(NASA-CR-141279) A CLASS OF FINITE DIMENSIONAL OPTIMAL NONLINEAR ESTIMATORS (Massachusetts Inst. of Tech.) 11 p HC $3.25 CSCL 12A

N75-15384

Unclas
G3/64 06536

September, 1974

ESL-P-569

A CLASS OF FINITE DIMENSIONAL OPTIMAL NONLINEAR ESTIMATORS

Steven I. Marcus* and Alan S. Willsky**

Electronic Systems Laboratory
Department of Electrical Engineering
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

Abstract

In this paper we consider classes of nonlinear systems for which the optimal (minimum variance) estimator is finite dimensional. Finite dimensional optimal nonlinear state estimators are derived for bilinear systems evolving on nilpotent and solvable Lie groups. These results are extended to other classes of systems involving polynomial nonlinearities. Finally, the concepts of exact differentials and path-independent integrals are used to derive optimal finite dimensional estimators for a further class of nonlinear systems.

1. INTRODUCTION

It is well known that the class of linear dynamical systems with linear observations and white Gaussian plant and observation noises is particularly appealing, because the optimal state estimator consists of a finite dimensional linear system (which is easily implemented with the aid of a digital computer). In general, the optimal (minimum variance) estimator for a finite dimensional nonlinear system consists of an infinite dimensional system of moment equations, and approximations must be made for practical implementation. Consequently, one is led to investigate subclasses of nonlinear systems which admit finite dimensional optimal estimators. Such an investigation not only identifies systems for which optimal estimation is computationally feasible, but it also provides valuable theoretical insight into the underlying structure of optimal estimation for general nonlinear systems.

There is, in fact, a class of nonlinear systems which possesses a great deal of structure -- the class of bilinear systems. Such systems have been studied by several authors [1]-[12], and many tools from the theories of Lie groups and differential geometry have proved to be quite useful. Estimation for bilinear systems on abelian Lie groups is discussed by Lo and Willsky [8], and some optimal finite dimensional estimators are derived; these results are generalized to a larger class of systems by Willsky [9].

In this paper we consider bilinear systems evolving on solvable and nilpotent Lie groups. For such systems, we will prove that the minimum variance estimator is finite dimensional [illegible]

---

* National Science Foundation Fellow.

** Assistant Professor of Electrical Engineering, M.I.T. The work of this author was supported in part by NASA under Grant NGL-22-009-124 and in part by NSF under Grant GK-42090.

stochastic system driven by the innovations. These results are extended to other systems with polynomial nonlinearities. In addition, the concepts of path-independent integrals and exact differential forms are used to prove the existence of finite dimensional estimators for another class of nonlinear systems.

2. NOTATION AND FORMULATION OF THE PROBLEM

The basic bilinear equation considered in this paper is

$$\dot{X}(t) = (A_0 + \sum_{i=1}^{N} u_i(t)A_i)X(t); \; X(0) = I \qquad (1)$$

where the A_i are given $k \times k$ matrices, X is a $k \times k$ matrix, u_i is the i^{th} componenent of u, and u is the n-dimensional "colored noise" process generated by the finite dimensional linear system

$$dx(t) = F(t)x(t)dt + Q^{1/2}(t)dw(t) \qquad (2)$$

$$u(t) = C(t)x(t) \qquad (3)$$

Here w is a standard Brownian motion process, $Q \geq 0$, $x(0)$ has a Gaussian distribution independent of w, and the pair $[F(t), Q^{1/2}(t)]$ is stabilizable [18]. The existence of a unique solution to (1),(2) is proved in [15],[16]. Although X by itself is not a Markov process, it can be shown that the pair (X,x) is Markov.

As in the deterministic case [1], the solution X of (1) evolves on a matrix Lie group. More specifically, we define $\mathscr{L} = \{A_i\}_{LA}$ to be the matrix Lie algebra generated by $\{A_i, i=0, 1,\ldots,N\}$; i.e., $\mathscr{L}$ is the smallest subspace of $k \times k$ matrices containing $\{A_i, i=0, 1,\ldots,N\}$ and closed under the commutor product $[P,Q] = PQ-QP$. We also define the matrix Lie group $G = \{\exp\mathscr{L}\}_G$ associated with $\mathscr{L}$ to be the smallest group (under matrix multiplication) containing $\exp L$ for all $L \in \mathscr{L}$. If $X(0) \in G$, then $X(t) \in G$ for all $t \geq 0$.

In the sequel we will be primarily concerned with systems in which $\mathscr{L}$ is a solvable or nilpotent Lie algebra; such systems evolve on solvable or nilpotent Lie groups.

Definition 1. A Lie algebra $\mathscr{L}$ is solvable if the derived series of ideals [23]

$$\mathscr{L}^{(0)} = \mathscr{L}$$

$$\mathscr{L}^{(n)} = \left[\mathscr{L}^{(n-1)}, \mathscr{L}^{(n-1)}\right] \triangleq \{[L_1,L_2] \mid L_1,L_2 \in \mathscr{L}^{(n-1)}\},$$

$n \geq 1$, terminates in $\{0\}$. $\mathscr{L}$ is nilpotent if the lower central series of ideals

$$\mathscr{L}^{0} = \mathscr{L}$$

$$\mathscr{L}^{n} = \left[\mathscr{L},\mathscr{L}^{n-1}\right] \triangleq \{[L_1,L_2] \mid L_1 \in \mathscr{L}, L_2 \in \mathscr{L}^{n-1}\}, n \geq 1$$

terminates in $\{0\}$. $\mathscr{L}$ is abelian if $\mathscr{L}^{(1)} = \{0\}$.

