STOCHASTIC OPTIMAL CONTROL

Thesis by

Benjamin N. Early

In Partial Fulfillment of the Requirements

For the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

1970

(Submitted December 3, 1969)

ACKNOWLEDGEMENT

The author greatly appreciates the guidance of his research advisor Professor Sridhar who suggested the problem of this investigation. He would also like to acknowledge Professor R. V. Langmuir and Professor C. H. Wilts for their support and encouragement.

During his graduate studies the author was a National Science Foundation Fellow.

ABSTRACT

H. J. Kushner has obtained the differential equation satisfied by the optimal feedback control law for a stochastic control system in which the plant dynamics and observations are perturbed by independent additive Gaussian white noise processes. However, the differentiation includes the first and second functional derivatives and, except for a restricted set of systems, is too complex to solve with present techniques.

This investigation studies the optimal control law for the open loop system and incorporates it in a suboptimal feedback control law. This suboptimal control law's performance is at least as good as that of the optimal control function and satisfies a differential equation involving only the first functional derivative. The solution of this equation is equivalent to solving two two-point boundary valued integro-partial differential equations. An approximate solution has advantages over the conventional approximate solution of Kushner's equation.

As a result of this study, well known results of deterministic optimal control are deduced from the analysis of optimal open loop control.

TABLE OF CONTENTS

INTRODUCTION

Since the notable contributions of Pontryagin, Bellman and Kalman to the field of deterministic control theory, researchers have wondered if some of these same concepts could be extended to the field of stochastic control theory. In particular, there has been considerable interest in the area of stochastic optimal control theory.

Historically, the works of Florentin [1] and Wonham [2] advanced the state of the art of stochastic optimal control theory for a restricted set of systems. A system of a more general nature was studied by Kushner [3]. His investigations culminated in the derivation of a functional differential equation for the optimal feedback control law. The purpose of this treatise is to expound upon the results of Kushner.

Stochastic control theory is concerned with the control of dynamical systems which in some sense are random. [2] The physical system to be controlled is called the plant. Although the output of the plant of a deterministic system is the state of the system, the plant output in general is not a realizable Markov process and consequently is not the state of the stochastic system. Appropriately, the state of a stochastic system is the probability density function of the plant output. The plant output will be referred to as the plant variable.

If the density function of the plant variable is derived from only the statistics of the plant, it is known as the a priori density function since theoretically it can be computed "off line". The state of the system may be updated if real time observations are taken of the plant variable. If such is the case, this a posteriori or conditional density function of the plant variable, conditioned on all past observations is the state of the system.

In this investigation, two types of control laws [31] are discussed. If the state of the system is determined using the observations, then the control law, which is a functional of the state, is called a feedback or closed loop control law. On the other hand, if the state of the system is mathematically independent of the observations, the control law is called an open loop control law. Consequently, given the open loop control law, the control at each instant of time can be determined a priori. Such a mapping is called a control function.

Because it is intuitively obvious that the plant can be controlled better if the control law is feedback, the author investigates the optimal feedback control problem. The formulation of the problem often found in the literature and presented in this dissertation is as follows. The plant and observational equations are perturbed by independent additive Gaussian white noise [4] processes. The criterion for the optimal feedback control law is to minimize a performance index - the expected value of a functional of the state and the control law. For this model of the system, Kushner [5] derived the integro-partial difference equation of the a posteriori density function. Bucy [38,39] and Mortensen [33] obtained an equivalent equation by an alternate method. In [3] Kushner deduced the functional differential equation for the optimal feedback control law. Also in [33] Mortensen rigorously derived such an equation, again for the alternate method. Since then, the problem of great interest has been the solution of the functional differential equation.

Florentin [1,39] derived the partial differential equation for the optimal feedback control law based upon perfect observations. This equation is more difficult to solve than the well known Hamilton-Jacobi equation [6] of optimal control theory. However, the problem presented by the functional differential equation is avoided. Another problem which simplifies the complex feedback equation arises if the system is linear and the initial state is Gaussian. Then the a posteriori density function is Gaussian for all time and reduces to the Kalman-Bucy filter [7], in which the conditional mean is a sufficient statistic. Wonham [2] utilized this fact in deriving the optimal feedback control law when the performance index is quadratic. His results show that the functional differential equation can be represented by a finite number of ordinary differential equations.

Other analytical or numerical results are lacking because of the complexity of the functional differential equation. A feeling for this difficulty can be seen when the conditional density function is represented by its mean and central moments. Then the control law is derived from an integro-partial differential equation of an infinite number of variables. This method was investigated in [3] and [8].

The purpose of this investigation is to study the functional differential equation derived by Kushner [3] of the optimal feedback control law and its solution and to try to find an engineering approximation to it that would advance the state of the art of stochastic control theory. In Chapter I the model of the system is described, the optimal feedback control problem is formulated, and the feedback equation is derived.

The solution of the functional differential equation is discussed in Chapter II. The "classical" problems of Florentin and Wonham are presented. Approximations that were inspired by Wonham's solution are given and are shown to suggest a study of the optimal open loop control problem.

In Chapter III the optimal open loop control problem is solved in terms of quantities analogous to the Lagrange multiplier (costate variable) and the Hamiltonian of deterministic optimal control theory. When there is no dynamical noise, these quantities are shown to be identical to each other and to imply the canonical equations.

A suboptimal feedback control law, discussed in Chapter IV, is motivated by the previous two chapters. It has the features of the usual approximations with the advantage that nonlinear functions do not have to be truncated in a Taylor series. A numerical example demonstrates the performance of the control law.

The impact of this investigation on the field of stochastic control theory is presented in Chapter V.

I. FORMULATION OF THE OPTIMAL FEEDBACK CONTROL PROBLEM

1.1. Introduction

This chapter is concerned with the presentation of the model and the optimum criterion of the control system. The model is characterized by the plant and observational equations that are perturbed by independent Gaussian white noise [4] processes. Historically, Kushner [5] derived the equation of the evolution of the state of the system - the a posteriori density function of the plant variable. Consequently, this conditional density function is used in defining the performance index of a control law as the expected value of the "cost" of a random run or job. Appropriately, the optimal feedback control law is defined as the control law which minimizes the performance index, and a derivation of the functional differential equation it satisfies is presented. Finally, the performance index of this feedback law is shown to be, as expected, at least as good as that of the optimal control function.

1.2. Model of the System

Let $x(t)$ be an n-vector where $0 \le t \le T$ for a fixed T. Then assume $\{x(t)\}$ is an incremental stochastic process with the following structure:

$$\delta x(t) = m(t,x(t),u(t))\Delta + G(t,x(t))\delta w(t) + O(\Delta^2) \ . \quad (1\text{-}1)$$

Equation (1-1) is known as the plant equation, and $x(t)$ is referred to as the <u>output of the plant</u> or the <u>plant variable</u>. Here $u(t)$, the control, is an n_1-vector, and $w(t)$ is an n_2-vector. The stochastic process $\{w(t)\}$ is a Wiener-Lévy [4] process with unit variance parameter, i.e., $w(t) - w(s) \sim N(0, I_{n_2}|t - s|)$. This process has stationary, independent increments [4,9] and is of a more general class of processes called Brownian motion [4,9].

The n_3-observational vector $y(t)$ of the plant variable has the property that $\{ \int_0^t y(s)ds \}$ is an incremental stochastic process that is described via the following definition. If $z(t) \triangleq \int_0^t y(s)ds$, then

$$\delta z(t) = h(t,x(t))\Delta + K(t)\delta v(t) + O(\Delta^2) \ . \qquad (1\text{-}2)$$

Here $v(t)$ is an n_4-vector where $\{v(t)\}$ is a Wiener-Lévy process with unit variance parameter and is independent of the $\{w(t)\}$ process.

Equations (1-1) and (1-2) comprise the equations defining the model of the system. For infinitesimal Δ, they are known as stochastic or Ito difference equations [9,10]. By dividing them by Δ and formally letting $\Delta \to 0$, they may be written symbolically as:

$$\dot{x}(t) = m(t,x(t),u(t)) + G(t,x(t))\zeta(t) \qquad (1\text{-}3)$$

$$y(t) = h(t,x(t)) + k(t)\eta(t) \ . \qquad (1\text{-}4)$$

The processes $\{\zeta(t)\}$ and $\{\eta(t)\}$ are independent Gaussian white noise [4] processes with autocorrelation functions [4] $I_{n_2}\delta(t - \tau)$ and $I_{n_4}\delta(t - \tau)$, respectively.

In (1-1) through (1-4) the functions G, h, and K may include the control $u(t)$ explicitly, but it was omitted without loss of generality.

1.3. A Posteriori Density Function

Often in control theory and in this investigation one is interested in controlling the present output of the plant. Because of the noise present in the model of the system, the plant variable, in general, is not a realizable Markov process and thus is not the state of the system. Therefore, before the control problem can be formulated, the state of the system is defined.

Assume that the a priori probability density function of $x(0)$ is $\Upsilon(x)$. Define

$$P(t,x)dx \triangleq \text{Prob. } [x(t) \in dx \mid y(s) \quad 0 \le s \le t] \ .$$

Then $P(t,x)$ is the <u>a posteriori</u> or <u>conditional density function</u> of $x(t)$ conditioned on all observations up to time t. Kushner [3] showed that $P(t,x)$ is a Markov process. Appropriately, the a posteriori density function is called the state of the system.

From (1-1) and (1-2) it follows [3] that $P(t,x)$

satisfies the integro-partial difference equation

$$\delta P(t,x) \triangleq P(t+\Delta,x) - P(t,x) = \mathcal{L}_+(t,u(t))P(t,x)\Delta$$

$$+ (\delta z(t) - \bar{h}(t)\Delta)^T r(t,x)P(t,x) + O(\Delta^2) \qquad (1\text{-}5)$$

with $P(0,x) = \Upsilon(x)$

where:

$$\mathcal{L}(t,u(t))\cdot \triangleq m(t,x,u(t))^T\nabla_x\cdot$$

$$+ \tfrac{1}{2}\,\mathrm{trace}\,[G(t,x)G(t,x)^T\nabla_x\nabla_x{}^T\cdot]$$

and its adjoint

$$\mathcal{L}_+(t,u(t))\cdot \triangleq -\nabla_x{}^T[m(t,x,u(t))\cdot]$$

$$+ \tfrac{1}{2}\nabla_x{}^T(\nabla_x{}^T[G(t,x)G(t,x)^T\cdot])^T$$

such that

$$\int A(x)\mathcal{L}B(x)dx = \int \mathcal{L}_+A(x)B(x)dx \qquad (1\text{-}6)$$

when $A(x)$ and $B(x)$ vanish fast enough as $|x|\to\infty$,

$$\mathcal{L}\cdot \triangleq \mathcal{L}(t)\cdot \triangleq \mathcal{L}(t,u(t))\cdot \quad \text{and}$$

$$\mathcal{L}_+\cdot \triangleq \mathcal{L}_+(t)\cdot \triangleq \mathcal{L}_+(t,u(t))\cdot$$

when the arguments are obvious,

$R(t) \triangleq [K(t)K(t)^T]^{-1}$, which is assumed to exist,

$$r(t,x) \triangleq R(t)[h(t,x) - \bar{h}(t)] \; , \quad \text{and}$$

$$\bar{h}(t) = \int h(t,x)P(t,x)dx \; .$$

Dividing (1-5) by Δ and formally letting $\Delta \to 0$, (1-5) may be written as

$$P_t(t,x) = \mathcal{L}_+ P(t,x) + (y(t) - \bar{h}(t))^T r(t,x)P(t,x) \tag{1-7}$$

with $P(0,x) = \Upsilon(x)$.

Equations (1-3) and (1-7) are not differential equations in the usual sense, but are called stochastic or Ito differential equations [9,10].

1.4. Formulation of the Optimal Policy

1.4.1. Determining the Criterion

A fairly general class of deterministic optimal control problems has a performance index of the form

$$\int_0^T f(s,x(s),u(s))ds + g(x(t)) \; .$$

The control $u(s)$ is limited to some prescribed set $\Omega(s)$ of admissible controls, a subset of Euclidean n_1-space. This statement will be omitted for brevity in further discussions.

Because $x(s)$ in the stochastic system is a random variable whose density function satisfies (1-7) , a natural criterion for selecting the optimal control is to minimize

$$\text{Exp}\left[\int_0^T f(s,x(s),u(s))ds + g(x(T))\right] \tag{1-8}$$

where Exp is a suitably defined expected value operator. That is, (1-8) is defined as

$$\underset{\substack{y(\tau) \\ 0 \le \tau \le T}}{\text{Exp}} \left[\int_0^T \int f(s,x,u(s))P(s,x)dxds + \int g(x)P(T,x)dx\right] . \tag{1-9}$$

Appendices A and B show that (1-9) is equivalent to

$$\int_0^T \int f(s,x,u(s))Q(s,x)dxds + \int g(x)Q(T,x)dx \tag{1-10}$$

where

$$Q_s(s,x) = \mathcal{L}_+(s,u(s))Q(s,x) \qquad Q(0,x) = \Upsilon(x) .$$

It is evident that a control function and not a feedback control law minimizes (1-10). This control function is called the <u>optimal control function</u>.

What is more desirable, as is intuitively obvious and as will be shown in 1.5., is to select the optimal control $u(t)$ as a functional of the state of the system - the a posteriori density function. The control law is then closed loop. [11]

1.4.2. Performance of Feedback Control

Let $\Gamma(t,P(t,x))$ be a control functional of $P(t,x)$. Then the performance index of the feedback control law Γ is

$$\underset{\substack{y(\tau)\\0\le\tau\le T}}{\mathrm{Exp}}\left[\int_0^T \int f(s,x,\Gamma(s,P(s,x)))P(s,x)dxds + \int g(x)P(T,x)dx\right]$$

where $u(s)$ is replaced by $\Gamma(s,P(s,x))$ in (1-7). Consider the imbedded system which starts at time t in a state $\varphi(x)$. Define $C(t,\varphi;\Gamma)$ as the performance index of the control law Γ for such a system. Then by the technique of invariant imbedding [12-14]

$$C(t,\varphi;\Gamma) \triangleq \underset{\substack{y(\tau)\\t\le\tau\le T}}{\mathrm{Exp}}\left[\int_t^T \int f(s,x,\Gamma(s,P(s,x)))P(s,x)dxds + \int g(x)P(T,x)dx\right] \qquad (1\text{-}11)$$

$$= \underset{t \le \tau \le t+\Delta}{\underset{y(\tau)}{\text{Exp}}} \left\{ \int_t^{t+\Delta} \int f(s,x,\Gamma(s,P(s,x)))P(s,x)dxds \right.$$

$$+ \underset{t+\Delta \le \tau \le T}{\underset{y(\tau)}{\text{Exp}}} \left[\int_{t+\Delta}^{T} \int f(s,x,\Gamma(s,P(s,x)))P(s,x)dxds \right.$$

$$\left.\left. + \int g(s)P(T,x)dx \right] \right\}$$

$$= \underset{\delta z(t)}{\text{Exp}} \left[\int f(t,x,\Gamma(t,\varphi))\varphi(x)dx\Delta + C(t+\Delta,\varphi+\delta\varphi;\Gamma) \right.$$

$$\left. + O(\Delta^2) \right] \tag{1-12}$$

where $\delta\varphi$ is defined by (1-5) .

Before expanding $C(t+\Delta,\varphi+\delta\varphi;\Gamma)$, the first and second functional derivatives are defined. [15]

Let φ , ψ_1 , and ψ_2 be functions. Then the first functional derivative of the functional $F(\varphi)$ with respect to ψ_1 is

$$F'(\varphi)(\psi_1) \triangleq \lim_{\alpha \to 0} \frac{F(\varphi+\alpha\psi_1) - F(\varphi)}{\alpha} \quad . \tag{1-13}$$

It follows from [15] that $F'(\varphi)(\cdot)$ is a linear operator. The second functional derivative of the functional $F(\varphi)$ with respect to ψ_1 and ψ_2 is

$$F''(\varphi)(\psi_1,\psi_2) \triangleq \lim_{\alpha\to 0} \frac{F'(\varphi+\alpha\psi_2)(\psi_1) - F'(\varphi)(\psi_1)}{\alpha} . \quad (1\text{-}14)$$

Hence $F'(\varphi)(\psi_1,\cdot)$ is a linear operator, and thus $F''(\varphi)(\cdot,\cdot)$ is a bilinear operator.

