ECONOMETRIC INSTITUTE

LECTURES ON INVARIANTS, REPRESENTATIONS AND LIE ALGEBRAS IN SYSTEMS AND CONTROL THEORY

M. HAZEWINKEL

REPRINT SERIES no. 372

This article appeared in "Lecture Notes in Mathematics", no. 1029, A.Dold and B. Eckmann (eds.) Springer Verlag, Berlin, 1983.

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

LECTURES ON INVARIANTS, REPRESENTATIONS AND LIE ALGEBRAS IN SYSTEMS AND CONTROL THEORY

Michiel HAZEWINKEL

The Math. Centre, P.O. Box 4079, 1009 AB Amsterdam

The general purpose of these three lectures is to explain to an audience assumed to consist mainly of pure mathematicians, algebraically oriented perhaps, some of the many mathematical problems (and their solutions) which arise in systems and control theory with maybe a little extra emphasis on unsolved problems. It was a pleasure and an honor to be invited to speak on this topic in the Séminaire d'Algèbre and I here record my feelings of indebtedness in this respect towards the organizer in this case, Mme Prof. Marie-Paule MALLIAVIN.

CONTENTS

Lecture 1. Invariants and moduli for linear systems and applications

- Systems and linear systems, 1.1.
- A selection of questions concerning systems, 1.6.
- Why one should study families of systems rather than single ones, 1.13.
- On realization theory, 1.14.
- Stochastic realization theory, 1.19.
- Moduli theorems. Invariants, 1.21.
- Systems with delays, 1.23.
- Continuous canonical forms, 1.26.

- A few open questions concerning the moduli space $M^{cr,co}_{m,n,p}$, 1.27.
- A wild problem, 1.28.
- Systems with special structure. Symmetry algebras, 1.29.

Lecture 2. Feed back invariants of systems, nilpotent matrices, Schubert cells, representations of the symmetric groups and holomorphic vectorbundles.

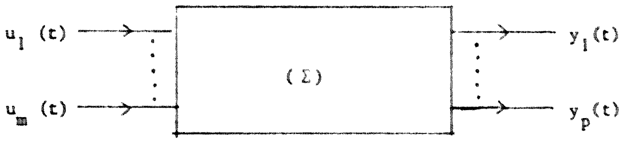
- Invariants under the feedback group, 2.1.
- Orbits of nilpotent matrices, 2.5.
- Holomorphic vector bundles over $\mathbb{P}^1(\mathbb{C})$, 2.6.
- The Hermann-Martin vector bundle of a system, 2.7.
- Schubert cells, 2.8.
- Representations of the symmetric group, 2.10.
- Some combinatorial occurences of the specialization order, 2.11.
- The specialization order in physics and chemistry, 2.12.

Lecture 3. Filtering problems and Lie algebras of differential operators and their representations.

- A non-representation theorem, 3.1.
- The recursive filtering problem, 3.3.
- The Duncan-Mortenson-Zakai equation, 3.11.
- The estimation Lie algebra, 3.14.
- Filters and homomorphisms of Lie algebras, 3.16.
- Robustness questions, 3.17.
- The cubic sensor, 3.18.
- Ideals of the estimation algebra, filters and approximate filters, 3.22.
- The Lie algebra $\underline{\ell s}_n$, 3.23.
- The representation defined by all Kalman-Bucy filters, 3.25.

Lecture 1. INVARIANTS AND MODULI FOR LINEAR SYSTEMS AND APPLICATIONS

1.1. Systems and linear systems. Very roughly a system is a device which accepts certain inputs : deterministic controls, stochastic noises or a mixture of the two, which processes these inputs and then produces certain outputs in response. The traditional diagrammatic picture is as follows :



where $(u_1(t),\ldots,u_m(t)) \in \mathbb{R}^n$ is an m-dimensional vector of inputs (depending on time t) and $(y_1(t),\ldots,y_p(t)) \in \mathbb{R}^p$ is a p-dimensional vector of outputs. It is easy to imagine systems with more general input and output spaces (than $\mathbb{R}^n$ and $\mathbb{R}^p$).

For example the machine suggested by picture (1.2) couldbe described by a set of differential equations :

$$(1.3) \qquad \dot{x} = f(x,u)\ ,\ y = h(x)\ ,\ x \in \mathbb{R}^n,\ u \in \mathbb{R}^m,\ y \in \mathbb{R}^p.$$

More generally x could evolve on a finite dimensional differentiable manifold M with $f(x,u)$ a family of vectorfields on M parametrised (differentiably) by $u \in \mathbb{R}^n$.

A particularly important class (for applications) of systems are the linear time invariant systems which are given by the equations :

$$(1.4) \qquad \dot{x} = Ax + Bu\ ,\ y = Cx\ ,\ x \in \mathbb{R}^n,\ u \in \mathbb{R}^m,\ y \in \mathbb{R}^p$$

where A,B,C are constant matrices of the appropriate dimensions, i.e. they are of sizes $n \times n$, $n \times m$ and $p \times n$, respectively. Equations (1.3) and (1.4) describe a continuous time model; equally important are discrete time models, which in the linear case look like :

$$(1.5) \qquad x_{t+1} = Ax_t + bu_t\ ,\ y_t = Cx_t.$$

(Sometimes one considers more general models than (1.4),(1.5) involving also a direct feed through term, so that then : $y = Cx + Du$, resp. $y_t = Cx_t + Du_t$; for the mathematical problems considered in this paper this makes little difference). It is to be observed that (1.5) makes sense over any ring ; also it is a fact that such systems over rings have real applications, e.g. in automata theory, picture processing.

In this first lecture we shall be exclusively concerned with linear systems and various (open) problems concerning them.

1.6. A selection of questions concerning systems. Systems as roughly described above arise e.g. as (tentative)(approximate) models of certain (ill understood) dynamic in put/out put phenomena (processes) like for instance economic development processes (or time series) and as models of devices involving controls (sometimes partly to be automated) like aeroplanes. Most of the questions discussed below receive content when viewed in the light of such examples.

A. Realization questions Given a device (1.3),(1.4) or (1.5) and a starting point $x(0) = x_o$ at time zero, the equations (1.3),(1.4) or (1.5) define an input/output operator taking input functions $u(t)$ to output functions $y(t)$. This operator describes what comes out of the device (i.e. is observed) when it is started at time zero in $x(0)$ and it is fed the input function $u(t)$.

E.g. if $\Sigma = (A,B,C)$ is the system (1.4) and $x(0) = 0 \in \mathbb{R}^n$,

then the corresponding input/output operator V_Σ is given by the convolution formula :

$$y(t) = \int_0^t C\, e^{(t-\tau)A}\, Bu(\tau)\, d\tau .$$

The basic "realization theory" question is now : given some input/output operator V when does there exist a system of type (1.3) or type (1.4) together with an initial state x_o such that the associated input/output operator is the pregiven operator V, or is a "best" approximation.

In this connection it should be remarked that there are often great advantages in having a model of say type (1.4). It should also be noted that the socalled state space models (1.3),(1.4),(1.5) are by no means the only way to specify a dynamic input/output relationship. Another way are the socalled ARMA models which in discrete time e.g. are specified by a relation-ship of the form :

$$(1.8) \qquad A_o y_t + A_1 y_{t-1} + \ldots + A_p y_{t-p} = B_o u_t + B_1 u_{t-1} + \ldots + B_q u_{t-q}.$$

Bits of realization theory will be discussed in sections 1.14 and 1.19 below.

B. Moduli problems. Invariants. As was remarked before if it is possible to realize a given input/output operator V by means of a linear system (1.4) and (1.5) it isfor many purposes advantageous to do so. However, there is a price. The input/output operator V_Σ, $\Sigma = (A,B,C)$, given by (1.7) in the continuous time case, does ..ot uniquely determine the triple of matrices (A,B,C). Indeed if S is an invertible $n \times n$ matrix then the triple :

$$(1.9) \qquad \Sigma^S = (A,B,C)^S = (S A S^{-1}, SB, CS^{-1})$$

gives exactly the same input/output operator (as follows immediately from (1.7)). The question immediately arises whether this is the only redundancy in $\Sigma = (A,B,C)$ vis-à-vis V_Σ. (Generically this is the case : cf. section 1.14 below). This leads to the following invariants and moduli problem.

Let $L_{m,n,p}(\mathbb{R})$ be the space of all triples (A,B,C) of real matrices of dimensions $n \times n$, $n \times m$, $p \times n$ respectively. Consider the action of $GL_n(\mathbb{R})$, the group of real invertible $n \times n$ matrices,on $L_{m,n,p}(\mathbb{R})$ given by (1.9). What are the invariants for this action ? To what extend does the quotient space $L_{m,n,p}(\mathbb{R})/GL_n(\mathbb{R})$ exist ? Is it a nice space in some sense ? Results concerning these questions can be found in 1.21 below. More generally for discrete time systems these questions are important over any ring (instead of $\mathbb{R}$).

