The linear systems Lie algebra, the Segal-Shale-Weil representation and all Kalman-Bucy filters

by

Michiel Hazewinkel
Centre for Mathematics and Computer Science
P.O. Box 4079, 1009 AB Amsterdam.

Abstract

Let ls_n be the real Lie algebra of all differential operators in n-variables $\sum c_{\alpha\beta}x^\alpha\partial^\beta/\partial x^\beta$, $c_{\alpha\beta}\in\mathbf{R}$ where the sum is over all multi-indices α, β such that $|\alpha|+|\beta|\leq 2$. This note describes a certain representation of ls_n by means of (nonlinear) vectorfields which in a certain sense is all Kalman-Bucy filters for $n$dimensional linear systems put together. This representation also turns out to be very closely related to the so-called Segal-Shale-Weil representation of the simple quotient sp_n of ls_n.

1. INTRODUCTION

Let ls_n be the Lie algebra of all differential operators in n variables $\sum c_{\alpha\beta}x^\alpha\partial^\beta/\partial x^\beta$, $c_{\alpha\beta}\in\mathbf{R}$, with $c_{\alpha\beta}=0$ if $|\alpha|+|\beta|>2$. Here $\alpha=(\alpha_1,\ldots,\alpha_n)$, $\alpha_i\in\mathbf{N}\cup\{0\}$ is a multiindex and $|\alpha|=\alpha_1+\cdots+\alpha_n$. The Lie bracket is the commutator difference $[D_1,D_2]=D_1D_2-D_2D_1]$.

Thus for example ls_1 has the basis 1, x, d/dx, $x\,d/dx$, d^2/dx^2, d^2/dx^2, x^2 and two examples of brackets are $[\frac{d}{dx},x]=1$, $[\frac{d^2}{dx^2},x^2]=4x\frac{d}{dx}+2$ as is easily checked by letting the left and right hand operators act on a test function $f(x)$. The dimension of ls_n is $2n^2+3n+1$.

Let $V(\mathbf{R}^N)$ be the Lie algebra of all smooth vectorfields on $\mathbf{R}^N$, that is the Lie algebra of all differential operators $\sum_{i=1}^{N}f_i(x)\frac{\partial}{\partial x_i}$ with $f_i(x)$ a smooth function of $x=(x_1,\ldots,x_N)$. In this note I describe a representation of ls_n in $V(\mathbf{R}^N)$, $N=\frac{1}{2}n(n+1)+n+1$, i.e. a homomorphism of Lie algebras $ls_n\to V(\mathbf{R}^N)$, which is essentially all Kalman-Bucy filters for n-dimensional linear differential systems put together, making this representation, so to speak, the universal grand Kalman-Bucy filter. Below in section 4 it is explained how this phrase must be interpreted.

Let h_n be the subspace of ls_n spanned by the operators $1,x_1,\ldots,x_n,\frac{\partial}{\partial x_1},\ldots,\frac{\partial}{\partial x_n}$. One easily checks that h_n is an ideal of ls_n and the quotient ls_n/h_n turns out to be the symplectic Lie algebra sp_n. There is a very famous (and somewhat mysterious) representation of sp_n called the Segal-Shale-Weil representation (or oscillator representation [7]), which is of importance in number theory [14],quantum mechanics [12,13], harmonic analysis and representation theory [8,11], Lagrangian mechanics [9]. There is a subalgebra of ls_n isomorphic to sp_n (the Levi-factor) and it now turns out that the representation obtained from the Kalman-Bucy filter representation by mapping $(P,m,x)\in\mathbf{R}^N$ (P a symmetric $n\times n$ matrix, m an n vector and c a scalar) to the unnormalized "density" $\exp(c+<2\pi im,x>-2\pi_2P(x))$ (where $P(x)$ is the quadratic form defined by the matrix P, is a real form of the Segal-Shale-Weil representation, i.e. the two become isomorphic after tensoring with $\mathbf{C}$. This strengthens and precizes the links between filtering and quantum mechanics as discussed in [10].

This note is a drastically shortened version of [4] of the same title, giving just the basic outlines.

