Sanjoy K. Mitter

EXISTENCE AND NON-EXISTENCE OF FINITE DIMENSIONAL FILTERS

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 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 focussing 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 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 here. In many senses, this view point has been remarkably successful,although the results obtained so far have been of a negative nature.

This expository paper describes the work done by Baras, Benes, Brockett, Clark, Davis, Hazewinkel, Hijab, Marcus, Ocone and Sussman and the work of the author (the most recent work of the author being joint work with Fleming). An account of these ideas may be found in the Proceedings of the Les Arcs Conference on *Stochastic Systems: The Mathematics of Filtering and*

Identification eds. M. Hazewinkel and J.C. Willems, D. Reidel Publ. Company, 1981 and in as yet unpublished work of the author. The programme outlining this approach can be found in the work of Brockett [2], [3] and in the author's paper [5]. See also the doctoral dissertation of D. Ocone [6], written under this author's direction.

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

We consider the signal-observation model:

$$(2.1)\quad \begin{cases} dx_t = f(x_t)\,dt + G(x_t)\,dw_t \;\; ; \;\; x(0) = x_0 \\ dy_t = h(x_t)\,dt + d\eta_t \;, \quad \text{where} \end{cases} \qquad 0 \leqslant t \leqslant 1$$

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^∞-functions). It is further assumed that the stochastic differential equation (1) has a global solution in the sense of Ito. It is also assumed that x_t and η_t are independent and $E\int_0^1 |h(x_t)|^2 dt < \infty$. For most of our considerations, the function $h(\cdot)$ will be a polynomial.

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

$$(2.2)\quad d\rho(t,x) = \left(\mathcal{L}_0^* - \frac{1}{2}\sum_{i=1}^{p} h_i^2(x)\right)\rho(t,x)\,dt + \sum_{i=1}^{p} h_i(x)\,\rho(t,x) \circ dy_t \,,$$

where

$$(2.3)\qquad \mathcal{L}_0^*\varphi = \sum_{i,j=1}^{n} \frac{d^2}{dx_i dx_j}\,(G(x)\,G'(x))_{ij}\,\varphi - \sum_{i=1}^{n} \frac{d}{dx_i}\,f_i(x)\,\varphi$$

and the $\circ$ denotes Stratonvich differential. It is imperative that we consider (2.2) as a Stratonvich 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 (2.2) 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 function than the filter typically is required to compute $E(\varphi(x_t)|\mathcal{F}_t^y)$, where $\mathcal{F}_t^y = \sigma\{y_s, 0 \leqslant s \leqslant t\}$. If we denote by $\hat{\varphi}_t \triangleq E(\varphi(x_t)|\mathcal{F}_t^y)$, then $\hat{\varphi}_t$ is obtained from $\rho(t,x)$ by integration:

$$(2.4) \qquad \hat{\varphi}_t = \int_{\mathbb{R}^n} \varphi(x)\rho(t,x)\,dx \Big/ \int_{\mathbb{R}^n} \varphi(t,x)\,dx$$

$\hat{\varphi}_t$ will be referred to as a "conditional statistic", and no matter what $\hat{\varphi}_t$ 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 $\hat{\varphi}_t$ (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 (2.2) so that we can evaluate $\rho(t,x)$ corresponding to any initial condition.

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

$$(2.5) \qquad d\xi_t = \alpha(\xi_t)\,dt + \beta(\xi_t) \circ dy_t$$

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

$$(2.6) \qquad \hat{\varphi}_t = \gamma(\xi_t) \; ,$$

which computes the conditional statistic. Equation (2.5) is to be interpreted in the Stratanovich sense for reasons we have mentioned above. We shall also assume that the stochastic dynamical system (2.5)-(2.6) is minimal in the sense of Sussmann (cf. later section).

3. Preliminaries. (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 $\mathcal{L}$ ove $\mathbb{R}$ is a *real Lie algebra,* if in addition to its vector space structure it posseses a product $\mathcal{L} \times \mathcal{L} \longrightarrow \mathcal{L} : (X,Y) \longrightarrow [X,Y]$ 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: $[X,[Y,Z]] + [Y,[Z,X]] + [Z,[X,Y]] = 0$.