Thus by Taylor's expansion (1-12) becomes

$$\begin{aligned} C(t,\varphi;\Gamma) = \underset{\delta z(t)}{\mathrm{Exp}} \Big\{ & \int f(t,x,\Gamma(t,\varphi))\varphi(x)dx\Delta + C(t,\varphi;\Gamma) \\ & + C_t(t,\varphi;\Gamma)\Delta + C'(t,\varphi;\Gamma)(\delta\varphi) \\ & + \tfrac{1}{2}C''(t,\varphi;\Gamma)(\delta\varphi,\delta\varphi) + O(\Delta^2) \Big\} \\ = & \int f(t,x,\Gamma(t,\varphi))\varphi(x)dx\Delta + C(t,\varphi;\Gamma) \\ & + C_t(t,\varphi;\Gamma)\Delta + C'(t,\varphi;\Gamma)(\mathcal{L}_+(t,\Gamma(t,\varphi))\varphi)\Delta \\ & + \tfrac{1}{2}\,\mathrm{trace}\,[R^{-1}(t)C''(t,\varphi;\Gamma)(r(t,x)\varphi,r(t,x)^{T}\varphi)]\,\Delta \\ & + O(\Delta^2) . \end{aligned} \quad (1\text{-}15)$$

Equation (1-15) follows from the linearity properties of the functional derivatives and from (1-5) , (A-2) and (A-3).

In the limit as $\Delta \to 0$ (1-15) implies

$$-C_t(t,\varphi;\Gamma) = \int f(t,x,\Gamma(t,\varphi))\varphi(x)dx + C'(t,\varphi;\Gamma)(\mathcal{L}_+(t,\Gamma(t,\varphi))\varphi)$$

$$+ \tfrac{1}{2}\ \text{trace}\ [R^{-1}(t)C''(t,\varphi;\Gamma)(r(t,x)\varphi, r(t,x)'\varphi)] \ . \qquad (1\text{-}16)$$

From (1-11) it is obvious that the initial condition for (1-16) is $C(T,\varphi;\Gamma) = \int g(x)\varphi(x)dx$.

1.4.3. Optimal Feedback Control

Thus $C(t,\varphi;\Gamma)$, the imbedded performance index of a control law Γ , satisfies (1-16). The problem of interest, however, is to derive the feedback control law that minimizes $C(t,\varphi;\Gamma)$. Such a control law is called the optimal feedback control law.

Define the performance index of the optimal feedback control law

$$J(t,\varphi) \triangleq \min_{\Gamma} \ C(t,\varphi;\Gamma) \ . \qquad (1\text{-}17)$$

Let Ξ be the optimal feedback control law and let $\{\Gamma(s,\cdot) = \Xi(s,\cdot) \mid t\ s\leq T\}$. Then independent of $\Gamma(t,\cdot)$ $J(t,\varphi) = C(t,\varphi;\Gamma)$. Likewise, $J_t = C_t$, $J' = C'$ and $J'' = C''$. Hence (1-17) and (1-16) imply

$$-J_t(t,\varphi) = \min_{\Gamma(t,\varphi)} \Big\{ \int f(t,x,\Gamma(t,\varphi))\varphi(x)dx$$

$$+ J(t,\varphi)(\mathcal{L}_+(t,\Gamma(t,\varphi))\varphi)$$

$$+ \frac{1}{2} \operatorname{trace} [R^{-1}(t)J''(t,\varphi)(r(t,x)\varphi, r(t,x)'\varphi)] \Big\} \qquad (1\text{-}18)$$

with $J(T,\varphi) = \int g(x)\varphi(x)dx$.

Equation (1-18) , which was derived by Kushner in [3], is the functional differential equation of the optimal feedback control law for the system of 1.2. Its solution is discussed in Chapter II.

1.5. Why Feedback Control?

Feedback control is desired because it is intuitively obvious that the state of a system can be controlled better if real time observations of the plant variable are utilized in determining the control rather than implementing a control function. This obvious but heretofore unproven fact can be shown via an alternate definition of $J(t,\varphi)$.

Let $S = \{s_0, s_1, \cdots, s_m\}$ be a partition of $[t,T]$, and let α_k be a point in the subinterval $[s_k, s_{k+1}]$. Define $\alpha \triangleq \{\alpha_0, \alpha_1, \cdots, \alpha_{m-1}\}$, and with $P(t,x) = \varphi(x)$ let

$$A(S,\alpha) \triangleq \prod_{i=1}^{m} \underset{u(\alpha_{m-i})}{\text{Min}} \underset{\delta z(\alpha_{m-i})}{\text{Exp}} \left[\sum_{k=0}^{m-1} \int f(\alpha_k, x, u(\alpha_k)) P(\alpha_k, x) dx \Delta s_k \right.$$

$$\left. + \int g(x) P(\alpha_{m-1}, x) dx \right] .$$

Then define $J(t,\varphi) \triangleq \lim_{S\subset[t,T]} A(S,\alpha)$ when the limit (defined in Appendix A) exists. The limit is denoted by

$$\underset{\substack{u(\tau),y(\tau)\\ t\le\tau\le T}}{\mathrm{MEixnp}}\left[\int_t^T \int f(s,x,u(s))P(s,x)dxds + \int g(x)P(T,x)dx\right] \quad (1\text{-}19)$$

with $P(t,x) = \varphi(x)$.

The operator MEixnp represents the sequences of operations Min Exp and symbolizes the structure of $A(S,\alpha)$.

By the technique of invariant imbedding

$$J(t,\varphi) = \underset{\substack{u(\tau),y(\tau)\\ t\le\tau\le T}}{\mathrm{MEixnp}}\left[\int_t^{t+\Delta}\int f(s,x,u(s))P(s,x)dxds\right.$$

$$\left. + \int_{t+\Delta}^{T}\int f(s,x,u(s))P(s,x)dxds + \int g(x)P(T,x)dx\right]$$

$$= \underset{\substack{u(\tau),y(\tau)\\ t\le\tau\le t+\Delta}}{\mathrm{MEixnp}}\left\{\int_t^{t+\Delta}\int f(s,x,u(s))P(s,x)dxds\right.$$

$$+ \underset{\substack{u(\tau),y(\tau)\\ t+\Delta\le\tau\le T}}{\mathrm{MEixnp}}\left[\int_{t+\Delta}^{T}\int f(s,x,u(s))P(s,x)dxds\right.$$

$$+ \int g(x)P(T,x)dx \Big] \Big\}$$

$$J(t,\varphi) = \underset{\substack{u(\tau),y(\tau) \\ t\le\tau\le t+\Delta}}{\text{MEixnp}} \Big\{ \int_t^{t+\Delta} \int f(s,x,u(s))P(s,x)dxds$$

$$+ J(t+\Delta,\varphi+\delta\varphi) \Big\}$$

$$= \underset{u(t)}{\text{Min}} \ \underset{\delta z(t)}{\text{Exp}} \Big\{ \int f(t,x,u(t))\varphi(x)dx\Delta$$

$$+ J(t+\Delta,\varphi+\delta\varphi) + 0(\Delta^2) \Big\} \qquad (1\text{-}20)$$

where $\delta\varphi$ is defined by (1-5) .

By analysis similar to that in (1-15) , (1-20) implies

$$- J_t(t,\varphi) = \underset{u(t)}{\text{Min}} \Big\{ \int f(t,x,u(t))\varphi(x)dx + J'(t,\varphi)(\mathcal{L}_+(t,u(t))\varphi)$$

$$+ \tfrac{1}{2}\text{trace}[R^{-1}(t) J''(t,\varphi)(r(t,x)\varphi, r(t,x)^T\varphi)] \Big\} \qquad (1\text{-}21)$$

with $J(T,\varphi) = \int g(x)P(T,x)dx$.

By (1-18) and (1-21) $J(t,\varphi)$ can be defined by (1-19) . Therefore, because of the relation

$$\underset{r}{\text{Exp}}\ \underset{v}{\text{Min}} \left[\,\cdot\,\right] \leq \underset{v}{\text{Min}}\ \underset{r}{\text{Exp}} \left[\,\cdot\,\right] ,$$

(1-21) implies by induction

$$J(t,\varphi) \leq \underset{\substack{u(\tau) \\ t\leq\tau\leq T}}{\text{Min}}\ \underset{y(\tau)}{\text{Exp}} \left[\int_t^T \int f(s,x,u(s))P(s,x)dxds \right.$$

$$\left. + \int g(x)P(T,x)dx \right] ,$$

the greater of which by Appendices A and B is the imbedded performance index of the optimal control function.

Thus, as has been conjectured, the performance of the optimal feedback control law is at least as good as that of the optimal control function.

II. SOLUTION OF THE OPTIMAL FEEDBACK CONTROL FUNCTIONAL DIFFERENTIAL EQUATION

2.1. Introduction

For the system presented in 1.2. the optimal feedback control law satisfies (1-18). This equation is analogous to the Hamilton-Jacobi [6] equation of deterministic optimal control theory. Because the latter, the simpler version, is difficult to solve both analytically and numerically, it is an extremely arduous task to solve the former, more complex version. Florentin's solution [1] shows that (1-18) reduces to a partial differential equation when the observations are perfect. Wonham's results show that the functional differential equation reduces to a finite system of ordinary differential equations when the system is linear and the performance index is quadratic. Other systems for which (1-18) reduces to a finite system of ordinary differential equations have been fruitlessly investigated. Of course, a system of ordinary differential equations may still be far from a solution. However, their solutions have been studied more and are understood better than those of other types of differential equations.

If the state of the system has a finite number of sufficient statistics, then an alternate approach is to represent the state by them. In addition, (1-18) reduces to a partial differential equation. The problem with an infinite number of sufficient statistics is discussed in [3] and [8] .

A numerical solution of the complex feedback equation was attempted for a simple first order nonlinear problem. Efforts were terminated because of the enormous amount of computational time needed.

A typical approximation of (1-18) arises from Wonham's solution. It is shown to suggest the studying of the optimal open loop control problem.

2.2. System with Perfect Observations

Florentin [1,39] derived the optimal feedback control law when there are perfect observations; that is, for a system with $n_3 = n$, $h(t,x(t)) = x(t)$ and $K(t) = 0$. Equation (1-7) then implies $P(t,x) = \delta(x-y(t))$.

By defining $\Lambda(t,c)$ to be the optimal feedback control law at time t given the observation $y(t) = c$, Florentin defined as the imbedded performance index

$$F(t,c) \triangleq \underset{\Lambda}{\text{Min}} \underset{\substack{y(\tau) \\ t\le\tau\le T}}{\text{Exp}} \left[\int_t^T f(s,y(s),\Lambda(s,y(s)))ds + g(y(T))\right] \quad (2\text{-}1)$$

with $y(t) = c$. Analogous to the method in 1.3.3., he derived from (2-1) the partial differential equation for the optimal feedback control law

$$- F_t(t,c) = \underset{\Lambda(t,c)}{\text{Min}} \left\{ f(t,c,\Lambda(t,c)) + \mathcal{L}(t,\Lambda(t,c))F(t,c) \right\} \quad (2\text{-}2)$$

with $F(T,c) = g(c)$. This equation may also be derived from (1-18).

Equation (2-2) is much simpler than (1-20) but more complex than the well known Hamilton-Jacobi equation [6] of deterministic optimal control theory. It has been shown to reduce to a finite set of ordinary differential equations only for a restricted class of systems.

2.3. Linear System with a Quadratic Performance Index

One of the restricted classes of systems for which (1-18) reduces to a finite system of ordinary differential equations is the linear system with a quadratic performance index. The set $\Omega(t)$ of admissible controls at time t is understood to be the Euclidean n_1-space. Such a system implies

$$m(t,x(t),u(t)) = A(t)x(t) + B(t)u(t) \ ,$$

$$G(t,x(t)) = G(t) \ ,$$

$$h(t,x(t)) = H(t)x(t) \ ,$$

$$f(t,x(t),u(t)) = \tfrac{1}{2}x(t)^T Q(t)x(t) + \tfrac{1}{2}u(t)^T C(t)u(t) ,$$

$$g(x(T)) = \tfrac{1}{2}x(T)^T Sx(T) ,$$

where matrices $A(t)$ is $n \times n$, $B(t)$ is $n \times n_1$, $G(t)$ is $n \times n_2$, $H(t)$ is $n_3 \times n$, $Q(t)$ and S are $n \times n$, symmetrical and positive semi-definite, and $C(t)$ is $n_1 \times n_1$, symmetrical and positive definite.

If $\Upsilon(x) = N(x,\mu_0,M_0)$, then (1-7) reduces [2] to the Kalman-Bucy filter[7], that is, $P(t,x) = N(x,\mu(t),M(t))$ where

$$\dot{\mu}(t) = A(t)\mu(t) + B(t)u(t) + M(t)H(t)^T R(t)[y(t) - H(t)\mu(t)] \qquad \mu(0) = \mu_0 \qquad (2\text{-}3)$$

and

$$\dot{M}(t) = A(t)M(t) + M(t)A(t)^T - M(t)H(t)^T R(t)H(t)M(t) + G(t)G(t)^T \qquad M(0) = M_0 \qquad (2\text{-}4)$$

Since $M(t)$ can be determined a priori, it is thought of as a function of time and not as a statistic. Thus $J(t,P(t,x))$ is a function of t and $\mu(t)$.

It can be easily shown that

$$J(t,\varphi) = \tfrac{1}{2}\int x^{T}\varphi(x)dxU(t)\int x\varphi(x)dx + \tfrac{1}{2}w(t) \qquad (2\text{-}5)$$

is a solution of (1-18) where $U(t)$ is the $n \times n$ symmetrical matrix that satisfies

$$\dot{U}(t) = - A(t)^{T}U(t) - U(t)A(t) + U(t)B(t)C^{-1}(t)B(t)^{T}U(t)$$

$$- Q(t) \qquad\qquad U(T) = S \qquad (2\text{-}6)$$

The optimal feedback control law which follows from (1-18) and (2-5) is

$$\Gamma(t,\varphi) = - C^{-1}(t)B(t)^{T}U(t)\int x\varphi(x)dx\ . \qquad (2\text{-}7)$$

Both (2-4) and (2-5) are matrix Riccati-equations [16] which may be solved a priori. Only (2-3) has to be solved in real time to yield the optimal feedback control, $\Gamma(t,N(\mu(t),M(t)))$, of (2-7) .

These are the results derived by Wonham [2], who approached the problem as discussed in 2.4. without using (1-18) .

2.4. Alternate Approach to the Optimal Policy

An alternate approach to deriving the optimal feedback control law is to represent the a posteriori density function by its mean and central moments. The conditional mean

$$\mu(t) = \int xP(t,x)dx \quad .$$

With the definitions

$$\alpha \triangleq (\alpha_1, \alpha_2, \cdots, \alpha_n) \quad ,$$

$$A \triangleq \{\alpha \mid \alpha_i \in \{0,1,2,\cdots\} \quad 1 \le i \le n\}$$

and

$$B \triangleq \{\alpha \mid \alpha \in A \text{ and } \sum_{i=1}^{n} \alpha_i < 2\}$$

The central moments are $\{\beta(t;\alpha) \mid \alpha \in A - B\}$ where

$$\beta(t;\alpha) = \int \prod_{i=1}^{n} (x_i - \mu_i(t))^{\alpha_i} P(t,x)dx \quad .$$

Now let $\lambda_1(t)$, $\lambda_2(t)$, $\lambda_3(t)$, $\cdots$ represent some ordering of the central moments and define

$$\lambda(t) \triangleq \mathrm{Col}_{i=1}^{\infty} \{\lambda_i(t)\}$$

as the central moment.

From (1-5) the ordinary difference equations for $\mu(t)$ and $\lambda(t)$ can be derived. With $\Sigma(t,\mu(t),\lambda(t))$ as a feedback control law, define

$$n(t,\mu(t),\lambda(t),\ \Sigma(t,\mu(t),\lambda(t))) \triangleq \int f(t,x,\Sigma(t,\mu(t),\lambda(t)))P(t,x)dx,$$

and

$$q(\mu(T),\lambda(T)) \triangleq \int g(x)P(T,x)dx \ .$$

The imbedded performance index for the optimal feedback control law of a process which begins at time t with a conditional mean c and central moment d is

$$L(t,c,d) = \underset{\substack{\Sigma \\ t\le\tau\le T}}{\text{Min}} \ \underset{y(\tau)}{\text{Exp}} \left\{ \int_t^T n(s,\mu(s),\lambda(s),\ \Sigma(s,\mu(s),\lambda(s)))ds \right.$$

$$\left. + \ q(\mu(T),\lambda(T)) \right\} \qquad (2\text{-}8)$$

with $\mu(t) = c$ and $\lambda(t) = d$.

A partial differential equation [3,8] for L can be derived from (2-8) analogous to the procedure presented in 1.3.3. But since d has an infinite number of components, the equation has an infinite number of variables. Thus only a truncation of it can be implemented.

