ECONOMETRIC INSTITUTE

NONEXISTENCE OF FINITE-DIMENSIONAL FILTERS
FOR CONDITIONAL STATISTICS OF THE
CUBIC SENSOR PROBLEM

M. HAZEWINKEL, S.I. MARCUS and
H.J. SUSSMANN

REPRINT SERIES no. 366

This article appeared in "Systems & Control Letters",
Vol. 3 (1983).

ERASMUS UNIVERSITY ROTTERDAM - P.O. BOX 1738 - 3000 DR ROTTERDAM - THE NETHERLANDS

Systems & Control Letters 3 (1983) 331–340
North-Holland
December 1983

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

M. HAZEWINKEL

Mathematical Centre, P.O. Box 4079, 1009 AB Amsterdam and Dept. Math., Erasmus Univ. Rotterdam, P.O. Box 1738, 3000 DR Rotterdam, The Netherlands

S.I. MARCUS *

Dept. Electrical Engineering, University of Texas at Austin, TX 78712, USA

H.J. SUSSMANN

Dept. Mathematics, Rutgers University, New Brunswick, NJ 08903, USA

Received 18 June 1983

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 expectation 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 cannot exist a recursive finite-dimensional filter for $\hat{\phi}$ driven by the observations.

Keywords: Cubic sensor, Recursive filter, Robust filtering, Weyl Lie algebra.

AMS Classification: Primary 93E11; Secondary 17B65, 57R25, 57R27, 60H15.

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

$$dx = dw, \quad x_0 = x^{\text{in}}, \tag{1.1}$$

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

$$dy = x^3 dt + dv, \quad y_0 = 0, \tag{1.2}$$

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

$$\begin{aligned} \hat{\phi}_t &= \widehat{\phi(x_t)} \\ &= E[\phi(x_t) : y_s, 0 \leq s \leq t]. \end{aligned} \tag{1.3}$$

By definition a smooth finite-dimensional recursive filter for $\hat{\phi}_t$ is a dynamical system on a smooth finite-dimensional manifold M governed by an equation

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

driven by the observation process, together with an output map

$$\gamma : M \to \mathbf{R} \tag{1.5}$$

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

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

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 purpose 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 [20,21], the details of the algebraic part of the proof can be found in [7]. An alternative much better and shorter proof of the hardest bit of the algebraic part will appear in [15].

The preprint version [8] of the present note contains some 9 pages more detail on the analytic and realization-theoretic parts.

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

* Supported in part by the National Science Foundation under Grant ECS-8022033 and in part by the Joint Services Electronics program under contract F-49620-77-C-0101.

sensor filtering problem is as follows.

(i) $(\Omega, \mathscr{A}, P)$ is a probability space.

(ii) $(\mathscr{A}_t : 0 \leqslant t)$ is an increasing family of σ-algebras.

(iii) (w, v) is a two-dimensional standard Wiener process adapted to the $\mathscr{A}_t$.

(iv) $x = \{x_t : t \geqslant 0\}$ is a process which satisfies $\mathrm{d}x = \mathrm{d}w$, i.e.

$$x_t = x_0 + w_t \quad \text{a.s. for each } t. \tag{2.1}$$

(v) x_0 is $\mathscr{A}_0$-measurable and has a finite fourth moment.

(vi) $\{y_t : t \geqslant 0\}$ is a process which satisfies $\mathrm{d}y = x^3 \mathrm{d}t + \mathrm{d}v$, i.e.

$$y_t = \int_0^t x_s^3 \, \mathrm{d}s + v_t \quad \text{a.s. for each } t. \tag{2.2}$$

(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, $t \geqslant 0$, be the σ-algebra generated by the y_s, $0 \leqslant s \leqslant t$ and let $\phi : \mathbb{R} \to \mathbb{R}$ be a Borel-measurable function. Then the filtering problem (for this particular ϕ) consists of determining

$$E[\phi(x_t)|Y_t].$$

2.4. *Smooth finite-dimensional filters.* Consider a (Fisk–Stratonivič) stochastic differential equation

$$\mathrm{d}z = \alpha(z)\mathrm{d}t + \beta(z)\mathrm{d}y, \quad z \in M, \tag{2.5}$$

