ON THE RELATIONSHIP BETWEEN LIE ALGEBRAS AND NONLINEAR ESTIMATION

Michiel Hazewinkel
Erasmus University Rotterdam
Econometric Institute
Burg. Oudlaan 50
Rotterdam, The Netherlands

Steven I. Marcus
Department of Electrical Engineering
The University of Texas at Austin
Austin, Texas 78712

Abstract

A Lie algebra $L(\Sigma)$ can be associated with each nonlinear filtering problem, and the realizability of $L(\Sigma)$ or quotients of $L(\Sigma)$ with vector fields on a finite dimensional manifold is related to the existence of finite dimensional recursive filters. In this paper the structure and realizability properties of $L(\Sigma)$ are analyzed for several interesting classes of problems. It is shown that, for certain nonlinear filtering problems, $L(\Sigma)$ is given by the Weyl algebra

$W_n = \mathbb{R} < x_1,\ldots,x_n, \frac{\partial}{\partial x_1},\ldots,\frac{\partial}{\partial x_n} >$. It is proved that neither W_n nor any quotient of W_n can be realized with C^∞ or analytic vector fields on a finite dimensional manifold, thus showing that for these problems, no statistic of the conditional density can be computed with a finite dimensional recursive filter. For another class of problems (including bilinear systems with linear observations), it is shown that $L(\Sigma)$ is a certain type of filtered Lie algebra; the implications of this property are discussed.

I. Introduction

This paper is concerned with the problem of recursively filtering the state x_t of a nonlinear stochastic system, given the past observations $z^t = \{z_s, 0 \le s \le t\}$. The systems we consider satisfy the Ito stochastic differential equations

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

where $x \in \mathbb{R}^n$, $w \in \mathbb{R}^m$, $z \in \mathbb{R}^p$, w and v are independent unit variance Wiener processes, and $R > 0$. The optimal (minimum-variance) estimate of x_t is of course the conditional mean $\hat{x}_t \triangleq E[x_t|z^t]$ (also denoted $\hat{x}_{t|t}$ or $E^t[x_t]$). Aside from the linear-Gaussian case in which the Kalman filter is optimal, there are very few known cases in which the conditional mean, or indeed <u>any</u> statistic of the conditional distribution, can be computed with a finite dimensional recursive filter (a number of these are summarized in [1]). More precisely, a <u>finite dimensional recursive filter</u> is a stochastic differential equation driven by the observations of the form

$$d\eta_t = a(\eta_t)dt + \sum_{i=1}^{p} b_i(\eta_t)dz_{it}, \qquad (1)$$

where η evolves on a finite dimensional manifold and a and $\{b_i\}$ are sufficiently smooth to insure existence and uniqueness (these conditions will be strengthened later). The conditional statistic $E[c(x_t)|z^t]$ is said to be <u>finite dimensionally computable</u> (FDC) if it can be computed "pointwise" as a function of the state of a finite dimensional recursive filter:

$$\hat{c}(x_t) \triangleq E[c(x_t)|z^t] = \gamma(\eta_t). \qquad (2)$$

As a practical matter, it is also useful to require that the combined estimator (1)-(2) yield a statistic $\hat{c}(x_t)$ which is a continuous function of z; we will comment on this later in this section.

Recently, Brockett [2],[3] and Mitter [4],[5] have shown that Lie algebras play an important role in nonlinear recursive estimation theory; the approach of Brockett [2] is the following. Consider the Zakai equation for an unnormalized conditional density $\rho(t,x)$ [6]:

$$d\rho(t,x) = L\rho(t,x)dt + \sum_{i=1}^{p} h_i(x)\rho(t,x)dz_{it} \qquad (3)$$

where z_i and h_i are the i^{th} components of z and h,

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

is the forward diffusion operator, and $\rho(t,x)$ is related to the conditional density $p(t,x)$ of x_t given z^t by

$$p(t,x) = \rho(t,x)\cdot(\int\rho(t,x)dx)^{-1}. \qquad (5)$$

Notice that (3) is a bilinear differential equation [7] in ρ, with z considered as the input. <u>Suppose</u> that, for some initial density, some statistic of the conditional distribution of x_t given z^t can be calculated with a finite dimensional recursive estimator of the form (1)-(2), where a, b_i, and γ are analytic. Of course, this statistic can also be obtained from $\rho(t,x)$ by

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

For the rest of the development, it is more convenient to write (1) and (3) in Fisk-Stratonovich form (so that they obey the ordinary rules of calculus and so that Lie algebraic calculations involving differential operators can be performed as usual):

$$d\eta_t = \tilde{a}(\eta_t)dt + \sum_{i=1}^{p} b_i(\eta_t)dz_{it} \qquad (7)$$

$$d\rho(t,x) = [L - \frac{1}{2}\sum_{i=1}^{p} h_i^2(x)]\rho(t,x)dt + \sum_{i=1}^{p} h_i(x)\rho(t,x)dz_{it} \qquad (8)$$

where the i^{th} component $\tilde{a}_i(\eta) = a_i(\eta) - \frac{1}{2}\sum_{j,k} b_{jk}(\eta) \cdot \frac{\partial b_{ik}}{\partial \eta_j}(\eta)$ (here b_{jk} is the k^{th} component of b_j).

