March, 1983 LIDS-P-1279

GEOMETRIC THEORY OF NONLINEAR FILTERING

by

Sanjoy K. Mitter
Department of Electrical Engineering and Computer Science
and
Laboratory for Information and Decision Systems
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

This research has been supported by the Air Force Office of Scientific Research under Grant AFOSR 77-3281D and AFOSR 82-0135. A preliminary version of this paper was written while the author was a CNR Visiting Professor (Mathematics Group, GNAFA) at the Istituto Matematico "Ulisse Dini," Universita di Firenze, Firenze, Italy.

To appear in Outils et Méthodes Mathématiques pour l'Automatiques, Vol. 3.

1. Introduction

Until quite recently, the basic approach to non-linear filtering theory was via the "innovations method," originally proposed by Kailath ca.1967 and subsequently rigorously developed by Fujisaki, Kallianpur and Kunita [1] in their seminal paper of 1972. The difficulty with this approach is that the innovations process is not, in general, explicitly computable (excepting in the well-known Kalman-Bucy case). To circumvent this difficulty, it was independently proposed by Brockett-Clark [2], Brockett [3], Mitter [4] that the construction of the filter be divided into two parts: (i) a universal filter which is the evolution equation describing the unnormalized conditional density, the Duncan-Mortensen-Zakai (D-M-Z) equation and (ii) a state-output map, which depends on the statistic to be computed, where the state of the filter is the unnormalized conditional density. The reason for focusing on the D-M-Z equation is that it is an infinite-dimensional bi-linear system driven by the incremental observation process, and a much simpler object than the conditional density equation (which is a non-linear equation) and can be treated using geometric ideas. Moreover, it was noticed by this author that this equation bears striking similarities to the equations arising in (Euclidean)-quantum mechanics and it was felt that many of the ideas and methods used there could be used in this context. The ideas and methods referred to here are the functional integration view of Feynman (for a modern exposition see Glimm-Jaffe [5]). In many senses, this viewpoint has been remarkably successful--although the results obtained so far have been of a negative nature. Nevertheless the recent work has given us a deeper understanding of the D-M-Z equation which was essential for progress in non-linear filtering, as well as in stochastic control. The variational interpretation of non-linear filtering given by Fleming-Mitter

[6], Mitter [7] and the work on the partially observable stochastic control problem by Fleming-Pardoux [8] can be considered to have arisen from the "state-space" interpretation given to the filter.

This is an expository paper and contains no original results. For rigorous derivation of some of the results presented here, the reader is referred to the doctoral dissertation of Ocone [9], Hazewinkel-Marcus [10] and Sussmann [11]. The interested reader may also read with profit Hazewinkel-Willems [12] and Mitter-Moro [13].

2. The Filtering Problem Considered, And the Basic Questions.

We consider the signal-observation model:

$$\text{(1)}\quad \begin{aligned} dx_t &= f(x_t)dt + G(x_t)dw_t \; ; \; x(0) = x_o \\ dy_t &= h(x_t)dt + d\eta_t \; , \text{ where} \end{aligned} \qquad 0 \le t \le 1$$

x, w and y are $\mathbb{R}^n$, $\mathbb{R}^m$ and $\mathbb{R}^p$-valued processes, and it is assumed that f, G and h are vector-valued, matrix-valued and vector-valued functions which are smooth (which mean C^∞-function). It is further assumed that the stochastic differential equation (1) has a global solution in the sense of Ito. It is further assumed that x_t and η_t are independent and $E\int_o^1 |h(x_t)|^2 dt < \infty$. For much of our considerations, the function $h(\cdot)$ will be a polynomial.

It is well-known that the unnormalized conditional density $\rho(t,x)$ (where we have suppressed the y( ) and ω-dependence) satisfies the D-M-Z equation:

$$\text{(2)}\quad d\rho(t,x) = \left(\mathscr{L}_o^* - \frac{1}{2}\sum_{i=1}^{p} h_i^2(x)\right) \rho(t,x)dt + \sum_{i=1}^{p} h_i(x)\rho(t,x)\circ dy_t \; ,$$

where

$$\text{(3)}\quad \mathscr{L}_o^* \phi = \sum_{i,j=1}^{n} \frac{d^2}{dx_i dx_j}\left(G(x)G'(x)\right)_{ij}\phi - \sum_{i=1}^{n} \frac{d.}{dx_i} f_i(x)\Phi$$

and the · denotes the Stratanovich differential. It is imperative that we consider (2) as a Stratanovich differential equation, since the Ito-integral, because it "points to the future," is not invariant under smooth diffemorphisms of the x-space, and we want to study equation (2) in an "invariant manner."

We think of $\rho(t,\cdot)$ as the "state" of the filter and is, what we have referred to before, as the universal part of the filter. If ϕ, say, is a bounded, continuous functional then the filter typically is required to compute $E(\phi(x_t)|\mathscr{F}_t^Y)$, where $\mathscr{F}_t^Y = \sigma\{y_s, o\leq s\leq t\}$. If we denote by $\hat{\phi}_t \underline{\Delta} E(\phi(x_t)|\mathscr{F}_t^Y)$, then $\hat{\phi}_t$ is obtained from $\rho(t,x)$ by integration:

(4) $\hat{\phi}_t = \int_{\mathbb{R}^n} \phi(x)\rho(t,x)dx / \int_{\mathbb{R}^n} \rho(t,x)dx$

$\hat{\phi}_t$ will be referred to as a "conditional statistic," and no matter what $\hat{\phi}_t$ we wish to compute, $\rho(t,x)$ serves as a "sufficient statistic."

