SMOOTHING FOR TIME-VARYING SYSTEMS USING MEASUREMENTS CONTAINING COLORED NOISE

By

R. K. Mehra and A. E. Bryson, Jr.

Technical Report No. 1

June 1967

Prepared under Grant NGR 22-007-068
Division of Engineering and Applied Physics
Harvard University Cambridge, Massachusetts

for

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

SMOOTHING FOR LINEAR TIME-VARYING SYSTEMS USING MEASUREMENTS CONTAINING COLORED NOISE

by

R. K. Mehra and A. E. Bryson, Jr.

Division of Engineering and Applied Physics
Harvard University
Cambridge, Mass.

Abstract

Kalman and Bucy (1961) derived the optimal filter for continuous linear dynamic systems where all measurements contain "white noise," i.e., noise with correlation times short compared to times of interest in the system. Bryson and Frazier (1962) described the corresponding optimal smoother. ("Filtering" involves making an estimate at a time t_1 using measurements made before t_1; "smoothing" involves making an estimate at a time t_1 using measurements made both before and after t_1.)

Bryson and Johansen (1965) described the optimal filter for the case where some measurements contained either no noise or "colored noise," i. e., noise with correlation times comparable to or larger than times of interest in the system. The present paper describes the optimal smoother for the case where some measurements contain either no noise or "colored noise." The problem is formulated as a calculus of variations problem with equality constraints, and is solved using the "sweep method" of McReynolds and Bryson (1965). The case of a single measurement containing colored noise is treated first. Then the general case and a simple example are described.

1. INTRODUCTION

The Kalman-Bucy filter [1] and the Bryson-Frazier smoother [2] presume that all measurements contain white noise. If some or all of the measurements contain time-correlated noise, or no noise at all, the correlation matrix of noise in the measurements (usually denoted as R) is singular. Since the inverse of R appears explicitly in the filtering and smoothing equations of Refs. [1] and [2], they cannot be used if R is singular. By augmenting the state of the system, the problem of time-correlated (or "colored") noise in the measurements can be reduced to a problem in which some of the measurements contain white noise and others are perfect.

The optimal filter for continuous linear dynamic systems using measurements containing time-correlated noise was described by Bryson and Johansen [3]. Unlike their approach, the present approach does not rely on any previous results, and, furthermore, it leads naturally to the smoothing results.

To elucidate the method, we shall first consider the case of a single measurement containing time-correlated noise. After combining measurement dynamics with system dynamics and augmenting the state of the system, this case reduces to the case of a system with a scalar perfect measurement.

2. STATEMENT OF THE PROBLEM

Consider a continuous linear dynamic system whose augmented state is represented by the n component state vector x.

Let the differential equations obtained by combining measurement dynamics with system dynamics be

$$\dot{x} = Fx + \Gamma u \qquad t_o \leq t \leq T, \tag{1}$$

where

$F(t) = n \times n$ matrix of functions

$\Gamma(t) = n \times r$ matrix of functions

$u(t) = r \times 1$ vector of random forcing functions (white noise*)

The mean value and correlation of u are given as

$$E\{u(t)\} = 0; \qquad E\{u(t)u^T(\tau)\} = Q(t)\,\delta(t-\tau),$$

where

$Q(t) = r \times r$ non-negative definite matrix

$E\{\ \}$ = expected value operator.

The initial conditions are also random with mean and covariance given as

---

*Note: If u is not white noise, we can construct a shaping filter for it and augment the state of the system.

$$E\{x(t_o)\} = 0; \qquad E\{x(t_o)x^T(t_o)\} = P(t_o).$$

It is assumed that u(t) is independent of $x(t_o)$.

$$E\{x(t_o)u^T(t)\} = 0 \qquad \text{for } t_o \leq t \leq T.$$

Let z(t) denote the scalar measurement made on the system continuously from t_o to T. It is linearly related to the state of the system as

$$z(t) = h^T x \qquad t_o \leq t \leq T, \tag{2}$$

where

$h(t) = n \times 1$ vector of functions.

The problem consists in finding the maximum likelihood estimates of $x(t_o)$, x(t), and u(t) for $t_o \leq t \leq T$ using $\{z(t), t_o \leq t \leq T\}$.

3. FORMULATION OF THE PROBLEM

Since x(t) is a Gauss-Markov random process, the minimum variance estimate, the maximum likelihood estimate and the min-max estimate are all equal [4]. Hence let us consider the maximum likelihood estimate in which we try to maximize the probability of x and u, given equations of motion (1) and the set of measurements (2).

The problem can be stated as follows.* Find $x(t_o)$ and u(t) to minimize

---

*Note: From here on x and u will denote smoothed estimates of the corresponding random variables. This simplifies the terminology. In the literature, it is common to denote these as $\hat{x}(t/T)$ and $\hat{u}(t/T)$.

$$J = \frac{1}{2} x^T(t_o) P^{-1}(t_o) x(t_o) + \frac{1}{2} \int_{t_o}^{T} u^T(t) Q^{-1}(t) u(t) dt \qquad (3)$$

subject to the constraints

$$\dot{x} = Fx + \Gamma u \qquad t_o \leq t \leq T$$

$$z = h^T x$$

where $z(t)$ is given. Note that e^{-J} is proportional to the joint probability density of $x(t_o)$ and $\{u(t), t_o \leq t \leq T\}$. Hence minimizing J subject to the measurement constraint is equivalent to maximizing the joint probability density function of $x(t_o)$ and $\{u(t), t_o \leq t \leq T\}$, given $\{z(t), t_o \leq t \leq T\}$. Once $x(t_o)$ and $u(t)$ are known, the state $x(t)$ for $t_o \leq t \leq T$ is determined from dynamic equations (1).

4. REDUCTION TO A STANDARD CALCULUS OF VARIATIONS PROBLEM

The smoothing problem has been formulated as a calculus of variations problem involving equality constraints on the state variables alone. One way to solve this problem, as outlined in [6], is to differentiate the equality constraint until it contains one or more components of u explicitly. In this way, the state variable equality constraints are converted into mixed equality constraints involving both the state and the control variables. Following the same method here, we differentiate the measurement and substitute for x from the dynamic equations.

$$z = h^T x$$

$$\dot{z} = \dot{h}^T x + h^T \dot{x}$$

$$= (\dot{h}^T + h^T F)x + h^T \Gamma u.$$

The correlation of the additive noise in the expression for $\dot{z}$ is $h^T \Gamma Q \Gamma^T h \delta(t-\tau)$ and, if it is not identically zero over the interval $t_o \leq t \leq T$, then $\dot{z}$ may be regarded as a measurement with additive white noise and $z(t_o) = h^T(t_o)\, x(t_o)$ as an initial condition. This would be called a first order perfect measurement.

Note further that z is just a linear combination of the state variables and it can be made a component of the state vector in place of any other component by properly transforming the state. As this component (z) is known from measurements, the dimension of the state to be estimated is effectively reduced by one. If we denote by x_1 the new state vector which has only (n - 1) components, it would satisfy an equation of the type

$$\dot{x}_1 = F_{11} x_1 + F_{1z} z + G_1 u \qquad \text{(n-1) equations.}$$

Then the n^{th} state equation is a "measurement" equation

$$\dot{z} = (\dot{h}^T + h^T F)x + h^T \Gamma u,$$

where F_{11}, F_{1z}, G_1 are obtained from F and Γ by suitable transformations. (These transformations will be discussed later.)

In short, the following two sets of equations are completely equivalent:

Set 1		Set 2	
$\dot{x} = Fx + \Gamma u$	n equations	$\dot{x}_1 = F_{11}x_1 + F_{1z}z + G_1 u$	(n-1) equations
$z = h^T x$	1 equation	$\dot{z} = (\dot{h}^T + h^T F)x + h^T \Gamma u$	1 equation
	for $t_o \leq t \leq T$	$z(t_o) = h^T(t_o)\, x(t_o)$	1 equation

We shall consider equation set 2 because it will yield a lower-order filter and eliminate the problem of singularity of the R matrix in set 1. (R denotes the correlation matrix of the noise in the measurements. In set 1, it is clear that R = 0 because the scalar measurement does not contain any noise.)