We state here two results conserning canonical representations of nilpotent and solvable Lie algebras which will be particularly germane to our study; the reader is referred to [23] for further properties of solvable and nilpotent Lie algebras and groups. Let $\mathscr{C}$ denote the complex numbers, $g\ell(n,\mathscr{C})$ denote the space of $n \times n$ matrices with complex entries, and $GL(n,\mathscr{C})$ denote the space of nonsingular complex $n \times n$ matrices.

Lemma 1 [23, p.214]: Let $\mathscr{L}$ be a Lie algebra of matrices in $g\ell(n,\mathscr{C})$. Then $\mathscr{L}$ is solvable if and only if there exists a matrix $P \in GL(n,\mathscr{C})$ such that, for all elements $A \in \mathscr{L}$, the matrix $B = PAP^{-1}$ is n upper triangular form ($b_{ij} = 0$ for $i > j$).

Lemma 2 [23, p.224]: Let $\mathscr{L}$ be a Lie algebra of matrices in $g\ell(n,\mathscr{C})$. Then $\mathscr{L}$ is nilpotent if and only if there exists a matrix $P \in GL(n,\mathscr{C})$ such that, for all elements $A \in \mathscr{L}$, the matrix $B = PAP^{-1}$ has the block diagonal form

$$\begin{bmatrix} \begin{bmatrix} \phi_1(A) & & * \\ 0 & \ddots & \\ & & \phi_1(A) \end{bmatrix} & & 0 \\ & \begin{bmatrix} \phi_2(A) & & * \\ 0 & \ddots & \\ & & \phi_2(A) \end{bmatrix} & \\ 0 & & \ddots \end{bmatrix}$$

The functions $\phi_k: \mathscr{L} \to \mathscr{C}$ are linear. Furthermore, $\phi_k([\mathscr{L},\mathscr{L}]) = 0$.

The block form for nilpotent Lie algebras will be called the nilpotent canonical form.

The first estimation problem to be discussed involves the state equations (1)-(3) and the p-dimensional observation process

$$dz(t) = H(t)x(t)dt + R^{1/2}(t)dv(t) \tag{4}$$

where $R > 0$ and v is a standard Brownian motion independent of w and $x(0)$. This observation process is of interest in the problem of estimating the attitude of a rotating rigid body by means of a strapdown inertial navigation system [5],[7].

The criterion for the optimal estimate $(\tilde{X}(t|t),\tilde{x}(t|t))$ will be the minimization of the conditional error covariance

$$E[(x(t) - \tilde{x}(t|t))'(x(t) - \tilde{x}(t|t))$$
$$+ tr\{(X(t)-\tilde{X}(t|t))'(X(t) - \tilde{X}(t|t))\} \mid z^t]$$

where "tr" denotes trace and $z^t = \{z(s), 0 \leq s \leq t\}$. It is well known [14] that the causal minimum-variance estimates are given by the conditional means

$$\hat{X}(t|t) = E^t[X(t)] \overset{\Delta}{=} E[X(t)|z^t]$$
$$\hat{x}(t|t) = E^t[x(t)] \overset{\Delta}{=} E[x(t)|z^t]$$

(we will use the three notations for conditional expectation interchangeably). The computation of $\hat{x}(t|t)$ is performed by the finite-dimensional (linear) Kalman filter; moreover, the conditional density of $x(t)$ given z^t is Gaussian with mean $\hat{x}(t|t)$ and nonrandom covariance $P(t)$ [14]. As remarked in Section I, the computation of $\hat{X}(t|t)$ requires in general an infinite dimensional system of equations. We will show in the succeeding sections that $\hat{X}(t|t)$ can be computed with a finite-dimensional nonlinear estimator if (1)-(3) evolves on certain nilpotent or solvable Lie groups.

3. REDUCTION OF THE GENERAL PROBLEM

In this section we show that some estimation problems on solvable Lie groups can be solved by considering an estimation problem on a particular nilpotent Lie group. The first lemma generalizes a result of Willsky [9] (the proof is analogous). For $A \in g\ell(k,\mathscr{C})$, we define [illegible]

$g\ell(k,\mathscr{C})$ by

$$ad_A(B) \overset{\Delta}{=} AB-BA \overset{\Delta}{=} [A,B] \tag{5}$$

The notation ad_A^i denotes the i^{th} power of the operator ad_A.

Lemma 3: Consider the equations (1)-(4), and let $\mathscr{L}_0$ be the ideal in $\mathscr{L} \overset{\Delta}{=} \{A_i\}_{LA}$ spanned by

$ad_{A_0}^i(A_j)$ $j=1,\ldots,N;$ $i=0,\ldots,k^2-1.$

Define the k x k matrix valued process

$$Y(t) = e^{-A_0 t} X(t) \tag{6}$$

Then there exists a matrix $D(t)$ such that Y satisfies

$$\dot{Y}(t) = \left[\sum_{i=1}^{M} H_i y_i(t)\right] Y(t) \tag{7}$$

where $\{H_1,\ldots,H_M\}$ is a basis for $\mathscr{L}_0$ and

$$y(t) = D(t)x(t) \tag{8}$$

In addition, $\hat{X}$ can be computed according to

$$\hat{X}(t|t) = e^{A_0 t}\hat{Y}(t|t) \tag{9}$$

Lemma 3 enables us, without loss of generality, to examine the estimation problem for $Y(t)$ evolving on the normal subgroup $G_0 = \{\exp \mathscr{L}_0\}_G$, rather than for $X(t)$ evolving on the full Lie group $G = \{\exp \mathscr{L}\}_G$. The particular case with which we will be concerned is that in which $\mathscr{L}$ is solvable and $\mathscr{L}_0$ is nilpotent. In fact, it can easily be shown that if $\mathscr{L}_0$ is nilpotent, then $\mathscr{L}$ must be solvable; however, the converse is not true. According to Lemma 3, for such systems we need only consider the case in which $A_o = 0$ and $\mathscr{L} = \mathscr{L}_o$ is nilpotent.