(ii)–(iii) hold $\forall X, Y, Z \in \mathcal{L}$ ∎

Example: $\mathcal{M}_n(\mathbb{R})$ = algebra of $n \times n$ 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 $\mathcal{M}_n(\mathbb{R})$. ■

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

$$Z_p f = (XY - YX)_p f = X_p(Yf) - Y_p(Xf) , \qquad f \in C^\infty(p)$$

(the C^∞-functions in a neighbourhood of p), and $[X, Y] = XY - YX$. ■

Both these examples will be useful to us later on.

Let $\mathcal{L}$ be a Lie algebra over $\mathbb{R}$ and let $\{X_1, \dots, X_n\}$ be a basis of $\mathcal{L}$ (as a vector space). There are uniquely determined constants $c_{rsp} \in \mathbb{R}$ $(1 \leqslant r, s, p \leqslant n)$ such that

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

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

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

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

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

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

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

In that case $\pi(\mathcal{L})$ is a *subalgebra* of $\mathcal{L}'$ and $\ker \pi$ is an ideal in $\mathcal{L}$. Conversely let $\mathcal{L}$ be a Lie algebra over $\mathbb{R}$ and $\mathcal{K}$ an *ideal* of $\mathcal{L}$. Let $\mathcal{L}' = \mathcal{L}/\mathcal{K}$ be the quotient vector space and $\pi: \mathcal{L} \longrightarrow \mathcal{L}'$ 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 $\mathcal{L}'$ a Lie algebra over $\mathbb{R}$ and π is then a homomorphism of $\mathcal{L}$ into $\mathcal{L}'$ with $\mathcal{K}$ as the kernel. $\mathcal{L}' = \mathcal{L}/K$ is called the *quotient* of $\mathcal{L}$ by K.

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

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

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 $\mathcal{U}$ (assumed finite dimensional) is a subalgebra of $gl(\mathcal{U})$, the Lie algebra of all endomorphisms of $\mathcal{U}$. ■

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

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

$$\pi : X \longmapsto \pi(X) : \mathcal{L} \longrightarrow gl(V) \qquad \text{(all endomorphisms of } V\text{)}$$

such that

(i) π is linear

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

For any $X \in \mathcal{L}$, let $\operatorname{ad} X$ denote the endomorphism of $\mathcal{L}$

$\operatorname{ad} X : Y \longmapsto [X, Y] \qquad (Y \in \mathcal{L})$.

$\operatorname{ad} X$ is a derivation of $\mathcal{L}$ and $X \longmapsto \operatorname{ad} X$ is a representation of $\mathcal{L}$ in $\mathcal{L}$, called the *adjoint representation*. ■

Let G be a topological group and at the same time a differentiable manifold. G is a Lie group if the mapping $(x,y) \longmapsto xy : G \times G \longrightarrow G$ and the mapping $x \longmapsto x^{-1} : G \longrightarrow G$ 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 $\mathcal{L}$ over $\mathbb{R}$ is said to be *nilpotent* if $\operatorname{ad} X$ is a nilpotent endomorphism of $\mathcal{L}$, $\forall X \in \mathcal{L}$. Let the dimension of $\mathcal{L}$ be m. Then there are ideals $\mathcal{I}_j$ of $\mathcal{L}$ such that

(i) $\dim \mathscr{I}_j = m - j$, $0 \leqslant j \leqslant m$
(ii) $\mathscr{I}_0 = \mathcal{L} \supseteq \mathscr{I}_1 \supseteq \ldots \supseteq \mathscr{I}_m = 0$ and
(iii) $[\mathcal{L}, \mathscr{I}_j] \subseteq \mathscr{I}_{j+1}$, $0 \leqslant j \leqslant m-1$.

