APPLICATION OF PONTRYAGIN'S MINIMUM PRINCIPLE TO FILTERING PROBLEMS

by

EDISON TACK-SHUEN TSE

SUBMITTED IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF
MASTER OF SCIENCE

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
June 1967

Signature of Author --
Department of Electrical Engineering, May, 1967

Certified by --
esis Supervisor

Accept ------------
Chairman, Departmental Committee on Graduate Students

APPLICATION OF PONTRYAGIN'S MINIMUM PRINCIPLE TO FILTERING PROBLEMS

by

EDISON TACK-SHUEN TSE

Submitted to the Department of Electrical Engineering in May, 1967 in partial fulfilment of the requirements for the degree of Master of Science.

ABSTRACT

The purpose of this thesis is to consider the formulation and solution of a class of linear and nonlinear stochastic problems pertaining to control and estimation. These stochastic problems are reformulated as essentially deterministic problems. The matrix minimum principle is used to derive necessary conditions for optimality; these are used to obtain optimal and suboptimal designs.

Thesis Supervisor: Michael Athans

Title: Associate Professor of Electrical Engineering

ACKNOWLEDGEMENT

I would like to express my sincere gratitude to Professor Michael Athans for his constant encouragement, friendly advice and sound criticism. The discussions with Professors Michael Athans, Fred C. Schweppe, Donald L. Snyder and Mr. Patrick Lee are most helpful and stimulating.

At this time I am happy to express my appreciation to Professor Lan Jen Chu for his encouragement, comment and support while acting as my Undergraduate Faculty Counselor. My interest in stochastic processes relating to engineering problems comes from the inspiring teaching of Professor Wilbur B. Davenport, Jr. and Professor Harry L. Van Trees, Jr.

I would like to also thank Mrs. Ionia Lewis for coordinating the thesis and Mrs. Clara Conover for typing it.

CONTENTS

CONTENTS (Contd.)

CONTENTS (Contd.)

LIST OF FIGURES

NOTE TO THE READER:

Throughout this thesis, unless otherwise defined, lower case letters will stand for column vectors (e.g., x, y); upper case letters will represent matrices (e.g., A, B); lower case letters with subscripts will denote components (e.g., x_i will be the i-th component of the vector x, a_{ij} will be the ij-th entry of the matrix A).

The transpose of a matrix A is denoted by A'. The transpose of a column vector, x, is a row vector and is denoted by x'.

Let A be an nxn square matrix; the trace of A is defined as

$$\mathrm{Tr}\, A = \sum_{i=1}^{n} a_{ii}$$

Let $H(x_{11}, x_{12}, \ldots, x_{nm})$ be a scalar function; we shall denote it by H(X). The gradient matrix is defined by

$$\frac{\partial H(X)}{\partial X} \triangleq \left(\frac{\partial H(x_{11}, x_{12}, \ldots, x_{nm})}{\partial x_{ij}} \right)$$

Let Ω be a metric space, the closure of Ω will be denoted by $\bar{\Omega}$.

Let V be an inner product space, if v_1, v_2 are in V, the inner product of these two elements will be denoted by $< v_1, v_2 >_V$.

M_{nm} will denote the set of all nXm matrices.

R_n will denote the product space of ordered n-tuples of real numbers, we shall denote the elements in R_n by column vector x.

CHAPTER I

INTRODUCTION

The general problem of engineering system design in the presence of uncertainty has received much attention. (See for example Refs. 12, 28.) It turns out that design problems can be formulated and viewed as special cases of general decision problems under uncertainty.[28*] In this thesis we shall approach these engineering problems from this point of view.

In a general decision problem, there are basically three nonempty sets: D, the action space; N, the states of uncertainty, and a well-ordered set O, the set of outcomes; furthermore, there is a mapping M

$$M : D \times N \rightarrow O \tag{1.1}$$

Loosely speaking, we are given the three sets, D, N and O, and the mapping M, and we are to find an element in D, such that, under all uncertainties, it will be in some sense more preferable than the other elements in D.[†] Unfortunately, few specific results can be obtained under this general formulation. If we attach more structure to the sets N, D and O, and also be more specific about the mapping M, then we may be able to find the "best" action among the elements of the action space D.

Let us consider two special cases that are of interest. In the first case, let us suppose that N consists of only one element, D is a set of measurable functions defined in the time interval $[t_o, t_1]$, O is a set of 3-tuples (x, y, y_o) where x and y are measurable functions defined in the time interval $[t_o, t_1]$ and y_o is a real number. The mapping M maps x into (x, y, y_o) where x, y and y_o are related by

---

*Superscripts refer to numbered items in the Bibliography.

†For more details see Ref. 28.

$$\left.\begin{aligned} \dot{y}(t) &= f_M(x, y, t) \\ y(t_o) &= y_o \end{aligned}\right\} \tag{1.2}$$

The elements in O are well-ordered, i.e., there is a functional from O into the set of real numbers R

$$J : O \rightarrow R \tag{1.3}$$

such that†

$$(x, y, y_o) \geq (\overline{x}, \overline{y}, \overline{y}_o) \text{ iff } J(x, y, y_o) \leq J(\overline{x}, \overline{y}, \overline{y}_o) \tag{1.4}$$

We immediately recognize this as a deterministic optimal control problem where we choose x to minimize $J(x, y, y_o)$ subject to the constraint (1.2). For this class of problem, the Pontryagin maximum (or minimum) principle gives the necessary conditions for optimality. (See Refs.1, 23 and 25) Another case is when $D = N =$ set of measurable functions defined in the interval $[t_o, t_1]$, O is a set of 4-tuples (x, n, y, y_o) where x, n, y are measurable function and y_o is a real number; the mapping M maps $(x, n) \epsilon D \times N$ into (x, n, y, y_o) where x, n, y, y_o satisfy

$$\left.\begin{aligned} \dot{y}(t) &= f_M(x, n, y, t) \\ y(t_o) &= y \end{aligned}\right\} \tag{1.5}$$

The set O is well-ordered, and

$$(x, n, y, y_o) \geq (\overline{x}, \overline{n}, \overline{y}, \overline{y}_o) \text{ iff } J(x, n, y, y_o) \leq J(\overline{x}, \overline{n}, \overline{y}, \overline{y}_o) \tag{1.6}$$

where J is again a functional which maps O into the set of real

---

†The ordering in O is defined in such a way that if o_1 in O is more preferable than o_2 in O, then $o_1 \geq o_2$.

numbers R. This class of problems we shall call optimal control problems under uncertainty. In this case, it is difficult to say which decision is more preferable, because it depends on the uncertainty which is unknown to us. If we suppose further that the uncertainty space N has a σ-algebra structure[21 and 11] and defined on it a probability measure, then it is <u>reasonable</u> to say that the decision x is more preferable than $\overline{x}$ if

$$\int_N J(x, n, y, y_o)\,d\mu \leq \int_N J(\overline{x}, \overline{n}, \overline{y}, \overline{y}_o)\,d\mu \tag{1.7}$$

where μ is the measure defined on the σ-algebra of N. This type of problem is refered to as a stochastic control problem.[13, 14, 30] The mathematical technique for solving this specific formulation of problem is rather involved because we are dealing with stochastic differential equations, and stochastic integrals instead of ordinary differential equations and integrals.[20, 29, 31] Yet, if we note that the stochastic control problem is but one special formulation of uncertainty problems, we can reformulate the problem to avoid the difficulties that may arise in stochastic problems. This thesis represents an attempt in this direction. We shall reformulate the problems as deterministic control problems, and the nature of uncertainty may be incorporated in the functional J or in the reformulation of the problem; in this manner we will be able to apply the Pontryagin maximum principle.

Pontryagin et al.[23] proved the maximum principle which is applicable to the minimization of a given functional subject to differential equation constraints. In Chapter II, we shall extend it so as it becomes applicable when the system is described by matrix differential equations. We shall refer to the result as the <u>matrix-minimum principle</u>.

In Chapter III, we shall investigate stochastic regulator problems. These problems had been investigated by Florentin[13] and Wonham.[30]

Their approach is to define a stochastic control problem and apply the technique of Dynamic Programming to obtain the optimal control. This further reduces to the problem of solving a partial differential equation.[13] Specific results can be obtained if the system is linear with quadratic criteria.[13]

We shall deviate from this conventional approach. First, we shall reformulate the stochastic problem as a deterministic problem; then, we shall apply the matrix minimum principle to obtain the optimal control. We shall see that the nature of uncertainty will be incorporated in the reformulation of the problem. The results obtained are found to be the same as those obtained by the Dynamic Programming approach. In this framework, we can easily solve the optimal regulator problem when observation noise is added to the output of the system. This problem has also been studied by Meier,[22] Roberts,[24] and Wonham[30] using a different approach from our own. We shall also find that the results obtained in the course of this research work agrees with theirs.

In Chapter IV, we shall study the nonlinear filtering problem. This problem has been investigated by Bucy,[8] Cox,[9] Kushner,[9,20] Snyder,[27] and Wonham.[31] The approach of these authors is the stochastic one. Few results of practical significance have been obtained;[19] this is due to the difficulties that arise in most stochastic processes. Other approaches have also been attempted.[7,10]

In this thesis, we shall use the decision approach; we shall formulate the problem as a deterministic control problem while the nature of uncertainty is incorporated in the cost functional. A justification for this formulation will be given. If the system is not very nonlinear, we shall see that the results obtained here agree with the results obtained by Snyder.[27]

In Chapter V, we shall work on a simple example on nonlinear filtering. We shall take a quadratic drag system and investigate the effect of the second order dynamic of the nonlinear system. This example also shows some of the advantages of the decision approach over the conventional stochastic approach.

In summary, the work is an attempt to study a new viewpoint toward design problems. Instead of worrying about the existing problems in stochastic processes, we shall concentrate on the physical behavior of the system under uncertainty. Extensions may be possible when the space of uncertainty does not have a σ-algebra structure. Also it is hoped that this study will provide more insight into the structure of the system to be designed.

CHAPTER II

CONDITIONS FOR OPTIMALITY

We shall be considering the mathematical problem of minimizing a given functional subject to a given matrix-differential equations. We can view this as an extension of the work by Pontryagin, et al.[23] because the nature of the proof depends heavily on their results. To formulate the problem vigorously, we shall state in a general context the definition of class of admissible control, even though the set of admissible controls we shall be considering is a particular class. We shall then state and prove the matrix minimum principle, which gives us a necessary condition for optimality. The local sufficient conditions will be stated and proved. The result of this work is, in fact, "buried" in Pontryagin's work, but because of the mathematical form of the results obtained we may view Pontryagin's minimum (or maximum) principle as a special case of the results obtained here.

A. CLASS OF ADMISSIBLE CONTROLS

In this section we shall define mathematically the class of functions that we are allowed to choose in our optimization problem. Physically, the class of functions that we are considering has the property of separating past and future.

Let Ω be an arbitrary subset of the r-dimensional vector space E_r. We shall call the set Ω the control region. Let u be an element in Ω; u is represented by the r-tuple $(u_1, \ldots, u_r)$. Consider a function $u(t)$ defined on the time interval $[t_o, t_1]$ with range in Ω; we shall call $u(t)$ a control. The control $u(t)$, $t_o \leq t \leq t_1$, is measurable if the set $T_U = \{t_o \leq t \leq t_1 : u(t) \epsilon U\}$ is measurable for every open set $U \subset E_r$ (see Ref. 21). The control $u(t)$, $t_o \leq t \leq t_1$, is bounded if the set of points $u(t)$, $t_o \leq t \leq t_1$, has a compact closure in E_r. We shall now give a precise definition for the class of admissible controls.

Definition 2.1: Let $D_{[t_o, t_1]}$ be a subset of the set of all controls defined on the interval $[t_o, t_1]$. $D_{[t_o, t_1]}$ is called a class of admissible controls if for any $u(t)$, $t_o \leq t \leq t$, in $D_{[t_o, t_1]}$, we have

(1) $u(t)$, $t_o \leq t \leq t_1$, is measurable and bounded.

(2) For $v \in \Omega$, the control defined by

$$u'(t) = \begin{cases} v & \text{for } t' < t \leq t'' \\ \\ u(t) & \text{for } t_o \leq t \leq t' \text{ or } t'' < t \leq t_1 \end{cases} \tag{2.3}$$

is also in $D_{[t_o, t_1]}$.

(3) For $\alpha > 0$, the control defined by

$$u''(t) = u(t-\alpha);\ t_o + \alpha \leq t \leq t_1 + \alpha \tag{2.4}$$

is also in $D_{[t_o, t_1]}$.

(4) $u(t)$, $t' \leq t \leq t''$, is in $D_{[t', t'']}$ where $t_o \leq t' \leq t'' \leq t_1$.

The class of all measurable and bounded controls and the class of all piecewise constant controls are examples of class of admissible controls. In this chapter, we consider the case where $E_r = R_r$; and $D_{[t_o, t_1]}$ is the set of all bounded piecewise continuous function $u(t)$, $t_o \leq t \leq t_1$, such that

(1) $\qquad u(t) \epsilon \Omega \qquad t_o \leq t \leq t_1$ (2.5)

(2) $\qquad u(t-) = u(t) \qquad t_o \leq t \leq t_1$ (2.6)

B. THE MATRIX MINIMUM PRINCIPLE: A NECESSARY CONDITION FOR OPTIMALITY

Consider the system of differential equations

$$\frac{dx_{ij}}{dt} = f_{ij}(x_{11}, x_{12} \ldots x_{nm}, u_{11}, u_{12}, \ldots u_{rs}) \quad \begin{array}{l} i=1, \ldots n \\ i=1, \ldots m \end{array} \tag{2.7}$$

or in the matrix form

$$\frac{dX}{dt} = F(X, U) \tag{2.8}$$

The functions $f_{ij}(X, U)$ and $\frac{\partial f_{ij}}{\partial x_{mn}}$ are assumed to be given and to be continuous on the direct product $R_{nm} \times \overline{\Omega}$. Consider the integral functional

$$J = \int_{t_o}^{t_1} f_o(X, U)\, dt \quad ; t_1 \text{ Free} \tag{2.9}$$

We wish to choose among all admissible controls those which will transfer $X(t)$ from $X(t_o) = X_o$ to $X(t_1) = X_1$, and which will give us a minimum for the functional (2.9). The desired controls are called optimal controls. We now state and prove a necessary condition for optimality.

<u>Theorem 2.1</u>: (Matrix minimum principle) Let $U^*(t)$ be an admissible control with range Ω, which transfers (X_o, t_o) to $S = \{X_1\} \times (T_1, T_2)$. Let $X^*(t)$ be the trajectory of (2.7) corresponding to $U^*(t)$ with $X^*(t_o) = X_o, X^*(t_1) = X_1$. In order that

$U^*(t)$ is optimal, it is necessary that there exist a scalar $p_o^* \geq 0$ and a nxm costate matrix $P^*(t)$ such that

(1) $X^*(t)$, $P^*(t)$ satisfies the canonical equations:

$$\dot{X}^*(t) = \frac{\partial H}{\partial P(t)}\bigg|_*$$

$$\dot{P}^*(t) = -\frac{\partial H}{\partial X(t)}\bigg|_* \tag{2.10}$$

Satisfying the boundary conditions

$$X^*(t_o) = X_o \ , \ X^*(t_1) = X_1 \tag{2.11}$$

with $$H \triangleq p_o f_o(X, U) + \mathrm{Tr}[F\ P'] \tag{2.12}$$

(2) $H(X^*(t), P^*(t), U, p_o^*)$ has an absolute minimum as a function of U over Ω at $U = U^*(t)$, $t_o \leq t \leq t_1$, i.e.,

$$H(X^*(t), P^*(t), U^*(t), p_o) = \min_{U \in \Omega} H(X^*(t), P^*(t), U, p_o^*) \tag{2.13}$$

(3) $$H(X^*(t), P^*(t), U^*(t), p_o^*) = 0 \quad t_o \leq t \leq t_1 \tag{2.14}$$

<u>Proof</u>: Consider a mapping $\sigma : M_{nm} \to R_{nm}$ defined by

$$\sigma(X) = \begin{bmatrix} x_{11} \\ x_{12} \\ \vdots \\ x_{1m} \\ x_{21} \\ \vdots \\ x_{nm} \end{bmatrix} \triangleq y \tag{2.15}$$

Clearly the mapping σ is one-to-one and onto the whole space R_{nm}. We shall consider $U(t)$ as a point in $\Omega \subset R_{rs}$ where $u_{ij}(t)$ is the ijth component. With these transformation, we have converted the original problem into an equivalent new problem: given

$$\frac{dy_i(t)}{dt} = g_i(y_1, \ldots, y_{nm}, u_{11}, \ldots u_{rs}) \quad i=1, \ldots, nm \tag{2.16}$$

where $g_1 \stackrel{\Delta}{=} f_{11},\ g_2 \stackrel{\Delta}{=} f_{12}, \ldots, g_{m+1} \stackrel{\Delta}{=} f_{21}, \ldots, g_{nm} \stackrel{\Delta}{=} f_{nm}$ (2.17)

$g_i(y(t), U(t))$ and $\frac{\partial g_i}{\partial y_j}$ are given and continuous

$$J = \int_{t_o}^{t_1} g_o(y(t), U(t))\, dt \tag{2.18}$$

where $$g_o(y(t), U(t)) \stackrel{\Delta}{=} f_o(X(t), U(t)) \tag{2.19}$$

We are to find an admissible control $U(t) \subset \Omega$, which will transfer $y(t)$ from $y(t_o) = y_o = \sigma(X_o)$ to $y_1 = \sigma(X_1)$ such that the functional (2.18) will be minimized. For this equivalent problem, Pontryagin has proved the minimum (or maximum) principle which gives the necessary condition for optimality: in order that $U^*(t)$ for the above special case to be optimal, it is necessary that there exists a scalar $p_o^* \geq 0$ and a costate vector $q^*(t)$ such that (see Ref. 1)

(1) y
$$\left.\begin{aligned} \dot{y}^*(t) &= \left.\frac{\partial H_o}{\partial q(t)}\right|_* \\ \dot{q}^*(t) &= -\left.\frac{\partial H_o}{\partial y(t)}\right|_* \end{aligned}\right\} \tag{2.20}$$

with boundary conditions

$$y^*(t_o) = y_o \ ; \ y^*(t_1) = y_1 \tag{2.21}$$

where $$H_o \overset{\Delta}{=} p_o g_o(y(t),\ U(t)) + \langle g, q\rangle_{R_{nm}} \tag{2.22}$$

(2) $$H_o(y^*(t), q^*(t), U^*(t), p_o) = \min_{U\epsilon\Omega} H_o(y^*(t), q^*(t), U, p_o^* \tag{2.23}$$

(3) $$H_o(y^*(t), q^*(t), U^*(t), p_o) = 0 \tag{2.24}$$

Since σ is one to one onto, σ^{-1} exists and it is a mapping from R_{nm} to M_{nm} where $\sigma^{-1} \circ \sigma(X) = X$. Let us define

$$P(t) = \sigma^{-1}(q(t)) \tag{2.25}$$

$$H(X(t), P(t), U(t), p_o) \overset{\Delta}{=} H_o(\sigma(X(t)), \sigma(P(t)),\ U(t),\ p_o)$$

$$= H_o(y(t),\ q(t),\ U(t),\ p_o) \tag{2.26}$$

Let $U^*(t)$ be the optimal control for the original problem; then it is also optimal for the equivalent problem; since $p_o^* \geq 0$ and $q^*(t)$ exist, this implies that $p_o^* \geq 0$ and $P^*(t)$ exist. The $q^*(t)$ satisfies (2.20), (2.21) and (2.22); we shall show that this implies $P^*(t)$ satisfies (2.10), (2.11) and (2.12). We rewrite (2.20) to obtain

$$\dot{y}_1^*(t) = \dot{x}_{11}(t) = \frac{\partial H_o}{\partial q_1}\bigg|_* = \frac{\partial H}{\partial p_{11}}\bigg|_*$$

$$\vdots$$

$$\dot{y}_m^*(t) = \dot{x}_{1m}(t) = \frac{\partial H_o}{\partial q_m}\bigg|_* = \frac{\partial H}{\partial p_{1m}}\bigg|_*$$

$$\vdots$$

$$\dot{y}_{nm}^*(t) = \dot{x}_{nm}(t) = \frac{\partial H_o}{\partial q_{nm}}\bigg|_* \tag{2.27}$$

We have a similar set of equations relating $p^*_{ij}(t)$ with $- \left.\frac{\partial H}{\partial x_{ij}}\right|_*$,
and thus we see that (2.20) yields the set of the matrix equations:

$$\left.\begin{aligned} \dot{X}^*(t) &= \left.\frac{\partial H}{\partial P(t)}\right|_* \\ \dot{P}^*(t) &= - \left.\frac{\partial H}{\partial X(t)}\right|_* \end{aligned}\right\} \tag{2.28}$$

Equation 2.21 implies that

$$X^*(t_o) = \sigma^{-1}(y(t_o)) = \sigma^{-1}(y_o) = X_o \; ; \; X^*(t_1) = X_1 \tag{2.29}$$

Equations 2.22 and 2.26 imply that

$$\begin{aligned} H(X(t),\ P(t),\ U(t),\ p_o) &= p_o g_o(y, U) + \sum_{i=1}^{nm} g_i\, q_i \\ &= p_o f_o(X, U) + \sum_{i, j}^{n, m} f_{ij}\, p_{ij} \\ &= p_o f_o(X, U) + \mathrm{Tr}\,[FP'] \end{aligned} \tag{2.30}$$

Also by (2.25), (2.26), we see that (2.23), (2.24) can be written as (2.13) and (2.14) respectively.

Q.E.D.

For the case where the target set $S = S_1 \times (T_1, T_2)$, $S_1 \subset M_{nm}$, we can establish the transversality conditions: the matrix $P^*(t)$ is transversal (normal) to S_1 at $X^*(t_1)$, i.e.,

$$< \sigma(P^*(t),\ \sigma(X^*(t_1)) > \ = \ \mathrm{Tr}\,[X^*(t_1)\ P^{*'}(t_1)] = 0$$

We shall now state without proof† the results of two control problems which we shall use in Chapters 3 and 4.

