Optimal Filtering and Filter Stability of Linear Stochastic Delay Systems

RAYMOND HON-SING KWONG, MEMBER, IEEE, AND ALAN S. WILLSKY, MEMBER, IEEE

Abstract—**Optimal filtering equations are obtained for very general linear stochastic delay systems. Stability of the optimal filter is studied in the case where there are no delays in the observations. Using the duality between linear filtering and control, asymptotic stability of the optimal filter is proved. Finally, the cascade of the optimal filter and the deterministic optimal quadratic control system is shown to be asymptotically stable as well.**

I. Introduction

In recent years, considerable attention has been paid to the optimal control of linear delay systems with quadratic cost. Optimal control over a finite as well as infinite time interval has been studied, and most of the analogs of the finite dimensional results have been established [1]–[6]. Linear delay systems constitute then an important class of systems where a virtually complete extension of the theory of finite dimensional linear quadratic control is available. It is natural, therefore, to consider the extension of the linear stochastic filtering and control theory for finite dimensional systems to delay system as well. Linear filtering for delay systems was first considered by Kwakernaak [7], who gave a formal derivation of the filter equations using the approach of Kalman and Bucy [8]. The linear filtering and stochastic control problem was then studied by Lindquist, who proved a duality theorem between estimation and control and a separation theorem of stochastic control [9]–[11]. However, he did not characterize the covariance of the optimal filter and his results did not clarify the structure of the optimal filter. In particular, it is not quite appropriate for the stability analysis of the optimal filter. Mitter and Vinter [12] studied the filtering problem from the viewpoint of stochastic evolution equations and discussed the filter stability problem. But they had to restrict their considerations to time-invariant systems and exclude point delays in the observations. In this paper, both the linear optimal filtering as well as stochastic control problems will be discussed. We will completely characterize the structure of the optimal filter for Gaussian noises, and under suitable conditions, obtain asymptotic stability of the optimal filter. As a by-product of this study, the precise relationships between optimal linear control with quadratic cost and optimal linear filtering are established. By putting together the deterministic optimal control law and the optimal filter, the asymptotic stability of the stochastic control system is also established. The main emphasis of the paper will be on the interpretation of the results in comparison to the finite dimensional case. Details of the proofs of the results are too lengthy to be incorporated here and will be published elsewhere [13].

Manuscript received March 30, 1976; revised November 7, 1976. Paper recommended by Y. Bar-Shalom, Chairman of the IEEE S-CS Stochastic Control Committee. This work was performed at the M.I.T. Electronic Systems Laboratory under NSF Grant GK41647 and NASA Ames Grant NGL-22-009-124, and at McGill University under National Research Council of Canada Grants A9067 and A3921, and at the Centre de Recherches Mathematiques, Universite de Montreal under Subvention FCAC du Ministere de l'Education du Quebec.

R. H. Kwong is with the Department of Electrical Engineering, McGill University and the Centre de Recherches Mathematiques, Universite de Montreal, Montreal, P. Q., Canada.

A. S. Willsky is with the Electronic Systems Laboratory and the Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA 02139.

II. Linear Stochastic Delay Systems

We shall study linear stochastic delay systems of the form

$$dx(t)=\int_{-\tau}^{0} d_\theta A(t,\theta)x(t+\theta)\,dt+F(t)\,dw(t) \quad (2.1)$$

$$x(\theta)=x_0(\theta), \qquad -\tau\leqslant\theta\leqslant 0$$

$$dz(t)=\int_{-\tau}^{0} d_\theta C(t,\theta)x(t+\theta)\,dt+N(t)\,dv(t) \quad (2.2)$$

$$z(t)=0, \quad t\leqslant 0.$$

The system process $x(t)$ takes values in R^n, the observation process $z(t)$ in R^p. The processes $w(t)$ and $v(t)$ are standard Wiener processes in R^m and R^p, respectively, completely independent of each other. The initial function x_0 is taken to be some Gaussian process on $[-\tau,0]$, completely independent of $w(t)$ and $v(t)$, with mean $\bar{x}_0(\theta)$ and $\operatorname{cov}[x_0(\theta);\ x_0(\xi)]=\Sigma_0(\theta,\xi)$. The maps $A(t,\theta)$ and $C(t,\theta)$ satisfy the same conditions as those stated in [14] for deterministic linear delay systems. The maps $F(t)$ and $N(t)$ are $n\times m$ and $p\times p$ matrix-valued continuous functions, respectively. Furthermore, $N(t)$ is assumed to be nonsingular. We shall often write $F(t)F'(t)=Q(t)$ and $N(t)N'(t)=R(t)$.

