November, 1982 LIDS-P-1256

LINEAR ESTIMATION OF BOUNDARY VALUE STOCHASTIC PROCESSES

PART II:

1-D SMOOTHING PROBLEMS

by

Milton B. Adams†

Alan S. Willsky*

Bernard C. Levy*

Abstract: This paper addresses the fixed-interval smoothing problem for linear two-point boundary value stochastic processes of the type introduced by Krener [5]. As these models are not Markovian, Kalman filtering and associated smoothing algorithms are not applicable. The smoothing problem for this class of noncausal processes is solved here by an application of the estimator solution which is developed in Part I [3] via the method of complementary models. For an n^{th} order model, this approach yields the smoother as a $2n^{th}$ order two-point boundary value problem. It is shown that this smoother can be realized in a stable two-filter form which is remarkably similar to two-filter smoothers for causal processes. In addition, expressions for the smoothing error and smoothing error covariance are developed. These equations are employed to perform a covariance analysis of estimating the temperature and heat flow in a cooling fin.

† C. S. Draper Laboratory, 555 Technology Square, Cambridge Ma. 02139 (also Dept. of Aero. and Astro.,MIT). The work of this author was supported by a Draper Fellowship.

* Dept. of Elec. Eng. and Comp. Sci. and L.I.D.S., MIT, Cambridge, Ma. 02139.
The work of these authors was supported in part by the National Science Foundation under Grant ECS-8012668.

SECTION 1

INTRODUCTION

Both linear filtering and linear smoothing for one-dimensional (1-D), nonstationary, causal processes have been extensively studied. Many of the classical solutions to these problems are discussed in the review paper by Kailath [1]. The derivations of these solutions have relied heavily on the Markovian nature of the models for these 1-D processes [2]. However, inasmuch as stochastic processes in higher dimensions (random fields) are typically noncausal and consequently are not Markovian in the usual sense, their estimators cannot be derived through a direct extension of these 1-D derivations. Thus linear estimation problems for noncausal processes require new approaches. One such new approach has been developed in Part I of this paper [3] where we have extended Weinert and Desai's [4] method of complementary models. This extension allows us to write solutions to estimation problems for a broad class of noncausal processes in one and higher dimensions. In this second part of the paper (Part II), we build upon this solution procedure in order to perform a detailed investigation of the smoothing problem for 1-D noncausal processes.

The processes that we consider are governed by the linear noncausal 1-D dynamic models introduced by Krener in [5]. In his study of these models, he has developed results on controllability, observability and minimality and has solved a deterministic linear control problem. In addition, he has posed the fixed-interval smoothing problem for these systems [6] and has derived integral equations for both the weighting pattern and error covariance of the optimal smoother. Working directly with these equations he has had success in obtaining a dynamic realization of the smoother for a special "stationary-cyclic" class of these models [7]. In this paper we begin by applying the solution for linear estimation of noncausal processes developed in Part I, and we obtain a differential realization for the optimal smoother and the smoothing error for the complete class of 1-D noncausal processes considered by Krener. For a noncausal process defined by an n^{th} order model, this solution takes the form of a $2n^{th}$ order two-point boundary value problem. Typically, solutions for this type of boundary value problem are given in the Green's function form [8], and the smoother implementation implied by this form is such that the estimate at each point in the interval of interest is obtained by numerical quadrature over the entire interval. As an alternative, in this paper we develop a two-filter implementation for our smoother which is remarkably similar to, and of nearly the same complexity as two-filter implementations developed for the fixed-interval smoother for causal processes [9,10]. As we will show, the advantage of such a two-filter form is that the estimate at each point in the interval is obtained through a linear combination of stable forward and stable backward recursions rather than numerical quadrature.

1.1 Outline

In Section 2 the linear stochastic differential equation and boundary conditions which define the noncausal 1-D process that we study are presented. Along with the model for this process, two forms of the general solution are outlined and the matrix differential equation governing the evolution of the process covariance is given. The fixed-interval smoothing problem for this model is described in Section 3. In Section 4 we formulate a two-filter implementation of the smoother by applying a decoupling transformation to the smoother dynamics which are specified by the complementary models solution. Transformations of this type have previously been applied to the smoother for causal processes by Kailath and Ljung [11] and Desai [12]. A discussion of the properties of the smoother for some special cases including causal processes and a class of systems related to Krener's [13] "separable" systems is given in Section 5. In Section 6 we apply our smoother solution to a noncausal model representing a cooling fin. Finally, Section 7 contains some concluding remarks.

SECTION 2

LINEAR STOCHASTIC TWO-POINT BOUNDARY VALUE PROCESS (TPBVP)

2.1 General Solution

The model for the one-dimensional stochastic process we consider here was introduced by Krener in [5]. The process is governed by an n^{th} order linear stochastic differential equation together with a specified two-point boundary condition. Accordingly, the process will be referred to as a linear stochastic two-point boundary value process or TPBVP. This linear boundary value process has been used to model a variety of space-time processes in temporal steady-state including the deflection of a beam under loading [8], the deflection of a rotating shaft [14] and the temperature distribution in a cooling fin [15]. (See the example in Section 6 of this paper.)

As we have shown in Part I [3], the formal structure of the linear stochastic differential equation governing the complementary process is defined by way of the structure of a related deterministic differential equation. For this reason, in Part I and here in Part II we find it convenient to employ the white noise formalism for representing linear stochastic differential equations. Let $u(t)$ be a $m\times 1$ white noise process with covariance parameter $Q(t)$. Let v be a $n\times 1$ random vector, independent of $u(t)$, with covariance matrix Π_V. The $n\times 1$ boundary value process $x(t)$ is governed on the interval $[0,T]$ by

$$\dot{x}(t) = A(t)x(t) + B(t)u(t) \tag{2.1a}$$

with boundary condition

$$v = V^{0}x(0) + V^{T}x(T) \quad . \tag{2.1b}$$

It will be assumed that A and B are continuous on $[0,T]$ and that all random variables are zero-mean since the contribution of any nonzero mean can be added separately by invoking superposition.

It is instructive to derive one form of the general solution for (2.1) as the approach we take in this derivation will be used later. The form of the solution which we obtain differs from the usual Green's function solution (e.g. see [5]). Specifically, this derivation which is posed in the terminology of linear systems theory highlights the role of a process which we will denote below by x^0. Let $\Phi(t,s)$ be the transition matrix associated with $A(t)$. If $x(0)$ were known, then $x(t)$ could be represented in the variation-of-constants form

$$x(t) = \Phi(t,0)x(0) + x^0(t) \tag{2.2a}$$

where $x^0(t)$ is the solution of (2.1a) with $x^0(0) = 0$:

$$x^0(t) = \int_0^t \Phi(t,s)B(s)u(s)ds \quad . \tag{2.2b}$$

Substituting from (2.2a) at $t = T$ into the boundary condition (2.1b), we can write

$$v - V^T x^0(t) = \left[V^0 + V^T\Phi(T,0)\right]x(0) \quad . \tag{2.3a}$$

For a well-posed problem, there will be a unique $x(0)$ for a given v and u on $[0,T]$. Thus well-posedness requires that the $n \times n$ matrix

$$F = V^0 + V^T\Phi(T,0) \tag{2.3b}$$

be nonsingular. With F invertible, we can solve for $x(0)$ as

$$x(0) = F^{-1}\left(v - V^T x^0(T)\right) \quad . \tag{2.3c}$$

Substituting $x(0)$ into (2.2a) gives the general solution for (2.1a,b) as

$$x(t) = \Phi(t,0)F^{-1}\left(v - V^T x^0(T)\right) + x^0(t) \quad . \tag{2.4}$$

The Green's function form of the general solution can be obtained from (2.4) by combining the two integrals representing $\Phi(t,0)F^{-1}V^Tx^0(T)$ and $x^0(t)$ into a single integral over $[0,T]$.

The noncausal nature of the TPBVP x(t) is clearly displayed if we correlate the value of x at t = 0 with future values of the input u:

$$E\{x(0)u'(t)\} = -F^{-1}V^{T}\Phi(T,t)B(t)Q(t) \qquad t \in [0,T] \quad . \tag{2.5}$$

Thus, the n^{th} order model in (2.1) is not Markovian, and consequently Kalman filtering and associated smoothing techniques are not directly applicable.

It is often the case for a TPBVP that the system dynamics matrix A in (2.1a) will have both positive and negative eigenvalues (see the example in Section 6). In these cases, when implementing a solution for $x^0(t)$ in (2.2b) as an initial value problem, the positive eigenvalues may cause numerical instabilities. Below, as an alternative, we present a second form for the general solution of (2.1) which leads to a numerically stable implementation. Consider the equivalent process obtained by transforming x as

$$\begin{bmatrix} x_f(t) \\ x_b(t) \end{bmatrix} = T(t)x(t) \tag{2.6a}$$

where the transformation matrix T(t) is chosen so that 1) the dynamics of the system model in (2.1) become decoupled[1]:

$$\begin{bmatrix} \dot{x}_f \\ \dot{x}_b \end{bmatrix} = \begin{bmatrix} A_f & 0 \\ 0 & A_b \end{bmatrix} \begin{bmatrix} x_f \\ x_b \end{bmatrix} + \begin{bmatrix} B_f \\ B_b \end{bmatrix} u \tag{2.6b}$$

and 2) A_f is exponentially stable in the forward direction and A_b is exponentially stable in the backward direction. For "time"-invariant systems this is always possible by assigning those modes associated with eigenvalues greater than or equal to zero to A_f and those less than zero to A_b. For time-varying dynamics, it may be difficult to determine the dynamics and boundary conditions for a transformation T(t) which transforms the system dynamics into this form. However, we will find that by invoking results obtained previously for smoothing solutions for causal processes we can overcome this difficulty for the systems of interest to us later in this paper.