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

$$z^{\text{in}} \in M, \quad \gamma : M \to \mathbb{R}. \tag{2.6}$$

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

$$\begin{aligned} &0 < T \leqslant \infty \quad \text{a.s.}, \\ &\{\omega | T(\omega) > t\} \in Y_t \quad \text{for } t \geqslant 0. \end{aligned} \tag{2.7}$$

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

$$\begin{aligned} &E[\phi(x_t)|Y_t] = \gamma(z_t) \\ &\quad \text{a.s. on } \{\omega | T(\omega) > t\}. \end{aligned} \tag{2.8}$$

2.9. *Statement of the theorem.* With these notions the main theorem of this note can be stated

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 $\beta \geqslant 0$ and $0 \leqslant r < 4$*

$$|\phi(x)| \leqslant \exp(\beta |x|^r), \quad -\infty < x < \infty. \tag{2.11}$$

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

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

$$\begin{aligned} &\mathrm{d}x_t = f(x_t)\mathrm{d}t + G(x_t)\mathrm{d}w_t, \\ &x^{\text{in}} = x_0, \quad x_t \in \mathbb{R}^n, \quad w_t \in \mathbb{R}^m, \end{aligned} \tag{3.2}$$

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

$$\mathrm{d}y_t = h(x_t)\mathrm{d}t + \mathrm{d}v_t, \quad y_t \in \mathbb{R}^p, \tag{3.3}$$

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

$$\hat{x}_t = E[x_t|Y_t] = E[x_t | y_s, 0 \leqslant s \leqslant t] \tag{3.4}$$

where Y_t is the σ-algebra generated by the y_s, $0 \leqslant s \leqslant t$. 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 (DMZ) equation

$$\begin{aligned} \mathrm{d}\rho(x, t) = \Bigg(& \tfrac{1}{2}\sum_{i,j} \frac{\partial^2}{\partial x_i \partial x_j}\big((GG^{\mathrm{T}})_{ij}\big) \\ & - \sum_i \frac{\partial}{\partial x_i} f_i - \tfrac{1}{2}\sum h_j^2 \Bigg) \rho(x, t)\mathrm{d}t \qquad (3.5) \\ & + \sum h_j \rho(x, t)\, \mathrm{d}y_{jt}, \end{aligned}$$

$\rho(x, 0) =$ density of x^{in},

where $h_j = h_j(x)$ is the j-th component of h, $(GG^{\mathrm{T}})_{ij}$ 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

$$\begin{aligned} \mathrm{d}\rho(x, t) = & \left(\frac{1}{2}\frac{\mathrm{d}^2}{\mathrm{d}x^2} - \frac{1}{2}x^6 \right) \rho(x, t)\mathrm{d}t \\ & + x^3 \rho(x, t)\mathrm{d}y. \end{aligned} \qquad (3.6)$$

3.7. *The homomorphism principle.* Now assume for a given $\phi: \mathbf{R}^n \to \mathbf{R}$ we have a smooth finite-dimensional filter

$$\begin{aligned} & \mathrm{d}z = \alpha(z)\mathrm{d}t + \sum_j \beta_j(z)\mathrm{d}y_j, \\ & z_0 = z^{\mathrm{in}}, \quad \gamma: \mathbf{R}^n \to \mathbf{R}, \end{aligned} \qquad (3.8)$$

to calculate the statistic

$$\hat{\phi}_t = E\left[\phi(x_t) | Y_t\right].$$

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 Stratonovič sense.

Then, very roughly, we have two ways to process an observation path

$$\gamma^{\omega}: s \to y_s(\omega), \quad 0 \leqslant s \leqslant t,$$

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

$$\Phi: \psi \to \left(\int \psi(x)\, \mathrm{d}x \right)^{-1} \int \psi(x)\phi(x)\, \mathrm{d}x. \qquad (3.9)$$

Assuming that (3.8) is observable, deterministic realization theory [16] then suggests 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 vector fields of (3.6) to the vector fields of (3.6) and which is compatible with the output maps γ and (3.9). The operators in (3.6) define linear vector fields 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

$$\mathrm{d}\rho = L_0\rho\mathrm{d}t + L_1\rho\mathrm{d}y_1 + \cdots + L_p\rho\mathrm{d}y_p.$$

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. [2] and [13] and the references in these two papers.