As in deterministic delay systems, the state of the system (2.1) is not the process $x(t)$ but rather a piece of the trajectory x_t [14]. The x_t process is derived from the $x(t)$ process and is defined on $[-\tau,0]$ by

$$x_t(\theta)=x(t+\theta), \qquad \theta\in[-\tau,0].$$

As such, x_t is a process taking values in a function space. Throughout this and the next section, we will choose the function space to be C, the space of continuous functions on $[-\tau,0]$. Accordingly, we will also assume the initial function x_0 to take values in C.

III. The Linear Filtering Problem

The filtering problem consists of characterizing the conditional mean of $x(t)$ given the observations up to time t. We shall use the notation $\hat{x}(t+\theta|t)$, $-\tau \leqslant \theta \leqslant 0$, to mean $E\{x(t+\theta)|z^t\}$, where z^t is the σ-algebra generated by $z(s)$, $0 \leqslant s \leqslant t$. The optimal filter equations for the system (2.1), (2.2) have been derived rigorously via probabilistic arguments in [13] and are stated below.

Theorem 3.1. The optimal filter for the systems (2.1), (2.2) defined on the interval $[0,T]$ is characterized as follows.

1) The conditional mean $\hat{x}(t|t)$ is generated by

$$d\hat{x}(t|t) = \int_{-\tau}^{0} d_\theta A(t,\theta)\hat{x}(t+\theta|t)\,dt + \int_{-\tau}^{0} P(t,0,\theta)\,d_\theta C'(t,\theta)R^{-1}(t)\,d\nu(t) \quad (3.1)$$

where $\nu(t)$ is the innovations given by

$$\nu(t) = z(t) - \int_0^t \int_{-\tau}^{0} d_\theta C(s,\theta)\hat{x}(s+\theta|s)\,ds.$$

2) The smoothed estimate $\hat{x}(t+\theta|t)$ is generated by

$$\hat{x}(t+\theta|t) = \hat{x}(t+\theta|t+\theta) + \int_{t+\theta}^{t}\int_{-\tau}^{0} P(s,t+\theta-s,\xi)\,d_\xi C'(s,\xi) \cdot R^{-1}(s)\,d\nu(s), \qquad -\tau \leqslant \theta \leqslant 0. \quad (3.2)$$

3) The function $P(t,\theta,\xi)$ has the interpretation of a conditional covariance, i.e., for $-\tau \leqslant \theta$, $\xi \leqslant 0$,

$$P(t,\theta,\xi) = E\{[x(t+\theta) - \hat{x}(t+\theta|t)] \cdot [x(t+\xi) - \hat{x}(t+\xi|t)]'|z^t\}.$$

It is a deterministic quantity and is characterized by the following coupled set of equations:

$$\frac{d}{dt}P(t,0,0) = \int_{-\tau}^{0} P(t,0,\theta)\,d_\theta A'(t,\theta) + \int_{-\tau}^{0} d_\theta A(t,\theta)P(t,\theta,0) - \int_{-\tau}^{0}\int_{-\tau}^{0} P(t,0,\theta)\,d_\theta C'(t,\theta)R^{-1}(t)\,d_\xi \cdot C(t,\xi)P(t,\xi,0) + Q(t) \quad (3.3)$$

$$\sqrt{2}P_\eta(t,\theta,0) = \int_{-\tau}^{0} P(t,\theta,\xi)\,d_\xi A'(t,\xi) - \int_{-\tau}^{0}\int_{-\tau}^{0} P(t,\theta,\xi)\,d_\xi C'(t,\xi) \cdot R^{-1}(t)\,d_\alpha C(t,\alpha)P(t,\alpha,0) \quad (3.4)$$

$$\sqrt{3}P_\sigma(t,\theta,\xi) = -\int_{-\tau}^{0}\int_{-\tau}^{0} P(t,\theta,\beta)\,d_\beta C'(t,\beta) \cdot R^{-1}(t)\,d_\alpha C(t,\alpha)P(t,\alpha,\xi) \quad (3.5)$$

where η is the unit vector in the $(1,-1,0)$ direction, σ the unit vector in the $(1,-1,-1)$ direction, and $P_\eta(t,\theta,0)$ and $P_\sigma(t,\theta,\xi)$ are the directional derivatives of $P(t,\theta,0)$ and $P(t,\theta,\xi)$ in the directions η and σ, respectively. The initial conditions are given by

$$\hat{x}(\theta|0) = \bar{x}_0(\theta), \qquad \theta \in [-\tau,0]$$

$$P(0,\theta,\xi) = \Sigma_0(\theta,\xi), \qquad -\tau \leqslant \theta, \xi \leqslant 0.$$

Remark 3.1: Similar equations for the optimal filter for systems with point delays only have been formally derived by Kwakernaak [7], but his proofs were not satisfactory. For example he used partial derivatives formally for the smoothed estimates and did not demonstrate that it is sufficient to specify the function $P(t,\theta,\xi)$ in the interval $0 \leqslant t \leqslant T$, $-\tau \leqslant \theta$, $\xi \leqslant 0$. This is related to the structure of delay systems, as the next remark will explain.