---

[1] When there is no risk of confusion we will often omit explicit reference to the independent variable, i.e. $A(t) \rightarrow A$.

The boundary condition for the transformed process will be written in the following partitioned form:

$$v = \left[V_f^0 \;\vdots\; V_b^0 \right] \begin{bmatrix} x_f(0) \\ x_b(0) \end{bmatrix} + \left[V_f^T \;\vdots\; V_b^T \right] \begin{bmatrix} x_f(T) \\ x_b(T) \end{bmatrix} \quad . \tag{2.6c}$$

The reason for our choice of subscripts f and b, denoting forward and backward respectively, will become apparent below.

If $x_f(0)$ and $x_b(T)$ were known, then we could solve for $x_f(t)$ and $x_b(t)$ as

$$x_f(t) = \Phi_f(t,0)x_f(0) + x_f^0(t) \tag{2.7a}$$

and

$$x_b(t) = \Phi_b(t,T)x_b(T) + x_b^0(t) \tag{2.7b}$$

where $x_f{}^0(t)$ is governed by (2.6b) with $x_f{}^0(0) = 0$ and $x_b{}^0(t)$ is governed by (2.7b) with $x_b{}^0(T) = 0$. Following a derivation similar to that used to obtain the general solution in (2.4), it can be shown that

$$\begin{bmatrix} x_f(t) \\ x_b(t) \end{bmatrix} = \left[\begin{array}{c:c} \Phi_f(t,0) & 0 \\ \hdashline 0 & \Phi_b(t,T) \end{array}\right] F_{fb}^{-1}\left(v - V_f^T x_f^0(T) - V_b^0 x_b^0(0)\right) + \begin{bmatrix} x_f^0(t) \\ x_b^0(t) \end{bmatrix} \tag{2.8}$$

where

$$F_{fb} = \left[V_f^0 + V_f^T\Phi_f(T,0) \;\vdots\; V_b^T + V_b^0\Phi_b(0,T) \right] \quad . \tag{2.9}$$

The TPBVP x is recovered from (2.8) by inverting (2.6a):

$$x(t) = T^{-1}(t) \begin{bmatrix} x_f(t) \\ x_b(t) \end{bmatrix} \quad . \tag{2.10}$$

As we will see, the two-filter form of the general solution in (2.8) is the foundation for the implementation of the estimator that we develop later in Section 4. The term two-filter is used to signify that the numerical solution of (2.8) requires the integration of a forward process $x_f{}^0$ and a backward process $x_b{}^0$.

2.2 Covariance of the TPBVP x(t)

By a direct calculation, it can be shown that the covariance of x(t)

$$P_x(t) = E\{x(t)x'(t)\} \tag{2.11a}$$

satisfies the differential equation

$$\begin{aligned}\dot{P}_x = AP_x + P_xA' + BQB' - BQB'\Phi'(T,t)V^{T'}F^{-1'}\Phi'(t,0) \\ - \Phi(t,0)F^{-1}V^T\Phi(T,t)BQB' \;;\end{aligned} \tag{2.11b}$$

$$P_x(0) = F^{-1}(\Pi_v + V^T\Pi^0(T)V^{T'})F^{-1} \tag{2.11c}$$

where Π^0 is governed by

$$\dot{\Pi}^0 = A\Pi^0 + \Pi^0A' + BQB' \quad ; \quad \Pi^0(0) = 0 \quad . \tag{2.11d}$$

An alternative expression for P_x which requires the solution of only one matrix differential equation can be derived from (2.4) as

$$\begin{aligned}P_x(t) = P_x^0(t) + \Phi(t,0)F^{-1}\left[\Pi_v + V^TP_x^0(T)V^{T'}\right]F^{-1'}\Phi'(t,0) \\ - \Phi(t,0)F^{-1}V^TP_x^0(t) - P_x^0(t)V^{T'}F^{-1'}\Phi'(t,0)\end{aligned} \tag{2.12a}$$

where $P_x^0(t)$ is the covariance of $x^0(t)$ satisfying

$$\dot{P}_x^0 = AP_x^0 + P_x^0A' + BQB' \quad ; \qquad P_x^0(0) = 0 \quad . \tag{2.12b}$$

An additional expression for P_x can be derived from the two-filter form of the general solution (equation (2.8)). However, because this expression is somewhat complex,we will wait until later in Section 4 to present it in the context of our examination of the estimation error covariance.

2.3 Green's Identity

It was shown in Part I that the differential realization for the estimator is written in terms of the operators which define the Green's Identity for the differential operator governing the dynamics of the process to be estimated. In the notation of Part I, the differential operator representing the dynamics in (2.1a) is

$$L:D(L) \rightarrow R(L) \quad ; \quad (Lx)(t) = \dot{x}(t) - A(t)x(t) \tag{2.13}$$

where D(L) is the space of once continuously differentiable n×1 vector functions on [0,T] and R(L) is the Hilbert space of square integrable n×1 vector functions on [0,T]. Let E be the 2n×2n matrix partitioned into n×n blocks with:

$$E = \begin{bmatrix} -I & 0 \\ 0 & I \end{bmatrix} \tag{2.14a}$$

and define the 2n×1 vector

$$x_b = \begin{bmatrix} x(0) \\ x(T) \end{bmatrix} \quad . \tag{2.14b}$$

The formal adjoint of the operator L is [16]

$$(L^{\dagger}\lambda)(t) = -\dot{\lambda}(t) - A(t)\lambda(t) \quad . \tag{2.14c}$$

Given these definitions, the Green's Identity for L on the interval [0,T] is obtained directly by integration by parts, yielding

$$\langle Lx, \lambda \rangle_{L_2^n[0,T]} = \langle x, L^{\dagger}\lambda \rangle_{L_2^n[0,T]} + \langle x_b, E\lambda_b \rangle_{\mathbf{R}^{2n}} \quad . \tag{2.15}$$

SECTION 3

PROBLEM STATEMENT

The fixed-interval smoothing problem for the noncausal process $x(t)$ defined earlier in Section 2.1 is stated as follows. Let $r(t)$ be a $p\times1$ white noise process uncorrelated with v and $u(t)$ and with continuous covariance parameter $R(t)$. Let $C(t)$ be a $p\times n$ matrix whose elements are continuous on $[0,T]$. The observations of $x(t)$ are given by the $p\times1$ vector stochastic process:

$$y(t) = C(t)x(t) + r(t) \quad . \tag{3.1}$$

In addition to the observation $y(t)$, we assume that there may be available a boundary observation y_b defined as follows. Let r_b be a $q\times1$ random vector uncorrelated with $r(t)$, $u(t)$ and v with covariance matrix Π_b. Define a $q\times2n$ matrix W partitioned into $q\times n$ blocks as

$$W = \left[W^0 \;\vdots\; W^T\right] \quad . \tag{3.2a}$$

The boundary observation is the $q\times1$ random vector:

$$y_b = Wx_b + r_b \quad . \tag{3.2b}$$

Define an $n\times2n$ matrix V as

$$V = \left[V^0 \;\vdots\; V^T\right] \tag{3.3a}$$

so that the boundary condition in (2.1b) can be written as

$$v = Vx_b \tag{3.3b}$$

A condition imposed in Part I is the assumption that the rows of W and the rows of V are linearly independent. The significance of this assumption is explained as follows. If, say, the i^{th} row of W were a linear combination of the rows of V:

$$W_i = M_i V \quad , \tag{3.4a}$$

then the i^{th} element of y_b could be written as

$$\begin{aligned} y_{b_i} &= M_i V x_b + r_{b_i} \\ &= M_i v + r_{b_i} \quad . \end{aligned} \tag{3.4b}$$

Thus, y_{b_i} in (3.4b) can be viewed as a measurement of the boundary condition v. Without loss of generality we can assume that y_b has been transformed so that the elements of r_b are mutually orthogonal. As such, y_{b_i} could be eliminated from the boundary observation to be used to update our knowledge of v. This relationship between y_b and v implies that the dimension of y_b is less than or equal to n, the dimension of v.

The concept of the boundary measurement has been introduced previously in a simpler form ($W^0 = 0$, $W^T = I$) into a smoothing problem for causal processes by Bryson and Hall [17]. They included a "post-flight" measurement and showed that this additional measurement results in a nonzero initial condition for the backward filter in the two-filter implementation of the causal smoother solution. Thus, the boundary measurement introduces additional symmetry into the structure of the two-filter solution. This type of boundary measurement has a natural analog in higher dimensions where measurements of a random field may often be made along the boundary of the region over which it is defined. For example, one might have observations of temperature on the surface of an object whose internal temperature distribution is of interest. Measurements of gravity at the surface of the earth or some other body provides another example.

Returning to the 1-D problem of interest here, the fixed-interval smoothing problem is to find the linear minimum variance estimate of the noncausal TPBVP x(t), t ε [0,T], given the complete observation set Y:

$$Y = \{y_b,\ y(t) : t \in [0,T]\} \quad . \tag{3.5}$$

SECTION 4

THE TPBVP SMOOTHER

4.0 Introduction

A direct application of the differential operator representation for the estimator developed in Part I immediately yields the TPBVP smoother as a $2n^{th}$ order boundary value process. Given this two-point boundary value process, we show how it can be transformed into a two-filter form as discussed in Section 2.1. In a similar manner, we also apply the results of Part I to write a $2n^{th}$ order boundary value representation of the smoothing error and use the same transformation to develope expressions for the error covariance.