Example 1: Assume that A_o is upper triangular, and $\{A_1,\ldots,A_n\}$ are strictly upper triangular (diagonal elements are zero). Then $\mathscr{L}$ is solvable and $\mathscr{L}_o$ is nilpotent.

Example 2: In the finite dimensional estimator [illegible]

By means of Lemma 2, the problem can be further reduced to the consideration of Lie algebras in nilpotent canonical form.

Lemma 4: Consider (1)-(4), where $A_o = 0$ and $\mathscr{L}$ is nilpotent. Then there exists a matrix $P \in GL(k,\mathscr{C})$ such that

$\hat{X}(t|t) = P\hat{Y}(t|t)$

where Y satisfies (7) and $\{H_1,\ldots,H_M\}$ are in nilpotent canonical form.

Finally, by means of the following trivial lemma, we reduce the problem to the consideration of one block in the nilpotent canonical form.

Lemma 5: Consider (1)-(4), where $A_o = 0$ and $\{A_1,\ldots,A_N\}$ are in nilpotent canonical form. Then X(t) has a block diagonal form conformable with that of $\{A_1,\ldots,A_N\}$.

Thus the system (1) can be viewed as the direct sum of a number of subsystems; for each k_i-dimensional subsystem, $\{A_1,\ldots,A_N\} \in gn(k_i)$ (here we have defined gn(m) to be the Lie subalgebra of $g\ell(m,\mathscr{C})$ consisting of the upper triangular matrices with equal diagonal elements). We now state the major theorem on finite dimensional estimation for such systems.

Theorem 1: Consider (1)-(4), where $A_o = 0$ and $\{A_1,\ldots,A_N\} \in gn(k)$. Then the conditional mean $\hat{X}(t|t)$ can be computed by a finite-dimensional system of nonlinear stochastic differential equations.

Theorem 1 is proved by induction: the case k=3 will be proved in Section IV, and the induction step is proved in [16]. We note that our result also includes the result in [8] as a special case.

4. PROOF FOR k=3

For simplicity of notation, the theorem will be proved for the case C(t) = I; the proof is precisely the same for arbitrary C(t). For k=3, we assume that n=N=4, and $A_o = 0$, $A_1 = I$,

$$A_2 = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad A_3 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad A_4 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

Then $\{A_1, A_2, A_3, A_4\}$ is a basis for gn(3). The solution of (1) can be expressed in closed form as

$$X(t) = \begin{bmatrix} e^{y_1(t)} & e^{y_1(t)}y_2(t) & e^{y_1(t)}[y_3(t)+\xi(t)] \\ 0 & e^{y_1(t)} & e^{y_1(t)}y_4(t) \\ 0 & 0 & e^{y_1(t)} \end{bmatrix} \tag{10}$$

where

$$y_i(t) = \int_0^t x_i(s)ds \tag{11}$$

and

$$\xi(t) = \int_0^t \int_0^{\sigma_1} x_2(\sigma_1)x_4(\sigma_2)d\sigma_2 d\sigma_1 \tag{12}$$

Remark: In the sequel we will apply, without further comment, a version of the Fubini theorem [17] which permits the interchange of integration and conditional expectation. Since we are dealing only with integrals of products of Gaussian random processes, the use of the Fubini theorem is easily justified [16].

It is evident from (10) that the computation of $\hat{X}(t|t)$ (in particular, $\hat{X}_{13}(t|t)$) requires the knowledge of the statistics of the entire process $\{x(s), 0 \leq s \leq t\}$ given z^t. Therefore we first define the conditional cross cross-covariance

$P(\sigma_1,\sigma_2,t) = E[(x(\sigma_1)-\hat{x}(\sigma_1|t))(x(\sigma_2)-\hat{x}(\sigma_2|t))'|z^t]$

and prove two crucial lemmas.

Lemma 6: The joint conditional density $p_{x(\sigma_1),x(\sigma_2)}(\nu,\nu'|z^t)$ is Gaussian with nonrandom conditional cross-covariance $P(\sigma_1,\sigma_2,t)$.

Proof: First, the conditional density is Gaussian because x^t and z^t are jointly Gaussian random processes. Assume $\sigma_1 > \sigma_2$; then

$$p_{x(\sigma_1),x(\sigma_2)}(\nu,\nu'|z^t)$$

$$= p_{x(\sigma_1)}(\nu|x(\sigma_2) = \nu',z^t)\ p_{x(\sigma_2)}(\nu'|z^t)$$

$$= p_{x(\sigma_1)}(\nu|x(\sigma_2) = \nu',z^t_{\sigma_2})\ p_{x(\sigma_2)}(\nu'|z^t) \qquad (13)$$

where $z^t_{\sigma_2} = \{z(s),\ \sigma_2 \le s \le t\}$.

Each of the densities in (13) is the result of a linear smoothing operation; hence, each is Gaussian with nonrandom covariance

$P_{\sigma_1|\sigma_2}(t)$ and $P(\sigma_2,\sigma_2,t)$, respectively [20].