One of the questions we want to try to answer in this paper is: when can $\hat{\phi}_t$ (corresponding to a given ϕ) be computed via a finite-dimensional filter? The other remark to be made is: we are interested in computing the fundamental solution of (2) so that we can evaluate $\rho(t,x)$ corresponding to any initial condition.

To proceed further, we need to make a definition. By a <u>finite-dimensional</u> filter for a conditional statistic $\hat{\phi}_t$, we mean a stochastic dynamical system driven by the observations:

(5) $d\xi_t = \alpha(\xi_t)dt + \beta(\xi_t)\circ dy_t$

defined on a finite-dimensional manifold M, so that $\xi_t \varepsilon M$, and $\alpha(\xi_t)$ and $\beta(\xi_t)$ are smooth vector fields on M, together with a smooth output map

(6) $\hat{\phi}_t = \gamma(\xi_t)$, which computes the

conditional statistic. Equation (5) is to be interpreted in the Stratanovich sense for reasons we have mentioned above. We shall also assume that the stochastic dynamical system (5)-(6) is minimal in the sense of Sussmann [14].

For the definitions and properties of Lie algebras and Lie Groups used in the sequel the reader is referred to the Appendix.

3. Lie Algebra of Operators Associated with the Filtering Problem

Consider the Lie algebra generated by the unbounded operators

$$\mathscr{L} = \mathscr{L}_o^* - \frac{1}{2}\sum_{i=1}^{p} h_i^2(x) \text{ and } h_i(x) \ , \ i = i, \ldots p,$$

where the operators $\mathscr{L}$ and $h_i(x)$ (the h_i considered as multiplication operators $\phi(x) \to h_i(x)\phi(x)$) act on some common dense invariant domain $\mathscr{D}$ (say $\mathscr{D} = C_o^\infty(\mathbb{R}^n)$ or $\mathscr{S}(\mathbb{R}^n)$).

This Lie algebra contains important information and if it is finite-dimensional then it is a guide that a finite dimensional universal filter for computing $\rho(t,x)$ may exist.

Care should be taken in interpreting this statement. Firstly, referring to the definition of a finite-dimensional filter in (5), there is a Lie algebra of vector fields associated with it which in general is infinite-dimensional. Therefore, the fact that the Lie algebra $\mathscr{L}A\{\mathscr{L}, h_1, \ldots, h_p\}$ is infinite-dimensional does not preclude the filtering problem having a finite-dimensional solution. Secondly, even if $\mathscr{L}A\{\mathscr{L}, h_1, \ldots, h_p\}$ is finite-dimensional it does not mean that a finite-dimensional filter exists. The reason for this is that constructing the filter requires integrating the Lie algebra and it is a well-known fact from the theory of Unitary representations of Lie Groups that not all Lie algebra representations extend to a Group representation (see the Appendix of this paper). However, it is still a good question to ask as to whether examples of filtering problems exist where the Lie algebra $\mathscr{L}A\{\mathscr{L}, h_1, \ldots h_p\}$ is finite-dimensional and also how big is this class. The answer to the first part of this question is positive but the answer to the second part of the question appears to be that this class is small.

Example 1: (Kalman Filtering)

(7) $$\begin{cases} dx_t = Ax_t dt + b dw_t \\ dy_t = c'x_t dt + d_t \end{cases}$$

A = n x n matrix
b = n x 1 matrix
c = n x 1 matrix

Then

(8) $$\begin{cases} \mathscr{L}_o^* = \frac{1}{2} \sum_{i,j=1}^{n} \frac{\partial^2}{\partial x_i \partial x_k} Q_{ij} - \sum_{i=1}^{n} \frac{\partial}{\partial x_i} (Ax)_i \text{ , and} \\ \mathscr{L} = \mathscr{L}_o^* - \frac{1}{2}(c'x)^2 \text{ , where} \\ Q = bb' \end{cases}$$

Define the Hamiltonian matrix

$$E = \begin{pmatrix} -A' & cc' \\ bb' & A \end{pmatrix} \text{, and the vector}$$

$$\alpha = \begin{pmatrix} c \\ 0 \end{pmatrix} \varepsilon \mathbb{R}^{2n}$$

and the controllability matrix

$W = [\alpha : E\alpha \vdots \ldots\ldots\ldots \vdots E^{2n-1}\alpha]$ and assume that W is non-singular.

Define $Z_1 = c'x$ and

$$Z_i = [\text{ad}\mathscr{L}]^{i-1} Z_1.$$

Then one can show that

(9) $$Z_i = \sum_{j=1}^{n} (E^{i-1}\alpha)_j x_j + \sum_{j=1}^{n} (E^{i-1}\alpha)_{j+n} \frac{\partial}{\partial x_j} \text{ , and}$$

(10) $$[Z_i, Z_j] = (E^{i-1}\alpha)' \begin{pmatrix} 0 & -I \\ I & 0 \end{pmatrix} (E^{j-1}\alpha),$$

We can then conclude that the Lie algebra of the filter

$\mathscr{F}$ = span $\{\mathscr{L}, Z_1, \ldots, Z_{zn}, I\}$, where the $Z_1, \ldots, Z_{2n}$ are independent by hypothesis. Hence, $\mathscr{F}$ has dimension 2n+2, and this algebra is isomorphic to the oscillator algebra of dimension 2n+2 (see the Appendix).

3.1 Invariance Properties of the Lie Algebra and the Benes Problem.

The filter algebra is invariant under certain transformations, namely, diffeomorphisms on the x-space and gauge transformations to be discussed below. These ideas are best discussed on an example.