Control Problem 1:

Given $$\dot{X}(t) = F(X,U) \tag{2.31}$$

and the cost functional

$$J = K(X(t_1)) + \int_{t_o}^{t_1} f_o(X,U)\,dt \tag{2.32}$$

where $K(X(t_1)$ is a scalar function. We are to choose $U(t)\epsilon\Omega$ which transfers $X(t)$ from $X(t_o) = X_o$ to $X(t_1)$ such that (2.31) is satisfied and the cost functional (2.32) is minimized.

Theorem 2.2: In order that $U^*(t)$ be optimal for the control problem 1, it is necessary that there exist a nonnegative constant p_o^* and a costate matrix function $P^*(t)$, such that (2.10), (2.11), (2.12), (2.13) and (2.14) are satisfied. In addition we must have

$$P^*(t_1) = \left.\frac{\partial K(X(t_1))}{\partial X(t_1)}\right|_{X^*(t_1)} \tag{2.33}$$

Control Problem 2:

Given $$\dot{X}(t) = F(X,U) \tag{2.34}$$

$$J = K(X(t_o)) + \int_{t_o}^{t_1} f_o(X\,U)\,dt \tag{2.35}$$

†For a proof in the vector case, see Ref. 1.

We are to choose $U(t) \epsilon \Omega$ which transfer $X(t)$ from $X(t_o)$ to $X(t_1)$ such that (2.34) is satisfied and the cost functional (2.35) is minimized.

Theorem 2.3: In order that $U^*(t)$ be optimal for the control problem 2, it is necessary that there exist a nonnegative constant p_o^* and a costate matrix function $P^*(t)$, such that (2.10), (2.11), (2.12), (2.13) and (2.14) are satisfied, in addition we must have

$$P^*(t_o) = - \left. \frac{\partial K(X(t_o))}{\partial X(t_o)} \right|_{X^*(t_o)} \tag{2.36}$$

C. A SUFFICIENT CONDITION FOR OPTIMALITY

We shall now establish a sufficient condition for local optimality under a stronger assumption on the cost functional.

Consider the control problem:

$$\dot{X}(t) = F(X, U, t) \tag{2.37}$$

with cost functional

$$J(X_o, U, t_o) = \int_{t_o}^{t_1} f_o(X(t), U(t), t)\, dt \tag{2.38}$$

We wish to choose among all admissible controls with range in Ω the one which transfers $X(t_o) = X_o$ to $X(t_1) = X_1 \epsilon S_1$ such that the functional (2.38) is minimized. It is shown in (2.38) explicitly that the functional depends on the initial states X_o, the initial time t_o and the particular control chosen.

Let $\hat{U}(t)$, $t_o \leq t \leq t_1$, be a particular control which transfers (X_o, t_o) to $S = S_1 \times (T_1, T_2)$; we shall denote the corresponding trajectory by $\hat{X}(t)$. We shall write

$$J[\hat{X}(t),\ t,\ \hat{U}] = \int_t^{t_1} f_o(\hat{X}(\tau),\ \hat{U}(\tau),\ \tau)\, d\tau \tag{2.39}$$

Since for every control $\hat{U}(t)$, $t_o \leq t \leq t_1$, there corresponds an unique trajectory $\hat{X}(t)$, $t_o \leq t \leq t_1$, we can write the functional (2.39) as a function of the states and time only, i.e., we can uniquely define

$$\hat{J}[\hat{X}(t),\ t] \triangleq J[\hat{X}(t),\ t, \hat{U}]\ ;\ t_o \leq t \leq t_1 \tag{2.40}$$

We shall assume that $\hat{J}[\hat{X}(t), t]$ is a continuously differentiable function with respect to its arguments on a region $A \subset M_{nm} \times (T_1, T_2)$ with $(\hat{X}(t), t) \epsilon A$; $t_o \leq t \leq t_1$.

Before proving the sufficient condition for optimality, we need some definitions to simplify the statements of the theorem.

Definition 2.2: Let $H(X, P, U\ p_o, t)$ be the Hamiltonian

$$H(X, P, U, p_o, t) = p_o f_o(X, U, t) + Tr[F\ P'] \tag{2.41}$$

If for each point (X, t) in A, the function $H(X, P, W, p_o, t)$ has a unique absolute minimum with respect to all $W \epsilon \Omega$ at $W = \tilde{U}(X, P, t)$, then we say that H is normal relative to A or that our control problem is normal relative to A, and $\tilde{U}(X, P, t)$ is the H-minimal control relative to A.

Definition 2.3: If H is normal relative to A and if $\tilde{U}(X, P, t)$ is the H-minimal control relative to A, we shall call the partial differential equation

$$\frac{\partial J(X, t)}{\partial t} + H[X, \frac{\partial T}{\partial X}(X, t),\ \tilde{U}(X\ \frac{\partial J}{\partial X}(X, t),\ t] = 0 \tag{2.42}$$

with boundary condition

$$J(X, t) = 0 \qquad \text{for} \quad (X, t) \epsilon S \tag{2.43}$$

the Hamilton-Jacobi (H-J) equation relative to A.

We shall state and prove the sufficient condition for local optimality. The technique of the proof is similar to that used in proving Theorem 2.1; for this reason, we shall outline the argument instead of giving the rigorous formal steps.

Theorem 2.4: Suppose that $A = A_1 \times (T_1, T_2)$; $A_1 \subset M_{nm}$, and that H is normal relative to A such that $\widetilde{U}(X, P, t)$ is the H-minimal control relative to A. Let $\hat{U}(t)$ be an admissible control such that:

(1) $\hat{U}$ transfers (X_o, t_o) to S

(2) If $\hat{X}(t)$ is the trajectory corresponding to $\hat{U}(t)$, then

$$(\hat{X}(t),\ t) \in A \qquad t_o \leq t \leq t_1 \tag{2.44}$$

(3) There is a solution $\hat{J}(X, t)$ of the H-J equation (2.42) satisfying the boundary condition (2.43), such that

$$\hat{U}(t) = \widetilde{U}[\hat{X}(t),\ \frac{\partial \hat{J}}{\partial X}(\hat{X}(t),\ t),\ t] \tag{2.45}$$

Then $\hat{U}$ is an optimal control relative to the set of controls U that generate trajectories lying entirely in A, and

$$J[\hat{X}(t), t, \hat{U}] = \hat{J}[\hat{X}(t),\ t] \qquad t_o \leq t \leq t_1 \tag{2.46}$$

Proof: Let $A_o = \sigma(A_1) \times (T_1, T_2)$, where σ is the same mapping defined by (2.15). Let H_o be defined as in (2.26). If H is normal relative to A, then H_o is normal relative to A_o. Through the mapping σ, we can rephrase our assumptions: we have that H_o is normal relative to A_o and that $\widetilde{U}(y, q, t)$ is the H-minimal control relative to A_o. $\hat{U}(y, q, t)$ is an admissible control such that

(1) $\hat{U}$ transfer (y_o, t) to $S_o = \sigma(S)$

(2) If $\hat{y}(t)$ is the trajectory corresponding to $\hat{U}(t)$, then

$$(\hat{y}(t),\ t) \in A_o \qquad t_o \le t \le t_1 \tag{2.47}$$

(3) There is a solution $\hat{J}(\underline{y}, t)$ of the H-J equation (2.42) satisfying the boundary condition (2.43) such that

$$\hat{U}(t) = \tilde{U}[\hat{y}(t),\ \frac{\partial \hat{J}}{\partial y}(\hat{y}(t),\ t),\ t] \tag{2.48}$$

$$\text{where} \quad y = \sigma(X)\ ;\quad q = \sigma(P)\ ;\quad y_o = \sigma(X_o) \tag{2.49}$$

Then we know (see Ref. 1) that $\hat{U}$ is an optimal control relative to the set of controls U that generate trajectories lying in A_o and

$$J[\hat{y}(t),\ t,\ \hat{U}] = \hat{J}[\hat{y}(t), t] \quad t_o \le t \le t_1 \tag{2.50}$$

This implies that $\hat{U}$ is an optimal control relative to the set of controls U that generate trajectories lying in A and

$$J[\hat{X}(t),\ t, \hat{U}] = \hat{J}[\hat{X}(t),\ t] \quad t_o \le t \le t_1 \tag{2.51}$$

Q.E.D.

If the region A is M_{nm} (T_1, T_2), then the theorem will give a global sufficient condition for optimality.

D. DISCUSSION OF RESULTS

1. The Matrix Minimum Principle:

From the nature of the proof, we see that its validity hinges on the construction of an one-to-one onto mapping σ from the space M_{nm} to R_{nm}. If we examine the mapping σ carefully, we see that it is a linear mapping. If we define an inner product in the space M_{nm} by[6]

$$\langle X, Y \rangle_{M_{nm}} \triangleq \mathrm{Tr}[XY'] \tag{2.52}$$

then we see that the mapping σ also preserves the inner product because we have

$$\langle \sigma(X),\ \sigma(Y) \rangle_{R_{nm}} = \sum_{i,j} x_{ij} y_{ij} = \mathrm{Tr}[XY']$$

$$= \langle X, Y \rangle_{M_{nm}} \tag{2.53}$$

This implies that the mapping σ preserves the metric distance between elements in M_{nm}. Thus the spaces M_{nm} and R_{nm} are algebraically and topologically equivalent. By comparing (2.8) and (2.16) we see that the differential structure of M_{nm} can be viewed through that of R_{nm}. Abstractly, the two spaces, M_{nm} and R_{nm}, have the same differential, algebraical and topological structures; thus the results that hold in one space must also hold in the other space and vice-versa. Notice that the statements of the matrix minimum principle are just expressing the Pontryagin's principle in terms of the elements and operations on the space M_{nm}.

The proof of the matrix minimum principle relies heavily on the fact that the mapping σ is nonsingular, as this guarantees the existence of $P^*(t)$ from (2.25). Let us consider the case where we required that the state matrix X is a symmetric matrix. Then, if we consider the range space of σ (acting on the set of all nxn symmetric matrices), then the range space will be a proper subspace of R_{nn}. Therefore, the existence of $q^*(t)$ in R_{nn} does not guarantee the existence of $\sigma^{-1}(q^*(t))$ in the original space for now the inverse mapping is not defined on the entire space R_{nn}, and $q^*(t)$ may not belong to the range space. Similarly the proof breaks down if we assume any functional relations between the entries of the matrix X. This means that we must view the space M_{nm} as a nm-dimensional vector space. This point is very important, since it tells us when the matrix minimum principle is applicable.

The matrix minimum principle gives a set of necessary conditions that an optimal control must satisfy. Usually we first find all the controls that will satisfy the necessary conditions; these controls are called extremal controls. Among these we are to find the one that is optimal. The general questions as to the existence and uniqueness of the optimal controls is still an open area of research.

2. The Sufficient Condition for Optimality

Theorem 2.4 is a local sufficiency condition. In the proof we specified the region A a priori; yet, in practice, we are to find a suitable region A in which the theorem may be applied.

Note that the H-J equation is a partial differential equation with given boundary conditions; in general, there are difficulties in solving this partial differential equation. For this reason, we shall mostly use the H-J equation as a check on the optimality of the extremal solutions obtained by the use of the matrix minimum principle.

We note that the constant p_o^* which appears in Theorem 2.1 is constrained to be nonnegative. In most cases which are of interest to us, p_o^* will be positive; (Ref. 1) for this reason, in the subsequent chapters we shall set $p_o^* = 1$.

E. FUTURE RESEARCH

From this approach, we see clearly that the validity of the minimum principle relies on the algebraical, topological and differential structures of the state space we are interested in. It seems (to the author) that the results may be able to extend to the case where the state space we are considering is an infinite (but countable) dimensional Hibert Space, on which a differential structure is defined.

CHAPTER III

NOISY STATE-REGULATOR PROBLEMS

In this chapter and the next, we shall apply the matrix minimum principle to two general types of problems of engineering importance. Suppose that we wish to regulate (i.e., bring near zero) the states of a linear dynamical system. We shall investigate the two cases:

1. systems with driving disturbance only
2. systems with driving disturbance and measurement noise

These problems had been studied using the Dynamic Programming approach.[13,30] Specific results for linear system with quadratic criteria had been obtained.[13,30] In this chapter, we shall approach the same problem from a different viewpoint. The results obtained in this work agree with the existing results; it will be seen that the new approach sheds more light on the structure of the system, and that it will be useful in the design of a suboptimal control system subject to some constraints imposed by the designer.

We shall now outline the general idea of our approach to these problems. By a reasonable argument, we make an assumption about the form of the optimal control with some undetermined parameters. The optimal control is completely specified if we set the parameters at their optimal values.

If the disturbances are modelled as white Gaussian noise, (see Appendix 1) then the states of the system turn out to be Gaussian; thus, one can completely describe the vector processes by their second and first moments. From the given system and the known statistics of the disturbance, we can write the dynamics of the second and first moments of the vector processes. The moments of the vector processes are regarded as the states of a new dynamical system. In so doing we have converted a stochastic problem into a deterministic problem so that the matrix minimum principle derived in Chapter II can be applied. We shall use the Hamilton-Jacobi equation to check the solution obtained by the application of the minimum principle. The minimum value of the cost functional shall be derived so as to analytically describe the performance of the system.

A. STATE-REGULATOR PROBLEM WITH DRIVING DISTURBANCE OF ZERO MEAN

Consider the plant

$$\dot{x}(t) = A(t)\, x(t) + B(t)\, u(t) + \xi(t) \tag{3.1}$$

where
$x(t)$ is an n-vector
$A(t)$ is an $n \times n$ matrix
$u(t)$ is an r-vector
$B(t)$ is an $n \times r$ matrix
$\xi(t)$ is an n-vector white Gaussian Process

We assume that we know the statistics of $\xi(t)$. For simplicity we shall assume that

$$E[\xi(t)] = 0 \; ; \quad E[\xi(t)\xi'(\tau)] = D(t)\delta(t-\tau) \tag{3.2}$$

and that

$$E[x(t_o)\xi'(t)] = 0 \qquad t_o \leq t \leq t_1 \tag{3.3}$$

We want to find $u(t)$, $t_o \leq t \leq t_1$, which will regulate the states to zero when $x(t_o) = x_o \neq 0$, with some constraints on the energy of the control function used. We choose a cost functional which will indicate the performance of the system, and also gives a simple form for the optimal control. A reasonable choice of the cost functional is[1]

$$J = E[\langle x(t_1),\, Fx(t_1)\rangle + \int_{t_o}^{t_1} \langle x(t),\, Q(t)\, x(t)\rangle + \langle u(t),\, R(t)\, u(t)\rangle dt] \tag{3.4}$$

where F and $Q(t)$ are $n \times n$ positive semidefinite matrices, and $R(t)$ is an $r \times r$ positive definite matrix.

1. Reformulation of the Problem:

We shall now fix the structural form of the optimal control. We assume here that there is no observation noise; thus, we assume that we know the exact states of the system at any instant of time. The states at any instant of time are what we need for describing completely the future behavior of the system with different control functions applied. The optimal control must be a function of maximal information. Thus we see that the optimal control at time t must be a function of the states at time t and of the time t, i.e.,

$$u(t) = g(x(t), t) \tag{3.5}$$

When the variance of the driving disturbance $\xi(t)$ approaches zero, this means that $\xi(t)$ becomes more and more deterministic and approaches its mean which is equal to zero. For this limiting case, we know[1] that the optimal control is generated via a linear feedback law, i.e.,

$$u(t) = M(t)\, x(t) \tag{3.6}$$

Since the effect of the disturbance $\xi(t')$, $t_o \leq t' \leq t$, is included in the state at time t, $x(t)$, it is reasonable to argue that the presence of disturbing noise should not change the form of the optimal control. For this reason, we <u>fix</u> the structural form of the optimal control to be a linear time-varying feedback law. In other words, we <u>assume</u> that the control is generated from the state by

$$u(t) = M(t)\, x(t) \tag{3.7}$$

where $M(t)$ is an $r \times n$ matrix with, as yet, undetermined entries.

Substituting (3.7) into (3.1), we have

$$\dot{x}(t) = [A(t) + B(t)\ M(t)]\ x(t) + \xi(t) \tag{3.8}$$

From (3.8) and the statistics of the noise (3.2), (3.3), we can write the dynamics of the second moment of $x(t)$. (see Appendix 2)

$$\dot{\Sigma}_x(t) = [A(t) + B(t)\,M(t)]\,\Sigma_x(t) + \Sigma_x(t)[A'(t) + M'(t)B'(t)] + D(t) \tag{3.9}$$

where

$$\Sigma_x(t) \triangleq E[x(t)x'(t)] \tag{3.10}$$

The cost functional (3.4) can now be written as

$$J = Tr[F\Sigma_x(t_1)] + \int_{t_o}^{t_1} Tr[Q(t)\Sigma_x(t) + M'(t)\,R(t)\,M(t)\,\Sigma_x(t)]\,dt \tag{3.11}$$

The new optimization problem we have to solve is: <u>Given the matrix differential equation</u> (3.9) <u>and the cost functional</u> (3.11) <u>choose M(t) (which is unconstrained) so that the functional</u> (3.11) <u>is minimized subject to the matrix differential equation</u> (3.9).

2. Application of the Matrix Minimum Principle

We have changed a physical problem into a mathematical problem for which the results obtained in Chapter II can be applied directly. We shall now apply the matrix minimum principle to derive an extremal matrix $M(t)$ which is now viewed as the control function for the new system (3.9). Toward this end, we form the Hamiltonian, in which $P_x(t)$ is the costate matrix associated with $\Sigma_x(t)$,

$$H = Tr[(A(t) + B(t)M(t))\Sigma_x P_x'(t) + \Sigma_x(t)(A'(t) + M'(t)B'(t))P_x'(t)$$

$$+ D(t)P_x'(t) + Q(t)\Sigma_x(t) + M'(t)R(t)M(t)\Sigma_x(t)] \tag{3.12}$$

The necessary condition that $\Sigma_x^*(t)$, $P_x^*(t)$ satisfy the canonical equations yields:

$$\dot{\Sigma}_x^*(t) = [A(t) + B(t)M^*(t)]\,\Sigma_x^*(t) + \Sigma_x^*(t)[A'(t) + M^{*\prime}(t)B'(t)] + D(t) \qquad (3.13)$$

$$-\dot{P}_x^*(t) = [A'(t) + M^*(t)B'(t)]\,P_x^*(t) + P_x^*(t)[A(t) + B(t)M^*(t)]$$

$$+Q(t) + M^{*\prime}(t)\; R(t)\; M^*(t) \qquad (3.14)$$

The transversity condition yields†

$$P_x^*(t_1) = \frac{\partial \mathrm{Tr}[F\Sigma_x(t_1)]}{\partial \Sigma_x(t_1)} = F \qquad (3.15)$$

A necessary condition for $M^*(t)$ to minimize the Hamiltonian is given by

$$\left.\frac{\partial H(\Sigma_x^*(t),\; P_x^*(t),\; M(t))}{\partial M(t)}\right|_{M(t)=M^*(t)} = 0 \qquad (3.16)$$

Differentiating (3.12) with respect to $M(t)$ and setting it to zero,

we have
$$M^*(t) = -R^{-1}(t)\; B'(t)\; P_x^*(t) \qquad (3.17)$$

By the assumption that $R(t)$ is positive definite, then indeed Eq. 3.16 gives us the $M^*(t)$ which minimizes the Hamiltonian. Substituting (3.17) into (3.14), we have the complete solution which satisfies the necessary conditions for optimality:

$$-\dot{P}_x^*(t) = A'(t)\,P_x^*(t) + P_x^*(t)\,A(t) - P_x^*(t)B(t)R^{-1}(t)B'(t)P_x^*(t) + Q(t) \qquad (3.18)$$

with boundary condition

$$P_x^*(t_1) = F \qquad (3.19)$$

---

†See Appendix 6 for formulae of Gradient Matrices.

The feedback gain is given by

$$M^*(t) = -R^{-1}(t)B'(t)\,P_x^*(t) \tag{3.20}$$

where $P_x^*(t)$ satisfies (3.18) and (3.19). We note that Eq. 3.18 is the familiar matrix Riccati differential equation.

3. <u>The Minimum Cost Functional</u>

We shall now show that

$$E[J(x(t),t)] \stackrel{\Delta}{=} J^*[\Sigma_x(t),t] = \mathrm{Tr}[P_x(t)\Sigma_x(t)] + f(t) \tag{3.21}$$

where
$$\begin{aligned} \dot{f}(t) &= -\mathrm{Tr}[D(t)P_x(t)] \\ f(t_1) &= 0 \end{aligned} \tag{3.22}$$

will satisfy the H-J equation with $M(t) = -R^{-1}(t)B(t)P_x(t)$ and $P_x(t)$ satisfies (3.18) and (3.19). Thus the solution obtained by the application of the matrix minimum principle satisfies the sufficient conditions and, therefore, is optimal. Clearly, the boundary condition is satisfied, because

$$J^*[\Sigma_x(t_1),t_1] = \mathrm{Tr}[P_x(t_1)\,\Sigma_x(t_1)] = \mathrm{Tr}[F\Sigma_x(t_1)] \tag{3.23}$$

Also we have

$$\frac{\partial J^*[\Sigma_x(t),t]}{\partial t} = \mathrm{Tr}[\dot{P}_x(t)\Sigma_x(t)] + \frac{df}{dt} \tag{3.24}$$

$$\frac{\partial J^*[\Sigma_x(t),t]}{\partial \Sigma_x(t)} = P_x(t) \tag{3.25}$$

By (3.24), (3.25), (3.22) and the relation $M(t) = -R^{-1}(t)B'(t)P_x(t)$ where $P_x(t)$ satisfies (3.18) and (3.19), we see that direct substitution yields

$$\frac{\partial J^*}{\partial t} + H[\Sigma_x(t), \frac{\partial J^*}{\partial \Sigma_x}, M(t)] = 0 \tag{3.26}$$

This shows that the cost functional defined by (3.21) and (3.22) is the minimum cost functional which describes the performance of the optimal system and that the optional control law is described by (3.20), (3.18) and (3.19).

