Stochastics, 1983, Vol. 11, pp. 65–101
0090-9491/83/1102-0065 $18.50/0

Current Algebras and the Identification Problem

P. S. KRISHNAPRASAD
University of Maryland, College Park, MD 20742

and

STEVEN I. MARCUS
University of Texas, Austin, Texas 78712

and

MICHIEL HAZEWINKEL
Mathematisch Centrum, Amsterdam, The Netherlands

(*Accepted for publication June 30, 1983*)

In this paper, we investigate the identification problem of linear system theory from the viewpoint of nonlinear filtering. Following the work of Brockett and Mitter, one associates in a natural way a certain (infinite dimensional) Lie algebra of differential operators known as the estimation algebra of the problem. For the identification problem the estimation algebra is a subalgebra of a *current algebra*. In this paper we study questions of representation and integrability of current algebras as they impinge upon the identification problem. A Wei–Norman type procedure for the associated Cauchy problem is developed which reveals a sequence of functionals of the observations that play the role of joint sufficient statistics for the identification problem.

1. INTRODUCTION

Consider the stochastic differential system:

$$
\begin{aligned}
d\theta &= 0 \\
dx_t &= A(\theta)x_t\,dt + b(\theta)\,dw_t \\
dy_t &= \langle c(\theta), x_t \rangle dt + dv_t.
\end{aligned}
\tag{1}
$$

Here $\{w_t\}$ and $\{v_t\}$ and independent, scalar, standard, Wiener processes and $\{x_t\}$ is an $\mathbb{R}^n$-valued process. We assume that θ takes values in a smooth manifold $\Theta \hookrightarrow \mathbb{R}^N$, and the map $\theta \to \Sigma(\theta) := (A(\theta), b(\theta), c(\theta))$ is a smooth map taking values in minimal triples. By the *identification problem* we mean the nonlinear filtering problem associated with Eq. (1); i.e. the problem of recursively computing conditional expectations of the form $\pi_t(\phi) \stackrel{\Delta}{=} E[\phi(x_t, \theta) | \mathscr{Y}_t]$ where $\mathscr{Y}_t$ is the σ-algebra generated by the observations $\{y_s : 0 \leqq s \leqq t\}$ and ϕ is a member of a suitable class of real-valued functions on $\mathbb{R}^n \times \Theta$.

It is now well-known that the solution to nonlinear filtering problems of the above type involves in an essential way a linear stochastic partial differential equation in the Ito sense known as the Duncan–Mortensen–Zakai equation (see the papers of Davis and Marcus [1] and Mitter [2] for overview and historical remarks). In the present context, this equation takes the form

$$d\rho = \mathscr{A}_0 \rho dt + \mathscr{B}_0 \rho dy_t, \tag{2}$$

where $\rho \stackrel{\Delta}{=} \rho(t, x, \theta)$ is the joint unnormalized conditional density of x_t and θ given $\mathscr{Y}_t$. The operators $\mathscr{A}_0$ and $\mathscr{B}_0$ are given by

$$\begin{aligned} \mathscr{A}_0 := \tfrac{1}{2}\langle b(\theta), \partial/\partial x\rangle^2 - tr(A(\theta)) \\ - \langle A(\theta)x, \partial/\partial x\rangle \end{aligned} \tag{3}$$

and

$$\mathscr{B}_0 := \langle c(\theta), x\rangle. \tag{4}$$

The Bayes formula, (Kallianpur [3]) implies that

$$\pi_t(\phi) = \sigma_t(\phi)/\sigma_t(1) \tag{5}$$

where

$$\sigma_t(\phi) = \iint_{\Theta\,\mathbb{R}^n} \phi(x, \theta)\rho(t, x, \theta)dx \cdot d\theta. \tag{6}$$

Further, if we let $Q(t, \theta)$ denote the (unnormalized) posterior density of θ given $\mathscr{Y}_t$, then it follows (see [4]) that.

$$dQ(t, \theta) = E[\langle c(\theta), x_t\rangle | \theta, \mathscr{Y}_t] \cdot Q(t, \theta) \cdot dy_t. \tag{7}$$

It is possible to give a pathwise interpretation of equation (2) (see Mitter [5], Davis [6]) by applying a time-dependent gauge transformation of the form,

$$\tilde{\rho}(t,x,\theta) \stackrel{\Delta}{=} \exp(-\langle c(\theta), x\rangle y_t)\cdot \rho(t,x,\theta). \tag{8}$$

Then (applying Ito's rule) $\tilde{\rho}(t,x)$ satisfies the deterministic partial differential equation

$$\frac{\partial \tilde{\rho}}{\partial t} = \left(\mathscr{L}_0 + y_t \mathscr{L}_1 + \frac{y_t^2}{2}\mathscr{L}_2\right)\tilde{\rho} \tag{9}$$

where,

$$\mathscr{L}_0 := \mathscr{A}_0 - \frac{\mathscr{B}_0^2}{2} \tag{10}$$

$\mathscr{L}_1$ and $\mathscr{L}_2$ are given by the commutation rules:

$$\mathscr{L}_1 = -[\mathscr{B}_0, \mathscr{A}_0] \tag{11}$$

$$\mathscr{L}_2 = [\mathscr{B}_0, [\mathscr{B}_0, \mathscr{A}_0]]. \tag{12}$$

The pathwise form (9) is most suitable for what follows and leads to geometrical investigations. *By the estimation algebra of the identification problem we mean the operator Lie algebra $\tilde{G}$ generated by $(\mathscr{A}_0 - \mathscr{B}_0^2)$ and $\mathscr{B}_0$.* For more general nonlinear filtering problems, estimation algebras analogous to $\tilde{G}$ have been emphasized by Brockett and Clark [7], Brockett [8–11], Mitter [12, 2, 5], Hazewinkel and Marcus [13] and others (see [14]) as being objects of central interest. In the papers [24, 15]) we give a classification theorem for identification problems in terms of $\tilde{G}$. See Theorem 1 below.

Our purpose in this paper is to make explicit the structure of the Lie algebra $\tilde{G}$ and certain associated representations. These representations, especially the nontrivial ones, play an important role in sensitivity equations for finite dimensional filters. Sensitivity equations are an essential part of various approximate maximum likelihood algorithms which are widely used in practice [50, 51].

Following Brockett and Mitter we view the Cauchy problem associated with Eq. (9) as a problem of integrating a Lie algebra representation. In the mathematical literature, this approach to solving p.d.e.'s appears in the work of Steinberg [43].

This leads to a Wei–Norman type representation for an infinite-dimensional Lie group associated to the Lie algebra $\tilde{G}$.

Finally, our calculations indicate what functionals of the observations constitute a set of joint sufficient statistics for the identification problem.

2. THE STRUCTURE OF THE ESTIMATION ALGEBRA $\tilde{G}$

To understand the structure of the estimation algebra $\tilde{G}$ it is well worth considering an example.

Example 1 Let

$$dx_t = \theta \cdot dw_t; \quad d\theta = 0 \qquad dy_t = x_t dt + dv_t.$$

Then

$$\mathscr{A}_0 - \frac{\mathscr{B}_0^2}{2} = \frac{\theta^2}{2}\frac{\partial^2}{\partial x^2} - \frac{x^2}{2}$$

and $\mathscr{B}_0 = x$. The Lie algebra $\tilde{G} = \{\mathscr{A}_0 - \mathscr{B}_0^2/2, \mathscr{B}_0\}_{\mathrm{L.A.}}$ is spanned by the set of operators

$$\left(\frac{\theta^2}{2}\frac{\partial^2}{\partial x^2} - \frac{x^2}{2}\right)$$

$$\{\theta^{2n} x\}_{n=0}^{\infty}, \quad \left\{\theta^{2n}\frac{\partial}{\partial x}\right\}_{n=1}^{\infty} \quad \text{and} \quad \{\theta^{2n} \cdot 1\}_{n=1}^{\infty}.$$

We then notice that $\tilde{G}$ is simply a Lie subalgebra with two generators of the infinite dimensional Lie algebra obtained by tensoring the polynomial ring $\mathbb{R}[\theta^2]$ with the six dimensional Lie algebra of operators

$$\mathrm{st}(1)=\left\{\frac{\partial^2}{\partial x^2}, x\frac{\partial}{\partial x}, \frac{\partial}{\partial x}, x^2, x, 1\right\}: \tilde{G}\subseteq\mathbb{R}[\theta^2]\otimes\mathrm{st}(1). \qquad \square$$

The general situation is very much as in this example. Consider the vector space (over the reals) of operators spanned by the set

$$\mathscr{S}:=\left\{\frac{\partial^2}{\partial x_i\partial x_j}, x_i\frac{\partial}{\partial x_j}, \frac{\partial}{\partial x_i}, x_i x_j, x_j, 1,\quad i=1,2,\ldots,n;\ \ j=1,2,\ldots,n\right\}. \tag{13}$$

Elements of $\mathscr{S}$ are assumed to act on $\mathscr{S}(\mathbb{R}^n)$ the Schwartz space of rapidly decreasing functions. This space of operators can be given the structure of a Lie algebra (of dimension $2n^2+3n+1$) under operator commutation (the commutation rules being

$$\left[\frac{\partial^2}{\partial x_i\partial x_j}, x_k\right]=\delta_{jk}\frac{\partial}{\partial x_i}+\delta_{ik}\frac{\partial}{\partial x_j}, \left[\frac{\partial^2}{\partial x_i\partial x_j}, x_l\frac{\partial}{\partial x_k}\right]$$
$$=\delta_{li}\frac{\partial^2}{\partial x_j\partial x_k}+\delta_{jl}\frac{\partial^2}{\partial x_i\partial x_k}\quad \text{etc.}$$

where δ_{jk} denotes the Kronecker symbol). We denote this Lie algebra as $\mathrm{st}(n)$. The structure of $\mathrm{st}(n)$ can be made quite explicit, as follows.

Let (V, B) be a symplectic vector space over the reals. Thus V is a vector space of dimension $2n$ and B: $V\times V\to\mathbb{R}$ is a nondegenerate skew-symmetric bilinear form. The direct sum $V\oplus\mathbb{R}$ can be given the structure of a Lie algebra as follows:

$$[\ ,\ ]:(V\oplus\mathbb{R})\times(V\oplus\mathbb{R})\to V\oplus\mathbb{R}\quad ((v,k),(v',k'))\to(0, B(v,v')).$$

We denote by $h(n)$ the above Lie algebra. The choice of a (symplectic) basis makes the matrix of B take the form

$$J=\begin{pmatrix}0 & I_n\\ -I_n & 0\end{pmatrix} \tag{14}$$

and thus $h(n)$ is nothing but the Heisenberg Lie algebra of dimension $2n+1$. The symplectic group $\mathrm{Sp}(B)$ acts on (V, B) as a group of automorphisms of the symplectic structure B and hence on $h(n)$. This immediately defines a semidirect sum $\mathrm{sp}(2n)\oplus_s h(n)$, of $\mathrm{sp}(2n)$ the symplectic Lie algebra (of matrices M with the property $MJ+JM'=0$) and $h(n)$.

