INFORMATION AND CONTROL 34, 324–338 (1977)

Fixed-Point Smoothing in Hilbert Spaces

S. OMATU, Y. TOMITA, AND T. SOEDA

Department of Information Science and Systems Engineering, Faculty of Engineering, University of Tokushima, Minamijosanjima 2-1, Tokushima, Japan

The fixed-point smoothing estimator and smoothing error covariance operator equations are derived for infinite-dimensional linear systems using both the Wiener–Hoph theory in Hilbert spaces developed by Falb and the abstract evolution theory. Since it is clear that the prediction problems can be solved by the same approach, the present results in conjunction with the work of Falb on filtering give a complete treatment of the infinite-dimensional linear estimation problem from the viewpoint of Wiener–Hoph theory. Finally, based on the optimal smoothing estimator in Hilbert space, the fixed-point smoothing estimator is derived for a linear distributed parameter system of parabolic type.

1. INTRODUCTION

In recent years a great deal of research has been devoted to the study of optimal filtering problems for a stochastic distributed parameter system. The techniques used for generating the filtering estimate include orthogonal projection theories (Curtain, 1975a, b; Falb, 1967; Sakawa, 1972; Tzafestas and Nightingale, 1968), maximum-likelihood approaches combined with the variational inequality or dynamic programming methods (Bensoussan, 1971, 1975; Meditch, 1971; Tzafestas and Nightingale, 1969), and the Bayesian approach (Kushner, 1970; Tzafestas, 1972).

On the other hand, the smoothing problems for distributed parameter systems have been of significant importance in various engineering fields and several techniques have been discussed up to now. Tzafestas and Nightingale (1968) and Meditch (1971) used the maximum-likelihood approach to derive an algorithm of fixed-interval smoothing estimates and Tzafestas (1972) considered the several kinds of smoothing problems by applying Kalman's limiting procedure. However, its validity has not been proved for infinite-dimensional spaces. Curtain (1975b) derived the fixed-interval smoothing estimator using both the innovation process and the Wiener–Hopf theory.

In this paper, we derive the fundamental equations of the fixed-point smoothing estimator for infinite-dimensional linear systems based on both the Wiener–Hopf theory in Hilbert space developed by Falb (1967) and the abstract evolution theory (Kato and Tanabe, 1962). As the theory relies heavily on these results,

ISSN 0019-9958

we give a symmary of them and then develop the fixed-point smoothing problems in Hilbert spaces. Finally, we briefly discuss the application of the infinite-dimensional theory to the smoothing problems for a parabolic type linear distributed parameter system with pointwise observation.

2. Problem Statement

Let $\mathscr{H}$, $\mathscr{K}$ be real Hilbert spaces, $(\Omega, \mathscr{S}, \mu)$ be an underlying probability space, and $\mathscr{L}(\mathscr{H}, \mathscr{K})$ be a family of bounded linear operators from $\mathscr{H}$ into $\mathscr{K}$. Consider the following infinite-dimensional linear systems:

$$
\begin{aligned}
dU(t, \omega) &= A(t)\, U(t, \omega)\, dt + B(t)\, dW(t, \omega), \qquad (2.1)\\
U(0, \omega) &= U_0(\omega), \qquad t \in T = [0, t_f],\ \omega \in \Omega,\\
dZ(t, \omega) &= C(t)\, U(t, \omega)\, dt + r(t)\, dV(t, \omega),\\
Z(0, \omega) &= 0, \qquad t \in T,\ \omega \in \Omega, \qquad (2.2)
\end{aligned}
$$

where $B(\cdot) \in L_\infty(T; \mathscr{L}(\mathscr{H}, \mathscr{H}))$, $C(\cdot) \in L_\infty(T; \mathscr{L}(\mathscr{H}, \mathscr{K}))$, $r(\cdot)$, $r^{-1}(\cdot) \in L_\infty(T; \mathscr{L}(\mathscr{K}, \mathscr{K}))$, and U_0 is an $\mathscr{H}$-valued random variable independent of $W(t, \omega)$ and $V(t, \omega)$ and has zero expectation and covariance operator P_0. $W(t, \omega)$ and $V(t, \omega)$ are independent Wiener processes on $\mathscr{H}$ and $\mathscr{K}$ with covariance operators $\mathscr{W}(t)$ and $\mathscr{V}(t)$, respectively, i.e.,

$$
\begin{aligned}
\operatorname{Cov}[W(t, \omega) - W(s, \omega), W(t, \omega) - W(s, \omega)] &= \int_s^t \mathscr{W}(\tau)\, d\tau,\\
\operatorname{Cov}[V(t, \omega) - V(s, \omega), V(t, \omega) - V(s, \omega)] &= \int_s^t \mathscr{V}(\tau)\, d\tau,
\end{aligned} \qquad (2.3)
$$

where $\mathscr{W}(t)$ and $\mathscr{V}(t)$ are nuclear, positive symmetric operators for almost all $t \in T$, $\int_T \operatorname{tr} \mathscr{W}(\tau)\, d\tau < \infty$ and the covariance operator $\operatorname{Cov}[\cdot, \cdot]$ is defined by

$$\operatorname{Cov}[x, y] = E\{x \circ y\} - E\{x\} \circ E\{y\},$$

where $E\{\cdot\}$ and $\operatorname{tr}[\cdot]$ denote the expectation and the trace of the operator, respectively, and $\circ$ is given by the following relation:

$$(x \circ y)z = x\langle y, z\rangle \qquad \text{for fixed } x \in \mathscr{K},\ y \in \mathscr{H}, \text{ and } \forall z \in \mathscr{H}. \qquad (2.4)$$