C. Feedback problems. Stabilization. A linear system (1.4) or (1.5) is said to be stable if, for all initial states $x(0)$, $x(t)$ goes to zero as $t \to \infty$ if $u(t) \equiv 0$. In case of continuous time (system (1.4)) this is the case if all eigenvalues of A have strictly negative real parts and in discrete time (system (1.5)) this is the case if all eigenvalues of A are less than one in absolute value. An important class of problems in system theory asks to what extent systemscan be stabilized or be caused to have other desirable properties by means of feeding back certain linear combinations of the state or outputs into theinputs.

Mathematically state space feedback is described by the following action of the additive group $M(m,n)$ of all $m \times n$ matrices on $L_{m,n,p}$:

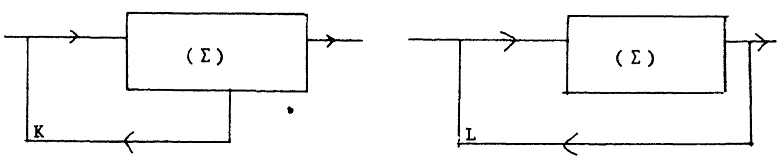
(1.10) $$(A,B,C)^K = (A + BK,B,C)\ , \ K \in M(m,n)$$

and output feedback is described by the action of $M(m,p)$ on $L_{m,n,p}$ given by :

(1.11) $$(A,B,C)^L = (A + BLC,B,C).$$

In block diagrams these feedback loops are depicted as below :

(1.12)

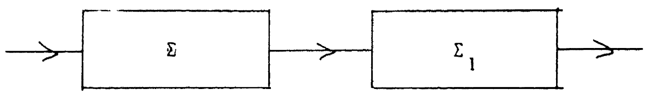


Typical problems are now : Which systems can be stabilized by state space feedback?

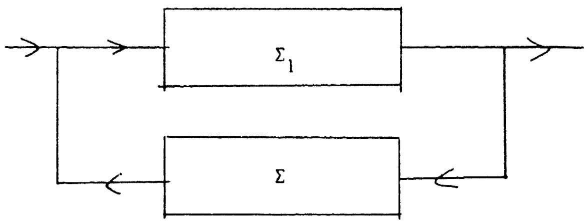
Completely solved by Wonham [1]. Which systems can be stabilized by output feedback ? This one is still essentially completely open ; for some recent results using Grassmann manifolds and intersection theory cf Byrnes [2].

Another problem could be : given an additional input channel through which undesired disturbances (or noise) enters the system. It is possibly to use state-space or output feedback in such a way that the disturbances do not show up in the outputs or such that (by employing larger and larger feedback matrices (= high feedback gain)) , the influence of these disturbances can be made as small as desired. Considerable and interesting work on this last problem has been done by Willems [36].

D. Model matching. Dynamic feedback. Another class of problem has to do with whether at certain points in an interconnected collection of linear dynamical systems there can be inserted a linear dynamical system (preferably of minimal dimension) such that certain properties hold. Consider for example two given systems Σ_1 and Σ_2 and the question of whether there exists a Σ such that the composed system :



has the same input/output operator as Σ_2. Or let there be given one system Σ_1 and consider the problem of constructing a system Σ such that the system with dynamic feedback loop :



is stable. There are many similar problems often involving much more complicated diagrams.

1.13. Why one should study families of systems rather than single ones.
In many cases with design problems as indicated under C and D above it will be the case that the given Σ_1 and Σ_2 are only imperfectly known. Or these systems may have certain parameters which can be adjusted to a variety of possible uses. In both cases the question arises how to solve these problems not for one system but for a family of systems (perhaps uniformly), and the question arises which of the single system solutions (if any) is continuous in the system parameters. Most of the problems mentioned above are largely open, even in the case of the largest family of them all, the one parametrized by the "quotient" $L_{m,n,p}/GL_n$. Still more reasons for studying families of systems rather that single ones can be found in [4,5,6].

1.14. On realization theory. Applying the Laplace transform to formula (1.7) yields:

$$Y(s) = T(s)\ U(s) \tag{1.15}$$

where $Y(s)$ and $U(s)$ are respectively the Laplace transforms of $y(t)$ and $u(t)$ and where $T(s)$, the socalled transfer function, is given by :

$$T(s) = C(sI - A)^{-1} B = H_0 s^{-1} + H_1 s^{-2} + H_2 s^{-3} + \dots, \quad H_i = CA^i B. \tag{1.16}$$

Thus one way to pose the realization question of 1.6.A above is to ask : given a sequence of $p \times m$ matrices $H_0, H_1, H_2, \dots$ when do there exist matrices A, B, C such that $H_i = CA^i B$, $i = 0,1,2,\dots$ The answer is as follows. Form the block Hankel matrix :

$$\mathcal{H} = \begin{pmatrix} H_0 & H_1 & H_2 & \cdots \\ H_1 & H_2 & H_3 & \cdots \\ H_2 & H_3 & H_4 & \cdots \\ \vdots & \vdots & \vdots & \end{pmatrix}$$

Then such (A,B,C) exist if and only if the rank of this matrix is finite. Moreover the minimal n for which there exists an $(A,B,C) \in L_{m,n,p}$ for which $H_i = CA^iB$, $i = 0,1,2,\ldots$ is equal to the rank of this Hankel matrix. These minimal dimensional realizations of the sequence $(H_o,H_1,H_2,\ldots)$ have two additional properties : they are completely observable (co) and completely reachable (cr). The abstract definitions of these two notions are as follows. The system $\Sigma = (A,B,C) \in L_{m,n,p}$ is cr iff the matrix :

$$R(A,B) = (B \mid AB \mid A^2B \mid \ldots\ldots \mid A^nB) \tag{1.17}$$

consisting of the blocks $B, AB,\ldots,A^nB$, is of rank n. Dually $\Sigma = (A,B,C)$ is completely observable if the matrix :

$$Q(A,C) = R(A^T,C^T)^T \tag{1.18}$$

has rank n. Here an upper T denotes transposes. These notions have the physical meanings their names suggest. Σ is cr if starting in $x(o) = 0$ at time 0 any state x can be reached by means of a suitable control function $u(t)$ and Σ is co if from the observations $y(t)$, $t \geq 0$, it can be seen whether two initial states $x(o)$, $x'(o)$ are different or not (assuming $u(t) \equiv 0$). It is also true that two realizations (A,B,C), $(A',B',C') \in L_{m,n,p}$ which are both cr and co yield the same input/output operator if and only if there is an $S \in GL_n$ such that $(A',B',C') = (A,B,C)^S$.

The fact that a minimal dimensional realization is cr and co follows immediately from the observation that $\mathcal{H} = Q(A,C)\,R(A,B)$ if (A,B,C) realizes $(H_o,H_1,H_2,\ldots)$.

For more details concerning the deterministic realization theory described above (which is due to Kalman) cf [7] and also [8].

1.19. Stochastic realization theory. If $u(t)$ in (1.4) is white noise, more precisely if we rewrite (1.4) as an Ito stochastic differential equation :

$$(1.20) \qquad dx = Axdt + Bdw_t \ , \ dy = Cxdt + dv_t$$

where w_t and v_t are independent unit variance Wiener processes also independent of $x(o)$, then $y(t)$ is a Gaussian stochastic process. The question now arises : given a (Gaussian) stochastic process, when does there exists a machine (system) (1.20) which generates this process. This belongs to the area of stochastic realization theory where there are still a good many open problems. For a recent survey cf. [9].

1.21. Moduli theorems. Invariants. Let $L^{cr}_{m,n,p}(\mathbb{R})$ (resp. $L^{co}_{m,n,p}(\mathbb{R})$) denotes the space of all cr (resp. co) triples (A,B,C) and let $L^{co,cr}_{m,n,p}(\mathbb{R})$ be the intersection of these two subspaces of $L_{m,n,p}(\mathbb{R})$. The basic theorem concerning the action of $Gl_n(\mathbb{R})$ on $L^{co,cr}_{m,n,p}(\mathbb{R})$ is the following :

1.22. Theorem. The quotient $L^{co,cr}_{m,n,p}(\mathbb{R})/GL_n(\mathbb{R}) = M^{co,cr}_{m,n,p}(\mathbb{R})$ exists and it is a smooth differentiable manifold of dimension $mn + np$. The projection $\pi : L^{co,cr}_{m,n,p}(\mathbb{R}) \to M^{co,cr}_{m,n,p}(\mathbb{R})$ is a principal $GL_n(\mathbb{R})$-fibre bundle which is trivial if and only if $m = 1$ or $p=1$. The manifold $M^{co,cr}_{m,n,p}(\mathbb{R})$ is never compact ; it is connected iff $mp \geq 2$. The map $(A,B,C) \to (H_o,\ldots,H_{2n})$, $H_i = CA^iB$, induces an embedding of differentiable manifolds $M^{co,cr}_{m,n,p}(\mathbb{R}) \to \mathbb{R}^{(2n+1)mp}$. (Actually H_{2n} is superfluous and can be calculated from $H_o,\ldots,H_{2n-1}$).

