December, 1983 LIDS-P-1349

SOME MATHEMATICAL CONNECTIONS BETWEEN NONLINEAR FILTERING AND QUANTUM PHYSICS

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

To appear in
Encyclopedia of Systems and Control,
Ed., M.G. Singh, Pergamon, 1984

This research has been supported by the Air Force Office of Scientific Research under Grant AFOSR-82-0135A.

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 1972) in their seminal paper. 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 the construction of the filter can 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 1981). 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 1982; Mitter 1983) and the work on the partially observable stochastic control problem by (Fleming-Pardoux 1982) can be considered to have arisen from the "state-space" interpretation given to the filter.

2. The Filtering Problem Considered, And the Basic Questions

We consider the signal-observation model:

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

x, w and y are $\mathbb{R}^n$, $\mathbb{R}^m$ and $\mathbb{R}^p$-valued processes, and it is assumed that b, ρ 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 $E\int_0^1 |h(x_t)|^2 dt < \infty$. 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:

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

where

$$\mathscr{L}^*\phi = \sum_{i,j=1}^{n} \frac{\partial^2}{\partial x_i \partial x_j} (\sigma(x)\sigma'(x))_{ij}\phi - \sum_{i=1}^{n} \frac{\partial}{dx_i} b_i(x)\phi \qquad (3)$$

and the $\cdot$ 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 $\hat{\phi}_t \underset{-}{\Delta} E(\phi(x_t)|\mathscr{F}_t^y)$, then $\hat{\phi}_t$ is obtained from $\rho(t,x)$ by integration:

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

$\hat{\phi}_t$ 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."

2. Pathwise Non-Linear Filtering and Analogy of the D-M-Z Equation to Schrödinger Equations

The D-M-Z equation bears a close resemblance to a Schrondinger equation with a random potential if we formally rewrite it as

$$\frac{d\rho}{dt}(t,x) = \mathscr{L}^*\rho(t,x) - \frac{1}{2}\left[\sum_{i=1}^{p} h_i^2(x) - 2\left(\sum_{i=1}^{p} h_i(x)(\dot{y}_t)_i\right)\right]\rho(t,x). \tag{5}$$

where $\dot{y}_t$ is the formal derivative of y_t. However since the operator $\rho(t,x) \to \sum_{i=1}^{p} h_i(x)(\dot{y}_t)_i\rho(t,x)$ is a multiplication operator we can transform this equation by utilizing a time-dependent gauge transformation. To simplify the notation, we assume y is scalar and in the sequal we use subscript x to denote partial derivative.

This leads us to ideas of pathwise nonlinear filtering (Clark 1978; Davis 1980; and Mitter 1980.

There is as yet no theory of non-linear filtering where the observations are:

$$Y(t) = h(x(t)) + \tilde{w}(t)^{(1)} \tag{6}$$

where $\tilde{w}$ is a physical wide-band noise and hence smooth. Define $Y(t) = \dot{y}(t)$ and $\tilde{w}(t) = \dot{\eta}(t)$ where $\cdot$ denotes differentiation. Then (7) can be written as:

$$dy(t) = h(x(t))dt + d\dot{\eta}(t), \text{ of} \tag{8}$$

$$y(t) = \int_0^t h(x(t))dt + \eta(t) \tag{9}$$

Equation (9) is a mathematical model of the physical observation (7) where the wide band noise $\tilde{w}(t)$ has been approximated as "white noise" $\dot{\eta}(t)$ and hence $\eta(t)$ is a Wiener process.

Now, if we wish to compute

$$E(\phi((x(t))|\mathscr{F}_t^y) = \text{Functional of } y \text{ a.s. Wiener measure}$$

then this filter does not accept the physical observation y. The idea is to at least construct a suitable version of the conditional expectation so that the performance of the filter as measured by the mean-square error remains close when the physical observation 'Y' is replaced by the mathematical model of the observation.

This is most easily done by eliminating the stochastic integral in (2) by a suitable transformation (gauge transformation in the language of physicists).

---

(1) To conform to a P.D.E. viewpoint we are writing processes as $\tilde{w}(t)$ etc. instead of $\tilde{w}_t$ etc.

