Nonexistence of finite dimensional filters for conditional statistics of the cubic sensor problem

by

M. Hazewinkel[*)], S.I. Marcus[**)] & H.J. Sussmann[***)]

ABSTRACT

Consider the cubic sensor $dx = dw$, $dy = x^3dt + dv$ where w, v are two independent brownian motions. Given a function $\phi(x)$ of the state x let $\hat{\phi}_t(x)$ denote the conditional expection given the observations y_s, $0 \leq s \leq t$. This paper consists of a rather detailed discussion and outline of proof of the theorem that for nonconstant ϕ there can not exist a recursive finite dimensional filter for $\hat{\phi}$ driven by the observations.

CONTENTS

KEY WORDS & PHRASES: *cubic sensor, recursive filter, robust filtering, Weyl Lie algebra*

---

*) Mathematical Centre, P.O.Box 4079, 1009 AB Amsterdam.

**) Dept. Electrical Engr. Univ. of Texas at Austin, Texas. Supported in past by the National Science Foundation under Grant ECS-8022033 and in past by the Joint Services Electronics Program under Contract F49620-77-c-0101.

***) Dept. Math., Rutgers Univ. New Brunswick, New Jersey.

1. INTRODUCTION

The cubic sensor problem is the problem of determining conditional statistics of the state of a one dimensional stochastic process $\{x_t: t \geq 0\}$ satisfying

$$(1.1) \qquad dx = dw, \quad x_0 = x^{in}$$

with w a Wiener process, independent of x^{in}, given the observation process $\{y_t: t \geq 0\}$ satisfying

$$(1.2) \qquad dy = x^3 dt + dv, \quad y_0 = 0$$

where v is another Wiener process independent of w and x^{in}. Given a smooth function $\phi: \mathbb{R} \to \mathbb{R}$ let $\hat{\phi}_t$ denote the conditional expection

$$(1.3) \qquad \hat{\phi}_t = \phi(x_t) = E[\phi(x_t) \mid y_s, 0 \leq s \leq t]$$

By definition a smooth finite dimensional recursive filter for ϕ_t is a dynamical system on a smooth finite dimensional manifold M governed by an equation

$$(1.4) \qquad dz = \alpha(z)dt + \beta(z)dy, \quad z_0 = z^{in}$$

driven by the observation process, together with an output map

$$(1.5) \qquad \gamma: M \to \mathbb{R}$$

such that, if z_t denotes the solution of (1.4),

$$(1.6) \qquad \gamma(z_t) = \hat{\phi}_t \quad \text{a.s.}$$

Roughly speaking one now has the theorem that for nonconstant ϕ such filters cannot exist. For a more precise statement of the theorem see 2.10 below.

It is the purpox of this note to give a fairly detailed outline of the proof of this theorem and to discuss the structure of the proof. That is the general principles underlying it. The full precise details of the analytic and realization theoretic parts of the proof will appear in [Sussmann 1983a, 1983b], the details of the algebraic part of the proof can be found in [Hazewinkel - Marcus, 1982]. An alternative much better and shorter proof of the hardest bit of the algebraic part will appear in [Stafford, 1983].

2. SYSTEM THEORETIC PART. I: PRECISE FORMULATION OF THE THEOREM

2.1 The setting

The precise system theoretic - probabilistic setting which we shall use for the cubic sensor filtering problem is as follows

(i) (Ω, A, P) is a probability space

(ii) $(A_t: 0 \leq t)$ is an increasing family of σ-algebras

(iii) (w,v) is a two-dimensional standard Wiener process adapted to the A_t.

(iv) $x = \{x_t: t \geq 0\}$ is a process which satisfies $dx = dw$, i.e.

(2.1) $$x_t = x_0 + w_t \qquad \text{a.s. for each } t$$

(v) x_0 is A_0-measurable and has a finite fourth moment

(vi) $\{y_t: t \geq 0\}$ is a process which satisfies $dy = x^3 dt + dv$, i.e.

(2.2) $$y_t = \int_0^t x_s^3 ds + v_t \qquad \text{a.s. for each } t$$

(vii) the processes v, w, x, y all have continuous sample paths, so that in particular (2.1) and (2.2) actually hold and not just almost surely. (More precisely one can always find if necessary modified versions of v, w, x, y such that (vii) (also) holds).

2.3. The filtering problem

Let y_t, $t \geq 0$ be the σ-algebra generated by the y_s, $0 \leq s \leq t$ and let $\phi: \mathbb{R} \to \mathbb{R}$ be a Borel measurable function. Then the filtering problem (for this particular ϕ) consists of determining $E[\phi(x_t) \mid Y_t]$.

2.4. Smooth finite dimensional filters

Consider a (Fisk-Stratonovič) stochastic differential equation

$$(2.5)\qquad dz = \alpha(z)dt + \beta(z)dy, \quad z \in M,$$

where M is a *finite dimensional* smooth manifold and α and β are smooth vectorfields on M. Let there also be given an initial state and a smooth output map

$$(2.6)\qquad z^{in} \in M, \quad \gamma: M \to \mathbb{R}.$$

The equation (2.5) together with the initial condition $z(0) = z^{in}$ has a solution $z = \{z_t: t \geq 0\}$ defined up to a stopping time T, which satisfies

$$(2.7)\qquad 0 < T \leq \infty \text{ a.s.}, \quad \{\omega \mid T(\omega) > t\} \in Y_t, \text{ for } t \geq 0.$$

Moreover there is a unique maximal solution, i.e. one for which the stopping time T is a.s. $\geq T_1$ if T_1 is the stopping time of an arbitrary other solution z_1. In the following $z = \{z_t: t \geq 0\}$ denotes such a maximal solution.

The system given by (2.5), (2.6) is now said to be a smooth finite dimensional filter for the cubic sensor (2.1) (i) - (vii) if for y equal to the observation process (2.2) the solution z of (2.5) satisfies

$$(2.8)\qquad E[\phi(x_t)|Y_t] = \gamma(z_t) \text{ a.s. on } \{\omega \mid T(\omega) > t\}.$$

2.9. Statement of the theorem

With these notions the main theorem of this note can be stated as:

2.10. THEOREM. *Consider the cubic sensor 2.1. (i) - (vii); i.e. assume that these conditions hold. Let* $\phi: \mathbb{R} \to \mathbb{R}$ *be a Borel measurable function which satisfies for some* $\beta \geq 0$ *and* $0 \leq r < 4$

$$(2.11)\qquad |\phi(x)| \leq \exp(\beta|x|^r), \quad -\infty < x < \infty.$$

Assume that ϕ is not almost everywhere equal to a constant. Then there exists no smooth finite dimensional filter for the conditional statistic $E[\phi(x_t) \mid Y_t]$.

3. SYSTEM THEORETIC PART. II: THE HOMOMORPHISM PRINCIPLE AND OUTLINE OF THE PROOF (HEURISTICS)

3.1. The Duncan-Mortensen-Zakai equation

Consider a nonlinear stochastic dynamical system

$$(3.2) \qquad dx_t = f(x_t)dt + G(x_t)dw_t, \quad x^{in} = x_0, \quad x_t \in \mathbb{R}^n, \quad w_t \in \mathbb{R}^m$$

where w_t is a standard Brownian motion independent of the initial random variable x^{in} and where f and G are appropriate vector valued and matrix valued functions. Let the observations be given by

$$(3.3) \qquad dy_t = h(x_t)dt + dv_t, \quad y_t \in \mathbb{R}^p$$

where v_t is another standard Brownian motion independent of w and x^{in}. Let $\hat{x}_t$ denote the conditional expectation

$$(3.4) \qquad \hat{x}_t = E[x_t \mid Y_t] = E[x_t \mid y_s, \ 0 \le s \le t]$$

where Y_t is the σ-algebra generated by the y_s, $0 \le s \le t$. Let p(x,t) be the density of $\hat{x}_t$ where it is assumed (for the purposes of this heuristic section) that p(x,t) exists and is sufficiently smooth as a function of x and t. Then an unnormalized version $\rho(x,t)$ satisfies the Duncan-Mortensen-Zakai equation

$$(3.5) \qquad d\rho(x,t) = (\frac{1}{2} \sum_{i,j} \frac{\partial^2}{\partial x_i \partial_j} ((GG^T)_{ij} - \sum_i \frac{\partial}{\partial x_i} f_i - \frac{1}{2} \sum h_j^2)\rho(x,t)dt$$

$$+ \sum h_j \rho(x,t)dy_{jt}, \quad \rho(x,0) = \text{density of } x^{in}$$

where $h_j = h_j(x)$ is the j-th component of h, $(GG^T)_{i,j}$ is the (i,j)-th entry of the product of the matrix G(x) with its transpose and $f_i = f_i(x)$ is the i-th component of f(x). The equation (3.5) is a stochastic partial differential equation in Fisk-Stratonovič form. In the case of the cubic sensor (2.1), (2.2) (or (1.1), (1.2)) the equation becomes

$$(3.6) \qquad d\rho(x,t) = (\frac{1}{2}\frac{d^2}{dx^2} - \frac{1}{2}x^6)\rho(x,t)dt + x^3\rho(x,t)dy$$

