Netherlands school of economics

ECONOMETRIC INSTITUTE

ON THE (INTERNAL) SYMMETRY GROUPS OF LINEAR DYNAMICAL SYSTEMS

M. HAZEWINKEL

REPORT 7811/M

ERASMUS UNIVERSITY ROTTERDAM, P.O. BOX 1738, ROTTERDAM, THE NETHERLANDS

ON THE (INTERNAL) SYMMETRY GROUPS OF LINEAR DYNAMICAL SYSTEMS

by

M. Hazewinkel

ABSTRACT

Let $\dot{x} = Fx + Gu, y = Hx$, $u \in \mathbb{R}^m$, $y \in \mathbb{R}^p$, $x \in \mathbb{R}^n$ be a linear dynamical system of state space dimension n with m inputs and p outputs. The input-output operator $f(\Sigma)$ associated to this system Σ, $u(t) \mapsto y(t) = \int_0^t He^{F(t-\tau)}Gu(\tau)d\tau$, is invariant under the following action of $GL_n(\mathbb{R})$: $(F,G,H)^S = (SFS^{-1}, SG, HS^{-1})$, $S \in GL_n(\mathbb{R})$. Thus the external description of Σ by means of the operator $f(\Sigma)$ is degenerate, much as e.g. in atomic physics an energy level may be degenerate. Or, again, there is an (internal) symmetry group, viz. $GL_n(\mathbb{R})$. This paper, which will be a chapter in a forthcoming book on "Groups in many body physics and systems" (to be published by Vieweg) is concerned with those aspects of the theory of linear dynamical systems which immediately relate to the presense of this symmetry group (or degeneracy). The paper is mainly expository, though it does contain some new results (e.g. on how to "split" the degeneracy mentioned above) and some new proofs.

CONTENTS

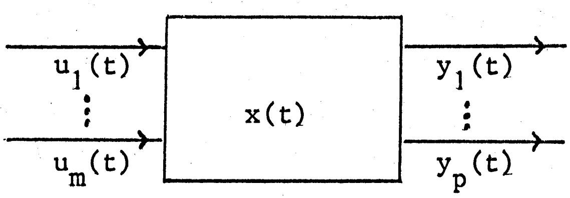
May 13, 1978 Preliminary and Confidential

1. INTRODUCTION AND STATEMENT OF THE MAIN DEFINITIONS AND RESULTS.

A time invariant linear dynamical system is a set of equations

$$\text{(1.1)}\qquad \begin{array}{lcl} \dot{x} = Fx + Gu & & x(t+1) = Fx(t) + Gu(t) \\ y = Hx & & y(t) = Hx(t) \\ \text{(continuous time)} & (\Sigma) & \text{(discrete time)} \end{array}$$

where $x \in X = \mathbb{R}^n$, $u \in U = \mathbb{R}^m$, $y \in Y = \mathbb{R}^p$ and where F, G, H are matrices with coefficients in $\mathbb{R}$ of the dimensions $n \times n$, $n \times m$, $p \times n$ respectively. We speak then of a system of dimension n, $\dim(\Sigma) = n$, with m inputs and p outputs. Of course the discrete time case also makes sense over any field k, (instead of $\mathbb{R}$). The spaces X,U,Y are respectively called state space, input space and output space. The usual picture is a "black box".

(1.2)


That is the system Σ is viewed as a machine which transforms an m-tuple of input or control functions $u_1(t), \ldots, u_m(t)$ into a p-tuple of output or observation functions $y_1(t), \ldots, y_p(t)$. The formulas expressing $y(t)$ in terms of the $u(t)$ are

$$\text{(1.3)}\qquad y(t) = He^{Ft}x(0) + \int_0^t He^{F(t-\tau)}Gu(\tau)d\tau,$$

$$y(t) = HF^t x(0) + \sum_{i=0}^{t-1} HF^{t-i-1}Gu(i)$$

where $x(0)$ is the state of the system at time 0 (and where we start putting in input at time $t = 0$). Thus the input-output behaviour of our box depends of course on the initial state $x(0)$. One is particularly interested in the input-output behaviour of Σ when $x(0) = 0$. We shall write $f(\Sigma)$ for the associated input-output operator. Thus

$$(1.4)\quad f(\Sigma): u(t) \mapsto \int_0^t He^{F(t-\tau)}Gu(\tau)d\tau,\quad f(\Sigma): u(t) \mapsto \sum_{i=0}^{t-1} HF^{t-i-1}Gu(i)$$

It is now an important fact that the input-output behaviour description of the machine (1.2) is degenerate much as, say, energy levels in atomic physics may be degenerate. More precisely the matrices F,G,H (and the initial state x(0)) depend on the choice of a basis in state space and from the input-output behaviour of the machine there is (without changing the machine) no way of deciding on a "canonical" basis for the state space $X = \mathbb{R}^n$. More mathematically we have the following. Let $GL_n(\mathbb{R})$ be the group of all invertible real $n \times n$ matrices and let $L_{m,n,p}(\mathbb{R})$ be the space of all triples of matrices (F,G,H) of dimensions $n \times n$, $n \times m$, $p \times n$ respectively. The group $GL_n(\mathbb{R})$ acts on $L_{m,n,p}(\mathbb{R})$ and $\mathbb{R}^n$ = space of initial states as

$$(1.5)\qquad (F,G,H)^S = (SFS^{-1}, SG, HS^{-1}),\ x(0)^S = Sx(0)$$

and as is easily checked the associated input-output behaviour of the corresponding machine as given by (1.3) and (1.4) is invariant under this action of $GL_n(\mathbb{R})$; i.e., in particular $f(\Sigma^S) = f(\Sigma)$. This action corresponds to base change in state space. Indeed if $x' = Sx$ and $\dot{x} = Fx + Gu$, $y = Hx$ then $S^{-1}\dot{x}' = FS^{-1}x' + Gu$, $y = HS^{-1}x'$ so that $\dot{x}' = SFS^{-1}x' + SGu$, $y = HS^{-1}x'$ and $x'(0) = Sx(0)$.

This chapter is concerned with those aspects of the theory of linear dynamical systems which are more or less directly related to the presence of the internal symmetry group $GL_n(\mathbb{R})$ of the internal description of linear dynamical systems by triples of matrices (cf. (1.1)) as compared to the degenerate external description by means of the operator $f(\Sigma)$ (or (1.3)). This is not really a research paper (though it does in fact contain a few new results) but rather a graduate level expository account of some of the material of [3 - 8] and immediately related matters.

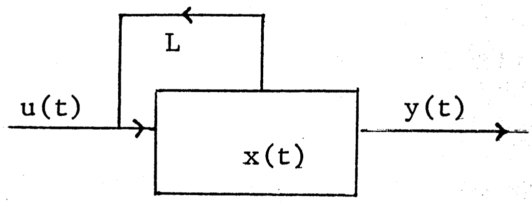
In the remaining part of this introduction we give a slightly informal description of most of the main results of sections 2-8 below.

We shall concentrate on the continuous time case.

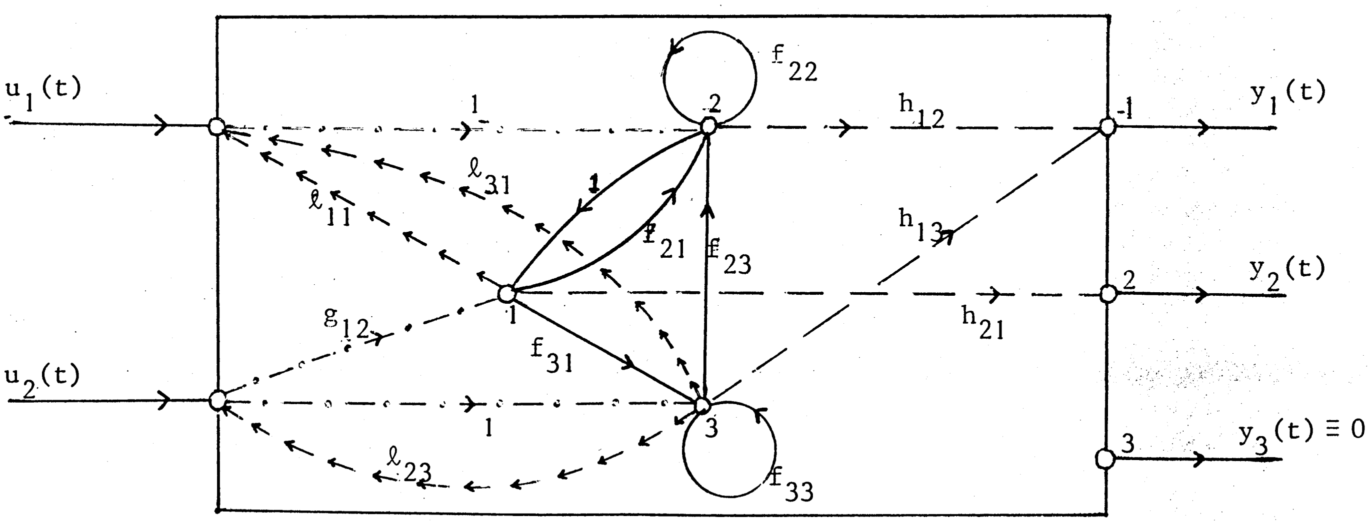
1.6. Feedback and how to resolve the external description degeneracy.

In the case of atomic physics a degenerate energy level may be split by means of, e.g., a suitable magnetic field. One can ask whether there exist something analogous in our case of degenerate external (=observable) descriptions of linear dynamical systems. There does in fact exist some such thing. It is called state space feedback. Consider the system (1.1). Introduction of state space feedback L changes it to the system $\Sigma(L)$

$$\dot{x} = (F+GL)x + Gu \qquad (1.7)$$
$$y = Hx$$



In thinking about these things the author has found it helpful to visualize a linear dynamical system with (variable) feedback as a set of n-integrators, 1, ..., n, interconnected by means of the matrix F, a set of m input ports connected to the integrators by means of the matrix G, a set of p output ports connected to the integrators by means of the matrix H and a set of connections from the integrators to the input ports (feedback) which maybe varied in strength by the experimentator (as in atomic physics the splitting magnetic field may be varied). Cf. also the picture below.



interconnections between the integrators as given by the matrix F.

$$F = \begin{pmatrix} 0 & 1 & 0 \\ f_{21} & f_{22} & f_{23} \\ f_{31} & 0 & f_{33} \end{pmatrix}$$

connections from the input ports to integrators as given by the matrix G

$$G = \begin{pmatrix} 0 & g_{12} \\ 1 & 0 \\ 0 & 1 \end{pmatrix}$$

connections from the integrators to the output ports as given by the matrix H

$$H = \begin{pmatrix} 0 & h_{12} & h_{13} \\ h_{21} & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

connections from the integrators to the input ports (can be varied in strength by the experimentator) as given by the matrix L

$$L = \begin{pmatrix} \ell_{11} & 0 & \ell_{13} \\ 0 & 0 & \ell_{23} \end{pmatrix}$$

Now let $\Sigma = (F,G,H)$ and $\Sigma' = (F',G',H')$ be two linear dynamical systems, and suppose that Σ and Σ' are completely reachable and completely observable. (This is an entirely natural restriction in this context, cf. 1.9 below; for a precise definition of these notions, cf. 2.1 below). Suppose that $\Sigma \neq \Sigma'$ but $f(\Sigma) = f(\Sigma')$. Let $\Sigma(L)$, $\Sigma'(L)$ be the systems obtained by introducing the feedback L, i.e $\Sigma(L) = (F+GL,G,H)$, $\Sigma'(L) = (F'+G'L,G',H')$. Then there is a suitable feedback matrix L, which can be taken arbitrarily small (so that $\Sigma(L)$ and $\Sigma'(L)$ are still completely reachable and observable) such that $f(\Sigma(L)) \neq f(\Sigma'(L))$. I.e. feedback splits the $GL_n(\mathbb{R})$ - degenerate external description of linear dynamical systems.

1.8. Realization theory. Let Σ be a linear dynamical system (1.1). Then, if we leave Σ unchanged, from our observations we can deduce the operator $f(\Sigma)$ or, equivalently, we can find the sequence of matrices $\mathcal{A}(\Sigma) = (A_o,A_1,A_2,\ldots)$, $A_i = HF^iG$. To obtain these use δ-functions and derivates of δ-functions as inputs. Another way to see this is to apply

Laplace transforms to (1.1). This gives

(1.9) $$s\hat{x}(s) = F\hat{x}(s) + G\hat{u}(s) \;,\; \hat{y}(s) = H\hat{x}(s)$$

so that the relation between the Laplace transforms $\hat{y}(s),\hat{u}(s)$ of the outputs $y(t)$ and inputs $u(t)$ is given by multiplication with the socalled transfer matrix $T(s)$

(1.10) $$\hat{y}(s) = T(s)\hat{u}(s) \;,\; T(s) = H(s-F)^{-1}G$$

The power series development of $T(s)$ in powers of s^{-1} (around $s = \infty$) is now

(1.11) $$T(s) = A_o s^{-1} + A_1 s^{-2} + A_2 s^{-3} + \dots$$

The question now naturally arises: when does a sequence of $p \times m$ matrices $\mathcal{A} = (A_o, A_1, \dots)$ come from a linear dynamical system (1.1), or, as we shall say, when is $\mathcal{A}$ <u>realizable</u>.

1.12. <u>Theorem</u> (i) If $\mathcal{A}$ is realizable by an n-dimensional system Σ then it is also realizable by an $n' \leq n$ dimensional system Σ' which is moreover completely reachable and completely observable.

(ii) The sequence $\mathcal{A}$ is realizable by an n dimensional system Σ if and only if $\mathrm{rank}(\mathcal{H}_s(\mathcal{A})) \leq n$ for all $s \in \mathbb{N} \cup \{0\}$.

Here $\mathcal{H}_s(\mathcal{A})$ is the block Hankel matrix

$$\mathcal{H}_s(\mathcal{A}) = \begin{pmatrix} A_o & A_1 & \cdots & A_s \\ A_1 & & \cdot & \vdots \\ \vdots & \cdot & & \vdots \\ A_s & \cdots & & A_{2s} \end{pmatrix}$$

1.13. <u>Invariants and the structure of</u> $M^{cr,co}_{m,n,p}(\mathbb{R}) = L^{co,cr}_{m,n,p}(\mathbb{R})/GL_n(\mathbb{R})$.

Let $L_{m,n,p}(\mathbb{R})$ be the space of all triples of matrices (F,G,H) of dimensions $n \times n$, $n \times m$, $p \times n$ respectively. The group $GL_n(\mathbb{R})$ acts on $L_{m,n,p}(\mathbb{R})$ as in (1.5). The input-output matrices $A_i = HF^iG$ are clearly invariants for this action and the question arises whether these are the only invariants. Here an invariant is defined as a function

$\rho\colon L_{m,n,p}(\mathbb{R}) \to \mathbb{R}$ (or possibly a function defined on an invariant open dense subset of $L_{m,n,p}(\mathbb{R})$) such that $\rho((F,G,H)^S) = \rho(F,G,H)$ for all triples (F,G,H) (in the open dense subset).

1.14. Theorem. Every invariant of $GL_n(\mathbb{R})$ acting on $L_{m,n,p}(\mathbb{R})$ is a function of the entries of $A_o, \ldots, A_{2n-1}$.

Let $L^{co,cr}_{m,n,p}(\mathbb{R})$ be the subspace of all triples $(F,G,H) \in L_{m,n,p}(\mathbb{R})$ which are both completely observable and completely reachable. This is an open and dense subspace of $L_{m,n,p}(\mathbb{R})$. On this subspace $GL_n(\mathbb{R})$ acts faithfully and a more precise version of theorem 1.14 describes the quotient space $M^{co,cr}_{m,n,p}(\mathbb{R}) = L^{co,cr}_{m,n,p}(\mathbb{R})/GL_n(\mathbb{R})$ explicitly and gives an algorithm for recovering (F,G,H) up-to-$GL_n(\mathbb{R})$-equivalence from $A_o, \ldots, A_{2n-1}$ (cf. 4.25 below). It turns out that $M^{co,cr}_{m,n,p}(\mathbb{R})$ is a smooth differentiable manifold and that the projection $L^{co,cr}_{m,n,p}(\mathbb{R}) \to M^{co,cr}_{m,n,p}(\mathbb{R})$ is a principal $GL_n(\mathbb{R})$-bundle (cf. 6.4 below).

1.15. Canonical forms. For many purposes (prediction, construction of feedbacks, identification and, not least, for proving theorems) an internal description of a black box by means of a triple of matrices (F,G,H) is preferable over knowledge of the input-output operator $f(\Sigma)$. As was remarked in section 1.13 above there do exist algorithms for calculating some $\Sigma = (F,G,H)$ which realizes $f(\Sigma)$ or $\mathcal{A}(\Sigma)$ from the matrices $A_o, \ldots, A_{2n-1}$. One such algorithm is described in 4.25 below. All these algorithms have the drawback that they are discontinuous in general. This is a nontrivial difficulty, because after all one calculates the (F,G,H) because one wants to use them as a basis for further calculations, design, predictions etc., and the $A_o, \ldots, A_{2n-1}$ are after all subject to (small) measurement errors. Thus the question arises whether there exist continuous methods of recovering (F,G,H) up-to-$GL_n(\mathbb{R})$-equivalence from $A_o, \ldots, A_{2n-1}$. Or, in other words, because $M^{co,cr}_{m,n,p}(\mathbb{R})$ is explicitly describable subspace of the space of all sequences of $2n$ $p \times m$ matrices and $M^{co,cr}_{m,n,p}(\mathbb{R}) = L^{co,cr}_{m,n,p}(\mathbb{R})/GL_n(\mathbb{R})$, the question arises whether there exist continuous canonical forms on $L^{co,cr}_{m,n,p}(\mathbb{R})$, where a continuous canonical form is defined as follows.