Remark 3.2: Equation (3.1) shows that the evolution of the conditional mean depends on the smoothed estimate $\hat{x}(t+\theta|t)$. If we interpret this formally as the process $E[x_t|t]$, we see that we need the "conditional mean" of the true state x_t to determine $\hat{x}(t|t)$. Similarly, we can think of $P(t,\theta,\xi)$ as the "covariance" of the true state x_t (formally, this corresponds to saying

$$E\{[x_t - E(x_t|t)](\theta)[x_t - E(x_t|t)]'(\xi)|_z t\} = P(t,\theta,\xi).)$$

Hence, while we would be primarily interested in the process $\hat{x}(t|t)$, we need also the smoothed estimate $\hat{x}(s|t)$ over the interval $t-\tau \leqslant s < t$. Similarly, while the estimation error covariance is provided by the function $P(t,0,0)$, we must characterize the entire function $P(t,\theta,\xi)$. This constitutes the complete information about the optimal filter.

Remark 3.3: For general linear stochastic delay systems of the form (2.1) and (2.2), the function $P(t,\theta,\xi)$ will not be continuously differentiable in (t,θ,ξ). A similar situation has been discussed in the optimal control of delay systems with quadratic cost [6]. This is why directional derivatives have to be used. Of course, if $P(t,\theta,\xi)$ were continuously differentiable,

$$\sqrt{3}P_\sigma(t,\theta,\xi) = \left(\frac{\partial}{\partial t} - \frac{\partial}{\partial\theta} - \frac{\partial}{\partial\xi}\right)P(t,\theta,\xi).$$

We have given the optimal filter equations for very general linear stochastic delay systems. For our stability analysis, however, we shall restrict our attention to a class of linear delay systems, those with delays in the dynamics only.

IV. Filter Stability for Systems with Delays in the Dynamics

In this section, we will discuss the stability of the optimal filter for the delay system

$$dx(t) = Ax(t)\,dt + Bx(t-\tau)\,d\tau + F\,dw(t) \quad (4.1)$$

$$x(\theta)=0, \qquad -\tau\leqslant\theta\leqslant 0$$
$$dz(t)=Cx(t)\,dt+N\,dv(t). \tag{4.2}$$

There are several reasons for restricting our attention to this, the simplest of linear stochastic delay systems. First, since we are interested in asymptotic stability, the assumption of zero initial condition entails no essential loss of generality since initial effects will "die" away. Second, the techniques used in establishing asymptotic stability can be readily generalized to situations with multiple and or distributed delays in the dynamics. Third, while delays in the observations can also be treated, such delays will introduce additional complications which will lengthen the present paper considerably. Fourth, for purposes of comparison with the finite dimensional case, the results for this simple system show most transparently where the similarities and differences occur. Finally, for this case, the matrix $P(t,\theta,\xi)$ can be shown [13] to be continuously differentiable in (t,θ,ξ). Hence, directional derivatives can be replaced by partial derivatives and technical complications can be kept to a minimum. On the other hand, most of the basic ideas and techniques will already be used in the discussion of this simple system. It is therefore deemed best at this stage to restrict our attention to the simplest case, and relegate the various generalizations and refinements, particularly the delays in the observations problem, to a future paper (see also [16]).

Applying (3.1) and (3.2) to the system (4.1), (4.2), yields

$$d\hat{x}(t|t)=\left[A\hat{x}(t|t)+B\hat{x}(t-\tau|t)\right]dt+P_0(t)C'R^{-1}\left[dz(t)-C\hat{x}(t|t)\,dt\right]$$
$$\hat{x}(t-\tau|t)=\hat{x}(t-\tau|t-\tau)+\int_{t-\tau}^{t}P_1(s,t-\tau-s)C'R^{-1}\left[dz(s)-C\hat{x}(s|s)\,ds\right].$$