3.7. <u>The homomorphism principle</u>

Now assume for a given $\phi\colon \mathbb{R}^n \to \mathbb{R}$ we have a smooth finite dimensional filter

$$(3.8) \qquad dz = \alpha(z)dt + \sum_j \beta_j(z)dy_j, \quad z_0 = z^{in}, \quad \gamma\colon \mathbb{R}^n \to \mathbb{R}$$

to calculate the statistic $\hat{\phi}_t = E[\phi(x_t) \mid Y_t]$. I.e. $\hat{\phi}_t = \gamma(z_t)$ a.s. if z_t is the solution of (3.8). The equation (3.8) is to be interpreted in the Statonovič sense.

Then, very roughly, we have two ways to process an observation path $y^\omega\colon s \to y_s(\omega)$, $0 \le s \le t$ to give the same result. One way is by means of the filter (3.8), the other way is by means of the infinite dimensional system (3.5) (defined on a suitable space of functions) coupled with the output map

$$(3.9) \qquad \Phi\colon \psi \to (\int\psi(x)dx)^{-1}\int\psi(x)\phi(x)dx$$

Assuming that (3.8) is observable, deterministic realization theory [Sussmann 1977] then suggest that there exists a smooth map F from the reachable part (from $\rho(x,0)$) of (3.6) to the reachable part of (3.8), which takes the vectorfields of (3.6) to the vectorfields of (3.6) and which is compatible with the output maps γ and (3.9). The operators in (3.6) define linear vectorfields in the state space of (3.6) (a space of functions). Let $L_0, L_1, \ldots, L_p$ be the operators occuring in (3.5) so that $d\rho = L_0\rho dt + L_1\rho dy_1 + \ldots + L_p\rho dy_p$. The Lie algebra of differential operators generated by $L_0, \ldots, L_p$ is called the *estimation Lie algebra*, and is denoted $L(\Sigma)$. The idea of studying this Lie algebra to find out things about filtering problems is apparently due to both Brockett and Mitter, cf e.g. [Brockett 1981] and Mitter [1981] and the references in these two papers.

Let $L \mapsto \tilde{L}$ be the map which assigns to an operator the corresponding linear vectorfield (analogous to the map which assigns to an $n \times n$ matrix $A = (a_{ij})$ the linear vector field $\Sigma a_{ij} x_i \frac{\partial}{\partial x_j}$ as $\mathbb{R}^n$). Then $L \mapsto -\tilde{L}$ is a homomorphism of Lie algebras. Further F induces a homomorphism of Lie algebras dF: $\tilde{L}_0 \to \alpha$, $\tilde{L}_i \to \beta_i$, $i = 1, \ldots, p$. Thus the existence of a finite dimensional filter should imply the existence of a homomorphism of Lie algebras $L(\Sigma) \to V(M)$ where $V(M)$ is the Lie algebra of smooth vectorfields on a smooth finite dimensional manifold M. This principle, originally enunciated by Brockett, has come to be called the homomorphism principle.

3.10. Pathwise filtering (robustness)

As it stands to remarks in 3.7 above are quite far from a proof of the homomorphism principle. First of all (3.6) and (3.8) are stochastic differential equations and as such they have solutions defined only almost everywhere. The first thing to do to remedy this situation is to show that these equations make sense and have solutions pathwise so that they can be interpreted as processing devices which accept an observation path $y: [0,t] \to \mathbb{R}^p$ and produce outputs $\hat{\phi}_t(y)$ as a result. Another reason for looking for pathwise robust versions which is most important for actual applications, lies in the observation that actual physical observation paths will be piece-wise differentiable and that the space of all such paths is of measure zero in the probability space of paths underlying (3.6) and (3.8). Cf. [Clark, 1978].

Another difficulty in using the remarks of 3.7 above to establish a general homomorphism principle lies in the fact that (3.6) evolves on an infinite dimensional state space. A different approach to the establishing of homomorphism principles (than the one used in this paper) is described in [Hijab, 1982].

3.11. On the proof of theorem 2.10

In this paper the following route is followed to establish the homomorphism principle for the case of the cubic sensor. First for suitable $\phi: \mathbb{R} \to \mathbb{R}$ it is established that there exists a robust pathwise version of the functional $\hat{\phi}_t$. More precisely if C_t is the space of continuous functions $[0,t] \to \mathbb{R}$ then it is shown that there exists a functional $\Delta_t^\phi: C_t \to \mathbb{R}$ such that (proposition 4.15)

$$(3.12) \qquad \hat{\phi}_t = \frac{\Delta_t^\phi(y)}{\Delta_t^1(y)} \qquad \text{a.s. if } y = y^\omega.$$

The next step is to show that $\Delta_t^\phi(y)$, $y \in C_t$ is given by a density $n_t(y)(x)$ so that $\Delta_t^\phi(y) = \int n_t(y)(x)\,\phi(x)dx$ and to show that $n_t(y)(x)$ is smooth (as a function of x).

The next step is to use that there exist (up to a stopping time) pathwise and robust solutions of stochastic differential equations like (3.8). Robustness of both (3.6) and (3.8) then gives the central equality (4.24) *anywhere* (not just a.s.), that is:

$$(3.13) \qquad \frac{\Delta_t^\phi(y)}{\Delta_t^1(y)} = \gamma(z_t(y)), \quad y \in C_t.$$

The next step is to prove results about the smoothness properties of the density $n_t(y)$ as a function of $t_1, \ldots, t_m$ for paths y such that $u = \dot{y}$ is of the bang-bang type: $u(s) = \bar{u}_m \in \mathbb{R}$ for $0 \le t < t_m$, equal to $\bar{u}_{m-1}$ for $t_m \le t < t_m + t_{m-1}$, etc. ... and to observe that $(t,x) \mapsto n_t(y)(x)$ satisfies the DMZ equation (3.6). This permits to write down and calculate the result of applying $\frac{\partial^m}{\partial t_1 \ldots \partial t_m}\big|_{t_1 = \ldots = t_m = 0}$ to both sides of (3.13) and gives a relation of the type

$$(3.14) \qquad (A(\bar{u}_m) \ldots A(\bar{u}_1)\gamma)(z) = \tilde{L}(\bar{u}_m) \ldots \tilde{L}(\bar{u}_1)\Phi(\psi_z)$$

where $A(\bar{u})$ is the vectorfield $\alpha + \bar{u}\beta$, $L(\bar{u})$ the operator $L_0 + \bar{u}L_1 = (\frac{1}{2}\frac{d^2}{dx^2} - \frac{1}{2}x^6) + \bar{u}x^3$ and $\tilde{L}(\bar{u})$ the linear vectorfield associated to $L(\bar{u})$ Φ the functional (3.9), and ψ_z a function corresponding to z, cf. 5.10.

A final realization theoretic argument having to do with reducing the filter dynamical system (3.8) to an equivalent observable and reachable system then establishes the homomorphism principle in the case of the cubic sensor and the fact that if the homomorphism is zero ϕ was a constant.