Here $\langle \cdot, \cdot\rangle$ denotes the inner product on $\mathscr{H}$ and $A(t)$ is a linear operator in $\mathscr{H}$, satisfying the following:

(i) the domain $\mathscr{D}(A)$ of $A(t)$ is dense in $\mathscr{H}$ and independent of t, and $A(t)$ is a closed operator;

(ii) $(\lambda\mathscr{I} - A(t))^{-1}$ exists for Re $\lambda \geqslant 0$ and $\|(\lambda\mathscr{I} - A(t))^{-1}\| \leqslant c/(1 + |\lambda|)$, where c is constant and $\mathscr{I}$ denotes the identity operator;

(iii) $\|(A(t) - A(\tau))\, A^{-1}(\tau)\| \leqslant C\, |t - \tau|^{\alpha}$ for all $t, \tau \geqslant 0$, where C and α are positive constants.

Then note that there exists the evolution operator $\mathscr{U}(t, s)$ with the following properties (Kato and Tanabe, 1962):

(1) $\mathscr{U}(t, s) : T \times T \to \mathscr{L}(\mathscr{H}, \mathscr{H})$ and is strongly continuous in s and t for $0 \leqslant s \leqslant t \leqslant t_f$;

(2) $\mathscr{U}(t, s) = \mathscr{U}(t, \tau)\, \mathscr{U}(\tau, s)$ if $0 \leqslant s \leqslant \tau \leqslant t \leqslant t_f$, and $\mathscr{U}(s, s) = \mathscr{I}$;

(3) $\mathscr{U}(t, s)$ is strongly continuously differentiable in t for $t > s$ and

$$\partial\mathscr{U}(t, s)/\partial t = A(t)\, \mathscr{U}(t, s), \tag{2.5}$$

where

$$\| A(t)\, \mathscr{U}(t, s)\| \leqslant c_1/| t - s | \qquad \text{for} \quad 0 \leqslant s < t \leqslant t_f \text{ and some constant } c_1 .$$

We assume from now on that the random variables which we consider have expected values and we often delete the explicit ω-dependence of these random variables.

The smoothing problem we propose is to find the best estimate of the state $U(t_1)$ based on observations $Z(s)$, $0 \leqslant s \leqslant t$, $t_1 < t$, which has the form

$$\hat{U}(t_1 \mid t) = \int_0^t K(t, s)\, dZ(s), \tag{2.6}$$

where $K(t, \cdot) \in L_2(T; \mathscr{L}(\mathscr{K}, \mathscr{H}))$ for almost all t and which minimizes

$$E\{\langle h, \tilde{U}(t_1 \mid t)\rangle^2\},\ \tilde{U}(t_1 \mid t) \triangleq U(t_1) - \hat{U}(t_1 \mid t), \qquad \forall h \in \mathscr{H}. \tag{2.7}$$

Note that setting $t_1 = t$ implies that (2.6) reduces to the filtering problem which has been discussed by Falb (1967), Curtain (1975a, b), and Bensoussan (1971).

3. Preliminary Results

Let us recapitulate the results for the filtering problems and the properties of stochastic integral of Ito type required in this paper which have been proved by Falb (1967).

Lemma 3.1. *If* $\mathrm{Cov}[W(t), U_0] = 0$, *then for* $\sigma < t$,

$$\frac{d}{dt}\, \mathrm{Cov}[U(t), \Delta Z(\sigma)] = \mathrm{Cov}[A(t)\, U(t), \Delta Z(\sigma)], \tag{3.1}$$

where

$$\Delta Z(\sigma) = Z(\sigma) - Z(0).$$

LEMMA 3.2. *If $\partial M(t, s)/\partial t$ exists and is regulated, then for $\sigma < t$*

$$\frac{d}{dt} \operatorname{Cov} \left[\int_0^t M(t, s)\, dZ(s), \Delta Z(\sigma)\right]$$

$$= \operatorname{Cov} \left[\int_0^t \frac{\partial M(t, s)}{\partial t}\, dZ(s) + M(t, t)\, C(t)\, U(t), \Delta Z(\sigma)\right], \qquad (3.2)$$

where $M(t, \cdot) \in L_2(T; \mathscr{L}(\mathscr{K}, \mathscr{H}))$ for almost all t.