As a corollary of theorem 1.22 it follows that the only invariants of $GL_n(\mathbb{R})$ acting on $L_{m,n,p}(\mathbb{R})$ are functions* of the entries of the $H_i = CA^iB$, $i = 0,1,2,\ldots,2n-1$. (These entries are of course obvious invariants). The relations between these invariants are the defining equations of the closure of $M^{co,cr}_{m,n,p}(\mathbb{R})$

in $\mathbb{R}^{2nmp}$. These are all determinantal identities and are given by the prescription that all $(n+1)\times(n+1)$ subdeterminants of the matrix (of block Hankel type) :

$$\begin{pmatrix} H_o & H_1 \cdots\cdots H_{n-1} & H_n \\ H_1 & H_2 \cdots\cdots H_n & H_{n+1} \\ \vdots & \vdots \qquad\quad \vdots & \vdots \\ H_{n-1} & H_n \quad\cdots\quad H_{2n-2} & H_{2n-1} \\ H_n & H_{n+1}\cdots H_{2n-1} & H_{2n} \end{pmatrix}$$

are zero.

As a matter of fact an even stronger theorem that 1.22 holds. It turns our that $M^{co,cr}_{m,n,p}$ (IR) is a fine moduli space, i.e. that there exists over $M^{co,cr}_{m,n,p}$ (IR) an universal family of co and cr systems from which every family can be obtained uniquely by pull-back. This family is defined on an n-dimensional vector bundle over $M^{co,cr}_{m,n,p}$ (IR) which is trivial if and only if $m=1$ or $p=1$. cf [10] or [5] for details.

In the next section we shall see how this fact can be used to say things about the realization theory of certain infinite dimensional systems.

There exists also an algebraic geometric version of theorem 1.22 which essentially says that there exists a scheme $M_{m,n,p}$ defined over Z of which $M^{co,cr}_{m,n,p}$ (IR) and $M^{co,cr}_{m,n,p}$ ($\mathbb{C}$) are the varieties of real and complex points.

One can of course also study pairs of matrices (A,B) under the action of GL_n given by : $(A,B)^S = (S\,A\,S^{-1}, SB)$. This is also of relevance to system and control theory (through to a lesser extent). Mathematically though things come out prettier and this particular problem fits in better with the existing techniques and theorems of geometric invariant theory. See the lecture notes by Tannenbaum [12].

1.23. Systems with delays. A linear system with delays is e.g. :

$$
\begin{aligned}
(1.24)\qquad \dot{x}_1(t) &= 2x_1(t) + x_1(t-\alpha_1) + 3x_2(t-\alpha_2) + u(t-\alpha_1-\alpha_2) + 2u(t) \\
\dot{x}_2(t) &= x_1(t-3\alpha_1) + 4\,x_2(t) + u(t-2\alpha_2) \\
y(t) &= 4\,x_1(t-\alpha_2) - 2x_2(t)
\end{aligned}
$$

where α_1,α_2 are two positive numbers (the delays) such that α_1,α_2 are independent over $\mathbb{Q}$.

The transfer function of this example, that is the Laplace transform of the corresponding input/output operator is a rational function in s, $e^{-\alpha_1 s}$, $e^{-\alpha_2 s}$. More precisely it is a strictly proper rational function in s with coefficients which are polynomials in $e^{-\alpha_1 s}$, $e^{-\alpha_2 s}$.

These are in principle infinite dimensional systems. (To predict for given inputs $u(t)$ the development of the system one needs data not finite dimensional like $x(o)$ but initial data which live in an infinite dimensional space, e.g. one needs the function $x(\tau)$ for $-\max(\alpha_1,\alpha_2) \leq \tau \leq 0$.).

Let σ_1,σ_2 denote the delay operators : $\sigma_i f(t) = f(t-\alpha_i)$, $i=1,2$. Then we can rewrite (1.24) in the form :

$$
(1.25)\qquad \dot{x} = \begin{pmatrix} 2+\sigma_1 & 3\sigma_2 \\ \sigma_1^3 & 4 \end{pmatrix} x + \begin{pmatrix} 2+\sigma_1 + \sigma_2 \\ \sigma_2^2 \end{pmatrix} u \ , \ y = (4\sigma_2-2)\,x
$$

and associate to this in turn the family of systems (1.25) parametrized by the complex parameters σ_1,σ_2. By this technique one can e.g. prove certain stabilization theorem for systems with delays, cf. [11] (and the survey paper [6] for a different proof of this same theorem). On spite of these results (which rely on the Quillen-Suslin theorem) most questions concerning stabilization and feedback for delay systems are sill open and many more results should come out of associating families of systems to them.

One can also pose the realization problem for delay system. Let there be given a matrix valued rational function $T(s)$ in $s, e^{-\alpha_1 s}, \ldots, e^{-\alpha_r s}$, where $\alpha_1, \ldots, \alpha_r$ are positive numbers linearly independent over $\mathbb{Q}$. Does there exists a linear delay system with delays $\alpha_1, \ldots, \alpha_r$ with this transfer function ? Here is a result concerning this. First because $s, e^{-\alpha_1 s}, \ldots, e^{-\alpha_r s}$ are algebraically independent there exists a uniquely determined rational function $T_\sigma(s)$ with coefficients in $\mathbb{R}[\sigma_1, \ldots, \sigma_r]$ such that substituting $e^{-\alpha_i s}$ for σ_i gives $T(s)$. For each complex vector $\bar{\sigma} = (\bar{\sigma}_1, \ldots, \bar{\sigma}_r)$ this gives an ordinary complex transfer function $T_{\bar{\sigma}}(s)$. These are all realizable. Suppose that the minimal realization dimension (called the Mac Millan degree) of $T_{\bar{\sigma}}(s)$ is n for all $\bar{\sigma} \in \mathbb{C}^r$. Then $\bar{\sigma} \mapsto$ (first $2n+1$ matrix coefficients of the s^{-1} power series development of $T_{\bar{\sigma}}(s)$) defines a continuous map $\mathbb{C}^r \to \mathbb{C}^{(2n+1)mp}$ whose image is in $M^{co,cr}_{m,n,p}(\mathbb{C}) \subset \mathbb{C}^{(2n+1)mp}$. Pulling back the universal family over $M^{co,cr}_{m,n,p}(\mathbb{C})$ gives a family of systems of $\mathbb{C}^r$, which is algebraic and defined over $\mathbb{R}$. By the Quillen-Suslin theorem the underling vector bundle is trivial which implies that there are matrices $A(\sigma)$, $B(\sigma)$, $C(\sigma)$ which coefficients which are polynomials over $\mathbb{R}[\sigma_1, \ldots, \sigma_r]$. Now reinterpret σ_i as the delay operator $\sigma_i f(t) = f(t-\alpha_i)$ to find the desired system with delays. cf [5,6] for more details.

Remark. Both the stabilization theorem and the realization theorem for delay systems mentioned above rely on the Quillen-Suslin theorem on the triviality of algebraic vector bundles over the affine spaces : $\mathrm{Spec}(k[X_1, \ldots X_r])$. This means that to calculate the desired feedback matrix (with delays) and the realizing delay system we need an algorithmic (effective) way of obtaining the Quillen-Suslin trivialization. I.e. a constructive proof is needed and this is so far missing.

1.26 Continuous canonical forms : To obtain the matrices (A,B,C) from measurement data $H_o,\dots H_{2n}$ certain choices have to be made because the $H_o,\dots,H_{2n}$ determine (A,B,C) only up to $GL_n(\mathbb{R})$-equivalence. In other words given certain (statistical) data on the input/output behaviour of a system which is assumed to be linear of Mac Millan degree n, the statistical problem of finding the best (A,B,C) which model the data is not well posed : there are redundant parameters to be eliminated. And the question arises whether this can be done in a continuous way. (For obvious reasons this is desirable). A continuous canonical form on $L^{co,cr}_{m,n,p}(\mathbb{R})$ is a continuous mapping $c : L^{co,cr}_{m,n,p}(\mathbb{R}) \to L^{co,cr}_{m,n,p}(\mathbb{R})$ such that (i) (A,B,C) and c(A,B,C) have the same input/output map and (ii) (A,B,C) and (A',B',C') have the same input/output operators iff c(A,B,C) = c(A',B',C'). The question now is whether continuous canonical forms exist. The answer is given by theorem 1.22 : such a canonical form exists iff p=1 or m=1.

Given the fact that in general no continuous canonical forms exist one wonders whether there exist discontinuous ones such that the discontinuities are everywhere bounded by a universal constant k (in norm), and how small k can be. This is completely open.

