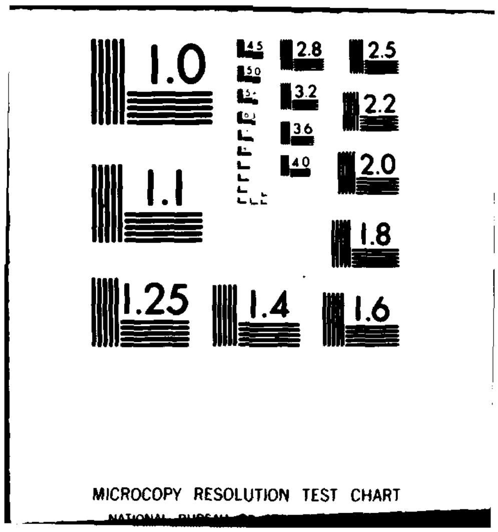
AD-A091 032 TEXAS UNIV AT AUSTIN DEPT OF ELECTRICAL ENGINEERING F/G 12/1
THE LIE ALGEBRAIC STRUCTURE OF A CLASS OF FINITE DIMENSIONAL NO--ETC(U)
JUL 80 C LIU, S I MARCUS F49620-77-C-0101
UNCLASSIFIED AFOSR-TR-80-0879 NL

END
DATE
FILMED
12-80
DTIC


1.0
4.5
5.0
2.8
2.5
3.2
2.2
3.6
4.0
2.0
1.1
1.8
1.25
1.4
1.6
MICROCOPY RESOLUTION TEST CHART

AFOSR-TR- 80-0879

(4) LEVEL II

AD A091032

THE LIE ALGEBRAIC STRUCTURE OF A CLASS OF FINITE DIMENSIONAL NONLINEAR FILTERS *

Chang-Huan Liu and Steven I. Marcus

Dept. of Electrical Engineering
University of Texas at Austin
Austin, Texas 78712

ABSTRACT.

We present an example of the application of Lie algebraic techniques to nonlinear estimation problems. The method relates the computation of the (unnormalized) conditional density and the computation of statistics with finite dimensional estimators. The general method is explained; for a particular example, the structures of the Lie algebras associated with the unnormalized conditional density equation and the finite dimensionally computable conditional moment equations are analyzed in detail. The relationship between these Lie algebras is studied, and the implications of these results are discussed.

*. Supported in part by the Air Force Office of Scientific Research (AFSC) under Grant. AFOSR-79-0025, in part by the National Science Foundation under Grant ENG 76-11106, and in part by the Joint Services Electronics Program under Contract F 49620-77-C-0101.

Filterdag Rotterdam 1980, M. Hazewinkel, Editor, Report of the Econometric Institute, Erasmus University, Rotterdam, 1980.

1. INTRODUCTION.

This paper is concerned with the optimal recursive estimation of the state x_t of a nonlinear stochastic system, given the past observations $z^t = \{z_s, 0 \le s \le t\}$. Specifically, we consider systems of the form

$$(1) \qquad \begin{aligned} dx_t &= f(x_t)dt + G(x_t)dw_t \\ dz_t &= h(x_t)dt + R_t^{\frac{1}{2}}dv_t \end{aligned}$$

where w and v are independent unit variance vector Wiener processes, f and h are vector-valued functions, G is a matrix-valued function, and R > 0. The optimal (minimum-variance) estimate is of course the conditional mean $\hat{x}_t = E[x_t|z^t]$ (also denoted $\hat{x}_{t|t}$ or $E^t[x_t]$); $\hat{x}_t$ satisfies the (Ito) stochastic differential equation [1] - [3]

$$(2) \qquad \begin{aligned} d\hat{x}_t = &[\hat{f}(x_t) - (\widehat{x_t h^T} - \hat{x}_t\hat{h}^T)R^{-1}(t)\hat{h}]dt \\ &+ (\widehat{x_t h^T} - \hat{x}_t\hat{h}^T)R^{-1}(t)dz_t \end{aligned}$$

where ^ denotes conditional expectation given z^t and h denotes $h(x_t)$. Also, the conditional probability density p(t,x) of x_t given z^t (we will assume that p(t,x) exists) satisfies the stochastic partial differential equation [3], [4]

$$(3) \qquad dp(t,x) = \mathcal{L}p(t,x)dt + (h(x)-\hat{h}(x))^T R^{-1}(t)(dz_t - \hat{h}(x)dt)p(t,x)$$

where

$$(4) \qquad \mathcal{L}(.) = -\sum_{i=1}^{n} \frac{\partial(.f_i)}{\partial x_i} + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} \frac{\partial^2(.(GG^T)ij)}{\partial x_i \partial x_j}$$

is the forward diffusion operator.

Notice that the differential equation (2) is not recursive, and indeed appears to involve an infinite dimensional computation in general. Aside from the linear-Gaussian case in which the Kalman filter is optimal, there are very few known cases in which the optimal estimator is fin dimensional (a number of these are summarized in [5]). However, in [6] - [8] we have shown that for certain classes of nonlinear stochastic systems in continuous and discrete time, the conditional mean can be computed with a recursive filter of fixed finite dimension. The typical nonlinear system in these classes consists of a linear system with linear measurements, which feeds forward into a nonlinear system described by a certain type of Volterra series expansion or by a bilinear system satisfying certain algebraic conditions. The major purpose of this paper is to consider these estimation problems from a new perspective, and to gain much deeper insight into their structure.

