SIAM J. CONTROL AND OPTIMIZATION
Vol. 16, No. 4, July 1978

0363-0129/78/1604-0009 $1.00/0

ESTIMATION AND FILTER STABILITY OF STOCHASTIC DELAY SYSTEMS*

RAYMOND H. KWONG† AND ALAN S. WILLSKY‡

Abstract. Linear and nonlinear filtering for stochastic delay systems are studied. A representation theorem for conditional moment functionals is obtained, which, in turn, is used to derive stochastic differential equations describing the optimal linear or nonlinear filter. A complete characterization of the optimal filter is given for linear systems with Gaussian noises. 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.

1. Introduction. In recent years, the control of delay differential systems has received considerable attention. Optimal control problems for both linear as well as nonlinear delay systems have been studied intensively. In particular, there is a rather well-developed theory for the optimal control of linear delay systems with a quadratic criterion [1]–[4]. In contrast, optimal filtering for delay systems has not yet received an in-depth study. There is very little literature on the filtering of nonlinear stochastic delay systems which takes into account the structure of such systems. The linear filtering problem on a finite interval has been studied by Kwakernaak [5], Lindquist [6], Mitter and Vinter [7], and recently by Bagchi [8]. Kwakernaak's derivations in [5] were formal; Lindquist [6] did not characterize the covariance of the optimal filter; and Mitter and Vinter [7] restricted their considerations to time-invariant systems and excluded point delays in their observation equations. Bagchi [8] recently gave a rigorous derivation of the filter equations for linear systems with only point delays, using martingale theory and functional analytic methods very different from those in this paper. Stability of the linear filter was also studied recently by Vinter [9], independently of our work. He used infinite dimensional filtering methods, again quite different from our approach. In this paper, we shall study the filtering problem for both nonlinear and linear delay systems. We give a representation theorem which characterizes conditional moment functionals of nonlinear delay systems. Under certain conditions, stochastic differential equations for conditional moment functionals can be derived from the representation theorem. We then specialize these results to obtain the filtering equations for general linear delay systems. We study the stability of the optimal filter in the case of time-invariant systems with no delays in the observations. Under suitable stabilizability and detectability assumptions, we prove that the optimal filter is asymptotically stable. Finally, we combine the linear deterministic control results and the linear filtering results to show that the closed-loop linear stochastic control system is also asymptotically stable.

* Received by the editors June 14, 1976, and in revised form July 27, 1977. This research was performed at the M.I.T. Electronic Systems Laboratory and supported by the National Science Foundation Grant GK41647 and NASA Ames Grant NGL-22-009-124, and at McGill University by National Research Council of Canada under Grants A9067 and A3921, and at the Centre de Recherches Mathématiques, Université de Montréal under subvention FCAC du Ministère de l'Education du Québec.

† Department of Electrical Engineering, McGill University and Centre de Recherches Mathématiques, Université de Montréal, Montreal, Quebec, Canada. Now at Department of Electrical Engineering, University of Toronto, Toronto, Ontario, Canada.

‡ Electronic Systems Laboratory and Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139.

2. Stochastic delay differential systems. We shall study the filtering problem for stochastic delay differential systems of the form

$$(2.1)\qquad \begin{aligned} &dx(t)=f(x_t, t)\,dt+F(t)\,dw(t), \qquad t\in[0, T],\\ &x(\theta)=x_0(\theta), \qquad \theta\in[-\tau, 0]. \end{aligned}$$

The observation equation is given by

$$(2.2)\qquad \begin{aligned} &dz(t)=h(x_t, t)\,dt+N(t)\,dv(t), \qquad t\in[0, T],\\ &z(t)=0, \qquad t\leqq 0. \end{aligned}$$

All stochastic processes are defined relative to a given probability space $(\Omega, \mathcal{S}, P)$ and on an interval of the form $[0, T]$. The system process $x(t)$ takes values in R^n, the observation process $z(t)$ in R^p. The process x_t is a function on $[-\tau, 0]$ derived from $x(t)$ and is defined by

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

Unless otherwise stated, we shall let the process x_t take values in $\mathcal{C}$, the space of R^n-valued continuous functions on $[-\tau, 0]$. For simplicity, we take $w(t)$ and $v(t)$ to be 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 random function on $[-\tau, 0]$, independent of $w(t)$ and $v(t)$. The maps f and h are respectively R^n and R^p-valued functionals defined on $\mathcal{C}\times[0, T]$. 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 also write $F(t)F'(t)=Q(t)$ and $N(t)N'(t)=R(t)$.

In order for our estimation problem to be well defined, we need conditions which guarantee existence and uniqueness of solutions to the stochastic functional differential equations (2.1) and (2.2). Such questions have been studied by various authors [10]–[12]. Following their work, we assume that the following conditions are satisfied:

(A1) $f(\phi, t)$ is Borel measurable on $\mathcal{C}\times[0, T]$;

(A2) there exists a bounded measure Γ on $[-\tau, 0]$ and a positive constant K such that for ϕ and ψ in $\mathcal{C}$

$$|f(\phi, t)-f(\psi, t)|\leqq K\int_{-\tau}^{0}|\phi(s)-\psi(s)|\,d\Gamma(s)$$

and

$$|f(\phi, t)|^2\leqq K\left[1+\int_{-\tau}^{0}|\phi(s)|^2\,d\Gamma(s)\right];$$

(A3) on the interval $[-\tau, 0]$, $x(t)$ is continuous with probability 1 with $E|x(\theta)|^4<\infty$, $-\tau\leqq\theta\leqq 0$;

(A4) $h(\phi, t)$ is Borel measurable on $\mathcal{C}\times[0, T]$;

(A5)
$$\int_0^T E[h(x_t, t)'h(x_t, t)]\,dt<\infty.$$

Under these assumptions, (2.1) and (2.2) can be shown ([10]–[12]) to have a solution which is continuous w.p.1 and has bounded second moment. Furthermore x_t is a Markov process.

Since linear stochastic delay systems admit a much more complete theory, we shall, for greater clarity in our exposition, use a different notation for such system. We write

$$dx(t)=a(x_t, t)\,dt+F(t)\,dw(t), \tag{2.3}$$

$$dz(t)=c(x_t, t)\,dt+N(t)\,dv(t), \tag{2.4}$$

where $a(x_t, t)$ and $c(x_t, t)$ are given by the Lebesgue–Stieltjes integrals

$$a(x_t, t)=\int_{-\tau}^{0} d_\theta A(t, \theta)x(t+\theta),$$

$$c(x_t, t)=\int_{-\tau}^{0} d_\theta C(t, \theta)x(t+\theta).$$

Here $A(t, \theta)$ is a function on $R\times R$ jointly measurable in (t, θ), continuous in t, of bounded variation in θ for each t, with $\mathrm{Var}_{[-\tau,0]}A(t, \cdot\,)\leqq m(t)$, a locally integrable function on R^n. Furthermore $A(t, \theta)=0$ for $\theta\geqq 0$, $A(t, \theta)=A(t, -\tau)$ for $\theta\leqq-\tau$, and it is continuous from the left in θ on $(-\tau, 0)$. The function $C(t, \theta)$ is assumed to satisfy similar conditions.

It is not difficult to show (see e.g. [6], [13]) that the linear stochastic delay system (2.3)–(2.4) has a unique solution (up to almost everywhere equivalence) given by the formula

$$x(t)=\Phi(t, 0)x_0(0)+\int_{-\tau}^{0} d_\beta\left\{\int_0^{\tau} \Phi(t, s)A(s, \beta-s)\,ds\right\}x_0(\beta)+\int_0^{t} \Phi(t, s)F(s)\,dw(s) \tag{2.5}$$

where $\Phi(t, s)$ is the fundamental matrix associated with the homogeneous delay differential system

$$\dot{x}(t)=a(x_t, t)$$

(see e.g. [14]).

3. A representation theorem for conditional moment functionals. In filtering problems for stochastic differential systems, one is usually interested in estimating some function ϕ of the system process $x(t)$ given the observations $z(s)$, $0\leqq s\leqq t$. It is well known that the optimal estimate with respect to a large class of criteria is the conditional expectation $E\{\phi(x(t))/z^t\}$ where z^t denotes the σ-algebra generated by the observations $z(s)$, $0\leqq s\leqq t$. We shall also write $E\{\phi(x(t))/z^t\}$ as $E^t\{\phi[x(t)]\}$, and we shall omit the qualification of almost sure equivalence for conditional expectations. Fujisaki et al [15] have given a stochastic differential equation for the evolution of $E^t[\phi(x(t))]$ for rather general stochastic systems, which include our delay model. Specifically, they showed that for the delay system (2.1) and (2.2)

$$\begin{aligned} dE^t[\phi(x(t))]&=E^t[\mathscr{L}_t\phi(x(t))]\,dt+\{E^t[\phi(x(t))h'(x_t, t)]-E^t[\phi(x(t))]E^t[h'(x_t, t)]\}\\ &\quad\cdot R^{-1}(t)[dz(t)-E^t(h(x_t, t))\,dt] \end{aligned} \tag{3.1}$$

where $\mathscr{L}_t$ is a differential operator (see (4.7)). However, the right hand side of (3.1) contains terms of the form $E^t[g(x_t, t)]$, which we shall call conditional moment functionals. It is not clear how one can obtain an equation for these conditional moment functionals from (3.1). Equation (3.1), therefore, does not constitute a complete solution. It appears that the fundamental quantities we need to calculate are the conditional moment functionals $E^t[\phi(x_t)]$ (see also §§ 4 and 5). In this section, we will derive a representation theorem for $E^t[\phi(x_t)]$.

Our derivation is based on the work of Kunita [16]. Kunita obtained a representation theorem under the assumptions:

(i) The signal process is a stationary Markov process with compact state space, and the functional h is independent of time and continuous.

(ii) The functional $\phi: \mathscr{C} \to R$ is bounded.