If $h^T \Gamma Q \Gamma^T h \equiv 0$ over $t_o \leq t \leq T$, we must differentiate z again and substitute for $\dot{x}$. If p differentiations of z are required to involve u, we shall call z a p^{th} order perfect measurement.

Let

$$\frac{d^p z}{dt^p} = z^{(p)} = z_2 = \ell^T x + Du,$$

where $D^T Q D \neq 0$ for $t_o \leq t \leq T$, ℓ is an $n \times 1$ vector, and D is a $1 \times r$ vector of functions. Now $z^{(p)}(t)$ may be regarded as a measurement with additive white noise and $z(t_o)$, $z^{(1)}(t_o)$, ..., $z^{(p-1)}(t_o)$ as initial conditions.

Let

$$x_2(t_o) = \begin{bmatrix} z(t_o) \\ z^{(1)}(t_o) \\ \vdots \\ z^{(p-1)}(t_o) \end{bmatrix}$$

$x_2(t_o)$ is a $p \times 1$ vector of initial conditions.

x_2 is related to x by

$$x_2 = Cx$$

where C is a $p \times n$ matrix of functions obtained via the differentiation process.

$$C = \begin{bmatrix} h^T \\ \dot{h}^T + h^T F \\ \ddot{h}^T + \dot{h}^T F + h^T \dot{F} + \dot{h}^T F + h^T F^2 \\ \vdots \end{bmatrix}$$

In this way we get the following new constraining equations:

$$z^{(p)} = z_2 = \ell^T x + D^T u \qquad \text{1 equation} \tag{4}$$

$$x_2(t_o) = C(t_o)\, x(t_o) \qquad \text{p equations} \tag{5}$$

In the case of a first order perfect measurement the dimension of the state vector was reduced by one. Similarly, we can reduce the dimension of the state vector by p in the case of a p^{th} order perfect measurement. To do this, we transform the state x of the system by a matrix M such that the new state vector has x_2 as part of its state representation. Let x_1 denote the remaining (n-p) components of the state vector.

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = Mx = \begin{bmatrix} M_1 \\ \hline M_2 \end{bmatrix} x. \tag{6}$$

M is $n \times n$; M_1 is $(n-p) \times n$; and M_2 is $p \times n$.

The choice of M is arbitrary to a certain extent, but one obvious choice for M is

$$M = \left[\begin{array}{c|c} I & 0 \\ \hline C_1 & C_2 \end{array}\right] = \begin{bmatrix} M_1 \\ \hline M_2 \end{bmatrix},$$

where C_1 and C_2 are obtained by partitioning matrix C $(p \times n)$ along the $(n-p)^{th}$ column so that C_1 is $p \times (n-p)$ and C_2 is $p \times p$.

$$C = [C_1 \mid C_2].$$

Moreover, C_2 is nonsingular. For the transformation (6) to be unique or one-to-one, M must be nonsingular and our present choice for M insures that M is nonsingular as long as C_2 is nonsingular. In fact,

$$M^{-1} = \left[\begin{array}{c|c} I & 0 \\ \hline -C_2^{-1}C_1 & C_2^{-1} \end{array}\right]$$

obtained from a formula given in Bodewig [5]. Differentiating (6),

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \dot{M}x + M\dot{x} = \dot{M}x + MFx + M\Gamma u$$

$$= (\dot{M}M^{-1} + MFM^{-1}) \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + M\Gamma u$$

$$\dot{x}_1 = F_{11}x_1 + F_{12}x_2 + G_1u \qquad (n-p) \text{ equations} \tag{7}$$

$$\dot{x}_2 = F_{21}x_1 + F_{22}x_2 + G_2u \qquad p \text{ equations} \tag{8}$$

where

$$\begin{bmatrix} F_{11} & F_{12} \\ F_{21} & F_{22} \end{bmatrix} = \dot{M}M^{-1} + MFM^{-1}$$

$$\begin{bmatrix} G_1 \\ G_2 \end{bmatrix} = M\Gamma.$$

F_{11} is $(n-p) \times (n-p)$; F_{22} is $p \times p$;

F_{12} is $(n-p) \times p$; G is $(n-p) \times r$;

F_{21} is $p \times (n-p)$; G_2 is $p \times r$.

p^{th} order measurement equation (4) becomes

$$z_2 = z^{(p)} = \ell^T M^{-1} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + Du.$$

Let

$$\ell^T M^{-1} = [H_1 \mid H_2],$$

where H_1 is a $1 \times (n-p)$ row vector and H_2 is a $1 \times p$ row vector.

$$z_2 = H_1 x_1 + H_2 x_2 + Du \qquad t_o \leq t \leq T \tag{9}$$

Also

$$x_2(t_o) = C(t_o)\, x(t_o). \tag{10}$$

p equations (8) are contained in equations (9) and (10) just as in the case of a first order perfect measurement. So (8) need not be considered while considering (7), (9), and (10). Thus we have reduced the dimension of the state by p as x_1 is only $(n-p) \times 1$. x_2 is completely determined from the measurements (in practice, by actual differentiation).

5. THE TWO POINT BOUNDARY VALUE PROBLEM

The problem may now be restated as follows: Find $x(t_o)$ and $u(t)$ to minimize

$$J = \frac{1}{2} x^T(t_o)\, P^{-1}(t_o)\, x(t_o) + \frac{1}{2} \int_{t_o}^{T} u^T(t)\, Q^{-1}(t)\, u(t)\, dt$$

subject to

$$\dot{x}_1 = F_{11}x_1 + F_{12}x_2 + G_1u \qquad t_o \leq t \leq T \qquad \text{(n-p) equations} \qquad (7)$$

$$z_2 = H_1x_1 + H_2x_2 + Du \qquad \text{1 equation} \qquad (9)$$

and

$$x_2(t_o) = C(t_o)\, x(t_o) \qquad \text{p equations} \qquad (10)$$

Note that the last two equations together imply

$$x_2(t) = C(t)\, x(t) \qquad \text{for } t_o \leq t \leq T.$$

These constraint equations may be adjoined to the criterion function J using undetermined multipliers:

$\lambda_1(t)$ = (n-p)$\times$1 vector of multiplier functions for (7)

$\mu(t)$ = a scalar multiplier functions for (9)

ν = p $\times$ 1 vector of multiplier constants for (10)

$$J = \frac{1}{2} x^T(t_o)\, P^{-1}(t_o)x(t_o) + \nu^T(x_2(t_o) - C(t_o)x(t_o))$$

$$+ \frac{1}{2} \int_{t_o}^{T} \{u^TQ^{-1}u + \lambda^T(-\dot{x}_1 + F_{11}x_1 + F_{12}x_2 + G_1u) + \mu(z_2 - H_1x_1 - H_2x_2 - Du)\}\, dt.$$

Considering variations in J due to infinitesimal variations in $\{u(t), t_o \leq t \leq T\}$ and $x(t_o)$ and the resulting variations in x_1, we get (notice that the variations in x_2 and z_2 are zero)

$$\delta J = x^T(t_o)\, P^{-1}(t_o)\, \delta x(t_o) + \nu^T(-C(t_o)\delta x(t_o))$$

$$+ \int_{t_o}^{T} \{u^T Q^{-1} \delta u + \lambda_1^T(-\delta\dot{x}_1 + F_{11}\delta x_1 + G_1 \delta u) + \mu(-H_1 \delta x_1 - D\delta u)\}\, dt.$$