Let $L \mapsto \tilde{L}$ be the map which assigns to an operator the corresponding linear vector field (analogous to the map which assigns to an $n \times n$ matrix $A = (a_{ij})$ the linear vector field

$$\sum a_{ij} x_i \frac{\partial}{\partial x_j}$$

as $\mathbf{R}^n$). Then $L \mapsto -\tilde{L}$ is a homomorphism of Lie algebras. Further F induces a homomorphism of Lie algebras

$$\mathrm{d}F: \tilde{L}_0 \to \alpha, \ \tilde{L}_i \to \beta_i, \quad i = 1, \ldots, p.$$

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 vector fields 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 the remarks in 3.7 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 $y: [0, t] \to \mathbf{R}^p$ 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 piecewise 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. [3]).

Still another difficulty in using the remarks of 3.7 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 [11].

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 (cf. 4.1 below)

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

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

$$\Delta_t^\phi(y) = \int n_t(y)(x)\phi(x)\mathrm{d}x$$

and to show that $n_t(y)(x)$ is smooth (as a function of x).

The next step is to use that there exists (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 (4.7) *anywhere* (not just a.s.), that is

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

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 $0 \leqslant t < t_m$, equal to $\bar{u}_{m-1}$ for $t_m \leqslant t < t_m + t_{m-1}$, 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

$$\left.\frac{\partial^m}{\partial t_1 \cdots \partial t_m}\right|_{t_1 = \cdots = t_m = 0}$$

to both sides of (3.13) and gives a relation of the type

$$\begin{aligned}(A(\bar{u}_m) \cdots A(\bar{u}_1)\gamma)(z) \\ = \tilde{L}(\bar{u}_m) \cdots \tilde{L}(\bar{u}_1)\Phi(\psi_z)\end{aligned} \tag{3.14}$$

where $A(\bar{u})$ is the vector field $\alpha + \bar{u}\beta$, $L(\bar{u})$ the operator

$$L_0 + \bar{u}L_1 = \left(\frac{1}{2}\frac{\mathrm{d}^2}{\mathrm{d}x^2} - \frac{1}{2}x^6\right) + \bar{u}x^3$$

and $\tilde{L}(\bar{u})$ the linear vector field associated to $L(\bar{u})$, Φ the functional (3.9), and ψ_z a function corresponding to z, cf. Section 5.

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 $L(\Sigma)$ for the cubic sensor. It turns out that $L(\Sigma)$ 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 vector fields on a smooth finite-dimensional manifold and $\alpha: W_1 \to V(M)$ a homomorphism of Lie algebras, then $\alpha = 0$.

4. Analytic part

$\mathscr{E}$ denotes the space of all functions $\phi: \mathbb{R} \to \mathbb{R}$ such that there exist constants $C \in \mathbb{R}$, $\alpha \in \mathbb{R}$ and r, $0 \leqslant r < 4$, such that

$$|\phi(x)| \leqslant C\exp(\alpha|x|^r)$$

for all $x \in \mathbb{R}$. And $\mathscr{F}$ denotes the space of all C^∞-functions $\phi: \mathbb{R} \to \mathbb{R}$ such that

$$\exp(\beta|x|^r)|\phi^{(k)}(x)|$$

is bounded for all $\beta \geqslant 0$, $0 \leqslant r < 4$, $k \in \mathbb{N} \cup \{0\}$. Finally C_t is the space of all continuous functions $y: [0, t] \to \mathbb{R}$ such that $y(0) = 0$.

4.1. *A robust version of* $\hat{\phi}_t = E[\phi(x_t)|y_t]$. There exists for all $\phi \in \mathscr{E}$ a functional

$$\Delta_t^\phi(y): C_t \to \mathbf{R}, \quad \Delta_t^\phi(y) = \langle N_t(y), \phi \rangle$$

such that $\Delta_t^\phi(y)$ is continuous, $\Delta_t^1(y) > 0$ and such that

$$\delta_t^\phi(y) = \frac{\Delta_t^\phi(y)}{\Delta_t^1(y)}, \quad \omega \mapsto \delta_t^\phi(y_t(\omega)), \tag{4.2}$$

is a version of $\hat{\phi}_t$. The formula for $\Delta_t^\phi(y)$ is provided by (the numerator of) the Kallianpur–Striebel formula (slightly modified by a partial integration to remove the $\mathrm{d}y_t$ term). By means of some explicit estimates on the terms occurring in the formula for $\Delta_t^\phi(y)$ it is shown that $\Delta_t^\phi(y)$ is continuous, and that $\Delta_t^\phi(h)$ has a density $n_t(y)$ for ϕ bounded.