Define $q(t,x)$ by

$$\rho(t,x) = \exp(h(x)y(t))q(t,x) \tag{10}$$

Then $q(t,x)$ satisfies the parabolic partial differential equation

$$\begin{cases} q_t = (L^y)^* q + \tilde{V}^y q, \quad \text{where} \\ L^y\phi = L\phi - y(t)a(x)h_x(x)\cdot\phi_x \qquad (a(x) = \sigma(x)\sigma'(x)) \\ V^y(t,x) = \frac{1}{2}h^2(x) - y(t)Lh(x) + \frac{1}{2}y^2(t)h_x(x)'a(x)h_x(x) \end{cases} \tag{11}$$

Equation (11) is the pathwise non-linear filtering equation and should be solved for each observation path y (which can be taken to be physical observation). Equation (11) can be written explicitly as

$$\begin{cases} q_t = \frac{1}{2}\operatorname{tr} a\ (x)q_{xx} + g^y(x,t)\ \cdot\ q_x + V^y(x,t)q \\ q(0,x) = p^0(x), \text{ the density of } x(0), \text{ where} \\ \begin{cases} g^y = -b + y(t)ah_x + \gamma, \quad \gamma_j = \sum_{i=1}^{n} \frac{\partial a_{ij}}{\partial x_j}, \quad j = 1,2,\dots n \\ V^y = \tilde{V}^y - \operatorname{div}(b - y(t)ah_x) + \frac{1}{2}\sum_{i,j=1}^{n} \frac{\partial^2 a_{ij}}{\partial x_i \partial x_j} \end{cases} \end{cases} \tag{12}$$

Equation (12) can be considered to be a rigorous version of equation (5).

3. Schrodinger Operators, Diffusion Operators and Time Reversibility

Under suitable hypotheses (e.g. uniform ellipicity, growth conditions on g^y, V^y bounded above) we can express the solution of (12) as a Feynman-Kac integral

$$q(t,x) = E_x[p^o(x_t)\exp(L_t)\exp(\int_0^t V^y(t,x_s)ds)] \tag{13}$$

where

$$L_t = \int_0^t a^{-1}(x_s)g^y(x_s,s)dw_s - \frac{1}{2}\int_0^t |a^{-1}(x_s)g^y(x_s,s)|^2 ds,$$

and where E_x denotes expectation with respect to the path space of ξ, and ξ satisfies

$$\begin{cases} d\xi_t = \sigma(\xi_t)dw_t \\ \xi_0 = x \end{cases} \tag{14}$$

We may ask whether the functional integration (13) can be reduced to quadratures. This leads us to consider the relation between Schrodinger Operators and Diffusion Operators or what is equivalent, the relation between the Feynman-Kac formula and the Girsanov Formula (cf. Simon 1979; Mitter 1980).

Let us suppose that $V: \mathbb{R}^n \to \mathbb{R}$, be measurable, bounded below and tends to $+\infty$ as $|x| \to \infty$ and consider the Schrodinger operator $H = -\Delta + V$ where Δ is the n-dimensional Laplacian. The H defines a self-adjoint operator on $L^2(\mathbb{R}^n; dx)$ which is bounded below and the lower bound λ of the spectrum of H is an eigenvalue of H. Let $\psi(x)$ be the corresponding eigenfunction of H, the so-called ground state and assume $\psi(x) > 0$. We normalize $\psi(x)$ i.e. $\int_{\mathbb{R}^n} |\psi(x)|^2 dx = 1$. Define the probability measure $d\mu = |\psi(x)|^2 dx$, and the unitary operator

$$U : L^2(\mathbb{R}^n; dx) \to L^2(\mathbb{R}^n; d\mu(x))$$

$$: f \to \psi^{-1} f \ .$$

If we define the Dirichlet form for f, g $\in C_c^\infty(\mathbb{R}^n)$

$$\delta(f,g) = \frac{1}{2} \int_{\mathbb{R}^n} \nabla f(x) \cdot \nabla g(x)\,dx \tag{15}$$

then a calculation shows

$$\delta(f,g) = (\mathscr{L}f,g)_\mu \tag{16}$$

where $(\ ,\ )_\mu$ denotes the scalar product in $L^2(\mathbb{R}^n; d\mu)$ and $\mathscr{L}$ is the diffusion operator (self-adjoint, positive)

$$\mathscr{L}\psi = -\frac{1}{2}\Delta\psi + \nabla b \cdot \nabla\psi \tag{17}$$

$$b = -\log \psi$$

Now since ψ satisfies

$$-\frac{1}{2}\Delta\psi(x) + V(x)\psi(x) = 0 \qquad \text{(assuming } \lambda=0\text{)} \tag{18}$$

we get

$$V(x) = \frac{1}{2}(|\nabla b(x)|^2 - \Delta b(x)), \tag{18b}$$

where differentials have to be interpreted in the sense of distributions.