Consider the filtering problem:

$$(11) \quad \begin{cases} x_t = w_t \\ dy_t = x_t dt + d\eta_t \end{cases}$$

A basis for the filter algebra $\mathscr{F}$ is

$\left\{\mathscr{L}, x, \frac{d}{dx}, I\right\}$, where

$\mathscr{L} = \frac{1}{2}\frac{d^2}{dx^2} - \frac{1}{2}x^2$ and this is the 4-dimensional oscillator algebra.

It is easy to see that if we perform a smooth change of coordinates $x \to \phi(x)$ then the Filter algebra gives rise to an isomorphic Lie algebra, and two filtering problems with isomorphic Lie algebras should have the same filter.

Now consider the example first treated by Benes [15],

$$(12) \quad \begin{cases} dx_t = f(x_t)dt + dw_t \\ dy_t = x_t dt + d\eta_t \end{cases}, \text{ where}$$

f is the solution of the Riccati equation:

$\frac{df}{dx} + f^2 = ax^2 + bx + c$, and the coefficients a,b,c are so chosen that the equation has a global solution on all of R. We want to show that by introducing gauge transformations, we can transform the filter algebra of (12) to one which is isomorphic to the 4-dimensional oscillator algebra. Hence, the Benes filtering problem is essentially the same as the Kalman filtering problem considered in example 1.

To see this, first note that for (12)

$$[\mathscr{L}, x] = \frac{d}{dx} - f \text{ , where the brackets are computed on } C_o^\infty(\mathbb{R}).$$

Now consider the commutative diagram:

$$\begin{array}{ccc} C_o^\infty(\mathbb{R}) & \xrightarrow{\;\frac{d}{dx}\;} & C_o^\infty(\mathbb{R}) \\ \Psi \big\downarrow & & \big\downarrow \Psi \\ C_o^\infty(\mathbb{R}) & \xrightarrow[\;\frac{d}{dx} - f\;]{} & C_o^\infty(\mathbb{R}) \end{array}$$

Here Ψ is the multiplication operator $\phi(x) \to \Psi(x)\phi(x)$ and it is assumed that Ψ is invertible. Then it is easy to see that

$$\Psi(x) = \exp \int_o^x f(z)dz.$$

Under the transformation Ψ, the operator $\mathscr{L}_o^* = \frac{1}{2}\frac{\partial^2}{\partial x^2} - \frac{\partial}{\partial x} f$ transforms to $\frac{1}{2}\frac{\partial^2}{\partial x^2} - \frac{1}{2}V(x)$, where $V(x) = \frac{df}{dx} + f^2$.

It is easy to see that the Filter algebra $\mathscr{F}$ is isomorphic to the Lie algebra with generators

$$\frac{1}{2}\frac{d^2}{dx^2} - \frac{1}{2}V(x) - \frac{1}{2}x^2, \; x$$

We now see that if $V(x)$ is a quadratic, then this Lie algebra is essentially the 4-dimensional oscillator algebra corresponding to the Kalman Filter in Example 1.

What we have done is to introduce the gauge transportation $\rho(t,x) \to \Psi^{-1}(x)\rho(t,x)$, where $\rho(t,x)$ is the solution of the D-M-Z equation and what we have shown is that the Filter algebra is invariant under this isomorphism.

However, for the class of scalar models considered in (12) with general drifts f, the Benes problem is the only one with a finite-dimensional Lie algebra (we restrict ourselves to diffusions defined on the whole real line). For further details on this point the reader should consult Ocone [9].

There is no difficulty in generalizing these considerations to the vector case, provided f is a gradient vector field.

3.2 The Weyl Algebras and the Cubic Sensor Problem.

The Weyl algebra W_n is the algebra of all polynomial differential operators $\mathbb{R}(x_1, \ldots, x_n, \frac{\partial}{\partial x_1}, \ldots, \frac{\partial}{\partial x_n})$.

A basis for W_n consists of all monomial expressions

$$\frac{x^\alpha \partial^\beta}{\partial x^\beta} = x_1^{\alpha_1} \cdots x_n^{\alpha_n} \frac{\partial^{\beta_1}}{\partial x_1^{\beta_1}} \cdots \frac{\partial^{\beta_n}}{\partial x_n^{\beta_n}}$$

where α, β range over all multiindices $\alpha = (\alpha_1, \ldots, \alpha_n), \beta = (\beta_1, \ldots \beta_n)$. W_n can be endowed with a Lie algebra structure in the usual way. The centre of W_n, that is the ideal $\mathscr{I} = \{Z \in W_n \mid [X,Z] = 0, \forall X \in W_n\}$ is the one-dimensional space $\mathbb{R} \cdot 1$ and the Lie algebra $W_n / \mathbb{R} \cdot 1$ is simple.

Consider the cubic sensor filtering problem:

$$\begin{cases} x_t = w_t \\ dy_t = x_t^3 \, dt + d\eta_t . \end{cases}$$

Then the filter algebra $\mathscr{F}$ generated by the operators

$$\mathscr{L} = \frac{1}{2} \frac{d^2}{dx^2} - \frac{1}{2} x^6, \text{ and } \mathscr{L}_1 = x^3 \text{ is the Weyl algebra } W_1/\mathbb{R}.$$

A proof of this can be constructed by performing calculations similar to that in Avez-Heslot [16].