The remaining algebraic part of the proof consists of two parts:

(i) a calculation of $L(\bar{z})$ for the cubic sensor. It turns out that $L(\bar{z})$ is in this case equal to the Heisenberg-Weyl algebra W_1 of all differential operators (any order) in x with polynomial coefficients.

(ii) the theorem that if $V(M)$ is the Lie algebra of smooth vectorfields on a smooth finite dimensional manifold and $\alpha: W_1 \to V(M)$ a homomorphism of Lie algebras then $\alpha = 0$.

4. ANALYTIC PART

4.1. The space of functions E

Let E denote the space of all Borel functions $\phi: \mathbb{R} \to \mathbb{R}$ such that there exists constants $C \in \mathbb{R}$, $\alpha \in \mathbb{R}$ and r, $0 \leq r < 4$ such that

$$|\phi(x)| \leq C \exp(\alpha|x|^r) \quad \text{for all } x \in \mathbb{R}. \tag{4.2}$$

The space $E(\alpha,r)$ is the normed space of all Borel functions for which

$$\|\phi\|_{(\alpha,r)} = \sup\{|\phi(x)| \exp(-\alpha|x|^r) \; ; \; x \in \mathbb{R}\} \tag{4.3}$$

is finite. The space E is the union of the $E(\alpha,r)$ and is topologized as such, i.e. as the inductive limit of the $E(\alpha,r)$.

4.2. Some bounds

Let C_t be the space of continuous functions $[0,t] \to \mathbb{R}$ such that $y(0) = 0$ and C_t^1 the functions of class C^1 in C_t, and $C_t^{1\#}$ the functions $y \in C_t$ which are piecewise C^1. Let H be the space of functions on Ω

$$H = \bigcap_{1 \leq p < \infty} L^p(\Omega, A, P). \tag{4.5}$$

For the processes x and w of 2.1, β, $r \in \mathbb{R}$, and a given $y \in C_t$ we define

$$(4.6) \qquad U(\beta,r,t,y) = \exp(\beta|x_t|^r + y(t)x_t^3 - 3\int_0^t y(s)x_s^2 dw_s - 3\int_0^t y(s)x_s ds - \frac{1}{2}\int_0^t x_s^6 ds).$$

The reason for considering this expression becomes clearer below. The $U(0,0,t,y)$ occur in a slightly modified version of the Kallianpur-Stiebel formula for $\hat{\Phi}_t$. The very slight modification (an integration by parts to remove dy_s) sees to it that the formula makes pathwise sense for continuous sample paths $y_s(\omega)$. The number 4 of 4.1 above gets into the picture as a result of wanting to keep $E[\phi(x_t)U|0,0,t,y)]$ bounded for bounded y and t.

4.7. <u>Lemma</u>. $U(\beta,r,t,y) \in \mathcal{H}$ *if* $0 \le \beta$, $0 \le r < 4$, $0 < t$, $y \in C_t$.

This is proved by straightforward estimates (and the Ito formula) using each time $-\frac{1}{8}\int_0^t x_s^6 ds$ to keep the contributions of each of the other terms in check.

4.8. <u>A robust version for $\hat{\Phi}_t = E[\phi(x_t) \mid Y_t]$</u>

Now let B denote the space of bounded Borel functions $\phi: \mathbb{R} \to \mathbb{R}$ endowed with the sup norm. Define for $\phi \in B$

$$(4.9) \qquad \langle N_t(y),\phi\rangle = E[\phi(x_t)U(0,0,t,y)].$$

This is well defined because $U(0,0,t,y) \in L^1(\Omega,A,P)$ and $\phi(x_t)$ is bounded. One now shows using 4.7 that the finite positive measure $N_t(y)$ has a density $n_t(y)(x) \in L^1(\mathbb{R})$. Now define for all $\phi \in E$

$$(4.10) \qquad \langle N_t(y),\phi\rangle = \int_{-\infty}^{+\infty} \phi(x)n_t(y)(x)dx$$

(which agrees (4.9) for $\phi \in B$). The functional $N_t(y): E \to \mathbb{R}$ satisfies

$$(4.11) \qquad \|\langle N_t(y),\phi\rangle\| \le C\|\phi\|_{(\beta,r)} \quad \text{for } \phi \in \varepsilon(\beta,r)$$

with C independent of y, t for bounded $\|y\|_t$ and t in a compact subset of $(0,\infty)$.

4.12. Proposition. $N_t(y)$ *depends continuously on* y, *for each fixed* t, $\phi \in E$.

Now define $\Delta_t^\phi \colon C_t \to \mathbb{R}$ by

$$\Delta_t^\phi(y) = \langle N_t(y), \phi \rangle. \tag{4.13}$$

Then $\Delta_t^\phi(y)$ is continuous and one shows that $\Delta_t^1(y) > 0$ for all $t > 0$ so that we can define

$$\delta_t^\phi(y) = \frac{\Delta_t^\phi(y)}{\Delta_t^1(y)}. \tag{4.14}$$

4.15. Proposition. *If* $\phi \in E$, $\omega \mapsto \delta_t^\phi(y_t(\omega))$, *where* $y_t(\omega)$ *is the* ω*-path of the process* y *of (2.1) is a version of* $\hat{\phi}_t = E[\phi(x_t) \mid y_t]$ *for* $t > 0$. *(So* δ_t^ϕ *is a robust pathwise version of* $\hat{\phi}_t$.*)*

This is proved via the Kallianpur-Stiebel formula which says in our case that if

$$k_t(\phi) = \int_\Omega \phi(x_t(\omega'))\exp\Big(\int_0^t x_s^3(\omega')dy_s(\omega) - \frac{1}{2}\int_0^t x_s^6(\omega')ds\Big)dP(\omega')$$

then

$$\hat{\phi}_t = \frac{k_t(\phi)}{k_t(1)} .$$

Indeed rewrite the term $\int_0^t x_s^3(\omega')dy_s(\omega)$ by means of partial integration as

$$\int_0^t x_s^3(\omega')dy_s(\omega) = x_t^3(\omega')y_t(\omega) - \int_0^t y_s(\omega)d(x_s^3(\omega')) =$$

$$= x_t^3(\omega')y_t(\omega) - 3\int_0^t y_s(\omega)x_s^2(\omega')dw_s(\omega') -$$

$$- 3\int_0^t y_s(\omega)x_s(\omega)ds$$

and it readily follows that $\int U(0,0,t,y_t(\omega))\phi(x_t(\omega'))dP(\omega') = \Delta_t^\phi(y)$ is a version of $k_t(\phi)$.

4.16. Smoothness properties of $n_t(y)(x)$

Let F be the space of all C^∞- functions $\phi: \mathbb{R} \to \mathbb{R}$ for which $\exp(\beta|x|^r)|\phi^{(k)}(x)|$ is bounded for all $\beta \geq 0$, $0 \leq r < 4$, $k \in \mathbb{N} \cup \{0\}$ and give F the topology defined by the family of norms

(4.17) $$||\phi||_{F(\beta,r,k)} = \sup\{\exp(\beta|x|^r)\ |\phi^{(k)}(x)|\}.$$

4.18. Lemma. *If* $x_0 \in H$ *then* $n_t(y) \in F$ *for all* $t > 0$, $y \in C_t$ *and* $n_t(y)(x) > 0$ *for all* $x \in \mathbb{R}$, $t > 0$ *for* y *differentiable.*

This is approached by considering the derivatives of $N_t(y)$ defined by $\langle N_t(y)',\phi\rangle = \langle N_t(y),\phi'\rangle$ for smooth $\phi : \mathbb{R} \to \mathbb{R}$.

4.19. Robustness for the filter

Now consider a stochastic differential equation with output map and initial condition driven by the observation process y_t

(4.20) $$dz = \alpha(z)dt + \beta(z)dy_t,\quad z(0) = z^{in},\quad z \mapsto \gamma(z),\quad z \in M$$

as we would have for a filter for $\hat{\phi}$ cf. 2.4 above. Equation (4.20) is to be interpreted in the Stratonovič sense. Let T be the stopping time for a maximal solution. Then, as was shown in [Sussman, 1978] these equations admit robust solutions in the following sense.