1.16. Definition. A continuous canonical form on a $GL_n(\mathbb{R})$-invariant subspace $L' \subset L_{m,n,p}(\mathbb{R})$ is a continuous map $c\colon L' \to L'$ such that

(i) $c((F,G,H)^S) = c((F,G,H))$ for all $(F,G,H) \in L'$,

(ii) if $c((F,G,H)) = c((F',G',H'))$ then there is a $S \in GL_n(\mathbb{R})$ such that $(F',G',H') = (F,G,H)^S$, and

(iii) for all $(F,G,H) \in L'$ there is an $S \in GL_n(\mathbb{R})$ such that $c(F,G,H) = (F,G,H)^S$.

For some additional remarks on the desirability of continuous canonical forms cf. [2] and also [15]. Also our proof of "feedback suspends degeneracy" theorem mentioned in 1.6 above is based on the use of a suitable canonical form. It turns out that there exist open dense subspaces $U_\alpha \subset L_{m,n,p}(\mathbb{R})$, which together cover $L^{co,cr}_{m,n,p}(\mathbb{R})$, on which canonical forms exist. Cf. 3.10 below. On the other hand

1.17. Theorem. There exists a continuous canonical form on all of $L^{co,cr}_{m,n,p}(\mathbb{R})$ if and only if $m = 1$ or $p = 1$.

1.18. On the geometry of $M^{co,cr}_{m,n,p}(\mathbb{R})$. Holes. Now suppose we have a black box (1.2) which is to be modelled by a linear dynamical system of dimension n. Then the input-output data give us a point of $M^{co,cr}_{m,n,p}(\mathbb{R})$ and as more and more data come in we find (ideally) a sequence of points in $M^{co,cr}_{m,n,p}(\mathbb{R})$ representing better and better linear dynamical system approximations to the given black box. The same thing happens when one is dealing with a slowly varying black box or linear dynamical system. If this sequence approaches a limit we have "identified" the black box. Unfortunately the space $M^{co,cr}_{m,n,p}(\mathbb{R})$ is never compact so that a sequence of points may fail to converge to anything whatever. There are holes in $M^{co,cr}_{m,n,p}(\mathbb{R})$. Consider for example the following family of 2-dimensional one input, one output systems

$$g_z = \begin{pmatrix}1\\1\end{pmatrix},\quad F_z = \begin{pmatrix}-z & -z\\ 0 & -z\end{pmatrix},\quad H_z = (z^2,0),\quad z = 1,2,3,\ \dots \tag{1.19}$$

Let $u(t)$, $0 \leq t \leq t_0$ be a smooth input function, then $y(t) = \lim_{z\to\infty} f(\Sigma_z)u(t)$ exists and is equal to $y(t) = \frac{d}{dt}u(t)$. This operator can not be of the form $f(\Sigma)$ for any system Σ of the form (1.1) (because the $f(\Sigma)$ are always bounded operators and $\frac{d}{dt}$ is an unbounded operator). A characteristic feature of this example is that the individual matrices F_z, G_z, H_z do not have limits as $z \to \infty$. (A not unexpected phenomenon, because after all we are taking quotients by the noncompact group $GL_n(\mathbb{R})$). This sort of situation is actually important in practise, e.g. in the study of very high gain state feedback systems $\dot{x} = Fx + Gu$, $u = cLx$, where c is a large scalar gain factor. Cf. [12].

Another type of hole in $M^{co,cr}_{m,n,p}(\mathbb{R})$ corresponds to lower dimensional systems, and in way these two holes and combinations of them are all the holes there are in the sense of the following definitions and theorems.

1.21. Definition. We shall say that a family of systems $\Sigma_z = (F_z, G_z, H_z)$ converges in input-output behaviour to an operator B if for every m-vector of smooth input functions u(t) we have $\lim_{z\to\infty} f(\Sigma_z)u(t) = Bu(t)$ uniformly in t on bounded t intervals.

1.22. Definition. A differential operator of order r is an operator of the form $u(t) \mapsto y(t) = Dy(t) = a_o u(t) + a_1 \frac{d}{dt} u(t) + \ldots + a_r \frac{d^r}{dt^r} u(t)$, where the $a_o, \ldots, a_r$ are $p \times m$ matrices with coefficients in $\mathbb{R}$, and $a_r \neq 0$. We write ord(D) for the order of D. By definition ord(0) = -1.

1.23. Theorem. Let $(\Sigma_z)_z$ be a family of systems in $L_{m,n,p}(\mathbb{R})$ which converges in input-output behaviour. Let B be the limit input-output operator. Then there exist a system Σ' and a differential operator D such that

$$Bu(t) = f(\Sigma')u(t) + Du(t)$$

and $\mathrm{ord}(D) + \dim(\Sigma') \leq n-1$.

1.24. Theorem. Let D be a linear differential operator and $\Sigma' \in L_{m,n,p}(\mathbb{R})$ and suppose that $\mathrm{ord}(D) + \dim(\Sigma') \leq n-1$. Then there exists a family of systems $(\Sigma_z)_z$, $\Sigma_z \in L^{co,cr}_{m,n,p}(\mathbb{R})$ such that for every smooth input vector u(t)

$$\lim_{z\to\infty} f(\Sigma_z)u(t) = f(\Sigma')u(t) + Du(t)$$

uniformly on bounded t-intervals.

1.25. Concluding introductory remarks.

Many of the results described above have their analogues in the discrete case and/or the time varying case, cf. [3-8, 9-11,14]. But not all. For instance the obvious analogues of theorems 1.23 and 1.24 fail utterly in the discrete time case. In this case $\lim_{z\to\infty} f(\Sigma_z)u(t)$ exists for all inputs u(t) if and only if the individual matrices $A_i(z) = H_z F_z^i G_z$ converge for $z \to \infty$. This means that in the case of input-output convergence the limit operator is necessarily of the form $f(\Sigma')$ for some, possibly lower dimensional, system Σ'. The same answer obtains in the continuous time case if besides input-output convergence one also requires that the F_z, G_z, H_z (or more generally the $A_i(z)$) remain bounded.

2. COMPLETE REACHABILITY AND COMPLETE OBSRVABILITY.

Let $(F,G,H) \in L_{m,n,p}(\mathbb{R})$ be a real linear dynamical system of state space dimension n, with m inputs and p outputs. We define

$$(2.1) \qquad R_s(F,G) = (G \;\; FG \;\; \ldots, \;\; F^sG), \; s = 0,1,2,\ldots, \; R(F,G) = R_n(F,G)$$

the $n \times (s+1)m$ matrices consisting of the blocks G, FG, ..., F^sG, and, dually

$$(2.2) \qquad Q_s(F,H) = \begin{pmatrix} H \\ HF \\ \vdots \\ HF^s \end{pmatrix}, \; s = 0,1,2,\ldots, \; Q(F,H) = Q_n(F,H)$$

We also define

$$(2.3) \qquad \mathcal{H}_s(F,G,H) = \mathcal{H}_s(\Sigma) = \begin{pmatrix} A_0 & A_1 & \ldots & A_s \\ A_1 & & \cdot & \vdots \\ \vdots & \cdot & & \vdots \\ A_s & & \ldots & A_{2s} \end{pmatrix} = Q_s(F,H)R_s(F,G),$$

$$s = 0,1,2,\ldots$$

where $A_i = HF^iG$, $i = 0,1,2,\ldots$.

It is useful to notice that

$$(2.4) \qquad R_k((F,G)^S) = SR_k(F,G), \; Q_k((F,H)^S) = Q_k(F,H)S^{-1}$$

where of course $(F,G)^S = (SFS^{-1},SG)$, $(F,H)^S = (SFS^{-1},HS^{-1})$. It follows that

$$(2.5) \qquad \mathcal{H}_k(\Sigma^S) = \mathcal{H}_k((F,G,H)^S) = \mathcal{H}_k((F,G,H)) = \mathcal{H}_k(\Sigma)$$

for all $S \in GL_n(\mathbb{R})$, which is of course also immediately clear from (2.3)

2.6. Definitions of complete reachability of complete observability.

The system $(F,G,H) \in L_{m,n,p}(\mathbb{R})$ is said to be completely reachable iff $\mathrm{rank}(R(F,G)) = n$. The system (F,G,H) is said to be completely observable iff $\mathrm{rank}(Q(F,H)) = n$. These are generic conditions; in fact the subspace $L^{co,cr}_{m,n,p}(\mathbb{R})$ of $L_{m,n,p}(\mathbb{R})$ consisting of all systems which are both completely reachable and completely observable is open and dense. We note that

2.10. Theorem. The pair of matrices (F,G), $F \in \mathbb{R}^{n\times n}$, $G \in \mathbb{R}^{n\times m}$ is completely reachable iff every symmetric set with multiplicities of size n occurs as the spectrum of $F + GL$ for a suitable (state feedback) matrix L.

I.e. the system (F,G,H) is cr iff we can by means of suitable state feedback arbitrarily reassign the poles of the system. For a proof cf., e.g., [18, section 2.2].

3. NICE SELECTIONS AND THE LOCAL STRUCTURE OF $L^{cr}_{m,n,p}(\mathbb{R})/GL_n(\mathbb{R})$.

3.1. Nice Selections. Let $(F,G,H) \in L_{m,n,p}(\mathbb{R})$. We use $I(n,m)$ to denote the ordered set of indices of the columns of the matrix $R(F,G)$. I.e. $I(n,m) = \{(i,j) \mid i = 0, \ldots, n;\ j = 1, \ldots, m\}$ with the ordening $(0,1) < (0,2) <\ldots< (0,m) < (1,1) <\ldots< (1,m) <\ldots< (n,1) <\ldots< (n,m)$. A nice selection $\alpha \subset I(n,m)$ is a subset of $I(n,m)$ of size $n = \dim\Sigma$ such that $(i,j) \in \alpha \Rightarrow (i-1,j) \in \alpha$ if $i \geq 1$. Pictorially we represent $I(n,m)$ as an $(n\times 1) \times m$ rectangular array of which the first row represents the indices of the columns of G, the second row the indices of the columns of FG,... etc.... We indicate the elements of a subset with crosses. The subset of the picture on the left is then a nice selection (m=4,n=5) and the subset α' of the picture on the right below is not a nice selection

$$\begin{array}{cccc} \cdot & \times & \cdot & \times \\ \cdot & \times & \cdot & \times \\ \cdot & \times & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{array} \qquad\qquad \begin{array}{cccc} \cdot & \cdot & \times & \cdot \\ \cdot & \times & \cdot & \times \\ \cdot & \cdot & \times & \times \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{array}$$

If β is a subset of $I(n,m)$ we denote with $R(F,G)_\beta$ the matrix obtained from $R(F,G)$ by removing all columns whose index is not in β. We use $L_{m,n}(\mathbb{R})$ to denote the space of all pairs of real matrices (F,G) of dimensions $n \times n$, $n \times m$ respectively.

3.2. Lemma. Let $(F,G) \in L_{m,n}(\mathbb{R})$ be a completely reachable pair of matrices. Then there is a nice selection α such that $R(F,G)_\alpha$ is invertible.

<u>Remark</u>. Complete reachability means that rank $R(F,G) = n$, so that there is in any case some subset β of size n of $I(n,m)$ such that $R(F,G)_\beta$ is invertible. The lemma says that in that case there is also a nice selection for which this holds.

<u>Proof of the lemma</u>. Define a nice subselection of $I(n,m)$ as any subset β (of size $\leq n$) such that $(i,j) \in \beta$, $i \geq 1 \Rightarrow (i-1,j) \in \beta$. Let α be a maximally large nice subselection of $I(n,m)$ such that the columns in $R(F,G)_\alpha$ are linearly independent. We shall show that $\mathrm{rank}(R(F,G)_\alpha) = \mathrm{rank}(R(F,G))$, which will prove the lemma because by assumption rank $R(F,G) = n$.
Let $\alpha = \{(0,j_1), \ldots, (i_1,j_1); \ldots; (0,j_s), \ldots, (i_s,j_s)\}$. Then by the maximality of α we know the columns of $R(F,G)$ with indices $(0,j)$, $j \in \{1,\ldots,m\} \setminus \{j_1,\ldots,j_s\}$ and the columns of $R(F,G)$ with indices (i_t+1,j_t), $t = 1, \ldots, s$ are linearly dependent on the columns of $R(F,G)_\alpha$. With induction assume that all columns with indices (i_t+k,j_t), $k \leq r$, $t = 1, \ldots, s$ and $(k-1,j)$, $k \leq r$, $j \in \{1,\ldots,m\} \setminus \{j_1,\ldots,j_s\}$ are linearly dependent on the columns of $R(F,G)_\alpha$. So we have relations

$$F^{r-1}g_j = \sum_{(i,j)\in\alpha} a(i,j)F^i g_j, \quad j \in \{1,\ldots,m\} \setminus \{j_1,\ldots,j_s\}$$

$$F^{i_t+r} g_{j_t} = \sum_{(i,j)\in\alpha} b(i,j)F^i g_j, \qquad t = 1, \ldots, s$$

where g_j denotes the j-th column of G. Multiplying on the left with F we find

$$F^r g_j = \sum_{(i,j)\in\alpha} a(i,j)F^{i+1} g_j$$
$$F^{i_t+r+1} g_{j_t} = \sum_{(i,j)\in\alpha} b(i,j)F^{i+1} g_j$$

We have already seen that the $F^{i+1}g_j$, $(i,j) \in \alpha$ are linear combinations of the column of $R(F,G)_\alpha$. It follows that also the $F^r g_j$ and $F^{i_t+r+1} g_{j_t}$ are linear combinations of the columns of $R(F,G)_\alpha$. This finishes the induction and hence the proof of the lemma.

3.3. <u>Successor indices</u>. Let $\alpha \subset I(n,m)$ be a nice selection. The successor indices of α are those elements $(i,j) \in I(n,m) \setminus \alpha$ for which $i = 0$ or for which $(i',j) \in \alpha$ for all $i' < i$ if $i \geq 1$. For every $j_o \in \{1,\ldots,m\}$ there is precisely one successor index of α of the form (i,j_o); this successor index is denoted $s(\alpha,j_o)$. In the picture below the successor indices of α are indiced by *'s (and the elements of α with x's).

Columns of G	$*$	$\times$	$*$	$\times$	x_1	e_1	x_3	e_2
Columns of $F\,G$	.	$\times$	.	$\times$	.	e_3	.	e_4
.	.	$\times$	.	$*$	.	e_5	.	x_4
.	.	$*$	.	.	.	x_2	.	
.	.	.	.	.	.	.	.	.
Columns of F^5G	.	.	.	.	.	.	.	.

3.4. <u>Lemma</u>. Let $\alpha \subset I(n,m)$ be a nice selection and $x_1, \ldots, x_m$ and m-tuple of n-vectors. Then there is precisely one pair $(F,G) \in L_{m,n}(\mathbb{R})$ such that

$$R(F,G)_\alpha = I_{n\times n}, \text{ the } n \times n \text{ unit matrix}$$

$$R(F,G)_{s(\alpha,j)} = x_j \text{ for all } j = 1, \ldots, m$$

Proof. Let f_i be the i-th column of the matrix F, $i = 1,2, \ldots, n$. Then in the example given above the values of the g_j, $j = 1, \ldots, m$ and f_i, $i = 1, \ldots, n$ can simply be read of from the diagram. One has in this case

$$g_1 = x_1,\ g_2 = e_1,\ g_3 = x_3,\ g_4 = e_2$$

$$f_1 = e_3,\ f_2 = e_4,\ f_3 = e_5,\ f_4 = x_4,\ f_5 = x_2$$

It is easy to see that this works in general and to write down the general proof though it tends to be notationally cumbersome.

3.5. <u>Local structure of $L^{cr}_{m,n,p}(\mathbb{R})/GL_n(\mathbb{R})$</u>. Let $\alpha \subset I(n,m)$ be a nice selection.

We define

$$U_\alpha = \{(F,G,H) \in L_{m,n,p}(\mathbb{R}) \mid \det R(F,G)_\alpha \neq 0\}$$

(3.6)

$$V_\alpha = \{(F,G,H) \in L_{m,n,p}(\mathbb{R}) \mid R(F,G)_\alpha = I_{n\times n}\}$$

3.7. Lemma. (i) $U_\alpha \simeq V_\alpha \times GL_n(\mathbb{R})$

(ii) $V_\alpha \simeq \mathbb{R}^{mn+np}$

Proof. (i) Let $(F,G,H) \in U_\alpha$. We assign to (F,G,H) the pair $((F,G,H)^S, S^{-1})$ where $S = R(F,G)_\alpha^{-1}$. Then $(F,G,H)^S \in V_\alpha$ because $R(SFS^{-1},SG) = SR(F,G)$ and hence $R(SFS^{-1},SG)_\alpha = SR(F,G)_\alpha$. Inversely given $((F,G,H),S) \in V_\alpha \times GL_n(\mathbb{R})$ we assign to it the element $(F,G,H)^S$. This proves (i).