3.3 Example with Pro-finite-dimensional Lie Algebra (cf. Hazewinkel-Marcus [10]).

Consider the filtering problem:

$$\begin{cases} x_t = W_t \\ d\xi_t = x_t^2 dt \\ dy_t = x_t dt + dv_t \end{cases}$$

In [17] it was shown that all conditional moments of ξ_t can be computed using recursive filters. For this problem $\mathscr{F}$ is generated by $-x^2 \frac{\partial}{\partial \xi} + \frac{1}{2} \frac{\partial^2}{\partial x^2} - \frac{1}{2} x^2 = \mathscr{L}$ and $x = \mathscr{L}_1$. A basis for $\mathscr{F}$ is given by $\mathscr{L}$ and $x \frac{\partial^i}{\partial \xi^i}$, $\frac{\partial}{\partial} \frac{\partial^i}{\partial \xi^i}$, $\frac{\partial^i}{\partial \xi_i}$, $i = o,1, \ldots$

Defining $\mathscr{I}_i$ to be the ideal generated by $x \frac{\partial^i}{\partial \xi^i}$, $i = 0,1,2, \ldots$ it can be shown $\mathscr{F}$ is a pro-finite-dimensional filtered Lie algebra, solvable and $\mathscr{F}/\mathscr{I}_i$ is finite-dimensionall and can be realized in terms of finite-dimensional filters corresponding to conditional statistics.

Remark 1.

Other examples of finite-dimensional filters can be constructed by combining the attributes of the Benes example considered in Section 3.1 and the example of section 3.3. Thus, in example 3.3 the process x_t may be replaced by

$$dx_t = f(x_t)dt + dw_t$$

where f satisfies $\frac{df}{dx} + f^2 = ax^2 + bx + c$, and a,b,c are chosen so that this equation has a global solution. Then it is shown in [18] that all conditional moments of ξ_t can be computed using finite-dimensional recursive filters.

Remark 2

The Lie-algebraic and representation approach to the filtering problem is really concerned with the "classification" question for filters. The actual construction of the filter can apparently be achieved using probabilistic techniques.

4. Existence and Nonexistence of Finite-dimensional Filters and the Homomorphism Ansatz of Brockett.

In Section 2 we have given the definition of a finite-dimensional filter. We would consider (5) and (6) as the description of a control system with inputs y_t and output $\hat{\phi}_t$. Furthermore, as we have said we may assume that (5) - (6) is minimal in the sense of Sussmann. We thus have two ways of computing $\hat{\phi}_t$ --one via (2) - (4) (D-M-Z equation) and the other via (5) - (6). The ansatz of Brockett says: Suppose there exists a finite-dimensional filter and consider the Lie algebra of vector fields generated by $\alpha(\xi_t)$ and $\beta(\xi_t)$ and call this Lie algebra $L(\Sigma)$. Then there must exist a non-trivial homomorphism between the Filter algebra $\mathscr{F}$ and $L(\Sigma)$ such that $\mathscr{L} \to \alpha$ and $h_i \to \beta_i$ where β_i is the i^{th} row of β.

Conversely, suppose that the Lie algebra $\mathscr{F}$ cannot be generated as the Lie algebra of vector-fields with smooth coefficients on some finite-dimensional manifold, then there exists no such homomorphism and hence no conditional statistic can be computed using a finite-dimensional filter.

The Brockett ansatz suggests a possible strategy for obtaining finite-dimensional filters for computing certain conditional statistics. Suppose, we are in the situation of Example 3.3, that is, the Lie algebra $\mathscr{F}$ is pro-finite dimensional. Since $\mathscr{F}/\mathscr{I}_i$ is finite-dimensional it has a faithful finite-dimensional representation (by Ado's theorem) and hence can be realized with linear vector fields on a finite-dimensional manifold which may give rise to a bilinear filter computing some conditional statistic. However, what statistic this filter computes is in general difficult to determine, and one has to resort to indirect and probabilistic techniques for this determination. One should also remark again that $\mathscr{F}$(or any of its quotients) need not be finite-dimensional for a finite-dimensional filter to exist.

4.1 Kalman Filter Revisited

It is instructive to view the Kalman filter in the light of the above discussion and solve explicitly the corresponding D-M-Z equation. We shall consider the special case where the Filter Lie algebra is generated by $\left\{\frac{1}{2}\frac{d^2}{dx^2} - \frac{1}{2}x^2, \frac{d}{dx}, x, I\right\}$. For a rigorous justification of the calculations which follow see Ocone [9].

The basic idea is to do the following formal calculation which needs to be justified.

Suppose that we want to solve the evolution equation

$$(13) \qquad \frac{d\rho}{dt} = L_1\rho + u(t)L_2\rho \text{ , where}$$

L_1 and L_2 are in general unbounded linear operators and $u(t)$ is a given continuous function. Let us assume that the Lie algebra of operators $A\{L_1, L_2\}$ has a finite set of generators $\{L_1, L_2, \ldots, L_d\}$. We try a solution

$$(14) \qquad \rho(t) = \exp\left(g_1(t)L_1\right)\exp\left(g_2(t)L_2\right) \ldots \exp\left(g_d(t)L_d\right)\rho(o)$$

where $\rho(o)$ is the initial condition. For ideas similar to this in the context of ordinary stochastic differential equations, see Kunita [19].