Integrating $\lambda_1^T \delta\dot{x}_1$ by parts

$$\delta J = [x^T(t_o)\, P^{-1}(t_o) - \nu^T C(t_o)]\, \delta x(t_o) - \lambda_1^T \delta x_1 \Big|_{t_o}^{T}$$

$$+ \int_{t_o}^{T} \{(u^T Q^{-1} + \lambda_1^T G_1 - \mu D)\delta u + (\dot{\lambda}_1^T + \lambda_1^T F_{11} - \mu H_1)\delta x_1\}\, dt.$$

But

$$x_1 = M_1 x$$

or

$$\delta x_1 = M_1 \delta x.$$

Substituting this above for $\delta x_1(t_o)$ and choosing multipliers ν, μ, and λ so that the coefficients of $\delta x(t_o)$ and $\delta x_1(t)$, $t_o \leq t \leq T$ vanish, we obtain

$$x^T(t_o)\, P^{-1}(t_o) - \nu^T C(t_o) + \lambda_1^T(t_o)\, M_1(t_o) = 0 \tag{11}$$

$$\lambda_1^T(T) = 0$$

$$\dot{\lambda}_1^T = -\lambda_1^T F_{11} + \mu H_1$$

$$\delta J = \int_{t_o}^{T} (u^T Q^{-1} + \lambda_1^T G_1 - \mu D)\; \delta u\;\; dt.$$

For stationarity, i.e., $\delta J = 0$ for arbitrary $\delta u(t)$, it follows that

$$u^T Q^{-1} + \lambda_1^T G_1 - \mu D = 0$$

or

$$u = Q\left(D^T \mu - G_1^T \lambda_1\right). \qquad (12)$$

Thus the <u>two point boundary value problem</u> is (substituting for u in x_1 and λ_1 equations)

$$\dot{x}_1 = F_{11} x_1 + F_{12} x_2 + G_1 Q\left(D^T \mu - G_1^T \lambda_1\right) \qquad (13)$$

$$\dot{\lambda}_1 = -F_{11}^T \lambda_1 + H_1^T \mu. \qquad (14)$$

From (11)

$$x(t_o) = P(t_o)\left[C^T \nu - M_1^T \lambda_1\right]_{t=t_o} \qquad (15)$$

or

$$x_1(t_o) = M_1(t_o)\, P(t_o)\left[C^T \nu - M_1^T \lambda_1\right]_{t=t_o}$$

$$\lambda_1(T) = 0,$$

where $\mu(t)$ and ν are determined from

$$z_2 = H_1 x_1 + H_2 x_2 + Du \qquad t_o \leqslant t \leqslant T \qquad (9)$$

$$x_2(t_o) = C(t_o)\, x(t_o). \qquad (10)$$

6. SOLUTION OF THE TWO POINT BOUNDARY VALUE PROBLEM

We can eliminate μ and ν using (9) and (10) due to linearity of the problem. Substituting in (10) from (15)

$$x_2(t_o) = CP\left[C^T\nu - M_1^T\lambda_1\right]_{t=t_o}$$

or

$$\nu = (CPC^T)^{-1}\left[x_2 + CPM_1^T\lambda_1\right]_{t=t_o}.$$

Putting ν back in (15)

$$x(t_o) = \left[PC^T(CPC^T)^{-1}\left(x_2 + CPM_1^T\lambda_1\right) - PM_1^T\lambda_1\right]_{t=t_o}$$

$$x_1(t_o) = M_1(t_o)\left\{PC^T(CPC^T)^{-1}x_2 - [P - PC^T(CPC^T)^{-1}CP]\,M_1^T\lambda_1\right\}_{t=t_o}.$$

Let

$$\hat{x}_1(t_o+) = \left(M_1PC^T(CPC^T)^{-1}x_2\right)_{t=t_o} \tag{16}$$

$$P_1(t_o+) = \left\{M_1[P - PC^T(CPC^T)^{-1}CP]\,M_1^T\right\}_{t=t_o}. \tag{17}$$

The reason for this particular notation will be explained shortly. Then

$$x_1(t_o) = \hat{x}_1(t_o+) - P_1(t_o+)\,\lambda_1(t_o). \tag{18}$$

Note that

$$x_2(t_o) = C(t_o)\,x(t_o)$$

$$= \left\{CPC^T(CPC^T)^{-1}x_2 - [CP - CPC^T(CPC^T)^{-1}]\,M_1^T\lambda_1\right\}_{t=t_o}$$

$$= x_2(t_o)$$

and

$$P(t_o+) = \{P - PC^T(CPC^T)^{-1} CP\}_{t=t_o}$$

$$P_2(t_o+) = E\left\{x_2 x_2^T\right\}_{t_o+} = (CPC^T)_{t=t_o} = 0.$$

As pointed out in [3], there is a simple explanation for these results in terms of the single stage estimation theory. As soon as the measurements $x_2(t_o)$ become available at (t_o+), we update our estimate of the state to $\hat{x}(t_o+)$ and the covariance matrix is correspondingly updated to $P(t_o+)$. Thus there are discontinuities in the state of the optimal filter at time $t = t_o$.

Now let us eliminate μ from equations (13) and (14) using (9). Substituting u from (12) in (9),

$$z_2 = H_1 x_1 + H_2 x_2 + DQ\left(D^T \mu - G_1^T \lambda_1\right)$$

or

$$\mu = (DQD^T)^{-1} \left(z_2 - H_1 x_1 - H_2 x_2 + DQG_1^T \lambda_1\right) \tag{19}$$

Putting this in (13), (14), we get

$$\begin{bmatrix} \dot{x}_1 \\ \dot{\lambda}_1 \end{bmatrix} = \begin{bmatrix} F_{11} - G_1 Q D^T R^{-1} H_1 & -G_1 Q G_1^T + G_1 Q D^T R^{-1} D Q G_1^T \\ -H_1^T R^{-1} H_1 & -F_{11}^T + H_1^T R^{-1} D Q G_1^T \end{bmatrix} \begin{bmatrix} x_1 \\ \lambda_1 \end{bmatrix}$$

$$+ \begin{bmatrix} F_{12} x_2 + G_1 Q D^T R^{-1} (z_2 - H_2 x_2) \\ H_1^T (R)^{-1} (z_2 - H_2 x_2) \end{bmatrix} \tag{20}$$

where $R = DQD^T$ is a scalar and denotes the covariance of noise in z_2.

There are various ways of solving linear two-point boundary value problems. The transition matrix $\Phi(t, t_o)$ can be calculated for this system.

Let

$$\Phi(t, t_o) = \begin{bmatrix} \Phi_{x_1x_1} & \Phi_{x_1\lambda_1} \\ \Phi_{\lambda_1x_1} & \Phi_{\lambda_1\lambda_1} \end{bmatrix}$$

Then

$$\lambda(T) = 0 = \Phi_{\lambda_1x_1}(T,t_o)\,x_1(t_o) + \Phi_{\lambda_1\lambda_1}(T,t_o)[P^{-1}(\hat{x}_1(t_o+)-x_1(t_o))]$$

$$+ \int_{t_o}^{T} \Big\{\Phi_{\lambda_1x_1}(T, t)[F_{12}x_2+G_1QD^TR^{-1}(z_2-H_2x_2)]$$

$$+ \Phi_{\lambda_1\lambda_1}(T, t)[H_1^TR^{-1}(z_2-H_2x_2)]\Big\}\, dt.$$

Solution of this gives $x_1(t_o)$. Thus $x_1(t_o)$ and $\lambda_1(t_o)$ are known and we can solve the smoothing problem.

However, it is much more interesting to solve these equations by sweep method as this also gives us the filtering results. In this method we effectively sweep the boundary conditions from one end to the other [2].