LEMMA 3.3. *Let $\Phi(t)$ and $\Psi(s)$ be elements of $L_2(T; \mathscr{L}(\mathscr{H}, \mathscr{H}))$. Then*

$$\operatorname{Cov} \left[\int_0^t \Phi(s)\, dW(s), \int_0^t \Psi(s)\, dW(s)\right] = \int_0^t \Phi(s)\, \mathscr{W}(s)\, \Psi^*(s)\, ds. \qquad (3.3)$$

Furthermore, if $\Psi(s)$ is an element of $L_2(T; \mathscr{L}(\mathscr{K}, \mathscr{K}))$, then

$$\operatorname{Cov} \left[\int_0^t \Phi(s)\, dW(s), \int_0^t \psi(s)\, dV(s)\right] = 0. \qquad (3.4)$$

LEMMA 3.4. *Suppose that there is a solution of the filtering problem of the form*

$$\hat{U}(t \mid t) = \int_0^t L(t, s)\, dZ(s) \qquad (3.5)$$

with $\partial L(t, s)/\partial t$ regulated, where $L(t, \cdot) \in L_2(T; \mathscr{L}(\mathscr{K}, \mathscr{H}))$ for almost all t. Then

$$\partial L(t, s)/\partial t = A(t)\, L(t, s) - L(t, t)\, C(t)\, L(t, s) \qquad \text{for} \quad 0 \leqslant s \leqslant t. \quad (3.6)$$

LEMMA 3.5. *Under the hypotheses of Lemma 3.4, $\hat{U}(t \mid t)$ satisfies the linear stochastic differential equation*

$$d\hat{U}(t \mid t) = (A(t) - K(t)\, C(t))\, \hat{U}(t \mid t)\, dt + K(t)\, C(t)\, U(t)\, dt + K(t)\, r(t)\, dV(t), \qquad (3.7)$$

where

$$K(t) = L(t, t) = P(t \mid t)\, C^*(t)\, R^{-1}(t),$$

$$P(t \mid t) = \Big[(\mathscr{U}(t, 0) - \phi(t, 0))P_0 + \int_0^t (\mathscr{U}(t, s) - \phi(t, s))\, Q(s)\, \mathscr{U}^*(0, s)\, ds\Big]\, \mathscr{U}^*(t, 0), \qquad (3.8)$$

$$\phi(t, s) \triangleq \int_s^t L(t, \tau)\, C(\tau)\, \mathscr{U}(\tau, s)\, d\tau, \tag{3.9}$$

$$Q(s) \triangleq B(s)\, \mathscr{W}(s)\, B^*(s),\ R(s) \triangleq r(s)\, \mathscr{V}(s)\, r^*(s).$$

*Here * denotes the adjoint of the operator.*

4. The Derivation of the Smoothing Estimator

Let us derive the optimal smoothing estimator using the lemmas in the preceding section. The following theorem, which involves the Wiener–Hoph equation, gives the basic necessary and sufficient condition for $\hat{U}(t_1 \mid t)$ to be the optimal solution of the problem.

THEOREM 4.1. *$\hat{U}(t_1 \mid t) = \int_0^t K(t, \tau)\, dZ(\tau)$ is a solution of the smoothing problem if and only if $E\{\tilde{U}(t_1 \mid t) \circ (Z(\sigma) - Z(\tau))\} = 0$ for all σ, τ such that $0 \leqslant \tau < \sigma < t$, or equivalently, if and only if*

$$\operatorname{Cov}[U(t_1), (Z(\sigma) - Z(\tau))] = \operatorname{Cov}\left[\int_0^t K(t, \tau)\, dZ(\tau), (Z(\sigma) - Z(\tau))\right]. \tag{4.1}$$

LEMMA 4.2. *Let $K(t, s)$ be the optimal smoothing kernel and $K(t, s) + N(t, s)$ satisfies the Wiener–Hopf equation. Then $N(t, s) = 0$ for any $0 \leqslant s \leqslant t$.*

Theorem 4.1 and Lemma 4.2 were proved by Falb (1967) and Curtain (1975) using the method of orthogonal projection theory in Hilbert spaces. Then the next theorem follows.

THEOREM 4.3. *Suppose that there is a solution of the smoothing problem of the form*

$$\hat{U}(t_1 \mid t) = \int_0^t K(t, \tau)\, dZ(\tau)$$

with $\partial K(t, \tau)/\partial t$ regulated. Then

$$\partial K(t, \tau)/\partial t = -K(t, t)\, C(t)\, L(t, \tau), \tag{4.2}$$

where $L(t, \tau)$ has been given by Lemma 3.4.

Proof. Since $\hat{U}(t_1 \mid t)$ is a solution of the smoothing problem, we have by virtue of Theorem 4.1,

$$\frac{d}{dt} \operatorname{Cov}[U(t_1), \Delta Z(\sigma)] = \frac{d}{dt} \operatorname{Cov}\left[\int_0^t K(t, \tau)\, dZ(\tau), \Delta Z(\sigma)\right].$$

It follows from Lemma 3.2 and Theorem 4.1 that

$$\operatorname{Cov}\left[\int_0^t \frac{\partial K(t,\tau)}{\partial t}\, dZ(\tau) + K(t,t)\, C(t)\, U(t),\, \Delta Z(\sigma)\right]$$

$$= \operatorname{Cov}\left[\int_0^t \left\{\frac{\partial K(t,\tau)}{\partial t} + K(t,t)\, C(t)\, L(t,\tau)\right\} dZ(\tau),\, \Delta Z(\sigma)\right] = 0.$$

Hence, from Lemma 4.2 we have

$$\partial K(t,\tau)/\partial t + K(t,t)\, C(t)\, L(t,\tau) = 0.$$

Thus, the theorem is established. Q.E.D.

This leads us to the following theorem.