Let ξ_t satisfy the stochastic differential equation

$$\begin{cases} d\xi_t = -\nabla b(\xi_t)\,dt + dw_t \\ \xi_o = x, \quad \text{where } w_t \text{ is standard Brownian motion.} \end{cases} \tag{19}$$

Define

$$\begin{aligned} L_t &= \exp\left(-\int_0^t \nabla b(\xi_s) \cdot d\xi_s - \frac{1}{2}\int_0^t |\nabla b(\xi_s)|^2 ds\right) \\ &= \exp\left(-\int_0^t \nabla b(\xi_s) \cdot dw_s + \frac{1}{2}\int_0^t |\nabla b(\xi_s)|^2 ds\right) \end{aligned} \tag{20}$$

which can be shown to be a $\mathcal{B}_t^{(1)}$ - martingale. Then if μ_ω denotes Wiener measure and if denote a new probability measure μ_ξ on the path space of ξ by

$$\frac{d\mu_\xi}{d\mu_\omega} = L_t \ ,$$

then from the Girsanov theorem ξ_t is Brownian motion under the measure μ_ξ and hence we can write the solution of

$$\begin{cases} \dfrac{d\rho}{ds} + \mathscr{L}\rho = 0, \quad 0 \leq s \leq t \\ \rho(t,x) = \psi(x) \end{cases} \tag{21}$$

as $\rho(s,x) = \tilde{E}_{sx}[\psi(\xi_t)]$ where $\tilde{E}$ denotes integration with respect to μ_ξ.

On the other hand, by the Generalized Ito-Differential Rule

$$db(\xi_t) = \nabla b(\xi_t)d\xi_t + \frac{1}{2}\Delta b(\xi_t)dt$$

and hence (20) reduces to

$$L_t = \exp(-b(\xi_t) + b(\xi_0) - \frac{1}{2}\int_0^t |\nabla b(\xi_s|^2 ds + \frac{1}{2}\int_0^t \Delta b(\xi_s)ds).$$

$$= \psi(\xi_0)^{-1}\psi(\xi_t)\exp(-\int_0^t V(\xi_s)ds)$$

and therefore

$$\psi(x) = E_x[\psi(\xi_t)\exp(-\int_0^t V(\xi_s)ds] \tag{22}$$

where E denotes expectation with respect to Wiener-measure and we have derived the Feynman-Kac Formula.

---

(1) $\beta_t = \sigma\{w_s | 0 \leq s \leq t\}$

Equation (19) denotes a stationary, reversible ($\xi(t)$ and $\xi(-t)$ are stochastically equivalent) Markov process with invariant measure μ. Thus with operators $H = -\Delta + V$ with V satisfying the hypotheses given above, we have a unique stationary, reversible Markov process intrinsically attached to it.

These ideas have a bearing on non-linear filtering (Benes 1981). Consider the scalar non-linear filtering problem

$$\begin{aligned} dx_t &= -b(x_t)\,dt + dw_t \\ dy_t &= x_t\,dt + d\eta_t \end{aligned} \tag{23}$$

and assume that

$$b(x_t) = f_x(x_t)$$

and f satisfies

$$\frac{1}{2}(|f_x|^2 - f_{xx}) = \frac{1}{2}x^2 \tag{24}$$

The D-M-Z equation for this problem is

$$d\rho(t,x) = (\mathscr{L}^* - \frac{1}{2}x^2)\rho(t,x)\,dt + x\rho(t,x)\cdot dy_t \tag{25}$$

We may write its solution as (using previous considerations)

$$\rho(t,x) = E_x[\exp(-f(x_t) + f(x_0) - \frac{1}{2}\int_0^t x_s^2 ds)\,. \tag{26}$$

$$\exp(\int_0^t x_s dy_s - \frac{1}{2}\int_0^t x_s^2 ds)]\,.$$

Because of the quadratic nature of the potential this function space integration can essentially be reduced to Gaussian integrals (cf. Mitter 1983 for an illuminating discussion). Indeed the filter for this problem is essentially a Kalman Filter.