This research was supported in part by the National Science Foundation under Grant ENG 76-11106 and in part by the Joint Services Electronics Program under Contract F49620-77-C-0101. A portion of this research was conducted while S. Marcus was visiting the Econometric Institute, Erasmus University, Rotterdam, The Netherlands.

0191-2216/80/0000-0066$00.75 © 1980 IEEE

The two systems (7),(2) and (8),(6) are thus two representations of the same mapping from "input" functions z to "outputs" $\hat{c}(x_t)$: (8),(6) via a bilinear infinite dimensional state equation, and (7),(2) via a nonlinear finite dimensional state equation. Motivated by the results of [8],[9] for finite dimensional state equations, the major observation of [2] is that, under appropriate hypotheses including minimality of the representation (7),(2), the Lie algebra F generated by $a,b_1,\ldots,b_p$ (under the commutator $[a,b] = \frac{\partial a}{\partial \eta} b - \frac{\partial b}{\partial \eta} a$) should be a homomorphic image of the Lie algebra $L(\Sigma)$ generated by $e_0 = L - \frac{1}{2}\sum_{i=1}^{p} h_i^2(x)$ and $e_i = h_i(x)$, $i=1,\ldots,p$ (under the commutator $[e_0,e_i] = e_0e_i - e_ie_0$), with $e_0 \to \tilde{a}$ and $e_i \to b_i$, $i=1,\ldots,p$. On the other hand, if there is a homomorphism ϕ of $L(\Sigma)$ onto a Lie algebra generated by p+1 complete vector fields $\tilde{a},b_1,\ldots,b_p$, on a finite dimensional manifold, then this is an indication that some conditional statistic may be computable by an estimator of the form (7),(2). It is not known in what generality such results are valid, especially for cases in which $L(\Sigma)$ is infinite dimensional, and much work remains to be done (the fact that existence of a finite dimensional filter implies the existence of a Lie algebra homomorphism has been made rigorous for a class of estimation problems, including some of those discussed in Section II, in [26]). However, it is clear that there is a strong relationship between the structure of $L(\Sigma)$ and the existence of finite dimensional filters. In this paper, we discuss the properties of $L(\Sigma)$ for some interesting classes of examples. These Lie algebraic calculations give some new insights into certain nonlinear estimation problems and some guidance in the search for finite dimensional estimators.

If $L(\Sigma)$ is finite dimensional (this seems to occur only in very special cases [5],[10]), a finite dimensional estimator can in some cases be constructed by integrating the Lie algebra representation. Indeed, if $L(\Sigma)$ or any of its quotients is finite dimensional, then by Ado's Theorem [11, p. 202] this Lie algebra has a faithful finite dimensional representation; thus it can be realized with linear vector fields on a finite dimensional manifold, resulting in a bilinear filter (see, e.g., [12] and [16] for examples). However, actually computing the mapping from $\rho(t,x)$ to $\hat{c}(x_t)$ (i.e., deciding which statistic the filter computes) is a difficult problem from this point of view; one must so far use other, more direct, methods to actually construct this mapping or to derive the filter for a particular conditional statistic (see, e.g., [14]-[17]). On the other hand, if $L(\Sigma)$ or its quotients are infinite dimensional, it is still possible that these Lie algebras can be realized by nonlinear vector fields on a finite dimensional manifold. Conditions under which this can be done is an unsolved problem in general; we show in Section II that this is not possible for certain classes of Lie algebras. However, to see that two vector fields on a finite dimensional manifold can generate an infinite dimensional Lie algebra, consider the vector fields $a = x^2 \frac{\partial}{\partial x}$ and $b = x^3 \frac{\partial}{\partial x}$ on a one-dimensional manifold; it is easy to see that a and b generate the infinite dimensional Lie algebra of vector fields of the form $x^2 p(x) \frac{\partial}{\partial x}$, where p is a polynomial.

