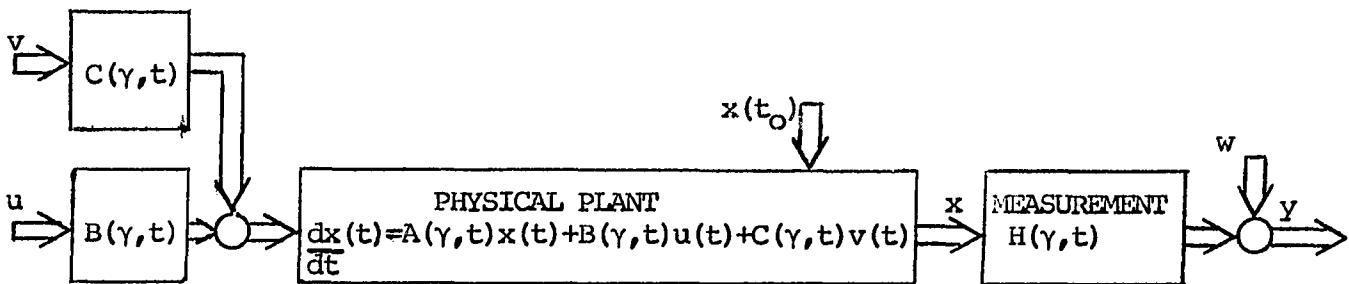
Figure 1.1 General linear problem

In this situation, the measurement ẏ corrupted by
noise ẇ, of the state x is to be driven by the control u,
which is optimal in some manner, while in the presence of
additional noise v, at the input to the plant. Simul-
taneously, the vector γ , which is the collection of all
elements of A, B, C, and H which are not known, is to be
determined in some way.

The optimal control solution can be carried out offline. Then the identification can be performed online, while the actual plant is being controlled.

The problem of Figure 1.1 is linear, however both the plant and measurement may actually be nonlinear. In this work, only linear problems are dealt with leaving nonlinear problems for future research. The identification scheme employed is discussed in Chapter 6. The stochastic optimal control problem is detailed in Chapters 2 and 3. Throughout the dissertation only continuous-time problems are studied, since stochastic discrete-time problems submit to some approaches which just do not carry over to continuous time. This point will be clarified in Chapter 2. Also the measurement y, will be assumed to be equal to the state x, within the dissertation. This corresponds to $w\equiv o$ and $H(\gamma,t)=I$. Having completed the solution of this problem, the extension to the case of noise corrupted measurement is possible. These modifications slightly change Figure 1.1 to the following.
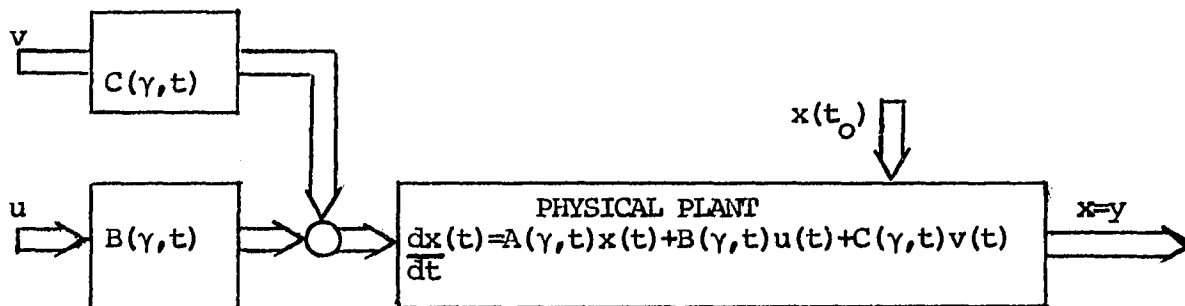
Figure 1.2 General linear problem

without measurement

The state equation is

$$\frac{dx}{dt}(t)=A(\gamma,t)x(t)+B(\gamma,t)u(t)+C(\gamma,t)v(t),t\in[t_o,t_f] \quad 1.1$$

where

x is an n vector

$A(\gamma,t)$ is an nxn matrix

u is an r vector

$B(\gamma,t)$ is an nxr matrix

v is an s vector

$C(\gamma,t)$ is an nxs matrix

and

γ is an m vector.

The state of the system is x and the measurement is y=x. The optimal control is u and the additive stochastic disturbance is v. The unknown gains, time constants, etc. which are the elements of the matrices A, B and C form the vector γ.

The overall problem of figure 1-2 essentially is the combination of two problems: stochastic optimal control and identification. In this work, differential game theory will be employed in the "optimal" part of the solution. Generally speaking background material required for identification problems is presented in the texts by Sage and Melsa (96) and Graupe (31). Discrete-time process identification is given in the text by Mendel (76). Identification

is the subject of the survey articles by Astrom and

Eykhoff (7) and Balakrishnan (14) and the entire special

issue of the December, 1974 IEEE Transaction on Automatic

Control. Deterministic optimal control background is

presented in the texts by Athans and Falb (9), Hsu and

Meyer (38), and Leitmann (67). The variational techniques

employed in this work are found in Kirk (45), Sage (95) and

Citron (22). Some of the proofs depend on a matrix

formulation similar to that in Athans (10). Stochastic

optimal control for some types of stochastic processes is

given in Astrom (8), and Sage and Melsa (97). The separa-

tion principle is found in Wonham (109). Various exten-

sions and alternatives to the separation principle exist

as in e.g. Athans (12). The original work on deterministic

differential game theory is in Isaacs (39). Other texts

are Blaquiere (18) and Friedman (27). Articles by Ho (33),

Kuo (50) and (52), and Kuo and Burbank (55) present solu-

tions to wide classes of two person, zero-sum differential

games. Pure stochastic differential games are discussed

in Behn and Ho (16) and by Willman in (108). Estimation

theory is needed for the study of stochastic optimal control,

and especially for the separation theorem, which states that

the optimal feedback control solution is obtained only after

the best estimate of the state has been found.

Texts on estimation theory are Bryson and Ho (20),

.and Sage and Melsa (97). The monograph by Lee (64) formulates estimation type problems concisely. The detailed class of problems will be developed in Chapters 2 and 3. References dealing with these details will be given in those chapters.

For problems of stochastic optimal control, the approaches used to date require estimation of the state and much a priori knowledge of the stochastic disturbances v and w. The linear estimation problem was first presented in Kalman (40). Many extensions have been made such as in Mehra (74) where errors in initial values of the variance are discussed. In Lee (64) the estimation problem is solved using a Bayesian approach. In Kushner (62), Sage and Melsa (97) and Meditch (73) linear and nonlinear estimation, and stochastic control are discussed. Astrom (8) uses the Ito formulation for the same type of problems.

The class of stochastic processes studied is the class of second order processes. Davis (23) and Boonton (19) deal with prediction and estimation for this class of inputs in the classical manner. A deterministic disturbance in a differential game is given by Krikelis and Rekasius (49). Their problem is a conflict of interest game between a deterministic optimal control and a deterministic disturbance. The differential game approach to be employed in this research considers a stochastic disturbance. It is

assumed that the effects of the stochastic disturbance
are approximated by its first two moments. The conti-
nuous time vector case is studied. In Yoshikawa (110) a
scalar two stage discrete stochastic differential game
is solved in terms of the complete probability density
function. The problem here is based on the assumption
that the density function can be approximated by its
first two moments, but a wider class of problems is con-
sidered than in (110). Differential games with imperfect
information are covered by Leondes and Pearson (69),
Leondes and Stuart (68) and Kushner and Chamberlain (61).
In these references noise corrupted measurements or in-
complete state situations are considered.

The complete probability density function for a class
of nonlinear Bayesian estimation problems is approximated
for the discrete time case only by Alspach (2), and
Sorensen and Alspach (105). In another article (1) they
consider non Gaussian Bayesian estimation. Sain and
Liberty (98) computationally obtain a density function for
a quadratic performance index.

The class of differential games studied here is
characterized as games of imperfect information in that
parameter identification is required. It is assumed though
that the stochastic disturbance can be approximated by its

first .two moments which is not the most general case.
However the problem of simultaneous optimal control and
identification included in this research is seen to be
an extension of known results even with this assumption.
The approach developed here is one in which the plant
is optimally controlled in the presence of any stochastic
process v, considering the effects of the disturbances as
follows. The first two moments of v are assumed to pro-
vide sufficient information to obtain physically workable
solutions to the overall problem. This assumption is
crucial to all that follows but it is noted that the
effects of any finite number of moments could be included
using the approach of this dissertation as a starting point.
The manner in which the first two moments propagate through
the plant or system then, is the means to describe how the
stochastic disturbance affects, in a degrading manner, the
optimal operation of the plant. These two moments are then
considered to be inputs to the same plant as the optimal
control u. There is apparent the conflict between the
effects of the optimal control u and the effects of the two
moments of the stochastic disturbance v. For each given
performance index, a quantitative measure of these effects
is obtained.

Using the conflict characterization, a differential
game is defined between "man" who chooses to optimally

control the plant versus "nature" who seeks to choose moments of stochastic disturbances which degrade optimal performance . The "man-nature" differential game approach to other kinds of problems appears e.g. in Kuo (51) and (53).

The approach to be employed using differential game theory with imperfect information in the sense that the parameter γ' is unknown is mentioned by Ho in (34) as a possible area for future research. This approach is formulated in this work for the first time as applied to the problems of figure 1.2 . One contribution then is the differential game approach formulation. Actually, the formulation is equivalent to an absolute-worst-case controller design problem as described in Ragade and Sarma (92).

The specific class of problems considered in the research is for stochastic processes with finite variances which are also continuous in the mean square sense with respect to time. To the author's knowledge, this class of stochastic processes with unknown γ has not been fully dealt with before in a control problem or in an identification problem. This class of stochastic process is best characterized by the term nonstationary since two independent time variables are required to mathematically describe the moments of the process. Nonstationary inputs are dealt with in Boonton (19), but the resulting integral equations are very difficult to deal with. Baggeroer (13) illustrates the

solution to similar integral equations as does Shinbrot
(104). Davis (23) formulates a class of nonstationary
prediction problems.

As will be seen in detail in Chapters 2 and 3, the
specific problem characterization of this work is one in
which the stochastic disturbances are non-stationary with
correlation times of the same order of magnitude as the
system or plant. This situation appears in many biomedical
monitoring and control problems. In particular, as seen in
Kawabata (41), EEG waveforms exhibit these properties. It
is speculated that the theory developed will permit much
more thorough analysis and understanding of any EEG related
phenomena, such as time series analysis of the waveforms,
modeling of the system generating the waveform, and/or con-
trol of the system generating the waveform by drugs. For
example, the state of a person's consciousness, sleeping,
awake, alert, etc. is influenced by drugs.

The therapeutic use of drugs and the covert use of
drugs are both problems requiring deeper understanding.
The effect of these drugs on EEG waveforms, and the modeling
of the system creating these waveforms would possibly enhance
this understanding from a new viewpoint. Several of the
previous points are considered in Nunez (82).

Basically this work presents an approach to the solution

of an optimal control problem in which stochastic distur-
bances are present. Techniques for on-line identification
are included. The approach is to consider an absolute-
worst-case situation as described mathematically with
differential game theory. The particular class of pro-
blems dealt with is for nonstationary, continuous-time
stochastic disturbances. These disturbances are typically
found in many biomedical areas, and especially in EEG related
phenomena.

Chapter 2 provides material needed to formulate the
specific problem. This formulation is presented in Chapter 3.
Chapters 4 and 5 provide necessary and sufficient conditions
for the solution of the "optimal" part of the problem.
Chapter 6 provides the solution to the identification por-
tion of the problem. In Chapter 7, examples are worked out
to illustrate the theory. The results are summarized in
Chapter 8 and the extensions possible are presented in
Chapter 9. The derivations of the state equations is given
in Appendix A. Appendix B contains derivations of perform-
ance index and endpoint condition transformations. The
proofs of the necessary condition theorems are in Appendix C,
and the proofs of the sufficient condition theorems in
Appendix D.

CHAPTER 2

STOCHASTIC PROCESSES

When a stochastic process is only a part of a larger
problem, some assumptions must be made about the stochastic
process and only then can the solution to the larger problem
begin. This chapter develops the main points required to
classify stochastic processes and in particular to delineate
the class of stochastic processes chosen for this study.

2.1 Discrete and Continuous Stochastic Processes

A stochastic process is, in words, a random phenomena
that changes with time, or some other parameter. The real,
scalar stochastic process $v(.,.)$, a family or ensemble of
functions, depends on the outcome of an event $w \in \Omega$, the
sample space and, a parameter t, usually assumed to be time
where $t \in [\ldots,-1,0,1,\ldots] \triangleq I$ for a discrete-time process, and
$t \, (-\infty,\infty) \triangleq E$ for a continuous-time process.

For fixed $t=t_j \in I$ (or E), $v(t_j,.)$ is a random variable.

For fixed $w=w_j \in \Omega$, $v(.,w_j)$ is a (deterministic) function
    of time.

For fixed $t=t_j$ and $w=w_j$, $v(t_j,w_j)$ is a number.

For each fixed t_j, the random variable $v(t_j,.)$ is
defined on a sample space Ω, where there exists a Borel field B
of subsets of Ω, and a probability measure P on B. The

probability space (Ω,B,P) is the basis for the measure-theoretic (axiomatic) study of probability, as discussed in Dubes (25). The sample space Ω, can in general, be either discrete, continuous, or mixed. That is, $\Omega=\{j;j\epsilon I\}$ or $\Omega=E$ or some combination. Similarly, $t_j\epsilon E$ or $t_j\epsilon I$. The stochastic process then may be a combination of discrete-time or continuous time with discrete, continuous, or mixed sample space. Considering only continuous sample spaces Ω, a quantitative description of the stochastic process is desired. Prabhu (90), states that a stochastic process $v(.,.)$ is statistically determined if the nth order joint distribution function $F(\xi_1,\xi_2,\dots,\xi_n;t_1,t_2,\dots,t_n)=$

$$P(v(t_1,.)\leq\xi_1;v(t_2,.)\leq\xi_2;\dots;v(t_n;)\leq\xi_n)$$

.2.1.1

is known for all n and $t_1,t_2,\dots,t_n$ where F satisfies the symmetry and compatibility conditions. For the discrete-time scalar stochastic process $v(t,.),t\epsilon[0,1,\dots,k],k<\infty$, there are finitely many distribution functions, and a complete statistical description of the process is possible with them. On any open or closed subset of the real line, however, there are infinitely many instants of time, hence infinitely many joint distribution functions and in theory, the continuous-time stochastic process can never be statistically determined. This point is of major significance in estimation theory where density functions have to be defined. It is possible in

theory to completely determine all the distribution
functions of a discrete-time process and utilize them,
whereas continuous-time problems do not submit to this
method of study. For this reason, the present research
deals with continuous-time dynamic and stochastic processes,
since the discrete situation can be studied later. However,
the reverse is not true for many kinds of analysis.

The first order distribution function

$$F(\xi_1;t_1) = P(v(t_1,\cdot) \leq \xi_1)$$

2.1.2

over the ensemble $v(t_1,\cdot)$, for all $t_1 \in [t_0,t_f]$ exists as
does the second order joint distribution function.

$$F(\xi_1,\xi_2;t_1,t_2) = P(v(t_1,\cdot) \leq \xi_1; v(t_1,\cdot) \leq \xi_2)$$

2.1.3

over the ensembles $v(t_1,.)$ and $v(t_2,.)$ for all
$t_1,t_2 \in [t_0,t_f]$. From these two functions much useful
information is available. Theorems involving only these
two functions and their properties appear in Bhat (17) and
Hoel (35). Of course, knowledge of these first two distri-
bution functions above does not imply the process $v(.,.)$ is
completely statistically determined. It is noted in Dubes
(25) and Sage and Melsa (97) that in many physical situations
knowledge of the first two distribution functions is all that
is necessary for satisfactory performance or results.
Techniques are available as in Parzen (86) for implementing
knowledge of the first $n, n < \infty$, distribution functions as an
approximation to full knowledge for all n.

2.2 Continuous Stochastic Processes

Next is an illustrative example from Astrom (8) which will be used to classify various types of continuous stochastic processes. This example delineates those properties of continuous stochastic processes most useful in the analysis of a larger problem. A set of "reasonable" assumptions might be the following as a starting point. The real scalar continuous-time stochastic process $v(t,.), t\epsilon[t_o, t_f] \hat{=} T$ which is a continuous random variable for each fixed $t\epsilon T$ should be

1)   second order, i.e. have finite variance,

2)   continuous in the mean square sense $\forall t\epsilon T$,

3)   a process such that $v(t,.)$ is independent of

$\qquad v(\tau,.) \forall t \neq \tau \dot{\epsilon} T$,

and

4)   zero mean.

It is shown in Astrom (8) that the mean-square value of $v(.,.), \Phi_{vv}(t,t) = E\{x^2(t)\} = 0 \ \forall t\epsilon T$.

$\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ 2.2.1

If this process is the input (forcing function) of a linear ordinary differential equation, then the difference between the solution with zero input and the solution with this input is zero in the mean square. According to Astrom (8), this is not a "sensible" stochastic state model. Remarks in Papoulis (84) lead to the same conclusion.

In order to describe a "sensible" state model, one or more of assumptions 1) through 4) must be relaxed. It is well known that 4) can be relaxed with no loss of generality in most circumstances. However there does remain the accountability for non-zero means.

If assumption 1) is dropped and in addition stationarity imposed, the process would be basically a white noise process, i.e., the variance is infinite.

Relaxing assumption 2) leads to the time derivative of the stochastic process being undefined (in the mean-square sense). If this process was the input to a linear O.D.E., then the derivatives in the equation would also be undefined. The Ito calculus or the Stratonovich calculus allows analytic treatment of this type of process. Basically, an independent interval type process exists. For normal stationary transition probability, a Weiner process is formed and it is well known that Weiner processes have no defined time derivatives (in the mean-square sense).

Assumption 3) could be dropped and nonstationary processes with mean square sense time derivatives and finite continuous variances are obtained. Such processes are sometimes called second order processes.

2.3 Classification

It is desired to grossly classify continuous-time stochastic properties by which of assumptions 1) through 3)

are relaxed. Whether assumption 4) holds or not will not affect the generality of any class.

2.3.1 Class 1.

This is the class of processes with infinite variances. If the process is also stationary it is a white noise process.

2.3.2 Class 2.

This is the class of independent increment processes. The best known example is the Wèiner process for which the transition probability is Gaussian and stationary. No member of this class of processes is continuous (mean square) in time, hence does not have a defined time derivative.

2.3.3 Class 3.

This is the class of second order processes. Such processes are nonstationary with finite variances and continuous (mean square) derivatives.

2.3.4 Class 4.

This is the class of stochastic processes with correlation times very much less than the smallest time constants of the system which they enter. Such a process $v(.,.)$ has a variance given by

$$V_{vv}(t,\tau) = Q(t)\delta(t-\tau).$$

$$2.3.4.1$$

This class of processes is discussed in Bryson and Ho (20) and has been treated in stochastic control and estimation

problems. However, at t=τ, the variance is infinite. As in (20), the Dirac delta function is the limit of a pulse with amplitude 1/2 ε of duration 2ε. This pulse can be thought of as having large variance for a very short time. If 2ε is much less than the smallest time constant of the system, the time correlation of the process dies out in times of order of magnitude relevant to the system. Essentially this implies the nonstationarity property is not of significance and only one time axis is required to describe the statistics of the process. This may not be true in all physical situations. Therefore, this work will deal with class 3 processes and it is noted that approximations other than those with impulses of class 4 processes are included as members of class 3 processes.

2.4 Conclusions

This chapter provides the reasoning used to select the class of processes studied in the research. Continuous-time processes are selected since discrete-time analyses do not always allow the extension to the continuous-time case, whereas the reverse is possible. Class 3) stochastic processes are chosen for study since they have not been studied to date in control or identification problems. In fact this class of processes has only been the subject of a few papers most of which formulate but do not solve problems. The other three classes of processes have been studied extensively.

CHAPTER 3

<u>PROBLEM FORMULATION</u>

In this chapter the approach to the solution of the
stochastic optimal portion of the problem begins. The
identification portion of the problem starts in Chapter 6.
Throughout Chapters 3, 4, and 5 all elements of γ are
assumed to be known. Therefore the notational dependence of
all variables on the vector γ of parameters to be identified,
will be eliminated until Chapter 6.

For any optimal problem, four data are needed as dis-
cussed in Lee and Markus (65). The dynamic system equations,
the performance index, endpoint conditions, and classes of
admissible controls must be specified. In addition if the
dynamic system has stochastic inputs, some assumptions must
be made regarding the class of processes. Also a priori
assumptions on the initial values of the moments of the
stochastic processes are needed.

From these four data and the assumptions about the
stochastic processes, the stochastic optimal problem is
formulated. It is sought to derive necessary and sufficient
conditions for solutions to be optimal. Also if possible
the existence and uniqueness of solutions is to be established.
Further, in many problems closed loop (feedback) control laws

are to be found. In this work, the existence of solutions, is
assumed.

3.1 Stochastic Formulation

The four data for the classes of stochastic optimal con-
trol problems considered follow. This description covers a
wide range of physical situations. An actual physical process
is what is being described.

3.1.1 Dynamic System

The dynamic system considered is the linear time varying
system given by

$$\frac{dx(t,.)}{dt}=A(t)x(t,.)+B(t)u(t)+C(t)v(t,.) \qquad \text{3.1.1.1}$$

for $t \in [t_o, t_f] \triangleq \Gamma_1$

where

x is an nx1 state vector

A is an nxn matrix

u is an rx1 optimal control vector

B is an nxr matrix

v is an sx1 stochastic disturbance vector

and

C is an nxs matrix.

The stochastic process v is assumed to be a Class 3)
type process defined in Chapter 2, usually called a second
order process. No assumptions are made about its distribution
(or density) functions. It is assumed that for each fixed

$t \in \Gamma_1$, $v(t,.)$ is a continuous random variable, which implies
a continuous sample space.

3.1.2 Performance Index

The scalar quantitative measure of performance termed
the performance index is mathematically given by

$$J=J(u) = K(x(t_f,.),t_f)+\int_{t_o}^{t_f} f_{00}(x(t,.),u(t),v(t,.),t)dt$$

3.1.2.1

for dynamic optimization problems. The functional J con-
tains two terms, a terminal cost term K, and an integral
cost term of the function f_{00}. Most performance indices
for stochastic optimal problems do not explicitly show
dependence on the stochastic input v. The effect of this
disturbance is assumed to be wholly contained in its effect
on the state x. It is seen from 3.1.1.1 that x is a
stochastic process as well as v. It is a priori assumed
that u is a deterministic function of time. From 3.1.2.1,
J is a random variable. As such, it is not a "sensible"
quantitative measure. In many stochastic problems, the random
variable J is made deterministic and is given by

$$J_D=J_D(u) = E\{J\},$$

3.1.2.2

where $E\{.\}$ denotes expectation. Other means of making J a
deterministic measure exist in the literature. All of these
require expectation, but not in the same way as shown in
3.1.2.2. For example, the measure

$$J_{D1} = J_{D1}(u) = E\{[J-E(J)]^2\}$$

3.1.2.3

is discussed in articles by Sain and/or Liberty (98), (99), (100), and (101). Other related formulations are given in Murphy (81), and Rekasius (93). In the articles by Pugachev (89), and Andreev (3), (4), (5), and (6), the problem of synthesizing optimal systems for a wide variety of performance criteria is investigated.

Typical deterministic measures are minimum energy, fuel, and time. These are for deterministic problems equivalent to $v(t,.)=0 \forall t \in \Gamma_1$.

They are

$$J_{ME}(u) = \langle x(t_f), Q(t_f)x(t_f) \rangle + \int_{t_o}^{t_f} \{\langle x, Rx \rangle + \langle u, Su \rangle\}dt$$

3.1.2.4

$$J_{MF}(u) = \int_{t_o}^{t_f} \{\sum_{j=1}^{r} s_j |u_j|\}dt$$

3.1.2.5

$$J_{MT}(u) = \int_{t_o}^{t_f} dt = t_f - t_o$$

3.1.2.6

These three performance indices are most widely used. Even if the noise v is present the same type of index is still used, except that the $E\{J\}$ or a function of $E\{J\}$ is required for a "sensible" measure. A slight generalization for the minimum energy criteria J_{ME} would explicitly include the energy due to the stochastic disturbance v, given in a general form by 3.1.2.1. The minimum energy, or quadratic, form of 3.1.2.1 is

$$J_{SME}(u) = E\{J_S\} = E\{<x(t_f,.), Q(t_f)x(t_f,.)>$$

$$+ \int_{t_o}^{t_f} \{<x, R_1 x> + <u, R_3 u> + <v, R_4 v>\}dt\}$$

$$3.1.2.7$$

In this case, minimum energy control of state is desired in the presence of minimum disturbance energy, where $J_{SME}(u)$ is a more general stochastic minimum energy criteria. Such criteria have indirectly been studied in maximum signal-to-noise ratio problems for stationary Gaussian white noise processes. A modern formulation of this problem is found in Holtzman (37) and Athans and Schweppe (12).

3.1.3 Endpoint Conditions

Normally in deterministic optimal problems, only the initial value of the state $x(t_o)$, the initial time t_o, the final state $x(t_f)$, and the finite final time t_f are required as known or as to be determined if not specified. These data are present in the transversality conditions when variational techniques are used.

When stochastic inputs are present, much more a priori information is required such as the mean value and variance of the initial state $x(t_o)$, and the correlation between the stochastic disturbance and the initial state. These data are explicitly required. For many problems, the distribution function (or density function) of the noise v and the state x must also be specified.

3.1.4 Classes of Admissible Controls

The last data needed to formulate a deterministic optimal control problem is the set U, in which the optimal control u belongs. Whether the set U is open or closed puts bounds or no bounds on the values u can have. Most minimum energy problems require U to be an open set and minimum fuel or time problems require U to be a closed set, i.e., u is bounded above and below. The physical application determines which condition is present in a specific example.

3.2 A Transformation

For stochastic optimal control problems, the four data of section 3.1 acquire characteristics which add to the degree of difficulty in completion of the solution as compared to deterministic optimal control problems. Basically the only stochastic optimal problems solved to date are linear ones for which the separation theorem holds. This restriction requires that a Kalman estimate be obtained in addition to the four data. In this work an alternate approach is given. As a first step, a transformation of the four data is made using mainly the expectation operator. This transformation changes the problem from stochastic to deterministic. The form of the data after transformation suggests a differential game approach as will be discussed in Section 3.3.

3.2.1 Dynamic System

The stochastic dynamic system governed by the state equation 3.1.1.1, is transformed directly by taking the expectations of both sides resulting in

$$\frac{d}{dt}\mu_x(t)=A(t)\mu_x(t)+B(t)u(t)+C(t)\mu_v(t)$$

for all $t\in T_1$

$$3.2.1.1$$

where

$$E\{x(t,.)\}\triangleq\mu_x(t)$$

$$3.2.1.2$$

and

$$E\{v(t,.)\}\triangleq\mu_v(t)$$

$$3.2.1.3$$

The mean μ_x, of the state x and the mean μ_v of the stochastic input v are deterministic functions of time. The control u is assumed a priori to be a deterministic function of time. The mean μ_x, is the first moment of the state x, and similarly for μ_v. As mentioned in Chapter 1, the effects of the stochastic disturbance v will be approximated by its first and second moments. The first moment propagates through the system as in equation 3.2.1.1. The complete description of the propagation of the second moment of v on the state x is given by the two equations,

$$\frac{\partial V}{\partial \tau}_{vx}(t,\tau) = A(\tau)V_{vx}(t,\tau) + C(\tau)V_{vv}(t,\tau)$$

$$\text{for } (t,\tau) \in \Gamma_1 x \Gamma_1 \hat{=} \Gamma \qquad\qquad 3.2.1.4$$

and

$$\frac{\partial}{\partial t} V_{xx}(t,\tau) = A(t)V_{xx}(t,\tau) + C(t)V_{vx}(t,\tau)$$

$$\text{for } (t,\tau) \in \Gamma \qquad\qquad 3.2.1.5$$

where

$$E\{[v(t,.) - E\{v(t,.)\}][v(\tau,.) - E\{v(\tau,.)\}]'\} \hat{=} V_{vv}(t,\tau)$$

$$\text{for } (t,\tau) \in \Gamma, \qquad\qquad 3.2.1.6$$

$$E\{[v(t,.) - E\{v(t,.)\}][x(\tau,.) - E\{x(\tau,.)\}]'\} \hat{=} V_{vx}(t,\tau)$$

$$\text{for } (t,\tau) \in \Gamma, \qquad\qquad 3.2.1.7$$

$$E\{[x(t,.) - E\{x(t,.)\}][x(\tau,.) - E\{x(\tau,.)\}]'\} \hat{=} V_{xx}(t,\tau)$$

$$\text{for } (t,\tau) \in \Gamma, \qquad\qquad 3.2.1.8$$

The covariance of v,

$$V_{vv}(t,\tau) \text{ is an sxs matrix}$$

and

$$V_{vv}(t,\tau) = \Phi_{vv}(t,\tau) - \mu_v(t)\mu_v'(t)$$

$$\text{for } (t,\tau) \in \Gamma \qquad\qquad 3.2.1.9$$

where

$$E\{v(t,.)v'(\tau,.)\} \hat{=} \Phi_{vv}(t,\tau)$$

$$\text{for } (t,\tau) \in \Gamma \qquad\qquad 3.2.1.10$$

and $\Phi_{vv}(t,\tau)$ is the sxs matrix called the autocorrelation of v.

The cross-covariance of v and x,

$V_{vx}(t,\tau)$ is an sxn matrix

and

$$V_{vx}(t,\tau)=\Phi_{vx}(t,\tau)-\mu_v(t)\mu_x^{'}(\tau)$$

$\quad$ for $(t,\tau)\epsilon\Gamma$

$\hspace{10cm}$ 3.2.1.11

where

$$E\{v(t,.)x^{'}(\tau,.)\}\triangleq\Phi_{vx}(t,\tau)$$

$\quad$ for $(t,\tau)\epsilon\Gamma$

$\hspace{10cm}$ 3.2.1.12

and $\Phi_{vx}(t,\tau)$ is the sxn matrix called the cross correlation of v and x.

And the covariance of x,

$V_{xx}(t,\tau)$ is an nxn matrix

and

$$V_{xx}(t,\tau)=\Phi_{xx}(t,\tau)-\mu_x(t)\mu_x^{'}(\tau)$$

$\quad$ for $(t,\tau)\epsilon\Gamma$

$\hspace{10cm}$ 3.2.1.13

where

$$E\{x(t,.)x^{'}(\tau,.)\}\triangleq\Phi_{xx}(t,\tau)$$

$\quad$ for $(t,\tau)\epsilon\Gamma$

$\hspace{10cm}$ 3.2.1.14

and $\Phi_{xx}(t,\tau)$ is the nxn matrix called the autocorrelation of x.

The following moments also appear and are defined below.

The variance of v

$V_{vv}(t,t)$ is an sxs matrix

and

$$V_{vv}(t,t) = \Phi_{vv}(t,t) - \mu_v(t)\mu_v^T(t)$$

$$\text{for } t \in \Gamma_1$$

where

$$E\{v(t,.)v^T(t,.)\} \triangleq \Phi_{vv}(t,t)$$

$$\text{for } t \in \Gamma_1$$

and $\Phi_{vv}(t,t)$ is the sxs matrix called the mean square value of v.

The cross variance of v and x

$V_{vx}(t,t)$ is an sxn matrix

and

$$V_{vx}(t,t) = \Phi_{vx}(t,t) - \mu_v(t)\mu_x^T(t)$$

$$\text{for } t \in \Gamma_1$$

where

$$E\{v(t,.)x^T(t,.)\} \triangleq \Phi_{vx}(t,t)$$

$$\text{for } t \in \Gamma_1$$

and $\Phi_{vx}(t,t)$ is the sxn matrix called the mean cross value of v and x.

The variance of x

$V_{xx}(t,t)$ is an nxn matrix

and

$$V_{xx}(t,t) = \Phi_{xx}(t,t) - \mu_x(t)\mu_x^T(t)$$

$$\text{for } t \in \Gamma_1$$

where

$$E\{x(t,.)x^{T}(t,.)\} \triangleq \Phi_{xx}(t,t)$$

$$\text{for } t \in \Gamma_1$$

and

$\Phi_{xx}(t,t)$ is the nxn matrix called the mean square value of x.

The detailed derivation of equations 3.2.1.1, 3.2.1.4, and 3.2.1.5 is contained in Appendix A. The derivation uses material from Papoulis (84), Lebedev (63), Pugachev (89), and Sage and Melsa (97). These equations describe how the first and second moments of a Class 3) stochastic process propagate through a linear system. Similar equations could be obtained for any other class of stochastic processes.

3.2.2 Performance Index

The transformation of performance indices was first performed in estimation problems. Typically, minimum variance criteria resulted for criteria quadratic in the state x. The major portion of criteria after transformation used in the literature depend on criteria which were originally quadratic. For example, if x is a stochastic process, then from 3.1.2.4

$$E\{J_{ME}\} = E\{<x(t_f,.),Qx(t_f,.)>$$

$$+\int_{t_o}^{t_f}\{<x,Rx>+<u,Su>\}dt\}$$

$$=\text{tr}[Q\Phi_{xx}(t_f,t_f)]+\int_{t_o}^{t_f}\{\text{tr}[R\Phi_{xx}(t,t)]+<u,Ru>\}dt$$

$$= \text{tr}[\, QV_{xx}(t_f,t_f) - Q\mu_x(t_f)\mu_x^{T}(t_f)\,]$$

$$+ \int_{t_o}^{t_f}\{\text{tr}\,[RV_{xx}(t,t) - R\mu_x(t)\mu_x^{T}(t)] + <u,Ru>\}dt$$

$$= \text{tr}[\, QV_{xx}(t_f,t_f)\,] - <\mu_x(t_f),Q\mu_x(t_f)>$$

$$+ \int_{t_o}^{t_f}\{\text{tr}\,[RV_{xx}(t,t)] - <\mu_x(t),R\mu_x(t)> + <u,Ru>\}dt.$$

3.2.2.1

Similarly, it can be shown that

$$E\{J_s\} = \text{tr}\,[Q(t_f)V_{xx}(t_f,t_f)] - <\mu_x(t_f),Q(t_f)\mu_x(t_f)>$$

$$+ \int_{t_o}^{t_f}\{\text{tr}\,[R_1V_{xx}(t,t)] - <\mu_x(t),R_1\mu_x(t)> + <u,R_3u>$$

$$+ \text{tr}[\,R_4V_{vv}(t,t)] - <\mu_v(t),R_4\mu_v(t)>\}dt$$

3.2.2.2

The other types of performance indices mentioned in the references of Section 3.1.2 are one of the following:

$$E\{J_{ME}^2\}, E\{[\,J_{ME}-E\{J_{ME}\}]^2\}, [E\,\{J_{ME}\}]^2$$

or for J_s,

$$E\{J_s^2\}, E\{[J_s-E\{J_s\}]^2\}, [E\{J_s\}]^2.$$

In some of those references, the equations are worked out, and they contain the first and second moments of v and x as variables, as do equations 3.2.2.1 and 3.2.2.2. Slight generalizations of these types of criteria are the ones chosen for this research. The transformations of these particular indices is contained in Appendix B. Four different criteria are used to

derive necessary conditions in Chapter 4. The physical
significance of these four criteria is also detailed in
Appendix B.

3.2.3 Endpoint Conditions

From the equations of the dynamic system and the per-
formance indices, it will be seen that a modification in the
form of the endpoint conditions is necessary. The fixed
endpoint case is used in Chapter 4 derivations of necessary
conditions, as well as fixed finite final time. Other kinds
of endpoint conditions, such as free final state and free
final time are left for future research.

3.2.4 Classes of Admissible Controls

In Section 3.3, the moments of v will be designated as
controls. Therefore, these moments must be members of a set
which is either open or closed. For Class 3) stochastic pro-
cesses there are no finite bounds on either the mean or the
covariance of v which exist due to the theory of stochastic
processes, so it is assumed that they belong to open sets.

3.3 Differential Game Approach

In this section the variables of equations 3.2.1.1,
3.2.1.4, and 3.2.1.5 are renamed for ease of notation and for
clarity in the formulation of the differential game. Then
the equations of the dynamic system, the performance index,
the endpoint conditions, and the classes of admissible controls

are specified in full such that necessary and sufficient
conditions can be obtained.

3.3.1 Formulation

Equation 3.2.1.1 is changed to

$$\frac{dz_1}{dt}(t) = A(t) z_1(t) + B(t) u_1(t) + C(t) u_2(t)$$

for all $t \in \Gamma_1$.

3.3.1.1

where

$$\mu_x(t) \hat{=} z_1(t) \quad \text{an nx1 vector,}$$

$$u(t) \hat{=} u_1(t) \quad \text{an rx1 vector,}$$

and

$$\mu_v(t) \hat{=} u_2(t) \quad \text{an sx1 vector.}$$

Equation 3.2.1.4 becomes

$$\frac{\partial}{\partial \tau} z_2(t,\tau) = A(\tau) z_2(t,\tau) + C(\tau) u_3(t,\tau)$$

for all $(t,\tau) \in \Gamma$

3.3.1.2

where

$$V_{vx}(t,\tau) \hat{=} z_2(t,\tau) \quad \text{an nxs matrix}$$

and

$$V_{vv}(t,\tau) \hat{=} u_3(t,\tau) \quad \text{an sxs matrix.}$$

Equation 3.2.1.5 becomes

$$\frac{\partial}{\partial t} z_3(t,\tau) = A(t) z_3(t,\tau) + C(t) z_2(t,\tau)$$

for all $(t,\tau) \in \Gamma$

3.3.1.3

where

$$V_{xx}(t,\tau) \hat{=} z_3(t,\tau) \quad \text{an nxn matrix.}$$

The state of the system becomes z_1, z_2, z_3 with the "man" chosen controls u_1 and the opposing "nature" controls u_2, u_3.