Let g be a Lie algebra of finite-dimension over $\mathbb{R}$ and write $\mathscr{D}g = [g, g]$. $\mathscr{D}g$ is a sub-algebra of g called the derived algebra. Define $\mathscr{D}^p g$ ($p \geqslant 0$) inductively by

$$\mathscr{D}^0 g = g$$
$$\mathscr{D}^p g = \mathscr{D}(\mathscr{D}^{p-1} g) \qquad (p \geqslant 1).$$

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

Examples

(i) Let $n \geqslant 0$ 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

$$[h, p_i] = q_i \ , \quad [h, q_i] = p_i \ , \quad [p_i, q_i] = z \ ,$$

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* if it has no non-trivial ideals.

An infinite dimensional Lie algebra $\mathcal{L}$ is called *profinite dimensional* and filtered if there exists a sequence of ideals $\mathscr{I}_1 \supset \mathscr{I}_2 \ldots$ such $\mathcal{L}/\mathscr{I}_i$ is finite-dimensional for all i and $\cap \mathscr{I}_i = \{0\}$.

3.1 - *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 $\mathscr{H}$. Let τ be such a representation. That is, $\tau: G \longrightarrow$ $\longrightarrow L(H)$ satisfies

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

The following problem of Group representation has been considered by Nelson [7] 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

$$\tau(\exp(tX)) = \exp(t\pi(X)) \qquad \forall X \in G \, .$$

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

$$\frac{d}{dt}\exp(t\pi(X))\varphi = \pi(X)\varphi \qquad \forall \varphi \in \mathscr{D}$$

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

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

$$\tau(\exp(t_1 X_1) \dots \exp(t_d X_d)) = \exp(t_1 \pi(X_1)) \dots \exp(t_d \pi(X_d)) \, .$$

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

$$\exp(tA_j)A_i = \sum_{n=0}^{\infty} \frac{t^n}{n!} [\mathrm{ad}\, A_j]^n A_i \exp(tA_j) \tag{3.1}$$

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

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 $\left\{-ix, \frac{d}{dx}, -i\right\}$ on $L^2(\mathbb{R}_+)$ with domain $C_0^\infty(\mathbb{R}_+)$ which does not extend to a unitary representation (since essential self-adjointness fails).

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

4. Lie Algebra of Operators Associated with the Filtering Problem.

Consider the unbounded operators

$$\mathcal{L} = \mathcal{L}_0^* - \frac{1}{2}\sum_{i=1}^{p} h_i^2(x) \text{ and } h_i(x) , \qquad i = 1, \dots, p ,$$

where the operators $h_i(x)$ are considered as multiplication operators $\varphi(x) \longmapsto h_i(x)\varphi(x)$, and $\mathcal{L}$ and h_i are defined on some common dense invariant domain $\mathcal{D}$ (say $\mathcal{D} = C_0^\infty(\mathbb{R}^n)$ or $\mathcal{S}(\mathbb{R}^n)$).

The Lie algebra of differential operators $\mathcal{F}$ generated by $\mathcal{L}$ and $h_1, \dots, h_p$ is called the *filter algebra.*

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 (it is not being said that if this Lie algebra is infinite--dimensional that no finite-dimensional filter exists).

Therefore the first question that arises is: are there examples of non--linear filtering problems with finite dimensional filter algebras? The second question is:

How large is this class? The answer to the first question is – yes –, but the answer to the second question appears to be is that this class is small.

Example 1 (*Kalman Filtering*)

(4.1) $$\begin{cases} dx_t = Ax_t dt + b\, dw_t \\ dy_t = c'x_t dt + d\eta_t \end{cases} \qquad \begin{aligned} A &= n \times n \text{ matrix} \\ b &= n \times 1 \text{ matrix} \\ c &= n \times 1 \text{ matrix} \end{aligned}$$

Then

(4.2) $$\begin{cases} \mathcal{L}_0^* - \frac{1}{2}\sum_{i,j=1}^{n} \frac{\partial^2}{\partial x_i \partial x_j} Q_{ij} - \sum_{i=1}^{n} \frac{\partial}{\partial x_i}(Ax)_i , \quad \text{and} \\ \mathcal{L}_0 = \mathcal{L}_0^* - \frac{1}{2}(c'x)^2 , \quad \text{where} \\ Q = bb' \end{cases}$$

Define the Hamiltonian matrix

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

and the vector

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

and the controllability matrix

$$W = [\alpha \vdots E\alpha \vdots \ldots\ldots \vdots E^{2n-1}\alpha]$$

and assume that W is non-singular.