Assertion (ii) follows immediately from lemma 3.4. Indeed, let $z \in \mathbb{R}^{mn+np}$ and view z as an m + p tuple of n-vectors $z = (x_1,\dots,x_m; y_1,\dots,y_p)$. Then there are unique F,G,H such that $R(F,G)_\alpha = I_{n\times n}$, $R(F,G)_{s(\alpha,j)} = x_j$, $h_\ell = y_\ell$ where h_ℓ is the ℓ-th row of H.

3.8. Local structure of $L^{co,cr}_{m,n,p}(\mathbb{R})/GL_n(\mathbb{R})$. Let again α be a nice selection. Then we define in addition.

(3.9) $$U_\alpha^{co} = U_\alpha \cap L^{co,cr}_{m,n,p}(\mathbb{R}),\ V_\alpha^{co} = V_\alpha \cap L^{co,cr}_{m,n,p}(\mathbb{R})$$

Then one has clearly that V_α^{co} is an open dense (algebraic) subset of V_α and that $U_\alpha^{co} \simeq V_\alpha^{co} \times GL_n(\mathbb{R})$.

3.10. The local nice selection canonical forms c_α. Lemma 3.7 defines us a (local) continuous canonical form on U_α for each nice selection α. It is

(3.11) $$c_\alpha((F,G,H)) = (F,G,H)^{S_\alpha} \in V_\alpha, S_\alpha = R(F,G)_\alpha^{-1}, (F,G,H) \in U_\alpha$$

The U_α are open dense subsets of $L^{cr}_{m,n,p}(\mathbb{R})$, and by lemma 3.2 the union of all the U_α, α a nice selection, covers all of $L^{cr}_{m,n,p}(\mathbb{R})$. This is thus a set of local canonical forms which can be useful in identification problems (it leads to statistically and numerically well posed problems, cf [15, section II].

3.11. The dual results. Dually we consider the set I(n,p) of all row indices of Q(F,H), which we also picture as an $(n+1) \times p$ array of dots. Now the first row represents the rows of H, the second row the rows of HF, ... A nice selection is defined as before and one has the obvious analogues of all the results given above. In particular if $(F,G,H) \in L^{co}_{m,n,p}(\mathbb{R})$ there is a nice selection $\beta \subset I(n,p)$ such that $Q(F,H)_\beta$ is invertible. Here $Q(F,H)_\beta$ is the matrix obtained from Q(F,H) by removing all rows whose index is not in β.

One also has of course local canonical forms $\bar{c}_\beta$ (defined on $\bar{U}_\beta$) for every nice selection $\beta \subset I(n,p)$:

$$\bar{c}_\beta((F,G,H)) = (F,G,H)^{S_\beta}, \; S_\beta = Q(F,H)_\beta \,, \; (F,G,H) \in \bar{U}_\beta \tag{3.12}$$

$$\bar{U}_\beta = \{(F,G,H) \in L_{m,n,p}(\mathbb{R}) \mid Q(F,H)_\beta \text{ is invertible}\} \tag{3.13}$$

4. REALIZATION THEORY.

Let $\mathcal{A} = (A_o, A_1, A_2, \ldots)$ be a sequence of $p \times m$ matrices. We shall say that the sequence $\mathcal{A}$ is realizable by an n-dimensional linear system if there exist a system $(F,G,H) \in L_{m,n,p}(\mathbb{R})$ such that $A_i = HF^iG$, $i = 0,1,2,\ldots$. It follows immediately from (the proof of) theorem 2.6 above that if $\mathcal{A}$ is realizable by means of (F,G,H), then there is also a possibly lower dimensional system $\Sigma' = (F',G',H') \in L^{co,cr}_{m,n',p}(\mathbb{R})$, $n' \leq n$. which also realizes $\mathcal{A}$ and which is moreover completely reachable and completely observable.

For each sequence of $p \times m$ matrices $\mathcal{A}$ we define the block Hankel matrices

$$\mathcal{H}_s(\mathcal{A}) = \begin{pmatrix} A_o & A_1 & \cdots & A_s \\ A_1 & & \cdot & \\ \vdots & \cdot & & \vdots \\ A_s & \cdots & & A_{2s} \end{pmatrix}, \; s = 0,1,2 \ldots \tag{4.1}$$

4.2. <u>Theorem.</u> The sequence of real $p \times m$ matrices $\mathcal{A} = (A_o, A_1, \ldots)$ is realizable by means of a completely reachable and completely observable n-dimensional system if and only if rank $\mathcal{H}_s(\mathcal{A}) = n$ for all large enough s. Moreover if both $\Sigma, \Sigma' \in L^{co,cr}_{m,n,p}(\mathbb{R})$ realize $\mathcal{A}$ then $\Sigma' = \Sigma^S$ for some $S \in GL_n(\mathbb{R})$.

This theorem will be proved below. First, however, we mention a consequence.

4.3. <u>Corollary</u>. If the sequence of $p \times m$ matrices $\mathcal{A}$ is such that rank $\mathcal{H}_s(\mathcal{A}) = n$ for all sufficiently large s, then rank $\mathcal{H}_s(\mathcal{A}) = n$ for all $s \geq n-1$.

Proof. If $\Sigma = (F,G,H)$ realizes $\mathcal{A}$ and Σ is co and cr and of dimension n, then rank $R_{n-1}(F,G) =$ rank $Q_{n-1}(F,H) = n$, so that rank $\mathcal{H}_{n-1}(\mathcal{A}) =$ rank$(R_{n-1}(F,G)Q_{n-1}(F,H)) = n$.

A first step in the proof of theorem 4.2 is now the following lemma which says that if rank $\mathcal{H}_s(\mathcal{A}) = n$ for all $s \geq r-1$, then the A_i for $i \geq 2r$ are uniquely determined by the 2r matrices $A_o, \ldots, A_{2r-1}$.

4.4. Lemma. Let $\mathcal{A} = (A_o, A_1, \ldots)$ be a series of $p \times m$ matrices such that rank $\mathcal{H}_s(\mathcal{A}) = n$ for all $s \geq r-1$. There are $m \times m$ matrices $S_o, \ldots, S_{r-1}$ and $p \times p$ matrices $T_o, \ldots, T_{r-1}$ such that for all $i = 0,1,2,\ldots$.

$$(4.5) \qquad A_{i+r} = A_i S_o + A_{i+1}S_1 + \ldots + A_{i+r-1}S_{r-1} =$$

$$= T_o A_i + T_1 A_{i+1} + \ldots + T_{r-1}A_{i+r-1}$$

Proof. Because rank $\mathcal{H}_{r-1}(\mathcal{A}) = n$ and rank $\mathcal{H}_r(\mathcal{A}) = n$ we have

$$n = \text{rank}\,\mathcal{H}_{r-1}(\mathcal{A}) = \text{rank}\left(\begin{array}{cccc|c} A_o & A_1 & \cdots & A_{r-1} & A_r \\ A_1 & & & \vdots & \vdots \\ \vdots & & & \vdots & \vdots \\ A_{r-1} & & \cdots & A_{2r-2} & A_{2r-1} \end{array}\right)$$

so that there are $m \times m$ matrices $S_o, \ldots, S_{r-1}$ such that

$$A_{i+r} = A_i S_o + \ldots + A_{i+r-1}S_{r-1}, \quad i = 0, \ldots, r-1$$

Similarly, it follows from

$$n = \text{rank}\,\mathcal{H}_{r-1}(\mathcal{A}) = \text{rank}\left(\begin{array}{ccc} A_o & \cdots & A_{r-1} \\ \vdots & & \vdots \\ A_{r-1} & \cdots & A_{2r-2} \\ \hline A_r & \cdots & A_{2r-1} \end{array}\right)$$

that there are matrices $T_o, \ldots, T_{r-1}$ such that

$$(4.6) \qquad A_{r+i} = T_o A_i + \ldots + T_{r-1}A_{i+r-1}, \quad i = 0, \ldots, r-1$$

Suppose with induction we have already proved (4.5) for $i \leq k-1$, $k \geq r$.

Consider the following submatrix of $\mathcal{H}_k(\mathcal{A})$

$$(4.7)\qquad \left(\begin{array}{cccc|ccc} A_o & A_1 & \cdots & A_{r-1} & A_r & \cdots & A_k \\ A_1 & & & \vdots & \vdots & & \vdots \\ \vdots & & & \vdots & \vdots & & \vdots \\ A_{r-1} & & \cdots & A_{2r-2} & A_{2r-1} & \cdots & A_{k+r-1} \\ \hline A_r & & \cdots & A_{2r-1} & A_{2r} & \cdots & A_{k+r} \end{array}\right)$$

Using the relations (4.5) for $i \leq k-1$ we see that the rank of 4.7 is equal to the rank of

$$(4.8)\qquad \left(\begin{array}{cccc|ccc} A_o & A_1 & \cdots & A_{r-1} & 0 & \cdots 0 & 0 \\ A_1 & & & \vdots & \vdots & \vdots & \vdots \\ \vdots & & & \vdots & \vdots & \vdots & \vdots \\ A_{r-1} & & \cdots & A_{2r-2} & 0 & \cdots 0 & 0 \\ \hline A_r & & \cdots & A_{2r-1} & 0 & \cdots 0 & X \end{array}\right)$$

where $X = A_{k+r} - S_oA_k - \ldots - S_{r-1}A_{k+r-1}$. Using (4.6) we see by means of row operations on (4.8) that the rank of (4.7) is also equal to the rank of

$$\left(\begin{array}{ccc|ccc} A_o & \cdots & A_{r-1} & 0 & \cdots 0 & 0 \\ \vdots & & \vdots & \vdots & \vdots & \vdots \\ A_{r-1} & & A_{2r-2} & 0 & \cdots 0 & 0 \\ \hline 0 & \cdots & 0 & 0 & \cdots 0 & X \end{array}\right)$$

Now the rank of (4.7) is $n = \operatorname{rank} \mathcal{H}_{r-1}(\mathcal{A})$. Hence $X = 0$ which proves the induction step. This proves the first half of (4.5); the second half is proved similarly.

More generally one has the following result (which we shall not need in the sequel).

*4.9. Lemma. Let $A_o, \ldots, A_s$ be a finite series of matrices and suppose there are $i,j \in \mathbb{N} \cup \{0\}$ such that $i + j = s - 1$ and

$$\operatorname{rank}\begin{pmatrix} A_o & \cdots & A_i \\ \vdots & & \vdots \\ A_j & \cdots & A_{i+j} \end{pmatrix} = \operatorname{rank}\left(\begin{array}{ccc|c} A_o & \cdots & A_i & A_{i+1} \\ \vdots & & \vdots & \vdots \\ A_j & \cdots & A_{i+j} & A_{i+j+1} \end{array}\right) = \operatorname{rank}\left(\begin{array}{ccc} A_o & \cdots & A_i \\ \vdots & & \vdots \\ A_j & \cdots & A_{i+j} \\ \hline A_{j+1} & \cdots & A_{i+j+1} \end{array}\right) = n$$

for some $n \in \mathbb{N} \cup \{0\}$, then there are unique $A_{s+1}, A_{s+2}, \ldots$ such that

$$\operatorname{rank} \mathcal{H}_t(\mathcal{A}) = n$$

for all $t \geq \max(i,j)$.

Proof. By hypothesis we know that there exist matrices $S_o, \ldots, S_i$ such that

$$A_{i+r+1} = A_r S_o + \ldots + A_{r+i} S_i \;, \; r = 0, \ldots, j \tag{4.10}$$

Now define A_t for $t > s$ by the formula

$$A_t = A_{t-i-1} S_o + \ldots + A_{t-1} S_i \tag{4.11}$$

Also by hypothesis we know that there exist $T_o, \ldots, T_j$ such that

$$A_{j+r+1} = T_o A_r + \ldots + T_j A_{j+r}, \; r = 0, \ldots, i \tag{4.12}$$

To prove that $\operatorname{rank} \mathcal{H}_t(\mathcal{A}) = n$ for all $t \geq \max(i,j)$ it now clearly suffices to show that (4.12) holds in fact for all $r \geq 0$. Suppose this has been proved for $r \leq q-1$, $q \geq i+1$. Consider the matrix

$$\left(\begin{array}{ccc|ccc} A_o & \cdots & A_i & A_{i+1} & \cdots & A_q \\ \vdots & & \vdots & \vdots & & \vdots \\ A_j & \cdots & A_{i+j} & A_{i+j+1} & \cdots & A_{j+q} \\ \hline A_{j+1} & \cdots & A_{i+j+1} & A_{i+j+2} & \cdots & A_{j+q+1} \end{array}\right) \tag{4.13}$$

By means of column operations, the hypothesis of the lemma, and (4.10) - (4.11) we see that the rank of the matrix (4.13) is n. Using row operations and (4.12) for $r \leq q-1$ (induction hypothesis) we see that the rank of (4.13) is equal to the rank of

$$(4.14)\qquad \left(\begin{array}{ccc|ccc} A_o & \cdots & A_i & A_{i+1} & \cdots & A_q \\ \vdots & & \vdots & \vdots & & \vdots \\ A_j & \cdots & A_{i+j} & A_{i+j+1} & \cdots & A_{j+q} \\ \hline 0 & \cdots & 0 & 0 & \cdots & 0\ X \end{array}\right)$$

where X is the matrix $A_{j+q+1} - T_oA_q - \ldots - T_jA_{j+q}$. Now use column operations and (4.10), (4.11) to see that the rank of (4.14) is equal to the rank of

$$(4.15)\qquad \left(\begin{array}{ccc|ccc} A_o & \cdots & A_i & 0 & \cdots & 0 \quad 0 \\ \vdots & & \vdots & \vdots & & \vdots \quad \vdots \\ A_j & \cdots & A_{i+j} & 0 & \cdots & 0 \quad 0 \\ \hline 0 & \cdots & 0 & 0 & \cdots & 0 \quad X \end{array}\right)$$

It follows that X = 0.

4.16. Proof of theorem 4.2 (first step: existence of a co and cr realization). Let $r \in \mathbb{N}$ be such that $r \geq n$ amd rank $\mathcal{H}_s(\mathcal{A}) = n$ for all $s \geq r-1$. We write

$$\mathcal{H} = \mathcal{H}_{r-1}(\mathcal{A}) = \begin{pmatrix} A_o & \cdots & A_{r-1} \\ \vdots & & \vdots \\ A_{r-1} & \cdots & A_{2r-2} \end{pmatrix}, \quad \mathcal{H}^{(k)} = \begin{pmatrix} A_k & \cdots & A_{r+k-1} \\ \vdots & & \vdots \\ A_{r+k-1} & \cdots & A_{2r+k-1} \end{pmatrix}$$

and for all $s,t \in \mathbb{N}$ we define

$$E_{s\times t} = (I_{s\times s} \,|\, 0_{s\times(t-s)}) \quad \text{if } s < t$$

$$E_{s\times s} = I_{s\times s} \quad \text{if } s = t$$

$$E_{s\times t} = \begin{pmatrix} I_{t\times t} \\ \hline 0_{(s-t)\times t} \end{pmatrix} \quad \text{if } s > t$$

where $I_{a\times a}$ is the $a \times a$ identity matrix and $0_{a\times b}$ is the $a \times b$ zero matrix. Because $\mathcal{H}$ is of rank n, there exist an invertible $pr \times pr$ matrix P and an invertible $mr \times mr$ matrix M such that

$$(4.17)\qquad P\mathcal{H}M = \left(\begin{array}{c|c} I_{n\times n} & 0_{n\times(mr-n)} \\ \hline 0_{(pr-n)\times n} & 0_{(pr-n)\times(mr-n)} \end{array}\right) = E_{pr\times n}E_{n\times mr}$$

Now define

$$(4.18) \qquad F = E_{n\times pr}P\mathcal{H}^{(1)}M\,E_{mr\times n}\ , \quad G = E_{n\times pr}P\mathcal{H}E_{mr\times m},$$

$$H = E_{p\times pr}\mathcal{H}M\,E_{mr\times n}$$

We claim that then (F,G,H) realizes $\mathcal{A}$, i.e. that

$$(4.19) \qquad A_i = HF^iG\ , \ i = 0,1,2, \ldots$$

To prove this we define

$$D = \begin{pmatrix} 0 & \cdots & 0 & S_o \\ I & \ddots & \vdots & \vdots \\ 0 & \ddots & 0 & \vdots \\ 0 & \cdots & 0\ \ I & S_{r-1} \end{pmatrix} \quad C = \begin{pmatrix} 0' & I' & 0' & \cdots & 0' \\ \vdots & & \ddots & \ddots & \vdots \\ \vdots & & & \ddots & 0' \\ 0' & \cdots & & 0' & I' \\ T_o & \cdots & \cdots & \cdots & T_{r-1}' \end{pmatrix}$$

where 0, I, 0', I' are respectively the m × m zero matrix, the m × m identity matrix, the p × p zero matrix and the p × p identity matrix and where the $S_o, \ldots, S_{r-1}$ and $T_o, \ldots, T_{r-1}$ are such that (4.5) holds for all i. Then

$$(4.20) \qquad \mathcal{H}^{(k)} = C^k\mathcal{H} = \mathcal{H}D^k, \ k = 1,2, \ldots$$

Let $\mathcal{H}^* = ME_{mr\times n}E_{n\times pr}P$. Then $\mathcal{H}^*$ is a pseudoinverse of $\mathcal{H}$ in that

$$(4.21) \qquad \mathcal{H}\,\mathcal{H}^*\mathcal{H} = \mathcal{H}$$

(Indeed using (4.17) we have $\mathcal{H}\mathcal{H}^*\mathcal{H} = P^{-1}E_{pr\times n}E_{n\times mr}M^{-1}ME_{mr\times n}\,E_{n\times pr}\,P\,P^{-1}E_{pr\times n}E_{n\times mr}M^{-1} = \mathcal{H}$ because $M^{-1}M = I$, $PP^{-1} = I$, $E_{n\times mr}E_{mr\times n} = I_{n\times n}$, $E_{n\times pr}E_{pr\times n} = I_{n\times n}$.).