A general performance index of the form

$$J = J(u_1, u_2, u_3) = \int_{t_0}^{t_f} \int_{t_0}^{t_f} f_2(z_1, z_2, z_3, u_1, u_2, u_3) \, dt \, d\tau$$

$$+ \int_{t_0}^{t_f} f_1(z_1, u_1, u_2) \, dt \qquad\qquad 3.3.1.4$$

will be employed. The "man" controls u_1 seek to minimize J while the "nature" controls u_2, u_3 seek to maximize J.

The initial and final times are fixed. The initial state is specified and given by

$$z_1(t_0) = \pi_{10}$$

$$z_2(t, t_0) = \pi_{20}(t)$$

$$z_3(t_0, \tau) = \pi_{30}(\tau). \qquad\qquad 3.3.1.5$$

The terminal state is also specified and given as

$$z_1(t_f) = \pi_{11}$$

$$z_2(t, t_f) = \pi_{21}(t)$$

$$z_3(t_f, \tau) = \pi_{31}(t). \qquad\qquad 3.3.1.6$$

Finally it is assumed that $u_1 \in U_1 = E^r$, i.e. u_1 takes values in r dimensional Euclidean space which is an open set. Similarly $u_2 \in U_2 = E^s$ and $u_3 \in U_3 = E^s \times E^s$.

3.3.2 Differential Game Theory

At this point some remarks about deterministic differential games are necessary as background for the next section. In the description of the four data for a one-sided optimal control problem given in Section 3.1 and 3.2 the dynamic system and performance index change considerably. The discussion of endpoint conditions and classes of admissible controls remains the same except more variables are present.

A two-person zero-sum differential game is a two-sided optimal control problem. The two sides are exemplified as two sets of controls u_1 and u_2, which drive dynamic systems with goals that are in conflict with each other. Mathematically, the dynamic system is

$$\frac{dx}{dt}(t) = A(t)x(t) + B(t)u_1(t) + C(t)u_2(t), t \in [t_o, t_f]$$

3.3.2.1

for a class of linear differential games. Equations 3.3.2.1 is sometimes obtained by combining the equations for two systems, one the "man" system given by

$$\frac{dx_m}{dt}(t) = A_m(t)x_m(t) + B_m(t)u_1(t), t \in [t_o, t_f]$$

3.3.2.2

and the other the "nature" system given by

$$\frac{dx_n}{dt}(t) = A_n(t)x_n(t) + C_n(t)u_2(t), t \in [t_o, t_f].$$

3.3.2.3

The performance index would include the effects of both controls, and for system 3.3.2.1 would functionally be

$$J_{DG} = J_{DG}(u_1, u_2) = K(x(t_f), t_f) + \int_{t_o}^{t_f} f(x, u_1, u_2) dt$$

3.3.2.4

a minimum energy type criteria is

$$J_{ME} = J_{ME}(u_1, u_2) = \langle x(t_f), R_1 x(t_f) \rangle + \int_{t_o}^{t_f} \{ \langle x, R_2 x \rangle + \langle u_1, R_3 u_1 \rangle$$

$$- \langle u_2, R_4 u_2 \rangle \} dt \qquad 3.3.2.5$$

The conflict is present since u_1 seeks to minimize J_{ME} while u_2 seeks to maximize J_{ME}, i.e. it is sought to simultaneously establish

$$\min_{u_1 \in U_1} \max_{u_2 \in U_2} J_{ME}(u_1, u_2) \qquad 3.3.2.6$$

where U_1 and U_2 are classes of admissible controls. The solution is sought by u_1 in the knowledge that u_2 seeks to maximize J_{ME} and vice versa. Therefore the optimal point is

$$J_{ME}(x, u_1^*, u_2) \leq J_{ME}(x^*, u_1^*, u_2^*) \leq J_{ME}(x, u_1, u_2^*) \qquad 3.3.2.7$$

which is the definition of a saddle point. From 3.3.2.7 it is seen that u_1 seeks to minimize J_{ME} knowing u_2^* is maximizing J_{ME} and vice versa.

These key points are needed to formulate the class of differential game problems in the next section.

3.3.3 Statement of the Problem

It is desired to find the pair of controls $(u_1; u_2, u_3)$ which establish a saddle point for the performance index J of equation 3.3.1.4 in the sense that

$$\min_{\substack{u_1 \in U_1}} \quad \max_{\substack{u_2 \in U_2 \\ u_3 \in U_3}} \quad J(u_1, u_2, u_3, t_o, t_f)$$

occurs, while transferring the initial state $z_1(t_o), z_2(t, t_o),$ $z_3(t_o, \tau)$ to fixed terminal state $z_1(t_f), z_2(t, t_f) z_3(t_f, \tau)$ in fixed finite time $t_f - t_o$ subject to the constraints given by 3.3.1.1, 3.3.1.2 and 3.3.1.3. Necessary conditions for four different specific problems are given in Chapter 4.

3.4 Conclusions

A stochastic optimal control problem for dynamic systems with Class 3) stochastic disturbances was given. The nature of the problem was transformed to deterministic using mainly the expectation operator. The effects of the stochastic disturbance after this transformation are assumed to be contained in the propagation of the first two moments through the dynamic system. The degrading effect of these moments is in conflict with the efforts to optimally control the dynamic system. This conflict situation was formally cast into a differential game. The various data needed to mathematically describe the differential game were specified and finally the statement of the problem in this context was given.

CHAPTER 4

NECESSARY CONDITIONS

Four different classes of problems are considered in this
chapter. For the first three classes, the dynamic system remains
the same but the performance index changes. For the fourth class of
problems the dynamic system also changes. Then sets of necessary
conditions for each problem are given.

4.1 Problem Statements

The four data required to define each class of problems are
specified in this section.

4.1.1 Problem 1.

It is desired to find the pair of controls $(u_1; u_2, u_3')$ which
establish a saddle point for the performance index

$$J(u_1, u_2, u_3') = -\frac{1}{2} \int_{t_0}^{t_f}\int_{t_0}^{t_f} \{ tr[R_1 z_3(t,\tau) z_3'(t,\tau) R_1'] +$$

$$tr[R_2 u_3(t,\tau) u_3'(t,\tau) R_2'] \} dt d\tau + \frac{1}{2} \int_{t_0}^{t_f} \{ <u_1(t), R_3 u_1(t)> - <u_2(t), R_4 u_2(t)> \} dt$$

where R_1 is assumed non-negative definite and symmetric and R_2, R_3,   4.1.1.1
and R_4 are assumed to be positive definite and symmetric, in the
sense that

$$\min_{\substack{u \in U_1 \\ u_1 \in U_1}} \max_{\substack{u_2 \in U_2 \\ u_3' \in U_3}} J(u_1, u_2, u_3')$$ 4.1.1.2

occurs, while transferring the specified initial state   4.1.1.3

$$z_1(t_0) = \pi_{10}$$

$$z_2(t,t_0) = \pi_{20}(t)$$

$$z_3(t_0,\tau) = \pi_{30}(\tau)$$

to the specified final state   4.1.1.4

$$z_1(t_f) = \pi_{11}$$

$$z_2'(t,t_f) = \pi_{21}(t)$$

$$z_3(t_f,\tau) = \pi_{31}(\tau)$$

in fixed finite time $t_f - t_0$, where the dynamic system is governed

by

$$\frac{dz_1}{dt}(t) = A(t)z_1(t) + B(t)u_1(t) + C(t)u_2(t), \forall t \in \Gamma_1 \qquad 4.1.1.5$$

$$\frac{\partial z_2'}{\partial \tau}(t,\tau) = A(\tau)z_2'(t,\tau) + C(\tau)u_3'(t,\tau), \forall (t,\tau) \in \Gamma \qquad 4.1.1.6$$

$$\frac{\partial z_3}{\partial t}(t,\tau) = A(t)z_3(t,\tau) + C(t)Z_2(t,\tau), \forall (t,\tau) \in \Gamma \qquad 4.1.1.7$$

and the admissible classes of controls are the open sets

$$u_1 \in U_1 = E^r$$

$$u_2 \in U_2 = E^s$$

$$u_3' \in U_3 = E^s \times E^s \qquad\qquad\qquad\qquad\qquad 4.1.1.8$$

4.1.2 Problem 2.

Problem 2 involves the functional performance index

$$J(u_1,u_2,u_3') = \int_{t_0}^{t_f}\int_{t_0}^{t_f} f_0(z_1(t),z_1(\tau),z_2'(t,\tau),z_3(t,\tau),u_1(t),u_1(\tau),u_2(t),u_2(\tau),$$

$$u_3'(t,\tau))dtd\tau \qquad 4.1.2.1$$

instead of 4.1.1.1, all other data remaining the same.

4.1.3 Problem 3.

Problem 3 involves the specific performance index

$$J(u_1, u_2, u_3') = \int_{t_o}^{t_f}\int_{t_o}^{t_f}\{f_{oo}(z_1'(t), z_1''(t), z_2'(t,\tau), z_3(t,\tau), u_1(t), u_1(\tau), u_2(t), u_2(\tau))$$
$$+ \sum_{k=1}^{s}\sum_{m=1}^{s} f_{km}(z_2', z_3)u_{3km}' \} \, dt d\tau$$

instead of 4.1.1.1, all other data remaining the same.   4.1.3.1

4.1.4 Problem 4

Consider a dynamic system described by

$$\frac{dz_1(t)}{dt} = A(t)z_1(t) + B(t)u_1(t) + C(t)u_2(t)$$ 4.1.4.1

$$\frac{dz_2(t)}{dt} = F(t)z_2(t) + z_2(t)A'(t) + u_3(t)C'(t)$$ 4.1.4.2

$$\frac{dz_3(t)}{dt} = A(t)z_3(t) + z_3(t)A'(t) + C(t)z_2(t) + z_2'(t)C'(t)$$ 4.1.4.3

with the functional

$$J(u_1, u_2, u_3) = \int_{t_o}^{t_f} tr \{\frac{d_1}{2} R_1(t)u_1(t)u_1'(t)$$

$$+ \frac{d_2}{4} R_2(t)[u_3(t) + u_2(t)u_2'(t)]^2 \}dt$$ 4.1.4.4

Given the state equations 4.1.4.1 through 4.1.4.3 and the
performance index 4.1.4.4, find the optimal strategy
$u_i \in U_i$, i = 1,2,3 such that

$$J(u_1^*, u_2, u_3) \leq J(u_1^*, u_2^*, u_3^*) \leq J(u_1, u_2^*, u_3^*)$$ 4.1.4.5

for all $u_i \in U_i$, i = 1,2,3

where

$$z_1(t_0) = z_{10}$$

$$z_2(t_0) = z_{20}$$

$$z_3(t_0) = z_{30} \qquad\qquad 4.1.4.6$$

are specified

and

$$d_1 + d_2 = 1$$

$$d_1 \geq 0$$

$$d_2 \geq 0 \qquad\qquad 4.1.4.7$$

4.2 Statement of Necessary Conditions

4.2.1 Theorem 4.1

In order that the pair of controls $(u_1; u_2, u_3')$ be extremal for Problem 1 of Section 4.1.1, it is necessary that there exists non-zero continuous functions $\lambda_1(t), t \epsilon \Gamma_1$, and $\lambda_2(t,\tau), \lambda_3(t,\tau)$, $(t,\tau) \in \Gamma$ which are solutions of

$$\frac{d\lambda^*(t)}{dt^1} = - A'(t)\lambda_1^*(t), \; \forall t \epsilon \Gamma_1 \qquad\qquad 4.2.1.1$$

$$\frac{\partial \lambda^*(t,\tau)}{\partial \tau^2} = - A'(\tau)\lambda_2^*(t,\tau) - \lambda_3^*(t,\tau)C(\tau), \; \forall (t,\tau) \in \Gamma \qquad\qquad 4.2.1.2$$

$$\frac{\partial \lambda^*(t,\tau)}{\partial \tau^3} = - A'(t)\lambda_3^*(t,\tau) - R_1'R_1 z_3^*(t,\tau), \; \forall (t,\tau) \in \Gamma \qquad\qquad 4.2.1.3$$

through which the trajectories governed by

$$\frac{dz}{dt}_1^*(t) = A(t)z_1^*(t) + B(t)u_1^*(t) + C(t)u_2^*(t), \ \forall t \in \Gamma_1 \qquad \text{4.2.1.4}$$

$$\frac{\partial z}{\partial \tau}_2^{\prime*}(t,\tau) = A(\tau)z_2^{\prime*}(t,\tau) + C(\tau)u_3^{\prime*}(t,\tau), \ \forall (t,\tau) \in \Gamma \qquad \text{4.2.1.5}$$

$$\frac{\partial z}{\partial t}_3^*(t,\tau) = A(t)z_3^*(t,\tau) + C(t)z_2^*(t,\tau), \ \forall (t,\tau) \in \Gamma \qquad \text{4.2.1.6}$$

are transferred from

$$z_1(t_o) = \pi_{10}$$

$$z_2^{\prime}(t,t_o) = \pi_{20}(t)$$

$$z_3(t_o,\tau) = \pi_{30}(\tau) \qquad \text{4.2.1.7}$$

to

$$z_1^{\prime}(t_f) = \pi_{11}$$

$$z_2^{\prime}(t,t_f) = \pi_{21}(t)$$

$$z_3(t_f,\tau) = \pi_{31}(\tau) \qquad \text{4.2.1.8}$$

in fixed finite time $t_f - t_o$, by the controls in the admissible

classes

$$u_1^*\epsilon U_1 = E^r$$

$$u_2^*\epsilon U_2 = E^s$$

$$u_3^{\prime*}\epsilon U_3 = E^s \ X \ E^s \qquad \text{4.2.1.9}$$

which must satisfy

$$u_1*(t) = -R_3^{-1} B'(t) \lambda_1*(t), \forall t \in \Gamma_1$$ \hfill 4.2.1.10

$$u_2*(t) = R_4^{-1} C'(t) \lambda_1*(t), \forall t \in \Gamma_1$$ \hfill 4.2.1.11

$$u_3'*(t,\tau) = C'(\tau)\lambda_2*(t,\tau)R_2^{-1}(R_2^{-1})', \forall (t,\tau) \in \Gamma$$ \hfill 4.2.1.12

where ( )* denotes extremal.

Remark:

After substitution of equations 4.2.1.10, 4.2.1.11 and

4.2.1.12 into 4.2.1.4 and 4.2.1.5, there are $2n(1 + s + n)$

differential equations with $n(1 + s + n)$ initial conditions and

$n(1 + s + n)$ final conditions, forming a two point boundary value

problem (TPBVP) with $2n(1 + s + n)$ differential equations and

$2n(1 + s + n)$ endpoint conditions. In theory then, it is possible

to solve for the extremal controls and trajectories from the set

of necessary conditions of Theorem 4.1.

4.2.2 Theorem 4.2

In order that the pair of controls $(u_1; u_2, u_3^{\prime})$ be extremal for Problem 2 of Section 4.1.2, it is necessary that there exists non-zero continuous functions $\lambda_1(t)$, $t \epsilon \Gamma_1$, and $\lambda_2(t,\tau), \lambda_3(t,\tau)$, $(t,\tau) \epsilon \Gamma$ which are solutions of

$$\int_{t_0}^{t_f} \{(\frac{\partial fo}{\partial z_1(t)})^* + A^{\prime}(t)\lambda_1^*(t) + \frac{d\lambda_1}{dt}^*(t)\} d\tau = 0, \quad \int_{t_0}^{t_f} \{(\frac{\partial fo}{\partial z_1(\tau)})^*\} dt = 0 \qquad 4.2.2.1$$

$$\frac{\partial \lambda_2}{\partial \tau}^*(t,\tau) = -A^{\prime}(\tau)\lambda_2^*(t,\tau) - \lambda_3^{\prime}*(t,\tau)C(\tau) - (\frac{\partial fo}{\partial z_2})_* \qquad 4.2.2.2$$

$$\frac{\partial \lambda_3}{\partial t}^*(t,\tau) = -A^{\prime}(\tau)\lambda_3^*(t,\tau) - (\frac{\partial fo}{\partial z_3})_* \qquad 4.2.2.3$$

through which the trajectories governed by

$$\frac{dz_1}{dt}^*(t) = A(t)z_1^*(t) + B(t)u_1^*(t) + C(t)u_2^*(t), \quad t \epsilon \Gamma_1 \qquad 4.2.2.4$$

$$\frac{\partial z_2^{\prime}}{\partial \tau}^*(t,\tau) = A(\tau)z_2^{\prime}*(t,\tau) + C(\tau)u_3^{\prime}*(t,\tau), \quad (t,\tau) \epsilon \Gamma \qquad 4.2.2.5$$

$$\frac{\partial z_3}{\partial t}^*(t,\tau) = A(t)z_3^*(t,\tau) + C(t)z_2^*(t,\tau), \quad (t,\tau) \epsilon \Gamma \qquad 4.2.2.6$$

are transferred from

$$z_1(t_0) = \pi_{10}$$

$$z_2^{\prime}(t,t_0) = \pi_{20}(t)$$

$$z_3(t_0,\tau) = \pi_{30}(\tau) \qquad 4.2.2.7$$

to

$$z_1(t_f) = \pi_{11}$$

$$z_2^{\prime}(t,t_f) = \pi_{21}(t)$$

$$z_3(t_f,\tau) = \pi_{31}(\tau) \qquad 4.2.2.8$$

in fixed finite time $t_f - t_o$, by the controls in the admissible classes

$u_1^* \epsilon U_1 = E^r$

$u_2^* \epsilon U_2 = E^S$

$u_3^* \epsilon U_3 = E^S x E^S$ 4.2.2.9

which must satisfy

$$\int_{t_o}^{t_f} \{ (\frac{\partial fo}{\partial u_1 (t)})_* + B'(t) \lambda_1^*(t) \} d\tau = 0, \quad \int_{t_o}^{t_f} \{ (\frac{\partial fo}{\partial u_1 (\tau)})_* \} dt = 0 \qquad 4.2.2.10$$

$$\int_{t_o}^{t_f} \{ (\frac{\partial fo}{\partial u_2 (\tau)})_* + C'(t) \lambda_1^*(t) \} d\tau = 0, \quad \int_{t_o}^{t_f} \{ (\frac{\partial fo}{\partial u_2 (\tau)})_* \ dt = 0 \qquad 4.2.2.11$$

$$(\frac{\partial fo}{\partial u_3})_* + C'(\tau) \lambda_2'^*(t, \tau) = 0 \qquad 4.2.2.12$$

where ( )* denotes extremal.

4.2.3 Theorem 4.3

In order that the pair of controls $(u_1; u_2, u_3')$ be extremal

for Problem 3 of Section 4.2.3, it is necessary that in addition to

Theorem 4.2, the following conditions

$$\frac{\partial}{\partial \tau} \frac{\partial H}{\partial u_3'} = 0 \qquad\qquad 4.2.3.1$$

$$\| \frac{\partial}{\partial u_3'} \{ \| \frac{\partial}{\partial \tau} \frac{\partial H}{\partial u_3'} \| \} \| = 0 \qquad\qquad 4.2.3.2$$

$$\| \frac{\partial}{\partial u_3'} \{ \| \frac{\partial^2}{\partial \tau^2} \frac{\partial H}{\partial u_3'} \| \} \| \geq 0 \qquad\qquad 4.2.3.3$$

hold where

$$H \triangleq f_{oo}(z_1(t), z_1(\tau), z_2'(t,\tau), z_3(t,\tau), u_1(t), u_1(\tau), u_2(t), u_2(\tau))$$

$$+ \sum_{k=1}^{s} \sum_{m=1}^{s} f_{km}(z_2'(t,\tau), z_3(t,\tau)) u_{3_{km}}'(t,\tau)$$

$$+ <\lambda_1(t), A(t)z_1(t) + B(t)u_1(t) + C(t)u_2(t)>$$

$$+tr [[A(\tau)z_2'(t,\tau) + C(\tau)u_3'(t,\tau)][\lambda_2'(t,\tau)]]$$

$$+tr [[A(t)z_3(t,\tau) + C(t)z_2(t,\tau)][\lambda_3'(t,\tau)]] \qquad 4.2.3.4$$

and where for G an $n \times m$ matrix

$$\| G \| \triangleq \sum_{i=1}^{n} \sum_{j=1}^{m} g_{ij} . \qquad\qquad 4.2.3.5$$

Remark:

The performance index for this class of problems is nonlinear

in z_1, z_2', z_3, u_1 and u_2 but linear in u_3' . Therefore, the optimal

controls u_1 and u_2 are found as nonsingular controls but $u_3^{'}$ is a singular control. The equations 4.2.3.1 and 4.2.3.2 allow the solution for $u_3^{'*}$ to be carried out. Equation 4.2.3.3 represents a strengthened necessary condition similar to the Legendre Clebsch condition of nonsingular problems. For this class of problems u_1, u_2 and $u_3^{'}$ are not bounded, i.e. they take values in open sets.

4.2.4. Theorem 4.4

In order that $u_i^*(t) \in U_i, t \in \Gamma_1, i = 1,2,3$ be the optimal strategies for Problem 4 of Section 4.1.4, it is necessary that there exist a nonzero vector function $\lambda_1^*(t)$ and nonzero matrix functions $\lambda_2^*(t)$ and $\lambda_3^*(t)$ such that:

a) $\lambda_1^*(t), \lambda_2^*(t), \lambda_3^*(t), z_1^*(t), z_2^*(t), z_3^*(t)$ are solutions of

$$\frac{dz_1^*(t)}{dt} = \left(\frac{\partial H}{\partial \lambda_1}\right)_* = A(t)z_1^*(t) + B(t)u_1^*(t) + C(t)u_2^*(t) \qquad 4.2.4.1$$

$$\frac{dz_2^*(t)}{dt} = \left(\frac{\partial H}{\partial \lambda_2}\right)_* = F(t)z_2^*(t) + z_2^*(t)A'(t) + u_3^*(t)C'(t) \qquad 4.2.4.2$$

$$\frac{dz_3^*(t)}{dt} = \left(\frac{\partial H}{\partial \lambda_3}\right)_* = A(t)z_3^*(t) + z_3^*(t)A'(t) + C(t)z_2^*(t)$$

$$+ z_2'^*(t)C'(t) \qquad 4.2.4.3$$

$$\frac{d\lambda_1^*(t)}{dt} = - \left(\frac{\partial H}{\partial z_1}\right)_* = - A'(t)\lambda_1^*(t) \qquad 4.2.4.4$$

$$\frac{d\lambda_2^*(t)}{dt} = - \left(\frac{\partial H}{\partial z_2}\right)_* = - [F'(t)\lambda_2^*(t) + \lambda_2^*(t)A(t) + C'(t)\lambda_3^*(t)$$

$$+ C'(t)\lambda_3'^*(t)] \qquad 4.2.4.5$$

$$\frac{d\lambda_3^*(t)}{dt} = - \left(\frac{\partial H}{\partial z_3}\right) = - [A'(t)\lambda_3^*(t) + \lambda_3^*(t)A(t)] \qquad 4.2.4.6$$

with the boundary conditions

$$z_1^*(t_o) = z_{10}$$

$$z_2^*(t_o) = z_{20}$$

$$z_3^*(t_o) = z_{30} \qquad 4.2.4.7$$

and

$$\lambda_1^*(t_f) = 0$$

$$\lambda_2^*(t_f) = 0$$

$$\lambda_3^*(t_f) = 0 \qquad 4.2.4.8$$

b) The extremal strategy $(u_1*(t),u_2*(t),u_3*(t))$ satisfies the following min-max principle

$H(z_1*(t),z_2*(t),z_3*(t),u_1*(t),u_2(t),u_3(t),\lambda_1*(t),\lambda_2*(t),\lambda_3*(t))$

$\leq H(z_1*(t),\ z_2*(t),\ z_3*(t),u_1*(t),u_2*(t),u_3*(t),\lambda_1*(t),\lambda_2*(t),\lambda_3*(t))$

$\leq H(z_1*(t),z_2*(t),z_3*(t),u_1(t),u_2*(t),u_3*(t),\lambda_1*(t),\lambda_2*(t),\lambda_3*(t))$

4.2.4.9

for all $u_i(t)\epsilon U_i, i = 1,2,3, t\epsilon\Gamma_1$

and c.)

$H\ (z_1*(t),z_2*(t),z_3*(t),u_1*(t),u_2*(t),u_3*(t),\lambda_1*(t),\lambda_2*(t),\lambda_3*(t))$

= constant   4.2.4.10

for all $t\epsilon\Gamma_1$.

4.3 Conclusions

Four theorems containing sets of necessary conditions for Problems 1 thru 4 were stated. The equations for the propogation of the first two monents through a linear time varing system are derived in Appendix A. The relation of the performance indices of the four problems to physical criteria is explored in Appendix B. Also in Appendix B the physical implications of the endpoint conditions are established. The proofs of Theorems 4.1, 4.2, 4.3 and 4.4 are given in Appendix C.

CHAPTER 5

SUFFICIENT CONDITIONS

Sufficient conditions are developed for Problem 1 using
the properties of convexity and concavity. The particular
performance index of equation 4.1.1.1 is shown to satisfy
these properties. A similar theorem is given for the
functional form of the performance index of Problem 2. No
theorem was developed for Problem 3 since this is a singular
problem and in general sufficient conditions are very diffi-
cult to develop for singular optimal problems. A set of
sufficient conditions for Problem 4 was derived using the
same techniques as in Problems 1 and 2 but is not included
in this chapter. The uniqueness of the optimal solution of
Problem 1 is established in Section 5.3.

5.1 Theorem 5.1

For Problem 1 of Section 4.1.1 it is sufficient that the
pair of external controls $(u_1; u_2, u_3^{'})$ are optimal in the sense
of establishing a saddle point defined by

$$J(z_1, z_2^{'}, z_3, u_1^{*}, u_2, u_3^{'}, \lambda_1^{*}, \lambda_2^{*}, \lambda 3^{*}) \leqq$$

$$J(z_1^{*}, z_2^{'*}, z_3^{*}, u_1^{*}, u_2^{*}, u_3^{'*}, \lambda_1^{*}, \lambda_2^{*}, \lambda_3^{*}) \leqq$$

$$J(z_1, z_2^{'}, z_3, u_1, u_2^{*}, u_3^{'*}, \lambda_1^{*}, \lambda_2^{*}, \lambda_3^{*})$$

5.1.1

for J where

$$J = J(z_1, z_2', z_3, u_1, u_2, u_3', \lambda_1, \lambda_2, \lambda_3) =$$

$$-\frac{1}{2} \int_{t_0}^{t_f} \int_{t_0}^{t_f} \{ \text{tr}[R_1 z_3(t,\tau) z_3'(t,\tau) R_1'] + \text{tr}[R_2 u_3(t,\tau) u_3'(t,\tau) R_1']$$

$$+ \text{tr}[[A(\tau) z_2'(t,\tau) + C(\tau) u_3'(t,\tau) - \frac{\partial z_2'}{\partial \tau}(t,\tau)][\lambda_2'(t,\tau)]]$$

$$+ \text{tr}[[A(t) z_3(t,\tau) + C(t) z_2(t,\tau) - \frac{\partial z_3}{\partial t}(t,\tau)][\lambda_3'(t,\tau)]] \} dt d\tau$$

$$+ \frac{1}{2} \int_{t_0}^{t_f} \{ <u_1(t), R_3 u_1(t)> - <u_2(t), R_4 u_2(t)>$$

$$+ <\lambda_1(t), A(t) z_1(t) + B(t) u_1(t) + C(t) u_2(t) - \frac{d}{dt} z_1(t)> \} dt$$

<div align="right">5.1.2</div>

or

$$J = J(z_1, z_2', z_3, u_1, u_2, u_3', \lambda_1, \lambda_2, \lambda_3) =$$

$$\int_{t_0}^{t_f} \int_{t_0}^{t_f} \Omega_2(z_2', z_3, u_3', \lambda_2, \lambda_3) dt d\tau + \int_{t_0}^{t_f} \Omega_1(z_1, u_1, u_2, \lambda_1) dt$$

<div align="right">5.1.3</div>

if the necessary conditions of Theorem 4.1 are satisfied.

Proof:

The proof follows immediately from Theorem 5.1.1 and Theorem 5.1.2.

5.1.1 Theorem 5.1.1

For Problem 1 of Section 4.1.1 it is sufficient that the pair of extremal controls $(u_1; u_2, u_3')$ are optimal in the sense of establishing a saddle point, as in 5.1.1, for the functional

form of the performance index J in 5.1.3 if in addition to the necessary conditions of Theorem 4.1 being satisfied,

Ω_1 is convex WRT u_1 and concave WRT u_2

5.1.1.1

and

Ω_2 is concave WRT u_3' and z_3.

5.1.1.2

Proof:

The complete proof is contained in Appendix D.1.

5.1.2 Theorem 5.1.2

It is both necessary and sufficient that the function Ω_1 is convex in u_1 and concave in u_2, and that the function Ω_2 is concave in z_3 and u_3' if:

R_1 is non-negative definite and symmetric

5.1.2.1

and

R_2, R_3, R_4 are positive definite and symmetric

5.1.2.2

Proof:

The complete proof is also contained in Appendix D.1.

5.2 Theorem 5.2

For Problem 2 of Section 4.1.2 it is sufficient that the pair of extremal controls $(u_1; u_2, u_3')$ are optimal in the sense of establishing a saddle point defined by equation 5.1.1 for J where

$$J = \int_{t_o}^{t_f} \int_{t_o}^{t_f} fo(z_1(t), z_1(\tau), z_2'(t,\tau), z_3(t,\tau), u_1(t), u_1(\tau),$$

$$u_2(t), u_2(\tau), u_3'(t,\tau)) dt d\tau \qquad 5.2.1$$

if in addition to the necessary conditions of Theorem 4.2 being satisfied

fo is convex in $u_1(t), u_1(\tau), z_1(t)$ and $z_1(\tau)$

$$5.2.2$$

and

fo is concave in $u_2(t), u_2(\tau), z_1(t), z_1(\tau), u_3'(t,\tau),$

$z_2'(t,\tau)$ and $z_3(t,\tau)$.

$$5.2.3$$

5.3 Uniqueness of Problem 1 Solution

The definiteness of the matrices R_2, R_3 and R_4 provide the means to establish equation 5.1.1 as strict inequalities. This means the extremal solution obtained from Theorem 4.1 is the only solution which is optimal. If the solution obtained from Theorem 4.1 is the only solution of the resulting TPBVP, then it is by the strict inequalities the unique solution. This point is clarified in Appendix D.1.

Since global convexity and concavity is used, global sufficient conditions result. Then if the TPBVP has only one solution, Theorem 4.1 provides necessary and sufficient conditions for unique global solutions for the specific performance index of equation 5.1.2.

53

These remarks do not apply for the functional form of the performance index given in equation 5.1.3, i.e., it is not shown to be the unique solution although it is the global optimal solution.

5.4 Conclusions

It is established that the necessary conditions of Theorem 4.1 are also sufficient conditions for Problem 1 of Section 4.1.1. This is accomplished by considering the properties of convexity and concavity in Theorem 5.1.1, and then proving Theorem 5.1.2 which gives the requirements on the weighting matrices of the performance index such that the convexity and concavity conditions are satisfied. In view of the results of these two theorems, the optimal solutions are the unique global optimal solutions if they exist.

In addition a sufficient condition theorem for Problem 2 is stated, also through the properties of convexity and concavity.

CHAPTER 6

IDENTIFICATION

The overall problem of simultaneous identification and control is now presented. It is the object of some techniques to identify while a process or plant is being controlled while other methods are used to identify only. Simultaneous identification and control require techniques suitable for "on-line" implementation. Methods for identification only are considered as "off-line" methods. It is the purpose of this chapter to study the "on-line" situation. Further, the simultaneous control is to be optimal or sub-optimal with respect to a given criteria. The control generated from Chapters 4 and 5 is to be implemented in such a way that if identification is complete, i.e. all parameters are known exactly, it would be the actual optimal control. If some parameters are not known exactly, then the control is actually sub-optimal. As the identification progresses in real time, the parameters are more closely known and if the dependence of the control on the parameters is updated, the sub-optimal control becomes closer to optimal. This scheme is carried out using a model reference adaptive control method as described in Section 6.3.

6.1 Background

The basic question of identifiability of parameters is discussed by several authors inthe December, 1974 Special Issue of the I.E.E.E. Transactions on Automatic Control. For the problems considered in this work, it will be assumed that the systems are controllable, observable, and in canonical form, hence identifiable.

Also, as previously mentioned, only parameter identification will be covered, since model identification would preclude the use of the simultaneous optimal control which is solved for offline. Further, the parameters are assumed to be constants.

Many different identification methods are currently known. Loosely-speaking, they can be classified in the following way.

First, frequency domain or spectrum analysis techniques are used in time series analysis. Also, random input methods are used which imply that all modes of the plant will be excited, hence identifiable. These first two methods are usually offline in that identification only is sought.

Secondly, when on-line identification and control is desired, the state space, or time domain representation is most often used. The two main classifications are model reference adaptive control and nonlinear estimation. As discussed in Sage (96) and Graupe (31), even if linear systems are to be identified, non-linear estimation must be employed. The reason for this is that the parameters to be identified are collected in a vector and adjoined to the usual state vector resulting in nonlinear state equations, since the parameters of the actual state matrix multiply the actual state variables. Depending on the a priori statistical information about the original state, measurement and inputs, various estimation schemes, such as Kalman filtering of a linearized representation of the nonlinear system can be used to identify the parameters. Further, the controlling input usually is not available as an optimal control, but must be chosen to aid the identification as mentioned in (96).

6.2 Model Reference - Gradient Approach

For two main reasons, the model reference adaptive control
method was chosen to be able to simultaneously identify and
optimally (or suboptimally) control. First, it is seen from
Section 6.1, that on-line controls available for optimization
are not included in most other identification methods. Second,
no methods are currently available when the stochastic disturbances
are Class 3), (i.e. second order) processes. For other kinds of
stochastic inputs, i.e. other than Class 3), model reference
adaptive control schemes have been studied. Many adaptive schemes
are not concerned with "optimal" control, but more with parameter
tracking or trajectory following.

In the model reference system used in this work several
issues arise which would significantly degrade performance in a
real - world application. The stability of the overall system of
plant, model, and adaptive loop with feedback control from the
model is questionable. Using Lyapunov stability theory Kuo(59)
derived conditions such that the overall system is stable. This
method to insure stability is applicable directly to the overall
system of Section 6.3, and that system can be shown to satisfy the
conditions given (59).

The stochastic input disturbance requires that an estimate,
hopefully optimal in some sense, of the state is available. In
the case where the complete state is available, it has been shown
by Kurtaran and Menachem (60) that the actual state is the best
estimate in the sense that it is the minimum error variance estimate.

For the case where only a noisy measurement of the state is available, a best estimateof the actual state is needed. This estimate has been derived and is discussed in Section 6.4. For plants with some parameters to be identified and only a noisy measurement of the state available, questions about sensitivity arise which are not covered here.

Finally, the parameter adjustment criteria was chosen as an integral square error criteria, since stability was easily established. The gradient or steepest descent algorithm as discussed in Kirk (45), Wilde and Beightler (107), and Sage (96) was modified and used such that a real time implementation could be readily obtained. It is seen that in a real world application this algorithm requires much less digital hardware than say a Kalman estimator, since no calculus is involved in the parameter adjustment, whereas in Kalman estimation a matrix Ricatti differential equation must be solved by the digital hardware.

6.3 Implementation

This section describes in detail how to implement the optimal control and simultaneously identify parameters in a model reference adaptive scheme. The controls used are those derived in Theorem 4.1 since closed loop strategies are readily obtainable and easily implemented.

6.3.1 Optimal Closed Loop Strategies

The necessary conditions of Section 4.2.1 imply the open loop problem solution, i.e. the control laws for u_1^* and u_2^* depend on end times and endpoint conditions only

$$u_1^*(t) = u_1^*(t_0, z_1(t_0), t_f, z_1(t_f)), \quad t \in \Gamma_1 \qquad 6.3.1.1$$

and

$$u_2^*(t) = u_2^*(t_0, z_1(t_0), t_f, z_1(t_f)), \quad t \in \Gamma_1 \qquad 6.3.1.2$$

In many situations it is more desired to form a closed loop control by feedback of the state $z_1(t)$, for all $t \in \Gamma_1$. These control laws are desired to be in the form

$$u_1^*(t) = u_1^*(t, z_1^*(t)), \quad t \in \Gamma_1 \qquad 6.3.1.3$$

and

$$u_2^*(t) = u_2^*(t, z_1^*(t)), \quad t \in \Gamma_1. \qquad 6.3.1.4$$

These would be nonlinear control laws. It is even more desired to find linear time - varying feedback gains for ease of implementation and stability purposes. These points are discussed in Athans and Falb (9) for deterministic optimal control problems and in Lee (64) and Wonham(109) for stochastic optimal control problems.

For this class of problems, control laws given by

$$u_1^*(t) = W_1(t) z_1^*(t), \quad t \in \Gamma_1 \qquad 6.3.1.5$$

and

$$u_2^*(t) = W_2(t) z_1^*(t), \quad t \in \Gamma_2 \qquad 6.3.1.6$$

are to be found.