Moreover one shows that $n_t(y)$ is smooth, that is in $\mathscr{F}$ if y is smooth. More importantly one shows that $n_t(y)$ as a family of densities depending on y is a smooth family in a certain technical sense. In particular this implies that if $y(t)$ is such that

$$\dot{y}(t) = \bar{u}_i \in \mathbf{R},$$
$$t_{i+1} + \cdots + t_m \leqslant t \leqslant t_i + \cdots + t_m,$$
$$i = 1, \ldots, m,$$
$$t_1 + \cdots + t_m = t, \qquad t_i \geqslant 0,$$

then $n_t(y)$ depends smoothly on $t_1, \ldots, t_m$ in the sense that $n_t(y)(x)$ is a jointly smooth function of $t_1, \ldots, t_m, x$.

Directly from the formula for $\Delta_t^\phi(y)$ one shows that $(t, x) \mapsto n_t(y)(x)$ satisfies the (DMZ) PDE (belonging to the cubic sensor)

$$\frac{\partial \rho}{\partial t} = \frac{1}{2}\frac{\partial^2 \rho}{\partial x^2} + \left(x^3 u(t) - \frac{x^6}{2}\right)\rho, \tag{4.3}$$
$$\rho(0, x) = n_0(x), \quad u = \dot{y}.$$

Note that we first establish existence and smoothness of $n_t(y)(x)$ and afterwards prove that it satisfies the DMZ equation.

Let $\exp(tL(\bar{u}))\psi$ denote the solution of (4.3) thus obtained with

$$\psi = n_0(x) = \text{density of } x^{\text{in}},$$
$$\dot{y}(t) = \bar{u}, \quad 0 \leqslant \tau \leqslant t,$$

and let

$$L(\bar{u}) = L_0 + \bar{u} L_1$$

where

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

Then it readily follows that

$$\frac{\partial^m}{\partial t^m} \exp(t_1 L(\bar{u}_1) \cdots t_m L(\bar{u}_m))$$
$$= L(\bar{u}_1)\exp(t_1 L(\bar{u}_1)) \cdots$$
$$\cdots L(\bar{u}_m)\exp(t_m L(\bar{u}_m))\psi. \tag{4.4}$$

4.5. *Robustness of the filter.* Now consider a stochastic differential on a manifold M (in the Stratonivič sense) with an output map and initial state driven by the (observation) process y_t

$$\mathrm{d}z = \alpha(z)\mathrm{d}t + \beta(z)\mathrm{d}y,$$
$$z(0) = z^{\text{in}}, \quad z \mapsto \gamma(z) \in \mathbf{R}, \quad z \in M \tag{4.6}$$

(such as we would have for a filter for $\hat{\phi}$, cf. 2.4 above). Let $y \in C_t$ be given (not necessarily differentiable). Then $\tau \to z(\tau)$, $0 \leqslant \tau \leqslant t$, is said to be a solution of (4.6) if there exists a neighbourhood U of y in C_t and a continuous map $\tilde{y} \mapsto z(\tilde{y})$, $U \subset C([0, t], M)$ such that $z(\tilde{y})$ in a solution of (4.6) in the usual sense of ODE's for all once differentiable $\tilde{y}$. It is now a theorem that up to a stopping time (4.6) admits solutions in this sense (where now $y_t(w)$ is an observation process), cf. [21] for details, cf. also [17] for the case that y_t is a Wiener process; the same techniques apply.