Thus the cross-covariance satisfies

$$P_{\sigma_1|\sigma_2}(t) = P(\sigma_1,\sigma_1,t) - P(\sigma_1,\sigma_2,t)\cdot$$
$$\cdot P^{-1}(\sigma_2,\sigma_2,t)\ P'(\sigma_1,\sigma_2,t)$$

and $P(\sigma_1,\sigma_2,t)$ is also nonrandom (here $P(\sigma_2,\sigma_2,t)$ is invertible because $[F,Q^{1/2}]$ is stabilizable).

Lemma 6 allows the <u>off-line</u> computation of of $P(\sigma_1,\sigma_2,t)$ via the equations of Kwakernaak [24] (for $\sigma_1 \le \sigma_2$)

$$P(\sigma_1,\sigma_2,t) = P(\sigma_1)\ \Psi'(\sigma_2,\ \sigma_1)$$
$$-P(\sigma_1)\ [\int_{\sigma_2}^{t} \Psi'(\tau,\ \sigma_1)\ H'(\tau)\ R^{-1}(\tau)\ .$$
$$.\ H(\tau)\ \Psi\ (\tau,\ \sigma_2)d\tau]P(\sigma_2) \qquad (14)$$

$$\dot{\Psi}(t) = [F(t) - P(t)H'(t)\ R^{-1}(t)H(t)]\ \Psi\ (t);$$
$$\Psi\ (0) = I \qquad (15)$$

where the Kalman filter error covariance matrix $P(t) = P(t,\ t,\ t)$ is computed by the Riccati equation

$$\dot{P}(t) = F(t)\ P(t) + P(t)\ F'(t) + Q(t)$$
$$- P(t)H'(t)\ R^{-1}(t)\ H(t)P(t) \qquad (16)$$

Lemma 7: The conditional cross-covariance satisfies

$$P(\sigma,\ t,\ t) = K(t,\ \sigma)\ P(t) \qquad (17)$$

where

$$\frac{d}{dt} K'(t,\sigma) = -[F'(t) + P^{-1}(t)\ Q(t)]\ K'(t,\ \sigma);$$
$$K'(\sigma,\ \sigma) = I \qquad (18)$$

Proof: Let $\tilde{P}(\sigma,t) = E[(x(\sigma) - \hat{x}(\sigma|\sigma))(x(t) - \hat{x}(t|t))']$, and consider

$$P(\sigma,t,t) - \tilde{P}(\sigma,t) = E^t[(\hat{x}(\sigma|\sigma) - \hat{x}(\sigma|t))$$
$$(x(t) - \hat{x}(t|t))'] \qquad (19)$$

Since $\hat{x}(\sigma|\sigma) - \hat{x}(\sigma|t)$ is measurable with respect to z^t, the projection theorem implies that (19) equals zero. The proof is concluded by noticing that $\tilde{P}(\sigma,t) = K(t,\sigma)\ P(t)$ [21].

Returning to the proof of Theorem 1 for k=3, we first augment the state of (2) with $y_i(t)$, $i=1,\ldots,4$ (if a particular y_i can be obtained as a linear combination of the x_i's, we need not augment the state with that y_i). Then the Kalman Filter for the system (2), (4), (11) generates $\hat{x}(t|t)$ and $\hat{y}(t|t)$. We define the 8x8 conditional covariance matrix

$$V(\sigma_1,\sigma_2,t) = \begin{bmatrix} P(\sigma_1,\sigma_2,t) & S(\sigma_1,\sigma_2,t) \\ S'(\sigma_1,\sigma_2,t) & T(\sigma_1,\sigma_2,t) \end{bmatrix}$$

where

$$S(\sigma_1,\sigma_2,t) \triangleq E^t[(x(\sigma_1) - \hat{x}(\sigma_1|t)) \cdot (y(\sigma_2) - \hat{y}(\sigma_2|t))'] = \int_0^{\sigma_2} P(\sigma_1,\tau,t)d\tau \tag{20}$$

$$T(\sigma_1,\sigma_2,t) \triangleq E^t[(y(\sigma_1) - y(\sigma_1|t)) \cdot (y(\sigma_2) - y(\sigma_2|t))'] = \int_0^{\sigma_1}\int_0^{\sigma_2} P(\tau_1,\tau_2,t)d\tau_1 d\tau_2 \tag{21}$$

We also define $T(t) \triangleq T(t,t,t)$, $S(t) \triangleq S(t,t,t)$, $V(t) \triangleq V(t,t,t)$.

The characteristic function of a Gaussian random vector x with mean m and covariance P is given by

$$M_x(u) = E[\exp(iu'x)] = \exp[i\,u'm - \tfrac{1}{2}u'Pu] \tag{22}$$

Thus (for $j=1,2,3$)

$$\hat{x}_{jj}(t|t) = M_{y_1(t)}(-i) = e^{\hat{y}_1(t|t) + \frac{1}{2}T_{11}(t)} \tag{23}$$

It can also be shown that

$$\hat{x}_{12}(t|t) = [\hat{y}_2(t|t) + T_{12}(t)]e^{\hat{y}_1(t|t) + \frac{1}{2}T_{11}(t)} \tag{24}$$

$$\hat{x}_{23}(t|t) = [\hat{y}_4(t|t) + T_{14}(t)]e^{\hat{y}_1(t|t) + \frac{1}{2}T_{11}(t)} \tag{25}$$

$$E^t[e^{y_1(t)}y_3(t)] = [\hat{y}_3(t|t) + T_{13}(t)] \cdot e^{\hat{y}_1(t|t) + \frac{1}{2}T_{11}(t)} \tag{26}$$

Since (23)-(26) represent instantaneous nonlinear functions of $\hat{y}(t|t)$, they can be computed with a finite dimensional estimator.