We shall extend his results by making only the assumptions:

(A6) $$E|\phi(x_t)|^2 < \infty,$$

(A7) $$\int_0^T E|\phi(x_t)h(x_t, t)|^2 \, dt < \infty.$$

THEOREM 3.1. *Suppose* (A1)–(A7) *hold. Then we have the following representation for the conditional expectation of* ϕ *given* z^t:

$$E^t[\phi(x_t)] = E[\phi(x_t)] + \int_0^t E^s\{E[\phi(x_t)|x_s][h'(x_s, s) - E^s(h'(x_s, s))]\}R^{-1}(s)\, d\nu(s) \tag{3.2}$$

where $\nu(t) = z(t) - \int_0^t E^s[h(x_s, s)]\, ds$ *is the innovations.*

Proof. We follow the approach of Kunita [16]. First suppose that ϕ is bounded. Let $\mathscr{G}^t$ denote the σ-algebra $\sigma\{x(s), v(s); s \leqq t\}$. Clearly $z^t \subset \mathscr{G}^t$. Moreover, by the Markov property of x_t and the independence of the x and v processes, $E[\phi(x_t)|\mathscr{G}^s] = E[\phi(x_t)|x_s]$, for $t \geqq s$.

By the assumptions of the theorem, all terms in (3.2) are in $L^2(\Omega, z^t, P)$. Thus, just as in [16], it is sufficient to verify that

$$\begin{aligned} &E\{[E^t(\phi(x_t)) - E(\phi(x_t))]Y_t\} \\ &\quad = E\left\{\int_0^t E^s[E(\phi(x_t)|x_s)[h'(x_s, s) - E^s(h'(x_s, s))]]R^{-1}(s)\, d\nu(s) Y_t\right\} \end{aligned} \tag{3.3}$$

for all Y_t represented as $\int_0^t g_s'\, d\nu(s)$, with g_s a jointly measurable and z^t-adapted process such that $\int_0^t E|g_s|^2\, ds < \infty$.

Using the independence of the x and v processes, we conclude, on following the same argument as Kunita [16], that

$$\begin{aligned} &E\left\{[E^t(\phi(x_t)) - E(\phi(x_t))] \int_0^t g_s'\, d\nu(s)\right\} \\ &\quad = \int_0^t E\{\phi(x_t) g_s'[h(x_s, s) - E^s(h(x_s, s))]\}\, ds \\ &\quad = \int_0^t E\{E[\phi(x_t)|\mathscr{G}^s] g_s'[h(x_s, s) - E^s(h(x_s, s))]\}\, ds \\ &\quad = \int_0^t E\{E[\phi(x_t)|x_s] g_s'[h(x_s, s) - E^s(h(x_s, s))]\}\, ds. \end{aligned} \tag{3.4}$$

Since

$$\begin{aligned} &E\left\{\int_0^t E^s[E(\phi(x_t)|x_s)[h'(x_s, s) - E^s(h'(x_s, s))]]R^{-1}(s)\, d\nu(s) \int_0^t g_s'\, d\nu(s)\right\} \\ &\quad = \int_0^t E\{E[\phi(x_t)|x_s][h'(x_s, s) - E^s(h'(x_s, s))] g_s\}\, ds \end{aligned} \tag{3.5}$$

we obtain, on combining (3.4) and (3.5), the desired equation (3.3). Thus the theorem is true if ϕ is bounded.

In the general case, let $\phi_N(x_t)=\phi(x_t)\chi_N$, where $\chi_N=1$ if $|\phi(x_t)|\leqq N$, and $\chi_N=0$ if $|\phi(x_t)|>N$. Clearly $E|\phi_N(x_t)-\phi(x_t)|\to 0$ as $N\to\infty$. Since ϕ_N is bounded, the above development yields

$$\begin{aligned} E^t[\phi_N(x_t)] &= E[\phi_N(x_t)] \\ &\quad + \int_0^t E^s\{E[\phi_N(x_t)|x_s][h(x_s,s)-E^s(h(x_s,s))]'\}R^{-1}(s)\,d\nu(s). \end{aligned} \tag{3.6}$$

By assumption (A7), the last term on the right hand side of (3.6) converges in probability to

$$\int_0^t E^s\{E[\phi(x_t)|x_s][h(x_s,s)-E^s(h(x_s,s))]'\}R^{-1}(s)\,d\nu(s)$$

(see, for example, [17]). Hence, on letting $N\to\infty$ in (3.6), we finally obtain (3.2). The proof is completed.

The following corollary is immediate.

COROLLARY 3.1. *The smoothed estimate* $E^t[x(t+\theta)]$, $-\tau\leqq\theta<0$, *is given by*

$$\begin{aligned} E^t[x(t+\theta)] &= E^{t+\theta}[x(t+\theta)] \\ &\quad + \int_{t+\theta}^t E^s\{x(t+\theta)[h'(x_s,s)-E^s(h'(x_s,s))]\}R^{-1}(s)\,d\nu(s). \end{aligned} \tag{3.7}$$

Remark 3.1. Theorem 3.1 remains true if we merely assume that the signal process is a Markov process with state space a separable complete metric space. The same proof goes through for this more general case. Theorem 3.1 corresponds to the special situation where the signal process is the Markov process x_t generated by the stochastic delay equation (2.1). Similar remarks also apply to Theorem 4.1 in the next section on stochastic differential equations for the nonlinear filtering problem.

4. Stochastic differential equations for nonlinear filtering of delay systems. While Theorem 3.1 gives an abstract representation for the optimal estimates, it is completely nonrecursive in the sense that knowledge of $E[\phi(x_t)/z^t]$ is of no use in determining $E[\phi(x_{t+\Delta})/z^{t+\Delta}]$. In fact, for every t, we must completely reprocess our past observations. For implementation and approximation purposes, one would like to obtain a stochastic differential equation for the evolution of $E[\phi(x_t)/z^t]$. In this section, we shall give conditions on ϕ under which we can obtain a stochastic differential equation for $E[\phi(x_t)/z^t]$. As we shall see, these conditions are intimately related to the (extended) infinitesimal generator of the Markov process x_t [15].

DEFINITION. A family of linear operators A_t, $t\in[0,T]$ defined on the space of real-valued measurable functions on $\mathscr{C}$ is called an *(extended) infinitesimal generator* if

$$E[\phi(x_t)/x_s]-\phi(x_s)=\int_s^t E[A_u\phi(x_u)/x_s]\,du \tag{4.1}$$

is satisfied for all $0\leqq s<t\leqq T$. We use the notation $D(A)$ to denote the space of all functionals ϕ satisfying $E|\phi(x_t)|^2<\infty$, $\int_0^T E|A_t\phi(x_t)|^2\,dt<\infty$, and (4.1).

Define the process $e_h(t)=h(x_t,t)-E^t[h(x_t,t)]$. Then we have

THEOREM 4.1. *Let the conditions of Theorem* 3.1 *be satisfied. In addition, let* ϕ *belong to* $D(A)$ *and suppose that*

$$\int_0^T E|[A_t\phi(x_t)]h(x_t,t)|^2\,dt<\infty.$$

Then the functional $E[\phi(x_t)|z^t]$ *satisfies the stochastic differential equation*

$$dE[\phi(x_t)|z^t]=E[A_t\phi(x_t)/z^t]\,dt+E[\phi(x_t)e_h'(t)/z^t]R^{-1}(t)\,d\nu(t). \tag{4.2}$$

Proof. For any $t\in[0,T]$ and $\varepsilon>0$, we have, by a simple calculation,

$$\begin{aligned}&E[\phi(x_{t+\varepsilon})/z^{t+\varepsilon}]-E[\phi(x_t)/z^t]\\&=E[\phi(x_{t+\varepsilon})-\phi(x_t)]+\int_0^t E\{E[\phi(x_{t+\varepsilon})-\phi(x_t)/x_s]e_h'(s)/z^s\}R^{-1}(s)\,d\nu(s)\\&\quad+\int_t^{t+\varepsilon}E\{E[\phi(x_{t+\varepsilon})/x_s]e_h'(s)/z^s\}R^{-1}(s)\,d\nu(s).\end{aligned} \tag{4.3}$$

Using (4.1), we get that

$$\begin{aligned}&\int_0^t E\{E[\phi(x_{t+\varepsilon})-\phi(x_t)/x_s]e_h'(s)/z^s\}R^{-1}(s)\,d\nu(s)\\&=\int_0^t E\left\{E\left\{\int_t^{t+\varepsilon}E[A_u\phi(x_u)/x_t]\,du/x_s\right\}e_h'(s)/z^s\right\}R^{-1}(s)\,d\nu(s)\\&=\int_0^t E\left\{\int_t^{t+\varepsilon}E[A_u\phi(x_u)/x_s]e_h'(s)/z^s\}R^{-1}(s)\,du\,d\nu(s)\right.\\&=\int_t^{t+\varepsilon}\int_0^t E\{E[A_u\phi(x_u)/x_s]e_h'(s)/z^s\}R^{-1}(s)\,d\nu(s)\,du\end{aligned} \tag{4.4}$$

with the last equality justified in view of the assumptions of the theorem. Similarly

$$\begin{aligned}&\int_t^{t+\varepsilon}E\{E[\phi(x_{t+\varepsilon})/x_s]e_h'(s)/z^s\}R^{-1}(s)\,d\nu(s)\\&=\int_t^{t+\varepsilon}E[\phi(x_s)e_h'(s)/z^s]R^{-1}(s)\,d\nu(s)\\&\quad+\int_t^{t+\varepsilon}\int_s^{t+\varepsilon}E\{E[A_u\phi(x_u)/x_s]e_h'(s)/z^s\}\,du\,R^{-1}(s)\,d\nu(s).\end{aligned} \tag{4.5}$$

Finally, using the representation theorem for $A_t\phi(x_t)$, we find that

$$\begin{aligned}E[\phi(x_{t+\varepsilon})-\phi(x_t)]&=\int_t^{t+\varepsilon}E[A_u\phi(x_u)]\,du\\&=\int_t^{t+\varepsilon}E[A_u\phi(x_u)/z^u]\,du\\&\quad-\int_t^{t+\varepsilon}\int_0^u E\{E[A_u\phi(x_u)/x_s]e_h'(s)/z^s\}R^{-1}(s)\,d\nu(s)\,du.\end{aligned} \tag{4.6}$$

Adding up (4.4) to (4.6) yields

$$\begin{aligned}&E[\phi(x_{t+\varepsilon})/z^{t+\varepsilon}]-E[\phi(x_t)/z^t]\\&=\int_t^{t+\varepsilon}E[A_u\phi(x_u)/z^u]\,du+\int_t^{t+\varepsilon}E[\phi(x_u)e_h'(u)/z^u]R^{-1}(u)\,d\nu(u)\end{aligned}$$

which is precisely (4.2).