We now first prove that

$$(4.22) \qquad E_{n\times pr}P\,C^k\mathcal{H}M\,E_{mr\times n} = F^k\ , \ k = 1,2, \ldots$$

In view of (4.20) this is the definition of F (cf. (4.18)) in the case $k = 1$. So assume (4.22) has been proved for $k \leq t$. We then have

$$E_{n\times pr} P\, C^{t+1} \mathcal{H} M\, E_{mr\times n} = E_{n\times pr} P\, C^{t} \mathcal{H} D\, M\, E_{mr\times n} \quad \text{(by (4.20))}$$

$$= E_{n\times pr} P\, C^{t} \mathcal{H}\, \mathcal{H}^* \mathcal{H}\, D\, M\, E_{mr\times n} \quad \text{(by (4.21))}$$

$$= E_{n\times pr} P\, C^{t} \mathcal{H} M\, E_{mr\times n} E_{n\times pr} P \mathcal{H} D M E_{mr\times n}$$

(by the definition of $\mathcal{H}^*$)

$$= F^t E_{n\times pr} P\, C \mathcal{H} M E_{mr\times n} \quad \text{(by the induction hypothesis and (4.20))}$$

$$= F^t F \quad \text{(by (4.20))}$$

We now have for all $k \geq 0$

$$A_k = E_{p\times pr}\, \mathcal{H}^{(k)} E_{mr\times m} \quad \text{(definition of } \mathcal{H}^{(k)})$$

$$= E_{p\times pr} C^k \mathcal{H} E_{mr\times m} \quad \text{(by (4.20))}$$

$$= E_{p\times pr} C^k \mathcal{H}\, \mathcal{H}^* \mathcal{H} E_{mr\times m} \quad \text{(by (4.21))}$$

$$= E_{p\times pr} C^k \mathcal{H} M\, E_{mr\times n} E_{n\times pr} P \mathcal{H} E_{mr\times m} \quad \text{(by the definition of } \mathcal{H}^*)$$

$$= E_{p\times pr} \mathcal{H} D^k M\, E_{mr\times n} G \quad \text{(by the definition of G and (4.20))}$$

$$= E_{p\times pr}\, \mathcal{H}\, \mathcal{H}^* \mathcal{H} D^k M E_{mr\times n} G \quad \text{(by (4.21))}$$

$$= E_{p\times pr} \mathcal{H} M\, E_{mr\times n} E_{n\times pr} P \mathcal{H} D^k M\, E_{mr\times n} G \quad \text{(by the definition of } \mathcal{H}^*)$$

$$= H\, E_{n\times pr} P\, C^k \mathcal{H} M\, E_{mr\times n} \quad \text{(by the definition of H and (4.20))}$$

$$= H\, F^k G \quad \text{(by (4.22))}$$

This proves the existence of an n-dimensional system $\Sigma = (F,G,H)$ which realizes $\mathcal{A}$. Now for all $s = 0,1,2,\ldots$

$$\mathcal{H}_s(\mathcal{A}) = Q_s(F,H)R_s(F,G)$$

where

$$Q_s(F,H) = \begin{pmatrix} H \\ HF \\ \vdots \\ HF^s \end{pmatrix}, \quad R_s(F,G) = (G \quad FG \quad \dots \quad F^sG)$$

Both $Q_s(F,H)$ and $R_s(F,G)$ have necessarily rank $\leq n$. It follows via the Cayley-Hamilton theorem that (F,G,H) is completely reachable and completely controllable, because rank $\mathcal{H}_s(\mathcal{A}) = n$ for $s \geq r-1$.

4.23. <u>Proof of the uniqueness statement of theorem 4.2.</u>

Let $\Sigma = (F,G,H)$ and $\bar{\Sigma} = (\bar{F},\bar{G},\bar{H})$ be two co and cr realizations of $\mathcal{A}$. Then $\dim(\Sigma) = \operatorname{rank}\mathcal{H}_{n-1}(\mathcal{A}) = \dim(\bar{\Sigma})$. By hypothesis we have

(4.24) $$A_i = HF^iG = \bar{H}\,\bar{F}^i\bar{G} \quad , i = 0,1,2, \dots$$

According to lemma 3.2 and 3.11 there exists a *nice selection* α of $I(n-1,m)$, the set of column indices of $R_{n-1}(F,G)$ and $\mathcal{H}_{n-1}(F,G,H)$, and there exists a *nice selection* β of $I(n-1,p)$, the set of row indices of $Q_{n-1}(F,H)$ and $\mathcal{H}_{n-1}(F,G,H)$, such that

$$\operatorname{rank}(R_{n-1}(F,G)_\alpha) = \operatorname{rank}(Q_{n-1}(F,H)_\beta) = n$$

Let $\mathcal{H}_{n-1}(F,G,H)_{\alpha,\beta}$ be the matrix obtained from $\mathcal{H}_{n-1}(F,G,H)$ by removing all rows whose index is not in β and all columns whose index is not in α. Then

$$\mathcal{H}_{n-1}(F,G,H)_{\alpha,\beta} = Q_{n-1}(F,H)_\beta\, R_{n-1}(F,G)_\alpha$$

so that $\mathcal{H}_{n-1}(F,G,H)_{\alpha,\beta}$ is an invertible $n \times n$ matrix. Also

$$\mathcal{H}_{n-1}(F,G,H)_{\alpha,\beta} = \mathcal{H}_{n-1}(\bar{F},\bar{G},\bar{H})_{\alpha,\beta} = Q_{n-1}(\bar{F},\bar{H})_\beta R_{n-1}(\bar{F},\bar{G})_\alpha$$

so that $Q_{n-1}(\bar{F},\bar{H})_\beta$ and $R_{n-1}(\bar{F},\bar{G})_\alpha$ are also invertible. Now let

$$\Sigma_1 = (F_1,G_1,H_1) = (F,G,H)^T, \quad T = Q_{n-1}(F,H)_\beta$$
$$\bar{\Sigma}_1 = (\bar{F}_1,\bar{G}_1,\bar{H}_1) = (\bar{F},\bar{G},\bar{H})^{\bar{T}}, \quad \bar{T} = Q_{n-1}(\bar{F},\bar{H})_\beta$$

Then of course Σ_1 and $\bar{\Sigma}_1$ also realize . Moreover, using (2.4) we see

$$Q_{n-1}(F_1,H_1)_\beta = I_n = Q_{n-1}(\bar{F}_1,\bar{H}_1)_\beta$$

It follows that

$$R(F_1,G_1) = \mathcal{H}_n(\Sigma_1)_\beta = \mathcal{H}_n(\Sigma)_\beta = \mathcal{H}_n(\bar{\Sigma})_\beta = \mathcal{H}_n(\bar{\Sigma}_1)_\beta = R(\bar{F}_1,\bar{G}_1)$$

and, in turn, this means that $F_1 = \bar{F}_1$ and $G_1 = \bar{G}_1$ by lemma (3.7) (i) combined with lemma (3.4). Further the matrix consisting of the first p rows of $\mathcal{H}_n(\Sigma_1) = \mathcal{H}_n(\bar{\Sigma}_1)$ is equal to

$$H_1R(F_1,G_1) = \bar{H}_1R(\bar{F}_1,\bar{G}_1)$$

so that also $H_1 = \bar{H}_1$ because $R(F_1,G_1) = R(\bar{F}_1,\bar{G}_1)$ is of rank n. This proves that indeed $\bar{\Sigma} = \Sigma^S$ with $S = \bar{T}^{-1}T$.

4.25. <u>A realization algorithm</u>. Now that we know that $\mathcal{A}$ is realizable by a co and cr system of dimension n iff rank $\mathcal{H}_s(\mathcal{A}) = n$ for all large enough s it is possible to give a rather easier algorithm for calculating a realization than the one used in 4.16 above (which is the algorithm of B.L. Ho). It goes as follows. Because $\mathcal{A}$ is realizable by a $\Sigma \in L^{co,cr}_{m,n,p}(\mathbb{R})$ there exist a nice selection $\alpha \subset I(n,m)$, the set of column indices of $R(F,G)$ and $\mathcal{H}_n(\Sigma)$, and a nice selection $\beta \subset I(n,p)$, the set of row indices of $Q(F,H)$ and $\mathcal{H}_n(\Sigma)$, such that

$$\mathcal{H}_n(\mathcal{A})_{\alpha,\beta} = S \tag{4.26}$$

is an invertible $n \times n$ matrix. Consider

$$S^{-1}\mathcal{H}_n(\mathcal{A})_\beta$$

This $n \times (n+1)m$ matrix is necessarily of the form $R(F,G)$ for some $(F,G) \in L^{cr}_{m,n}(\mathbb{R})$ and moreover by (4.26)

$$(S^{-1}\mathcal{H}_n(\mathcal{A})_\beta)_\alpha = I_n$$

so that F,G can simply be written down from $S^{-1}\mathcal{H}_n(\mathcal{A})_\beta$ as in the proof of lemma 3.4. The matrix H is now obtained as the matrix consisting of the first p rows of $\mathcal{H}_n(\mathcal{A})_\alpha$.

After choosing α, this algorithm describes the unique triple (F,G,H) which realizes $\mathcal{A}$ such that moreover $R(F,G)_\alpha = I_n$.

*4.27. Relation with rational functions.

Suppose that $\mathcal{H}_k(\mathcal{A})$ is of rank n for all sufficiently large k. Then by theorem 4.2 the sequence $\mathcal{A}$ is realizable. Using Laplace transforms (cf. 1.8 above) we see that this means that the $p \times m$ matrix of power series $\sum_{i=o}^{\infty} A_i s^{-i-1}$ is in fact a matrix of rational functions.

(4.28) $$\sum_{i=o}^{\infty} A_i s^{-i-1} = (s^n - a_{n-1}s^{n-1} - \dots - a_1 s - a_o)^{-1}B(s) =$$
$$= d(s)^{-1}B(s)$$

where B(s) is a $p \times m$ matrix of polynomials in s of degree $\leq n-1$.

Inversely if

(4.29) $$\sum_{i=o}^{\infty} A_i s^{-i-1} = d'(s)^{-1}B'(s)$$

for a matrix of polynomials B'(s) and a polynomial d'(s) = $= s^r - a'_{r-1}s^{r-1} - \dots - a'_1 s - a'_o$ with r = degree(d'(s)) > degree B'(s), then

$$A_{i+r} = a'_o A_i + a'_1 A_{i+1} + \dots + a'_{r-1}A_{i+r-1}$$

for all i = 0,1,2,... . And this, in turn implies that

$$\text{rank}\,\mathcal{H}_k(\mathcal{A}) = \text{rank}\,\mathcal{H}_{r-1}(\mathcal{A})$$

for all $k \geq r-1$, so that $\mathcal{A}$ is realizable. It follows that $\mathcal{A}$ is realizable iff $\Sigma A_i s^{-i-1}$ represents a rational function which goes to zero as $s \to \infty$.

5. FEEDBACK SPLITS THE EXTERNAL DESCRIPTION DEGENERACY.

In this section we shall prove the result described in section 1.6 To do this we first discuss still another local canonical form.

5.1. <u>The Kronecker nice selection of a system</u>. Let $(F,G,H) \in L^{cr}_{m,n,p}(\mathbb{R})$. We proceed as follows to obtain a "first" nice selection κ such that $(F,G,H) \in U_\kappa$.

Consider the set of column indices $I(m,n)$ in the order $(0,1) < (0,2) <\ldots< (0,m) < (1,1) <\ldots< (1,m) <\ldots< (n,1) <\ldots< (n,m)$. For each (i,j) we set $(i,j) \in \kappa \Leftrightarrow F^i g_j$ is linear independent of the $F^{i'} g_{j'}$ with $(i',j') < (i,j)$. We shall call the subset κ of $I(n,m)$ thus obtained, the Kronecker selection of (F,G,H) and denote it with $\kappa(F,G,H)$. It is obvious that κ has n elements if $(F,G,H) \in L^{cr}_{m,n,p}(\mathbb{R})$.

5.2. <u>Lemma</u>. The Kronecker selection κ defined above is a nice selection.

Proof. Let $(i,j) \in \kappa$ and suppose $i \geq 1$. Suppose that $(i',j) \notin \kappa$, $i' < i$. This means that there is a relation

$$F^{i'} g_j = \sum_{(k,\ell)<(i',j)} b(k,\ell) F^k g_\ell$$

Multiplying with $F^{i-i'}$ on the left one obtains

$$F^i g_j = \sum_{(k,\ell)<(i',j)} b(k,\ell) F^{i-i'+k} g_\ell$$

showing that $F^i g_j$ is linearly dependent on the $F^s g_{j'}$ with $(s,j') < (i,j)$. A contradiction, q. e.d.

5.3. <u>Lemma</u>. Let $(F,G,H) \in L^{cr}_{m,n,p}(\mathbb{R})$ and $S \in GL_n(\mathbb{R})$, then

$$\kappa(F,G,H) = \kappa((F,G,H)^S)$$

5.4. <u>Lemma</u>. Let $(F,G,H) \in L^{cr}_{m,n,p}(\mathbb{R})$ and let L be an $m \times n$ matrix. Then

$$\kappa(F,G,H) = \kappa(F+GL,G,H)$$

The proof of lemma 5.3 is immediate. As to lemma 5.4 we define

$$(5.5)\quad \begin{aligned} X_o(\Sigma) &= \text{subspace of } X = \mathbb{R}^n \text{ generated by } g_1, \ldots, g_m \\ X_1(\Sigma) &= \text{subspace of } X = \mathbb{R}^n \text{ generated by } g_1, \ldots, g_m, F_{g_1}, \ldots, F_{g_m} \\ &\;\vdots \\ X_n(\Sigma) &= \text{subspace of } X \in \mathbb{R}^n \text{ generated by } g_1, \ldots, g_m, \\ &\qquad F_{g_1}, \ldots, F_{g_m}, \ldots, F^n g_1, \ldots, F^n g_m \end{aligned}$$

Let $\Sigma(L) = (F+GL,G,H)$ and let $\hat{F} = F + GL$. Then one easily obtains by induction that

$$(5.6)\qquad X_i(\Sigma(L)) = X_i(\Sigma), \quad i = 0, \ldots, n$$

and that

$$(5.7)\qquad \hat{F}^i g_j \equiv F^i g_j \mod X^{i-1}(\Sigma), \; i = 0,1, \ldots, n$$

(where, by definition, $X^{-1}(\Sigma) = \{0\}$). Lemma 5.4 is an immediate consequence of (5.7). (Note that a basis for $X^i(\Sigma)$ is formed by the vectors $F^k g_\ell$ with $(k,\ell) \in \kappa(\Sigma)$ and $k \leq i$; the classes of the $F^k g_\ell$ with $(k,\ell) \in \kappa(\Sigma)$, $k = i$ are a basis for the quotient space $X^i(\Sigma)/X^{i-1}(\Sigma)$, $i = 0, \ldots, n$).

If $\Sigma = (F,G,H) \in L^{cr,co}_{m,n,p}(\mathbb{R})$ then $\kappa(F,G,H)$ can be calculated from $\mathcal{H}_n(F,G,H)$. Indeed in that case $Q(F,H)$ is of rank n. Therefore, because $\mathcal{H}_n(F,G,H) = Q(F,H)R(F,G)$, the dependency relations between the columns of $\mathcal{H}_n(F,G,H)$ and between the columns of $R(F,G)$ are exactly the same.

5.8. Remark. If $(F,G,H) \in L^{cr}_{m,n,p}(\mathbb{R})$ then also $(F+GL,G,H) \in L^{cr}_{m,n,p}(\mathbb{R})$ as is easily checked. But if $(F,G,H) \in L^{co}_{m,n,p}(\mathbb{R})$, then $(F+GL,G,H)$ need not also be completely observable. Though of course this will be the case for sufficiently small L (because $L^{co}_{m,n,p}(\mathbb{R})$ is an open subset of $L_{m,n,p}(\mathbb{R})$).