Differentiating the above, we get

$$\frac{d\rho}{dt} = \dot{g}_1(t)L_1\rho + \dot{g}_2(t)\exp\left(g_1(t)L_1\right)L_2\exp\left(g_2(t)L_2\right) \ldots \exp\left(g_d(t)L_d\right)\rho(o)$$
$$+ \dot{g}_d(t)\exp\left(g_1(t)L_1\right) \ldots L_d\exp\left(g_d(t)L_d\right)\rho(o).$$

Now, we use the Campbell-Baker-Hausdorff formula: for $1 \le i, j \le d$,

$\exp(tL_j)L_i = \sum_{m=1}^{d} c_m^{i,j}(t)L_m\exp(tL_j)$ repeatedly to obtain

$$(15)\qquad \frac{d\rho}{dt} = F_1\Big(g(t),\dot g(t)\Big)L_1\rho + \ldots + F_d\Big(g(t),\dot g(t)\Big)L_d\rho$$

for some non-linear functions F_i of $g(t) = \Big(g_1(t), \ldots, g_d(t)\Big)$ and $\dot g(t)$.
For (15) to define a solution of (13), we need

$$F_1\Big(g(t),\dot g(t)\Big) = 1$$
$$F_2\Big(g(t),\dot g(t)\Big) = u(t)$$
$$F_j\Big(g(t),\dot g(t)\Big) = 0 \qquad \text{for } j>2.$$

For the Kalman-filter problem considered, one gets (formally)

$$\dot g_1(t) = 1$$
$$\dot y(t) = \dot g_2(t)\cosh g_1(t) + \dot g_3(t)\sinh g_1(t)$$
$$0 = \dot g_2(t)\sinh g_1(t) + \dot g_3(t)\cosh g_1(t)$$
$$0 = \dot g_4(t) - \dot g_3(t)g_2(t)$$
$$g_i(0) = 0\ ,\ i = 1, 2, \ldots 4.$$

One can explicitly solve the above set of equations to obtain

$$g_2(t) = \int_0^t \cosh(s)\,dy(s)$$
$$g_3(t) = -\int_0^t \sinh(s)\,dy(s)$$
$$g_4(t) = \int_0^t (\sinh s)(\cosh s)\,ds - \int_0^t g_2(s)\sinh(s)\,dy(s)$$

where we have now used stochastic integrals.

Substituting the above in (14) and using

$$(e^{tL_1}\phi)(s) = \int_{-\infty}^{\infty} G(x,y,t,)\phi(y)\,dy\ ,\ t \ge 0, \text{ where}$$

$$G(x,y,t) = (2\pi\sinh t)^{-\frac{1}{2}}\exp\left[-\frac{1}{2}(\coth)(x^2+y^2) + xy/\sinh t\right],$$

one gets

$$\rho(x,t) = \int_{-\infty}^{\infty} k(z,t)\exp\Big(-\tfrac{1}{2}p^{-1}(t)[x-m(t)]^2\Big)\rho_0(z)\,dz,$$

where $p(t) = \tanh t$

$$m(t) = \frac{z}{\cosh t} + \int_0^t \frac{\sinh s}{\cosh t}\,dy(s)$$

(and $k(z,t)$ is a function which can be computed), which is the familiar Kalman-filter solution.

The essential point in proving the above results rigorously is to note that $-\frac{1}{2}\frac{d^2}{dx^2} + \frac{1}{2}x^2$ generates a positivity-preserving Hypercontractive semigroup and that the operators $-\frac{1}{2}\frac{d^2}{dx^2} + \frac{1}{2}x^2$, x, $\frac{d}{dx}$ have a common dense set of analytic vectors.

Finally, since the Lie algebra corresponding to the Kalman filter is solvable (14) is a global representation for the solution.

We remark that the Benes problem considered in Section 3.1 can be integrated in exactly the same fashion.

Note also that this method computes the fundamental solution of the D-M-Z equation and hence these ideas can be applied to solve Kalman filtering problems with non-Gaussian initial conditions.

4.2 Non-Existence of Finite-Dimensional Filters

In an earlier part of this section we have suggested a strategy for obtaining finite-dimensional filters when the Lie algebra of the filter has a "good" ideal-structure using the Brockett Homomorphism Ansatz. We have also remarked how the same ansatz may lead to negative results.

Now, in section 3.2 we have shown that for the cubic-sensor problem the Lie algebra of the filter is isomorphic to the $W_1/\mathbb{R}$. In [10], Hazewinkel and Marcus have shown that $W_1/\mathbb{R}$ cannot be realized as the Lie algebra of vector fields with smooth coefficients on a finite-dimensional smooth manifold. On the other hand, Sussmann [11] has shown that if there is a finite-dimensional filter for a conditional statistic, then there exists a non-zero homomorphism of Lie algebras according to the Brockett prescription. Some further work combining these two ideas shows that no conditional statistic for the cubic-sensor problem can be computed using finite-dimensional filters.

We conjecture that essentially similar results can be proved for the following class of filtering problems:

$$\begin{cases} dx_t = f(x_t)dt + dw_t \\ dy_t = x_t dt + dy_t \end{cases}$$

Suppose that f satisfies:

$\frac{df}{dx} + f^2 = V(x)$, where V(x) is an even-positive polynomial. Then the Lie algebra for this filtering problem is an algebra which is isomorphic to the Weyl algebra $W_1/\mathbb{R}$, and hence all the above results of this section will hold.

4.3 Some Recent Positive Results

There have been some recent positive results using the Lie-algebra formalism. One such result is concerned with the asymptotic expansion in ε of the unnormalized conditional-density for the filtering problem

$$dx_t = ax_t dt + dw_t$$

$$dy_t = [x_t + \varepsilon (x_t)^k]dt + dy_t, \quad k \geq 1$$

$$y_o^\varepsilon = 0; \; \rho_o(x) \text{ Gaussian },$$

where ε is some small positive answer.

For this class of problems it has been shown [20], [21] that the various terms in the formal asymptotic expansion of $\rho^\varepsilon(t,x)$ can be computed by finite-dimensional filters using the ideas developed in this section.