If a statistic $\hat{c}(x_t)$ is finite dimensionally computable, the Lie algebraic approach gives some insight into the continuity of the estimator. Since there is a Lie algebra homomorphism as discussed above, the vector fields $b_1,\ldots,b_p$ are homomorphic images of the operators $e_1,\ldots,e_p$ which all commute with each other (these are just multiplication operators). Thus $b_1,\ldots,b_p$ also commute, and the results of [18] imply that the filter (7) represents a continuous map (in the C^0 and L_p topologies) from the space of "inputs" z to the solutions η. Hence, the estimator (7),(2) gives a continuous map from z to $c(x_t)$; this is a very useful property, indicating the "robustness" of the filter (see also [19],[20]).

Brockett and Clark [13] used this approach to study the estimation of a finite state Markov process observed in additive Brownian motion; the Lie algebraic approach led to the discovery of new low dimensional filters for the conditional distribution, even in some cases when the number of states is arbitrarily large. In [2], the Lie algebraic approach is explicitly carried out and analyzed for the problem in which f and h are linear and G is constant. In that case, the Lie algebra $L(\Sigma)$ 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 [21], a similar analysis is carried out for an example of the class of estimation problems considered in [14]-[16]; for this class of nonlinear stochastic systems, the conditional mean (and all conditional moments) of x_t given z^t are finite dimensionally computable. For this example, the Lie algebra $L(\Sigma)$ is infinite dimensional but has many finite dimensional quotients (the Lie algebras of the finite dimensional filters), and these are analyzed in detail in [21]. These last two examples, as well as the example of Benes [17], are special cases of the class considered in Section III.

In Section II, we consider estimation problems for which $L(\Sigma)$ is the Weyl algebra W_n. A number of examples are given and useful properties of the Weyl algebra are derived; some of these results have been obtained independently by Mitter [5]. The major results of Section II are that neither W_n nor any quotient of W_n can be realized by vector fields with either C^∞ or formal power series coefficients on a finite dimensional manifold; this shows that for these problems, no statistic of the conditional density can be computed with a finite dimensional recursive filter. Most of the results in this paper will be stated without proof; for the proofs, see [27].

II. The Weyl Algebras W_n

The Weyl algebra W_n [22],[23, Chapter 1] is the algebra of all polynomial differential operators; i.e., $W_n = \mathbb{R} < x_1,\ldots,x_n; \frac{\partial}{\partial x_1},\ldots,\frac{\partial}{\partial x_n} >$. A basis for W_n consists of all monomial expressions

$$e_{\alpha,\beta} \triangleq x^\alpha \frac{\partial^\beta}{\partial x^\beta} \triangleq x_1^{\alpha_1} \cdots x_n^{\alpha_n} \frac{\partial^{\beta_1}}{\partial x_1^{\beta_1}} \cdots \frac{\partial^{\beta_n}}{\partial x_n^{\beta_n}} \tag{9}$$

where α,β range over all multiindices $\alpha = (\alpha_1,\ldots,\alpha_n)$, $\beta = (\beta_1,\ldots,\beta_n)$, $\alpha,\beta \in \mathbb{N} \cup \{0\}$ (the non-negative integers). W_n is a Lie algebra under the Lie bracket; as an example, we state the general formula for W_1:

$$[x^i \frac{\partial^j}{\partial x^j}, x^k \frac{\partial^\ell}{\partial x^\ell}] = \sum_{r=1}^{j} \binom{j}{r}\binom{k}{r} r!\, x^{i+k-r} \frac{\partial^{j+\ell-r}}{\partial x^{j+\ell-r}} - \sum_{s=1}^{\ell} \binom{\ell}{s}\binom{i}{s} s!\, x^{i+k-s} \frac{\partial^{j+\ell-s}}{\partial x^{j+\ell-s}} \quad (10)$$

where $\binom{j}{r} = \frac{j!}{(j-r)!r!}$ is the binomial coefficient and we have used the convention that $\binom{j}{r} = 0$ if $r<0$ or $j<r$. The center of W_n (i.e., the ideal of all elements $Z \in W_n$ such that $[X,Z]=0$ for all $X \in W_n$) is the one-dimensional space $\mathbb{R}\cdot 1$ with basis $\{1\}$ [22, p. 148]. Our first result is the simplicity of the Lie algebra $W_n/\mathbb{R}\cdot 1$; this is of course stronger than showing that W_n is simple as an associative algebra [22, p. 148]. Our proof follows that of Avez and Heslot [24] for the Lie algebra P_n of polynomials under the Poisson bracket. A number of the following results are common to P_n and W_n, but these two Lie algebras are not isomorphic (this is basically because the expression in P_n corresponding to (10) would retain only the terms for r=1 and s=1). Hence, one must be careful in literally interpreting results proved for P_n in the context of W_n [30].