Define $Z_1 = c'x$ and

$$Z_i = [\mathrm{ad}\, \mathcal{L}_0]^{i-1} Z_1 \ .$$

Then one can show that

$$Z_i = \sum_{j=1}^{n} (E^{i-1}\alpha)_j x_j + \sum_{j=1}^{n} (E^{i-1}\alpha)_{j+n} \frac{\partial}{\partial x_j} \ , \tag{4.3}$$

and

$$[Z_i, Z_j] = (E^{i-1}\alpha)' \begin{pmatrix} 0 & -I \\ I & 0 \end{pmatrix} (E^{j-1}\alpha) \ . \tag{4.4}$$

We can then conclude that the filter algebra $\mathscr{F} = \mathrm{span}\,\{\mathcal{L}_0, Z_1, \ldots, Z_{2n}, I\}$, 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 $2n + 2$.

4.1 - *Invariance Properties of the Lie Algebra and the Benes Problem.*

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

Consider the filtering problem:

$$\begin{cases} x_t = w_t \\ dy_t = x_t dt + d\eta_t \end{cases} \tag{4.5}$$

A basis for the filter algebra is

$$\left\{ \mathcal{L}_0 \, , \, x \, , \, \frac{d}{dx} \, , \, I \right\} \, , \quad \text{where}$$

$\mathcal{L}_0 = \frac{1}{2}\frac{d^2}{dx^2} - \frac{1}{2}x^2$ and this is the 4-dimensional oscillator algebra.

It is easy to see that if we perform a smooth change of coordinates

$x \mapsto \varphi(x)$ then the Filter algebra gives rise to an isomorphic Lie algebra.

Now consider the example first treated by Benes [10],

$$(4.6) \qquad \begin{cases} dx_t = f(x_t)\,dt + dw_t \\ dy_t = x_t\,dt + d\eta_t \end{cases}, \quad \text{where}$$

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 $\mathbb{R}$. We want to show that by introducing gauge transformations, we can transform the filter algebra of (4.6) 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 (4.6)

$$[\mathcal{L}_0, x] = \frac{d}{dx} - f\,,$$

where the brackets are computed on $C_0^\infty(\mathbb{R})$.

Now consider the commutative diagram:

$$\begin{array}{ccc} C_0^\infty(\mathbb{R}) & \xrightarrow{\quad \frac{d}{dx} \quad} & C_0^\infty(\mathbb{R}) \\ \psi \big\downarrow & & \big\downarrow \Psi \\ C_0^\infty(\mathbb{R}) & \xrightarrow[\quad \frac{d}{dx} - f \quad]{} & C_0^\infty(\mathbb{R}) \end{array}$$

Here ψ is the multiplication operator $\varphi(x) \longrightarrow \psi(x)\varphi(x)$ and it is assumed that ψ is invertible. Then it is easy to see that

$$\psi(x) = \exp\left(\int_0^x f(z)\,dz\right).$$

Under the transformation ψ, the operator $\mathcal{L}_0^* = \frac{1}{2}\frac{\partial^2}{\partial x^2} - \frac{\partial}{\partial x} f$ transforms to $\frac{1}{2}\frac{\partial^2}{\partial x^2} - \frac{1}{2}V(x)$, where $V(x) = f_x + f^2$.

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

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

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 transformation

$$\rho(t,x) \mapsto \psi^{-1}(x)\,\rho(t,x),$$

where $\rho(t,x)$ is the solution of the Zakai equation and what we have shown in that the Filter algebra is invariant under this isomorphism.

However, for the class of models considered in (4.6) with general drifts f, the Benes problem is the only one with a finite-dimensional Lie algebra (diffusions defined on the whole real line).

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

4.2 - *The Weyl Algebras and the Cubic Sensor Problem.*

The Weyl algebra W_n is the algebra of all polynomial differential operators $\mathbb{R}\langle x_1, \dots, x_n, \frac{\partial}{\partial x_1}, \dots, \frac{\partial}{\partial x_n}\rangle$.