One can show that (see Kirillov [16], Lion and Vergne [17]) that

$$\mathrm{st}(n) \cong \mathrm{sp}(2n)\oplus_s h(n). \tag{15}$$

The algebra $\mathrm{st}(n)$ has the following faithful matrix representation as a subalgebra of $\mathrm{sp}(2n+2)$:

$$\xi \in \mathrm{st}(n) \to A(\xi) = \left[\begin{array}{c|cc|c} 0 & p' & q' & d \\ \hline 0 & \tilde{A} & \tilde{B} & q \\ 0 & \tilde{C} & -\tilde{A}' & -p \\ \hline 0 & 0 & 0 & 0 \end{array}\right] \tag{16}$$

where $\tilde{A}, \tilde{B}, \tilde{C}$, are $n \times n$ matrices $\tilde{B}=\tilde{B}'$, $\tilde{C}=\tilde{C}'$ and $p, q \in \mathbb{R}^n$ and $d \in \mathbb{R}$. (See appendix 1 for the explicit isomorphism between $\mathrm{st}(n)$ as a Lie algebra of differential operators and as a matrix subalgebra of $\mathrm{sp}(2n+2)$.) Since $\mathrm{sp}(2n)$ is simple and $h(n)$ is nilpotent, it follows that Eq. (15) gives the Levi decomposition of $\mathrm{st}(n)$.

Suppose for a moment that θ is a *known constant* (equivalently Θ is a 1-point manifold). This is then the setting of linear filtering and in this case $\tilde{G}=\{\mathscr{A}_0-\mathscr{B}_0^2/2, \mathscr{B}_0\}_{\mathrm{L.A.}} \subset \mathrm{st}(n)$ is solvable and the whole situation is quite well understood (see Brockett [8, 11], Ocone [18]).

In the setting of the identification problem however, θ should be treated as a variable and for each θ, $(\mathscr{A}_0-\mathscr{B}_0^2/2)$ and $\mathscr{B}_0$ take values in $\mathrm{st}(n)$. From the smooth dependence on θ of the triple $(A(\theta), b(\theta), c(\theta))$, it follows that $(\mathscr{A}_0-\mathscr{B}_0^2/2)$ and $\mathscr{B}_0$ are smooth maps from Θ into $\mathrm{st}(n)$. The following general viewpoint is essential.

Let M be a smooth finite-dimensional manifold and let L be a finite-dimensional Lie algebra (over the reals) with the usual topology. The space $L_M \triangleq C^\infty(M; L)$ of smooth maps from M into L can be given the structure of a Lie algebra in following way: given $\phi, \psi \in L_M$, define

$$[\ ,\ ]_M : L_M \times L_M \to L_M \quad [\phi, \psi]_M(p) = [\phi(p), \psi(p)], \quad p \in M.$$

We call L_M with the Lie algrebra structure $[\cdot,\cdot]_M$ defined above, a current algebra.

Remark 2.1 Current algebras play a fundamental role in the physics of quantum fields [19] and in the geometric theory of Yang–Mills fields [20]. Elsewhere in mathematics they are studied in the guise of local Lie algebras or Lie algebra bundles [21, 22]. The following is immediate.

PROPOSITION 2.1 *For the identification problem, the estimation algebra $\tilde{G}$ generated by the operators*

$$\mathscr{A}_0-\frac{\mathscr{B}_0^2}{2}=\frac{1}{2}\left\langle b(\theta),\frac{\partial}{\partial x}\right\rangle^2-\left\langle\frac{\partial}{\partial x},A(\theta)x\right\rangle-\frac{\langle c(\theta),x\rangle^2}{2}$$

and

$$\mathscr{B}_0=\langle c(\theta),x\rangle,$$

is a solvable *subalgebra of the current algebra $C^\infty(\Theta;\mathrm{st}(n))$.* □

If Θ is a finite set then clearly $C^\infty(\Theta;\mathrm{st}(n))$ is finite dimensional and so is $\tilde{G}$ (this feature has been exploited by Hijab [49]). In general, when $\dim\Theta\geqq 1$, the following proposition is of interest.

PROPOSITION 2.2 *Assume that $b(\theta)\equiv 0$ (no driving noise), and Θ is a smooth connected manifold of* $\dim\geqq 1$. *Then one of the following situations holds:*

a) *the map $\theta\to A(\theta)$ is* nonconstant *and $\tilde{G}$ is isomorphic to the "shift" Lie algebra with basis $\{X_0,X_1,X_2,\dots\}$ and commutation relations*

$$[X_0,X_i]=X_{i+1}\quad i=1,2,\dots\qquad [X_i,X_j]=0\qquad i,j\geqq 1$$

b) *the map $\theta\to A(\theta)$ is* constant *and $\tilde{G}$ is finite dimensional and isomorphic to a Lie algebra with basis $\{X_0,X_1,\dots,X_k\}$ and commutation relations*

$$[X_0,X_i]=X_{i+1}\quad i=1,2,\dots,k-1$$

$$[X_0,X_k]=\sum_{i=1}^{k-1}p_iX_i\qquad [X_i,X_j]=0\quad i,j\geqq 1$$

where p_i are constants.

If $b(\theta) \not\equiv 0$ then $\tilde{G}$ is necessarily infinite-dimensional.

Proof It is a calculation [see 15]. □

Remark 2.2 The conclusions of Proposition 2.2 are not affected if in our stochastic differential equation model we introduce an additional, known, deterministic input term with known constant coefficients.

3. SOBOLEV LIE GROUPS ASSOCIATED TO $\tilde{G}$

It has been pointed out elsewhere [8, 2, 18, 23, 24, 43] and [25] that the Cauchy problem associated with equations of the form (9) may be viewed as a problem of integrating a Lie algebra representation. To pursue this further one should be able to associate Lie groups to $\tilde{G}$. Since $\tilde{G}$ is infinite-dimensional this question merits careful study. We carry out such a study for general current algebras and later specialize to the cases motivated by the identification problem.

Let M be a compact Riemannian manifold of dimension r with volume element dm. Let L be a Lie algebra (over $\mathbb{R}$) of dimension $n<\infty$. We can always view L as a subalgebra of the general linear Lie algebra $gl(m;\mathbb{R})$ for some $m>n$ (Ado's theorem).

Hypothesis (H) Let $G=\{\exp(L)\}_G \subset gl(m;\mathbb{R})$ be the smallest Lie group containing the exponential of elements of L. We assume that G is a closed subset of $gl(m;\mathbb{R})$.

Define the spaces of smooth maps $\mathscr{R}=C^\infty(M;gl(m,\mathbb{R}))$; $\mathscr{L}=C^\infty(M;L)$; $\mathscr{G}=C^\infty(M;G)$. The space $\mathscr{R}$ has the structure of an infinite-dimensional algebra under pointwise multiplication and

$$\mathscr{L}\subset\mathscr{R}, \quad \mathscr{G}\subset\mathscr{R}.$$

One can construct Sobolev completions of $\mathscr{L}$ and $\mathscr{G}$ as follows. Let $\{(U_\alpha, \varphi_\alpha)\}$ be a finite covering of C^∞ charts for the manifold M. Let $(x_1,\ldots,x_r)$ denote the associated local coordinates. For $l=(l_1,\ldots,l_r)$ an r-multi-index and $f\in\mathscr{R}$, let

$$D^l(f\circ\varphi_\alpha^{-1})=(\partial^{|l|}(f\circ\varphi_\alpha^{-1})/\partial x_1^{l_1}\ldots\partial x_r^{l_r}) \tag{17}$$

(an $m \times m$ matrix of partial derivatives w.r.t. the coordinates x_i). One can now define, for $f_1, f_2 \in \mathscr{R}$

$$\|f_1 - f_2\|_k = \left[\sum_\alpha \int_{\varphi_\alpha(U_\alpha)} dm. \sum_{|l|=0}^{k} |D^l((f_1 - f_2) \circ \varphi_\alpha^{-1})|^2 \right]^{1/2} \tag{18}$$

where

$$|f|^2 = \operatorname{tr}(f'f). \tag{19}$$

(Here $k > r/2$). We call $\|\cdot\|_k$ the Sobolev k-norm and we let $\mathscr{R}_k$, $\mathscr{L}_k$ and $\mathscr{G}_k$ respectively denote the (Sobolev) completions of $\mathscr{R}$, $\mathscr{L}$ and $\mathscr{G}$ in the norm $\|\cdot\|_k$. By our hypothesis (H) and by the condition $k > r/2$, $\mathscr{G}_k$ is closed in $\mathscr{R}_k$. The condition $k > r/2$ guarantees that the definitions of $\mathscr{R}_k$, $\mathscr{L}_k$ and $\mathscr{G}_k$ do not depend on the choice of charts ([26, 27, 28]). Furthermore by the Sobolev theorem, $\mathscr{R}_k$ is a Banach algebra (a matrix Schauder ring) and hence the group operation

$$\cdot : \mathscr{G}_k \times \mathscr{G}_k \to \mathscr{G}_k \qquad (f_1, f_2) \to f_1 f_2$$

where $(f_1 f_2)(m) = f_1(m) \cdot f_2(m)$, $m \in M$ is continuous. Thus $\mathscr{G}_k$ is a *topological group*.

Similarly the bracket operation,

$$[\ ,\ ] : \mathscr{L}_k \times \mathscr{L}_k \to \mathscr{L}_k \qquad (f_1, f_2) \to [f_1, f_2]$$

where $[f_1, f_2](m) = [f_1(m), f_2(m)]$ is continuous. Now the next step is to give $\mathscr{G}_k$ the structure of a Lie group and then identity $\mathscr{L}_k$ as the Lie algebra (tangent space at the identity) of the infinite-dimensional Lie group $\mathscr{G}_k$. The basic idea is to use the exponential map. Define,

$$\exp : \mathscr{L}_k \to \mathscr{G}_k \qquad f \to \exp f$$

$$(\exp f)(m) = \exp f(m) \quad m \in M. \tag{20}$$

We can now appeal to the following version of the "ω-Lemma" [27, 29, 30], (a basic result in global nonlinear analysis).