THEOREM 4.4. *Under the hypotheses of Theorem* 4.3, $\hat{U}(t_1 \mid t)$ *satisfies the following linear stochastic differential equation*;

$$d\hat{U}(t_1 \mid t) = K(t,t)[dZ(t) - C(t)\, \hat{U}(t \mid t)\, dt]. \tag{4.3}$$

Proof. Applying the standard Fubini theorem (Falb, 1967) and Theorem 4.3 yields

$$-\int_0^t K(s,s)\, C(s)\, \hat{U}(s \mid s)\, ds = -\int_0^t K(s,s)\, C(s) \left[\int_0^s L(s,\tau)\, dZ(\tau)\right] ds$$

$$= \int_0^t \left[\int_\tau^t \frac{\partial K(s,\tau)}{\partial s}\, ds\right] dZ(\tau) = \int_0^t [K(t,\tau) - K(\tau,\tau)]\, dZ(\tau)$$

$$= \hat{U}(t_1 \mid t) - \int_0^t K(\tau,\tau)\, dZ(\tau).$$

Hence, the theorem follows from this relation by direct differentiation. Q.E.D.

COROLLARY 4.5. *Under the hypotheses of Theorem* 4.3, $\tilde{U}(t_1 \mid t)$ *satisfies the linear stochastic differential equation*

$$d\tilde{U}(t_1 \mid t) = -K(t,t)[dZ(t) - C(t)\, \hat{U}(t \mid t)\, dt]. \tag{4.4}$$

Then we have the following theorem.

THEOREM 4.6. *Suppose that the conditions of Theorem* 4.3 *are satisfied. Then we have*

$$K(t,t) = B(t_1 \mid t)\, C^*(t)\, R^{-1}(t), \qquad t_1 < t,$$
$$R(t) = r(t)\, \mathscr{V}(t)\, r^*(t),$$

where $B(t_1 \mid t)$ *is a solution of the following equation with the initial condition* $B(t_1 \mid t_1) = P(t_1 \mid t_1) = \mathrm{Cov}[\tilde{U}(t_1 \mid t_1), \tilde{U}(t_1 \mid t_1)]$;

$$dB(t_1 \mid t)/dt = B(t_1 \mid t)[A^*(t) - C^*(t) R^{-1}(t) C(t) P(t \mid t)]. \tag{4.5}$$

Here $\tilde{U}(t \mid t) = U(t) - \hat{U}(t \mid t)$ *and* $P(t \mid t)$ *is given by Lemma* 3.5.

Proof. Let us set

$$y(\sigma) \triangleq \int_0^\sigma C(s) U(s)\, ds = \Delta Z(\sigma) - \int_0^\sigma r(s)\, dV(s).$$

Then we have by the direct computation

$$\frac{d}{d\sigma} \mathrm{Cov}[\tilde{U}(t_1 \mid t), y(\sigma)] = \mathrm{Cov}[\tilde{U}(t_1 \mid t), C(\sigma) U(\sigma)] = \mathrm{Cov}[\tilde{U}(t_1 \mid t), U(\sigma)]\, C^*(\sigma)$$

and by Theorem 4.1

$$\mathrm{Cov}[\tilde{U}(t_1 \mid t), y(\sigma)] = \mathrm{Cov}\left[U(t_1) - \hat{U}(t_1 \mid t), \Delta Z(\sigma) - \int_0^\sigma r(s)\, dV(s)\right]$$

$$= \mathrm{Cov}\left[\hat{U}(t_1 \mid t), \int_0^\sigma r(s)\, dV(s)\right] = \int_0^\sigma K(t, s) R(s)\, ds.$$

Hence, it follows that for $\sigma < t$

$$K(t, \sigma) R(\sigma) = V(t_1 \mid \sigma) C^*(\sigma),$$

where

$$V(t_1 \mid \sigma) = \mathrm{Cov}[\tilde{U}(t_1 \mid t), U(\sigma)].$$

Since any regulated function is equivalent in an almost everywhere sense to a function continuous on the left, we can take limits as σ approaches t from below in the above relation and thus deduce that

$$K(t, t) = V(t_1 \mid t) C^*(t) R^{-1}(t). \tag{4.6}$$

Let us now derive the time evolution of $V(t_1 \mid t)$. Setting $\Psi(t, s) = \int_0^t K(t, \tau) \times C(\tau) \mathscr{U}(\tau, s)\, d\tau$ and using Theorem 4.3 yields

$$\partial\Psi(t, s)/\partial t = K(t, t) C(t)(\mathscr{U}(t, s) - \phi(t, s)), \tag{4.7}$$

where $\phi(t, s)$ is given by (3.9).