4. Variational Interpretation of Non-Linear Filtering

We now give a stochastic variational interpretation of non-linear filtering in the spirit of the work of Feynman (Feynman-Hibbs 1965). We do this by associating a stochastic control problem with the D-M-Z equation. This section provides a justification of the ideas of Kalman on the duality between filtering and control. The original ideas of this section are due to Fleming-Mitter (cf. Fleming-Mitter (1982); Pardoux (1981) and Bensoussan (1982)). We follow the exposition of Pardoux and for simplicity consider the scalar case.

Let us formally denote the differential dy_s as $\dot{y}_s ds$ and consider the D-M-Z equation

$$\begin{cases} \dfrac{\partial \rho}{\partial t} = (L^* - \frac{1}{2})\rho(t,x) + h(x)\rho(t,x)y_s \\ \rho(0,x) = \rho_0(x) \end{cases} \tag{27}$$

Now $\rho(t,x)$ admits the factorization

$$\rho(t,x) = \ell(t,x)p(t,x)$$

where $p(t,x)$ is the density of the x-process and $\ell(t,x)$ is the likelihood function given by

$$\ell(t,x) = E(\exp(-\int_0^t h(x_s)\dot{y}_s ds - \frac{1}{2}\int_0^t |h(x_s)|^2 ds)|x_t = x) \tag{28}$$

Then a calculation shows that ℓ satisfies the equation

$$\begin{cases} \dfrac{\partial \ell}{\partial t} = (\tilde{L} - \frac{1}{2}h^2)\ell + h\dot{y}_s \ell \\ \ell(0) = 1 \end{cases} \tag{29}$$

where

$$\tilde{L} = \frac{1}{2}\, a(x) \frac{\partial^2}{\partial x^2} + \tilde{b} \frac{\partial}{\partial x} \qquad \text{with}$$

$$\tilde{b} = \frac{1}{p} \frac{\partial}{\partial x} (a(x)p) - b.$$

Here $\tilde{L}$ is the infinitesimal generator of the time-reversed x-process with $x_t = x$.

Now consider the transformation

$$\tilde{L}(t,x) = \exp(-S(t,x)).$$

Then S satisfies the Bellman equation

$$\begin{cases} \dfrac{\partial S}{\partial t} = \tilde{L}S - \frac{1}{2}\, a \left(\dfrac{\partial S}{\partial x}\right)^2 + \frac{1}{2}\,|h-\dot{y}_s|^2 - \dot{y}_s^2 \\ S(0) = 0\ . \end{cases} \tag{30}$$

Denote by $\tilde{x}$ the reverse x-Markov process conditioned on $x_t = x$. Then (30) corresponds to the following stochastic control problem:

$$\begin{cases} d\tilde{x}_s + (\tilde{b}(\tilde{x}_s) + u_s)ds + \sigma(\tilde{x}_s)\overset{(1)}{\oplus} d\tilde{w}_s = 0, \quad s \leq t \\ \tilde{x}_t = x \end{cases} \tag{31}$$

where the control u_s is to be chosen as a Markovian feedback control to minimize the cost function

---

(1) $\oplus$ denotes the backward Ito-differential.

$$E\left[\int_0^t \left(a^{-1}(\tilde{x}_s)|u_s|^2 + \frac{1}{2}|h(\tilde{x}_s) - \dot{y}_s|^2 - \dot{y}_s^2\right)ds\right], \tag{32}$$

and S(t,x) is the optimal value function of this stochastic control problem.

In the situation that the dynamics are linear and the observation map is linear we have a linear filtering problem and the stochastic control problem (31)-(32) corresponds to a linear-quadratic-gaussian problem with full observations. But the theory of this problem is essentially the same as the linear-quadratic deterministic optimal control problem. This explains in a clear manner the duality principle first enunciated by Kalman.

This stochastic variational interpretation can be effectively used to construct maximum a-posteriori density filters and maximum likelihood filters and allow us to give a derivation of the Extended-Kalman filter (Mitter 1983).