*5.9. The Kronecker control invariants. The invariant $\kappa(F,G,H)$ depends only on F and G, so that we can also write $\kappa(F,G)$. For each $j = 1, \ldots, m$, let k_j be the number of elements (i,ℓ) in $\kappa(F,G)$ such that $\ell = j$. Let $\kappa_1(F,G) \geq \ldots \geq \kappa_{m'}(F,G)$, $m' = \text{rank}(G)$, be the sequence of those k_j which are $\neq 0$ ordered with respect to size. It follows from lemma's 5.3 and 5.4 that the $\kappa_i(F,G)$ are invariant for the transformations

(5.10) $(F,G) \mapsto (F,G)^S = (SFS^{-1},SG)$ (base change in state space)

(5.11) $(F,G) \mapsto (F+GL,G)$ (feedback)

One easily checks that the $\kappa_i(F,G)$ are also invariant under

(5.12) $(F,G) \mapsto (F,GT)$, $T \in GL_m(\mathbb{R})$ (base change in input space)

This can e.g. be seen as follows. Let $\lambda_i(\Sigma) = \dim X^i(\Sigma) - \dim X^{i-1}(\Sigma)$ for $i = 0,1,\ldots,n$. Consider an rectangular array of $(n+1) \times m$ boxes with the rows labelled $0, \ldots, n$. Now put a cross in the first $\lambda_i(\Sigma)$ boxes of row i for $i = 0, \ldots, n$. Then $\kappa_j(\Sigma)$, $j = 1, \ldots, m'$ is the number of crosses in column j of the array. Obviously the $\lambda_i(\Sigma)$ do not change under a transformation of type (5.12), proving that also the $\kappa_j(F,G)$ are invariant under 5.12.

The group generated by all these transformations is called the <u>feedback group</u>. Thus the $\kappa_i(F,G)$ are invariants of the feedback group acting on $L^{cr}_{m,n}(\mathbb{R})$. It now turns out that these are in fact the only invariants. I.e. if (F,G), $(\bar{F},\bar{G}) \in L^{cr}_{m,n}(\mathbb{R})$ and $\kappa_i(F,G) = \kappa_i(\bar{F},\bar{G})$, $i = 1, \ldots, m'$, then $(\bar{F},\bar{G})$ can be obtained from (F,G) by means of a series of transformations from (5.10) - (5.12). Cf. [11] for a proof , or cf. 5.30 below.

The $\kappa_i(F,G)$ are also identifiable with Kronecker's minimal column indices of the singular matrix pencil $(zI_n - F \mid G)$, cf [11].

Still another way to view the $\kappa_i(F,G)$ is as follows.

Consider the transfer matrix $T(s) = H(sI_n-F)^{-1}G$ of the linear dynamical system $\Sigma = (F,G,H)$ considered as a $p \times m$ matrix valued function of the complex variable s. One can now prove (cf. [14]).

<u>Theorem</u>. There ~~exist~~ matrices $N(s)$ and $D(s)$ of polynomial functions of s such that (i) $T(s) = N(s)D(s)^{-1}$, (ii) there exist matrices of polynomials such that $X(s)N(s) + Y(s)D(s) = I_m$, (iii) $N(s)$ and $D(s)$ are unique up to multiplication on the right by a unit from the ring of polynomial $m \times m$ matrices. Moreover degree$(\det D(s)) = n = \dim(\Sigma)$.

Now for each $s \in \mathbb{C}$, one defines

$$\phi_\Sigma(s) = \{(N(s)u,\ D(s)u) \mid u \in \mathbb{C}^m\} \subset \mathbb{C}^{p+m}$$

If $s \in \mathbb{C}$ is such that $D(s)^{-1}$ exists, then also $\phi_\Sigma(s) = \{(T(s)u,u) | u \in \mathbb{C}^m\} \subset \mathbb{C}^{p+m}$. In any case $\phi_\Sigma(s)$ is a p-dimensional subspace of $\mathbb{C}^{p+m}$. In addition one defines $\phi_\Sigma(\infty) = \{(0,u) | u \in \mathbb{C}^m\} \subset \mathbb{C}^{p+m}$, which is entirely natural because $\lim\limits_{s\to\infty} T(s) = 0$. This gives a continuous map of the Riemann sphere $\mathbb{C} \cup \{\infty\} = S^2$ to the Grassmann manifold $G_{m,p+m}(\mathbb{C})$ of m-planes in p + m space. Let $\xi_m \to G_{m,p+m}(\mathbb{C})$ be the canonical complex vector bundle whose fibre over $z \in G_{m,p+m}(\mathbb{C})$ is the m-plane represented by z. Pulling back ξ_m along ϕ_Σ gives us a holomorphic complex vector bundle $\xi(\Sigma)$ over S^2.

Now holomorphic vectorbundles over the sphere S^2 have been classified by Grothendieck. The classification result is: every holomorphic vector-bundle over S^2 is isomorphic to a direct sum of line bundles and line bundles are classified by their degrees.

It now turns out that the numbers classifying $\xi(\Sigma)$, the bundle over S^2 defined by the system Σ, are precisely the $\kappa_i(\Sigma)$, $i = 1, \ldots, m$, where $\kappa_i(\Sigma) = 0$ for $i > m' = \text{rank}(G)$. One also recovers $n = \dim(\Sigma)$ as the intersection number of $\phi_\Sigma(S^2)$ with a hyperplane in $G_{m,m+p}(\mathbb{C})$.

These observations are due to Clyde Martin and Bob Hermann, cf. [13].

As we have seen the $\kappa_i(\Sigma)$ are invariants for the transformations (5.10), (5.11), (5.12). Being defined in terms of F and G alone they are also obviously invariant under base change in output space: $(F,G,H) \mapsto (F,G,SH)$, $S \in GL_p(\mathbb{R})$. The $\kappa_i(\Sigma)$ are, however, definitely not a full set of invariants for the group G acting on $L_{m,n,p}(\mathbb{R})$, where G is the group generated by base changes in state space, input space and output space and the feedback transformations.

5.13. <u>The canonical input base change matrix</u> $T(\Sigma)$.

Let $\Sigma = (F,G,H) \in L^{cr}_{m,n,p}(\mathbb{R})$ and let $\kappa = \kappa(\Sigma)$ be the Kronecker nice selection of Σ. Let $(i,j) = s(\kappa,j)$ be a successor index of κ. By the definition of κ we have a unique expression of the form

$$F^i g_j = \sum_{\substack{(i,j')\in\kappa \\ j'<j}} a_j(j') F^i g_{j'} + \sum_{\substack{(k,\ell)\in\kappa \\ k<i}} a(k,\ell) F^k g_\ell \tag{5.14}$$

Now define recursively

$$(5.15)\qquad \hat{g}_j = g_j - \sum_{j'<j} a_j(j')g_{j'},\ \hat{G} = (\hat{g}_1, \ldots, \hat{g}_m)$$

and

$$(5.16)\quad T(\Sigma) = (b_{jk}),\ b_{jk} = 1 \text{ if } j = k,\ b_{jk} = -a_k(j), \text{ if } j < k,$$
$$b_{jk} = 0 \text{ if } j > k,$$

then $\hat{G} = GT(\Sigma)$, and $T(\Sigma)$ is an upper triangular matrix of determinant 1.

5.17. <u>Lemma</u>. Let $\Sigma \in (F,G,H) \in L^{cr}_{m,n,p}(\mathbb{R})$, then

$$T(\Sigma) = T(\Sigma^S),\ T(\Sigma(L)) = T(\Sigma)$$

for all $S \in GL_n(\mathbb{R})$ and all feedback matrices $L \in \mathbb{R}^{m\times n}$.

Proof. Obvious. (Use (5.7)).

5.18. <u>Example</u>. Let $m = 5$, $n = 9$, and let $(F,G,H) \in L^{cr}_{5,9,p}(\mathbb{R})$ have Kronecker selection $\kappa(F,G,H)$ equal to

$$\kappa = \begin{matrix} \times & \times & . & \times & \times \\ \times & \times & . & \times & . \\ . & \times & . & . & . \\ . & \times & . & . & . \\ . & . & . & . & . \end{matrix}$$

where we have omitted the last five rows of dots.

Then $T(\Sigma)$ is an upper triangular matrix of the form

$$T(\Sigma) = \begin{pmatrix} 1 & 0 & * & * & * \\ 0 & 1 & * & 0 & * \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & * \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Note that $T(\Sigma)^{-1}$ is of precisely the same form.

This is a general phenomon. Indeed by (5.14) and (5.15) (cf. also example (5.18)) $\hat{g}_j$ is of the form

$$(5.19) \qquad \hat{g}_j = g_j + \sum_{\substack{k_i > k_j \\ i<j}} b_{ij} g_i \quad , \; T(\Sigma) = (b_{ij})$$

So that $b_{ij} = 0$ unless $i = j$ (and then $b_{ij} = 1$) or $i < j$ and $k_i > k_j$
Let $t_1, \ldots, t_m$ be the columns of $T(\Sigma)$ and $e_1, \ldots, e_m$ the standard basis for $\mathbb{R}^m$. Then

$$(5.20) \qquad t_j = e_j + \sum_{\substack{k_i > k_j \\ i<j}} b_{ij} e_i$$

Using induction with respect an ordening of the $\{1, \ldots, m\}$ satisfying $i < j \Rightarrow k_i \geq k_j$ it readily follows that

$$e_j = t_j + \sum_{\substack{i<j \\ k_i > k_j}} b'_{ij} t_i$$

which proves that $T(\Sigma)^{-1}$ also has zero entries at all spots (i,j) with $i > j$ or $i < j$ and $k_i \leq k_j$.

5.21. <u>The block companion canonical form.</u> Let κ be a nice selection. We are going to construct a canonical form on the subspace W_κ of all $\Sigma \in L^{cr,co}_{m,n,p}(\mathbb{R})$ with $\kappa(\Sigma) = \kappa$. We shall do this only in full detail for the case that κ is the nice selection of example 5.18. This special case is, however, general enough to see that this construction works in general. Let $(F,G,H) \in W_\kappa$ and let $\hat{G} = GT(\Sigma)$. Now consider the system $(F,\hat{G},H)$ which is also in W_κ as is easily checked. This system has the property that for each successor index $s(\kappa,j) = (i,j)$ of κ with $i \neq 0$ we have

$$(5.22) \qquad F^i \hat{g}_j = \sum_{\substack{(k,\ell) \in \kappa \\ k<i}} a'(k,\ell) F^k \hat{g}_\ell$$

(i.e. $T(F,\hat{G},H) = I_m$). Indeed , using (5.14)

$$F^i \hat{g}_j = F^i g_j - \sum_{j'<j} a_j(j') F^i g_{j'} = \sum_{\substack{(k,\ell) \in \kappa \\ k<i}} a(k,\ell) F^k g_\ell = \sum_{\substack{(k,\ell) \in \kappa \\ k<i}} a'(k,\ell) F^k \hat{g}_\ell$$

(Cf. (5.5)). Now define a new basis for $\mathbb{R}^n$ as follows. Let $\kappa = \{(0,j_1), \ldots, (i_1,j_1); \ldots; (0,j_r), \ldots, (i_r,j_r)$. Then $k_t = i_t + 1$, $t = 1, \ldots, r$, and $k_1 + \ldots + k_r = n$. For the successor indices $s(\kappa,j\ ) = (k_t,j_t)$, $t = 1, \ldots, r$, write

$$F^{k_t}\hat{g}_{j_t} = - \sum_{\substack{(k,\ell)\in\kappa \\ k<k_t}} b_t(k,\ell)F^k\hat{g}_\ell \tag{5.23}$$

Setting $b_t(k,\ell) = 0$ for all $(k,\ell) \notin \kappa$ we now define a new basis for $\mathbb{R}^n$ by

$$e_1 = F^{k_1-1}\hat{g}_{j_1} + \sum_{i=1}^{t} b_1(k_1-1,j_t)F^{k_1-2}\hat{g}_{j_t} + \ldots + \sum_{i=1}^{t} b_1(1,j_t)\hat{g}_{j_t}$$

$$e_2 = F^{k_1-2}\hat{g}_{j_1} + \sum_{i=1}^{t} b_1(k_1-1,j_t)F^{k_1-3}\hat{g}_{j_t} + \ldots + \sum_{i=1}^{t} b_1(1,j_t)\hat{g}_{j_t}$$

$$\vdots$$

$$e_{k_1} = \hat{g}_{j_1}$$

(5.24)

$$e_{k_1+1} = F^{k_2-1}\hat{g}_{j_2} + \sum_{i=1}^{t} b_2(k_2-1,j_t)F^{k_2-2}\hat{g}_{j_t} + \ldots + \sum_{i=1}^{t} b_2(1,j_t)\hat{g}_{j_t}$$

$$\vdots$$

$$e_{k_1+k_2} = \hat{g}_{j_2}$$

$$\vdots$$

$$e_{k_1+\ldots+k_r} = \hat{g}_{j_r}$$

Let $X_o \subset \mathbb{R}^n$ be the space spanned by the vectors $\hat{g}_{j_1}, \ldots, \hat{g}_{j_t}$ i.e. $X_o = X_o(F,\hat{G},H) = X_o(\Sigma)$. Then we see from (5.23) that for the vectors defined by (5.24) above we have

$$F_{e_1} \in X(\hat{G}),\ F(e_i) \equiv e_{i-1} \bmod X_o \text{ for } i = k_1, k_1-1, \ldots, 2$$

$$F_{e_{k_1+1}} \in X(\hat{G}),\ F(e_i) \equiv e_{i-1} \bmod X_o \text{ for } i = k_1+k_2, \ldots, k_1+2$$

$$\vdots$$

$$F_{e_{k_1+\ldots+k_{r-1}+1}} \in X(\hat{G}),\ F(e_i) \equiv e_{i-1} \bmod X_o \text{ for } i = k_1+\ldots+k_r, \ldots, k_1+\ldots+ \\ + k_{r-1} + 2$$

It follows that with respect to the basis $e_1, \ldots, e_n$, F and $\hat{G}$ are of the form

$$(5.25)\quad F = \begin{array}{c}\left(\begin{array}{c|c|c|c}
\begin{matrix}0&1&0&\cdots&0\\ \vdots&&\ddots&\ddots&\vdots\\ \vdots&&&\ddots&0\\ 0&\cdots&\cdots&0&1\\ *&*&\cdots&*&*\end{matrix} &
\begin{matrix}0&\cdots&0\\ \vdots&&\vdots\\ \vdots&&\vdots\\ 0&\cdots&0\\ *&\cdots&*\end{matrix} & \cdots &
\begin{matrix}0&\cdots&0\\ \vdots&&\vdots\\ \vdots&&\vdots\\ 0&\cdots&0\\ *&\cdots&*\end{matrix} \\ \hline
\begin{matrix}0&\cdots&0\\ \vdots&&\vdots\\ \vdots&&\vdots\\ 0&\cdots&0\\ *&\cdots&*\end{matrix} &
\begin{matrix}0&1&0&\cdots&0\\ \vdots&&\ddots&\ddots&\vdots\\ \vdots&&&\ddots&0\\ 0&\cdots&\cdots&0&1\\ *&\cdots&\cdots&*&*\end{matrix} & \cdots &
\begin{matrix}0&\cdots&0\\ \vdots&&\vdots\\ \vdots&&\vdots\\ 0&\cdots&0\\ *&\cdots&*\end{matrix} \\ \hline
\vdots & \vdots & & \vdots \\ \hline
\begin{matrix}0&\cdots&0\\ \vdots&&\vdots\\ \vdots&&\vdots\\ 0&\cdots&0\\ *&\cdots&*\end{matrix} &
\begin{matrix}0&\cdots&0\\ \vdots&&\vdots\\ \vdots&&\vdots\\ 0&\cdots&0\\ *&\cdots&*\end{matrix} & \cdots &
\begin{matrix}0&1&0&\cdots&0\\ \vdots&&\ddots&\ddots&\vdots\\ \vdots&&&\ddots&0\\ 0&\cdots&\cdots&0&1\\ *&\cdots&\cdots&*&*\end{matrix}
\end{array}\right)
\begin{array}{l}\left.\vphantom{\begin{matrix}0\\0\\0\\0\\0\end{matrix}}\right\} k_1\\ \left.\vphantom{\begin{matrix}0\\0\\0\\0\\0\end{matrix}}\right\} k_2\\ \vdots \\ \left.\vphantom{\begin{matrix}0\\0\\0\\0\\0\end{matrix}}\right\} k_r\end{array}
\\
\quad\underbrace{\hphantom{0\ 1\ 0\ \cdots\ 0}}_{k_1}\quad\underbrace{\hphantom{0\ 1\ 0\ \cdots\ 0}}_{k_2}\quad\cdots\quad\underbrace{\hphantom{0\ 1\ 0\ \cdots\ 0}}_{k_r}\end{array}$$

$$(5.26)\quad \begin{aligned}
&\hat{G} = (\hat{g}_1, \hat{g}_2, \ldots, \hat{g}_m), \text{ with}\\
&\hat{g}_{j_1} = e_{k_1},\ \hat{g}_{j_2} = e_{k_1+k_2},\ \ldots,\ \hat{g}_{j_r} = e_{k_1+\ldots+k_r} = e_n,\\
&\hat{g}_j = 0 \text{ for } j \in \{1,\ldots,m\} \setminus \{j_1,\ldots,j_r\}
\end{aligned}$$

In particular in the case that κ is the nice selection of example 5.18 we see that with respect to the basis $e_1, \ldots, e_n$ defined by 5.24 the matrices F and G take the form (cf. 5.18, the inverse of $T(\Sigma)$ is of the same form as $T(\Sigma)$).

$$
F' = \left(\begin{array}{cc|cccc|cc|c}
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
a_1 & a_2 & a_3 & a_4 & a_5 & a_6 & a_1 & a_8 & a_9 \\
\hline
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
b_1 & b_2 & b_3 & b_4 & b_5 & b_6 & b_1 & b_8 & b_9 \\
\hline
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
c_1 & c_2 & c_3 & c_4 & c_5 & c_6 & c_7 & c_8 & c_9 \\
\hline
d_1 & d_2 & d_3 & d_4 & d_5 & d_6 & d_7 & d_8 & d_9
\end{array}\right)
$$

(5.27)

$$
G' = \begin{pmatrix}
0 & 0 & 0 & 0 & 0 \\
1 & 0 & * & 0 & * \\
0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 \\
0 & 1 & * & * & * \\
0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & * \\
0 & 0 & 0 & 0 & 1
\end{pmatrix}
$$

This does not yet define a canonical form on W_K. True, for every $\Sigma \in W_K$ there exists an $S \in GL_n(\mathbb{R})$ such that $(F,G)^S$ takes the form (5.27). But for two pairs $(F,G) \neq (\bar{F},\bar{G})$, both of the form (5.27), there may very well exists an $S \neq I_n$ such that $(F,G)^S = (\bar{F},\bar{G})$.