4.1 A Differential Realization for the Smoother

Let the $2n\times 2n$ matrix H be given by

$$H = \left[\begin{array}{c:c} A & BQB' \\ \hdashline C'R^{-1}C & -A' \end{array}\right] \quad . \tag{4.1a}$$

Let the $2n\times p$ matrix G be given by

$$G = \left[\begin{array}{c} 0 \\ \hdashline -C'R^{-1} \end{array}\right] \quad . \tag{4.1b}$$

Then substituting into (5.25a) of Part I, it can be shown that the smoother dynamics are given by the $2n^{th}$ order differential equation

$$\begin{bmatrix} \dot{\hat{x}} \\ \dot{\hat{\lambda}} \end{bmatrix} = H \begin{bmatrix} \hat{x} \\ \hat{\lambda} \end{bmatrix} + Gy \quad . \tag{4.2}$$

To obtain an expression for the boundary condition for this differential equation, first define two 2n×2n matrices

$$V_{x\lambda}^{0} = \left[\begin{array}{c|c} V^{0'}\Pi_v^{-1}V^{0} + W^{0'}\Pi_b^{-1}W^{0} & -I \\ \hline V^{T'}\Pi_v^{-1}V^{0} + W^{T'}\Pi_b^{-1}W^{0} & 0 \end{array}\right] \tag{4.3a}$$

and

$$V_{x\lambda}^{T} = \left[\begin{array}{c|c} V^{0'}\Pi_v^{-1}V^{T} + W^{0'}\Pi_b^{-1}W^{T} & 0 \\ \hline V^{T'}\Pi_v^{-1}V^{T} + W^{T'}\Pi_b^{-1}W^{T} & I \end{array}\right] . \tag{4.3b}$$

Then from (5.25b) in Part I, with the transpose of the matrices V and W identified as the operator adjoints V^* and W^*, the boundary condition for the smoother can be shown to be given by

$$\left[\begin{array}{c} W^{0'}\Pi_b^{-1}y_b \\ \hline W^{T'}\Pi_b^{-1}y_b \end{array}\right] = V_{x\lambda}^{0}\begin{bmatrix} \hat{x}(0) \\ \hat{\lambda}(0) \end{bmatrix} + V_{x\lambda}^{T}\begin{bmatrix} \hat{x}(T) \\ \hat{\lambda}(T) \end{bmatrix} . \tag{4.3c}$$

4.2 Hamiltonian Diagonalization

The solution of the $2n^{th}$ order boundary value process in (4.2) and (4.4) could be implemented by either of the two forms of the general solution derived in Section 2.1. However, by considering the "time"-invariant case we can anticipate, as discussed in that section, that there may be numerical instabilities associated with the first of those methods. In the time-invariant case the eigenvalues of the 2n×2n Hamiltonian[1] matrix H defined in (4.1a) are symmetric about the imaginary axis [19], i.e. there are n eigenvalues in each of the left and right half planes. Thus, for the time-invariant case, the right half plane eigenvalues will result in numerical instabilities for the unidirectional implementation suggested by (2.4). Recall that these stability problems can be avoided in general by transforming the smoother dynamics into the stable forward/backward form in (2.8). To achieve this second form, we need a transformation which diagonalizes the dynamics of H into two n×n blocks, one stable in the forward

[1] The terminology Hamiltonian is employed for historical reasons [18].

direction and the other backwards stable. As discussed below, this transformation is readily obtained by adapting results from previous studies of the smoother for causal processes.

Since the dynamics of our smoother as represented by H are identical to those of the smoother for causal processes as originally derived by Bryson and Frazier [20], any transformation which results in a two-filter smoother for causal processes will also diagonalize our smoother. As mentioned earlier, these diagonalizing transformations have been studied in [11] and [12], see also [21]. However, choosing a diagonalizing transformation for our problem requires special considerations not encountered in the causal case. First, because the two-point boundary condition provides incomplete information for both the initial and final values of the process, we will choose a transformation which corresponds to a two-filter solution for causal processes with both filters in information form. Second, as we will see, it is important to choose the boundary conditions properly for the Riccati differential equations which govern the time-varying elements of the diagonalizing transformation. In particular, the choice that we make here leads to an explicit representation for both the smoother and smoothing error covariance in terms of a single critical variable. With the smoother in this form we will be able to interpret some special cases in the next section. Finally, as discussed later, our choice of diagonalizing transformation and corresponding boundary conditions makes it possible to formulate a numerically stable two-filter form for our smoother which is remarkably similar to two-filter smoothers for causal processes.

Define the time-varying transformation T(t) as the 2n×2n matrix partitioned in n×n blocks as

$$T(t) = \left[\begin{array}{c:c} \theta_f(t) & -I \\ \hdashline \theta_b(t) & I \end{array}\right] \quad . \tag{4.4a}$$

Let the transformed process be denoted by

$$q(t) = \begin{bmatrix} q_f(t) \\ q_b(t) \end{bmatrix} = T(t) \begin{bmatrix} \hat{x}(t) \\ \hat{\lambda}(t) \end{bmatrix} \quad . \tag{4.4b}$$

Also define

$$\dot{H}_q = \dot{T}T^{-1} + THT^{-1} \tag{4.5a}$$

and

$$G_q = TG \tag{4.5b}$$

so that the dynamics of the transformed process can be written as

$$\begin{bmatrix} \dot{q}_f \\ \dot{q}_b \end{bmatrix} = H_q \begin{bmatrix} q_f \\ q_b \end{bmatrix} + G_q y \quad . \tag{4.5c}$$

If we use the following form for the inverse of T:

$$T^{-1}(t) = \left[\begin{array}{c:c} I & I \\ \hdashline -\theta_b(t) & \theta_f(t) \end{array}\right] \left[\begin{array}{c:c} P_s(t) & 0 \\ \hdashline 0 & P_s(t) \end{array}\right] \tag{4.6a}$$

where

$$P_s(t) = \left[\theta_f(t) + \theta_b(t)\right]^{-1} \quad , \tag{4.6b}$$

and if we choose the dynamics for θ_f and θ_b as

$$-\dot{\theta}_f = \theta_f A + A'\theta_f + \theta_f BQB'\theta_f - C'R^{-1}C \tag{4.6c}$$

and

$$-\dot{\theta}_b = \theta_b A + A'\theta_b - \theta_b BQB'\theta_b + C'R^{-1}C \quad , \tag{4.6d}$$

then carrying out the calculation in (4.5a), it can be shown that H_q is diagonalized with diagonal blocks

$$H_f = -[A' + \theta_f BQB'] \tag{4.6e}$$

and

$$H_b = -[A' - \theta_b BQB'] \quad . \tag{4.6f}$$

Thus the dynamics of q_f and q_b are decoupled and are given by

$$\dot{q}_f = H_f q_f + C'R^{-1}y \tag{4.7a}$$

and

$$\dot{q}_b = H_b q_b - C'R^{-1}y \quad . \tag{4.7b}$$

If we assume for time-invariant dynamics that $\{A,B\}$ is stabilizable and that $\{A,C\}$ is detectible and for time-varying dynamics that $\{A,B\}$ is uniformly completely controllable and $\{A,C\}$ is uniformly completely reconstructable, then the invertibility of P_s in (4.6b) is guaranteed if both $\theta_f(0)$ and $\theta_b(T)$ are nonnegative definite [19]. Furthermore, these conditions guarantee that θ_f and θ_b and their derivatives are bounded and that H_f and H_b are forward and backward stable respectively.

Under the transformation (4.4a), the boundary condition (4.3c) becomes

$$\begin{bmatrix} W^{0'}\Pi_b^{-1}y_b \\ W^{T'}\Pi_b^{-1}y_b \end{bmatrix} = V_q^0 \begin{bmatrix} q_f(0) \\ q_b(0) \end{bmatrix} + V_q^T \begin{bmatrix} q_f(T) \\ q_b(T) \end{bmatrix} \tag{4.8a}$$

where

$$V_q^0 = V_{x\lambda}^0 T^{-1}(0) \tag{4.8b}$$

and

$$V_q^T = V_{x\lambda}^T T^{-1}(T) \quad . \tag{4.8c}$$

To simplify the expressions for the boundary value coefficient matrices in (4.8b) and (4.8c), choose the following nonnegative definite initial and final conditions for the Riccati equations (4.6c) and (4.6d):

$$\theta_f(0) = V^{0'}\Pi_v^{-1}V^0 + W^{0'}\Pi_b^{-1}W^0 \tag{4.9a}$$

and

$$\theta_b(T) = V^{T'}\Pi_v^{-1}V^T + W^{T'}\Pi_b^{-1}W^T \quad . \tag{4.9b}$$

Then defining θ_c as the following n×n matrix:

$$\theta_c = V^{T'}\Pi_v^{-1}V^0 + W^{T'}\Pi_b^{-1}W^0 \quad , \tag{4.10}$$

it can be shown that the boundary value coefficient matrices can be written as

$$V_q^0 = \begin{bmatrix} I & \vdots & 0 \\ - & \vdots & - \\ \theta_c P_s(0) & \vdots & \theta_c P_s(0) \end{bmatrix} \equiv \begin{bmatrix} V_f^0 & \vdots & V_b^0 \end{bmatrix} \tag{4.10a}$$

and

$$V_q^T = \begin{bmatrix} \theta_c' P_s(T) & \vdots & \theta_c' P_s(T) \\ - & \vdots & - \\ 0 & \vdots & I \end{bmatrix} \equiv \begin{bmatrix} V_f^T & \vdots & V_b^T \end{bmatrix} \quad . \tag{4.10b}$$

Since the dynamics of q_f and q_b are decoupled, the only coupling between the two enters through the boundary condition. By our choice of initial and final conditions for the Riccati equations, we have been able to display this coupling solely as a function of the matrix θ_c.