We note that

$$\hat{U}(t_1 \mid t) = \int_0^t K(t, \tau)\, dZ(\tau) = \int_0^{t_1} L(t_1, \tau)\, dZ(\tau) + \int_{t_1}^t K(t, \tau)\, dZ(\tau)$$
$$= \hat{U}(t_1 \mid t_1) + \int_{t_1}^t K(t, \tau)\, dZ(\tau), \tag{4.8}$$

and

$$dZ(\tau) = C(\tau)\, \mathscr{U}(\tau, t_1)\, U(t_1)\, d\tau + r(\tau)\, dV(\tau)$$
$$+ \left[\int_{t_1}^{\tau} C(\tau)\, \mathscr{U}(\tau, s)\, B(s)\, dW(s)\right] d\tau, \qquad \text{for} \quad t_1 \leqslant s \leqslant \tau \leqslant t.$$

Then we have

$$\int_{t_1}^t K(t, \tau)\, dZ(\tau) = \Psi(t, t_1)\, U(t_1) + \int_{t_1}^t \Psi(t, s)\, B(s)\, dW(s)$$
$$+ \int_{t_1}^t K(t, \tau)\, r(\tau)\, dV(\tau). \tag{4.9}$$

Hence, it follows that

$$\hat{U}(t_1 \mid t) = \hat{U}(t_1 \mid t_1) + \Psi(t, t_1)\, U(t_1)$$
$$+ \int_{t_1}^t \Psi(t, \tau)\, B(\tau)\, dW(\tau) + \int_{t_1}^t K(t, \tau)\, r(\tau)\, dV(\tau).$$

On the other hand, it follows that

$$U(t) = \mathscr{U}(t, t_1)\, U(t_1) + \int_{t_1}^t \mathscr{U}(t, \tau)\, B(\tau)\, dW(\tau).$$

Hence, from Lemma 3.3 we have

$$\operatorname{Cov}[\hat{U}(t_1 \mid t), U(t)] = (\operatorname{Cov}[\hat{U}(t_1 \mid t_1), U(t_1)] + \Psi(t, t_1) \operatorname{Cov}[U(t_1), U(t_1)]$$
$$+ \int_{t_1}^t \Psi(t, s)\, Q(s)\, \mathscr{U}^*(t_1, s)\, ds)\, \mathscr{U}^*(t, t_1). \tag{4.10}$$

Continuing in the same vein, we have

$$\operatorname{Cov}[U(t_1), U(t)] = \operatorname{Cov}[U(t_1), U(t_1)]\, \mathscr{U}^*(t, t_1). \tag{4.11}$$

It follows from (4.10) and (4.11) that

$$V(t_1 \mid t) = (P(t_1 \mid t_1) - \Psi(t, t_1)P_1 - \int_{t_1}^{t} \Psi(t, s)\, Q(s)\, \mathscr{U}^*(t_1, s)\, ds)\, \mathscr{U}^*(t, t_1), \tag{4.12}$$

where

$$P_1 \triangleq \mathrm{Cov}[U(t_1), U(t_1)].$$

Hence, it follows from (4.12) that

$$\frac{dV(t_1 \mid t)}{dt} = V(t_1 \mid t)\, A^*(t) - \left(\frac{\partial \Psi(t, t_1)}{\partial t} P_1 + \int_{t_1}^{t} \frac{\partial \Psi(t, s)}{\partial t} Q(s)\, \mathscr{U}^*(t_1, s)\, ds\right) \mathscr{U}^*(t, t_1). \tag{4.13}$$

Substituting (4.7) into (4.13) and taking into consideration of (3.8), it follows that

$$\begin{aligned} dV(t_1 \mid t)/dt &= V(t_1 \mid t)\, A^*(t) - K(t, t)\, C(t) \Big[(\mathscr{U}(t, t_1) - \phi(t, t_1))P_1 \\ &\quad + \int_{t_1}^{t} (\mathscr{U}(t, s) - \phi(t, s))\, Q(s)\, \mathscr{U}^*(t_1, s)\, ds\Big]\, \mathscr{U}^*(t, t_1) \\ &= V(t_1 \mid t)\, A^*(t) - K(t, t)\, C(t)\, P(t \mid t). \end{aligned}$$

Furthermore, from (4.12) we have

$$V(t_1 \mid t_1) = P(t_1 \mid t_1).$$

Since the uniqueness of the solution for (4.5) is clear, using $K(t, t)\, C(t) = V(t_1 \mid t)\, C^*(t)\, R^{-1}(t)\, C(t)$ yields

$$B(t_1 \mid t) = V(t_1 \mid t) \qquad \text{for any} \quad t > t_1 .$$

Thus, the theorem is established. Q.E.D.

Bearing (4.5) in mind, we have the following corollary.

COROLLARY 4.7. *Under the hypotheses of Theorem* 4.6, $B^*(t_1 \mid t)$ is given by

$$B^*(t_1 \mid t) = (\mathscr{U}(t, t_1) - \phi(t, t_1))\, P(t_1 \mid t_1), \tag{4.14}$$

where $\mathscr{U}(t, t_1)$ *and* $\phi(t, t_1)$ *are defined by* (2.5) *and* (3.9), *respectively.*

Proof. From (2.5) and (3.9), we have

$$\frac{\partial \mathscr{U}(t, t_1)}{\partial t} = A(t)\, \mathscr{U}(t, t_1),$$

$$\frac{\partial \phi(t, t_1)}{\partial t} = L(t, t)\, C(t)\, \mathscr{U}(t, t_1) + \int_{t_1}^{t} \frac{\partial L(t, \tau)}{\partial t}\, C(\tau)\, \mathscr{U}(\tau, t_1)\, d\tau.$$