Here we have, for convenience in later comparison with optimal control results, defined $P_0(t)=P(t,0,0)$, $P_1(t,\theta)=P(t,\theta,0)$ and $P_2(t,\theta,\xi)=P(t,\theta,\xi)$. Combining these two equations we conclude that the optimal estimate $\hat{x}(t|t)$ is generated by a stochastic delay differential equation driven by the observations

$$\begin{aligned}d\hat{x}(t|t)=&\left[A\hat{x}(t|t)+B\hat{x}(t-\tau|t-\tau)\right]dt\\&+P_0(t)C'R^{-1}\left[dz(t)-C\hat{x}(t|t)\,dt\right]\\&+\int_{t-\tau}^{t}BP_1(s,t-\tau-s)C'R^{-1}\left[dz(s)-C\hat{x}(s|s)\,ds\right]dt\end{aligned} \tag{4.3}$$
$$\hat{x}(\theta|0)=0, \qquad -\tau\leqslant\theta\leqslant 0.$$

Furthermore, from (3.3)–(3.5), we obtain the following equations for the matrices $P_0(t)$, $P_1(t,\theta)$, and $P_2(t,\theta,\xi)$:

$$\frac{d}{dt}P_0(t)=AP_0(t)+P(t)A'-P_0(t)C'R^{-1}CP_0(t)+Q+BP_1(t,-\tau)+P_1'(t,-\tau)B' \tag{4.4}$$

$$\left(\frac{\partial}{\partial t}-\frac{\partial}{\partial\theta}\right)P_1(t,\theta)=P_1(t,\theta)\left[A'-C'R^{-1}CP_0(t)\right]+P_2(t,\theta,-\tau)B' \tag{4.5}$$

$$\left(\frac{\partial}{\partial t}-\frac{\partial}{\partial\theta}-\frac{\partial}{\partial\xi}\right)P_2(t,\theta,\xi)=-P_1(t,\theta)C'R^{-1}CP_1'(t,\xi) \tag{4.6}$$

with

$$\begin{gathered}P_0(0)=P_1(0,\theta)=P_2(0,\theta,\xi)=0\\P_1(t,0)=P_0(t)\\P_2(t,\theta,0)=P_1(t,\theta)\\P_0(t)=P_0'(t),\ P_2(t,\theta,\xi)=P_2'(t,\xi,\theta), \qquad -\tau\leqslant\theta,\xi\leqslant 0.\end{gathered} \tag{4.7}$$

In establishing stability of the optimal filter, we shall make essential use of the duality between optimal filtering of linear stochastic delay systems and optimal control of linear delay systems with quadratic cost. We begin therefore by summarizing the optimal control results [1]–[6]. Consider the system

$$\begin{gathered}\dot{x}(t)=Ax(t)+Bx(t-\tau)+Cu(t)\\x(\theta)=x_0(\theta).\end{gathered} \tag{4.8}$$

Since the following results hold for x_0 in C or in the product space $R^n\times L_2$ [4], we shall not worry about the specific choice of function space. The cost functional is given by

$$J_T(u,x_0)=\int_0^T\left[x'(t)Mx(t)+u'(t)Su(t)\right]dt$$

where M and S are symmetric matrices of appropriate dimensions, $M\geqslant 0$, $S>0$. When $T<\infty$, the optimal control is given by

$$u^*(t)=-S^{-1}C'K_0(t)x(t)-S^{-1}C'\int_{-\tau}^{0}K_1(t,\theta)x(t+\theta)\,d\theta. \tag{4.9}$$

The feedback gains satisfy the following coupled set of partial differential equations:

$$\frac{d}{dt}K_0(t)=-A'K_0(t)-K_0(t)A+K_0(t)CS^{-1}C'K_0(t)-M-K_1(t,0)-K_1'(t,0) \tag{4.10}$$

$$\left(\frac{\partial}{\partial t}-\frac{\partial}{\partial\theta}\right)K_1(t,\theta)=-\left[A'-K_0(t)CS^{-1}C'\right]K_1(t,\theta)-K_2(t,0,\theta) \tag{4.11}$$

$$\left(\frac{\partial}{\partial t}-\frac{\partial}{\partial\theta}-\frac{\partial}{\partial\xi}\right)K_2(t,\theta,\xi)=K_1'(t,\theta)CS^{-1}C'K_1(t,\xi) \tag{4.12}$$

with

$$K_0(T)=K_1(T,\theta)=K_2(T,\theta,\xi)=0, \qquad -\tau\leqslant\theta,\xi\leqslant 0$$

$$K_1(t,-\tau)=K_0(t)B$$
$$K_2(t,-\tau,\theta)=B'K_1(t,\theta) \tag{4.13}$$
$$K_0(t)=K_0'(t),\ K_2(t,\theta,\xi)=K_2'(t,\xi,\theta).$$

To discuss the infinite time control problem, we need the concepts of stabilizability and detectability. As in the finite dimensional case, stabilizability corresponds to the existence of a linear feedback operator on the complete state x_t which stabilizes the system, and detectability corresponds to stabilizability of the adjoint system. For our purposes, we do not need the most general setup for these notions. The following definitions will suffice.