The smoothed estimate of x is recovered by inverting T(t) in (4.4b) so that we obtain

$$\hat{x}(t) = P_s(t)\left[q_f(t) + q_b(t)\right] \quad . \tag{4.11}$$

Following (2.8), an explicit expression for the two-filter solution for q_f and q_b is formulated as follows. Let $q_f{}^0$ and $q_b{}^0$ be governed by

(4.7a) and (4.7b) respectively with boundary conditions: $q_f^0(0) = 0$ and $q_b^0(T) = 0$. Define F_{fb} and Φ_{fb} as the 2n×2n matrices

$$F_{fb} = \left[V_f^0 + V_f^T \Phi_f(T,0) \;\vdots\; V_b^T + V_b^0 \Phi_b(0,T)\right]$$

$$= \left[\begin{array}{c:c} I + \theta_c' P_s(T)\Phi_f(T,0) & \theta_c' P_s(T) \\ \hdashline \bar{\theta}_c P_s(0) & I + \bar{\theta}_c \bar{P}_s(0)\bar{\Phi}_b(0,T) \end{array}\right] \tag{4.12}$$

and

$$\Phi_{fb}(t) = \left[\begin{array}{c:c} \Phi_f(t,0) & 0 \\ \hdashline 0 & \Phi_b(t,T) \end{array}\right] \quad . \tag{4.13}$$

Then the two-filter solution for q(t) is given by

$$\begin{bmatrix} q_f(t) \\ q_b(t) \end{bmatrix} = \Phi_{fb}(t) F_{fb}^{-1} \left\{ \begin{bmatrix} W^{0'} \\ W^{T'} \end{bmatrix} \Pi_b^{-1} y_b - \begin{bmatrix} \theta_c' P_s(T) q_f^0(T) \\ \theta_c P_s(0) q_b^0(0) \end{bmatrix} \right\} + \begin{bmatrix} q_f^0(t) \\ q_b^0(t) \end{bmatrix} \quad . \tag{4.14}$$

The computational complexity of the noncausal smoother implementation suggested by (4.11) and (4.14) is nearly the same as that of the two-filter smoothers for causal processes such as the Mayne-Fraser form [9,10]. We note, however, that before q_f and q_b can be evaluated for any $t \in [0,T]$, both q_f^0 and q_b^0 must be computed and stored along with P_s and Φ_{fb} for the entire interval [0,T]. Thus, the required storage exceeds that of the smoother for causal processes. Indeed, the Mayne-Fraser solution and ours differ significantly in one aspect. That is, for our smoother the contribution of the forward filter to the smoothed estimate at some point t depends not only on past observations, as does the Mayne-Fraser solution, but also on future observations through the term $\theta_c' P_s(T) q_f^0(T)$ in (4.14). A similar statement applies for the backward process.

4.3 Smoothing Error

From (5.36) in Part I, the differential realization of the smoothing error is

$$\begin{bmatrix} \dot{\tilde{x}} \\ \dot{\hat{-\lambda}} \end{bmatrix} = H \begin{bmatrix} \tilde{x} \\ \hat{-\lambda} \end{bmatrix} + \begin{bmatrix} Bu \\ - - - \\ C'R^{-1}r \end{bmatrix} \tag{4.15b}$$

with boundary condition (from (5.33) of Part I)

$$v_e = \begin{bmatrix} V^{0'} & W^{0'} \\ V^{T'} & W^{T'} \end{bmatrix} \begin{bmatrix} \Pi_v^{-1} & 0 \\ 0 & \Pi_b^{-1} \end{bmatrix} \begin{bmatrix} v \\ -r_b \end{bmatrix} = V_x^0 \begin{bmatrix} \tilde{x}(0) \\ \hat{-\lambda}(0) \end{bmatrix} + V_x^T \begin{bmatrix} \tilde{x}(T) \\ \hat{-\lambda}(T) \end{bmatrix} \quad . \tag{4.15b}$$

The same diagonalizing transformation in (4.4a) can be applied to the error dynamics with the result that, as we will see, the error covariance can be computed from many of the same quantities required for computing the smoothed estimate.

In a manner similar to (4.4b) let

$$e(t) = \begin{bmatrix} e_f(t) \\ e_b(t) \end{bmatrix} = T(t) \begin{bmatrix} \tilde{x}(t) \\ \hat{-\lambda}(t) \end{bmatrix} \quad . \tag{4.16}$$

Then the smoothing error is

$$\tilde{x}(t) = P_s(t)\left[e_f(t) + e_b(t)\right] \tag{4.17}$$

where e_f and e_b satisfy the decoupled dynamics

$$\dot{e}_f = H_f e_f + \left[\theta_f B \; \vdots \; -C'R^{-1}\right] \begin{bmatrix} u \\ r \end{bmatrix} \tag{4.18a}$$

and

$$\dot{e}_b = H_b e_b + \left[\theta_b B \; \vdots \; C'R^{-1}\right] \begin{bmatrix} u \\ r \end{bmatrix} \quad . \tag{4.18b}$$

Under this transformation the boundary condition takes the form (see (4.10a,b))

$$v_e = \left[V_f^0 \;\vdots\; V_b^0\right] \begin{bmatrix} e_f(0) \\ e_b(0) \end{bmatrix} + \left[V_f^T \;\vdots\; V_b^T\right] \begin{bmatrix} e_f(T) \\ e_b(T) \end{bmatrix} \quad . \tag{4.19}$$

Below we develop an expression for the error covariance. Let

$$\Sigma_f(t) = E\{e_f(t)e_f'(t)\} \; , \tag{4.20a}$$

$$\Sigma_b(t) = E\{e_b(t)e_b'(t)\} \tag{4.20b}$$

and

$$\Sigma_{fb}(t) = E\{e_f(t)e_b'(t)\} \quad . \tag{4.20c}$$

The covariance of the smoothing error can be written directly from (4.17) as

$$\begin{aligned} P(t) &= E\left\{\tilde{x}(t)\tilde{x}'(t)\right\} \\ &= P_s(t)\left[\Sigma_f(t) + \Sigma_b(t) + \Sigma_{fb}(t) + \Sigma_{fb}'(t)\right]P_s(t) \quad . \end{aligned} \tag{4.21}$$

We derive expressions for each of the individual covariances in (4.20) by expressing e(t) in the two-filter form of (2.8). Accordingly, let e_f^0 and e_b^0 be governed by (4.18a) and (4.18b) respectively with boundary conditions: $e_f^0(0) = 0$ and $e_b^0(T) = 0$. Then e_f and e_b can be written as

$$\begin{bmatrix} e_f(t) \\ e_b(t) \end{bmatrix} = \Phi_{fb}(t)F_{fb}^{-1}\left\{ v_e - \begin{bmatrix} \theta_c' P_s(T)e_f^0(T) \\ \theta_c P_s(0)e_b^0(0) \end{bmatrix} \right\} + \begin{bmatrix} e_f^0(t) \\ e_b^0(t) \end{bmatrix} \quad . \tag{4.22}$$

Thus the covariances in (4.20) can be expressed in terms of the covariance of v_e and the covariances and cross-covariance of e_f^0 and e_b^0.

First note from (4.15b) that the covariance of v_e is given by

$$\Pi_{v_e} = E\{v_e v_e'\} = \begin{bmatrix} \theta_f(0) & \vdots & \theta_c' \\ \cdots & & \cdots \\ \theta_c & \vdots & \theta_b(T) \end{bmatrix} \quad . \tag{4.23}$$

The covariance of e_f^0

$$\Sigma_f^0(t) = E\{e_f^0(t)e_f^{0'}(t)\} \tag{4.24a}$$

satisfies

$$\dot{\Sigma}_f^0 = H_f\Sigma_f^0 + \Sigma_f^0 H_f' + \theta_f BQB'\theta_f + C'R^{-1}C \quad ; \quad \Sigma_f^0(0) = 0 \quad . \tag{4.24b}$$

Similarly, the covariance matrix for e_b^0 satisfies

$$\dot{\Sigma}_b^0 = H_b\Sigma_b^0 + \Sigma_b^0 H_b' - \theta_b BQB'\theta_b - C'R^{-1}C \quad ; \quad \Sigma_b^0(T) = 0 \quad . \tag{4.25}$$

To obtain an expression for the cross-correlation:

$$E\left\{e_f^0(t)e_b^{0'}(\tau)\right\} = \begin{cases} \Sigma_{fb}^0(t,\tau), & t > \tau \\ 0 \quad , & t \leq \tau \end{cases} \quad , \tag{4.26}$$

first define

$$\dot{\Pi}_{fb}^0 = H_f\Pi_{fb}^0 + \Pi_{fb}^0 H_b' + \theta_f BQB'\theta_b - C'R^{-1}C \quad ; \quad \Pi_{fb}^0(0) = 0 \quad . \tag{4.27}$$