2. STRUCTURE OF THE LINEAR SYSTEMS LIE ALGEBRA ls_n

The Lie-algebras ls_n and h_n were defined in Section 1 above. The symplectic Lie algebra sp_n consists by definition of all $2n\times 2n$ matrices M such that $MJ+JM^T=0$ where J is the matrix $J=\begin{pmatrix}0 & I\\ -I & 0\end{pmatrix}$. (The Lie bracket in sp_n is the commutator difference $[M,M'] = MM' \cdot M'M$. Let $E_{i,j}$ be the $2n \times 2n$ matrix with a 1 at location (i,j) and zeros elsewhere. Then

$$x_i x_j \to E_{i,n+j} + E_{j,n+i}, \quad x_i \frac{\partial}{\partial x_j} \to E_{i,j} - E_{n+j,n+i}$$

$$\frac{\partial^2}{\partial s_i \partial x_j} \to E_{n+i,j} + E_{n+j,i}, \quad h_n \to 0$$

is a surjective homomorphism of Lie algebras $ls_n \to sp_n$ with kernel h_n. An isomorphic copy of sp_n in ls_n (a Levi factor) is spanned by the operators

$$x_i x_j, \quad \frac{\partial^2}{\partial x_i \partial x_j}, \quad x_i \frac{\partial}{\partial x_j} + \frac{1}{2}\delta_{i,j}$$

where $\delta_{i,j}$ denotes the Kronecker delta.

3. THE FILTER ANTI-REPRESENTATION OF ls_n

Let $N = 1/2\, n(n+1) + n + 1$ and denote a point in $\mathbf{R}^N$ as a triple $(P,m,c,)$ with P a symmetric $n \times n$ matrix, m an n-vector and c a scalar. Consider the Lie algebra $V(\mathbf{R}^N)$ of smooth vectorfields as $\mathbf{R}^N$ (cf. Section 1) and consider the homomorphism of real vector spaces

$$k : ls_n \to V(\mathbf{R}^N)$$

defined by the formulas

$$1 \to \frac{\partial}{\partial c} \tag{3.1}$$

$$x_i \to m_i \frac{\partial}{\partial c} + \sum_{t=1}^{n} P_{it} \frac{\partial}{\partial m_t} \tag{3.2}$$

$$\frac{\partial}{\partial x_i} \to -\frac{\partial}{\partial m_i} \tag{3.3}$$

$$x_i x_j \to (m_i m_j + P_{ij}) \frac{\partial}{\partial c} + \sum_t (m_i P_{jt} + m_j P_{it}) \frac{\partial}{\partial m_t} \tag{3.4}$$

$$+ \sum_{s,t} P_{is} P_{jt} \frac{\partial}{\partial P_{st}} + \sum_t P_{it} P_{jt} \frac{\partial}{\partial P_{tt}}$$

$$x_i \frac{\partial}{\partial x_j} \to -m_i \frac{\partial}{\partial m_j} \delta_{ij} \frac{\partial}{\partial c} P_{ij} \frac{\partial}{\partial P_{jj}} \sum_t P_{it} \frac{\partial}{\partial P_{jt}} \tag{3.5}$$

$$\frac{\partial^2}{\partial x_i \partial x_j} \to \frac{\partial}{\partial P_{ij}} \text{ if } i \neq j, \quad \frac{\partial^2}{\partial x_i^2} \to 2 \frac{\partial}{\partial P_{ii}} \tag{3.6}$$

Theorem. The vector space homomorphism $k : ls_n \to V(\mathbf{R}^N)$ defined by the formulae (3.1)-(3.6) is an injective anti-homomorphism of Lie algebras; i.e. it satisfies $k[D,D'] = [k(D'), k(D)]$ for all $D, D' \in ls_n$.

By changing the minus sign in formulae (3.3) and (3.5) to a plus sign one finds a homomorphism of Lie algebras, i.e. a representation. And by replacing all terms in (3.2)-(3.6) involving a $\frac{\partial}{\partial c}$ with zero one obtains an antihomomorphism

$$k' : ls_n \to V(\mathbf{R}^{N-1})$$

with kernel $\mathbf{R}.1$ (i.e. only multiples of the identity operator are mapped to zero). It is this last anti-homomorphism which is all Kalman-Bucy filters put together. (The $\frac{\partial}{\partial c}$-terms have to do with normalization.)