Consider the equation for $y \in C_t$

(4.21) $$dz = \alpha(z)dt + \beta(z)dy,\quad z(0) = z^{in}.$$

A curve $z: \tau \to z(\tau)$, $0 \leq \tau \leq t$ is said to be a solution of (4.21) if there exists a neighborhood U of y in C_t with the property that there is a continuous map $U \to C([0,t],M)$ $\tilde{y} \to z(\tilde{y})$ to the space of continuous curves in M such that $z(\tilde{y})$ is a solution of (4.21) in the usual sense for all $\tilde{y} \in U \cap C_t^1$ (so that the equation can be written as a usual differential equation) and $z(y) = z$.

With this notion of solutions the robustness result is:

4.22. THEOREM [Sussman, 1978]. *(i) Given any continuous* $y: [0,\infty) \to \mathbb{R}$, $y(0) = 0$, *there exists a time* $T(y) > 0$ *such that there is a unique solution* $\tau \to z(y)(\tau)$ *of (4.21). If* $T(y) < \infty$ *then* $\{z(y)(t): 0 \le t < T(y)\}$ *is not relatively compact on* M.

(ii) If y *is a Wiener process with continuous sample paths defined on* (Ω,A,P) *and if* $y^{\omega}(t) = y_t(\omega)$, *then* $\omega \to T(y^{\omega})$ *is a version of the stopping time up to which the Stratonovič solution of (4.20) is defined and* $\omega \to z(y^{\omega})(t)$, $0 \le t < T(y^{\omega})$ *is a version of the solution* z_t *for each* $t > 0$.

In our setting y_t is not a Wiener process, but the same techniquess apply, and the same results hold.

In other words up to a stopping time, solutions of (4.20) exist pathwise, they are continuous as a function of the path and hence can be calculated as limits of solutions to the corresponding nonstochastic differential equations (4.21) for (piecewise) differential continuous y.

·23. Everywhere equality of the robust filter output and the robust DMZ output and consequences

Now let (4.20) be a smooth filter for Φ in the sense of section 2.4 ≥ and let $\phi \in E$. Choose the robust version of Φ_t, i.e. the map $\delta_t^{\phi}(y^{\omega})$ and choose the robust solution $\omega \to z(y^{\omega})(t)$ of (4.20). The fact (4.20) is a filter for Φ_t says by definition that $\delta_t^{\phi}(y^{\omega}) = \gamma(z(y^{\omega})(t))$ almost all ω such that $T(y^{\omega}) > t$. The robustness of the two versions readily implies that

$$\delta_t^{\phi}(y) = \gamma(z(y)(t)) \qquad)$$

s everywhere whenever $t > 0$, $y \in C_t$, $T(y) > t$.

4.25. Smoothness properties of the family of densities $n_t(y)$

When y is piecewise C^1 and the initial probability density ν^{in} is in F the study of the measure $N_t(y)$ is much easier. (By modifying the data (Ω,A,P), etc. it can actually be arranged that ν^{in} is in F essentially by replacing ν^{in} with the density at a slightly later time $\tau < t$). Now integration by parts gives

$$(4.26) \qquad U(0,0,t,y) = \exp\left(\int_0^t x_s^3 \dot{y}(s)ds - \tfrac{1}{2} \int_0^t x_s^6 ds\right)$$

Now let ϕ: $\mathbb{R} \to \mathbb{R}$ be of class C^2, then the differential of $\phi(x_t)U(0,0,t,y)$ can easily be computed to be

$$(4.27) \qquad [\phi'(x_t)dw_t + \tfrac{1}{2}\phi''(x_t)dt + (x_t^3\, \dot{y}(t) - \tfrac{1}{2}x_t^6)\phi(x_t)]U(0,0,t,y)$$

so that (if, say, ϕ has compact support), y_τ denoting the restriction of y to $[0,\tau]$

$$(4.28) \qquad E[\phi(x_t)U(0,0,t,y) - \phi(x_\tau)U(0,0,\tau,y_\tau)] =$$

$$\int_\tau^t \left(\frac{1}{2}\frac{d^2}{dx^2} + x^3\dot{y}(t) - \frac{x^6}{2}\right)\phi(x))\Big|_{x=x_s} U(0,0,s,y_s)ds$$

and this in turn says that the densities $n_t(y)$ of $N_t(y)$, i.e. the functions $(t,x) \to n_t(y)(x)$ satisfy the partial differential equation

$$(4.29) \qquad \frac{\partial\rho}{\partial t} = \frac{1}{2}\frac{\partial^2\rho}{\partial x^2} + (x^3u(t) - \frac{x^6}{2})\rho, \quad \rho(0,x) = n_0(x)$$

where n_0, the initial density is in F and $u = \dot{y}$. One has

4.30. LEMMA. *Let u be piecewise continuous an $[0,T]$, and for each $n_0 \in F$ let $\rho_{n_0,t}$ be the function $x \to \rho(t,x)$ where ρ solves (4.29) then $(n_0,t) \to \rho_{n_0,t}$: $F \times [0,T] \to F$ is continuous.*

Now let

(4.31) $$L_0 = \frac{1}{2}\frac{d^2}{dx^2} - \frac{x^6}{2}, \quad L_1 = x^3,$$

considered as (differential) operators $F \to F$. For each constant $\bar{u}$ let $L(\bar{u}) = L_0 + \bar{u}L_1$ and let $\exp(tL(\bar{u}))\psi$ for $\psi \in F$ denote the solution of (4.29) with $u(\cdot) \equiv \bar{u}$, $n_0 = \psi$.

Let $K \subset \mathbb{R}^n$ be a convex subset with nonempty interior. A family $\{\psi(v): v \in K\}$ of elements of F is said to depend smoothly on v if $(x,v) \to \psi(v)(x)$ is a C^∞ function on $K \times \mathbb{R}$ and for each $\underline{m} = (m_1,\ldots,m_n)$, $v \mapsto (\frac{\partial^{\underline{m}}}{\partial v^{\underline{m}}}\psi)(v)$, $v \in K$ takes values in F and is a continuous map $K \to F$. One then has

4.32. <u>LEMMA</u>. $\{\exp(tL(\bar{u}))\psi(v)\}$ *depends smoothly on* (v,t) *if* $\psi(v), v \in K$ *is a smooth family.*

4.33. <u>Corollary</u>. *Let* $\bar{u}_1,\ldots,\bar{u}_m \in \mathbb{R}$. *Then if* $\psi \in F$ *the family* $\{\exp(t_1L(\bar{u}_1)) \ldots \exp(t_mL(\bar{u}_m))\psi : (t_1,\ldots,t_m) \in [0,\infty]^m\}$ *depends continuously on* $t_1,\ldots,t_m$; *moreover for each* $\underline{\mu} = (\underline{\mu}_1,\ldots,\underline{\mu}_m)$ *we have*

(4.34) $$\frac{\partial^{\underline{\mu}}}{\partial t^{\underline{\mu}}}(\exp(t_1L(\bar{u}_1)) \ldots \exp(t_mL(\bar{u}_m))\psi) =$$

$$L(\bar{u}_1)^{\mu_1}\exp(t_1L(\bar{u}_1))L(\bar{u}_2)^{\mu_2}\exp(t_2L(\bar{u}_2)) \ldots L(\bar{u}_m)^{\mu_m}\exp(t_mL(\bar{u}_m))\psi.$$

SYSTEM THEORETIC PART III: REALIZATION THEORY

1. <u>Some differential topology on F</u>

Let U be an open subset of the space of smooth functions F. A map $U \to \mathbb{R}$ is said to be of class C^∞ if the function $\nu \to \Phi(\psi(\nu))$ is C^∞ in the ual sense for every family $\{\psi(v) : v \in K\}$ depending smoothly on v in the nse described in section 4 above. This class of functions is denoted $C^\infty(U)$. λ is a continuous linear functioned on F then λ (restricted to any U) is class C^∞. Note that $C^\infty(U)$ is closed under pointwise multiplication and ivision by functions in $C^\infty(U)$ which are everywhere nonzero.