Let

$$x_1(t) = \hat{x}_1(t) - P_1(t)\,\lambda_1(t). \tag{21}$$

(This form is suggested by the boundary condition at t_o.) Differentiating and substituting from (20)

$$(F_{11}-G_1QD^TR^{-1}H_1)\hat{x}_1 + F_{12}x_2 + G_1QD^TR^{-1}(z_2-H_2x_2) - \dot{\hat{x}}_1 - P_1H_1^TR^{-1}H_1\hat{x}_1$$

$$+ P_1H_1^T(R)^{-1}(z_2-H_2x_2)$$

$$= \Big[-\dot{P}_1+P_1F_{11}^T-P_1H_1^TR^{-1}DQG_1^T+G_1QG_1^T-G_1QD^TR^{-1}DQG_1^T+F_{11}P_1$$

$$-G_1QD^TR^{-1}H_1P_1-P_1H_1^TR^{-1}H_1P_1\Big]\lambda_1.$$

Setting the coefficient of λ_1 equal to zero,

$$\dot{P}_1 = P_1F_{11}^T + F_{11}P_1 + G_1QG_1^T - (P_1H_1^T+G_1QD^T)\,R^{-1}(P_1H_1^T+G_1QD^T)^T.$$

Let

$$K = (P_1H_1^T+G_1QD^T)\,R^{-1} \tag{22}$$

$$\boxed{\dot{P}_1 = P_1F_{11}^T + F_{11}P_1 + G_1QG_1^T - KRK^T} \tag{23}$$

$$\boxed{\dot{\hat{x}}_1 = F_{11}\hat{x}_1 + F_{12}x_2 - KH_1\hat{x}_1 + K(z_2-H_2x_2)} \tag{24}$$

(22-24) are the equations of the optimal filter for the case of perfect measurements. In practice, we cannot obtain z_2 by differentiating x_2 as z_2 contains white noise. So we should eliminate z_2 from these expressions.

$$z_2 = z^{(p)} = \frac{d}{dt}\,z^{(p-1)}.$$

Let $\quad x_1^* = \hat{x}_1 - Kz^{(p-1)}$

$$\dot{x}_1^* = \dot{\hat{x}}_1 - \dot{K}z^{(p-1)} - Kz_2.$$

Substituting for $\dot{\hat{x}}_1$,

$$\dot{x}_1^* = F_{11}\hat{x}_1 + F_{12}x_2 - KH_1\hat{x}_1 - \dot{K}z^{(p-1)} - KH_2x_2$$

$$\boxed{\dot{x}_1^* = (F_{11}-KH_1)\,\hat{x}_1 + (F_{12}-KH_2)\,x_2 - \dot{K}z^{(p-1)}.} \tag{25}$$

This can be implemented as shown in Fig. 1.

Equations (22), (23), (24) look very much like the Kalman Filtering Equations, but they are of lower order.

Using the filtering equations, we can get $\hat{x}_1(T)$ which is the same as the smoothed estimate $x_1(T)$ because it uses all the measurements $z(t)$, $t_o \leq t \leq T$. Knowing $x_1(T)$ and $\lambda_1(T) = 0$, we can integrate (20) backwards for $x_1(t)$ and $\lambda_1(t)$. Then $u(t)$ is calculated from (12). The procedure mentioned above can be used to eliminate z_2 from the smoothing equations also.

Let

$$\begin{bmatrix} x_1' \\ \lambda_1' \end{bmatrix} = \begin{bmatrix} x_1 \\ \lambda_1 \end{bmatrix} - \begin{bmatrix} G_1QD^TR^{-1} \\ H_1^TR^{-1} \end{bmatrix} z^{(p-1)}$$

$$\begin{bmatrix} \dot{x}_1' \\ \dot{\lambda}_1' \end{bmatrix} = \begin{bmatrix} F_{11} - G_1QD^TR^{-1}H_1 & -G_1QG_1^T + G_1QD^TR^{-1}DQG_1^T \\ -H_1^TR^{-1}H_1 & -F_{11}^T + H_1^TR^{-1}DQG_1^T \end{bmatrix} \begin{bmatrix} x_1 \\ \lambda_1 \end{bmatrix}$$

$$- \begin{bmatrix} \dot{G}_1QD^TR^{-1} + G_1\dot{Q}D^TR^{-1} - G_1QD^TR^{-1}\dot{R}R^{-1} \\ \dot{H}_1^TR^{-1} - H_1^TR^{-1}\dot{R}R^{-1} \end{bmatrix} z^{(p-1)}$$

$$+ \begin{bmatrix} F_{12} - G_1QD^TR^{-1}H_2 \\ -H_1^TR^{-1}H_2 \end{bmatrix} x_2$$