Substituting the variation of constants integral expressions for the processes in the expectation in (4.26), it can be shown that for $t > \tau$:

$$\Sigma_{fb}^0(t,\tau) = \Phi_f(t,\tau)\Pi_{fb}^{0'}(\tau) - \Pi_{fb}^0(t)\Phi_b'(\tau,t) \tag{4.28a}$$

and that

$$\Sigma_{bf}^0(\tau,t) = \Sigma_{fb}^{0'}(t,\tau) \quad . \tag{4.28b}$$

Finally, combining these identities we can express

$$\Sigma_e(t) = E\{e(t)e'(t)\} = \begin{bmatrix} \Sigma_f(t) & \Sigma_{fb}(t) \\ \Sigma_{fb}'(t) & \Sigma_b(t) \end{bmatrix} \tag{4.29}$$

as

$$\Sigma_e(t) = \Phi_{fb}(t)F_{fb}^{-1}\left[\begin{array}{c|c} \theta_f(0) + \theta_c' P_s(T)\Sigma_f^0(T)P_s(T)\theta_c & \theta_c' P_s(T)\Sigma_{fb}^0(T,0)P_s(0)\theta_c' + \theta_c' \\ \hline \theta_c P_s(0)\Sigma_{bf}^0(0,T)P_s(T)\theta_c + \theta_c & \theta_b(T) + \theta_c P_s(0)\Sigma_b^0(0)P_s(0)\theta_c' \end{array}\right]\Phi_{fb}(t)F_{fb}^{-1}$$

$$+ \begin{bmatrix} \Sigma_f^0(t) & 0 \\ 0 & \Sigma_b^0(t) \end{bmatrix} - \Phi_{fb}(t)F_{fb}^{-1}\left[\begin{array}{c|c} \theta_c' P_s(T)\Phi_f(T,t)\Sigma_f^0(t) & \theta_c' P_s(T)\Sigma_{fb}^0(T,t) \\ \hline \theta_c P_s(0)\Sigma_{bf}^0(0,t) & \theta_c P_s(0)\Phi_b(0,t)\Sigma_b^0(t) \end{array}\right]$$

$$- \left[\begin{array}{c|c} \Sigma_f^0(t)\Phi_f'(T,t)P_s(T)\theta_c & \Sigma_{fb}^0(t,0)P_s(0)\theta_c' \\ \hline \Sigma_{bf}^0(t,T)P_s(T)\theta_c & \Sigma_b^0(t)\Phi_b'(0,t)P_s(0)\theta_c' \end{array}\right] F_{fb}^{-1'}\Phi_{fb}'(t) \quad . \tag{4.30}$$

Next, note that it can be shown that the solutions of (4.24) and (4.25) are related to θ_f and θ_b in (4.6c) and (4.6d) by

$$\Sigma_f^0(t) = \theta_f(t) - \Phi_f(t,0)\theta_f(0)\Phi_f'(t,0) \tag{4.31a}$$

and

$$\Sigma_b^0(t) = \theta_b(t) - \Phi_b(t,T)\theta_b(T)\Phi_b'(t,T) \quad . \tag{4.31b}$$

That is,

$$\Sigma_f^0(t) = \theta_f^0(t) \qquad \text{and} \qquad \Sigma_b^0(t) = \theta_b^0(t) \quad . \tag{4.31c}$$

When $\Sigma_f{}^0$ and $\Sigma_b{}^0$ are replaced in (4.30) by the expressions in (4.31a) and (4.31b), it can be seen that the only computation required in excess of that already performed for the smoother solution is the integration of $\Pi^0{}_{fb}$ in (4.27).

Although the expression for the covariance in (4.30) may seem forbidding, it does explicitly display the dependence of Σ_e on θ_c. In the next section we discuss a special class of problems for which θ_c is zero. As a preview to that discussion, we note that when $\theta_c = 0$,

i) $F_{fb} = I$

and

$$\text{ii)}\quad \Sigma_e(t) = \Phi_{fb}(t)\left[\begin{array}{c:c} \theta_f(0) & 0 \\ \hdashline 0 & \theta_b(T) \end{array}\right]\Phi_{fb}'(t) + \left[\begin{array}{c:c} \Sigma_f^0(t) & 0 \\ \hdashline 0 & \Sigma_b^0(t) \end{array}\right] .$$

Substituting from (4.29), $\Sigma_e(t)$ for this case becomes simply

$$\Sigma_e(t) = \left[\begin{array}{c:c} \theta_f(t) & 0 \\ \hdashline 0 & \theta_b(t) \end{array}\right]$$

which implies that the forward and backward error processes e_f and e_b are orthogonal and that the smoothing error covariance in (4.21) is

$$P(t) = P_s(t) = \left[\theta_f(t) + \theta_b(t)\right]^{-1} .$$

Also, when θ_c is zero, the noncausal contributions of the forward and backward processes $q_f{}^0$ and $q_b{}^0$ to the smoothed estimate are eliminated (see (4.14)). Note that all of these are also properties of the two-filter smoothers for causal processes [2]. In the next section we will show that for the case when θ_c is zero, q_f and q_b can be interpreted as the forward and backward information vectors for a causal process smoother with special nonzero boundary values for θ_f and θ_b.

SECTION 5

SPECIAL CASES

5.0 Introduction

In the first part of this section we discuss some properties of the smoother for a class of noncausal processes with special boundary conditions and boundary observations. A subset of this class was first studied by Krener [13]. Here we show for this class that the smoother described in the previous section is equivalent to a previously derived smoother for causal processes. The last topic of the section is alternative transformations which lead to two of the popular forms of the smoother for causal processes, namely the Mayne-Fraser and the Rauch-Tung-Striebel. The former belongs to the class of diagonalizing transformations studied by Kailath and Ljung [11] and Desai [12] and the latter is a triangularizing transformation [21].

5.1 Separable Systems

In the context of 1-D linear stochastic TPBVPs, Krener first introduced the terminology separable to describe a class of n^{th} order noncausal stationary processes which are, in fact, n^{th} order Markov, i.e. their evolution can be described by an n^{th} order linear stochastic differential equation with a prescribed initial condition which is orthogonal to future inputs. Recall that, in general, the boundary value representation for noncausal processes which we presented in Section 2.1 is not a Markov model. Along with stationarity, Krener's criteria for separability includes a block-diagonal form for Π_V and the orthogonality condition: $V^{T'}V^0 = 0$. In fact, the slightly less restrictive condition

$$V^{T'}\Pi_V^{-1}V^0 = 0 \tag{5.1}$$

could have been imposed. In [21], the stationarity condition was shown to be unnecessary so that (5.1) is both necessary and sufficient for the existence of an n^{th} order Markov model. With respect to the smoothing problem, the existence of such a model implies that when there is no boundary measurement, any of the smoothers for causal processes can be applied directly to the Markov model. Here we will extend the notion of separability to include cases for which there is a boundary measurement and say that a system is separable if

$$\theta_c = V^{T'} \Pi_v^{-1} V^0 + W^{T'} \Pi_b^{-1} W^0 \tag{5.2}$$

is zero. Note that this condition is compatible with Krener's original condition when there is no boundary measurement ($W^0 = W^T = 0$).

When θ_c is zero, the boundary condition in (4.8) becomes decoupled (see (4.10)) and F_{fb} in (4.12) becomes the identity so that q_f and q_b are completely decoupled with boundary conditions

$$q_f(0) = W^{0'} \Pi_b^{-1} y_b \tag{5.3a}$$

and

$$q_b(T) = W^{T'} \Pi_b^{-1} y_b \quad . \tag{5.3b}$$

Based on this observation, we can interpret the smoother for the separable case as being equivalent to Bryson and Hall's [17] problem with a "post-flight" measurement as follows.

Here we consider the information in the boundary condition v and observation y_b when combined into a single measurement:

$$\begin{bmatrix} 0 \\ y_b \end{bmatrix} = \left[\begin{array}{c:c} V^0 & V^T \\ \hdashline W^0 & W^T \end{array}\right] \begin{bmatrix} x(0) \\ x(T) \end{bmatrix} + \begin{bmatrix} -v \\ r_b \end{bmatrix} \quad . \tag{5.4}$$

This information will be viewed in the form of an information vector [22]. An information vector is used to store information about a random vector when the apriori uncertainty for that random vector (or at least some of its components) is infinite, i.e. it is totally unkown. When sufficient measurement information has been gathered so that the error-covariance matrix for the random vector becomes finite, the stored information in the form of the information vector can be transformed by the inverse of the covariance matrix (the information matrix) to produce a finite error-variance estiatmte of the random vector. In (5.4) above we have posed the boundary condition for $\{x(0),x(T)\}$ as a measurement. In this way we can consider the aproiri information as totally uncertain. Since v and r_b are orthogonal random variables, it can be shown that the information matrix ψ_x and information vector i_x associated with (5.4) are

$$\psi_x = \left[\begin{array}{c:c} \theta_f(0) & \theta_c' \\ \hdashline \theta_c & \theta_b(T) \end{array}\right] \tag{5.5a}$$

and

$$i_x = \left[\begin{array}{c:c} V^{0'} & W^{0'} \\ \hdashline V^{T'} & W^{T'} \end{array}\right] \left[\begin{array}{c:c} \Pi_v^{-1} & 0 \\ \hdashline 0 & \Pi_b^{-1} \end{array}\right] \begin{bmatrix} 0 \\ y_b \end{bmatrix} = \begin{bmatrix} W^{0'} \\ W^{T'} \end{bmatrix} \Pi_b^{-1} y_b \tag{5.5b}$$

where $\theta_f(0)$ and $\theta_b(T)$ are given by (4.9a,b). Separability is thus the case when the information about x(0) contained in the combined boundary measurement (5.4) is orthogonal to that for x(T), i.e. ψ_x is block-diagonal. By considering (5.3a) as the initial value for an information form Kalman filter for x(t) with associated information matrix $\theta_f(0)$ and by considering (5.3b) as the information vector corresponding to a "post-flight" measurement with associated information matrix $\theta_b(T)$, we find that separability is equivalent to a causal process with (possibly) incomplete information about its initial value plus a post-flight measurement. Finally, we remark that from (5.2) we see that even when (5.1) is not satisfied it is still possible to achieve separability if the boundary measurement is designed so that $W^{T'}\Pi_b^{-1}W^0$ cancels $V^{T'}\Pi_v^{-1}V^0$.