Definition 4.1: The linear delay system

$$\dot{x}(t)=Ax(t)+Bx(t-\tau)+Cu(t)$$

is said to be stabilizable if there exist a constant matrix K_0 and a continuous matrix-valued function $K_1(\theta)$, $-\tau\leqslant\theta\leqslant 0$, such that the system

$$\dot{x}(t)=Ax(t)+Bx(t-\tau)+CK_0x(t)+C\int_{-\tau}^{0}K_1(\theta)x(t+\theta)\,d\theta$$

is asymptotically stable. We then also say that (A,B,C) is stabilizable.

Definition 4.2: The delay system

$$\dot{x}(t)=Ax(t)+Bx(t-\tau)$$
$$z(t)=Hx(t)$$

is said to be detectable if there exist a constant matrix L_0 and a continuous matrix-valued function $L_1(\theta)$, $-\tau\leqslant\theta\leqslant 0$, such that

$$\dot{x}(t)=Ax(t)+Bx(t-\tau)+L_0z(t)+\int_{-\tau}^{0}L_1(\theta)z(t+\theta)\,d\theta$$

is asymptotically stable. We then also say that (A,B,H) is detectable.

The properties of stabilizability and detectability for delay systems are discussed further in [6] and [13], to which the reader is referred.

We can now state the result concerning the infinite time quadratic control problem obtained essentially in [5], [6], and [15]. Let $M=H'H$.

Proposition 4.1: Assume that (A,B,C) is stabilizable and (A,B,H) is detectable. Then the gains $K_0(t)$, $K_1(t,\theta)$ and $K_2(t,\theta,\xi)$, for each fixed $t<T$, converge to K_0, $K_1(\theta)$, and $K_2(\theta,\xi)$, respectively, as $T\to\infty$. The optimal control law for the infinite time problem is given by

$$u^*(t)=-S^{-1}C'K_0x(t)-\int_{-\tau}^{0}S^{-1}C'K_1(\theta)x(t+\theta)\,d\theta. \tag{4.14}$$

Furthermore, the optimal closed-loop system is asymptotically stable. The matrices K_0, $K_1(\theta)$, and $K_2(\theta,\xi)$ satisfy the obvious equations.

Remark 4.1: Proposition 4.1 was proved earlier in [5] and [6] under the assumption that $M>0$. This was relaxed in [13] to requiring (A,B,H) to be observable [13]. Subsequently, the work of Zabczyk [15] became known to the author. The fact that the weaker condition of detectability is sufficient can be deduced from the results of [15].

Remark 4.2: If we substitute the control law (4.9) into (4.8), we see that the resulting delay system is very similar to the optimal filter (4.3) with the observations set to zero. Furthermore, the covariance equations (4.4)–(4.7) are also very similar to the control gain equations (4.10)–(4.13). This strongly suggests that duality relations between optimal quadratic control and optimal linear filtering exist, and that proving results for one case allows us to prove results for the other case also. We shall see that this is indeed true.

Proposition 4.2 ([9],[13]): Consider the optimal filtering problem over the interval $[0,T]$ for the system defined by (4.1), (4.2). Define the dual control system by

$$\dot{y}(t)=-A'y(t)-B'y(t+\tau)-C'u(t) \tag{4.15}$$

with

$$y(T)=b,\quad y(s)=0,\qquad s>T. \tag{4.16}$$

The dual control problem is defined to be to minimize

$$J_T(b,u)=\int_0^T\left[y'(t)Qy(t)+u'(t)Ru(t)\right]dt \tag{4.17}$$

where $Q=FF'\geqslant 0$ and $R=NN'>0$. As before, let the optimal linear least squares estimate for $x(T)$ be $\hat{x}(T|T)$, and let the optimal control for the dual problem be u_T. Then $b'\hat{x}(T|T)$ is related to u_T by

$$b'\hat{x}(T|T)=-\int_0^T u_T'(s)\,dz(s). \tag{4.18}$$

Remark 4.3: Proposition 4.2 as stated was first proved in [9]. One of us gave a somewhat different proof in [13] using a certain bilinear form in the theory of delay differential systems. This second proof appears to be simpler and perhaps of independent interest.