A basis for W_n consists of all monomial expressions

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

where α, β range over all multiindices $\alpha = (\alpha_1, \dots, \alpha_n)$, $\beta = (\beta_1, \dots, \beta_n)$. W_n can be endowed with a Lie algebra structure in the usual way. The centre of W_n, that is the ideal $\mathcal{I} = \{Z \in W_n \mid [X,Z] = 0,\ \forall X \in W_n\}$ 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:

$$(4.7) \qquad \begin{cases} x_t = w_t \\ dy_t = x_t^3\,dt + d\eta_t\,. \end{cases}$$

Then the filter algebra $\mathcal{F}$ generated by the operators $\mathcal{L}_0 = \frac{1}{2}\frac{d^2}{dx^2} - \frac{1}{2}x^6$, and $\mathcal{L}_1 = x^3$ is the Weyl algebra W_1.

4.3 - *Example with Pro-finite-dimensional Lie Algebra.*

Consider the filtering problem:

$$(4.8)\qquad \begin{cases} x_t &= w_t \\ d\xi_t &= x_t^2\,dt \\ dy_t &= x_t\,dt + dv_t \end{cases}$$

In [8] it was shown that all conditional moments of ξ_t can be computed using recursive filters. For this problem $\mathcal{F}$ is generated by $-x^2 \dfrac{\partial}{\partial \xi} + \dfrac{1}{2}\dfrac{\partial^2}{\partial x^2} - \dfrac{1}{2}x^2 = \mathcal{L}_0$ and $x = \mathcal{L}_1$. A basis for $\mathcal{F}$ is given by $\mathcal{L}_0$ and $\left\{ x\dfrac{\partial^i}{\partial \xi^i}, \dfrac{\partial}{\partial x}\dfrac{\partial^i}{\partial \xi^i}, \dfrac{\partial^i}{\partial \xi^i}, i = 0, 1, \dots, 2 \right\}$. Defining $\mathcal{J}_i$ to be the ideal generated by $x\dfrac{\partial^i}{\partial \xi^i}$, $i = 0, 1, 2, \dots$ it can be shown $\mathcal{F}$ is a pro-finite-dimensional filtered Lie algebra, solvable and $\mathcal{F}/\mathcal{J}_i$ is finite-dimensional and can be realized in terms of finite-dimensional filters corresponding to conditional statistics.

5. The Homomorphism Ansatz of Brockett.

In Section 2 we have given the definition of a finite-dimensional filter. We could consider (2.5) and (2.6) as the description of a control system with inputs $\dot{y}_t$ and output $\hat{\varphi}_t$. Furthermore, we may assume that (2.5)-(2.6) is *minimal* in the sense of Sussmann [9]. We thus have two ways of computing $\hat{\varphi}_t$-one via (2.2)-(2.4) (Zakai equation) and the other via (2.5)-(2.6). 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 exists a non-trivial homomorphism between $\mathcal{F}$ and $L(\Sigma)$ such that $\mathcal{L}_0 \mapsto \alpha$ and $b_i \mapsto \beta_i$ where β_i is the i^{th} row of β.

Conversely, suppose that the Lie algebra $\mathcal{F}$ 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 *no conditional statistic* can be computed using a finite-dimensional filter.

5.1 - *The Kalman Filter Revisited.*

It is instructive to view the Kalman Filter in the light of Brockett's Ansatz and to solve explicitly the Zakai equation. We write the solution as:

(5.1) $\rho(t,x) = [(\exp(t\mathcal{L}_0)\exp(g_1(t)Z_1)\ldots\exp(g_{2n}(t)Z_{2n})\exp g_{2n+1}(t))\rho_0](x)$

and solve for the $g_i(t)$.