In fact, it is now not difficult to check that if S is an $n \times n$ matrix of the form

$$
S = \left(\begin{array}{cc|cccc|cc|c}
1 & 0 & s_{13} & s_{14} & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & s_{13} & s_{14} & 0 & 0 & 0 & 0 \\
\hline
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
\hline
0 & 0 & s_{73} & s_{74} & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & s_{73} & s_{74} & 0 & 0 & 1 & 0 \\
\hline
s_{91} & 0 & s_{93} & s_{94} & s_{94} & 0 & s_{97} & 0 & 1
\end{array}\right)
$$

then SG = G and SFS^{-1} is of the same general form as F, if F and G are of the form (5.27). Choosing s_{13}, s_{14}, s_{73}, s_{74}, s_{91}, s_{93}, s_{94}, s_{95} and s_{97} judiciously we see that for every $\Sigma = (F,G,H) \in W_K$, there exists a $S \in GL_n(\mathbb{R})$ such that SFS^{-1} and SG take the forms

$$SFS^{-1} = \left(\begin{array}{cc|cccc|cc|c}
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
a_1 & a_2 & a_3 & a_4 & 0 & 0 & a_7 & a_8 & a_9 \\ \hline
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
b_1 & b_2 & b_3 & b_4 & b_5 & b_6 & b_7 & b_8 & b_9 \\ \hline
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
c_1 & c_2 & c_3 & c_4 & 0 & 0 & c_7 & c_8 & c_9 \\ \hline
d_1 & 0 & d_3 & 0 & 0 & 0 & d_7 & 0 & d_9
\end{array}\right)$$

(5.28)

$$SG = \begin{pmatrix}
0 & 0 & 0 & 0 & 0 \\
1 & 0 & c_{13} & 0 & c_{15} \\
0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 \\
0 & 1 & c_{23} & c_{24} & c_{25} \\
0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & c_{45} \\
0 & 0 & 0 & 0 & 1
\end{pmatrix}$$

where

$$T(\Sigma)^{-1} = \begin{pmatrix}
1 & 0 & c_{13} & 0 & c_{15} \\
0 & 1 & c_{23} & c_{24} & c_{25} \\
0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 1 & c_{45} \\
0 & 0 & 0 & 0 & 1
\end{pmatrix}$$

The general pattern should be clear: the off-diagonal blocks have zero's in the last row iff there are more columns than rows, in fact in that case the last row ends with (number of columns) - (number of rows) zero's; the structure of the diagonal blocks is clear.

Now suppose that (F',G',H') and (F'',G'',H'') are two systems such that $(F',G')^S = (F'',G'')$ for some S and such that (F',G') and (F'',G'') are both of the forms (5.28). One checks easily that then necessarily $S = I_n$. We have shown

5.29. Proposition. Let κ be the nice selection of example 5.18. Then for every $\Sigma = (F,G,H) \in W_\kappa$ there is precisely one $S \in GL_n(\mathbb{R})$ such that SFS^{-1} and SG have the forms (5.28).

This means in particular (in view of the results of section 4 above) that if $\Sigma \in W_\kappa \cap L^{co,cr}_{n,m,p}(\mathbb{R})$, then the numbers $a_1, \ldots, a_4, a_7, \ldots, a_9, b_1, \ldots, b_9, c_1, \ldots, c_4, c_7, \ldots, c_9, d_1, d_3, d_7, d_9$ can be calculated from $f(\Sigma)$ (or $A_o, \ldots, A_{2n-1}$). Of course these results hold quite generally for all nice selections κ. We note that in general W_κ is not an open subspace of $L^{cr}_{n,m,p}(\mathbb{R})$. In fact $W_\kappa/GL_n(\mathbb{R})$ is a linear subspace of $U_k/GL_n(\mathbb{R}) = \mathbb{R}^{mn+np} \simeq V_\kappa$. In case κ is the nice selection of example 5.18 the codimension of $W_\kappa/GL_n(\mathbb{R})$ in $U_\kappa/GL_n(\mathbb{R})$ is 12. (This number can immediately be read off from κ: g_3 linear dependent on g_1, g_2 causes $9 - 2 = 7$ linear restrictions; Fg_5 linearly dependent on g_1, g_2, g_4, g_5, Fg_1, Fg_2, Fg_4 causes $9 - 7 = 2$ extra linear restrictions; F^2g_1 linearly dependent on $g_1, \ldots, g_5$, Fg_1, Fg_2, Fg_4 causes $9 - 7 = 2$ more linear restrictions; and finally F^2g_4 dependent on $g_1, \ldots, g_5$, Fg_1, Fg_2, Fg_4, F^2g_2 causes $9 - 8 = 1$ more linear restriction; $7+2+2+1=12$).

*5.30. Using the results above, it is now easy to prove that the $\kappa_1(F,G), \ldots, \kappa_{m'}(F,G)$ are the only invariants of the feedback group acting on $L^{cr}_{m,n}(\mathbb{R})$. Indeed, we have already shown that the $\kappa_i(F,G)$, $i = 1, \ldots, m'$ are invariants.

Inversely, using first of all a transformation of type (5.12) we can see to it that (F,GT) has $k_1 \geq k_2 \geq \ldots \geq k_m$, and then $\kappa_1(F,G) = k_1, \ldots, \kappa_{m'}(F,G) = k_{m'}$, $k_i = 0$ for $i > m'$. Then, using transformations of type (5.10) and (5.12), we can change (F,GT) into a pair (F',G') with F' and G' of the type (5.25), (5.26). A final transformation of type (5.11) then changes F' into a matrix of type (5.25) with all stars equal to zero. The final pair (F'',G'') thus obtained depends only on the numbers $\kappa_1(F,G), \ldots, \kappa_{m'}(F,G)$.

q.e.d.

5.31. Feedback breaks all symmetry. We are now in a position to prove the result mentioned in 1.6 that feedback splits the degenerate external description of systems. We shall certainly have proved this if we have proved.

5.32. Theorem. Let $\Sigma \in L^{co,cr}_{m,n,p}(\mathbb{R})$. Then Σ is completely determined by the input-output maps $f(\Sigma(L))$ for small L. More precisely let $\Sigma = (F,G,H)$ and $A_i(L) = H(F+GL)^iG$ for $i = 0,1, \ldots, 2n-1$. Then the entries of $A_i(L)$ are differentiable functions of L, and F, G and H can be calculated from $A_o, \ldots, A_{2n-1}$ and the numbers

$$\left.\frac{\partial A_i(L)}{\partial \ell_{jk}}\right|_{L=o}, \quad i = 0, \ldots, 2n-1,\ j = 1, \ldots, m,\ k = 1, \ldots, n.$$

Proof. Let $\kappa = \kappa(\Sigma)$. Recall that κ can be calculated from $A_o, \ldots, A_{2n-1}$ (because Σ is co and cr). Now assume that κ is the nice selection of example 5.18. (This is sufficiently general, I hope, to make it clear that the theorem holds in general). Let $\Sigma' = (F',G',H')$ be the block companion canonical form of (F,G,H) (Σ' is obtained as follows: first calculate any realization $\Sigma'' = (F'',G'',H'')$ of $A_o, \ldots, A_{2n-1}$, e.g. by means of the algorithm of 4.25 above and then put Σ'' in block companion canonical form as in 5.21 above).
Then

$$\Sigma' = \Sigma^{S^{-1}}$$

for a certain $S \in GL_n(\mathbb{R})$, and it remains to calculate S. With this aim in mind we examine $\Sigma(L) = (F+GL,G,H)$ and its block companion canonical form. Consider

$$\Sigma(L)^{S^{-1}} = (S^{-1}FS+S^{-1}GLS,\ S^{-1}G,HS)$$

$$= (F'+G'LS,G',H')$$

Now assume that L is of the form

$$(5.33) \qquad L = \begin{pmatrix} 0 & \cdot & \cdot & \cdot & 0 \\ \ell_{21} & \cdot & \cdot & \cdot & \ell_{29} \\ 0 & \cdot & \cdot & \cdot & 0 \\ 0 & \cdot & \cdot & \cdot & 0 \\ 0 & \cdot & \cdot & \cdot & 0 \end{pmatrix}$$

Then if F' is of the form (5.28) we see that if $S = (s_{ij})$

$$F' + G'LS = \left(\begin{array}{cc|cccc|cc|c} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ a_1 & a_2 & a_3 & a_4 & 0 & 0 & a_7 & a_8 & a_9 \\ \hline 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ b_1' & b_2' & b_3' & b_4' & b_5' & b'_6 & b_7' & b_8' & b_9' \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ c_1 & c_2 & c_3 & c_4 & 0 & 0 & c_7 & c_8 & c_9 \\ \hline d_1 & 0 & d_3 & 0 & 0 & 0 & d_7 & 0 & d_9 \end{array}\right)$$

with $b_i' = b_i(L) = b_i + \sum_{j=1}^{9} \ell_{2j} s_{ji}$, $i = 1, \ldots, 9$. Thus the block companion canonical from of $\Sigma(L)$ is always $\Sigma(L)^{S^{-1}}$ if L is of the form (5.33). Note that the number of the row which has nonzero entries is determined by $\kappa(\Sigma)$; it is the smallest i for which k_i is maximal; note also that if j is such that k_j is maximal then the j-th vector of G' is always the $(k_1+\ldots+k_j)$-th standard basis vector (cf.just below 5.19).

So to find S we proceed as follows. Calculate the block companion canonical forms of $\Sigma(L)$ from $A_o(L), \ldots, A_{2n-1}(L)$ for small L. (This can be done because for small enough L, $\Sigma(L)$ is still co). This gives us in particular the functions $b_i(L)$. Then

$$s_{ji} = \left.\frac{\partial b_i(L)}{\partial \ell_{2j}}\right|_{L=0}.$$

This determines S and gives us Σ as $\Sigma = (\Sigma')^S$. q.e.d.

6. DESCRIPTION OF $L^{co,cr}_{m,n,p}(\mathbb{R})/GL_n(\mathbb{R})$. INVARIANTS.

6.1. Local structure of $L^{co,cr}_{m,n,p}(\mathbb{R})$. Let $\alpha \subset I(n,m)$ be a nice selection. We recall that $U_\alpha = \{(F,G,H) \in L_{m,n,p}(\mathbb{R}) \mid \det R(F,G)_\alpha \neq 0\}$, that $V_\alpha = \{(F,G,H) \in L_{m,n,p}(\mathbb{R}) \mid R(F,G)_\alpha = I_n\}$ and that $U_\alpha/GL_n(\mathbb{R}) \simeq V_\alpha \simeq \mathbb{R}^{nm+np}$, cf. section 3.

For each $x \in \mathbb{R}^{nm+np}$ let $(F_\alpha(x), G_\alpha(x), H_\alpha(x)) \in V_\alpha$ be the unique system corresponding to x according to the isomorphism of 3.7 above.

6.2. The quotient manifold $M^{cr}_{m,n,p}(\mathbb{R}) = L^{cr}_{m,n,p}(\mathbb{R})/GL_n(\mathbb{R})$. Now that we know what $U_\alpha/GL_n(\mathbb{R})$ looks like it is not difficult to describe $L^{cr}_{m,n,p}(\mathbb{R})/GL_n(\mathbb{R})$. (Recall that the union of the U_α for α nice covers $L^{cr}_{m,n,p}(\mathbb{R})$). We only need to figure out how the $V_\alpha \simeq \mathbb{R}^{mn+np}$ should be glued together. This is not particularly difficult because if $(F,G,H)^S = (F',G',H')$ for some S and $(F,G,H) \in U_\alpha$ then $S = R(F',G')_\alpha R(F,G)_\alpha^{-1}$. It follows that the quotient space $M^{cr}_{m,n,p}(\mathbb{R}) = L^{cr}_{m,n,p}(\mathbb{R})/GL_n(\mathbb{R})$ can be constructed as follows.

For each nice selection α let $\bar{V}_\alpha = \mathbb{R}^{mn+np}$ and for each second nice selection β let

$$\bar{V}_{\alpha\beta} = \{x \in \bar{V}_\alpha \mid \det R(F_\alpha(x), G_\alpha(x))_\beta \neq 0\}$$

We define

$$\phi_{\alpha\beta} : \bar{V}_{\alpha\beta} \to \bar{V}_{\beta\alpha}$$

by the formula

$$(6.3) \qquad \phi_{\alpha\beta}(x) = y \iff R(F_\alpha(x), G_\alpha(x))_\beta^{-1} R(F_\alpha(x), G_\alpha(x)) = R(F_\beta(y), G_\beta(y)).$$

Let $M^{cr}_{m,n,p}(\mathbb{R})$ be the topological space obtained by glueing together the $\bar{V}_\alpha$ by means of the isomorphisms $\phi_{\alpha\beta}$.

Then $M^{cr}_{m,n,p}(\mathbb{R}) = L^{cr}_{m,n,p}(\mathbb{R})/GL_n(\mathbb{R})$. If we denote also with $\bar{V}_\alpha$ the isomorphic image of $\bar{V}_\alpha$ in $M^{cr}_{m,n,p}(\mathbb{R})$ then the quotient map

$\pi: L^{cr}_{m,n,p}(\mathbb{R}) \to M^{cr}_{m,n,p}(\mathbb{R})$ can be described as follows. For each $\Sigma = (F,G,H) \in L^{cr}_{m,n,p}(\mathbb{R})$, choose a nice selection α such that $\Sigma \in U_\alpha$. Then $\pi(\Sigma) = x \in \bar{V}_\alpha \subset M^{cr}_{m,n,p}(\mathbb{R})$ where x is such that $\Sigma^S = (F_\alpha(x), G_\alpha(x), H_\alpha(x))$ with $S = R(F,G)^{-1}_\alpha$.

6.4. Theorem. $M^{cr}_{m,n,p}(\mathbb{R})$ is a differentiable manifold and $\pi : L^{cr}_{m,n,p}(\mathbb{R}) \to M^{cr}_{m,n,p}(\mathbb{R})$ is a principal $GL_n(\mathbb{R})$ fibre bundle.

For a proof, cf. [5].

6.5. The quotient manifold $M^{co,cr}_{m,n,p}(\mathbb{R}) = L^{co,cr}_{m,n,p}(\mathbb{R})/GL_n(\mathbb{R})$. Let $M^{co,cr}_{m,n,p}(\mathbb{R}) = \pi(L^{co,cr}_{m,n,p}(\mathbb{R}))$. Then $M^{co,cr}_{m,n,p}(\mathbb{R})$ is an open submanifold of $M^{cr}_{m,n,p}(\mathbb{R})$. It can be described as follows. For each nice selection α let $\bar{V}^{co}_\alpha = \{x \in \bar{V}_\alpha \mid (F_\alpha(x), G_\alpha(x), H_\alpha(x))$ is completely observable$\}$, and for each nice selection β let $\bar{V}^{co}_{\alpha\beta} = \bar{V}^{co}_\alpha \cap \bar{V}_{\alpha\beta}$. Then $\phi_{\alpha\beta}(\bar{V}^{co}_{\alpha\beta}) = \bar{V}^{co}_{\beta\alpha}$ and $M^{co,cr}_{m,n,p}(\mathbb{R})$ is the differentiable manifold obtained by glueing together the $\bar{V}^{co}_\alpha$ by means of the isomorphisms $\phi_{\alpha\beta}: \bar{V}^{co}_{\alpha\beta} \to \bar{V}^{co}_{\beta\alpha}$.