Let L be a continuous linear operator on F, then L defines a "linear vectorfield"

$\tilde{L}: C^\infty(F) \to C^\infty(F)$ (and $C^\infty(U) \to C^\infty(U)$ for each U) defined by

$$(5.2)\qquad (\tilde{L}\,\Phi)(\psi) = \frac{d}{dt}\Big|_{t=0} \Phi(\psi + tL\psi)$$

This is completely analogous to the map which assigns to an $n \times n$ matrix $A = (a_{ij})$ the "linear vector field" $\Sigma\, a_{ij} x_i \frac{\partial}{\partial x_j}$. It is totally routine to check that

$$(5.3)\qquad [L_1, L_2]^\sim = -[\tilde{L}_1, \tilde{L}_2]$$

5.4. <u>LEMMA</u>. *Let* $\{\psi(t) : 0 \le t < \varepsilon\} \subset U$ *depend smoothly on* t *and let* $\dot{\psi}$ *be the t-derivative of* ψ. *Then for all* $\Phi \in C^\infty(U)$

$$(5.5)\qquad \frac{d}{dt}\Big|_{t=0} \Phi(\psi(t)) = \frac{d}{dt}\, \Phi(\psi(0) + t\,\dot{\psi}(0)).$$

In particular if L *is a continuous linear operator on* F *such that* $L\psi(0) = \dot{\psi}(0)$, *then*

$$(5.6)\qquad (\tilde{L}\,\Phi)(\psi(0)) = \frac{d}{dt}\Big|_{t=0} \Phi(\psi(t)).$$

Now let $U \subset F$ be the set of all $\psi \in F$ such that $\int_{-\infty}^{+\infty} \psi(x)dx > 0$ and let $\Phi: U \to \mathbb{R}$ be given by the kind of formula occuring in our conditional expectation expressions

$$(5.7)\qquad \Phi(\psi) = \frac{\int \phi(x)\psi(x)dx}{\int \psi(x)dx}.$$

For the smooth families $e^{tL(\bar{u})}\psi$, where $L(\bar{u})$ is as in 4.32 above, one finds

$$(5.8)\qquad (\tilde{L}(\bar{u})\Phi)(\psi) = \frac{d}{dt}\, \Phi(e^{tL(\bar{u})}\psi)$$

and repeating this

$$(5.9)\qquad \tilde{L}(\bar{u}_m) \ldots \tilde{L}(\bar{u}_1)\Phi(\psi) =$$

$$\frac{\partial^m}{\partial t_1 \partial t_2 \ldots \partial t_m}\Big|_{t_1 = \ldots = t_m = 0} \Phi(e^{t_1 L(\bar{u}_1)} \ldots e^{t_m L(\bar{u}_m)}\psi).$$

5.10. The Lie-algebraic inplications of the existence of a smooth filter

Now let us repeat these remarks for the more familiar case of vectorfields α, β on a smooth finite dimensional manifold M with for each $\bar{u} \in \mathbb{R}$, $A(\bar{u})$ the vectorfield $\alpha + \bar{u}\beta$. Let $\pi((\bar{u}_1,t_1),(\bar{u}_2,t_2),\ldots,(\bar{u}_m,t_m));z)$ be the result of letting z evolve on M along $A(\bar{u}_m)$ during time t_m, then along $A(\bar{u}_{m-1})$ during time t_{m-1}, ... Let $\gamma: M \to \mathbb{R}$ be a smooth function. We have of course

(5.11) $$(A(\bar{u})\gamma)(z) = \frac{d}{dt}\Big|_{t=0} \gamma(\pi((\bar{u},t);z))$$

and

(5.12) $$(A(\bar{u}_m)\ldots A(\bar{u}_1)\gamma)(z) =$$

$$\frac{\partial^m}{\partial t_1 \ldots \partial t_m}\Big|_{t_1=\ldots=t_m=0} \gamma(\pi((\bar{u}_1,t_1),\ldots,(\bar{u}_m,t_m);z).$$

Let $R \subset M$ be the set of all points in M which can be reached from z^{in} v means of these bang-bang-bang controls in time $< \bar{T}$ i.e. R is the set of l $\pi((\bar{u}_1,t_1),\ldots,(\bar{u}_m,t_m);z^{in})$ with $\Sigma\, t_i < \bar{T}$. Let $z \in R$ and choose a bang-ıg control which steers z^{in} to z in time $\tau < \bar{T}$; let $\psi_z \in F$ be the solution the "control version" of the $\mathcal{DMZ}$ equation (4.29), with initial condition the density of ν^{in}. Then $\psi_z \in U$, because $\Delta_t^1(y) > 0$ (cf. just below position 4.12). Now let $\bar{u}_1,\ldots,\bar{u}_m$, $t_1,\ldots,t_m$ satisfy $|\bar{u}_i| = 1$, $_1,\ldots,t_m| < \bar{T} - \tau$, $t_i \geq 0$ and assume that (α,β,γ) on M define a smooth lter for a given $\phi \in E$ in the sense of subsection 2.4. Let Φ be the ɔrresponding functional (5.7). Then by (4.24) we have

ı.13) $$\Phi(e^{t_1 L(\bar{u}_1)} \ldots e^{t_m L(\bar{u}_m)} \psi_z) = \gamma(\pi((\bar{u}_1,t_1)),\ldots,(\bar{u}_m,t_m));z)$$

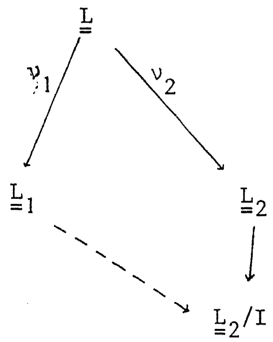
and this was really the whole reason for establishing formula (4.24), that ıs the reason why we needed to prove the existence of a robust pathwise version $\hat{\phi}_t$).

Now let $\underline{\underline{A}}$ denote the free associative algebra on two generators a_-, a_+. Let $\underline{\underline{A}}_1$ be the associative algebra (under composition) of linear maps

$C^{\infty}(U) \to C^{\infty}(U)$ generated by $\tilde{L}(-1)$, $\tilde{L}(1)$ and let $\underline{\underline{A}}_2$ be the associative algebra (again under composition) of differential operators on M. Homomorphisms of associative algebras $\nu_1 : \underline{\underline{A}} \to \underline{\underline{A}}_1$, $\nu_2 : \underline{\underline{A}} \to \underline{\underline{A}}_2$ are defined by $\nu_1(a_-) = \tilde{L}(-1)$, $\nu_1(a_+) = \tilde{L}(1)$, $\nu_2(a_-) = A(-1)$, $\nu_2(a_+) = A(1)$.

Let $\underline{\underline{L}}$ denote the free Lie-algebra on the generators a_-, a_+ (viewed as a subalgebra of A) and let $\underline{\underline{L}}_1$, $\underline{\underline{L}}_2$ denote the Lie algebras generated respectively by $\tilde{L}(-1)$, $\tilde{L}(1)$ and $A(-1)$, $A(1)$ (as subalgebras of $\underline{\underline{A}}_1$ and $\underline{\underline{A}}_2$). Then of course we have induced homomorphisms $\nu_i : L \to L_i$

(5.14)



Let I denote the set of those vectorfields $V \in \underline{\underline{L}}_2$ such that

(5.15) $$([V_1,[V_2,[\ldots[V_m,V]\ldots]]\gamma)(z) = 0 \text{ for all } V_1,\ldots,V_m \in \underline{\underline{L}}_2, \quad m \in \mathbb{N} \cup \{0\}$$

5.16. <u>LEMMA</u>. I *is an ideal and if* $a \in \underline{\underline{L}}$ *is such that* $\nu_1(a) = 0$ *then* $\nu_2(a) \in I$.