Remark 4.1. Theorem 4.1 is a generalization to systems with delays of the usual formula for conditional moments of ordinary diffusion processes. While the form of the stochastic differential equation is exactly the same as that for diffusion processes, here we need to know the structure of the infinitesimal generator of the x_t process. We know from (4.1) that functionals of the form $\phi[x(t+\theta)]$, $\theta \in (-\tau, 0)$ do not belong to $D(A)$, since $x(t)$ is not in general differentiable. Hence it is not possible to derive a stochastic differential equation for a functional of the form $\phi[x(t+\theta)]$. Indeed, a complete characterization of the operators A_t is not known, although certain special classes of functionals which are in the domain of A_t have been stated in Kushner [12]. We mention these results to illustrate Theorem 4.1.

Case 1. Suppose the functional $\phi(x_t) = \phi[x(t)]$, and is twice continuously differentiable in its argument. Then

$$A_t\phi[x(t)] \equiv \mathscr{L}_t\phi[x(t)] = f(x_t, t)'\phi_x[x(t)] + \tfrac{1}{2}\operatorname{tr} Q(t)\phi_{xx}[x(t)] \tag{4.7}$$

where ϕ_x is the n-vector whose ith component is $(\partial\phi/\partial x_i)[x(t)]$. In particular

$$dE^t[x(t)] = E^t[f(x_t, t)]\,dt + \{E^t[x(t)h'(x_t, t)] - E^t[x(t)]E^t[h'(x_t, t)]\}R^{-1}(t)\,d\nu(t) \tag{4.8}$$

In this case, (4.2) reduces to the well-known results of Fujisaki et al. [15].

Case 2. Let $\phi(x_t) = \int_{-\tau}^{0} \psi(\theta)g[x(t+\theta), x(t)]\,d\theta$, where ψ is continuously differentiable on $[-\tau, 0]$, and g is twice continuously differentiable in its second argument. Then

$$\begin{aligned} A_t\phi(x_t) = {} & \psi(0)g[x(t), x(t)] - \psi(-\tau)g[x(t-\tau), x(t)] \\ & - \int_{-\tau}^{0} \dot{\psi}(\theta)g[x(t+\theta), x(t)]\,d\theta + \int_{-\tau}^{0} \psi(\theta)\mathscr{L}_t g[x(t+\theta), x(t)]\,d\theta \end{aligned} \tag{4.9}$$

where $\mathscr{L}_t$ is the operator defined in Case 1 and acts on g as a function of $x(t)$ only.

Case 3. Let $\phi(x_t) = D[F(x_t)]$ where D is a twice continuously differentiable real-valued function, and $F(x_t) = \int_{-\tau}^{0} \psi(\theta)g[x(t+\theta), x(t)]\,d\theta$ is the type of functional described in Case 2. Then

$$A_t\phi(x_t) = D_\alpha(\alpha)|_{\alpha=F(x_t)}A_tF(x_t) + \tfrac{1}{2}D_{\alpha\alpha}(\alpha)|_{\alpha=F(x_t)} \cdot G$$

where

$$G = \int_{-\tau}^{0}\int_{-\tau}^{0} \psi(\theta)\psi(\eta)\sum_{i,j} g_{\beta_i}[x(t+\theta), x(t)]g_{\beta_j}[x(t+\theta), x(t)]Q_{ij}(t)\,d\theta\,d\eta$$

and g_{β_i} denotes partial differentiation of g with respect to the ith component of the second argument.

From the above special cases, we can see that basically we need twice continuous differentiability of ϕ with respect to the dependence on $x(t)$, and Fréchet differentiability with respect to the dependence on the piece of the trajectory x_t. As discussed before, this rules out functionals of the form $\phi[x(t+\theta)]$, $\theta \in [-\tau, 0)$. Hence for nonlinear systems with point delays, any attempt in deriving stochastic differential equations for conditional moment functionals will have to face the difficulty of functionals not being in the domain of the generator of the Markov process x_t. For example, if the observation process is of the form

$$dz(t) = \{h_1[x(t)] + h_2[x(t-\tau)]\}\,dt + dv(t)$$

$\phi[x(t)]h_2[x(t-\tau)]$ will not be in the domain of A_t. On the other hand, in order to analyze (4.2), we do need to calculate the conditional expectation for

$$\phi[x(t)]h_2[x(t-\tau)]$$

Of course, there are many physical problems (for example, radar problems with spread targets [23]) where the observations are of the form $h(x_t)=\int_{-\tau}^{0}\psi(\theta)H[x(t+\theta),x(t)]\,d\theta$. Moreover, one can approximate point delays by distributed delays of the above form. This will allow us to write a stochastic differential equation for $\phi[x(t)]h(x_t)$. However, we will then get the unknown $A_t\phi[x(t)]h(x_t)$ in our equation for $\phi[x(t)]h(x_t)$. If $\psi(-\tau)\neq 0$, $A_t\phi[x(t)]h(x_t)$ will contain a term with point delay (see Case 2 above), and we are faced with the same problems as before. In general, if the functionals involved are in the domain of A_t^i, $i=1,\cdots,n$, we can write n coupled stochastic differential equations involving the moment functionals, just as in the diffusion process case. It should be clear from the above discussion that this puts rather severe restrictions on the functionals involved.

There is, however, one special case where the optimal filter can be completely specified even when there are point delays in the system. This is the linear case with Gaussian distributions and will be treated next.

5. Optimal filtering of linear stochastic delay systems. Consider the linear stochastic delay system defined by

$$\begin{aligned} dx(t)&=a(x_t,t)\,dt+F(t)\,dw(t),\\ x(\theta)&=x_0(\theta),\qquad \theta\in[-\tau,0];\end{aligned}\tag{5.1}$$

$$dz(t)=c(x_t,t)\,dt+N(t)\,dv(t)\tag{5.2}$$

where $a(\cdot,\cdot)$ and $c(\cdot,\cdot)$ are as described in § 2. Also, we take x_0 to be a Gaussian process on $[-\tau,0]$ with mean $\bar{x}_0(\theta)$ and cov $[x_0(\theta);x_0(\xi)]=\sum_0(\theta,\xi)$. By an argument similar to the case without delays, it is readily seen that the conditional distribution of $x(t+\theta)$, for any $\theta\in[-\tau,0]$, given $z(s)$, $0\leqq s\leqq t$, is Gaussian. We shall write $\hat{x}(t+\theta/t)$ to denote $E(x(t+\theta)/z^t\}$, $\theta\in[-\tau,0]$. Using (4.8), we immediately obtain the following stochastic differential equation for the conditional mean

$$\begin{aligned} d\hat{x}(t/t)=&\int_{-\tau}^{0}d_\theta A(t,\theta)\hat{x}(t+\theta/t)\,dt\\ &+\left[\int_{-\tau}^{0}E^t(x(t)x(t+\theta)')\,d_\theta C'(t,\theta)\right.\\ &\left.-\int_{-\tau}^{0}\hat{x}(t/t)\hat{x}(t+\theta/t)'\,d_\theta C'(t,\theta)\right]R^{-1}(t)\,d\nu(t)\end{aligned}\tag{5.3}$$

with the innovations $\nu(t)$ given by

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

Define the "smoothed" conditional error covariance as

$$P(t,\theta,\xi)=E^t\{[x(t+\theta)-\hat{x}(t+\theta/t)][x(t+\xi)-\hat{x}(t+\xi/t)]'\},\quad -\tau\leqq\theta,\quad \xi\leqq 0.$$

Then (5.3) can be rewritten as

$$d\hat{x}(t/t)=\int_{-\tau}^{0}d_\theta\,A(t,\theta)\hat{x}(t+\theta/t)\,dt+\int_{-\tau}^{0}P(t,0,\theta)\,d_\theta\,C'(t,\theta)R^{-1}(t)\,d\nu(t).\tag{5.4}$$

To evaluate the unknown terms on the right hand side of (5.4), we use (3.16) to write the smoothed estimate as

$$\hat{x}(t+\theta/t)=\hat{x}(t+\theta/t+\theta)+\int_{t+\theta}^{t}\int_{-\tau}^{0}P(s,t+\theta-s,\xi)\,d_\xi C(s,\xi)'R^{-1}(s)\,d\nu(s) \quad \text{for } \theta\in[-\tau,0]. \tag{5.5}$$

An inspection of (5.4) and (5.5) shows that the optimal linear filter is completely characterized by $\hat{x}(t+\theta/t)$, $-\tau\leqq\theta\leqq 0$, and the "smoothed" error covariance function $P(t,\theta,\xi)$. It remains only to derive appropriate equations for $P(t,\theta,\xi)$. Since the processes x and z are jointly Gaussian, the error process $x(t+\theta)-\hat{x}(t+\theta|t)$ is independent of $z(s)$, $s\leqq t$ (see, for example, Bagchi [8]). Hence $P(t,\theta,\xi)$ is independent of the observations and equals $E\{[x(t+\theta)-\hat{x}(t+\theta|t)][x(t+\xi)-\hat{x}(t+\xi|t)]'\}$. This fact will enable us to simplify the derivations of the equations for $P(t,\theta,\xi)$. The next theorem summarizes the complete structure of the optimal filter.