We now have two representations of $b'\hat{x}(T|T)$, one directly from (4.3)–(4.7), the other indirectly from (4.17). Our strategy is to compare the two representations and identity the control and filter gains appropriately. This will enable us to exploit the known results of the optimal control problem to conclude filter stability. The reason for adopting this somewhat indirect method of proof is that it is usually much easier to prove results for optimal control owing to its variational interpretation. In this case, this basically enables us to avoid analysing the coupled equations (4.4)–(4.7) directly.

To this end, we introduce the matrices $\tilde{K}_0(t)$, $\tilde{K}_1(t,\theta)$, and $\tilde{K}_2(t,\theta,\xi)$. These are the optimal gains corresponding to the dual control problems (4.15)–(4.17). They satisfy equations very similar to those of (4.10)–(4.13), with suitable changes in the variables and with A,B,C replaced by A',B',C', respectively. The optimal closed-loop system is given by

$$\dot{y}(t) = -\left[A' - C'R^{-1}C\tilde{K}_0(t)\right]y(t) - B'y(t+\tau) + \int_0^{\tau} C'R^{-1}C\tilde{K}_1(t,\theta)y(t+\theta)\,d\theta. \quad (4.19)$$

But then the equations for $\tilde{K}_0(t)$, $\tilde{K}_1(t,\theta)$, and $\tilde{K}_2(t,\theta,\xi)$ look exactly like those for $P_0(t)$, $P_1(t,\theta)$, and $P_2(t,\theta,\xi)$. In fact, the following result establishes their precise relationships [13].

Lemma 4.1: Let the fundamental matrix associated with the system

$$\dot{x}(t) = \left[A - P_0(t)C'R^{-1}C\right]x(t) + Bx(t-\tau) - B\int_{-\tau}^{0} P_1(t+\theta, -\theta-\tau)C'R^{-1}Cx(t+\theta)\,d\theta \quad (4.20)$$

[i.e., the homogeneous part of (4.3)] be $\Phi(t,s)$ and let the fundamental matrix associated with (4.19) be $Y(t,s)$. Then we have, for $0 \leqslant t \leqslant T$, $0 \leqslant \theta, \xi \leqslant \tau$,

$$P_0(t) = \tilde{K}_0(t) \quad (4.21)$$

$$P_1'(t,\theta-\tau)B' = \tilde{K}_1(t,\theta) \quad (4.22)$$

$$BP_2(t,\theta-\tau,\xi-\tau)B' = \tilde{K}_2(t,\theta,\xi). \quad (4.23)$$

Moreover, (4.19) and (4.20) are adjoints of each other [15] so that

$$\Phi(t,s) = Y'(s,t).$$

We have now everything for proving asymptotic stability of the optimal filter. We state it as Theorem 4.1.

Theorem 4.1: Consider the system defined by (4.1) and (4.2). Suppose (A,B,C) is detectable and (A,B,F) is stabilizable. Then the gains of the optimal filter defined by (4.3)–(4.7) converge, and the steady-state optimal filter is asymptotically stable.

Remark 4.4: Theorem 4.1 is not the most general form of the filter stability results. For example, if there are delays in the observations, the dual control problem corresponds to systems with delays both in the state as well as control. To study the infinite time quadratic control problem for such systems, we have to extend our notion of stabilizability and detectability. Such and other generalizations will be taken up in a future paper.

V. Stochastic Control of Linear Delay Systems

We can now combine the results for optimal control with quadratic cost and optimal linear filtering to obtain a stable stochastic control system. The stochastic control system is given by

$$dx(t) = \left[Ax(t) + Bx(t-\tau) + Gu(t)\right]dt + F\,dw(t) \quad (5.1)$$

$$x(\theta) = x_0(\theta)$$

$$dz(t) = Cx(t)\,dt + N\,dv(t) \quad (5.2)$$

and the criterion is given by

$$J_T = E\int_0^T \left[x'(t)Mx(t) + u'(t)Su(t)\right]dt. \quad (5.3)$$

The objective is to choose a control law u in a suitable set of admissible laws such that the cost functional (5.3) is minimized. Following Lindquist [10], [11] we let $z_0(t)$ be the process generated by (5.2) with $u(t) \equiv 0$. Define the set U_0 of admissible control laws to be the class of measurable processes satisfying the following conditions.

1) For each t, $u(t)$ is measurable with respect to $\sigma\{z(s), 0 \leqslant s \leqslant t\}$, the σ-algebra generated by $z(s)$, $0 \leqslant s \leqslant t$.

2) For each $u \in U_0$, there exist unique solutions to (5.1) and (5.2).

3) $\int_0^T E|u(t)|^2\,dt < \infty$.