Using (3.6) yields

$$\frac{\partial \phi(t, t_1)}{\partial t} = A(t)\, \phi(t, t_1) + L(t, t)\, C(t)\, (\mathscr{U}(t, t_1) - \phi(t, t_1)).$$

Hence,

$$\left(\frac{\partial \mathscr{U}(t, t_1)}{\partial t} - \frac{\partial \phi(t, t_1)}{\partial t}\right) P(t_1 \mid t_1)$$

$$= A(t)(\mathscr{U}(t, t_1) - \phi(t, t_1))\, P(t_1 \mid t_1) - L(t, t)\, C(t)(\mathscr{U}(t, t_1) - \phi(t, t_1))\, P(t_1 \mid t_1).$$

Letting $N(t_1 \mid t) = (\mathscr{U}(t, t_1) - \phi(t, t_1))\, P(t_1 \mid t_1)$,

$$\frac{dN(t_1 \mid t)}{dt} = (A(t) - L(t, t)\, C(t))\, N(t_1 \mid t)$$

and using the relation of Lemma 3.5 given by

$$L(t, t)\, C(t) = P(t \mid t)\, C^*(t)\, R^{-1}(t)\, C(t),$$

it follows that

$$\frac{dN^*(t_1 \mid t)}{dt} = N^*(t_1 \mid t)(A^*(t) - C^*(t)\, R^{-1}(t)\, C(t)\, P(t \mid t)).$$

Furthermore, from the definition of $N(t_1 \mid t)$ we have

$$N^*(t_1 \mid t_1) = P(t_1 \mid t_1).$$

Thus, it follows from (4.5) that $N(t_1 \mid t) = B^*(t_1 \mid t)$ and the corollary is established. Q.E.D.

To obtain the relation for the smoothing error covariance operator, we first prove the following lemma.

LEMMA 4.8. *Suppose that* $\hat{U}(t_1 \mid t)$ *and* $\tilde{U}(t_1 \mid t)$ *are defined by* (2.6) *and* (2.7), *respectively. Then we have*

(a) $\mathrm{Cov}[\tilde{U}(t_1 \mid t), \hat{U}(t_1 \mid t)] = -\Psi(t, t_1)\, \mathrm{Cov}[U(t_1), \hat{U}(t_1 \mid t_1)]$,

and

(b) $\mathrm{Cov}[\tilde{U}(t_1 \mid t), U(t_1)] = P(t_1 \mid t_1) - \Psi(t, t_1)\, \mathrm{Cov}[U(t_1), U(t_1)]$.

Proof. (a) It follows from (4.8) that

$$\tilde{U}(t_1 \mid t) = U(t_1) - \hat{U}(t_1 \mid t) = \tilde{U}(t_1 \mid t_1) - \int_{t_1}^{t} K(t, \tau)\, dZ(\tau),$$

and

$$\begin{aligned} \mathrm{Cov}[\tilde{U}(t_1 \mid t), \hat{U}(t_1 \mid t_1)] &= \mathrm{Cov}[\tilde{U}(t_1 \mid t), \hat{U}(t_1 \mid t_1)] \\ &\quad + \mathrm{Cov}\left[\tilde{U}(t_1 \mid t), \int_{t_1}^{t} K(t, \tau)\, dZ(\tau)\right]. \end{aligned}$$

Using (4.9) yields

$$\mathrm{Cov}[\tilde{U}(t_1 \mid t), \hat{U}(t_1 \mid t)] = -\Psi(t, t_1)\, \mathrm{Cov}[U(t_1), \hat{U}(t_1 \mid t_1)], \tag{4.15}$$

and

$$\begin{aligned} &\mathrm{Cov}\left[\tilde{U}(t_1 \mid t), \int_{t_1}^{t} K(t, \tau)\, dZ(\tau)\right] \\ &\quad = P(t_1 \mid t_1)\, \Psi^*(t, t_1) - \Psi(t, t_1)\, \mathrm{Cov}[U(t_1), U(t_1)]\, \Psi^*(t, t_1) \\ &\qquad - \int_{t_1}^{t} \Psi(t, \tau)\, Q(\tau)\, \Psi^*(t, \tau)\, d\tau - \int_{t_1}^{t} K(t, \tau)\, R(\tau)\, K^*(t, \tau)\, d\tau. \end{aligned}$$

Substituting $K(t, \tau)\, R(\tau)$ given by (4.6) into the above relation, we have

$$\mathrm{Cov}\left[\tilde{U}(t_1 \mid t), \int_{t_1}^{t} K(t, \tau)\, dZ(\tau)\right] = 0. \tag{4.16}$$

Hence, from (4.15) and (4.16), part (a) of the lemma is established.

(b) From (4.8) and (4.9) we have

$$\begin{aligned} \mathrm{Cov}[\tilde{U}(t_1 \mid t), U(t_1)] &= \mathrm{Cov}[\tilde{U}(t_1 \mid t_1), U(t_1)] \\ &\quad - \mathrm{Cov}\left[\int_{t_1}^{t} K(t, \tau)\, dZ(\tau), U(t_1)\right] \\ &= P(t_1 \mid t_1) - \Psi(t, t_1)\, \mathrm{Cov}[U(t_1), U(t_1)]. \end{aligned}$$

Thus, part (b) of the lemma is established. Q.E.D.