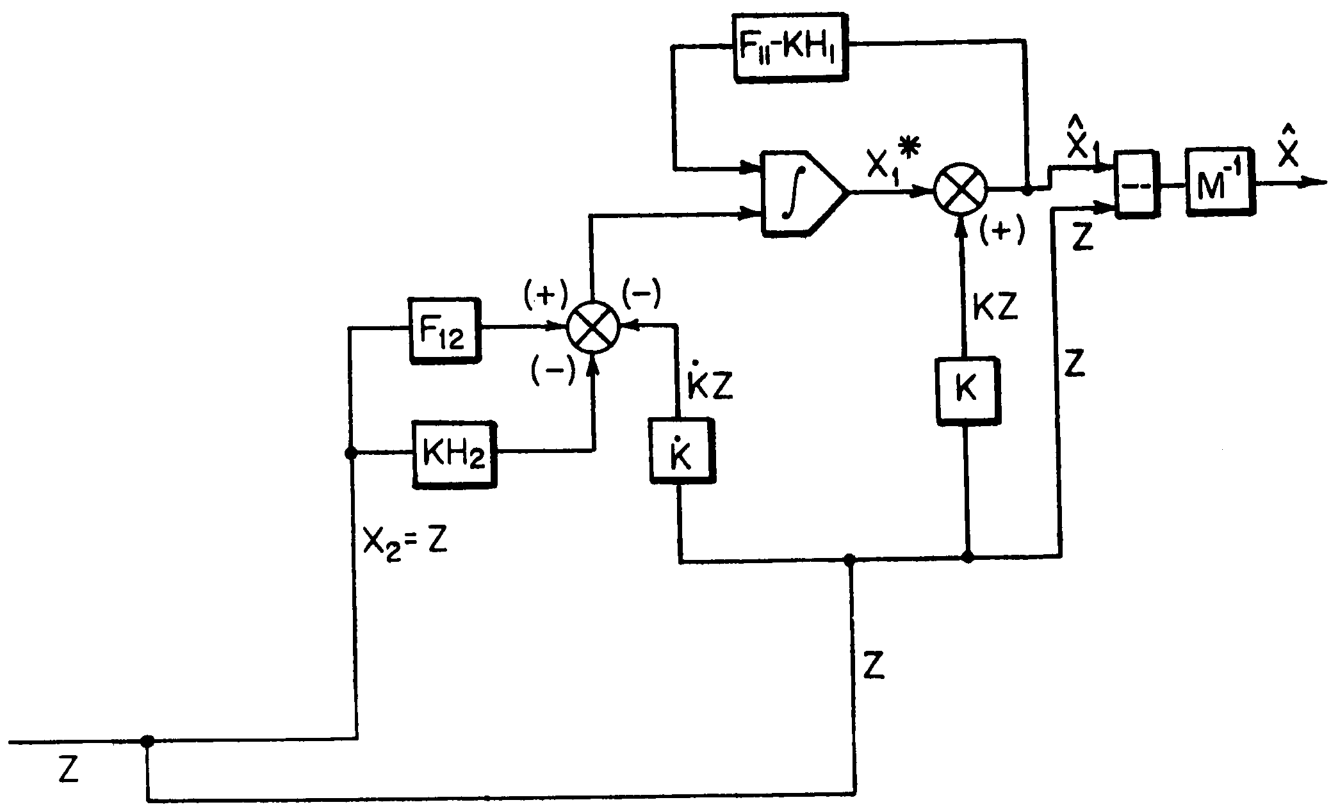

FIG. 1 OPTIMAL FILTER FOR THE CASE OF A FIRST ORDER PERFECT MEASUREMENT

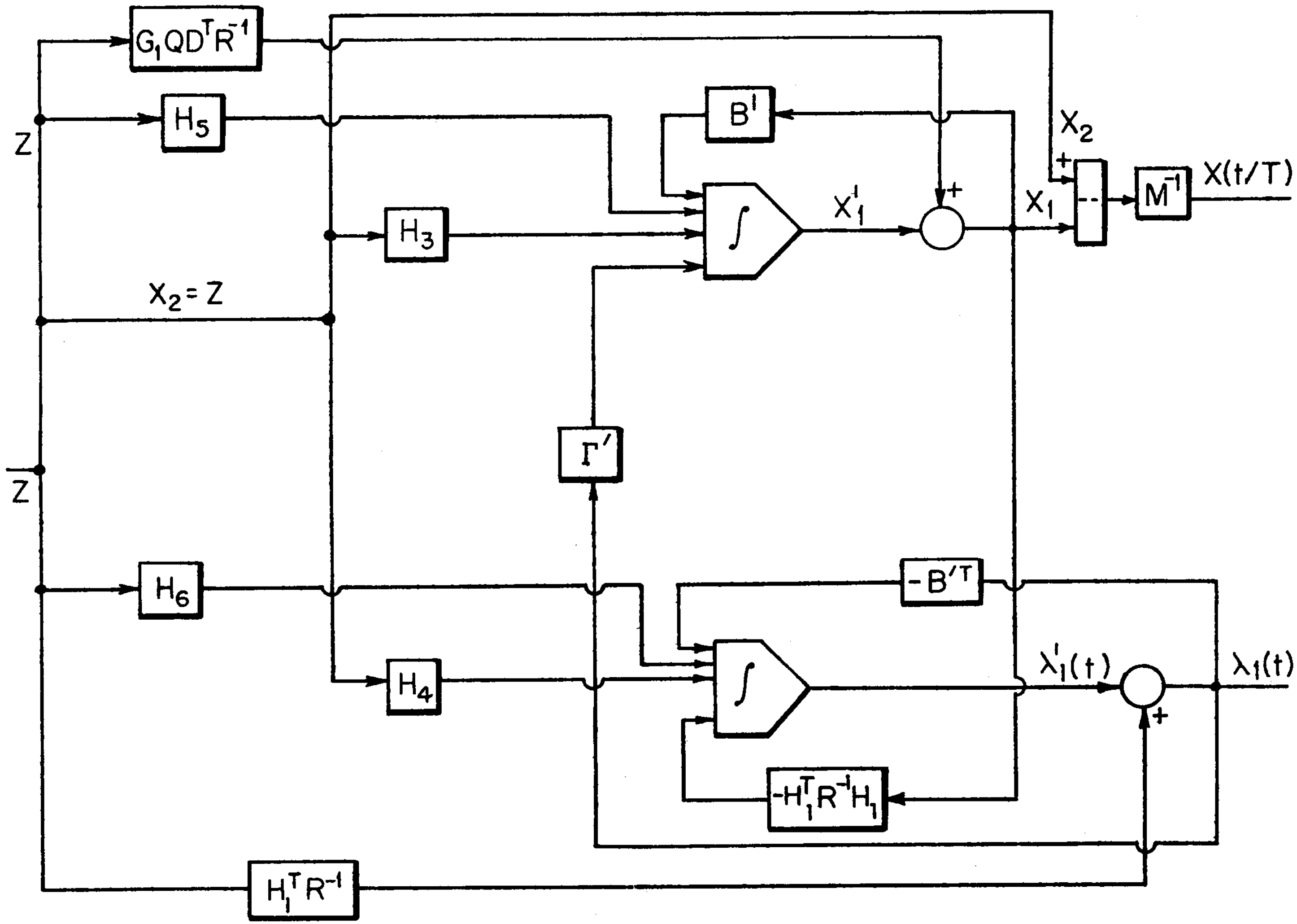

FIG. 2 OPTIMAL SMOOTHER FOR THE CASE OF A FIRST ORDER PERFECT MEASUREMENT

or
$$\begin{bmatrix} \dot{x}_1' \\ \dot{\lambda}_1' \end{bmatrix} = \begin{bmatrix} B' & \Gamma' \\ -H_1^T R^{-1} H_1 & -B'^T \end{bmatrix} \begin{bmatrix} x_1 \\ \lambda_1 \end{bmatrix} + \begin{bmatrix} H_3 \\ H_4 \end{bmatrix} x_2 + \begin{bmatrix} H_5 \\ H_6 \end{bmatrix} z^{(p-1)}, \qquad (26)$$

where B', Γ', H_3, H_4, H_5, H_6 are easily defined by correspondence. The estimate of u is calculated from (12) which, on substitution of μ, gives

$$u = Q[D^T(DQD^T)^{-1}(z_2 - H_1x_1 - H_2x_2 + DQG_1^T\lambda_1) - G_1^T\lambda_1].$$

Equation (26) can be implemented as shown on Fig. 2.

7. ERROR COVARIANCE MATRIX OF SMOOTHED ESTIMATES

In this section, we shall derive the analogs of Bryson-Frazier [2], and Rauch-Tung and Striebel [7], for the case of time-correlated noise in the measurements.

Let

e_s = error in smoothed estimates

e_f = error in filtered estimates.

Let

$$E\left\{\begin{bmatrix} e_s \\ e_f \end{bmatrix}\begin{bmatrix} e_s^T & e_f^T \end{bmatrix}\right\} = \begin{bmatrix} P_{ss} & P_{sf} \\ P_{sf}^T & P_1 \end{bmatrix}$$

Expressions for e_s and e_f can be obtained by subtracting equations of motion (7) from the smoothing equations (20) and the filtering equations (24).

$$\begin{bmatrix} \dot{e}_s \\ \dot{e}_f \end{bmatrix} = \begin{bmatrix} \mathcal{F} + \mathcal{G}P_1^{-1} & -\mathcal{G}P_1^{-1} \\ 0 & \mathcal{F} - P_1H_1^TR^{-1}H_1 \end{bmatrix} \begin{bmatrix} e_s \\ e_f \end{bmatrix} - \begin{bmatrix} G_1 - KD + P_1H_1^TR^{-1}D \\ G_1 - KD \end{bmatrix} u$$

where $\mathcal{F} = F_{11} - G_1QD^TR^{-1}H_1$ and $\mathcal{G} = G_1(Q - QD^TR^{-1}DQ)G_1^T$.

$$\begin{bmatrix} \dot{P}_{ss} & \dot{P}_{sf} \\ \dot{P}_{sf}^T & \dot{P}_1 \end{bmatrix} = \begin{bmatrix} \mathscr{F} + \mathscr{G}P_1^{-1} & -\mathscr{G}P_1^{-1} \\ 0 & \mathscr{F} - P_1H_1^TR^{-1}H_1 \end{bmatrix} \begin{bmatrix} P_{ss} & P_{sf} \\ P_{sf}^T & P_1 \end{bmatrix}$$

$$+ \begin{bmatrix} P_{ss} & P_{sf} \\ P_{sf}^T & P_1 \end{bmatrix} \begin{bmatrix} \left(\mathscr{F} + \mathscr{G}P_1^{-1}\right)^T & 0 \\ -P_1^{-1}\mathscr{G}^T & \left(\mathscr{F} - P_1H_1^TR^{-1}H_1\right)^T \end{bmatrix}$$

$$+ \begin{bmatrix} \left(G_1 - KD + P_1H_1^TR^{-1}D\right) Q \left(G_1 - KD + P_1H_1^TR^{-1}D\right)^T, & \left(G_1 - KD + P_1H_1^TR^{-1}D\right) Q \left(G_1 - KD\right)^T \\ \left(G_1 - KD\right) Q \left(G_1 - KD + P_1H_1^TR^{-1}D\right)^T, & \left(G_1 - KD\right) Q \left(G_1 - KD\right)^T \end{bmatrix}$$

The $\dot{P}_1$ equation leads to the same equation as obtained earlier for $\dot{P}_1$, i.e., equation (23).

$$\dot{P}_1 = F_{11}P_1 + P_1F_{11}^T + G_1QG_1^T - KRK^T \qquad (23)$$

$$\dot{P}_{sf} = \left(\mathscr{F} + \mathscr{G}P_1^{-1}\right) P_{sf} - \mathscr{G} + P_{sf}\left(\mathscr{F} - P_1H_1^TR^{-1}H_1\right)^T$$

$$+ \left(G_1 - KD + P_1H_1^TR^{-1}D\right) Q(G - KD)^T.$$

But $e_s(T) = e_f(T)$ because $x_1(T) = \hat{x}_1(T)$.