We close this section with a remark on the identification problem for linear stochastic dynamical systems. These problems can be viewed as non-linear filtering problems and lead to Lie algebras which are known as "current-algebras" in mathematical physics. The integration of these Lie algebras in a rigorous manner has recently been done in the work of

Hazewinkel-Krishnaprasad-Marcus [22].

5. Non-linear Filtering and Hamilton-Jacobi-Bellman Theory.

An entirely different geometric approach to non-linear filtering arises by giving the D-M-Z equation a stochastic control interpretation via an exponential transformation. This was done in joint work with Wendell Fleming [6]. The exponential transformation $\rho(t,x) = \exp(-S(t,x))$ leads to Hamilton-Jacobi-Bellman equation for $S(t,x)$. It has been shown in [7] that one is interested in maximum a-posteriori probability filters or maximum-likelihood filters then these filters can be constructed using S (or equivalently) ρ. The assumption that S is a Morse-function (with parameters) leads to an interesting geometric theory for non-linear filtering. This will be developed elsewhere.

REFERENCES

1. M. Fujisaki, G. Kallianpur and H. Kunita: Stochastic Differential Equations for the Nonlinear Filtering Problem, Osaka J. of Math., Vol. 1, pp. 19-40, 1972.
2. R.W. Brockett and J.M.C. Clark: The Geometry of the Conditional Density Equation, in Analysis and Optimization of Stochastic Systems, O.L.R Jacobs et al, eds., Academic Press, New York, pp. 299-309, 1980.
3. R.W. Brockett: Remarks on Finite Dimensional Nonlinear Estimation in Analyse des Systemes. Asterisque, Vol. 75-76, pp. 47-55, 1980.
4. S.K. Mitter: Filtering Theory and Quantum Fields, in Analyse des Systemes, Asterisque, Vol. 75-76, pp. 199-205, 1980.
5. J. Glimm and A. Jaffe: Quantum Physics, Springer-Verlag, Berlin-New York, 1981.
6. W. Fleming and S.K. Mitter: Optimal Control and Nonlinear Filtering for Nondegenerate Diffusion Processes, Stochastics, 8, pp. 63-78, 1982.
7. S.K. Mitter: Approximations to Nonlinear Filtering, to appear in Nonlinear Stochastic Systems, eds. R.S. Bucy and J. Moura, Reidel, Dordrecht, 1983.
8. W. Fleming and E. Pardoux: Existence of Optimal Controls for Partially-Observed Diffusions, SIAM J. on Control, 20, pp. 261-285, 1982.
9. D. Ocone: Topics in Nonlinear Filtering Theory, Doctoral Dissertation, M.I.T., June 1980.
10. M. Hazewinkel and S.I. Marcus: On Lie Algebras and Finite Dimensional Filtering, Stochastics, 7, pp. 29-62, 1982.
11. H.J. Sussmann: Rigorous Results on the Cubic Sensor Problem, in M. Hazewinkel and J.C. Willems (eds), Stochastic Systems: The Mathematics of Filtering and Identification and Applications, Nato Advanced Study Institute Series, Reidel, Dordrecht, 1981.

12. M. Hazewinkel and J.C. Willems (eds): Stochastic Systems: The Mathematics of Filtering and Identification and Applications, Nato Advanced Study Institute Series, Reidel, Dordrecht, 1981.

13. S.K. Mitter and A. Moro (eds): Nonlinear Filtering and Stochastic Control, Proceedings, Cortona 1981, Springer Lecture Notes in Mathematics, 972, Springer-Verlag, Berlin, New York, 1983.

14. H.J. Sussmann: Existence and Uniqueness of Minimal Realizations of Nonlinear Systems, Math. Syst. Theory, Vol. 10, pp. 263-284, 1977.

15. V.E. Benes: Exact Finite Dimensional Filters for Certain Diffusions with Non-Linear Drift, Stochastics 5, pp. 65-92, 1981.

16. 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.

17. S.I. Marcus, S.K. Mitter and D. Ocone: Finite Dimensional Nonlinear Estimation for a Class of Systems in Continuous and Discrete Time, in Analysis and Optimization of Stochastic Systems, eds. O.L.R. Jacobs, et al, Academic Press, New York, pp. 387-406, 1980.

18. D. Ocone, J.S. Baras and S.I. Marcus: Explicit Filters for Diffusions with Certain Nonlinear Drifts, Stochastics, 8, pp. 1-16, 1982.

19. H. Kunita: On the Decomposition of Solutions of Stochastic Differential Equations, in Stochastic Integrals, Springer Lecture Notes in Mathematics 851, Springer-Verlag, Berlin, 1981.

20. S.I. Marcus, C.H. Liu and G.L. Blankenship: Lie Algebraic and Approximation Methods for Some Nonlinear Filtering Problems, in Advances in Filtering and Optimal Stochastic Control, eds. W.H. Fleming and L.G. Gorostiza, Springer-Verlag, 1982, pp. 225-238.

21. H.J. Sussmann: Approximate Finite-Dimensional Filters for Some Nonlinear Problems, Stochastics, 7, pp. 183-204, 1982.

22. M. Hazewinkel, P.S. Krishnaprasad and S.I. Marcus: Current Algebras and the Identification Problem, Technical Report, University of Maryland, Department of Electrical Engineering, 1982.

APPENDIX

On Lie Algebras, Lie Groups and Representations

For most of this paper, the C^∞-manifold we will be interested in is $\mathbb{R}^n$ (which is covered by a single coordinate system).