THEOREM 5.1. *The optimal filter for the system* (5.1)–(5.2) *is characterized as follows*:

(i) *The conditional mean* $\hat{x}(t/t)$ *satisfies* (5.4).

(i) *The smoothed estimate* $\hat{x}(t+\theta/t)$ *satisfies* (5.5).

(iii) *The smoothed error covariance* $P(t,\theta,\xi)$ *satisfies the equations*

$$\begin{aligned}\frac{d}{dt}P(t,0,0)=&\int_{-\tau}^{0}P(t,0,\theta)\,d_\theta A'(t,\theta)+\int_{-\tau}^{0}d_\theta A(t,\theta)P(t,\theta,0)\\&-\int_{-\tau}^{0}\int_{-\tau}^{0}P(t,0,\theta)\,d_\theta C'(t,\theta)R^{-1}(d)\,d_\xi C(t,\xi)P(t,\xi,0)+Q(t),\end{aligned}\tag{5.6}$$

$$\begin{aligned}\sqrt{2}P_\eta(t,\theta,0)=&\int_{-\tau}^{0}P(t,\theta,\xi)\,d_\xi A'(t,\xi)-\int_{-\tau}^{0}\int_{-\tau}^{0}P(t,\theta,\xi)\,d_\xi C'(t,\xi)R^{-1}(t)\\&\cdot d_\alpha C(t,\alpha)P(t,\alpha,0),\end{aligned}\tag{5.7}$$

$$\sqrt{3}P_\sigma(t,\theta,\xi)=-\int_{-\tau}^{0}\int_{-\tau}^{0}P(t,\theta,\beta)\,d_\beta C'(t,\beta)R^{-1}(t)\,d_\alpha C(t,\alpha)P(t,\alpha,\xi)\tag{5.8}$$

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*

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

$$P(0,\theta,\xi)=\textstyle\sum_0(\theta,\xi),\qquad -\tau\leqq\theta,\quad \xi\leqq 0.$$

Proof. See Appendix A.

Remark 5.1. Equations similar to those of (5.6)–(5.8) for $P(t,\theta,\xi)$ were formally derived by Kwakernaak in [5], and rigorously rederived by Bagchi in [8], for systems with point delays only. Instead of directional derivatives, they used partial derivations with respect to the variables θ and ξ. In the general case, however, $P(t,\theta,\xi)$ will not be continuously differentiable in (t,θ,ξ). This is why directional derivatives have to be used. A similar situation has already been noted in the quadratic optimal control problem for linear delay systems [20].

In the special case where $x_0\equiv 0$, $a(x_t,t)=Ax(t)+Bx(t-\tau)$, $c(x_t,t)=Cx(t)$, Q, R constant matrices, it can be shown directly or by exploiting the connection between linear optimal filtering and optimal control with quadratic criterion (see § 6) that

$P(t, \theta, \xi)$ is in fact continuously differentiable in (t, θ, ξ). When we compare the solutions to the linear optimal control and optimal linear filtering problems in our study of filter stability, it will be helpful to use the notation $P_0(t) = P(t, 0, 0)$, $P_1(t, \theta) = P(t, \theta, 0)$, and $P_2(t, \theta, \xi) = P(t, \theta, \xi)$. In this case, the optimal filter is given by the equations

$$d\hat{x}(t/t) = A\hat{x}(t/t)\,dt + B\hat{x}(t-\tau/t)\,dt + P_0(t)C'R^{-1}[dz(t) - C\hat{x}(t/t)\,dt]; \tag{5.9}$$

$$\hat{x}(t-\tau/t) = \hat{x}(t-\tau/t-\tau) + \int_{t-\tau}^{t} P_1(s, t-\tau-s)C'R^{-1}[dz(s) - \hat{C}x(s/s)\,ds], \qquad \hat{x}(\theta/0) = 0, \quad -\tau \leqq \theta \leqq 0; \tag{5.10}$$

$$\frac{d}{dt}P_0(t) = AP_0(t) + P_0(t)A' - P_0(t)C'R^{-1}CP_0(t) + Q + BP_1(t, -\tau) + P_1'(t, -\tau)B'; \tag{5.11}$$

$$\left(\frac{\partial}{\partial t} - \frac{\partial}{\partial \theta}\right)P_1(t, \theta) = P_1(t, \theta)[A' - C'R^{-1}CP_0(t)] + P_2(t, \theta, -\tau)B'; \tag{5.12}$$

$$\left(\frac{\partial}{\partial t} - \frac{\partial}{\partial \theta} - \frac{\partial}{\partial \xi}\right)P_2(t, \theta, \xi) = -P_1(t, \theta)C'R^{-1}CP_1(t, \xi); \tag{5.13}$$

with

$$\begin{aligned} &P_0(0) = P_1(0, \theta) = P_2(0, \theta, \xi) = 0, \\ &P_1(t, 0) = P_0(t), \qquad P_2(t, \theta, 0) = P_1(t, \theta), \\ &P_0(t) = P_0'(t), \qquad P_2(t, \theta, \xi) = P_2'(t, \xi, \theta). \end{aligned} \tag{5.14}$$

Notice that in this special case where there are no delays in the observations, $\hat{x}(t/t)$ depends only on $\hat{x}(s/s)$, $t-\tau \leqq s \leqq t$, and from (5.9) and (5.10), we can obtain the following explicit stochastic delay equation for $\hat{x}(t/t)$:

$$\begin{aligned} d\hat{x}(t|t) &= [A\hat{x}(t|t) + B\hat{x}(t-\tau|t-\tau)]\,dt + P_0(t)C'R^{-1}[dz(t) - C\hat{x}(t|t)\,dt] \\ &\quad + \int_{t-\tau}^{t} BP_1(s, t-\tau-s)C'R^{-1}[dz(s) - C\hat{x}(s|s)\,ds]\,dt, \\ \hat{x}(\theta|0) &= 0, \qquad -\tau \leqq \theta \leqq 0. \end{aligned} \tag{5.15}$$

This will not be true if we have delays in the observations (see the discussions in [18] and [24]).

6. Stability of linear optimal filters and control systems. In this section, we study the stability of optimal linear filters and stochastic control systems for linear delay systems. We shall concentrate on the filters defined by (5.9)–(5.13). The extension to the case with multiple delays in the system dynamics is straightforward. However, the situation for systems with delays in the observations is much more complicated and will be treated separately in a forthcoming paper. In our analysis, we make essential use of the duality between optimal filtering of linear stochastic delay systems and optimal control of linear delay systems with quadratic cost. These results complete the extension of the well-known linear quadratic Gaussian theory to systems with delays in the dynamics.

We begin by summarizing the results for the optimal control of linear delay systems with quadratic cost [1]–[4]. Consider the system

$$\frac{dx}{dt} = Ax(t) + Bx(t-\tau) + Cu(t), \tag{6.1}$$
$$x(\theta) = x_0(\theta), \qquad \theta \in [-\tau, 0].$$

In previous sections, we have used the space $\mathscr{C}$ as our state space. For this system, however, we may allow the initial function x_0 to lie in the larger space $R^n \times L^2$ (see [4] or [20]). The admissible control set U is the set of R^m-valued L_2 functions on $[0, T]$. The cost functional is given by

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

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

$$u^*(t) = -S^{-1}C'K_0(t)x(t) - S^{-1}C' \int_{-\tau}^0 K_1(t, \theta)x(t+\theta)\, d\theta. \tag{6.2}$$

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

$$\frac{d}{dt}K_0(t) = -A'K_0(t) - K_0(t)A + K_0(t)CS^{-1}C'K_0(t) - M - K_1(t, 0) - K_1'(t, 0), \tag{6.3}$$

$$\left(\frac{\partial}{\partial t} - \frac{\partial}{\partial \theta}\right)K_1(t, \theta) = -[A' - K_0(t)CS^{-1}C']K_1(t, \theta) - K_2(t, 0, \theta), \tag{6.4}$$

$$\left(\frac{\partial}{\partial t} - \frac{\partial}{\partial \theta} - \frac{\partial}{\partial \xi}\right)K_2(t, \theta, \xi) = K_1'(t, \theta)CS^{-1}C'K_1(t, \xi), \tag{6.5}$$

with

$$K_0(T) = K_1(T, \theta) = K_2(T, \theta, \xi) = 0 \qquad -\tau \leqq \theta, \quad \xi \leqq 0,$$
$$K_1(t, -\tau) = K_0(t)B, \tag{6.6}$$
$$K_2(t, -\tau, \theta) = B'K_1(t, \theta),$$
$$K_0(t) = K_0'(t), \qquad K_2(t, \theta, \xi) = K_2(t, \xi, \theta)'.$$

The optimal cost can be expressed as

$$J_T^*(x_0) = x_0'(0)K_0(0)x_0(0) + \int_{-\tau}^0 x_0'(0)K_1(0, \theta)x_0(\theta)\, d\theta \tag{6.7}$$
$$+ \int_{-\tau}^0 x_0'(\theta)K_1'(0, \theta)x_0(0)\, d\theta + \int_{-\tau}^0 \int_{-\tau}^0 x_0'(\theta)K_2(0, \theta, \xi)x_0(\xi)\, d\theta\, d\xi.$$

We now consider the infinite time control problem, i.e., $T = \infty$. To discuss this problem, we need some condition to ensure that the optimal cost will be finite. The relevant concepts are those of stabilizability and detectability. These definitions for the case of delay systems are given below.