From Theorem 4.2.1,

$$u_1^*(t) = -R_3^{-1}B'(t)\lambda_1^*(t), \ t \in \Gamma_1 \qquad\qquad 6.3.1.7$$

and

$$u_2^*(t) = R_4^{-1}C'(t)\lambda_1^*(t), \ t \in \Gamma_1 \qquad\qquad 6.3.1.8$$

so letting

$$\lambda_1^*(t) = K(t) \, z_1^*(t), \ t \in \Gamma_1 \qquad\qquad 6.3.1.9$$

where

K (t) is an n x n matrix

would result in

$$u_1^*(t) = -R_3^{-1} B'(t) \, K(t) \, z_1^*(t), \ t \in \Gamma_1 \qquad\qquad 6.3.1.10$$

giving

$$W_1(t) = -R_3^{-1}B'(t)K(t), \ t \in \Gamma_1 \qquad\qquad 6.3.1.11$$

and

$$u_2^*(t) = R_4^{-1}C'(t)K(t)z_1^*(t), \ t \in \Gamma_1 \qquad\qquad 6.3.1.12$$

giving

$$W_2(t) = R_4^{-1}C'(t)K(t), \ t \in \Gamma_1 \qquad\qquad 6.3.1.13$$

Differentiating 6.3.1.9 gives

$$\frac{d\lambda_1^*}{dt} = \frac{dK}{dt}(t) \, z_1^*(t) + K(t) \frac{dz_1^*}{dt}(t), \ t \in \Gamma_1 \qquad\qquad 6.3.1.14$$

Substituting from the necessary conditions of Theorem 4.2.1 it can

be shown that K(t) must satisfy

$$\frac{dK}{dt}(t) = -A'(t) \, K(t) - K(t) \, A(t) - K(t) \, D(t) \, K(t), \ t \in \Gamma_1$$

$$6.3.1.15$$

where

$$D(t) = -B(t)R_3^{-1}B'(t) + C(t)R_4^{-1}C'(t), t \in \Gamma_1 \qquad \qquad 6.3.1.16$$

and further $K(t) = K'(t)$ if $D(t)$ is symmetric as is shown in Kirk (45) and Ogata (83).

The initial conditions required for the solution of equation 6.3.1.15 are obtained from equation 6.3.1.9 at $t = t_0$.

This control law gives a closed loop feedback solution which is solved offline and implemented on-line either by storing the function $K(t) \forall t \in \Gamma_1$ or by on line simulation. This is a very useful and practical means for generating an optimal control, and even more useful when simultaneous on-line identification is required.

6.3.2 Model Reference Identification and Simultaneous Updated Sub-Optimal Control

The overall method used is explained by considering Figure 6.3.2.1

I.C.

v →

PHYSICAL PLANT
$$\dot{x}=Ax+Bu_1+Cv$$

x

u_1

e

$W\int$ J

I.C.

MODEL
$$\dot{z}_1=A(\gamma)z_1+B(\gamma)u_1+C(\gamma)u_2$$

z_1

u_2

γ_{NEW}

$R_4^{-1}C'(\gamma)$

$-R_3^{-1}B(\gamma)$

Digital hardware used to generate a new γ after each sampling period

I.C.

$\int$ K

γ_{NEW}

$A(\gamma)$

$A'(\gamma)$

$D(\gamma)$
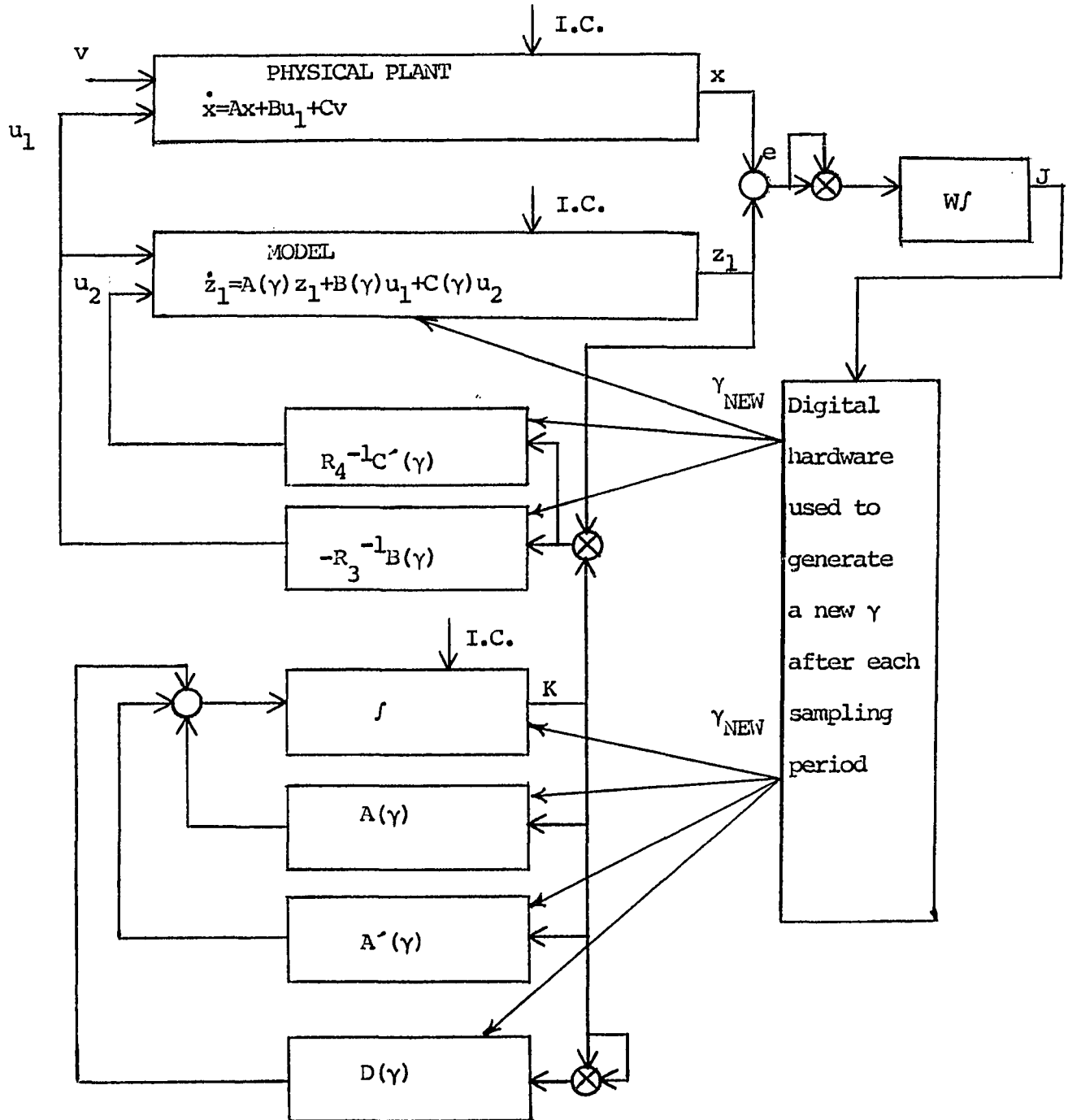
Figure 6.3.2.1 Model reference adaptive system

As is seen, the only digital block is that one used to generate a new set of values for the parameters being identified. The closed loop feedback controls can be simulated using analog equipment.

The steepest descent algorithm is carried out by the digital block. The rate of convergence, region of convergence, sampling rate, and weights are all very involved problems in themselves. For a particular real-world application, these problems are generally dealt with after several trial runs. No comprehensive theory exists for predetermined solutions. After completion of this step in some manner a workable, near-optimal system is obtained.

For a completely identified system, the closed loop controller u_i would be optimal for both the model and the plant with respect to the criteria of Theorem 4.1. If some parameters are not identified exactly, then suboptimal control is applied while identification is being carried out. By adjusting the parameters in the analog simulation of the feedback gain K and model z_i, the control u_i is updated and as the parameters become closer and closer to their true values, the suboptimal controls become closer and closer to the optimal controls.

The actual plant is governed by

$$\frac{dx(t,\cdot)}{dt} = A(t)x(t,\cdot) + B(t)u_1(t) + C(t)v(t,\cdot), \quad t \in \Gamma_1 \qquad 6.3.2.1$$

where some parameters of A, B, and C are unknown.

The model is

$$\frac{dz_1(t)}{dt} = A(\gamma,t)z_1(t) + B(\gamma,t)u_1(t) + C(\gamma,t)u_2(t), \quad t \in \Gamma_1 \qquad 6.3.2.2$$

where γ is a vector of the unknown parameters.

In equation 6.3.2.2, the controls $u_1(t)$ and $u_2(t)$ are simulated

in terms of the time-varying feedback gain matrix $K(t)$ governed by

$$\frac{dK(t)}{dt} = -A'(\gamma,t)K(t) - K(t)A(\gamma,t) - K(t)D(\gamma,t)K(t), t \in \Gamma_1 \qquad 6.3.2.3$$

where

$$D(\gamma,t) = -B(\gamma,t)R_3^{-1}B'(\gamma,t) + C(\gamma,t)R_4^{-1}C(\gamma,t). \qquad 6.3.2.4$$

Define the error between plant and model state as

$$e(\gamma,t) = z_1(t) - x(t) \qquad 6.3.2.5$$

and the weighted criteria used to generate updated values of the

parameters as

$$J_I(\gamma,i) = E\{\int_{t_i}^{t_{i+1}} <e(\gamma,t),W_i e(\gamma,t)>dt\} \qquad 6.3.2.6$$

where W_i is a constant positive definite weighting matrix,

t_i, $t_i + 1 \in \Gamma_1$, $i = 1$, ---, N where N is the number of sampling

intervals.

The method of steepest descent is used to numerically minimize

$J_I(\gamma)$ with respect to γ such that the system will be identified, i.e. the model parameters are equal to the actual plant parameters when $J_I(\gamma)$ becomes sufficiently small, in comparison to some preselected numerical value. The function is minimized by evaluating the slope or gradient at a given initial point, and then moving in the direction of steepest slope to a new point. The process is carried out numerically until $J_I(\gamma)$ is less than some preselected value, or until no further decrease in $J_I(\gamma)$ can be obtained. The slope is approximately numerically as

$$\frac{\partial J}{\partial \gamma} \simeq \frac{\Delta J}{\Delta \gamma}$$
6.3.2.7

for each iteration where

$$\Delta J_I(\gamma) = J_I(\gamma, i+1) - J_I(\gamma, i)$$
6.3.2.8

and

$$\Delta \gamma = \gamma_{ORIGINAL} - \gamma_{NET\ CHANGE}$$
6.3.2.9

and $\gamma_{ORIGINAL}$ is the initial guess at the parameters used in solving the "worst-case" optimal control problem offline.

At each iteration a new value of γ is obtained from

$$\gamma_i = \gamma_{i+1} - \Delta \gamma$$
6.3.2.10

and

$$\Delta \gamma = \theta \frac{\partial J_I(\gamma, i+1)}{\partial \gamma}$$
6.3.2.11

where θ is a weighting factor chosen to aid convergence. As each new value of γ is obtained, the model generating the state z_i and feedback gain K are updated.

The overall system is suboptimally controlled until identification is complete, then optimally controlled from that point on. This overall system is an adaptive system using a reference model, or a model-reference adaptive control system. A complete digital simulation of the overall system was written and is discussed in Chapter 7.

6.4 Optimal Estimation with Second Order Stochastic Disturbances

In this section an estimate is obtained for the case where only a noisy measurement of the state, not the state itself, is available. The results could be used in conjunction with the identification procedure of the first portion of this chapter, where the estimate $\hat{x}$, replaces the state x in the model reference system. If this was to be done, a sensitivity analysis would be necessary before implementation. Assuming identification was complete, i.e. all parameters were known, the estimate $\hat{x}$ could be used for many purposes, just as the estimate from a Kalman filter is used.

Consider a plant with stochastic disturbance

$$\frac{dx(t,.)}{dt} = A(t)x(t,.) + B(t)u(t) + C(t)v(t,.) \qquad 6.4.1$$

and with noisy measurement

$$z(t,.) = H(t)x(t,.) + w(t,.). \qquad 6.4.2$$

It is desired to find an optimal filter such that $\hat{x}$ is the best estimate, in the sense of minimum error covariance, where the filter is constrained by

$$\frac{d\hat{x}(t)}{dt} = F(t)\hat{x}(t) + G(t)z(t) + D(t)u(t) \qquad 6.4.3$$

where F, G, and D are to be determined. The filter is also to be unbiased. For this condition,

$$F(t) = A(t) - G(t)H(t) \qquad 6.4.4$$

and

$$D(t) = B(t) \qquad 6.4.5$$

where the means $\mu_v(t)$ and $\mu_w(t)$ must be either zero or known a priori. If $v(t)$ and $w(t)$ have essentially the same a priori data known that is discussed in Appendix B, then the optimal value of $G(t)$ can be found and the estimate $\hat{x}$ generated as given in:

Theorem 6.1

For the conditions of 6.4.1 through 6.4.5 the optimal

estimate $\hat{x}$ is given by

$$\frac{d\hat{x}}{dt}(t) = A(t)\hat{x}(t) + G(t)[z(t)-H(t)\hat{x}(t)] + B(t)u(t) \qquad 6.4.6$$

where $G(t)$ is determined by and must satisfy

$$G(t)=V_{w\tilde{x}}(t,\tau)[V_{w\tilde{x}}(t,\tau)H(\tau)+V_{ww}(t,\tau)]^{-1} +S(t) \qquad 6.4.7$$

and where $V_{w\tilde{x}}(t,\tau)$ is found from

$$\frac{\partial V_{w\tilde{x}}}{\partial \tau}(t,\tau)=V_{w\tilde{x}}(t,\tau)[A(\tau)-G(\tau)H(\tau)]^{'}+V_{wv}(t,\tau)C^{'}(\tau)-V_{ww}(t,\tau)G^{'}(\tau) \qquad 6.4.8$$

Proof: The proof leading to equations 6.4.6 through 6.4.8 is

developed in Appendix F. The method is similar to the calculus

of variations derivation of the Kalman filter given in Sage (97).

Remark: The actual a priori statistical data, the differential

equations describing the evolution of the error covariance, and

the performance index used are all detailed in Appendix F.

In Appendix F it is shown that $S(t)$ is a weighting matrix. The

error $x(t)$ is defined as

$$\tilde{x}(t,.) = x(t,.) -\hat{x}(t) \qquad 6.4.9$$

The results given above are just briefly quoted to show

the preliminary work done in extending the "worst case" control

theorems of Chapter 4 and 5, and the identification method of the

previous sections of Chapter 6.

6.5 Conclusions

In this chapter, the model reference adaptive control system

used to simultaneously identify parameters and suboptimally control

a physical plant is given. The real-time or on-line implementation

is possible and in fact provided the reasoning on which the choice

of this method was made. The parameters are identified using

steepest descent or gradient methods suitable for on-line use.

The preliminary estimation results of the last section are included to show the extension made to the case where only a noisy measurement of the actual system state is available.

CHAPTER 7

APPLICATIONS

Consider a physical plant, e.g. a cardiac cell, which can be mathematically described by linear ordinary differential equations with time varying coefficients. Assuming the parameters, i.e. gains and time constants of the equations are known then it is desired to optimally control the plant with respect to some measure of performance. If, in addition, the plant is subjected to stochastic disturbances, the optimal control of the plant becomes much more difficult. Assuming some of the plant parameters are not known further complicates the implementation of some optimal scheme.

This chapter illustrates through example, a technique for the simultaneous optimal control and parameter identification of a physical plant in the presence of stochastic inputs.

The theory required for "worst case" optimal control in a stochastic environment was presented in Chapters 4 and 5. All the computations can be done off-line before the operation of the plant begins. As the plant operation progresses the unknown parameters are updated. These updated parameters are fed to the model in such a way that the "optimal" controls are generated from these new values of the parameters after each update. Until the identification is complete, though, actual optimal control is not possible, though qualitatively "good" suboptimal control is actually obtained over the time interval of interest as can be seen by comparing actual and non-identified trajectories.

7.1 An Illustrative Example

An illustrative example showing in detail the optimal computations and then a simulation of simultaneous sub-optimal control and identification is given. The simulation was completely performed on a digital computer using the Fortran language. Ideally, a hybrid computer simulation would perhaps be more suitable as a means to illustrate all the various aspects of the theory, but large enough facilities were not available. The Fortran program listings are contained in Appendix E. The particular example is specified and set up in Section 7.1.1 . The offline optimal computations are detailed in Section 7.1.2 . The results of the simulation of implementing both the identification and control are given in Section 7.1.3.

7.1.1 Setup

A particular example was simulated on the digital computer. The

actual, real-world physical plant is described in this section.

The general equation for the plant is

$$\frac{dx(t,.)}{dt}=A(\gamma,t)x(t,.)+B(\gamma,t)u(t)+C(\gamma,t)v(t,.),t\in\Gamma_1 \qquad 7.1.1.1$$

Choosing n=r=s=1 gives x,u, and v as scalars, hence

$$\frac{dx(t,.)}{dt}=a(\gamma)x(t,.) + b(\gamma) u (t) + c(\gamma)v(t,.).t\in\Gamma_1 \qquad 7.1.1.2$$

Assuming c is known, a(Y) = a and b(Y) = b yields

$$\gamma= \begin{bmatrix} a \\ b \end{bmatrix} \qquad 7.1.1.3$$

and omitting the dependence of a and b on Y,

$$\frac{dx(t,.)}{dt} = ax(t,.)+bu(t)+cv(t,.)t\in\Gamma_1 \qquad 7.1.1.4$$

where a and b are to be identified.   Equation 7.1.1.4

in analog computer form is
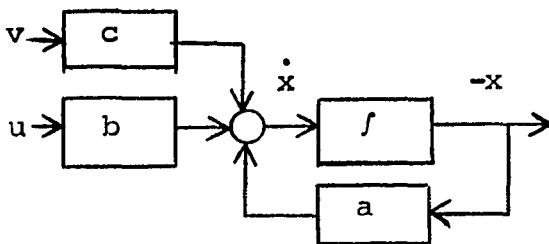
Figure 7.1.1.1 Analog computer diagram of example

which physically is a lag with time constant $1/|a|$,

and with the sum of deterministic control u with gain b and

stochastic disturbance v with gain c as input.

It is desired to identify the parameters a and b while

simultaneously optimally or at least sub optimally controlling the

plant in the presence of the stochastic disturbance v, which

is assumed to be a Class 3), i.e. second order, stochastic process.

In order to solve the optimal control problem offline,

initial guesses at the values of the parameters are required.

In the particular example simulated these were

$$a = -3.0$$

$$b = 2.0 \hspace{6cm} 7.1.1.5$$

Also, the other data required was selected as

$$c = 1.0$$

$$t_o = 0.0$$

$$t_f = 1.0 \hspace{6cm} 7.1.1.6$$

More data is required for the optimal solution and is specified

in the next section.

7.1.2 Optimal Solution

For purposes of obtaining the optimal problem solution,

the physical plant is described by

$$\frac{dx(t,.)}{dt} = -3.0x(t,.)+2.0u(t)+1.0v(t,.) \quad ,t\in[0.0,1.0] \hspace{2cm} 7.1.2.1$$

as specified in the previous section.

Using the theory of Appendix A, the first two moments

of the state $x(t,.)$ become

$$\frac{dz_1(t)}{dt} = -3.0z_1(t) + 2u_1(t)+u_2(t) ,t\in[0.0,1.0] \hspace{2cm} 7.1.2.2$$

$$\frac{\partial z_2(t,\tau)}{\partial \tau} = -3.0z_2(t,\tau) + u_3(t,\tau), \quad (t,\tau)\in[0.0,1.0]\times[0.0,1.0] \hspace{1cm} 7.1.2.3$$

$$\frac{\partial z_3(t,\tau)}{\partial t} = -3.0z_3(t,\tau)+z_2(t,\tau), \quad (t,\tau)\in[0.0,1.0]\times[0.0,1.0] \hspace{0.5cm} 7.1.2.4$$

The variables z_1, z_2, z_3, u_1, u_2 and u_3 are all scalars.

The weighting matrices were chosen as

$$R_1 = 0.0$$
$$R_2 = \sqrt{2}$$
$$R_3 = 1.0$$
$$R_4 = 1.0 \hspace{5cm} 7.1.2.5$$

giving

$$J(u_1, u_2, u_3) = 1/2 \int_{t_o}^{t_f} \{u_1^2 - u_2^2\} dt - 1/2 \int_{t_o}^{t_f} \int_{t_o}^{t_f} \{2u_3^2\} dt d\tau \qquad 7.1.2.6$$

The initial states used were

$$z_1(t_o) = z_1(0) = 5.0 \qquad\qquad 7.1.2.7$$

corresponding to $\pi_1 = 5.0$ where this is the measurable prior mean.

It was desired to drive the state (in terms of z_1) to

$$z_1(t_f) = z_1(1) = 0.1 \qquad\qquad 7.1.2.8$$

The initial cross covariance chosen was

$$z_2(t, t_o) = 1.0 e^{-t} \qquad\qquad 7.1.2.9$$

and the initial covariance

$$z_3(t_o, \tau) = 10.0 e^{-\tau} \qquad\qquad 7.1.2.10$$

which correspond to a cross variance

$$z_2(t_o, t_o) = 1.0$$

and a variance

$$z_3(t_o, t_o) = 10.0$$

which are measurable as is the time constant $T = 1.0$

used in 7.1.2.9 and 7.1.2.10 for the exponential correlation

distribution discussed in Appendix B. The final time endpoints

are discussed later. For the above data, application of

equations 4.2.1.1 through 4.2.1.12 yield

$$\frac{dz_1^*(t)}{dt} = -3.0z_1^*(t)+2.0u_1^*(t)+u_2^*(t)$$

$$\frac{\partial z_2^*(t,\tau)}{\partial \tau} = -3.0z_2^*(t,\tau)+u_3^*(t,\tau)$$

$$\frac{\partial z_3^*(t,\tau)}{\partial t} = -3.0z_3^*(t,\tau)+z_2^*(t,\tau) \qquad 7.1.2.11$$

$$\frac{d\lambda_1^*(t)}{dt} = 3.0\lambda_1^*(t)$$

$$\frac{\partial \lambda_2^*(t,\tau)}{\partial \tau} = 3.0\lambda_2^*(t,\tau)-\lambda_3^*(t,\tau)$$

$$\frac{\partial \lambda_3^*(t,\tau)}{\partial t} = 3.0\lambda_3^*(t,\tau) \qquad 7.1.2.12$$

$$u_1^*(t) = -2.0\lambda_1^*(t)$$

$$u_2^*(t) = \lambda_1^*(t)$$

$$u_3^*(t,\tau) = 0.5\lambda_2^*(t,\tau) \qquad 7.1.2.13$$

Putting 7.1.2.13 into 7.1.2.11 yields the following
2n (1+s+n) = 6 differential equations with the four
endpoint conditions of 7.1.2.7 through 7.1.2.10,

$$\frac{dz_1^*(t)}{dt} = -3.0z_1^*(t)-3.0\lambda_1^*(t)$$

$$\frac{\partial z_2^*(t,\tau)}{\partial \tau} = -3.0z_2^*(t,\tau)+0.5\lambda_2^*(t,\tau)$$

$$\frac{\partial z_3^*(t,\tau)}{\partial t} = -3.0z_3^*(t,\tau)+z_2^*(t,\tau)$$

$$\frac{d\lambda_1^*(t)}{dt} = 3.0\lambda_1^*(t)$$

$$\frac{\partial\lambda_2^*(t,\tau)}{\partial\tau} = 3.0\lambda_2^*(t,\tau)-\lambda_3^*(t,\tau)$$

$$\frac{\partial\lambda_3^*(t,\tau)}{\partial t} = 3.0\lambda_3^*(t,\tau) \qquad\qquad 7.1.2.14$$

Two more final time endpoint conditions are required.

These conditions can be established by assuming the following

form for the initial conditions of λ_2 and λ_3

$$\lambda_2(t,t_o) = Le^{-t}$$

$$\lambda_3(t_o,\tau) = Ke^{-\tau} \qquad\qquad 7.1.2.15$$

With this form the differential equations can be solved in

terms of the constants L and K. These constants can be

determined algebraically by choosing $z_2(t,t_f)$ and $z_3(t_f,\tau)$ as

functions of t and τ respectively, realizing that z_2 and z_3

will have solutions due to their respective transition

matrices and initial time endpoint conditions as well as those

of λ_2 and λ_3. From a physical viewpoint, it would be desired

to drive $z_2(t_f,t_f)$ and $z_3(t_f,t_f)$ to values much smaller than the

initial cross variance and variance. Specifying the final cross

variance and variance as

$$z_2(t_f,t_f) = 0.688$$

$$z_3(t_f,t_f) = 0.303 \qquad\qquad 7.1.2.16$$

requires

$$z_2(t, t_f) = 0.061e^{-t} + 0.033e^{3t}$$

$$z_3(t_f, \tau) = 0.480e^{-\tau} + 0.170e^{-3\tau} + 0.589e^{3\tau} \qquad 7.1.2.17$$

where the constants K and L are algebraically determined from

7.1.2.16 and 7.1.2.17 as

$$L = 0.00694$$

$$K = -0.333 \qquad 7.1.2.18$$

The solutions for the differential equations may be obtained

by three methods. From Appendix A, the solutions are directly

found if the transition matrix is known. Assuming the form of

equation 7.1.2.15 the two-dimensional Laplace transform technique

in Kuo (54) can be applied. Digital computer simulation of the

$\frac{1}{2}n(1+s+n)$ differential equations could also be used. Algebraic

solutions for the constants in 7.1.2.15 is used for the first

two methods above. Shooting techniques would allow these constants

to be evaluated on a digital computer.

The complete analytic solution is

$$z_1^*(t) = 5.0e-t - 0.015\sinh 3t$$

$$z_2^*(t, \tau) = 1.0e^{-t}e^{-3\tau} + 0.00347e^{3t}e^{-3\tau} + 0.000578e^{-t}e^{3\tau}$$

$$+0.00173e^{3t}e^{3\tau} - 0.00520e^{3t}e^{-\tau}$$

$$z_3^*(t, \tau) = 10.0e^{-3t}e^{-\tau} - 0.500e^{-3t}e^{-3\tau} - 0.000578e^{-3t}e^{3\tau}$$

$$+0.499e^{-t}e^{-3\tau} + 0.000578e^{3t}e^{-3\tau} + 0.000289e^{-t}e^{3\tau}$$

$$+0.000289e^{3t}e^{3\tau} - 0.000868e^{3t}e^{-\tau}$$

$$\lambda_1^*(t) = 0.0150e^{3t}$$

$$\lambda_2^*(t,\tau) = 0.00694e^{-t}e^{3\tau} + 0.0208e^{3t}e^{3\tau} - 0.0208e^{3t}e^{-\tau}$$

$$\lambda_3^*(t,\tau) = 0.0833e^{3t}e^{-\tau}$$

$$u_1^*(t) = -0.0300e^{3t}$$

$$u_2^*(t) = 0.0150e^{3t}$$

$$u_3^*(t,\tau) = 0.00347e^{-t}e^{3\tau} + 0.0104e^{3t}e^{3\tau} - 0.0104e^{3t}e^{-\tau} \qquad 7.1.2.19$$

The value of the performance index J is

$$J(u_1^*, u_2^*, u_3^*) = -0.452$$

From Section 6.3.1, the analytic solution for the time varying gain is

$$K(t) = 0.09/ (29.955e^{-6t} + 0.045) \qquad 7.1.2.20$$

where for this problem

$$\frac{dK(t)}{dt} = 6K(t) + 3K^2(t) \qquad 7.1.2.21$$

and

$$K(t_o) = 0.003 \qquad 7.1.2.22$$

All these solutions were obtained and plotted using various digital techniques and are presented graphically in the following figures.
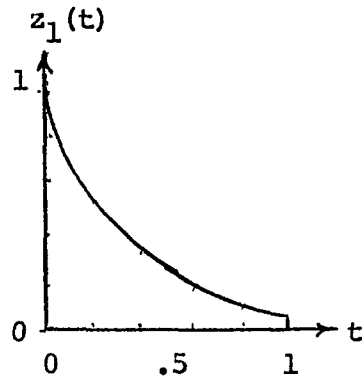
Figure 7.1.2.1 Plot of z_1
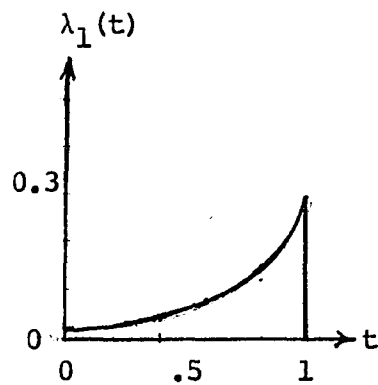
Figure 7.1.2.2 Plot of λ_1
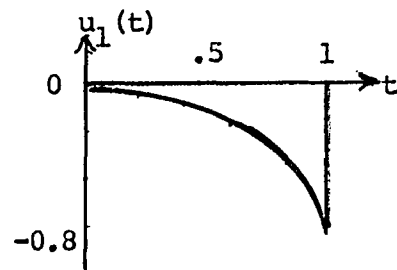
Figure 7.1.2.3 Plot of u_1
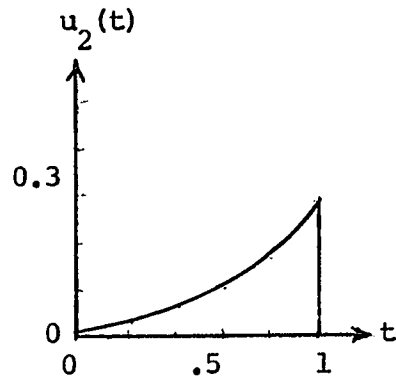
Figure 7.1.2.4 Plot of u_2
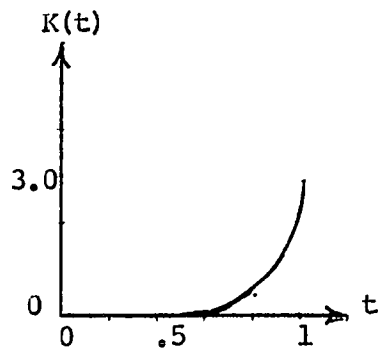
Figure 7.1.2.5 Plot of K
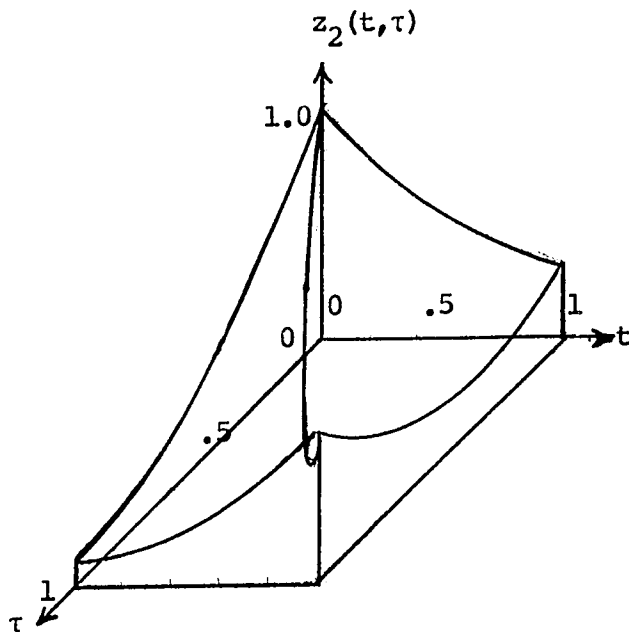
Figure 7.1.2.6 Plot of z_2
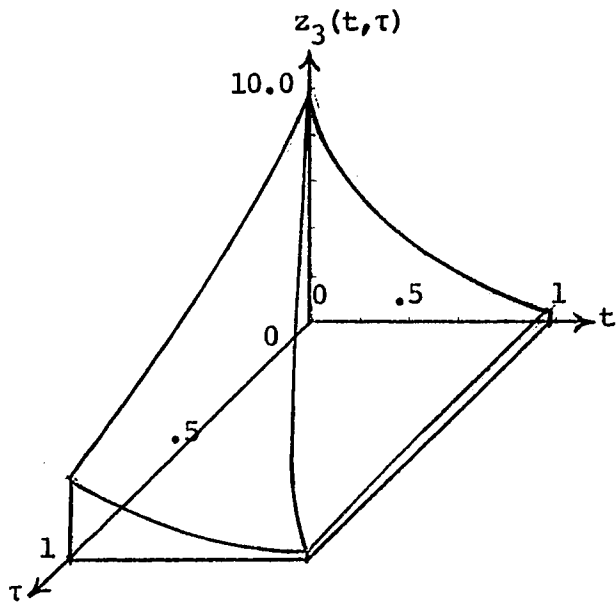
Figure 7.1.2.7 Plot of z_3
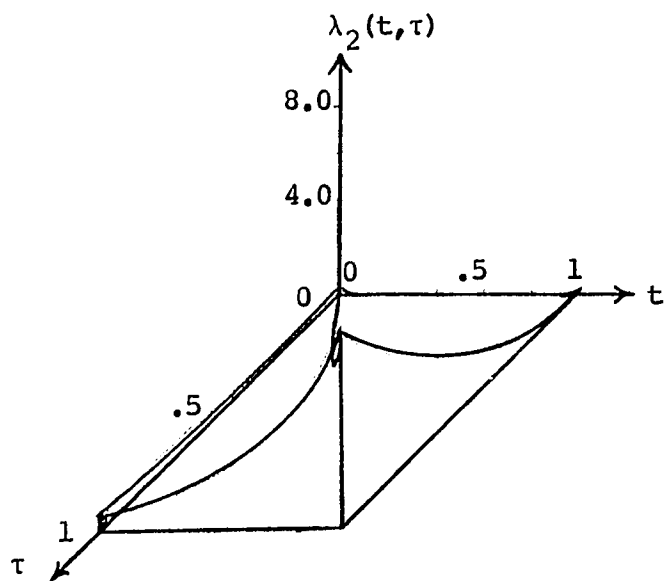
Figure 7.1.2.8 Plot of λ_2
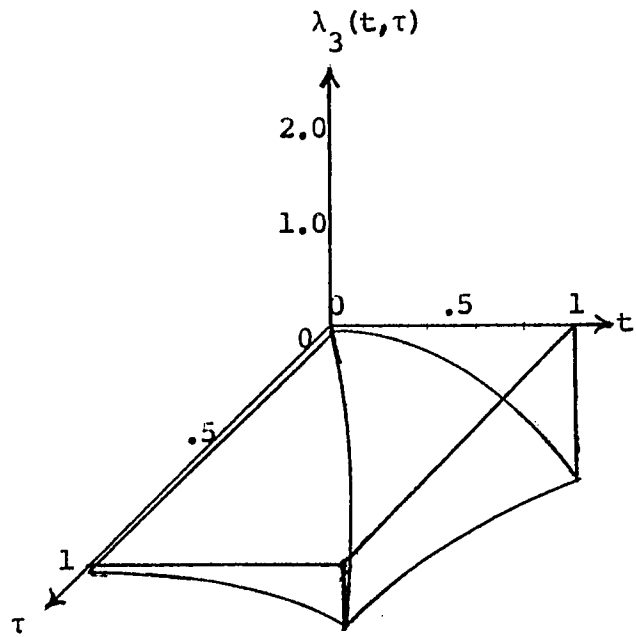
Figure 7.1.2.9 Plot of λ_3
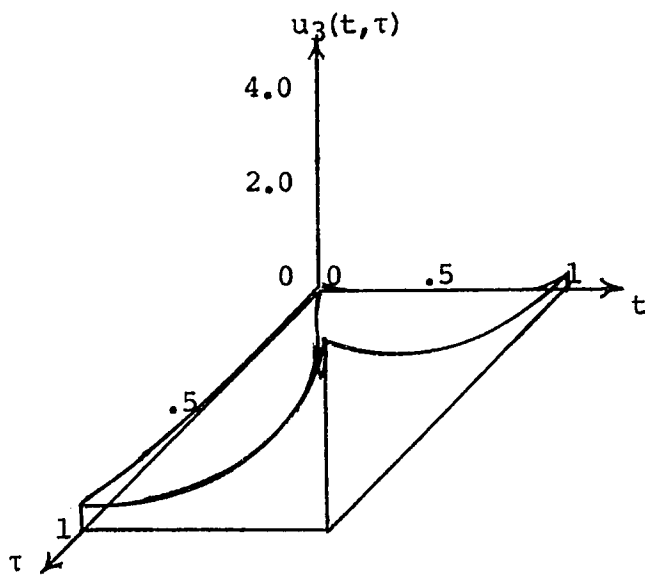
Figure 7.1.2.10 Plot of u_3

7.1.3 Simultaneous Identification Solution

The complete model reference system of Figure 6.3.2.1 was

simulated and run on a digital computer. The results of the

previous section were implemented as shown in Figure 6.3.2.1 with

allowance to change the parameters a and b. A continuous

system was approximated by dividing the interval [0.0,1.0] into

400 sub-intervals for integration purposes. Every ten

sub-intervals, the error was numerically sampled, and a resulting

change in parameters a and b calculated. This corresponds

to there being 40 sampling intervals in one second, therefore

convergence must be fast enough to come to completion before

40 changes occur. Similarly, overshoot of the minimization of

the functional must be prohibited. These resulted in selection of

a heuristic scaling of the factor θ such that smaller percentage

changes occured as the percent change in the error function

decreased. The stopping criteria was selected as 0.05% of the

value of the error during the first sampling period. This

corresponds roughly to a gradient of less than 0.083. that is

the magnitude of the gradient of the function at this stopping

point is very small. The problem was run assuming initial guesses of

$$a = 3.0$$
$$b = 2.0 \hspace{4cm} 7.1.3.1$$

with actual parameters of

$$a_{PLT} = -4.0$$
$$b_{PLT} = 3.0 \hspace{4cm} 7.1.3.2$$

For the above particular set of actual plant parameters, the

identification procedure resulted in

a = -3.89707

b = 3.11534   7.1.3.3

after 0.6 seconds corresponding to 24 identification sub-intervals

having elapsed. For purposes of convergence, the weight W of the

integral square error measure was assigned as

W = 500   7.1.3.4

after several trial runs established the range of W such that

overshoot did not occur, yet convergence did progress rapidly

enough such that the stopping criteria was met in less that 40

sampling intervals. The results of the overall simulation are

presented graphically in the following figures obtained with

digital plotting routines.