The new perspective, originally proposed by Brockett [9] (see also [10], [11]), takes the following approach to the general estimation problem (1) (we assume for simplicity that z is a scalar). Instead of studying the equation (3) for the conditional density, we consider the Zakai equation for an unnormalized conditional density $\rho(t,x)$ [12]:

$$(5) \qquad d\rho(t,x) = \mathcal{L}\rho(t,x)dt + h(x)\rho(t,x)dz_t$$

where $\rho(t,x)$ is related to $p(t,x)$ by the normalization

$$(6) \qquad p(t,x) = \rho(t,x).(\int\rho(t,x)dx)^{-1}.$$

The Zakai equation (5) is much simpler than (3); indeed, (5) is a bilinear differential equation [13] in ρ, with z considered as the input. This is the first clue that the Lie algebraic and differential geometric techniques developed for finite dimensional systems of this type may be brought to bear here. Suppose that some statistic of the conditional distribution of x_t given z^t can be calculated with a finite dimensional recursive estimator of the form

$$(7) \qquad d\eta_t = a(\eta_t)dt + b(\eta_t)dz_t$$

$$(8) \qquad E[c(x_t)|z^t] = \gamma(\eta_t)$$

where η evolves on a finite dimensional manifold, and a and b are suitably smooth. Of course, this statistic can also be obtained from $\rho(t,x)$: by

(9) $\quad E[c(x_t)|z^t] = \int c(x)\rho(t,x)dx\ (\int\rho(t,x)dx)^{-1}$

For Lie-algebraic calculations, it is more convenient to write (5) and (3) in Fisk-Stratonovich form (so that they obey the ordinary rules of calculus)

(10) $\quad d\eta_t = \tilde{a}(\eta_t)dt + b(\eta_t)dz_t$

(11) $\quad d\rho(t,x) = [\mathcal{L} - \frac{1}{2} h^2(x)]\rho(t,x)dt + h(x)\rho(t,x)dz_t$

where the $i^{\underline{th}}$ component $\tilde{a}_i(\eta) = a_i(\eta) - \frac{1}{2}\sum_j b_j(\eta)\frac{\partial b_i}{\partial \eta_j}(\eta)$

(Beginning with (10), all equations will be in Fisk-Stratonovich form, unless otherwise indicated). The two systems (10), (8) and (11), (9) are thus two representations of the same mapping from "input" functions z to "outputs" $E[c(x_t)|z^t]$: (11), (9) via a bilinear infinite dimensional state equation, and (10), (8) via a nonlinear finite dimensional state equation. Generalizing the results of [14], [15] to infinite dimensional state equations, the major assertion of [9] is that, under appropriate hypotheses, the Lie algebra F generated by $\tilde{a}$ and b (under the commutator $[\tilde{a},b] = \frac{\partial b}{\partial \eta}\tilde{a} - \frac{\partial \tilde{a}}{\partial \eta} b$) is a homomorphic image of the Lie algebra L generated by $A_o = \mathcal{L} - \frac{1}{2} h^2(x)$ and $B_o = h(x)$ (under the commutator $[A_o,B_o] = A_oB_o-B_oA_o$). Conversely, any homomorphism of L onto a Lie algebra generated by two complete vector fields on a finite dimensional manifold allows the computation of some information about the conditional density with a finite dimensional estimator of the form (10).

In [9], this approach is explicitly carried out and analyzed for the problem in which f, G and h (1) in are all linear. In that case, the Lie algebra L of the Zakai equation is finite dimensional and the unnormalized conditional density can in fact be computed with a finite dimensional estimator, the Kalman filter. In this paper, we carry out a similar analysis for the simplest example of the class considered in [6] - [8]. For this example, all conditional moments of the state can be computed with finite dimensional estimators; the Lie algebra L is <u>infinite</u> dimensional but has many finite dimensional homomorphic images (the Lie algebras of the finite dimensional estimators), thus yielding a very interesting structure. The example to be considered has state equations

$$dx_t = dw_t \tag{12}$$
$$dy_t = x_t^2 dt$$

with observations

$$dz_t = x_t dt + dv_t \tag{13}$$

where v and w are unit variance Wiener processes, $\{x_o, y_o, v, w\}$ are independent, and x_o is Gaussian. The computation of $\hat{x}_t$ is of course straightforward by means of the Kalman filter, but the computation of $\hat{y}_t$ requires a nonlinear estimator.

2. THE LIE ALGEBRA OF THE UNNORMALIZED CONDITIONAL DENSITY EQUATION.

For the system (12) - (13), the equation (5) in Fisk-Stratonovich form is

$$d\rho(t,x) = \left(-x^2\frac{\partial}{\partial y} + \frac{1}{2}\frac{\partial^2}{\partial x^2} - \frac{1}{2}x^2\right)\rho(t,x)dt + x\rho(t,x)dz_t, \tag{14}$$

so the Lie algebra L is generated by $A_o = -x^2\frac{\partial}{\partial y} + \frac{1}{2}\frac{\partial^2}{\partial x^2} - \frac{1}{2}x^2$ and $B_o = x$.

The following theorem is straightforward to prove.

<u>Structure theorem 1:</u>

(i) The Lie algebra L generated by A_o and B_o has as basis the elements A_o and B_i, C_i, D_i, $i = 0,1,2, \ldots$, where

$$B_i = x\frac{\partial^i}{\partial y^i} \qquad i = 0,1,2,\ldots$$

$$C_i = \frac{\partial}{\partial x}\frac{\partial^i}{\partial y^i} \qquad i = 0,1,2,\ldots$$

$$D_i = \frac{\partial^i}{\partial y^i} \qquad i = 0,1,2,\ldots$$

(ii) The commutation relations are given by

$$[A_o, B_i] = C_i, \quad \forall\, i$$

$$[A_o, C_i] = B_i + 2B_{i+1}, \quad \forall\, i$$

$$[A_o, D_j] = [B_i, D_j] = [C_i, D_j] = 0, \quad \forall\, i,j$$

$$[B_i, C_j] = -D_{i+j}, \quad \forall\, i,j$$

$$[B_i, B_j] = [C_i, C_j] = 0, \quad \forall\, i \neq j$$

(iii) The center of L is $\{D_i,\ i = 0,1,2,\ldots\}$.