Denoting the stopping time with T it readily follows that if (4.6) is a filter for a cubic sensor then

$$\delta_t^\phi(y) = \gamma(z(y)(t)) \tag{4.7}$$

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

5. System-theoretic part III. Realization theory

5.1. *Some differential topology on* $\mathscr{F}$. Let $U \subset \mathscr{F}$ be the open subset of all ψ such that $\int\psi(x)\,\mathrm{d}x > 0$. Let $\Phi: U \to \mathbf{R}$ be a functional of the form

$$\Phi(\psi) = \left(\int\psi(x)\,\mathrm{d}x\right)^{-1}\int\phi(x)\psi(x)\,\mathrm{d}x.$$

Then Φ is smooth in the sense that it takes a smooth family of densities depending on a finite number of parameters into a smooth function of these parameters. In particular because $n_t(y)$ is a smooth family, we have that $\delta_t^{\phi}(y)$ is a smooth function of $t_1, \ldots, t_m$ for y of the form described just above (4.3) in the previous section.

To a continuous (linear) operator $L: \mathscr{F} \to \mathscr{F}$ one naturally associates a (linear) vector field $\tilde{L}$ on $\mathscr{F}$ defined by the formula

$$(\tilde{L}\Phi)\psi = \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} \Phi(\psi + tL\psi), \quad \Phi: \mathscr{F} \supset U \to \mathbb{R},$$

and the map $L \mapsto \tilde{L}$ is an anti-homomorphism of Lie algebras, i.e. $[\tilde{L}_1, \tilde{L}_2] = [L_1, L_2]^{\sim}$.

Using the $\tilde{L}$ one has for smooth functionals Φ

$$\frac{\partial^m}{\partial t_1 \cdots \partial t_m}\bigg|_{t_1 = \cdots = t_m = 0} \Phi(\mathrm{e}^{t_1 L(\bar{u}_1) \cdots t_m L(\bar{u}_m)}\psi) = \tilde{L}(\bar{u}_m) \cdots \tilde{L}(\bar{u}_1)\Phi(\psi). \tag{5.2}$$

5.3. *Lie-algebraic implications of the existence of a smooth filter.* The (much easier and well known) analogue of (5.2) for a system (4.6) evolving on a smooth finite-dimensional manifold M is

$$\frac{\partial^m}{\partial t_1 \cdots \partial t_m}\bigg|_{t_1 = \cdots = t_m = 0} \gamma(\pi(\bar{u}_1, t_1) \cdots \pi(\bar{u}_m, t_m); z) = (A(\bar{u}_m) \cdots A(\bar{u}_1)\gamma)(z) \tag{5.4}$$

where

$$\gamma(\pi(\bar{u}_1, t_1) \cdots (\bar{u}_m, t_m); z)$$

is the point of M reached at time $t = t_1 + \cdots + t_m$ by starting in z and evolving along

$$\dot{z} = \alpha(z) + u(t)\beta(z)$$

with

$$u(t) = \bar{u}_i \quad \text{for } t_{i+1} + \cdots + t_m \leqslant t \leqslant t_i + \cdots + t_m.$$

Here $A(\bar{u})$ is the vector field $\alpha(z) + \bar{u}\beta(z)$.

Let $\mathbf{L}_1$ be the Lie algebra generated $\tilde{L}(-1)$, $\tilde{L}(1)$ and $\mathbf{L}_2$ the Lie algebra generated by $A(-1)$ and $A(1)$. Let I denote the ideal in $\mathbf{L}_2$ consisting of the vector fields V such that

$$[V_1, [V_2, [\ldots[V_m, V]\ldots]\gamma(z) = 0 \quad \text{for all } V_1, \ldots, V_m \in \mathbf{L}_2. \tag{5.5}$$

Combining (5.2), (5.4) and (4.7) it follows that $\tilde{L}(-1) \mapsto A(-1)$, $\tilde{L}(1) \mapsto A(2)$ defines a homomorphism of Lie algebras

$$\mathbf{L}_1 \to \mathbf{L}_2 I. \tag{5.6}$$

One now uses fairly standard realization-theoretic arguments to show that (for suitable z) $\mathbf{L}_2 I$ is (locally near z) the Lie algebra of vector fields of the reachable and observable sub-quotient M_3 of M. Thus from the existence of a smooth filter for the cubic sensor the existence results of a homomorphism of Lie algebras

$$\nu: \mathbf{L}_1 \to V(M_3) \tag{5.7}$$

for some smooth finite-dimensional manifold.