1.27 A few open questions concerning $M^{co,cr}_{m,n,p}$. As was stated above a continuous canonical form usually does not exist on $L^{co,cr}_{m,n,p}(\mathbb{R})$. Discontinuous ones do of course exist. So this approach to get rid of the superfluous parameters does not seem to work very well. It seems much more natural to eliminate the redundant parameters by going to the quotient $M^{co,cr}_{m,n,p}(\mathbb{R})$ and to view identification of a system as walking around on $M^{co,cr}_{m,n,p}(\mathbb{R})$ getting closer and closer to the true (a best approximating) system as more and more measurement data come in. With this in mind one would like to know much more about $M^{co,cr}_{m,n,p}(\mathbb{R})$ than we do at present. For instance : it is complete in some metric which agrees with a natural concept of convergence of input/output operators ? (For some initial results in such questions cf. also [4]. One would also like the Riemannian metric on $M^{co,cr}_{m,n,p}(\mathbb{R})$ to agree with the statistics of the situation perhaps in the following sense. Consider the model (1.20) for two different triples (A,B,C),(A',B',C'). Feed these systems the same white noise w_t starting in the same initial point x(o). There result two different random processes y(t), y'(t). The distance between (A,B,C) and (A',B',C') now should have much to do with the amount of information which y(t) carries about y'(t) and vice versa (perhaps for small t only).

1.28 <u>A wild problem</u>. The action of GL_n on triples (A,B,C) corresponds to "base change in state space" ; that is if $\dot{x} = Ax + Bu$, $y = Cx$ it corresponds to a transformation $x \to Sx$. In some settings it seems entirely natural to admit also base change in input space and output space. This leads to an action of $Gl_n \times Gl_m \times Gl_p$ on triples given by :

$$
(1.29) \qquad \begin{aligned}
(A,B,C)^S &= (SAS^{-1}, SB, CS^{-1}),\ S \in GL_n \\
(A,B,C,)^T &= (A,\ BT^{-1},\ C)\ ,\ T \in GL_m \\
(A,B,C)^U &= (A,B,UC)\ ,\ u \in GL_p
\end{aligned}
$$

and there arises the problem of studying and describing the "quotient" $L^{co,cr}_{m,n,p} / GL_n \times Gl_m \times GL_p$. This is the problem of describing all the representations of the diagramm :

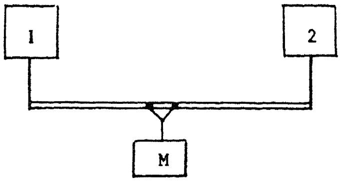
$$
\begin{tikzcd}
\bullet \arrow[r] & \bullet \arrow[r] \arrow[loop above] & \bullet
\end{tikzcd}
$$

in the sense of quiver theory. It is also a wild problem in the technical sense of the word. The proceedings [13] contain much information on this branch of representation theory of algebras.

1.29 <u>Systems with special structure. Symmetry algebras</u>. Often dynamical systems are composed of several (identical) subsystems interconnected in various ways. As an example one might have a linear control system $\dot{x} = Ax + Bu$ with A and B given by :

(1.30) $$A = \begin{pmatrix} F & H \\ -H & F \end{pmatrix}, \; B : \begin{pmatrix} G & 0 \\ 0 & G \end{pmatrix}.$$

This represents a linear model of two helicopters connected with a rigid beam as sketched below :



where M is a load to be lifted which is too heavy for a single helicopter. Then $\dot{x} = Fx + Gu$ is a linear model of a single helicopter and H represents the interaction dynamics.

A problem is now e.g. to find a feedback matrix of the form :

(1.31) $$K = \begin{pmatrix} K_1 & 0 \\ 0 & K_1 \end{pmatrix}$$

which stabilizes the compound system. Or in any case to do this by means of a feedback which preserves the special structure of the matrix A. One approach to such questions is as follows. Given a class of systems like (1.30) the symetry algebra R is defined as consisting of all elements $(S;T)$ in $M(2n,2n) \times M(2m,2m)$ (here $M(q,q)$ is the matrix algebra of all $q \times q$ matrices) such that :

$$SA = AS \qquad , \quad SB = BT.$$

$\mathbb{R}^{2n}$ and $\mathbb{R}^{2m}$ are natural $M(2n,2n) \times M(2m,2m)$-modules and R is the maximal subalgebra for which A and B are R-module homomorphisms. In the example under consideration R turns out to be $\mathbb{R}[i] \simeq \mathbb{C}$. The algebra R is the <u>symmetry algebra</u> for this class of systems.

One can show that every associative finite dimensional algebra can arise as the symmetry algebra of some class of systems with special structure [21]. The extra requirement that the feedbacks preserve the special structure now becomes that K : state space $\rightarrow$ input space be a homomorphism of R-modules. In this example that means that K must be of the form :

$$\begin{pmatrix} K_1 & -K_2 \\ K_2 & H_1 \end{pmatrix} ,$$

which is still not what is required. This can be taken care of by a second larger symmetry algebra $R' \supset R$ and requiring that K be a R'-module homomorphism.

I remark that output feedback problems can also be put in this framework indicating that these problems of special structure preserving feedback will probably be quite hard.

Indeed in the example under consideration, it seems likely that there exist examples with the following properties :

(i) (A,B) is completely reachable

(ii) (F,G) is completely reachable

(iii) There exists a number t such that for every feedback K of the form (1.31) with $\|K\| > t$ the system $(A + BK, B)$ is unstable.

For more details concerning this topic of linear systems with special structure and decentralized control, cf [21,35].

Lecture 2. FEEDBACK INVARIANTS OF SYSTEMS, NILPOTENT MATRICES , SCHUBE REPRESENTATIONS OF THE SYMMETRIC GROUPS AND HOLOMORPHIC VECTOR-BUNDLES.

2.1. Invariants and the feedback group. Let $L_{m,n}$ denote the space of of matrices (A,B) of dimension $n \times n$ and $n \times m$ respectively. $L^{cr}_{m,}$ subspace of all completely reachable pairs. The feedback group $F_{n,m}$ a $L_{m,n}$ is generated by base change in state space, base change in input state space feedback. More precisely $F_{n,m}$ is the closed subgroup of consisting of all matrices of the form :

$$g = \begin{pmatrix} S & 0 \\ K & T \end{pmatrix}, \quad S \in GL_n, \quad T \in GL_m, \quad K \in M(m,n)$$

acting on $L_{m,n}$ by :

$$(2.2) \qquad (A,B)^g = (SAS^{-1} + SBTS^{-1}K, \ SBT).$$

The subspace $L^{cr}_{m,n}$ is stable under $F_{n,m}$.

Now consider an array of dots of dimensions $m \times (n+1)$ as bel

$$\begin{array}{cccc} \bullet & \bullet & \bullet & \cdots \\ \bullet & \bullet & \bullet & \cdots \\ \bullet & \bullet & \bullet & \cdots \\ B & AB & A^2B & \end{array}$$

The first column represents the columns of B. The second, the columns of Now for a given $(A,B) \in L^{cr}_{m,n}$ play the following game. Go down the first put a cross whenever the corresponding column vector of B is not in th generated by the previous vectors. (For the first column vector this sub is the zero subspace) ; continue with the second column of dots. The res for instance be as below ($m = 4$, $n = 7$)

$$\begin{array}{cccccccc} \times & \times & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \times & \times & \times & \times & \cdot & \cdot & \cdot & \cdot \\ \times & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{array}$$

which, if D_i denotes the i-th column vector of the matrix D, means that :

$B_1 \neq 0$, $B_2 \in \langle B_1 \rangle$, $B_3 \notin \langle B_1, B_2 \rangle$, $B_4 \notin \langle B_1, B_2, B_3 \rangle$, $(AB)_1 \notin \langle B_1, B_2, B_3, B_4 \rangle$, $(AB)_2 \in \langle B_1, \ldots, B_4, (AB)_1 \rangle$, $(AB)_3 \notin \langle B_1, \ldots, (AB)_2 \rangle$, $(AB)_4 \in \langle B_1, \ldots, (AB)_3 \rangle$, $(A^2B)_1 \in \langle B_1, \ldots, (AB)_4 \rangle$, ...

It is an easy lemma to show that if there is a cross anywhere then to the left of it there are necessarily crosses. Also the total number of crosses is equal to n iff $(A,B) \in L^{cr}_{m,n}$ (by the definition of cr). It follows that the pattern of crosses is defined by m integers ≥ 0 giving the number of crosses in each line. This sequence of integers $\tilde{\kappa}(A,B) = (\tilde{\kappa}_1(A,B), \tilde{\kappa}_2(A,B), \ldots, \tilde{\kappa}_m(A,B))$ is called the <u>Kronecker selection</u> of the pair (A,B). In the example $\tilde{\kappa} = (2,0,4,1)$. The <u>Kronecker indices</u> $\kappa(A,B)$ of the pair (A,B) are the same integers arranged in decreasing order of magnitude. Thus in the example $\kappa = (4,2,1,0)$. The final zeros are often omitted. More details, including an explanation of why these invariants are named after Kronecker are in [6].