First the convergence of the parameters to the true values

is pictured.
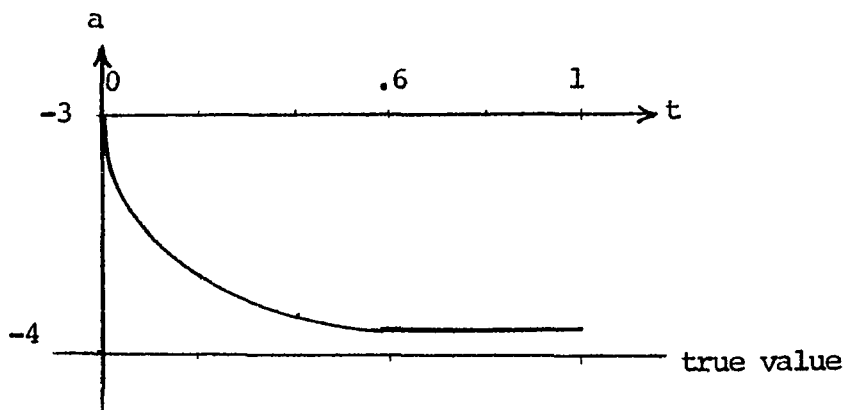
Figure 7.1.3.1   Identification of a
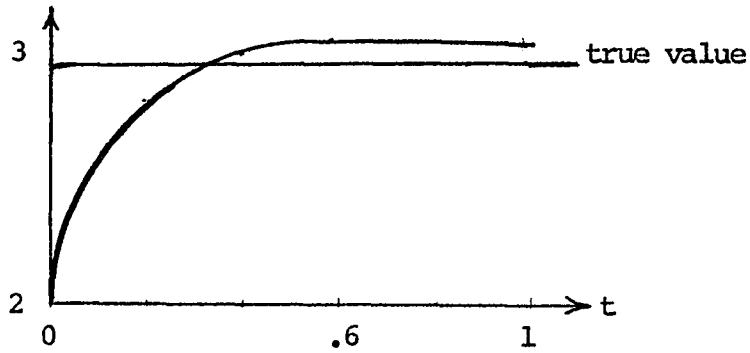
Figure 7.1.3.2   Identification of b

Using the actual time varying values of a and b in Figure 7.1.3.1 , the state of the model is
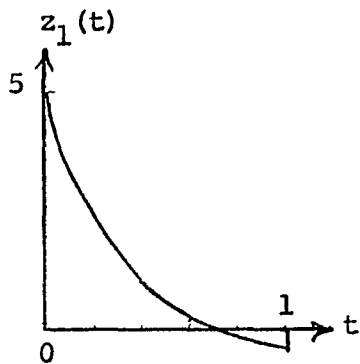
Figure 7.1.3.3.

Trajectory of z_1 from model reference simulation

whereas the state of the model from the optimal solution

run with the same values as the true parameters and with

the same initial conditions as would be used offline is
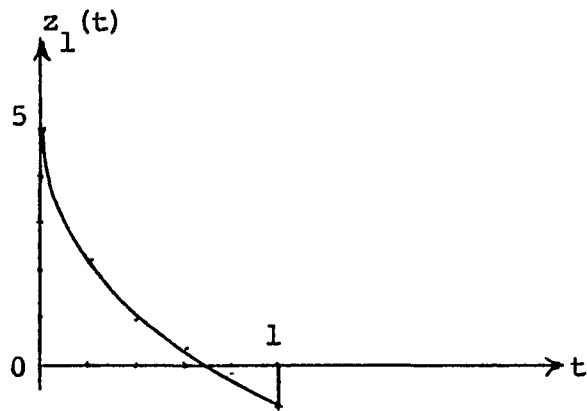
Figure 7.1.3.4 Trajectory of z_1 from optimal solution

Similarly the adaptive scheme provides a suboptimal control u_1 as
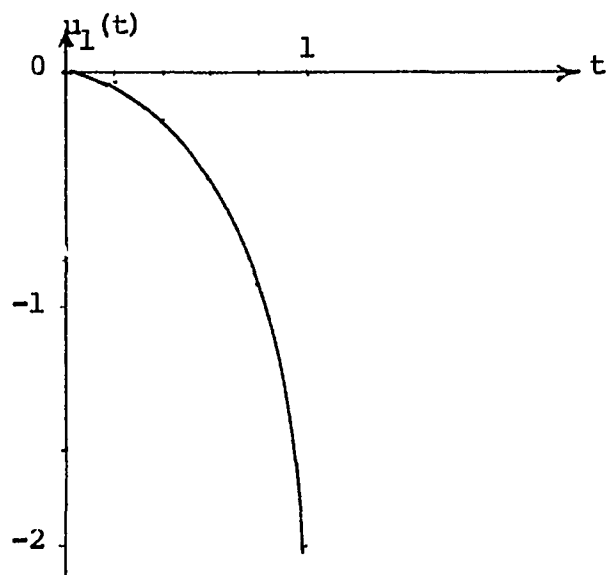
Figure 7.1.3.5 Trajectory of u_1 from model reference

simulation

whereas the optimal control u_1 from the same run as Figure 7.1.3.4 is
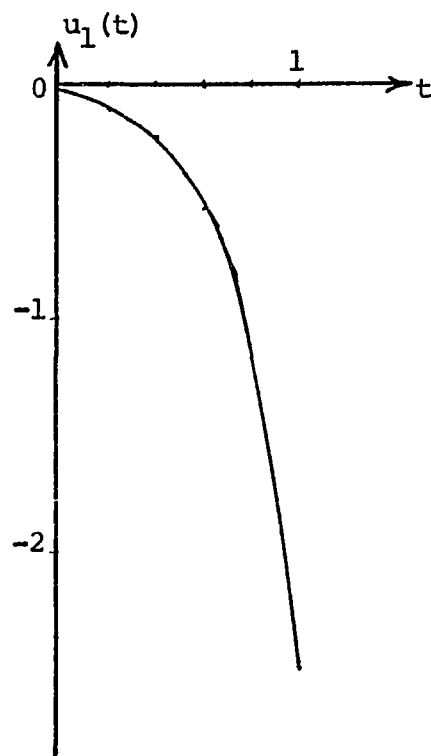
Figure 7.1.3.6 Trajectory of u_1 from optimal solution

Comparing the last four figures it is seen that "qualitatively good" suboptimal control is obtained in terms of the closeness of trajectories of the state z_1 and control u_1 from the model reference scheme and the optimal run with true values substituted. Therefore in this sense the model reference scheme is close to optimal or suboptimal with regard to the criteria of Chapters 4 and 5.

An example of the stochastic process v is shown below
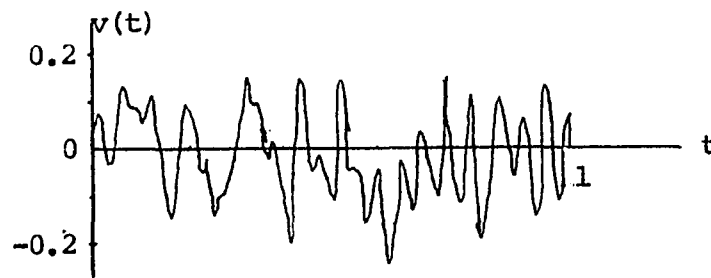
Figure 7.1.3.7 Plot of noise v used in model reference simulation

The typical effects of the noise v on a steepest descent path are illustrated next.
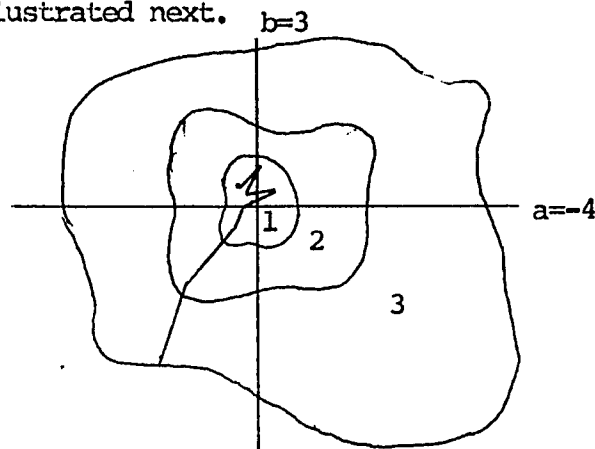
Figure 7.1.3.8

Regions of parameter space

Region 1 shows the area close to the minimum where the parameters are close. In this region the noise effects are most pronounced. Region 2 is where the noise effects are observed but not quantitatively significant with respect to the gradient search. In this region the parameters are far enough unequal such that the gradient procedure continues with no randomness. In Region 3, the parameters are so far apart that the error due to this is very much greater than the error due to the noise. The outer bound of Region 3 is the limit of the region of convergence. These regions were not numerically established but could have been by just executing many runs with various data. Generally speaking, the circle for Regions 1 and 2 could be determined in terms of the norm of the variance of the noise v as is mentioned in Bryson and Ho (20).

In the actual results obtained, different noise sequences were generated for runs with the same unknown parameters. The average value of parameters as identified from the runs was

$a_{AVG} = -3.89266$

$b_{AVG} = 3.112383$

In the next two figures are samples of the identification of
a and b for several different noise sequences
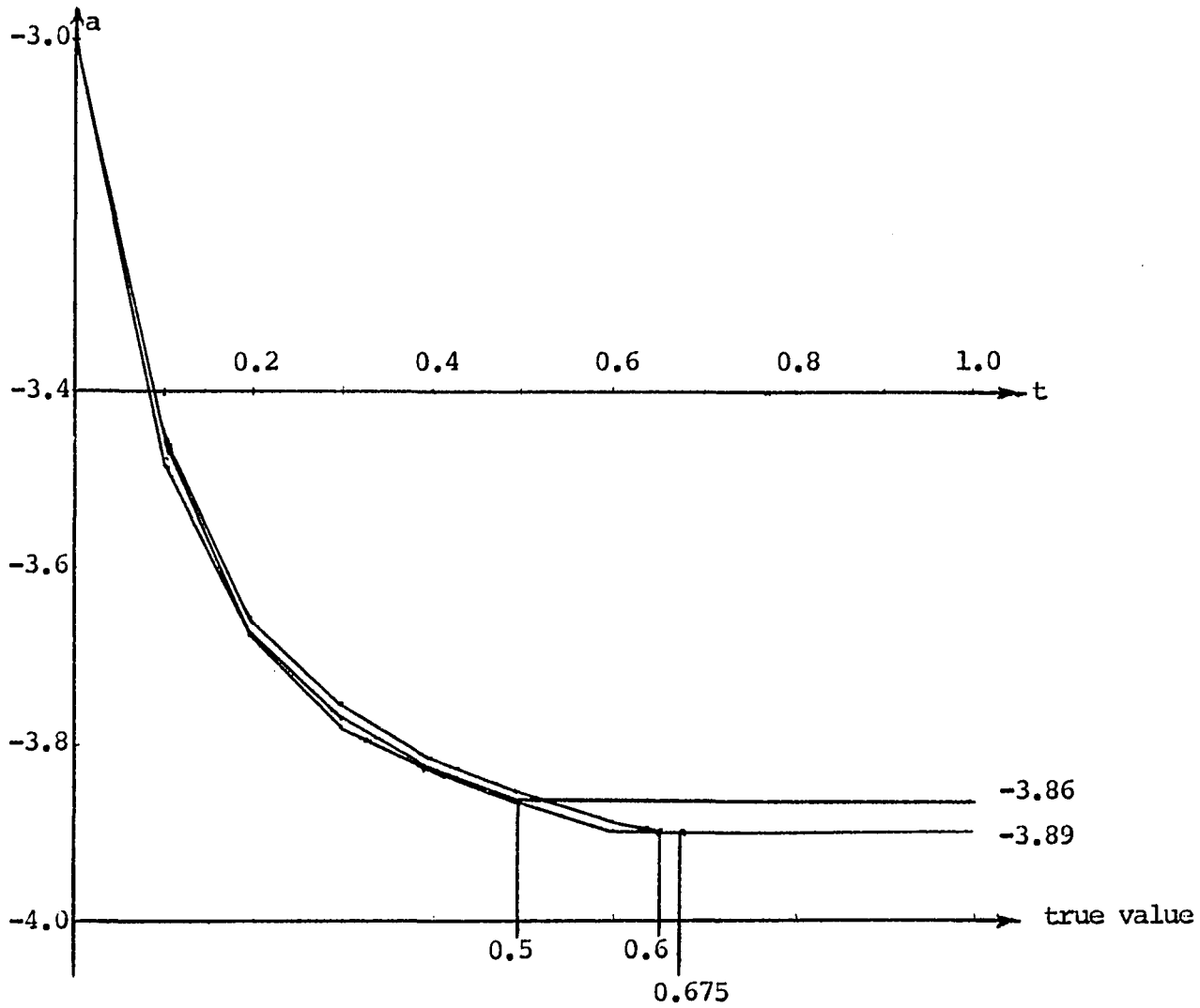
Figure 7.1.3.9 Identification of a for other noise inputs
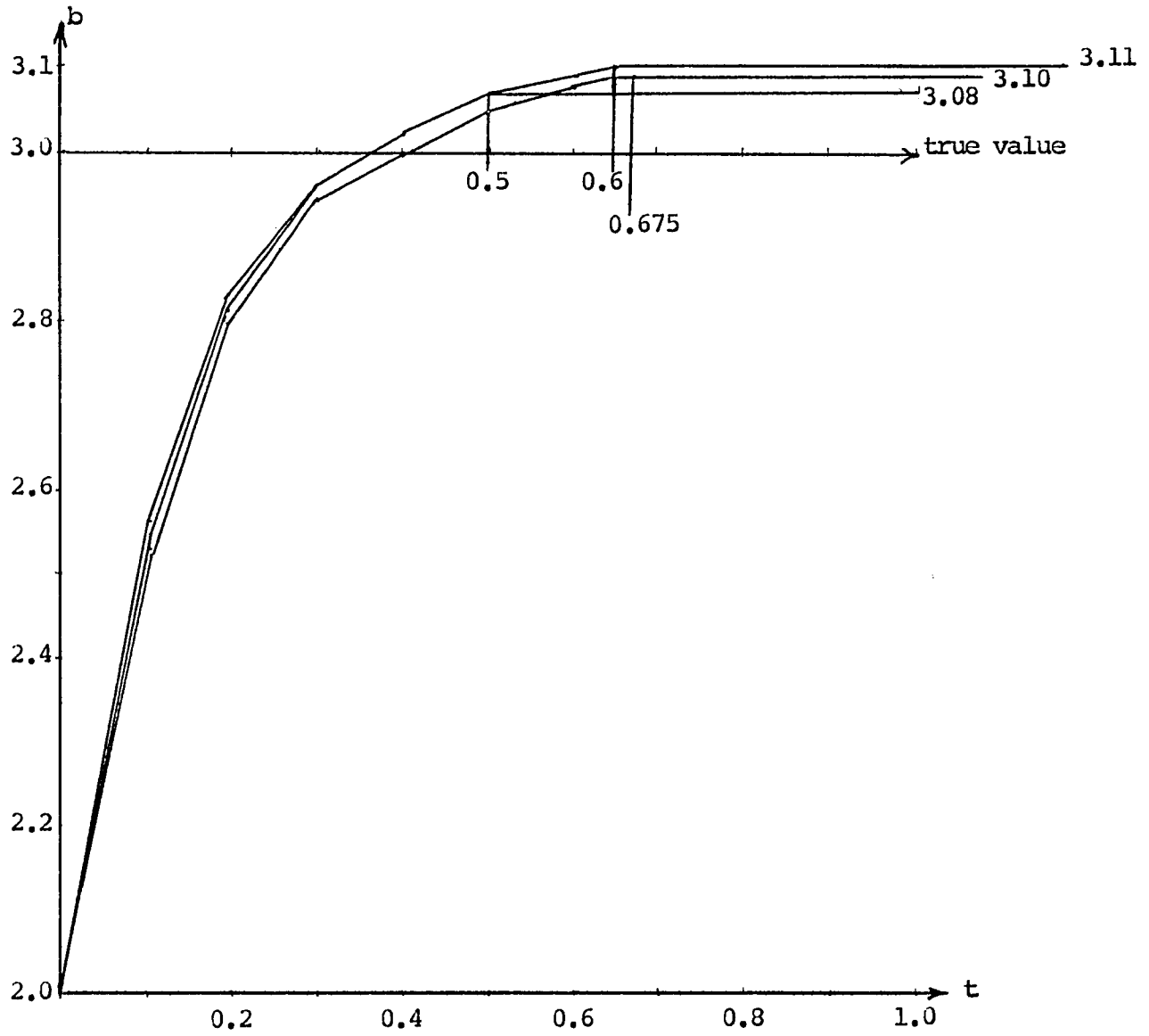
Figure 7.1.3.10 Identification of b for other noise inputs

The results of runs with different actual parameter
values are given in Figures 7.1.3.11 and 7.1.3.12 .
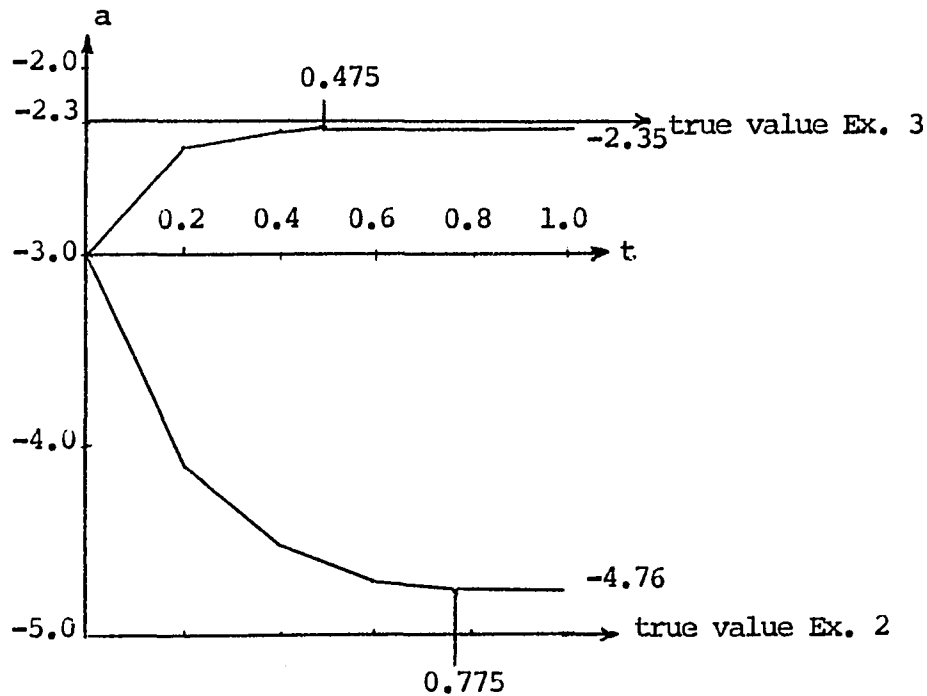
Figure 7.1.3.11 Identification of a for other values
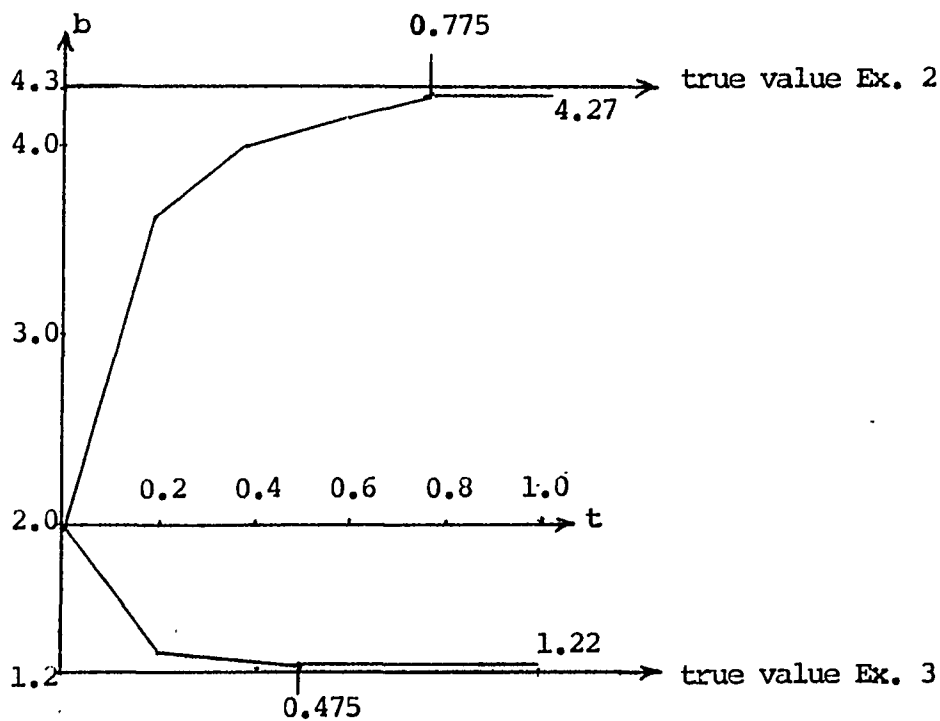
Figure 7.1.3.12 Identification of b for other values

As mentioned in Chapter 6, the difficult problems of
predetermining optimal sampling rates, rates of convergence,
regions of convergence and weights have not been solved explicitly,
but rather through establishing workable values by several trial runs.

The flowcharts and program listings of the adaptive
simulation are given in Appendix E.

7.2   Conclusions

The illustrative example presented in this chapter was
completly solved with respect to all aspects required for
simultaneous on-line identification and control. The
same techniques could be applied in principle to any problem
such that Theorems 4.1, 4.2, 5.1, and 5.2 are applicable.
That is, the case for vector state, control, and stochastic
disturbance can be solved.

CHAPTER 8

CONCLUSIONS

The study of second order stochastic processes
as input disturbances to linear time-varying systems
was dealt with in a number of ways. This class of
stochastic processes had been virtually unstudied
previously. .The main advantages of assuming these
disturbances are physically present, are that the
processes are continuous in the mean-square sense and
the first two moments are continuous and finite. As
such, these moments and the actual processes are time
integrable without any of the difficulties and limi-
tations encountered in white noise or colored noise
stochastic processes whose covariances can only be
written with Dirac delta functions. Further, the con-
tinuous time case can be treated independently of the
discrete time case in many areas where this is not true
for white noise or colored noise processes.

The first method of study was to approach a stochastic optimal control problem in a way such that optimal control was possible regardless of disturbance, by approximately the effects of the disturbance by the first two moments and casting the resulting moment equations into the form of a differential game. This approach was taken as a means to finding the "worst-case" optimal controls in the sense that optimal control was found for any set of first two moments of the stochastic disturbance. The initial studies of this approach led to both singular and non-singular performance indices. Also, the determination of optimal performance weighting constants was performed, corresponding to a greater degree of freedom in optimal system synthesis than is normally allowed.

The next area studied was to establish a method suitable for implementation in real-time for an actual

physical plant, that simultaneously identified system parameters and optimally or at least sub-optimally controlled the system. The particular technique used was a model reference adaptive system. The main advantage of this system is that it could be established in a recursive manner with an algorithm that is relatively easy to implement in terms of hardware and software. A complete illustrative example was given as simulated on a digital computer. Both identification using steepest descent and "worst case" suboptimal control were obtained. The overall model reference system can be shown to satisfy Lyapunov stability criteria, and the simulation verified that this stability did exist. The particular problems associated with gradient minimization, such as rate of convergence and region of convergence, were solved by several trial runs of the simulation rather than by explicit analytical techniques.

Finally, a best in the sense of minimum error variance, linear estimate was obtained for linear time-varying systems

with a stochastic input which was a second order process,
as well as a measurement including another second order
stochastic process possibly correlated with the input
process. This estimate is actually an unbiased minimum
variance estimate and was derived using variational
techniques similar to those used in deriving the "worst
case" optimal controls.

The three aspects of control, identification, and
estimation were studied for linear stochastic systems.
The stochastic processes utilized and studied throughout
were second order processes, characterized mainly by
having finite continuous covariances which was a very
useful analytic property.

CHAPTER 9

AREAS FOR FUTURE RESEARCH

The classes of "worst case" optimal control problems, i.e., the differential games could be extended to cover a slightly wider class in several ways. The addition of a terminal cost term to the performance index and the inclusion of a measurement of the state are examples of the possible extensions.

The study of nonlinear systems or linearized nonlinear systems is a major area for further investigation. In several ways, the differential game-moment treatment is more amenable to linearized nonlinear systems than present methods due to the presence already of both means and covariances.

The model reference adaptive control system has inherent in it several interesting side issues. Sensitivity, stability and overall adaptive optimality are possible points to consider further. The special problem of a true optimal control having a component for control only and a component for identification only may possibly be looked at in a setting very similar to the "worst case" situation already studied.

Similarly, the use of the "worst case" covariances and cross-covariances could be compared to lack of a priori statistical knowledge in the minimum error variance estimation problem.

Finally, a good solid area seemingly ripe to explore using the techniques for optimal control, identification and estimation for second order processes is in the study of various EEG phenomena. It is the fervent desire of this author to try out the different theories on several facets of these particular practical problems.

APPENDIX A.

DERIVATION OF SYSTEM EQUATIONS

The equations governing the propagation of the first two moments of the Class 3) stochastic process v through the dynamic system

$$\frac{d}{dt} x(t,\cdot) = A(t) \ x \ (t,\cdot) + B(t)u(t) + C(t)v(t,\cdot), \ t \in [t_0, t_f] \triangleq \Gamma_1 \quad \text{A-1}$$

are derived. Here

    x is nx1

    A is nxn

    B is nxr

    u is rx1

    C is rxs

    v is sx1 .

As is well known, the solution of A-1 is

$$x(t,\cdot) = \Phi(t,t_0) \ x \ (t_0,\cdot) + \int_{t_0}^{t} \Phi(t,\tau) [B(\tau)u(\tau) + C(\tau)v(\tau,\cdot)]d\tau, t \in \Gamma_1.$$

$$\text{A-2}$$

where

$\Phi(t,t_0)$ is the transition matrix found from

$$\frac{d}{dt} \Phi(t,t_0) = A(t)\Phi(t,t_0) \ , t \in \Gamma_1 \quad \text{A-3}$$

with initial condition

$$\Phi(t_0,t_0) = I. \quad \text{A-4}$$

Taking the expectation of A-1 with

$$E\{x(t,\cdot)\} \triangleq \mu_x(t) \text{ and } E\{v(t,\cdot)\} \triangleq \mu_v(t), \ E\{\frac{d}{dt} x(t,\cdot)\} = \frac{d}{dt} E\{x(t,\cdot)\} =$$

$$\frac{d}{dt} \mu_x(t) = E\{A(t)x(t,\cdot) + B(t)u(t) + C(t)\mu_v(t)\} =$$

$$A(t) \ E\{x(t,\cdot)\} + B(t)u(t) + C(t) \ E\{v(t,\cdot)\}$$

or finally

$$\frac{d}{dt} \mu_x(t) = A(t)\mu_x(t) + B(t)u(t) + C(t)\mu_v(t), t \in \Gamma_1 \qquad \text{A-5}$$

The solution of equation A-5 is given by taking the expectation

of A-2 and is

$$E\{x(t,\cdot)\} = \mu_x(t) =$$

$$E\{\Phi(t,t_0) \ x(t_0,\cdot) + \int_{t_0}^{t} \Phi(t,\tau)[B(\tau)u(\tau) + C(\tau)v(\tau,\cdot)]d\tau\} =$$

$$\Phi(t,t_0)E\{x(t_0,\cdot)\} + \int_{t_0}^{t} \Phi(t,\tau)[B(\tau)u(\tau) + C(\tau) \ E\{v(\tau,\cdot)\}]d\tau, t \in \Gamma_1,$$

or finally with $E\{x(t_0)\} \triangleq \mu_x(t_0)$, the prior mean, $\mu_x(t) =$

$$\Phi(t,t_0) \ \mu_x(t_0) + \int_{t_0}^{t} \Phi(t,\tau)[B(\tau)u(\tau) + C(\tau)\mu_v(\tau)]d\tau, t \in \Gamma_1. \qquad \text{A-6}$$

In applications either A-6 or A-5 would be used to determine

$\mu_x(t), t \in \Gamma_1$ depending on the specific case as mentioned in Sage and

Melsa (97).

The cross-covariance $V_{vx}(t,\tau)$, an nxs matrix, is found by differentiating

the expectation

$$\frac{\partial}{\partial \tau} E \{[v(t,\cdot) - \mu_v(t)] \ [x(\tau,\cdot) - \mu_x(\tau)]'\} = \frac{\partial}{\partial \tau} V_{vx}(t,\tau)$$

and using A-1 and A-5 evaluated at $t = \tau$ for all $\tau \in \Gamma_1$, which after

transposing become

$$\frac{d}{dt} x'(\tau,\cdot) = x'(\tau,\cdot)A'(\tau) + u'(\tau)B'(\tau) + v'(\tau,\cdot)C'(\tau), \tau \in \Gamma_1 \qquad \text{A-7}$$

and

$$\frac{d}{dt} \mu_x'(\tau) = \mu_x'(\tau)A'(\tau) + u'(\tau)B'(\tau) + \mu_v'(\tau)C'(\tau), \tau \in \Gamma_1 \qquad \text{A-8}$$

Then

$$\frac{\partial}{\partial \tau} E\{[v(t,\cdot) - \mu_v(t)] [x(\tau,\cdot) - \mu_x(\tau)]'\} = \frac{\partial}{\partial \tau} V_{vx}(t,\tau)$$

$$E\{[v(t,\cdot) - \mu_v(t)] [\frac{d}{d\tau} x'(\tau,\cdot) - \frac{d}{d\tau} \mu_x'(\tau)]\} =$$

$$E\{[v(t,\cdot) - \mu_v(t)] [x'(\tau,\cdot)A'(\tau) + u'(\tau)B'(\tau) + v'(\tau,\cdot)C'(\tau)$$

$$- \mu_x'(\tau)A'(\tau) - u'(\tau)B'(\tau) - \mu_v'(\tau)C'(\tau)]\} =$$

$$E\{[v(t,\cdot) - \mu_v(t)] \quad [x'(\tau,\cdot) - \mu_x'(\tau)]A'(\tau)$$

$$+ [v'(\tau,\cdot) - \mu_v'(\tau)]C'(\tau) \} =$$

$$V_{vx}(t,\tau)A'(\tau) + V_{vv}(t,\tau) C'(\tau), (t,\tau) \in \Gamma_1 x \Gamma_1 \triangleq \Gamma$$

or finally

$$\frac{\partial}{\partial \tau} V_{vx}(t,) = V_{vx}(t,\tau)A'(\tau) + V_{vv}(t,\tau)C'(\tau), (t,\tau) \in \Gamma. \qquad \text{A-9}$$

The solution of A-9 can be determined by using A-2 and A-6
and forming

$$E\{[v(t,\cdot) - \mu_v(t)] [x(\tau,\cdot) - \mu_x(\tau)]'\} = V_{vx}(t,\tau) =$$

$$E\{[v(t,\cdot) - \mu_v(t)] [x'(t_0,\cdot)\Phi'(\tau,t_0) + \int_{t_0}^{t} [v'(s,\cdot)C'(s)$$

$$+u'(s)B'(s)] x \Phi'(\tau,s)ds - \mu_x'(t_0)\Phi'(\tau,t_0) - \int_{t_0}^{t} [\mu_v'(s)C'(s)$$

$$+ u'(s)B'(s)] x \Phi'(\tau,s)ds\} =$$

$$E\{[v(t;) - \mu_v(t)] [x'(t_0;) - \mu_x'(t_0)]\Phi'(\tau,t_0) +$$

$$\int_{t_o}^{t} [v'(s,\cdot) - \mu_v'(s)]C'(s)\Phi'(\tau,s)ds]\} =$$

$$V_{vx}(t,t_o)\Phi'(\tau,t_o) + \int_{t_o}^{t} V_{vv}(t,s)C'(s)\Phi'(\tau,s)ds, (t,\tau)\in \Gamma$$

or finally

$$V_{vx}(t,\tau) = V_{vx}(t,t_o)\Phi'(\tau,t_o) + \int_{t_o}^{t} V_{vv}(t,s)C'(s)\Phi'(\tau,s)ds, (t,\tau) \in \Gamma. \quad \text{A-10}$$

Taking the transpose of A-9 and A-10 gives

$$\frac{\partial}{\partial \tau} V_{vx}'(t,\tau) = A(\tau)V_{vx}'(t,\tau) + C(\tau)V_{vv}'(t,\tau), (t,\tau)\in \Gamma \qquad \text{A-11}$$

and

$$V_{vx}'(t,\tau) = \Phi(\tau,t_o)V_{vx}'(t,t_o) + \int_{t_o}^{t} \Phi(\tau,s)C(s)V_{vv}(t,s)ds, (t,\tau)\in\Gamma \qquad \text{A-12}$$

The other possible cross-covariance is

$$V_{vx}(\tau,t) \triangleq E\{[v(\tau,\cdot) - \mu_v(\tau)] [x(t,\cdot) - \mu_x(t)]'\}$$

and from the definitions

$$V_{vx}(\tau,t) = V_{vx}'(t,\tau)$$

hence can be determined by interchanging t and τ upon solution of either A-9 or A-10.

For the solution of A-9 or A-10, the function of t at the boundary $V_{vx}(t,t_o)$ must be known. This function represents the a priori knowledge of the randomness of the state at initial time correlated with the stochastic input.

The equations of the auto-covariance or covariance of x are found by differentiating the expectation

$$\frac{\partial}{\partial t} E \{[x(t,\cdot) - \mu_x(t)] [x(\tau,\cdot) - \mu_x(\tau)]\} \triangleq \frac{\partial}{\partial t} V_{xx}(t,\tau) \qquad A-13$$

using A-1 and A-6 gives

$$\frac{\partial}{\partial t} E \{[x(t,\cdot) - \mu_x(t)] [x(\tau,\cdot) - \mu_x(\tau)]'\} = \frac{\partial}{\partial t} V_{xx}(t,\tau) =$$

$$E \{\frac{d}{dt}[x(t,\cdot) - \mu_x(t)] [x(\tau,\cdot) - \mu_x(\tau)]'\} =$$

$$E \{[A(t)x(t,\cdot) + B(t) u(t) + C(t)v(t,\cdot) - A(t)\mu_x(t) - B(t) u(t)$$

$$- C(t)\mu_v(t)] [x(\tau,\cdot) - \mu_x(\tau)]'\} =$$

$$E \{ A(t)[x(t,\cdot) - \mu_x(t)] + C(t) [v(t,\cdot) - \mu_v(t)] \quad [x(\tau,\cdot) - \mu_x(\tau)]'\}$$
$$(t,\tau)\in \Gamma$$

or finally

$$\frac{\partial}{\partial t} V_{xx}(t,\tau) = A(t)V_{xx}(t,\tau) + C(t)V_{vx}(t,\tau), \quad (t,\tau)\in \Gamma. \qquad A-14$$

The solution of A-14 can be found by using A-2 in

$$E \{[x(t,\cdot) - \mu_x(t)] [x(\tau,\cdot) - \mu_x(\tau)]\} =$$

$$E \{[\Phi(t,t_0)x(t_0,\cdot) + \int_{t_0}^{t} \Phi(t,s) [B(s)u (s) + C(s)v(s,\cdot)]ds$$

$$- \Phi(t,t_0)\mu_x(t_0) - \int_{t_0}^{t} \Phi(t,s) [B(s)u(s) + C(s)\mu_v(s)]ds \quad X$$

$$[x(\tau,\cdot) - \mu_x(\tau)]\} =$$

$$E \{ \Phi(t,t_0) [x(t_0,\cdot) - \mu_x(t_0)] + \int_{t_0}^{t} \Phi(t,s)C(s) [v(s,\cdot) - \mu_v(s)]ds \quad X$$

$$[x(\tau,\cdot) - \mu_x(\tau)]\} =$$

$$\Phi(t,t_0)V_{xx}(t_0,t) + \int_{t_0}^{t} \Phi(t,s)C(s)V_{vx}(s,\tau)ds, (t,\tau) \in \Gamma$$

or finally

$$V_{xx}(t,\tau) = \Phi(t,t_0)V_{xx}(t_0,\tau) + \int_{t_0}^{t} \Phi(t,s)C(s)V_{vx}(s,\tau)d\tau, \quad (t,\tau)\in\Gamma \ . \ \text{A-15}$$

Here the prior correlation of x at t_0 with itself for all τ must be known to complete the solution, i.e. $V_{xx}(t_0,\tau)$ is known.