5. Geometric Theory of Nonlinear Filtering

In the introduction we have suggested that the fact that the D-M-Z equation is an infinite-dimensional bilinear equation allows us to develop a geometrical theory of non-linear filtering. This geometrical theory, originally independently suggested by Brockett and Clark (1980), Brockett (1980), Mitter (1980; (1) & (2)) was motivated by the desire to measure the complexity of nonlinear filters and to discover whether finite-dimensional filters existed for non-linear problems. The present exposition follows Mitter (Mitter(2) 1983).

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

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

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

$$\hat{\phi}_t = \gamma(\xi_t),$$

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 1977).

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

5.1 Lie Algebra of Operators Associated with the Filtering Problem

Consider the Lie algebra generated by the unbounded operators

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

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 $\mathscr{L}A\{L, h_1, \ldots, h_p\}$ is infinite-dimensional does not preclude the filtering problem having a finite-dimensional solution. Secondly, even if $\mathscr{L}A\{L, h_1, \ldots, h_p\}$ 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 $\mathcal{L}A\{L, h_1, \ldots h_p\}$ 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)

$$\begin{cases} dx_t = Ax_t dt + bdw_t & A = n \times n \text{ matrix} \\ & b = n \times 1 \text{ matrix} \\ dy_t = c'x_t dt + d_t & c = n \times 1 \text{ matrix} \end{cases} \tag{33}$$

Then

$$\begin{cases} \mathcal{L}^* = \frac{1}{2} \sum_{i,j=1}^{n} \frac{\partial^2}{\partial x_i \partial x_k} Q_{ij} - \sum_{i=1}^{n} \frac{\partial}{\partial x_i} (Ax)_i \text{ , and} \\ L = \mathcal{L}^* - \frac{1}{2} (c'x)^2 \text{ , where} \\ Q = bb' . \end{cases} \tag{34}$$

Define the Hamiltonian matrix

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

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

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

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

Then one can show that

$$Z_i = \sum_{j=1}^{n} (E^{i-1}\alpha)_j \, x_j + \sum_{j=1}^{n} (E^{i-1}\alpha)_{j+n} \frac{\partial}{\partial x_j}\ , \text{ and} \tag{35}$$

$$[Z_i, Z_j] = (E^{i-1}\alpha)' \begin{pmatrix} 0 & -I \\ I & 0 \end{pmatrix} (E^{j-1}\alpha), \tag{36}$$

$\mathscr{F} = \text{span}\,\{L, Z_1, \ldots, Z_{2n}, I\}$, 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 (see the Appendix).

5.2. 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:

$$\begin{cases} x_t = w_t \\ dy_t = x_t dt + d\eta_t \end{cases} \tag{37}$$

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

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

$L = \frac{1}{2}\frac{d^2}{dx^2} - \frac{1}{2}x^2$ 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 (loc.at).

$$\begin{cases} dx_t = f(x_t)dt + dw_t \\ dy_t = x_t dt + d\eta_t\ , \text{ where} \end{cases} \tag{38}$$

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 (38) 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 (38)

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

Now consider the commutative diagram:

$$\begin{tikzcd}
C_o^\infty(\mathbb{R}) \arrow[r, "\frac{d}{dx}"] \arrow[d, "\Psi"'] & C_o^\infty(\mathbb{R}) \arrow[d, "\Psi"] \\
C_o^\infty(\mathbb{R}) \arrow[r, "\frac{d}{dx} - f"'] & C_o^\infty(\mathbb{R})
\end{tikzcd}$$

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 $\mathscr{L}^* = \frac{1}{2}\frac{\partial^2}{\partial x^2} - \frac{\partial}{\partial x} f$ 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 transformation

$\rho(t,x) \to \Psi^{-1}(x)\rho(t,x)$, 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 (Ocone, 1980).

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

5.3 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

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

A proof of this can be constructed by performing calculations similar to that in Avez-Heslot (Avez-Heslot 1979).