For the linear system Wonham noted that $\lambda(t)$ was independent of the observations and control and absorbed it in $L(t,c)$ from which he derived the optimal control law.

2.5. Numerical Solution

The partial differential equation suggested in 2.4. is an awkward equation to solve numerically. Because of its infinite number of variables, it reduces to a cumbersome sequence of ordinary differential equations where each equation contains an infinite number of variables. However, the functional differential equation (1-18), which is more complex conceptually, can be reduced to a sequence of ordinary differential equations, each with a finite number of variables. There is a natural ordering of this sequence which is suggested after the following definitions.

With $\varphi(x)$ fixed, define the scalar quantities

$$A_1(s) \triangleq J(s,\varphi)$$

$$A_2(s) \triangleq \Gamma(s,\varphi)$$

$$A_3(s) \triangleq J'(s,\varphi)(\mathcal{L}_+\varphi)$$

$$A_{3+\frac{i(i-1)}{2}+j}(s) \triangleq J''(s,\varphi)(r_i(s,x)\varphi, r_j(s,x)\varphi)$$

for $1 \le j \le i$ and $1 \le i \le n$.

By defining $\{A_1\}$ as the first step and $\left\{A_2, A_3, \cdots, A_{3+\frac{n(n+1)}{2}}\right\}$ as the second step, $\dot{A}_1$ can be determined by (1-18) as a function of the second step. By taking the necessary functional derivatives of both sides of (1-18) , the time derivatives of each element of the second step can be derived as functions of the second step and extra variables that define the third step. This process is continued where the time derivatives of the k^{th} step are derived via (1-18) as a function of variables defined in the first k steps and extra variables that define the $(k+1)^{st}$ step.

Here a natural sequence of ordinary differential equations is evolved from the functional differential equation and offers a more concise numerical approach than the partial differential equation referred to in 2.4.

A computer program was written to generate the sequence of ordinary differential equations for a simple first-order nonlinear system. On the IBM 360/75 computer only the differential equations of the first three steps (47) were derived because of the magnitude of time needed to derive those of the fourth step consisting of about 400 terms.

However, the solution of three steps of equations with $[0,T]$ partitioned into 50 parts takes approximately

400 minutes and thus was not attempted. The numerical solution resulting from just three steps of equations was attempted for the first order system of 2.3., and the answer was incorrect by several orders of magnitude.

Assuming that four steps of ordinary differential equations yield a satisfactory numerical solution, they must be solved in real time for each $P(t,x)$ for each increment of time. Hence on one of the fastest computers of its generation, the increment of real time for a simple problem can be no smaller than 200 minutes!

Obviously, except in a few rare cases, the determination of the stochastic optimal control is impractical. Consequéntly, either faster computers must be built or else simpler but effective suboptimal schemes must be used.

2.6. Approximate Solution

2.6.1. In General

Because of the awesome amount of real computational time to solve (1-18) numerically, given $P(t,x)$, and to solve (1-7) to determine $P(t,x)$, simple suboptimal solutions are desirable. One approximation is to expand all nonlinear functions in a Taylor series about some a priori nominal trajectory of the state of the system. By neglecting second order terms, a system of equations like those in 2.3. are derived. Consequently, a suboptimal control law is obtained using Wonham's solution. However, it has been recognized [17] that unacceptable performance may follow, possibly because second order terms are significant. Thus, such expansions are not discussed in this treatise.

Other approximations of (1-18) stem from the approach of 2.4. First, $P(t,x)$ is approximated by a Gaussian density function $N(x,\hat{\mu}(t),\hat{M}(t))$ where $\hat{\mu}(t)$ and $\hat{M}(t)$ are generated by a nonlinear filter. At least eight nonlinear filters [18-27] for the system of 1.2. have been published. Next $\hat{M}(t)$ is approximated by some $\hat{\hat{M}}(t)$ that can be determined a priori. Lastly, since a Dirac delta function is a simpler form for $P(t,x)$, the latter is approximated by

$$P(t,\hat{\mu}(t),x) \triangleq \begin{cases} N(x,\hat{\mu}(t),\hat{\hat{M}}(t)) \\ \quad\text{or} \\ \delta(x-\hat{\mu}(t)) \quad . \end{cases} \tag{2-9}$$

Thus $\hat{\mu}(t)$ is a sufficient statistic for $\hat{P}(t,\hat{\mu}(t),x)$.

Appropriately, the imbedded performance index $J(t,P(t,x))$ is approximated by $L(t,\hat{\mu}(t))$. Let $\Lambda(t,\hat{\mu}(t))$ be the corresponding feedback control law.

Such a control law is called suboptimal when it is derived from an approximation of the optimal control law.

By a derivation suggested in 2.4. and analogous to that for (1-20) ,

$$L(t,\nu) = \underset{\Lambda(t,\nu)}{\text{Min}} \; \underset{\delta z(t)}{\text{Exp}} \left\{ \int f(t,x,\Lambda(t,\nu))\hat{P}(t,\nu,x)dx\Delta \right.$$

$$\left. + L(t+\Delta,\nu+\delta\nu) + O(\Delta^2) \right\} . \qquad (2\text{-}10)$$

The increment in ν is derived from the nonlinear filter used. Most nonlinear filters are of the form

$$\dot{\hat{\mu}}(t) = a(t,\hat{\mu}(t),u(t)) + F(t,\hat{\mu}(t))(y(t) - h(t,\hat{\mu}(t))) \qquad (2\text{-}11)$$

where $F(t,\cdot)$ is a $n \times n_3$ matrix function of t , $\hat{M}(t)$, and $\nabla_{\cdot} h(t,\cdot)^{T}$. Notice that if h is linear, F is independent of $\hat{\mu}(t)$.

Continuing from (2-10), since

$$L(t+\Delta,\nu+\delta\nu) = L(t,\nu) + L_t(t,\nu)\Delta + \delta\nu^{T}\nabla_{\nu}L(t,\nu)$$

$$+ \tfrac{1}{2}\delta\nu^{T}[\nabla_{\nu}\nabla_{\nu}^{T}L(t,\nu)]\delta\nu + O(\Delta^2) ,$$

then by (2-11) , (2-10) leads to

$$- L_t(t,\nu) = \underset{\Lambda(t,\nu)}{\text{Min}} \Big\{ \int f(t,x,\Lambda(t,\nu))\hat{P}(t,\nu,x)dx$$

$$+ a(t,\nu,\Lambda(t,\nu))^T \nabla_\nu L(t,\nu)$$

$$+ \tfrac{1}{2}\text{trace}[R^{-1}(t)F(t,\nu)^T[\nabla_\nu \nabla_\nu{}^T L(t,\nu)]F(t,\nu)] \Big\} \qquad (2\text{-}12)$$

with $L(t,\nu) = \int g(x)\hat{P}(T,\nu,x)dx$.

Next $L(t,\nu)$ is approximated by a quadratic polynomial in ν : $\frac{1}{2}\nu^T S_1(t)\nu + S_2(t)^T \nu + S_3(t)$ where $S_1(t)$ is an n x n symmetrical matrix, and $S_2(t)$ is an n-vector. Equation (2-12) then becomes

$$- \tfrac{1}{2}\nu^T \dot{S}_1 \nu - \dot{S}_2{}^T \nu - \dot{S}_3 = \underset{\Lambda(t,\nu)}{\text{Min}} \Big\{ \int f(t,x,\Lambda(t,\nu))\hat{P}(t,\nu,x)dx$$

$$+ a(t,\nu,\Lambda(t,\nu))^T (S_1 \nu + S_2)$$

$$+ \tfrac{1}{2}\text{trace}[R^{-1}(t)F(t,\nu)^T S_1 F(t,\nu)] \Big\} \qquad (2\text{-}13)$$

with $L(T,\nu) = \int g(x)\hat{P}(T,\nu,x)dx$.

The last approximation before obtaining the suboptimal feedback control law is to expand the right side of (2-13) and its initial condition in a Taylor series in ν where terms of degree three or more are truncated. Call these quadratic polynomials $A(t,\nu)$ and $B(\nu)$, respectively. The resulting equation is

$$-\tfrac{1}{2}\nu^T\dot{S}_1\nu - \dot{S}_2^T\nu - \dot{S}_3 = A(t,\nu) \tag{2-14}$$

with $\quad \tfrac{1}{2}\nu^T S_1(T)\nu + S_2^T(T)\nu + S_3(T) = B(\nu)$.

As a result, the ordinary differential equations for $S_1(t)$, $S_2(t)$ and $S_3(t)$ follow. Since $S_i(t)$ is independent of ν , it can be computed a priori.

This standard suboptimal scheme, which is a function of $\frac{n^2+3n+2}{2}$ components of S_1 , S_2 and S_3 , yields the optimal feedback control law for the linear system with a quadratic performance index.

2.6.2. Linear Observations

A typical approximate solution of (1-18) was described in 2.6.1. When the observations are linear, the suboptimal feedback control law is the same as a suboptimal open loop control law. This will be shown after this comment on linear observations.

The observations are linear if $h(t,x(t)) = H(t)x(t)$ where $H(t)$ is an $n_3 \times n$ matrix. Notice that if there exists a vector function $g(x(t))$ that has an n-vector function inverse g^{-1}, and that if the transformation $c(t) = g(x(t))$ implies that $h(t,x(t)) = h(t,g^{-1}(c(t)))$ is linear in $c(t)$, then redefining the state of the system to be $c(t)$ will yield linear observations.

The significance of having linear observations in the suboptimal feedback control problem is that F and consequently the last term in (2-13) are then independent of ν. Consequently, S_1 and S_2 are independent of the last term, which effects only S_3. If h and K do not depend explicitly on the control, then the suboptimal control law is a function of t, ν, S_1, and S_2 but not of S_3. Thus, for determining the suboptimal control, the last term in (2-13) may be deleted.

Thus, if the observations are linear and $J(t,\varphi)$ is approximated by the quadratic

$$L(t,\int x\varphi(x)dx) = \tfrac{1}{2}\int x^{T}\varphi(x)dx S_1(t)\int x\varphi(x)dx$$

$$+ S_2^{T}(t)\int x\varphi(x)dx + S_3(t) ,$$

then $R^{-1}(t)J''(t,\varphi)(r(t,x)\varphi, r(t,x)'\varphi)$ may be deleted from (1-18) without altering the suboptimal feedback control law. The equation resulting from (1-18) is

$$- J_t(t,\varphi) = \underset{\Gamma(t,\varphi)}{\text{Min}} \left\{ \int f(t,x,\Gamma(t,\varphi))\varphi(x)dx \right.$$

$$\left. + J'(t,\varphi)(\mathcal{L}_+(t,\Gamma(t,\varphi))\varphi) \right\} \qquad (2\text{-}15)$$

with $J(T,\varphi) = \int g(x)\varphi(x)dx$.

Notice that equation (2-15) is derived from (1-18) when the observations are independent of the plant variable. Consequently, (2-15) is the equation for optimal open loop control. Appropriately, in the search of "better" approximations of (1-18) , the optimal open loop control problem [31,32] is studied in Chapter III.

III. OPTIMAL OPEN LOOP CONTROL

3.1. Introduction

The optimal open loop control problem is studied because the conventional suboptimal feedback control law can be derived from it. The optimal open loop control law, which can be deduced from (1-18), is derived using dynamic programming [28]. This approach leads to the definitions of the open loop or a priori stochastic analog of the Lagrange multiplier (costate variable), the Hamiltonian and Pontryagin's maximum principle [6]. A feature, which is desirable but uproven for the feedback system, exists for the open loop system and is as follows. If the system is linear and the functions f and g are polynomials in the plant variable, then the optimal control law can be represented by a finite system of ordinary differential equations.

3.2. Open Loop Policy via Dynamic Programming

If there are no observations of the plant variable, then $h(t,x(t))$ is independent of $x(t)$, and the system of 1.2. is said to be an open loop control system since the control can be determined a priori. Therefore, $\bar{h}(t) = h(t,x(t))$, and hence $r(t,x(t)) = 0$. Consequently, (1-7) reduces to $P_t(t,x) = \mathcal{L}_+P(t,x)$ which is known as the Fokker-Planck [9] equation.

If the a priori density function of $x(0)$ is $\Upsilon(x)$, define $Q(t,x)$ as the a priori density function of $x(t)$. Then

$$Q_t(t,x) = \mathcal{L}_+(t,u(t))Q(t,x) \qquad Q(0,x) = \Upsilon(x) . \qquad (3\text{-}1)$$

Consequently, when $h(t,x(t)) = h(t)$ in the model of the system the state of the system is the a priori density function.

With the open loop control system defined, the optimal control function $\{\theta(s) \mid 0 \leq s \leq T\}$ is chosen to minimize the performance index

$$\mathrm{Exp}\left[\int_0^T f(s,x(s),u(s))ds + g(x(T))\right]$$

$$\triangleq \int_0^T \int f(s,x,u(s))Q(s,x)dxds + \int g(x)Q(T,x)dx . \qquad (3\text{-}2)$$

Notice that the optimal control function is mathematically independent of the observations.

Consider the imbedded performance index for the control function u

$$\int_t^T \int f(s,x,u(s))Q(s,x)dxds + \int g(x)Q(T,x)dx . \qquad (3\text{-}3)$$

Define

$$V(t,\varphi) \triangleq \underset{\substack{u(\tau)\\ t\le\tau\le T}}{\text{Min}} \; [(3\text{-}3)] \qquad Q(t,x) = \varphi(x) \; . \qquad (3\text{-}4)$$

Notice that as in (1-9) Appendices A and B show that

$$V(t,\varphi) = \underset{\substack{u(\tau)\\ t\le\tau\le T}}{\text{Min}} \; \underset{y(\tau)}{\text{Exp}} \left[\int_t^T \int f(s,x,u(s))P(s,x)dxds \right.$$

$$\left. + \int g(x)P(T,x)dx \right] \qquad P(t,x) = \varphi(x) \; .$$

Continuing, by dynamic programming [28]

$$V(t,\varphi) = \underset{\substack{u(\tau)\\ t\le\tau\le t+\Delta}}{\text{Min}} \left\{ \int_t^{t+\Delta} \int f(s,x,u(s))Q(s,x)dxds \right.$$

$$\left. + V(t+\Delta,\varphi+\delta\varphi) \right\}$$

$$= \underset{u(t)}{\text{Min}} \left\{ \int f(t,x,u(t))\varphi(x)dx\Delta + V(t,\varphi) + V_t(t,\varphi)\Delta \right.$$

$$\left. + V'(t,\varphi)(\mathcal{L}_+\varphi)\Delta + O(\Delta^2) \right\} \qquad (3\text{-}5)$$

where $\delta\varphi$ is derived from (3-1) .

Equation (3-5) then reduces as $\Delta \to 0$ to

$$- V_t(t,\varphi) = \underset{u(t)}{\text{Min}} \left\{ \int f(t,x,u(t))\varphi(x)dx + V'(t,\varphi)(\mathcal{L}_+\varphi) \right\} \quad (3\text{-}6)$$

with $V(T,\varphi) = \int g(x)\varphi(x)dx$.

By (3-6) the optimal open loop control law is a function of t and φ , i.e., $\Psi(t,\varphi)$. By replacing φ by $Q(t,x)$, which is independent of the observations, the optimal control function is

$$\theta(t) = \Psi(t,Q(t,x)) \; . \quad (3\text{-}7)$$

3.3. Toward the A Priori Stochastic Hamiltonian

It is well known that in deterministic optimal control theory [6] the solution of the Hamilton-Jacobi equation, which is a partial differential equation of the first order, is equivalent to solving the canonical equations- $2n$ ordinary differential equations where n is the order of the state equation. Lur'e [36] proved the existence of a similar result for the functional differential equation of the first kind as typified by (3-6). He also showed, as Mortensen [34,35] and Wang [37] point out, that the solution of such an equation is equivalent to solving two partial differential equations of n independent variables. However, the partial differential equations were not explicitly exhibited. Derivations which follow produce these equations for (3-6) and show how (3-6) relates to the a priori stochastic Hamiltonian and maximum principle.

Theorem: For $t<s$ let $\{u(\tau) \mid t\le\tau\le s\}$, $c(s,x)$ and $Q(t,x)$ be given. Define $q(s,x,s) \triangleq c(s,x)$ and $q_\sigma(\sigma,x,s) \triangleq -\mathcal{L}(\sigma)q(\sigma,x,s)$ for $t\le\sigma\le s$.