Equations A-5, A-9, and A-14 are the differential equation representation of how the moments of v propagate through a system given by A-1. Equations A-6, A-10, and A-15 are the solutions of A-5, A-9, and A-14 with the specified initial conditions. For either representation these equations are for v a Class 3 stochastic process. They are the state-space form of the nth order differential equations in Papoulis (84) or the operator equations of Lebedev (63). In Sage and Melsa (97), Bryson and Ho (20), and Astrom (8), the same type of derivation is given for other classes of stochastic processes.

It is shown by various theorems in Bhat (17), Hoel (35), and Prabhu (90) that the interchange of operators required for the derivations in this appendix is valid for Class 3 stochatic processes. Further all the specified derivatives and integrals are shown to exist in these same three references.

It also can be shown that differentiating A-6, A-10, and A-15 gives A-5, A-9, and A-14 respectively; since in the derivations the

solution set was not obtained by directly integrating the differential equations. More will be said about the initial conditions in Appendix B.

The complete set in differential form is collected below. The mean of x is

$$\frac{d}{dt}\mu_x(t) = A(t)\mu_x(t) + B(t)u(t) + C(t)\mu_v(t), \quad t\epsilon\Gamma_1 \qquad \text{A-16}$$

where

μ_x is nx1

A is nxn

B is nxr

u is rx1

C is nxs

μ_v is sx1.

The cross variance of x and v is

$$\frac{\partial}{\partial\tau}V'_{vx}(t,\tau) = A(\tau)V'_{vx}(t,\tau) + C(\tau)V'_{vv}(t,\tau), \quad (t,\tau)\epsilon\Gamma \qquad \text{A-17}$$

where

V_{vx} is sxn

V'_{vx} is nxs

V_{vv} is sxs

V'_{vv} is sxs.

The auto-covariance of x is

$$\frac{\partial}{\partial t}V_{xx}(t,\tau) = A(t)V_{xx}(t,\tau) + C(t)V_{vx}(t,\tau), \quad (t,\tau)\epsilon\Gamma \qquad \text{A-18}$$

where V_{xx} is nxn.

It is now desired to derive the equivalent equations for Class 4)

stochastic inputs, which are characterized by having covariance.

$$V_W(t,\tau) = \Psi(t) \; \delta(t-\tau) \qquad\qquad \text{A-19}$$

where $\delta(t-\tau)$ is the scalar symmetric Dirac delta function

defined by

$$\int_{t_o}^{t_f} f(s)\delta(s-\tau)ds = \begin{cases} 0 & t_o > \tau > t_f \\ f(\tau) & t_o < \tau < t_f \\ f(t_o)/2 & \tau = t_o \\ f(t_f)/2 & \tau = t_f \end{cases} \qquad \text{A-20}$$

For the system of A-1 with $v(t,.)$ having covariance given

above, the mean of v propagates the same as for Class 3)

processes and is given by equations A-5 and A-6.

The cross covariance is

$$V_{xv}(t,\tau) = E\{[x(t,.) - \mu_x(t)] \; [v(\tau,.) - \mu_v(\tau)]'\}$$

and can be evaluated by post multiplying equation A-6 minus

A-2 by $[v(\tau,.) - \mu_v(\tau)]'$ and taking the expectation,

$$x(t,.)-\mu_x(t) = \Phi(t,t_o)[x(t_o,.)-\mu_x(t_o)]+\int_{t_0}^{t} \Phi(t,s)C(s)[v(s,.)-\mu_v(s)]ds$$

$$\text{A-21}$$

giving

$$V_{xv}(t,\tau)=\Phi(t,t_o)E\{[x(t_o,.)-\mu_x(t_o)][v(\tau,.)-\mu_v(\tau)]'\}$$

$$+\int_{t_o}^{t} \Phi(t,s)C(s)E\{[v(s,.)-\mu_v(s)][v(\tau,.)-\mu_v(\tau)]'\}ds \quad \text{A-22}$$

Normally it is assumed that $x(t_o,.)$ is uncorrelated with

$v(\tau,.), \tau \in \Gamma_1$, then

$$V_{xv}(t,\tau) = \int_{t_o}^{t} \Phi(t,s)C(s)V_{vv}(s,\tau)ds \ , \quad (t,\tau) \in \Gamma \qquad \text{A-23}$$

and knowing the form of $V_{vv}(s,t)$ from A-19 gives

$$V_{xv}(t,\tau) = \int_{t_o}^{t} \Phi(t,s)C(s)\delta(s-\tau)ds \qquad \text{A-24}$$

From A-20, this becomes

$$V_{xv}(t,\tau) = \begin{cases} 0 & t_o < t < \tau \\ C(\tau)\Psi(\tau)/2 & t_o < t = \tau \\ \Phi(t,\tau)C(\tau)\Psi(\tau) & t_o < \tau < t \end{cases} \qquad \text{A-25}$$

and it is clear that a discontinuity occurs at $t = \tau$.

Post multiplying A-21 by $[x(\tau,.)-\mu_x(\tau)]'$ and

taking the expectation gives

$$V_{xv}(t,\tau) = E\{[x(t,.)-\mu_x(t)][x(\tau,.)-\mu_x(\tau)]'\} = \Phi(t,t_o)V_{xx}(t_o,t_o)\Phi'(\tau,t_o)$$

$$+\Phi(t,t_o)E\{[x(t_o,.)-\mu_x(t_o)][\int_{t_o}^{\tau}\Phi(\tau,s)C(s)[v(s,.)-\mu_v(s)]ds]'\}$$

$$+E\{[\int_{t_o}^{t}\Phi(t,s)C(s)[v(s,.)-\mu_v(s)]][\Phi(\tau,t_o)[x(t_o,.)-\mu_x(t_o)]]'ds\}$$

$$+E\{[\int_{t_o}^{t}\Phi(t,s)C(s)[v(s,.)-\mu_v(s)]ds][\int_{t_o}^{\tau}\Phi(\tau,\sigma)C(\sigma)[v(\sigma,.)-\mu_v(\sigma)]d\sigma]'\}$$
$$\text{A-26}$$

Again it is assumed that $x(t_o,.)$ and $v(\tau,.)$ are uncorrelated

resulting in the second and third terms going to zero.

Rearranging,

$$V_{xx}(t,\tau) = \Phi(t,t_o)V_{xx}(t_o,t_o)\Phi'(\tau,t_o)$$

$$+\int_{t_o}^{t}\int_{t_o}^{\tau}\Phi(t,s)C(s)E\{[v(s,.)-\mu_v(s)][v(\sigma,.)-\mu_v(\sigma)]'\}C'(\sigma)\Phi'(t,\sigma)d\sigma ds$$
$$\text{A-27}$$

and substituting A-19 gives

$$V_{xx}(t,\tau) = \Phi(t,t_o)V_{xx}(t_o,t_o)\Phi^{*}(\tau,t_o)$$

$$+\int_{t_o}^{t}\int_{t_o}^{\tau}\Phi(t,s)C(s)\Psi(s)\delta(s-\sigma)C^{*}(\sigma)\Phi^{*}(\tau,\sigma)d\sigma ds \qquad A-28$$

In this last expression the order in which the double integration is carried out must be selected carefully. If $t > \tau$, then first integrate with respect to s in order to obtain a range where the delta function exists. If $\tau > t$, then first integrate with respect to σ for the same reason.

Rewriting A-28 with this reasoning gives

$$V_{xx}(t,\tau) = \Phi(t,t_o)V_{xx}(t_o,t_o)\Phi^{*}(\tau,t_o)$$

$$+ \int_{t_o}^{\min[t,\tau]}\Phi(t,\eta)C(\eta)\Psi(\eta)C^{*}(\eta)\Phi^{*}(\tau,\eta)d\eta$$

$$A-29$$

Restricting attention to only the case where $t = \tau$ gives

$$V_{xx}(t,\tau) = \Phi(t,t_o)V_{xx}(t_o,t_o)\Phi^{*}(\tau,t_o)$$

$$+ \int_{t_o}^{t}\Phi(t,\eta)C(\eta)\Psi(\eta)C^{*}(\eta)\Phi^{*}(t,\eta)d\eta$$

$$A-30$$

Now, A-30 is the solution of an ordinary differential equation which can be obtained by differentiating and using Leibnitz's rule for differentiation under an integral which is

$$\frac{\partial}{\partial t}\int_{\alpha(t)}^{\beta(t)}f(t,\tau)d\tau = \int_{\alpha(t)}^{\beta(t)}\frac{\partial f(t,\tau)}{\partial t}d\tau + f(t,\beta)\frac{d\beta(t)}{dt} - f(t,\alpha)\frac{d\alpha(t)}{dt}$$

$$A-31$$

Using A-31, A-30 becomes

$$\frac{dV_{xx}}{dt}(t,t)=\frac{d\Phi}{dt}(t,t_0)V_{xx}(t_0,t_0)\Phi^-(t,t_0)+\Phi(t,t_0)V_{xx}(t_0,t_0)\frac{d\Phi^-}{dt}(t,t_0)$$

$$+\int_{t_0}^{t}\frac{d\Phi}{dt}(t,\eta)C(\eta)\Psi(\eta)C^-(\eta)\Phi^-(t,\eta)d\eta+\int_{t_0}^{t}\Phi(t,\eta)C(\eta)\Psi(\eta)C^-(\eta)\frac{d\Phi^-}{dt}(t,\eta)d\eta$$

$$+\Phi(t,t)C(t)\Psi(t)C^-(t)\acute{\Phi}(t,t)\frac{dt}{dt}-0 \qquad\qquad A-32$$

But $\Phi(t,t)=I$ and $\frac{d\Phi}{dt}(t,\tau)=A(t)\Phi(t,\tau)$, $t\in\Gamma_1$

which gives

$$\frac{dV_{xx}}{dt}(t,t)=A(t)[\Phi(t,t_0)V_{xx}(t_0,t_0)\Phi^-(t,t_0)]+[\Phi(t,t_0)V_{xx}(t_0,t_0)\Phi^-(t,t_0]A^-(t)$$

$$+A(t)[\int_{t_0}^{t}\Phi(t,\eta)C(\eta)\Psi(\eta)C^-(\eta)\Phi^-(t,\eta)d\eta]+[\int_{t_0}^{t}\Phi(t,\eta)C(\eta)\Psi(\eta)C^-(\eta)\Phi^-(t,\eta)d\eta]A^-(t)$$

$$+C(t)\Psi(t)C^-(t) \qquad\qquad A-33$$

and from A-30

$$\frac{dV_{xx}}{dt}(t,t)=A(t)V_{xx}(t,t)+V_{xx}(t,t)A^-(t)+C(t)\Psi(t)C^-(t), t\in\Gamma_1 \quad A-34$$

Equation A-34 is used throughout much of the literature dealing with estimation theory. The major difference between this equation and equations A-17 and A-18 which hold for Class 3) stochastic processes is seen in that equation A-34 holds only in the plane with $t=\tau$, but equations A-17 and A-18 hold throughout the square $(t,\tau)\in[t_0,t_f]X[t_0,t_f]$. Further, as remarked on page 106, $V_{xv}(t,\tau)|t=\tau$ is a point of discontinuity in the square, and equation A-34 includes the effects of $V_{xv}(t,t)$ hence includes the effects of this discontinuity, in fact A-34 only is true along a line where $V_{xv}(t,\tau)$ is everywhere discontinuous.

Therefore, it is seen that much more information is implied by equations A-17 and A-18 for Class 3) processes than by equation A-34 for Class 4) processes.

An alternate derivation of equation A-34 can be performed by differentiating the definition of the covariance of x, as

$$\frac{d}{dt} E\{[x(t,.)-\mu_x(t)][x(t,.)-\mu_x(t)]\} = \frac{dV_{xx}}{dt}(t,t) =$$

$$E\{[\frac{dx(t,.)}{dt}-\frac{d\mu}{dt}_x(t)][x(t,.)-_x(t)]\}+E\{[x(t,.)-\mu_x(t)][\frac{dx(t,.)}{dt}-\mu_x(t)]^\prime\}=$$

$$E\{[A(t)[x(t,.)-\mu_x(t)]+C(t)[v(t,.)-\mu_v(t)]][x(t,.)-\mu_x(t)]^\prime\}$$

$$+E\{[x(t,.)-\mu_x(t)][[v(t,.)-\mu_v(t)]^\prime C^\prime(t)+[x(t,.)-\mu_x(t)]^\prime A^\prime(t)]\}=$$

$$A(t)V_{xx}(t,t)+C(t)V_{vx}(t,t)+V_{xv}(t,t)C^\prime(t)+V_{xx}(t,t)A^\prime(t)$$

A-35

This equation shows the explicit dependence of $V_{xx}(t,t)$ on

$V_{xv}(t,t)$

From A-25,

$$V_{xv}(t,t)=C(t)\Psi(t)/2$$

and

$$V_{vx}(t,t)=V_{xv}^\prime(t,t)=\Psi^\prime(t)C^\prime(t)/2$$

therefore, substituting in A-35 gives

$$\frac{dV_{xx}}{dt}(t,t)=A(t)V_{xx}(t,t)+V_{xx}(t,t)A^\prime(t)+C(t)\Psi^\prime(t)C^\prime(t)/2+C(t)\Psi(t)C^\prime(t)/2$$

A-36

and since $\Psi(t)$ is symmetric, A-36 is the same as A-34.

Evaluating A-30 at t=τ and premultiplying by $\Phi(t,\tau)$ gives

$$V_{xx}(\tau,\tau)=\Phi(t,\tau)\Phi(\tau,t_o)V_{xx}(t_o,t_o)\Phi^\prime(\tau,t_o)+$$

$$\int_{t_o}^{\tau}\Phi(t,\tau)\Phi(\tau,\eta)C(\eta)\Psi(\eta)C^\prime(\eta)\Phi^\prime(\tau,\eta)d\eta$$

A-37

but since $\Phi(t,\tau)\Phi(\tau,\beta)=\Phi(t,\beta)$ for $t>\tau$ if it is restricted to $t>\tau$,

A-37 is equal to equation A-29. Hence for $t>\tau$,

$$V_{xx}(t,\tau)=\Phi(t,\tau)V_{xx}(\tau,\tau) \qquad\qquad A\text{-}38$$

and similarly for $\tau>t$,

$$V_{xx}(t,\tau)=V_{xx}(t,t)\Phi'(\tau,t) \qquad\qquad A\text{-}39$$

These last two equations enable more information to be obtained

since they permit the covariance of x to be found over the whole

square.

However, there is no convenient differential equation form

for A-38 and A-39, and the transition matrix is required. The

major differences then in the equations for Class 3) and Class 4)

stochastic processes are:

1) All the first two moments of state and input are

continuous everywhere in the square for Class 3) processes

but not for Class 4),

2) The evaluation of all the first two moments of state and

input for Class 3) processes does not require knowledge of the

transition matrix whereas this knowledge is required for Class

4) processes,

3) For Class 4) processes, two separate evaluations are

required for the cases $t>\tau$ and $\tau>t$, after obtaining all data at

$t=\tau$, but this separation of the square is not at all needed for

Class 3) processes,

and

4) The equations developed for Class 4) processes only hold

for "white" or "not time correlated" inputs whereas the equations

for the Class 3) processes allow any type of correlation, including
memory. "Non-white" or "time correlated" models can be obtained
for Class 4) inputs with "white" noise driving a prefilter
becoming the input to the plant.

The "colored noise" representation is

$$\frac{dx(t,.)}{dt}=A(t)x(t,.)+B(t)u(t)+C(t)v(t,.)$$
<div align="right">A-40</div>

where

$$v(t,.)=\Lambda(t)\gamma(t,.)+\beta(t,.)$$
<div align="right">A-41</div>

and

$$\frac{d\gamma(t,.)}{dt}=\Xi(t)\gamma(t,.)+T(t)\alpha(t,.)$$
<div align="right">A-42</div>

with

$$V_{\alpha\beta}(t,\tau)=0 \quad (t,\tau)\in\Gamma$$

$$V_{\beta\alpha}(t,\tau)=0 \quad (t,\tau)\in\Gamma$$

$$V_{\beta\beta}(t,\tau)=\Psi_{\beta}(t)\delta(t-\tau)$$

$$V_{\alpha\alpha}(t,\tau)=\Psi_{\alpha}(t)\delta(t-\tau)$$

Since $\alpha(t,.)$ is "white noise", the covariance of γ is found
using A-29, as

$$V_{\gamma\gamma}(t,\tau)=\Phi_{\gamma}(t,t_{o})V_{\gamma\gamma}(t_{o},t_{o})\Phi^{\sim}(\tau,t_{o})$$

$$+\int_{t_{o}}^{min[t,\tau]}\Phi_{\gamma}(t,\eta)T(\eta)\Psi_{\alpha}(\eta)T(\eta)\Phi_{\gamma}^{\sim}(\tau,\eta)d\eta$$
<div align="right">A-43</div>

where

$$\frac{d\Phi_{\gamma}(t,t_{o})}{dt}=\Xi(t)\Phi_{\gamma}(t,t_{o})$$

and

$$\Phi_{\gamma}(t_o, t_o) = I$$

From A-43, the explicit dependence of $V_{\gamma\gamma}$ on t and τ is exhibited by the τ argument in the transition matrix Φ_{γ}. Then the state covariance V_{xx} will depend on τ through this means. The complete solution is obtained from the augmented model

$$\frac{d}{dt}\begin{bmatrix} x(t,.) \\ \gamma(t,.) \end{bmatrix} = \begin{bmatrix} A(t) & C(t)\Lambda(t) \\ 0 & \Xi(t) \end{bmatrix}\begin{bmatrix} x(t,.) \\ \gamma(t,.) \end{bmatrix} + \begin{bmatrix} B(t) & 0 \\ 0 & 0 \end{bmatrix} u(t)$$

$$+ \begin{bmatrix} C(t) & 0 \\ 0 & T(t) \end{bmatrix}\begin{bmatrix} \beta(t,.) \\ \alpha(t,.) \end{bmatrix} \qquad \text{A-44}$$

rewritten with obvious definitions as

$$\frac{dx^*(t,.)}{dt} = A^*(t)x^*(t,.) + B^*(t)u(t) + C^*(t)v^*(t,.) \qquad \text{A-45}$$

with v^* a "white noise" process. The covariance of x^* is found using A-29 as

$$V_{x^*x^*}(t,\tau) = \Phi^*(t,t_o)V_{x^*x^*}(t_o,t_o)\Phi^{*\prime}(\tau,t_o)$$

$$+ \int_{t_o}^{\min[t,\tau]} \Phi^*(t,\eta)C^*(\eta)\Psi_{v^*}(\eta)C^{*\prime}(\eta)\Phi^{*\prime}(\tau,\eta)d\eta \qquad \text{A-46}$$

where

$$\frac{d\Phi^*(t,t_o)}{dt} = A^*(t)\Phi^*(t,t_o)$$

and

$$\Phi^*(t_o, t_o) = I$$

and

$$V_{v^*v^*}(t, \tau) = \begin{bmatrix} \Psi_\beta(t) & 0 \\ 0 & \Psi_\alpha(t) \end{bmatrix} \qquad \delta(t-\tau) = \Psi_{v^*}(t)\,\delta(t-\tau)$$

The solution of A-46 enables the retrieval of the covariance of the original state x driven by "colored noise" v to be obtained.

The "colored noise" v with covariance depending on t and τ is obtained through the artificial use of the prefilter of equation A-42. It is obvious then that $\Lambda(t)$, $\Xi(t)$ and $T(t)$ must be linear time varying or linear constant matrices. Nonlinear memory type elements could not be present. Therefore, only processes with Markov properties can be treated. This restriction is not present in the development of the second moment equations for Class 3) processes.

APPENDIX B

TRANSFORMATIONS OF PERFORMANCE INDEX AND ENDPOINT CONDITIONS

B.1 Generalized Performance Index

The typical "minimum energy" performance index including all
energies present in a physical plant was obtained in Section 3.1.2
as the expectation of

$$J_S = \langle x(t_f,\cdot), Q(t_f) \ x \ (t_f,\cdot) \rangle$$

$$+ \int_{t_o}^{t_f} \{\langle x(t,\cdot), R_1 x(t,\cdot) \rangle + \langle u(t), R_2 u(t) \rangle + \langle v(t,\cdot), R_3 v(t,\cdot) \rangle\} dt \qquad \text{B.1-1}$$

Redefine B.1-1 as

$$J_S = K(t_f, x(t_f)) + J_{xs} + J_{us} + J_{vs} \qquad\qquad \text{B.1-2}$$

where the various terms are obviously defined. For x a stochastic
process, the criteria most often used in the literature is

$$J_{m,x} \triangleq E\{J_{xs}\} = E\{\int_{t_o}^{t_f} \langle x(t,\cdot), R_1 x(t,\cdot) \rangle dt \qquad\qquad \text{B.1-3}$$

which is a measure of the mean value of state energy. However,
other means to make J_{xs} a deterministic number appear. Sain and
Liberty (98) use the minimum variance value of state energy

$$J_{mv,x} \triangleq E\{[J_{xs} - \{J_{xs}\}]^2\} \qquad\qquad \text{B.1-4}$$

Pugachev (91) and Andreev (5) use the other measures

$$J_{ms,x} \triangleq E\{J_{xs}^2\} \qquad\qquad \text{B.1-5}$$

a minimum mean square value of state energy and

$$J_{s,x} \triangleq [E\{J_{xs}\}]^2$$ B.1-6

a minimum mean-squared value of state energy. Murphy (81)

introduces arbitrary weighting of these measures. Rekasius (93)

and Sherman(103) define further possible modifications. The

original work of Kalman and Bucy (40) in estimation theory, and

its extensions such as in Mehra (74), Sage and Melsa (97), Bryson

and Ho (20), Astrom (8) and Kushner (61), all use the mean value

of cost.

A more general performance index would be the weighed sum of

these four, defined as

$$IP_x \triangleq a_{1x}J_{m,x} + a_{2x}J_{mv,x} + a_{3x}J_{ms,x} + a_{4x}J_{s,x}$$ B.1-7

where a_{ix}, $i = 1, \cdots, 4$ are constants.
Similarly, define

$$IP_v \triangleq a_{1v}J_{m,v} + a_{2v}J_{mv,v} + a_{3v}J_{ms,v} + a_{4v}J_{s,v}$$ B.1-8
and

$$IP_u \triangleq a_{1u}J_{m,u} + a_{2u}J_{mv,u} + a_{3u}J_{ms,u} + a_{4u}J_{s,u}$$ B.1-9
and finally the overall measure

$$IP \triangleq IP_x + IP_v + IP_u$$ B.1-10

Several interesting sidelights arise when all of the terms are

collected as in B.1-10. Consideration of this measure may result

in solutions to the propositions of Guillemin (29) about finding

Nature's error criteria.

The optimal selection of the a_{ix}, a_{iv} and a_{iu}, $i = 1, \cdots, 4$ is discussed with respect to optimal system synthesis by Andreyev (6) and others. A particular version of this type of problem is included as Problem 4).

The separation of state, control and disturbance energies precludes the occurence of cross-terms between these variables. There would be many cross terms present if the four measures of J_s in B.1-1 were taken and summed. These cross terms are not considered here as is done in most of the literature.

B.2 Performance Indices of Problems 1 and 2

The choice of

$$a_{3x} = -\frac{1}{2}$$

$$a_{3v} = -\frac{1}{2}$$

$$a_{4u} = +\frac{1}{2}$$

$$a_{4v} = +\frac{1}{2} \qquad\qquad\qquad \text{B.2-1}$$

and the rest of the a's zero in B.1-10 yields a performance index related to the Problem 1 performance index, though not exactly equal. The relation is established through the inequality

$$-\int_{t_o}^{t_f}\int_{t_o}^{t_f} tr\ R_2 u_3(t,\tau)u_3'(t,\tau)R_2'dtd\tau \geqq -\int_{t_o}^{t_f}\int_{t_o}^{t_f} tr\ R_2 u_3(t,t)\,tr\ R_2 u_3(\tau,\tau)dtd\tau$$

$$\text{B.2-2}$$

which can be shown to hold using some basic properties of covariances as in Papoulis (84), some theorems on positive definite matrices as in Hohn (36) and Graybill (32), and some inequalities as in Mitrinovic (77), (78), Beckenbach and Bellman (15) and Marcus and Minc (72).

In addition to the terms of the problem 1 performance index, a term linear in tr $R_2 u_3^{\sim}(t,t)$ will arise. Using the theory in Kleindorfer and Kleindorfer (46) and Athans (11), the cost of this term can be shown to be included in the RHS of B.2-2.

Alternately, the performance index of problem 1 can be treated as a function of other indices as in Petrov (87) and Andreyev (6) yielding a similar relation.

A detailed analysis of B.1-10 would show that all terms would be functions of the first and second moments of state and disturbance. Including all of these moments under a double integral yields the functional performance index of Problem 2.

B.3 Transformation of Problem 3 Performance Index

The performance index of Problem 3, equation 4.1.3.1 is

$$J(u_1,u_2,u_3^{\sim}) = \int_{t_o}^{t_f}\int_{t_o}^{t_f} \{ f_{oo}(z_1(t),z_1(\tau),z_2^{\sim}(t,\tau),z_3(t,\tau),u_1(t),u_1(\tau),u_2(t),u_2(\tau))$$

$$+ \sum_{k=1}^{s}\sum_{m=1}^{s} f_{km}(z_2^{\sim},z_3)u_{3km}^{\sim} \}dtd\tau \qquad\qquad\qquad B.3\text{-}1$$

where B.3-1 represents a functional linear in u_3^- and nonlinear

in z_1, z_2^-, z_3, u_1 and u_2. The problem then becomes singular with

respect to u_3^-, but remains nonsingular with respect to u_1 and u_2.

Since the state equations are also linear in u_3^-, the Hamiltonian, H

is linear in u_3^- and the necessary condition of equation 4.2.2.12

does not enable u_3^{-*} to be found. Therefore the higher order

necessary conditions of Theorem 4.3 had to be developed.

A typical performance index in the form of equation 4.1.3.1

arises with the choice of

$$a_{4x} = - \frac{1}{2}$$

$$a_{4u} = + \frac{1}{2}$$

$$a_{3v} = - \frac{1}{2}$$

in the generalized measure of equation B.1-10. With the aid of

inequality B.2-2, a specific measure singular, hence

linear, in u_3^- and non-singular in u_1 and u_2, but nonlinear in z_3 is

$$J(u_1,u_2,u_3^-) = \int_{t_o}^{t_f}\int_{t_o}^{t_f} \{-1/2\mathrm{tr}[R_1 z_3(t,\tau)z_3^-(t,\tau)R_1^-] + 1/2<u_1(t),R_2 u_1(t)>$$

$$-1/2\mathrm{tr}[R_3 u_3^-(t,\tau)]-1/2<u_2(t),R_3 u_2(\tau)>\}dtd\tau \qquad B.3-2$$

Other terms in the form of $f_{1k}(z_2,z_3)u_{3_{1k}}$ would arise naturally

by considering the cross terms which were neglected in

developing equation B.1-10 .

Since B.3-1 is linear in u_3, Problem 3 is a singular control problem and much more difficult to solve than problems with quadratic criteria. This fact led to the formulation of Problem 2) and Problem 1).

B.4 Endpoint Conditions

The choice of performance indices for problems 1), 2) and 3) constrains the problems to be fixed endpoint problems. If a terminal cost term were present, free endpoint conditions would arise, but this case is not covered in this research.

Therefore it becomes necessary to specify the state z_1, z_2', z_3 at t_0 and at fixed finite final time t_f as

$$z_1(t_0) = E\{x(t_0,\cdot)\} = \mu_x(t_0) \qquad \text{B.4-1}$$

$$z_1(t_f) = E\{x(t_f,\cdot)\} = \mu_x(t_f) \qquad \text{B.4-2}$$

$$z_2'(t,t_0) = E\{[v(t,\cdot) - \mu_v(t)][x(t_0,\cdot) - \mu_x(t_0)]'\} \qquad \text{B.4-3}$$

$$z_2'(t,t_f) = E\{[v(t,\cdot) - \mu_v(t)][x(t_f,\cdot) - \mu_x(t_f)]'\} \qquad \text{B.4-4}$$

$$z_3(t_0,\tau) = E\{[x(t_0,\cdot) - \mu_x(t_0)][x(\tau,\cdot) - \mu_x(\tau)]'\} \qquad \text{B.4-5}$$

$$z_3(t_f,\tau) = E\{[x(t_f,\cdot) - \mu_x(t_f)][x(\tau,\cdot) - \mu_x(\tau)]'\} \qquad \text{B.4-6}$$

Graphically B.4-1 through B. 4-6 are given in the following figures. In these figures the functions plotted are scalars, but a similar figure would apply for each element of the vector-matrix state. Assume

$$z_3(t,\tau) = \Phi(t)e^{-\frac{|t-\tau|}{T}} \qquad \text{B.4-7}$$

and

$$\Phi(t) = z_3(t,t) \qquad \text{B.4-8}$$
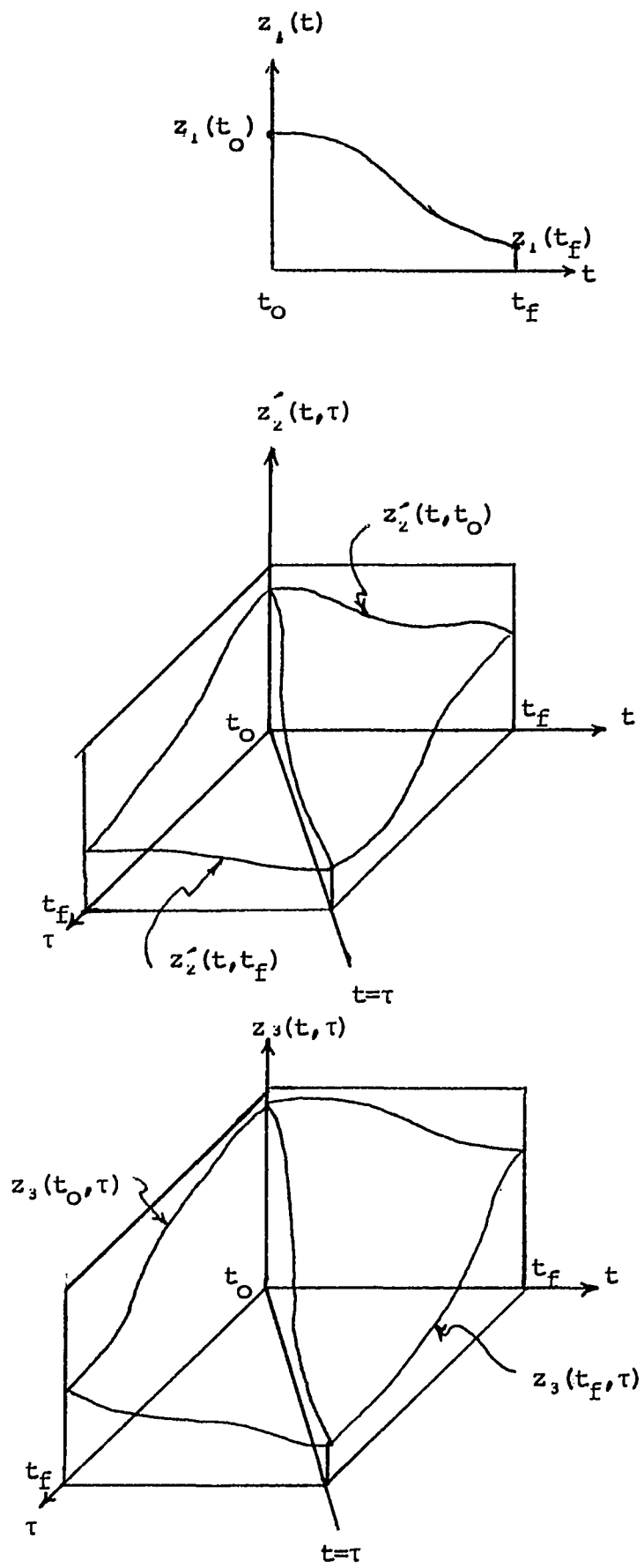
Figure B.4.1 Endpoint conditions

then

$$z_3(t_0, \tau) = \Phi(t_0) e^{\dfrac{-|t_0 - \tau|}{T}}$$

and for $\tau \geq t_0$

$$z_3(t_0, \tau) = \Phi(t_0) e^{\dfrac{-(\tau - t_0)}{T}}$$

$$= \Phi(t_0) e^{\dfrac{t_0}{T}} e^{-\dfrac{\tau}{T}}$$

$$= K e^{-\dfrac{\tau}{T}} \qquad\qquad\qquad \text{B.4-9}$$

and this is the form assumed in example 1 for $T = 1$. Similar remarks hold for $z_2'(t, t_0)$. Then since $z_2'(t_0, t_0)$ and $z_3(t_0, t_0)$ can be measured as is described in Kalman (40), and T also can be measured for a physical process, with the form in B.4-9, the initial point conditions are known.

The conditions at t_f result from the initial condition response and response due to the transition matrices of $A(t)$ and $- A'(t)$. It would be desired to drive $z_3(t_f, t_f)$ and $z_2(t_f, t_f)$ to a smaller value than at t_0, and this can be accomplished through the above analysis and selection of initial conditions on λ_2 and λ_3 which satisfy the TPBVP. The mean value $z_1(t_0)$ can be measured hence is assumed given.

NECESSARY CONDITIONS

It was proposed that the geometric approach of Kuo in (50) and (52), and Leitmann(67) be employed in deriving necessary conditions. The choice of Class 3) stochastic disturbances, and the resulting form of equations for the dynamic system and performance index after transformation precluded the use of the geometric approach. It was chosen to employ variational techniques instead for all four classes of problems. Standard forms of variational approaches are found in Kirk(45), Citron (22), Bryson and Ho (20) and Athans and Falb (9). Since some of the state and control variables are matrices rather than vectors, the results of standard forms of variational approaches could not be directly applied. Instead, matrix variations had to be defined and the entire proof had to be carried out. A formal extension of Pontriagin's minimum principle, found in Pontriagin et al (88) , to the matrix variable case was given by Athans in (10), but since the form of the equations included dependence on two independent variables, t and τ, the result of that derivation could not be directly applied either. The dynamic optimisation of criteria and constraints with more than one independent variable is briefly described for the vector case in Gottfried and Weisman (30). Introductory distributed parameter optimal control theory is covered in Sage(95), but this the results of this theory cannot be directly applied either.

C.1 Proof of Theorem 4.1:

For Problem 1 of Section 4.1.1 defined by equations 4.1.1.2
through 4.1.1.8 with the performance index of 4.1.1.1, define

$\lambda_1(t)$ an nx1 vector

$\lambda_2(t,\tau)$ an nxs matrix

$\lambda_3(t,\tau)$ an nxn matrix

and form the augmented performance index

$$Ja = J (z_1,z_2',z_3,u_1,u_2,u_3',\lambda_1,\lambda_2,\lambda_3) =$$

$$-\frac{1}{2} \int_{t_o}^{t_f}\int_{t_o}^{t_f} \{ tr[R_1 z_3(t,\tau)z_3'(t,\tau)R'] + tr[R_2 u_3(t,\tau)u_3'(t,\tau)R_2']$$

$$+ tr [[A(\tau)z_2'(t,\tau) + C(\tau)u_3'(t,\tau) - \frac{\partial z_2'}{\partial \tau}(t,\tau)] [\lambda_2'(t,\tau)]]$$

$$+ tr [[A(t)z_3'(t,\tau) + C(t)z_2'(t,\tau) - \frac{\partial z_3}{\partial t}(t,\tau)] [\lambda_3'(t,\tau)]] \}dtd\tau$$

$$+ \frac{1}{2} \int_{t_o}^{t_f} \{<u_1(t),R_3 u_1(t) > - < u_2(t),R_4 u_2(t)>$$

$$+ <\lambda_1(t),A(t)z_1(t) + B(t)u_1(t) + C(t)u_2(t) - \frac{dz_1(t)}{dt}>\} dt$$

C.1-1

noting that Ja = J of 4.1.1.1. .