Theorem 1: The Lie algebra $W_n/\mathbb{R}\cdot 1$ is simple; i.e., it has no ideals other than $\{0\}$ and $W_n/\mathbb{R}\cdot 1$. Equivalently, the only ideals of W_n are $\{0\}$, $\mathbb{R}\cdot 1$, and W_n.

This theorem basically shows that if W_n occurs as the Lie algebra $L(\Sigma)$ for some estimation problem, then either the unnormalized conditional density itself is finite dimensionally computable or no statistic at all is finite dimensionally computable. The next two theorems complete the argument by showing that in fact neither W_n nor its quotients can be realized by vector fields on a finite dimensional manifold.

Let $\hat{V}_m$ be the Lie algebra of vector fields

$$\hat{V}_m \triangleq \{ \sum_{i=1}^{m} f_i(x_1,\dots,x_m) \frac{\partial}{\partial x_i} \}$$ with (formal) power series coefficients $f_i \in \mathbb{R}[[x_1,\dots,x_m]]$, and let $V(M)$ be the Lie algebra of C^∞-vector fields on a C^∞-manifold M.

Theorem 2: Fix $n \neq 0$. Then there are no non-zero homomorphisms from W_n to $\hat{V}_m$ or from $W_n/\mathbb{R}\cdot 1$ to $\hat{V}_m$ for any m.

Theorem 3: Fix $n \neq 0$. Then there are no non-zero homomorphisms from W_n to $V(M)$ or from $W_n/\mathbb{R}\cdot 1$ to $V(M)$ for any finite dimensional C^∞-manifold M.

These results show (assuming the appropriate analog of the results of [2],[8]) that if a system Σ has estimation algebra $L(\Sigma) = W_n$ for some n, then neither the conditional density of x_t given z^t nor any nonzero statistic of the conditional density can be computed with a finite dimensional filter of the form (7) with a and $\{b_i\}$ C^∞ or analytic. We will give several examples of such systems, but first we present a general method for showing that $L(\Sigma) = W_n$, the proof of which is similar to that of [24] for Poisson brackets.

Theorem 4: The Lie algebra W_n is generated by the elements

$$x_i,\ \frac{\partial^2}{\partial x_i^2},\ x_i^2 \frac{\partial}{\partial x_i},\quad i=1,\dots,n;\ \text{and}\ x_i x_{i+1},\ i=1,\dots,n-1.$$

Theorem 4 provides a relatively systematic method for showing that $L(\Sigma) = W_n$ for a particular estimation problem: one need only show that by taking repeated Lie brackets of $L - \frac{1}{2}\Sigma h_i^2$ and $\{h_i\}$, the generating elements of W_n given in Theorem 4 are obtained. Notice that if n=1, the generating elements are x, $\frac{\partial^2}{\partial x^2}$, and $x^2 \frac{\partial}{\partial x}$. Some interesting examples are the following.

Example 1 (the cubic sensor problem [5],[25]): Consider the system

$$dx_t = dw_t$$
$$dz_t = x_t^3 dt + dv_t.$$

The Lie algebra $L(\Sigma)$ is generated by the operators

$$e_0 = \frac{1}{2}\frac{\partial^2}{\partial x^2} - \frac{1}{2}x^6,\quad e_1 = x^3.$$

We can compute a sequence of Lie brackets to obtain a sequence of elements $e_i \in L(\Sigma)$, eventually obtaining the desired generators of W_n:

$$[e_0,e_1] = 3x^2\frac{\partial}{\partial x} + 3x \Rightarrow e_2 = x^2\frac{\partial}{\partial x} + x$$

$$ad^k_{e_2} e_1 = 3\cdot 4\cdots(k+2)x^{k+3} \Rightarrow x^k \in L(\Sigma),\ k \geq 3$$

(where $ad^0_{e_2} e_1 = e_1$ and $ad^{k+1}_{e_0} e_1 = [e_0, ad^k_{e_0} e_1]$).

Combined with e_0, $x^6 \in L(\Sigma)$ implies that $e_3 = \frac{\partial^2}{\partial x^2} \in L(\Sigma)$. Continuing,

$$[e_3,e_2] = 4x\frac{\partial^2}{\partial x^2} + 4\frac{\partial}{\partial x} \Rightarrow e_4 = x\frac{\partial^2}{\partial x^2} + \frac{\partial}{\partial x}$$

$$[e_4,e_2] = 3x^2\frac{\partial^2}{\partial x^2} + 6x\frac{\partial}{\partial x} + 1 \Rightarrow e_5 = 3x^2\frac{\partial^2}{\partial x^2} + 6x\frac{\partial}{\partial x} + 1$$

$$[e_4,e_1] = 6x^3\frac{\partial}{\partial x} + 9x^2 \Rightarrow e_6 = 2x^3\frac{\partial}{\partial x} + 3x^2$$