Now consider $\xi(t)$ (see (12)); the approach here will be the reduction of this problem to [illegible] of terms of the form

$$\hat{\alpha}(t|t) = E^t[\alpha(t)] \triangleq E^t\left[\int_0^t W(t,\sigma)\,x(\sigma)d\sigma\right] \tag{27}$$

[illegible]

of the form

$$W(t,\sigma) = U(t)V(\sigma) \tag{28}$$

then $\alpha(t)$ is the output of a finite dimensional linear system driven by $x(t)$ [19], and $\hat{\alpha}(t|t)$ can be computed by the <u>linear</u> finite dimensional Kalman filter for the augmented state $(x'(t)\ \alpha'(t))'$.

It can easily be shown that

$$E^t[e^{y_1(t)}x_2(\sigma_1)\,x_4(\sigma_2)] = e^{\hat{y}_1(t|t) + \frac{1}{2}T_{11}(t)}[P_{24}(\sigma_1,\sigma_2,t) + (\hat{x}_2(\sigma_1|t) + S_{12}(t,\sigma_1,t))(\hat{x}_4(\sigma_2|t) + S_{14}(t,\sigma_2,t))] \tag{29}$$

Hence

$$E^t[e^{y_1(t)}\xi(t)] = e^{\hat{y}_1(t|t) + \frac{1}{2}T_{11}(t)} \cdot \Bigg\{E^t\left[\int_0^t\int_0^{\sigma_1} x_2(\sigma_1)\,x_4(\sigma_2)\,d\sigma_2 d\sigma_1\right] + E^t\left[\int_0^t\int_0^{\sigma_1} S_{12}(t,\sigma_1,t)\,x_4(\sigma_2)\,d\sigma_2\,d\sigma_1\right] + E^t\left[\int_0^t\int_0^{\sigma_1} S_{14}(t,\sigma_2,t)\,x_2(\sigma_1)\,d\sigma_2 d\sigma_1\right] + \int_0^t\int_0^{\sigma_1} S_{12}(t,\sigma_1,t)\,S_{14}(t,\sigma_2,t)\Bigg\} \tag{30}$$

Consider the second term in (30). Lemma 7 implies that $V(\sigma,t,t)$ can be written as

$$V(\sigma,t,t) = L(t,\sigma)\,V(t)$$

where

$$\frac{d}{dt}L'(t,\sigma) = \begin{bmatrix} F'(t) + [illegible]_{11}(t)\,[illegible](t) & I \\ [illegible]'(t)\,[illegible](t) & [illegible] \end{bmatrix} L'(t,\sigma);$$

$$V^{-1}(t) \triangleq U(t) = \begin{bmatrix} U_{11}(t) & U_{12}(t) \\ U_{12}'(t) & U_{22}(t) \end{bmatrix}$$

Since S_{12} is the (1,6) element of V, we have that

$$E^t\left[\int_0^t\int_0^{\sigma_1} S_{12}(t,\sigma_1,t)\ x_4(\sigma_2)\ d\sigma_2 d\sigma_1\right]$$

$$= E^t\left[\int_0^t S_{12}(t,\sigma_1,t)\ y_4(\sigma_2)\ d\sigma_1\right]$$

$$= E^t\left[\int_0^t e_6'\ L(t,\sigma)\ y_4(\sigma)d\sigma\right]V(t)e_1$$

$$\triangleq E^t[\delta'(t)]\ V(t)e_1 \qquad (32)$$

(where e_j is the j^{th} unit vector) can be computed with a finite dimensional linear estimator.

Similar reasoning implies that the third term in (30)

$$E^t\left[\int_0^t\int_0^{\sigma_1} S_{14}(t,\sigma_2,t)\ x_2(\sigma_1)\ d\sigma_2 d\sigma_1\right]$$

$$= E^t\left[\int_0^t\left(\int_0^{\sigma_1} e_3'\ L(t,\sigma_2)d\sigma_2\right) x_2(\sigma_1)d\sigma_1\right]\cdot$$

$$\cdot V(t)e_1 \triangleq E^t[\varepsilon'(t)]\ V(t)e_1 \qquad (33)$$

Can be computed with a finite dimensional linear estimator.

Let γ be defined by

$$d\gamma(t) = x_2(t)\ y_4(t)\ dt;\quad \gamma(0) = 0 \qquad (34)$$

Then the nonlinear filtering equation [22] yields

$$d\hat{\gamma}(t|t) = E^t[x_2(t)\ y_4(t)]\ dt$$

$$+ \{E^t[\gamma(t)\ x'(t)] - \hat{\gamma}(t|t)\ \hat{x}'(t|t)\}\cdot$$

$$\cdot H'(t)\ R^{-1}(t)\ [dz(t)-H(t)\ \hat{x}(t|t)dt] \qquad (35)$$

The first term in (35) (the drift term) can be written as

$$E^t[x_2(t)\ y_4(t)] = \hat{x}_2(t|t)\ \hat{y}_4(t|t) + S_{24}(t) \qquad (36)$$

which is computable "finite dimensionally". We have that the gain term

$$E^t[\gamma(t)\ x'(t)] - \hat{\gamma}(t|t)\ \hat{x}'(t|t)$$

$$= \int_0^t\int_0^{\sigma_1}\Big\{E^t[x_2(\sigma_1)x_4(\sigma_2)x'(t)]$$

$$- \hat{x}'(t|t)E^t[x_2(\sigma_1)x_4(\sigma_2)]\Big\}\ d\sigma_2 d\sigma_1$$

$$= \int_0^t\int_0^{\sigma_1}\Big\{K_2(t,\sigma_1)\quad E^t[x_4(\sigma_2)]$$

$$+ K_4(t,\sigma_2)\quad E^t[x_2(\sigma_1)]\Big\}\ d\sigma_2 d\sigma_1\ \cdot P(t)$$