(iv) Every ideal of L has finite codimension; i.e., for any ideal I, the quotient L/I is finite dimensional.

(v) Let I_j be the ideal generated by B_j, with basis $\{B_i, C_i, D_i;\ i \geq j\}$. Then $I_o \supset I_1 \supset \ldots$ and $\bigcap_j I_j = \{0\}$, so that the canonical map $\pi: A \to \bigoplus_j A/I_j$ is injective.

(vi) L is the semidirect sum [18] of A_o and the nilpotent ideal I_o; hence L is solvable.

In light of the remarks in the previous section, it should be expected that many statistics of the conditional distribution can be computed with finite dimensional estimators, since there are an infinite number of finite dimensional quotients(homomorphic images) L/I . By Ado's theorem, these can be realized by bilinear systems. However, we will present a slightly different realization of the sequence of quotients in (vi) above: L/I_1 is realized by the Kalman filter for $\hat{x}_t$ (L/I_1 is the oscillator algebra [9] - [11]), and L/I_j $(j \geq 2)$ is realized by the estimator which computes $\hat{x}_t$ and $\widehat{y^i_t} = E[y^i_t | z^t]$ $(i = 1,2,\ldots,j-1)$. Of course, the dimension of L/I_j increases with j, so we will only present the estimator equations for $j = 4$ in the next section. Other sequences of quotients possessing the property (vi) can also be realized (e.g., those generated by the $\{C_j\}$), but those realizations do not have as natural an interpretation in terms of conditional moments.

The properties (iv) and (v) of the structure theorem are useful for an "estimation algebra" to possess, in the following sense: they basically say that L has enough finite dimensional quotients that it is determined by their direct sum. Translating this into an estimation context via the reasoning of the previous section, if we can realize all the quotients with finite dimensionally computable statistics, then these properties give us hope of being able to approximate the conditional density (or conditional characteristic function) with a convergent series of functions of these statistics, even if the conditional density cannot be computed exactly by a finite dimensional estimator.

3. THE LIE ALGEBRA OF THE FINITE DIMENSIONAL ESTIMATOR.

The method of [6] for computing the finite dimensional estimator for $\hat{y}_t$ systematically uses the estimation equation (2) and the fact that the conditional density of $x_\cdot$ given $z_\cdot$ is Gaussian to express higher order moments in terms of lower. This procedure can also be applied to obtain equations for higher order conditional moments of y for the estimation problem (12) - (13). The first three conditional moments of y_t, together with $\hat{x}_t$ and the necessary auxiliary filter states are computed recursively by the finite dimensional estimator (in Fisk-Stratonovich form, with explicit time-dependent notation omitted):

$$(15)\quad d\begin{bmatrix} \hat{x} \\ \hat{\xi} \\ \hat{y} \\ \hat{\theta} \\ \widehat{y^2} \\ \hat{\phi} \\ \widehat{y^3} \\ \\ \\ \\ t \end{bmatrix} = \underbrace{\begin{bmatrix} -\hat{x}P \\ \hat{x}(1-P_{12}) - \hat{\xi}P^{-1} \\ \hat{x}^2 - 2\hat{x}\hat{\xi}P + P - PP_{12} \\ \hat{x}(P_{12}-P_{13}) + \hat{\xi}P(1-P_{12}) - \hat{\theta}P^{-1} \\ 2\hat{x}^2\hat{y} + 2\hat{y}P + 8\hat{x}\hat{\xi}P + 2PP_{12} - 4\hat{x}\hat{\xi}\hat{y}P - 8\hat{x}\hat{\theta}P - 2\hat{y}PP_{12} - 4\hat{\xi}^2P^2 - 4PP_{13} \\ \hat{x}(P_{13}-P_{14}) + \hat{\xi}P(P_{12}-P_{13}) + \hat{\theta}(P-PP_{12}) - \hat{\phi}P^{-1} \\ 3\hat{x}^2\widehat{y^2} + 3\widehat{y^2}P + 24\hat{x}\hat{\xi}\hat{y}P + 48\hat{x}\hat{\theta}P + 24\hat{\xi}^2P^2 + 6\hat{y}PP_{12} \\ + 24PP_{13} - 3\widehat{y^2}PP_{12} - 48\hat{\xi}\hat{\theta}P^2 - 12\hat{y}PP_{13} - 12\hat{\xi}^2\hat{y}P^2 - 24PP_{14} \\ - 6\hat{x}\hat{\xi}\widehat{y^2}P - 24\hat{x}\hat{\theta}\hat{y}P - 48\hat{x}\hat{\phi}P \\ \\ 1 \end{bmatrix}}_{a_0} dt$$

$$
\begin{bmatrix} P \\ P_{12} \\ 2\hat{\xi}P \\ P_{13} \\ 4\hat{\xi}\hat{y}P + 8\hat{\theta}P \\ P_{14} \\ 6\hat{\xi}\widehat{y^2}P + 24\hat{\theta}\hat{y}P + 48\hat{\phi}P \\ 0 \end{bmatrix} dz \quad ; \quad \underbrace{\begin{bmatrix} \hat{x}_o \\ \hat{\xi}_o \\ \hat{y}_o \\ \hat{\theta}_o \\ \widehat{y^2_o} \\ \hat{\phi}_o \\ \widehat{y^3_o} \\ t_o \end{bmatrix}}_{b_o} = \begin{bmatrix} E[x_o] \\ 0 \\ E[y_o] \\ 0 \\ E[y_o^2] \\ 0 \\ E[y_o^3] \\ 0 \end{bmatrix}
$$

where the nonrandom conditional covariance equations are

$$
\begin{aligned}
&\dot{P} = 1 - P^2 \\
&\dot{P}_{12} = P - (P+P^{-1})P_{12} \\
(16) \qquad &\dot{P}_{13} = 2PP_{12} - PP_{12}^2 - (P+P^{-1})P_{13} \\
&\dot{P}_{14} = 2PP_{13} + PP_{12}^2P_{13} - (P+P^{-1})P_{14} \\
&P(0) = \mathrm{cov}(x_o) \neq 0; \; P_{12}(0) = P_{13}(0) = P_{14}(0) = 0
\end{aligned}
$$