5.4 Example with Pro-finite-dimensional Lie Algebra (cf. Hazewinkel-Marcus 1982).

Consider the filtering problem:

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

It can be 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 $x\frac{\partial^i}{\partial\xi^i}$, $\frac{\partial}{\partial}\frac{\partial^i}{\partial\xi^i}$, $\frac{\partial^i}{\partial\xi_i}$, $i = 0,1, \ldots$

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 5.2 and the example of section 5.4. Thus, in example 5.4 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 can be shown that all conditional moments of ξ_t can be computed using finite-dimensional recursive filters (Ocone-Baras-Marcus 1982).

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.

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

Earlier we have given the definition of a finite-dimensional filter. We would consider this definition to be the description of a control system with inputs y_t and output $\hat{\phi}_t$. Furthermore, as we have said we may assume that this representation is minimal in the sense of Sussmann. We thus have two ways of computing $\hat{\phi}_t$--one via the D-M-Z equation and the other via the control system. 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 anti-homomorphism between the Filter algebra $\mathcal{F}$ and $L(\Sigma)$ such that $L \to \alpha$ and $h_i \to \beta_i$ where β_i is the i^{th} row of β.

Conversely, suppose that the Lie algebra $\mathcal{F}$ 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 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 5.4, that is, the Lie algebra $\mathcal{F}$ is pro-finite dimensional. Since $\mathcal{F}/\mathcal{I}_i$ 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 $\mathcal{F}$ or any of its quotients) need not be finite-dimensional for a finite-dimensional filter to exist.

5.6 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 1980.

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

Suppose that we want to solve the evolution equation

$$\frac{d\rho}{dt} = L_1\rho + u(t)L_2\rho , \text{ where} \tag{39}$$

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

$$\rho(t) = \exp\left(g_1(t)L_1\right)\exp\left(g_2(t)L_2\right) \ldots \exp\left(g_d(t)L_d\right)\rho(o) \tag{40}$$

where $\rho(o)$ is the initial condition. For ideas similar to this in the context of ordinary stochastic differential equations, see (Kunita 1981).

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

$$\frac{d\rho}{dt} = F_1\left(g(t),\dot{g}(t)\right)L_1\rho + \ldots + F_d\left(g(t),\dot{g}(t)\right)L_d\rho \tag{41}$$

for some non-linear functions F_i of $g(t) = \left(g_1(t), \ \ldots \ , \ g_d(t)\right)$ and $\dot{g}(t)$.

For (41) to define a solution of (39), we need

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

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 (40) and using

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

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

one gets

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

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 (40) 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.

5.7 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 $W_1/\mathbb{R}$. Now Hazewinkel and Marcus (loc.cit) 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 (Sussmann 1981) 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.

5.8 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

$$dx_t = ax_t dt + dw_t$$

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

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

where ε is some small positive answer.

For this class of problems it has been shown (Sussmann 1982) 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 Hazelwinkel-Krishnaprasad-Marcus (1983).

Marcus (1983) (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 (40) 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.

5.9 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 $W_1/\mathbb{R}$. Now Hazewinkel and Marcus (loc.at.) 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 1981) 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.

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:

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

(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

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

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

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

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

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 $X \varepsilon \mathscr{L}$ let adX denote the endomorphism of $\mathscr{L}$

$adX : Y \to [X,Y] \quad (Y \varepsilon \mathscr{L})$.

adX is a derivation of $\mathscr{L}$ and $X \to adX$ 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 $(x,y) \to xy : GxG \to G$ 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 $\mathscr{L}$, $\forall X \varepsilon \mathscr{L}$. Let the dimension of $\mathscr{L}$ be m. Then there are ideals $\mathscr{I}_j$ of $\mathscr{L}$ such that (i) dim $\mathscr{I}_j = m-j$, $0 \leq j \leq m$.

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

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 $\mathscr{D}^p g$ $(p \geq o)$ inductively by

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

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

We then get a sequence $\mathcal{D}^0 g \supseteq \mathcal{D}^1 g \supseteq \ldots$ of subalgebras of g. g is said to be solvable if $\mathcal{D}^p g = 0$ 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 $\mathcal{V}$. Define a Lie algebra structure on $\mathcal{V}$ by $[p_i, q_i] = [q_i, p_i] = Z$, the other brackets being zero. This nilpotent Lie algebra $\mathcal{N}$ 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 $\mathcal{N}$.

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

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 $\mathcal{D}$. 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

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

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.