$[e_3,e_6] = 12x^2\frac{\partial^2}{\partial x^2} + 24x\frac{\partial}{\partial x} + 6$, which combined with e_5 implies that $e_7 = 1$ and $e_8 = x^2\frac{\partial^2}{\partial x^2} + 2x\frac{\partial}{\partial x}$ are in $L(\Sigma)$. A few more calculations will complete the demonstration:

$$[e_3,e_8] = 4x\frac{\partial^3}{\partial x^3} + 6\frac{\partial^2}{\partial x^2} \Rightarrow e_9 = x\frac{\partial^3}{\partial x^3}$$

$$[e_1,e_8] = -6x^4\frac{\partial}{\partial x} - 12x^3 \Rightarrow e_{10} = x^4\frac{\partial}{\partial x}$$

$$[e_2,e_9] = -5x^2\frac{\partial^3}{\partial x^3} - 9x\frac{\partial^2}{\partial x^2} \Rightarrow e_{11} = 5x^2\frac{\partial^3}{\partial x^3} + 9x\frac{\partial^2}{\partial x^2}$$

$$[e_3,e_{10}] = 8x^3\frac{\partial^2}{\partial x^2} + 12x^2\frac{\partial}{\partial x} \Rightarrow e_{12} = 2x^3\frac{\partial^2}{\partial x^2} + 3x^2\frac{\partial}{\partial x}$$

$$[e_3,e_{12}] = 12x^2 \frac{\partial^3}{\partial x^3} + 24x \frac{\partial^2}{\partial x^2} + 6 \frac{\partial}{\partial x}$$

$$\Rightarrow e_{13} = 2x^2 \frac{\partial^3}{\partial x^3} + 4x \frac{\partial^2}{\partial x^2} + \frac{\partial}{\partial x}$$

Now e_{13}, e_{11}, and e_4 are all linear combinations of the elements $x^2 \frac{\partial^3}{\partial x^3}$, $x \frac{\partial^2}{\partial x^2}$, and $\frac{\partial}{\partial x}$, and the coefficient matrix

$$\begin{bmatrix} 0 & 1 & 1 \\ 5 & 9 & 0 \\ 2 & 4 & 1 \end{bmatrix}$$

is nonsingular. It follows that $L(\Sigma)$ contains $e_{14} = \frac{\partial}{\partial x}$, $e_{15} = x \frac{\partial^2}{\partial x^2}$, and $e_{16} = x^2 \frac{\partial^3}{\partial x^3}$. Finally,

$$[e_{14},e_1] = 3x^2 \Rightarrow e_{17} = x^2$$

$$[e_{14},e_{17}] = 2x \Rightarrow e_{18} = x$$

which combined with e_2 gives $x^2 \frac{\partial}{\partial x} \varepsilon L$; thus by Theorem 4 $L(\Sigma) = W_1$. This example is in the class studied in [26], for which the ideas of [2] are made rigorous. Thus we have in fact shown that no conditional statistic is finite dimensionally computable for the cubic sensor.

Analogous computation of selected Lie brackets and the use of Theorem 4 yields similar results for the following examples.

Example 2: For the system

$$dx_t = x_t^3 dt + dw_t$$

$$dz_t = x_t dt + dv_t,$$

$L(\Sigma)$ is generated by $\frac{1}{2} \frac{\partial^2}{\partial x^2} - x^3 \frac{\partial}{\partial x} - \frac{7}{2} x^2$ and x, and $L(\Sigma) = W_1$.

Example 3 (mixed linear-bilinear type): Consider the system with state equations

$$dx_t = dw_{1t}$$

$$dy_t = x_t dt + x_t dw_{2t}$$

with observations

$$dz_t = y_t dt + dv_t.$$

$L(\Sigma)$ is generated by $\frac{1}{2} \frac{\partial^2}{\partial x^2} + \frac{1}{2} x^2 \frac{\partial^2}{\partial y^2} - x \frac{\partial}{\partial y} - \frac{1}{2} y^2$ and y, and $L(\Sigma) = W_2$. The same result is obtained if the $x_t dt$ term is absent in the y equation; in that case we have a multiple Wiener integral of Brownian motion observed in Brownian motion noise.

Example 4: Consider the system with state equations

$$dx_t = dw_t$$

$$dy_t = x_t^2 dt$$

and observations

$$dz_{1t} = x_t dt + dv_{1t}$$

$$dz_{2t} = y_t dt + dv_{2t}.$$

$L(\Sigma)$ is generated by $\frac{1}{2} \frac{\partial^2}{\partial x^2} - x^2 \frac{\partial}{\partial y} - \frac{1}{2} x^2 - \frac{1}{2} y^2$, x, and y; it is easily shown that $L(\Sigma) = W_2$. This is the example studied in [21], but here we have the additional observation z_2; the relationship between these examples will be examined in the next section.