The estimator (15) is obtained by first augmenting the state x with auxiliary states ξ, θ, and ϕ; then the Kalman filter for the linear system with states $[x,\xi,\theta,\phi]$ and observations z computes $[\hat{x},\hat{\xi},\hat{\theta},\hat{\phi}]$. In addition, $[P,P_{12},P_{13},P_{14}]$ is the first row of the Kalman filter error covariance matrix; (16) is obtained by selecting the corresponding components of the Riccati equation. Then $\hat{y}$, $\widehat{y^2}$, and $\widehat{y^3}$ are seen, after tedious calculations, to be computed by the given equations (some of the calculations are presented in the Appendix, in order to illustrate the method). The filter state is augmented with t in order to make (15) time-invariant thus facilitating the use of Lie algebraic techniques. The filter (15) can be viewed as a cascade of linear filters [19]: $[\hat{x},\hat{\xi},\hat{\theta},\hat{\phi},t]$ satisfies a linear equation; some of these states then feed forward and can be viewed as parameters in a linear equation for $\hat{y}$; the states $\hat{x},\hat{\xi},\hat{\theta},\hat{y},t$ then feed forward as parameters into a linear equation for $\widehat{y^2}$; etc. This structure is typical of the class of finite dimensional estimators derived in [6] - [8].

In order to study the structure of the estimation problem as discussed in section 1, we must analyze the Lie algebra F generated by a_o and b_o in (15). The structure of the class of problems of [6] is analyzed from a different point of view in [21].

<u>Structure theorem 2:</u>

(i) F has as basis the elements a_o; b_i, c_i, $i = 0,1,2,3$; d_i, $i = 1,2,3$, where a_o and b_o are given in (15) and

$$c_o = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad b_1 = \begin{bmatrix} 0 \\ 1 \\ 2\hat{x} \\ P_{12} \\ 4\hat{x}\hat{y} + 8\hat{\xi}P \\ P_{13} \\ 6\hat{x}\hat{y}^2 + 24\hat{\xi}\hat{y}P + 48\hat{\theta}P \\ 0 \end{bmatrix}, \quad c_1 = \begin{bmatrix} 0 \\ P^{-1} \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix},$$

$$b_2 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 8\hat{x} \\ P_{12} \\ 24\hat{x}\hat{y} + 48\hat{\xi}P \\ 0 \end{bmatrix}, \quad c_2 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ P^{-1} \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad b_3 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 48\hat{x} \\ 0 \end{bmatrix}, \quad c_3 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ P^{-1} \\ 0 \\ 0 \end{bmatrix}$$

$$d_1 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 2\hat{y} \\ 0 \\ 3\hat{y}^2 \\ 0 \end{bmatrix}, \quad d_2 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 3\hat{y} \\ 0 \end{bmatrix}, \quad d_3 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}$$

(ii) The commutation relations are given by

$$[a_o,b_i] = c_i \quad , \quad i = 0,1,2,3$$

$$[a_o,c_i] = b_i - b_{i+1} \, , \quad i = 0,1,2$$

$$[a_o,c_3] = b_3$$

$$[b_i,c_j] = \begin{cases} -2d_{i+j} & i+j=1 \\ -8d_{i+j} & i+j=2 \\ -48d_{i+j} & i+j=3 \\ 0 & \text{otherwise} \end{cases}$$

$$[a_o,d_j] = [b_i,d_j] = [c_i,d_j] = 0, \ \forall \ i,j$$

(iii) Let $\tilde{I}_4$ be the ideal in L with basis B_i, C_i, D_i, $i \geq 4$ <u>and</u> D_o. Then F is isomorphic to $L/\tilde{I}_4$; hence, F is also solvable.

(iv) The isomorphism ϕ between L and $F / \tilde{I}_4$ is given by:

$\phi(A_o) = a_o$; $\phi(B_i) = (-\frac{1}{2})^i b_i$,

$\phi(C_i) = (-\frac{1}{2})^i c_i$, $i = 0,1,2,3$; $\phi(D_i) = (-1)^i (i!) d_i$; $i = 1,2,3$;

$\phi(E) = 0$, $E \in \tilde{I}_4$.

(v) F is the semidirect sum of a_o and the nilpotent ideal generated by b_c.

<u>Remarks</u>:

(i) The estimator (15) is not quite a realization of L / I_4, since D_o is also in the kernel of the homomorphism (i.e., the ideal $\tilde{I}_4$). However, a finite dimensional estimator realizing L/ I_4 (or L/I_j; for any;) is easily obtained by augmenting (15) with the equation for the normalization factor α_t for $\rho(t,x)$ (the denominator of (6)) which satisfies (in Ito form)

$$d\alpha_t = \hat{x}_t \, \alpha_t \, dz_t$$

or (in Fisk-Stratonovich form)

(17) $$d\alpha_t = -\frac{1}{2}(\hat{x}_t^2 + P_t)\alpha_t dt + \hat{x}_t \alpha_t dt$$

If (17) is augmented at the end of (15), the Lie algebra generated by a_o and b_o has the same commutation relations as in (ii) above, except that

$$[b_o, c_o] = \begin{bmatrix} 0 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ 0 \\ \alpha \end{bmatrix} \triangleq d_o$$

and d_o commutes with all the other elements. Thus a realization of L/I_4 is an easy modification of (15), so we will concentrate on (15).