Define Ω_2 as the scalar function

$$\Omega_2(z_2',z_3, \frac{\partial z_2'}{\partial \tau}, \frac{\partial z_3}{\partial \tau}, u_3',\lambda_2,\lambda_3) = -\frac{1}{2} tr\ R_1 z_3 z_3'R_1'$$

$$-\frac{1}{2} tr\ R_2 u_3 u_3'R_2 + tr\ [[Az_2' + Cu_3' - \frac{\partial z_2}{\partial \tau}\ ] [\lambda_2'\ ]]$$

Then

$$\delta J = \int_{t_0}^{t_f}\int_{t_0}^{t_f} \delta\Omega_2 \, dt d\tau + \int_{t_0}^{t_f} \delta\Omega_1 \, dt \qquad\qquad C.1\text{-}7$$

The variation of a scalar function $g(A)$ of an $n\times m$ matrix A is

$$\delta g(A) = tr\ [\delta A\ (\frac{\partial g(A)}{\partial A})^{\hat{}}]_* \qquad\qquad C.1\text{-}8$$

where

$$\frac{\partial g(A)}{\partial A} \qquad\qquad C.1\text{-}9$$

is an $n\times m$ gradient matrix and

$$\delta A = A - A^* \qquad\qquad C.1\text{-}10$$

Using C.1-8,

$$\delta\Omega_2 = tr\ [\delta z_2^{\hat{}}\ (\frac{\partial\Omega_2}{\partial z_2^{\hat{}}})^{\hat{}}]_* + tr\ [\delta z_3\ (\frac{\partial\Omega_2}{\partial z_3})^{\hat{}}]_*$$

$$+ tr\ [\delta\frac{\partial z_2^{\hat{}}}{\partial\tau}(\frac{\partial\Omega_2}{\partial(\frac{\partial z_2^{\hat{}}}{\partial\tau})})^{\hat{}}]_* + tr\ [\ \delta\frac{\partial z_3}{\partial t}\ (\frac{\partial\Omega_2}{\partial(\frac{\partial z_3}{\partial t})})^{\hat{}}\ ]_*$$

$$+ tr\ [\delta u_3^{\hat{}}(\frac{\partial\Omega_2}{\partial u_3^{\hat{}}})^{\hat{}}]_* + tr\ [\delta\lambda_2\ (\frac{\partial\Omega_2}{\partial\lambda_2})^{\hat{}}]_* + tr\ [\delta\lambda_3\ (\frac{\partial\Omega_2}{\partial\lambda_3})^{\hat{}}\ ]_* \qquad C.1\text{-}11$$

$$\delta\Omega_1 = <\delta z_1\ ,\ \frac{(\partial\Omega_1)}{\partial z_1}_*> + < \delta\frac{dz_1}{dt}\ ,\ (\frac{\partial\Omega_1}{\partial(\frac{dz_1}{dt})})_*>$$

$$+<\delta u_1\ ,\ \frac{(\partial\Omega_1)}{\partial u_1}_*>+< \delta u_2\ ,\frac{(\partial\Omega_1)}{\partial u_2}_*> + <\delta\lambda_1\ ,\frac{(\partial\Omega_1)}{\partial\lambda_1}_*> \qquad C.1\text{-}12$$

It is seen from C.1-2 and C.1-3 that

$$\frac{\partial\Omega}{\partial(\frac{dz_1}{dt})}\bigg|_* = -\ \lambda_1^*$$

$$C.1\text{-}13$$

$$\frac{\partial \Omega_2}{\partial(\dot{\partial z_2})}\bigg|_{\frac{\partial z_2}{\partial \tau}}\bigg|_* \doteq -\lambda_2^*$$ C.1-14

$$\frac{\partial \Omega_2}{\partial(\dot{\partial z_3})}\bigg|_{\frac{\partial z_3}{\partial t}}\bigg|_* = -\lambda_3^*$$ C.1-15

These last three equations are substituted in C.1-11 and C.1-12 and then those two equations put in C.1-7 gives

$$\delta J = \int_{t_o}^{t_f}\int_{t_o}^{t_f} \{ \text{tr}[\delta z_2^{\prime}(\frac{\partial \Omega_2}{\partial z_2^{\prime}})^{\prime}]_* + \text{tr}\ [\delta z_3(\frac{\partial \Omega_2}{\partial z_3})^{\prime}]_*$$

$$+ \text{tr}\ [\frac{\delta \partial z_3^{\prime}}{\partial \tau}\ (-\lambda_2^{\prime})_*] + \text{tr}\ [\frac{\delta \partial z_3}{\partial t}\ (-\lambda_3^{\prime})_*]$$

$$+ \text{tr}\ [\delta u_3^{\prime}\ (\frac{\partial \Omega_2}{\partial u_3^{\prime}})^{\prime}]_* + \text{tr}\ [\delta \lambda_2\ (\frac{\partial \Omega_2}{\partial \lambda_2})^{\prime}]_*$$

$$+ \text{tr}\ [\delta \lambda_3\ (\frac{\partial \Omega_2}{\partial \lambda_3})^{\prime}]_* \} \ dt d\tau$$

$$+ \int_{t_o}^{t_f} \{<\delta z_1, (\frac{\partial \Omega_1}{\partial z_1})_*>-<\delta dz_1, \lambda_1^*>+< \delta u_1, (\frac{\partial \Omega_1}{\partial u_1})_*$$

$$+ <\delta u_2, (\frac{\partial \Omega_1}{\partial u_2})_*> + <\delta \lambda_1, (\frac{\partial \Omega_1}{\partial \lambda_1})_*> \} \ dt$$ C.1-16

Using integration by parts it can be shown that

$$- \int_{t_o}^{t_f}<\frac{\delta dz_1}{dt}, \lambda_1^*> \ dt = \int_{t_o}^{t_f} <\delta z_1, \frac{d\lambda_1^*}{dt}>dt - <\delta z_1(t_f), \lambda_1^*(t_f)$$

$$+ <\delta z_i(t_o), \ \lambda_i^*(t_o)> \qquad \qquad \text{C.1-17}$$

and

$$\int_{t_o}^{t_f}\int_{t_o}^{t_f} tr[\delta \frac{\partial z_2'}{\partial \tau} [-\lambda_2'^*]] \ dt d\tau = \int_{t_o}^{t_f}\int_{t_o}^{t_f} tr[\delta z_2' \frac{\partial \lambda_2'^*}{\partial \tau}] \ dt d\tau$$

$$- \int_{t_o}^{t_f} tr[\delta z_2'(t,t_f)\lambda_2'^*(t,t_f)] \ dt + \int_{t_o}^{t_f} tr[\delta z_2'(t,t_o) \ \lambda_2'^*(t,t_o)] \ dt \qquad \text{C.1-18}$$

and

$$\int_{t_o}^{t_f}\int_{t_o}^{t_f} tr[\delta \frac{\partial z_3}{\partial t} [-\lambda_3'^*]] \ dt d\tau = \int_{t_o}^{t_f}\int_{t_o}^{t_f} tr[\delta z_3 \frac{\partial \lambda_3'^*}{\partial t}] \ dt d\tau$$

$$- \int_{t_o}^{t_f} tr[\delta z_3(t_f,\tau) \ \lambda_3'^*(t_f,\tau)] d\tau + \int_{t_o}^{t_f} tr[\delta z_3(t_o,\tau)\lambda_3'^*(t_o,\tau)] \ d\tau \qquad \text{C.1-19}$$

These last three equations are substituted into C.1-16. Terms
multiplied by the same arbitrary variation are collected and the
fundamental theorem of the calculus of variations is applied, i. e.
if $\delta J = o$, then the coefficients multiplying each arbitrary variation
must be zero. This results in

$$\frac{(\partial \Omega_2)'}{\partial z_2'} * + \frac{\partial \lambda_2'^*}{\partial \tau} = o$$

$$\frac{(\partial \Omega_2)'}{\partial z_3} * + \frac{\partial \lambda_3'^*}{\partial t} = o$$

$$\frac{(\partial \Omega_i)}{\partial z_i} * + \frac{d\lambda_i}{dt}^* = o$$

$$\frac{(\partial \Omega_2)'}{\partial u_3'} * = o$$

$$\left(\frac{\partial \Omega}{\partial u_1^1}\right)_* = 0$$

$$\left(\frac{\partial \Omega}{\partial u_2^1}\right)_* = 0$$

$$\left(\frac{\partial \Omega}{\partial \acute{\lambda}_2^2}\right)_* = 0$$

$$\left(\frac{\partial \Omega}{\partial \acute{\lambda}_3^2}\right)_* = 0$$

$$\left(\frac{\partial \Omega}{\partial \acute{\lambda}_1^1}\right)_* = 0 \qquad\qquad\qquad\qquad\qquad C.1\text{-}20$$

and

$$\delta z_1(t_0) = \delta z_1(t_f) = 0$$

$$\delta z_2'(t,t_0) = \delta z_2'(t,t_f) = 0$$

$$\delta z_3(t_0,\tau) = \delta z_3(t_f,\tau) = 0 \qquad\qquad\qquad\qquad C.1\text{-}21$$

Equation C.1-21 implies the specified endpoints of equations 4.1.1.3 and 4.1.1.4.

Equations C.1-20 became, for the functions Ω_2 and Ω_1 of C.1-2 and C.1-3,

$$\frac{d\lambda^*(t)}{dt^1} = - A'(t)\lambda_1^*(t), \ \forall t \in \Gamma_1 \qquad\qquad\qquad C.1\text{-}22$$

$$\frac{\partial \lambda^*}{\partial \tau^2}(t,\tau) = - A'(\tau)\ \lambda_2^*(t,\tau) - \lambda_3'^*(t,\tau)\ C(\tau), \ \forall (t,\tau) \in \Gamma \qquad C.1\text{-}23$$

$$\frac{\partial \lambda^*}{\partial t^3}(t,\tau) = - A'(t)\ \lambda_3^*(t,\tau) - R'R_1 z_3^*(t,\tau), \ \forall (t,\tau) \in \Gamma \qquad C.1\text{-}24$$

$$\frac{dz^*(t)}{dt^1} = A(t)z_1^*(t) + B(t)u_1^*(t) + C(t)u_2^*(t), \ \forall t \in \Gamma_1 \qquad C.1\text{-}25$$

$$\frac{\partial z_2'^*}{\partial t^2}(t,\tau) = A_2(\tau)z_2'^*(t,\tau) + C_3(\tau)u_3'^*(t,\tau), \forall (t,\tau) \in \Gamma \qquad \text{C.1-26}$$

$$\frac{\partial z_3^*}{\partial t^3}(t,\tau) = A_3(t) \ z_3^*(t,\tau) + C(t)z_2^*(t,\tau), \forall (t,\tau) \in \Gamma \qquad \text{C.1-27}$$

$$u_1^*(t) = - R_3^{-1}B'(t)\lambda_1^*(t), \forall t \in \Gamma_1 \qquad \text{C.1-28}$$

$$u_2^*(t) = R_4^{-1}C'(t)\lambda_1^*(t), \forall t \in \Gamma_1 \qquad \text{C.1-29}$$

$$u_3'^*(t,\tau) = C'(\tau)\lambda_2^*(t,\tau) (R_2^{-1})(R_2^{-1}), \forall (t,\tau) \in \Gamma \qquad \text{C.1-30}$$

which are the necessary conditions stated in Theorem 4.1. It is seen that the inverses of R_2, R_3 and R_4 are required, but not for R_1, hence

R_1 is assumed to be non-negative definite

and

R_2, R_3, R_4 are assumed to be positive definite.

Remark: Substitution of equations C.1-28, C.1-29 and C.1-30 into C.1-25 and C.1-26 result in 2n(1 + s + n) differential equations to be solved. There are n(1 + s + n) endpoint conditions specified by 4.1.1.3 and n(1 + s + n) more by 4.1.1.4. So the complete solution of the problem requires the solution of a 2n(1 + s + n) dimension two point boundary value problem (TPBVP) since half of the endpoint conditions are specified at t_o and the other half at t_f.

C.2 Proof of Theorem 4.2

For Problem 2 of Section 4.1.2 defined by equations 4.1.1.2

through 4.1.1.8 with the performance index of 4.1.2.1 define

$\lambda_1(t)$ an n x 1 vector

$\lambda_2(t,\tau)$ an n x s matrix

$\lambda_3(t,\tau)$ an n x n matrix

and form the augmented performance index

$$Ja = \int_{t_0}^{t_f}\int_{t_0}^{t_f}\{fo(z_1(t),z_1(\tau),z_2'(t,\tau),z_3(t,\tau),u_1(t),u_1(\tau),u_2(t),u_2(\tau),u_3'(t,\tau))$$

$$+ \text{tr}\ [[A(\tau)z_2'(t,\tau)+C(\tau)u_3'(t,\tau)-\frac{\partial z_2'}{\partial \tau}(t,\tau)][\lambda_2'(t,\tau)]]$$

$$+ \text{tr}\ [[A(t)z_3(t,\tau)C(t)z_2(t,\tau)-\frac{\partial z_3}{\partial t}(t,\tau)][\lambda_3'(t,\tau)]]$$

$$+<A(t)z_1(t)+B(t)u_1(t)+C(t)u_2(t)-\frac{dz_1}{dt}(t),\lambda_1(t)>\}dtd\tau \qquad \text{C.2-1}$$

or

$$Ja = \int_{t_0}^{t_f}\int_{t_0}^{t_f}\Omega\ dtd\tau$$

where

$$\Omega = \Omega(z_1(t),z_1(\tau),z_2'(t,\tau),z_3(t,\tau),\frac{dz_1}{dt}(t),\frac{\partial z_2'}{\partial \tau}(t,\tau),\frac{\partial z_3}{\partial t}(t,\tau),$$

$$u_1(t),u_1(\tau),u_2(t),u_2(\tau),u_3'(t,\tau),\lambda_1(t),\lambda_2(t,\tau),\lambda_3(t,\tau)) \qquad \text{C.2-2}$$

and Ω is defined in C.2-1, and it is noted that Ja=J of equation

4.1.2.1.

In addition to the variations in C.1-5, define

$$\delta z_1(\tau) = \delta z_1(t)\ \Big|_{t=\tau}$$

$$\delta u_1(\tau) = \delta u_1(t)\ \Big|_{t=\tau}$$

$$\delta u_2(\tau) = \delta u_2(t)\ \Big|_{t=\tau} \qquad \text{C.2-3}$$

Then,

$$\delta Ja = \delta J = \int_{t_o}^{t_f}\int_{t_o}^{t_f} \delta\,\Omega\,\,dtd\tau \qquad\qquad\text{C.2-4}$$

where

$$\delta\Omega = tr\delta z_2^{'}\left(\frac{\partial\Omega}{\partial z_2^{'}}\right)^{'} + tr\delta z_3\left(\frac{\partial\Omega}{\partial z_3}\right)^{'} + tr\delta\,\frac{\partial z_2^{'}}{\partial\tau}\left(\frac{\partial\Omega}{\partial\left(\frac{\partial z_2^{'}}{\partial\tau}\right)}\right)^{'}$$

$$+ tr\,\delta\,\frac{\partial z_3}{\partial t}\left(\frac{\partial\Omega}{\partial\left(\frac{\partial z_3}{\partial t}\right)}\right)^{'} + tr\delta u_3^{'}\left(\frac{\partial\Omega}{\partial u_3}\right)^{'} + tr\,\delta\lambda_2\left(\frac{\partial\Omega}{\partial\lambda_2}\right)^{'}$$

$$+ tr\,\delta\lambda_3\left(\frac{\partial\Omega}{\partial\lambda_3}\right)^{'} + <\delta z_1(t),\frac{\partial\Omega}{\partial z_1(t)}> + <\delta z_1(\tau),\frac{\partial\Omega}{\partial z_1(\tau)}>$$

$$+<\delta\frac{dz_1(t)}{dt},\frac{\partial\Omega}{\partial\left(\frac{dz_1(t)}{dt}\right)}> +<\delta u_1(t),\frac{\partial\Omega}{\partial u_1(t)}> + <\delta u_1(\tau),\frac{\partial\Omega}{\partial u_1(\tau)}>$$

$$+<\delta u_2(t),\frac{\partial\Omega}{\partial u_2(t)}> + <\delta u_2(\tau),\frac{\partial\Omega}{\partial u_2(\tau)}> + <\delta\lambda_1(t),\frac{\partial\Omega}{\partial\lambda_1(t)}>\}dtd\tau \qquad\text{C.2-5}$$

Since fo does not depend on $\frac{dz_1(t)}{dt}$, $\frac{\partial z_2^{'}(t,\tau)}{\partial\tau}$ or $\frac{\partial z_3(t,\tau)}{\partial t}$

$$\left(\frac{\partial\Omega}{\partial\left(\frac{\partial z_2^{'}}{\partial\tau}\right)}\right)^{'} = -\lambda_2^{'}(t,\tau) \qquad\qquad\text{C.2-6}$$

$$\left(\frac{\partial\Omega}{\partial\left(\frac{\partial z_3}{\partial t}\right)}\right)^{'} = -\lambda_3^{'}(t,\tau) \qquad\qquad\text{C.2-7}$$

$$\frac{\partial\Omega}{\partial\left(\frac{dz_1}{dt}\right)} = -\lambda_1(t) \qquad\qquad\text{C.2-8}$$

Substituting C.2-6, C.2-7, and C.2-8 into C.2-5, and then that equation

into C.2-4 results in 3 terms to be integrated by parts similar to

C.1-17, C.1-18, and C.1-19. After this step, terms are collected,

δJ is set to zero, and the fundamental theorem of the calculus of

variations applied, which results in the following necessary conditions

$$\int_{t_o}^{t_f}\{\partial \frac{fo}{\partial z_1(t)} + A^\prime(t)\lambda_1(t) + \frac{d\lambda_1}{dt}(t)\}d\tau=0, \quad \int_{t_o}^{t_f}\{\frac{\partial fo}{\partial z_1(\tau)}\}dt=0 \qquad \text{C.2-9}$$

$$\frac{\partial \lambda_2}{\partial \tau}(t,\tau)= -A^\prime(\tau)\lambda_2(t,\tau)-\lambda_3^\prime(t,\tau)C(\tau)-\frac{\partial fo}{\partial z_2^\prime} \qquad \text{C.2-10}$$

$$\frac{\partial \lambda_3}{\partial t}(t,\tau)= -A^\prime(t)\lambda_3(t,\tau)-\frac{\partial fo}{\partial z_3} \qquad \text{C.2-11}$$

$$\frac{dz_1}{dt}(t)= A(t)z_1(t)+B(t)u_1(t)+C(t)u_2(t) \qquad \text{C.2-12}$$

$$\frac{\partial z_2^\prime}{\partial \tau}(t,\tau)= A(\tau)z_2^\prime(t,\tau)+C(\tau)u_3^\prime(t,\tau) \qquad \text{C.2-13}$$

$$\frac{\partial z_3}{\partial t}(t,\tau)= A(t)z_3(t,\tau)+C(t)z_2(t,\tau) \qquad \text{C.2-14}$$

$$\int_{t_o}^{t_f}\{\frac{\partial fo}{\partial u_1(t)} + B^\prime(t)\lambda_1(t)\}d\tau=0, \quad \int_{t_o}^{t_f}\{\frac{\partial fo}{\partial u_1(\tau)}\}dt=0 \qquad \text{C.2-15}$$

$$\int_{t_o}^{t_f}\{\frac{\partial fo}{\partial u_2(t)} + C^\prime(t)\lambda_1(t)\}d\tau=0, \quad \int_{t_o}^{t_f}\{\frac{\partial fo}{\partial u_2(\tau)}\}dt=0 \qquad \text{C.2-16}$$

$$\frac{\partial fo}{\partial u_3^\prime} + C^\prime(\tau)\lambda_2^\prime(t,\tau)=0 \qquad \text{C.2-17}$$

with the endpoint conditions of C.1-21.

Remark: The equations C.2-9, C.2-15, and C.2-16 can be shown to not present constraint difficulties for the following examples. Only the case for C.2-15 is given.

Consider four cases.

Case 1: $fo(u_1(t),u_1(\tau))=<u_1(t),R_3u_1(t)>,R_3=R_3^\prime$

$$\frac{\partial fo}{\partial u_1(\tau)}=0$$

$$\frac{\partial fo}{\partial u_1(t)}=2R_3u_1(t)$$

and C.2-15 becomes

$$u_1(t)= -\frac{1}{2} R_3^{-1} B^\prime(t)\lambda_1(t)$$

Case 2: $fo(u_\perp(t), u_\perp(\tau)) = <u_\perp(t), R_3 u_\perp(t)> W(t)W(\tau), R_3 = R_3'$

where $W(.)$ is a scalar, time-varying weighting function

Then

$$\frac{\partial fo}{\partial u_\perp(\tau)} = 0$$

$$\frac{\partial fo}{\partial u_\perp(t)} = 2R_3 W(t)W(\tau) u_\perp(t)$$

and C.2-15 becomes

$$u_\perp(t) = \frac{-2(t_f - t_o)}{W(t) \int_{t_o}^{t_f} W(\tau) d\tau} R_3^{-1} B'(t) \lambda_\perp(t)$$

Case 3: $fo(u_\perp(t), u_\perp(\tau)) = <u_\perp(t), R_3 u_\perp(\tau)>, R_3 = R_3'$

Then

$$\frac{\partial fo}{\partial u_\perp(\tau)} = R_3' u_\perp(t)$$

$$\frac{\partial fo}{\partial u_\perp(t)} = R_3 u_\perp(\tau)$$

and C.2-15 becomes

$$\int_{t_o}^{t_f} u(\tau) d\tau = \frac{-(t_f - t_o)}{2} R_3^{-1} B'(t) \lambda_\perp(t)$$

Case 4: $fo(u_\perp(t), u_\perp(\tau)) = <u_\perp(t), R_3(t,\tau) u_\perp(\tau)>$

with $R_3(t,\tau) = R_3'(\tau, t)$

then C.2-15 becomes

$$\int_{t_o}^{t_f} R_3(t,\tau) u(\tau) d\tau = \frac{-(t_f - t_o)}{2} B'(t) \lambda_\perp(t)$$

C.3 Proof of Theorem 4.3

The proof is an extension of derivations in Bryson and Ho (20) and Gabasov and Kirillova (28). In (20) the actual derivation is carried out for a one sided problem with a scalar control in one independent variable, t. The result is

$$\frac{d}{dt}\frac{\partial H}{\partial u} = 0 \qquad\qquad C.3\text{-}1$$

$$\frac{\partial}{\partial u}\frac{d}{dt}\frac{\partial H}{\partial u} = 0 \qquad\qquad C.3\text{-}2$$

$$\frac{\partial}{\partial u}\left\{\frac{d^2}{dt^2}\frac{\partial H}{\partial u}\right\} \leq 0 \qquad\qquad C.3\text{-}3$$

where H is the Hamiltonian for the one-sided problem and minimization of cost with respect to the scalar u is carried out. The form of C.3-3 is similar to the form of the Legendre-Clebsch condition

$$H_{uu} \geq 0 \qquad\qquad C.3\text{-}4$$

For singular problems, the equality holds in C.3-4 therefore no information is obtained from this condition. For nonsingular problems C.3-4 implies the sign of the second variation of the augmented performance index thereby indicating whether minimization or maximization is achieved.

When equality holds in C.3-4, the test for maximization or minimization is provided by C.3-3.

The results for a vector control u are quoted in (28) and were derived in Kelley (43) and Kelley, Kopp and Moyer (42). The equations of Theorem 4.3 are a generalization of the above for a matrix control $u_3^{<}$ in two independent variables, t and τ where maximization with respect to $u_3^{<}$ is desired.

The proof is extremely lengthy but similar in form to the derivation in Bryson and Ho (20). The second variation is found, and then the optimal state. equations and costate equations are perturbed resulting in the comparison of neighboring trajectories, i.e. variations in the state and costate are examined. It is seen that equations 4.2.3.1 through 4.2.3.3 are equivalent to C.3-1 through C.3-3 if $u_3^{<}$ is restricted to be a scalar function of one independent variable, t realizing that maximization rather than minimization with respect to $u_3^{<}$ is sought.

C.4 Proof of Theorem 4.4

Define the Hamiltonian

$$H(z_1, z_2, z_3, u_1, u_2, u_3, \lambda_1, \lambda_2, \lambda_3) =$$

$$\frac{d_1}{4} u_1'(t) R_1(t) u_1(t) + \frac{d_2}{4} \operatorname{tr}\{R_2(t)[u_3(t) + u_2(t)u_2'(t)]^2\}$$

$$+ \lambda_1'(t)[A(t)z_1(t) + B(t)u_1(t) + C(t)u_2(t)]$$

$$+ \operatorname{tr}\{[F(t)z_2(t) + z_2(t)A'(t) + u_3(t)C'(t)]\lambda_2'(t)\}$$

$$+ \operatorname{tr}\{[A(t)z_3(t) + z_3(t)A'(t) + C(t)z_2(t) + z_2'(t)C')t)]\lambda_3'(t)\} \quad \text{C.4-1}$$

where the vector $\lambda_1(t)$ and the matrixes $\lambda_2(t), \lambda_3(t)$ are Lagrange

multipliers. With the aid of the matrix maximum principle of Athans

(10), the following result is obtained. In order that $u_i*(t) \in U_i, t \in \Gamma_i$,

$i = 1,2,3$ be the optimal strategies for Problem 4, it is necessary that

there exist a nonzero vector function $\lambda_1*(t)$ and nonzero matrix functions

$\lambda_2*(t)$ and $\lambda_3*(t)$ such that conditions a), b) and c) of Theorem 4.4

hold, as given in Section 4.2.4.

Proof:

Define the variations

$$z_i(t) = z_i*(t) + \delta z_i(t)$$

$$u_i(t) = u_i*(t) + \delta u_i(t) \quad \text{C.4-2}$$

and using the standard variational approach, it can be shown that

$$\delta Ja = \int_{t_0}^{t_f} \{ [(\frac{\partial H}{\partial z_1})_* + \frac{d\lambda_1}{dt}*]'\delta z_1(t) + \operatorname{tr}[\delta z_2'((\frac{\partial H}{\partial z_2})_* + \frac{d\lambda_2}{dt}*)]$$

$$+ \operatorname{tr}[\delta z_3'((\frac{\partial H}{\partial z_3})_* + \frac{d\lambda_3}{dt}*)] - \lambda_1'*(t_f)\delta z_1(t_f)$$

$$- \operatorname{tr}[\lambda_2*(t_f) \delta z_2(t_f)] - \operatorname{tr}[\lambda_3*(t_f) \delta z_3(t_f)] \quad \text{C.4-3}$$

where $(\cdot)_*$ denotes that the function $(\cdot)$ is evaluated at

$z_i = z_i^*, u_i = u_i^*$ and $\lambda_i = \lambda_i^*,\ i = 1,2,3$. If the state equations in 4.2.4.1 through 4.2.4.3 are satisfied, and $\lambda_i^*(t)$, $i = 1,2,3$ are selected so that the coefficients of $\delta z_i(t)$, $i = 1,2,3$ in the integral are identically zero, and the boundary conditions of 4.2.4.7 and 4.2.4.8 are satisfied, then

$$\delta J_a = \int_{t_o}^{t_f} \{ (\tfrac{\partial H}{\partial u_1})_*^{\prime} \delta u_1 + (\tfrac{\partial H}{\partial u_2})_* \delta u_2 + tr[\delta u_3^{\prime}(\tfrac{\partial H}{\partial u_3})_*]\} dt \qquad C.4\text{-}4$$

To the first order approximation

$$(\tfrac{\partial H}{\partial u})_*^{\prime} \delta u_1(t) = H(z_1^*,z_2^*,z_3^*,u_1^* + \delta u_1, u_2^*, u_3^*, \lambda_1^*, \lambda_2^*, \lambda_3^*) - H^* \qquad C.4\text{-}5$$

Therefore, with $u_1 = u_1^* + \delta u_1, u_2 = u_2^*, u_3 = u_3^*$

$$\delta J_a = \int_{t_o}^{t_f} (\tfrac{\partial H}{\partial u_1})^{\prime} \delta u_1 dt =$$

$$\int_{t_o}^{t_f} \{H_1 - H^*\} dt \qquad C.4\text{-}6$$

Define a sufficiently small neighborhood of u_i^* as

$$\delta U_i = \{\delta u_i :\ ||\delta u_i|| < \beta_i \text{ and } u_i^* + \delta u_i \in U_i\} \qquad C.4\text{-}7$$

for $i = 1,2,3$ where the β_i are positive constants. For u_1^* to be a minimizing strategy it is necessary that

$$\delta J_a(u_1^*, \delta u_1) = \int_{t_o}^{t_f} \{H_1 - H^*\} dt \geq 0 \qquad C.4\text{-}8$$

for all $\delta u_1 \in \delta U_1$. It can be shown as in Kirk (45) that in order for C.4-8 to be satisfied for all $\delta u_1 \in \delta U_1$, it is necessary that

$$H_1 \geq H^* \qquad C.4\text{-}9$$

In a similar manner, with $u_1 = u_1^*$, $u_2 = u_2^* + \delta u_2$, $u_3 = u_3^*$ + δu_3 where u_2 and u_3 are maximizing strategies, it can be shown that

$$H_2 \leqq H^*$$ C.4-10

and combining C.4-9 and C.4-10 yields conditions b) and c) of

Theorem 4.4.

APPENDIX D

SUFFICIENT CONDITIONS

In one-sided deterministic optimal control problems sufficient
conditions determine whether the extremal solution obtained by the
necessary conditions is a minimizing or maximizing solution. The
technique often used is to see if the second variation of the augmented
performance index is positive definite implying a minimizing extremal
solution or negative definite implying a maximizing solution. In
this approach, the restrictions on the variations lead to local
sufficient conditions. This means the extremal solution is compared
only to neighboring values of this solution. Global solutions can be
obtained by comparing the value of the solution to all possible values,
not only those due to neighboring solutions. The theory of convex functions
is most often used to establish global sufficient conditions.

In two-sided deterministic optimal control problems, which are
two player zero-sum differential games, the second variation technique
is not as useful as the notions of convexity and concavity. The latter
techniques were given for nonlinear programming problems in Saaty and
Bram (94) and Wilde and Beightler (107). The approach of convex and
concave functions was used to establish global sufficient conditions
for a wide class of differential games in Kuo and Burbank (55) and
Kuo (52).

The definition of a saddle point for the functional performance
index of equation 5.1.3 is

$$J(z_1, z_2', z_3, u_1^*, u_2, u_3', \lambda_1^*, \lambda_2^*, \lambda_3^*) \triangleq J_2 \leq$$

$$J(z_1^*, z_2'^*, z_3^*, u_1^*, u_2^*, u_3'^*, \lambda_1^*, \lambda_2^*, \lambda_3^*) \triangleq J^* \leq$$

$$J(z_1, z_2', z_3, u_1, u_2^*, u_3'^*, \lambda_1^*, \lambda_2^*, \lambda_3^*) \triangleq J_1 \qquad \text{D.1-1}$$

where now ( )* denotes optimal, not extremal. This equation can be written as the two separate inequalities

$$J_z - J^* \leq 0 \qquad\qquad \text{D.1-2}$$

and

$$J_i - J^* \geq 0 \qquad\qquad \text{D.1-3}$$

Recalling that it is sought to establish

$$\min_{\substack{u_i \in U_i}} \ \max_{\substack{u_z \in U_z \\ u_j' \in U_j}} \ J(u_i; u_z, u_j') \qquad\qquad \text{D.1-4}$$

the inequalities can be described. J_z is J evaluated at $u_i = u_i^*$, the optimal condition, and at u_z, u_j' some non optimal conditions. The resulting trajectories for z_i, z_z', and z_j will not be optimal since u_z, u_j' are not. Since u_z and u_j' seek to maximize J, when they are not optimal the value of J should be less than the value J* which is the maximum attainable at all optimal conditions.

Similarly, J_i is at optimal u_z^*, $u_j'^*$ and non-optimal u_i. Since u_i minimizes, the value of J at non-optimal u_i should be greater than the minimum attainable at all optimal conditions, J*. If the inequalities D.1-2 can be established for all non-optimal u_z, u_j' and D.1-3 for all non-optimal u_i, then global sufficient conditions have been found.

D.1 Proof of Theorems 5.1.1 and 5.1.2

Proof of Theorem 5.1.1:

First half:

It is sought to establish D.1-2 for the augmented performance index of equation 5.1.2. From D.1-1,

$$J_2 - J^* = \int_{t_o}^{t_f}\int_{t_o}^{t_f}\{\Omega_2(z_2',z_3,\tfrac{\partial z_2'}{\partial \tau},\tfrac{\partial z_3}{\partial t},u_3',\lambda_2^*,\lambda_3^*) -$$

$$\Omega_2(z_2'^*,z_3^*,\tfrac{\partial z_2'^*}{\partial \tau},\tfrac{\partial z_3^*}{\partial t},u_3'^*,\lambda_2^*,\lambda_3^*)\}dt d\tau$$

$$+\int_{t_o}^{t_f}\{\Omega_1(z_1,\tfrac{dz_1}{dt},u_1^*,u_2,\lambda_1^*) -\Omega_1(z_1^*,\tfrac{dz_1^*}{dt},u_1^*,u_2^*,\lambda_1^*)\}dt \qquad \text{D.1-5}$$

where Ω_2 and Ω_1 are defined in equations C.1-2 and C.1-3. Using the properties of concavity it can be shown that

$$\Omega_2(z_2',z_3,\tfrac{\partial z_2'}{\partial \tau},\tfrac{\partial z_3}{\partial t},u_3',\lambda_2^*,\lambda_3^*)-\Omega_2(z_2'^*,z_3^*,\tfrac{\partial z_2'^*}{\partial \tau},\tfrac{\partial z_3^*}{\partial t},u_3'^*,\lambda_2^*,\lambda_3^*)\leq$$

$$tr[\delta z_2'(\tfrac{\partial \Omega_2}{\partial z_2'})^*]+tr[\delta(\tfrac{\partial z_2'}{\partial \tau})(\tfrac{\partial \Omega_2}{\partial(\partial z_2'/\partial \tau)})^*]$$

$$+tr[\delta\tfrac{\partial z_3}{\partial t}(\tfrac{\partial \Omega_2}{\partial(\partial z_3/\partial t)})^*]+tr[\delta z_3(\tfrac{\partial \Omega_2}{\partial z_3})^*]+tr[\delta u_3'(\tfrac{\partial \Omega_2}{\partial u_3})^*] \qquad \text{D.1-6}$$

But, from C.1-20 and C.1-30

$$(\tfrac{\partial \Omega_2}{\partial z_2'})_*' = \lambda_2'^*(t,\tau)A(\tau)+C'(t)\lambda_3^*(t,\tau)$$

$$(\tfrac{\partial \Omega_2}{\partial(\partial z_2'/\partial \tau)})_*' = -\lambda_2'^*(t,\tau)$$

$$(\tfrac{\partial \Omega_2}{\partial(\partial z_3/\partial t)})_*' = -\lambda_3'^*(t,\tau)$$

$$(\tfrac{\partial \Omega_2}{\partial z_3})_*' = -z_3'^*(t,\tau)R_1'R_1+\lambda_3'^*(t,\tau)A(t)$$

and

$$(\tfrac{\partial \Omega_2}{\partial u_3})_*' = [C'(\tau)\lambda_2'^*(t,\tau)-u_3'^*(t,\tau)R_2'R_2]'=0$$

After substituting, D.1-6 becomes

$$\Omega_2(z_2',z_3,\tfrac{\partial z_2'}{\partial \tau},\tfrac{\partial z_3}{\partial t},u_3',\lambda_2^*,\lambda_3^*)-\Omega_2(z_2'^*,z_3^*,\tfrac{\partial z'^*}{\partial \tau},\tfrac{\partial z_3^*}{\partial t},u_3'^*,\lambda_2^*,\lambda_3^*)\leq$$

$$+tr[\delta z_2'[\lambda_2'^*(t,\tau)A(\tau)+C'(\tau)\lambda_3^*(t,\tau)]]-tr[\delta(\tfrac{\partial z_2'}{\partial \tau})\lambda_2'^*(t,\tau)]$$

$$-\text{tr}[\delta(\tfrac{\partial z_3}{\partial t})\lambda_3'^{*}(t,\tau)]+\text{tr}[\delta z_3[-z_3'^{*}(t,\tau)R_1'R_1+\lambda_3'^{*}(t,\tau)A(t)]] \qquad \text{D.1-7}$$

Using the endpoint conditions

$$\delta z_2'(t,t_o)=\delta z_2'(t,t_f)=\delta z_3(t_o,\tau)=\delta z_3(t_f,\tau)=0$$

and integration by parts it can be shown that

$$-\int_{t_o}^{t_f}\int_{t_o}^{t}\text{tr}[\delta(\tfrac{\partial z_2'}{\partial \tau})\lambda_2'^{*}(t,\tau)]dtd\tau=\int_{t_o}^{t_f}\int_{t_o}^{t}\text{tr}[\delta z_2'\tfrac{\partial \lambda_2'^{*}}{\partial \tau}]dtd\tau \qquad \text{D.1-8}$$

and

$$-\int_{t_o}^{t_f}\int_{t_o}^{t}\text{tr}[\delta(\tfrac{\partial z_3}{\partial t})\lambda_3'^{*}(t,\tau)]dtd\tau=\int_{t_o}^{t_f}\int_{t_o}^{t_f}\text{tr}[\delta z_3\tfrac{\partial \lambda_3'^{*}}{\partial t}]dtd\tau \qquad \text{D.1-9}$$

Substituting D.1-7 into D.1-5, and then substituting D.1-8 and

D.1-9 gives

$$\int_{t_o}^{t_f}\int_{t_o}^{t}\{\Omega_z(z_2',z_3,\tfrac{\partial z_2'}{\partial \tau},\tfrac{\partial z_3}{\partial t},u_3',\lambda_2{}^{*},\lambda_3{}^{*})$$

$$-\Omega_z(z_2'^{*},z_3{}^{*},\tfrac{\partial z_2'^{*}}{\partial \tau},\tfrac{\partial z_3}{\partial t}{}^{*},u_3'^{*},\lambda_2{}^{*},\lambda_3{}^{*})\}dtd\tau \leq 0 \qquad \text{D.1-10}$$

in view of C.1-23 and C.1-24.

Using the properties of concavity, and integration by parts

$$\int_{t_o}^{t_f}\{\Omega_1(z_1,\tfrac{dz_1}{dt},u_1{}^{*},u_2,\lambda_1{}^{*})-\Omega_1(z_1{}^{*},\tfrac{dz_1}{dt}{}^{*},u_1{}^{*},u_2{}^{*},\lambda_1{}^{*})\}dt \leq$$

$$\int_{t_o}^{t_f}\{<\delta z_1,(\tfrac{\partial \Omega_1}{\partial z_1})_{*}>+<\delta\tfrac{dz_1}{dt},(\tfrac{\partial \Omega_1}{\partial(\tfrac{dz_1}{dt})})_{*}>+<\delta u_2,(\tfrac{\partial \Omega_2}{\partial u_2})_{*}>\}dt=$$

$$\int_{t_o}^{t_f}<\delta z_1,A'(t)\lambda_1{}^{*}(t)+\tfrac{d\lambda_1}{dt}{}^{*}>+<\delta u_2,-R_4 u_2{}^{*}(t)+C'(t)\lambda_1{}^{*}(t)>\}dt=0 \qquad \text{D.1-11}$$

in view of C.1-22, C.1-25 and C.1-29. From D.1-10 and D.1-11,

D.1-2 is established.