4. KALMAN-BUCY FILTERS AND THE ANTI-REPRESENTATION k'

Now consider a linear dynamical system driven by white noise

$$dx_t = Ax_t + Bdw_t, \quad dy_t = Cx_t\,dt + dv_t, \tag{4.1}$$

$$x_t \in \mathbf{R}^n,\ u_t \in \mathbf{R}^m,\ v_t \in \mathbf{R}^p,\ y_t \in \mathbf{R}^p$$

The Kalman-Bucy filter for such a system is a set of equations

$$\binom{dm}{dP} = \alpha(m,P)dt + \beta_1(m,P)dy_{1t} + \cdots + \beta_p(m,P)dy_{p,t} \tag{4.2}$$

For instance in one of the simplest cases:

$$dx_t = du_t, \quad dy_t = x_t dt + dv_t$$

one has

$$dP = (1-P^2)dt, \quad dm = P(dy_t - mdt)$$

so that the vector fields $\alpha(m,P)$ and $\beta(m,P)$ are respectively equal to

$$\alpha(m,P) = (1-P^2)\frac{\partial}{\partial P} - mP\frac{\partial}{\partial m}$$

$$\beta(m,P) = P\frac{\partial}{\partial m}$$

The relation between the Kalman-Bucy filter of a system (4.1) and the anti-representation k' is now as follows. Consider the Duncan-Mortensen-Zakai equation of (4.1), which is satisfied by an unnormalized version $\rho(x,t)$ of the density of $x_t = E[x_t | y_s, 0 \leq s \leq t]$:

$$d\rho = L\rho dt + \sum_{i=1}^{p} (Cx)_i dy_{it}$$

Here L is the second order differential operator

$$L = \frac{1}{2}\sum (BB^T)_{i,j}\frac{\partial^2}{\partial x_i \partial x_j} - \sum \frac{\partial}{\partial x_i}(Ax)_i - \frac{1}{2}\sum (Cx)_i^2 \tag{4.3}$$

(Here $(Cx)_i$ is the i-th component of the vector Cx and $(BB^T)_{i,j}$ is the i,j entry of the product of the matrix B with its transpose.)

Theorem The α vectorfield of the Kalman-Bucy filter (4.2) of the system 4.1 is given by $k'(L)$ where L is given by (4.3) and th β vector fields in (4.2) are equal to the $k'((Cx)_i)$.

This is essentially proved by the straightforward calculation [4]. This result also establishes for linear systems the (anti-)homomorphism principle of filtering, a powerful principle due to Brockett and Clark [1]. The proof in [2] of this principle for single-input single-output systems is wrong.

5. THE SEGAL-SHALE-WEIL REPRESENTATION

One way to obtain this representation is via the Stone-Von Neumann uniqueness theorem. Let H_n be the Heisenberg group $\mathbf{R}^n \times \mathbf{R}^n \times S^1$, where S^1 is the unit circle in $\mathbf{C}$ with the multiplication $(x,y,z)(x',y',z') = (x+x', y+y', zz' \exp(-2\pi i < x,y >))$. The Lie-algebra of H_n is $h_n = \mathbf{R}^n \times \mathbf{R}^n \times \mathbf{R}$. The Lie bracket of h_n restricted to $\mathbf{R}^n \times \mathbf{R}^n \times \{0\}$ defines a bilinear form on $\mathbf{R}^n$ given by the matrix J of the first few lines of section 2 above.

The symplectic group SP_n is the group of all $2n \times 2n$ matrices S such that $MJM^T = J$ so that we can view SP_n as a group of automorphisms of h_n and H_n which moreover leaves the centre $S^1 \subset H_n$ invariant. Now let $\rho: H_n \to L^2(\mathbf{R}^n)$ be the Schrödinger representation of H_n. Let $g \in SP_n$, then $h \to \rho(g(h))$ is also a unitary representation of H_n which by (the Weyl form of) the Stone-von Neumann uniqueness theorem is unitarily equivalent to ρ. This associates a unitary operator $w(g)$ to each $g \in SP_n$ which is unique up to a scalar. It turns out that the scalars can be fixed up to so as to define a unitary representation of the universal 2-fold covering $\tilde{SP}_n$ of sp_n.