5.2 Alternative Transformations

As Kailath and Ljung [11] have noted, there exists a family of transformations which diagonalize the Hamiltonian H. In addition to diagonalization, there are other special structures for the smoother dynamics which lead to smoother implementations which may also be of interest. For example, here we present both a diagonalizing and a triangularizing transformation each with appropriate boundary conditions so that their application results in the Mayne-Fraser and Rauch-Tung-Striebel smoothers respectively for causal processes.

I) Mayne-Fraser

The Mayne-Fraser two-filter smoother is obtained by choosing the transformation

$$T(t) = \left[\begin{array}{c:c} I & -P(t) \\ \hline \theta_b(t) & I \end{array}\right] \tag{5.6a}$$

where P satisfies

$$\dot{P} = AP + PA' + BQB' - PC'R^{-1}CP \quad ; \quad P(0) = \Pi_v \tag{5.6b}$$

and θ_b satisfies (4.9b) with boundary condition $\theta_b(T) = 0$.

II) Rauch-Tung-Striebel

As an alternative to diagonalization, the smoother dynamics are triangularized by applying the transformation

$$T(t) = \left[\begin{array}{c:c} 0 & I \\ \hline I & -P(t) \end{array}\right] \tag{5.7}$$

with the dynamics and boundary condition of P given by (5.6b). With this transformation, the Hamiltonian dynamics become block-triangular yielding the Rauch-Tung-Striebel smoother for causal processes.

SECTION 6

EXAMPLE: THIN ROD HEAT EXCHANGER

6.0 Introduction

Thin rods or fins are commonly used as the medium for dissipating heat from some primary source to a coolant fluid which passes over the rods [15]. We will consider the temporal steady-state heat transfer for the two configurations depicted in Figures 6.1a and 6.1b[1]. That is, we will be looking at the heat distribution for some snapshot in time when temporal variations have settled out.

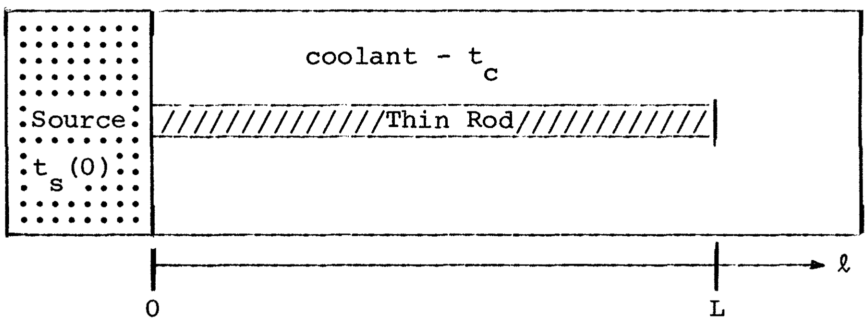


Figure 6.1a) Thin Rod Case

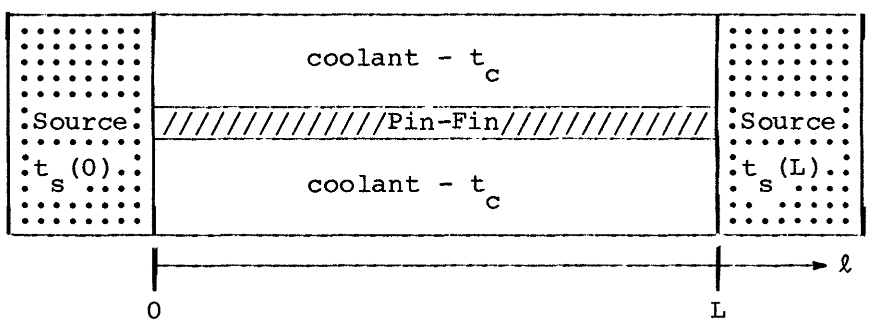


Figure 6.1b) Pin-Fin Case

In this section we present a probabilistic two-point boundary value representation for the steady-state temperature distribution and heat flow along the rod for these two cases. The corresponding deterministic TPBVP models for these configurations in temporal steady-state can be found in most

[1] Temperatures are denoted by lower case t and the independent variable, length along the rod, by ℓ.

introductory texts on heat transfer such as [15] or [23]. Following the discussion of these models, some numerical results for a covariance analysis of the TPBVP smoother as applied to these cases are presented.

6.1 The Dynamics

As is typically done [23], it will be assumed that the rod is sufficiently thin so that in temporal steady-state the temperature of the rod can be considered constant throughout any cross-section. Given this assumption, the spatial dynamics of the temporal steady-state temperature and heat flow are derived by balancing the rod-to-coolant heat energy exchange with the along-rod heat energy conduction.

For our probabalistic approach, the coolant temperature along the rod, $t_c(\ell)$, will be modelled as a constant ambient value plus a white noise fluctuation:

$$t_c(\ell) = t_{amb} + \eta(\ell) \tag{6.1}$$

$$E\{\eta(\ell)\eta(s)\} = Q\delta(\ell-s)$$

The fluctuation is meant to account for both spatial and temporal variations in coolant temperature. Note that $\eta(\ell)$ might be a second order process which could be modelled as the output of shaping filter and incorporated into our state model below via state augmentation. We have used white noise here for simplicity in presentation.

One state variable, $t(\ell)$, is defined as the difference between the rod temperature and the coolant ambient:

$$t(\ell) = t_{rod}(\ell) - t_{amb} \tag{6.2}$$

The other state variable is the derivative of $t(\ell)$:

$$\dot{t}(\ell) = \frac{dt(\ell)}{d\ell} \quad . \tag{6.3}$$

Defining

k = thermal conductivity of the rod (Btu/(hr ft F))
A = cross-sectional area of the rod (sq ft)
p = rod perimeter (ft)
h = rod-coolant heat transfer coefficient (Btu/(sq ft hr F))

and

$$m^2 = hp/kA \quad ,$$

the state dynamics with t in degrees F are given by

$$\begin{bmatrix} \dot{t} \\ \ddot{t} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ m^2 & 0 \end{bmatrix} \begin{bmatrix} t \\ \dot{t} \end{bmatrix} + \begin{bmatrix} 0 \\ m^2 \end{bmatrix} \eta \quad . \tag{6.4}$$

The heat flow at any point along the rod is given by [23]

$$q(\ell) = -kA\dot{t}(\ell) \quad \text{(Btu/hr)} \quad . \tag{6.5}$$

6.2 Measurement Model

The dynamics in (6.4) are common to both the thin rod and pin-fin configurations. Before discussing their boundary conditions, we describe the measurement which is assumed to be available for both cases. Let

$$y(\ell) = [1:0] \begin{bmatrix} t(\ell) \\ \dot{t}(\ell) \end{bmatrix} + r(\ell) \tag{6.6}$$

$$E\{r(\ell)r(s)\} = R\delta(\ell-s)$$

represent a noisy measurement of temperature along the rod. One could conceive of these measurements as being obtained optically by infra-red techniques. Here we have modelled the measurement noise as white, while in practice optical measurements might also contain some noncausal blurring which could be accounted for via state augmentation.

6.3 Boundary Conditions

The two cases depicted in Figure 6.1 are distinguishable through their boundary conditions. The boundary condition for the thin rod case in Figure

6.1a is determined by a) the temperature of the rod at the source:

$$
\begin{aligned}
t(0) &= t_s \\
&= t_m + v_t(0)
\end{aligned} \tag{6.7a}
$$

where t_m is an a priori mean, and $v_t(0)$ is a zero mean variation about t_m with variance $\sigma_t^2(0)$; and by b) equating conduction and convection at the end of the rod:

$$
v_q(L) = h'A\left[t(L) - t_{amb}\right] + kA\dot{t}(L) \tag{6.7b}
$$

where h' is the coefficient of heat transfer through the end of the rod and $v_q(L)$ is a zero mean random variable with variance σ_q^2 used to compensate for errors in determining k and h'.

Thus, we have the following boundary condition for the thin rod case:

$$
\begin{bmatrix} (t_m - t_{amb}) + v_t(0) \\ h'At_{amb} + v_q(L) \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} t(0) \\ \dot{t}(0) \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ Ah' & Ak \end{bmatrix} \begin{bmatrix} t(L) \\ \dot{t}(L) \end{bmatrix} \quad . \tag{6.7c}
$$

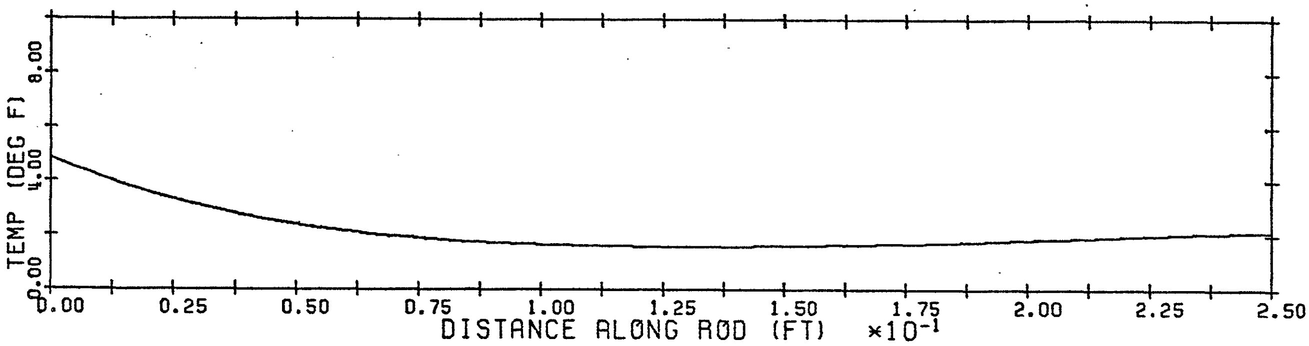
Note that when $v_t(0)$ and $v_q(L)$ are uncorrelated, (6.7c) satisfies the separability condition (5.2).