Then

$$\int c(s,x)Q(s,x)dx = \int q(t,x,s)Q(t,x)dx .$$

Proof: Define

$$p(s,\xi;\sigma,x)d\xi \triangleq \text{Prob.}[x(s)\epsilon d\xi \mid x(\sigma)=x] .$$

Then $p(s,\xi;\sigma,x)$ is the <u>transition density function</u> of the Markov process $\{x(t)\}$ and satisfies Kolmogorov's <u>backward equation</u> [9]

$$p_\sigma(s,\xi;\sigma,x) = -\mathcal{L}(\sigma)p(s,\xi;\sigma,x)$$

and the Fokker-Planck or <u>forward equation</u> [9]

$$p_s(s,\xi;\sigma,x) = \mathcal{L}_+(s)p(s,\xi;\sigma,x) .$$

The independent variables of the operators $\mathcal{L}(\sigma)$ and $\mathcal{L}_+(s)$ are x and ξ, respectively.

By the properties of conditional density functions

$$Q(s,\xi) = \int p(s,\xi;\sigma,x)Q(\sigma,x)dx . \qquad (3\text{-}8)$$

Define $q(\sigma,x,s) \triangleq \int c(s,\xi)p(s,\xi;\sigma,x)d\xi$. (3-9)

Then $q_\sigma(\sigma,x,s) = \int c(s,\xi)p_\sigma(s,\xi;\sigma,x)d\xi$

$= -\int c(s,\xi)\mathcal{L}(\sigma)p(s,\xi;\sigma,x)d\xi$

$= -\mathcal{L}(\sigma)\int c(s,\xi)p(s,\xi;\sigma,x)d\xi = -\mathcal{L}(\sigma)q(\sigma,x,s).$

Therefore, by (3-8) and (3-9) with $\sigma=t$

$$\int c(s,\xi)Q(s,\xi)d\xi = \int c(s,\xi)\int p(s,\xi;t,x)Q(t,x)dxd\xi$$

$$= \iint c(s,\xi)p(s,\xi;t,x)d\xi Q(t,x)dx$$

$$= \int q(t,x,s)Q(t,x)dx \ . \quad \blacksquare$$

3.4. A Priori Stochastic Operand

Recall that the optimal control function θ minimizes (3-3) with $Q(0,x) = \Upsilon(x)$. Define $c_i(s,x)$ and $q_i(t,x,s)$, analogous to c and q in the theorem of 3.3., such that

$$c_1(s,x) \triangleq f(s,x,\theta(s)) \quad \text{and}$$

$$c_2(T,x) \triangleq g(x) \ .$$

Then by (3-4),

$$V(t,Q(t,x)) = \int_t^T \int c_1(s,x)Q(s,x)dxds + \int c_2(T,x)Q(T,x)dx$$

$$= \int_t^T \int q_1(t,x,s)Q(t,x)dxds + \int q_2(t,x,T)Q(t,x)dx$$

$$V(t,Q(t,x)) = \int \left[\int_t^T q_1(t,x,s)ds + q_2(t,x,T) \right] Q(t,x)dx \ .$$

Define as the a priori stochastic operand

$$E(t,x) \triangleq E(t,x,Q(t,x))$$

$$\triangleq \int_t^T q_1(t,x,s)ds + q_2(t,x,T) \ . \qquad (3\text{-}10)$$

Then

$$V(t,Q(t,x)) = \int E(t,x)Q(t,x)dx \ . \qquad (3\text{-}11)$$

Clearly, $E(T,x) = q_2(T,x,T) = g(x)$.

By differentiating (3-10) with respect to t , the following integro-partial differential equation is derived:

$$E_t(t,x) = - q_1(t,x,t) - \int_t^T \mathcal{L}(t)q_1(t,x,s)ds - \mathcal{L}(t)q_2(t,x,T)$$

$$= - f(t,x,\theta(t)) - \mathcal{L}\int_t^T q_1(t,x,s)ds - \mathcal{L}q_2(t,x,T)$$

$$= - f(t,x,\theta(t)) - \mathcal{L}E(t,x) \ . \qquad (3\text{-}12)$$

If $f(t,x,\theta(t))$ and $g(x)$ are positive semi-definite, then Appendix C shows that $E(t,x)$ is also positive semi-definite.

If the optimal control function θ is known, then (3-12) can be solved in backward time.

More generally, if the a priori density function at time t is φ, then by (3-10) $E(t,x) = E(t,x,\varphi)$ and hence $E(t+\Delta,x,\varphi+\delta\varphi) = E(t+\Delta,x)$ where $\delta\varphi$ is defined by (3-1).

By (3-7) and (3-12)

$$E(t+\Delta,x) = E(t,x) - f(t,x,\Psi(t,\varphi))\Delta - \mathcal{L}(t,\Psi(t,\varphi))E(t,x)\Delta$$
$$+ O(\Delta^2) ,$$

and by (1-13)

$$E(t+\Delta,x,\varphi+\delta\varphi) = E(t,x,\varphi) + E_t(t,x,\varphi)\Delta$$
$$+ E'(t,x,\varphi)(\mathcal{L}_+(t,\Psi(t,\varphi)))\Delta + O(\Delta^2) .$$

These results lead to the following functional differential equation as $\Delta \to 0$:

$$E_t(t,x,\varphi) = - E'(t,x,\varphi)(\mathcal{L}_+(t,\Psi(t,\varphi))) - f(t,x,\Psi(t,\varphi))$$
$$- \mathcal{L}(t,\Psi(t,\varphi))E(t,x,\varphi) \qquad (3\text{-}13)$$

with $E(T,x,\varphi) = \int g(x)\varphi(x)dx$.

Thus

$$E_t(t,x) = E_\tau(\tau,x,Q(t,x))\big|_{\tau=t} + E'(t,x,Q(t,x))(\mathcal{L}_+Q(t,x)) .$$

3.5. A Priori Stochastic Hamiltonian and Maximum Principle

By (3-10) and (3-11)

$$V(t,\varphi) = \int E(t,x,\varphi)\varphi(x)dx . \tag{3-14}$$

To determine $V'(t,\varphi)(\psi)$ for an arbitrary function $\psi(x)$, consider

$$V(t,\varphi+\alpha\psi) = \int E(t,x,\varphi+\alpha\psi)\varphi(x)dx$$

$$+ \alpha\int E(t,x,\varphi+\alpha\psi)\psi(x)dx . \tag{3-15}$$

Associated with $V(t,\varphi+\alpha\psi)$ and $E(t,x,\varphi+\alpha\psi)$ is a control function $\gamma(s) \triangleq \gamma(s;\psi,\alpha)$ which minimizes (3-3) with $Q(t,x) = \varphi(x) + \alpha\psi(x)$. But by the theorem of 3.3. $\int E(t,x,\varphi+\alpha\psi)\varphi(x)dx$ is equal to the expression (3-3) with $u(s) = \gamma(s)$ and $Q(t,x) = \varphi(x)$. And since under these conditions (3-3) is at its minimum when $\gamma(s) = \theta(s)$ or when $\alpha = 0$, then

$$\frac{\partial}{\partial\alpha}\int E(t,x,\varphi+\alpha\psi)\varphi(x)dx\Big|_{\alpha=0} = 0 . \tag{3-16}$$

By (1-13) $V'(t,\varphi)(\psi) = \frac{\partial}{\partial\alpha} V(t,\varphi+\alpha\psi)\big|_{\alpha=0}$.

Consequently, by (3-15) and (3-16)

$$V'(t,\varphi)(\psi) = \int E(t,x,\varphi)\psi(x)dx \quad . \tag{3-17}$$

Now define as the <u>a priori stochastic Hamiltonian</u>

$$H(t,\varphi,E,u(t)) \triangleq \int f(t,x,u(t))\varphi(x)dx$$

$$+ \int \mathcal{L}(t,u(t))E(t,x,\varphi)\varphi(x)dx \quad . \tag{3-18}$$

It follows that

$$\underset{u(t)}{\text{Min}}\ H(t,\varphi,E,u(t))$$

$$= \underset{u(t)}{\text{Min}} \left\{ \int f(t,x,u(t))\varphi(x)dx + \int \mathcal{L}E(t,x,\varphi)\varphi(x)dx \right\}$$

$$= \text{by (1-6)} \ \underset{u(t)}{\text{Min}} \left\{ \int f(t,x,u(t))\varphi(x)dx \right.$$

$$\left. + \int E(t,x,\varphi)\mathcal{L}_{+}\varphi(x)dx \right\}$$

$$= \text{by (3-18)} \ \underset{u(t)}{\text{Min}} \left\{ \int f(t,x,u(t))\varphi(x)dx + V'(t,\varphi)(\mathcal{L}_{+}\varphi) \right\}$$

$$= \text{by (3-6)} \ - V_t(t,\varphi) \ .$$

Thus by (3-6), the optimal control $\theta(t)$ satisfies

$$H(t,\varphi,E,\theta(t)) = \underset{u(t)}{\text{Min}}\ H(t,\varphi,E,u(t)) \ , \qquad (3\text{-}19)$$

which is the a priori stochastic maximum principle.

3.6. Summary

The functional differential equation (3-6) which yields the optimal control function $\theta(t) = \Psi(t,Q(t,x))$ reduces to solving two two-point boundary valued integro-partial differential equations:

$$Q_t(t,x) = \mathcal{L}_+(t,\theta(t))Q(t,x) \qquad Q(0,x) = \Upsilon(x) \qquad (3\text{-}20a)$$

and

$$E_t(t,x) = - f(t,x,\theta(t)) - \mathcal{L}(t,\theta(t))E(t,x) \qquad (3\text{-}21a)$$

with $E(T,x) = g(x)$

where

$$H(t,Q,E,\theta(t)) = \underset{u(t)}{\text{Min}}\ H(t,Q,E,u(t)) \ . \qquad (3\text{-}22a)$$

These equations result from (3-1) , (3-12) and (3-19) and represent the solution of the optimal open loop control problem. By (3-11) its performance index is

$$\int E(0,x)\Upsilon(x)dx \ .$$

Notice that $E(t,x,\varphi)$ can be found by solving the initial valued equation (3-13) or by solving these two-point boundary valued equations:

$$Q_s(s,x) = \mathcal{L}_+(s,\theta(s))Q(s,x) \qquad Q(t,x) = \varphi(x) \tag{3-20b}$$

and

$$E_s(s,x) = -f(s,x,\theta(s)) - \mathcal{L}(s,\theta(s))E(s,x) \tag{3-21b}$$

with $E(T,x) = g(x)$

where

$$H(s,Q,E,\theta(s)) = \underset{u(s)}{\text{Min}}\ H(s,Q,E,u(s)) \ . \tag{3-22b}$$

Consequently, $E(t,x,\varphi) = E(t,x)$.

<u>3.7. Linear System</u>

A class of problems which has been of interest is the linear system with a non-quadratic performance index. While (1-18) has not been reduced to a finite system of ordinary differential equations, (3-6) via the results in 3.6. can be reduced when f and g are polynomials in x. This can be seen after the following definitions.

A form of degree p in x is the sum

$$\sum_{i_1=1}^{n} \cdots \sum_{i_p=1}^{n} a_{i_1,\cdots,i_p} x_{i_1} \cdots x_{i_p} \not\equiv 0 \ .$$

A polynomial of degree q is the sum of forms the maximum degree of which is q. Define a constant as a polynomial of degree 0.

If $f(t,x,\theta(t))$ is a polynomial of degree r_1 and $g(x)$ is a polynomial of degree r_2, then a polynomial of degree $r_3 = \max\{r_1,r_2\}$ is a solution of (3-21). First of all, the boundary condition of (3-21) can be satisfied with such a polynomial. Secondly, since the plant equation is linear, $\mathcal{L}(t,\theta(t))E(t,x)$ is a polynomial of degree r_3.

Thus both sides of (3-21) is a polynomial of degree r_3 and $E(t,x)$ can be represented by a finite number of ordinary differential equations of the coefficients of powers of x.

As is well known, [7] (3-20) can be represented by a finite number of ordinary differential equations.

3.8. Deterministic System

The equations of 3.6. are applied to the deterministic system to relate the a priori stochastic Hamiltonian, operand, and maximum principle to the deterministic

Hamiltonian, the Lagrange multiplier, and Pontryagin's maximum principle. [6] Because there is no dynamical noise, $G(t,x(t)) = 0$.

Let $c_0 \triangleq x(0)$, $c(t) \triangleq x(t)$, and $c(T)$ be free. Let $u(t)$ be the optimal control function.

Then

$$\dot{c}(t) = m(t,c(t),u(t)) \quad ,$$

$$Q(t,x) = \delta(x-c(t))$$

and

$$\Upsilon(x) = \delta(x-c_0) \; .$$

Define

$$\lambda(t) \triangleq \int \nabla_x E(t,x) Q(t,x) dx = \nabla_c E(t,c) \; .$$

From (3-21a) it follows that

$$\frac{\partial}{\partial t} \nabla_x E(t,x) = -\nabla_x f(t,x,u(t)) - \nabla_x \mathcal{L} E(t,x)$$

$$= -\nabla_x f(t,x,u(t)) - [\nabla_x m(t,x,u(t))^T] \nabla_x E(t,x)$$

$$-[\nabla_x \nabla_x^T E(t,x)] m(t,x,u(t)) \; . \quad (3\text{-}23)$$

Therefore,

$$\dot{\lambda}(t) = \int\left[\frac{\partial}{\partial t}\,\nabla_x E(t,x)\right]Q(t,x)dx + \int\nabla_x E(t,x)Q_t(t,x)dx \ . \qquad (3\text{-}24)$$

But by (3-23)

$$\int\left[\frac{\partial}{\partial t}\,\nabla_x E(t,x)\right]Q(t,x)dx$$

$$= -\nabla_c f(t,c,u(t)) - [\nabla_c m(t,c,u(t))^T]\nabla_c E(t,c)$$

$$- [\nabla_c \nabla_c{}^T E(t,c)]m(t,c,u(t)) \ , \qquad (3\text{-}25)$$

and by (3-20a)

$$\int\nabla_x E(t,x)Q_t(t,x)dx = \int\nabla_x E(t,x)\mathcal{L}_+ Q(t,x)dx$$

$$= \int\mathcal{L}[\nabla_x E(t,x)]Q(t,x)dx$$

$$= \int[\nabla_x \nabla_x{}^T E(t,x)]m(t,x,u(t))Q(t,x)dx$$

$$= [\nabla_c \nabla_c{}^T E(t,c)]m(t,c,u(t)) \ . \qquad (3\text{-}26)$$

Thus by (3-24) through (3-26) ,

$$\dot{\lambda}(t) = -\nabla_c f(t,c,u(t)) - [\nabla_c m(t,c,u(t))^T]\nabla_c E(t,c)$$

$$= -\nabla_c f(t,c,u(t)) - \nabla_c m(t,c,u(t))^T \lambda(t) . \qquad (3\text{-}27)$$

By 3-22a the a priori stochastic Hamiltonian

$$H(t,Q,E,u(t)) = \int f(t,x,u(t))Q(t,x)dx + \int \mathcal{L}E(t,x)Q(t,x)dx$$

$$= f(t,c,u(t)) + m(t,c,u(t))^T \nabla_c E(t,c)$$

$$= f(t,c,u(t)) + m(t,c,u(t))^T \lambda(t) \qquad (3\text{-}28)$$

equals the deterministic Hamiltonian [6] $H(t,c,\lambda,u(t))$ where $\lambda(t)$ is the Lagrange multiplier [6].

The a priori stochastic maximum principal $H(t,Q,E,u(t)) = \underset{v(t)}{\text{Min}}\; H(t,Q,E,v(t))$ is equivalent to $H(t,c,\lambda,u(t)) = \underset{v(t)}{\text{Min}}\; H(t,c,\lambda,v(t))$, which is Pontryagin's maximum principle [6].

Since

$$\nabla_c H(t,c,\lambda,u(t)) = \nabla_c f(t,c,u(t)) + \nabla_c m(t,c,u(t))^T \lambda(t) ,$$

and $E(T,x) = g(x)$, it follows from (3-27) and (3-28) that

$$\dot{c}(t) = \nabla_{\lambda} H(t,c(t),\lambda(t),u(t)) \qquad c(0) = c_0$$

and

$$\dot{\lambda}(t) = -\nabla_c H(t,c(t),\lambda(t),u(t)) \qquad \lambda(T) = \nabla_c g(c(T))$$

which are the canonical equations [6].

Thus the optimal open loop control solution implies the deterministic solution. With no dynamical noise, the expected value of the gradient of the a priori stochastic operand is the Lagrange multiplier, which is the gradient of the imbedded performance index $E(t,c)$, the a priori stochastic Hamiltonian is identical to the deterministic Hamiltonian, and the a priori maximum principle is Pontryagin's maximum principle.

Since $E(t,x)$ in (3-28) is operated on by $\mathcal{L}$, it was named an operand rather than a multiplier.