It follows that there is a homomorphism of Lie algebras $L_1 \to L_2/I$ making diagram (5.14) commutative.
The lemma is proved by combining (5.12) and (5.9) with (5.13) and this is why we needed to establish smoothness properties of families like $\exp(tL(\bar{u})\psi)$.

5.17. Foliations and such

The last step in this section is now to show that $\underline{\underline{L}}_2/I$, or more precisely a suitable quotient, is (isomorphic to) a subalgebra of a Lie algebra of vectorfields in a smooth finite dimensional manifold (a subquotient manifold of M) for suitable z. Let S be a set of smooth vectorfields on M. For each $z \in M$ consider $S(z) = \{V(z) : V \in S\}$. For $\underline{\underline{L}}_2$ and I, $\underline{\underline{L}}_2(z)$ and $I(z)$ are vectorspaces. Let k be the maximum of the dim $\underline{\underline{L}}_2(z)$ for $z \in R$ and k_0 the maximum of the dim $I(z)$, for $z \in R$ and dim $\underline{\underline{L}}_2(z) = k$. Choose a $\bar{z}$ in the relative interior of R such that dim $\underline{\underline{L}}_2(\bar{z}) = k$, dim $I(\bar{z}) = k_0$, and choose a neighborhood N of $\bar{z}$ (in M) such that dim $\underline{\underline{L}}_2(z) \geq k$, dim $I(z) \geq k_0$ for $z \in N$. (This can be done (obviously)). Suppose $\bar{z}$ can be reached from z^{in} in time $\tau < \bar{T}$. Now let M_0 be a connected submanifold of N of which all points are reachable from $\bar{z}$ in time $< \bar{T} - \tau - \delta$ for some $\delta > 0$ and which is maximal in dimension in the set of all such manifolds. Then α, β are tangent to M_0 and $\bar{z}$ (as is easily checked). Also (becaume $M_0 \subset N$ and $M_0 \subset R$) we have dim $(\underline{\underline{L}}_2(z)) = k$ for all $z \in M_0$ so that by Frobenius theorem there exists a submanifold M_1 of M_0 whose tangent space at each point $z \in M_1$ is precisely $\underline{\underline{L}}_2(z)$. One then also has that dim $I(z) = k_0$ for all $z \in M_1$ so that I has integral manifolds M_2 locally near $\bar{z}$. M_1 is then foliated by the integral manifolds of I so that M_1 locally near $\bar{z}$ looks like $M_1 \simeq M_2 \times M_3$. The Lie algebra of vectorfields of M_3 is then isomorphic to the quotient $\underline{\underline{L}}_2|_{M_1}/I|_{M_1}$.

So That by restriction to M_1 the dotted arrow is diagram 5.14 gives a homomorphism of Lie algebras

$$\underline{\underline{L}}_1 \to V(M_3) \tag{5.18}$$

5.19. Proposition. *Assume that the homomorphism of Lie algebras (5.18) is zero and assume moreover that* $\underline{\underline{L}}_1$ *contains all the operators* $L_k = \frac{d}{dx} x^k$, $k = 0,1,\ldots$ *Then* ϕ *is a constant almost everywhere.*

This is seen as follows. This homomorphism is zero iff $k_0 = k$ so that for $z \in R$, $V\gamma(z) = 0$ for all $V \in \underline{\underline{L}}_2$, which gives $(\tilde{L}\Phi)(\psi_2) = 0$ for all $L \in \underline{\underline{L}}_2$. Now calculate $(\tilde{L}\Phi)(\psi_2)$ using formula (5.7) for $\Phi(\psi_2)$ to find that

(5.20) $\quad <\phi,L\psi_z> <1,\psi_z> = <\phi,\psi_z> <1,L\psi_z>$ for all $L \in \underline{\underline{L}}_2$.

As $<1,L_k\psi_z> = 0$ this gives

(5.21) $$\int \phi(x)[x^k\psi_z'(x) + kx^{k-1}\psi_z(x)]dx = 0, \quad k = 0,1,2,\ldots .$$

From this, using that $\phi(x)\psi_z'(x)$ and $\phi(x)\psi_2(x)$ are bounded by $e^{\beta|x|^r}$ for some $\beta > 0$, $r < 4$, one sees by considering the Fourier transforms of $\phi(x)\psi_z'(x)$ and $\phi(x)\psi_z(x)$ that $\phi'\psi_z = 0$ and as ψ_2 never vanishes that ϕ is constant.

6. ALGEBRAIC PART

6.1. The Weyl Lie algebras W_n

The Weyl Lie algebra W_n is the algebra of all differential operators (any order) in $\frac{\partial}{\partial x_1},\ldots,\frac{\partial}{\partial x_n}$ with polynomial coefficients. The Lie brackett operation is of course the commutator $[D_1,D_2] = D_1D_2 - D_2D_1$. A basis for W_1 (as a vector space over $\mathbb{R}$) constists of the operators

(6.2) $$x^i \frac{\partial^j}{\partial x^j}, \quad i,j = 0,1,2,\ldots$$

(where of course $x^i \frac{\partial^0}{\partial x^0} = x^i$, $x^0 \frac{\partial^j}{\partial x^j} = \frac{\partial^j}{\partial x^j}$, $x^0 \frac{\partial^0}{\partial x^0} = 1$). One has for example

$$[\frac{\partial^2}{\partial x^2},x^2] = 4x\frac{\partial}{\partial x} + 2$$

as is easily verified by calculating

$$[\frac{\partial^2}{\partial x^2},x^2]f(x) = \frac{\partial^2}{\partial x^2}(x^2f(x)) - x^2\frac{\partial^2}{\partial x^2}(f(x))$$

for an arbitrary test function (polynomial) $f(x)$.

Some easy facts (theorems) concerning the Weyl Lie algebras W_n are (cf. [Hazewinkel-Marcus, 1981] for proofs):

6.3. Proposition. *The Lie algebra* W_n *is generated (as a Lie algebra) by the elements* x_i, $\partial^2/\partial x_i^2$, $x_i^2(\partial/\partial x_i)$, $i = 1,\dots,n$; $x_i x_{i-1}$, $i = 2,\dots,n$. *In particular* W_1 *is generated by* x, $\partial^2/\partial x^2$, $x^2(\partial/\partial x)$.

6.4. Proposition. *The only nontrivial ideal of* W_n *is the one-dimensional ideal* $\mathbb{R}1$ *of scalar multiples of the identity operator.*

If M is a C^∞ differentable manifold let $V(M)$ denote the Lie-algebra of all C^∞ vectorfields on M (i.e. the Lie algebra of all derivations on the ring of smooth functions on M). If $M = \mathbb{R}^n$, $V(\mathbb{R}^n)$ is the Lie algebra of all differential operators of the form

$$\sum_{i=1}^{n} g_i(x_1,\dots,x_n)\frac{\partial}{\partial x_i}$$

with $g_i(x_1,\dots,x_n)$ a smooth function on $\mathbb{R}^n$.

A deep fact concerning the Weyl Lie algebras W_n is now

6.5. Theorem. *Let* M *be a finite dimensional smooth manifold. Then there are no nonzero homomorphisms of Lie algebras* $W_n \to V(M)$ or $W_n/\mathbb{R}1 \to V(M)$ *for* $n \geq 1$.

The original proof of this theorem ([Hazewinkel-Marcus, 1981] is long and computational. Fortunately there now exists a much better proof (about two pages) of the main and most difficult part [Stafford, 1982], essentially based on the observation that the associative algebra W_1 cannot have left ideals of finite codimension. For some more remarks about the proof cf. 6.8 below.