The boundary condition for the pin-fin case in Figure 6.1b is obtained from (6.7a) at both $\ell = 0$ and $\ell = L$:

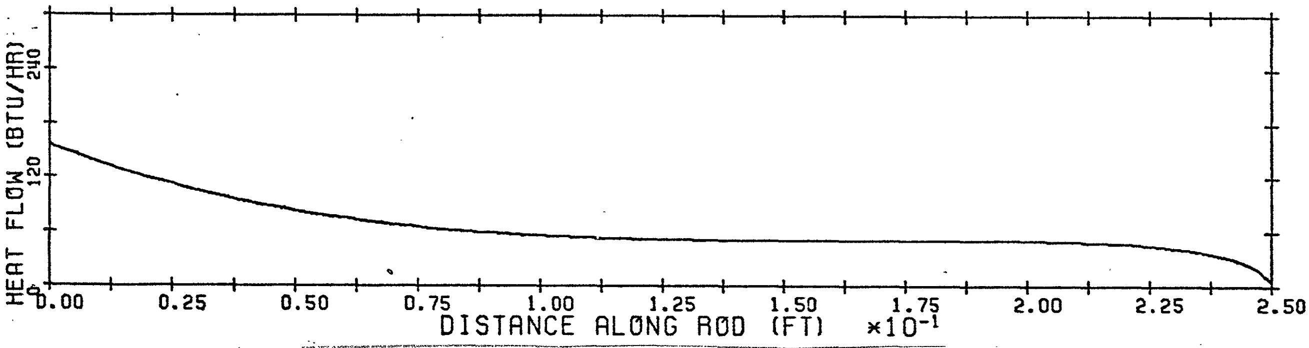
$$
\begin{bmatrix} (t_m - t_{amb}) + v_t(0) \\ (t_m - t_{amb}) + v_t(L) \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} t(0) \\ \dot{t}(0) \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} t(L) \\ \dot{t}(L) \end{bmatrix} \quad . \tag{6.8}
$$

Similar to the thin rod case, if $v_t(0)$ and $v_t(L)$ are uncorrelated, then (6.8) would represent a separable case. However, in many pin-fin configurations, the physical proximity of the two ends of the fin will result in the variations $v_t(0)$ and $v_t(L)$ being correlated. For example, consider the correlated case represented by

$$
\Pi_v = E\left\{ \begin{bmatrix} v_t(0) \\ v_t(L) \end{bmatrix} \left[v_t(0) \vdots v_t(L) \right] \right\} = \begin{bmatrix} \sigma_t^2 & \rho\sigma_t^2 \\ \rho\sigma_t^2 & \sigma_t^2 \end{bmatrix} \quad . \tag{6.9a}
$$

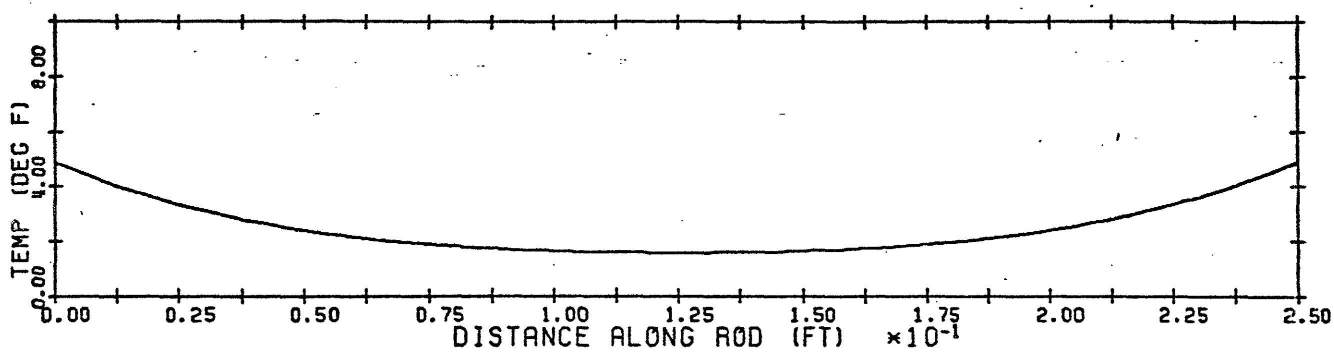


6.2(a)

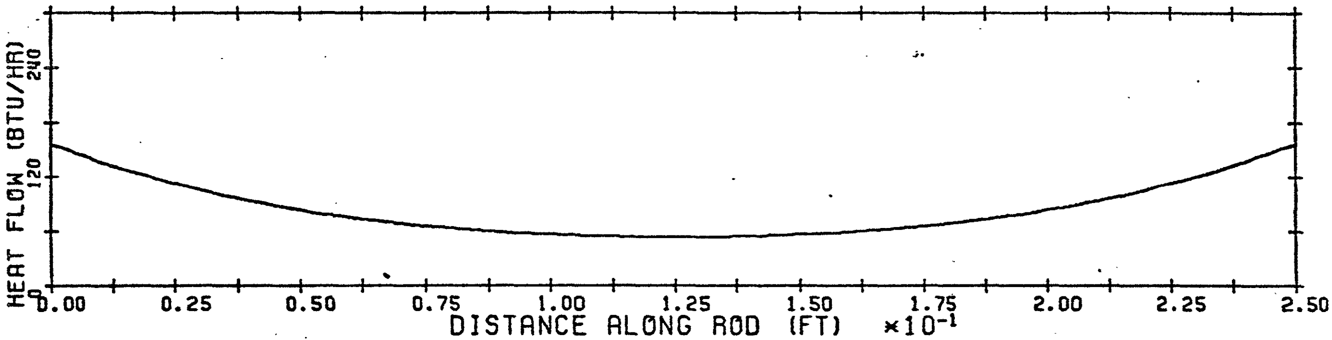


6.2(b)

Figure 6.2 Thin Rod Smoothing Error Standard Deviations: Example 1



6.3(a)



6.3(b)

Figure 6.3 Pin-Fin Smoothing Error Standard Deviations: Example 2, $\rho=0$

In this case due to the nonzero correlation ρ, θ_c is nonzero:

$$\theta_c = V^{T'} \Pi_v^{-1} V^0 = \left[\begin{array}{c:c} -\rho\sigma_t^{-2}(1-\rho^2)^{-1} & 0 \\ \hdashline 0 & 0 \end{array}\right] \tag{6.9b}$$

resulting in a nonseparable case.

Numerical Results

Error covariance results are presented for the three examples. The first is a thin rod case and the last two are pin-fin cases. For one pin-fin case the correlation ρ in (6.9) is assumed to be zero and for the other ρ is assumed nonzero. For all three examples we assume a 0.25 ft long copper rod with outer diameter 0.1 ft: $L = 0.25$ ft, $D_0 = 0.1$ ft and $k = 280$ Btu/(hr ft F). The coolant is water at 100 degrees F passing over the rod at a velocity of 5 ft/sec. These conditions correspond to a Reynolds number $R_e \simeq 6.75\times10^5$, a Prandtl number $P_r \simeq 4.52$ and a coefficient of heat transfer for the water of $k_w = 0.364$ Btu/(hr ft F). Applying an approximation from [23], the water-to-rod convective heat transfer coefficient is

$$h \simeq \frac{0.0263\ k_w R_e^{0.805}\ P_r^{0.31}}{D_0}$$

$$= 1180 \text{ Btu/(sq ft hr F)} \quad .$$

We will assume a process noise variance parameter $q = 1$ F/ft^2 and a measurement noise variance parameter $R = 1$ F/ft. Table 6.1 lists the uncertainties associated with the boundary conditions for the three examples.

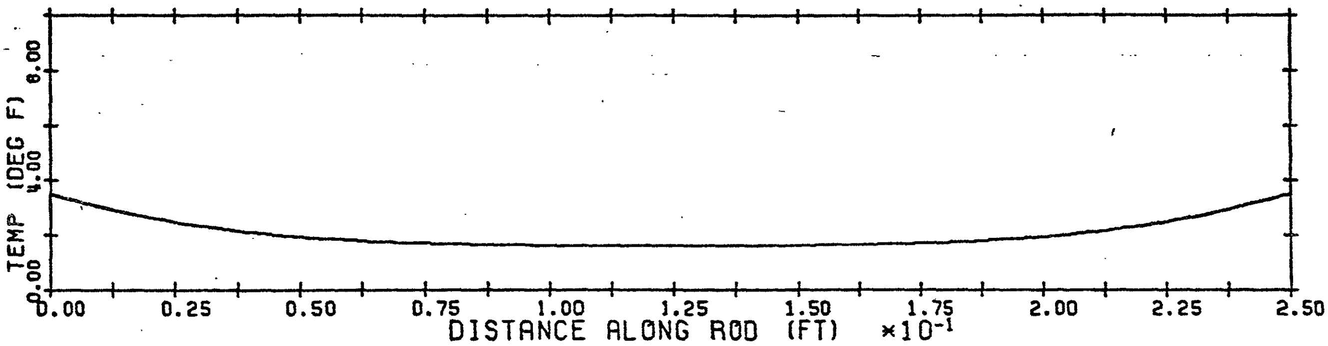
Example	$\sigma_t(0)$ (F)	$\sigma_t(L)$ (F)	$\sigma_q(L)$ (Btu/hr)	ρ
1. Thin rod	10.0	-	5.0	-
2. Pin-fin	10.0	10.0	-	0.0
3. Pin-fin	10.0	10.0	-	0.99

Plots of the results of the covariance analyses are presented in Figures 6.2, 6.3 and 6.4. Part a) of each figure shows the standard deviation in the smoothing error for temperature along the rod in degrees F. Part b) of each depicts the standard deviation of the heat flow in Btu/hr which has been calculated by scaling the uncertainty in $dt/d\ell$ as indicated in (6.5).