We shall say that a vector space $\mathscr{L}$ over $\mathbb{R}$ is a real Lie algebra, if in addition to its vector space structure it possesses a product $\mathscr{L} \times \mathscr{L} \to \mathscr{L}$: $(X,Y) \to [X,Y]$ which has the following properties:

(i) it is bilinear over $\mathbb{R}$

(ii) it is skew commutative : $[X,Y] + [Y,X] = 0$

(iii) it satisfies the Jacobi identity:

$$[X,[Y,Z]] = [Y,[Z,X]] + [Z,[X,Y]] = 0$$

$X,Y,Z \varepsilon \mathscr{L}$

Example: $M_n(\mathbb{R})$ = algebra of n x n matrices over $\mathbb{R}$.

If we denote by $[X,Y] = XY - YX$, where XY is the usual matrix product, then this commutator defines a

Lie algebra structure on $M_n(\mathbb{R})$.

Example: Let $\mathscr{X}(M)$ denote the C^∞-vector fields on a C^∞-manifold M. $\mathscr{X}(M)$ is a vector space over R and a $C^\infty(M)$ module. (Recall, a vector field X on M is a mapping: $M \to T_p(M)$: $p \to x$ where $p \varepsilon M$ and $T_p(M)$ is the tangent space to the point p at M). We can give a Lie algebra structure to $\mathscr{X}(M)$ by defining:

$$\mathscr{Z}_p f = (XY - YX)_p f = X_p(Yf) - Y_p(Xf) \ , \ f \varepsilon C^\infty(p)$$

(the C^∞- functions in a neighborhood of p), and

$$[X,Y] = XY - YX.$$

Both of these examples will be useful to us later on.

Let $\mathscr{L}$ be a Lie algebra over $\mathbb{R}$ and let $\{X_1, \ldots, X_n\}$ be a basis of $\mathscr{L}$ (as a vector space). There are uniquely determined constants

$c_{rsp} \varepsilon \mathbb{R}$ $(1 \leq r,s,p \leq n)$ such that

$$[X_r, X_s] = \sum_{1 \le p \le n} c_{rsp} X_p$$

The c_{rsp} are called the structure constants of $\mathscr{L}$ relative to the basis $\{X_1, \ldots, X_n\}$. From the definition of a Lie algebra:

(i) $c_{rsp} + c_{srp} = 0 \quad (1 \le r,s,p \le n)$

(ii) $\sum_{1 \le p \le n} (c_{rsp}c_{ptu} + c_{stp}c_{pru} + c_{trp}c_{psu}) = 0 \quad (1 \le r,s,t,u \le n).$

Let $\mathscr{L}$ be a Lie algebra over $\mathbb{R}$. Given two linear subspaces M,N of $\mathscr{L}$ we denote by [M,N] the linear space spanned by [X,Y], $X \in M$ and $Y \in N$. A linear subspace K of $\mathscr{L}$ is called a sub-algebra if $[K,K] \subseteq K$, an ideal if $[\mathscr{L},K] \subseteq K$.

If $\mathscr{L}$ and $\mathscr{L}'$ are Lie algebras over $\mathbb{R}$ and $\pi : \mathscr{L} \to \mathscr{L}' : X \to \pi(X)$, a linear map, π is called a homomorphism if it preserves brackets:

$$[\pi(X), \pi(Y)] = \pi([X,Y]) \quad (X, Y \in \mathscr{L}).$$

In that case $\pi(\mathscr{L})$ is a subalgebra of $\mathscr{L}'$ and ker π is an ideal in $\mathscr{L}$. Conversely, let $\mathscr{L}$ be a Lie algebra over R and K an ideal of $\mathscr{L}$. Let $\mathscr{L}' = \mathscr{L}/K$ be the quotient vector space and $\pi : \mathscr{L} \to \mathscr{L}'$ the canonical linear map. For $X' = \pi(X)$ and $Y' = \pi(Y)$, let

$$[X', Y'] = \pi([X,Y]).$$

This mapping is well-defined and makes $\mathscr{L}'$ a Lie algebra over $\mathbb{R}$ and π is then a homomorphism of $\mathscr{L}$ into $\mathscr{L}'$ with K as the kernel. $\mathscr{L}' = \mathscr{L}/K$ is called the quotient of $\mathscr{L}$ by K.

Let $\mathscr{U}$ be any algebra over $\mathbb{R}$, whose multiplication is bilinear but not necessarily associative. An endomorphism D of $\mathscr{U}$ (considered as a vector space) is called a derivation if

$$D(ab) = (Da)\, b + a\, (Db) \quad a, b \in \mathscr{U}$$

If D_1 and D_2 are derivations so is $[D_1, D_2] = D_1D_2 - D_2D_1$
The set of all derivations on $\mathscr{U}$ (assumed finite dimensional) is a subalgebra of $gl(\mathscr{U})$, the Lie algebra of all endomorphisms of $\mathscr{U}$.

For us the notion of a representation of a Lie algebra is very important.

Let $\mathscr{L}$ be a Lie algebra over $\mathbb{R}$ and V a vector space over $\mathbb{R}$, not necessarily finite dimensional. By a <u>representation</u> of $\mathscr{L}$ in V we mean a map.

$\pi : X \to \pi(X) : \mathscr{L} \to gl(V)$ (all endomorphisms of V), such that

(i) π is linear

(ii) $\pi([X,Y]) = \pi(X)\pi(Y) - \pi(Y)\pi(X)$.

For any $X \in \mathscr{L}$ let adX denote the endomorphism of $\mathscr{L}$

$\mathrm{ad}X : Y \to [X,Y] \quad (Y \in \mathscr{L})$.

adX is a derivation of $\mathscr{L}$ and $X \to \mathrm{ad}X$ is a representation of $\mathscr{L}$ in $\mathscr{L}$, called the <u>adjoint representation</u>.