Second half:

It is sought to establish D.1-3 for the augmented performance

index of 5.1.2. From D.1-1,

$$J_1 - J^* = \int_{t_o}^{t_f} \int_{t_o}^{t_f} \{\Omega_2(z_2', z_3, \frac{\partial z_2'}{\partial \tau}, \frac{\partial z_3}{\partial t}, u_3'^*, \lambda_2^*, \lambda_3^*) -$$

$$\Omega_2(z_2'^*, z_3^*, \frac{\partial z_2'^*}{\partial \tau}, \frac{\partial z_3^*}{\partial t}, u_3'^*, \lambda_2^*, \lambda_3^*)\} dt d\tau$$

$$+ \int_{t_o}^{t_f} \{\Omega_1(z_1, \frac{dz_1}{dt}, u_1, u_2^*, \lambda_1^*) - \Omega_1(z_1^*, \frac{dz_1^*}{dt}, u_1^*, u_2^*, \lambda_1^*)\} dt \qquad \text{D.1-12}$$

and from C.1-26 and C.1-27, if u_3' is $u_3'^*$, then

$z_2', z_3, \frac{\partial z_2'}{\partial \tau}$, and $\frac{\partial z_3}{\partial t}$ are evaluated at optimal conditions.

Therefore D.1-12 goes to

$$J_1 - J^* = \int_{t_o}^{t_f} \{\Omega_1(z_1, \frac{dz_1}{dt}, u_1, u_2^*, \lambda_1^*) - \Omega_1(z_1^*, \frac{dz_1^*}{dt}, u_1^*, u_2^*, \lambda_1^*)\} dt \qquad \text{D.1-13}$$

Using the properties of convexity it can be shown that

$$\Omega_1(z_1, \frac{dz_1}{dt}, u_1, u_2^*, \lambda_1^*) - \Omega_1(z_1^*, \frac{dz_1^*}{dt}, u_1^*, u_2^*, \lambda_1^*) \geq$$

$$<\delta u_1, (\frac{\partial \Omega_1}{\partial u_1})_*> + <\delta z_1, (\frac{\partial \Omega_1}{\partial z_1})_*> + <\delta (\frac{dz_1}{dt}), (\frac{\partial \Omega_1}{\partial (\frac{dz_1}{dt})})_*> \qquad \text{D.1-14}$$

From C.1-20 and C.1-28

$$(\frac{\partial \Omega_1}{\partial u})_* = 0 \qquad \text{D.1-15}$$

then substituting D.1-15 into D.1-14 and the result into

D.1-13 gives

$$J_1 - J^* \geq \int_{t_o}^{t_f} \{<\delta z_1, (\frac{\partial \Omega_1}{\partial z_1})_*> + <\delta (\frac{dz_1}{dt}), (\frac{\partial \Omega_1}{\partial (\frac{dz_1}{dt})})_*>\} dt \qquad \text{D.1-16}$$

Using the endpoint condition

$$\delta z_1(t_o) = \delta z_1(t_f) = 0$$

and integration by parts, it can be shown that

$$J_i - J^* \geq \int_{t_o}^{t_f} <\delta z_i, (\frac{\partial \Omega_i}{\partial z_i})_* + \frac{d\lambda_i}{dt}^*>dt=0 \qquad \text{D.1-17}$$

in view of C.1-22 and C.1-25, thereby establishing D.1-3.
This completes the proof of Theorem 5.1.1.

Proof of Theorem 5.1.2:

The statement of Theorem 5.1.1 as given by equations 5.1.1.1
and 5.1.1.2 follows directly from the proof of Theorem 5.1.1 in this
Appendix.

In Saaty and Bram (94) and Parthasarathy and Raghavan (85)
it is shown that the necessary and sufficient conditions such that
5.1.1.1 holds are that

R_3, R_4 are non-negative definite and symmetric.   D.1-18

Since R_3 and R_4 are assumed to be positive definite and symmetric
for Theorem 4.1, they are non-negative definite, hence 5.1.1.1 holds
true.

After a lengthy rearrangement of the s x s matrix u_j^* into an
s^2 x 1 vector of the columns of u_j^* and similar rearrangements of
z_3, R_1 and R_2, it is seen that the

tr $R_2 u_3 u_3^* R_2^*$ and tr $R_1 z_3 z_3^* R_1^*$ functions are

equivalent to inner products of the columns of u_j^* and z_3. Using
theorems from Hohn (36) and Graybill (32), the weighting matrices
of these inner products can be proven to be positive definite or
positive semidefinite if R_1 and R_2 are positive definite or positive
semidefinite respectively. Then the same approach is used to
establish 5.1.1.2 as was done for 5.1.1.1, the result being

R_1, R_2 are non-negative definite and symmetric.   D.1-19

Since R_1 is assumed positive semidefinite and R_2 positive definite,
and both symmetric, from Theorem 4.1, equation 5.1.1.2 holds true,
and the proof of Theorem 5.1.2 is complete.

Remark:

The positive definiteness of R_2, R_3 and R_4 establish D.1-2
and D.1-3 as strict inequalities, hence only one optimal solution
can be obtained, if there is only one solution to the $2n(1 + s + n)$
dimensional T P B V P.

Remark:

The sufficient conditions are global sufficient conditions
since the convexity and concavity requirements are global. Therefore,
if only one solution of the T P B V P exists, it is the unique global
optimal solution.

D.2 Proof of Theorem 5.2

The details of this proof are similar to the proof of Theorem 5.1.2 except that terms in the variations $\delta z_1(\tau)$, $\delta u_1(\tau)$ and $\delta u_2(\tau)$ are included. As in the proof of Theorem 5.1.2 which includes the necessary conditions of Theorem 4.1, this proof requires all the necessary conditions of Theorem 4.2. The convexity and concavity requirements are that

f_o is convex in $u_1(t), u_1(\tau), z_1(t)$ and $z_1(\tau)$

and

f_o is concave in $u_2(t), u_2(\tau), z_1(t), z_1(\tau), u_3'(t,\tau), z_2'(t,\tau)$

and $z_3(t,\tau)$.

APPENDIX E

COMPUTER PROGRAMS

There were two main problems solved. The "worst case" optimal control problem, i.e. the differential game of Chapters 4 and 5, and the parameter identification problem of Chapter 6. The four programs OPTIML, ANLPLT, FITTPT and VALUE were written to complete the solution of the differential game. The program PROB1P was written to simulate the overall model reference adaptive system described in Chapter 6. This program included both identification and suboptimal control. A description and a listing of each program is contained in this appendix. Flow-charts for the programs OPTIML and PROB1P are included. No flowcharts are given for the other three programs since they only contain straight forward calculations.

The assistance of Dr. F. Russel in creating a set of files which greatly eased the whole programming effort is acknowledged and was deeply appreciated. This set of files provided means to edit and execute any of the above five programs from a remote terminal in the interactive mode and thereby saved much time and physical labor.

The complete analytic solution to the optimal problem was obtained by hand using two methods. First, the solution form of the equations presented in Appendix A was used. Also, the two dimensional Laplace transform technique of Kuo (54) was extended and applied. Both methods yielded identical results. The complete solution was differentiated by hand insuring that the results actually did satisfy the differential equations and

endpoint conditions of Theorem 4.1 . The convexity and concavity requirements of Theorem 5.1 were satisfied, hence the solution obtained was the unique, optimal, global solution.

The program OPTIML simulated the differential equations of all the variables which were functions of t only. Numerical integration of the ordinary differential equations was carried out with double precision arithmetic. The results agreed exactly with those obtained by hand. A program to numerically evaluate the functions of two independent variables was not written directly, but could be in terms of either of the methods used to obtain the solutions by hand. Shooting techniques could be employed to solve the TPBVP of these variables rather than the actual algebraic hand calculations which were made.

Both open loop and closed versions of OPTIML were run, the only

difference being the form of the differential equations. In

Figure E.2 is the program listing and Figure E.3.1 through

E.3.4 contain the actual plots obtained by computer.

```
┌─────────────────────────────────────────┐
│                                           │
│         Read in data                      │
│                                           │
│   Matrix elements,times,                  │
│                                           │
│   Initial conditions,tolerances           │
│                                           │
└─────────────────────────────────────────┘
                    │
                    ▼
┌─────────────────────────────────────────┐
│                                           │
│       Obtain initial conditions           │
│                                           │
│   For λ or K                              │
│                                           │
└─────────────────────────────────────────┘
                    │
                    ▼
┌─────────────────────────────────────────┐
│                                           │
│         Integrate                         │
│                                           │
│   O.D.E.'s over Γ₁  storing               │
│                                           │
│   Data as integration proceeds.           │
│                                           │
│   Using DDESP,FUNC1,OUT1.                 │
│                                           │
└─────────────────────────────────────────┘
                    │
                    ▼
┌─────────────────────────────────────────┐
│                                           │
│       Retrieve data                       │
│                                           │
│   Plot out trajectories.                  │
│                                           │
│   Using PLTVAR                            │
│                                           │
└─────────────────────────────────────────┘
                    │
                    ▼
              ┌───────────┐
              │           │
              │    END    │
              │           │
              └───────────┘
```

Integrate O.D.E.'s over Γ_1 storing
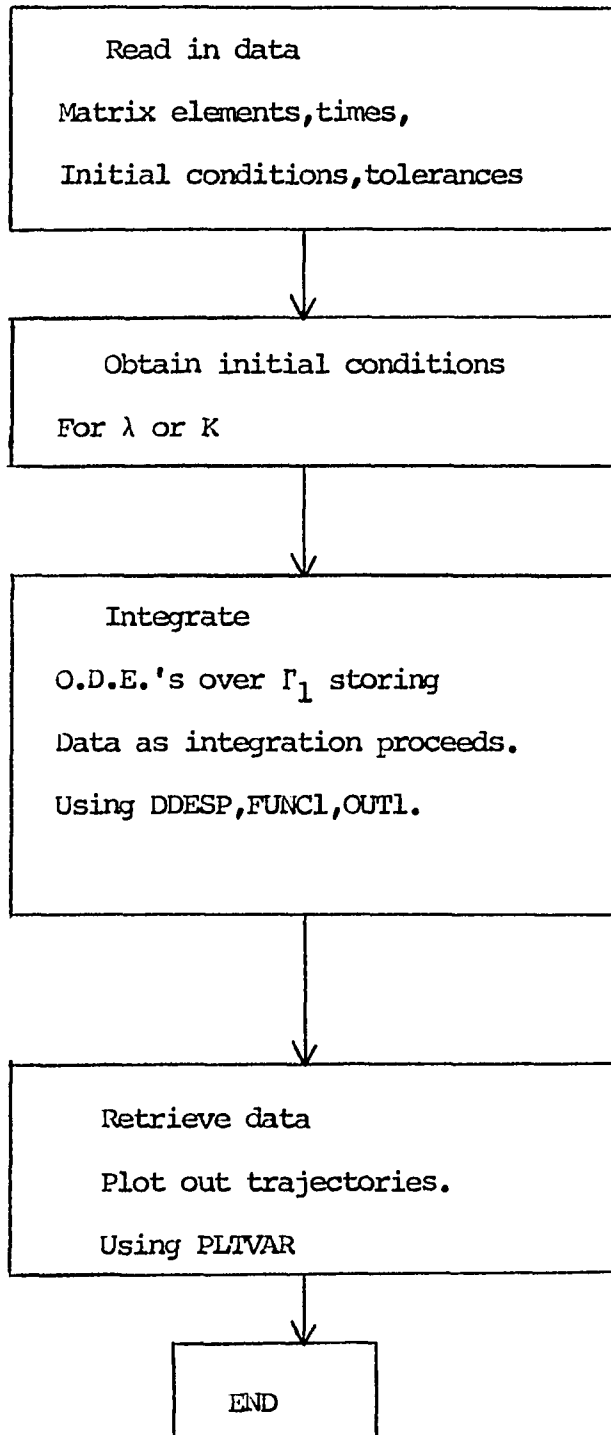
Figure E.1 Flowchart for optimal problem

```
      PROGRAM OPTIML
C     H. BURBANK  E.E.DEPT.
C                      PROBLEM 1 WITH PLOTS
      DOUBLE PRECISION YSTART,XSTART,XEND,H,EP,SP,Y,DY,X,TSTART,TEND
      DOUBLE PRECISION A,B,C,R3,R4
      EXTERNAL FUNC1,DSE,OUT1
      DIMENSION Y(2),DY(2),YSTART(2)
      DIMENSION  Z(2),F(51),TIME(51),U1(51),U2(51)
      REAL XX,TIME,F,U1,U2
      COMMON NOFNS,A,B,C,R3,R4
      NOFNS=0
      REWIND 40
C
      N=2
C
      TSTART=0.
C
      TEND=1.
C
      H=0.02
C
      EP=0.0001
C
      A=-3.
      B=2.
      C=1.
       R3=1.
      R4=1.
      DD=-B*(1./R3)*B+C*(1./R4)*C
C
C                                   Y(1)=Z1
C                                   Y(2)=LAMBDA1
C    Z1(0)=5.
C    Z1(1)=0.1
      PI1=5.
      PI2=0.1
      YSTART(1)=PI1
      YSTART(2)=(-A*(PI2-PI1*(DEXP(A*TEND))))/(DD*DSINH(-A*TEND))
      WRITE(6,10)  (YSTART(I),I=1,2)
   10 FORMAT(5X,'IC,S',2E15.6)
C
      SP=0.02
C
      CALL DDESP(SP,FUNC1,N,YSTART,TSTART,TEND,H,EP,DSE,OUT1,&99)
      GO TO 98
   99 CALL DERROR
   98 WRITE(6,5) NOFNS
    5 FORMAT(1H ,5X,36HTOTAL NO OF FUNCTION EVALUATIONS IS ,I6)
      ENDFILE 40
      DO 100 K=1,2
```

```
      REWIND 40
      DO 200  L=1,51
      READ(40,30) XX,(Z(I),I=1,2)
  30  FORMAT(3E15.6)
      F(L)=Z(K)
 200  TIME(L)=XX
      CALL PLTVAR(TIME,F,51)
      GO TO (1000,1001),K
1000  WRITE(6,301)
 301  FORMAT(20X,'F(T)=Z1(T)')
      GO TO 100
1001  WRITE(6,302)
 302  FORMAT(20X,'F(T)=LAMBDA1(T)')
 100  CONTINUE
      DO 735 J=1,51
      U1(J)=SNGL(-B/R3)*F(J)
 735  U2(J)=SNGL(C/R4)*F(J)
      CALL PLTVAR(TIME,U1,51)
      WRITE(6,305)
 305  FORMAT(20X,'F(T)=U1(T)')
      CALL PLTVAR(TIME,U2,51)
      WRITE(6,306)
 306  FORMAT(20X,'F(T)=U2(T)')
      STOP
      END
```

```
SUBROUTINE FUNC1(Y,T,DY)
DOUBLE PRECISION A,B,C,Y,T,DY,R3,R4,DD
DIMENSION Y(2),DY(2)
COMMON NOFNS,A,B,C,R3,R4
DD=-B*(1./R3)*B+C*(1./R4)*C
DY(1)=A*Y(1)+DD*Y(2)
DY(2)=-A*Y(2)
NOFNS=NOFNS+1
RETURN
END
```

```
      SUBROUTINE OUT1(Y,DY,N,X,SPTYPE,*)
      LOGICAL SPTYPE
      REAL Z,XX
      DOUBLE PRECISION Y,DY,X
      DIMENSION Y(2),DY(2),Z(2)
      IF(.NOT.SPTYPE) GO TO 100
      DO 20 K=1,N
  20  Z(K)=SNGL(Y(K))
      XX=SNGL(X)
      WRITE(40,10) XX,(Z(K),K=1,2)
  10  FORMAT(3E15.6)
 100  RETURN
      END
```

```
      SUBROUTINE PLTVAR(T,F,NP)
      REAL*4 LINE( 51),F(1),T(1)
      DATA BLANK,DOT,STAR,DASH,ZERO/1H ,1H.,1H*,1H-,1H0/
      DO 101 J=1,51
101   LINE(J)=DOT
      DO 127 L=1,10
127   LINE( 5*L+1)=DASH
      LINE(26)=ZERO
      LINE(1)=DASH
      PRINT 800
800   FORMAT(1H1)
      PRINT 102, LINE
102   FORMAT(1X , 51A1,5X,'TIME',7X,'F(T)')
      DO 103 J=1,51
103   LINE (J)=BLANK
      LINE(26)=DOT
      PM=ABS(F(1))
      DO 104 M=2,NP
      IF(PM .GE. ABS(F(M))) GO TO 104
105   PM=ABS(F(M))
104   CONTINUE
      IF(PM .EQ. 0.0) PM=1.0
      DO 107 M=1,NP
      J=25.0*F(M)/PM+26.5
      LINE(J)=STAR
      PRINT 106, LINE,T(M),F(M)
106   FORMAT(1X , 51A1,F10.4,E15.6)
      LINE(J)=BLANK
      LINE(26)=DOT
107   CONTINUE
      RETURN
      END
```

Figure E.2 Listing of program OPTIML

F(T)=Z1(T)

Figure E.3.1 Plot of z_1
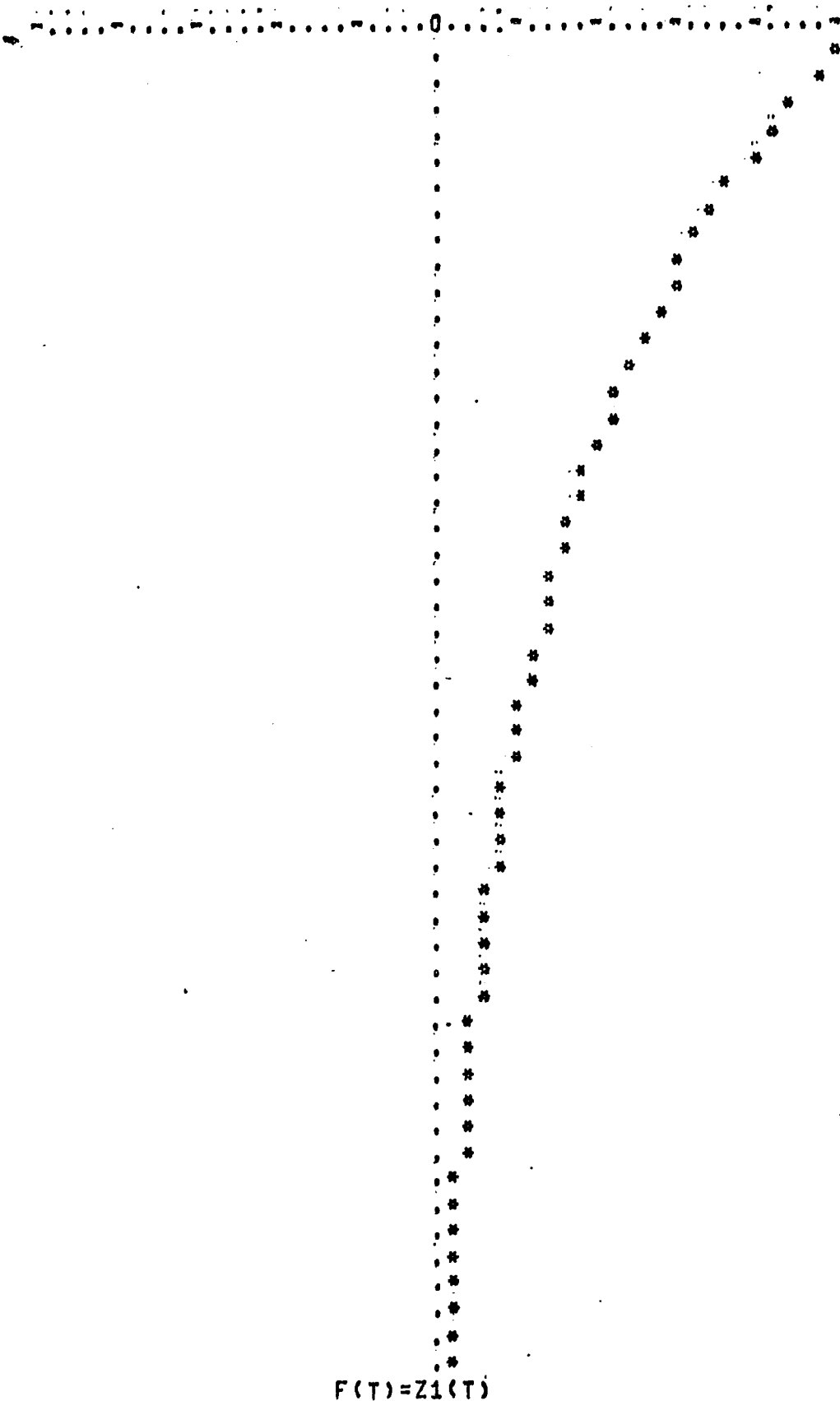
TIME	F(T)
0.0000	0.500000E
0.0200	0.470793E
0.0400	0.443281E
0.0600	0.417366E
0.0800	0.392954E
0.1000	0.369956E
0.1200	0.348291E
0.1400	0.327880E
0.1600	0.308650E
0.1800	0.290532E
0.2000	0.273459E
0.2200	0.257372E
0.2400	0.242211E
0.2600	0.227922E
0.2800	0.214454E
0.3000	0.201759E
0.3200	0.189790E
0.3400	0.178504E
0.3600	0.167861E
0.3800	0.157823E
0.4000	0.148353E
0.4200	0.139417E
0.4400	0.130984E
0.4600	0.123022E
0.4800	0.115503E
0.5000	0.108399E
0.5200	0.101687E
0.5400	0.953403E
0.5600	0.893371E
0.5800	0.836556E
0.6000	0.782753E
0.6200	0.731770E
0.6400	0.683421E
0.6600	0.637534E
0.6800	0.593942E
0.7000	0.552489E
0.7200	0.513026E
0.7400	0.475410E
0.7600	0.439507E
0.7800	0.405186E
0.8000	0.372324E
0.8200	0.340802E
0.8400	0.310508E
0.8600	0.281333E
0.8800	0.253170E
0.9000	0.225919E
0.9200	0.199481E
0.9400	0.173762E
0.9600	0.148669E
0.9800	0.124111E
1.0000	0.100000E

TIME	F(T)
0.0000	0.148670E
0.0200	0.157863E
0.0400	0.167625E
0.0600	0.177990E
0.0800	0.188996E
0.1000	0.200683E
0.1200	0.213092E
0.1400	0.226269E
0.1600	0.240261E
0.1800	0.255118E
0.2000	0.270894E
0.2200	0.287645E
0.2400	0.305432E
0.2600	0.324319E
0.2800	0.344373E
0.3000	0.365668E
0.3200	0.388280E
0.3400	0.412290E
0.3600	0.437784E
0.3800	0.464855E
0.4000	0.493600E
0.4200	0.524123E
0.4400	0.556533E
0.4600	0.590947E
0.4800	0.627489E
0.5000	0.666291E
0.5200	0.707492E
0.5400	0.751241E
0.5600	0.797695E
0.5800	0.847021E
0.6000	0.899398E
0.6200	0.955014E
0.6400	0.101407E
0.6600	0.107678E
0.6800	0.114336E
0.7000	0.121406E
0.7200	0.128913E
0.7400	0.136885E
0.7600	0.145350E
0.7800	0.154337E
0.8000	0.163881E
0.8200	0.174015E
0.8400	0.184775E
0.8600	0.196201E
0.8800	0.208334E
0.9000	0.221216E
0.9200	0.234896E
0.9400	0.249421E
0.9600	0.264844E
0.9800	0.281221E
1.0000	0.298611E

F(T)=LAMBDA1(T)

Figure E.3.2 Plot of λ_1

F(T)=U1(T)

Figure E.3.3 Plot of u_1

TIME	F(T)
0.0000	-0.297340l
0.0200	-0.315726l
0.0400	-0.335250l
0.0600	-0.355980l
0.0800	-0.377992l
0.1000	-0.401366l
0.1200	-0.426184l
0.1400	-0.452538l
0.1600	-0.480522l
0.1800	-0.510236l
0.2000	-0.541788l
0.2200	-0.575290l
0.2400	-0.610864l
0.2600	-0.648638l
0.2800	-0.688746l
0.3000	-0.731336l
0.3200	-0.776560l
0.3400	-0.824580l
0.3600	-0.875568l
0.3800	-0.929710l
0.4000	-0.987200l
0.4200	-0.104825l
0.4400	-0.113307l
0.4600	-0.118189l
0.4800	-0.125498l
0.5000	-0.133258l
0.5200	-0.141498l
0.5400	-0.150248l
0.5600	-0.159539l
0.5800	-0.169404l
0.6000	-0.179880l
0.6200	-0.191003l
0.6400	-0.202814l
0.6600	-0.215356l
0.6800	-0.228672l
0.7000	-0.242812l
0.7200	-0.257826l
0.7400	-0.273770l
0.7600	-0.290700l
0.7800	-0.308674l
0.8000	-0.327762l
0.8200	-0.348030l
0.8400	-0.369550l
0.8600	-0.392402l
0.8800	-0.416668l
0.9000	-0.442432l
0.9200	-0.469792l
0.9400	-0.498842l
0.9600	-0.529688l
0.9800	-0.562442l
1.0000	-0.597222l

TIME	F(T)
0.0000	0.148670E-
0.0200	0.157863E-
0.0400	0.167625E-
0.0600	0.177990E-
0.0800	0.188996E-
0.1000	0.200683E-
0.1200	0.213092E-
0.1400	0.226269E-
0.1600	0.240261E-
0.1800	0.255118E-
0.2000	0.270894E-
0.2200	0.287645E-
0.2400	0.305432E-
0.2600	0.324319E-
0.2800	0.344373E-
0.3000	0.365668E-
0.3200	0.388280E-
0.3400	0.412290E-
0.3600	0.437784E-
0.3800	0.464855E-
0.4000	0.493600E-
0.4200	0.524123E-
0.4400	0.556533E-
0.4600	0.590947E-
0.4800	0.627489E-
0.5000	0.666291E-
0.5200	0.707492E-
0.5400	0.751241E-
0.5600	0.797695E-
0.5800	0.847021E-
0.6000	0.899398E-
0.6200	0.955014E-
0.6400	0.101407E
0.6600	0.107678E
0.6800	0.114336E
0.7000	0.121406E
0.7200	0.128913E
0.7400	0.136885E
0.7600	0.145350E
0.7800	0.154337E
0.8000	0.163881E
0.8200	0.174015E
0.8400	0.184775E
0.8600	0.196201E
0.8800	0.208334E
0.9000	0.221216E
0.9200	0.234896E
0.9400	0.249421E
0.9600	0.264844E
0.9800	0.281221E
1.0000	0.298611E

F(T)=U2(T)

Figure E.3.4 Plot of u_2

The programs ANLPLT and PlTTPT were written to obtain computer plots of all the states and controls. Program VALUE evaluated the performance index, J for optimal states and controls. They are listed in Figures E.4, E.5, and E.6. The results of these programs are given in Chapter 7.

```
      PROGRAM ANLPLT
      DIMENSION Y(5),YY(51),TT(51)
      REAL*4 LMBD1
      T=0.
      TF=1.
      DELTAT=(TF-T)/50.
      PI1=5.
      PI2=0.1
      A=-3.
      B=2.
      C=1.
      R3=1.
      R4=1.
      DD=-B*(1./R3)*B+C*(1./R4)*C
      ETA=(-A)*(PI2-PI1*EXP(-3.))/(DD*SINH(3.))
      REWIND 30
100   Z1=PI1*EXP(-3.*T)+(DD/(-A))*ETA*(SINH(-A*T))
      Y(1)=Z1
      LMBD1=ETA*EXP(-A*T)
      Y(2)=LMBD1
      U1=(-B/R3)*LMBD1
      Y(3)=U1
      U2=(C/R4)*LMBD1
      Y(4)=U2
      G=(-A*ETA*2.)/(ETA*DD-(A*PI1*2.+ETA*DD)*EXP(-6.*T))
      Y(5)=G
      WRITE(30,10) T,(Y(I),I=1,5)
10    FORMAT(6E15.6)
      T=T+DELTAT
      IF(T.LT.TF) GO TO 100
      ENDFILE 30
      DO 5 K=1,5
      REWIND 30
      DO 6 J=1,51
      READ(30,10) T,(Y(I),I=1,5)
      YY(J)=Y(K)
6     TT(J)=T
5     CALL PLTVAR(TT,YY,51)
      STOP
      END
```

Figure E.4 Listing of program ANLPLT

```
      PROGRAM PITTPT
C     H. BURBANK   E.E.DEPT.
      REAL LMBD2,LMBD3
      DIMENSION Y(5),YY(51),TT(51)
      EK=-1./12.
      EL=1./144.
      EM=1.
      EN=10.
      WRITE(6,1) EK,EL,EM,EN
    1 FORMAT(5X,'K=',F10.5,'L=',F10.5,'M=',F10.5,'N=',F10.5)
      AA=EM-EL/12.
      AB=-EK/24.
      AC=EL/12.
      AD=-EK/48.
      AE=EK/16.
      WRITE(6,2) AA,AB,AC,AD,AE
    2 FORMAT(5X,'AA=',E15.6,'AB=',E15.6,'AC=',E15.6,'AD=',E15.6,'AE=',
     1E15.6)
      AF=EN-EK/96.
      AG=-0.5*(EM-(EL/12.)-(EK/144.))
      AH=-1.*((EL/24.)-(EK/288.))
      AI=0.5*(EM-(EL/12.))
      WRITE(6,3) AF,AG,AH,AI
    3 FORMAT(5X,'AF=',E15.6,'AG=',E15.6,'AH=',E15.6,'AI=',E15.6)
      AJ=-EK/144.
      AK=EL/24.
      AL=-EK/288.
      AM=EK/96.
      WRITE(6,4) AJ,AK,AL,AM
    4 FORMAT(5X,'AJ=',E15.6,'AK=',E15.6,'AL=',E15.6,'AM=',E15.6)
      AN=EL
      AO=-EK/4.
      AP=EK/4.
      WRITE(6,7) AN,AO,AP
    7 FORMAT(5X,'AN=',E15.6,'AO=',E15.6,'AP=',E15.6)
      AQ=EK
      WRITE(6,8) AQ
    8 FORMAT(5X,'AQ=',E15.6)
      C1=(EN-(EK/96.))*EXP(-3.)+(EK/96.)*EXP(3.)
      C2=-0.5*(EM-(EL/12.)-(EK/144.))*EXP(-3.)+0.5*(EM-(EL/12))*
     1   EXP(-1.)-(EK/144.)*EXP(3.)
      C3=-1.*((EL/24.)-(EK/288.))*EXP(-3.)+(EL/24.)*EXP(-1.)-(EK/288.)*
     1   EXP(3.)
      D1=(EM-(EL/12.))*EXP(-3.)+(EL/12.)*EXP(3.)
      D2=-1.*(EK/24.)*EXP(-3.)-1.*(EK/48.)*EXP(3.)+(EK/16.)*EXP(-1.)
      WRITE(6,10) D1,D2
   10 FORMAT(5X,'D,S',2E15.6)
      WRITE(6,11) C1,C2,C3
   11 FORMAT(5X,'C,S',3E15.6)
      WRITE(6,13)
```

```
   13 FORMAT(5X,'Z2(T,1)=D1*EXP(-T)+D2*EXP(3*T)')
      WRITE(6,14)
   14 FORMAT(5X,'Z3(1,TAU)=C1*EXP(-TAU)+C2*EXP(-3*TAU)+C3*EXP(3*TAU)')
      TAU=0.
   90 T=0.
      REWIND 27
      TF=1.
      DELTAT=(TF-T)/50.
  100 EP1T=EXP(T)
      EM1T=EXP(-T)
      EP3T=EXP(3.*T)
      EM3T=EXP(-3.*T)
      EP1TAU=EXP(TAU)
      EM1TAU=EXP(-TAU)
      EP3TAU=EXP(3.*TAU)
      EM3TAU=EXP(-3.*TAU)
      Z2=AA*EM1T*EM3TAU+AB*EP3T*EM3TAU+AC*EM1T*EP3TAU+AD*EP3T*EP3TAU
     1    +AE*EP3T*EM1TAU
      Z3=AF*EM3T*EM1TAU+AG*EM3T*EM3TAU+AH*EM3T*EP3TAU+AI*EM1T*EM3TAU
     1    +AJ*EP3T*EM3TAU+AK*EM1T*EP3TAU+AL*EP3T*EP3TAU+AM*EP3T*EM1TAU
      LMBD2=AN*EM1T*EP3TAU+AO*EP3T*EP3TAU+AP*EP3T*EM1TAU
      LMBD3=AQ*EP3T*EM1TAU
      U3=LMBD2/2.
      Y(1)=Z2
      Y(2)=Z3
      Y(3)=LMBD2
      Y(4)=LMBD3
      Y(5)=U3
      WRITE(27,12) T,(Y(I),I=1,5)
   12 FORMAT(6E15.6)
      T=T+DELTAT
      IF(T.LT.TF) GO TO 100
      ENDFILE 27
      DO 5 K=1,5
      REWIND 27
      DO 6 J=1,51
      READ(27,12) T,(Y(I),I=1,5)
      YY(J)=Y(K)
    6 TT(J)=T
      CALL PLTVAR(TT,YY,51)
      GO TO(1001,1002,1003,1004,1005),K
 1001 WRITE(6,301)
  301 FORMAT(20X,'F(T)=Z2(T,TAU)')
      WRITE(6,500) TAU
  500 FORMAT(20X,'TAU=',F10.4)
      GO TO 5
 1002 WRITE(6,302)
  302 FORMAT(20X,'F(T)=Z3(T,TAU)')
      WRITE(6,500) TAU
      GO TO 5
```

```
1003 WRITE(6,303)
 303 FORMAT(20X,'F(T)=LAMBDA2(T,TAU)')
     WRITE(6,500) TAU
     GO TO 5
1004 WRITE(6,304)
 304 FORMAT(20X,'F(T)=LAMBDA3(T,TAU)')
     WRITE(6,500) TAU
     GO TO 5
1005 WRITE(6,305)
 305 FORMAT(20X,'F(T)=U3(T,TAU)')
     WRITE(6,500) TAU
   5 CONTINUE
     ENDFILE 27
     TAU=TAU+0.2
     IF(TAU.LT.TF) GO TO 90
     STOP
     END
```

Figure E.5 Listing of program PlTTPT

```
      PROGRAM VALUE
C     H. BURBANK   E.E.DEPT.
      F(A)=(EXP(A)-1.)/A
      A=-3.
      B=2.
      C=1.
      R3=1.
      R4=1.
      DD=-B*(1./R3)*B+C*(1./R4)*C
      PI1=5.
      PI2=0.1
      EK=-1./12.
      EL=1./144.
      EM=1.
      EN=10.
      RT =(-A)*(PI2-PI1*EXP(-3.))/(DD*SINH(3.))
      WRITE(6,2) RT
    2 FORMAT(5X,'RT=',E15.6)
      V=0.5*RT*RT*F(6.)
      VV=-0.25
      VA=VV*EL*EL*F(-2.)*F(6.)
      VB=VV*(EK*EK/16.)*F(6.)*F(6.)
      VC=VV*(EK*EK/16.)*F(6.)*F(-2.)
      VD=VV*(-EL*EK/4.0)*F(2.)*F(6.)*2.
      VE=VV*(EL*EK/4.)*F(2.)*F(2.)*2.
      VF=VV*(-EK*EK/16.)*F(6.)*F(2.)*2.
      WRITE(6,3)V
      WRITE(6,3)VA
      WRITE(6,3)VB
      WRITE(6,3)VC
      WRITE(6,3)VD
      WRITE(6,3)VE
      WRITE(6,3)VF
    3  FORMAT(5X,E15.6)
      VALUE=V+VA+VB+VC+VD+VE+VF
      WRITE(6,1) VALUE
    1 FORMAT(5X,'VALUE=',E15.6)
      STOP
      END
```