3.9. Canonical Equations in Function Space

Before concluding this chapter, it should be noted that the a priori density function and the a priori stochastic operand satisfy the canonical equations in function space as mentioned by Lur'e [36], Mortensen [33,34] and Wang [37].

But first by the Riesz Representation theorem [15], there exists a functional $L(\varphi,x)$ for the linear operator in (1-13) such that

$$F'(\varphi)(\psi) = \int L(\varphi,x)\psi(x)dx \tag{3-29}$$

for all functions ψ. Define

$$\frac{\delta F}{\delta\varphi} \triangleq L(\varphi,x) . \tag{3-30}$$

Then $\frac{\delta F}{\delta\varphi}$ is the <u>Frechet derivature of F at</u> φ. Mortensen discusses and gives examples of the derivature in [34].

Thus by (3-18) with $F(E) = H(t,Q,E,\theta(t))$,

$$F'(E)(\psi) = \int \mathcal{L}\psi(x)Q(t,x)dx$$

$$= \text{by (1-6)} \int \psi(x)\mathcal{L}_+Q(t,x)dx.$$

Consequently by (3-29) and (3-30)

$$\frac{\delta H}{\delta E} = \mathcal{L}_+Q(t,x) . \tag{3-31}$$

Likewise with $F(Q) = H(t,Q,E,\theta(t))$,

$$F'(Q)(\psi) = \int f(t,x,\theta(t))\psi(x)dx + \int \mathcal{L}E(t,x)\psi(x)dx.$$

Similarly, by (3-29) and (3-30)

$$\frac{\delta H}{\delta Q} = f(t,x,\theta(t)) + \mathcal{L}E(t,x) \quad . \tag{3-32}$$

Indeed, by (3-20a), (3-21a), (3-31) and (3-32)

$$Q_t(t,x) = \frac{\delta H}{\delta E}$$

and

$$E_t(t,x) = - \frac{\delta H}{\delta Q}$$

which are the canonical equations in function space for the a priori stochastic optimal control problem.

IV. OPEN LOOP CONTROL LAW

4.1. Introduction

It was shown in 1.5. that the performance index of optimal feedback control is at least as good as that of optimal open loop control. Consequently, the performance index of any suboptimal feedback control law proposed should also be no worse than that of the open loop control, or else the trouble of taking real time observations is wasted.

Such a suboptimal scheme is inspired by the results of studying the optimal open loop control system. Unfortunately, the scheme is not easily implemented. However, an approximation of it is easily implemented and has advantages over the typical suboptimal schemes discussed in 2.6.1. The results of a simple numerical problem, in which the approximate suboptimal control law is superior, is given.

4.2. Definition

While both the optimal open loop and feedback control laws satisfy functional differential equations, the former can be derived from two integro-partial differential equations. Because of the attractiveness of the solution of the open loop problem over that of the closed loop problem, the optimal open loop control law is proposed as a suboptimal feedback control law. [31,32] By applying it to the a posteriori instead of the a priori density function, it will be a feedback control law and henceforth will be referred to as the <u>optimal open loop control law</u>. By (3-7) this suboptimal feedback control law is Ψ, and the suboptimal feedback control $\hat{u}(t) = \Psi(t,P(t,x))$. Alternatively, by 3.6. $\hat{u}(t)$ satisfies

$$H(t,P,E,\hat{u}(t)) = \min_{u(t)} H(t,P,E,u(t))$$

where H is the a priori stochastic Hamiltonian. Thus $\hat{u}(t)$ minimizes

$$H(t,P,E,u(t)) = \int f(t,x,u(t))P(t,x)dx$$

$$+ \int \mathcal{L}(t,u(t))E(t,x,P(t,x))P(t,x)dx \ . \quad (4\text{-}1)$$

Here, $E(t,x,P(t,x))$ can be determined by (3-20b) through (3-22b).

<u>4.3. Performance of the Optimal Open Loop Control Law</u>

In 1.5. it was shown that the performance index, $J(0,T)$, of the optimal feedback control law is less than or equal to $V(0,T)$ of the optimal control function. The same inequality is shown to be true for the performance index, $c(0,T)$, of the optimal open loop control law and $V(0,T)$.

Let $S = \{s_0,s_1,\cdots,s_m\}$ be a partition of $[0,T]$

and $P(0,x;0) \triangleq \Upsilon(x)$. By induction define for $0 \le k \le m-1$ and $s_k \le s \le T$ $\theta(s;k)$, $Q(s,x;k)$, and $P(s,x;k)$ such that : $Q(s_k,x;k) \triangleq P(s_k,x;k)$ and

$Q_s(s,x;k) \triangleq \mathcal{L}_+(s,\theta(s;k))Q(s,x;k)$ where $\theta(s;k)$ minimizes $H(s,Q(s,x;k),E(s,x,Q(s,x;k)),u(t))$ (thus, $\theta(s;k)$ is the optimal control function for a process which starts at time s_k in a state $P(s_k,x;k)$; $Q(s,x;k)$ is the corresponding a priori density function; in particular, $\theta(s;0)$ and $Q(s,x;0)$ equal $\theta(s)$ and $Q(s,x)$, respectively, of 3.6.); $P(s,x;k)$ satisfies (1-7) with $u(s) = \theta(s;k)$, and $P(s_{k+1},x;k+1) \triangleq P(s_{k+1},x;k)$ (hence $P(s,x;k)$ is the conditional density function driven by the control function $\theta(s;k)$).

Define for $0 \le \ell \le m$

$$B(S,\ell) \triangleq \sum_{k=0}^{\ell-1} \int_{s_k}^{s_{k+1}} \int f(s,x,\theta(s;k))P(s,x;k)dxds$$

$$+\int_{s_\ell}^{s_m} \int f(s,x,\theta(s;\ell))P(s,x;\ell)dxds$$

$$+ \int g(x)P(s_m,x;\ell)dx$$

Then $\underset{0\le\tau\le T}{\overset{\text{Exp}}{y(\tau)}} B(S,0)$ is the performance index of the optimal control function, and $\underset{0\le\tau\le T}{\overset{\text{Exp}}{y(\tau)}} \underset{S\subset[0,T]}{\overset{\text{lim}}{}} B(S,m)$ is the performance index of the optimal open loop control law. Therefore for $\ell<m$

$$B(S,\ell) - B(S,\ell+1) = \int_{s_\ell}^{s_m}\int f(s,x,\theta(s;\ell))P(s,x;\ell)dxds$$

$$+ \int g(x)P(s_m,x;\ell)dx - \left[\int_{s_\ell}^{s_{\ell+1}}\int f(s,x,\theta(s;\ell))P(s,x;\ell)dxds\right.$$

$$+ \int_{s_{\ell+1}}^{s_m}\int f(s,x,\theta(s;\ell+1))P(s,x;\ell+1)dxds$$

$$\left.+ \int g(x)P(s_m,x;\ell+1)dx\right]$$

$$= \int_{s_{\ell+1}}^{s_m}\int f(s,x,\theta(s;\ell))P(s,x;\ell)dxds + \int g(x)P(s_m,x;\ell)dx$$

$$-\left[\int_{s_{\ell+1}}^{s_m}\int f(s,x,\theta(s;\ell+1))P(s,x;\ell+1)dxds\right.$$

$$\left.+ \int g(x)P(s_m,x;\ell+1)dx\right] .$$

For $1\le\ell\le m$ define $Q(s,x:\ell)$ to satisfy (3-1) with $u(s) = \theta(s;\ell-1)$ and $Q(s_\ell, x:\ell) = P(s_\ell, x;\ell)$.

Then by Appendices A and B

$$\underset{s_{\ell+1}\le\tau\le T}{\underset{y(\tau)}{\mathrm{Exp}}} [B(S,\ell) - B(S,\ell+1)]$$

$$= \int_{s_{\ell+1}}^{s_m} \int f(s,x,\theta(s;\ell))Q(s,x:\ell+1)dxds + \int g(x)Q(s_m,x:\ell+1)dx$$

$$-\left[\int_{s_{\ell+1}}^{s_m} \int f(s,x,\theta(s;\ell+1))Q(s,x;\ell+1)dxds\right.$$

$$\left. + \int g(x)Q(s_m,x;\ell+1)dx\right]$$

$$= \int_{s_{\ell+1}}^{s_m} \int f(s,x,\theta(s;\ell))Q(s,x:\ell+1)dxds + \int g(x)Q(s_m,x:\ell+1)dx$$

$$-V(s_{\ell+1}, P(s_{\ell+1},x;\ell+1)) \quad . \tag{4-2}$$

Since the minimum of (3-3) with $T = s_m$, $t = s_{\ell+1}$ and $Q(t,x) = P(s_{\ell+1},x;\ell+1)$ is $V(s_{\ell+1}, P(s_{\ell+1},x;\ell+1))$, then by (4-2)

$$\underset{s_{\ell+1}\le\tau\le T}{\underset{y(\tau)}{\mathrm{Exp}}} [B(S,\ell) - B(S,\ell+1)] \ge 0 \quad .$$

Hence $\underset{0\le\tau\le T}{\overset{\text{Exp}}{y(\tau)}} [B(S,\ell) - B(S,\ell+1)] \ge 0$.

Thus

$$V(0,T) = \underset{0\le\tau\le T}{\overset{\text{Exp}}{y(\tau)}} B(S,0)$$

$$= \underset{0\le\tau\le T}{\overset{\text{Exp}}{y(\tau)}} \left\{ \sum_{\ell=0}^{m-1} [B(S,\ell) - B(S,\ell+1)] + B(S,m) \right\}$$

$$\ge \underset{0\le\tau\le T}{\overset{\text{Exp}}{y(\tau)}} B(S,m) \quad ,$$

and hence

$$V(0,T) \ge \underset{S\subset[0,T]}{\text{lim}} \underset{0\le\tau\le T}{\overset{\text{Exp}}{y(\tau)}} B(S,m)$$

$$= \underset{0\le\tau\le T}{\overset{\text{Exp}}{y(\tau)}} \underset{S\subset[0,T]}{\text{lim}} B(S,m)$$

$$= C(0,T) \quad . \quad \blacksquare$$

As hoped for the performance of the optimal open loop control law is at least as good as that of the optimal control function. The performance indices of specific forms of the model of the system in nondecreasing order are those of the deterministic system driven by its optimal control function, of the stochastic system with perfect observations driven by the optimal feedback control law and of the stochastic systems driven by the optimal feedback control law, driven by the optimal open loop control law and driven by its optimal control function.

Thus an upper bound of the ratio of the performance index of the optimal open loop control law to that of the optimal feedback control law is the ratio of the performance index of the optimal control function of the stochastic system to that of the deterministic system, the latter pair being the easier to compute.

Also, as discussed in 3.7., when the plant equation is linear and f and g are polynomials in x, then the optimal open loop control law reduces to a finite system of ordinary differential equations.

4.4. Linear System with a Quadratic Performance Index

Since (1-18) reduces to a finite system of ordinary differential equations when the system is linear and the performance index is quadratic, a criterion for any suboptimal feedback control law is that it satisfy this optimal result. With such a system defined in 2.3., this property is shown true for the optimal open loop control law. [31]

By (4-1) $\hat{u}(t)$ minimizes

$$H(t,\varphi,E,u(t))$$

$$= \int \tfrac{1}{2}[x^T Q(t)x + u(t)^T C(t)u(t)]\varphi(x)dx$$

$$+ \int \{ [A(t)x + B(t)u(t)]^T \nabla_x E(t,x,\varphi)$$

$$+ \tfrac{1}{2}\mathrm{trace}[G(t)G(t)^T \nabla_x \nabla_x^T E(t,x,\varphi)] \} \varphi(x)dx \ . \tag{4-3}$$

Differentiating (4-3) with respect to $u(t)$ yields

$$C(t)u(t) + B(t)^T \int \nabla_x E(t,x,\varphi)\varphi(x)dx$$

which implies

$$\hat{u}(t) = -C^{-1}(t)B(t)^T \int \nabla_x E(t,x,\varphi)\varphi(x)dx \ . \tag{4-4}$$

A quadratic form for the a priori stochastic operand is attempted, i.e.,

$$E(t,x,\varphi) = \tfrac{1}{2}x^T D_1(t)x + x^T D_2(t)\int x\varphi(x)dx + D_3(t,\varphi)$$

where $D_1(t)$ and $D_2(t)$ are n x n symmetrical matrices.

Thus the initial condition $E(T,x,\varphi) = \frac{1}{2}x^{\intercal}Sx$ implies $D_1(T) = S$, $D_2(T) = 0$ and $D_3(T,\varphi) = 0$.

Also,

$$\nabla_x E(t,x,\varphi) = D_1(t)x + D_2(t)\int x\varphi(x)dx ,$$

$$\nabla_x \nabla_x{}^{\intercal} E(t,x,\varphi) = D_1(t) ,$$

and

$$E'(t,x,\varphi)(\psi) = x^{\intercal}D_2(t)\int x\psi(x)dx + D_3{}'(t,\varphi)(\psi) .$$

Consequently, by (4-4)

$$\hat{u}(t) = -C^{-1}(t)B(t)^{\intercal}[D_1(t) + D_2(t)]\int x\varphi(x)dx . \tag{4-5}$$

Equation (3-13) implies

$$\begin{aligned}
&\tfrac{1}{2}x^{\intercal}\dot{D}_1(t)x + x^{\intercal}\dot{D}_2(t)\int x\varphi(x)dx + D_{3_t}(t,\varphi) \\
&= -x^{\intercal}D_2(t)\int x\mathcal{L}_{+}\varphi(x)dx - D_3{}'(t,\varphi)(\mathcal{L}_{+}\varphi) - \tfrac{1}{2}x^{\intercal}Q(t)x \\
&\quad -\tfrac{1}{2}\int x^{\intercal}\varphi(x)dx[D_1(t) + D_2(t)]B(t)C^{-1}(t)B(t)^{\intercal}[D_1(t) \\
&\quad + D_2(t)]\int x\varphi(x)dx - \{A(t)x - B(t)C^{-1}(t)B(t)^{\intercal}[D_1(t) \\
&\quad + D_2(t)]\int x\varphi(x)dx\}^{\intercal}[D_1(t)x + D_2(t)\int x\varphi(x)dx \\
&\quad - \tfrac{1}{2}\mathrm{trace}[G(t)G(t)^{\intercal}D_1(t)] .
\end{aligned} \tag{4-6}$$

Notice that by (1-6)

$$\int x\mathcal{L}_{+}\varphi(x)dx = \int \mathcal{L}x\varphi(x)dx$$

$$= \{A(t)-B(t)C^{-1}(t)B(t)'[D_1(t) + D_2(t)]\}\int x\varphi(x)dx \; .$$

Equating coefficients of like powers of x in (4-6) yields for $D_1(t)$ and $D_2(t)$

$$\dot{D}_1(t) = -Q(t) - A(t)'D_1(t) - D_1(t)A(t) \tag{4-7}$$

$$\dot{D}_2(t)\int x\varphi(x)dx$$

$$= \{-D_2(t)A(t) + D_2(t)B(t)C^{-1}(t)B(t)'[D_1(t) + D_2(t)]$$

$$-A(t)'D_2(t) + D_1(t)B(t)C^{-1}(t)B(t)'[D_1(t)$$

$$+D_2(t)]\}\int x\varphi(x)dx \; . \tag{4-8}$$

The functional equation for $D_3(t,\varphi)$ is irrelevant to the control law.

Since (4-8) holds for any $\varphi(x)$, $\int x\varphi(x)dx$ may be cancelled from both sides of the equation.

Let $U(t) = D_1(t) + D_2(t)$. Then by (4-5)

$$\hat{u}(t) = -C^{-1}(t)B(t)'U(t)\int x\varphi(x)dx , \tag{4-9}$$

and by (4-7) and (4-8) $U(t)$ satisfies (2-6) .

Consequently, the optimal open loop control law (4-9) is the same as (2-7) , the optimal closed loop control law derived by Wonham [2]. Thus for one of the limited class of systems for which (1-18) can be represented as a finite system of ordinary differential equations, the optimal open loop control law produces the optimal performance index.

4.5. Approximation of the Optimal Open Loop Control Law

The optimal open loop control law is derived after the a priori stochastic operand $E(t,x,P(t,x))$ is determined. For each $P(t,x)$, $t\epsilon[0,T]$ the two-point boundary valued equations of (3-20b) and (3-21b) must be solved. Because this is a very tedious task, an approximation of $E(t,x,P(t,x))$ is desired.

Let $F(d(t),x,\varphi)$ be linear in the vector $d(t)$, which is chosen to minimize

$$\int \left[F(d(t),x,Q(t,x)) - E(t,x,Q(t,x)) \right]^2 Q(t,x)dx , \tag{4-10}$$

where the a priori density function $Q(t,x)$ satisfies

(3-1) with $u(t) = \theta(t) = \Psi(t,Q(t,x))$, which is defined in (3-7) .