The final result in this section is:

5.8. Lemma. *Assume that the homomorphism ν of Lie algebras of (5.7) is zero and assume that $\mathbf{L}_1$ contains all the operators*

$$\frac{\mathrm{d}}{\mathrm{d}x}x^k, \quad k = 0, 1, 2, \ldots.$$

Then ϕ is constant almost everywhere.

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 bracket operation is of course the commutator

$$[D_1, D] = D_1 D_2 - D_2 D_1.$$

A basis for W_1 (as a vector space over $\mathbb{R}$) consists of the operators

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

where of course

$$x^i \frac{\partial^0}{\partial x^0} = x^i, \quad x^0 \frac{\partial^j}{\partial x^j} = \frac{\partial^j}{\partial x^j}, \quad x^0 \frac{\partial^0}{\partial x^0} = 1.$$

One has for example

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

as is easily verified by calculating

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

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

Some easy facts (theorems) concerning the Weyl Lie algebras W_n are (cf. [7] 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^∞ vector fields 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 $n \geqslant 1$.*

The original proof of this theorem [6] is long and computational. Fortunately there now exists a much better proof (about two pages) of the main and most difficult part [15], 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

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

Calculating $[L_0, L_1]$ gives

$$C = 3x^2\frac{\mathrm{d}}{\mathrm{d}x} + 3x.$$

Let $\mathrm{ad}_C(-) = [C, -]$. Then $(\mathrm{ad}_C)^3 = B = \text{Const.}\ x^6$ which combined with A gives as that $(\mathrm{d}^2/\mathrm{d}x^2) \in L(\Sigma)$. To show that also $x^2(\mathrm{d}/\mathrm{d}x) \in L(\Sigma)$ requires the calculation of some more brackets (about 15 of them). For the details cf. [6]. Then

$$x,\ x^2\frac{\mathrm{d}}{\mathrm{d}x},\ \frac{\mathrm{d}^2}{\mathrm{d}x^2} \in L(\Sigma)$$

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

$$\mathrm{d}x_t = \mathrm{d}w_t,\quad \mathrm{d}y_t = x_t\mathrm{d}t + \varepsilon x_t^3 + \mathrm{d}v_t$$

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

$$\mathrm{d}x_t = f(x_t)\mathrm{d}t + G(x_t)\mathrm{d}t,$$
$$\mathrm{d}y_t = h(x_t)\mathrm{d}t + \mathrm{d}v_t,$$

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

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

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

$$\alpha: W_n \to V(M) \tag{6.10}$$

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 algebras

$$\hat{\alpha}: W_n \to \hat{V}_m \tag{6.11}$$

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 $\mathbf{R}1$ (cf. Proposition 6.4) the existence of a nonzero $\hat{\alpha}: W_1 \to \hat{V}_m$ implies that W_1 or $W_1/\mathbf{R}1$ can be embedded in $\hat{V}_m$.

The Lie algebra $\hat{V}_m$ carries a filtration

$$\hat{V}_m = L_{-1} \supset L_0 \supset L_1 \supset \cdots$$

where the L_i are sub-Lie algebras. This filtration has the following properties:

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

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

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

where dim means dimension of real vector spaces.

Indeed let

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

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

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

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

If there were an embedding $W_1 \to \hat{V}_m$ or $W_1/\mathbf{R}1 \to \hat{V}_m$, the Lie algebra W_1 or $W_1/\mathbf{R}1$ would inherit a similar filtration satisfying (6.12)–(6.15). One can now show, essentially by brute force calculations, that W_1 and $W_1/\mathbf{R}1$ do not admit such filtrations. Or much better one observes that (6.12) and (6.14) say that L_i, $i = 0, 1, 2, \ldots$, is a subalgebra of finite codimension and applies Toby Stafford's result [15] that W_1 has no such sub-Lie algebras.