Figure E.6 Listing of program VALUE

The combined identification and suboptimal control was
verified by the simulation of the model reference adaptive
system of Chapter 6 by the program P R O B 1 P. The flow
chart is contained in Figure E.7. The listing of P R O B 1 P
is given in Figure E.8 . The actual results are contained in
Chapter 7. Several examples were run for different sets
of unknown parameters.

```
┌─────────────────────┐
│    Read in data     │
└─────────────────────┘
           │
           ▼
┌─────────────────────┐
│    Obtain I.C.'s     │
└─────────────────────┘
           │
           ▼
┌─────────────────────┐
│   Integrate for one │
│   Sampling period   │
└─────────────────────┘
           │
           ▼
┌─────────────────────┐
│ First sampling      │──── Y ───► Initialize J_I, etc.
│ period?             │
└─────────────────────┘
           │ N
           ▼
┌─────────────────────┐
│     IFLAG=0?        │──── N ───►
└─────────────────────┘
           │ Y
           ▼
┌─────────────────────┐
│    End of Γ_1?      │──── Y ───►  END
└─────────────────────┘
           │ N
           ▼
┌─────────────────────┐
│  Percent change in  │           ┌──────────┐
│  J_I< stopping      │── Y ────► │ IFLAG=1  │
│  criteria?          │           └──────────┘
└─────────────────────┘
           │ N
           ▼
┌─────────────────────┐
│  Evaluate  ΔJ,Δγ    │
│  and a new γ        │
└─────────────────────┘
```
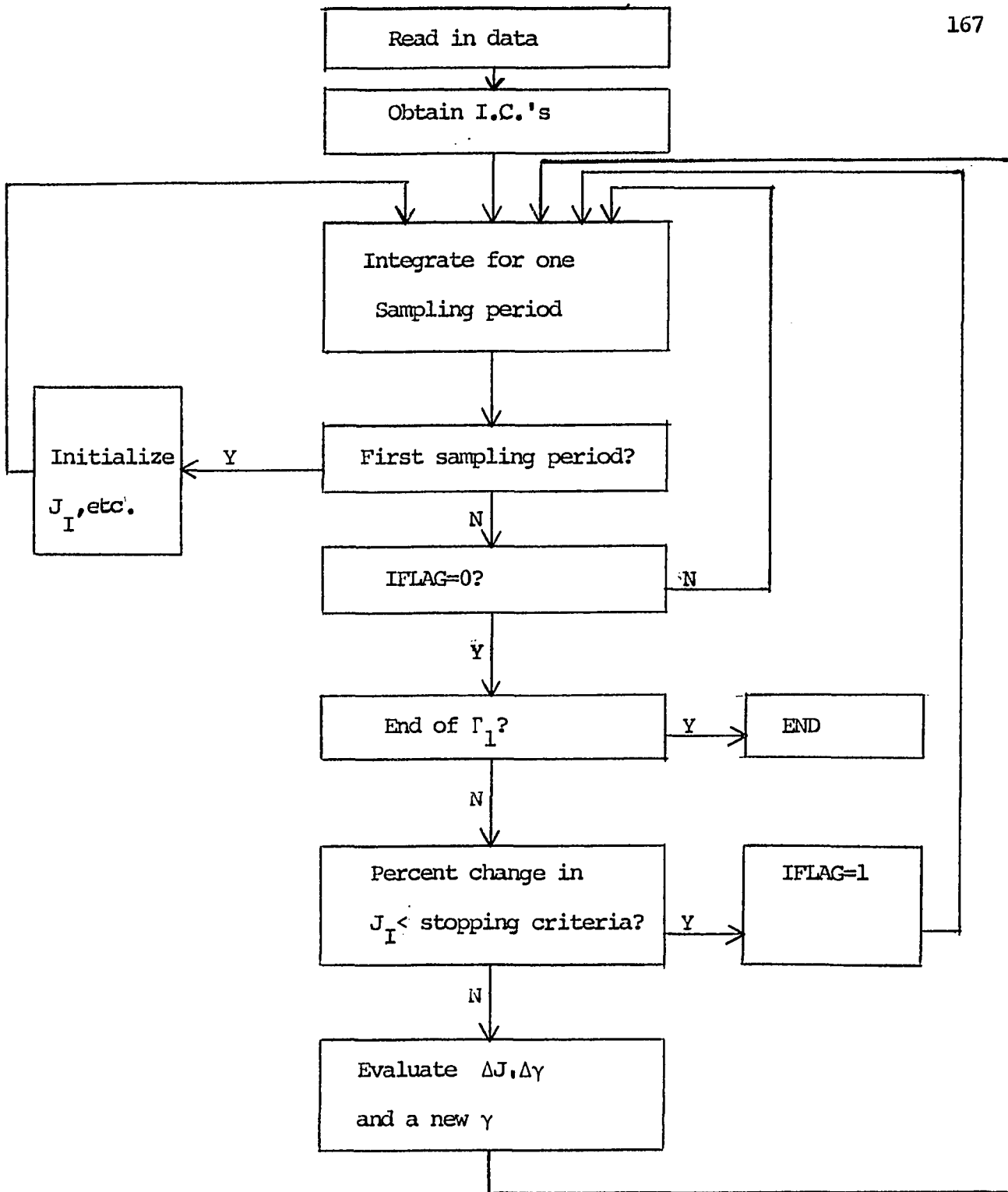
Figure E.7 Flowchart for identification problem

```
      PROGRAM PROB1P
C     H. BURBANK   E.E.DEPT.
C                     PROBLEM 1 WITH PLOTS
      DOUBLE PRECISION YSTART,XSTART,XEND,H,EP,SP,Y,DY,X,TSTART,TEND
      DOUBLE PRECISION A,B,C,R3,R4,APLT,BPLT,CPLT,V,DD
       DOUBLE PRECISION PTS,DELTAT,PI1,PI2,ASTART,BSTART,TOTA,TOTB
      DOUBLE PRECISION TINT,TSTRNU,TENDNU,PTSINT,TCHG,PJNORM
      DOUBLE PRECISION STATE,DIFF,SUM,MEAS,MEASOL,XX,MEAS1
      DOUBLE PRECISION DELA,DELB,DELJ,SCALEK,PJWRTA,PJWRTB,PCTCHG
      DIMENSION STATE(3),DIFF(30)
      EXTERNAL FUNC1,DSE,OUT1
      DIMENSION Y(3),DY(3),YSTART(3)
      DIMENSION  Z(3),F(501),TIME(501),U1(501),U2(501),TIMEL(501)
       DIMENSION FF(501)
      DIMENSION VNOISE(4)
      REAL TIME,F,U1,U2
      COMMON NOFNS,A,B,C,R3,R4,APLT,BPLT,CPLT,V
      V=0.
      APLT=-4.
      BPLT=3.0
      CPLT=1.0
      MEAS=0.
      IFLAG=0
      NOFNS=0
      NTIMES=0
      REWIND 50
      REWIND 51
      REWIND 52
C NPTS IS NO,S DELTA T,S PLUS 1 FOR 0.
C LPTS IS NO. OF POINTS PRINTED OUT
       NPTS=440
       LPTS=440
       NMULT=(NPTS-1)/(LPTS-1)
       WRITE(6,21) NMULT
   21  FORMAT(5X,6HNMULT=,I6)
        PTS=400.
        PI1=5.
        PI2=0.1
C
      N=3
C
      TSTART=0.
C
      TEND=1.
C
       DELTAT=(TEND-TSTART)/PTS
       H=DELTAT
       SP=DELTAT
C
C
```

```
      TINT=40.
      NTINT=40
      TSTRNU=TSTART
      TCHG=(TEND-TSTART)/TINT
      TENDNU=TCHG
      PTSINT=PTS/TINT
      NPTSIT=(NPTS)/NTINT
      WRITE(6,22)PTSINT,NPTSIT
   22 FORMAT(5X,'PTS PER INT',E15.6,5X,I4)
C
      EP=0.002
C
      A=-3.
      B=2.
      ASTART=A
      BSTART=B
      C=1.0
      R3=1.
      R4=1.
      DD=-B*(1./R3)*B+C*(1./R4)*C
C
C                                   Y(1)=Z1
C                                   Y(2)=LAMBDA1
C                                   Y(3)=XPLANT
C   Z1(0)=PI1
C   Z1(1)=PI2
C   XPLANT(0)=PI1
      YSTART(1)=PI1
      YSTART(2)=(-A*(PI2-PI1*(DEXP(A*TEND))))/(DD*DSINH(-A*TEND))
      YSTART(3)=PI1
      WRITE(6,10) (YSTART(I),I=1,3)
   10 FORMAT(5X,'IC,S',3E15.6)
C
C
10000 CALL DDESP(SP,FUNC1,N,YSTART,TSTRNU,TENDNU,H,EP,DSE,OUT1,&99)
      GO TO 98
   99 CALL DERROR
      GO TO 9999
   98 TSTRNU=TSTRNU+TCHG
      TENDNU=TENDNU+TCHG
      IF(TENDNU.GT.TEND) GO TO 9900
      ENDFILE 51
      REWIND 51
      DO 400 KL=1,NPTSIT
  400 READ(51,31) XX,(YSTART(I),I=1,3)
   31 FORMAT(4D25.16)
      ENDFILE 51
      REWIND 51
      WRITE(6,23) TSTRNU,TENDNU
   23 FORMAT(5X,'NEXT TIME INTERVAL IS FROM',E15.6,
```

```
      1        'TO',E15.6,'WITH')
        WRITE(6,10) (YSTART(I),I=1,3)
        YSTART(1)=YSTART(3)
C
C
C
C
C          PARAM CHG HERE
C
C
C
C
        IF (IFLAG.GT.0) GO TO 10000
        MEASOL=MEAS
        DO 1300 J=1,NPTSIT
        READ(51,31) XX,(STATE(I),I=1,3)
        DIFF(J)=(STATE(1)-STATE(3))**2
 1300   WRITE(6,40)DIFF(J)
   40   FORMAT(5X,D25.16)
        ENDFILE 51
        REWIND 51
C
        SUM=0.5*(DIFF(1)+DIFF(NPTSIT))
        NPTSIM=NPTSIT-1
        DO 1301 J=2,NPTSIM
 1301   SUM=SUM+DIFF(J)
        MEAS=SUM*DELTAT*20.*25.
        WRITE(6,41)SUM
   41   FORMAT(5X,'SUM=',D25.16)
        WRITE(6,1302)MEAS
        IF (NTIMES.EQ.0) MEAS1=MEAS
         PCTCHG=(MEAS/MEAS1)*100.
         WRITE(6,1312)PCTCHG
        PCTCHG=ABS(PCTCHG)
 1312    FORMAT(5X,'PERCENT CHANGE IN MEAS=',E15.6)
 1302   FORMAT(5X,'MEAS=',E15.6)
        NTIMES=NTIMES+1
        IF(NTIMES.GT.1) GO TO 1303
        MEAS1=MEAS
        DELA=0.1*A
        DELB=0.1*B
        TOTA=DELA
        TOTB=DELB
        A=A+DELA
        B=B+DELB
         WRITE(6,1304)A,B
 1304   FORMAT(5X,'A=',E15.6,'B=',E15.6)
        GO TO 10000
 1303   IF (PCTCHG.GT.20.) GO TO 1305
         IF (PCTCHG.GT.15.) GO TO 1306
        IF (PCTCHG.GT.10.) GO TO 1307
```

```
      IF (PCTCHG.GT.5.) GO TO 1308
      IF(PCTCHG.GT.2.) GO TO 1313
      IF(PCTCHG.GT.1.) GO TO 1314
      IF(PCTCHG.GT.0.1) GO TO 1315
      IF(PCTCHG.GT.0.05) GO TO 1316
      IFLAG=1
       GO TO 10000
 1305 SCALEK=1.0
      GO TO 1309
 1306 SCALEK=0.9
      GO TO 1309
 1307 SCALEK=0.8
      GO TO 1309
 1308 SCALEK=0.6
      GO TO 1309
 1313 SCALEK=0.5
      GO TO 1309
 1314 SCALEK=0.4
      GO TO 1309
 1315 SCALEK=0.2
      GO TO 1309
 1316 SCALEK=0.1
 1309 WRITE(6,1310)SCALEK
 1310 FORMAT(5X,'SCALEK=',E15.6)
      TOTA=ASTART-A
      TOTB=BSTART-B
      DELJ=MEAS-MEAS1
      PJWRTA=DELJ/TOTA
      PJWRTB=DELJ/TOTB
      PJNORM=DSQRT((PJWRTA**2)+(PJWRTB**2))
      WRITE(6,1319)PJNORM
 1319 FORMAT(5X,'PJNORM=',E15.6)
      WRITE(6,1311)DELJ,PJWRTA,PJWRTB
 1311 FORMAT(5X,'DELJ=',E15.6,'PJWRTA=',E15.6,'PJWRTB=',E15.6)
      DELA=SCALEK*PJWRTA*0.8
      DELB=SCALEK*PJWRTB*1.3
      A=A+DELA
      B=B+DELB
      WRITE(6,1304)A,B
      GO TO 10000
C
C
 9900 WRITE(6,5) NOFNS
    5 FORMAT(1H ,5X,36HTOTAL NO OF FUNCTION EVALUATIONS IS ,I6)
      ENDFILE 50
      DO 100 K=1,2
      REWIND 50
      DO 200  L=1,NPTS
      READ(50,30) XX,(Z(I),I=1,3)
   30 FORMAT(4E15.6)
```

```
        F(L)=Z(K)
  200 TIME(L)=XX
        DO 205 JJ=1,LPTS
        LL=1+(JJ-1)*NMULT
        FF(JJ)=F(LL)
  205       TIMEL(JJ)=TIME(LL)
        CALL PLTVAR(TIMEL,FF,LPTS)
        GO TO (1000,1001),K
 1000 WRITE(6,301)
  301 FORMAT(20X,'F(T)=Z1(T)')
        GO TO 100
 1001 WRITE(6,302)
  302 FORMAT(20X,'F(T)=LAMBDA1(T)')
  100 CONTINUE
        DO 735 J=1,LPTS
        U1(J)=SNGL(-B/R3)*FF(J)
  735 U2(J)=SNGL(C/R4)*FF(J)
        CALL PLTVAR(TIMEL,U1,LPTS)
        WRITE(6,305)
  305 FORMAT(20X,'F(T)=U1(T)')
        CALL PLTVAR(TIMEL,U2,LPTS)
        WRITE(6,306)
  306 FORMAT(20X,'F(T)=U2(T)')
 9998 ENDFILE 52
        REWIND 52
        DO 500 K=1,NPTS
        READ(52,530)WW
  500 U2(K)=WW
        CALL PLTVAR(TIMEL,U2,NPTS)
        WRITE(6,531)
  531 FORMAT(20X,'F(T)=V(T)')
  530 FORMAT(E15.6)
 9999 STOP
        END
```

```
SUBROUTINE FUNC1(Y,T,DY)
DOUBLE PRECISION A,B,C,Y,T,DY,R3,R4,DD,APLT,BPLT,CPLT,V
DIMENSION Y(3),DY(3),VNOISE(4)
COMMON NOFNS,A,B,C,R3,R4,APLT,BPLT,CPLT,V
CALL RANGEN(VNOISE,4)
V=VNOISE(4)
DD=-B*(1./R3)*B+C*(1./R4)*C
DY(1)=A*Y(1)+DD*Y(2)
DY(2)=-A*Y(2)
DY(3)=APLT*Y(3)+BPLT*(-1./R3)*B*Y(2)+CPLT*V
NOFNS=NOFNS+1
RETURN
END
```

```
      SUBROUTINE RANGEN(VNOISE,N)
      DIMENSION VNOISE(20)
      CALL RANORM(VNOISE,4)
      DO 1 K=1,N
    1 VNOISE(K)=VNOISE(K)/6.
      RETURN
      END
```

```
      SUBROUTINE OUT1(Y,DY,N,X,SPTYPE,*)
         DOUBLE PRECISION A,B,C,R3,R4,APLT,BPLT,CPLT,V
      COMMON NOFNS,A,B,C,R3,R4,APLT,BPLT,CPLT,V
      LOGICAL SPTYPE
      REAL Z,XX
      DOUBLE PRECISION Y,DY,X
      DIMENSION Y(3),DY(3),Z(3)
      IF(.NOT.SPTYPE) GO TO 100
      DO 20 K=1,N
   20 Z(K)=SNGL(Y(K))
      XX=SNGL(X)
      WRITE(50,10) XX,(Z(K),K=1,3)
      WRITE(51,11)X,(Y(K),K=1,3)
   11 FORMAT(4D25.16)
   10 FORMAT(4E15.6)
      WRITE(52,12)V
   12 FORMAT(E15.6)
  100 RETURN
      END
```

Figure E.8 Listing of program PROB1P

APPENDIX F

MINIMUM COVARIANCE ESTIMATION

For a plant

$$\frac{dx(t,.)}{dt} = A(t)x(t,.) + B(t)\ u(t) + C(t)v(t,.) \qquad \text{F.1}$$

with measurement

$$z(t,.) = H(t)x(t,.) + w(t,.) \qquad \text{F.2}$$

it is desired to obtain an optimal filter constrained by

$$\frac{d\hat{x}(t)}{dt} = F(t)\ \hat{x}(t) + G\ (t)\ z(t) + D(t)u(t) \qquad \text{F.3}$$

where F, G, and D are to be found. Define the error

$$\tilde{x}(t,.) = x(t,.) - \hat{x}(t) \qquad \text{F.4}$$

then

$$\frac{d\tilde{x}(t,.)}{dt} = [A(t) - G(t)\ H(t) - F(t)]x(t.,) + F(t)\tilde{x}(t.,)$$

$$+C(t)v(t,.) - G(t)w(t,.) + [B(t) - D(t)]u(t) \qquad \text{F.5}$$

If the filter is to be unbiased, then

$$E\{\ \tilde{x}(t,.)\} = E\{\ \frac{d\tilde{x}(t,.)}{dt}\} = 0 \ \forall\ t \in \Gamma_1 \qquad \text{F.6}$$

Equation F.6 holds if

$$D(t) = B(t) \qquad \text{F.7}$$

and

$$F(t) = A(t) - G(t)\ H(t) \qquad \text{F.8}$$

and if $\mu_v(t) = \mu_w(t) = 0 \forall t \in \Gamma_1$ or if not zero they are known

$\forall t$ and can be subtracted out. With these conditions substituted,

F.5 becomes

$$\frac{d\hat{x}(t)}{dt} = A(t)\hat{x}(t) + G(t)\ [z(t) - H(t)\hat{x}(t)] + B(t)u(t) \qquad \text{F.9}$$

from which an optimal unbiased estimate can be generated once

G(t) is found. The matrix G(t) is chosen to minimize the error

covariances V_{xx}, V_{wx}, and V_{vx} at final times, i.e. at t_f,

according to the weighted performance measure

$$J(G) = \int_{t_o}^{t_f}\int_{t_o}^{t_f} \mathrm{tr}\{\partial^2 \frac{V_{xx}(t,\tau)}{\partial t \partial \tau} + R(t)\frac{\partial V_{xx}(t,\tau)}{\partial \tau}$$

$$+ S(t)\frac{\partial V_{wx}(t,\tau)}{\partial \tau} + Q(t)\frac{\partial V_{vx}(t,\tau)}{\partial \tau}\} dt d\tau$$

F.10
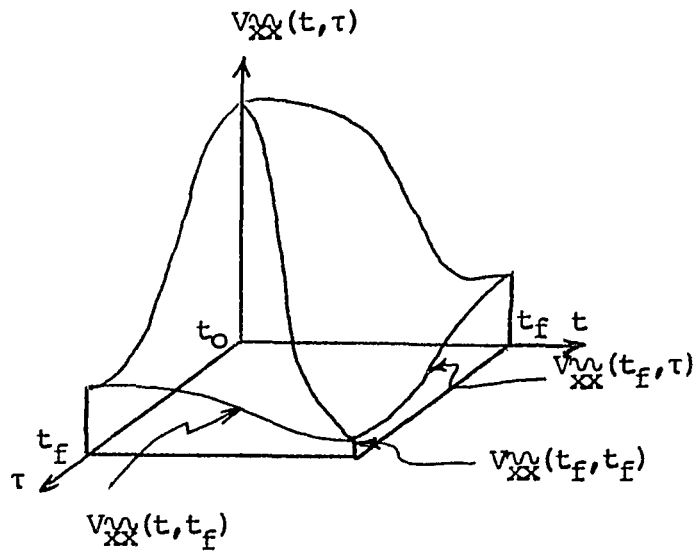
The physical meaning of J is determined from the following diagram.

Figure F.1

Boundary cost of error covariance

Since

$$V_{xx}(t_f,t_f) - V_{xx}(t_o,t_o) = \int_{t_o}^{t_f}\int_{t_o}^{t_f} \partial^2 \frac{V_{xx}(t,\tau)}{\partial t \partial \tau} dt d\tau \qquad \text{F.11}$$

and

$$\int_{t_o}^{t_f} R(t)V_{xx}(t,t_f)dt = \int_{t_o}^{t_f}\int_{t_o}^{t_f} R(t)\frac{\partial V_{xx}(t,\tau)}{\partial \tau}dt d\tau \qquad \text{F.12}$$

the terms in F.10 are seen to be terminal cost terms. Similar
relations hold for all four terms in J.

The a priori data required is

$$V_{vv}(t,\tau), \quad V_{vw}(t,\tau) \quad \text{and} \quad V_{wv}(t,\tau) \quad \forall (t,\tau) \in \Gamma \qquad \text{F.13}$$

and

$$V_{\tilde{x}\tilde{x}}(t,t_o), V_{w\tilde{x}}(t,t_o), V_{v\tilde{x}}(t,t_o), \quad \text{and} \quad V_{\tilde{x}\tilde{x}}(t_o,t_o) \forall t \in \Gamma_1 \qquad \text{F.14}$$

It can be shown by analysis similar to that in Appendix A,
that the error covariance must satisfy

$$\frac{\partial V_{\tilde{x}\tilde{x}}}{\partial t}(t,\tau) = [A(t)-G(t)H(t)]V_{\tilde{x}\tilde{x}}(t,\tau) + C(t)V_{v\tilde{x}}(t,\tau) - G(t)V_{w\tilde{x}}(t,\tau)$$

$$\text{F.15}$$

Similarly,

$$\frac{\partial V_{v\tilde{x}}}{\partial \tau}(t,\tau) = V_{v\tilde{x}}(t,\tau)[A(\tau)-G(\tau)H(\tau)]' + V_{vv}(t,\tau)C'(\tau) - V_{vw}(t,\tau)G'(\tau)$$

$$\text{F.16}$$

and

$$\frac{\partial V_{w\tilde{x}}}{\partial \tau}(t,\tau) = V_{w\tilde{x}}(t,\tau)[A(\tau)-G(\tau)H(\tau)]' + V_{wv}(t,\tau)C'(\tau) - V_{ww}(t,\tau)G'(\tau)$$

$$\text{F.17}$$

Substituting F.15 through F.17 into F.10 and choosing the weights

$$R(t) = G(t)H(t)$$

$$Q(t) = C(t) \qquad \text{F.18}$$

with $S(t)$ arbitrary, an unconstrained dynamic optimization
problem results. Applying variational techniques similar to
those used in Appendices C and D, Theorem 6.1 can be proven.

14. Balakrishnan, A. V. and Peterka, V., "Identification in Automatic Control Systems," _Automatica_, Vol. 5, pp. 817-829. Pergamon Press, Great Britain, 1969.

15. Beckenbach, E. and Bellman, R., "An Introduction to Inequalities," Random House, Conn., 1961.

16. Behn, R. D. and Y. C. Ho, "On a Class of Linear Stochastic Differential Games," _IEEE Trans. on A. C._, Vol. AC-13, June 1968.

17. Bhat, U. Naroyan, "Elements of Applied Stochastic Processes," John Wiley and Sons, Inc., N. Y., 1972.

18. Blaquiere, A. Gerard F. and Leitmann, G., "Quantitative and Qualitative Games," Academic Press, New York and London, 1969.

19. Booton, R. C., "An Optimization Theory for Time-Varying Linear Systems with Nonstationary Statistical Inputs," _Proceedings of the IRE_, August, 1952.

20. Bryson, A. and Ho, "Applied Optimal Control," Ginn and Company, Mass., 1969.

21. Carlson, A. Bruce, "Communication Systems," McGraw-Hill Book Co. New York, 1968.

22. Citron, S. J., "Elements of Optimal Control," Halt, Rinehart and Winston, Inc., New York, 1969.

23. Davis, R. C., "On the Theory of Prediction of Nonstationary Stochastic Processes," _Journal of Applied Physics_, Vol. 23, No. 9. September, 1952.

24. Ditkin, V. A. and Prudnikov, A. P., "Operational Calculus in Two Variables and Its Applications," Pergamon Press (Int. Series of Monographs in Pure and Applied Math), New York, 1962.

25. Dubes, R., "The Theory of Applied Probability," Prentice-Hall Inc., Englewood Cliffs, N. J., 1968.

26. Eggleston, H. G., "Convexity," Cambridge University Press, London, 1958.

27. Friedman, A., "Differential Games," Wiley-Interscience, N. Y., 1971.

28. Gabasov, R., and F. M. Kirillova, "High Order Necessary Conditions for Optimality," _SIAM J. Control_, Vol. 10, No. 1, p. 127, February, 1972.

29. Guillemin, E. A., "What is Natures Error Criteria?" IRE Trans. of P. G. on C. T. Vol. CT-1, 1954.

30. Gottfried, Byron and Weisman ,Joel, "Introduction to Optimization Theory," Prentice Hall, Inc. Englewood Cliffs, New Jersey, 1973.

31. Graupe, D., "Identification of Systems," Van Nostrand Reinhold Co., New York, 1972.

32. Graybill, Franklin A., "Introduction to Matrices with Application in Statistics," Wadsworth Publishing Co. Inc., Belmont, California, 1969.

33. Ho, Y. C., Bryson, A. and Baron, S., "Differential Games and Optimal Pursuit-Evasion Strategies," IEEE Trans. on A. C., AC 10, No. 4, October, 1965.

34. Ho, Y. C., "On Deterministic and Stochastic Differential Games," Studies in Optimization, SIAM, Philadelphia, 1970.

35. Hoel, P., Port, S. and Stone, C., "Introduction to Stochastic Processes," Houghton Mifflin Co., Boston, 1972.

36. Hohn, F., "Elementary Matrix Algebra," MacMillon Co., N. Y., 1964.

37. Holtzman, J. M., "Signal-Noise Ratio Maximization Using the Pontryagin Maximum Principle," The Bell System Technical Journal, March, 1966.

38. Hsu, J. C. and A. U. Meyer, "Modern Control Principles and Applications," McGraw-Hill Book Co., N. Y., 1968.

39. Isaacs, R., "Differential Games," John Wiley, New York, 1965.

40. Kalman, R. E. and R. S. Bucy, "New Results in Linear Filtering and Prediction Theory," J. Basic Eng., Vol. 83, 1961.

41. Kawabata, Nobuo, "A Nonstationary Analysis of the Electroencephologram," IEEE Transactions on Biomedical Engineering. Vol. BME 20, No. 6, Nov., 1973.

42. Kelley, Kopp and Moyer, "Singular Extremals," Topics in Optimization, Vol. 1, N. Y., Academic Press, 1966.

43. Kelley, Henry J., "A Second Variation Test for Singular Extremals," AIAAJ. 2, (380-382), 1964.

44. Kipniak, Walerian, "Dynamic Optimization and Control," The M.I.T. Press and John Wiley and Sons, Inc., N. Y., 1961.

45. Kirk, Donald E., "Optimal Control Theory," Prentice-Hall Inc., N. J., 1970.

46. Kleindorfer, Geo., Paul Kleindorfer, "Quadratic Performance Criteria with Linear Terms in Discrete-Time Control," IEEE Trans. on Automatic Control, June, 1967.

47. Kopp, Richard F., and H. Gardner Moyer, "Necessary Conditions for Singular Extremals," AIAA Journal, Vol. 3, No. 8, August, 1965.

48. Kramer, Land Athans M., "On the Application of Deterministic Optimization Methods to Stochastic Control Problems," IEEE Trans. on Automatic Control, Vol. AC-19, No. 1, February, 1974, p. 22.

49. Krikelis, N. J. and Z. V. Rekasius, "On the Solution of the Optimal Linear Control Problems under Conflict of Interest," IEEE Trans. on Automatic Control, Vol. AC-16, April, 1971.

50. Kuo, Marshall C. Y., "The Geometric Aspects of a Class of Differential Games, I Necessary Condtions," International Journal of Control, Vol. 10, No. 4, 1969.

51. Kuo, Marshall C. Y., "Minimax Principle for the Multiple Target Problem," Proceedings of the Seventh Annual Princeton Conference on Information Sciences and Systems. Princeton, N. J., March, 1973.

52. Kuo, Marshall C. Y., "The Geometric Aspects of a Class of Differential Games, II. Sufficient Conditions," International Journal of Control, Vol. 10, No. 4, 1969.

53. Kuo, Marshall, C. Y., "A Game Approach to the Design of Optimal Linear Controllers with Inaccessible State Variables," Proc. Second Asilomar Conf. on Circuits and Systems, Nov., 1968.

54. Kuo, Marshall C. Y., "Analysis of Nonstationary Processes from the State Variable Approach," Eighth Annual Princeton Conference on Information Sciences and Systems, Princeton, N. J., March, 1974.

55. Kuo, Marshall C. Y. and Harry M. Burbank, "Sufficient Conditions for a Class of Differential Games," Proc. of the First International Conf. on the Theory and Applications of Differential Games, Amherst, Mass., 1969.

56. Kuo, Marshall C. Y., and Harry Burbank, "A Class of Stochastic Optimal Control Problems Under Conflict of Interest," Proceedings of the 1975 Conference on Information Sciences and Systems, Baltimore, Maryland, 1975.

57. Kuo, Marshall C. Y., and Harry Burbank, "Optimal Estimation and Identification with Second Order Stochastic Processes," to be submitted for publication I.E.E.E. Tran. on Automatic Control, 1975.

58. Kuo, Marshall C.Y., and Harry Burbank,"A Class of Stochastic Controls with Vector Criteria , " to be submitted for publication, International Journal of Control, 1975.

59. Kuo, Marshall C.Y., Doctoral Dissertation, University of Michigan, Michigan, 1964.

60. Kurtaran, B. F. and Menachem, S., "Optimal Instantaneous Output Feedback Controllers for Linear Stochastic Systems," International Journal of Control, Vol. 19, No. 4, 1974.

61. Kushner, H. J. and S. G. Chamberlain, "Finite State Stochastic Games: Existence Theorems and Computational Procedures," IEEE Trans. on Automatic Control, Vol. AC-14, No. 3, June, 1969.

62. Kushner, Harold, "Introduction to Stochastic Control," Holt, Rinehart and Winston, Inc., New York, 1971.

63. Lebedev, V. L. "Random Processes in Electrical and Mechanical Systems," U. S. Dept. of Commerce, Washington, D. C., 1958.

64. Lee, Robert, "Optimal Estimation, Identification and Control," The M.I.T. Press, Cambridge, Massachusetts, 1964.

65. Lee, E. B. and L. Markus, "Foundations of Optimal Control Theory," John Wiley and Sons, New York, 1967.

66. Lee, J. S., "Optimal Linear Filter and Its Relation to Singular Optimal Control," Preprints of 11th JACC, p. 313, June, 1970.

67. Leitmann, C., "An Introduction to Optimum Control," McGraw-Hill Book Co., N. Y., 1966.

68. Leondes, C. T. and D. M. Stuart, "Stochastic Differential Games with Complexity-Constrained Strategies," IEEE Trans. on Automatic Control, February, 1972.

69. Leondes, Cornelius and J. O. Pearson, "A Minimax Filter for Systems with Large Plant Uncertainties," IEEE Trans. on Automatic Control, April, 1972.

70. Lewis, T. O. and P. L. Odell, "Estimation in Linear Models," Prentice-Hall, Inc., Englewood Cliffs, N. J., 1971.

71. Liberty, Stanley R. and Michael Sain, "Minimal Variance Feedback Controllers: Initial Studies of Solutions and Properties," Proc. Sixth Allerton Conf. on Circuit and System Theory, pp. 408-417, October, 1968.

72. Marcus, M. and Minc, H., "A Survey of Matrix Theory and Matrix Inequalities," Allyn & Bacon, Inc., Boston, 1964.

73. Meditch, J. S., "Stochastic Optimal-Linear Estimation and Control," McGraw-Hill Book Co., New York, 1969.

74. Mehra, Raman K., "On-Line Identification of Linear Dynamic Systems with Applications to Kalman Filtering," IEEE Trans. on Automatic Control, Vol. AC-16, No. 1, February, 1971.

75. Melsa, J. and Sage, A., "An Introduction to Probability and Stochastic Processes," Prentice-Hall, Inc., Englewood Cliffs, N. J., 1973.

76. Mendel, J., "Discrete Techniques of Parameter Estimation," Marcel Dekker, Inc., N. Y., 1973.

77. Mitrinovic, D. S., Barnes, E. S., Marsh, D.C.B. and Radok, J.R.M., "Elementary Inequalities," P. Noordhoff Ltd., Gronigen, The Netherlands, 1964.

78. Mitrinovic, D. S., "Analytic Inequalities," Springer-Verlag, N. Y., 1970.

79. Moore, J. B., "A Note on a Singular Optimal Control Problem," Automatica, Vol. 5, pp. 857-858, Pergamon Press, Great Britain, 1969.

80. Moylan, P. J. and J. B. Moore, "Generalizations of Singular Optimal Control Theory," Automatica, Vol. 7, pp. 591-598, Pergamon Press, Great Britain, 1971.

81. Murphy, G. J., "Mean-Weighted Square-Error Criterion for Optimum Filtering of Nonstationary Random Processes," IRE Trans. on Automatic Control, Vol. AC-6, No. 2, pp. 211-216, May, 1961.

82. Nunez, Paul, "Wavelike Properties of the Alpha Rhythm," IEEE Trans. on Biomedical Engineering, Vol. BME-21, No. 6, November, 1974.

185

83. Ogata, K., "State Space Analysis of Control Systems," Prentice-Hall, Inc., Englewood Cliffs, N. J., 1967.

84. Papoulis, Athanasios, "Probability, Random Variables and Stochastic Processes," McGraw-Hill, N. Y., 1965.

85. Parthasarathy, T. and Raghaven, T.E.S., "Some Topics in Two-Person Games," American Elsenier Publishing Co., Inc., N. Y. 1971.

86. Parzen, E., "Stochastic Processes," Holden-Day Inc., San Francisco, 1962.

87. Petrov, Iu. P., "Variational Methods in Optimum Control Theory," Academic Press, New York & London, 1968.

88. Pontryagin, L. S., V. Boltyanskii, R. Gamkrelidge, and E. Mishchenke, "The Mathematical Theory of Optimal Processes," Interscience Publishers, Inc., N. Y., 1962.

89. Pugachev, V. S., "Theory of Random Functions," Addison-Wesley Publishing Co., Reading, 1965.

90. Prabhu, N. V., "Stochastic Processes," The Macmillan Co., N. Y., 1965.

91. Pugachev, V. S., "The Determination of an Optimal System by Some Arbitrary Criterion," Automation and Remote Control (U.S.S.R.) pp. 513-532, December 16, 1957.

92. Ragade, R. and I. Sarma, "Game Theoretic Approach to Optimal Control in the Presence of Uncertainty," Transactions on Automatic Control, Vol. AC-12, No. 4, August, 1967.

93. Rekasius, Z. V., "A General Performance Index for Analytical Design of Control Systems," IRE Transactions on Automatic Control, Vol. AC-6, pp. 217-222, 1961.

94. Saaty, Thomas L. and Bram, J., "Nonlinear Mathematics," McGraw-Hill Book Co., New York, 1964.

95. Sage, A. P., "Optimum Systems Control," Prentice-Hall, Englewood Cliffs, N. J., 1968.

96. Sage, Andrew P. and Melsa, James L., "System Identification," Academic Press, N. Y., 1971.

97. Sage, A. and Melsa, J., "Estimation Theory with Applications to Communications and Control," McGraw-Hill Book Co., New York, 1971.

98. Sain, M. and S. Liberty, "Performance-Measure Densities for a Class of LQG Control Systems," IEEE Trans. on Automatic Control, Vol. AC-16, No. 5, October, 1971.

99. Sain, M. K., "Relative Costs of Mean and Variance Control for a Class of Linear, Noisy Systems," Proc. Third Allerton Conf. on Circuit and System Theory, pp. 121-129, October, 1965.

100. Sain, M. K. and Stanley R. Liberty, "Some New Computational Results for Minimum Variance Controllers," Proc. Seventh Allerton Conf. on Circuit and System Theory, October, 1969.

101. Sain, M. K., "Control of Linear Systems According to the Minimal Variance Criterion - A New Approach to the Disturbance Problem," IEEE Trans. on Automatic Control, Vol. AC-11, pp. 118-121, January, 1966.

102. Schweppe, Fred C., "Uncertain Dynamic Systems," Prentice-Hall, Englewood Cliffs, 1973.

103. Sherman, Seymour, "Non-Mean-Square Error Criteria," IRE Trans. on Information Theory, Vol. IT-4, No. 3, p. 125, 1958.

104. Shinbrot, M., "A Generalization of a Method for the Solution of the Integral Equation Arising in Optimization of Time-Varying Linear Systems with Nonstationary Inputs," IRE Transactions on Information Theory, December, 1957.

105. Sorenson, H. W. and D. L. Alspach, "Recursive Bayesian Estimation Using Gaussian Sums," Automatica, Vol. 7, pp. 465-479, Pergamon Press, Great Britain, 1971.

106. Stratonovich, R. L., "Topics in the Theory of Random Noise," Gordon and Breach, New York & London, 1963.

107. Wilde, D. and Beightler, C., "Foundations of Optimization," Prentice-Hall, Inc., Englewood Cliffs, N. J., 1967.

108. Willman, Warren, "Conflict and Decentralization in Stochastic Differential Games," Differential Games: Theory and Applications, 1970 Joint Automatic Control Conference, Georgia Institute of Technology, Atlanta, 1970.

109. Wonham, W. M., "On the Separation Theorem of Stochastic Control," SIAM J. Control, Vol. 6, 1968.

110. Yoshikawa, T., "An Example of Stochastic Multistage Games with Noisy State Observation," IEEE Trans. on Automatic Control, Vol. AC-15, August, 1970.

VITA

Harry Burbank was born in   on   .
He received his B.S.E.E. in June, 1967, and his M.S.E.E. in
June, 1969 from Newark College of Engineering. He began
work on this dissertation in 1969 and remained at Newark
College of Engineering full time until June, 1971. During this
time he received support from the National Science Foundation,
and from an N.C.E. Alumni Association grant. He was then
employed by the Bendix Corporation, Teterboro, New Jersey
until September, 1974. At that time he became an Instructor
in the Electrical Engineering Department of Lafayette College,
Easton, Pennsylvania, where he is currently employed.