DEFINITION 6.1. The system

$$\dot{x}(t) = Ax(t) + Bx(t-\tau) + Cu(t) \tag{6.8}$$

is said to be *stabilizable* if there exist matrices L_0, L_1, and $L_2(\theta)$, $\theta \in [-\tau, 0]$, with L_2 strongly measurable and bounded, such that

$$\dot{x}(t) = (A + CL_0)x(t) + (B + CL_1)x(t-\tau) + \int_{-\tau}^{0} CL_2(\theta)x(t+\theta)\,d\theta \tag{6.9}$$

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

DEFINITION 6.2. The system

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bx(t-\tau), \\ z(t) &= Cx(t) \end{aligned} \tag{6.10}$$

is said to be *detectable* if there exist matrices K_0, K_1, and $K_2(\theta)$, $\theta \in [-\tau, 0]$, with K_2 strongly measurable and bounded, such that

$$\dot{x}(t) = Ax(t) + Bx(t-\tau) + K_0 z(t) + K_1 z(t-\tau) + \int_{-\tau}^{0} K_2(\theta) z(t+\theta)\,d\theta \tag{6.11}$$

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

The following proposition can be easily proved from the above definitions [18].

PROPOSITION 6.1. *The system* (6.10) *is detectable if and only if the "adjoint" system (which runs backwards in time)*

$$\dot{y}(t) = -A'y(t) - B'y(t+\tau) - C'u(t) \tag{6.12}$$

is stabilizable.

The properties of stabilizability and detectability, and their relationships to controllability and observability, are further discussed in [18], [25], [26], to which the reader is referred.

We can now state the result concerning the infinite time quadratic control problem. Let $M = H'H$.

PROPOSITION 6.2 ([19], [20], [21]). *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 $T \to \infty$ in the following sense:*

$$\begin{aligned} K_1(t, \cdot) &\to K_1(\cdot) && \textit{strongly in } L_2[-\tau, 0], \\ K_2(t, \cdot, \cdot) &\to K_2(\cdot, \cdot) && \textit{strongly in } L_2[-\tau, 0] \times L_2[-\tau, 0]. \end{aligned}$$

The optimal control law for the infinite time problem is given by

$$u^*(t) = -S^{-1}C'K_0 x(t) - \int_{-\tau}^{0} S^{-1}C'K_1(\theta)x(t+\theta)\,d\theta \tag{6.13}$$

where K_0, $K_1(\theta)$ and $K_2(\theta, \xi)$ satisfy the following set of equations

$$A'K_0 + K_0 A - K_0 CS^{-1}C'K_0 + M + K_1'(0) + K_1(0) = 0, \tag{6.14}$$

$$\frac{d}{d\theta}K_1(\theta) = [A' - K_0 CS^{-1}C']K_1(\theta) + K_2(0, \theta), \tag{6.15}$$

$$\left(\frac{\partial}{\partial\theta} + \frac{\partial}{\partial\xi}\right)K_2(\theta, \xi) = -K_1'(\theta)CS^{-1}C'K_1(\xi), \tag{6.16}$$

with

$$\begin{aligned} K_1(-\tau) &= K_0 B, & K_2(\theta, -\tau) &= K_1'(\theta)B, \\ K_0 &= K_0', & K_2(\theta, \xi) &= K_2'(\xi, \theta). \end{aligned} \tag{6.17}$$

Furthermore, the optimal closed-loop system is asymptotically stable with the optimal *cost given by*

$$J_\infty^*(x_0)=x_0'(0)K_0x_0(0)+\int_{-\tau}^0 x_0'(0)K_1(\theta)x_0(\theta)\,d\theta \tag{6.18}$$
$$+\int_{-\tau}^0 x_0'(\theta)K_1'(\theta)x_0(0)\,d\theta+\int_{-\tau}^0\int_{-\tau}^0 x_0'(\theta)K_2(\theta,\xi)x_0(\xi)\,d\theta\,d\xi.$$

Remark 6.1. Proposition 6.2 is an extension of the result of [19] and [20] where the matrix M is assumed to be positive definite. One of the authors first proved in [18] that the condition $M>0$ can be relaxed to (A, B, H) observable. Subsequently, the work of Zabczyk [21] became known to the authors and the present conclusions, assuming the still weaker condition of detectability, can be obtained from the results of [21].

To connect the optimal control result of Proposition 6.2 with those of optimal filtering, we need the following duality theorem which can be deduced from the work of Lindquist [13].

PROPOSITION 6.3. *Consider the optimal filtering problem over the interval* $[0, T]$ *for the system*

$$dx(t)=[Ax(t)+Bx(t-\tau)]\,dt+F\,dw(t), \tag{6.19}$$
$$x(\theta)=0,\qquad \theta\leqq 0;$$
$$dz(t)=Cx(t)\,dt+N\,dv(t). \tag{6.20}$$

Define the dual control system by

$$\dot y(t)=-A'y(t)-B'y(t+\tau)-C'u(t) \tag{6.21}$$

with

$$y(T)=b,\quad y(s)=0,\quad s>T. \tag{6.22}$$

The dual control problem is defined to be to minimize

$$J_T(b,u)=\int_0^T [y'(t)Qy(t)+u'(t)Ru(t)]\,dt \tag{6.23}$$

where $Q=FF'\geqq 0$ *and* $R=NN'>0$. *Let the optimal linear least squares estimate of* $x(T)$ *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*

$$b'\hat x(T/T)=-\int_0^T u_T'(s)\,dz(s). \tag{6.24}$$

We now have two representations of $b'\hat x(T/T)$, one directly from (5.9)–(5.13), the other indirectly from (6.24). Our strategy is to compare the two representations and identify the control and filter gains appropriately. This will enable us to exploit the known results of the optimal control problem to conclude filter stability. We begin by stating the following lemma.

LEMMA 6.1. *The conditional mean of* $x(T)$, *denoted by* $\hat x(T/T)$, *is given either by*

$$\hat x(T/T)=\int_0^T\Big\{\Phi(T,s)P_0(s)C'R^{-1} \tag{6.25}$$
$$+\int_0^{\min(\tau,T-s)}\Phi(T,s+\theta)BP_1(s,\theta-\tau)C'R^{-1}\,d\theta\Big\}\,dz(s)$$

or by

$$\hat{x}(T/T)=\int_0^T\left\{Y'(t,T)\tilde{K}_0(t)C'R^{-1}+\int_0^{\min(\tau,T-t)}Y(t+\theta,T)'\tilde{K}_1(t,\theta)C'R^{-1}\,d\theta\right\}dz(t). \tag{6.26}$$

Here $P_0(t)$, $P_1(t,\theta)$, *and* $P_2(t,\theta,\xi)$ *are given by* (5.11)–(5.14) *and* $\Phi(t,s)$ *is the fundamental matrix* [14] *associated with the delay equation*

$$\dot{x}(t)=[A-P_0(t)C'R^{-1}C]x(t)+Bx(t-\tau)-B\int_{-\tau}^0 P_1(t+\theta,-\theta-\tau)C'R^{-1}Cx(t+\theta)\,d\theta. \tag{6.27}$$

The functions $\tilde{K}_0(t)$, $\tilde{K}_1(t,\theta)$ *and* $\tilde{K}_2(t,\theta,\xi)$ *satisfy the equations*

$$\dot{\tilde{K}}_0(t)=A\tilde{K}_0(t)+\tilde{K}_0(t)A'-\tilde{K}_0(t)C'R^{-1}C\tilde{K}_0(t)+Q+\tilde{K}_1(t,0)+\tilde{K}_1'(t,0), \tag{6.28}$$

$$\left(\frac{\partial}{\partial t}-\frac{\partial}{\partial\theta}\right)\tilde{K}_1(t,\theta)=[A-\tilde{K}_0(t)C'R^{-1}C]\tilde{K}_1(t,\theta)+\tilde{K}_2(t,0,\theta), \tag{6.29}$$

$$\left(\frac{\partial}{\partial t}-\frac{\partial}{\partial\theta}-\frac{\partial}{\partial\xi}\right)\tilde{K}_2(t,\theta,\xi)=-\tilde{K}_1'(t,\theta)C'R^{-1}C\tilde{K}_1(t,\xi) \tag{6.30}$$

with

$$\begin{gathered}\tilde{K}_0(0)=\tilde{K}_1(0,\theta)=\tilde{K}_2(0,\theta,\xi)=0,\qquad 0\leqq\theta,\quad\xi\leqq\tau,\\ \tilde{K}_1(t,\tau)=\tilde{K}_0(t)B',\\ \tilde{K}_2(t,\tau,\xi)=B\tilde{K}_1(t,\xi),\\ \tilde{K}_0(t)=\tilde{K}_0'(t),\qquad \tilde{K}_2(t,\theta,\xi)=\tilde{K}_2'(t,\xi,\theta)\end{gathered} \tag{6.31}$$

and $Y(t,s)$ *is the fundamental matrix associated with the system*

$$\frac{dy}{dt}=-[A'-C'R^{-1}C\tilde{K}_0(t)]y(t)-B'y(t+\tau)+\int_0^\tau C'R^{-1}C\tilde{K}_1(t,\theta)y(t+\theta)\,d\theta. \tag{6.32}$$

Proof. This simply involves solving for $\hat{x}(T/T)$ explicitly from (5.9)–(5.13) and from (6.24) and making appropriate changes of variables. For details, the reader may consult [18].