ω-LEMMA *Let M be a compact manifold of dimension d and let $H^s(M; \mathbb{R}^m)$ and $H^s(M; \mathbb{R}^p)$ respectively denote the Sobolev spaces of maps (of order s) from M into $\mathbb{R}^m$ and $\mathbb{R}^p$. Assume $s > d/2$. Then for* *any C^∞ map $\phi: \mathbb{R}^m \to \mathbb{R}^p$, the map*

$$\tilde{\phi}: H^s(M, \mathbb{R}^m) \to H^s(M, \mathbb{R}^p)$$

defined by

$$(\tilde{\phi} f)(m) = \phi(f(m))$$

is a C^∞ map of Sobolev spaces. □

In the present case, since $\mathscr{G}_k \subset \mathscr{R}_k$ we conclude that,

$$\exp: \mathscr{L}_k \to \mathscr{G}_k \text{ is a smooth map.}$$

It can be shown further that the differential of exp is the identity map at the origin. Hence by the inverse function theorem, there is a sufficiently small neighborhood $V_k(0)$ of the origin of $\mathscr{L}_k$ which is mapped diffeomorphically onto a neighborhood of the identity in $\mathscr{G}_k$. Translating $\exp(V_k(0))$ by right multiplication by elements of $\mathscr{G}_k$, we obtain a covering of $\mathscr{G}_k$ by C^∞ charts. We have thus proved.

PROPOSITION 3.1 *The topological group $\mathscr{G}_k$ is a Lie group and $\mathscr{L}_k$ is its Lie algebra.* □

Remark 3.1 The above result appears in study of Yang–Mills fields [31, 32, 33] for the restricted case of G (the gauge group) being a compact Lie group. Here we essentially use the fact that for any right-invariant Riemannian metric on a Lie group, the exponential map is global (i.e. the geodesics are extendable for all time). Further, the ω-lemma furnishes the essential step in the construction of the C^∞ structure on $\mathscr{G}_k$.

We emphasize that whereas in the setting of Yang-Mills fields the gauge group G is compact (Su(n) etc.), for the identification problem G is not compact. In fact we let,

$$L = \mathrm{st}(n) \qquad G = \{\exp(\mathrm{st}(n))\}_G.$$

From Kirillov [16, 34] it follows that $G = \mathrm{St}(n)_+$ is the connected component of the identity in the subgroup $\mathrm{St}(n)$ of $\mathrm{Sp}(2n+2)$ that leaves fixed a nonzero vector in $\mathbb{R}^{2n+2}$. Hence it satisfies the hypothesis (H) of this section. We may take $m = 2n+2$.

In the notation of this section, let, $M=\Theta$ the parameter space ($\dim\Theta=r$) with a fixed Riemannian metric and associated volume element $d\theta$. The methods of this section apply and we obtain a whole family of Sobolev Lie groups $\{\mathscr{G}_k : k>r/2\}$ associated to the current algebra $C^\infty(\Theta;\mathrm{st}(n))\subset C^\infty(\Theta;\mathrm{sp}(2n+2))$. The estimation algebra $\tilde{G}\subset C^\infty(\Theta;\mathrm{st}(n))$ can be completed in the Sobolev norm $\|\cdot\|_k$ and we obtain a closed subgroup of $\mathscr{G}_k$ with Lie algebra = completion of $\tilde{G}$ in the $\|\cdot\|_k$ norm. We denote this Sobolev Lie group as G_k^Θ. Thus it is now possible to associate a family $\{G_k^\Theta : k$ integer, $k>r/2\}$ of Sobolev Lie groups to a given identification problem, – specified by Θ, the parameterization $(A(\theta), b(\theta), c(\theta))$ and the Riemannian volume element $d\theta$.

Remark 3.2 Diffeomorphisms of Θ that preserve the volume element $d\theta$ result in isomorphic G_k^Θ.

Before we close this section we discuss an alternative approach to current algebras and current groups. Given,

a) A Lie algebra L (over $\mathbb{R}$ say) with bracket [,], and dimension $=n$.

b) A commutative ring F with unit 1; consider the tensor-product space $L_F=L\otimes F$. L_F can be given the structure $[\ ,\ ]_F$ of a Lie algebra as follows:

$$[\ ,\ ]_F : L_F\times L_F\to L_F$$

$$[X_1\otimes f_1, X_2\otimes f_2]_F=[X_1,X_2]\otimes f_1 f_2$$

$$X_i\in L, f_i\in F.$$

If F = the ring of C^∞ real-valued functions on a manifold M then L_F can be identified with the current algebra $C^\infty(M;L)$. The identification may be given as follows. Choose a basis $\{X_1, X_2, \ldots, X_n\}$ for L. Then any $\phi\in C^\infty(M;L)$ may be represented as

$$\phi(p)=\phi_1(p)X_1+\ldots+\phi_n(p)X_n,\quad p\in M$$

for some uniquely determined $\phi_i\in F$. Hence the required identification is

$$\phi\leftrightarrow\sum_{i=1}^{n} X_i\otimes\phi_i. \tag{22}$$

Furthermore if $H^k(M;\mathbb{R})$ are the Sobolev completions of F, then for $k>r/2$ we also have the identifications of the Sobolev Lie algebras $\mathscr{L}_k$ with the tensor products $L\otimes H^k(M;\mathbb{R})$.

This latter approach to current algebras is especially useful in treating purely algebraic matters. In the next section we discuss representations of the estimation algebra.

4. REPRESENTATIONS OF THE ESTIMATION ALGEBRA OF THE IDENTIFICATION PROBLEM

In this section we initiate a study of the representations of the estimation algebra $\tilde{G}$. Our idea is to construct representations of the current algebra $C^\infty(\Theta;\mathrm{st}(n))$ and restrict these representations to the subalgebra $\tilde{G}$. We focus on two types of representations.

Type (I) Representations by differential operators on $C^\infty(\Theta;\mathscr{S}(\mathbb{R}^n))$ the space of smooth maps from Θ into the Schwartz space of $\mathbb{R}^n$.

Type (II) Representations by vector fields on smooth manifolds.

The motivation for studying representations of Type (I) derives from the fact that $\tilde{G}$ is to begin with given as the Lie algebra generated by the differential operators $(\mathscr{A}_0-\mathscr{B}_0^2/2)$ and $\mathscr{B}_0$ and these in fact act on $C^\infty(\Theta;\mathscr{S}(\mathbb{R}^n))$. In more general terms we have the *imaginary spinor representation*. (See Appendix 1).

$$T_{-\sqrt{-1}}:\mathrm{st}(n)\to\mathrm{End}(\mathscr{S}(\mathbb{R}^n)).$$

Associated to this we have the current algebra representation

$$\hat{T}_{-\sqrt{-1}}:C^\infty(\Theta;\mathrm{st}(n))\to\mathrm{End}(C^\infty(\Theta;\mathscr{S}(\mathbb{R}^n))$$

defined by

$$((\hat{T}_{-\sqrt{-1}}\phi)f)(\theta)=(T_{-\sqrt{-1}}\phi(\theta))f(\theta) \tag{23}$$

where $f\in C^\infty(\Theta;\mathscr{S}(\mathbb{R}^n))$.

The representation $\hat{T}_{-\sqrt{-1}}$ is now restricted to $\tilde{G}$ and we obtain $\mathscr{A}_0-\mathscr{B}_0^2/2$ and $\mathscr{B}_0$ as generators. Now representations of the form

(23) are *trivial* in the sense that the operators $\hat{T}_{\ldots}\,\phi$ do *not* depend on the derivatives of ϕ. In what follows we demonstrate that representations that are *nontrivial* (in the sense that the operator associated to ϕ does depend on the derivatives of ϕ) are of interest. We then proceed to construct such representations. To keep the notation from getting too messy we focus on the case when Θ is a connected compact subset of the real line.

For reasons of Taylor series expansions and approximation, it is important to know how $\tilde{\rho}$ satisfying Eq. (9) depends on θ. One can write down the parabolic system

$$\frac{\partial}{\partial t}\begin{bmatrix} \dfrac{\partial \tilde{\rho}}{\partial \theta} \\ \\ \tilde{\rho} \end{bmatrix} = \left[\begin{array}{c|c} \Delta(\theta) & \dfrac{\partial \Delta}{\partial \theta} \\ \hline 0 & \Delta(\theta) \end{array}\right] \begin{bmatrix} \dfrac{\partial \tilde{\rho}}{\partial \theta} \\ \\ \tilde{\rho} \end{bmatrix} \tag{24}$$

where

$$\Delta(\Theta) = \mathscr{L}_0 + y_t \mathscr{L}_1 + \frac{y_t^2}{2} \mathscr{L}_2$$

and

$$\frac{\partial \Delta}{\partial \theta} = \frac{\partial \mathscr{L}_0}{\partial \theta} + y_t \frac{\partial \mathscr{L}_1}{\partial \theta} + \frac{y_t^2}{2} \frac{\partial \mathscr{L}_2}{\partial \theta}. \tag{25}$$

We shall see that Eq. (24) is intimately connected to the above-mentioned *nontrivial representations*.

Motivated by earlier work of Gell-Mann and others, Robert Hermann gave in a series of basic papers the following scheme for constructing nontrivial representations of current algebras [35, 36, 37]:

Let L be a Lie algebra with basis $\{A_1, \ldots, A_n\}$ and a faithful representation

$$\varphi: L \to \mathrm{End}(V) \qquad A_i \to \varphi(A_i) = D_i.$$

Let $\{C_{jk}^i\}$ be the associated structure constants of L. Let $F = C_0^\infty(\mathbb{R})$ be the ring of smooth functions on the real line with compact support. Then the following map (for any vector space W)

$$\hat{\varphi}: L\otimes F \to \mathrm{End}(W\otimes F)$$

$$A_i \otimes f \to D_i \otimes f + d_i \otimes \frac{\partial f}{\partial x} \quad i=1,2,\ldots,n \tag{26}$$

is a faithful representation of the current algebra $L\otimes F$ provided that the following conditions hold:

$$[d_i, d_j]=0 \qquad [D_i, d_j]=\sum_{k=1} C^k_{ij} d_k \quad j,i=1,2,\ldots,n. \tag{27}$$

The proof of these conditions follows by a direct calculation requiring that $\hat{\varphi}$ be a homomorphism. Clearly the operators $\{d_i\}$ form an abelian Lie algebra with the operators $\{D_i\}$ acting on the set $\{d_i\}$ by adjoint representation. Notice further that by introducing a parameter μ in (26)

$$A_i \otimes f \to D_i \otimes f + \mu d_i \otimes \frac{\partial f}{\partial x} \tag{26μ}$$

one obtains a family $\{\hat{\varphi}_\mu\}$ of *nontrivial representations* of $L\otimes F$ that may be viewed as a deformation of the trivial representation $\hat{\varphi}_0$.

For a complete understanding of the above scheme in its full generality, one needs more sophisticated tools (jet bundles, Lie algebra cohomology). In this paper we follow the more elementary methods of Parthasarathy–Schmidt [38].