B. STATE-REGULATOR PROBLEMS WITH DRIVING DISTURBANCE OF NONZERO MEAN

We now consider the plant (3.1) with a driving disturbance which has a nonzero mean, i.e., we assume

$$E[\xi(t)] = m(t) \tag{3.27}$$

Let us define

$$\xi_o(t) = \xi(t) - m(t) \tag{3.28}$$

We assume that

$$E[(\xi(t)-m(t))(\xi(t)-m(t))'] = E[\xi_o(t)\xi_o'(t)] = D(t)\delta(t-\tau) \tag{3.29}$$

Thus we can write (3.1) as

$$\dot{x}(t) = A(t)x(t) + B(t)u(t) + m(t) + \xi_o(t) \tag{3.30}$$

where $m(t)$ is now a deterministic quantity, and the statistics of $\xi_o(t)$ are

$$E[\xi_o(t)] = 0 \quad ; \quad E[\xi_o(t)\xi_o'(t)] = D(t)\delta(t-\tau) \tag{3.31}$$

The problem is to choose $u(t)$ such that the functional

$$J = E[\langle x(t_1), Fx(t_1)\rangle + \int_{t_o}^{t_1} \langle x(t), Q(t)x(t)\rangle + \langle u(t), R(t)u(t)\rangle dt] \tag{3.32}$$

is minimized subject to (3.30).

We see that mean of the disturbing noise effects the states of the system in a deterministic way; we would expect that the form of the optimal control must contain a part which is independent of the observed states, but only dependent on the structure of the system and the statistics of the disturbance.

1. Reformulation of the Problem

Let us decompose the state into

$$x(t) = x_r(t) + x_d(t) \tag{3.33}$$

and the control $u(t)$ into

$$u(t) = u_r(t) + u_d(t) \tag{3.34}$$

where $x_r(t)$, $x_d(t)$, $u_r(t)$, $u_d(t)$ satisfy

$$\dot{x}_r(t) = A(t)x_r(t) + B(t)u_r(t) + \xi_o(t) \tag{3.35}$$

$$\dot{x}_d(t) = A(t)x_d(t) + B(t)u_d(t) + m(t) \tag{3.36}$$

By comparing (3.35) with (3.1), (3.31) with (3.2), (3.3), we can use the same argument to fix the form of the optimal control $u_r(t)$ as a linear feedback; Eq. 3.36 gives the form of the optimal control $u_d(t)$ which is a linear feedback plus a function (predetermined) of time. Since the structure of systems (3.35) and (3.36) are the same, the feedback gain for each case is the same.† Thus we can write

$$u(t) = K(t)x_r(t) + K(t)x_d(t) + k(t) = K(t)x(t) + k(t) \tag{3.37}$$

† See discussion in Section 3.4.

The term $k(t)$ is necessary to cancel off the effect due to the mean of the disturbance noise. Substituting (3.37) into (3.30), we have

$$\dot{x}(t) = (A(t) + B(t)\ K(t))\ x(t) + B(t)\ k(t) + m(t) + \xi_o(t) \tag{3.38}$$

By the results of Appendix 2, (3.38) gives the dynamics of the first and second moment of $x(t)$, they are

$$\dot{m}_x(t) = (A(t) + B(t)\ K(t))m_x(t) + B(t)\ k(t) + m(t) \tag{3.39}$$

$$\dot{\Sigma}_x(t) = (A(t) + B(t)\ K(t))\Sigma_x(t) + \Sigma_x(t)(A'(t) + K'(t)B'(t)$$
$$+ m_x(t)(k'(t)B'(t) + m'(t)) + (B(t)k(t) + m(t))m_x'(t) + D(t) \tag{3.40}$$

where

$$m_x(t) \triangleq E[x(t)] \ ; \ \Sigma_x(t) \triangleq E[x(t)x'(t)] \tag{3.41}$$

The functional (3.32) can be written as

$$J = Tr[F\Sigma_x(t_1) + \int_{t_o}^{t_1} Q(t)\Sigma_x(t) + K'(t)R(t)K(t)\Sigma_x(t) + R(t)k(t)k'(t)$$
$$+ R(t)K(t)m_x(t)k'(t) + R(t)k(t)m_x'(t)K'(t)]\ dt \tag{3.42}$$

The new optimization problem we have to solve is: given the set of Eqs. 3.39, 3.40 and the cost functional (3.42) choose $K(t)$, $k(t)$ (which are unconstrained) so that the functional (3.42) is minimized subject to the differential equations (3.39) and (3.40).

2. Application of the Matrix Minimum Principle

We have formulated a physical problem in the mathematical framework in which the matrix minimum principle can be applied.

Let $P_x(t)$, $p_x(t)$ be the costate variables of $\Sigma_x(t)$ and $m_x(t)$, respectively. We form the Hamiltonian,

$$H = \mathrm{Tr}[\dot{\Sigma}_x(t)P_x'(t) + \dot{m}_x p_x'(t) + Q(t)\Sigma_x(t) + K'(t)R(t)K(t)\Sigma_x(t) + R(t)k(t)k'(t)$$

$$+ R(t)K(t)m_x(t)k'(t) + R(t)k(t)m_x'(t)K'(t)] \tag{3.43}$$

The canonical equations for the costates are:

$$-\dot{P}_x^*(t) = (A'(t) + K^{*'}(t)B'(t))P_x^*(t) + P_x^*(t)(A(t) + B(t)K^*(t)) + Q(t)$$

$$+ K^{*'}(t)R(t)K^*(t) \tag{3.44}$$

$$-\dot{p}_x^*(t) = 2P_x^*(t)B(t)k^*(t) + 2P_x^*(t)m(t) + (A'(t) + K^{*'}(t)B'(t))p_x^*(t)$$

$$+2K^{*'}(t)R(t)k^*(t) \tag{3.45}$$

The boundary conditions are

$$P_x^*(t_1) = F \; ; \; p_x^*(t_1) = 0 \tag{3.46}$$

From (3.44) and (3.46), we deduce that $P_x^*(t)$ is symmetric. To find the necessary condition for the minimization of $H(\Sigma_x^*(t), m_x^*(t), P_x^*(t), p_x^*(t), k(t), K(t))$ with respect to $k(t)$ and $K(t)$, we set

$$\left.\frac{\partial H}{\partial k(t)}\right|_{k(t) = k^*(t)} = 0 \tag{3.47}$$

$$\left.\frac{\partial H}{\partial K(t)}\right|_{K(t)=K^*(t)} = 0 \tag{3.48}$$

Using the fact that $P_x^*(t)$ is symmetric, Eqs. 3.47 and 3.48 give

$$2B'(t)P_x^*(t)m_x^*(t) + B'(t)p_x^*(t) + 2R(t)k^*(t) + 2R(t)K^*(t)m_x^*(t) = 0 \tag{3.49}$$

$$2B'(t)P_x^*(t)\Sigma_x^*(t) + 2R(t)K^*(t)\Sigma_x^*(t) + B'(t)p_x^*(t)m_x^*(t)$$

$$+ 2R(t)k^*(t)m_x^{*'}(t) = 0 \qquad (3.50)$$

Equations 3.50 and 3.49 give

$$k^*(t) = -\frac{1}{2}R^{-1}B'(t)p_x^*(t) \qquad (3.51)$$

$$K^*(t) = -R^{-1}(t)B'(t)P_x^*(t) \qquad (3.52)$$

Substituting (3.51), (3.52) into (3.44) and (3.45), we have specified $P_x^*(t)$, $p_x^*(t)$ uniquely

$$-\dot{P}_x^*(t) = A'(t)P_x^*(t) + P_x^*(t)A(t) - P_x^*(t)B(t)R^{-1}(t)B'(t)P_x^*(t) + Q(t) \qquad (3.53)$$

$$-\dot{p}_x^*(t) = 2P_x^*(t)m(t) + A'(t)p_x^*(t) - P_x^*(t)B(t)R^{-1}(t)B'(t)p_x^*(t) \qquad (3.54)$$

with boundary conditions

$$P_x^*(t_1) = F \ ; \quad p_x(t_1) = 0 \qquad (3.55)$$

Together with Eqs. 3.37, 3.51 and 3.52, we have specified the extremal control uniquely. (See Fig. 1.)

3. Minimum Cost Functional

We shall show that the extremal control obtained by the application of the matrix minimum principle is the optimal control.

Consider the functional,

$$J^*[\Sigma_x(t), m_x(t), t] = \mathrm{Tr}[P_x(t)\Sigma_x(t)] + \mathrm{Tr}[p_x(t)m_x'(t)] + f(t) \qquad (3.56)$$

where

$$\left.\begin{aligned} -\dot{f}(t) &= -\frac{1}{4}p_x'(t)B'(t)R^{-1}(t)B(t)p_x(t) + \mathrm{Tr}[D(t)P_x(t)] + m'(t)p_x(t) \\ f(t_1) &= 0 \end{aligned}\right\} \qquad (3.57)$$

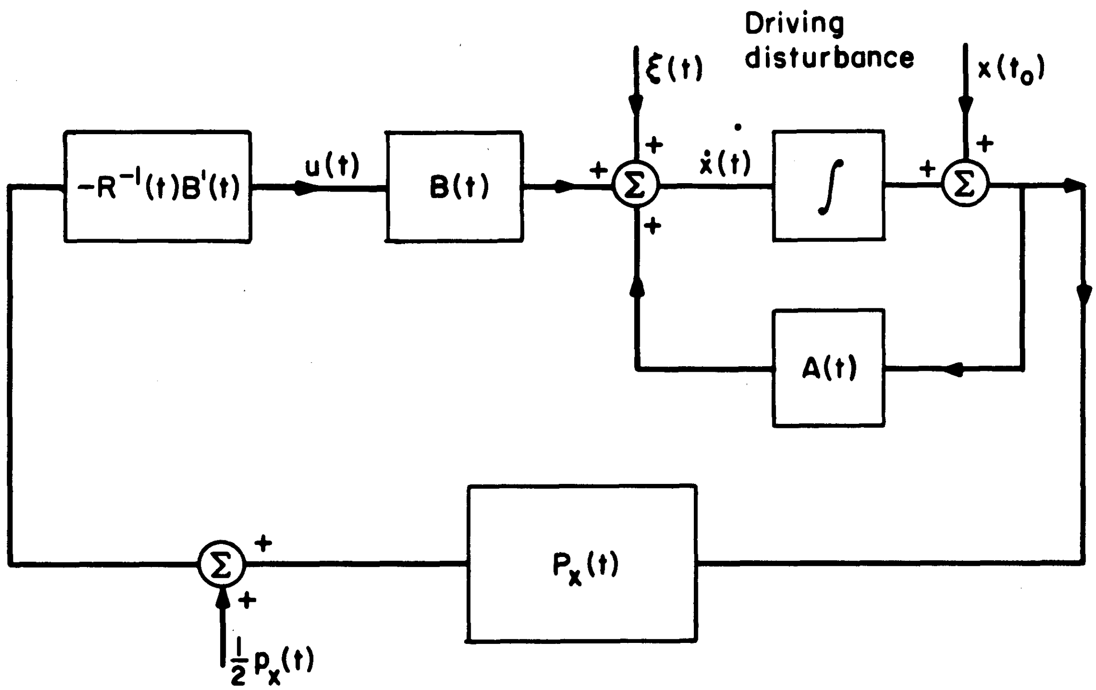

Fig. 1 State Regulator Problem with Driving Disturbance

$$-\dot{P}_x(t) = A'(t)P_x(t) + P_x(t)A(t) - P_x(t)B(t)R^{-1}(t)B'(t)P_x(t) + Q(t)$$

$$-\dot{p}_x(t) = 2P_x(t)m(t) + A'(t)p_x(t) - P_x(t)B(t)R^{-1}(t)B'(t)p_x(t)$$

$$P_x(t_1) = F \quad ; \quad p_x(t_1) = 0$$

It can be shown† that (3.56) with (3.57) will satisfy the H-J equation when $K(t)$, $k(t)$, $P_x(t)$, $p_x(t)$ satisfy (3.52), (3.51) (3.53), (3.54) and (3.55). This implies that the extremal solution satisfies the necessary and sufficient conditions for optimality and thus the control specified by (3.37), (3.51), 3.52), (3.53), (3.54) and (3.55) is the optimal control.

C. NOISY STATE REGULATOR PROBLEM WITH OBSERVATION NOISE AND DRIVING DISTURBANCE

In many cases we do not have exact observation of the states of the system; rather, we observe the output of the system in the presence of measurement noise. In this section, we shall consider the special case when the system is linear and the observation noise is white and Gaussian.

Consider the plant

$$\dot{x}(t) = A(t)x(t) + B(t)\,u(t) + \xi(t) \tag{3.58}$$

$$y(t) = C(t)\,x(t) \tag{3.59}$$

where $x(t)$ is an n-vector

$u(t)$ is an r-vector

$A(t)$ is an $n \times n$ matrix

$B(t)$ is an $n \times r$ matrix

$\xi(t)$ is an n-vector white Gaussian process

$C(t)$ is an $m \times n$ matrix

$y(t)$ is an m-vector

We observe

$$s(t) = v(t) + y(t) \tag{3.60}$$

where $v(t)$ is m-vector white Gaussian process. We shall assume

---

†See Appendix 5

that we know the statistics of the white Gaussian processes and the initial condition of the plant.

$$E[\xi(t)] = 0 \; ; \; E[\xi(t)\xi'(\tau)] = D(t)\delta(t-\tau) \tag{3.61}$$

$$E[v(t)] = 0 \; ; \; E[v(t)v'(\tau)] = N(t)\delta(t-\tau) \tag{3.62}$$

$$E[\xi(t)v'(\tau)] = 0; \; E[x(t_o)\xi'(t)] = 0 \; ; \; E[x(t_o)v'(t)] = 0 \tag{3.63}$$

$$E[x(t_o)] = x_o \; ; \; E[x(t_o) - x_o)(x(t_o) - x_o)'] = \Sigma_o \tag{3.64}$$

The cost functional is

$$J = E[\langle x(t_1), fx(t_1)\rangle + \int_{t_o}^{t_1} \langle x(t), Q(t)x(t)\rangle + \langle u(t), R(t)u(t)\rangle \, dt \tag{3.65}$$

The problem is to find $u(t)$ to minimize the cost functional (3.64) subject to the Eq. 3.58. This problem though looks similar to the problem in Section 3.1, it bares a significant difference, namely we don't have exact knowledge of states, thus the simple linear feedback control derived in Section 3.1 cannot be optimal in this case. We nonetheless shall follow a similar procedure to that used in the previous sections to find the optimal control law.

1. Reformulation of Problem

We first try to fix the form of the optimal control from physical arguments. The maximal information we have about the state of the system at time t is the observation $s(t')$, for $t_o \leq t' \leq t$; in order that the control be optimal, it must be a function of the entire observation $s(t')$, $t_o \leq t' \leq t$. We wish to incorporate all the available information up to time t in some function of time t; in other words, we like to construct a function $w(t)$ which summarizes all the information contained in $s(t')$, $t_o \leq t' \leq t$, and deduce that the optimal control is of the form

$$u(t) = g(w(t), t) \; ; \; t_o \leq t \leq t_1 \tag{3.66}$$

Consider a dynamical system with $s(t)$ as the input and $w(t)$ as the state of the system. Clearly $w(t)$ will summarize all the information of $s(t')$, $t_o \leq t' \leq t$. Since $s(t)$ is Gaussian, we also require $w(t)$ to be Gaussian, thus we shall fix the dynamical system to be a linear system with parameter matrices $U(t)$, the feedback gain, and $V(t)$, the forward gain. Therefore we can fix the form of the optimal control to be

$$u(t) = M(t)\, w(t) \tag{3.67}$$

where

$$\dot{w}(t) = U(t)w(t) + V(t)s(t) \tag{3.68}$$

We shall assume that $E[w(t)] = E[x(t)]$ and that the dimension of $w(t)$ equal to the dimension of $x(t)$.† We can view this as a constraint on the form of the system II (see Fig. 2); or, intuitively, we can think of the dynamical system II acting as a filter giving the unbiased estimate of the present state of the plant we are controlling.

Equation 3.58 implies that‡

$$\dot{m}_x(t) = A(t)m_x(t) + B(t)m_u(t) \tag{3.69}$$

where

$$m_x(t) = E[x(t)] \; ; \; m_u(t) = E[u(t)] \tag{3.70}$$

Equations 3.59, 3.62, 3.63 and 3.68 imply

$$\dot{m}_w(t) = U(t)m_w(t) + V(t)C(t)m_x(t) \tag{3.71}$$

---

†See Appendix 7.

‡See Appendix 2.

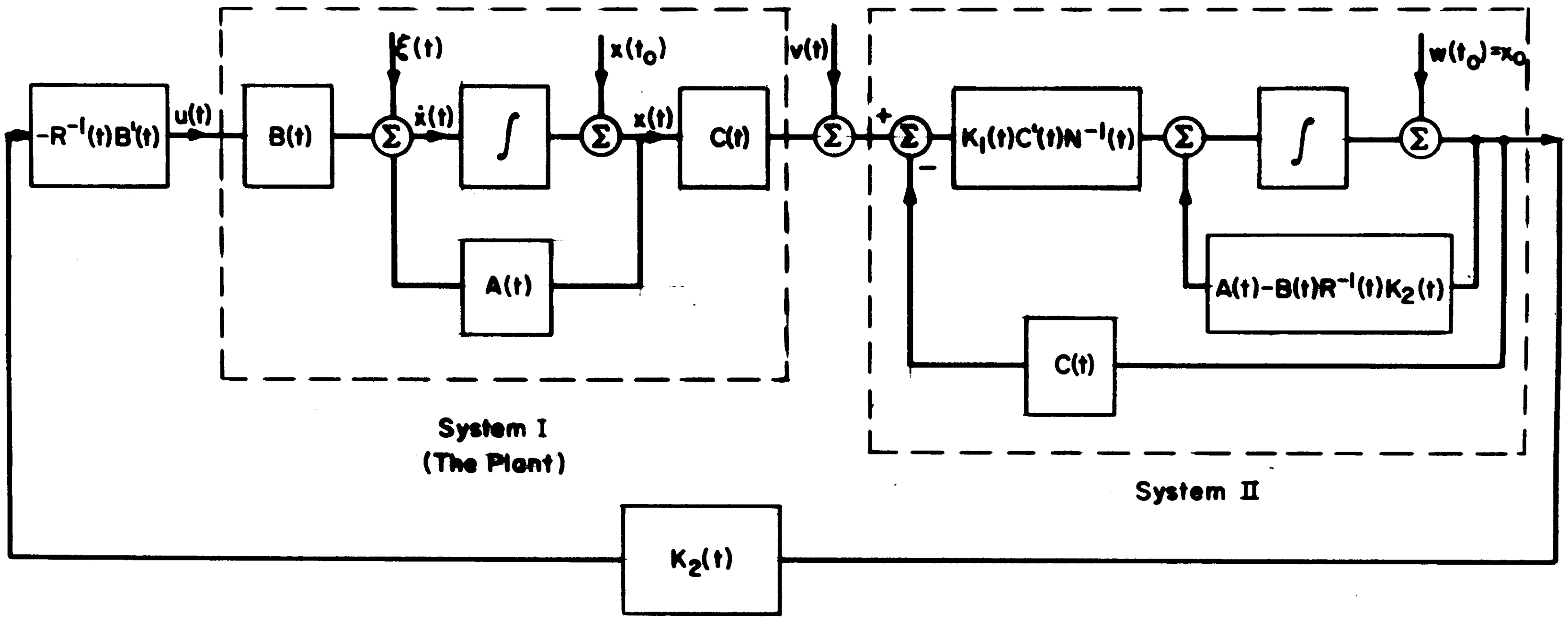

Fig. 2 State-Regulator Problem in the Presence of Disturbing and Observation Noise

$$\dot{K}_1(t) = A(t)K_1(t) + K_1(t)A'(t) - K_1(t)C'(t)N^{-1}(t)C(t)K_1(t) + D(t)$$

$$-\dot{K}_2(t) = A'(t)K_2(t) + K_2(t)A(t) - K_2(t)B(t)R^{-1}(t)B'(t)K_2(t) + Q(t)$$

$$K_1(t_o) = \Sigma_o \;;\; K_2(t_1) = F$$

where

$$m_w(t) = E[\,w(t)\,] \tag{3.72}$$

By Eq. 3.67, (3.69) yields

$$\dot{m}_x(t) = A(t)m_x(t) + B(t)M(t)m_w(t) \tag{3.73}$$

The unbiasness assumption gives us

$$[\,A(t) - U(t) - V(t)C(t) + B(t)M(t)\,]\,m_x(t) = 0 \tag{3.74}$$

This is true for all $m_x(t) \neq 0$, thus (3.74) implies

$$U(t) = A(t) - V(t)C(t) + B(t)M(t) \tag{3.75}$$

We see that the unbiasness assumption places a constraint on the parameters of the dynamical system II by Eq. 3.75. Let us define the composite vectors

$$z(t) = \begin{bmatrix} x(t) \\ \cdots\cdot \\ w(t) \end{bmatrix} \quad ; \quad \eta(t) \begin{bmatrix} \xi(t) \\ \cdots \\ v(t) \end{bmatrix} \tag{3.76}$$

Equations 3.58, 3.59, 3.60, 3.67, 3.68 and 3.75 imply that

$$z(t) = S(t)\; z(t) + Y(t)\; \eta(t) \tag{3.77}$$

where

$$S(t) = \left[\begin{array}{c:c} A(t) & B(t)M(t) \\ \cdots\cdots & \cdots\cdots\cdots\cdots\cdots\cdots \\ V(t)\,C(t) & A(t) - V(t)C(t) + B(t)M(t) \end{array}\right] \tag{3.78}$$

$$Y(t) = \left[\begin{array}{c:c} I & O \\ \cdots & \cdots \\ O & V(t) \end{array}\right] \tag{3.79}$$

and the statistics of $\eta(t)$ are

$$E[\eta(t)] = 0 \quad ; \quad E[\eta(t)\eta'(\tau)] = X(t)\ \delta(t-\tau) \tag{3.80}$$

where

$$X(t) = \begin{bmatrix} D(t) & \vdots & O \\ \cdots & \vdots & \cdots \\ O & \vdots & N(t) \end{bmatrix} \tag{3.81}$$

and

$$E[z(t_o)\eta'(t)] = 0 \quad ; \quad t_o \leq t \leq t_1 \tag{3.82}$$

By Appendix 2, Eqs. 3.77 together with 3.81, 3.82 imply that

$$\dot{\Sigma}_z(t) = S(t)\ \Sigma_z(t) + \Sigma_z(t)\ S'(t) + Y(t)\ X(t)\ Y'(t) \tag{3.83}$$

where we define

$$\Sigma_z(t) = \begin{bmatrix} \Sigma_x(t) & \vdots & \Sigma_{xw}(t) \\ \cdots & \vdots & \cdots \\ \Sigma_{wx}(t) & \vdots & \Sigma_w(t) \end{bmatrix} \tag{3.84}$$

The cost functional (3.64) can be written as

$$J = \mathrm{Tr}[F\Sigma_x(t_1) + \int_{t_o}^{t_1} Q(t)\Sigma_x(t) + M'(t)R(t)M(t)\Sigma_w(t)\ dt] \tag{3.85}$$

The new optimization problem is: <u>given</u> (3.83) <u>and the cost functional</u> (3.85), <u>choose</u> M(t), V(t) <u>such that</u> (3.85) <u>is minimized subject to the matrix differential equation</u> (3.83).