III. Pro-Finite Dimensional Filtered Lie Algebras

A Lie algebra L is defined to be a pro-finite dimensional filtered Lie algebra if L has a decreasing sequence of ideals $L = L_{-1} \supset L_0 \supset L_1 \supset \ldots$ such that

(a) $\cap L_i = 0$

(b) L/L_i is a finite dimensional Lie algebra for all i.

The terminology is analogous to that of pro-finite groups [28]. Notice that (a) implies that there is an injection from L to $\bigoplus_i L/L_i$. In the context of the estimation problem, this would correspond to $L(\Sigma)$ having an infinite number of finite dimensional quotients; if each of these can be realized with a recursively filterable statistic, then the injectivity of the map makes it reasonable to conjecture that these statistics represent some type of power series expansion of the conditional density. Of course, in addition to those discussed in Section I, other difficult technical questions such as moment determinacy will also be relevant here, but the structure of the Lie algebra should provide some guidance as to possible successful approaches to the problem and some insight into the structure of the resulting approximations.

Example 5 [21]: A simple example of the class considered in [14]-[16] is given by the state equations

$$dx_t = dw_t$$

$$dy_t = x_t^2 dt$$

and the observations

$$dz_t = x_t dt + dv_t$$

with x_0 Gaussian. The computation of $\hat{x}_t$ is of course straightforward by means of the Kalman filter; however, as shown in [14]-[16], all conditional moments of y_t can also be computed recursively with finite dimensional filters. $L(\Sigma)$ is generated by $e_0 = -x^2 \frac{\partial}{\partial y} + \frac{1}{2} \frac{\partial^2}{\partial x^2} - \frac{1}{2} x^2$ and $e_1 = x$; as shown in [21], a basis for $L(\Sigma)$ is given by e_0 and $\{x \frac{\partial^i}{\partial y^i}, \frac{\partial}{\partial x} \frac{\partial^i}{\partial y^i}, \frac{\partial^i}{\partial y^i}\ ;\ i=0,1,2,\ldots\}$. Defining L_i to be the ideal generated by $x \frac{\partial^i}{\partial y^i}$, $i=0,1,2,\ldots$, it is easy to see that $L(\Sigma)$ is a pro-finite dimensional filtered Lie algebra, and realizations of the $L(\Sigma)/L_i$ in terms of recursively filterable statistics are given in [21]. In addition, $L(\Sigma)$ is solvable [21].

A similar analysis for systems of the form of Example 5, with x_t^2 replaced by a general monomial x_t^p has also been done [31]; for $p > 2$, a similar but more complex Lie algebraic structure is exhibited. It is interesting to compare Example 5 with Example 4, which is the same except for the additional observation $dz_{2t} = y_t dt + dv_{2t}$; in that case $L(\Sigma) = W_2$, so that no conditional statistic can be computed exactly with a finite dimensional filter. However, it is probable that, due to the additional observation, a suboptimal approximate filter (such as the Extended Kalman Filter) for the conditional mean of y_t will result in lower mean-square error than the optimal filter which computes $\hat{y}_t$ in Example 5. Thus some care must be taken in interpreting the Lie algebraic structure of a nonlinear estimation problem; this structure has direct implications on the exact computation of conditional statistics, but its implications for approximate filtering remains to be investigated.

Example 6 (degree increasing operators and bilinear systems): Consider a system of the form (Σ), and suppose that f, G, and h are analytic with $f(0) = 0$ and $G(0) = 0$, so that the power series expansions of f and G around zero are of the form

$$f(x) = \sum_{|\alpha| \geq 1} f_\alpha x^\alpha, \quad G(x) = \sum_{|\alpha| \geq 1} G_\alpha x^\alpha, \qquad (11)$$

where $|\alpha| = \alpha_1 + \ldots + \alpha_n$. It follows that

$$G(x)G'(x) = \sum_{|\alpha| \geq 2} \tilde{G}_\alpha(x) x^\alpha.$$

An example of such systems is the class of bilinear systems

$$dx_t = Ax_t + \sum_{i=1}^{p} B_i x_t dw_t^i \qquad (12)$$

$$dz_t = Cx_t dt + dv_t.$$

Another example is

$$dx_t = x_t dt + \sin x_t \, dw_t$$

$$dz_t = h(x_t)dt + dv_t$$

with h analytic; in general, a wide variety of examples can be found.