(ii) The property (v) is typical of a cascade of linear systems.

(iii) One of the conditions in [9] for the existence of a Lie algebra homomorphism from L to the Lie algebra of a finite dimensional estimator is that the estimator be a "minimal" realization in a certain sense. If we consider the output of (15) to be $\widehat{y^3}$ and consider this realization of the input-output map from z to $\widehat{y^3}$, then it can be verified by the methods of [15] that the realization is locally weakly controllable and locally weakly observable. This implies that there is no other realization with lower dimension; it is in this sense that the statistics $\hat{\xi}$, $\hat{\theta}$, $\hat{\phi}$ are necessary for the computation of $\widehat{y^3}$.

Images of L under homomorphisms with successively larger kernels can be realized by using only certain of the equations in the finite dimensional estimator (15); that is, some subset of the equations (15) will generate a Lie algebra isomorphic to L/I. Let Υ_j denote the ideal with basis D_o and B_i, C_i, D_i, $i \geq j$; we will also use the notation that, e.g., $\tilde{\Upsilon}_j \oplus D_\ell$ denotes the ideal with basis the above elements and D_ℓ (which is in the center of L). Realizations of some of the many possible quotients are summarized in the following table, which gives the quotient along with the set of states of a finite dimensional estimator which realizes it (the filter states satisfy the corresponding equations in (15)). For example, $L/\tilde{I}_3$ is realized by (15) with the equations for $\hat{\phi}$ and $\widehat{y^3}$ omitted, with all the other filter states retained.

QUOTIENT	REALIZATION
$L/\tilde{I}_4$	$\hat{x}, \hat{\xi}, \hat{y}, \hat{\Theta}, \widehat{y^2}, \hat{\phi}, \widehat{y^3}, t$
$L/(\tilde{I}_4 \oplus D_3)$	$\hat{x}, \hat{\xi}, \hat{y}, \hat{\Theta}, \widehat{y^2}, \hat{\phi}, t$
$L/(\tilde{I}_4 \oplus D_2 \oplus D_3)$	$\hat{x}, \hat{\xi}, \hat{y}, \hat{\Theta}, \hat{\phi}, t$
$L/(\tilde{I}_4 \oplus D_1 \oplus D_2 \oplus D_3)$	$\hat{x}, \hat{\xi}, \hat{\Theta}, \hat{\phi}, t$
$L/\tilde{I}_3$	$\hat{x}, \hat{\xi}, \hat{y}, \hat{\Theta}, \widehat{y^2}, t$
$L/(\tilde{I}_3 \oplus D_2)$	$\hat{x}, \hat{\xi}, \hat{y}, \hat{\Theta}, t$
$L/(\tilde{I}_3 \oplus D_1 \oplus D_2)$	$\hat{x}, \hat{\xi}, \hat{\Theta}, t$
$L/\tilde{I}_2$	$\hat{x}, \hat{\xi}, \hat{y}, t$
$L/(\tilde{I}_2 \oplus D_1)$	$\hat{x}, \hat{\xi}, t$
$L/\tilde{I}_1$	$\hat{x}, t$

Table 1. Realization of some quotients.

The results of Table 1 follow from two observations: first, that if I and J are ideals of L such that $I \subset J$, then J/I is an ideal of L/I and $(L/I)/(J/I)$ is naturally isomorphic to L/J [16,p.8] (e.g., $I = \tilde{I}_4$ and $J = \tilde{I}_3$). Also, it is clear that if one defines homomorphisms from $L/\tilde{I}_4$ to the quotients in Table 1 by the canonical map, then the image can be realized by (15) with certain equations omitted. For example, it is clear that sending $d_3 \to 0$ can be accomplished by eliminating the equation for $\widehat{y^3}$; each d_i thus represents, in some sense, $\widehat{y^i}$. Notice, in particular, that $A/\tilde{I}_1$ is realized by the Kalman filter for $\hat{x}$.

Other interesting quotients are obtained by homomorphisms which send other elements of the center of $L/\tilde{I}_4$, say just d_1, to zero. However, such a quotient is more difficult to realize by an estimator, since the realization is not obtained by merely eliminating certain equations. For these quotients, the following result leads to a realization.

Proposition 1: Let F be the Lie algebra generated by two n-dimensional vector fields a and b. Assume that there is an element d in the center of F and a constant n-vector β such that $\beta'd = 1$ (prime denotes transpose). Then the mapping ϕ with $\phi(a) = a - (\beta'a)d$ and $\phi(b) = b - (\beta'b)d$ extends to a Lie algebra homomorphism with $\phi(f) = f - (\beta'f)d$ for all $f \in F$, $\phi(d) = 0$, and $\phi(F)$ isomorphic to $F/\{d\}$.

<u>Proof</u>: We must show that, for $f,g \in F$,

$$\phi([f,g]) = [f,g] - (\beta'[f,g])d = [\phi(f), \phi(g)].$$

Now, since $\beta' f$ and $\beta' g$ are functions (not constants),

$$\begin{aligned}
[\phi(f), \phi(g)] &= [f-(\beta'f)d, g-(\beta'g)d] \\
&= [f,g] - [(\beta'f)d, g] - [f, (\beta'g)d] + [(\beta'f)d, (\beta'g)d] \\
&= [f,g] - \{(\beta'f)[d,g] - g(\beta'f)d\} \\
&- \{(\beta'g)[f,d] + f(\beta'g)d\} \\
&+ \{(\beta'f)(\beta'g)[d,d] + (\beta'f)d(\beta'g)d - (\beta'g)d(\beta'f)d\}
\end{aligned}$$

Notice that, for any $f \in F$, $\beta'[f,d] = 0$ and $\partial(\beta'd)/\partial x = 0$ imply that

$$d(\beta'f) = \frac{\partial(\beta'f)}{\partial x} d = \frac{\partial(\beta'd)}{\partial x} f = 0.$$

Thus

$$\begin{aligned}
[\phi(f), \phi(g)] &= [f,g] - \{-g(\beta'f) + f(\beta'g)\}d \\
&= [f,g] - (\beta'[f,g])d.
\end{aligned}$$

Note finally that $\phi(d) = d - (\beta'd)d = d - d = 0$.