Note that (4.15) means that the usual orthogonality between $\tilde{U}(t_1 \mid t)$ and $\hat{U}(t_1 \mid t)$ does not hold although (4.16) shows that $\tilde{U}(t_1 \mid t)$ is orthogonal to $\int_{t_1}^{t} K(t, \tau)\, dZ(\tau)$.

From this lemma, we have the following theorem concerned with the smoothing error covariance operator $P(t_1 \mid t) = \mathrm{Cov}[\tilde{U}(t_1 \mid t), \tilde{U}(t_1 \mid t)]$.

THEOREM 4.9. *The error covariance operator* $P(t_1 \mid t)$ *can be given by*

$$P(t_1 \mid t) = P(t_1 \mid t_1) - \Psi(t, t_1)\, P(t_1 \mid t_1). \tag{4.17}$$

Furthermore, as the differential form of (4.17) *we have*

$$\frac{dP(t_1 \mid t)}{dt} = -B(t_1 \mid t)\, C^*(t)\, R^{-1}(t)\, C(t)\, B(t_1 \mid t) \tag{4.18}$$

with the initial condition $P(t_1 \mid t_1)$.

Proof. From the definition of $P(t_1 \mid t)$, we have

$$P(t_1 \mid t) = \mathrm{Cov}[\tilde{U}(t_1 \mid t), U(t_1)] - \mathrm{Cov}[\tilde{U}(t_1 \mid t), \hat{U}(t_1 \mid t)],$$

and using Lemma 4.8 yields

$$\begin{aligned} P(t_1 \mid t) &= P(t_1 \mid t_1) - \Psi(t, t_1)\, \mathrm{Cov}[U(t_1), \tilde{U}(t_1 \mid t_1)] \\ &= P(t_1 \mid t_1) - \Psi(t, t_1)\, P(t_1 \mid t_1). \end{aligned}$$

Hence, (4.17) is established.

Differentiating (4.17) with respect to time, we have

$$\frac{dP(t_1 \mid t)}{dt} = -\frac{\partial \Psi(t, t_1)}{\partial t} P(t_1 \mid t_1),$$

and it follows from (4.7) that

$$\frac{dP(t_1 \mid t)}{dt} = -K(t, t)\, C(t)(\mathscr{U}(t, t_1) - \phi(t, t_1))\, P(t_1 \mid t_1).$$

Using Corollary 4.7 and (4.6) yields

$$\frac{dP(t_1 \mid t)}{dt} = -B(t_1 \mid t)\, C^*(t)\, R^{-1}(t)\, C(t)\, B^*(t_1 \mid t).$$

Since it is clear that the initial condition of $P(t_1 \mid t)$ is $P(t_1 \mid t_1)$, (4.18) is established. Q.E.D.

If the underlying spaces are finite-dimensional, then $A(\cdot)$, $B(\cdot)$, $C(\cdot)$, and $r(\cdot)$ become matrices, and the results derived here can be reduced to those of the lumped parameter systems (Meditch, 1969; Kailath and Frost, 1968) obtained by the Kalman's limiting procedure or the innovation approach.

5. An Application to a Distributed Parameter System

We now briefly given an application of the results in the preceding sections to a smoothing problem for a distributed parameter system of parabolic type with pointwise observations. Let $A(t)$ be an elliptic differential operator $\sum_{i,j=1}^{m} \partial/\partial x_i(a_{ij}(t, x)\, \partial/\partial x_j)$ of order 2 in a bounded m-dimensional domain D and assume that the $a_{ij}(t, x)$ and the boundary of D, denoted by ∂D, are sufficiently smooth. The domain $\mathscr{D}(A)$ of $A(t)$ consists of all the smooth functions satisfying the Dirichlet boundary conditions or, in fact, any set of regular boundary conditions. Then $A(t)$ can be extended into a closed operator in $L^p(D)$ for any $1 < p < \infty$ (Yosida, 1968) and it satisfies conditions (i)–(iii) of Section 2.

Hence, we can apply Theorems 4.4, 4.6, and 4.9 to the following distributed parameter system;

$$\frac{\partial U(t, x)}{\partial t} = \sum_{i,i=1}^{m} \frac{\partial}{\partial x_i}\left(a_{ij}(t, x)\, \frac{\partial U(t, x)}{\partial x_j}\right) + B(t, x)\, \dot{W}(t, x), \tag{5.1}$$

$$\begin{aligned} \Gamma U(t, \xi) &= 0, \quad \xi \in \partial D, \\ U(0, x) &= U_0(x), \\ \dot{Z}(t) &= U_n(t) + \dot{V}(t), \end{aligned} \tag{5.2}$$

where

$$V(t) = (v_1(t), \ldots, v_n(t))', \; Z(t) = (Z_1(t), \ldots, z_n(t)',$$

and

$$U_n(t) = \left(\int_{D_1} U(t, x)\, dx, \ldots, \int_{D_n} U(t, x)\, dx\right)'.$$