2. Application of the Matrix Minimum Principle

Let the costate matrix of $\Sigma_z(t)$ be $P_z(t)$ where

$$P_z(t) = \begin{bmatrix} P_x(t) & \vdots & P_{xw}(t) \\ \cdots & \vdots & \cdots \\ P_{wx}(t) & \vdots & P_w(t) \end{bmatrix} \tag{3.86}$$

Therefore the costate matrices for $\Sigma_x(t)$, $\Sigma_w(t)$, $\Sigma_{wx}(t)$ and $\Sigma_{xw}(t)$ are $P_x(t)$, $P_w(t)$, $P_{wx}(t)$ and $P_{xw}(t)$ respectively. We form the Hamiltonian

$$H = Tr[\dot{\Sigma}_z(t)\ P_z'(t) + Q(t)\ \Sigma_x(t) + M'(t)R(t)M(t)\ \Sigma_w(t)] \tag{3.87}$$

The canonical equations for the costate yields:

$$-\dot{P}_x^* = A'(t)P_x^*(t) + C'(t)V^{*'}(t)\ P_{wx}^*(t) + P_x^*(t)A(t) + P_{xw}^*(t)V^*(t)C(t) + Q(t) \tag{3.88}$$

$$\begin{aligned}-\dot{P}_{xw}^*(t) = {} & A'(t)P_{xw}^*(t) + C'(t)V^{*'}(t)P_w^*(t) + P_x^*(t)B(t)M^*(t) + P_{xw}^*(t)A(t) \\ & -P_{xw}^*(t)V^*(t)C(t) + P_{xw}^*B(t)M^*(t)\end{aligned} \tag{3.89}$$

$$\begin{aligned}-\dot{P}_{wx}^*(t) = {} & A'(t)P_{wx}^* - C'(t)V^{*'}(t)P_{wx}^*(t) + M^{*'}(t)B'(t)P_{wx}^*(t) \\ & +M^{*'}(t)B'(t)P_x^*(t) + P_{wx}^*(t)A(t) + P_w^*(t)\ V^*(t)C(t)\end{aligned} \tag{3.90}$$

$$\begin{aligned}-\dot{P}_w^*(t) = {} & M^{*'}(t)B'(t)P_{xw}^*(t) + A'(t)P_w^*(t) - C'(t)V^{*'}(t)P_w^*(t) \\ & +M^{*'}(t)B'(t)P_w^*(t) + P_{wx}^*(t)B(t)M^*(t) + P_w^*(t)A(t) \\ & -P_w^*(t)V^*(t)C(t) + P_w^*(t)B(t)M^*(t) + M^{*'}(t)R(t)M^*(t)\end{aligned} \tag{3.91}$$

The transversality conditions give the boundary conditions

$$P_x^*(t_1) = F\ ;\quad P_{xw}^*(t_1) = P_{xw}^*(t_1) = P_w^*(t_1) = 0 \tag{3.92}$$

By Eqs. 3.89, 3.90, 3.91 and 3.92, we deduce that

$$P_{wx}^*(t) = P_{xw}^{*'}(t)\ ;\quad P_w^*(t) = P_w^{*'}(t) \tag{3.93}$$

To find the necessary conditions for the minimization of the Hamiltonian, $H(\Sigma_z^*(t), P_z^*(t), M(t), V(t))$ with respect to $V(t)$ and $M(t)$ independently, we set

$$\frac{\partial H}{\partial V(t)}[\Sigma_z^*(t),\ P_z^*(t),\ M(t),\ V(t)]\Bigg|_* = 0 \tag{3.94}$$

$$\frac{\partial H}{\partial M(t)}[\Sigma_z^*(t),\ P_z^*(t),\ M(t),\ V(t)]\Bigg|_* = 0 \tag{3.95}$$

Differentiating (3.78) with respect to $V(t)$ and $M(t)$, respectively, (3.94) and (3.95) yield

$$P_{wx}^*(t)\Sigma_x^*(t)C'(t) - P_{wx}^*(t)\Sigma_{xw}^*(t)C'(t) + P_w^*(t)\Sigma_{wx}^*(t)C'(t) - P_w^*(t)\Sigma_w^*(t)C'(t)$$

$$+P_w^*(t)\ V^*(t)\ N(t) = 0 \tag{3.96}$$

$$B'(t)P_x^*(t)\Sigma_{xw}^*(t) + B'(t)P_{xw}^*(t)\Sigma_w^*(t) + B'(t)P_{wx}^*(t)\Sigma_{xw}^*(t)$$

$$+ B'(t)P_w^*(t)\Sigma_w^*(t) + R(t)M^*(t)\Sigma_w^*(t) = 0 \tag{3.97}$$

By assumption, $N(t)$ and $R(t)$ are positive definite, therefore the Eqs. 3.94 and 3.95 give the values of $M^*(t)$ and $V^*(t)$ which minimize the Hamiltonian.

By assumption $\Sigma_w^*(t)$, $N(t)$, $R(t)$ are nonsingular, and Eqs. 3.91 and 3.92 imply that $P_w^*(t)$ is nonnegative definite;† thus Eqs. 3.96 and 3.97 yield

$$V^*(t) = P_w^{*-1}(t)P_{wx}(t)[\Sigma_{xw}^*(t) - \Sigma_x^*(t)]\,C'(t)N^{-1}(t) + [\,\Sigma_w^*(t) - \Sigma_{wx}^*(t)]\,C'(t)N^{-1}(t) \tag{3.98}$$

---

†If $M^*(t)$ is nonzero, then $P_w^*(t)$ is positive definite for $t \neq t_o$.

$$M^*(t) = -R^{-1}(t)B'(t)[P_x^*(t) + P_{wx}^*(t)]\,\Sigma_{xw}^*(t)\Sigma_w^{*-1}(t)\; -R^{-1}(t)B'(t)[P_{xw}^*(t) + P_w^*(t)] \tag{3.99}$$

Substituting (3.98) and (3.99) into (3.90), (3.91), (3.88), (3.83) we have a set of matrix differential equations† of $\Sigma_z^*(t)$ and $P_z^*(t)$ with boundary conditions

$$\Sigma_z^*(t_o) = \left[\begin{array}{c:c} \Sigma_o + x_o x_o' & x_o x_o' \\ \cdots\cdots & \cdots \\ x_o x_o' & x_o x_o' \end{array}\right] \tag{3.100}$$

$$P_z^*(t_1) = \left[\begin{array}{c:c} F & O \\ \cdots & \cdots \\ O & O \end{array}\right] \tag{3.101}$$

Formally, we can solve the resulting matrix differential equations for $\Sigma_z^*(t)$ and $P_z^*(t)$, and substitute the solution into (3.98) and (3.99) to obtain $V^*(t)$ and $M^*(t)$. However, it can be shown that‡ the solution that satisfies the resulting matrix differential equation and the corresponding boundary conditions is of the form

$$\Sigma_{xw}^*(t) = \Sigma_w^*(t) \quad ; \quad P_{xw}^*(t) = -P_{xw}^*(t) \tag{3.102}$$

By the uniqueness theorem, (3.102) is the form of the solution we are interested in. Substituting (3.102) into (3.97) and (3.98), we have

$$V^*(t) = [\,\Sigma_x^*(t) - \Sigma_w^*(t)]\,C'(t)N^{-1}(t) \tag{3.103}$$

$$M^*(t) = -R^{-1}(t)B'(t)[\,P_x^*(t) - P_w^*(t)] \tag{3.104}$$

---

†See Appendix 5, Eqs. A.5.16, A.5.17, A.5.18 and A.5.19.

‡See Appendix 5.

Let us define

$$K_1(t) = \Sigma_x^*(t) - \Sigma_w^*(t) \tag{3.105}$$

$$K_2(t) = P_x^*(t) - P_w^*(t) \tag{3.106}$$

Differentiating (3.105) and using (3.77), (3.103) and (3.104) we have

$$\dot{K}_1(t) = A(t)\,K_1(t) + K_1(t)\,A'(t) - K_1(t)C'(t)N^{-1}(t)\,C(t)\,K_1(t) + D(t) \tag{3.107}$$

Differentiating (3.106) and using (3.88), (3.91), (3.103) and (3.104) we have

$$-\dot{K}_2(t) = A'(t)\,K_2(t) + K_2(t)\,A(t) - K_2(t)\,B(t)\,R^{-1}(t)\,B'(t)\,K_2(t) + Q(t) \tag{3.108}$$

From Eqs. 3.100, 3.101, 3.105 and 3.106 we have the boundary conditions for $K_1(t)$ and $K_2(t)$.

$$K_1(t_o) = \Sigma_o \quad ; \quad K_2(t_1) = F \tag{3.109}$$

Using (3.105) and (3.106), the matrices $V^*(t)$ and $M^*(t)$ can be written as

$$V^*(t) = K_1(t)C'(t)N^{-1}(t) \tag{3.110}$$

$$M^*(t) = -R^{-1}(t)\,B'(t)\,K_2(t) \tag{3.111}$$

The extremal control which satisfies the necessary conditions for optimality is specified by (3.67), (3.68), (3.75), (3.107) (3.108), (3.110) and (3.111). (See Fig. 2)

3. <u>The Minimum Cost Functional</u>

Let

$$J^*[\Sigma_z(t), t] = Tr[K_2(t)\Sigma_x(t)] + f(t) \tag{3.112}$$

where

$$\left.\begin{aligned} -\dot{f}(t) &= \mathrm{Tr}[\,K_1(t)\,K_2(t)\,B(t)\,R^{-1}(t)\,B'(t)\,K_2(t) + D(t)\,K_2(t)] \\ f(t_1) &= 0 \end{aligned}\right\} \qquad (3.113)$$

and $K_1(t)$, $K_2(t)$ satisfy (3.107), (3.108) and (3.109). If $V(t)$ and $M(t)$ are chosen as given by (3.110) and (3.111), then by direct substitution, it can be shown† that (3.112), (3.113) satisfy the Hamilton-Jacobi equation and boundary condition. Thus the extremal solution satisfies the sufficient condition for optimality.

D. DISCUSSION OF RESULTS

1. State-Regulator Problem with Driving Disturbance of Zero-Mean

We find that the optimum linear feedback control for this problem is described by (see Eqs. 3.7, 3.17, 3.18 and 3.19)

$$u(t) = -R^{-1}(t)\,B'(t)\,P_x(t)\,x(t) \qquad (3.114)$$

where $P_x(t)$ satisfies

$$\left.\begin{aligned} -\dot{P}_x(t) &= A'(t)\,P_x(t) + P_x(t)\,A(t) - P_x(t)\,B(t)\,R^{-1}(t)\,B'(t)\,P_x(t) + Q(t) \\ P_x(t_1) &= F \end{aligned}\right\} \qquad (3.115)$$

First we notice that the linear feedback gain depends only on the structure of the plant and the weighting matrices, F, $R(t)$, $Q(t)$. Once the structure of the plant is specified and the weighting matrices are appropriately chosen, we can implement the linear feedback that will give us the optimal performance consistent with the given mathematical problem. Comparing the results obtained here with the case

---

†See Appendix 5

where there is no disturbance noise,[1] we see that the feedback gain is the same for both problems when the weighting matrices chosen in both problems are the same. If the weighting matrices are given and the dynamics of the system is known, then the results obtained in Section A of this chapter imply that we can design the optimum feedback gain by simply ignoring the effect of the noise. The driving disturbance does not effect the structure of the feedback gain, it only effects the minimum cost functional (Eqs. 3.21, 3.22).

2. State-Regulator Problem with Driving Disturbance of Nonzero Mean

For this problem, we find that the optimal control is described by (see Eqs. 3.37, 3.51, 3.52, 3.53, 3.54 and 3.55)

$$u(t) = -R^{-1}(t)B'(t)P_x(t)x(t) - \frac{1}{2}R^{-1}(t)B'(t)p_x(t) \qquad (3.116)$$

where $P_x(t)$ and $p_x(t)$ satisfy

$$-\dot{P}_x(t) = A'(t)P_x(t) + P_x(t)A(t) - P_x(t)B(t)R^{-1}(t)B'(t)P_x(t) + Q(t) \qquad (3.117)$$

$$-\dot{p}_x(t) = 2P_x(t)m(t) + A'(t)p_x(t) - P_x(t)B(t)R^{-1}(t)B'(t)p_x(t) \qquad (3.118)$$

with boundary conditions

$$P_x(t_1) = F \quad ; \quad p_x(t_1) = 0 \qquad (3.119)$$

Comparing (3.114) and (3.116), we see that we have a biasing term, $-\frac{1}{2}R^{-1}(t)B'(t)p_x(t)$, in (3.116) which appears because the mean of the disturbing noise is nonzero. Physically, we may view the biasing term as cancelling the effect of the mean of the noise. We also see that the structure of the optimal control depends on the structure of the plant, the weighting matrices and the first moment of the noise process. Once all these are specified, we can implement the control by a linear feedback gain plus a biasing term described by (3.116), (3.117), (3.118) and (3.119). The second moment of the noise process only effects the minimum cost functional.

3. State-Regulator Problem with Observation Noise and Driving Disturbance

The optimal control for this problem is described by (see Eqs. 3.67, 3.68, 3.75, 3.107, 3.108, 3.110 and 3.111)

$$u(t) = -R^{-1}(t)\, B'(t)\, K_2(t)\, w(t) \tag{3.120}$$

$$\dot{w}(t) = [\,A(t) - K_1(t)C'(t)N^{-1}(t)\, C(t) - B(t)\, R^{-1}(t)\, B'(t)\, K_2(t)]\, w(t)$$

$$+ K_1(t)\, C'(t)\, N^{-1}(t)\, s(t) \tag{3.121}$$

where $K_1(t)$ and $K_2(t)$ satisfy the differential equations

$$\dot{K}_1(t) = A(t)\, K_1(t) + K_1(t)\, A'(t) - K_1(t)\, C'(t)N^{-1}(t)\, C(t)\, K_1(t) + D(t) \tag{3.122}$$

$$-\dot{K}_2(t) = A'(t)\, K_2(t) + K_2(t)\, A(t) - K_2(t)\, B(t)\, R^{-1}(t)\, B'(t)\, K_2(t) + Q(t) \tag{3.123}$$

with boundary conditions

$$K_1(t_o) = \Sigma_o \;\; ; \;\; K_2(t_1) = F \tag{3.124}$$

We see that the optimal control is generated by a linear feedback of the state $w(t)$; note that $w(t)$ is the best estimate of the state $x(t)$ in the minimum mean square error sense. (see Ref. 18, 22, 24 and 30). Also note that Eqs. 3.122 and 3.123 are decoupled; this means that we can solve the problem by two separate steps: we first estimate the states and then find the optimal feedback gain feeding back the best state-estimate.[24] Thus, in solving this problem we simultaneously estimate and control. The minimum functional specified by (3.112) and (3.113) depends on the structure dynamics of the plant, the statistics of the disturbance noise and measurement noise, and the weighting matrices. Comparing Eqs. 3.112, 3.113, 3.108, 3.109 with Eqs. 3.21, 3.22, 3.18 and 3.19, we see that the minimal cost functional for the two problems (one with and the other without observation noise) are very similar, and the minimal functional for the two problems is the same in the limit when the covariance of the observation noise becomes smaller and smaller.

From the results we see that the linear feedback gain matrix only depends on the structure of the plant and the weighting matrices. The states to be fed back represent the maximum information about the past history of the plant to be controlled. Florentine[13] also investigated the state-regulator problems with disturbance noise only. Comparing Eq. 5.25 in Ref. 13 and Eqs. 3.57, 3.54 and 3.53, we see that there is an error in Eq. 5.25[†] of Ref. 13. The structure of the optimal control for state-regulator problem with measurement and disturbance noises agree with that obtained by Roberts[24] and Wonham.[30] In our results, we also obtain the minimal cost functional, (Eqs. 3.112 and 3.113).

E. DISCUSSION OF OUR APPROACH

The advantage of this approach is the simplicity of method even when the complexity of problem increases. The procedures are the same:

1. Reformulation of a stochastic problem into a deterministic system.
2. Application of the matrix minimum principle.
3. Evaluation of the minimal functional.

---

† Using the notation in this chapter, the second equation of (5.25) in Ref. 13 reads

$$-\dot{p}_x^{*'}(t) = -p_x^{*'}(t)\, B(t)\, R^{-1}(t)\, B(t) + p_x^{*'}(t)\, A(t) + 2m'(t)\, p_x^{*}(t)$$

Usually most of the work is in step one; in steps two and three, it only involves stright-forward manipulation. In the reformulation, usually we get a new system described by a matrix different equation. The matrix-state is a covariance matrix of a vector process; thus the matrix-state must be a symmetric matrix. In order for the matrix minimum principle to be applicable, we must not impose this constraint on the matrix-state.† We can bypass this apparent difficulty by assuming that the matrix-state is arbitrary and that its symmetry is only a consequence of the form of the matrix differential equation and of the specified boundary condition. It is only with this approach that we can apply with confidence the formulae for gradient matrices given in Appendix 6.

The disadvantage of this approach lies in the fact that we can only deal with problems for which we have a complete deterministic description of the random process by means of a finite set of ordinary differential equations. For all problems with linear plants and functionals of quadratic form, this approach is applicable. But for more general problems, this approach may fail, e.g., if the system is nonlinear or the functional is not of quadratic form, we may not be able to reformulate the problem in terms of the deterministic quantities describing the overall dynamics of the system. In such problems, the Dynamic Programming approach gives some theoretical results, but unfortunately few practical results can be obtained for this class of problems. Thus it seems that the approach used here, though somewhat limited, deserves its merits.

F. FUTURE RESEARCH

By a simple extension we can solve the problem of regulating the states when we have colored observation noise. Using this approach, we can tackle the problems when we have constraints on the entries of the feedback gain matrix. For example, we may impose

---

†See Chapter 2, Section D.1.

magnitude constraints on the gains $m_{ij}(t)$, $i=1,\ldots,n$, $j=1,\ldots,r$ of the type

$$\alpha_{ij} \leq m_{ij}(t) \leq \beta_{ij} \tag{3.125}$$

For this problem we can still use the matrix minimum principle to find the optimal linear feedback under these constraints. This approach is directly applicable to the design of an optimal cascade compensator, feedback compensator and many fixed configuration design problems in which we can adjust the parameters so as to minimize a cost functional with quadratic criteria.

CHAPTER IV

NONLINEAR STATE ESTIMATION WITH QUADRATIC CRITERIA

The state estimation problem of a nonlinear system is of interest both to system control engineers and communication engineers. During the past few years, many papers have been published which addressed themselves to the nonlinear filtering problems seeking an optimal or suboptimal filter which will generate the best estimate of state variables in a minimum mean square error sense.[8, 9, 19, 20, 27 and 31] It is well known[18] that the mean value of the random state vector of the system, conditioned on the Borel field induced by observation, will yield the best mean square error estimate. If the plant is nonlinear, the difficulty seems to lie in the explicit calculation of the conditional mean of the random state vector of the system. Even if we can mathematically write out the structure of the optimal filter, further difficulties arise due to the fact that, in many cases, the filter we obtain is of infinite dimension.[20, 19] This means that we must approximate the optimal filter by a suboptimal filter of finite dimensionality. In the approximation procedure, further difficulties can arise because we are approximating random variables instead of real numbers. The validity of the familiar technique--such as Taylor series expansion--is questionable.

In this chapter we shall attempt to obtain some results for the nonlinear state estimation problem. In view of the difficulties that we have to face in following a conventional approach, we shall attack the problem from a different direction. In so doing, we shall be able to bypass the Ito equation and the convergence of integrals to stochastic integrals. Rather, we shall first formulate the problem as a deterministic optimization problem so that we can apply the minimum principle. Then, using the imbedding technique, we shall approximate the solution for the nonlinear filter.