$\therefore \quad P_{sf}(T) = P_1(T).$

Substituting above and simplifying

$$\dot{P}_{sf}(T) = F_{11}P_1 + P_1F_{11}^T + G_1QG_1^T - KRK^T$$

$\therefore \quad P_{sf}(t) = P_1(t).$

Using these results, we can obtain an expression for $\dot{P}_{ss}$.

$$\dot{P}_{ss} = \left(\mathcal{F} + \mathcal{G}P_1^{-1}\right) P_{ss} - \mathcal{G} + P_{ss}\left(\mathcal{F} + \mathcal{G}P_1^{-1}\right)^T - \mathcal{G}^T$$

$$+ \left(G_1 - KD + P_1 H_1^T R^{-1} D\right) Q \left(G_1 - KD + P_1 H_1^T R^{-1} D\right)^T$$

or

$$\dot{P}_{ss} = \left(\mathcal{F} + \mathcal{G}P_1^{-1}\right) P_{ss} + P_{ss}\left(\mathcal{F} + \mathcal{G}P_1^{-1}\right)^T - \mathcal{G} \tag{27}$$

Equation (27) is the equivalent of the Rauch-Tung-Striebel Formula. To derive the equivalent of the Bryson-Frazier formula, let

$$P_{ss} = P_1 + P_1 \Lambda P_1 \tag{28}$$

where Λ is an $(n-p) \times (n-p)$ matrix to be determined.

$$\dot{P}_{ss} = \dot{P}_1 + \dot{P}_1 \Lambda P_1 + P_1 \dot{\Lambda} P_1 + P_1 \Lambda \dot{P}_1 .$$

Substituting for $\dot{P}_{ss}$ from (27) and $\dot{P}_1$ from (23), we get

$$\dot{\Lambda} = \Lambda\left(\mathcal{F} - P_1 H_1^T R^{-1} H_1\right) + \left(\mathcal{F} - P_1 H_1^T R^{-1} H_1\right)^T \Lambda - H_1^T R^{-1} H_1 \tag{29}$$

8. OPTIMAL SMOOTHER AS COMBINATION OF TWO OPTIMAL FILTERS

We shall now show that the optimal smoother is a weighted combination of two Kalman Filters, one a forward filter which uses $\{z(\alpha), t_o \leq \alpha \leq t\}$, and the other a backward filter which uses $\{z(\alpha), t \leq \alpha \leq T\}$ with covariance at T equal to infinity [8]. Let

$\hat{x}_B(t)$ = filtering estimate of backward filter

$P_B(t)$ = covariance of error in the backward filter.

Also, let

$$E = P_1^{-1}; \qquad B = P_B^{-1}.$$

We shall prove the following results:

$$P_{ss}^{-1} = E + B \tag{30}$$

$$x_1(t) = P_{ss}[E\hat{x}_1 + B\hat{x}_B] \tag{31}$$

With boundary conditions,

$$E(t_o+) = P_1^{-1}(t_o+), \qquad B(T) = 0$$
$$\hat{x}_1(t_o+) = \hat{x}_1(t_o+), \qquad \hat{x}_B(T) = 0 \tag{32}$$

E and B satisfy the following equations:

$$\dot{E} = -E\mathcal{F} - \mathcal{F}^T E - E\mathcal{G}E + H_1^T R^{-1} H_1$$

$$\dot{B} = -B\mathcal{F} - \mathcal{F}^T B + B\mathcal{G}B - H_1^T R^{-1} H_1$$

$$\therefore \qquad \dot{P}_{ss}^{-1} = \dot{E} + \dot{B} = -P_{ss}^{-1}\mathcal{F} - \mathcal{F}^T P_{ss}^{-1} - E\mathcal{G}E + \left(P_{ss}^{-1} - E\right)\mathcal{G}\left(P_{ss}^{-1} - E\right)$$

$$\dot{P}_{ss}^{-1} = -P_{ss}^{-1}\left(\mathcal{F} + \mathcal{G}E\right) + \left(\mathcal{F} + \mathcal{G}E\right)^T P_{ss}^{-1} + P_{ss}^{-1}\mathcal{G}P_{ss}^{-1}$$

or

$$\dot{P}_{ss} = -P_{ss}\dot{P}_{ss}^{-1}P_{ss} = \left(\mathcal{F} + \mathcal{G}P_1^{-1}\right)^T P_{ss} + P_{ss}\left(\mathcal{F} + \mathcal{G}P_1^{-1}\right) - \mathcal{G}$$

which is the same as (27).

Also, at final time $P_{ss}^{-1}(T) = E(T)$ and $B(T) = 0$. Hence (30) is proved.

$$\dot{\hat{x}}_1 = \mathcal{F}\hat{x}_1 + F_{12}x_2 + P_1H_1^TR^{-1}(z_2-H_2x_2-H_1\hat{x}_1) + G_1QD^TR^{-1}(z_2-H_2x_2)$$

$$\dot{\hat{x}}_B = \mathcal{F}\hat{x}_B + F_{12}x_2 - P_BH_1^TR^{-1}(z_2-H_2x_2-H_1\hat{x}_B) + G_1QD^TR^{-1}(z_2-H_2x_2).$$

From (31)

$$\dot{x}_1 = \dot{P}_{ss}[E\hat{x}_1+B\hat{x}_B] + P_{ss}[\dot{E}\hat{x}_1+\dot{B}\hat{x}_B] + P_{ss}[E\dot{\hat{x}}_1+\dot{B}\hat{x}_B+B\dot{\hat{x}}_B].$$

Substituting for $\dot{P}_{ss}$, P_{ss}, B, and $\hat{x}_B$ from (27), (30), (31) we get

$$\dot{x}_1 = \left(\mathcal{F} + \mathcal{G}P_1^{-1}\right)x_1 + F_{12}x_2 - \mathcal{G}P_1^{-1}\hat{x}_1 + F_{12}x_2 + G_1QD^TR^{-1}(z_2-H_2x_2)$$

which is the same equation as (20) if we substitute

$$\lambda_1 = P_1^{-1}(\hat{x}_1 - x_1).$$

Furthermore, $x_1(T) = \hat{x}_1(T)$ and $\hat{x}_B(T) = 0$. Hence (31) is proved for all t.