Given any Lie algebra L one can define its *Liebnitz extension* L_n as simply the $(n+1)$-fold cartesian product with a new bracket $[\cdot,\cdot]_n$ defined as follows:

$$[X, Y]_n = Z$$

where $X=(X_0, X_1, \ldots, X_n)$, $Y=(Y_0, Y_1, \ldots, Y_n)$ and $Z=(Z_0, Z_1, \ldots, Z_n)$ $\in L_n$

$$Z_0=[X_0, Y_0]$$

$$Z_k=\sum_{r=0}^{k}\binom{k}{r}[X_r, Y_{k-r}] \quad k=1,2,\ldots,n. \tag{28}$$

Elements of the form $(0, X_1, \ldots, X_n)$ in L_n form a nilpotent subalgebra of L_n, denoted on L_n^0. In case $n=1$, then L_1^0 is abelian and is simply the underlying vector space of L with the natural adjoint action defined on it.

Now let $C_0^\infty(\mathbb{R}; L_n) = L_n \otimes C_0^\infty(\mathbb{R})$ be the current algebra of compactly supported smooth maps of $\mathbb{R}$ into L_n. Then one has:

THEOREM 4.1 [Parthasarathy–Schmidt] *The map* $\Pi_n : C_0^\infty(\mathbb{R}, L) \to C_0^\infty(\mathbb{R}; L_n)$ *defined by*

$$\Pi_n(f) = (f, f^1, \ldots, f^{(n)}) \tag{29}$$

is an isomorphism into $C_0^\infty(\mathbb{R}; L_n)$ *of the current algebra* $C_0^\infty(\mathbb{R}; L)$.

Proof The proof follows from the observation,

$$\frac{d^k}{dx^k}[f, g] = \sum_{r=0}^{k} \binom{k}{r} [f^{(r)}, g^{(k-r)}]$$

and the definition (28) of the Leibnitz extension commutation relations. □

Now if V is a (possibly infinite-dimensional) vector space and if

$$\psi : L \to \mathrm{End}(V)$$

is a (faithful) representation, then the map

$$\psi_n : L_n \to \mathrm{End}(V \times V \times \ldots \times V) \quad n+1 \text{ times}$$

$$(X_0, X_1, \ldots, X_n) \to \begin{bmatrix} \psi(X_0) & \frac{\psi(X_1)}{1!} & \cdots & \frac{\psi(X_n)}{n!} \\ & \ddots & \ddots & \vdots \\ & & \ddots & \frac{\psi(X_1)}{1!} \\ 0 & & & \psi(X_0) \end{bmatrix} \tag{30}$$

is a (faithful) representation of the Leibnitz extension. The proof is a calculation and is left to the reader.

Now take the *trivial* representation $\hat{\psi}_n$ of $C_0^\infty(\mathbb{R}; L)^n)$ and the composition $\hat{\psi}_n \circ \Pi_n$. It follows from Theorem 4.1 that

$$\hat{\psi}_n \circ \Pi_n : C_0^\infty(\mathbb{R}; L) \to \mathrm{End}(V \times \ldots \times V)$$

is a (faithful) representation of the current algebra $C_0^\infty(\mathbb{R}; L)$. Moreover it is a *nontrivial representation* since it depends on derivatives of elements of $C_0^\infty(\mathbb{R}; L)$. Explicitly,

$$\hat{\psi}_n \circ \Pi_n(A \otimes f) = \begin{bmatrix} \psi(A) \otimes f & \dfrac{\psi(A) \otimes f'}{1!} & \cdots & \dfrac{\psi(A) \otimes f^{(n)}}{n!} \\ & \ddots & \ddots & \vdots \\ & & \ddots & \dfrac{\psi(A) \otimes f'}{1!} \\ 0 & & & \psi(A) \otimes f \end{bmatrix} \tag{31}$$

Now Hermann's scheme (26) follows as a special case of the above construction if we let:

$$n = 1 \quad \text{(first Leibnitz extension)}$$

$$W = V \times V$$

$$D_i = \begin{bmatrix} \psi(A_i) & 0 \\ 0 & \psi(A_i) \end{bmatrix} \tag{32}$$

$$d_i = \begin{bmatrix} 0 & \psi(A_i) \\ 0 & 0 \end{bmatrix}. \tag{33}$$

The choices (32), (33), automatically satisfy Hermann's conditions (27).

Returning to the identification problem we let $L = \mathrm{sl}(n)$ and $\psi = T_{-\frac{N}{2}-1}$ the imaginary spinor representation. Then restricting the constructions of this section to $\tilde{G} \subset C_0^\infty(\mathbb{R}; \mathrm{sl}(n))$ yields a whole family of $\{\hat{\psi}_k \circ \Pi_k : k = 1, \ldots\}$ of *nontrivial* representations of $\tilde{G}$. The family can be further enlarged by introducing deformation parameters as in (26μ). It should now be clear that the parabolic system (24) precisely corresponds to the member of this family associated with the first Leibnitz extension.

Example 4.1 Once again consider the identification problem of Example 2.1.

$$\mathscr{A}_0 = \frac{\theta^2}{2}\frac{\partial^2}{\partial x^2}; \quad \mathscr{B}_0 = x.$$

$\tilde{G}$ = the Lie algebra of operators generated by $\mathscr{A}_{00} = (\theta^2/2)(\partial^2/\partial x^2) - (x^2/2)$ and $\mathscr{B}_0$. The first Leibnitz extension yields the *nontrivial representation* of $\tilde{G}$ generated by the pair of operators on $(C^\infty(\Theta; \mathscr{S}(\mathbb{R}^1) \times \mathscr{S}(\mathbb{R}^1))$

$$\mathscr{A}_{00}^1 = \begin{bmatrix} \frac{\theta^2}{2}\frac{\partial^2}{\partial x^2} - \frac{x^2}{2} & \theta\frac{\partial^2}{\partial x^2} \\ 0 & \frac{\theta^2}{2}\frac{\partial^2}{\partial x^2} - \frac{x^2}{2} \end{bmatrix}$$

and

$$\mathscr{B}_0^1 = \begin{bmatrix} x & 0 \\ 0 & x \end{bmatrix}.$$

Thus

$$\{\mathscr{A}_{00}, \mathscr{B}_0\}_{\text{L.A.}} \cong \{\mathscr{A}_{00}^1, \mathscr{B}_0^1\}_{\text{L.A.}}! \qquad \square$$

We now turn to representations of Type II. The primary motivation here is what we call the *homomorphism principle* of nonlinear filtering theory, an idea due to Brockett [8].

Suppose for a given ϕ there exists a finite-dimensional stochastic differential system (in the Ito sense) of the form:

$$dz_t = f(z_t)\,dt + g(z_t)\,dy_t \tag{34}$$

$$\pi_t(\phi) = h(z_t). \tag{35}$$

Such recursive estimators are of obvious practical interest. One might view the pair of Eqs. (34)–(35) and the pair of Eqs. (2) and (5) as defining alternative realizations of the same input-output map $y_t \to \Pi_t(\phi)$. Brockett argued in [8] that under technical hypothesis there should be a homomorphism from the Lie algebra $\tilde{G} = \{\mathscr{A}_0 - \mathscr{B}_0^2/2, \mathscr{B}_0\}_{\mathrm{L.A.}}$ to the Lie algebra of vector fields generated by $(f - \frac{1}{2}(\partial g/\partial z)g)$ and g. This homomorphism principle has been verified in several situations (Brockett [8, 9]; Ocone [18]; Liu-Marcus [39]; Benes [40]). The question of existence of such homomorphisms (Type II representations) is thus of interest in connection with the existence of optimal finite-dimensional recursive estimators of the form (34)–(35) for nontrivial statistics. Hazewinkel and Marcus [13] have isolated classes of nonlinear filtering problems for which the appropriate estimation algebra admits no such Type II representation.

One of the results is that the Lie algebra $\tilde{G}$ of the identification problem admits faithful Type II representations, and further, that the homomorphism principle is verified.

First let $\{A_1, A_2, \ldots, A_n\}$ be a basis for a finite-dimensional Lie algebra L and let $\Phi: L \to \mathrm{Vect}(N)$ be a (faithful) representation of L, where $\mathrm{Vect}(N)$ is the Lie algebra of smooth vector fields on a finite-dimensional manifold N (recall Ado's theorem). Let $Y_i = \Phi(A_i)$, $i = 1, 2, 3, \ldots, n$. Let $P = M \times N$ and $\pi: P \to M$ be the canonical projection.

Then the map

$$\Phi_c: C^\infty(M; L) \to \mathrm{Vect}(P) \qquad \sum_{i=1}^{n} A_i \otimes f_i \to \sum_{i=1}^{n} (\pi^* f_i) Y_i$$

(where $\pi^* f_i$ is the pull-back of f_i) is a (faithful) representation of the current algebra $C^\infty(M; L)$ as a Lie algebra of vertical vector fields on P.

One can choose N to be the connected simply connected Lie group associated to L and $\Phi: L \to \mathrm{Vect}(N)$ the natural representation of L as the Lie algebra of left-invariant vector fields on N.

Specialize the above construction to $M = \Theta$ the parameter manifold and $L = \mathrm{st}(n)$ and restrict the representation Φ_c to the estimation algebra $\tilde{G}$ of the identification problem. In this way, we obtain Type II representations.