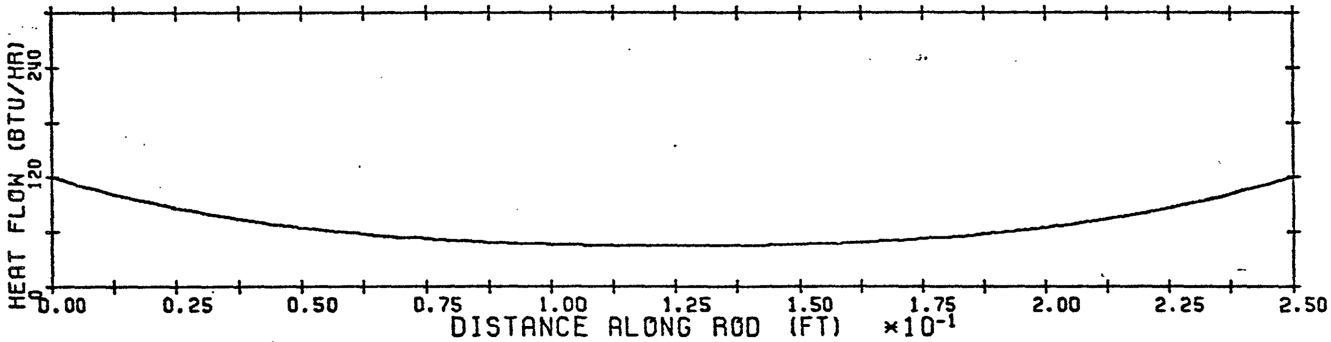
The results for the thin rod case in Figure 6.2 show that the heat flow uncertainty at the end of the rod, $\ell = 0.25$ ft, drops off to the boundary condition of 5 Btu/hr. In contrast, no such drop is seen for the pin-fin cases in Figures 6.3 and 6.4, for which the boundary condition is specified in terms of the temperature at both ends of the rod. Comparing between the pin-fin cases, we find that the highly correlated nonseparable case of example 3 has a larger reduction in uncertainty at the ends of the rod than does the separable case of example 2. In effect, the correlation allows the estimate at each end of the rod to utilize the information available at the opposite end. Comparing among all three examples, we find that the uncertainties at the midpoint of the rods, $\ell = 0.125$ ft, are about the same for all three cases. In fact, under the stabilizability and detectability conditions statedin Section 4, it can be shown for space-invariant cases and for very large smoothing intervals that the smoothing error covariance in the middle of the interval approaches

$$P_{ss} = \left[\Sigma^{0}_{f,ss} + \Sigma^{0}_{b,ss}\right]^{-1} ,$$

where ss denotes spatial steady-state values. Note that this expression for the steady-state error covariance is independent of both the structure and value of the smoother's boundary condition i.e. the steady-state covariance is the same for both causal and noncausal processes.



6.4(a)



6.4(b)

Figure 6.4 Pin-Fin Smoothing Error Standard Deviations: Example 3, $\rho=0.99$

SECTION 7

CONCLUSIONS

An internal differential realization of the fixed-interval smoother for a 1-D, n^{th} order noncausal two-point boundary value stochastic process (TPBVP) has been obtained by applying the method of complementary models developed in Part I [3], the companion to this paper. This representation for the TPBVP smoother has been shown to have the same $2n^{th}$ order Hamiltonian dynamics as the fixed-interval smoother for causal processes. The boundary condition for the TPBVP smoother, however, has been found to be more complex than that for the causal process smoother. By applying a time-varying diagonalizing transformation much like those employed by Kailath and Ljung [11] for causal processes, we have formulated a numerically stable n^{th} order two-filter implementation. The simplicity of this two-filter form is achieved by employing an information form for the diagonalizing transformation with carefully chosen boundary conditions for the differential equations governing its elements. The significant difference between our two-filter implementation and that for causal processes is that in the noncausal case the smoothed estimate at a given point in the interval is a noncausal function of each of the forward and backward processes (see (4.11) and (4.14)).

Our work in Part I has also provided a recipe for writing a differential realization for the smoothing error. Through an application of the same diagonalizing transformation, we have derived a two-filter representation for the smoothing error as well. From this representation, we have formulated an expression for the error covariance which is a function of the solutions of forward and backward Riccati equations (as in the causal process case) along with the solution of one additional matrix differential equation.

We have also discussed the application of the TPBVP smoother to a special class of noncausal processes which we refer to as separable, following the terminology introduced by Krener [13]. We have shown that separability can be interpreted in terms of the information contained in the two-point boundary condition v in (3.3b) and the boundary observation y_b in (3.2b). In particular, if the part of this information which pertains to the value of the process at the beginning of the smoothing interval, $x(0)$, is uncorrelated with the information about the process value at the end of the interval, $x(T)$, then the system is separable. The smoother for this class of systems is shown to be equivalent to a special form of a previously derived smoother for causal processes with "post-flight" measurements [17].

As discussed in Part I, differential realizations for estimators of both discrete and continuous parameter multidimensional stochastic processes can be formulated as well by the method of complementary models. As yet, the problems associated with the implementation of these estimators have not been completely solved, and we are currently pursuing answers to some of them.

REFERENCES

[1] Kailath, T.K., "A View of Three Decades of Linear Filtering Theory," IEEE Trans. Inf. Th., IT-20, No. 2, pp. 146-181, Mar. 1974.

[2] Wall, J., et al., "On the Fixed-Interval Smoothing Problem," Stochastics, V. 5, pp. 1-41, 1981.

[3] Adams, M.B., Willsky, A.S. and Levy, B.C., "Linear Estimation of Boundary Value Stochastic Processes, Part I: The Role and Construction of Complementary Models," submitted to IEEE Trans. Aut. Cont.

[4] Weinert, H.L. and Desai, U.B., "On Complementary Models and Fixed-Interval Smoothing," IEEE Trans. Aut. Cont., V. AC-26, Aug. 1981

[5] Krener, A.J., "Boundary Value Linear Systems," Asterique, V. 75-76, pp. 149-165, 1980.

[6] Krener, A.J., "Acausal Linear Systems," Proceedings 18th IEEE CDC, Ft. Lauderdale, 1981.

[7] Krener, A.J., "Smoothing of Stationary Cyclic Processes," Proceedings, Conference on Mathematical Theory of Networks and Systems, Santa Monica, CA, 1981.

[8] Greenberg, M.D., Applications of Green's Functions in Science and Engineering, Prentice-Hall, Englewood Cliffs, NJ, 1976

[9] Fraser, D.C., "A New Technique for Optimal Smoothing of Data," Ph.D. Dissertation, Dept. of Aero. and Astro., MIT, 1967.

[10] Mayne, D.Q., "A Solution of the Smoothing Problem for Linear Dynamical Systems," Automatica, V. 4, pp. 73-92, 1966

[11] Kailath, T. and Ljung, L., "Two-Filter Smoothing Formulas by Diagonalization of the Hamiltonian Equations," submitted to Int. J. of Cont.

[12] Desai, U.B., "Congruence Relations Among Quadratic Minimization Problems," Ph.D. Dissertation, Electrical Engineering Dept., The Johns Hopkins University, Sept. 1979.

[13] Krener, A.J., Lecture, Laboratory for Signals and Systems, Gif-Sur-Yvette, France, Jan. 1981.

[14] Hildebrandt, F.B., Advanced Calculus for Applications, Prentice-Hall, Englewood Cliffs, NJ, 1976.

[15] Myers, G.E., Analytical Methods in Conduction and Heat Transfer, McGraw-Hill Book Co., NY, 1971.

[16] Coddington, E.A. and Levinson, N., Theory of Ordinary Differential Equations, McGraw-Hill Book Co., NY, 1955.

[17] Bryson, A.E. and Hall, W.E., "Optimal Control and Filtering Synthesis by Eigenvector Decomposition," SUDAAR Rept. 436, Stanford University, Dept. of Aero. and Astro., Nov. 1971.

[18] Rozonoer, L.I., "L.S. Pontryagin's Maximum Principle in the Theory of Optimum Systems," Avtomatika i Telemekhanika, V. 20, No. 12, Dec. 1959, pp. 1561-1578.

[19] Kwakernaak, H. and Sivan, R., Linear Optimal Control Systems, John Wiley and Sons, NY, 1972.

[20] Bryson, A.E. and Frazier, M., "Smoothing for Linear and Nonlinear Dynamic Systems," TDR 63-119, Aero. Sys. Div., Wright-Patterson AFB, pp. 353-364, 1964.

[21] Adams, M.B., "Estimation of Noncausal Processes in One and Two Dimensions," Doctoral Thesis Proposal, C.S. Draper Lab. Rept. CSDL-T-770, Feb. 1982.

[22] Schweppe, F.C., Uncertain Dynamic Systems, Prentice-Hall, Englewood Cliffs, NJ, 1973.

[23] Kreith, H., Principles of Heat Transfer, Third Edition, Intext Press, Inc., NY, 1973.