6.6. $M^{co,cr}_{m,n,p}(\mathbb{R})$ as a submanifold of $\mathbb{R}^{2nmp}$. Let $(F,G,H) \in L^{co,cr}_{n,m,p}(\mathbb{R})$. We associate to (F,G,H) to sequence of $2n$ $p \times m$ matrices $(A_o, \dots, A_{2n-1}) \in \mathbb{R}^{2nmp}$, where $A_i = HF^iG$, $i = 0, \dots, 2n-1$. The results of section 4 above (realization theory) prove that this map is injective and prove that its image consists of those elements $(A_o, \dots, A_{2n-1}) \in \mathbb{R}^{2nmp}$ such that rank $\mathcal{H}_{n-1}(\mathcal{A}) = $ rank $\mathcal{H}_n(\mathcal{A}) = n$. We thus obtain $M^{co,cr}_{m,n,p}(\mathbb{R})$ as a (nonsingular algebraic) smooth submanifold of $\mathbb{R}^{2nmp}$.

6.7. Invariants. By definition a smooth invariant for $GL_n(\mathbb{R})$ acting on $L_{m,n,p}(\mathbb{R})$ is a smooth function $f: U \to \mathbb{R}$, defined on an open dense subset $U \subset L_{m,n,p}(\mathbb{R})$ such that $f(\Sigma) = f(\Sigma^S)$ for all $\Sigma \in U$ and $S \in GL_n(\mathbb{R})$ such that $\Sigma^S \in U$.

Now $L^{co,cr}_{m,n,p}(\mathbb{R})$ is open and dense in $L_{m,n,p}(\mathbb{R})$. It now follows from 6.6 that every invariant can be written as a smooth function of the entries of the invariant matrix valued functions $A_o, \dots, A_{2n-1}$ on $L_{m,n,p}(\mathbb{R})$.

7. ON THE (NON) EXISTENCE OF CANONICAL FORMS.

7.1. Canonical forms.

Let L' be a $GL_n(\mathbb{R})$-invariant subspace of $L_{m,n,p}(\mathbb{R})$. A canonical form for $GL_n(\mathbb{R})$ acting on L' is a mapping c: L' → L' such that the following three properties hold

(7.2) $c(\Sigma^S) = c(\Sigma)$ for all $\Sigma \in L'$, $S \in GL_n(\mathbb{R})$

(7.3) for all $\Sigma \in L'$ there is an $S \in GL_n(\mathbb{R})$ such that $c(\Sigma) = \Sigma^S$

(7.4) $c(\Sigma) = c(\Sigma') \Rightarrow \exists S \in GL_n(\mathbb{R})$ such that $\Sigma' = \Sigma^S$

(Note that (7.4) is implied by (7.3)).

Thus a canonical form selects precisely one element out of each orbit of $GL_n(\mathbb{R})$ acting on L'. We speak of a continuous canonical form if c is continuous.

Of course, there exist canonical forms on, say $L^{co,cr}_{m,n,p}(\mathbb{R})$, e.g. the following one, $\bar{c}_\kappa$: $L^{co,cr}_{m,n,p}(\mathbb{R}) \to L^{co,cr}_{m,n,p}(\mathbb{R})$ which is defined as follows: let $\Sigma \in L^{co,cr}_{m,n,p}(\mathbb{R})$, calculate $\kappa(\Sigma)$ and let $\bar{c}_\kappa(\Sigma)$ be the block companion canonical form of Σ as described in section 5.21 above.

This canonical form is not continuous, however (,though still quite useful, as we saw in section 5.31). As we argued in 1.15 above, for some purposes it would be desirable to have a continuous canonical form (cf, also[2]). In this connection let us also remark that the Jordan canonical form for square matrices under similarity transformations $(M \mapsto SMS^{-1})$ is also not continuous, and this causes a number of unpleasant numerical difficulties, cf. [16].

*7.5. Continuous canonical forms and sections. Let L' be a $GL_n(\mathbb{R})$-invariant subspace of $L^{cr}_{m,n,p}(\mathbb{R})$. Let $M' = \pi(L') \subset M^{cr}_{m,n,p}(\mathbb{R})$ be the image of L' under the projection π (cf. 6.2 above). Now let c: L' → L' be a continuous canonical form on L'. Then $c(\Sigma^S) = c(\Sigma)$ for all $\Sigma \in L'$ so that c factorizes through M' to define a continuous map s: M' → L' such that $c = s \circ \pi$. Because of (7.3) we have $\pi \circ c = \pi$ so that $\pi = \pi \circ s \circ \pi$. Because π is surjective it follows that $\pi \circ s = id_n$, so that s is a continuous section of the (principal $GL_n(\mathbb{R})$) fibre bundle π: L' → M'. Inversely let s: M' → L' be a continuous section of π. Then $s \circ \pi$: L' → L' is a continuous canonical form on L'.

7.6. <u>(Non)existence of global canonical forms</u>. In this section we shall prove theorem 1.17 which says that there exists a continuous canonical form on all of $L^{cr,co}_{m,n,p}(\mathbb{R})$ if and only if m = 1 or p = 1.

First suppose that m = 1. Then there is only one nice selection in I(n,m), viz. ((0,1), (1,1), ..., (n-1,1)). We have already seen that there exists a continuous canonical form $c_\alpha : U_\alpha \to U_\alpha$ for all nice selections α. (cf. 3.10). This proves the theorem for m = 1. The case p = 1 is treated similarly (cf. 3.11). It remains to prove that there is no continuous canonical form on $L^{co,cr}_{m,n,p}(\mathbb{R})$ if $m \geq 2$ and $p \geq 2$. To do this we construct two families of linear dynamical systems as follows for all $a \in \mathbb{R}$, $b \in \mathbb{R}$ (We assume $n \geq 2$; if n = 1 the examples must be modified somewhat).

$$G_1(a) = \left(\begin{array}{cc|ccc} a & 1 & 0 & \cdots & 0 \\ 1 & 1 & 0 & \cdots & 0 \\ \hline 2 & 1 & & & \\ \vdots & \vdots & & B & \\ 2 & 1 & & & \end{array}\right) \qquad G_2(b) = \left(\begin{array}{cc|ccc} 1 & b & 0 & \cdots & 0 \\ 1 & 1 & 0 & \cdots & 0 \\ \hline 2 & 1 & & & \\ \vdots & \vdots & & B & \\ 2 & 1 & & & \end{array}\right)$$

where B is some (constant) (n-2) × (m-2) matrix with coefficients in $\mathbb{R}$

$$F_1(a) = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & n \end{pmatrix} = F_2(b)$$

$$H_1(a) = \left(\begin{array}{cc|ccc} y_1(a) & 1 & 2 & \cdots & 2 \\ y_2(a) & 1 & 1 & \cdots & 1 \\ \hline 0 & 0 & & & \\ \vdots & \vdots & & C & \\ 0 & 0 & & & \end{array}\right) \qquad H_2(b) = \left(\begin{array}{cc|ccc} x_1(b) & 1 & 2 & \cdots & 2 \\ x_2(b) & 1 & 1 & \cdots & 1 \\ \hline 0 & 0 & & & \\ \vdots & \vdots & & C & \\ 0 & 0 & & & \end{array}\right)$$

where C is some (constant) real (p-2) × (n-2) matrix. Here the continuous functions

$y_1(a), y_2(a), x_1(b), x_2(b)$ are e.g. $y_1(a) = a$ for $|a| \leq 1$, $y_1(a) = a^{-1}$ for $|a| \geq 1$, $y_2(a) = \exp(-a^2)$, $x_1(b) = 1$ for $|b| \leq 1$, $x_1(b) = b^{-2}$ for $|b| \geq 1$, $x_2(b) = b^{-1}\exp(-b^{-2})$ for $b \neq 0$, $x_2(0) = 0$. The precise form of these functions is not important. What is important is that they are continuous, that $x_1(b) = b^{-1}y_1(b^{-1})$, $x_2(b) = b^{-1}y_2(b^{-1})$ for all $b \neq 0$ and that $y_2(a) \neq 0$ for all a and $x_1(b) \neq 0$ for all b.

For all $b \neq 0$ let $T(b)$ be the matrix

$$T(b) = \begin{pmatrix} b & 0 & \cdots & 0 \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{pmatrix} \tag{7.7}$$

Let $\Sigma_1(a) = (F_1(a), G_1(a), H_1(a))$, $\Sigma_2(b) = (F_2(b), G_2(b), H_2(b))$. Then one easily checks that

$$ab = 1 \Rightarrow \Sigma_1(a)^{T(b)} = \Sigma_2(b) \tag{7.8}$$

Note also that $\Sigma_1(a), \Sigma_2(b) \in L^{co,cr}_{m,n,p}(\mathbb{R})$ for all $a,b \in \mathbb{R}$; in fact

$$\Sigma_1(a) \in U_\alpha,\ \alpha = ((0,2), (1,2), \ldots, (n-1,2)) \text{ for all } a \in \mathbb{R} \tag{7.9}$$

$$\Sigma_2(b) \in U_\beta,\ \beta = ((0,1), (1,1), \ldots, (n-1,1)) \text{ for all } b \in \mathbb{R} \tag{7.10}$$

which proves the complete reachability. The complete observability is seen similarly.

Now suppose that c is a continuous canonical form on $L^{co,cr}_{m,n,p}(\mathbb{R})$. Let $c(\Sigma_1(a)) = (\bar{F}_1(a), \bar{G}_1(a), \bar{H}_1(a))$, $c(\Sigma_2(b)) = (\bar{F}_2(b), \bar{G}_2(b), \bar{H}_2(b))$. Let $S(a)$ be such that $c(\Sigma_1(a)) = \Sigma_1(a)^{S(a)}$ and let $\bar{S}(b)$ be such that $c(\Sigma_2(b)) = \Sigma_2(b)^{\bar{S}(b)}$.

It follows from (7.9) and (7.10) that

$$\begin{aligned} S(a) &= R(\bar{F}_1(a), \bar{G}_1(a))_\alpha\, R(F_1(a), G_1(a))_\alpha^{-1} \\ \bar{S}(b) &= R(\bar{F}_2(b), \bar{G}_2(b))_\beta\, R(F_2(b), G_2(b))_\beta^{-1} \end{aligned} \tag{7.11}$$

Consequently $S(a)$ and $\bar{S}(b)$ are (unique and are) continuous functions of a and b.

Now take $a = b = 1$. Then $ab = 1$ and $T(b) = I_n$ so that (cf(7.7), (7.8) and (7.11)) $S(1) = \bar{S}(1)$. It follows from this and the continuity of $S(a)$ and $\bar{S}(b)$ that we must have

$$(7.12) \qquad \operatorname{sign}(\det S(a)) = \operatorname{sign}(\det \bar{S}(b)) \quad \text{for all } a,b \in \mathbb{R}$$

Now take $a = b = -1$. Then $ab = 1$ and we have, using (7.8),

$$\begin{aligned} \Sigma_1(-1)^{\bar{S}(-1)T(-1)} &= (\Sigma_1(-1)^{T(-1)})^{\bar{S}(-1)} \\ &= \Sigma_2(-1)^{\bar{S}(-1)} = c(\Sigma_2(-1)) \\ &= c(\Sigma_1(-1)) = \Sigma_1(-1)^{S(-1)} \end{aligned}$$

It follows that $S(-1) = \bar{S}(-1)T(-1)$, and hence by (7.7), that

$$\det(S(-1)) = -\det(\bar{S}(-1))$$

which contradicts (7.12). This proves that there does not exists a continuous canonical form on $L^{co,cr}_{m,n,p}(\mathbb{R})$ if $m \geq 2$ and $p \geq 2$.

*7.13. <u>Acknowledgement and remarks</u>. By choosing the matrices B and C in $G_1(a), G_2(b), H_1(a), H_2(b)$ judiciously we can also ensure that $\operatorname{rank}(G_1(a) = m = \operatorname{rank} G_2(b)$ if $m < n$ and $\operatorname{rank} H_1(a) = p = \operatorname{rank} H_2(b)$ if $p < n$.

As we have seen in 7.5 above there exists a continuous canonical form on $L^{co,cr}_{m,n,p}(\mathbb{R})$ if and only if the principal $GL_n(\mathbb{R})$ fibre bundle $\pi : L^{co,cr}_{m,n,p}(\mathbb{R}) \to M^{co,cr}_{m,n,p}(\mathbb{R})$ admits a section. This, in turn is the case if and only if this bundle is trivial. The example on which the proof in 7.6 above is based precisely the same example we used in

[5] to prove the fibre bundle π is in fact nontrivial if $p \geq 2$ and $m \geq 2$, and from this point of view the example appears somewhat less "ad hoc" than in the present setting. The idea of using the example to prove nonexistence as done above is due to R.E. Kalman.

8. ON THE GEOMETRY OF $M^{co,cr}_{m,n,p}(\mathbb{R})$. HOLES AND (PARTIAL) COMPACTIFICATIONS.

As we have seen in the introduction (cf. 1.19) the differentiable manifold $M^{co,cr}_{m,n,p}(\mathbb{R})$ is full of holes, a situation which is undesirable in certain situations. In this section we prove theorems 1.23 and 1.24 but, for the sake of simplicity only in the case $m = 1$ or $p = 1$.

8.1. <u>An addendum to realization theory.</u> Let $T(s) = d(s)^{-1}b(s)$ be a rational function, with degree $d(s) = n >$ degree $b(s)$. Then we know by 4.27 that there is a one input one output system Σ with transfer function $T_\Sigma(s)$. We claim that we can see to it that $\dim(\Sigma) \leq n$. Indeed if

$$T_\Sigma(s) = a_o s^{-1} + a_1 s^{-2} + a_2 s^{-3} + \dots$$

then, if $d(s) = s^n - d_{n-1}s^{n-1} - \dots - d_1 s - d_o$, we have

$$a_{i+n} = d_o a_i + d_1 a_{i+1} + \dots + d_{n-1} a_{i+n-1}$$

for all $i \geq 0$. It follows that if $\mathcal{A} = (a_o, a_1, a_2, \dots)$, then rank $\mathcal{H}_r(\mathcal{A})$ = rank $\mathcal{H}_{n-1}(\mathcal{A})$ for all $r \geq n-1$. But $\mathcal{H}_{n-1}(\mathcal{A})$ is an $n \times n$ matrix and hence rank $\mathcal{H}_r(\mathcal{A}) \leq n$ for all s, which by section 4 means that there is a realization of $\mathcal{A}$ (or $T(s)$) of dimension $\leq n$.

It follows that a cr and co system Σ of dimension n has a transfer function $T_\Sigma(s) = d(s)^{-1}b(s)$ with degree $(d(s)) = n$ and no common factors in $d(s)$ and $b(s)$, and inversely if $T(s) = d(s)^{-1}b(s)$, degree $b(s) < n =$ degree $(d(s))$, and $b(s)$ and $d(s)$ have no common factors, then all n-dimensional realizations of $T(s)$ are co and cr.

Indeed if $d(s)$ and $b(s)$ have a common factor, then $T_\Sigma(s) = d'(s)^{-1}b'(s)$ with degree $(d'(s)) \leq n-1$ and it follows as above that rank $\mathcal{H}_r(\mathcal{A}) \leq n-1$ so that Σ is not cr and co. Inversely if Σ is not cr and co there is a Σ' of dimension $\leq n-1$ which also realizes $\mathcal{A}$ so that $T(s) = T_{\Sigma'}(s) = h'(sI-F')^{-1}g' = \det(sI-F')^{-1}B(s) = d'(s)^{-1}B(s)$ with degree$(d'(s)) \leq n-1$.

*8.2. There is a more input, more output version of 8.1. But it is not perhaps the most obvious possibility. E.g. the lowest dimensional realization of $s^{-1}\begin{pmatrix}1 & 2\\ 1 & 1\end{pmatrix}$ has dimension 2. The right generalization is: Let $T(s) = D(s)^{-1}N(s)$, where $D(s)$ and $N(s)$ are as in the theorem mentioned in section 5.9. Then there is a co and cr realization of $T(s)$ of dimension degree $(\det(D(s)))$.

8.3. <u>Theorem</u>. Let $D = a_o + a_1 \frac{d}{dt} + \ldots + a_{n-1} \frac{d^{n-1}}{dt^{n-1}}$, $a_i \in \mathbb{R}$ be a differential operator of order $\leq n-1$. Then there exists a family of systems $(\Sigma_z)_z \subset L^{co,cr}_{1,n,1}(\mathbb{R})$ such that the $f(\Sigma_z)$ converge to D in the sense of definition 1.21.

To prove this theorem we need to do some exercises concerning differentiation, determinants and partial integration. They are

(8.4) Let $k \in \mathbb{Z}$, $k \geq -1$ and let $B_{n,k}$ be the $n \times n$ matrix with (i,j)-th entry equal to the binomial coefficient $\binom{i+j+k}{i+k+1}$. Then $\det(B_{n,k}) = 1$.

(8.5) Let $u^{(i)}(t) = \dfrac{d^i u(t)}{dt^i}$. Then $\int_o^t z^n e^{-z(t-\tau)} u(\tau)d\tau =$

$$= z^{n-1}u(t) + \ldots + (-1)^{n-1}u^{(n-1)}(t) + O(z^{-1})$$

where O is the Landau symbol.