6.6. The Lie algebra of the cubic sensor

According to section 2 above the estimation Lie algebra $L(\Sigma)$ of the cubic sensor is generated by the two operators

$$L_0 = \frac{1}{2}\frac{d^2}{dx^2} - \frac{1}{2}x^6, \quad L_1 = x^3.$$

Calculating $[L_0,L_1]$ gives $C = 3x^2\frac{d}{dx} + 3x$. Let $\mathrm{ad}_C(-) = [C,-]$. Then

$(\mathrm{ad}_C)^3 B = C^{st}_{\cdot} x^6$ which combined with A gives as that $(d^2/dx^2) \in L(\Sigma)$. To show that also $x^2 \frac{d}{dx} \in L(\Sigma)$ requires the calculation of some more bracketts (about 15 of them). For the details cf. [Hazewinkel-Marcus, 1981]. Then $x, x^2\frac{d}{dx}, \frac{d^2}{dx^2} \in L(\bar{z})$ which by proposition 6.3 implies:

6.7. Theorem. *The estimation Lie algebra* $L(\Sigma)$ *of the cubic sensor is equal to the Weyl Lie algebra* W_1.

In a similar manner one can e.g. show that the estimation Lie algebra of the system $dx_t = dw_t$, $dy_t = (x_t+\varepsilon x_t^3)dt + dv_t$ is equal to W_1 for all $\varepsilon \neq 0$. It seems highly likely that this is a generic phenomenon i.e. that the estimation Lie algebra of a system of the form $dx_t = f(x_t)dt + G(x_t)dt$, $dy_t = h(x_t)dt + dv_t$ with $x \in \mathbb{R}^n$ and f, G and h polynomial is equal to W_n for almost all (in the Zariski topology sense) polynomials f, G, h.

6.8. Outline of the proof of the nonembedding theorem 6.5

Let $\hat{V}_n$ be the Lie algebra of all expressions

$$\sum_{i=1}^{n} f_i(x_1,\dots,x_n)\frac{\partial}{\partial x_i} \qquad (6.9)$$

where $f_1(x),\dots,f_n(x)$ are formal power series in $x_1,\dots,x_n$. (No convergenc properties are required). Suppose that

$$\alpha: W_n \to V(M) \qquad (6.10)$$

is a nonzero homomorphism of Lie algebras into some V(M) with M finite dimensional. Then there is a $D \in W_n$ and an $m \in M$ such that the tangent vector $\alpha(D)(m) \neq 0$. Now take formal Taylor series of the $\alpha(D)$ around m (with respect to local coordinates at m) to find a nonzero homomorphism of Lie algebra

$$\hat{\alpha}: W_n \to \hat{V}_m \qquad (6.11)$$

where $m = \dim(M)$.

Observe that W_1 is a sub-algebra of W_n (consisting of all differential operators not involving x_i, $i \geq 2$, and $\partial/\partial x_i$, $i \geq 2$) so that it suffices to prove theorem 6.5 for the case $n = 1$.

Because the only nontrivial ideal of W_1 is $\mathbb{R}1$ (cf. proposition 6.4) the existence of a nonzero $\hat{\alpha}: W_1 \to \hat{V}_m$ implies that W_1 or $W_1/\mathbb{R}1$ can be embedded in $\hat{V}_m$.

The Lie-algebra $\hat{V}_m$ carries a filtration $\hat{V}_m = L_{-1} \supset L_0 \supset L_1 \supset \dots$ where the L_i are sub-Lie-algebras. This filtration has the following properties

$$[L_i, L_j] \subset [L_{i+j}] \tag{6.12}$$

$$\bigcap_{i=-1}^{\infty} L_i = \{0\} \tag{6.13}$$

$$\dim(L_{-1}/L_i) < \infty, \quad i = -1, 0, 1, \dots \tag{6.14}$$

where "dim" means dimension of real vectorspaces.

Indeed let

$$f_i(x_1, \dots, x_n) = \sum_{\nu} a_{i,\nu} x^{\nu} \tag{6.15}$$

$\nu = (\nu_1, \dots, \nu_m)$, $\nu_i \in \mathbb{N} \cup \{0\}$ a multi index, be the explicit power series for $f_i(x)$. Then $L_j \subset \hat{V}_m$ consists of all formal vectorfields (6.15) for which

$$a_{i,\nu} = 0 \quad \text{for all } \nu \text{ with } |\nu| \leq j \tag{6.16}$$

where $|\nu| = \nu_1 + \dots + \nu_m$.

If there were an embedding $W_1 \to \hat{V}_m$ or $W_1/\mathbb{R}1 \to \hat{V}_m$ the Lie algebra W_1 or $W_1/\mathbb{R}1$ would interit a similar filtration satisfying (6.12) - (6.15). One can now show, essentially by brute force calculations that W_1 and $W_1/\mathbb{R}1$ do not admit such filtrations. Or much better one observes that (6.12) and (6.14) say that L_i $i = 0,1,2,\dots$ is a subalgebra of finite codimension and applies Toby Stafford's result, loc. cit. that W_1 has no such sub-Lie-algebras.

7. PUTTING IT ALL TOGETHER AND CONCLUDING REMARKS

To condude let us spell out the main steps of the argument leading to theorem 2.10 and finish the proof together with some comments as to the generalizability of the various steps.

We start with a stochastic system, in particular the cubic sensor

$$(7.1) \qquad dx = dw, \quad x(0) = x^{in}, \quad dy = x^3dt + dv$$

described more precisely in 2.1 and with a reasonable function ϕ of the state of which we want to compute the conditional expectation $\hat{\phi}_t$.

The first step now is to show that there exists a pathwise and robust version of $\hat{\phi}_t$. More precisely it was shown in section 4 that there exist a functional

$$(7.2) \qquad \delta_t^{\phi}(y) = \frac{\Delta_t^{\phi}(y)}{\Delta_t^{1}(y)}, \quad \Delta_t^{\phi}(y) = \langle N_t(y),\phi\rangle$$

such that the measures $N_t(y)$ depends continuously on the path $y: [0,t] \to \mathbb{R}$, such that $\Delta_t^1(y) > 0$ all $t > 0$, such that the density $n_t(y)$ is smooth and such that for $y(t) = y_t(\omega) =: y^{\omega}(t)$ a sample path of (7.1) then

$$(7.3) \qquad \hat{\phi}_t(\omega) = \delta_t^{\phi}(y^{\omega}).$$

From this we also obtained in the case of the cubic sensor that $n_t(y)(x)$ as a function of (t,x) satisfies the (control version) of the DMZ equation

$$(7.4) \qquad \frac{\partial}{\partial t} n_t(y)(x) = (\frac{1}{2}\frac{\partial^2}{\partial x^2} - \frac{1}{2}x^6)n_t(y)(x) + n_t(y)(x)\dot{y}(t)x^3$$

for piecewise differentiable functions $y: [0,t] \to \mathbb{R}$. And we showed that the family of densities $n_t(y)$, as a function of t, is smooth in the sense described in 4.25. Actually a more precise statement is needed, we need smoothness as a function of $t_1,\dots,t_m$ if $\dot{y} = u$ with u a bang-bang control of the type $u(t) = \bar{u}_i \in \mathbb{R}$ for $t_1 + \dots + t_{i-1} \leq t < t_1 + \dots + t_i$, $|u_i| = 1$.

This whole bit is the part of the proof that seems most resistant to generalization. At present at least this requires reasonable growth bounds on the exponentials occurring in the Kallianpur-Stiebel formula (that is the explicit pathwise expressions for $\Delta_t^{\phi}(y)$). In particular let us call a family Φ_t of continuous maps $C_t \to \mathbb{R}$ a path-wise version of $\hat{\Phi}_t$, if $\omega \mapsto \Phi_t(y^{\omega,t})$, $y^{\omega,t}(s) = y_s(\omega)$, $0 \le s \le t$ is a version of $\hat{\Phi}_t$. Then it is not at all clear that path-wise versions exist for arbitrary nonlinear filtering problems.