Let $M = \mathbb{R}[[x_1,\ldots,x_n]]$ be the module of all (formal) power series in $x_1,\ldots,x_n$, and define the submodules

$$M_i = \{\Sigma\, a_\alpha x^\alpha | a_\alpha = 0 \text{ for } |\alpha| \leq i\}, \; i=0,1,2,\ldots,$$

so that, e.g., M_0 consists of those power series with zero constant term. If Σ is a system satisfying the condition (11), it follows that for all i, the forward diffusion operator (4) satisfies

$$\mathcal{L} M_i \subset M_i;$$

hence,

$$(\mathcal{L} - \tfrac{1}{2}h^2 x)M_i \subset M_i$$

and of course

$$h(x)M_i \subset M_i.$$

Since the two generators of $L(\Sigma)$ thus leave M_i invariant, it is obvious that $L(\Sigma)M_i \subset M_i$; thus, each element of $L(\Sigma)$ can only increase (or leave the same) the degree of the first term in the power series expansion of an element of M. Let

$$L_i = \{X \in L(\Sigma) | XM \subset M_{i+1}\}, \; i=-1,0,1,2,\ldots.$$

Then L_i is an ideal in $L(\Sigma)$ and we have an induced representation

$$\rho_i: \; L/L_i \to \mathrm{End}(M/M_{i+1}).$$

Because M/M_{i+1} is finite dimensional, so is L/L_i, since ρ_i is injective (by definition of L_i). It is obvious that $\cap L_i = \{0\}$; thus $L(\Sigma)$ is a pro-finite dimensional filtered Lie algebra, with filtration L_i. One additional structural feature of this filtration is that L_0/L_i is a nilpotent Lie algebra for i=1,2,...; also, L_i/L_{i+1} is abelian for all $i \geq 0$. The nilpotency of the L_0/L_i is a property also possessed by the filtration of Example 5.

Since many systems can be well approximated by bilinear ones, these results may have important implications for approximate nonlinear filtering. We close this section with two interesting examples of this class; the first is a bilinear system of the form (12), but in which some elements of A are also unknown and must be estimated. The second is an angle modulation problem.

Example 7 (bilinear system with unknown parameter): The simplest example of this type is

$$dx_t = \alpha_t x_t dt + x_t dw_t$$

$$d\alpha_t = 0$$

$$dz_t = x_t dt + dv_t$$

Here both the state x_t and parameter α are to be estimated recursively. The Lie algebra $L(\Sigma)$ is generated by $\frac{1}{2}x^2 \frac{\partial^2}{\partial x^2} + 2x\frac{\partial}{\partial x} + 1 - \alpha x \frac{\partial}{\partial x} - \alpha - \frac{1}{2}x^2$ and x. Both of these operators are "degree increasing" when operating on $\mathbb{R}[[x,\alpha]]$, so $L(\Sigma)$ is a pro-finite dimensional filtered Lie algebra.

Example 8 (angle modulation without process noise): Consider the problem of observing

$$dz_{1t} = \sin(\omega t + \theta)dt + dv_{1t}$$

$$dz_{2t} = \cos(\omega t + \theta)dt + dv_{2t}$$

where ω and θ are constant random variables to be estimated. To place this problem in the present framework, we have the three state equations

$$\dot{\omega} = 0$$

$$\dot{\theta} = 0$$

$$\dot{t} = 1$$

The Lie algebra $L(\Sigma)$ is generated by $e_0 = \frac{\partial}{\partial t} - \frac{1}{2}$, $e_1 = \sin(\omega t + \theta)$, and $f_1 = \cos(\omega t + \theta)$. It is easily shown that $L(\Sigma)$ has basis elements $e_0, e_i = \omega^i \sin(\omega t + \theta)$, $f_i = \omega^i \cos(\omega t + \theta)$, i=0,1,2,.... The nonzero commutation relations are $[e_0, e_i] = f_{i+1}$, $[e_0, f_i] = -e_{i+1}$. Hence $L(\Sigma)$ is a pro-finite dimensional filtered Lie algebra, with filtration $\{L_i\}$, where L_i is the ideal generated by e_{i+1} and f_{i+1}, i=0,1,2,.... Phase-lock loops are often used for filtering problems such as this, but the form of the optimal estimator is unknown. This calculation suggests that an infinite number of statistics of the conditional density may be finite dimensionally computable.