(8.6) Let $\phi(\tau) = (t-\tau)^m u(\tau)$, $\phi^{(i)}(\tau) = \dfrac{d^i\phi(\tau)}{d\tau^i}$. Then $\phi^{(i)}(t) = 0$ for $i < m$ and

$$\phi^{(i)}(t) = (-1)^m i(i-1)\ldots(i-m+1)u^{(i-m)}(t) \text{ if } i \geq m.$$

And finally, combining (8.5) and (8.6),

(8.7) $$\int_o^t e^{-z(t-\tau)} z^n (t-\tau)^m u(\tau)d\tau = (-1)^m m! \sum_{i=m+1}^{n} (-1)^{i+1} z^{n-i}\binom{i-1}{m} u^{(i-1-m)}(t) + O(z^{-1})$$

8.8. <u>Proof of theorem 8.3</u>. We consider the following family of n dimensional systems (with one output and one input),

$$g_z = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ z^m \end{pmatrix}, \quad F_z = \begin{pmatrix} -z & z & 0 & \cdots & 0 \\ 0 & -z & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & z \\ 0 & \cdots & & 0 & -z \end{pmatrix}, \quad h_z = (0,\ldots,0,x_m,\ldots,x_1)$$

where the $x_1, \ldots, x_m$, $m \leq n$, are same still to be determined real numbers. One calculates

$$e^{sF_z} = \begin{pmatrix} 1 & sz & \frac{s^2z^2}{2!} & \cdots & \frac{(sz)^{n-1}}{(n-1)!} \\ 0 & 1 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \frac{s^2z^2}{2!} \\ \vdots & & \ddots & 1 & sz \\ 0 & \cdots & \cdots & 0 & 1 \end{pmatrix}$$

Hence

$$h_z e^{(t-\tau)F_z} g_z = \sum_{i=1}^{m} x_i z^{m+i} (i!)^{-1} (t-\tau)^i e^{-z(t-\tau)}$$

and, using (8.7),

$$\int_o^t h_z e^{(t-\tau)F_z} g_z u(\tau)d\tau = \sum_{i=1}^{m} (i!)^{-1} x_i \sum_{j=i+1}^{m+i} (-1)^i (i!)(-1)^{j+1} \binom{j-1}{i} z^{m+i-j}$$

$$u^{(j-i-1)}(t) + 0(z^{-1})$$

$$= \sum_{\ell=o}^{m-1} (-1)^{m-\ell+1} z^{\ell} \left(\sum_{i=1}^{m} x_i \binom{m+i-\ell-1}{i} \right) u^{(m-\ell-1)}(t) + 0(z^{-1})$$

Now, by (8.4) we know that $\det(\binom{m+i-\ell-1}{i}_{i,\ell}) = 1$, so that we can choose $x_1, \ldots, x_m$ in such a way that

$$\int_o^t h_z e^{(t-\tau)F_z} g_z u(\tau)d\tau = a_{m-1} u^{(m-1)}(t) + 0(z^{-1})$$

where a_{m-1} is any pregiven real number.

It follows that $\lim_{z\to\infty} f(\Sigma_z) = a_{m-1} \frac{d^{m-1}}{dt^{m-1}}$

Let $\Sigma_z(i) = (F_z(i), g_z(i), h_z(i))$, $i = 0, \ldots, n-1$ be systems constructed as above with limiting input/output operator equal to $a_i \dfrac{d^i}{dt^i}$. Now consider the n^2-dimensional systems $\hat{\Sigma}_z$ defined by

$$\hat{F}_z = \begin{pmatrix} F_z(0) & 0 & \cdots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & F_z(n-1) \end{pmatrix}, \quad \hat{g}_z = \begin{pmatrix} g_z(0) \\ \vdots \\ g_z(n-1) \end{pmatrix}, \quad \hat{h}_z = (h_z(0), \ldots, h_z(n-1))$$

Then clearly $\lim\limits_{z\to\infty} f(\hat{\Sigma}_z) = D$. Let $T_z^{(i)}(s)$ be the transfer function of $\Sigma_z(i)$. Then for certain polynomials $B_z^{(i)}(s)$ we have

$$(8.9) \qquad T_z^{(i)}(s) = d_z(s)^{-1} B_z^{(i)}(s), \quad d_z(s) \ \underline{\text{independent of } i}$$

The transfer function of $\hat{\Sigma}_z$ is clearly equal to

$$(8.10) \qquad T_z(s) = \sum_{i=o}^{n-1} T_z^{(i)}(s) = d_z(s)^{-1} B_z(s), \quad B_z(s) = \sum_{i=o}^{n-1} B_z^{(i)}(s)$$

By 8.1 it follows from (8.10) that $T_z(s)$ can also be realized by an n-dimensional system, Σ'_z. Then also $\lim\limits_{z\to\infty} f(\Sigma'_z) = D$. Finally we can change Σ'_z slightly to Σ_z for all z to find a family $(\Sigma_z)_z \subset L^{co,cr}_{1,n,1}(\mathbb{R})$ such that $\lim\limits_{z\to\infty} f(\Sigma_z) = D$. This proves the theorem.

8.11. Corollary. Let Σ' be a system of dimension i and let D be a differential operator of order $n - i - 1$ (where order(0)=-1). Then there exists a family $(\Sigma_z)_z \subset L^{co,cr}_{1,n,1}(\mathbb{R})$ such that $\lim\limits_{z\to\infty} f(\Sigma_z) = D + f(\Sigma')$.

Proof. Let $\Sigma''_z = (F''_z, g''_z, h''_z)$ be a family in $L_{1,n-i,1}(\mathbb{R})$ such that $\lim\limits_{z\to\infty} f(\Sigma''_z) = D$. Let $\Sigma' = (F', g', h')$. Let $\hat{\Sigma}_z$ be the n-dimensional system defined by the triple of matrices

$$\hat{F}_z = \begin{pmatrix} F''_z & 0 \\ 0 & F' \end{pmatrix}, \quad g_z = \begin{pmatrix} g''_z \\ g' \end{pmatrix}, \quad h_z = (h''_z, h')$$

Then $\lim\limits_{z\to\infty} f(\hat{\Sigma}_z) = D + f(\Sigma')$. Now perturb $\hat{\Sigma}_z$ slightly for each z to Σ_z, to find a completely reachable and completely observable family $(\Sigma_z)_z$ such that $\lim_{z\to\infty} f(\Sigma_z) = D + f(\Sigma')$.

8.12. Theorem. Let $(\Sigma_z)_z \subset L_{1,n,1}(\mathbb{R})$ be a family of systems which converges in input-output behaviour in the sense of definition 1.21. Then there exist a system Σ' and a differential operator D such that $\dim(\Sigma') + \mathrm{ord}(D) \leq n-1$ and $\lim_{z\to\infty} f(\Sigma_z) = f(\Sigma') + D$

Proof. Consider the relation

$$y_z(t) = f(\Sigma_z)u(t)$$

for smooth input functions $u(t)$. Let $\hat{u}(s)$ and $\hat{y}_z(s)$ be the Laplace transforms of $u(t)$ and $y_z(t)$. Then we have

$$\hat{y}_z(s) = T_z(s)\hat{u}(s)$$

where $T_z(s)$ is the transferfunction of Σ_z. Because the $f(\Sigma_z)$ converge as $z \to \infty$ (in the sense of definition 1.21), and because the Laplace transform is continuous, it follows that there is a rational function $T(s) = d(s)^{-1}b(s)$ with degree $d(s) \leq n$, degree $b(s) \leq n-1$ such that

$$\lim_{z\to\infty} T_z(s) = T(s)$$

pointwise in s for all but finitely many s. Write

$$T(s) = e_o + e_1 s + \dots + e_{n-i-1}s^{n-i-1} + \frac{b'(s)}{d'(s)}$$

with degree $d'(s) = i$, degree$(b'(s)) < i$. Let Σ' be a system of dimension $\leq i$ with transfer function equal to $d'(s)^{-1}b'(s)$ and let

D be the differential operator $e_o + e_1 \frac{d}{dt} + \dots + e_{n-i-1} \frac{d^{n-i-1}}{dt^{n-i-1}}$.

The Laplace transform of the relation

$$y(t) = f(\Sigma')u(t) + Du(t)$$

for smooth input functions $u(t)$, is

$$\hat{y}(s) = T(s)\hat{u}(s)$$

Because the Laplace transform is injective (on smooth functions) it follows that

$$\lim_{z\to\infty} f(\Sigma_z) = f(\Sigma') + D$$

*8.13. Remarks on compactification, desingularization, symmetry breaking, etc.

The more input, more output versions of theorems 8.3 and 8.12 are also true. To prove them it is more convenient to use another technique which is based on a continuity property of the inverse Laplace transform for certain sequences of functions. (The inverse Laplace transform is certainly not continuous in general; also it is perfectly possible to have a sequence of systems Σ_z such that their transfer functions $T_z(s)$ converge for $z \to \infty$, but such that the $f(\Sigma_z)$ do not converge, e.g. $T_z(s) = z(z-s)^{-1}$).

Let Σ be a co and cr system of dimension n with one input and one output. Let T(s)

$$T(s) = \frac{b_{n-1}s^{n-1}+\ldots+b_1s+b_o}{s^n+d_{n-1}s^{n-1}+\ldots+d_1s+d_o} = \frac{b(s)}{d(s)}$$

be the transfer function of Σ. Assign to T(s) the point

$$(b_o:\ldots:b_{n-1}:d_o:\ldots:d_{n-1}:1) \in \mathbb{P}^{2n}(\mathbb{R})$$

real projective space of dimension 2n. This defines an embedding of $M^{co,cr}_{1,n,1}(\mathbb{R})$ into $\mathbb{P}^{2n}(\mathbb{R})$. The image is obviously dense so that $\mathbb{P}^{2n}(\mathbb{R})$ is a smooth compactification of $M^{co,cr}_{1,n,1}(\mathbb{R})$.

Let $\bar{M}_{1,n,1}(\mathbb{R})$ be the subspace of $\mathbb{P}^{2n}(\mathbb{R})$ consisting of those points $(x_o:\ldots:x_{n-1}:y_o:y_1:\ldots:y_n) \in \mathbb{P}^{2n}(\mathbb{R})$ for which at least one y_i, $i = 0, \ldots, n$ is different from zero. For these points

$$\frac{x_o+x_1s+\ldots+x_{n-1}s^{n-1}}{y_o+y_1s+\ldots+y_ns^n}$$

has meaning and this rational function is then the transfer function of a generalized linear dynamical system:

$$\text{(8.14)} \qquad \begin{aligned} \dot{x} &= Fx + Gu \\ y &= Hx + Du \end{aligned}$$

where D is a differential operator. (The points in $\mathbb{P}^{2n}(\mathbb{R}) \setminus \bar{M}_{1,n,1}$ corresponds to "systems" which tend to give infinite outputs for finite inputs; they are interpretable, however, in terms of correspondences $y(t) \mapsto u(t)$).

Further let $\hat{M}_{1,n,1}$, consist of those $(x_o:\ldots:x_{n-1}:y_o:\ldots:y_n)$ for which if $y_i = 0$ for $i \geq r$ then also $x_{i-1} = 0$, $i \geq r$. For these points the D in (8.14) is zero and these points thus yield transfer functions of systems of dimension $\leq n$. (But many points in $\hat{M}_{1,n,1}$ have the same transfer functions). Assigning to a point in $\hat{M}_{1,n,1}$ the first 2n + 1 coefficients of

$$\frac{x_o+x_1s+\ldots+x_{n-1}s^{n-1}}{y_o+y_1s+\ldots+y_ns^n} = a_os^{-1} + a_1s^{-2} + a_2s^{-3} + \ldots$$

we find the following situation

$$\begin{array}{ccc} M^{co,cr}_{1,n,1} & \subset & \hat{M}_{1,n,1} \\ \downarrow{\scriptstyle \mathcal{H}} & & \downarrow{\scriptstyle \hat{\mathcal{H}}} \\ \mathbb{R}^{2n+1} & = & \mathbb{R}^{2n+1} \end{array}$$

Here $\mathcal{H}$ is an embedding and its image is the subspace of all sequences $\mathcal{A} = (a_o,\ldots,a_{2n})$ such that rank $\mathcal{H}_{n-1}(\mathcal{A})$ = rank $\mathcal{H}_n(\mathcal{A})$ = n. The image of $\hat{\mathcal{H}}$ is the space of all sequences $\mathcal{A}$ such that rank $\mathcal{H}_n(\mathcal{A})$ = rank $\mathcal{H}_{i-1}(\mathcal{A})$ = i for some $i \leq n$. This is a singular submanifold of $\mathbb{R}^{2n+1}$ and $\hat{\mathcal{H}}$ is a resolution of singularities.

The points of $(\hat{M}_{1,n,1} \setminus M^{co,cr}_{1,n,1})$ correspond to transfer functions of lower dimensional co and cr systems. If a sequence $x_z \in M^{co,cr}_{1,n,1}$

converges to such a point, the internal symmetry group $GL_n(\mathbb{R})$ of x_z suddenly contracts to some $GL_m(\mathbb{R}) \subset GL_n(\mathbb{R})$ with $m < n$.

REFERENCES.

1. J.L. Casti, Dynamical Systems and their Applications: Linear Theory, Acad. Pr., 1977.
2. M.J. Denham, Canonical Forms for the Identification of Multivariable Linear Systems, IEEE Trans. Automatic Control 19,6(1974), 646-656.
3. M. Hazewinkel, R.E. Kalman, Moduli and Canonical Forms for Linear Dynamical Systems, Report 7504, Econometric Institute, Erasmus Univ. Rotterdam, 1975.
4. M. Hazewinkel, R.E. Kalman, On Invariants, Canonical Forms and Moduli for Linear, Constant, Finite Dimensional, Dynamical Systems. In: Proc. CNR-CISM Symp. Algebraic System Theory (Udine 1975). Lect. Notes Economics and Math. Systems 131(1976), Springer, 48-60.
5. M. Hazewinkel, Moduli and Canonical Forms for Linear Dynamical Systems II: The Topological Case, J. Math. System Theory 10(1977), 363-385.
6. M. Hazewinkel, Moduli and Canonical Forms for Linear Dynamical Systems III: The Algebraic-Geometric Case. In: C. Martin, R. Hermann(eds), The 1976 AMES Research Centre (NASA) Conf. on Geometric Control, Math. Sci. Press, 1977, 291-360.
7. M. Hazewinkel, Degenerating Families of Linear Dynamical Systems I: Proc. 1977 IEEE Conf. on Decision and Control (New Orleans), 258-264.
8. M. Hazewinkel, Invariants, Moduli and Canonical Forms for Linear Time-varying Dynamical Systems, Report 7725, Econometric Inst., Erasmus Univ. Rotterdam, 1977 (Submitted IEEE Trans Circuits and Systems).
9. R.E. Kalman, Lectures on Controllability and Observability. In: G. Evangilisti (ed), Controllability and Observability (CIME, 1968), Edizione Cremonese, 1969, 1-151.
10. R.E. Kalman, P.L. Falb, M.A. Arbib, Topics in Mathematical Systems Theory, McGraw-Hill, 1969.
11. R.E. Kalman, Kronecker Invariants and Feedback. In: L. Weiss (ed), Ordinary Differential Equations (1971, NRL-MRC conference), Acad. Press, 1972, 459-471.

12. Kar-Keung, D. Young, P.V. Kokotovic, V.I. Utkin, A Singular Perturbation Analysis of high-gain Feedback Systems. IEEE Trans. Automatic Control 22, 6 (1977), 931-938.

13. C. Martin, R. Hermann, Applications of Algebraic Geometry to Systems Theory: The McMillan degree and Kronecker indices of transfer functions as Topological and Holomorphic System Invariants, Preprint, 1977.

14. H.H. Rosenbrock, State Space and Multivariable Theory, Nelson, 1970.

15. J.C. Willems, K. Glover, Parametrizations of Linear Dynamical Systems: Canonical Forms and Identifiability, IEEE Trans. Automatic Control 19, 6(1974), 640-646.

16. J.H. Wilkinson, G.H. Golub, Ill-conditioned Eigensystems and Computation of the Jordan Canonical Form, SIAM Review 18, 4(1976), 578-619.

17. W.A. Wolowich, Linear Multivariable Systems, Springer, 1974.

18. W. Murray Wonham, Linear Multivariable Control. A geometric Approach, Lect. Notes Economics and Math. Systems 101, Springer, 1974.

LIST OF REPORTS 1978

7800 "List of Reprints, nos 200-208; Abstracts of Reports Second Half 1977".

7801/S "Conjugate TT-Variation and Process Inversion", by L. de Haan and S.J. Resnick.

7802/S "General Quadratic Forms in Normal Variates", by C. Dubbelman.

7803/S "On Bahadur's Representation of Sample Quantiles", by L. de Haan and E. Taconis-Haantjes.

7804/0 "Experiments with a Reduction Method for Nonlinear Programming Based on a Restricted Lagrangian", by G. v.d. Hoek.

7805/S "Derivatives of Regularly Varying Functions in IR^d and Domains of Attraction of Stable Distributions", by L. de Haan and S.I. Resnick.

7806/E "Estimation and Testing of Alternative Production Function Models", by S. Schim van der Loeff and R. Harkema.

7807/E "Money Illusion and Aggregation Bias", by J. van Daal.

7808/M "Aspects of Elliptic Curves: An Introduction", by R.J. Stroeker.

7809/E "Analytical Utility Functions Underlying Fractional Expenditure Allocation Models", by W.H. Somermeyer and J. van Daal.

7810/M " A New Proof of Cartier's Third Theorem", by M. Hazewinkel.

7811/M "On the (Internal) Symmetry Groups of Linear Dynamical Systems", by M. Hazewinkel.