Equation (4-10) is minimal when

$$\int \Big[F(d(t),x,Q(t,x)) - E(t,x,Q(t,x))\Big]$$

$$\cdot \nabla_d F(d(t),x,Q(t,x))Q(t,x)dx = 0 \ . \qquad (4\text{-}11)$$

Differentiating (4-11) with respect to t yields by (3-13) and by the linearity of $d(t)$ in F

$$\int \Big\{ \Big[[\nabla_d F(d(t),x,Q(t,x))]' \dot{d}(t) + F'(d(t),x,Q(t,x))(\mathcal{L}_+ Q(t,x))$$

$$+ f(t,x,\theta(t)) + \mathcal{L}E(t,x,Q(t,x))\Big] \nabla_d F(d(t),x,Q(t,x))Q(t,x)$$

$$+ [F(d(t),x,Q(t,x)) - E(t,x,Q(t,x))]$$

$$\cdot [\nabla_d F'(d(t),x,Q(t,x))(\mathcal{L}_+ Q(t,x))Q(t,x)$$

$$+ \nabla_d F(d(t),x,Q(t,x))\mathcal{L}_+ Q(t,x)] \Big\} dx = 0 \ . \qquad (4\text{-}12)$$

By applying the adjoint of $\mathcal{L}_+$ as in (1-6) , and noting that for scalar functions $A(x)$ and $B(x)$

$$\mathcal{L}[A(x)B(x)] = A(x)\mathcal{L}B(x) + B(x)\mathcal{L}A(x) + \mathcal{L}_p(A(x),B(x))$$

where

$$\mathcal{L}_p(A(x),B(x)) \triangleq [\nabla_x A(x)]^T G(t,x)G(t,x)^T \nabla_x B(x) ,$$

(4-12) becomes

$$\int \nabla_d F(d(t),x,Q(t,x))[\nabla_d F(d(t),x,Q(t,x))]^T Q(t,x)dx\dot{d}(t)$$

$$= -\int\Big\{\Big[F'(d(t),x,Q(t,x))(\mathcal{L}_+Q(t,x)) + f(t,x,\theta(t))$$

$$+ \mathcal{L}F(d(t),x,Q(t,x))\Big]\nabla_d F(d(t),x,Q(t,x))$$

$$+ [F(d(t),x,Q(t,x))-E(t,x,Q(t,x))]$$

$$\cdot [\nabla_d F'(d(t),x,Q(t,x))(\mathcal{L}_+Q(t,x)) + \mathcal{L}\nabla_d F(d(t),x,Q(t,x))]$$

$$+ \mathcal{L}_p[F(d(t),x,Q(t,x))-E(t,x,Q(t,x)),\nabla_d F(d(t),x,Q(t,x))]\Big\}$$

$$\cdot Q(t,x)dx . \qquad (4\text{-}13)$$

If $E(t,x,Q(t,x))$ is approximated by $F(d(t),x,Q(t,x))$, (4-13) simplifies to

$$\int \nabla_d F(d(t),x,Q(t,x))[\nabla_d F(d(t),x,Q(t,x))]' Q(t,x)dx\dot{d}(t)$$

$$= -\int [F'(d(t),x,Q(t,x))(\mathcal{L}_+(t,\theta(t))Q(t,x)) + f(t,x,\theta(t))$$

$$+ \mathcal{L}(t,\theta(t))F(d(t),x,Q(t,x))]\nabla_d F(d(t),x,Q(t,x))Q(t,x)dx$$

with

$$\int \left\{ F(d(T),x,Q(T,x))-g(x) \right\} \nabla_d F(d(T),x,Q(T,x))Q(T,x)dx=0. \quad (4\text{-}14)$$

In 2.4. $J(t,\varphi)$ is approximated by a quadratic polynomial in $\int x\varphi(x)dx$. Analogously, $E(t,x,\varphi)$ is approximated by a quadratic polynomial in x :

$$F(d(t),x,\varphi) \triangleq \tfrac{1}{2}x'D_1(t)x + x'D_2(t)\int x\varphi(x)dx + D_3(t) \quad (4\text{-}15)$$

where $D_1(t)$ and $D_2(t)$ are $n \times n$ matrices, the former being symmetrical; $d(t)$ is the $\frac{3n^2+n+2}{2}$ - vector of the components of $D_1(t)$, $D_2(t)$ and $D_3(t)$, where n is the order of the plant equation.

Note that by (1-6) $F'(d(t),x,\varphi)(\mathcal{L}_+\varphi)$ in (4-15) equals

$$x^{T}D_2(t)\int \mathfrak{L}x\varphi(x)dx = x^{T}D_2(t)\int m(t,x,\theta(t))\varphi(x)dx \ .$$

An approximation of the optimal open loop control law is now defined. As in 2.6. $P(t,x)$ is approximated by $\hat{P}(t,\hat{\mu}(t),x)$ which is defined in (2-9). Let $\hat{\hat{\mu}}(t)$ be the a priori estimate of $\hat{\mu}(t)$ such that

$$\hat{Q}(t,x) \triangleq \begin{cases} N(x,\hat{\hat{\mu}}(t),\hat{\hat{M}}(t)) \\ \qquad \text{or} \\ \delta(x-\hat{\hat{\mu}}(t)) \end{cases}$$

Next, define $\hat{\theta}(t)$ as the approximate control function which minimizes $H(t,\hat{Q}(t,x),F(d(t),x,\hat{Q}(t,x)),u(t))$, where H is defined in (3-18) and F in (4-15). With Q and θ replaced by $\hat{Q}$ and $\hat{\theta}$, respectively, in (4-14), $d(t)$, that is, $D_1(t)$, $D_2(t)$ and $D_3(t)$, are determined a priori.

Finally, define $\hat{\hat{u}}(t)$ to be the control of the suboptimal open loop control law such that

$$\hat{\hat{u}}(t) \quad \text{minimizes} \quad H(t,\hat{P}(t,\hat{\mu}(t),x),F(d(t),x,\hat{P}(t,\hat{\mu}(t),x),u(t)) \ .$$

Since the function d can be determined a priori, only the evolution of the nonlinear filter $\hat{P}(t,x)$ has to be determined in real time to derive the suboptimal open loop control $\hat{\hat{u}}(t)$. As can be seen from (4-11), (4-14) and 4.4., $\hat{\hat{u}}(t)$ is the optimal feedback control (4-9) when the system is linear and the performance index is quadratic. Thus the suboptimal open loop control law has some of the properties of the conventional suboptimal feedback control law of (2-14).

An advantage the suboptimal open loop control law has over the conventional suboptimal feedback control law is that it can be determined without truncating the Taylor series expansions of m, G and f in (4-14). However, a disadvantage is that its parameters $D_1(t)$, $D_2(t)$ and $D_3(t)$ consist of $\frac{3n^2+n+2}{2}$ elements as compared with $\frac{n^2+3n+2}{2}$ elements of $S_1(t)$, $S_2(t)$ and $S_3(t)$ of (2-14), where n is the order of the plant equation.

Before the merits and limitations of the suboptimal open loop control law are given in detail, the optimal and suboptimal open loop control laws are demonstrated in a numerical example.

4.6. An Example

4.6.1. Formulation of the Problem

It is now appropriate to apply the suboptimal open loop control law to a system other than the linear system that has a quadratic performance index and to compare its performance index and computation time with those of the optimal feedback control law and the suboptimal feedback control law of (2-14).

In order to produce a simple system from which to calculate the performance index of the optimal feedback control law, a linear system was chosen:

$$\dot{x}(t) = -x(t) + u(t) + .3\zeta(t) \tag{4-16}$$

$$y(t) = x(t) + .3\eta(t) \ . \tag{4-17}$$

With minimum control u with $|u|<\infty$ the desired performance is to drive the state of the system close to the origin by the time $t=1$ where controls $|u|\geq 1$ are penalized more than controls $|u|<1$. A trade-off type of performance index was defined with

$$f(t,x,u(t)) = u^4(t)$$

and

$g(x) = x^2$.

Equations (2-3), (2-4), (4-16) and (4-17) imply the filter

$$\dot{\mu} = -\mu + u + \frac{100}{9} M(y-\mu) \qquad \mu(0) = \mu_0 \tag{4-18}$$

$$\dot{M} = -2M - \frac{100}{9} M^2 + .09 \qquad M(0) = M_0 \tag{4-19}$$

where $P(t,x) = N(x,\mu(t),M(t))$.

4.6.2. Optimal Control Function

The open loop filter for (4-16) and (4-17) as derived from (3-20a) or, equivalently, from (2-3) and (2-4) with $H(t) = 0$ is

$$\dot{m} = -m + \theta \qquad m(0) = \mu_0 \tag{4-20}$$

$$\dot{P} = -2P + .09 \qquad P(0) = M_0 \ . \tag{4-21}$$

where $Q(t,x) = N(x,m(t),P(t))$.

By 3.7. $E(t,x) = \frac{1}{2}d_1x^2 + d_2x + d_3$ is a solution of (3-21a). Therefore,

$$\tfrac{1}{2}\dot{d}_1x^2 + \dot{d}_2x + \dot{d}_3 = -\theta^4 - (-x+\theta)(d_1x+d_2) - \tfrac{1}{2}(.09)d_1 \tag{4-22}$$

with $d_1(1) = 2$, $d_2(1) = 0$, and $d_3(1) = 0$.

By (3-22a)

$$\frac{\partial}{\partial\theta}\int\left\{\theta^4 + (-x+\theta)(d_1x+d_2) + \tfrac{1}{2}(.09)d_1\right\}N(x,m,P)dx = 0 .$$

Consequently, the optimal control function

$$\theta = -\left(\frac{d_1m + d_2}{4}\right)^{1/3} .$$

Thus, by (4-22) the optimal open loop system is represented by the following differential equations

$$\dot{m} = -m-\left(\frac{d_1m + d_2}{4}\right)^{1/3} \qquad m(0) = \mu_0$$

$$\dot{P} = -2P + .09 \qquad P(0) = M_0$$

$$\dot{d}_1 = 2d_1 \qquad d_1(1) = 2$$

$$\dot{d}_2 = d_1\left(\frac{d_1m + d_2}{4}\right)^{1/3} + d_2 \qquad d_2(1) = 0$$

$$\dot{d}_3 = -\left(\frac{d_1m + d_2}{4}\right)^{4/3} + d_2\left(\frac{d_1m + d_2}{4}\right)^{1/3} - .045d_1$$

with $d_3(1) = 0$.

The performance index is

$$\int E(0,x)N(x,\mu_0,M_0)dx = \tfrac{1}{2}d_1(0)(\mu_0{}^2+M_0) + d_2(0)\mu_0 + d_3(0) .$$

4.6.3. Suboptimal Open Loop Control Law

With the open loop filter (4-20) and (4-21) apply

$$F(t,x,Q(t,x)) = \tfrac{1}{2}d_1x^2 + d_2x\int xQ(t,x)dx + d_3$$

to (4-14) or equivalently, try it as a solution of (3-21a). Then

$$\theta = -\left(\frac{(d_1+d_2)m}{4}\right)^{1/3}$$

$$\dot{m} = -m-\left(\frac{(d_1+d_2)m}{4}\right)^{1/3} \qquad m(0) = \mu_0$$

$$\dot{P} = -2P + .09 \qquad P(0) = M_0$$

$$\dot{d}_1 = 2d_1 \qquad d_1(1) = 2$$

$$\dot{d}_2 = 2d_2+\left[\frac{(d_1+d_2)^4}{4m^2}\right]^{1/3} \qquad d_2(1) = 0$$

$$\dot{d}_3 = -\left(\frac{(d_1+d_2)m}{4}\right)^{4/3} - md_2\left(\frac{(d_1+d_2)m}{4}\right)^{1/3} - .045d_1$$

with $d_3(1) = 0$.

The suboptimal open loop control law is

$$\hat{\mu}(t) = -\left(\frac{[d_1(t)+d_2(t)]\mu(t)}{4}\right)^{1/3} .$$

The performance index $L(0,\mu_0)$ of this control law is derived from 2.4. where

$$-L_t(t,\mu) = \left[\frac{(d_1+d_2)\mu}{4}\right]^{4/3} + L_\mu(t,\mu)\left\{-\mu-\left[\frac{(d_1+d_2)\mu}{4}\right]^{1/3}\right\}$$

$$+ \frac{50}{9} M^2(t)L_{\mu\mu}(t,\mu) \qquad\qquad L(1,\mu) = \mu^2+M(1).$$

<u>4.6.4. Optimal Open Loop Control Law</u>

By (4-1) the control $\hat{u}(t)$ of the optimal open loop control law minimizes $H(t,N(x,\mu,M),E,u(t))$. With the solutions of 4.6.2. and with (3-20b), (3-21b) and (3-22b), the optimal open loop control law is represented by the following system of differential equations

$$m_s(t,s) = -m(t,s) -\left(\frac{d_1(t,s)m(t,s)+d_2(t,x)}{4}\right)^{1/3}$$

with $m(t,t) = \mu(t)$

$$P_s(t,s) = -2P(t,s) + .09 \qquad P(t,t) = M(t)$$

$$d_{1_s}(t,s) = 2d_1(t,s) \qquad d_1(t,1) = 2$$

$$d_{2_s}(t,s) = d_1(t,s)\left(\frac{d_1(t,s)m(t,s)+d_2(t,s)}{4}\right)^{1/3} + d_2(t,s)$$

with $d_2(t,1) = 0$

$$d_{3_s}(t,s) = -\left(\frac{d_1(t,s)m(t,s)+d_2(t,s)}{4}\right)^{4/3}$$

$$+ d_2(t,s)\left(\frac{d_1(t,s)m(t,s)+d_2(t,s)}{4}\right)^{1/3} -.045d_1(t,s)$$

with $d_3(t,1) = 0$

where

$$\hat{\mu}(t) = -\left(\frac{d_1(t,t)\mu(t)+d_2(t,t)}{4}\right)^{1/3} .$$

From 2.4. and 4.2. $\hat{\mu}(t) = -[\frac{1}{4}W_\mu(t,\mu(t))]^{1/3}$

where the performance index of the optimal control function satisfies

$$-W_t(t,\mu) = -3(\tfrac{1}{4}W_\mu(t,\mu))^{1/3} - \mu W_\mu(t,\mu) \qquad (4\text{-}23)$$

with $W(1,\mu) = \mu^2 + P(0,1)$.

Similarly, by 2.4. the performance index $L(0,\mu_0)$ of the optimal open loop dontrol law satisfies

$$-L_t(t,\mu) = (\tfrac{1}{4}W_\mu(t,\mu))^{4/3} + L_\mu(t,\mu)[-\mu-(\tfrac{1}{4}W_\mu(t,\mu))^{1/3}]$$

$$+ \frac{50}{9} M^2(t)L_{\mu\mu}(t,\mu) \qquad L(1,\mu) = \mu^2 + M(1).$$

4.6.5. Optimal Feedback Control Law

By (1-18) and 2.4. the optimal feedback control law is

$$u(t) = -(\tfrac{1}{4}L_\mu(t,\mu(t)))^{1/3} ,$$

and its performance index $L(0,\mu_0)$ satisfies

$$-L_t(t,\mu) = -3(\tfrac{1}{4}L_\mu(t,\mu))^{4/3} - \mu L_\mu(t,\mu)$$

$$+ \frac{50}{9} M^2(t)L_{\mu\mu}(t,\mu) \qquad L(1,\mu) = \mu^2 + M(1). \quad (4\text{-}24)$$

4.6.6. Optimal Feedback Control Law with Perfect Observations

By (2-2) the optimal feedback control law of a stochastic system with perfect observations is

$$u(t) = -(\tfrac{1}{4}F_y(t,y(t)))^{1/3}$$

where its performance index $F(0,\mu_0)$ satisfies

$$-F_t(t,y) = -3(\tfrac{1}{4}F_y(t,y))^{4/3} - yF_y(t,y) + .045F_{yy}(t,y)$$

with $F(1,y) = y^2$.

4.6.7. Results

The comparison of optimal control of specific forms of the model of the system of 1.2. is presented in Table 1. There the mean computational time for calculating the control is real time needed on the IBM 360/75 computer for one increment; the interval [0,1] was divided into 100 steps. Because of the cube root in (4-24) the suboptimal scheme of (2-14) cannot be applied.

The deterministic and open loop performance indices depend on only five ordinary differential equations; there was no problem in the convergence of their solutions. However, the other performance indices

TABLE 1.