Thus to each $(A,B) \in L^{cr}_{m,n}$ there is associated a partition of n, the Kronecker indices $\kappa(A,B)$.

2.3. <u>Theorem</u> (Kalman, Brunovsky). The $\kappa(A,B)$ are invariants under $F_{n,m}$. They are also the only invariants. I.e. $(A,B)^g = (A',B')$ for some $g \in F_{n,m}$ iff $\kappa(A,B) = \kappa(A',B')$. All partitions of n occur as a $\kappa(A,B)$.

The discrete set of all partitions of n thus is equal to the quotient (as a set) $L^{cr}_{m,n} / F_{n,m}$ and it inherits a topology from $L^{cr}_{m,n}$ in case we are working over $\mathbb{R}$ or $\mathbb{C}$ (or in fact any field using the Zariski topology in those cases). This is a partial order on the finite set of all partitions of n which turns out to be the following :

(2.4) $\lambda = (\lambda_1,\dots,\lambda_m) > \mu = (\mu_1,\dots,\mu_m) \leftrightarrow \sum_{i=1}^{r} \lambda_i \leq \sum_{i=1}^{r} \lambda_i$ for $r = 1,\dots,m$.

This is an ordering which I like to call the specialization ordering and which occurs under different names also in several other parts in mathematics. Thus the question arises whether it is an accident that the same ordering occurs all over the place or whether there are deeper connections. The latter possibility turns out to hold and the rest of this second lecture is devoted to describing some of these other occurences and some of the connections between them. Most of what follows (and more) can be found in more detail in [17].

The "degeneration of systems theorem" stated above which says that the specialization order on partitions of n is the quotient topology on $L^{cr}_{m,n}/F_{n,m}$ is relevant for control theory in that it tells us how the control structure of a system can suddenly change under deformation (system failure).

Let me insert here a few words on how one can prove Byrnes' theorem on the stabilization of feedback systems. Let Σ be such a system, cf. 1.23, and let Σ_σ be the associated family of systems parametrized by $\sigma \in \mathbb{C}^r$. The theorem says that if $\kappa(\Sigma_\sigma)$ is constant as a function of σ then Σ can be stabilized by means of a state feedback law (which has delays). To prove this one, first shows that for polynomial families over $\mathbb{C}^r$ the constancy of $\kappa(\Sigma_\sigma)$ implies the constancy of $\tilde{\kappa}(\Sigma_\sigma)$. This uses the Quillen-Suslin theorem. Then for the space of all systems with the same Kronecker selection there does exist a continuous, indeed algebraic, canonical form (with respect to the action of GL_n) [8] and in terms of this canonical form the stabilizing matrix can be written down immediately.

2.5. Orbits of nilpotent matrices. Let N_n be the space of all complex nilpotent matrices of size $n \times n$. Consider the action of $GL_n(\mathbb{C})$ on N_n by similarity, i.e. $N^S = S N S^{-1}$. The orbits are classified by partitions of n (Jordan canonical form) and the Gerstenhaber-Hesselink theorem says that if $0(\kappa)$ denotes the orbit classified by the partition κ then $\overline{0(\kappa)} \supset 0(\lambda) \leftrightarrow \kappa < \lambda$ (in the specialization order). The connection with Kronecher indices is as follows. For every $N \in N_n$ let $s(N) = \{(A,B) \in L^{cr}_{m,n}(\mathbb{C}) : N^i A^{i-i} B = 0 \text{ for all } i \in \{1,\ldots,n\}\}$ and for every $(A,B) \in L^{cr}_{m,n}(\mathbb{C})$ let $m(A,B) = \{N \in N_n : N^i A^{i-1} B = 0, i=1,\ldots n\}$. Then s and m induce mutually inverse bijections from the set of closed orbits of N_n under $GL_n(\mathbb{C})$ and $L^{cr}_{m,n}(\mathbb{C})$ under $F_{n,m}(\mathbb{C})$.

2.6. Holomorphic vector-bundles over $\mathbb{P}^1(\mathbb{C})$. Let E be an m-dimensional holomorphic (or algebraic) vector-bundle over the Riemann sphere $\mathbb{P}^1(\mathbb{C})$. According to Grothendieck, E splits as a sum of line bundles : $E \simeq L_1 \oplus \ldots \oplus L_m$ and in turn these line bundles are classified by an integer (their first chern number). Thus vectorbundles E over $\mathbb{P}^1(\mathbb{C})$ are classified by decreasing sequences of integers $K(E) = (K_1(E),\ldots,K_m(E))$. For a completely elementary proof of this fact cf. [14]. The bundle E is ample if $K_i(E) \geq 0$ all i.

Now consider a holomorphic family E_t of m-dimensional vector-bundles over $\mathbb{P}^1(\mathbb{C})$ with $\sum_i K_i(E_t)$ constant. Then according to a theorem of Shatz $K(E_o) < K(E_t)$ for t small and conversely if κ and λ are two partitions of n and $\kappa < \lambda$, then there exists a family E_t of line bundles over $\mathbb{P}^1(\mathbb{C})$ such that $K(E_o) = \kappa$ and $K(E_t) = \lambda$ for $t \neq 0$.

2.7. The Herman-Martin vectorbundle of a system. Let $(A,B) \in L^{cr}_{m,n}$ and let $\mathcal{G}_n(\mathbb{C}^{n+m})$ be the Grassmann manifold of all n-dimensional subspaces of $\mathbb{C}^{n+m}$. The map $\varphi_{(A,B)} : \mathbb{P}^1(\mathbb{C}) \to \mathcal{G}_n(\mathbb{C}^{n+m})$ is defined as follows. For each $s \neq \infty$ in $\mathbb{P}^1(\mathbb{C})$, let $\varphi_{(A,B)}(s)$ be the point in $\mathcal{G}_n(\mathbb{C}^{n+m})$ represented by the $n \times (n+m)$ matrix :

$$(sI - A \,\vdots\, B)$$

and to $s = \infty \in \mathbb{P}^1(\mathbb{C})$ associate the point of $\mathcal{G}_n(\mathbb{C}^{n+m})$ represented by $(I_n \,\vdots\, 0)$. It is not difficult to check that this defines an holomorphic map $\mathbb{P}^1(\mathbb{C}) \to \mathcal{G}_n(\mathbb{C}^{n+m})$.

Now let ξ_m be the very ample universal bundle over $\mathcal{G}_n(\mathbb{C}^{n+m})$ whose fibre over $x \in \mathcal{G}_n(\mathbb{C}^{n+m})$ is the quotient space $\mathbb{C}^{n+m}/x$.

The Hermann-Martin bundle of a completely reachable pair $(A,B) \in L^{cr}_{m,n}(\mathbb{C})$ is the induced bundle $\varphi^!_{(A,B)}\ \xi_m$ over $\mathbb{P}^1(\mathbb{C})$ and they prove that $K_i(\varphi^!_{(A,B)}\ \xi_m) = \kappa_i(A,B)$ $i=1,\ldots,m$, which explains why the same ordering occurs for families of vector bundles under isomorphism and families of systems under feedback.

2.8 Schubert cells. Let $\mathcal{A} = (A_1,\ldots,A_n)$ be a sequence of subspaces $0 \neq A_1 \underset{\neq}{\subset} A_2 \underset{\neq}{\subset} A_3 \underset{\neq}{\subset} \ldots \underset{\neq}{\subset} A_n \subset \mathbb{C}^{n+m}$ of $\mathbb{C}^{n+m}$. The closed Schubert cell determined by $\mathcal{A}$ is defined by :

$$SC\mathcal{A} = \{x \in \mathcal{G}_n(\mathbb{C}^{n+m}) \ :\ \dim\,(x \cap A_i) \geq i\}.$$

In particular if $0 < \tau_1 < \ldots < \tau_n \leq n+m$ is a sequence of increasing natural numbers $\leq n+m$, we write $SC(\tau)$ for the Schubert cell determined by $(\mathbb{C}^{\tau_1},\ldots,\mathbb{C}^{\tau_n})$ where $\mathbb{C}^j \subset \mathbb{C}^{n+m}$ is the subspace of all vectors whose last $n+m-j$ coordinates are zero. It is easy to check that $SC(\tau) \supset SC(\sigma)$ if $\tau_i \geq \sigma_i$ for all i.

Now let κ be an ($\leq m$ part) partition of n, $\kappa = (\kappa_1,\ldots,\kappa_m)$, $\kappa_j \geq 0$. To κ associate the following sequence $\tau(\kappa)$ of n natural numbers :

$$\underbrace{2,3,\ldots,\kappa_1+1}_{\kappa_1}\ ,\ \underbrace{\kappa_1+3,\ldots,\kappa_1+\kappa_2+2}_{\kappa_2},\ \ldots,\ \underbrace{\kappa_1+\ldots+\kappa_{m-1}+m-1,\ldots,\kappa_1+\ldots+\kappa_m+m}_{\kappa_m}$$

one verifies immediately that $\kappa > \lambda \leftrightarrow \tau_i(\kappa) \geq \tau_i(\lambda)$, $i=1,\ldots,n$.