Here Γ denotes the boundary operators, $\cdot$ shows the formal derivative with respect to time t, and $D_1, \ldots, D_n$ are bounded closed subsets of D. The rigorous meaning of (5.1) has been studied by Bensoussan (1971). For simplicity of the expression, we assume that $R(t) = rI$, $r > 0$, where I is the identity matrix. Then the optimal smoothing estimator is given by

$$\begin{aligned} \frac{\partial \hat{U}(t_1, x \mid t)}{\partial t} &= \frac{1}{r} \sum_{k=1}^{n} \int_{D_k} B(t_1, x, y \mid t)\, dy \left[\dot{z}_k(t) - \int_{D_k} \hat{U}(t, x \mid t)\, dx\right], \\ \Gamma \hat{U}(t_1, \xi \mid t) &= 0, \; \xi \in \partial D, \end{aligned} \tag{5.3}$$

where

$$\frac{\partial B(t_1, x, y \mid t)}{\partial t} = \sum_{i,i=1}^{m} \frac{\partial}{\partial x_i} \left(a_{ij}(t, x) \frac{\partial B(t_1, x, y \mid t)}{\partial x_j} \right)$$

$$- \frac{1}{r} \sum_{k=1}^{n} \int_{D_k} B(t_1, x, y \mid t)\, dy \int_{D_k} P(t, x, y \mid t)\, dx, \tag{5.4}$$

$$\Gamma B(t_1, x, \xi \mid t) = 0, \ \xi \in \partial D,$$

$$B(t_1, x, y \mid t_1) = P(t_1, x, y \mid t_1).$$

Furthermore, the smoothing error covariance function is given by

$$\frac{dP(t_1, x, y \mid t)}{dt} = - \frac{1}{r} \sum_{k=1}^{n} \int_{D_k} B(t_1, x, y \mid t)\, dy \int_{D_k} B(t_1, x, y \mid t)\, dx, \tag{5.5}$$

$$\Gamma P(t_1, x, \xi \mid t) = 0, \ \xi \in \partial D.$$

6. Conclusions

We have derived the optimal fixed-point smoothing estimator in Hilbert spaces. Our results rely heavily on abstract evolution theory and the filtering theory in infinite dimensions developed by Falb (1967). Finally, we have obtained the fixed-point smoothing estimator for a distributed parameter system using the results of the Hilbert spaces. Analogous problems were solved previously (Tzafestas, 1972) by the Bayesian approach combined with the Kalman's limiting procedure. The advantage of the present paper is that the derivations are based on the mathematically rigorous treatment than those of Tzafestas (1972), since the present approach neither necessitates the Kalman's formal procedure nor requires the delta function and the Gaussian measure on Hilbert spaces. Although Curtain's method (Curtain, 1965b) for fixed-interval smoothing problems is very similar to our approach and can be applied to the derivation of the fixed-point smoothing estimator, our method is simpler than that of Curtain's work because we need not use the innovation process.

Furthermore, Lemma 4.8 tells us that the remarkable difference exists between the filtering and the smoothing problems; that is, the former possesses the orthogonality between the estimate and the estimation error (Falb, 1967) but the latter does not. A very important problem for future research is the question of the stability of the fixed-point smoothing error covariance operator. This leads us to studies of stochastic observability since if the error covariance operator goes to null, then the state can be determined in a mean-squared sense based on the observed data.

Received: December 31, 1975; revised: June 24, 1976

References

Bensoussan, A. (1971), "Filtrage optimal des sytèmes linéaires," Dunod, Paris.

Bensoussan, A. (1975), Optimal control of stochastic linear distributed parameter systems, *SIAM J. Contr.* **13**, 904–926.

Curtain, R. F. (1975a), Infinite-dimensional filtering, *SIAM J. Contr.* **13**, 89–104.

Curtain, R. F. (1975b), "Infinite-dimensional estimation theory for linear systems," Report No. 38, Control Theory Centre, University of Warwick, England, 1–51.

Falb, P. L. (1967), Infinite-dimensional filtering: The Kalman–Bucy filter in Hilbert space, *Inform. Contr.* **11**, 102–137.

Kailath, T., and Frost, P. (1968), An innovation approach to least-squares estimation, II: Linear smoothing in a additive white noise, *IEEE Trans. Automatic Contr.* **AC-13**, 655–660.

Kato, T., and Tanabe, H. (1962), On the abstract evolution equation, *Osaka J. Math.* **14**, 107–133.

Kushner, H. J. (1970), Filtering for linear distributed parameter systems, *SIAM J. Contr.* **8**, 346–359.

Meditch, J. S. (1969), "Stochastic Optimal Linear Estimation and Control," McGraw-Hill, New York.

Meditch, J. S. (1971), Least-square filtering and smoothing for linear distributed parameter systems, *Automatica* **7**, 315–322.

Sakawa, Y. (1972), Optimal filtering in linear distributed-parameter systems, *Int. J. Contr.* **16**, 115–127.

Tzafestas, S. G. (1972), Bayesian approach to distributed-parameter filtering and smoothing, *Int. J. Contr.* **15**, 273–295.

Tzafestas, S. G., and Nightingale, J. M. (1968), Optimal filtering, smoothing and prediction in linear distributed-parameter systems, *Proc. IEEE* **115**, 1027–1093.

Tzafestas, S. G., and Nightingale, J. M. (1969), Maximum-likelihood approach to the optimal filtering of distributed-parameter systems, *Proc. IEEE* **116**, 1085–1093.

Yosida, K. (1968), "Functional Analysis," Springer-Verlag, New York.