Let G be a topological group and at the same time a differentiable manifold. G is a Lie group if the mapping $(x,y) \to xy : G\mathrm{x}G \to G$ and the mapping $x \to x^{-1} : G \to G$ are both C^∞-mappings.

Given a Lie group G there is an essentially unique way to define its Lie algebra. Conversely, every finite-dimensional Lie algebra is the Lie algebra of some simply connected Lie group.

In filtering theory some special Lie algebras seem to arise. We give the basic definitions for three such Lie algebras.

A Lie algebra $\mathscr{L}$ over R is said to be <u>nilpotent</u> if adX is a nilpotent endomorphism of $\mathscr{L}$, $\forall X \in \mathscr{L}$. Let the dimension of $\mathscr{L}$ be m. Then there are ideals $\mathscr{I}_j$ of $\mathscr{L}$ such that (i) $\dim \mathscr{I}_j = m-j$, $0 \le j \le m$.

(ii) $\mathscr{I}_o = \mathscr{L} \supseteq \mathscr{I}_1 \supseteq \ldots . \mathscr{I}_m = 0$ and (iii) $[\mathscr{L}, \mathscr{I}_j] \subseteq \mathscr{I}_{j+1}$, $0 \le j \le m-1$.

Let g be a Lie algebra of finite-dimension over $\mathbb{R}$ and write $\mathscr{D}g = [g,g]$. $\mathscr{D}g$ is a subalgebra of g called the derived algebra. Define $\mathscr{D}^p g$ $(p \ge o)$ inductively by

$$\mathscr{D}^\circ g = g$$

$$\mathscr{D}^p g = \mathscr{D}\mathscr{D}^{p-1} \alpha) \quad (p \ge 1).$$

We then get a sequence $\mathscr{D}^0 g \supseteq \mathscr{D}^1 g \supseteq \ldots$ of subalgebras of g. g is said to be solvable if $\mathscr{D}^p g = 0$ for some $p \geq 1$.

Examples

(i) Let $n \geq 0$ and let $(p_1, \ldots, p_n, q_1, \ldots q_n, z)$ be a basis for a real vector space $\mathscr{V}$. Define a Lie algebra structure on $\mathscr{V}$ by $[p_i, q_i] = [q_i, p_i] = Z$, the other brackets being zero. This nilpotent Lie algebra $\mathscr{N}$ is the so-called Heisenberg algebra.

(ii) The real Lie algebra with basis $(h, p_1, \ldots, p_n, q_1, \ldots q_n, Z)$ satisfying the bracket relations

$[h, p_i] = q_i$, $[h, q_i] = p_i$, $[p_i q_i] = Z$, the other brackets being zero is a solvable Lie algebra, the so-called oscillator algebra. Its derived algebra is the Heisenberg algebra $\mathscr{N}$.

A Lie algebra is called simple is it has no nontrivial ideals. An infinite dimensional Lie algebra $\mathscr{L}$ is called pro-finite dimensional and filtered if there exists a sequence of ideals $\mathscr{I}_1 \supset \mathscr{I}_2 \ldots$ such $\mathscr{L}/\mathscr{I}_i$ is finite-dimensional for all i and $\cap \mathscr{I}_i = \{0\}$.

Infinite-Dimensional Representations

Let g be a finite dimensional Lie algebra and G its associated simply connected Lie group. Let H be a complex Hilbert space (generally infinite-dimensional). We are interested in representations of g by means of linear operators on H with a common dense invariant domain $\mathscr{D}$. Let π denote this representation.

Similarly, we are also interested in representations of G as bounded linear operators on H. Let τ be such a representation. That is, $\tau : G \to L(H)$ satisfies

$$\tau(g_1 g_2) = \tau(g_1)\, \tau(g_2) \quad , \; g_1, g_2 \in G.$$

The following problem of Group representation has been considered by Nelson and others. Given a representation π of g on H when does there exist a group representation (strongly continuous) τ of G on H such that

$$\tau\left(\exp\ (tX)\right) = \exp\ \left(t\pi(X)\right)\ \forall X\varepsilon G$$

Here exp $\left(t\pi(x)\right)$ is the strongly continuous group generated by $\pi(X)$ in the sense that

$$\frac{d\cdot}{dt}\ \exp\left(t\pi(x)\right)\phi = \pi(X)\phi\quad \forall\ \phi\varepsilon\mathcal{D}$$

and exp (tX) is the exponential mapping, mapping the Lie algebra g into the Lie group G.

Let $X_1, \ldots X_d$ be a basis for g. A method for constructing τ locally is to define

$$\tau\left(\exp(t_1X_1)\ \ldots\ \exp(t_dS_d)\right) = \exp\left(t_1\pi(X_1)\right)\ \ldots\ \exp\ \left(t_d\pi(x_d)\right)$$

A sufficient condition for this to work is that the operator identity

$$(3.1)\quad \exp(tA_j)A_i = \sum_{n=o}^{\infty}\frac{t^n}{n!}[adA_j]^nA_i\exp(tA_j)$$

holds for $A_j = \pi(X_j)$, $1\leq j$, $j\leq d$.

It is a well known fact, that many Lie algebra representations do not extend to Group representations. An example is the representation of the Heisenberg algebra consisting of three basis elements by the operators $\{-ix, \frac{d.}{dx}, -i\}$ on $L^2(\mathbb{R}_+)$ with domain $C_o^\infty(\mathbb{R}_+)$ which does not extend to a unitary representation (since essential self-adjointness fails).

Although in filtering theory we are not interested in unitary group representation, nevertheless these ideas will serve as a guide for integrating the Lie algebras arising in filtering theory.