References

1. J. H. Van Schuppen, "Stochastic filtering theory: a discussion of concepts, methods and results," in *Stochastic Control Theory and Stochastic Differential Systems*, M. Kolhmann and W. Vogel, eds., New York: Springer-Verlag, 1979.
2. R. W. Brockett, "Remarks on finite dimensional nonlinear estimation," presented at the Conference on Algebraic and Geometric Methods in System Theory, Bordeaux, France, September 1978; to appear in *Asterisque*, 1980.
3. R. W. Brockett, "Classification and equivalence in estimation theory," *Proc. 1979 IEEE Conf. on Decision and Control*, Ft. Lauderdale, December 1979.
4. S. K. Mitter, "Filtering theory and quantum fields," presented at the Conference on Algebraic and Geometric Methods in System Theory, Bordeaux, France, September 1978; to appear in *Asterisque*, 1980.
5. S. K. Mitter, "On the analogy between mathematical problems of non-linear filtering and quantum physics," to appear in *Richerche di Automatica.*
6. M. Zakai, "On the optimal filtering of diffusion processes," *Z. Wahr. Verw. Geb.*, Vol. 11, 1969, pp. 230-243.
7. R. W. Brockett, "System theory on group manifolds and coset spaces," *SIAM J. Control*, Vol. 10, 1972, pp. 265-284.
8. H. J. Sussmann, "Existence and uniqueness of minimal realizations of nonlinear systems," *Math. Systems Theory*, Vol. 10, 1977, pp. 263-284.
9. R. Hermann and A. J. Krener, "Nonlinear controllability and observability," *IEEE Trans. Automatic Control*, Vol. AC-22, Oct. 1977, pp. 728-740.
10. D. Ocone, "Nonlinear filtering problems with finite dimensional estimation algebras," *Proc. 1980 Joint Automatic Control Conf.*, San Francisco, Aug. 1980.
11. N. Jacobson, *Lie Algebras*. New York: Wiley-Interscience, 1962.
12. S. D. Chikte and J. T.-H. Lo, "Optimal filters for bilinear systems with nilpotent Lie algebras," *IEEE Trans. Automatic Control*, Vol. AC-24, Dec. 1979, pp. 948-953.
13. R. W. Brockett and J. M. C. Clark, "The geometry of the conditional density equation," *Proc. Int. Conf. on Analysis and Optimization of Stochastic Systems*, Oxford, Sept. 6-8, 1978.
14. 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.
15. S. I. Marcus, "Optimal nonlinear estimation for a class of discrete-time stochastic systems," *IEEE Trans. Auto. Control*, Vol. AC-24, April 1979, pp. 297-302.
16. S. I. Marcus, S. K. Mitter, and D. Ocone, "Finite dimensional nonlinear estimation for a class of systems in continuous and discrete time," *Proc. Int. Conf. on Analysis and Optimization of Stochastic Systems*, Oxford, Sept. 6-8, 1978.
17. 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, Dec. 1979; also, to appear in *Stochastics.*
18. M. I. Freedman and J. C. Willems, "Smooth representation of systems with differentiated inputs," *IEEE Trans. Autom. Control*, Vol. AC-23, Feb. 1978, pp. 16-21.
19. J. M. C. Clark, "The design of robust approximations to the stochastic differential equations of nonlinear filtering," in *Communication Systems and Random Process Theory*, ed. J. K. Skwirzynski, NATO Advanced Study Institute Series, Alphen aan den Rijn: Sijthoff and Noordhoff, 1978.
20. M. H. A. Davis, "On a multiplicative functional transformation arising in nonlinear filtering theory," submitted to *Z. Wahr. Verw. Geb.*
21. C.-H. Liu and S. I. Marcus, "The Lie algebraic structure of a class of finite dimensional nonlinear filters," in *Filterdag Rotterdam 1980*, M. Hazewinkel (ed.), Report of the Econometric Inst., Erasmus University, Rotterdam, 1980.
22. J. Dixmier, *Enveloping Algebras*. Amsterdam, North-Holland, 1977.
23. J.-E. Björk, *Rings of Differential Operators*. Amsterdam: North-Holland, 1979.
24. A. Avez and A. Heslot, "L'algèbre de Lie des polynômes en les coordonnées canoniques munie du crochet de Poisson," *C. R. Acad. Sci. Paris A*, T. 288, 7 Mai 1979, pp. 831-833.
25. R. S. Bucy and J. Pages," A priori error bounds for the cubic sensor problem," *IEEE Trans. Autom. Control*, Vol. AC-23, Feb. 1978, pp. 88-91.
26. M. Hazewinkel, S. I. Marcus, and H. Sussmann, "Nonexistence of exact finite dimensional filters for the cubic sensor problem," in preparation.
27. M. Hazewinkel and S. I. Marcus, "On Lie algebras and finite dimensional filtering," preprint, 1980, submitted to *Stochastics.*
28. J.-P. Serre, *Cohomologie Galoisienne*, Lecture Notes in Math. 5. New York: Springer-Verlag, 1964.