This result can be applied, for example, to $F = L/\tilde{\mathcal{I}}_4$ and d_1, since d_1 is in the center and the third component of d_1 equals 1 (thus $\beta = [0\ 0\ 1\ 0\ \dots\ 0]'$). The proposition implies that if we implement (15) with a_o, b_o replaced by $a_o - (\beta' a_o)d_1$ and $b_o - (\beta' b_o)d_1$, respectively, then the resulting estimator (call it (15')) will generate a Lie algebra isomorphic to $L/(\tilde{\mathcal{I}}_4 \oplus D_1)$. Notice that this transformation (due to the form of d_1) eliminates the $\hat{y}$ equation, modifies the $\widehat{y^2}$ and $\widehat{y^3}$ equations, and does not affect the others. From another point of view, the right-hand side of (15) is transformed from $a_o dt + b_o dz_t$ to

$$(18) \quad \begin{aligned}
&a_o dt + b_o dz_t - d_1[(\beta' a_o)dt + (\beta' b_o)dz_t] \\
&= a_o dt + b_o dz_t - d_1 d\hat{y}_t
\end{aligned}$$

Denoting the statistics in this estimator which replace $\widehat{y^2}$ and $\widehat{y^3}$ by $\widetilde{y^2}$ and $\widetilde{y^3}$, respectively, we see from (18) and the form of d_1 that

$$d\widetilde{y_t^2} = d\widehat{y_t^2} - 2\hat{y}_t d\hat{y}_t = d\widehat{y_t^2} - d(\hat{y}_t)^2;$$

thus this estimator computes the conditional second <u>central</u> moment $E[(y_t-\hat{y}_t)^2/z^t]$, rather than the second moment $\widehat{y_t^2}$. However,

$$\begin{aligned} d\widetilde{y_t^3} &= d\widehat{y_t^3} - 3\widehat{y_t^2}\,d\hat{y}_t \\ &= (24\hat{x}\hat{\xi}\hat{y}P + 48\hat{x}\hat{\Theta}P + 24\hat{\xi}^2P^2 + 6\hat{y}PP_{12} + 24PP_{13} - 48\hat{\xi}\hat{\Theta}P^2 \\ &\quad - 12\hat{y}PP_{13} - 12\hat{\xi}^2\hat{y}P^2 - 24PP_{14} - 24\hat{x}\hat{y}\hat{\Theta}P - 48\hat{x}\hat{\phi}P)dt \\ &\quad + (24\hat{\Theta}\hat{y}P + 48\hat{\phi}P)dz_t \end{aligned}$$

which is not the equation for any easily recognized statistic of the conditional distribution of y_t given z^t. On the other hand, the results of [17] - [18] imply that, since there is a Lie algebra homomorphism from the Lie algebra F of (15) to that of (15') and the isotropy subalgebra of F is {0} at every point, then there is (at least locally) an analytic map λ that carries solutions of (15) into those of (15'). We have already seen that λ takes the components $\hat{x}$, $\hat{\xi}$, $\hat{\Theta}$, $\hat{\phi}$, t into themselves, $\lambda(\widehat{y_t^2}) = E[(y-\hat{y}_t)^2/z^t]$, and $\lambda(\hat{y}_t) = 0$. The image $\lambda(\widehat{y_t^3})$ is difficult to compute, although a method is given in [17]; to first order for small t, $\widetilde{y_t^3} \simeq \widehat{y_t^3} - 3\widehat{y_o^2}(\hat{y}_t-\hat{y}_o)$, but more complete calculations are very involved.

4. CONCLUSIONS.

We have presented one example of the method proposed in [9] for using Lie algebraic techniques to study nonlinear estimation problems (a similar analysis can of course be done for other problems in the class discussed in [6] - [8]). This method clarifies the relationship between the computation of the (unnormalized) conditional density and the finite dimensional computation of certain statistics of the conditional distribution (in this case, the conditional moments). Although moments of any order can be computed by a finite dimensional estimator in this example, it is unresolved whether the same is true of the conditional density. That is, the Lie algebra of the Zakai equation (5) is infinite dimensional, but that certainly does not preclude its being isomorphic to a Lie algebra generated by two vector fields on a finite dimensional manifold (which would be the case if it could be computed in terms of finite dimensionally computable sufficient statistics). However, since moments of all orders can be calculated, it may be possible (modulo questions such as moment determinacy) to approximate the conditional density to any desired degree of accuracy by means of a series in the finite dimensionally computable statistics.

On the other hand, the Lie algebra of the Zakai equation may have very few ideals, in which case there may be no statistics which are "more easily" computable than the unnormalized conditional density. Examples of both types and further analysis along these lines will be presented in [22]. Finally, we should warn that Lie algebraic conditions do not always present the whole picture; as discussed in [20], one must essentially be able to "integrate" the abstract Lie algebra representations obtained in order to actually construct the estimator, and this is not always possible (see [23] for one further class of systems for which this <u>is</u> possible).

ACKNOWLEDGEMENT.

The second author would like to thank M. Hazewinkel for many stimulating discussions, and for providing him the opportunity to develop some of these ideas while visiting the Econometric Institute, Erasmus University Rotterdam, Rotterdam.

APPENDIX.

DERIVATION OF FINITE DIMENSIONAL ESTIMATOR.