A. FORMULATION OF THE CONTROL PROBLEM

Consider a plant

$$\dot{x}(t) = f(x(t)) + \xi(t) \tag{4.1}$$

where $x(t)$ is an n-vector

$\xi(t)$ is a Gaussian white noise n-vector

We assume that we know the statistics of the driving disturbance and some partial information about the initial value of the states, i.e.,

$$E[\xi(t)] = 0 \;\; ; \;\; E[\xi(t)\xi'(\tau)] = D(t)\,\delta(t-\tau) \tag{4.2}$$

$$E[x(t_o)] = x_o \;\; ; \;\; E[(x(t_o) - x_o)(x(t_o) - x_o)'] = \Sigma_o \tag{4.3}$$

In many cases, we cannot observe the state of the system but only the output of the system in the presence of noise,

$$s(t) = h(x(t)) + v(t) \tag{4.4}$$

where $s(t)$ is the observation m-vector

$v(t)$ is the white noise Gaussian random m-vector

We assume that $\xi(t)$ and $v(t)$ are uncorrelated, and the statistics of $v(t)$ are given by

$$E[v(t)] = 0 \;\; ; \;\; E[v(t)v'(t)] = N(t)\,\delta(t-\tau) \tag{4.5}$$

The problem we have to solve is to estimate the states of the system at time t base on the available information up to time t, i.e., the observation vector $s(t')$, $t_o \leq t' \leq t$. In the conventional approach, we view $\xi(t)$, $v(t)$, $x(t)$ and $s(t)$ as random processes; if we want to have a best estimate in the mean square sense, we need a filter which will produce the conditional mean of $x(t)$ from the observation data, $s(t')$, $t_o \leq t' \leq t$. In this thesis, we shall deviate from this approach. We shall view $s(t')$, $t_o \leq t' \leq t$, as a real function (not a random process) which is what we observed. We shall view $v(t)$ and $\xi(t)$ as sample functions from the two ensembles. The exact value of the sample functions are unknown, but somehow we have some information about the regularity (i.e., the statistics) of the functions. We shall construct a meaningful functional which we want to minimize so as to give a good estimate of the states.

Suppose that we know exactly the initial states of the system and the exact sample function $\xi(t)$, $t_o \leq t \leq t_1$; then, clearly, we can reconstruct the estimate of $x(t)$ (which is exactly equal to the state $x(t)$) from

$$\dot{w}(t) = f(w(t)) + \xi(t) \tag{4.5}$$

$$w(t_o) = x(t_o) \tag{4.6}$$

But the problem is that we know neither the exact sample function $\xi(t)$, $t_o \leq t \leq t_1$, nor the initial state $x(t_o)$. Let us guess the initial value of $x(t_o)$ as w_o, and the sample function as $u(t)$, $t_o \leq t \leq t$. Let $w(t)$ be a vector that satisfies the equation

$$\dot{w}(t) = f(w(t)) + u(t) \tag{4.7}$$

$$w(t_o) = w_o \tag{4.8}$$

If $u(t)$ and w_o are close to $\xi(t)$ and $x(t_o)$, respectively, then there is good reason to suppose that $w(t)$ will be close to $x(t)$. For this reason, we shall say that $w(t)$ is an estimate of $x(t)$.

Next we define

$$\hat{v}(t) = s(t) - h(w(t)) \; ; \; t_o \leq t \leq t_1 \tag{4.9}$$

If $w(t)$ is exactly the state of the system, $\hat{v}(t)$ will be exactly the sample value of the observation noise $v(t)$ at time t. Thus, if $w(t)$ is close to $x(t)$, then $\hat{v}(t)$ will be close to the true value of the sample function $v(t)$. In this sense, we can view $\hat{v}(t)$, $t_o \leq t \leq t_1$, as the guessed sample function for $v(t)$.

We suppose that we are given some partial information about $\xi(t)$, $x(t_o)$ and $v(t)$, namely, their means and variances. We shall make full use of the available information to obtain a good guess for the sample function $\xi(t)$, $t_o \leq t \leq t_1$, and the initial value w_o. Toward this end, we construct the functional

$$J = (w_o - x_o)'\Sigma_o^{-1}(w_o - x_o) + \int_{t_o}^{t_1} \{\hat{v}'(t)N^{-1}(t)\hat{v}(t) + u'(t)D^{-1}(t)u(t)\}dt \tag{4.10}$$

If we want to have a high probability that the actual sample function $\xi(t)$, $t_o \leq t \leq t_1$, and $x(t_o)$ are "close" to the guessed sample function $u(t)$ and initial value w_o, then we need to minimize the functional (4.10).† We have now the control problem:

Given the plant

$$\dot{w}(t) = f(w(t)) + u(t) \tag{4.11}$$

and the functional

$$J = (w_o - x_o)'\Sigma_o^{-1}(w_o - x_o) + \int_{t_o}^{t_1} \{(s(t) - h(w(t)))'N^{-1}(t)(s(t) - h(w(t))) + u'(t)D^{-1}(t)u(t)\}dt \tag{4.12}$$

†See Appendix 3.

We are to choose w_o, and $u(t)$, $t_o \leq t \leq t_1$, such that $w(t_o) = w_o$ and (4.11) is satisfied while minimizing the functional (4.12).

B. APPLICATION OF THE MINIMUM PRINCIPLE

Let $p(t)$ be the costate corresponding to $w(t)$; the Hamiltonian for the optimal control problem

$$H = p'(t)f(w(t)) + p'(t)u(t) + (s(t) - h(w(t))'N^{-1}(t)(s(t) - h(w(t)) + u'(t)D^{-1}(t)u(t) \tag{4.13}$$

The canonical equations are

$$-\dot{p}^*(t) = \frac{\partial H}{\partial w}\bigg|_* = \left(\frac{\partial f(w^*(t))}{\partial w^*(t)}\right)' p^*(t) + 2\left(\frac{2h(w^*(t))}{\partial w^*(t)}\right)' N^{-1}(t)\,(h(w^*(t)) - s(t)) \tag{4.14}$$

$$\dot{w}^*(t) = \frac{\partial H}{\partial p}\bigg|_* = f(w^*(t)) + u^*(t) \tag{4.15}$$

with boundary conditions:

$$w(t_o) = w_o \quad ; \quad p^*(t_o) = -2\,\Sigma_o^{-1}(w_o - x_o) \ ; \quad p^*(t_1) = 0 \tag{4.16}$$

To minimize H with respect to $u(t)$, we set

$$\frac{\partial H}{\partial u}\bigg|_* = 0 \tag{4.17}$$

to obtain the optimal $u^*(t)$ given by

$$u^*(t) = -\frac{1}{2}D(t)\,p^*(t) \tag{4.18}$$

The assumption that $D(t)$ is positive definite implies that $u^*(t)$ is a minimum of H. Substituting (4.18) into (4.15), we have the set of canonical equations

$$-\dot{p}^*(t) = \left(\frac{\partial f(w^*(t)}{\partial w^*(t)}\right)' p^*(t) + 2\left(\frac{\partial h(w^*(t)}{\partial w^*(t)}\right)' N^{-1}(t)(h(w^*(t)) - s(t)) \tag{4.19}$$

$$\dot{w}^*(t) = f(w^*(t)) - \frac{1}{2} D(t)p^*(t) \tag{4.20}$$

with boundary conditions

$$w^*(t_o) = w_o \ ; \ p^*(t_o) = -2\Sigma_o^{-1}(w_o - x_o) \ ; \ p^*(t_1) = 0 \tag{4.21}$$

To find the best estimate of the state for time t, $t_o \leq t \leq t_1$, based on the available data, $s(t)$ for $t_o \leq t \leq t_1$, we have to solve the set of differential equations (4.19), (4.20) subject to the boundary conditions (4.21).

C. SEQUENTIAL ESTIMATION

In practice, we would like to have an estimate of the state for time $t \geq t_o$. The results obtained in the last section provide a method for obtaining the best interval estimation of the state after we have observed $s(t)$, $t_o \leq t \leq t_1$. If the observation inverval increases, say the terminal time is t_2, $t_2 > t_1$, then to get the best interval estimate based on the observation $s(t)$, $t_o \leq t \leq t_2$, we have to solve the set of differential equations (4.19), (4.20) with the boundary conditions

$$w^*(t_o) = w_o \ ; \ p^*(t_o) = -2\Sigma_o^{-1}(w_o - x_o) \ ; \ p^*(t_2) = 0 \tag{4.22}$$

However, in filtering, we may be interested in an estimator which will give us the best point estimate, i.e., at any time t_1 we would like the filter to generate $w^*(t_1)$ and we may not be interested in $w^*(t)$, $t_o \leq t < t_1$. So, at any time t_1, we want the estimator to construct the best estimated state $w^*(t_1)$ based on the available data up to time t_1. Because of the nature of this estimator, the process is called a sequential estimation.

Suppose that we seek the interval estimate of the state after observing $s(t)$, $t_o \leq t \leq t_1$. After solving (4.19), (4.20) and (4.21), we have the solution for $w^*(t)$.

$$w^*(t) = r_{t_1}(t) \quad ; \quad t_o \leq t \leq t_1 \tag{4.23}$$

The subscript t_1 is used to denote that the terminal time is t_1, (see Fig. 3). The function that we are interested in is $r_\tau(\tau)$, $\tau \geq t_o$. We shall find the differential equation that $r_\tau(\tau)$, $\tau \geq t_o$, has to satisfy; this will give us the structural form of the optimum sequential estimator. To do this, we shall employ the imbedding technique which gives us the desired solution $r_\tau(\tau)$.

Consider the set of Eqs. 4.14, 4.15; we shall denote the value of the solution at the terminal time τ by

$$r_\tau(\tau) = r(c, \tau) \text{ when } p^*(\tau) = c \tag{4.24}$$

From Appendix 4, we observe that $r(c, \tau)$ and $H^*(\tau, r, c)$ satisfy the partial different equation

$$\frac{\partial r}{\partial \tau} - \frac{\partial r}{\partial c} \frac{\partial H^*(\tau, r, c)}{\partial r} = \frac{\partial H^*(\tau, r, c)}{\partial c} \tag{4.25}$$

We shall solve (4.25) for $r(c, \tau)$, and the solution that we are interested in is imbedded in the solution for (4.25), i.e., $r(0, \tau)$. If the system is linear, then we can solve (4.25) by the separation of variable method and obtain the Kalman-Bucy filter.† However, for general nonlinear systems, we may not solve (4.25) easily; even if the solution of (4.25) is possible, the estimator obtained may be of infinite dimensionality. Thus, for a practical solution, we shall have to obtain an approximation to the optimal estimator.

---

†See Section E in this chapter.

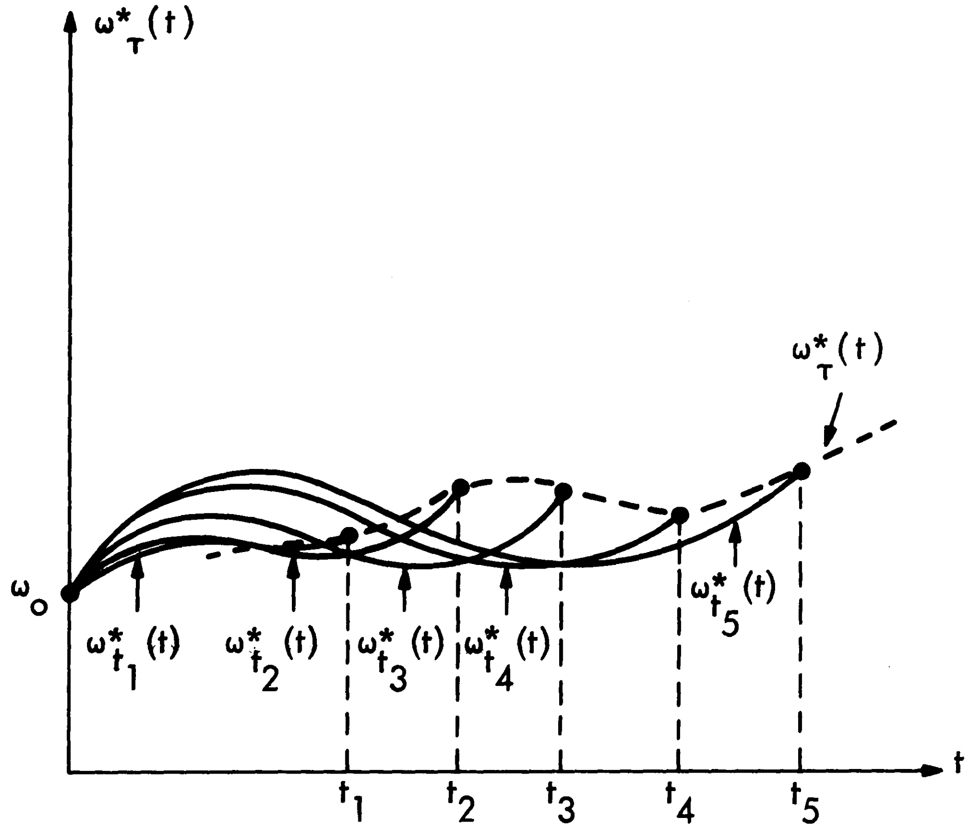

Fig. 3 Best Sequential Estimate of a Nonlinear Dynamical System

D. APPROXIMATION OF THE OPTIMUM NONLINEAR ESTIMATOR

We shall assume that $r(c, t)$ which satisfies (4.25) has a Taylor series expansion about $c=0$. We have for $i=1, \ldots, n$

$$r_i(c, \tau) = r_i(0, \tau) + \left(\frac{\partial r_i}{\partial c}\bigg|_{c=0}\right)' c + \sum_{j,k} \frac{1}{2} \frac{\partial^2 r_i}{\partial c_j \partial c_k}\bigg|_{c_j=c_k=0} c_j c_k + \ldots \tag{4.26}$$

We can substitute Eq. 4.26 into Eq. 4.25 and equate the coefficients of terms of the same order of c; in so doing, we have a set (usually infinite set) of ordinary different equations. The exact solution can be obtained by solving the resulting set of differential equations. However, we shall approximate the solution $r(c, \tau)$ by truncating the series (4.26). In this chapter we shall consider terms up to the first order of c.

Let us approximate $r_i(c, \tau)$ by

$$r_i(c, \tau) = a_i(\tau) + \sum_{j=1}^{n} b_{ij}(\tau) c_j \tag{4.27}$$

If we write Eq.4.27 in vector form, we have

$$r(c, \tau) = a(\tau) + B(\tau) c \tag{4.28}$$

Differentiating (4.13) and using Eq. 4.28, we have

$$\begin{aligned}\frac{\partial H(\tau, r, c)}{\partial r} = \left(\frac{\partial f(a(\tau))}{\partial a(\tau)}\right)' c - 2\left(\frac{\partial h(a(\tau))}{\partial a(\tau)}\right)' N^{-1}(\tau)(s(\tau) - h(a(\tau)) \\ -2\left\{\frac{\partial}{\partial a(\tau)}\left[\left(\frac{\partial h(a(\tau))}{\partial a(\tau)}\right)' N^{-1}(\tau)(s(\tau) - h(a(\tau))\right]\right\} B(\tau) c \\ + o(c\, c')\end{aligned} \tag{4.29}$$

$$\frac{\partial H(\tau, r, c)}{\partial c} = f(a(\tau)) + \frac{\partial f(a(\tau))}{\partial a(\tau)} B(\tau)c - \frac{1}{2} D(\tau)c \tag{4.30}$$

Substituting (4.28), (4.29), (4.30) into (4.25) and equating coefficients of zero and first order of c, give

$$\frac{da(\tau)}{d\tau} = f(a(\tau)) - 2B(\tau)\left(\frac{\partial h(a(\tau))}{\partial a(\tau)}\right)' N^{-1}(\tau)(s(\tau) - h(a(\tau)) \tag{4.31}$$

$$\frac{dB(\tau)}{d\tau} = B(\tau)\left(\frac{\partial f}{\partial a}\right)' + \left(\frac{\partial f}{\partial a}\right) B(\tau) - \frac{1}{2} D(\tau)$$
$$-2 B(\tau)\left\{\frac{\partial}{\partial a}\left[\left(\frac{\partial h}{\partial a}\right)' N^{-1}(\tau)(s(\tau) - h(a(\tau))\right]\right\} B(\tau) \tag{4.32}$$

By definition we have

$$r(-2\,\Sigma_o^{-1}(w_o - x_o),\ t_o) = w_o \tag{4.33}$$

Equation 4.28 implies that

$$a(t_o) - B(t_o)\ 2\Sigma_o^{-1}(w_o - x_o) = w_o \tag{4.34}$$

By setting

$$a(t_o) = x_o \quad , \quad B(t_o) = -\frac{1}{2}\Sigma_o \tag{4.35}$$

we see that (4.34) is satisfied. The solution that is of interest to us is $r(0,\tau)$; by (4.28), we have

$$r(0,\tau) = a(\tau) \tag{4.36}$$

Thus $a(\tau)$ is the best estimate of the state at time τ, $\tau \geq t_o$.

Define

$$K(\tau) = 2B(\tau) \tag{4.37}$$

Then Eqs. 4.31 and 4.32 give us the set of differential equations

$$\frac{da(\tau)}{d\tau} = f(a(\tau)) + K(\tau)\left(\frac{\partial h(a(\tau))}{\partial a}\right)' N^{-1}(\tau)(s(\tau) - h(a(\tau)) \qquad (4.38)$$

$$\frac{dK(\tau)}{d\tau} = K(\tau)\left(\frac{\partial f(a(\tau))}{\partial a}\right)' + \left(\frac{\partial f(a(\tau))}{\partial a}\right) K(\tau) + D(\tau)$$
$$+ K(\tau)\left\{\frac{\partial}{\partial a}\left[\left(\frac{\partial h(a(\tau))}{\partial a}\right)' N^{-1}(\tau)(s(\tau) - h(a(\tau))\right]\right\}K(\tau) \qquad (4.39)$$

and (4.35) and (4.37) give the boundary conditions

$$a(t_o) = x_o \quad ; \quad K(t_o) = \Sigma_o \qquad (4.40)$$

Equations 4.38, 4.39 and 4.40 provide us with an approximation to the optimal sequential estimator; the best sequential estimate of $x(\tau)$ is $a(\tau)$, $\tau \geq t_o$.

E. DISCUSSION OF RESULTS

We have derived an approximation to the optimal sequential estimator which is described by

$$\dot{a}(t) = f(a(t)) + K(t)\left(\frac{\partial h(a(t))}{\partial a}\right)' N^{-1}(t)(s(t) - h(a(t)) \qquad (4.41)$$

$$\dot{K}(t) = K(t)\left(\frac{\partial f(a(t))}{\partial a}\right)' + \left(\frac{\partial f(a(t))}{\partial a}\right) K(t) + D(t)$$
$$+K(t)\left\{\frac{\partial}{\partial a}\left[\left(\frac{\partial h(a(t)}{\partial a}\right)' N^{-1}(t)(s(t) - h(a(t))\right]\right\}K(t) \qquad (4.42)$$

with initial conditions

$$a(t_o) = x_o \quad ; \quad K(t_o) = \Sigma_o \qquad (4.43)$$

where $a(t)$ is the estimate of $x(t)$ at time t. The results obtained here agree with those obtained by Snyder.[27] The matrix $K(t)$ agrees with the error covariance matrix in Snyder's work. We see that Eqs. 4.41 and 4.42 are coupled and dependent on the observation $s(\tau)$, $t_o \leq \tau \leq t$. If the system is linear, i.e., if

$$f(x(t)) = A(t)\,x(t) \;\; ; \;\; h(x(t)) = W(t)\,x(t) \tag{4.44}$$

then (4.41), (4.42) and (4.43) reduce to

$$\dot{a}(t) = A(t)a(t) + K(t)W'(t)N^{-1}(t)(s(t) - W(t)a(t)) \tag{4.45}$$

$$\dot{K}(t) = K(t)A'(t) + A(t)K(t) - K(t)W'(t)N^{-1}(t)W(t)K(t) + D(t) \tag{4.46}$$

with initial conditions

$$a(t_o) = x_o \;\; ; \;\; K(t_o) = \Sigma_o \tag{4.47}$$

which we recognize to be the Kalman-Bucy filter. In fact, if $f(x(t))$ and $h(x(t))$ are linear and are given by (4.44), we have

$$\frac{\partial H(\tau, r, c)}{2r} = \left(\frac{\partial A(\tau)r(c,\tau)}{2r}\right)' c + 2\left(\frac{\partial W(\tau)r(c,\tau)}{\partial r}\right)' N^{-1}(\tau)(W(\tau)r(c,\tau) - s(\tau))$$

$$= A'(\tau)\,c + 2W'(\tau)N^{-1}(\tau)(W(\tau)r(c,\tau) - s(\tau)) \tag{4.48}$$

$$\frac{\partial H(\tau, r, c)}{\partial c} = A(\tau)\,r(c,\tau) - \frac{1}{2}D(\tau)c \tag{4.49}$$

Let

$$r(c,\tau) = a(\tau) - \frac{1}{2}K(\tau)\,c \tag{4.50}$$

where $a(\tau)$, $K(\tau)$ satisfy Eqs. 4.45, 4.46 and 4.47; then, by direct substitution, we see that $r(c,\tau)$, expressed by (4.50), satisfies the partial differential equation (4.25) and the initial conditions (4.33). This implies that in the linear case the Kalman-Bucy filter represents the exact solution of the optimum sequential estimator.

Equation 4.36 tells us that at any instant of time, the state of the filter (Eq. 4.41) is an approximation of the best estimate of the state of the nonlinear dynamical system. We see from Eqs. 4.41 and 4.42 that the filter constructs $a(t)$ based on all past observation

$s(\tau)$, $t_o \leq \tau \leq t$. The driven term of the filter is dependent on two terms, $K(t)$ and the effect of the observation noise. When $N(t)$ is large, the driven term of Eq. 4.41 becomes small, which implies that the filter pays less attention to the observation; when $N(t)$ is small, then the observation term is taken more into consideration. Comparing with Snyder's work, we see that matrix $K(t)$ is the error covariance matrix, thus if $K(t)$ is small (small estimation error) we can weigh the observation less. If $K(t)$ is large, we need more observations to decrease the estimation error.