4) For each $u \in U_0$, $\sigma\{z(s), 0 \leqslant s \leqslant t\} = \sigma\{z_0(s), 0 \leqslant s \leqslant t\}$.

We shall not dwell on the reasons for this choice of U_0 as the set of admissible laws, but refer the reader to [10], [11], and [13] for additional comments. The following version of the separation theorem has been proved by Lindquist [10].

Proposition 5.1: The optimal control for the problem defined immediately above is given by

$$u^*(t) = -S^{-1}G'K_0(t)\hat{x}(t|t) - S^{-1}G'\int_{-\tau}^{0} K_1(t,\theta)\hat{x}(t+\theta|t)\,d\theta \quad (5.4)$$

where, as before, $K_0(t)$ and $K_1(t,\theta)$ are the optimal gains for the deterministic optimal control problem and $\hat{x}(s|t)$, $t-\tau \leqslant s \leqslant t$, is the conditional expectation of $x(s)$ given $z(\sigma)$, $0 \leqslant \sigma \leqslant t$.

The expression for the optimal cost has been obtained in [13] and is given by Lemma 5.1.

Lemma 5.1: Corresponding to the optimal law (5.4), the optimal cost is given by

$$\begin{aligned} J^* = EV(x_0) &+ \int_0^T \operatorname{tr} FF'K_0(t)\,dt \\ &+ \int_0^T \operatorname{tr}\Bigg\{ K_0(t)GS^{-1}G'K_0(t)P_0(t) \\ &+ \int_{-\tau}^{0} K_1'(t,\theta)GS^{-1}G'K_0(t)P_1'(t,\theta)\,d\theta \\ &+ \int_{-\tau}^{0} K_0(t)GS^{-1}G'K_1(t,\theta)P_1(t,\theta)\,d\theta \\ &+ \int_{-\tau}^{0}\int_{-\tau}^{0} K_1'(t,\theta)GS^{-1}G'K_1(t,\xi) \\ &\cdot P_2(t,\xi,\theta)\,d\theta\,d\xi \Bigg\}\,dt \end{aligned} \quad (5.5)$$

where

$$V(x_t)=x'(t)K_0(t)x(t)+\int_{-\tau}^{0}x'(t)K_1(t,\theta)x(t+\theta)\,d\theta+\int_{-\tau}^{0}x'(t+\theta)K_1'(t,\theta)x(t)\,d\theta+\int_{-\tau}^{0}\int_{-\tau}^{0}x'(t+\theta)K_2(t,\theta,\xi)x(t+\xi)\,d\theta\,d\xi. \quad (5.6)$$

Remark 5.1: Notice that in (5.4), the conditional mean of the true state x_t is used. This vividly demonstrates the need to use the true state in the design of feedback control. Note also the striking similarity in the structure of the optimal cost in this case as compared to the finite dimensional case: there is a term due to initial conditions, a term due to the noise in the system dynamics, and terms due to the estimation error.

We turn our attention now to the properties of closed-loop system under the stationary version of the law (5.4). This is now just a matter of putting together the results we have developed.

Theorem 5.1: Let $M=H'H$. Suppose (A,B,G) and (A,B,F) are stabilizable, and (A,B,C) and (A,B,H) are detectable. Then the control law

$$u(t)=-S^{-1}G'K_0\hat{x}(t|t)-S^{-1}G'\int_{-\tau}^{0}K_1(\theta)\hat{x}(t+\theta|t)\,d\theta \quad (5.7)$$

where $\hat{x}(t+\theta|t)$, $-\tau\leqslant\theta\leqslant 0$, is generated by the steady-state filter and K_0, $K_1(\theta)$ are given by the deterministic optimal stationary control law, gives rise to an asymptotically stable closed-loop system.

VI. Conclusions

We have treated the problem of filtering and control for linear stochastic delay systems. The case with delays in the dynamics only is studied in detail, with particular emphasis on the stability of the optimal systems. These results show very clearly where the analogies with the finite dimensional case lie and what are the additional complications due to the presence of delays. The structure of delay systems in thus brought into focus. The study also indicates that a complete linear-quadratic-Gaussian theory for delay systems can be developed. It is hoped that this paper will serve as a stimulus to the development of a successful design method for this important class of systems.

References

[1] H. J. Kushner and D. I. Barnea, "On the control of a linear functional differential equation with quadratic cost," *SIAM J. Contr.*, vol. 8, pp. 257–272, 1970.

[2] Y. Alekal, P. Brunovsky, D. H. Chyung, and E. B. Lee, "The quadratic problem for systems with time delays," *IEEE Trans. Automat. Contr.*, vol. AC-16, pp. 673–688, Dec. 1971.