In what follows we construct a class of Type II representations that are intimately related to Kalman filtering and lead to the computation of conditional statistics. First we recall the following result of Brockett [11]: Once again treat θ as a constant and consider the linear filtering problem, with the associated Kalman–Bucy filter equations for state estimation:

$$dz_t = (A - Pcc^T)z_t\,dt + Pc\,dy_t \qquad \frac{dP}{dt} = AP + PA^T + bb^T - Pcc^TP. \tag{36}$$

For this system of equations, one obtains a pair of vector fields (on a manifold of dimension $n(n+1)/2+n$),

$$a_0 = \begin{bmatrix} (A-Pcc^T)z \\ AP+PA^T+bb^T-Pcc^TP \end{bmatrix}, \quad b_0 = \begin{bmatrix} Pc \\ 0 \end{bmatrix}.$$

Brockett showed that the Lie algebra of vector fields $\{a_0, b_0\}_{\text{L.A.}}$ is a homomorphic image of the estimation algebra of the filtering problem, the homomorphism being specified by

$$\mathscr{A}_{00} = \tfrac{1}{2}\langle b, \partial/\partial x\rangle^2 - \langle \partial/\partial x, Ax\rangle - \frac{\langle c, x\rangle^2}{2} \to a_0$$

$$\mathscr{B}_0 = \langle c, x\rangle \to b_0.$$

The homomorphism has a kernel consisting of the set of operators of multiplication by a constant. The kernel simply arises due to the fact that the Duncan–Mortensen–Zakai equation computes the *unnormalized* conditional density. To get rid of the kernel one should then append an equation to (36) for computing the normalization $\sigma_t(1)$. It can be verified that (for Gaussian initial conditions) the following does the job:

$$ds = \frac{\langle c, z\rangle^2}{2}\,dt - \langle c, z\rangle\,dy_t \qquad s_0 = 0 \tag{37}$$

$$\sigma_t(1) = e^{-s_t}. \tag{38}$$

Taking Eqs. (36) and (37) together, we now define a new pair of vector fields (on a space of dimension $n+n(n+1)/2+1$).

$$\tilde{a}_0 = \begin{bmatrix} (A-Pcc^T)z \\ AP+PA^T+bb^T-Pcc^TP \\ \dfrac{\langle c,z\rangle^2}{2} \end{bmatrix}; \quad \tilde{b}_0 = \begin{bmatrix} Pc \\ 0 \\ -\langle c,z\rangle \end{bmatrix} \tag{39}$$

THEOREM 4.2 *The Lie algebra of vector fields generated by $\tilde{a}_0$ and $\tilde{b}_0$ is isomorphic to the estimation algebra of the linear filtering problem, the isomorphism being given by,*

$$\mathscr{A}_{00} \to \tilde{a}_0; \mathscr{B}_0 \to b. \tag{40}$$

We now use the above result to produce a faithful representation of Type II for the estimation algebra $\tilde{G}$ of the identification problem. Treat θ now as a variable. Consider the system of embedding equations.

$$\begin{aligned} d\theta &= 0 \\ d\bar{z} &= (A(\theta)-\bar{P}c(\theta)c^T(\theta))\bar{z}.dt+\bar{P}c(\theta)dy_t \\ \frac{d\bar{P}}{dt} &= A(\theta)\bar{P}+\bar{P}A^T(\theta)+b(\theta)b^T(\theta)-\bar{P}c(\theta)c^T(\theta)\bar{P} \\ d\bar{s} &= \tfrac{1}{2}\langle c(\theta),\bar{z}\rangle^2 dt-\langle c(\theta),\bar{z}\rangle dy_t. \end{aligned} \tag{41}$$

The system of Eqs. (41) evolves on a manifold which looks locally like $\Theta \times \mathbb{R}^{n(n+1)/2+n+1}$. Associate with (41) a pair of vector fields (first order differential operators).

$$\begin{aligned} a_0^* = {} & \langle (A(\theta)-\bar{P}c(\theta)c^T(\theta))\bar{z}, \partial/\partial\bar{z}\rangle \\ & +\operatorname{tr}((A(\theta)\bar{P}+\bar{P}A^T(\theta)+b(\theta)b^T(\theta)-\bar{P}c(\theta)c^T(\theta)\bar{P}).\partial/\partial\bar{P}) \\ & +\tfrac{1}{2}\langle c(\theta),\bar{z}\rangle^2\frac{\partial}{\partial\bar{s}} \end{aligned} \tag{42}$$

and

$$b_0^* = \langle \bar{P}c(\theta), \partial/\partial \bar{z}\rangle - \langle c(\theta), \bar{z}\rangle \partial/\partial \bar{s}. \tag{43}$$

(Here $\partial/\partial \bar{P} = [\partial/\partial \bar{P}_{ij}] = (\partial/\partial \bar{P})^T = n \times n$ symmetric matrix of differential operators.) Consider the Lie algebra of vector fields generated by a_0^* and b_0^*. Since a_0^* and b_0^* are vertical vector fields with respect to the fibering $\Theta \times \mathbb{R}^{n(n+1)/2+n+1} \to \Theta$, so is every vector field in this Lie algebra. Using Theorem 4.2 and the nonoccurence of differential operators involving $\partial/\partial\theta$ we conclude:

Corollary 4.2 *The map*

$$\Phi_K : \tilde{G} \to \mathrm{Vect}(\Theta \times \mathbb{R}^{n+1+n(n+1)/2}),$$

defined by

$$\mathscr{A}_{00} \to a_0^* \qquad \mathscr{B}_0 \to b_0^*$$

is a faithful representation of the Lie algebra of the identification problem as a Lie algebra of vertical vector fields on a finite dimensional manifold fibered over Θ.

Remark 4.1 Detailed proof of Theorem 4.2 involve tedious Lie bracket calculations. These are only slight modifications of Brockett's calculations in [11] to take into account the normalization Eqs. (37) or (41) and hence are omitted. See also Hazewinkel [52].

Example 4.2 Consider again the model

$$dx_t = \theta dw_t; d\theta = 0 \qquad dy_t = x_t dt + dv_t$$

with

$$\mathscr{A}_{00} = \frac{\theta^2}{2}\frac{\partial^2}{\partial x^2} - \frac{x^2}{2} \quad \text{and} \quad \mathscr{B}_0 = x.$$

The estimation algebra $\tilde{G} = \{\mathscr{A}_{00}, \mathscr{B}_0\}_{\mathrm{L.A.}}$. The embedding Eqs. (41) take the form

$$d\theta = 0 \qquad d\bar{p} = (\theta^2 - \bar{p}^2)dt \qquad d\bar{z} = -\bar{p}\bar{z}dt + \bar{p}dy_t \qquad d\bar{s} = \bar{z}^2{}_2 dt - \bar{z}dy_t.$$

Then,

$$\Phi_K(\mathscr{A}_{00}) = a_0^* \qquad = (\theta^2 - \bar{p}^2)\frac{\partial}{\partial \bar{p}} + (-\bar{p}\bar{z})\frac{\partial}{\partial \bar{z}} + \frac{\bar{z}^2}{2}\frac{\partial}{\partial \bar{s}}$$

$$\Phi_K(\mathscr{B}_0) = b_0^* \qquad = \bar{p}\frac{\partial}{\partial \bar{z}} + (-\bar{z})\frac{\partial}{\partial \bar{s}}.$$

The induced map on Lie brackets is given by.,

$$\Phi_K\left(\theta^{2k}\frac{\partial}{\partial x}\right) = \theta^{2k}\frac{\partial}{\partial \bar{s}} \quad k = 0, 1, 2, \ldots$$

$$\Phi_K(\theta^{2k}x) = \theta^{2k}\left(\bar{p}\frac{\partial}{\partial \bar{z}} - \frac{\partial}{\partial \bar{s}}\right) \quad k = 1, 2, \ldots$$

$$\Phi_K(\theta^{2k}.1) = \theta^{2k}\partial/\partial \bar{s} \quad k = 1, 2, \ldots. \qquad \square$$

The embedding Eqs. (41) have the following statistical interpretation. Assume that the initial conditions for the conditional density take the form

$$\rho_0(x, \theta) = (2\pi \det(\Sigma(\theta)))^{-n/2} \cdot \exp\left\{-(x - \mu(\theta))^T \times \frac{\Sigma(\theta)^{-1}}{2}(x - \mu(\theta))\right\} \cdot Q_0(\theta) \tag{44}$$

where $\theta \to (\mu(\theta), \Sigma(\theta), Q_0(\theta)$ is a smooth map. $\Sigma(\theta)$ is a positive definite $n \times n$ matrix for every $\theta \in \Theta$ and $Q_0(\theta) > 0$. Suppose that the system of Eqs. (41) is initialized at

$$(\theta_0, x, P_0, s_0) = (\theta_0, \mu(\theta_0), \Sigma(\theta_0), -\ln Q_0(\theta_0)). \tag{45}$$

Append to the system (41) the output equation

$$\bar{Q}_t = e^{-\bar{s}_t}. \tag{46}$$

Now if Eq. (41) is solved with the initial condition (45), one can show by differentiating $\bar{Q}_t$ that $\bar{Q}_t$ satisfies Eq. (7). In other words, the system (41)–(46) is a finite-dimensional recursive estimator for the unnormalized posterior density $\bar{Q}_t = Q(t, \theta)$ *evaluated at* θ_0. We have thus verified the homomorphism principle of Brockett: that finite dimensional recursive estimators for conditional statistics must involve Lie algebras of vector fields that are homomorphic images of the Lie algebra of operators associated with the unnormalized conditional density equation. Notice that, although $Q(t, \theta_0)$ can be computed by a finite dimensional filter for every θ_0, the computation of the entire posterior density function $Q(t, .)$ appears to admit no finite-dimensional filters in general (unless Θ the parameter set is finite).

The Type II representations above for the estimation algebra $\tilde{G}$ once again have the feature that the vector field associated to an element $\phi \in \tilde{G}$ does not depend on the derivatives of ϕ. In other words, the representation of Corollary 4.2 is a trivial representation of the estimation algebra $\tilde{G}$. One can construct *nontrivial representations* of Type II by taking the same approach as we did with Type I representations—use Leibnitz extensions. This amounts to taking the *θ-sensitivity equations* associated to (41). Sensitivity equations are used in practical implementations of maximum likelihood identification algorithms [50, 51]. For dim $\Theta = 1$, we obtain,

$$d\theta = 0$$

$$dz = (A(\theta) - \bar{P}c(\theta)c^T(\theta))z\,dt + \bar{P}c(\theta)\,dy_t$$

$$\begin{aligned} dz_1 &= (A(\theta) - \bar{P}c(\theta)c^T(\theta))z_1\,dt \\ &\quad + \left(\frac{\partial A}{\partial \theta} - \bar{P}\frac{\partial}{\partial \theta}(c(\theta)c^T(\theta))\right)z.dt \\ &\quad - \bar{P}_1.c(\theta)c^T(\theta)z.dt + \bar{P}_1 c(\theta)\,dy_t + \bar{P}.\frac{\partial c}{\partial \theta}dy_t \end{aligned}$$

$$\frac{d\bar{P}}{dt} = A(\theta)\bar{P} + \bar{P}A^T(\theta) + b(\theta)b^T(\theta) - \bar{P}c(\theta)c^T(\theta)\bar{P}$$

$$\frac{d\bar{P}_1}{dt} = A(\theta)\bar{P}_1 + \frac{\partial A}{\partial\theta}\bar{P} + \bar{P}_1 A^T(\theta) + \bar{P}\frac{\partial A^T}{\partial\theta}$$

$$+ \frac{\partial}{\partial\theta}(b(\theta)b^T(\theta)) - \bar{P}_1 c(\theta)c^T(\theta)\bar{P}$$

$$- \bar{P}c(\theta)c^T(\theta)\bar{P}_1 - \bar{P}\frac{\partial}{\partial\theta}(c(\theta)c^T(\theta))\bar{P}$$

$$d\bar{s} = \tfrac{1}{2}\langle c(\theta), \bar{z}\rangle^2 dt - \langle c(\theta), \bar{z}\rangle dy_t$$

$$d\bar{s}_1 = \langle c(\theta), \bar{z}\rangle \left\{ \langle c(\theta), \bar{z}_1\rangle dt + \left\langle \frac{\partial c}{\partial\theta}, \bar{z}\right\rangle dt \right\}$$

$$- \langle c(\theta), \bar{z}_1\rangle dy_t - \left\langle \frac{\partial c}{\partial\theta}, \bar{z}\right\rangle dy_t. \tag{41.θ}$$

In Eq. (41.θ) z_1 and P_1 are to be interpreted as $\partial z/\partial\theta$ and $\partial P/\partial\theta$ respectively.