First we note [6, Appendix B] that if $x = [x_1 \ \ldots \ x_k]'$ is a Gaussian random vector with mean m and covariance P, then

$$E[x_1 \ldots x_k] = E[x_k]E[x_1 \ldots x_{k-1}] + \sum_{\alpha_1=1}^{k} P_{k\alpha_1} E[x_{\alpha_2} \ \ldots x_{\alpha_{k-1}}]$$

$$(A.1) \qquad = m_1 \ldots m_k + \sum_{(\alpha_1,\alpha_2)} P_{\alpha_1\alpha_2} m_{\alpha_3} \ldots m_{\alpha_k}$$

$$+ \sum_{(\alpha_1,\alpha_2,\alpha_3,\alpha_4)} P_{\alpha_1\alpha_2} P_{\alpha_3\alpha_4} m_{\alpha_5} \ldots m_{\alpha_k} + \ldots$$

where each set $\{\alpha_i, \ i = 1,\ldots,k\}$ is a permutation of $\{1,\ldots,k\}$ and the sums in (A.1) are over all possible combinations of pairs of the $\{\alpha_i\}$. Now, x in the problem (12) - (13) is conditionally Gaussian with conditional cross-covariance defined by (for $\sigma_1, \ \sigma_2 \le t$)

$$P(\sigma_1,\sigma_2,t) = E[(x_{\sigma_1} - \hat{x}_{\sigma_1|t})(x_{\sigma_2} - \hat{x}_{\sigma_2|t})|z^t],$$

where $\hat{x}_{\sigma|t} = E[x_\sigma|z^t]$; using the results of [6, section 2] it can be shown that

$$(A.2) \qquad \frac{d}{dt} P(\sigma_1,\sigma_2,t) = -P(\sigma_1,t,t)P(\sigma_2,t,t)$$

$$(A.3) \qquad P(\sigma,t,t) = K(t,\sigma)P_t$$

$$(A.4) \qquad \frac{d}{dt} K(t,\sigma) = -P_t^{-1}K(t,\sigma) \ ; \ K(\sigma,\sigma) = 1$$

where $P_t = P(t,t,t)$ is the solution of the Riccati equation (16).

The conditional mean $\hat{y}_t$ satisfies equation (2) in Ito form:

$$(A.5) \qquad d\hat{y}_t = E^t[x_t^2]dt + \{E^t[y_t x_t] - \hat{y}_t \hat{x}_t\}[dz_t - \hat{x}_t dt]$$

But $E^t[x_t^2] = \hat{x}_t^2 + P_t$, and using (A.1), (A.3), and (A.4),

$$E^t[y_t x_t] - \hat{y}_t\hat{x}_t = \int_0^t (E^t[x_s^2 x_t] - E^t[x_s^2]\hat{x}_t)ds$$
$$= \int_0^t 2P(s,t,t)\hat{x}_s ds$$
$$= 2\hat{\xi}_t P_t$$

where ξ satisfies

$$\dot{\xi}_t = x_t - \xi_t P_t^{-1}, \xi_o = 0.$$

Thus the Kalman filter for the system with states x, ξ and observations z computes $\hat{x}$, $\hat{\xi}$, and $\hat{y}$ is computed according to (A.5), thus yielding the first three equations in (15) and P, P_{12} in (16) (once they have been converted to Fisk-Stratonovich form).

Furthermore, since $dy_t^2 = 2y_t dy_t = 2y_t x_t^2 dt$, equation (2) yields

(A.6) $$d\widehat{y_t^2} = 2E^t[y_t x_t^2]dt + \{E^t[y_t^2 x_t] - \widehat{y_t^2}\hat{x}_t\}(dz_t - \hat{x}_t dt).$$

Using (A.1),

$$E^t[y_t x_t^2] = \int_0^t E^t[x_t^2 x_s^2]ds$$
$$= \hat{x}_t^2\hat{y}_t + \hat{y}_t P + 4\hat{x}_t\hat{\xi}_t P + 2PP_{12}.$$

Also, (A.1) - (A.4) imply that

$$E^t[y_t^2 x_t] - \widehat{y_t^2}\hat{x}_t$$
$$= 4\int_0^t\int_0^t P(s,t,t)E^t[x_s x_\tau^2]dsd\tau$$
$$= 4\int_0^t\int_0^t P(s,t,t)[\hat{x}_{s|\tau}E^t[x_\tau^2] + 2P(s,\tau,t)\hat{x}_{\tau|t}]dsd\tau$$
$$= 4\{(\int_0^t P(s,t,t)\hat{x}_{s|t}ds)\hat{y}_t + 2E^t[\int_0^t P(s,t,t)(\int_0^t P(s,\tau,t)x_\tau d\tau)ds]\}$$
$$= 4(\hat{\xi}_t\hat{y}_t P + 2\hat{\theta}_t P)$$

where Θ_t satisfies

$$\dot{\Theta}_t = P_{12}x_t + \xi_t(P-PP_{12}) - P^{-1}\Theta_t; \quad \Theta_o = 0$$

The Kalman-Bucy filter for the state equations of x, ξ, and Θ with observation z computes $\hat{x}_t$, $\hat{\xi}_t$, $\hat{\Theta}_t$, and $\widehat{y_t^2}$ is computed according to (A.6). After correction terms are added, these result in the first five equations in (15), and the first three in (16). The third moment $\widehat{y^3}$ (and higher moments) are computed in a similar manner.

REFERENCES.