Now suppose that there exists a smooth finite dimensional filter for $\hat{\Phi}_t$. That is a smooth dynamical system

$$(7.5) \qquad dz = \alpha(z) + \beta(z)dy, \quad \gamma\colon M \to \mathbb{R}, \quad z(0) = z^{in}$$

such that if $z_y(t)$ denotes the solution of (7.5) then

$$(7.6) \qquad \gamma(z_y(t)) = \hat{\Phi}_t = \delta_t^{\phi}(y)$$

almost surely. As described in 4.19 above up to a stopping time there also exists a robust pathwise version of the solutions of (7.5) so that $z_y(t)$ exists for all continuous y and so that (7.6) holds always. Now let $L_0 = \frac{1}{2}\frac{d^2}{dx^2} - \frac{1}{2}x^6$, $L_3 = x^3$, $L(\bar{u}) = L_0 + uL_1$. The next step is to show smoothness of

$$(7.7) \qquad e^{t_1 L(\bar{u}_s)} \cdots e^{t_m L(\bar{u}_m)} \psi$$

for smooth ψ as a function of $t_1, \ldots, t_m$, and to calculate $\partial^m/\partial t_1 \ldots \partial t_m$ of (7.7). The result being formula (4.34).

The next thing is to reinterpret a differential operator on F as a linear vectorfield $\tilde{L}$ on F by giving meaning to $\tilde{L}\Phi$ for Φ a functional $F \to \mathbb{R}$ for instance a functional of the form $\delta_t^{\phi}(y)$.

This permits us to give meaning to expressions like

$$(7.8) \qquad \frac{\partial^m}{\partial t_1 \ldots \partial t_m} \delta_t^{\phi}(y) \Big|_{t_1 = \ldots = t_m = 0}, \quad t = t_1 + \ldots + t_m$$

for $y \in C_t^1$ with $\dot{y} = u$ a bang-bang function. The same operator can be applied to the left handside of (7.6) and as both sides depend smoothly on $t_1,\ldots,t_m$ there results from (7.6) an equality of the type

$$(7.9) \qquad (A(\bar{u}_m)\ldots A(\bar{u}_1)\gamma)(z) = (\tilde{L}(\bar{u}_m)\ldots\tilde{L}(\bar{u}_1)\Phi)(\psi_z)$$

where $z \in M$ and $\psi_z \in F$ are corresponding quantities in that they result from feeding in the same control function $\dot{y}(t)$ to the evolution equations for z and ψ respectively.

This relation in turn using some techniques familiar from nonlinear realization theory (essentially restriction to the completely reachable and observable subquotient of M) then implies that there is a homomorphism of Lie algebras from the Lie algebra $L(\Sigma)$ generated by L_0 and L_1 to a Lie algebra of smooth vectorfields. Moreover under the rather inelegant extra assumption that $L(\Sigma)$ contains the operator $\frac{d}{dx}x^k$ we showed that ϕ must have been constant if this homomorphism of Lie algebras is zero. (Proposition 5.19).

The final part is algebra and shows (i) that $L(\Sigma) = W_1$ so that in particular $\frac{d}{dx}x^k \in L(\Sigma)$ for all $k = 0,1,\ldots$ and (ii) there are no nonzero homomorphisms of Lie algebras $W_1 \to V(M_1)$ for M_1 a smooth finite dimensional manifold. Thus both hypothesis of proposition 5.19 are fulfilled and ϕ is a constant. This proves the main theorem 2.10.

It seems by now clear [Hazewinkel-Marcus, 1981b] that the statement $L(\Sigma) = W_k$, k = dim (state space) will turn out to hold for a great many systems (though anything like a general proof for certain classes of systems is lacking). The system theoretic part of the argument is also quite general. The main difficulty of obtaining similar more general results lies thus in generalizing the analytic part or finding suitable subsitutes for establishing the homomorphism principle, perhaps as in [Hijab, 1982].

It should also be stressed that the main theorem 2.10 of this paper only says things about exact filters; it says nothing about approximate filters. On the other hand it seems clear that the Kalman-Bucy filter for $\hat{x}_t$ for

$$(7.10) \qquad dx = dw, \quad dy = xdt + dv$$

should for small ε give reasonable approximate results for

$$(7.11) \qquad dx = dw, \quad dy = (x+\varepsilon x^3)dt + dv.$$

Yet the estimation Lie algebra of (7.11) is for $\varepsilon \neq 0$ also equal to W_1 (a somewhat more tedious calculation cf. [Hazewinkel, 1981]) and the arguments of this paper can be repeated word for word (practically) to show that (7.11) does not admit smooth finite dimensional filters (for non-constant statistics). Positive results that the Kalman-Bucy filter of (7.10) does give an approximation to $\hat{x}_t$ for (7.11) are contained in loc. cit. [Sussmann, 1982], and [Blankenship - Liu - Marcus, 1983].

It is possible that results on approximate filters can be obtained by considering $L(\Sigma)$ not as a bare Lie algebra but as a Lie algebra with two distinguished generations L_0, L_1 which permits us to consider also the Lie algebra $L_s(\Sigma)$ generated by sL_0, sL_1 (where s is an extra variable) and to consider statements like $L_s(\Sigma)$ is close to $L_s(\Sigma')$ module s^t.

REFERENCES

BLANKENSHIP, G.L., C.-H. LIU, and S.I. MARCUS, 1983, *Asymptotic expansion and Lie algebras for some nonlinear filtering problems*, IEEE Trans. Aut. Control 28 (1983).

'ROCKETT, R.W., 1981, *Nonlinear systems and nonlinear estimation theory*, In: [Hazewinkel-Willems, 1981], 441-478.

.ARK, J.M.C., 1978, *The design of robust approximations to the stochastic differential equations of nonlinear filtering*, In: J.K. Skwirzynski (ed.), Communication systems and random process theory, Sijthoff and Noordhoff, 1978.

AVIS, M.H.A., 1982, *A pathwise solution to the equations of nonlinear filtering*, Teor. Verojatnost i. Prim. 27 (1982), 1, 160-167.

IAZEWINKEL, M., 1981, *On deformations, approximations and nonlinear filtering*, Systems and Control Letters 1 (1981).

HAZEWINKEL, M. & S.I. MARCUS, 1981, *Some facts and speculations on the role of Lie-algebras in nonlinear filtering*, In [Hazewinkel-Willems, 1981], 591-604.

HAZEWINKEL, M. & J.C. WILLEMS (eds), 1981, *Stochastic systems: the mathematics of filtering and identification and application,* Reidel Publ. Cy, 1981.

HAZEWINKEL, M. & S.I. MARCUS, 1982, *On Lie algebras and finite dimensional filtering,* Stochastics 7 (1982), 29-62.

HERMANN, R., 1963, *On the accessibility problem in control theory,* In: Int. Symp. on nonlinear differential equations and nonlinear mechanics, Acad. Pr., 1963, 325-332.

HIJAB, O., 1982, *Finite dimensional causal functionals of brownian motion,* To appear in Proc. NATO-ASI Nonlinear stochastic problems (Algarve, 1982), Reidel Publ. Cy.

MARCUS, S.I., S.K. MITTER & D. OCONE, 1980, *Finite dimensional nonlinear estimation for a class of systems in continuous and discrete time,* In: Proc. Int. Conf. on Analysis and Opt. of Stoch. Systems, Oxford 1978, Acad. Pr., 1980.

MITTER, S.K., 1981, *Nonlinear filtering and stochastic mechanics,* In: [Hazewinkel-Willems, 1981], 479-504.

NAGANO, S.K., 1966, *Linear differential systems with singularities and an application to transitive Lie algebras*, J. Math. Soc. Japan 18 (1966), 398-404.

STAFFORD, 1983, *On a result of Hazewinkel and Marcus,* to appear.

SUSSMANN, H.J., 1977, *Existence an uniqueness of minimal realizations of nonlinear systems*, Math. Systems Theory 10 (1977), 263-284.

SUSSMANN, H.J., 1978, *On the gap between deterministic and stochastic ordinary differential equations*, Ann. of Prob. 6 (1978), 19-41.

SUSSMANN, H.J., 1981, *Rigorous results on the cubic sensor problem*, In: [Hazewinkel-Willems, 1981], 637-648.

SUSSMANN, H.J., 1982, *Approximate finite-dimensional filters for some nonlinear problems,* Systems and control letters, to appear.

SUSSMANN, H.J., 1983a, *Rigorous results on robust nonlinear filtering,* to appear.

SUSSMANN, H.J., 1983b, *Nonexistence of finite dimensional filters for the cubic sensor problem,* to appear.