$$= \left\{E^t\left[\int_0^t K\ (t,\sigma_1)\quad y_4(\sigma_1)d\sigma_1\right]\right.$$

$$\left. + E\left[\int_0^t x_2(\sigma_1)\int_0^{\sigma_1} K_4(t,\sigma_2)\quad d\sigma_2 d\sigma_1\right]\right\}\cdot P(t)$$

$$\triangleq \{E^t[\beta'(t)] + E^t[\alpha'(t)]\}P(t) \qquad (37)$$

where K_i denotes the $i^{\underline{th}}$ row of K. Since $P(\sigma_1,t,t)$ and $\int_0^{\sigma_1} P(\sigma_2,t,t)$ are both separable, (37) and hence (35) can be computed finite dimensionally. The proof for k=3 is now complete.

The optimal estimator for the 3x3 nilpotent system (10) consists of a Kalman filter for the augmented state consisting of x, y, α, β, δ, and ε (defined in (2), (11), (37), (37), (32), and (33) respectively with observations (4), together with the nonlinear stochastic equation (35). These are followed by nonlinear transformations as shown in (23)-(26) and (30).

A block diagram illustrating the nonlinear estimator is shown in Figure 1; notice that the estimator is driven by the innovations process $d\nu(t) = dz(t) - H(t)\ \hat{x}(t|t)\ dt$.

5. GENERALIZATIONS

[illegible]

Theorem 2: Consider (1)-(4), where $\mathscr{L}$ is solvable and $\mathscr{L}_0$ is nilpotent. Then the conditional mean $\hat{x}(t|t)$ can be computed with a finite dimensional system of nonlinear stochastic differential equations.

If, however, $\mathscr{L}$ is solvable and $\mathscr{L}_0$ is not nilpotent, the optimal estimator will be infinite dimensional. For example, if k=2 we must compute the conditional expectation

$$E^t\left[\int_0^t x_2(s)\ \exp\left(\int_s^t x_1(\tau)d\tau + \int_0^s x_3(\tau)d\tau\right)ds\right.$$

(see [16] for further details).

A proof similar to that for Theorem 2 yields a further generalization.

Theorem 3: Consider the equations (2)-(4) and

$$\dot{X}(t) = (A_0(t) + \sum_{i=1}^{N} u_i(t)A_i)X(t);\ X(0) = I \qquad (38)$$

Let $\mathscr{L} \triangleq \{A_1,\ldots,A_N,\ A_0(t)\ (\forall t)\}_{LA}$, and let $\mathscr{L}_0$ be the ideal in $\mathscr{L}$ generated by $\{A_1,\ldots,A_N\}$. Assume that $\mathscr{L}_0$ is nilpotent. Then the conditional mean $\hat{x}(t|t)$ can be computed with a finite dimensional system of nonlinear stochastic differential equations.

Remark: Notice that if $A_0(t)$ is time-varying, the nilpotency of $\mathscr{L}_0$ does not imply that $X(t)$ evolves on a solvable Lie group.

Theorem 1 can also be extended to other systems with polynomial nonlinearities (the proof is similar)

Theorem 4: Consider the linear system described by (2) and (4), and define

$$\gamma(t) = \int_0^t \int_0^{\sigma_1} \cdots \int_0^{\sigma_{k-1}} x_{j_1}(\sigma_{m_1}) \ldots x_{j_i}(\sigma_{m_i}) \cdot M_1(\sigma_{n_1}) \ldots M_\ell(\sigma_{n_\ell})d\sigma_1 \ldots d\sigma_k$$

where $\{M_i\}$ are arbitrary deterministic matrix-valued functions. The subscripts $\{j_\alpha\}$, $\{m_\alpha\}$, $\{n_\alpha\}$ are not necessarily distinct, and i and ℓ are not necessarily less than or equal to k. Then $\hat{\gamma}(t|t)$ and $E^t\left[e^{x_{j_{i+1}}(t)}\gamma(t)\right]$ can be computed with a finite dimensional system of nonlinear stochastic differential equations.

6. PATH-INDEPENDENT INTEGRALS

The results of this section are based upon the work of Brockett [19] and Gruber [13] on stability analysis using exact differentials and path-independent integrals. We will consider equation (2) with $x(t) \in R^n$, in which it will be assumed that each component $x_i(t)$ is m_i - times mean-square (m.s) differentiable (m_i may be equal to zero for some i); let k_j be the number of components of x which are at least j-times m.s. differentiable. Also, we define S^j to be the $k_j \times n$ selector matrix which selects the components of x that are at least j-times m.s. differentiable:

$$S^j_{1,i_1} = 1 \text{ if } m_{i_1} \geq j \text{ and } m_\ell < j \text{ for } 1 \leq \ell \leq i_1 - 1$$

$$S^j_{p,i_p} = 1 \text{ if } m_{i_p} \geq j \text{ and } m_\ell < j \text{ for } i_{p-1}+1 \leq \ell \leq i_p - 1, \text{ for } p>1$$

$$S^j_{p,q} = 0 \quad \text{otherwise}$$

Finally, let

$$\tilde{x}^{(j)} = \frac{d^j}{dt^j}\left[S^j x(t)\right] \qquad (39)$$

If $m \triangleq \max_{1\leq i\leq n} (m_i)$ and f is a continuous function, the random process

[illegible]

is said to be <u>independent of path</u> (in the mean-square) if there exists a function g such that

$$\gamma(t) = g(x(t),\ldots,\tilde{x}^{(m-1)}(t),x(0),\ldots,\tilde{x}^{(m-1)}(0)) \tag{41}$$

where equality in (41) is in the mean-square sense (these definitions could also be placed in the "almost sure" framework).