We rewrite (5.1) in a more convenient form using:

$$(5.2)\qquad (\exp(sZ_i)f\cdot)(x) = \exp\left[\frac{1}{2}(E^{i-1}\alpha)_1'(E^{i-1}\alpha)_2 s^2 + \right.$$
$$\left. + \{(E^{i-1}\alpha\cdot)_1' x\}s\right] \times f(x + (E^{i-1}\alpha\cdot)_2 s)$$

where $(E^{i-1}\alpha\cdot)_1$ are the first n entries of $E^{i-1}\alpha$ and $(E^{i-1}\alpha\cdot)_2$ are the remaining n entries.

Using this we get:

$$\text{R.H.S. of (5.1)} = \exp(l(t))\exp\left(\sum_{i=1}^{2n} g_i(t)(E^{i-1}\alpha)_1' x\right)\rho_0\left(x + \sum_{i=1}^{2n} g_i(t)(E^{i-1}\alpha\cdot)_2\right)$$

where

$$l(t) = g_{2n+1}(t) - \frac{1}{2}\sum_{i=1}^{2n} g_i^2(t)(E^{i-1}\alpha)_1'(E^{i-1}\alpha)_2$$
$$- \sum_{i=1}^{2n}\sum_{j=i+1}^{2n} g_i(t)g_j(t)(E^{j-1}\alpha)_1'(E^{j-1}\alpha\cdot)_2$$

Thus if we define $v(t) = \sum_{i=1}^{2n} g_i(t)E^{i-1}\alpha = \begin{pmatrix} v_1(t) \\ v_2(t)\end{pmatrix}$, where

$$v_1(t) = \sum_{i=1}^{2n} g_i(t)(E^{i-1}\alpha)_1$$

$$v_2(t) = \sum_{i=1}^{2n} g_i(t)(E^{i-1}\alpha)_2 ,$$

then we can write

$$(5.3)\qquad \rho(t,x) = (\exp(t\mathcal{L}_0)[\exp(l(t))\exp(v_1'(t)y).\rho_0(y + v_2(t))].)(x) .$$

Now differential equations for $g_i(t)$ can be obtained by using the Baker-Campbell-Hausdorff formula and formal differentiation of both sides of (5.1) and equating coefficients. Also from (5.3) we see that

$$Z(t) = \begin{pmatrix} t \\ v(t) \\ l(t) \end{pmatrix} \quad \text{is a sufficient statistic.}$$

Finally using the differential equations for $g_i(t)$ we obtain:

$$\dot{Z}(t) = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \dot{y}_t \begin{pmatrix} 0 \\ \exp(-Z_1(t)E)\alpha \cdot \\ -[Z_{n+2}(t) \dots Z_{2n+1}(t)](\exp(-Z_1(t)E))_1 \end{pmatrix}$$

This system computes $\rho(t,x)$.

For later use: define

$$F = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \qquad G_1 = \begin{pmatrix} 0 \\ \exp(-Z_1(t)E)\alpha \cdot \\ -[Z_{n+2}(t) \dots Z_{2n+1}(t)](\exp^{-Z_1(t)E})_1 \end{pmatrix}$$

considered as vector fields.

Define

$$G_i = [\text{ad } F]^{i-1} G_1 .$$

The required homomorphism is

$$\begin{aligned} \mathcal{L}_0 &\longrightarrow F \\ Z_i &\longrightarrow G_i \qquad 1 \leqslant i \leqslant 2n \\ I &\longrightarrow \frac{\partial}{\partial Z_{2n+2}} \end{aligned}$$

which in fact is an isomorphism.

As we have mentioned in section 3.1, the crucial question in making the above results rigorous is the Campbell-Baker-Hausdorff formula (3.1) for operators. For the problem considered, this can be made rigorous by using the properties of the semi-group $e^{t\mathcal{L}_0}$ which has strong contractive properties.

We have referred to the homomorphism theorem as an Ansatz, because it has not been rigorously proved in general since it is not known whether the Sussman minimal realization theorem for bilinear system [9] extends to an infinite-dimensional situation.

Finally, since the Lie algebra of the Kalman filter is solvable, the representation (5.1) is global. It should be emphasized that the method advocated in this section provides a finite-dimensional statistic for the fundamental solution of the Zakai equation and hence in a certain sense the Kalman Filtering problem has a finite-dimensional sufficient statistic even for non-gaussian initial conditions.