BIBLIOGRAPHY

1. Avez A. and Heslot A. 1979 L'algèbre de lie des polynômes en les coordonnées canoniques munie du crochet de poisson, C.R. Acad. Sci., Paris A, T-288, pp. 831-833.

2. V. Benes 1981. Exact finite dimensional filters for certain diffusions with non-linear drift, Stochastics 5, pp. 65-92.

3. A. Bensoussan 1982. On the equation of nonlinear filtering, unpublished manuscript.

4. R.W. Brockett 1980. Remarks on finite dimensional nonlinear estimation in Analyse des Systemes. Asterisque, 75-76, pp. 199-205.

5. Brockett R.W. and Clark J.M.C. 1980. 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.

6. Clark J.M.C. 1978. The design of robust approximations to the stochastic differential equations of nonlinear filtering, in Communication Systems and Random Process Theory, J.K. Skwirzynski, ed., Sithoff and Noordhoff.

7. Davis M.H.A. 1980. On a multiplicative functional transformation arising in nonlinear filtering theory, Z. Wahr. verw Gebiete 54, pp. 125-139.

8. Feynman R.P. and Hibbs A. 1965. Quantum Mechanics and Path Integrals, McGraw Hill, New York.

9. Fleming W.H. and Mitter S.K. 1982. Optimal control and nonlinear filtering for nondegenerate diffusion processes, Stochastics, 8, pp. 63-78.

10. Fleming W.H. and Pardoux E. 1982. Existence of optimal controls for

partially-observed diffusions, SIAM J. on Control, 20, pp. 261-285.

11. Fujisaki M., Gallianpur G. and Kunita H. 1972. Stochastic differential equations for the nonlinear filtering problem, Osaka J. of Math., Vol. 1, pp. 19-40.

12. Glimm J. and Jaffe A. 1981. Quantum Physics, Springer-Verlag, Berlin-New York.

13. Hazewinkel M. and Marcus S.I. 1982. On lie algebras and finite dimensional filtering, Stochastics, 7, pp. 26-62.

14. Hazewinkel M., Krishnaprasad P.A., and Marcus S.I. 1982. Current algebras and the identification problem, Technical Report, University of Maryland, Department of Electrical Engineering.

15. Kunita H. 1981. On the decomposition of solutions of stochastic differential equations, in Stochastic Integrals, Springer Lecture Notes in Mathematics 851, Springer-Verlag, Berlin.

16. Marcus S.I., Mitter, S.K., and Ocone D. 1980. Finite dimensional nonlinear estimation for a class of systems in continuous and discrete time, in Analysis and Optimization of Stochastic Systems, O.L.R. Jacobs et al, eds., Academic Press, New York, pp. 387-406.

17. Mitter S.K. 1980. Filtering theory and quantum fields, in Analyse des Systemes, Asterisque, Vol. 75-76, pp. 199-205.

18. Mitter S.K. 1981. Non-linear filtering and stochastic mechanics in Stochastic Systems: The Mathematics of Filtering and Identification and Applications, eds. M. Hazewinkel and J.C. Willems, D. Reidel Publishing Co., pp. 479-503.

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

20. Mitter S.K. 1983. Approximations to nonlinear filtering, in Nonlinear Stochastic Systems, eds. R.S. Bucy and J. Moura, Reidel, Dordrecht, 1983.

21. Mitter S.K. 1983. Geometric theory of nonlinear filtering, in Methodes et Outils de Mathematiques en Theorie des Systems.

22. Pardoux E. 1981. The solution of the nonlinear filtering equation as a likelihood function, in Proceedings of the 20th IEEE Conference on Decision and Control, San Diego, California.

23. Ocone D. 1980. Topics in nonlinear filtering theory, Doctoral Dissertation, MIT, June 1980.

24. Ocone D., Baras J.S. and Marcus S.I. 1982. Explicit filters for diffusions with certain nonlinear drifts, Stochastics, 8, pp. 1-16.

25. Sussmann H.J. 1977. Existence and uniqueness of minimal realizations of nonlinear systems, Meth. Syst. Theory, 10, pp. 263-284.

26. Sussmann H.J. 1981. 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.

27. Sussmann H.J. 1982. Approximate finite-dimensional filters for some nonlinear problems, Stochastics, 7, pp. 183-204.