	$\mu_0=1.$ $M_0=0.$ $m(1)=.14586$ $M(1)=.034724$		$\mu_0=1.$ $M_0=1.$ $m(1)=.14586$ $M(1)=.049762$	
	I	II	I	II
Deterministic System with Optimal Control Function	.039304	.01	.039304	.01
Stochastic System with Perfect Observations with Optimal Feedback Control Law	.067949	60	.067949	60
Stochastic System with Optimal Feedback Control Law	.074175	60	.16525	60
Stochastic System with Optimal Open Loop Control Law	.074175	.5	.16614	.5
Stochastic System with Suboptimal Open Loop Control Law	.074186	.01	.16759	.01
Stochastic System with Optimal Control Function	.078335	.01	.21367	.01

I: Performance Index

II: Mean Computational Time for Control (seconds) are represented by an infinite number of ordinary differential equations. With 87.5^K words of memory in the computer, they could be represented by only nine differential equations. Thus their values are not very accurate. However, the performance indices of the optimal control function calculated via (4-23) differed by less than 5% from those determined from 4.6.2. Consequently, the calculations of the performance indices represented by an infinite sequence are estimated to be 5% in error.

Since the performance index of a suboptimal feedback control law must be better than that of the optimal control function and is desired to be close to that of the optimal feedback control law, a measure of the closeness is defined.

Let $C(0,T)$ be the performance index of a feedback control law. Recall that $J(0,T)$ and $V(0,T)$ are the performance indices of the optimal feedback control law and the optimal control function, respectively. Define the ratio

$$\frac{C(0,T) - J(0,T)}{V(0,T) - J(0,T)} \tag{4-25}$$

to be the <u>deficiency</u> of the feedback control law. Consequently, the optimal deficiency is zero with the worst deficiency being 1.

Therefore, for $\mu_0 = 1$ and $M_0 = 0$, the deficiency of the suboptimal open loop control law is .265% whereas its computation is .0167% of that of the optimal feedback control law. For the other run with $\mu_0 = 1$ and $M_0 = 1$, the deficiency is 4.84% with the same computational advantage. Sample runs of the suboptimal open loop control law are graphically presented in Figure 2. The a priori functions d_1 and d_2 of 4.6.3., which constitute the suboptimal open loop control law and the optimal open loop mean and control function, are shown in Figure 1.

As shown in Table 1, the optimal open loop control law has a better performance index for this example than does its approximation. However, its computation is about 50 times as much.

The purpose of this example is to compare the performance indices of the suboptimal and optimal open loop control laws with that of the optimal feedback control law. It was shown that the former are superior to the conventional suboptimal control law (2-14) in that the latter cannot be applied. All results of this chapter are now summarized.

4.7. Conclusions

The discussion in 2.6.2. show that the suboptimal

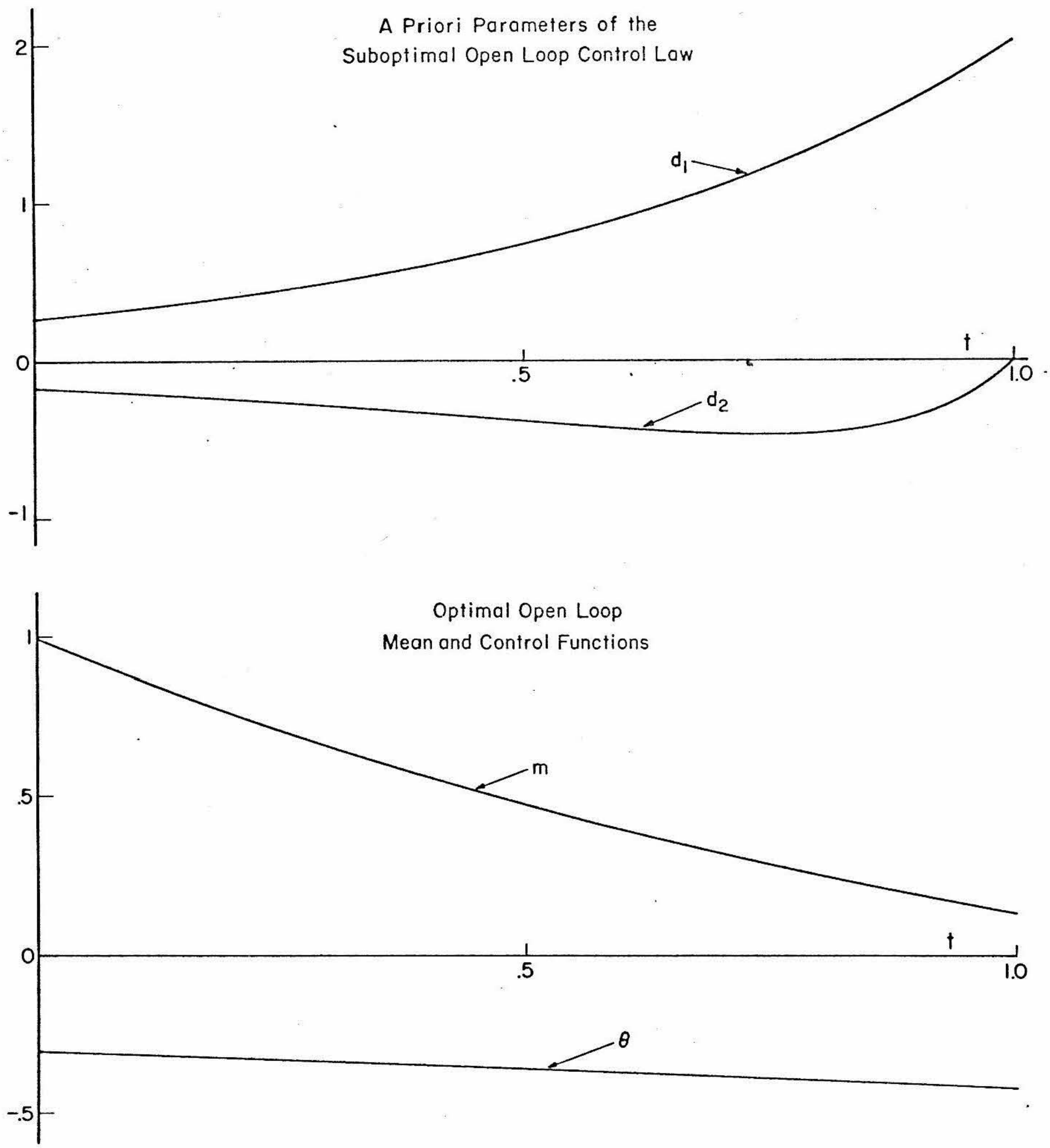
A Priori Parameters of the
Suboptimal Open Loop Control Law
2
1
0
-1
.5
1.0
t
d1
d2
Optimal Open Loop
Mean and Control Functions
1
.5
0
-.5
.5
1.0
t
m
θ

Figure 1.

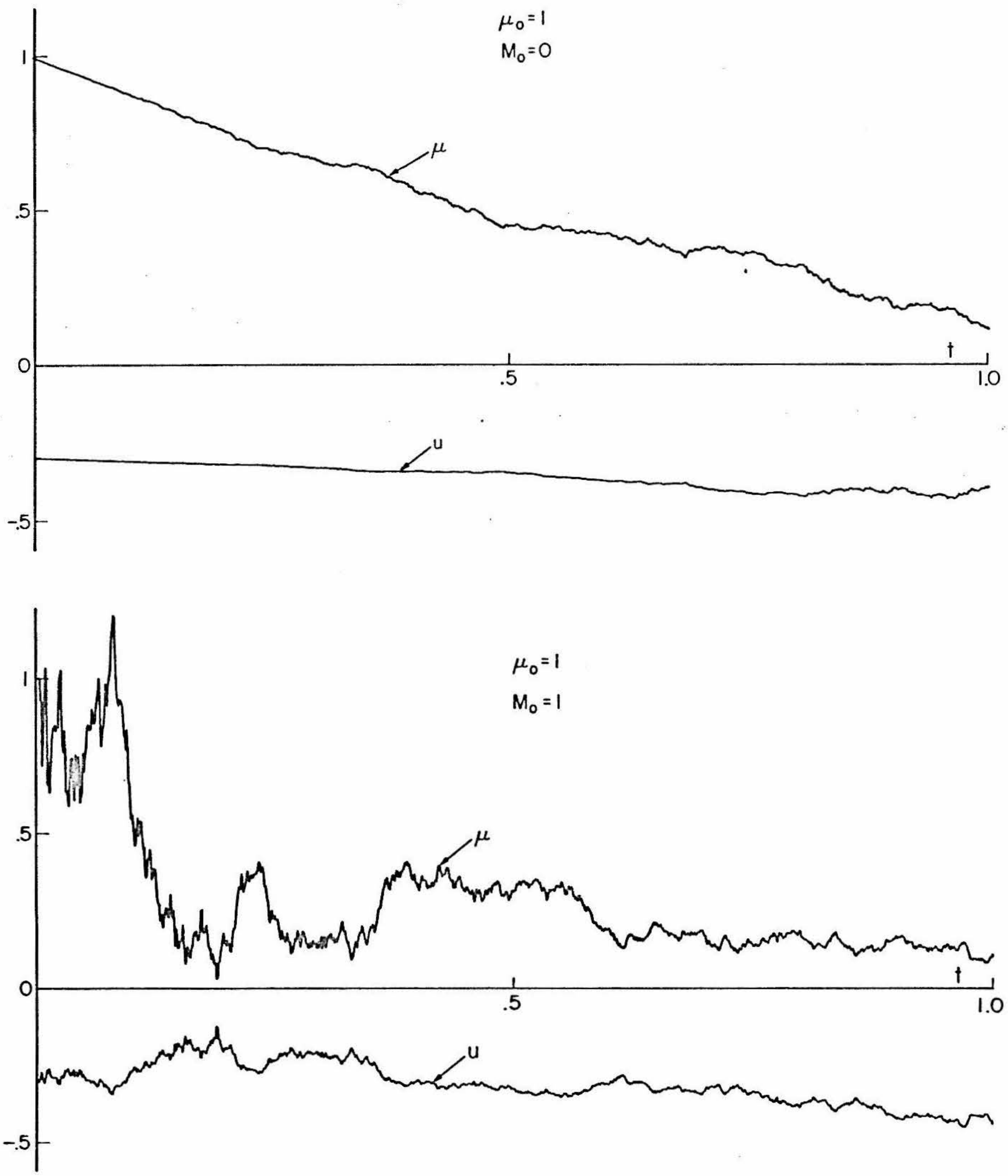

Figure 2. Sample Mean and Control Functions Generated by the Suboptimal Open Loop Control Law

feedback control law of (2-14) can be derived from the solution of the optimal open loop control problem when the system has linear observations. Consequently, the optimal open loop control law was studied in terms of a suboptimal feedback control law. As desired, its performance index was shown to be at least as good as that of the optimal control function. Because this control law is not easy to implement, an approximation of it was defined and is called the suboptimal open loop control law. The resulting control law is the main contribution of this investigation to stochastic control theory. Its properties are compared with the conventional suboptimal feedback control law of (2-14).

Both the suboptimal open loop control law and that of (2-14) are feedback, both approximate the a posteriori density function by means of a nonlinear filter, both are functions of other parameters that are determined a priori and both are optimal when the system is linear and the performance index is quadratic.

The advantages of the former over the latter is that its a priori parameters can be determined without the necessity of truncating the expansions of m, G and f, that these functions need not be analytic in the state variable and that the control law need not be an analytic function. These advantages stem from the versatility of the a priori stochastic operand in that it is a function of three variables whereas J in (1-18) is a function of only two variables.

However, because of the extra degree of freedom, a trade-off occurs between the two feedback control laws in that the suboptimal open loop control law requires $\frac{3n^2+n+2}{2}$ a priori parameters while the conventional scheme of (2-14) requires only $\frac{n^2+3n+2}{2}$, where n is the order of the plant equation.

In the example of 4.6. the suboptimal open loop control law is demonstrated to be superior to the optimal feedback control law for most systems because of its low deficiency, as defined in (4-25) and short computational time.

V. CONCLUSIONS

This investigation explores an area of stochastic optimal control. Its purpose is to study the solution of the functional differential equation of the optimal feedback control law of a system where the noise processes of the plant and observational equations are additive, independent Gaussian white noise processes. This complex equation arises from the fact that the best description of the output of the plant is a function - the a posteriori density function of the plant variable conditioned on all past observations.

Using an equivalent definition for the criterion of optimal feedback control, the author proved that the optimal feedback performance index is superior to that of the optimal control function. With the solution of the functional differential equation justified, it was applied to a simple first order nonlinear system. However, the solution required an overwhelming computational time of 400 minutes on the IBM 360/75 computer. This result accentuated the futility of implementing optimal feedback control with present techniques. Such a result had been suspected because the conditional density function is in general an infinite dimensional vector.

Consequently, in practice, one approximates the a posteriori density function by means of a nonlinear filter. Based upon this simplification, a suboptimal feedback control law is derived. It is a function of the nonlinear filter and $\frac{n^2+3n+2}{2}$ a priori parameters, where n is the order of the plant equation.

The author showed that the conventional suboptimal feedback control law could be derived from the solution of the optimal open loop control problem when the system has linear observations. Consequently, its solution was studied in hopes of discovering a better suboptimal feedback control law.

An investigation of the optimal open loop control problem brought to light some remarkable results! First, the functional differential equation for the optimal open loop control law was reduced to two two-point boundary valued integro-partial differential equations which satisfy the canonical equations in function space. The existence of these results were discussed by Lur'e [36], Mortensen [34,35] and Wang [37]. More interestingly, when the system is deterministic, these equations reduce to the canonical equations of deterministic optimal control theory [6]. Hence, the author defines the a priori stochastic operand, Hamiltonian and maximum principle which imply the deterministic Lagrange multiplier (costate variable) and Hamiltonian and Pontryagin's maximum principle, respectively. Thus, a priori stochastic optimal control theory, that is, optimal open loop control theory, was found to encompass deterministic optimal control theory in a very interesting way.

A feature of the optimal control function, yet to be shown for the optimal feedback control law, is that it can be derived from a finite system of ordinary differential equations whenever the system is linear and the performance index is a polynomial.

Appropriately, the author studies the optimal open loop control law as a suboptimal control law. [31,32] As a must for suboptimal schemes, its performance index is shown to be at least as good as that of the optimal control function. Because the optimal open loop control law is not easily implemented for all systems, it is approximated by what the author calls the suboptimal open loop control law. This suboptimal feedback control law is the author's main contribution to stochastic control theory. Its advantages and limitations are summarized.

An unwritten necessity satisfied by the suboptimal open loop control law, is that it is optimal when the system is linear and the performance index is quadratic. An advantage it has over the conventional suboptimal feedback control law is that its a priori parameters can be determined without the necessity of truncating the expansion of any nonlinear functions of the system. This feature is enhanced by the dependence of the a priori stochastic operand on three variables instead of only two as in the functional differential equation for the optimal feedback control law. Also, neither the control law nor nonlinear functions need be analytic functions. Its disadvantage over the conventional suboptimal control law is that it is characterized by $\frac{3n^2+n+2}{2}$ a priori parameters whereas the latter requires only $\frac{n^2+3n+2}{2}$, where n is the order of the plant equation. Thus a trade-off between the two suboptimal feedback control laws exists.

Through the use of a numerical example, the performance of the suboptimal open loop control law is shown to be near optimal, and its computation is 6000 times less. The results of this example are particularly noteworthy since the conventional suboptimal control law could not be applied.

Consequently, the concept of the suboptimal open loop control law advances the state of the art of stochastic control. Future efforts in stochastic optimal control are centered around simplifying or reducing the functional differential equation for optimal feedback control to a finite system of ordinary differential equations. Also, a simple analytical method of comparing suboptimal feedback control laws is desired in order to select the superior one given the model of the system.

NOTATION

$\triangleq$ indicates an equality by definition

$\delta a(t) \triangleq a(t+\Delta) - a(t)$

$O(\varepsilon)$ is understood to be a matrix function of ε and possibly of other variables

$N(a,A)$ or $N(\cdot,a,A)$ represents the normal density function of an m-vector random variable with an m-vector mean a and an $m \times m$-covariance matrix A.

$I_m \triangleq$ the $m \times m$ identity matrix

$da \triangleq [a, a+da]$ or its Euclidean m-space volume where a is an m-vector

A^T is the transposition of the matrix A.