The Benes filtering problems, being "gauge equivalent" to the Kalman Filtering problem is amenable to rigorous treatment using these same ideas.

5.2 - *Non-existence of Finite Dimensional Filters.*

Hazewinkel and Marcus [loc. cit., Les Arcs Proceedings] have rigorously shown that the Weyl algebra W_n cannot be realized as the Lie algebra of vector fields with smooth coefficients on a finite dimensional manifold. For the cubic sensor problem Sussman [loc. cit., Les Arcs Proceedings] has shown that if there exists a finite-dimensional filter for a conditional statistic then there exists a homomorphism according to the Brockett Ansatz. Combining these two results one can conclude: there exists no finite dimensional filter for computing any conditional statistic for the cubic sensor problem.

Sussman's result uses the pathwise formulation of non-linear filtering as originally proposed by Clark and subsequently developed by Davis [cf. Les Arcs Proceedings]. This is required to prove existence, uniqueness and continuity with respect to observation $y_{(.)}$ of the Zakai equation.

In recent work of the author (jointly with W. Fleming) existence, uniqueness and smoothness with respect to $y_{(.)}$ of the pathwise version of the Zakai equation has been proved for scalar observations $h(\cdot)$ which are polynomial, and for certain vector observations which may be unbounded. This work provides a stochastic control interpretation of the pathwise version of the Zakai equations, and also proves existence and uniqueness of the Zakai equation for certain unbounded observations [11].

REFERENCES

[1] M. Fujisaki, G. Kallianpur, H. Kunita: Stochastic Differential Equations of Non--linear filtering, Osaka J. of Math., Vol. 9, 1972, 19-40.

[2] R.W. Brockett and J.M.C. Clark: On the Geometry of the Conditional Density Equation; in *Analysis and Optimization of Stochastic Systems*, eds. O.L.R. Jacobs et al., Acad. Press, 1980.

[3] R.W. Brockett: Remarks on Finite Dimensional Non-linear Estimation, Asterisque, 1980 (Bordeaux 1978).

[4] S.K. Mitter: Filtering Theory and Quantum Fields, Analyse des Systèmes, Asterisque, Vol. 75-76, 1980 (Bordeaux 1978).

[5] S.K. Mitter: On the Analogy between Mathematical Problems of Nonlinear Filtering and Quantum Physics, Ricerche di Automatica, 10 (1980), 163-216.

[6] D. Ocone: *Topics in Nonlinear Filtering Theory*, Ph.D thesis M.I.T., June 1980.

[7] E. Nelson: Analytic Vectors, Annals of Math., 70, 1959, 572-615.

[8] S. Marcus, S.K. Mitter and D. Ocone: Finite Dimensional Non-linear Estimation in Continuous and Discrete Time, in *Analysis and Optimization of Stochastic Systems,* eds. O.L.R. Jacobs et al., Acad. Press, 1980.

[9] H. Sussmann: Existence and Uniqueness of Minimal Realizations of Non-linear Systems, Math. Syst. Th., 10, 1977, 263-284.

[10] V.E. Benes: Exact Finite Dimensional Filters for Certain Diffusions with Non-linear Drift, Stochastics 5 (1981), 65-92.

[11] W. Fleming and S.K. Mitter: Optimal Stochastic Control and Pathwise Filtering of Non-degenerate Diffusions, to appear in Stochastics.

Added in Proof.

M. Hazewinkel and S.I. Marcus, On Lie Algebras and Finite Dimensional Filtering, Stochastics 7 (1982), 29-62.

S.K. Mitter, Lectures on Nonlinear Filtering and Stochastic Control, *Proceedings of the CIME Course on Nonlinear Filtering and Stochastic Control,* Cortona, Italy, July 1981. To be published by Springer-Verlag, Berlin, N.Y., 1982.

This research has been supported by the Air Force Office of Scientific Research under Grant AFOSR 77-3281 B.

This paper was written while the author was a CNR Visiting Professor (Mathematics Group, GNAFA) at the Ist. Mat. U. Dini, Università di Firenze, Firenze, Italy.

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