F. DISCUSSION OF OUR APPROACH

The approach that we used in this chapter differs from the conventional one. Instead of looking at a stochastic state estimation problem, we view it as a decision problem under uncertainty. The space of uncertainty (the disturbance noise $\xi(t)$ and observation noise $v(t)$) has σ-algebra structure, but we formulate the problem as a deterministic control problem where the nature of uncertainty is reflected in the cost functional (4.12). One of the obvious advantages of this approach is that we can avoid the Ito calculus and the interpreation of a stochastic differential equation. In this approach, we also bypass the calculation of the conditional mean of the random state vector of the system. Another advantage of this approach is that we have avoided the difficulties of approximating a random variable and the consideration of convergence of stochastic integrals.[19,29] With an assumption on the analyticity of the function $r(c, \tau)$, we can use the technique of Taylor series expansion to approximate the exact solution. Instead of worrying about the mathematical consistency, which often arises in the stochastic approach, we can concentrate on the physical dynamical structure of the suboptimal sequential estimator.

The disadvantage of this approach is the lack of a performance index which will give us the performance level. By comparing our result to Snyder's work, we see that the matrix $K(t)$ is the error covariance matrix; however, we have not shown mathematically that the matrix $K(t)$ is indeed the error covariance matrix.

We can also view Σ_o^{-1}, $N^{-1}(t)$, and $D^{-1}(t)$ in the cost functional as weighting matrices. With this viewpoint, the same formulation and calculating procedure can be applied to cases where there is no σ-algebra structure on the space of uncertainty; the matrices Σ_o^{-1}, $N^{-1}(t)$, $D^{-1}(t)$ indicate the partial information of the uncertainty. Thus, in cases where we may not have the statistics of uncertainty, we can choose $N^{-1}(t)$, $D^{-1}(t)$ and Σ_o^{-1} which reflect our intuitive knowledge about the uncertainty; one constraint in our choice is that these matrices must be positive definite and symmetric for they must correspond to covariance matrices. If there is no disturbance noise, we can conveniently choose $D^{-1}(t)$ equal to the identity matrix.

G. FURTHER RESEARCH

In this chapter, we approximated the best estimator up to the first order effect of the nonlinear system. For systems which are not very nonlinear, this approximation yields satisfactory results,[27] but for a system which is very nonlinear, the higher order effects may be of importance.[19] If we approximate the solution $r(c, \tau)$ of Eq. 4.25 by

$$r_i(c,\tau) = a_i(\tau) + \sum_j b_{ij}(\tau)c_j + \sum_{j,k} \ell^i_{jk}(\tau)c_jc_k \tag{4.51}$$

then we will obtain an approximation of the filter which will include second-order effects. In the next chapter, we shall derive this second-order approximation for a simple quadratic drag system. Additional research is undergoing in trying to establish results for more general nonlinear systems.

It was pointed out in the last section that the matrix $K(t)$ seems to correspond to the estimation error covariance matrix. It would be of importance if we could interpret the significance of the matrix $K(t)$, for this would interrelate the stochastic approach and our approach.

We can also solve the lag-filtering problem in a similar manner. Instead of finding out the differential equation for $r_T(\tau)$, we find the differential equation for $r_T(\tau - \alpha)$, where α is a fixed positive real number; it seems that we can employ the same technique for the lag-filtering problem.

CHAPTER V

NONLINEAR ESTIMATION OF A FIRST ORDER QUADRATIC DRAG DYNAMICAL SYSTEM

In Chapter IV, we approximated the optimum sequential estimator by considering only the zero and first order effects of the dynamical system. If we approximate the solution for $r(c, \tau)$ (Eq. 4.25) as a nth-degree polynomial in c, where $n \geq 2$, the suboptimal filter would be affected by the second and higher order terms arising in the linearization of the original dynamical system. We can say that the system is not "very nonlinear" when we can ignore terms higher than the first order. To estimate the states of a not very nonlinear plant, the approximated filter given in Chapter IV (Eqs. 4.41, 4.42 and 4.43) should give good performance. But for many systems of interest, we cannot ignore the second order effects. In this chapter, we shall work out a simple example to investigate the effect of the second order terms to the structure of the suboptimal filter. We shall not present the most general case because of the resultant complicated equations as they may tend to obscure the applicability of the approximation method. Throughout this example, we shall examine any advantages of our approach as compared to the conventional stochastic approach.

In Section A, we state the problem; in Section B, we derive a suboptimal filter by considering only the zero and first order effects; in Section C, we shall obtain an approximation to the filter by considering the zero, first and second order effects; in Section D, we compare and discuss the differences between the two derived suboptimal filters.

A. STATE ESTIMATION OF A QUADRATIC DRAG DYNAMICAL SYSTEM

Consider a plant (mass subject to quadratic drag)

$$\dot{x}(t) = -hx^2(t) \; ; \quad x(t) > 0 \; ; \quad h > 0 \tag{5.1}$$

where $x(t)$ (the velocity) is a scalar. We shall assume that the initial value of $x(t)$, $x(t_o)$, is unknown, but that we know the statistics of $x(t_o)$

$$E[x(t_o)] = x_o \; ; \quad E[(x(t_o)-x)^2] = g \tag{5.2}$$

assume that we observe the scalar quantity $s(t)$

$$s(t) = x(t) + v(t) \tag{5.3}$$

where $v(t)$ is Gaussian white noise with statistics

$$E[v(t)] = 0 \; ; \quad E[v(t)v(\tau)] = q(t)\ \delta(t-\tau) \tag{5.4}$$

We wish to estimate the state of the system (5.1) at time t based on the observation $s(t')$, $t_o \leq t' \leq t$.

We now convert this problem into a control problem: (See Chapter IV)

Given

$$\dot{w}(t) = -hw^2(t) + u(t) \tag{5.5}$$

$$w(t_o) = w_o \tag{5.6}$$

and the cost functional

$$J = (w_o - x_o)^2 g^{-1} + \int_{t_o}^{t_1} u^2(t) + (\delta(t) - x(t))^2 q^{-1}(t)\ dt \tag{5.7}$$

We are to choose $u(t)$ so that the cost functional (5.7) is minimized subject to Eqs. 5.5 and 5.6.

B. APPROXIMATION OF OPTIMUM NONLINEAR FILTER: I

To get the best interval estimation after observing $s(t)$, $t_o \leq t \leq t_1$, we have to solve the canonical equations (see Chapter IV)

$$\begin{aligned} \dot{w}^*(t) &= -hw^{*2}(t) - \frac{1}{2}p^*(t) \\ \dot{p}^*(t) &= 2hw^*(t) + (s(t) - w^*(t))q^{-1}(t) \end{aligned} \tag{5.8}$$

with boundary conditions

$$p^*(t_1) = 0 \; ; \; p(t_o) = -2g^{-1}(w_o - x_o) \; ; \; w(t_o) = w_o \tag{5.9}$$

Using the same notations as in Chapter IV, we let

$$w^*(\tau) = r(c,\tau) \quad \text{when} \quad p(\tau) = c \tag{5.10}$$

To get the structure of the optimum sequential estimator, we shall have to solve for $r(c,\tau)$ which satisfies the partial differential equation:†

$$\frac{\partial r(c,\tau)}{\partial \tau} - \frac{\partial r(c,\tau)}{\partial c}\frac{\partial H(\tau,r,c)}{\partial r} = \frac{\partial H(\tau,r,c)}{\partial c} \tag{5.11}$$

where $H(\tau, r, c)$ is the Hamiltonian

$$H(\tau,r,c) = -hr^2(c,\tau)c - \frac{1}{4}c^2 + (s(\tau) - r(c,\tau))^2 q^{-1}(\tau) \tag{5.12}$$

and the solution we are interested in is $r(0,\tau)$.

If we approximate the solution $r(c,\tau)$ of (5.11) by a first degree polynomial in c,

$$r(c,\tau) = a(\tau) + b(\tau)\, c \tag{5.13}$$

---

†See Chapter IV.

then we have the suboptimal filter described by the following equations: (See Eqs. 4.41, 4.42 and 4.43)

$$\dot{a}(t) = -ha^2(t) + k(t)(s(t)-a(t))q^{-1}(t) \tag{5.14}$$

$$\dot{k}(t) = -4ha(t)k(t) - k^2(t)q^{-1}(t) + 1 \tag{5.15}$$

with boundary conditions:

$$a(t_o) = x_o \ ; \ k(t_o) = g \tag{5.16}$$

The scalar $a(t)$ is the sequentially estimated state of the plant (5.1). (See Fig. 4)

C. APPROXIMATION OF OPTIMUM NONLINEAR FILTER: II

We shall now find an approximation of the optimum sequential estimator which will include the second order effect of the plant (5.1). Let us approximate the solution $r(c,\tau)$ of (5.11) by a second degree polynomial in c,

$$r(c,\tau) = a(\tau) + b_1(\tau)\, c + b_2(\tau)c^2 \tag{5.17}$$

Differentiating the Hamiltonian (4.12) and using (5.17), we have

$$\frac{\partial H}{\partial r}(\tau, r, c) = -2[s(\tau)-a(\tau)]q^{-1}(\tau) - 2[ha(\tau)-b_1(\tau)q^{-1}(\tau)]c$$

$$-2[hb_1(\tau) - b_2(\tau)q^{-1}(\tau)]c^2 + \circ(c^2) \tag{5.18}$$

$$\frac{\partial H}{\partial c}(\tau, 4, c) = -ha^2(\tau) - [2ha(\tau)b_1(\tau) + \tfrac{1}{2}]c - h[b_1^2(\tau) + 2a(\tau)b_2(\tau)]c^2$$

$$+\circ(c^3) \tag{5.19}$$

Substituting (5.17), (5.18) and (5.19) into (5.11), and by equating the coefficients of the zero, first and second order terms of c we obtain

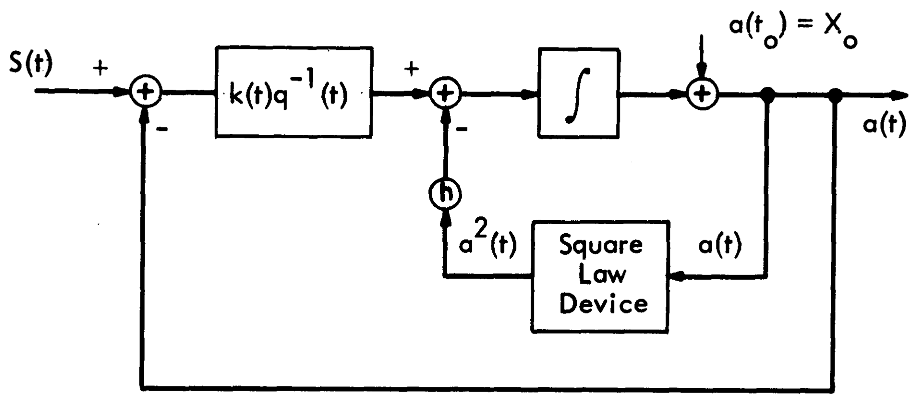

Fig. 4 Approximation of Optimum Nonlinear Filter of a First Order Quadratic Drag Dynamical System

Filter I: The gain k(t) satisfies

$$\dot{k}(t) = -4ha(t)k(t) - k^2(t)q^{-1}(t) + 1 \; , \quad k(t_0) = g$$

Filter II: The gain k(t) satisfies

$$\dot{k}(t) = -4ha(t) - k^2(t)q^{-1}(t) + 1 + 2\ell(t)(s(t) - a(t))q^{-1}(t)$$

$$\dot{\ell}(t) = -3hk^2(t) - 6ha(t)\ell(t) - 3\ell(t)k(t)q^{-1}(t)$$

$$k(t_0) = g \; ; \quad \ell(t_0) = 0$$

$$\frac{da(\tau)}{d\tau} = -ha^2(\tau) - 2b_1(\tau)(s(\tau) - a(\tau))\, q^{-1}(\tau) \tag{5.20}$$

$$\frac{db_1(\tau)}{d\tau} = -4ha(\tau)b_1(\tau) + 2b_1^2(\tau)q^{-1}(\tau) - \frac{1}{2} - 4b_2(\tau)(s(\tau) - a(\tau)q^{-1}(\tau) \tag{5.21}$$

$$\frac{db_2(\tau)}{d\tau} = -3hb_1^2(\tau) + 6b_1(\tau)b_2(\tau)q^{-1}(\tau) - 6ha(\tau)b_2(\tau) \tag{5.22}$$

Equations 5.9, 5.10 and 5.17 imply that

$$a(t_o) + b_1(t_o)[-2g^{-1}(w_o - x_o)] + b_2(t_o)[-2g^{-1}(w_o - x_o)]^2 = w_o \tag{5.23}$$

If we set

$$a(t_o) = x_o\ ,\quad b_1(t_o) = -\frac{1}{2}g\ ,\quad b_2(t_o) = 0 \tag{5.24}$$

we see that (5.23) is satisfied. The solution we are interested in is $r(0,\tau)$, and by (5.17), we have

$$r(0,\tau) = a(\tau) \tag{5.25}$$

Thus $a(\tau)$ is the best estimate of the state at time τ, $\tau \geq t_o$.

Define

$$k(t) = -2b_1(t)\ ,\quad \ell(t) = 4b_2(t) \tag{5.26}$$

We have the suboptimal filter described by the equations:

$$\dot{a}(t) = -ha^2(t) + k(t)(s(t) - a(t))q^{-1}(t) \tag{5.27}$$

$$\dot{k}(t) = -4ha(t) - k^2(t)q^{-1}(t) + 1 + 2\ell(t)(s(t) - a(t))q^{-1}(t) \tag{5.28}$$

$$\dot{\ell}(t) = -3hk^2(t) - 6ha(t)\ell(t) - 3\ell(t)k(t)q^{-1}(t) \tag{5.29}$$

with initial conditions

$$a(t_o) = x_o \ ; \ k(t_o) = g \ ; \ \ell(t_o) = 0 \qquad (5.30)$$

D. DISCUSSION OF RESULTS

We obtained two approximations of the otpimum sequential estimator of the state of the plant (5.1). One includes the effects of the zero and first order terms of the linearization of the plant (5.1), and is described by (See Eqs. 5.14, 5.15 and 5.16)

$$\dot{a}(t) = -ha^2(t) + k(t)(s(5) - a(t))q^{-1}(t) \qquad (5.31)$$

$$\dot{k}(t) = -4ha(t)k(t) - k^2(t)q^{-1}(t) + 1 \qquad (5.32)$$

with initial conditions:

$$a(t_o) = x_o \ ; \ k(t_o) = g \qquad (5.33)$$

The other approximation includes the effects of the zero, first and second order terms of linearization of the plant (5.1), and is described by (see Eqs. 5.27, 5.28, 5.29 and 5.30)

$$\dot{a}(t) = -ha^2(t) + k(t)(s(t) - a(t))\, q^{-1}(t) \qquad (5.34)$$

$$\dot{k}(t) = -4ha(t) - k^2(t)q^{-1}(t) + 1 + 2\ell(t)(s(t) - a(t))q^{-1}(t) \qquad (5.35)$$

$$\dot{\ell}(t) = -3hk^2(t) - 6ha(t)\ell(t) - 3\ell(t)k(t)q^{-1}(t) \qquad (5.36)$$

with initial conditions

$$a(t_o) = x_o \ ; \ k(t_o) = g \ ; \ \ell(t_o) = 0 \qquad (5.37)$$

Comparing (5.31) and (5.34), we notice that the structural form of the two approximations of optimum filter is the same; the only difference is the instantaneous value of the forward gain $k(t)$, which in the first case satisfies (5.32) and (5.33) while in the second case satisfies (5.35) (5.36) and (5.37). Equations 5.32 and 5.35 differ by a term $2\ell(t)(s(t)-a(t))q^{-1}(t)$. The term $2\ell(t)(s(t)-a(t))q^{-1}(t)$ depends on the effect of observation noise and the value $\ell(t)$. From (5.36) and (5.37) we see that if h is very small, $\ell(t)$ will stay near zero; if h is so large that the term $-3hk^2(t)$ is dominant then $\ell(t)$ will not stay near zero. Thus, if the system is quite nonlinear, i.e., h is large, we would expect that the filter II would give a better performance than filter I.

From this example we also see how perturbation method can be applied to obtain a more accurate approximation of the optimum sequential filter. Difficulties regarding stochastic convergence are avoided.

E. FURTHER RESEARCH

A research that must be undertaken should deal with the simulation of the two filters obtained in Section B and C. A comparison of the results may provide insight regarding the effects of the correction term. We would also work out and simulate some simple examples of second order plants, e.g., Van Der Pol equation, and compare with the results obtained by the stochastic approach (Ref. 19).

CHAPTER VI

CONCLUSIONS

In this thesis, it has been demonstrated that one can view certain engineering design problems as decision problems. The main mathematical tool--the matrix minimum principle--was developed in Chapter II. It is an extension of the Pontryagin's maximum (or minimum) principle. In this manner, we formulate and solve the noisy state regulator problem in Chapter III, and a nonlinear filtering problem in Chapter IV. The results in these chapters agree with those obtained by the stochastic approach. In Chapter V, we solve a nonlinear filtering problem for a simple quadratic drag plant. The results obtained differ from those derived using the conventional stochastic approach.

The approach used in this thesis can be extended to tackle general design problems with uncertainty when the statistics of the uncertainty may or may not be available. It is the opinion of the author that this approach will be of significance in the design of adaptive control problems.

APPENDIX 1

MODEL OF CONTINUOUS WHITE NOISE

In an engineering context, the notion of a white noise process is often used without any mathematical rigor. Because of its fictitious nature, we sometimes find it difficult to justify rigorously some of the statements about white noise. In this appendix we shall define white noise process as a sample limit of a sequence of sets of independent random variables.

Let $\{x(n\Delta):n=0, 1, \ldots N\}$ be a sequence of random variables; we shall define the representation of the sequence by

$$x_{\Delta, N}(t) = x(n\Delta) \quad \text{when} \quad n\Delta \leq t < (n+1)\Delta \tag{A.1.1}$$

$x_{\Delta, N}(t)$ is defined on $t\epsilon[0, N\Delta]$. We can now define the notion of sample convergence and sample limit.

Definition A.1.1: Let $\{(x(n\Delta_i): n=1, \ldots, N_i) \,|\, \Delta_i N_i=T\}$ $i=1, 2, \ldots$ be a sequence of sets of random variables. We say that the sequence is sample convergent if there exists a random process $x(t)$ where t is defined in $[0, T]$, such that

$$\lim_{\substack{\Delta_i \to 0 \\ N_i \to \infty}} P[\,|x_{\Delta_i N_i}(t) - x(t)| < \epsilon] \xrightarrow{a.e.} 1 \ ; \epsilon > 0, \ t\epsilon[0, T] \tag{A.1.2}$$

We shall call $x(t)$ the sample limit of $\{x(n\Delta_i):n=1, \ldots N_i \,|\, \Delta_i N_i=T\}$.

The motivation for this definition arises by considering the time-sampling of a random process. By virtue of the separability of a random process,† we can describe a random process by a sequence of countable sets of random variables.

---

†See Ref. 21 and 11.

We shall now define a white noise process:

Definition A.1.2: A random process $\{x(t): t\epsilon[0, T]\}$ is called a white noise process if there exists a sample sequence $\{(x(n\Delta_i): n=1, \ldots, N_i \mid \Delta_i N_i = T\}$ $i=1, 2, \ldots, \infty$ with properties

1. $$E[x(n\Delta_i] = m(n\Delta_i) \quad i=1, \ldots \infty \tag{A.1.3}$$

2. $$\mathrm{Cov}[x(n\Delta_i)\, x\,(m\Delta_i)] = \Delta_i^{-1} Q(n\Delta_i)\, \delta_{nm}\delta_{ij}{}^{\dagger} \quad i, j=1, \ldots \infty \tag{A.1.4}$$

such that $x(t)$ is the sample limit of the sample sequence.

If we let $n\Delta_i \to t$, we see that we have the properties of white noise process:

$$E(x(t)) = \lim_{\substack{\Delta_i \to 0 \\ N_i \to \infty}} E[x_{\Delta_i, N_i}(t)] = m(t) \tag{A.1.5}$$

$$\mathrm{Cov}[x(t)x(\tau)] = \lim_{\substack{\Delta_i \to 0 \\ N_i \to \infty}} \mathrm{Cov}[x_{\Delta_i N_i}(t)\, x_{\Delta_i N_i}(\tau)]$$

$$= Q(t)\delta(t-\tau) \tag{A.1.6}$$

We can extend this definition to a vector white noise process $x(t)$ with the properties

$$E[x(t)] = m(t) \tag{A.1.7}$$

$$\mathrm{Cov}[x(t)x'(\tau)] = Q(t)\delta(t-\tau) \tag{A.1.8}$$

where $m(t)$ is an n-vector

$Q(t)$ is an $n \times n$ positive symmetric matrix

---

† δ_{ij} is the Kronecker deta: $\delta_{ij}=1$ for $i=j$, $\delta_{ij}=0$ if $i \neq j$

Definition A.1.3: A random process $\{x(t): t \in [0, T]\}$ is called a Gaussian white noise process if

1. $\{x(t): t \in [0, T]\}$ is a white noise process

2. The sample sequence $\{x(n\Delta_i) : n=1, \ldots N_i \mid \Delta_i N_i = T\}$ $i=1, \ldots \infty$ is a sequence of sets of Gaussian random variable.