[3] A. Manitius, "Optimal control of time-lag systems with quadratic performance indexes," presented at the 4th Int. Fed. Automatic Control Cong., Warsaw, 1969, paper 13.2.

[4] M. C. Delfour and S. K. Mitter, "Controllability, observability, and optimal feedback control of affine hereditary differential systems," *SIAM J. Contr.*, vol. 10, pp. 298–328, May 1972.

[5] R. Datko, "Unconstrained control problems with quadratic cost," *SIAM J. Contr.*, vol. 11, pp. 32–52, 1973.

[6] M. C. Delfour, C. McCalla, and S. K. Mitter, "Stability and the infinite-time quadratic cost problem for linear hereditary differential systems," *SIAM J. Contr.*, vol. 13, pp. 48–88, Jan. 1975.

[7] H. Kwakernaak, "Optimal filtering in linear systems with time delays," *IEEE Trans. Automat. Contr.*, vol. AC-12, pp. 169–173, Apr. 1967.

[8] R. E. Kalman and R. S. Bucy, "New results in linear filtering and prediction theory," *J. Basic Eng.*, vol. 83, pp. 95–108, Mar. 1961.

[9] A. Lindquist, "A theorem on duality between estimation and control for linear stochastic systems with time delay," *J. Math. Anal. Appl.*, vol. 37, pp. 516–536, 1972.

[10] ——, "Optimal control of linear stochastic systems with applications to time lag systems," *Inform. Sci.*, vol. 5, pp. 81–126, 1973.

[11] ——, "On feedback control of linear stochastic systems," *SIAM J. Contr.*, vol. 11, pp. 323–343, May 1973.

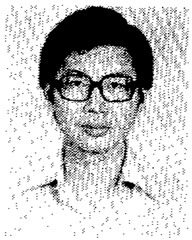
[12] S. K. Mitter and R. B. Vinter, "Filtering for linear stochastic hereditary differential systems," presented at the Int. Symp. Control Theory, Numerical Methods and Computer Systems Modelling, I.R.I.A., Rocquencourt, France, June 1974.

[13] R. H. Kwong, "Structural properties and estimation of delay systems," Electron. Syst. Lab. Massachusetts Inst. Technol., Cambridge, Rep. ESL-R-614, Sept. 1975.

[14] J. K. Hale, *Functional Differential Equations*. New York: Springer, 1971.

[15] J. Zabczyk, "Remarks on the algebraic Riccati equation in Hilbert space," *Appl. Math. Optimiz.*, to be published.

[16] R. H. Kwong, "The linear quadratic Gaussian problem for systems with delays in the state, control, and observations," in *Proc. 14th Allerton Conf. Circuit and System Theory*, Sept. 29–Oct. 1, 1976.



Raymond Hon-Sing Kwong (S'71–M'75) was born in Hong Kong on September 20, 1949. He received the S.B., S.M., and Ph.D. degrees in electrical engineering from the Massachusetts Institute of Technology, Cambridge, in 1971, 1972, and 1975, respectively.

From 1971 to 1975, he held positions as a Research Assistant, a Teaching Assistant, and a Department Fellow in the Electrical Engineering Department at M.I.T. He is now a Visiting Assistant Professor of Electrical Engineering at McGill University and a Research Associate at the Centre de Recherches Mathematiques, Université de Montréal, Montréal, Québec, Canada.

Dr. Kwong is a member of Sigma Xi, Tau Beta Pi, and Eta Kappa Nu. His current research interests are in the areas of estimation and stochastic control, identification and adaptive control, and hereditary systems.

Alan S. Willsky (S'70–M'73) was born in Newark, NJ, on March 16, 1948. He received the S.B. degree in aeronautics and astronautics and the Ph.D. degree in instrumentation and control from the Massachusetts Institute of Technology, Cambridge, in 1969 and 1973, respectively.

From 1969 through 1973 he held a Fannie and John Hertz Foundation Doctoral Fellowship. He is presently an Associate Professor in the Department of Electrical Engineering and Computer Science at M.I.T. and is Assistant Director of the M.I.T. Electronic Systems Laboratory. His present research interests are in the areas of nonlinear stochastic processes, algebraic system theory, failure detection and system reliability, and the application of system-theoretic concepts to problems in digital aircraft control and biomedical data processing.

Dr. Willsky is a member of Sigma Xi, Sigma Gamma Tau, and Tau Beta Pi. In May 1975 he received an award for outstanding teaching from the M.I.T. Graduate Student Council, and in August 1975 he was the recipient of the American Automatic Control Council's Donald P. Eckman Award.