The relations between Schubert cells and completely reachable pairs are mediated by the Hermann-Martin map $\varphi_{(A,B)} : \mathbb{P}^1(\mathbb{C}) \to \mathcal{G}_n(\mathbb{C}^{n+m})$. The results are as follows :

2.9. Theorem : Let $(A,B) \in L^{cr}_{m,n}(\mathbb{C})$ and let $\kappa = \kappa(A,B)$. Then there exists a sequence of subspaces $\mathcal{A} = (A_1,\ldots,A_m)$ of $\mathbb{C}^{n+m}$ such that $\dim A_i = \tau_i(x)$ and such that $\mathrm{Im}(\varphi_{(A,B)}) \subset SC(\mathcal{A})$. Conversely if $SC(\mathcal{B})$ is a closed Schubert cell such that $\mathrm{Im}(\varphi_{(A,B)}) \subset SC(B)$, then $\dim B_i \geq \tau_i(\kappa)$.

This shows that the degeneration (or specialization) of the Kronecker indices relates to the closure ordering of Schubert cells and links this order with the Bruhat order (or Bernstein-Gelfand-Gelfand order) on the symmetric groups S_n.

2.10. Representations of the symmetric group.

Let S_n denotes the symmetric group of all permutations of n letters. For a partition κ of n let S_κ denote the socalled Young subgroup $S_{\kappa_1} \times S_{\kappa_2} \times \ldots S_{\kappa_m} \subset S_n$. Finally let $\rho(\kappa)$ be the representation of S_n obtained by inducing the trivial representation of S_κ up to S_n. It is a theorem of Young Snapper, Lam, Liebler-Vitale that $\rho(\kappa)$ is a direct summand of $\rho(\lambda)$ iff $\kappa < \lambda$. For a completely elementary proof see [15].

There is a natural connection of this result with the result discussed before due to Kraft (with further developments by de Concini-Procesi). It goes as follows. Let $O(\kappa)$ be the orbit of nilpotent matrices under similarity classified by the partition κ. Let $\overline{O(\kappa)} \subset M(n,n)$ be its closure. Let $\mathcal{D}$ be the closed subvariety of diagonal matrices of $M(n,n)$. The set theoretic intersection of $\mathcal{D}$

and $\overline{O(\kappa)}$ is the zero matrix, but the scheme theoretic intersection need not be trivial. It is the spectrum of a finite dimensional local algebra $A(\kappa)$ over $\mathbb{C}$. Both $\mathcal{D}$ and $O(\kappa)$ are invariant under S_n(viewed in the natural way as a subgroup of $Gl_n(\mathbb{C})$, so $A(\kappa)$ carries a representation of S_n. This is the representation $\rho(\kappa)$.

The manifold interrelated occurences of the specialization order are not exhausted by what has been said above (indeed a few more appear below). This particular one (between nilpotent matrices and representations of S_n) also occurs in another diagram of functorial relations involving such things as the Springer representations, irreducible quotients of Verma modules, The Jantzen conjectures recently proved by A. Joseph, work of Kazhdan-Lusztig and work of Gelfand-Mac Pherson (which again links up Schubert cells). Clearly there is much room for further work here (and much is, in fact, in progress).

2.11. Some combinatorial occurences of the specialization order. In combinatorics the specialization order turns up in connection with such theorems as the marriage theorem, the theorem that every doubly stochastic matrix is a convex linear combination of permutation matrices, and the existense of (0,1)-matrices with prescribed row and column sums(Gale-Ryser theorem). A doubly stochastic matric is a matrix consisting of ≥ 0 elements such that all rows and columns sum to 1. One manifestation of the specialization ordering is that $\kappa > \lambda$ iff there exists a doubly stochastic matrix M such that $\kappa = M\lambda$. Another one involves Muirhead's inequality which is a far reaching generalization of the well known arithmetic mean geometric mean inequality. The latter corresponds to the extreme partition specialization ordering relation $(1,1,\ldots,1) > (n,0,\ldots,0)$. Cf [16] and [17].

2.12. The specialization ordering in physics and chemistry. Consider a thermodynamical process governed by a master equation. Then the vector $\rho = (\rho_1,\ldots,\rho_m,\ldots)$of probabilities that a particle be in state i evolves in such a way that $\rho(t) > \rho(t')$ if $t \geq t'$, ([18],[19]), a statement that is a good deal stronger than the statement that the entropy must always increase. There is a good deal more to be said on these relations between stochasticity and this particular partial order involving e.g. W^{*}-algebras, the Von Neumann-Murray classification of pairs of projectors and non commutative probability theory. c.f. [20].

Lecture 3. FILTERING AND LIE ALGEBRAS OF DIFFERENTIABLE OPERATORS AND THEIR REPRESENTATIONS.

3.1. A non representation theorem. Let W_n be the vector space of all differential operators in n variables with polynomial coefficients. As an associative algebra $W_n = \mathbb{R}< x_1,\dots,x_n, \frac{\partial}{\partial x_1},\dots,\frac{\partial}{\partial x_n} >$ the associative algebra generated by $2n$ symbols subject to the relations $[x_i,x_j] = [\frac{\partial}{\partial x_i}, \frac{\partial}{\partial x_j}] = 0$, $[\frac{\partial}{\partial x_i}, x_j] = \delta_{ij}$. In this section W_n will usually be considered as a Lie algebra under the commutator product $[D_1,D_2] = D_1D_2 - D_2D_1$. Thus for example in W_1 we have :

$$[\frac{\partial}{\partial x}, x] = 1 \ , \ [\frac{\partial^2}{\partial x^2}, x^2] = 4x\frac{\partial}{\partial x} + 2, \ [x\frac{\partial^2}{\partial^2 x^2}, x\frac{\partial}{\partial x}] = x\frac{\partial^2}{\partial x^2} .$$

There are many unsolved questions concerning the Lie algebras W_n. E.g. what is $\mathrm{Aut}(W_n)$, what are (up to isomorphism) the maximal subalgebras of W_n. An elementary fact concerning the W_n is that the only non trivial ideal is the ideal $\mathbb{R}.1$ consisting of all scalar multiples of the identity operator.

Let M be a finite dimensional smooth manifold ; and let $V(M)$ denote the Lie algebra of smooth vector fields on M. A basic non representation theorem concerning the W_n states :

3.2. Theorem, [22]. Let $\alpha : W_n \to V(M)$ or $W_n/\mathbb{R}.1 \to V(M)$ be an homomorphism of Lie algebras where $n \geq 1$ and M finite dimensional. Then $\alpha = 0$.

Below I shall first try to indicate how this theorem applies to filtering problems and then proceed to discuss related matters linking the Kalman-Bucy filter of linear system theory to the Segal-Shale-Weil representation of quantum field theory and number theory.

Before doing so let me remark that the present proof of the theorem ([22]) is highly computational and that a more conceptual proof would be a very desirable thing to have. Such a proof has now been given by Toby Stafford (remark added in proof).

3.3. The recursive filtering problem. Consider an Ito stochastic dynamical system of the form :

$$(3.4) \qquad dx_t = f(x_t)dt + g(x_t)dw_t \ , \ dy_t = h(x_t)dt + dv_t, \ x_t \in \mathbb{R}^n, \ y_t \in \mathbb{R}^p$$

where f,g,h are vector and matrix valued functions of x of the right dimensions and where $w_t \in \mathbb{R}^m$, $v_t \in \mathbb{R}^p$ are independant Wiener noise processes also independant of the initial random variable x_o.

The filtering problem is to find the best estimate $\hat{x}_t$ of x_t given the observations y_s, $0 \leq s \leq t$. Mathematically (but not constructively or computationally) the answer is given by the conditional expectation $\hat{x}_t = E[x_t | y_s, 0 \leq s \leq t]$.

More precisely what we would like to have is a recursive finite dimensional filter for calculating $\hat{x}_t$ or for calculating the best estimate $\widehat{\varphi(x_t)}$ of some interesting function $\varphi(x_t)$ of x_t. By definition such a filter is a "system" of the form :

$$d\xi_t = \alpha(\xi_t)\,dt + \sum_r \beta_r(\xi_t)\,dy_{rt},\ \varphi(x_t) = \gamma(\xi_t)\ ,\ \xi_t \in M \tag{3.5}$$

where α, β_j are known vector fields on the finite dimensional manifold M, y_{rt} is the r-th component of y_t, and γ is a known function on M. The stochastic DE in (3.5) is assumed to be in Fisk-Stratonovic form.