First we consider the case $m=1$. If g is twice continuously differentiable, and $x(0)$ is known, then $\hat{\gamma}(t|t)$ satisfies the nonlinear filtering equation [22]

$$\begin{aligned} d\hat{\gamma}(t|t) = {} & \{E^t[g_x'(x(t))F(t)x(t)] \\ & + \tfrac{1}{2}E^t[\mathrm{tr}(Q(t)g_{xx}(x(t)))]\}dt \\ & + \{E^t[\gamma(t)x'(t)] - \hat{\gamma}(t|t)\hat{x}'(t|t)\}\cdot \\ & \cdot H'(t)R^{-1}(t)[dz(t)-H(t)\hat{x}(t|t)dt] \end{aligned} \tag{42}$$

where g_x is the gradient and g_{xx} is the matrix of second partials. Since x and its derivatives are Gaussian, it is easy to see that (42) can be computed in terms of the conditional mean and covariance of x; thus the estimate $\hat{\gamma}(t|t)$ of a path-independent integral can be computed with a finite dimensional nonlinear estimator. This result can obviously be extended to the case in which $m > 1$.

<u>Example 3</u>: Define y_i as in (11). If

$$\gamma(t) = \int_0^t [x_i(\sigma)y_j(\sigma) + x_j(\sigma)y_i(\sigma)]d\sigma$$

$$= y_i(t)y_j(t) - y_i(0)y_j(0)$$

then $\hat{\gamma}(t|t)$ is finite dimensionally computable.

<u>Example 4</u>: Assume $x(0) = 0$, and let

$$\gamma(t) = \int_0^t < \dot{T}(\sigma)S^1x(\sigma) + T(\sigma)\tilde{x}^{(1)}(\sigma), \; T(\sigma)S^1x(\sigma) > d\sigma$$

where "<,>" denotes inner product. Then

$$\gamma(t) = < T(t)S^1x(t),\; T(t)S^1x(t) >$$

and $\hat{\gamma}(t|t)$ is finite dimensionally computable. A simple extension of these ideas is the following.

<u>Theorem 5</u>: Consider the linear system described by (2) and (4), where $x(0)$ is known. Let

$$\beta(t) = \gamma(t) + \delta(t) \tag{43}$$

where γ is a path-independent integral defined by (40), and $E^t[\delta(t)] = 0$. Then $\hat{\gamma}(t|t)$ can be computed with a finite dimensional system of nonlinear stochastic differential equations.

Finally, we state a theorem concerning the integrals of quadratic differentials, which is based on a result of Gruber [13].

<u>Theorem 6</u>: Consider the linear system described by (2) and (4), where $x(0)$ is known. Define the scalar differential operator

$$p(D) = \frac{d^m}{dt^m} + \sum_{i=0}^{m-1} p_i \frac{d^i}{dt^i} \tag{44}$$

$$p(s) = s^m + \sum_{i=0}^{m-1} p_0 s^i \tag{45}$$

and the $k_m \times k_m$ matrix differential operator

$$Q(D) = \sum_{i=0}^{q} Q_i \frac{d^i}{dt^i} \tag{46}$$

$$Q(s) = \sum_{i=0}^{q} Q_i s^i \tag{47}$$

where $q < m$. Assume that the matrix $R(s)$ satisfies

$$p(s)\,Q'(s) + p(-s)\,Q(s) = R'(-s)\,R(s) \tag{48}$$

Then

$$\gamma(t) = \int_0^t 2 < p(D)S^m x(\sigma),\; Q(D)S^m x(\sigma) > - < R(D)S^m x(\sigma),\; R(D)S^m x(\sigma) > d\sigma$$

is independent of path, and $\hat{\gamma}(t|t)$ can be computed with a finite dimensional system of nonlinear stochastic differential equations.

<u>Acknowledgement</u>: The authors would like to thank Professor Roger Brockett of Harvard University for many helpful discussions and for suggesting the use of path-independent integrals in the present context.

References

1. R.W. Brockett, "System theory on group manifolds and coset spaces," vol. 10, pp. 265-284, May, 1972.

2. R.W. Brockett, "Lie algebras and Lie groups in control theory," in Geometric Methods in System Theory, D.Q. Mayne and R.W. Brockett, Eds. The Netherlands: Reidel, 1973.

3. R.W. Brockett, "Lie theory and control systems on spheres," SIAM J. Appl. Math., vol. 25, pp. 213-225, Sept. 1973.

4. A.S. Willsky and S.I. Marcus, "Analysis of bilinear noise models in circuits and devices," Monograph of the Colloquium on the Application of Lie Group Theory to Nonlinear Network Problems, 1974 IEEE Int. Symp. on Circuits and Systems, San Francisco, April 1974.