Next, we relate the various quantities involved in (6.25) and (6.26) in

LEMMA 6.2. *The optimal filter gains* (5.11)–(5.14) *are related to the optimal control gains* (6.28)–(6.31) *for the dual problem by*

$$P_0(t)=\tilde{K}_0(t), \tag{6.33}$$

$$P_1'(t,\theta-\tau)B'=\tilde{K}_1(t,\theta), \tag{6.34}$$

$$BP_2(t,\theta-\tau,\xi-\tau)B'=\tilde{K}_2(t,\theta,\xi),\qquad 0\leqq t\leqq T,\quad 0\leqq\theta,\quad\xi\leqq\tau. \tag{6.35}$$

The systems (6.27) *and* (6.32) *are adjoints of each other* [14] *so that*

$$\Phi(t,s)=Y'(s,t). \tag{6.36}$$

Proof. For proving (6.33)–(6.35), we simply verify that they satisfy the same equations and boundary conditions. By uniqueness of the optimal control and optimal filter, we conclude that (6.33)–(6.35) hold. Substituting these results into (6.32), we see that $Y(t, s)$ is the fundamental matrix of

$$\dot{y}(t)=-[A'-C'R^{-1}CP_0(t)]y(t)-B'y(t+\tau) + \int_{-\tau}^{0} C'R^{-1}CP_1'(t, -\theta-\tau)B'y(t-\theta)\, d\theta. \tag{6.37}$$

But (6.37) is precisely the adjoint equation [14] to (6.27), and it is well-known [14] that $\Phi(t, s) = Y'(s, t)$.

We are now ready to prove asymptotic stability of the optimal filter.

THEOREM 6.1. *Consider the system defined by* (6.19)–(6.20). *Suppose* (A, B, C) *is detectable and* (A, B, F) *is stabilizable. Then the gains of the optimal filter defined by* (5.7)–(5.11) *converge, and the steady state optimal filter is asymptotically stable.*

Proof. Proposition 6.1 shows that the dual system (A', B', C') defined by (6.21) is stabilizable and (A', B', F') is detectable. Proposition 6.2 then shows that the gains $\tilde{K}_0(t)$, $\tilde{K}_1(t, \theta)$, $\tilde{K}_2(t, \theta, \xi)$ for the dual control problem, as given by (6.28)–(6.31) converge to $\tilde{K}_0$, $\tilde{K}_1(\theta)$, $\tilde{K}_2(\theta, \xi)$ respectively as $t \to \infty$. By Lemma 6.2, we conclude that as $t \to \infty$, $P_0(t) \to P_0$, $BP_1(t, \theta) \to BP_1(\theta)$, and $BP_2(t, \theta, \xi)B' \to BP_2(\theta, \xi)B'$, where

$$AP_0+P_0A'-P_0C'R^{-1}CP_0+Q+BP_1(-\tau)+P_1'(-\tau)B'=0, \tag{6.38}$$

$$\frac{d}{d\theta}BP_1(\theta)=-BP_1(\theta)[A'-C'R^{-1}CP_0]-BP_2(\theta, -\tau)B', \tag{6.39}$$

$$\left(\frac{\partial}{\partial\theta}+\frac{\partial}{\partial\xi}\right)BP_2(\theta, \xi)B'=BP_1(\theta)C'R^{-1}CP_1'(\xi)B' \tag{6.40}$$

with

$$P_1(0)=P_0, \qquad P_2(\theta, 0)=P_1(\theta), \qquad P_0=P_0', \qquad P_2(\theta, \xi)=P_2'(\xi, \theta). \tag{6.41}$$

In view of (5.9) and (5.10), stability of the steady state filter is then governed by the stability of the equation

$$\frac{d}{dt}x(t)=[A-P_0C'R^{-1}C]x(t)+Bx(t-\tau)-B\int_{-\tau}^{0} P_1(-\theta-\tau)C'R^{-1}Cx(t+\theta)\, d\theta \tag{6.42}$$

But the adjoint to (6.42) is given by

$$\dot{y}(t)=-[A'-C'R^{-1}CP_0]y(t)-B'y(t+\tau)+\int_{0}^{\tau} C'R^{-1}CP_1'(\theta-\tau)B'y(t+\theta)\, d\theta. \tag{6.43}$$

By Lemma 6.2 again, this corresponds to the closed-loop optimal system for the dual control problem. Proposition 6.2 then shows that (6.43) is asymptotically stable. Hence, the system defined by (6.42), being the adjoint of that of (6.43), is asymptotically stable as well.

Remark 6.2. Theorem 6.1 is not the most general form of the filter stability result for delay systems. Generalization to cases where we can have delays in the observations, random initial conditions, etc., will be treated in a forthcoming paper (see also [24]).

Remark 6.3. Vinter [9] has independently obtained a similar filter stability result using infinite dimensional filtering methods quite different from ours. In addition to the conclusion given in Theorem 6.1, he also proved that $\|\tilde{T}_{t,0}\| \to 0$ as $t \to \infty$, where $\tilde{T}_{t,s}$ is the evolution operator connected with the error process for the *time-varying* filter (5.9)–(5.14). His arguments can be readily adapted to our setting to prove the same result.

7. 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 stochastic control scheme which is asymptotically stable. To that end, we define the stochastic control problem as that of minimizing the cost functional

$$J_T(u, x_0) = E\int_0^T [x'(t)Mx(t) + u'(t)Su(t)]\, dt \tag{7.1}$$

for u in some set of admissible control laws, subject to the constraint

$$dx(t) = [Ax(t) + Bx(t-\tau)]\, dt + Gu(t)\, dt + F\, dw(t), \tag{7.2}$$

$$x(\theta) = x_0(\theta), \qquad \theta \in [-\tau, 0],$$

$$dz(t) = Cx(t)\, dt + N\, dv(t). \tag{7.3}$$

We can evidently write

$$z(t) = z_0(t) + \int_0^t \int_0^s \Phi(s, \sigma)Gu(\sigma)\, d\sigma\, ds. \tag{7.4}$$

Define the set U_0 consisting of the class of processes $u(t)$ satisfying the following conditions:

(i) $u(t)$ is measurable with respect to $\sigma\{z(s), 0 \leqq s \leqq t\}$, i.e., there is a Borel measurable function Π such that $u(t) = \Pi(t; z(s), 0 \leqq s \leqq t)$.

(ii) For each $u \in U_0$, the feedback system, obtained by using $\Pi(t; z(s), 0 \leqq s \leqq t)$ for $u(t)$ in (7.2) and (7.3), has a unique solution.

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

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

We shall take U_0 to be the set of admissible controls. For a discussion on this choice, see [18], [22].

The following result has been proved by Lindquist [6].

PROPOSITION 7.1. *The problem of determining $u \in U_0$ so as to minimize* (7.1) *has the following solution*

$$u^*(t) = -S^{-1}G'K_0(t)\hat{x}(t|t) - S^{-1}G'\int_{-\tau}^0 K_1(t, \theta)\hat{x}(t+\theta|t)\, d\theta \tag{7.5}$$

where $K_0(t)$ and $K_1(t, \theta)$ are the optimal gains for the deterministic optimal control problem and are given by (6.3)–(6.6), *and $\hat{x}(s|t)$, $t-\tau \leqq s \leqq t$, is the conditional expectation of $x(s)$ given $z(\sigma)$, $0 \leqq \sigma \leqq t$.*

We now give the expression for the optimal cost, obtained in [18].

LEMMA 7.1. *Corresponding to the optimal control* (7.5), *the optimal cost associated* *with the stochastic control problem* (7.1)–(7.4) *is given by*

$$
\begin{aligned}
J^* = EV(x_0) &+ \int_0^T \operatorname{tr} FF'K_0(t)\,dt \\
&+ \int_0^T \operatorname{tr}\Big\{K_0(t)GS^{-1}G'K_0(t)P_0(t) \\
&\qquad + \int_{-\tau}^0 K_1'(t,\theta)GS^{-1}G'K_0(t)P_1'(t,\theta)\,d\theta \\
&\qquad + \int_{-\tau}^0 K_0(t)GS^{-1}G'K_1(t,\theta)P_1(t,\theta)\,d\theta \\
&\qquad + \int_{-\tau}^0\int_{-\tau}^0 K_1'(t,\theta)GS^{-1}G'K_1(t,\xi)P_2(t,\xi,\theta)\,d\theta\,d\xi\Big\}\,dt
\end{aligned}
\tag{7.6}
$$

where

$$
\begin{aligned}
V(x_t) = x'(t)K_0(t)x(t) &+ \int_{-\tau}^0 x'(t)K_1(t,\theta)x(t+\theta)\,d\theta \\
&+ \int_{-\tau}^0 x'(t+\theta)K_1'(t,\theta)x(t)\,d\theta + \int_{-\tau}^0\int_{-\tau}^0 x'(t+\theta)K_2(t,\theta,\xi)x(t+\xi)\,d\theta\,d\xi.
\end{aligned}
\tag{7.7}
$$

Proof. See Appendix B.

We turn our attention now to the stochastic control system defined by using the steady state version of (7.5). The behavior of the closed-loop system under this law is summarized in the following theorem.