To the system (41.θ) we associate a pair of vector fields,

$$a_{01}^* = \langle (A(\theta) - \bar{P}c(\theta)c^T(\theta))\bar{z}, \partial/\partial\bar{z}\rangle$$

$$- \langle \bar{P}_1 c(\theta)c^T(\theta)\bar{z}, \partial/\partial\bar{z}_1\rangle$$

$$+ \Big\langle (A(\theta) - \bar{P}c(\theta)c^T(\theta))\bar{z}_1$$

$$+ \left(\frac{\partial A}{\partial\theta} - \bar{P}\frac{\partial}{\partial\theta}(c(\theta)c^T(\theta))\right)\bar{z}, \frac{\partial}{\partial z_1}\Big\rangle$$

$$+ \operatorname{tr}((A(\theta)\bar{P} + \bar{P}A^T(\theta) + b(\theta)b^T(\theta)$$

$$- \bar{P}c(\theta)c^T(\theta)\bar{P})\partial/\partial\bar{P})$$

$$+ \operatorname{tr}\Bigg(\Bigg(A(\theta)\bar{P}_1 + \frac{\partial A}{\partial\theta}\bar{P} + \bar{P}_1 A^T(\theta) + \bar{P}\frac{\partial A^T}{\partial\theta}$$

$$+ \partial/\partial\theta(b(\theta)b^T(\theta)) - \bar{P}_1 c(\theta)c^T(\theta)\bar{P}$$

$$-\bar{P}c(\theta)c^T(\theta)\bar{P}_1 - \bar{P}\frac{\partial}{\partial\theta}(c(\theta)c^T(\theta))\bar{P}\Big)\partial/\partial\bar{P}\Big)$$

$$+\tfrac{1}{2}\langle c(\theta), z\rangle^2 \partial/\partial \bar{s}$$

$$+\langle c(\theta), z\rangle\left\{\langle c(\theta), z_1\rangle + \left\langle\frac{\partial c}{\partial\theta}, z\right\rangle\right\}\partial/\partial\bar{s}_1$$

$$b_{01}^* = \langle\bar{P}c(\theta), \partial/\partial z\rangle + \langle\bar{P}_1 c(\theta), \partial/\partial z_1\rangle$$

$$+\left\langle\bar{P}\frac{\partial c}{\partial\theta}, \partial/\partial z_1\right\rangle - \langle c(\theta), z\rangle\partial/\partial\bar{s}$$

$$-\left\{\langle c(\theta), z_1\rangle + \left\langle\frac{\partial c}{\partial\theta}, z\right\rangle\right\}\partial/\partial\bar{s}_1.$$

The vectorfields a_{01}^* and b_{01}^* may be viewed as *partial prolongations* of the vector fields a_0^* and b_0^*.

Using essentially the same arguments as those following Theorem 4.1 we can show:

THEOREM 4.3 *The map*

$$\Phi_k^1 : \tilde{G} \to \mathrm{Vect}(\mathbb{R}^{2(n+1+n(n+1)/2)} \times \Theta)$$

defined by

$$\mathscr{A}_{00} \to a_{01}^* \qquad \mathscr{B}_0 \to b_{01}^*$$

extends to an isomorphism of Lie algebras. □

Higher order Leibnitz extensions Φ_k^r may be constructed in a similar manner.

In the next section we use Lie algebraic techniques for solving the basic initial-value problems arising in the identification problem.

5. ON INTEGRATING THE ESTIMATION ALGEBRA $\tilde{G}$

The primary Cauchy problem of interest in this paper is the one associated to Eqs. (9):

$$\frac{\partial \tilde{\rho}}{\partial t} = \left(\mathscr{L}_0 + y_t \mathscr{L}_1 + \frac{y_t^2}{2}\mathscr{L}_2\right)\tilde{\rho}$$
$$\tilde{\rho}(0, x, \theta) = \tilde{\rho}_0(x, \theta), \tag{47}$$

for a suitable class of prior joint probability densities $\tilde{\rho}_0$. The operators $\mathscr{L}_0$, $\mathscr{L}_1$ and $\mathscr{L}_2$ all lie in the estimation algebra $\tilde{G}$. By the integration problem for $\tilde{G}$, we mean the problem of constructing a fundamental solution to Cauchy problems of the type (47) above. In the situation where Θ is a 1-point manifold, (the linear filtering problem), Brockett observed that this is equivalent to the problem of constructing canonical coordinates of the second kind in a neighborhood of the identity in G, the finite dimensional, connected, simply connected Lie group associated to $\tilde{G}$. Hence it is natural to look for a (Wei–Norman) representation,

$$\rho(t, \cdot) = \exp(g_1(t)\mathscr{A}^1)\ldots\exp(g_\nu(t)\mathscr{A}^\nu)\rho_0$$

where $\mathscr{A}^1, \ldots, \mathscr{A}^\nu$ span $\tilde{G}$.

This approach to initial-value problems appears explicitly in the work of Wei and Norman [41, 42] and in the more recent paper of Steinberg [43]. In fact Steinberg's paper is very relevant to our problems since he devotes most attention to the Lie algebra st(n) of this paper. Brockett's paper [11] contains a nice exposition of this circle of ideas and applications to filtering. Also relevant is the thesis of Ocone [23] and the papers [18, 44].

We note first that solving (47) is equivalent to constructing the fundamental solution of,

$$\frac{\partial \rho}{\partial t} = (\mathscr{A}_0 - \mathscr{B}_0^2/2)\rho + \dot{y}_t \mathscr{B}_0 \rho$$
$$\rho(0, x) = \rho_0(x) \tag{49}$$

due to the relationship (gauge transformation) given by Eq. (8). Henceforth we treat only Cauchy problems of the type (49). Further, the coefficients $g_i(\cdot)$ in the Wei–Norman representation (48) are given by a system of ordinary differential equations obtained by substitution of the formula (48) in (49) and taking into account the relations (of Baker–Campbell–Hausdorff–Zassenhaus).

$$e^{t\mathscr{A}^i}\mathscr{A}^j=\left(\sum_{m=0}^{\infty}\frac{t^m}{m!}(ad\mathscr{A}_i)^m\mathscr{A}^j\right)e^{t\mathscr{A}^i}. \tag{50}$$

To illustrate,

Example 5.1 (Ocone [18, 23]) Let $\Theta=\{1\}$. Then a solution to the Cauchy problem,

$$\frac{\partial\rho}{\partial t}=\left(\frac{1}{2}\frac{\partial^2}{\partial x^2}-\frac{x^2}{2}\right)\rho+x\dot{y}_t\rho$$

$$\rho(0,x)=\rho_0(x)\qquad\rho_0\in L_2(\mathbb{R})$$

takes the form,

$$\rho(t,\cdot)=\exp(g_1(t)\mathscr{A}^1)\ldots\exp(g_4(t)\mathscr{A}^4)\rho_0$$

where,

$$\mathscr{A}^1=\frac{1}{2}\frac{\partial^2}{\partial x^2}-\frac{x^2}{2}\qquad\mathscr{A}^2=x\qquad\mathscr{A}^3=\frac{\partial}{\partial x}\qquad\mathscr{A}^4=1. \tag{51}$$

and the g_i's satisfy the Wei–Norman equations:

$$\dot{g}_1=1\qquad\dot{g}_2=\cosh(g_1)\dot{y}\qquad\dot{g}_3=-\sinh(g_1)\dot{y}\qquad\dot{g}_4=g_2\dot{g}_3 \tag{52}$$

and $g_i(0)=0$, $i=1,2,3,4$.

The system (52) may be solved by quadrature and $\{g_2,g_3\}$ constitutes the *joint-sufficient statistics for this linear filtering problem.* In particular, the inequality $g_1(t)=t\geqq 0$ for $t\geqq 0$ is compatible with the fact that $\mathscr{A}^1$ only generates a semigroup. In fact using the Mehler formula (see e.g. Davies [45], Ocone [18]) one can write down the expression:

$$\rho(t,x)=\int_{-\infty}^{\infty}\frac{1}{2\pi\sinh(t)}e^{-1/2\coth(x^2+z^2)t}e^{xz/\sinh(t)}$$

$$\cdot e^{g_4(t)}e^{g_2(t)z}\cdot\rho_0(g_3(t)+z)\,dz. \tag{53}$$

In order to ensure the validity of the basic formula (50) one needs a common, dense (in $L_2(\mathbb{R})$), invariant, set of analytic vectors for the Lie algebra spanned by $\mathscr{A}^1$, $\mathscr{A}^2$, $\mathscr{A}^3$ and $\mathscr{A}^4$. Such a set is constructed as the linear span of eigenvectors of the operator $\mathscr{A}^1$ (see Ocone [18]). □

The approach of Wei–Norman–Steinberg is originally set up for finite dimensional Lie algebras. However, it is now possible to extend it to infinite dimensional current algebras and their subalgebras. In the present context $\{\mathscr{A}^1,\mathscr{A}^2,\ldots,\mathscr{A}^\nu\}$ would be a basis for $\mathrm{st}(n)$ and the g_i's would be functions of time t as well as the coordinate θ on the underlying parameter manifold Θ. The functions g_i play the role of canonical coordinates of the second kind in a neighborhood of the identity on $C^\infty(\Theta;\mathrm{St}_+(n))$ or one of the associated Sobolev Lie groups. To illustrate, consider our favorite example.