5. A.S. Willsky and S.I. Marcus, "Estimation for Bilinear Stochastic Systems," M.I.T., Electronic Systems Laboratory Report ESL-R-544; presented at the U.S. - Italy Seminar on Variable Structure Systems, Portland, Oregon, May 1974.

6. J.T. Lo, "Bilinear stochastic systems and finite dimensional sensor orbits," Proc. of the Eighth Princeton Conf. on Information Sciences and Systems, Princeton, N.J., March, 1974.

7. A.S. Willsky, Dynamical Systems Defined on Groups: Structural Properties and Estimation, Ph.D. Thesis, Dept. of Aeronautics and Astronautics, M.I.T., Cambridge, Mass. June 1973.

8. J.T. Lo and A.S. Willsky, "Estimation for rotational processes with one degree of freedom: Parts, I,II,III," IEEE Trans. Automat. Contr., to appear.

9. A.S. Willsky, "Some estimation problems on Lie groups," in Geometric Methods in System Theory, D.Q. Mayne and R.W. Brockett, Eds. The Netherlands: Reidel, 1973.

10. A.S. Willsky, "Fourier series and estimation on the circle with applications to synchronous communication: Part I,II," IEEE Trans. Inform. Theory, vol. IT-20, Sept. 1974.

11. C. Bruni, G. DiPillo, and G. Koch, "Bilinear systems: an appealing class of 'nearly linear' systems in theory and applications," IEEE Trans. Automat. Contr., Vol. AC-19, pp. 334-348, Aug. 1974.

12. G.L. Blankenship, "Perturbation theory for stochastic ordinary differential equations with applications to optical waveguide analysis," Monograph of the Colloquium on the Application of Lie Group Theory to Nonlinear Network Problems, 1974 IEEE Int. Symp. on Circuits and Systems, San Francisco, April 1974.

13. M. Gruber, Stability Analysis Using Exact Differentials, Ph.D. Thesis, Dept. of Electrical Engineering, M.I.T., Cambridge, Mass., 1968.

14. A.H. Jazwinski, Stochastic Processes and Filtering Theory, New York: Academic Press, 1970.

15. D.N. Martin, Stability Criteria for Systems with Colored Multiplicative Noise, Ph.D. Thesis, Dept. of Electrical Engineering, M.I.T., Cambridge, Mass., June 1974.

16. S.I. Marcus, Estimation and Analysis of Nonlinear Stochastic Systems, Ph.D. Thesis, Dept. of Electrical Engineering, M.I.T., Cambridge, Mass., June 1975.

17. W. Rudin, Real and Complex Analysis, New York: McGraw-Hill, 1966.

18. W.M Wonham, "On a matrix Riccati equation of stochastic control," SIAM J. Control, vol. 6, pp. 681-697, 1968.

19. R.W. Brockett, Finite Dimensional Linear Systems, New York: John Wiley and Sons, 1970.

20. C.T. Leondes, J.B. Peller, and E.B. Stear, "Nonlinear smoothing theory," IEEE Trans. Syst. Sci. Cyber, vol. SSC-6, p.63, 1970.

21. T. Kailath and P. Frost, "An innovations approach to least-squares estimation, Part II: Linear smoothing in additive white noise," IEEE Trans. Automat. Contr., vol. AC-13, pp. 655-660, 1968.

22. M. Fujisaki, G. Kallianpur, and H. Kunita, "Stochastic differential equations for the nonlinear filtering problem," Osaka J. Math., vol. 9, pp. 19-40, 1972.

23. A.A. Sagle and R.E. Walde, Introduction to Lie Groups and Lie Algebras, New York: Academic Press, 1973.

24. H. Kwakernaak, "Optimal filtering in linear systems with time delays," IEEE Trans. Automat. Contrl., vol. AC-12, pp.169-173, 1967.

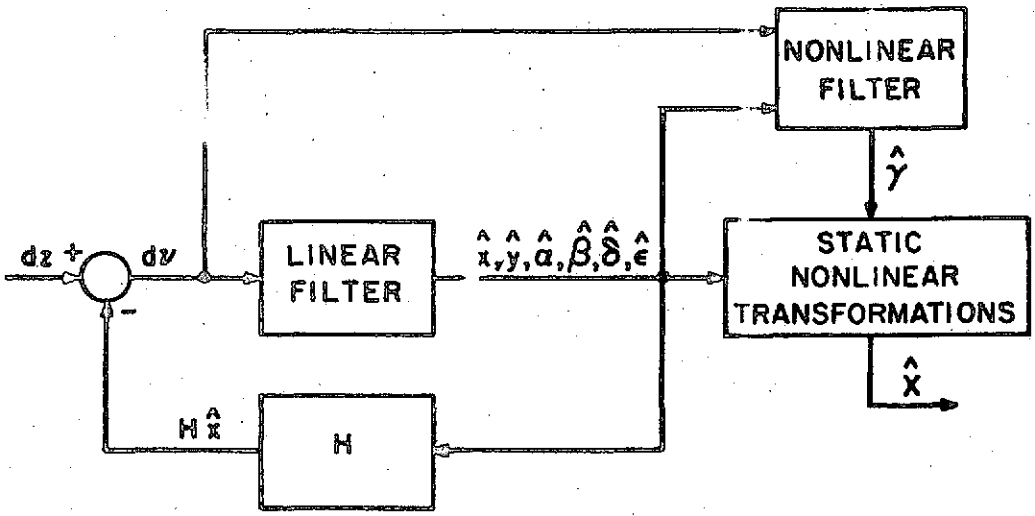

Fig. 1 Block Diagram for Optimal Nonlinear Estimation on gn(3)