THEOREM 7.1. *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*

$$
u(t) = -S^{-1}G'K_0\hat{x}(t|t) - S^{-1}G'\int_{-\tau}^0 K_1(\theta)\hat{x}(t+\theta|t)\,d\theta
\tag{7.8}
$$

where $\hat{x}(t+\theta|t)$, $-\tau \leqq \theta \leqq 0$, is generated by the steady state filter of Theorem 6.1, *and K_0, $K_1(\theta)$ are given by the deterministic stationary control law of Proposition* 6.2, *gives rise to an asymptotically stable closed-loop system. Furthermore, the cost "rate"*

$$
J_r = \lim_{T\to\infty} \frac{1}{T} E\Big\{\int_0^T [x'(t)Mx(t) + u'(t)Su(t)]\,dt\Big\}
\tag{7.9}
$$

associated with the above law is given by

$$
\begin{aligned}
J_r = \operatorname{tr} FF'K_0 + \operatorname{tr}\Big\{K_0GS^{-1}G'K_0P_0 &\\
&+ \int_{-\tau}^0 K_1'(\theta)GS^{-1}G'K_0P_1'(\theta)\,d\theta \\
&+ \int_{-\tau}^0 K_0GS^{-1}G'K_1(\theta)P_1(\theta)\,d\theta \\
&+ \int_{-\tau}^0\int_{-\tau}^0 K_1'(\theta)GS^{-1}G'K_1(\xi)P_2(\xi,\theta)\,d\theta\,d\xi\Big\}.
\end{aligned}
\tag{7.10}
$$

Proof. Stabilizability of (A, B, G) and detectability of (A, B, H) ensure that K_0, $K_1(\theta)$, $K_2(\theta, \xi)$ are well defined and that the solutions of the system

$$\dot{x}(t) = (A - GS^{-1}G'K_0)x(t) + Bx(t-\tau) - \int_{-\tau}^{0} GS^{-1}G'K_1(\theta)x(t+\theta)\,d\theta \tag{7.11}$$

are asymptotically stable (see Proposition 6.2). Detectability of (A, B, C) and stabilizability of (A, B, F) guarantee that the steady state filter is well defined and asymptotically stable (see Theorem 6.1). The closed-loop system is defined by the coupled set of equations

$$\begin{aligned} dx(t) = {} & [Ax(t) + Bx(t-\tau)]\,dt - GS^{-1}G'K_0\hat{x}(t|t)\,dt \\ & - \int_{-\tau}^{0} GS^{-1}G'K_1(\theta)\hat{x}(t+\theta|t)\,d\theta\,dt + F\,dw(t), \end{aligned} \tag{7.12}$$

$$\begin{aligned} d\hat{x}(t|t) = {} & A\hat{x}(t|t)\,dt + B\hat{x}(t-\tau|t-\tau)\,dt \\ & + P_0C'R^{-1}[dz(t) - C\hat{x}(t|t)\,dt] \\ & + B\int_{t-\tau}^{t} P_1(t-s-\tau)C'R^{-1}[dz(s) - C\hat{x}(s/s)\,ds]\,dt. \end{aligned} \tag{7.13}$$

Let the estimation error $e(t+\theta/t)$, $-\tau \leqq \theta \leqq 0$, be defined as $e(t+\theta/t) = x(t+\theta) - \hat{x}(t+\theta/t)$. We then get

$$\begin{aligned} dx(t) = {} & (A - GS^{-1}G'K_0)x(t)\,dt + Bx(t-\tau)\,dt \\ & - \int_{-\tau}^{0} GS^{-1}G'K_1(\theta)x(t+\theta)\,d\theta\,dt + F\,dw(t) \\ & + GS^{-1}G'K_0e(t|t)\,dt + \int_{-\tau}^{0} GS^{-1}G'K_1(\theta)e(t+\theta|t)\,d\theta\,dt \end{aligned} \tag{7.14}$$

and

$$\begin{aligned} de(t|t) = {} & (A - P_0C'R^{-1}C)e(t|t)\,dt + Be(t-\tau|t-\tau)\,dt \\ & + \int_{-\tau}^{0} BP_1(-\theta-\tau)C'R^{-1}Ce(t+\theta|t+\theta)\,d\theta\,dt \\ & + F\,dw(t) - P_0C'R^{-1}\,dv(t) - B\int_{t-\tau}^{t} P_1(t-s-\tau)C'R^{-1}\,dv(s)\,dt. \end{aligned} \tag{7.15}$$

Since (7.15) is decoupled from (7.14), the stability properties of the closed-loop system are precisely those of (7.11) and the steady state optimal filter. Since both of these are asymptotically stable as a consequence of our assumptions, the closed-loop stochastic control system is asymptotically stable as well. The expression for J_r follows readily from Lemma 7.1.

8. Concluding remarks. We have treated the problem of filtering and control for stochastic delay systems. The general filtering problem is studied for both linear and nonlinear stochastic delay systems. A representation theorem for conditional moment functionals is given, which forms the basis for derivations of stochastic differential equations describing the optimal linear or nonlinear filter. For linear systems with Gaussian initial conditions and noises, the optimal filter is completely specified by the equations derived for the conditional mean and covariance functions. The linear time-invariant case with delays in the system dynamics is investigated in detail, with particular emphasis on the stability of the optimal filter and stochastic control system. These results, together with those on deterministic optimal control or linear delay systems with quadratic cost, give a rather complete linear-quadratic-Gaussian theory for this class of delay systems. In a forthcoming paper, we will extend this theory to systems with delays in the control and delays in the observations.

Appendix A.

Proof of Theorem 5.1. It is only necessary to derive the equations for $P(t, \theta, \xi)$. For systems with point delays only, Bagchi [8] derived the equations for $P(t, \theta, \xi)$ using properties of the innovations and martingale theory. Although our approach is different, we shall use some of his results to simplify our derivations (a more complicated proof was given in [18]).

It is easy to see that

$$P(t, \theta, \xi) = E[x(t+\theta)x'(t+\xi)] - E[\hat{x}(t+\theta|t)\hat{x}(t+\xi|t)']. \tag{A.1}$$

Using (5.5), we obtain

$$\begin{aligned} &E[\hat{x}(t+\theta|t)\hat{x}(t+\xi|t)'] \\ &\quad = E\Big\{\Big[\hat{x}(t+\theta|t+\theta) + \int_{t+\theta}^{t}\int_{-\tau}^{0} P(s, t+\theta-s, \beta)d_\beta C(s, \beta)'R^{-1}(s)\,d\nu(s)\Big] \\ &\qquad \cdot \Big[\hat{x}(t+\xi|t+\xi) + \int_{t+\xi}^{t}\int_{-\tau}^{0} P(\sigma, t+\xi-\sigma, \alpha)d_\alpha C(\sigma, \alpha)'R^{-1}(\sigma)\,d\nu(\sigma)\Big]\Big\}. \end{aligned} \tag{A.2}$$

For any ε such that $-\tau \leqq \theta+\varepsilon \leqq 0$, $-\tau \leqq \xi+\varepsilon \leqq 0$, we get, using (A.1) and (A.2), that

$$\begin{aligned} &P(t, \theta, \xi) - P(t-\varepsilon, \theta+\varepsilon, \xi+\varepsilon) \\ &\quad = -E\Big\{\Big[\int_{t-\varepsilon}^{t}\int_{-\tau}^{0} P(s, t+\theta-s, \beta)d_\beta C(s, \beta)'R^{-1}(s)\,d\nu(s)\Big]\hat{x}(t+\xi|t+\xi)'\Big\} \\ &\qquad -E\Big\{\hat{x}(t+\theta|t+\theta)\Big[\int_{t-\xi}^{t}\int_{-\tau}^{0} P(\alpha, t+\xi-\sigma, \alpha)d_\alpha C(\sigma, \alpha)'R^{-1}(\sigma)\,d\nu(\sigma)\Big]'\Big\} \\ &\qquad -E\Big\{\Big[\int_{t+\theta}^{t}\int_{-\tau}^{0} P(s, t+\theta-s, \beta)d_\beta C(s, \beta)'R^{-1}(s)\,d\nu(s)\Big] \\ &\qquad\quad \cdot \Big[\int_{t+\xi}^{t}\int_{-\tau}^{0} P(\sigma, t+\xi-\sigma, \alpha)d_\alpha C(\sigma, \alpha)'R^{-1}(\sigma)\,d\nu(\sigma)\Big]'\Big\} \\ &\qquad +E\Big\{\Big[\int_{t+\theta}^{t-\varepsilon}\int_{-\tau}^{0} P(s, t+\theta-s, \beta)d_\beta C(s, \beta)'R^{-1}(s)\,d\nu(s)\Big] \\ &\qquad\quad \cdot \Big[\int_{t+\xi}^{t-\varepsilon}\int_{-\tau}^{0} P(\sigma, t+\xi-\sigma, \alpha)d_\alpha C(\sigma, \alpha)'R^{-1}(\sigma)\,d\nu(\sigma)\Big]'\Big\}. \end{aligned} \tag{A.3}$$

Using the fact that $E[\nu(t)-\nu(s)|z^s]=0$ as in [8], we see that the first two terms in (A.3) vanish. By the same argument, the last two terms can be easily simplified to yield

$$
\begin{aligned}
&P(t,\theta,\xi)-P(t-\xi,\theta+\varepsilon,\xi+\varepsilon)\\
&\quad=-E\Big\{\Big[\int_{t-\varepsilon}^{t}\int_{-\tau}^{0}P(s,t+\theta-s,\beta)d_\beta C(s,\beta)'R^{-1}(s)\,d\nu(s)\Big]\\
&\qquad\cdot\Big[\int_{t-\varepsilon}^{t}\int_{-\tau}^{0}P(\sigma,t+\xi-\sigma,\alpha)d_\alpha C(\sigma,\alpha)'R^{-1}(\sigma)\,d\nu(\sigma)\Big]'\Big\}\\
&\quad=-\int_{t-\varepsilon}^{t}\int_{-\tau}^{0}P(s,t+\theta-s,\beta)d_\beta C(s,\beta)'R^{-1}(s)\int_{-\tau}^{0}d_\alpha C(s,\alpha)P(s,\alpha,t+\xi-s)\,ds.
\end{aligned}
\tag{A.4}
$$

Since $P(t, \theta, \xi)$ is clearly continuous in (t, θ, ξ), we may divide (A.4) throughout by ε and let ε go to 0. This gives (5.8). The same arguments apply to the derivations of (5.6) and (5.7). Finally, the initial conditions follow immediately from the properties of conditional expectations.

Appendix B.