$\nabla_a \triangleq \mathrm{Col}_{i=1}^{m} \left\{ \frac{\partial}{\partial a_i} \cdot \right\}$ where a is an m-vector

$\prod_{i=1}^{m} a_i b \triangleq a_m[a_{m-1}[\cdots a_2[a_1 b]\cdots]]$

$\delta(a) \triangleq \prod_{i=1}^{m} \delta(a_i)$ where a is an m-vector and δ is the Dirac delta function

$\Delta a_k \triangleq a_{k+1} - a_k$

$\Delta a_k^2 \quad \triangleq \quad (\Delta a_k)^2$

A power of a $\triangleq$ a term of the form $\prod_{i=1}^{m} a_i^{\alpha_i}$ for the m-vector a where α_i is a nonnegative integer

APPENDIX A.

Definition: Let $S = \{s_0, s_1, \cdots, s_m\}$ be a partition of $[t,T]$ and let α_k be a point in the subinterval $[s_k, s_{k+1}]$. Let $\alpha \triangleq \{\alpha_0, \alpha_1, \cdots, \alpha_{m-1}\}$, and let $A(s,\alpha)$ be some function of S and α. Then $B \triangleq \lim_{S\subset[t,T]} A(S,\alpha)$ if for every $\epsilon>0$, there exists a partition S_ϵ of $[t,T]$ such that for every partition $S \supset S_\epsilon$ and for every choice of α, it follows that $|A(S,\alpha) - B| < \epsilon$.

Definition: Let

$$A(s,\alpha) = \prod_{i=1}^{m} \underset{\delta z(\alpha_{m-i})}{\mathrm{Exp}} \left[\sum_{k=0}^{m-1} \int f(\alpha_k, x, \Gamma(\alpha_k, P(\alpha_k, x))) \cdot P(\alpha_k, x)\,dx\,\Delta s_k + \int g(x) P(\alpha_{m-1}, x)\,dx \right] . \tag{A-1}$$

Then

$$\underset{\substack{y(\tau)\\ t\le\tau\le T}}{\mathrm{Exp}} \left[\int_t^T \int f(s,x,\Gamma(s,P(s,x)))P(s,x)\,dx\,ds + \int g(x)P(T,x)\,dx \right]$$

$$\triangleq \lim_{S\subset[t,T]} A(S,\alpha) .$$

Notice that the joint conditional probability density function $q(t,\xi,x)$ of $\delta z(t)$ and $x(t)$ can be approximated from (1-2). By (1-2)

$$q(t,\xi|x)d\xi$$

$$= \text{Prob}\ [\delta z(t)\epsilon d\xi|x(t)=x] \sim N(\xi, h(t,x)\Delta, R^{-1}(t)\Delta)d\xi,$$

that is,

$$q(t,\xi|x) \approx N(\xi, h(t,x)\Delta, R^{-1}(t)\Delta)\ .$$

Thus $q(t,\xi) = \int q(t,\xi|x)P(t,x)dx$. Observe that

$$\underset{\delta z(t)}{\text{Exp}}\ [\delta z(t)] = \int \xi q(t,\xi)d\xi = \bar{h}(t)\Delta + O(\Delta^2)\ , \qquad \text{(A-2)}$$

that

$$\underset{\delta z(t)}{\text{Exp}}\ [\delta z(t)\delta z(t)^{\tau}] = \int \xi\xi^{\tau} q(t,\xi)d\xi = R^{-1}(t)\Delta + O(\Delta^2) \quad \text{(A-3)}$$

and that higher moments are $O(\Delta^2)$.

Theorem: Let $u(s)$ be a control function. Then

$$\underset{\substack{y(\tau)\\ t\le\tau\le T}}{\text{Exp}} \left[\int_t^T \int f(s,x,u(s))P(s,x)dxds + \int g(x)P(T,x)dx\right]$$

$$= \int_t^T \int f(s,x,u(s))Q(s,x)dxds + \int g(x)Q(T,x)dx$$

where $Q_s(s,x) = \mathcal{L}_+(s,u(s))Q(s,x)$ with $Q(t,x) = P(t,x)$.

Proof: Since $u(s)$ is a control function, it is independent of $\{y(\tau) \mid t \le \tau \le T\}$. Then with $\Gamma(s,P(s,x)) = u(s)$ (A-1) can be written

$$A(S,\alpha) = \sum_{k=0}^{m-1} \int f(\alpha_k,x,u(\alpha_k)) \prod_{i=1}^{m} \underset{\delta z(\alpha_{m-i})}{\mathrm{Exp}} P(\alpha_k,x)dx\Delta s_k$$

$$+ \int g(x) \prod_{i=1}^{m} \underset{\delta z(\alpha_{m-i})}{\mathrm{Exp}} P(\alpha_{m-1},x)dx .$$

Define $E_\alpha \cdot \triangleq \prod_{i=1}^{m} \underset{\delta z(\alpha_{m-i})}{\mathrm{Exp}} \cdot$, and let

$Q(\alpha_0,x;\alpha) = P(\alpha_0,x)$ where

$$Q_s(s,x;\alpha) = \mathcal{L}_+(s,u(s))Q(s,x;\alpha)$$

Then the following is proven.

Lemma: $E_\alpha P(\alpha_k,x) = Q(\alpha_k,x;\alpha) + \sum_{i=0}^{k-1} 0(\Delta\alpha_i^2)$

for $0 \le k \le m-1$.

Proof: The lemma is true by definition for $k=0$. Assume it is true for $0 \le k = \ell < m-1$.

Then by (1-5)

$$P(\alpha_{\ell+1}) = P(\alpha_\ell, x) + \mathcal{L}_+(\alpha_\ell)P(\alpha_\ell, x)\Delta\alpha_\ell$$

$$+ (\delta z(\alpha_\ell) - \bar{h}(\alpha_\ell)\Delta\alpha_\ell)^T r(\alpha_\ell, x)P(\alpha_\ell, x)$$

$$+ O(\Delta\alpha_\ell^{\,2}) .$$

By (A-2)

$$E_\alpha P(\alpha_{\ell+1}, x) = E_\alpha P(\alpha_\ell, x) + E_\alpha \mathcal{L}_+(\alpha_\ell)P(\alpha_\ell, x)\Delta\alpha_\ell$$

$$+ O(\Delta\alpha_\ell^{\,2}) . \qquad \text{(A-4)}$$

Interchanging the integration and differentiation operations [29] implies

$$E_\alpha \mathcal{L}_+(\alpha_\ell)P(\alpha_\ell, x) = \mathcal{L}_+(\alpha_\ell)E_\alpha P(\alpha_\ell, x) .$$

Thus, by induction, (A-4) implies

$$E_\alpha P(\alpha_{\ell+1}, x) = Q(\alpha_\ell, x; \alpha) + \mathcal{L}_+(\alpha_\ell)Q(\alpha_\ell, x; \alpha)\Delta\alpha_\ell + \sum_{i=0}^{\ell} O(\Delta\alpha_i^2)$$

$$= Q(\alpha_{\ell+1}, x; \alpha) + \sum_{i=1}^{\ell} O(\Delta\alpha_i^2) .$$

Thus ℓ implies $\ell+1$. ◻

Now let $h = \max_{0 \le i \le m-1} \{\Delta\alpha_i, \Delta s_i\}$. Then

$$A(S,\alpha) = \sum_{k=0}^{m-1}\left[\int f(\alpha_k, x, u(\alpha_k))Q(\alpha_k, x;\alpha)dx + \sum_{i=0}^{k-1} O(\Delta\alpha_i^2)\right]\Delta s_k$$

$$+\int g(x)Q(\alpha_{m-1}, x;\alpha)dx + \sum_{i=0}^{m-2} O(\Delta\alpha_i^2) .$$

But

$$\sum_{k=0}^{m-1}\sum_{i=0}^{k-1} O(\Delta\alpha_i^2)\Delta s_k = \sum_{k=0}^{m-1}\sum_{i=0}^{k-1} O(h^3) = O(h) ,$$

and

$$\sum_{i=0}^{m-2} O(\Delta\alpha_i^2) = \sum_{i=0}^{m-2} O(h^2) = O(h) .$$

Thus

$$A(S,\alpha) = \sum_{k=0}^{m-1}\int f(\alpha_k, x, u(\alpha_k))Q(\alpha_k, x;\alpha)dx\Delta s_k$$

$$+ \int g(x)Q(\alpha_{m-1}, x;\alpha)dx + O(h) .$$

Therefore,

$$\lim_{S\subset[t,T]} A(S,\alpha) = \int_t^T\int f(s,x,u(s))Q(s,x)dx + \int g(x)Q(T,x)dx. \blacksquare$$

APPENDIX B.

The results of Appendix A are extended to the stochastic system with perfect observations.

Equations (1-5) and (1-7) assume that $[K(t)K(t)']^{-1}$ exists. When there are perfect observations, it does not exist, but then a difference equation for $P(t,x)$ is not needed since $n_3 = n$, $h(t,x(t)) = x(t)$, $K(t) = 0$ and $P(t,x) = \delta(x-y(t))$.

Consequently, with u as a control function in (A-1),

$$\lim_{S\subset[t,T]} A(S,\alpha) = \underset{\substack{y(\tau)\\ t\le\tau\le T}}{\text{Exp}} \left[\int_t^T f(s,y(s),u(s))ds+g(y(T))\right] . \quad \text{(B-1)}$$

Since $y(t) = x(t)$, (B-1) is equivalent to

$$\int_t^T \int f(s,x,u(s))Q(s,x)dxds + \int g(x)Q(T,x)dx$$

with $Q(t,x) = P(t,x)$.

APPENDIX C.

The a priori stochastic operand $E(t,x)$ is positive semi-definite if $f(t,x,\theta(t))$ and $g(x)$ are positive semi-definite. Here is the proof.

Assume that the optimal control function θ is known and that t_0 and x_0 are given.

Let $Q(t,x)$ satisfy

$$Q_t(t,x) = \mathcal{L}_+(t,\theta(t))Q(t,x) \qquad Q(t_0,x) = \delta(x-x_0) .$$

Then by the theorem of 3.3.

$$E(t_0,x_0) = \int E(t_0,x)Q(t_0,x)dx$$

$$= \int_{t_0}^{T}\int f(s,x,\theta(s))Q(s,x)dxds$$

$$+ \int g(x)Q(T,x)dx \geq 0 . \quad \blacksquare$$

REFERENCES

[1] J. J. Florentin, "Optimal Control of Continuous Time, Markov, Stochastic Systems," Journal of Electronics and Control, Vol. 10, 1961, pp. 473-488.

[2] W. M. Wonham, "Stochastic Problems in Optimal Control," RIAS Technical Report 63-14, May, 1963.

[3] H. J. Kushner, "On the Dynamical Equations of Conditional Probability Density Functions, with Applications to Optimal Stochastic Control Theory," Journal of Mathematical Analysis and Applications, Vol. 8, 1964, pp. 332-344.

[4] A. Papoulis, Probability, Random Variables and Stochastic Processes, McGraw-Hill, Inc., New York, 1965.

[5] H. J. Kushner, "On the Differential Equations Satisfied by Conditional Probability Densities of Markov Processes, with Applications," Journal SIAM Control, Vol. 2, No. 1, 1964, pp. 106-119.

[6] L. S. Pontryagin et al, "The Mathematical Theory of Optimal Processes," Interscience Publishers, John Wiley and Sons, 1962.

[7] R. E. Kalman and R. S. Bucy, "New Results in Linear Filtering and Prediction Theory," Journal of Basic Engineering, Trans. A. S. M. E., Series D, Vol. 83, No. 1, March, 1961, pp. 95-108.

[8] R. Sridhar, "Fifth Quarterly Report to Jet Propulsion Laboratory," Contract No. 950670, April, 1965, pp. 16-18.

[9] J. L. Doob, Stochastic Processes, John Wiley and Sons, Inc., New York, 1953.

[10] K. Ito, "On Stochastic Differential Equations," Mem. Amer. Math. Soc., No. 4, 1951.

[11] R. Bellman, Adaptive Control Processes: A Guided Tour, Princeton, 1961.

[12] R. Bellman, H. Kagiwada and R. Kalaba, "Invariant Imbedding and the Numerical Integration of Boundary-Value Problems for Unstable Linear Systems of Ordinary Differential Equations," Comm. ACM, Vol. 10, No. 2, Feb., 1967, pp. 100-102.

[13] R. Bellman, H. Kagiwada, R. Kalaba and S. Ueno, "Invariant Imbedding and the Computational of Internal Fields for Transport Processes," Journal of Mathematical Analysis and Applications, Vol. 12, 1965, pp. 541-548.

[14] R. Bellman and R. Kalaba, "On the Principle of Invariant Imbedding and Propagation through Inhomogeneous Media," Proceedings of the National Academy of Sciences U. S., Vol. 42, 1956, pp.629-632.

[15] L. V. Kantorovich and G. P. Akilov, Functional Analysis in Normed Spaces, Translated by D. E. Brown, The MacMillan Co., New York, 1964.

[16] J. J. Levin, "On the Matrix Riccati-equation," Proc. Amer. Math. Soc., Vol. 10, 1959, pp. 519-524.

[17] C. W. Merriam III, Optimization Theory and the Design of Feedback Control Systems, McGraw-Hill, New York, 1964.

[18] L. Schwartz and E. B. Stear, "A Computational Comparison of Several Nonlinear Filters," IEEE Trans. Automatic Control (Short Papers), Vol. AC-13, February, 1968, pp.83-86.

[19] L. Schwartz and R. W. Bass, "Extensions to Optimal Multichannel Nonlinear Filtering," Hughes Aircraft Co., Culver City, Calif., Report SSD60220R, February 21, 1966

[20] J. R. Fisher, "Conditional Probability Density Functions and Optimal Nonlinear Estimation," Ph. D. Thesis, Department of Engineering, University of California, Los Angeles, Calif., 1966.

[21] A. H. Jazwinsky, "Stochastic Processes with Application to Filtering Theory," Analytical Mechanical Associates, Inc., Report 66-6, Final Rept. Contract NAS 5-9298, November, 1966.

[22] R. W. Bass, V. D. Norum and L. Schwartz, "Optimal Multichannel Nonlinear Filtering," J. Math. Anal. Appl., Vol. 16, October, 1966, pp. 162-164.

[23] A. H. Jazwinsky, "Filtering for Nonlinear Dynamical Systems," IEEE Trans. Automatic Control (Correspondence), Vol. AC-11, October, 1966, pp. 765-766.

[24] D. M. Detchmendy and R. Sridhar, "Sequential Estimation of States and Parameters in Noisy Nonlinear Dynamical Systems," Preprints, 6th Joint Automatic Control Conf., Troy, New York, June, 1965, pp. 56-63.

[25] H. Cox, "Estimation of State Variables Via Dynamic Programming," Preprints, 5th Joint Automatic Control Conf., Stanford, Calif., June, 1964, pp. 376-381.

[26] A. H. Jazwinsky, "Nonlinear Filtering with Discrete Observations," AIAA Paper 66-38, presented at the AIAA 3rd Aerospace Sciences Meeting, New York, January, 1966.

[27] I. A. Gura and L. J. Henrikson, "A Unified Approach to Nonlinear Estimation," Aerospace Corporation Report No. TR-0200(4306-02)-02, January 15, 1969.

[28] R. Bellman, Dynamic Programming, Princeton University Press, 1957.

[29] T. M. Apostol, Mathematical Analysis, Addison Wesley, 1964.

[30] J. N. Franklin, Matrix Theory, Prentice-Hall, Inc., 1968.

[31] S. E. Dreyfus, "Some Types of Optimal Control of Stochastic Systems," J. Soc. Ind. Appl. Math. Ser. A. Control 2, No. 1, 1964, pp. 120-134.

[32] M. Aoki, Optimization of Stochastic Systems, Academic Press, New York, 1967.

[33] R. E. Mortensen, "Optimal Control of Continuous-time Stochastic Systems," Ph. D. Thesis, University of California, January, 1966.

[34] R. E. Mortensen, "A Priori Open Loop Control of Continuous Time Stochastic Systems," Int. J. Control, Vol. 3, No. 2, 1966, pp. 113-127.

[35] R. E. Mortensen, "Stochastic Optimal Control with Noisy Observations," Int. J. Control, No. 5, 1966, pp. 455-464.

[36] K. A. Lur'e, "On the Hamilton-Jacobi Method in Variational Problems of Partial Differential Equations," PMM, Vol. 27, No. 2, 1963, pp. 255-264.

[37] P. K. C. Wang, "Control of Distributed Parameter Systems," Advances in Control Systems, Vol. 1 (edited by C. T. Leondes), Academic Press, New York, 1964, pp. 75-172.

[38] R. S. Bucy, "Nonlinear Filtering," IEEE Trans. Aut. Control (1965), p. 198.

[39] R. S. Bucy and P. D. Joseph, Filtering For Stochastic Processes with Applications to Guidance, Interscience Publishers, New York, 1968.