9. EXAMPLE OF A SIMPLE INTEGRATOR WITH EXPONENTIALLY CORRELATED NOISE IN MEASUREMENTS

$$\dot{x} = u \qquad (33)$$

x, u, z, m all scalars

$$z = x + m \qquad 0 \leq t \leq T$$

$$E[m(t)] = 0, \qquad E[m(t)m(\tau)] = r\exp(-b|t-\tau|)$$

$$E[u(t)] = 0, \qquad E[u(t)u(\tau)] = q\delta(t-\tau).$$

m(t) can be produced by passing white noise through a first order filter.

$$\dot{m} = -bm + bw$$

where

$$E[w(t)] = 0, \qquad E[w(t)w(\tau)] = \frac{2r}{b}\,\delta(t-\tau)$$

$$E[m(0)] = 0, \qquad E[m^2(0)] = r.$$

There is no correlation between u(t), w(t), x(0), and m(0). In this problem, the augmented state equations are

$$\begin{bmatrix} \dot{x} \\ \dot{m} \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & -b \end{bmatrix} \begin{bmatrix} x \\ m \end{bmatrix} + \begin{bmatrix} u \\ bw \end{bmatrix}$$

$$z = [1 \quad 1] \begin{bmatrix} x \\ m \end{bmatrix}$$

$$\dot{z} = \dot{x} + \dot{m} = u - bm + bw = u - b(z-x) + bw$$

or

$$\dot{z} = bx - bz + [1 \ , \ 1] \begin{bmatrix} u \\ bw \end{bmatrix}.$$

$$\therefore \qquad x_2 = z \quad \text{and} \quad z_2 = \dot{z}.$$

Equation (33) corresponds to equation (7).

$$Q = \begin{bmatrix} q & 0 \\ 0 & 2rb \end{bmatrix}; \quad x_1 = x, \quad G_1 = [1 \quad 0], \quad F_{11} = 0, \quad F_{12} = 0$$

$$D = [1 \ , \ 1], \quad H_1 = b, \quad H_2 = -b$$

$$R = DQD^T = q + 2rb; \quad \mathscr{F} = -\frac{qb}{R}, \quad \mathscr{G} = q - \frac{q^2}{R}.$$

Forward Filter

Let

p_x = covariance of error in filtering estimate of x.

$K = \frac{1}{R}(p_x b + q)$.

(23) gives

$$\therefore \quad \dot{p}_x = q - \frac{(p_x b + q)^2}{R}.$$

(17) gives

$$p_x(0+) = p_x(0) - \frac{p_x^2(0)}{p_x(0) + r} = \frac{p_x(0)}{1 + \frac{p_x(0)}{r}}.$$

These equations can be solved analytically.

$$\boxed{p_x = -\frac{q}{b} + \frac{\sqrt{qR}}{b} \cdot \frac{1 + C\, e^{-2at}}{1 - C\, e^{-2at}}} \tag{34}$$

where

$$C = \frac{p_x(0+) + \frac{q}{b} - \frac{\sqrt{qR}}{b}}{p_x(0+) + \frac{q}{b} + \frac{\sqrt{qR}}{b}}$$

$$a = b\sqrt{\frac{q}{R}}.$$

$$\boxed{\hat{x}(0+) = \frac{p_x(0)}{p_x(0) + r}\, z(0)} \tag{35}$$

$$\dot{\hat{x}} = K(z_2 + bz - b\hat{x}) \qquad \text{where } z_2 = \dot{z}.$$

Let

$$\hat{x} = x^* + Kz$$

$$\dot{\hat{x}} = \dot{x}^* + \dot{K}z + Kz_2.$$

$$\therefore \quad \boxed{\dot{x}^* = Kb(z-\hat{x}) - \dot{K}z} \tag{36}$$

where

$$\dot{K} = \frac{b}{R}\,\dot{p}_x = \frac{b}{R}\,(q - K^2R)$$

or

$$\dot{K} = \frac{bq}{R} - \frac{b}{R^2}\,(p_x b + q)^2.$$

<u>Backward Filter</u>

Let

p_b = error covariance of backward filter estimates.

$$\dot{p}_b = -q + \frac{(p_b b - q)^2}{R} \quad , \quad p_b(T) = \infty.$$

Solution of this equation gives

$$p_b = \frac{q}{b} + \frac{\sqrt{qR}}{b}\,\frac{1 + e^{-2\alpha(T-t)}}{1 - e^{-2\alpha(T-t)}}. \tag{37}$$

Similarly,

$$\dot{\hat{x}}_B = \frac{(p_b b - q)}{R}\,(b\hat{x}_B - z_2 - bz), \qquad \hat{x}_B(T) = 0. \tag{38}$$

Optimal Smoother

If p_s denotes the error covariance of smoothing estimate of x (denoted by x_s),

$$p_s = \frac{p_x p_b}{p_x + p_b}$$

$$x_s = p_s\left(\frac{\hat{x}}{p_x} + \frac{\hat{x}_B}{p_B}\right).$$

Figure 3 shows curves for p_x, p_b, and p_s for a particular value of b and large time-interval T. Note that both p_x and p_b reach steady state after some time.

For t such that

$$2\alpha t \gg 1, \qquad p_x = -\frac{q}{b} + \frac{\sqrt{qR}}{b}$$

$$2\alpha(T-t) \gg 1, \qquad p_b = \frac{q}{b} + \frac{\sqrt{qR}}{b}.$$

So p_s has a constant value in the middle and two transients at the

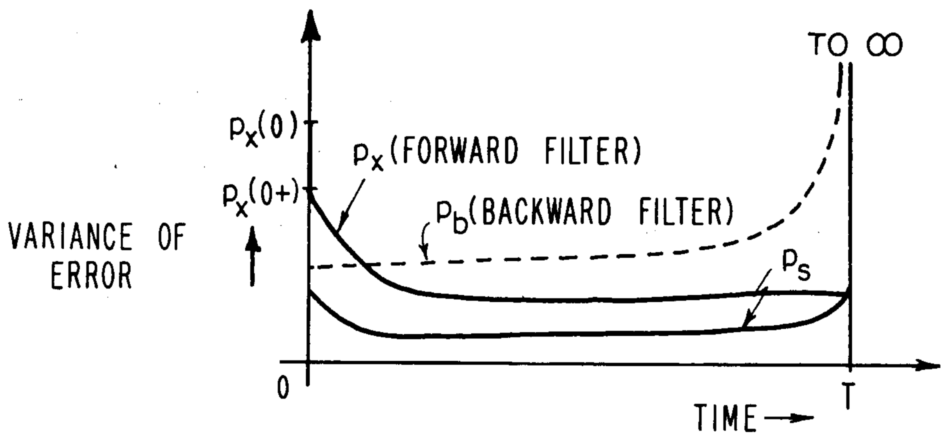

FIG. 3 SHOWING p_x, p_b, p_s VS TIME

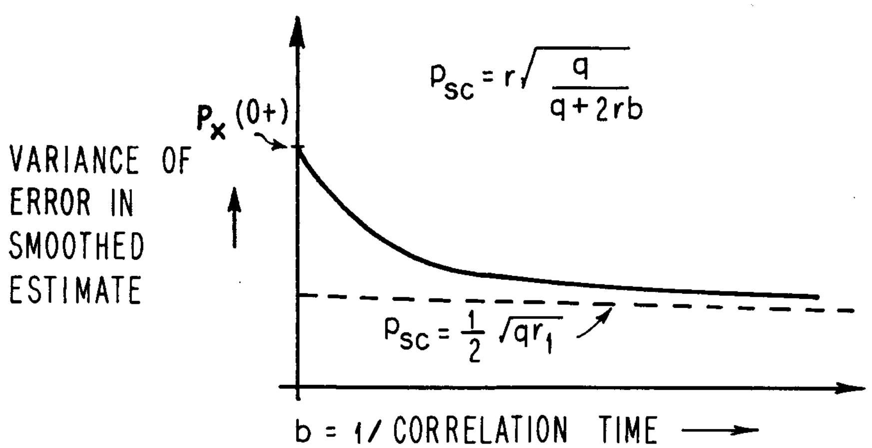

FIG. 4 SHOWING SMOOTHING VARIANCE VS. (1/CORRELATION TIME)

ends. The constant value of p_s denoted as p_{sc} can be calculated easily.

$$p_{sc} = \frac{(q/b+\sqrt{qR}/b)(-q/b+\sqrt{qR}/b)}{2\sqrt{qR}/b} = r\sqrt{\frac{q}{q+2rb}}.$$

Figure 4 shows p_{sc} vs b. It can be seen that p_{sc} has maximum value for b = 0 (bias error) and minimum value for b = ∞ (white noise). This means that smoothing is most effective for the white noise case.