Example 5.2 Let

$$\mathscr{A}^1=\frac{\theta^2}{2}\frac{\partial^2}{\partial x^2}-\frac{x^2}{2},\ \mathscr{A}^2=x,\ \mathscr{A}^3=\theta^2\frac{\partial}{\partial x}\quad\text{and}\quad\mathscr{A}^4=\theta^2\frac{\partial}{\partial x}.$$

Then, for the *Cauchy problem*,

$$\frac{\partial\rho}{\partial t}=(\mathscr{A}^1+\dot{y}_t\mathscr{A}^2)\rho$$

$$\rho(0,x,\theta)=\rho_0=\rho_0(x,\theta), \tag{54}$$

since the associated estimation algebra is spanned by the set of operators

$$\left\{\frac{\theta^2}{2}\frac{\partial^2}{\partial x^2}-\frac{x^2}{2},\ \theta^{2n}x,\ \theta^{2n+2}\frac{\partial}{\partial x},\ \theta^{2n+2}1;\quad n=0,1,2,\ldots\right\}.$$

we seek a representation of the form.

$$\rho(t,x,\theta)=\exp(g_1(t,\theta)\mathscr{A}^1)\exp\left(\sum_{k=0}^{\infty} g_2^k(t)\theta^{2k}\mathscr{A}^2\right)$$

$$\cdot\exp\left(\sum_{k=0}^{\infty} g_3^k(t)\theta^{2k}\mathscr{A}^3\right)\cdot\exp\left(\sum_{k=0}^{\infty} g_4^k(t)\theta^{2k}\mathscr{A}^4\right)\rho_0. \qquad (56)$$

Equivalently, we look for a representation of the form

$$\rho(t,x,\theta)=\exp(g_1(t,\theta)\mathscr{A}^1)\cdot\exp(g_2(t,\theta)\mathscr{A}^2)$$

$$\cdot\exp(g_3(t,\theta)\mathscr{A}^3)\cdot\exp(g_4(t,\theta)\mathscr{A}^4)\rho_0. \qquad (57)$$

In Eq. (57) the g_i's are to determined by substitution into (54). This step yields a system of first order partial differential equations.

$$\begin{aligned}
\frac{\partial g_1}{\partial t}(t,\theta)&=1\\
\frac{\partial g_2}{\partial t}(t,\theta)&=\cosh(g_1\theta)\dot{y}_t\\
\frac{\partial g_3}{\partial t}(t,\theta)&=-\frac{1}{\theta}\sinh(g_1\theta)\dot{y}_t\\
\frac{\partial g_4}{\partial t}(t,\theta)&=\frac{\partial g_3}{\partial t}(t,\theta)\cdot g_2(t,\theta),
\end{aligned} \qquad (58)$$

and $g_i(0,\theta)=0$ for $i=1, 2, 3, 4$ and $\theta\in\Theta$. Now suppose that Θ is a bounded set and $0\notin$ closure (Θ). Then using (58) our Cauchy problem may be explicitly solved and using a scaled version of the

Mehler formula, we have a representation,

$$\rho(t,x,\theta)=\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi\sinh(|\theta|t)}}$$
$$\cdot\exp\left(-\tfrac{1}{2}\coth^2\left(\frac{|x|^2}{|\theta|}+z\right)\cdot t|\theta|\right)$$
$$\cdot\exp\left(\frac{xz}{(\sqrt{|\theta|\sinh(|\theta|t)}}\right) \tag{59}$$
$$\cdot\exp(g_4(t,\theta)\theta^2)\cdot\exp(g_2(t,\theta)\cdot\sqrt{|\theta|}z)$$
$$\cdot\rho_0(g_3(t,\theta)\theta^2\sqrt{|\theta|}z,\theta)dz$$

where,

$$\rho_0\in\Theta\times L_2(\Theta\times\mathbb{R}). \qquad \square$$

Our purpose in following through this exercise is to illustrate that in the identification problem, even though the estimation algebra is infinite dimensional it is possible to solve the integration problem in a manner that is a natural generalization of the Wei–Norman–Steinberg technique. This happens precisely because the estimation algebra is a current algebra and the construction of canonical coordinates of the second kind is rigorously justified by the results of section 3.

Obtaining explicit formulas analogous to (59) for general linear system identification problems with many state variables is another matter. The essential complexity is in obtaining analogues of the Mehler formula and constructing analytic vectors. The details are extremely tedious, but the calculations of Steinberg [43], (see pages 418–423 of his paper especially Theorem 7.5) using Lie-transforms show the main steps.

6. ON SUFFICIENT STATISTICS FOR THE IDENTIFICATION PROBLEM

In his paper [46], Giorgio Picci investigated the relationships between the problem of identifiability of a linear system driven by deterministic inputs with white-noise corrupted measurements. In particular, he also isolated the *maximally-identifiable parameters* and *minimal sufficient statistics* associated with this problem. The former are precisely the Markov parameters $\{c(\theta)\mathscr{A}^k(\theta)b(\theta), k=0,1,2,\ldots,2n-1\}$ and the latter are simply a sequence of input dependent functionals of the observations that completely determine the likelihood-ratio [47, 48].

Our setting of the identification problem as a nonlinear-filtering problem differs from Picci's in three essential ways:

a) we treat here the joint state and parameter identification problem;

b) the fundamental solution associated to the Cauchy problem (49) and not the likelihood ratio contains the full information;

c) the input is white noise and not a known deterministic function of time.

Consider the functions g_2 and g_3 of Example 5.2 determined by the differential Eqs. (58). Expand the solutions to these equations:

$$g_2(t,\theta)=\sum_{k=0}^{\infty}\theta^{2k}\int_0^t\frac{\sigma^{2k}}{(2k)!}\dot{y}_\sigma d\sigma$$
$$g_3(t,\theta)=-\sum_{k=0}^{\infty}\theta^{2k}\int_0^t\frac{\sigma^{2k+1}}{(2k+1)!}\dot{y}_\sigma d\sigma. \tag{60}$$

It follows that all the information contained in the observations $\{y_\sigma: 0\leqq\sigma\leqq t\}$ concerning the joint unnormalized conditional density $\rho(t,x,\theta)$ is contained in the sequence,

$$T\triangleq\left\{\int_0^t\frac{\sigma^k}{k!}\dot{y}_\sigma d\sigma; k=0,1,2,\ldots\right\}. \tag{61}$$

Hence the sequence T is a joint sufficient statistic for the identification problem (Example 5.2) viewed as a nonlinear-filtering problem.

There is some evidence to believe that the sequence T (or some variation of it) is *universal* in the sense that it does not depend on the underlying parameterization or state space dimension of the problem. This is borne out in the driving-noise free case by our results [15]. Making this precise would entail elaborate calculations of the type mentioned at the end of section 5.

Our calculations show the difficulty in explicitly computing the conditional density. The statistics themselves are generated by an infinite dimensional bilinear system.

7. CONCLUSIONS

In this paper, we have examined the structure of the estimation Lie algebra $\tilde{G}$ of the identification problem. This Lie algebra and its representations arise in the computation of the conditional density and in the study of finite dimensional recursive filters for this problem. It is shown that $\tilde{G}$ is a solvable, infinite dimensional subalgebra of the current algebra $C^\infty(\Theta; \mathrm{st}(n))$. Although $\tilde{G}$ is infinite dimensional we are able to associate with it a family of Sobolev Lie groups $\{G_k^\Theta\}$. The conditional density is computed by solving the Cauchy problem (47) or (49) for G_k^Θ, which is related to finding canonical coordinates of the second kind in a neighborhood of the identity. We indicate how this is done.

Motivated by the search for finite dimensional recursive filters, we have constructed representations of the current algebra $C^\infty(\Theta; \mathrm{st}(n))$ and restricted these to the subalgebra $\tilde{G}$. Since the Kalman–Bucy filter solves the filtering problem for known θ, it is reasonable to attempt a Taylor expansion of the unnormalized conditional density $\tilde{\rho}(t, x, \theta)$ about a known $\hat{\theta}_0$ or a current estimate $\hat{\theta}_t$[50]. 52]: the equations for computing $\tilde{\rho}$ and $\partial\tilde{\rho}/\partial\theta$ yield a realization of nontrivial Type I representation of $\tilde{G}$. On the other hand, a realization of a Type II representation of $\tilde{G}$ is given by the finite dimensional filter that computes the unnormalized a posterior density $Q(t, \theta)$ evaluated at a point. Further nontrivial Type II representations are realized by augmenting these filters with the derivatives with respect to θ of the states of these filters; these may also be useful in approximations and in other computations [50]. Finally, the inherent difficulty and structure of the joint state and parameter estimation problem has been emphasized by providing an infinite set of finite dimensionally computable sufficient statistics for the conditional density.

Given the structure of the identification problem developed in this paper, the key remaining question is that of translating some of the structural properties into actual approximation methods for the identification problem. For example, Taylor series expansions of the unnormalized conditional density and some type of truncation of the sufficient statistics of section 6 should be pursued.

Acknowledgement

We thank our colleagues Roger Brockett, Sanjoy Mitter, John Baras and Gilmer Blankenship for kind support and encouragement in pursuing this line of work. The influence of the work of Robert Hermann on current algebras as they appear in physics should be apparent. One of us (P.S.K.) would also like to acknowledge the generous hospitality of Erasmus University during a recent visit at the invitation of Michiel Hazewinkel.

Partial support for this work was provided by the Air Force Office of Scientific Research under grant AFOSR79-0025, by the Department of Energy under Contract DEAC01-80-RA50420-A001 and by National Science Foundation under Grant ECS-81-18138.

References

[1] M. H. A. Davis and S. I. Marcus, An introduction to nonlinear filtering. In M. Hazewinkel and J. C. Willems (Ed.), *Stochastic Systems: The Mathematics of Filtering and Identification and Applications*, Reidel, Dordrecht, 1981.

[2] S. K. Mitter, On the analogy between mathematical problems of nonlinear filtering and quantum physics, *Richerche di Automatica* **10** (1980), 163–216.

[3] G. Kallianpur, *Stochastic Filtering Theory*, Springer-Verlag, New York, 1980.

[4] P. S. Krishnaprasad and S. I. Marcus, Identification and tracking: a class of nonlinear filtering problems. In Proc. 1981 *Joint Automatic Control Conference*, Charlottesville, 1981.

[5] S. K. Mitter, Nonlinear filtering and stochastic mechanics. In M. Hazewinkel and J. C. Willems (Ed.), *Stochastic Systems: The Mathematics of Filtering and Identification and Applications*, Reidel, Dordrecht, 1981.

[6] M. H. A. Davis, Pathwise nonlinear filtering. In M. Hazewinkel and J. C. Willems (Ed.), *Stochastic Systems: The Mathematics of Filtering and Identification and Applications*, Reidel, Dordrecht, 1981.

[7] R. W. Brockett and J. M. C. Clark, The geometry of the conditional density equation. In O. L. R. Jacobs (Ed.), *Analysis and Optimization of Stochastic Systems*, Academic Press, New York, 1978, pp. 299–309.

[8] R. W. Brockett, Remarks on finite dimensional nonlinear estimation. Lobry (Ed.), *Analyse des Systems*, Bordeaux (1980), Asterisque 75, 76, 1978.

[9] R. W. Brockett, Classification and equivalence in estimation theory. In *Proc. 18th IEEE Conf. on Decision and Control*, Ft. Lauderdale, 1979, pp. 172–174.

[10] R. W. Brockett, Estimation theory and the representation of Lie algebras. In *Proc. 19th Conf. on Decision and Control*, Albuquerque, 1980.

[11] R. W. Brockett, Nonlinear systems and nonlinear estimation theory. In M. Hazewinkel and J. C. Willems (Ed.), *Stochastic Systems: The Mathematics of Filtering and Identification and Applications*, Reidel, Dordrecht, 1981.