APPENDIX 2

LINEAR SYSTEMS DRIVEN BY WHITE NOISE PROCESSES

Consider a linear system

$$\dot{x}(t) = F(t)x(t) + G(t)\, w(t) \tag{A.2.1}$$

where $x(t)$ is an n-vector

$F(t)$ is an nxn matrice

$G(t)$ is an nxr matrix

$w(t)$ is an r-vector white noise process with

$$E[w(t)] = m(t) \tag{A.2.2}$$

$$E[(w(t)-m(t))(w(\tau)-m(\tau))'] = Q(t)\,\delta(t-\tau) \tag{A.2.3}$$

We shall interpret the process $x(t)$ as a sample limit of $\{(x(n\Delta_i) : n=0, \dots, N_i) \mid \Delta_i N_i = T\}$ $i=1, \dots \infty$ where $x(n\Delta_i)$ satisfies

$$x(n\Delta_i + \Delta_i) - x(n\Delta_i) = \Delta_i F(n\Delta_i)x(n\Delta_i) + G(n\Delta_i)w(n\Delta_i)\Delta_i \tag{A.2.4}$$

where $\{(w(n\Delta_i) : n=0, \dots, N_i) \mid \Delta_i N_i = T\}$ $i=1, \dots \infty$ is the sample sequence of $w(t)$. We shall assume that

$$E[x(0)] = x_o\, ; \quad E[(x(0) - x_o)(x(0) - x_o)'] = \Sigma_o \tag{A.2.5}$$

$$E[(w(n\Delta_i) - m(n\Delta_i))(x(0) - x_o)'] = 0 \tag{A.2.6}$$

By induction it is easily seen that

$$E[(w(n\Delta_i) - m(n\Delta_i))(x(\ell\Delta_i) - m_x(\ell\Delta_i)'] = 0 \tag{A.2.7}$$

$$\ell = 1, \dots N_i$$

where

$$m_x(\ell\Delta_i) \stackrel{\Delta}{=} E[x(\ell\Delta_i)] \tag{A.2.8}$$

Taking the expectation on both sides of (A.2.4) and taking sample limit, we have

$$\dot{m}_x(t) = F(t)\, m_x(t) + G(t)m(t) \tag{A.2.9}$$

Equations A.2.4, A.2.7, A.2.9 and A.1.4) imply that

$$E[(x(n\Delta_i+\Delta_i) - m_x(n\Delta_i+\Delta_i))(x(n\Delta_i+\Delta_i) - m_x(n\Delta_i+\Delta_i)'{]}$$

$$=E[(x(n\Delta_i)-m_x(n\Delta_i))(x(n\Delta_i)-m_x(n\Delta_i))']+G(n\Delta_i)Q(n\Delta_i)G'(n\Delta_i)\Delta_i$$

$$+F(n\Delta_i)E[(x(n\Delta_i)-m_x(n\Delta_i))(x(n\Delta_i)-m_x(n\Delta_i)'{]}\,\Delta_i$$

$$+E[(x(n\Delta_i)-m_x(n\Delta_i))(x(n\Delta_i)-m_x(n\Delta_i))']\,F'(n\Delta_i)\Delta_i \tag{A.2.10}$$

Let us define

$$\Sigma_x(t) \stackrel{\Delta}{=} E[(x(t)-m_x(t))(x(t)-m_x(t))'] \tag{A.2.11}$$

Divide both sides of (A.2.10) by Δ_i and taking sample limit

$\Delta_i \rightarrow 0$, $N_i \rightarrow \infty$, $n\Delta_i \rightarrow t$, then (A.2.10) implies

$$\dot{\Sigma}_x(t) = F(t)\,\Sigma_x(t) + \Sigma_x(t)\,F'(t) + G(t)\,Q(t)\,G'(t) \tag{A.2.12}$$

APPENDIX 3

ESTIMATION OF GAUSSIAN WHITE NOISE PROCESSES

Let $\xi(t)$ and $v(t)$ be two independent scalar[†] Gaussian white noise processes with properties:

1. $$E[\xi(t)] = E[v(t)] = 0 \qquad (A.3.1)$$

2. $$E[\xi(t)\xi(\tau)] = D(t)\delta(t-\tau) \qquad (A.3.2)$$

$$E[v(t)v(\tau)] = N(t)\ \delta(t-\tau) \qquad (A.3.3)$$

$$E[v(t)\xi(\tau)] = 0 \qquad (A.3.4)$$

We wish to find the best estimate of $\xi(t)$ and $v(t)$ in the following sense: we pick real functions $u(t)$ and $\hat{v}(t)$, $t\epsilon[0,T]$, from the set of all Borel measurable functions defined on the probability space such that

$$\Pr[\,|\xi(t)-u(t)| < \epsilon_1\ ;\ |v(t)-\hat{v}(t)| < \epsilon_2,\ t\epsilon[0,T]\,] = a \qquad (A.3.5)$$

will be as large as possible for fixed ϵ_1, ϵ_2. Essentially this means that we pick $u(t)$, $\hat{v}(t)$, $t\epsilon[0,T]$, so that the actual sample functions, $\xi(t)$ and $v(t)$, will have a high probability to be "near" the best estimated functions, $u(t)$ and $\hat{v}(t)$.

We shall assume that the functions, $u(t)$ and $\hat{v}(t)$, we can select are related by

$$\hat{v}(t) = g(u(t))\ ,\ 0 \le t \le T \qquad (A.3.6)$$

[†] For simplicity we let the processes with scalar quantity, the same argument holds for vector processes.

With these assumptions, the problem of selecting the best estimate becomes nontrivial; this is the case in many physical problems, in which we can only choose $u(t)$ freely while $v(t)$ is related to $u(t)$ by some functional.† Thus the problem reduces to selecting $u(t)$ so that (A.3.5) is maximized. We would like to find the best $u(t)$ using all the available information about the processes. To avoid the difficulty of discussing the estimation of white noise process, we shall represent the processes, $\xi(t)$ and $v(t)$, by their corresponding sample sequences, and find the best estimate for the sample sequences. The best estimate of a white noise process is then viewed as the sample limit of the best estimated sample sequence.

Let $\{(\xi(n\Delta_i) : n=0, \ldots, N_i)(N_i\Delta_i=T\}$, $\{(v(n\Delta_i) : n=0, \ldots, N_i \mid N_i\Delta_i=T\}$ be the sample sequences for $\xi(t)$ and $v(t)$ respectively. Let us select a sequence of sets of real numbers $\{(u(n\Delta_i) : n=0, \ldots N_i \mid N_i\Delta_i=T\}$ $i=1, \ldots \infty$, and

$$\hat{v}(n\Delta_i) = g(u(n\Delta_i)) \tag{A.3.7}$$

With these choices, let us define a_i by

$$a_i=\Pr[\ |\xi(n\Delta_i)-u(n\Delta_i)|<\epsilon;\ \ |v(n\Delta_i)-\hat{v}(n\Delta_i)|<\epsilon, n=0, \ldots N_i] \tag{A.3.8}$$

First we note that a_i is always finite for $i=1, \ldots \infty$. Thus the mathematical problem of maximizing (A.3.8), subject to (A.3.7), is meaningful. Suppose that for each $i=1, \ldots \infty$, we find the real sequence $\{(u^*(n\Delta_i) : n=1, \ldots, N_i) \mid \Delta_i N_i=T\}$ which maximizes (A.3.8) subject to (A.3.7); we have the sequence of functions $\{u_{\Delta_i, N_i}(t)\}$ $i=1. \ldots \infty$ defined by (A.1.1). If the sample limit exists, then $u^*(t)$ will maximize $a = \lim\limits_{\substack{\Delta_i \to 0 \\ N_i \to \infty}} a_i$ where

†See Chapter IV, Section A.

$$u^*(t) = \lim_{\substack{\Delta_i \to 0 \\ N_i \to \infty}} u^*_{\Delta_i, N_i}(t) \tag{A.3.8}$$

and $\hat{v}^*(t)$ is the best estimate for $v(t)$ where

$$\hat{v}^*(t) = \lim_{\substack{\Delta_i \to 0 \\ N_i \to 0}} g(u^*_{\Delta_i, N_i}(t)) \tag{A.3.9}$$

By the assumptions of the processes (A.3.1), (A.3.2), (A.3.3) and (A.3.4), we have the expression for a_i:

$$a_i = \int_{-\epsilon}^{\epsilon} K \exp\{-\sum_{n=0}^{N_i} [\,\xi(n\Delta_i) - u(n\Delta_i))^2 D^{-1}(n\Delta_i) \dotplus (v(n\Delta_i)$$

$$-\hat{v}(n\Delta_i))^2 N(n\Delta_i)]\,\Delta_i\} d\xi(0) \ldots d\xi(N_i\Delta_i) \ldots$$

$$\ldots dv(N_i\Delta_i) \tag{A.3.10}$$

where K is the normalizing constant. The geometrical interpretation of (A.3.10) is that a_i is equal to the area under the strip around zero of a multidimensional Gaussian density function with nonzero mean.† Thus the problem of maximizing a_i is equivalent to the minimization of the expression

$$J_i = \sum_{n=0}^{N_i} \{u^2(n\Delta_i) D^{-1}(n\Delta_i) + \hat{v}^2(n\Delta_i) N^{-1}(n\Delta_i)\}\Delta_i \tag{A.3.11}$$

---

†See Fig. 5.

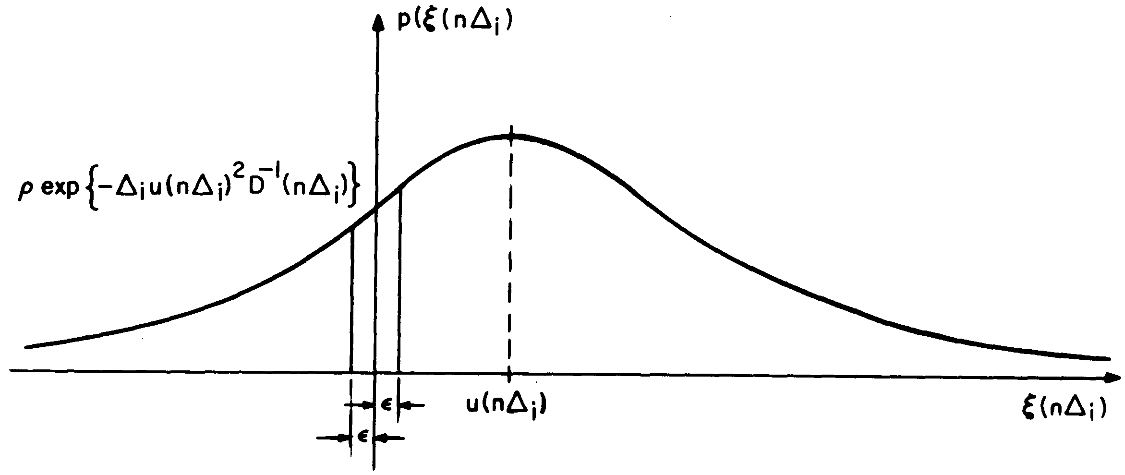

Fig. 5 Projection of a_i onto the $p(\xi(n\Delta_i))$-$\xi(n\Delta_i)$ plane (Eq. A.3.10)

We see that to maximize the shaded area by choosing $u(n\Delta_i)$, it is equivalent to maximize $\rho\ \exp\{-\Delta_i u(n\Delta_i)^2 D^{-1}(n\Delta_i)\}$ where ρ is the normalized constant; and also is equivalent to minimize the value $\Delta_i u(n\Delta_i)^2 D^{-1}(n\Delta_i)$.

Taking the sample limit, we are to choose $u(t)$ so as to minimize the functional

$$J = \lim_{\substack{\Delta_i \to 0 \\ N_i \to \infty}} J_i = \int_0^T (u^2(t)D^{-1}(t) + \hat{v}^2(t)N^{-1}(t))dt \tag{A.3.12}$$

under the constraint

$$\hat{v}(t) = g(u(t)) \tag{A.3.13}$$

If we are to choose $u(t)$ and $\hat{v}(t)$ freely, obviously we should choose $u(t) = \hat{v}(t) = 0$ since $D(t)$ and $N(t)$ are positive real numbers; but with the constraint (A.3.13), the minimization problem becomes nontrivial.

Similar arguments can be carried out for estimating two vector Gaussian white noise processes under a functional constraint

$$\hat{v}(t) = g(u(t)) \tag{A.3.14}$$

where $u(t)$, $v(t)$ are vector Gaussian white noise processes. The functional we want to minimize is

$$J = \int_0^T (u'(t)D^{-1}(t)u(t) + \hat{v}'(t)N^{-1}(t)\hat{v}(t))dt \tag{A.3.15}$$

We may view the cost functional (A.3.15) as a quadratic criteria with the information matrices (i.e., the inverse of covariance matrices) as the weighting matrices. This appendix justifies the physical significance of the functional (A.3.15).

APPENDIX 4

INVARIANT IMBEDDING

Consider the set of vector differential equations

$$\dot{w}(t) = f(t, w, p) \tag{A.4.1}$$

$$\dot{p}(t) = g(t, w, p) \tag{A.4.2}$$

where $w(t)$ and $p(t)$ are n-vectors, subject to the boundary conditions

$$p(t_o) = a \quad ; \quad p(\tau) = c \tag{A.4.3}$$

Let us denote the corresponding solution for $w(t)$ by

$$w_\tau(t) = r(c, t) \quad ; \quad t_o \leq t \leq \tau$$

where the subscript τ is used to denote that the terminal time is τ. We shall now find a partial differential equation which $r(c, \tau)$ must satisfy, for any given $\tau \geq t_o$.

By Eq. A.4.3, we have

$$r(c, \tau) = r(p(\tau), \tau) \tag{A.4.4}$$

Taking total derivatives with respect to τ we obtain

$$\frac{dr(p(\tau), \tau)}{d\tau} = \frac{\partial r(p(\tau), \tau)}{\partial \tau} + \frac{\partial r(p(\tau), \tau)}{\partial p(\tau)} \frac{dp(\tau)}{d\tau} \tag{A.4.5}$$

Substituting (A.4.1) and (A.4.2) into (A.4.5), and using (A.4.3) we find the partial differential equation

$$f(\tau, r, c) = \frac{\partial r}{\partial c} g(\tau, r, c) + \frac{\partial r}{\partial \tau} \tag{A.4.6}$$

For the special case when

$$\dot{w}(t) = \frac{\partial H(t, w, p)}{\partial p} \ ; \ \dot{p}(t) = - \frac{\partial H(t, w, p)}{\partial p} \tag{A.4.7}$$

We see that $r(c, \tau)$ will satisfy the partial differential equation

$$\frac{\partial r}{\partial \tau} - \frac{\partial r}{\partial c} \frac{\partial H(\tau, r, c)}{\partial r} = \frac{\partial H(\tau, r, c)}{\partial c} \tag{A.4.8}$$

The solution $r(c, \tau)$ of (A.4.8) is the terminal value $w_T(\tau)$ expressed as a function of the terminal time $\tau \geq t_o$, and the corresponding terminal condition for $p(\tau)$.

APPENDIX 5

CALCULATIONS

A.5.1 MINIMUM FUNCTIONAL FOR STATE-REGULATOR PROBLEM WITH DRIVING DISTURBANCE OF NONZERO MEAN

We shall show that

$$J^*[\Sigma_x(t), m_x(t), t] = Tr[P_x(t)\Sigma_x(t) + p_x(t)m_x'(t)] + f(t) \tag{A.5.1}$$

where

$$-\dot{f}(t) = -\frac{1}{4}p_x'(t)B'(t)R^{-1}(t)B(t)p_x(t) + Tr[D(t)P_x(t)] + m'(t)p_x(t) \tag{A.5.2}$$

$$f(t_1) = 0$$

will satisfy the Hamilton-Jacobi equation when $K(t)$, $k(t)$, $P_x(t)$, $p_x(t)$ satisfy

$$K(t) = -R^{-1}(t)B'(t)P_x(t) \tag{A.5.3}$$

$$k(t) = -\frac{1}{2}R^{-1}(t)B'(t)p_x(t) \tag{A.5.4}$$

$$-\dot{P}_x(t) = A'(t)P_x(t) + P_x(t)A(t) - P_x(t)B(t)R^{-1}(t)B'(t)P_x(t) + Q(t) \tag{A.5.5}$$

$$-\dot{p}_x(t) = 2P_x(t)m(t) + A'(t)p_x(t) - P_x(t)B(t)R^{-1}(t)B'(t)p_x(t) \tag{A.5.6}$$

$$P_x(t_1) = F \quad ; \quad p_x(t_1) = 0 \tag{A.5.7}$$

Taking partial derivatives of (A.5.1) with respect to t, $\Sigma_x(t)$ and $m_x(t)$ we have

$$\frac{\partial J^*[\Sigma_x(t), m_x(t), t]}{\partial t} = Tr[\dot{p}_x(t)\Sigma_x(t)] + m_x'(t)\dot{p}_x(t) + \dot{f}(t) \tag{A.5.8}$$

$$\frac{\partial J^*[\Sigma_x(t), m_x(t), t]}{\partial \Sigma_x(t)} = P_x(t) \tag{A.5.9}$$

$$\frac{\partial J^*[\Sigma_x(t), m_x(t), t]}{\partial m_x(t)} = p_x(t) \tag{A.5.10}$$

Using (A.5.3), (A.5.4), (A.5.9) and (A.5.10), we have the Hamiltonian†

$$H(\Sigma_x(t), m_x(t), \frac{\partial J^*}{\partial \Sigma_x}, \frac{\partial J^*}{\partial m_x}) = \mathrm{Tr}[\dot{\Sigma}_x(t)P_x(t)+Q(t)\Sigma_x(t)$$

$$+P_x(t)B(t)R^{-1}(t)B'(t)P_x(t)\Sigma_x(t)] + \dot{m}_x'(t)p_x(t)$$

$$+ \frac{1}{4}p_x'(t)B(t)R^{-1}(t)B'(t)p_x(t)+ \frac{1}{2}p_x'(t)B(t)R^{-1}(t)B'(t)m_x(t)$$

$$+ \frac{1}{2}m_x'(t)P_x(t)B(t)R^{-1}(t)B'(t)p_x(t) \tag{A.5.11}$$

Substituting (A.5.3) and (A.5.4) into (3.39) and (3.40) we have

$$\dot{m}_x(t) = A(t)m_x(t) - B(t)R^{-1}(t)B'(t)P_x(t)m_x(t)$$

$$- \frac{1}{2}B(t)R^{-1}(t)B'(t)p_x(t) + m(t) \tag{A.5.12}$$

$$\dot{\Sigma}_x(t) = A(t)\Sigma_x(t) - B(t)R^{-1}(t)B'(t)P_x(t)\Sigma_x(t) + \Sigma_x(t)A'(t)$$

$$-\Sigma_x(t)P_x(t)B(t)R^{-1}(t)B'(t) - \frac{1}{2}m_x(t)p_x'(t)B(t)R^{-1}(t)B'(t) + m_x(t)m'(t)$$

$$-\frac{1}{2}B(t)R^{-1}(t)B'(t)p_x(t)m_x'(t) + m(t)m_x'(t) + D(t) \tag{A.5.13}$$

---

†See Eq. 3.43 .

Substituting (A.5.9) into (A.5.11), Eqs. (A.5.8) and (A.5.11) imply

$$\frac{\partial J^*}{\partial t} + H(\Sigma_x(t), m_x(t), \frac{\partial J^*}{\partial \Sigma_x}, \frac{\partial J^*}{\partial m_x}) = \dot{f}(t) - \frac{1}{4} p_x'(t) B'(t) R^{-1}(t) B(t) p_x(t)$$

$$+ \mathrm{Tr}[D(t)P_x(t)] + m'(t)p_x(t) \qquad (A.5.14)$$

By Eq. A.5.2, the right-hand side is equal to zero; thus, the extremal solution satisfies the Hamilton-Jacobi equation. The boundary condition is also satisfied, for we have

$$J^*[\Sigma_x(t_1), m_x(t_1), t_1] = \mathrm{Tr}[P_x(t_1)\Sigma_x(t_1) + p_x(t_1)m_x'(t_1)] + f(t_1)$$

$$= \mathrm{Tr}[F\Sigma_x(t_1)] \qquad (A.5.15)$$

Therefore, the extremal solution specified by (A.5.3), (A.5.4), (A.5.5) (A.5.6) and (A.5.7) satisfies the sufficient conditions for optimality and thus is the optimal solution.

A.5.2 SOLUTION FOR THE SET OF MATRIX DIFFERENTIAL EQUATIONS OF $\Sigma^*_z(t)$ and $P^*_z(t)$ †

We have the set of matrix differential equations‡

---

†See Chapter III, Section 3.3.2.