Bias Error, b = 0

$$b \to 0, \quad R \to q, \quad K \to 1.$$

$$\dot{p}_x = 0, \quad \dot{p}_b = 0,$$

so p_x and p_b are constant. But

$$p_x(0+) = \frac{p_x(0)}{1+\frac{p_x(0)}{r}}; \qquad p_b(T) = \infty.$$

This means that randomness is only due to initial uncertainty and backward filter gives no information.

$$\dot{\hat{x}} = z_2 = \dot{z}$$

$$\therefore \quad \hat{x} = z + \hat{x}(0+) - z(0) = z + \frac{r}{p_x(0)+r}\, z(0).$$

$\frac{r}{p_x(0)+r}\, z(0)$ is the initial guess of the bias error and it is added to all subsequent measurements. Thus $p_x(t)$ remains constant at $p_x(0+)$.

It is clear that for this case, smoothing estimates are exactly the same as filtering estimates. $p_s = p_x$; $x_s = \hat{x}$. So we do not gain anything by smoothing the results.

White Noise, $b \to \infty$ and $\frac{2r}{b} \to r_1$

r_1 is the area under the delta function representing spectral density of white noise m(t).

$$\frac{R}{b^2} = \frac{q + 2br}{b^2} \to r_1$$

$$\alpha = b\sqrt{\frac{q}{R}} \to \sqrt{\frac{q}{r_1}}$$

$$p_x(0+) = \frac{rp_x(0)}{r + p_x(0)} = \frac{\frac{2r}{b}\, p_x(0)}{\frac{2r}{b} + \frac{2}{b}\, p_x(0)} \to p_x(0)$$

$$p_x \to \sqrt{qr_1}\ \frac{1 + C\, e^{-2\sqrt{q/r_1}\, t}}{1 - C\, e^{-2\sqrt{q/r_1}\, t}}$$

where

$$C = \frac{p_x(0) - \sqrt{qr_1}}{p_x(0) + \sqrt{qr_1}}$$

$$\hat{x}(0+) = 0, \quad K \to 0 \quad \text{but} \quad Kb \to \frac{p_x}{r_1}$$

$$\therefore \quad \dot{\hat{x}} = \frac{p_x}{r_1}(z-\hat{x}).$$

This is the usual Kalman Filter. For the backward filter,

$$p_b = \sqrt{qr_1}\,\frac{1 + e^{-2\sqrt{q/r_1}\,(T-t)}}{1 - e^{-2\sqrt{q/r_1}\,(T-t)}}$$

$$\dot{\hat{x}}_B = \frac{p_b}{r_1}(\hat{x}_B - z) \qquad \hat{x}_B(T) = 0.$$

The asymptotic values of p_b and p_x are the same, viz., $\sqrt{qr_1}$.

$$\therefore \qquad p_{sc} = \frac{1}{2}\sqrt{qr_1}.$$

For constant values of p_s, the smoothing estimate is the mean of the filtering estimates

$$x_s = \frac{\hat{x} + \hat{x}_B}{2}.$$

10. GENERAL CASE

For the general case, manipulations get very involved, but the results are essentially similar. We shall only state the problem and give the final results.

Augmented state and measurement equations are

$$\dot{x} = Fx + \Gamma u$$

$$z_1 = H_1 x + w$$

$$y = Cx,$$

where

z_1 = q × 1 vector of measurements containing white noise

y = r × 1 vector of perfect measurements

w = q × 1 vector of white noise errors in z_1

$$E\{u(t)\} = 0; \quad E\{u(t)u^T(\tau)\} = Q(t)\ \delta(t-\tau)$$

$$E\{w(t)\} = 0; \quad E\{w(t)w^T(\tau)\} = R_1(t)\ \delta(t-\tau)$$

$$E\{u(t)w^T(t)\} = S_1(t); \quad E\{x(t_o)\} = 0; \quad E\{x(t_o)x^T(t_o)\} = P(t_o)$$

$$E\{w(t)x^T(t_o)\} = 0.$$

The smoothing problem can be stated as follows: Minimize

$$J = \frac{1}{2}\, x^T(t_o)\, P^{-1}(t_o)\, x(t_o) + \frac{1}{2}\int_{t_o}^{T} \begin{pmatrix} z_1 - H_1 x \\ u \end{pmatrix}^T \begin{pmatrix} R_1 & S_1^T \\ S_1 & Q \end{pmatrix}^{-1} \begin{pmatrix} z_1 - H_1 x \\ u \end{pmatrix} dt$$

subject to

$$\dot{x} = Fx + \Gamma u$$

$$y = Cx.$$

Filtering results obtained are

$$\dot{P}_1 = F_{11}P_1 + P_1F_{11}^T + G_1QG_1^T - K_1RK_1^T$$

$$\dot{\hat{x}}_1 = F_{11}\hat{x}_1 + F_{12}x_2 + K_1(z - H\hat{x}_1)$$

where

$$K_1 = (P_1 H^T + G_1 S) R^{-1}; \qquad H_1 M^{-1} = [H_3 \mid H_4]$$

$$H_5 = AF_{21}, \qquad H_6 = AF_{22}$$

where A is a matrix of ones and zeros.

$$H = \begin{bmatrix} H_3 \\ \hline H_5 \end{bmatrix}, \qquad S = [S_1 \mid QD^T], \qquad R = \left[\begin{array}{c|c} R_1 & S_1 D^T \\ \hline DS_1 & DQD^T \end{array}\right]$$

and

$$z = \begin{bmatrix} z_1 - H_4 x_2 \\ \hline z_2 - H_6 x_2 \end{bmatrix}.$$

The smoothing equations are the same as before with

$$\mathcal{F} = F_{11} - G_1 (Q - SR^{-1} S^T) \Delta^{-1} S_1 R_1^{-1} H_3 - G_1 \Delta D^T (D \Delta D^T)^{-1} H_5$$

$$\mathcal{G} = G_1 (Q - SR^{-1} S^T) G_1^T$$

where

$$\Delta = Q - S_1 R_1^{-1} S_1^T.$$

REFERENCES

[1] Kalman, R. E., and R. S. Bucy, "New Results in Linear Filtering and Prediction Theory," ASME Trans. (J. of Basic Engineering), Vol. 83D, March 1961, pp. 95-108.

[2] Bryson, A. E., and M. Frazier, "Smoothing for Linear and Non-linear Dynamic Systems," Proc. of the Optimum Systems Synthesis Conf., U.S. Air Force Technical Report ASD-TDR-063-119, February 1963.

[3] Bryson, A. E., and D. E. Johansen, "Linear Filtering for Time-Varying Systems Using Measurements Containing Coloured Noise," IEEE Trans. on Automatic Control, January 1965.

[4] Ho, Y. C., and R. C. Lee, "A Bayesian Approach to Problems in Stochastic Estimation and Control," IEEE Transactions on Automatic Control, Vol. AC-9, No. 4, October 1964.

[5] E. Bodewig, Matrix Calculus, North-Holland Publishing Co., Amsterdam; Interscience Publishers, Inc., New York, 1956.

[6] Bryson, A. E., and Y. C. Ho, "Optimal Programming, Estimation and Control," Lecture Notes, Harvard University, 1966-1967.

[7] Rauch, Tung, and Striebel, "Maximum Likelihood Estimates of Linear Dynamic Systems," AIAA Journal, Vol. 3, No. 8, pp. 1445-1450, August 1965.

[8] Fraser, Donald C., "On the Application of Optimal Linear Smoothing Techniques to Linear and Nonlinear Dynamic Systems," Ph. D. Thesis, Massachusetts Institute of Technology, January 1967.

[9] A. E. Bryson and S R. McReynolds, "A Successive Sweep Method for Solving Optimal Programming Problems," Sixth JACC, Troy, N. Y., June 1965.