There is a partial description of this Segal-Shale-Weil representation as follows (cf. e.g. [6]). The elements

$$\begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}, \begin{pmatrix} A & 0 \\ 0 & (A^{-1})^T \end{pmatrix}, \begin{pmatrix} I & N \\ 0 & I \end{pmatrix}, \quad N \text{ symmetric}$$

of SP_n act respectively as:

The Fourier transform $L^2(\mathbf{R}^n) \to L^2(\mathbf{R}^n)$

$$f(x) \to |\det{}' A|^{1/2} f(A^T x)$$

$$f(x) \to \exp(\pi i N(x)) f(x)$$

From this it is not difficult to calculate that the derived representation on the smooth vectors $S(\mathbf{R}^n)$ (Schwartz space) is the one given by the oprators

$$i x_k x_j,\ i\frac{\partial^2}{\partial x_k \partial x_j},\ x_k\frac{\partial}{\partial x_j} + \frac{1}{2}\delta_{kj}$$

which after a transformation $x_k \to \sqrt{i}x_k$ is precisely the Levi factor isomorphic to sp_n in ls_n. This already shows the relationship between linear filtering and the Segal-Shale-Weil representation.

More precisely associate to $(m,P) \in \mathbf{R}^{N-1}$, P positive definite, the normal density with covariance P and mean m. The induced vector fields on the image of $\{(m,P) \in \mathbf{R}^{N-1} \mid P$ positive definite$\}$ in $S(\mathbf{R}^n)$ are linear and they are the linear vector fields corresponding to the operators in ls_n on $S(\mathbf{R}^n)$. They are obviously extendable to all of $S(\mathbf{R}^n)$ and this consititutes a more precise version of the relationship.

References

[1] R.W. Brockett, J.M.C. Clark, The geometry of the conditional density equations, Proc of the Oxford Conf. on Stochastic Systems 1978, Acad. Press 1980.

[2] R.W. Brockett, Geometric methods in stochastic control and estimation. In [5], 441-478.

[3] M. Hazewinkel, S.I. Marcus, On Lie algebras and nonlinear filtering, Stochastics **7** (1982), 29-62.

[4] M. Hazewinkel, The linear systems Lie algebra, the Segal-Shale-Weil representation and all Kalman Bucy filters. Report 8130, Econometric Inst., Erasmus Univ. Rotterdam. Submitted Amer. J. Math.

[5] M. Hazewinkel, J.C. Willems (eds), Stochastic systems: the mathematics of filtering and identification and applications, Reidel Publ. G., 1981.

[6] R.E. Howe, On the role of the Heisenberg group in harmonic analysis, Bull. Amer. Math. Soc. **3** (1980), 821-844.

[7] R.E. Howe, ϑ-series and invariant theory, Proc. Symp. Pure Math. **33**, 1, 275-285, Amer. Math. Soc. 1980.

[8] M. Kashiwara, M. Vergne, On the Segal-Shale-Weil representation and harmonic polynomials, Inv. Math. **44** (1978), 1-47.

[9] G. Lion, M. Vergne, The Weil representation, Maslow index and Theta series, Birkhäuser, 1980.

[10] S.K. Mitter, Nonlinear filtering and stochastic mechanics. In [5], 479-504. And references therein.

[11] S. Rallis, G. Schiffmann, Weil representation I. Memoirs Amer. Math. Soc. **231** (1980).

[12] I.E. Segal, Transforms for operators and symplectic automorphisms over a locally compact abelian group, Math. Scand. **13** (1963), 31-43.

[13] D. Shale, Linear symmetries of free boson fields, Trans. Amer. Math. Soc. **103** (1962), 149-167.

[14] A. Weil, Sur certains groupes d'operateurs unitaires, Collected Papers Vol.III 1-71, Springer, 1980; Comments: ibid 443-447.