‡See Eq. 3.83, 3.88, 3.89, 3.90, 3.91, 3.100 and 3.101.

$$
\frac{d}{dt}\left[\begin{array}{c|c} \Sigma_{x}^{*}(t) & \Sigma_{xw}^{*}(t) \\ \hline \Sigma_{wx}^{*}(t) & \Sigma_{w}^{*}(t) \end{array}\right] = \left[\begin{array}{c|c} A(t) & B(t)M^{*}(t) \\ \hline V^{*}(t)C(t) & \begin{array}{c} A(t)-V^{*}(t)C(t) \\ +B(t)M^{*}(t) \end{array} \end{array}\right] \left[\begin{array}{c|c} \Sigma_{x}^{*}(t) & \Sigma_{xw}^{*}(t) \\ \hline \Sigma_{wx}^{*}(t) & \Sigma_{w}^{*}(t) \end{array}\right]
$$

$$
+\left[\begin{array}{c|c} \Sigma_{x}^{*}(t) & \Sigma_{xw}^{*}(t) \\ \hline \Sigma_{wx}^{*}(t) & \Sigma_{w}^{*}(t) \end{array}\right] \left[\begin{array}{c|c} A'(t) & C'(t)V^{*'}(t) \\ \hline M^{*'}(t)B'(t) & \begin{array}{c} A'(t)-C'(t)V^{*'}(t) \\ +M^{*'}(t)B'(t) \end{array} \end{array}\right]
$$

$$
+\left[\begin{array}{c|c} D(t) & 0 \\ \hline 0 & V^{*}(t)N(t)V^{*'}(t) \end{array}\right] \tag{A.5.16}
$$

$$
-\frac{d}{dt}\left[\begin{array}{c|c} P_{x}^{*}(t) & P_{xw}^{*}(t) \\ \hline P_{wx}^{*}(t) & P_{w}^{*}(t) \end{array}\right] = \left[\begin{array}{c|c} P_{x}^{*}(t) & P_{xw}^{*}(t) \\ \hline P_{wx}^{*}(t) & P_{w}^{*}(t) \end{array}\right] \left[\begin{array}{c|c} A(t) & B(t)M^{*}(t) \\ \hline V^{*}(t)C(t) & \begin{array}{c} A(t)-V^{*}(t)C(t) \\ +B(t)M^{*}(t) \end{array} \end{array}\right]
$$

$$
+\left[\begin{array}{c|c} A'(t) & C'(t)V^{*'}(t) \\ \hline M^{*'}(t)B'(t) & \begin{array}{c} A'(t)-C'(t)V^{*'}(t) \\ +M^{*'}(t)B'(t) \end{array} \end{array}\right] \left[\begin{array}{c|c} P_{x}^{*}(t) & P_{xw}^{*}(t) \\ \hline P_{wx}^{*}(t) & P_{w}^{*}(t) \end{array}\right]
$$

$$
+\left[\begin{array}{c|c} Q(t) & 0 \\ \hline 0 & M^{*'}(t)R(t)M^{*}(t) \end{array}\right] \tag{A.5.17}
$$

$$\left[\begin{array}{c|c} \Sigma_x^*(t_o) & \Sigma_{xw}^*(t_o) \\ \hline \Sigma_{wx}^*(t_o) & \Sigma_w^*(t_o) \end{array}\right] = \left[\begin{array}{c|c} \Sigma_o + x_o x_o' & x_o x_o' \\ \hline x_o x_o' & x_o x_o' \end{array}\right] \tag{A.5.18}$$

$$\left[\begin{array}{c|c} P_x^*(t_1) & P_{xw}^*(t_1) \\ \hline P_{wx}^*(t_1) & P_w^*(t_1) \end{array}\right] = \left[\begin{array}{c|c} F & 0 \\ \hline 0 & 0 \end{array}\right] \tag{A.5.19}$$

where $V^*(t)$ and $M^*(t)$ are given by (see Eq. 3.90, 3.99)

$$V^*(t) = P_w^{*-1}(t)P_{wx}(t)[\Sigma_{xw}^*(t)-\Sigma_x^*(t)]C'(t)N^{-1}(t)+[\Sigma_w^*(t)-\Sigma_{wx}^*(t)]C'(t)N^{-1}(t) \tag{A.5.20}$$

$$M^*(t)=-R^{-1}(t)B'(t)[P_x^*(t)+P_{wx}^*(t)]\,\Sigma_{xw}^*(t)\Sigma_w^{*-1}(t)-R^{-1}(t)B'(t)[P_{xw}^*(t)+P_w^*(t)] \tag{A.5.21}$$

Equations A.5.16 and A.5.17 give us a set of matrix differential equations relating $\Sigma_x^*(t)$, $\Sigma_{xw}^*(t)$, $\Sigma_{wx}^*(t)$, $\Sigma_w^*(t)$, $P_x^*(t)$, $P_{xw}^*(t)$, $P_{wx}^*(t)$, and $P_w(t)$; the boundary conditions are given by (A.5.18) and (A.5.19). Instead of solving explicitly this set of matrix differential equations we shall sow that the solution must be of the form

$$\Sigma_{xw}^* = \Sigma_w^*(t) \ ; \quad P_{xw}^*(t) = -P_w^*(t) \tag{A.5.22}$$

Consider the set of matrix differential equations

$$\dot{Z}_1(t) = A(t)\ Z_1(t) + Z_1(t)A'(t) \tag{A.5.23}$$

$$-\dot{Z}_2(t) = A'(t)\ Z_2(t) + Z_2(t)\ A(t) \tag{A.5.24}$$

with boundary conditions

$$Z_1(t_o) = 0 \quad ; \quad Z_2(t_1) = 0 \tag{A.5.25}$$

Clearly the zero matrix is the solution for $Z_1(t)$ and $Z_2(t)$ which satisfies (A.5.23), (A.5.24) and (A.5.25). We shall now show that if $\Sigma^*_{xw}(t)$, $\Sigma^*_w(t)$, $P^*_{xw}(t)$ and $P^*_w(t)$ satisfy (A.5.22), then the matrix differential equations (A.5.16), (A.5.17) and the boundary conditions (A.5.18), (A.5.19) imply that $(\Sigma^*_w(t) - \Sigma^*_{xw}(t))$ satisfies (A.5.23) and (A.5.25), while $(P^*_w(t) + P^*_{wx}(t))$ satisfies (A.5.24) and (A.5.25). In view of this, by the Uniqueness Theorem, (A.5.22) must be the form of the solution for (A.5.16), (A.5.17), (A.5.18) and (A.5.19).

If (A.5.22) holds, then by direct substitution, (A.5.16), (A.5.17) (A.5.20) and (A.5.21) imply

$$\begin{aligned}\dot{\Sigma}^*_w(t) = {} & A(t)\Sigma^*_w(t) + \Sigma^*_w(t)A'(t) - B(t)R^{-1}(t)B'(t)[P^*_x(t) - P^*_w(t)]\Sigma^*_w(t) \\ & -\Sigma^*_w(t)[P^*_x(t) - P^*_w(t)]B(t)R^{-1}(t)B'(t) \\ & +[\Sigma^*_x(t) - \Sigma^*_w(t)]C'(t)N^{-1}(t)C(t)[\Sigma^*_x(t) - \Sigma^*_w(t)\end{aligned} \tag{A.5.26}$$

$$\begin{aligned}\dot{\Sigma}^*_{xw}(t) = {} & A(t)\Sigma^*_{xw}(t) + \Sigma^*_{xw}(t)A'(t) - B(t)R^{-1}(t)B'(t)[P^*_x(t) - P^*_w(t)]\Sigma^*_w(t) \\ & -\Sigma^*_w(t)[P^*_x(t) - P^*_w(t)]B(t)R^{-1}(t)B'(t) \\ & +[\Sigma^*_x(t) - \Sigma^*_w(t)]C'(t)N^{-1}(t)C(t)[\Sigma^*_x(t) - \Sigma^*_w(t)]\end{aligned} \tag{A.5.27}$$

$$\begin{aligned}-\dot{P}^*_{wx}(t) = {} & A'(t)P^*_{wx}(t) + P^*_{wx}(t)A(t) - [P^*_x(t) - P^*_w(t)]B(t)R^{-1}(t)B'(t)[P^*_x(t) - P^*_w(t)] \\ & +C'(t)N^{-1}(t)C(t)[\Sigma^*_x(t) - \Sigma^*_w(t)]P^*_w(t) + P^*_w(t)[\Sigma^*_x(t) - \Sigma^*_w(t)]C'(t)N^{-1}(t)C(t)\end{aligned} \tag{A.5.28}$$

$$-\dot{P}^*_w(t)=A'(t)P^*_w(t)+P^*_w(t)A(t)+[P^*_x(t)-P^*_w(t)]\,B(t)R^{-1}(t)B(t)]P^*_x(t)-P^*_w(t)]$$

$$-C'(t)N^{-1}(t)C(t)[\Sigma^*_x(t)-\Sigma^*_w(t)]\,P^*_w(t)-P^*_w(t)[\Sigma^*_x(t)-\Sigma^*_w(t)]\,C'(t)N^{-1}(t)C(t) \tag{A.5.29}$$

By inspection, we see that $(\Sigma^*_w(t) - \Sigma^*_{xw}(t))$ satisfies (A.5.23) and $(P^*_{wx}(t) + P^*_w(t))$ satisfies (A.5.24). Equations A.5.18 and A.5.19 give

$$\Sigma^*_w(t_o) - \Sigma^*_{xw}(t_o) = 0 \;\; ; \;\; P^*_w(t_1) + P^*_{xw}(t_1) = 0 \tag{A.5.30}$$

A.5.3 MINIMAL FUNCTIONAL FOR STATE-REGULATOR PROBLEM WITH DRIVING DISTURBANCE AND MEASUREMENT NOISE

We shall show that

$$J^*[\Sigma_z(t),\ t] = \mathrm{Tr}[K_2(t)\Sigma_x(t)] + f(t) \tag{A.5.31}$$

where

$$\left.\begin{aligned} -\dot{f}(t) &= \mathrm{Tr}[K_1(t)K_2(t)B(t)R^{-1}(t)B'(t)K_2(t)+D(t)K_2(t)] \\ f(t_1) &= 0 \end{aligned}\right\} \tag{A.5.32}$$

will satisfy the Hamilton-Jacobi equation when $V(t)$, $M(t)$, $K_1(t)$, $K_2(t)$ satisfy

$$V(t) = K_1(t)C'(t)N^{-1}(t) \tag{A.5.33}$$

$$M(t) = -R^{-1}(t)B'(t)K_2(t) \tag{A.5.34}$$

$$\dot{K}_1(t) = A(t)K_1(t)+K_1(t)A'(t)-K_1(t)C'(t)N^{-1}(t)C(t)K_1(t)+D(t) \tag{A.5.35}$$

$$-\dot{K}_2(t) = A'(t)K_2(t)+K_2(t)A(t)-K_2(t)B(t)R^{-1}(t)B'(t)K_2(t)+Q(t) \tag{A.5.36}$$

$$K_1(t_o) = \Sigma_o \ ; \quad K_2(t_1) = F \tag{A.5.37}$$

Taking partial derivatives of (A.5.31) with respect to t and $\Sigma_z(t)$, we have

$$\frac{\partial J^*[\Sigma_z(t), t]}{\partial t} = \mathrm{Tr}[\dot{K}_2(t)\Sigma_x(t)] + \dot{f}(t) \tag{A.5.38}$$

$$\frac{\partial J^*[\Sigma_z(t), t]}{\partial \Sigma_z(t)} = \left[\begin{array}{c|c} K_2(t) & 0 \\ \hline 0 & 0 \end{array}\right] \tag{A.5.39}$$

Using (A.5.33), (A.5.34), we have the Hamiltonian[†]

$$\begin{aligned} H(\Sigma_z(t), \frac{\partial J^*}{\partial \Sigma_z}) = \mathrm{Tr}[&A(t)\Sigma_x(t)K_2(t) - B(t)R^{-1}(t)B'(t)K_2(t)\Sigma_w(t)K_2(t) \\ &+\Sigma_x(t)A'(t)K_2(t) - \Sigma_w(t)K_2(t)B(t)R^{-1}(t)B'(t)K_2(t) + D(t)K_2(t) \\ &+Q(t)\Sigma_x(t) + K_2(t)B(t)R^{-1}(t)B'(t)K_2(t)\Sigma_w] \end{aligned} \tag{A.5.40}$$

By (A.5.36), we have

$$\begin{aligned} \frac{\partial J^*}{\partial t} + H(\Sigma_z(t), \frac{\partial J^*}{\partial \Sigma_z}) = \mathrm{Tr}[&(\Sigma_x(t) - \Sigma_w(t))K_2(t)B(t)R^{-1}(t)B'(t)K_2 + D(t)K_2(t)] \\ &+\dot{f}(t) \end{aligned} \tag{A.5.41}$$

Substituting (A.5.33) and (A.5.34) into (3.83), we have the matrix differential equations for $\Sigma_x(t)$ and $\Sigma_w(t)$

$$\dot{\Sigma}_x(t) = A(t)\Sigma_x(t) - B(t)R^{-1}(t)B'(t)K_2(t)\Sigma_w(t) + \Sigma_x(t)A'(t) + D(t) \tag{A.5.42}$$

---

† See Eq. 3.87.

$$\dot{\Sigma}_w(t)=A(t)\Sigma_w(t) + \Sigma_w(t)A'(t) - B(t)R^{-1}(t)B'(t)K_2(t)\Sigma_w(t)$$

$$-\Sigma_w(t)K_2(t)B(t)R^{-1}(t)B'(t)+K_1(t)C'(t)N^{-1}(t)C(t)K_1(t) \tag{A.5.43}$$

Subtracting (A.5.43) from (A.5.42), we see that $Z(t) \triangleq (\Sigma_x(t)-\Sigma_w(t))$ satisfies the matrix differential equation

$$\dot{\Sigma}(t)=A(t)Z(t)+Z(t)A'(t)-K_1(t)C'(t)N^{-1}(t)C(t)K_1(t)+D(t) \tag{A.5.44}$$

with boundary condition (see Eq. A.5.18),

$$Z(t_o) = \Sigma_x(t_o) - \Sigma_w(t_o) = \Sigma_o \tag{A.5.45}$$

Define

$$Z_1(t) = Z(t) - K_1(t) \tag{A.5.46}$$

By (A.5.35), (A.5.37) and (A.5.45), $Z_1(t)$ satisfies

$$\dot{Z}_1(t) = A(t)Z_1(t) + Z_1(t)A'(t) \tag{A.5.47}$$

with boundary condition

$$Z_1(t_o) = Z(t_o) - K_1(t_o) = 0 \tag{A.5.48}$$

Since the zero matrix is the solution, we conclude that

$$K_1(t) = Z(t) = \Sigma_x(t) - \Sigma_w(t) \tag{A.5.49}$$

Substituting (A.5.49) into (A.5.41), we have

$$\frac{\partial J^*}{\partial t} + H(\Sigma_z(t), \frac{\partial J^*}{\partial \Sigma_z}) = Tr[K_1(t)K_2(t)B(t)R^{-1}(t)B'(t)K_2(t)+D(t)K_2(t)]$$

$$+\dot{f}(t) \tag{A.5.50}$$

By (A.5.32) we see that the right hand side is equal to zero. We also have the boundary condition for $J^*[\Sigma_z(t), t]$

$$J^*[\Sigma_z(t_1), t_1] = Tr[K_2(t_1)\Sigma_x(t_1)] + f(t_1) = Tr[F\Sigma_x(t_1)] \tag{A.5.51}$$

Thus the extremal solution satisfies the sufficient conditions for optimality and is, therefore, optimal.

APPENDIX 6

GRADIENT MATRICES

Let X be an $n \times n$ matrice with element x_{ij}. Let $f(\cdot)$ be a scalar-valued function of the x_{ij}

$$f(X) \stackrel{\Delta}{=} f(x_{11}, x_{12}, \ldots, x1_n, x_{21}, \ldots, x_{nn}) \tag{A.6.1}$$

The gradient matrix of $f(X)$ with respect to X is defined to be the matrix

$$\frac{\partial f(X)}{\partial X} \stackrel{\Delta}{=} \left(\frac{\partial f(X)}{\partial x_{ij}}\right) \tag{A.6.2}$$

The trace of a nxn matrix X is denoted by

$$\mathrm{Tr}[X] = \sum_{i=1}^{n} x_{ii} \tag{A.6.3}$$

Straight forward computations yield the following formulae†

$$\frac{\partial}{\partial X} \mathrm{Tr}[X] = I \tag{A.6.4}$$

$$\frac{\partial}{\partial X} \mathrm{Tr}[AX] = A' \tag{A.6.5}$$

$$\frac{\partial}{\partial X} \mathrm{Tr}[AX'] = A \tag{A.6.6}$$

$$\frac{\partial}{\partial X} \mathrm{Tr}[AXB] = A'B' \tag{A.6.7}$$

$$\frac{\partial}{\partial X} \mathrm{Tr}[AX'B] = BA \tag{A.6.8}$$

---

†For a more extensive table of gradient matrices, see Ref. 2.

$$\frac{\partial}{\partial X} \mathrm{Tr}[AXBX] = A'X'B' + B'X'A' \tag{A.6.9}$$

$$\frac{\partial}{\partial X} \mathrm{Tr}[AXBX'] = A'XB' + AXB \tag{A.6.10}$$

APPENDIX 7

ASSUMPTIONS ON THE SYSTEM II†

In the state-regulator problem with noisy observation, we make an assumption on the dimension of the system II and also that the mean of the state of system II equals to that of the state of the plant which we are regulating. In this appendix we discuss these assumptions.

In Chapter III, Section C.1, we chose the dimension, n, of $w(t)$ to be equal to the dimension of the plant, and we required that

$$E[w(t)] = E[x(t)] \quad \text{for all } t \tag{A.7.1}$$

Essentially this assumption is equivalent to the assumption of sufficient statistics used in other approaches.‡ In essence, we are viewing the system II as a filter which generates an unbiased estimate of the state, $x(t)$. The mathematical justification of these assumptions is still the subject of current research.

Let us suppose now that the dimension of the control is less than the dimension of the state of the plant, i.e., $r < n$. Thus, the control at any time t is an element of the R_r space, and the state $w(t)$ is an element in R_m, where m is the dimension of the system II. The feedback gain represents a linear transformation from R_m to R_r. Suppose that for optimum performance, the optimum control at any time t will lie in a subset of R_r, then if we let the range space of the linear transformation equal to the whole space of R_r, we can still achieve the optimum performance by appropriately choosing the control $u(t)$ in R_r. For the range space to be equal to R_r, we

---

†Refer to Chapter III, Section C.1.

‡See Ref. 30.

require that $m \geq r$. But, the complexity of the system will increase as the dimension of the state increases; therefore we shall set $m=r$ to ensure that we can still get the best performance. If r is very much smaller than n, then we have decreased the dimension of the system II considerably. Under this assumption, how do we solve the regulator problem in the presence of observation noise? This is still a subject for further research. The answer to this problem will justify (A.7.1)† and the assumption of sufficient statistics.

---

†The unbiasness assumption implies the assumption on the dimension of system II. See Ref. 4.

BIBLIOGRAPHY

1. Athans, M., and Falb, P., Optimal Control: An Introduction to the Theory and Its Application, McGraw-Hill, New York, 1966.

2. Athans, M., and Schweppe, F.C., "Gradient Matrices and Matrix Caclulations," Lincoln Lab. Technical Report, 1965-53.

3. Athans, M. and Schweppe, F. C., "On Optimal Waveform via Control Theoretic Concepts," 1966 JACC, University of Washington, August, 1966.

4. Athans, M. and Tse, E., "A Direct Derivation of the Kalman-Bucy Filter Using the Maximum Principle," Paper ESL-P-300, February, 1967.

5. Baggeroer, A. B., "Maximum Posteriori Interval Estimation," WESCON Paper 7.3; August, 1966.

6. Bellman, R., Introduction to Matrix Analysis, McGraw-Hill Book Company, Inc., New York, 1960.

7. Bryson, A.E. and Frazier, M., "Smoothing for Linear and Nonlinear Dynamic Systems," USAF. Tech. Report ASD-TDR-63-119; February, 1966.

8. Bucy, R. S., "Nonlinear Filtering Theory," IEEE Trans. on Automatic Control, 10, 1965.

9. Cox, H., "On the Estimation of State Variables and Parameters for Noisy System," IEEE Trans. on Automatic Control, 9, 1964.

10. Detchmendy, D. M. and Sridhar, R., "Sequential Estimation of States and Parameters in Noisy Nonlinear Dynamical System," JACC Preprints, 1965, pp. 56-63.

11. Doob, J. L., Stochastic Processes, Wiley and Sons, Inc., New York, 1953.

12. Feldbaum, A.A., Optimal Control Systems, Academic Press, New York, 1965.

13. Florentin, J. J., "Optimal Control of Continuous Time, Markov Stochastic Systems," Journal of Electronic and Control, No. 10, 1961.

BIBLIOGRAPHY (Contd.)

14. Fuller, A.T., "Optimization of Nonlinear Control Systems With Random Inputs," Journal of Electronics and Control, No. 9, 1960.

15. Fuller, A.T., "Optimization of a Saturating Control System With a Brownian Motion Input," Journal of Electronics and Control, No. 10, 1961.

16. Herstein, I.N., Topics in Algebra, Blaisdell Publishing Company, 1964.

17. Hu, S.T., Elements of General Topology, Holden-Day, Inc., 1964.

18. Kalman, R.E., and Bucy, R.S., "New Results in Linear and Prediction Theory," ASME J. Basic Eng. Vol. D-83, 1961.

19. Kushner, H. J., "Approximations to Optimal Nonlinear Filters," (to appear in JACC, 1967).

20. Kushner, H. J., "On the Differential Equations Satisfied by Conditional Probability Densities of Markov Processes, With Application," SIAM J. on Control, 2, 1964.

21. Loeve, M., Probability Theory, Van Nostrand, New York, 3rd edition, 1963.

22. Meier, L., "Combined Optimal Control and Estimation Theory," NASA Contractor Report NASA CR-426, April, 1966.

23. Pontryagin, L. S., et al, The Mathematical Theory of Optimal Processes, Translation: Interscience, 1962.

24. Roberts, A.P., "Optimum Linear Control in the Presence of Coloured Measurement Noise," International Journal of Control, Vol. 3, No. 5, May, 1966.

25. Rozonoer, L. I., "The L. S. Pontryagin Maximum Principle in the Theory of Optimal Control, I, II, III," Automation and Remote Control, 20, 1959.

26. Rudin, W., Principles of Mathematical Analysis, 2nd edition, McGraw-Hill Book Company, New York, 1964.

27. Snyder, D.L., "The State Variable Approach to Continuous Estimation," Ph.D. Thesis, Department of Electrical Engineering, M.I.T., Cambridge, Mass., February, 1966.

BIBLIOGRAPHY (Contd.)

28. Witsenhausen, H. S., "Minimax Controls of Uncertain Systems," Ph.D. Thesis, Department of Electrical Engineering, M.I.T., Cambridge, Mass., May, 1966.

29. Wong, E. and Zakai, M., "On the Convergence of Ordinary Integrals to Stochastic Integrals," Ann. Math. Stat., 36, 1965.

30. Wonham, W. M., "Stochastic Problems in Optimal Control," IEEE International Convention Record, Part 2, 1963.

31. Wonham, W. M., "Some Applications of Stochastic Differential Equations to Optimal Nonlinear Filtering," RIAS TR-64-3, February, 1964.

32. Zadeh, L. A., and Desoer, C. A., Linear System Theory: The State Space Approach, McGraw-Hill, Inc., 1963.