7. Putting it all together and concluding remarks

To conclude 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

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

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 exists a functional

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

such that the measures $N_t(y)$ depend continuously on the path $y: [0, t] \to \mathbf{R}$, such that $\Delta_t^1(y) > 0$ for 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),

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

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

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

for piecewise differentiable functions $y: [0, t) \to \mathbf{R}$. And we showed that the family of densities $n_t(y)$, as a function of t, is smooth as a function of $t_1, \ldots, t_m$ if $\dot{y} = u$ with u a bang-bang control of the type $u(t) = \bar{u}_i \in \mathbf{R}$ for

$$t_1 + \cdots + t_{i-1} \leq t \leq t_1 + \cdots + t_i, \quad |u_i| = 1.$$

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–Striebel formula (that is the explicit pathwise expressions for $\Delta_t^{\phi}(y)$). In particular let us call a family Φ_t of continuous maps $C_t \to \mathbf{R}$ a pathwise version of $\hat{\phi}_t$, if

$$\omega \to \Phi_t(y^{\omega,t}),$$

$$y^{\omega,t}(s) = y_s(\omega), \quad 0 \leqslant s \leqslant t,$$

is a version of $\dot{\phi}_t$. Then it is not at all clear that pathwise versions exist for arbitrary nonlinear filtering problems.

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

$$\mathrm{d}z = \alpha(z) + \beta(z)\mathrm{d}y, \quad \gamma : M \to \mathbf{R}, \quad z(0) = z^{\text{in}}, \tag{7.5}$$

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

$$\gamma(z_y(t)) = \dot{\phi}_t = \delta_t^{\circ}(y) \tag{7.6}$$

almost surely. As described in 4.5 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

$$L_0 = \frac{1}{2}\frac{\mathrm{d}^2}{\mathrm{d}x^2} - \tfrac{1}{2}x^2, \quad L_1 = x$$

and

$$L(\bar{u}) = L_0 + uL_1.$$

The next step is to use the smoothness of

$$\mathrm{e}^{t_1 L(\bar{u}_1)} \cdots \mathrm{e}^{t_m L(\bar{u}_m)}\psi \tag{7.7}$$

for smooth ψ as a function of $t_1, \ldots, t_m$, that is the smoothness of $n_t(y)(x)$, and to calculate $\partial^m/\partial t_1 \cdots \partial t_m$ of (7.7). The result being formula (4.4).

The next thing is to reinterpret a differential operator on $\mathscr{F}$ as a linear vector field $\hat{L}$ on $\mathscr{F}$ by giving meaning to $\hat{L}\Phi$ for Φ a functional $\mathscr{F} \to \mathbf{R}$, for instance a functional of the form $\delta_t^{\circ}(y)$.

This permits us to give meaning to expressions like

$$\frac{\partial^m}{\partial t_1 \cdots \partial t_m}\delta_t^{\circ}(y)|_{t_1 = \cdots = t_m = 0}, \quad t = t_1 + \cdots + t_m, \tag{7.8}$$

for $y \in C_t$ with $\dot{y} = u$ a bang-bang function. The same operator can be applied to the left-hand side 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

$$(A(\bar{u}_m) \cdots A(\bar{u}_1)\gamma)(z)$$
$$= (\hat{L}(\bar{u}_m) \cdots \hat{L}(\bar{u}_1)\Phi)(\psi_z) \tag{7.9}$$

where $z \in M$ and $\psi_z \in \mathscr{F}$ are corresponding quantities in that they result from feeding in the same control function $y(t)$ 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 vector fields. Moreover under the rather inelegant extra assumption that $L(\Sigma)$ contains the operator $(\mathrm{d}/\mathrm{d}x)x^k$ it can be shown that ϕ must have been constant if this homomorphism of Lie algebras is zero (Lemma 5.8).

The final part is algebra and shows (i) that $L(\Sigma) = W_1^{\prime}$ so that in particular $(\mathrm{d}/\mathrm{d}x)x^k \in L(\Sigma)$ for all $k = 0, 1, \ldots$ and (ii) there are no nonzero homomorphisms of Lie algebras $W_1^{\prime} \to V(M_1)$ for M_1 a smooth finite-dimensional manifold. Thus both hypotheses of Lemma 5.3 are fulfilled and ϕ is a constant. This proves the main theorem 2.10.

It seems by now clear [6] that the statement

$$L(\Sigma) = W_k^{\prime}, \quad k = \dim(\text{state space})$$

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 substitutes for establishing the homomorphism principle, perhaps as in [11].