Such filters for $\hat{x}_t$ exist in the case of linear stochastic systems :

$$dx_t = Ax_t\,dt + B\,dw_t\ ,\ dy = Cx_t dt + dv_t. \tag{3.6}$$

In this case there is a filter for $\hat{x}_t$ (the Kalman-Bucy filter) which evolves on $\mathbb{R}^N$, $N = n + \frac{1}{2}n(n+1)$, as follows. A point ξ in $\mathbb{R}^N$ is interpreted as a pair (n,P) with $m \in \mathbb{R}^n$ and P a symmetric $n \times n$ matrix. The evolution equations for $\xi = (m,P)$ are now :

$$\begin{aligned} dP_t &= (AP_t + P_t A^T + BB^T - P_t C^T C P_t)dt \\ dm_t &= Am_t\,dt + P_t C^T (dy_t - Cm_t\,dt) \end{aligned} \tag{3.7}$$

where the upper T stands for transposes. The output map γ is the projection $(m,P) \to m$. For an introduction to the Kalman-Bucy filter cf. e.g. [3]. E.g. in the case of the simplest possible non zero system : Wiener noise linearly observed :

$$dx_t = dw_t\ ,\ dy_t = x_t\,dt + dv_t \tag{3.8}$$

the Kalman-Bucy filter is given by the equations :

$$dP = (1-P^2)dt\ ,\ dm = P dy_t - mPdt \tag{3.9}$$

so that the vectorfields α and β of this filter are equal to :

$$\alpha = (1-P^2)\frac{\partial}{\partial P} - mP\frac{\partial}{\partial m}\ ,\quad \beta = P\frac{\partial}{\partial m} \tag{3.10}$$

The Kalman-Bucy filter is of enormous importance for applications. But not all phenomena can be modelled well by linear systems and thus ever since 1961 (the year of birth of the Kalman-Bucy filter) there has been a search for recursive filters in non linear situations. Recently a new approach to this question has been initiated by Brockett and Mitter ([23],[24]) which I shall sketch now.

3.11. <u>The Duncan-Mortenson-Zakai equation</u>. The first ingredient in this new approach is the socalled Duncan-Mortenson-Zakai equation. Let $p(x,t)$ be the density of $\hat{x}_t$ (assumed to exist). Then an unnormalized version $\rho(x,t)$ of $p(x,t)$ (i.e. $\rho(x,t) = n(t)\, p(x,t)$ for some, unknown, function $n(t)$) satisfies the DMZ equation (in Fisk-Stratonovic form) :

$$(3.12) \qquad d\rho(x,t) = \mathcal{L}\rho(x,t)\,dt + \sum_{r=1}^{p} h_r(x)\ \rho(x,t)\,dy_{rt}$$

where $\mathcal{L}$ is the second order differential Fokker-Planck operator defined by :

$$(3.13) \qquad \mathcal{L}(\bullet) = \frac{1}{2} \sum_{i,j} \frac{\partial^2}{\partial x_i \partial x_j} ((g\, g^T)_{ij}\bullet) - \sum_i \frac{\partial}{\partial x_i} (f_i \bullet) - \frac{1}{2} \sum_r h_r\ (x)^2 \bullet\ ,$$

where y_{rt}, $h_r(x)$, $(gg^T)_{i,j}$, f_i are respectively the r-th, r-th , (i,j)-th, i-th components of y_t, $h(x)$, gg^T, f. E.g in case of the system (3.8) the operator becomes :

$$\frac{1}{2}\ \frac{d^2}{dx^2} - \frac{1}{2}\, x^2.$$

<u>The estimation Lie algebra</u>. Now $\mathcal{L}$ and multiplication by $h_r(x)$ are linear ors on the space of functions. Thus apart from being infinite dimensional looks like a socalled bilinear system, that is a system of the form :

$$\dot{x} = Ax + \sum_r (B_r\, x)\ u_r.$$

For this type of system it is known that the Lie algebra generated by the matrices A and $B_1,\dots,B_r$ has much to say about how hard it is to write down solutions of (3.15).

This is one bit of motivation for considering the Lie algebra of differential operators generated by the operators occuring in the DMZ equation, and to define $L(\Sigma)$, the estimation Lie algebra of a system (Σ) of type (3.4) as the Lie algebra generated by the operator $\mathcal{L}$ of (3.13) and the operators "multiplication with $h_r(x)$."

In the case of the example (3.8) $L(\Sigma)$ is generated by $\frac{1}{2}\frac{d^2}{dx^2} - \frac{1}{2}x^2$ and x so that $L(\Sigma)$ is the socalled oscillator Lie algebra with basis $\frac{1}{2}\frac{d^2}{dx^2} - \frac{1}{2}x^2$, x, $\frac{d}{dx}$, 1.

3.16. Filters and homomorphisms of Lie-algebras. Now assume that there exists a filter (3.5) for $\hat{\varphi}(x_t)$. Then we have two ways of calculating $\hat{\varphi}(x_t)$, an infinite dimensional one and the assumed finite dimensional one. The infinite dimensional filter consists of the DMZ equation (3.12) combined with the output map :

$$\hat{\varphi}(x_t) = \frac{\int \varphi(x)\ \rho(x,t)\ dx}{\int \rho(x,t)\ dx}$$

which is a perfectly good output map γ on the space of all unnormalized densities.

Thus we have two machines which when fed the same data y_s, $0 \leq s \leq t$, produce the same results $\widehat{\varphi(x_t)}$. If both were finite dimensional realization theory tells us roughly that there exists a differentiable map of the reacheable part of the one system to the other system taking the vector fields of the one to the ones of the other. It is not unreasonable to expect that a similar result holds for (certain) infinite dimensional systems and in certain cases this has been proved to be the case [25,26]. This implies in particular that there is an homomorphism of Lie algebras of the Lie algebra generated by the vectorfields of the first system to the Lie algebra generated by the vectorfields of the second system.

In our particular case this means that the existence of a filter (3.5) would imply that the map $\mathcal{L} \mapsto \alpha$, $h_r(x) \to \beta_r$ defines an anti-homomorphism of Lie algebras $L(\Sigma) \mapsto V(M)$. It becomes an anti-homomorphism because the map which assigns to a linear operator on a linear space the linear vector-field defined by that operator ($(a_{ij}) \to \Sigma\, a_{ij}\, x_j \frac{\partial}{\partial x_i}$ on $\mathbb{R}^n$) is an anti-isomorphism.

Thus for example the existence of the Kalman-Bucy filter implies in the case of system (3.8) that :

$$\frac{1}{2}\frac{d^2}{dx^2} - \frac{1}{2}x^2 \mapsto (1-P^2)\frac{\partial}{\partial P} - m\,P\frac{\partial}{\partial m}\ ,\ x \to P\frac{\partial}{\partial m}$$, defines an anti-homomorphism of Lie algebras from the oscillator algebra to the Lie algebra $V(\mathbb{R}^2)$. This is easily verified.

3.17. Robustness questions. As it stands the DMZ equation (3.12) is a stochastic partial differential equation and thus its solution need not be defined for every sample path $y_t(w)$. Yet in practice we will have one particular sample path which moreover will belong to the class of functions of bounded variation which is of measure zero. What is needed is a transformed version of (3.12) which makes sense for each separate sample path and which moreover is continuous with respect to varying sample paths : a robust version. In fact for a proof of the theorem that a filter induces an homomorphism of Lie algebras $L(\Sigma) \to V(M)$ along the lines used in [25] for certain special cases we seem to need smooth dependence of the solutions on the sample path.

3.18. The cubic sensor. The cubic sensor is the one dimensional system :

$$dx_t = dw_t \ , \ dy_t = x_t^3 \, dt + dv_t. \tag{3.19}$$

3.20 Theorem ([22]. L(cubic sensor) = W_1

That is the estimation Lie algebra of the cubic sensor is maximally large. This seems to be a generic phenomenon and I conjecture the following. Consider only polynomial f, g, h in (3.4), then the estimation Lie algebra will generically be equal to all of W_n.

In the case of the cubic sensor we can show that a filter would give rise to an homomorphism of Lie algebras [25],[26] and thus it follows from theorem 3.2 that :

3.21. Corollary. The only statistics of the cubic sensor which can be computed by a finite dimensional recursive filter are constants.

3.22. Ideals of the estimation algebra, filters, and approximate filters. One also expects the structure of the estimation algebra to help in finding actual filters ; that is one expects that suitable homomorphisms of Lie algebras : $L(\Sigma) \to V(M)$ will indeed give rise to filters. Thus it becomes relevant whether perhaps $L(\Sigma)$ has a series of ideals $I_1 \supset I_2 \supset I_3 \supset \ldots$ such that $\cap I_i = \{0\}$ and $L(\Sigma)/I_i$ is finite dimensional for all i. This is e.g. the case if the system (3.4) is analytic and $f(0) = g(0) = 0$,[22].