Proof of Lemma 7.1. We apply the Itô differential rule to the function $V(x_t)$ defined in (7.7). We calculate the first and second terms to illustrate the computations involved:

$$
\begin{aligned}
d[x'(t)K_0(t)x(t)]&=[dx'(t)]K_0(t)x(t)\\
&\quad+x'(t)[dK_0(t)]x(t)+x'(t)K_0(t)[dx(t)]+\operatorname{tr}FF'K_0(t)\,dt\\
&=x'(t-\tau)B'K_0(t)x(t)\,dt+u'(t)G'K_0(t)x(t)\,dt\\
&\quad+dw'(t)F'K_0(t)x(t)\,dt+x'(t)K_0(t)Bx(t-\tau)\,dt\\
&\quad+x'(t)K_0(t)Gu(t)\,dt+x'(t)K_0(t)F\,dw(t)-x'(t)Mx(t)\,dt\\
&\quad+x'(t)K_0(t)GR^{-1}G'K_0(t)x(t)\,dt-x'(t)K_1'(t,0)x(t)\,dt\\
&\quad-x'(t)K_1(t,0)x(t)\,dt+\operatorname{tr}FF'K_0(t)\,dt,
\end{aligned}
$$

$$
\begin{aligned}
&d_t\Big[\int_{-\tau}^{0}x'(t)K_1(t,\theta)x(t+\theta)\,d\theta\Big]\\
&\quad=d_t\Big[x'(t)\int_{t-\tau}^{t}K_1(t,\sigma-t)x(\sigma)\,d\sigma\Big]\\
&\quad=\{[x'(t)A'+x'(t-\tau)B'+u'(t)G']\,dt+dw'(t)F'\}\int_{-\tau}^{0}K_1(t,\theta)x(t+\theta)\,d\theta\\
&\qquad+x'(t)K_1(t,0)x(t)\,dt-x'(t)K_1(t,-\tau)x(t-\tau)\,dt\\
&\qquad+x'(t)\int_{t-\tau}^{t}d_tK_1(t,\sigma-t)x(\sigma)\,d\sigma\,dt\\
&\quad=[x'(t)A'+x'(t-\tau)B'+u'(t)G']\int_{-\tau}^{0}K_1(t,\theta)x(t+\theta)\,d\theta\,dt\\
&\qquad+dw'(t)F'\int_{-\tau}^{0}K_1(t,\theta)x(t+\theta)\,d\theta+x'(t)K_1(t,0)x(t)\,dt\\
&\qquad-x'(t)K_1(t,-\tau)x(t-\tau)\,dt+x'(t)\int_{-\tau}^{0}\Big[\Big(\frac{\partial}{\partial t}-\frac{\partial}{\partial\theta}\Big)K_1(t,\theta)\Big]x(t+\theta)\,d\theta\,dt.
\end{aligned}
$$

Similar calculations for the last two terms on the right hand side of (7.7) yield the following expression

$$\begin{aligned} dV(x_t) = {} & V_1(t)\,dt + dw'(t)F'K_0(t)x(t)\,dt + x'(t)K_0(t)F\,dw(t) \\ & + \operatorname{tr} FF'K_0(t)\,dt + dw'(t)F'\int_{-\tau}^{0} K_1(t,\theta)x(t+\theta)\,d\theta \\ & + \int_{-\tau}^{0} x'(t+\theta)K_1'(t,\theta)\,d\theta\,F\,dw(t) - x'(t)Mx(t) - u'(t)Su(t)\,dt \end{aligned} \tag{B.1}$$

where

$$\begin{aligned} V_1(t) = {} & \left[u(t) + S^{-1}G'K_0(t)x(t) + \int_{-\tau}^{0} S^{-1}G'K_1(t,\theta)x(t+\theta)\,d\theta\right]' \\ & \cdot S\left[u(t) + S^{-1}G'K_0(t)x(t) + \int_{-\tau}^{0} S^{-1}G'K_1(t,\xi)x(t+\xi)\,d\xi\right]. \end{aligned} \tag{B.2}$$

Using the boundary conditions at T for $K_0(t)$, $K_1(t, \theta)$ and $K_2(t, \theta, \xi)$, we see that $V(x_T) = 0$. Therefore, integrating (B.1) from 0 to T and taking expectations, we get

$$E\int_0^T [x'(t)Mx(t) + u'(t)Su(t)]\,dt = EV(x_0) + E\int_0^T V_1(t)\,dt + \int_0^T \operatorname{tr} FF'K_0(t)\,dt. \tag{B.3}$$

Now

$$E\int_0^T V_1(t)\,dt = \int_0^T EV_1(t)\,dt = \int_0^T E\{E[V_1(t)/z^t]\}\,dt$$

by the use of Fubini's theorem and properties of conditional expectations. Substituting the control law in (7.5) into (B.2), we get that

$$\begin{aligned} E[V_1(t)/z^t] = \operatorname{tr}\Big\{ & K_0(t)G_1R^{-1}G'K_0(t)P_0(t) \\ & + \int_{-\tau}^{0} K_1'(t,\theta)GR^{-1}G'K_0(t)P_1'(t,\theta)\,d\theta \\ & + \int_{-\tau}^{0} K_0(t)GR^{-1}G'K_1(t,\theta)P_1(t,\theta)\,d\theta \\ & + \int_{-\tau}^{0}\int_{-\tau}^{0} K_1'(t,\theta)GR^{-1}G'K_1(t,\xi)P_2(t,\theta,\xi)\,d\theta\,d\xi\Big\}. \end{aligned} \tag{B.4}$$

$E\{V_1(t)/z^t\}$ is now seen to be a deterministic function and hence equal to $EV_1(t)$. Substituting (B.4) into (B.3) yields the conclusion of the lemma.

Acknowledgments. We would like to thank the referees for drawing our attention to the work of Kunita [16] and Bagchi [8], and for helpful suggestions which led to general improvements in the presentation. The above references, in particular, enabled us to reduce the complicated proofs of Theorem 3.1 and 5.1 given in [18] to the simple forms here in this paper.

REFERENCES

[1] H. J. KUSHNER AND D. I. BARNEA, *On the control of a linear functional differential equation with quadratic cost*, this Journal, 8 (1970), pp. 257–272.

[2] Y. ALEKAL, P. BRUNOVSKY, D. H. CHYUNG AND E. B. LEE, *The quadratic problem for systems with time delays*, IEEE Trans. Automatic Control, AC-16 (1971), pp. 673–688.

[3] A. MANITIUS, *Optimal control of time-lag systems with quadratic performance indexes*, IV Congress of International Federation of Automatic Control, Warsaw, 1969, paper 13.2.

[4] M. C. DELFOUR AND S. K. MITTER, *Controllability, observability and optimal feedback control of affine hereditary differential systems*, this Journal, 10 (1972), pp. 298–328.

[5] H. KWAKERNAAK, *Optimal filtering in linear systems with time delays*, IEEE Trans. Automatic Control, AC-12 (1967), pp. 169–173.

[6] A. LINDQUIST, *Optimal control of linear stochastic systems with applications to time lag systems*, Information Sci., 5 (1973), pp. 81–126.

[7] S. K. MITTER AND R. B. VINTER, *Filtering for linear stochastic hereditary differential systems*, Intern. Symp. Control Theory, Numerical Methods, and Computer Systems Modelling, Institut de Recherche d'Informatique et d'Automatique, Rocquencourt, France, June 1974.

[8] A. BAGCHI, *A martingale approach to state estimation in delay-differential systems*, J. Math. Anal. Appl., 56 (1976), pp. 195–210.

[9] R. B. VINTER, *Filter stability for stochastic evolution equations*, this Journal, 15 (1977), pp. 465–485.

[10] K. ITO AND M. NISIO, *On stationary solutions of a stochastic differential equation*, Kyoto J. Math., 4 (1964), pp. 1–75.

[11] W. H. FLEMING AND M. NISIO, *On the existence of optimal stochastic controls*, J. Math. Mech., 15 (1966), pp. 777–794.

[12] H. J. KUSHNER, *On the stability of processes defined by stochastic differential-difference equations*, J. Differential Equations, 4 (1968), pp. 424–443.

[13] A. LINDQUIST, *A theorem on duality between estimation and control for linear stochastic systems with time delay*, J. Math. Anal. Appl., 37 (1972), pp. 516–536.

[14] J. K. HALE, *Functional Differential Equations*, Springer-Verlag, New York, 1971.

[15] M. FUJISAKI, G. KALLIANPUR AND H. KUNITA, *Stochastic differential equations for the nonlinear filtering problem*, Osaka J. Math., 9 (1972), pp. 19–40.

[16] H. KUNITA, *Asymptotic behavior of the nonlinear filtering errors of Markov processes*, J. Multivariate Anal., 1 (1971), pp. 365–393.

[17] I. I. GIKHMAN AND A. V. SKOROKHOD, *Introduction to the Theory of Random Processes*, W. B. Saunders, London, 1969.

[18] R. H. KWONG, *Structural properties and estimation of delay systems*, Electronic Systems Lab. rep. ESL-R-614, Massachusetts Inst. of Tech., Cambridge, MA, Sept. 1975.

[19] R. DATKO, *Unconstrained control problems with quadratic cost*, this Journal, 11 (1973), pp. 32–52.

[20] M. C. DELFOUR, C. MCCALLA AND S. K. MITTER, *Stability and the infinite-time quadratic cost problem for linear hereditary differential systems*, this Journal, 13 (1975), pp. 48–88.

[21] J. ZABCZYK, *Remarks on the algebraic Riccati equation in Hilbert space*, Appl. Math. and Optimization, 2 (1976), pp. 251–258.

[22] A. LINDQUIST, *On feedback control of linear stochastic systems*, this Journal, 11 (1973), pp. 323–343.

[23] H. L. VAN TREES, *Detection, Estimation, and Modulation Theory, Part III*, John Wiley, New York, 1971.

[24] R. H. KWONG, *The linear quadratic Gaussian problem for systems with delays in the state, control and observations*, Proc. 14th Allerton Conf. Circuit and System Theory, Univ. of Illinois, Sept. 29–Oct. 1, 1976, pp. 545–549.

[25] A. MANITIUS AND R. TRIGGIANI, *Function space controllability of linear retarded systems: a derivation from abstract operator conditions*, Centre de Recherches Mathématiques rep. CRM-605, Université de Montréal, Montréal, Québec, Canada, March 1976.

[26] A. MANITIUS, *Controllability, observability and stabilizability of retarded systems*, Proc. IEEE Conf. Decision and Control (Clearwater, FL, Dec. 1976), IEEE, New York, 1977, pp. 752–758.