[12] S. K. Mitter, Modeling for stochastic systems and quantum fields. In *Proc. 17th IEEE Conf. on Decision and Control*, San Diego, 1978.

[13] M. Hazewinkel and S. I. Marcus, On Lie algebras and finite-dimensional filtering, submitted to *Stochastics*.

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

[15] P. S. Krishnaprasad and S. I. Marcus, Some nonlinear filtering problems arising in recursive identification. In M. Hazewinkel and J. C. Willems (Ed.), *Stochastic Systems: The Mathematics of Filtering and Identification and Applications*, Reidel, Dordrecht, 1981.

[16] A. A. Kirillov, *Elements of the Theory of Representations*, Springer-Verlag, Berlin, 1976.

[17] G. Lion and M. Vergne, *The Weil Representation, Maslov Index and Theta Series*, Birkhauser, Boston, 1980.

[18] D. Ocone, Nonlinear filtering problems with finite dimensional Lie algebras. *In Proc. 1980 Joint Automatic Control Conference*, San Francisco, 1980.

[19] R. F. Streater, Current commutation relations, continuous tensor products and infinitely divisible group representations, *Rend Sculoa Internazionale Fis. E. Fermi*, **11** (1969), 247–263.

[20 M. Daniel and C. M. Viallet, The geometrical setting of gauge theories of the Yang–Mills type, *Reviews of Modern Physics* **52** (1980).

[21] A. A. Kirillov, Local Lie algebras, *Usp. Mat. Nauk* **81** (1976), 57–76.

[22] R. Hermann, *Topics in the Mathematics of Quantum Mechanics*: Interdisciplinary Mathematics Vol. VI, Math. Sci. Press, Brookline, 1973.

[23] D. Ocone, *Topics in Nonlinear Filtering Theory*, Ph.D. Thesis, M.I.T., Cambridge, 1980.

[24] P. S. Krishnaprasad and S. I. Marcus, On the Lie algebra of the identification problem. *In Proc. IFAC Symposium on Digital Control*, New Dehli, 1982.

[25] I. E. Segal, The cauchy problem for the Yang–Mills equations, *Journal of Functional Analysis* **33** (1979), 175–194.

[26] R. S. Palais, *Seminar on the Atiyah–Singer Index Theorem*, Annals of Mathematics Studies No. 57, Princeton University Press, Princeton, 1965.

[27] R. S. Palais, *Foundations of Global Nonlinear Analysis*, W. A. Benjamin, New York, 1968.

[28] J. Eels, A setting for global analysis, *Bulletin of the American Math. Soc.* **72** (1966), 751–807.

[29] J. Marsden, *Applications of Global Analysis in Mathematical Physics*, Publish or Perish, Boston, 1974.

[30] D. Ebin, The manifold of Riemannian Metrics. *In Proc. Symp. Pure Math., XI*, American Math. Soc., Providence, 1970.

[31] M. S. Narasimhan and T. R. Ramadas, Geometry of SU(2) gauge fields, *Commun. Math. Phys.* **67** (1979), 121–136.

[32] P. K. Mitter and C. M. Viallet, On the bundle of connections and the gauge orbit manifold in Yang–Mills theory, *Commun. Math. Phys.* **79** (1981), 457–472.

[33] P. K. Mitter, Geometry of the space of gauge orbits and the Yang Mills dynamical system. In G. T. Hooft et al. (ed.) *Recent Developments in Gauge Theories* (Cargese School), Plenum Press, New York, 1980.

[34] A. A. Kirillov, Characters of unitary representations of Lie groups. Reduction Theorems, *Funkcional. Anal. i Prilozen* **3** No. 1 (1969), 36–37.

[35] R. Hermann, *Lie Algebras and Quantum Mechanics*, W. A. Benjamin, New York, 1970.

[36] R. Hermann, Infinite dimensional Lie algebras and current algebra. In V. Bargmann (Ed.), *Group Representations in Physics Batelle Rencontres*, (Lectures Notes in Physics Vol. 6). Springer-Verlag, New York, 1969.

[37] R. Hermann, Current algebras, the Sugawara model and differential geometry, *J. Math. Phys.* **11, No. 6** (1970), 1825–1829.

[38] K. R. Parthasarathy and K. Schmidt, A new method for constructing factorizable representations for current groups and current algebras. *Commun. Math. Phys.* **50** (1976), 167–175.

[39] C.-H. Liu and S. I. Marcus, The Lie algebraic structure of a class of finite dimensional filters. In C. I. Byrnes and C. F. Martin (Ed.), *Lectures in Applied Mathematics, 18*, American Mathematical Society, Providence, 1980.

[40] V. E. Benes, Exact finite dimensional filters for certain diffusions with nonlinear drift, *Stochastics* **5** (1981), 65–92.

[41] J. Wei and E. Norman, On the global representation of solutions of linear differential equations as a product of exponentials, *Proc. Amer. Math. Soc.* April (1964), 327–334.

[42] J. Wei and E. Norman, On the Lie algebraic solution of linear differential equations, *J. Math. Phys.* **4** No. 4 (1963), 575–581.

[43] S. Steinberg, Application of the Lie algebraic formulas of Baker, Campbell, Hausdorff, and Zassenhaus to the calculation of explicit solutions of partial differential equations. *Journal of Differential Equations* **26** (1977), 404–434.

[44] D. Ocone, Finite dimensional estimation algebras in nonlinear filtering. In M. Hazewinkel and J. C. Willems (Ed.), *Stochastic Systems: The Mathematics of Filtering and Identification and Applications*, Reidel, Dordracht, (1981).

[45] E. B. Davies, *One-Parameter Semigroups*, Academic Press, London, 1980.

[46] G. Picci, Some connections between the theory of sufficient statistics and the identifiability problem. *SIAM J. Appl. Math.* **33** No. 3 (1977), 383–398.

[47] T. Kadota and L. A. Shepp, Conditions for absolute continuity between a certain pair of probability measures. *Z. Wahrscheinlichkeitstheorie und Verw. Giebiete* **16** (1970), 250–260.

[48] T. Kailath and M. Zakai, Absolute continuity and Radon-Nikodym derivatives between certain probability measures, *Ann. Math. Statistics* **42**, (1971), 130–140.

[49] O. Hijab, The adaptive LQG problem—part I, *IEEE Trans. Aut. Contr., Vol. AC-28, No. 2*, 1983, pp. 171–179.

[50] R. K. Mehra, Some recent advances in system identification. *Proc. 4th IFAC Symposium on Identification*, Tbilisi, UUSR, 1976.

[51] T. Soderstrom, A comparative study of recursive identification methods. *Report 7427, Lund Institute of Technology*, 1974.

[52] P. S. Krishnaprasad, M. Hazewinkel and B. Hanzon, Approximation methods for nonlinear filtering problems arising in system identification. *Proc. IEEE Conf. Decision and Control*, IEEE, NY, 1981.

[53] M. Hazewinkel, The linear systems Lie algebra, the Segal–Shale–Weil representation and all Kalman filters. *Report 81, Econometric Institute, Erasmus University, Rotterdam, The Netherlands*, 1981.

Appendix 1

Consider the algebra $A_n(\mathbb{R})$ with unit 1 over the reals generated by the elements $p_1, \ldots, p_n, q_1, \ldots, q_n$ and governed by the relations.

$$\begin{gathered} p_i p_j = p_j p_i ; q_i q_j = q_j q_i \\ p_i q_j - q_j p_i = \delta^i_j 1. \end{gathered} \tag{A.1}$$

The algebra $A_n(\mathbb{R})$ is known as the *Weyl algebra*. By defining the bracket of any two elements x, y in the Weyl algebra to be

$$[x, y] = xy - yx,$$

one introduces the structure of an infinite-dimensional Lie algebra over $A_n(\mathbb{R})$. The set of polynomials of total degree ≤ 2 in the variables $p_1, \ldots, p_n, q_1, \ldots, q_n$ is a Lie subalgebra of $A_n(\mathbb{R})$. This subalgebra is isomorphic to $\mathrm{st}(n)$ (see section 2). For example, with $n = 1$ this isomorphism is given by,

$$\begin{vmatrix} 0 & 0 & 0 & 0 \\ 0 & a & b & 0 \\ 0 & c & -a & 0 \\ 0 & 0 & 0 & 0 \end{vmatrix} \leftrightarrow \tfrac{1}{2}[a(pq+qp) + bq^2 - cp^2] \tag{A.2}$$

and

$$\begin{vmatrix} 0 & x & y & z \\ 0 & 0 & 0 & y \\ 0 & 0 & 0 & -x \\ 0 & 0 & 0 & 0 \end{vmatrix} \leftrightarrow 2(xp + yq) + z \cdot 1. \tag{A.3}$$

Now as mentioned in section 2, $\mathrm{st}(n) = \mathrm{sp}(2n) \oplus_s h(n)$ where $h(n)$ is the Heisenberg algebra and $\mathrm{sp}(2n)$ is the symplectic algebra. Let $H(n)$

denote the connected simply connected Lie group associated to $h(n)$. The group $H(n)$ admits a series of irreducible unitary representations U_λ, $\lambda \in \mathbb{R} - \{0\}$ acting on $L_2(\mathbb{R}^n)$ which is given infinitesimally by the assignments

$$\begin{aligned} p_j &\leftrightarrow \frac{\partial}{\partial x_j} \\ q_j &\leftrightarrow \sqrt{-1}\,\lambda x_j \\ 1 &\leftrightarrow \sqrt{-1}\,\lambda. \end{aligned} \tag{A.4}$$

We can extend the representation (A.4) of $h(n)$ to a representation of $\mathrm{st}(n)$ by the additional assignments

$$\begin{aligned} p_j p_k &\leftrightarrow \frac{1}{\sqrt{-1}\,\lambda} \frac{\partial^2}{\partial x_j \cdot \partial x_k} \\ q_j p_k &\leftrightarrow x_j \frac{\partial}{\partial x_k} \\ q_j q_k &\leftrightarrow \sqrt{-1}\,\lambda x_j x_k. \end{aligned} \tag{A.5}$$

The representation of $\mathrm{st}(n)$ given by (A.4) and (A.5) integrates (by Nelson's criterion) to give a unitary representation T_λ of $\widetilde{\mathrm{St}}(n)$ the connected simply connected Lie group associated to $\mathrm{st}(n)$. We call this the spinor representation. The representation $T_{-\sqrt{-1}}$ of the Lie algebra $\mathrm{st}(n)$ of section 2 is the analytic continuation of the representation T_1. Of course it does not integrate to give a group representation since $\partial^2/\partial x^2$ generates only a semigroup. (See the work of Kirillov [16, 34] for more details.)