1. H.J. Kushner, "Dynamical Equations for Optimal Nonlinear Filtering", J. Diff. Equations, Vol. 3, 1967, pp. 179-190.
2. M. Fujisaki, G. Kallianpur, and H. Kunita", Stochastic Differential Equations for the Nonlinear Filtering Problem". Osaka J. Math., Vol. 1, 1972, pp. 19-40.
3. R.S. Liptser and A.N. Shiryayev, Statistics of Random Processes I. New York: Springer-Verlag, 1977.
4. H.J. Kushner, "On the Dynamical Equations of Conditional Probability Functions with Application to Optimal Stochastic Control Theory," J. Math. Anal. Appl., Vol. 8, 1964, pp. 332-344.
5. J.H. Van Schuppen, "Stochastic Filtering Theory: A Discussion of Concepts, Methods, and Results", in Stochastic Control Theory and Stochastic Differential Systems, M. Kohlmann and W. Vogel, Eds. New York, Springer-Verlag, 1979.
6. S.I. Marcus and A.S. Willsky, "Algebraic Structure and Finite Dimensional Nonlinear Estimation", SIAM J. Math. Anal., Vol. 9, April 1978, pp. 312-327.
7. S.I. Marcus, "Optimal Nonlinear Estimation for a Class of Discrete-Time Stochastic Systems", IEEE Trans. Automat. Contr., Vol. AC-24, April 1979, pp. 297-302.
8. S.I. Marcus, S.K. Mitter, and D. Ocone, "Finite Dimensional Nonlinear Estimation for a Class of Systems in Continuous and Discrete Time, "Proc. of Int. Conf. on Analysis and Optimization of Stochastic Systems, Oxford, Sept. 6-8, 1978.
9. R.W. Brockett, "Remarks on Finite Dimensional Nonlinear Estimation, Journees sur l'Analyse des Systèmes, Bordeaux, 1978.
10. R.W. Brockett, "Classification and Equivalence in Estimation Theory," Proc. 1979 IEEE Conf. on Decision and Control, Ft. Lauderdale, December 1979.
11. S.K. Mitter, "Modeling for Stochastic Systems and Quantum Fields," Proc. 1978 IEEE Conf. on Decision and Control, San Diego, January 1979.
12. M. Zakai, "On the Optimal Filtering of Diffusion Processes," Z. Wahr. Verw. Geb., Vol. 11, 1969, pp. 230-243.
13. R.W. Brockett, "System Theory on Group Manifolds and Coset Spaces," SIAM J. Control, Vol. 10, 1972, pp. 265-284.

14. H.J. Sussmann, "Existence and Uniqueness of Minimal Realizations of Nonlinear Systems, "Math. Systems Theory, Vol. 10, 1976/1977.

15. R. Hermann and A.J. Krener, "Nonlinear Controllability and Observability, "IEEE Trans. Automat. Contr., Vol. AC-22, October 1977, pp. 728-740.

16. J.E. Humphreys, Introduction to Lie Algebras and Representation Theory. New York: Springer-Verlag, 1972.

17. A.J. Krener, "On the equivalence of Control Systems and the Linearization of Nonlinear Systems, "SIAM J. Control, Vol.11, 1973, pp. 670-676.

18. A.J. Krener, "A Decomposition Theory for Differentiable Systems," SIAM J. Control, Vol. 15, 1977, pp. 813-829.

19. E.D. Sontag, Polynomial Response Maps. New York: Springer-Verlag, 1979.

20. S.K. Mitter, in Richerche di Automatica, to appear.

21. S.D. Chikte and J.T.-H. Lo, "Optimal Filters for Bilinear Systems with Nilpotent Lie Algebras, "IEEE Trans. Automat. Contr., Vol. AC-24, December 1979, pp. 948-953.

22. M. Hazewinkel, C.-H. Liu, and S.I. Marcus, "Some Examples of Lie Algebraic Structure in Nonlinear Estimation, " to be presented at the 1980 Joint Automatic Control Conference, San Francisco, August 1980.

23. V.E. Benes, "Exact Finite Dimensional Filters for Certain Diffusions with Nonlinear Drift," presented at the 1979 IEEE Conf. on Decision and Control, Ft. Lauderdale, December 1979.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: AFOSR-TR-80-0879 (18) (19)

2. GOVT ACCESSION NO.: AD-A091032

3. RECIPIENT'S CATALOG NUMBER:

4. TITLE (and Subtitle): (6) THE LIE ALGEBRAIC STRUCTURE OF A CLASS OF FINITE DIMENSIONAL NONLINEAR FILTERS.

5. TYPE OF REPORT & PERIOD COVERED: (9) Interim rept.

6. PERFORMING ORG. REPORT NUMBER: (15) F49620-77-C-0101

7. AUTHOR(s): (10) Chang-Huan/Liu ~~and~~ Steven I./Marcus

8. CONTRACT OR GRANT NUMBER(s): AFOSR-79-0025

9. PERFORMING ORGANIZATION NAME AND ADDRESS:
The University of Texas at Austin
Department of Electrical Engineering
Austin, Texas 78712

10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: 61102F (16) 2304/A1 (17) A1

11. CONTROLLING OFFICE NAME AND ADDRESS:
Air Force Office of Scientific Research/NM
Bolling AFB, Washington, DC 20332

12. REPORT DATE: (11) 23 July 1980

13. NUMBER OF PAGES: 20

14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office): (12) 23

15. SECURITY CLASS. (of this report): UNCLASSIFIED

15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

Appeared in Filterdag Rotterdam 1980.

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Nonlinear Filtering, Lie Algebras

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

We present an example of the application of Lie algebraic techniques to nonlinear estimation problems. The method relates the computation of the (unnormalized) conditional density and the computation of statistics with finite dimensional estimators. The general method is explained; for a particular example, the structures of the Lie algebras associated with the unnormalized conditional density equation and the finite dimensionally computable conditional moment equations are analyzed in detail. The

DD FORM 1473 401997 UNCLASSIFIED

#20. Continued:

relationship between these Lie algebras is studied, and the implications of these results are discussed.