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 x_t for

$$\mathrm{d}x = \mathrm{d}w, \quad \mathrm{d}y = x\mathrm{d}t + \mathrm{d}v \tag{7.10}$$

should for small ε give reasonable approximate results for

$$\mathrm{d}x = \mathrm{d}w, \quad \mathrm{d}y = (x + \varepsilon x^3)\mathrm{d}t + \mathrm{d}v \tag{7.11}$$

Yet the estimation Lie algebra of (7.11) is for $\varepsilon \neq 0$ also equal to $W_1^{\prime}$ (a somewhat more tedious calculation, cf. [5]) 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 nonconstant 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 [5,19,1].

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 $L_s(\Sigma')$ modulo s^t.

References

[1] G.L. Blankenship, C.-H. Liu and S.I. Marcus, Asymptotic expansions and Lie algebras for some nonlinear filtering problems, *IEEE Trans. Automat. Control* 28 (1983).

[2] R.W. Brockett, Nonlinear systems and nonlinear estimation theory, in: [9], pp. 441–478.

[3] J.M.C. Clark, 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, Alphen a/d Rijn, 1978).

[4] M.H.A. Davis, A pathwise solution to the equations of nonlinear filtering, *Teor. Verojatnost i. Prim.* 27 (1) (1982) 160–167.

[5] M. Hazewinkel, On deformations, approximations and nonlinear filtering, *Systems Control Lett.* 1 (1981) 32–36.

[6] M. Hazewinkel and S.I. Marcus, Some facts and speculations on the role of Lie-algebras in nonlinear filtering, in: [9], pp. 591–604.

[7] M. Hazewinkel and S.I. Marcus, On Lie algebras and finite dimensional filtering, *Stochastics* 7 (1982) 29–62.

[8] M. Hazewinkel, S.I. Marcus and H.J. Sussmann, Nonexistence of finite dimensional filters for conditional statistics of the cubic sensor problem, Report MB178, 1983, Mathematical Centre Amsterdam (also to appear in *Coll. ENST-CNET, Developpements Recents dans le Filtrage et le Controle des Processus Aléatoires*, Lect. Notes in Inform. Sci. and Control, Springer).

[9] M. Hazewinkel and J.C. Willems, Editors, *Stochastic Systems: The Mathematics of Filtering and Identification and Application* (Reidel, Dordrecht, 1981).

[10] R. Hermann, On the accessibility problem in control theory, in: *Internat. Symp. on Nonlinear Differential Equations and Nonlinear Mechanics* (Academic Press, New York, 1963), pp. 325–332.

[11] O. Hijab, Finite dimensional causal functionals of Brownian motion, in: *Proc. NATO-ASI Nonlinear Stochastic Problems*, Algarve, 1982 (Reidel, Dordrecht, 1983).

[12] S.I. Marcus, S.K. Mitter and D. Ocone, Finite dimensional nonlinear estimation for a class of systems in continuous and discrete time, in: *Proc. Internat. Conf. on Analysis and Optim. of Stoch. Systems*, Oxford, 1978 (Academic Press, New York, 1980).

[13] S.K. Mitter, Nonlinear filtering and stochastic mechanics, in: [9], pp. 479–504.

[14] S.K. Nagano, Linear differential systems with singularities and an application to transitive Lie algebras, *J. Math. Soc. Japan* 18 (1966) 398–404.

[15] T. Stafford, On a result of Hazewinkel and Marcus, to appear.

[16] H.J. Sussmann, Existence an uniqueness of minimal realizations of nonlinear systems, *Math. Systems Theory* 10 (1977) 263–284.

[17] H.J. Sussmann, On the gap between deterministic and stochastic ordinary differential equations, *Ann. Probab* 6 (1978) 19–41.

[18] H.J. Sussmann, Rigorous results on the cubic sensor problem, in: [9], pp. 637–648.

[19] H.J. Sussmann, Approximate finite-dimensional filters for some nonlinear problems, *Systems Control Lett.*, to appear.

[20] H.J. Sussmann, Rigorous results on robust nonlinear filtering, *Stochastics*, to appear.

[21] H.J. Sussmann, Nonexistence of finite-dimensional filters for the cubic sensor problem, to appear.