A largely unexplored question is whether $L(\Sigma)$ also contains information on approximate filters. One could expect e.g. approximate filters to have to do with partial homomorphisms of Lie algebras ; that is linear mappings which respect Lie bracketts up to a certain order. The easiest way to formalize this is perhaps as follows. The Lie algebra $L(\Sigma)$ is not just a Lie algebra but a Lie algebra with prescribed generators $\mathcal{L}$, $h_1(x),\dots,h_r(x)$. Introducing an extra variable s we can consider the filtered Lie algebra $L_s(\Sigma)$ generated by the operators $s\mathcal{L}, sh_1(x),\dots,sh_r(x)$ and consider homomorphisms of this Lie algebra. If one can show that the DMZ equation (or rather a robust version) is stable in the sense that the higher bracketts between the generators have but small influence one would expect representations of the finite dimensional quotients $L_s(\Sigma)$ mod s^n to be relevant for approximate recursive filters. Some positive evidence in this direction is contained in [27] and [28] for certain cubic sensor like systems and in [29] where it is shown that the extended Kalman filter in a particular case corresponds to partial homomorphisms of Lie algebras.

3.23 The Lie algebra $\underline{ls}_n$. Let $\underline{ls}_n \subset W_n$ denote the Lie algebra spanned by all differential operators :

$$x^a \frac{\partial^b}{\partial x^b} \tag{3.24}$$

with $|a| + |b| \leq 2$. Here a,b are multiindices and $\partial = \partial_1+\dots+\partial_n$, $\partial_i \in \mathbb{N}\cup\{0\}$. It is easily checked that this vector space is a sub-Lie-algebra of W_n. It is also a maximal sub-Lie-algebra. I call it the linear systems Lie algebra for reasons which will become clear below. The elements (3.24) with $|a|+|b| \leq 1$ span an ideal $\underline{h}_n$ in $\underline{ls}_n$ which is of course the n-dimensional Heisenberg Lie algebra and the quotient is easily shown to be isomorphic to the symplectic Lie algebra $\underline{sp}_n$, so that there is an exact sequence :

$$0 \longrightarrow \underline{h}_n \longrightarrow \underline{ls}_n \longrightarrow \underline{sp}_n \longrightarrow 0$$

.25. The representation defined by all Kalman-Bucy filters Consider a linear ystem (3.6). One easily checks that in this case the operators occuring in the Z equation are all in $\underline{ls}_n$. The estimation Lie algebras of linear systems (A,B,C) are quite small in that $L(A,B,C) \cap \underline{h}_n$ is always an ideal of codimension 1 in the estimation Lie algebra $L(A,B,C)$. But for varying A,B,C the $L(A,B,C)$ do span all of $\underline{ls}_n$.

Now the Kalman-Bucy filter for a linear system (A,B,C) defines an anti-homomorphism of Lie algebras $L(A,B,C) \longrightarrow V(\mathbb{R}^N)$, $N = n+\frac{1}{2}n(n+1)$, and by adding one extra dimension (for the normalization factor essentially) one can lift this to injective anti-representation :

$$\rho_{(A,B,C)} : L(A,B,C) \longrightarrow V(\mathbb{R}^{N+1})$$

3.26. Theorem ([30]). The anti-representations $\rho_{(A,B,C)}$ fit together to define an anti-representation of all $\underline{ls}_n$.

This gives in particular a representation of $\underline{sp}_n$ in $V(\mathbb{R}^{N+1})$ via a Levi factor of $\underline{sp}_n$ in $\underline{ls}_n$. It now turns out,[30], that this representation is closely related to the socalled Segal-Shale-Weil representation ([31-33]) of quantum field theory. One is a complex version of the other, which also throws extra light on why the Kalman-Bucy representation can not be integrated in all directions (only a certain cone) while the SSW representation (also called oscillator representation) can be integrated to a representation of the simply connected cover $\tilde{Sp}_n$ of Sp_n.

REFERENCES

[1] M. WONHAM : On pole assignment in multi-input controllable linear systems IEEE Trans. AC 12 (1967), 660-665.

[2] C. BYRNES : Stabilizability of multivariable systems and the Ljusternik-Snirel'man category of real grassmannians, preprint 1981.

[3] J.C. WILLEMS : The LQG-problem : a brief tutorial introduction. In [34] 29-44.

[4] M. HAZEWINKEL : On families of linear systems : degeneration phenomena. In C. Byrnes, C.F. Martin (eds), Linear system theory, AMS, 1980, 157-190.

[5] M. HAZEWINKEL : (Fine) moduli (spaces) for linear systems : what they are and what they are good for, In : C. Byrnes, C.F. Martin (eds), Geometric methods for the theory of linear systems, Reidel, 1980, 125-193

[6] M. HAZEWINKEL : A partial survey of the uses of algebraic geometry in systems and control theory. In : Symposia Math. INDAM. 24, Acad. Press, 1981. 245-292.

[7] R.E. KALMAN . P.L. Falb. M.A. Arbib, Topics in mathematical system theory, Mac Graw-Hill, 1969.

[8] M. HAZEWINKEL : On the (internal) symmetry group of linear dynamical systems. In P. Kramer, M. Dal Cin (eds), Groups, systems and many-body physics, Vieweg, 1980, 362-404.

[9] A. LINDQUIST-G. PICCI : State space models for gaussian stochastic processes. In [34], 141-168.

[10] M. HAZEWINKEL : Moduli and canonical forms for linear dynamical systems II : the topological case, Math. Syst. Th. 10 (1977), 363-385.

[11] C. BYRNES : On the control of certain deterministic infinite dimensional systems by algebro-geometric techniques, Amer J. Math. 100 (1978), 1333-1381

[12] A. TANNEMBAUM : Invariance and system theory, Lect. Notes Math. 845, Springer, 1981.

[13] V. DLAB-P. GABRIEL (eds) : Representation theory I, II, Lect. Notes Math. 831, 832 , Springer, 1981.

[14] M. HAZEWINKEL - C.F. MARTIN : A short elementary proof of Grothendieck's theorem on algebraic vectorbundles over the projective line, J. pure and appl. Algebra 25 (1982), 207-212.

[15] M. HAZEWINKEL, A.C.F. Vorst, On the Snapper, Lübler-Vidale, Lam theorem on permutation representations of the symmetric groups, J. pure and appl. Algebra 23 (1982), 29-32.

[16] L. HARPER-G.C. ROTA : Matching theory : an introduction. In : P. Ney (ed.), Adv. in Probability, Vol. 1, Marcel Dekker, 1971, 171-215.

[17] M. HAZEWINKEL - C.F. MARTIN : Representations of the symmetric group, the specialization order, systems and Grassmann manifolds, to appear Ens. Math.

[18] E. RUCH : The diagram lattice as structural principle. Theor. Chim. Acta 38 (1975), 167-183.

[19] E. RUCH-A. MEAD : The principle of increasing mixing character and some of its consequences, Theor. Chim. Acta 41 (1976), 95-117.

[20] P.M. ALBERTI-A. UHLMANN : Stochasticity and partial order, Reidel, 1982.

[21] M. HAZEWINKEL-C.F. MARTIN : On decentralization, symmetry and special structure in linear systems, submitted J. Pure and Appl. Algebra.

[22] M. HAZEWINKEL-S.I. MARCUS : On Lie algebras and finite dimensional filtering. Stochastics 7 (1982), 29-62.

[23] R.W. BROCKETT : Non-linear systems and non-linear estimation theory. In [34] , 441-478.

[24] S. K. MITTER, Non-linear filtering ans stochastic mechanics. In [34], 479-504.

[25] M. HAZEWINKEL-S.I. MARCUS-H.J. SUSSMANN : Non existence of finite dimensional filters for conditional statistics of the cubic sensor problem, preprint, 1982.

[26] H.J. SUSSMANN : Rigorous results on the cubic sensor problem, In [34], 637-648.

[27] M. HAZEWINKEL : On deformation, approximations and non-linear filtering, Systems and control letters 1 (1981), 32-36.

[28] H.J. SUSSMAN : Approximate finite dimensional filters for some non-linear problems, Systems and Control letters, to appear.

[29] B. HANZON-M. HAZEWINKEL : On identification of linear systems and the estimation Lie algebra of the associated non-linear filtering problem. In : Proc. 6-th IFAC Conf. on Identification, Washington, June 1982.

[30] M. HAZEWINKEL : The linear systems Lie algebra, the Segal-Shale-Weil representation and all Kalman-Bucy filters, Subm. Amer. J. Math.

[31] D. SHALE : Linear symmetries of free boson fields. Trans. Amer. Math. Soc. 103 (1962), 149-167.

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

[33] A. WEIL : Sur certains groupes d'opérateurs unitaires. Coll. works Vol.III 1-71, Springer, 1980.

[34] M. HAZEWINKEL-J.C. WILLEMS (eds.), Stochastic systems : the mathematics of filtering and identification and applications, Reidel, 1981.

[35] M. HAZEWINKEL-C.F. Martin : Symmetric linear systems : an application of algebraic system theory, submitted IEEE Trans. AC.

[36] J.C. WILLEMS : Almost invariant subspaces : an approach to high gain feedback design, I, II, IEEE Trans. AC 26 (1981), 235-252, ibid to appear.