Lectures in Applied Mathematics
Volume 18, 1980

ON FAMILIES OF LINEAR SYSTEMS: DEGENERATION PHENOMENA

Michiel Hazewinkel

ABSTRACT. In this paper we study families of linear dynamical systems $\dot{x} = Fx + Gu$, $y = Hx + Ju$, where the matrices F,G,H,J depend on a parameter c. Let V_c be the associated input/output operator. Then this paper contains results about what operators can arise as limits of the V_c as $c \to \infty$.

1. INTRODUCTION. This paper is concerned with an aspect of the theory of *families* of linear dynamical systems rather than single systems, viz. degeneration phenomena. As such it is part of a general program (briefly discussed in [Haz 3]) which consists of trying to carry through for families of systems (and hence systems over rings) all the nice results and constructions which one has for single systems over fields (or finding out how and why these results and constructions break down in this more general setting). This includes a systematic investigation of which constructions are continuous in the system parameters; that is, which constructions and calculations are stable (more or less) with respect to small perturbations or errors in the system parameters, a topic which obviously deserves at least some attention in a world full of uncertain measurements. And, in turn, this topic includes trying to find out what may happen to systems and associated objects when certain parameters go to zero (or infinity, or ...), which is the topic of this paper.

Still more motivation for studying families rather than single systems can be found in [Haz 3] and some results concerning other aspects of the theory of families (than the degeneration phenomena discussed below) can be found in [Haz 4] (fine moduli spaces, continuous canonical forms) and [HP] (pointwise-local-global isomorphism problems).

Here we discuss degeneration phenomena. That is, suppose there is given a family of systems

$$\Sigma(c):\ \dot{x} = Fx + Gu, \quad y = Hx + Ju \tag{1.1}$$

where the matrices F, G, H, J depend on a parameter c. What can be said about the limit as $c \to \infty$. For example let V_c be the input/output operator of $\Sigma(c)$

$$V_c:\ u(t) \mapsto y(t) = \int_0^t He^{F(t-\tau)}Gu(\tau)d\tau \tag{1.2}$$

and suppose that as $c \to \infty$ the operators V_c converge (in some suitable sense) to some operator V. What can be said about V? E.g. can V still be viewed as the input/output operator of some sort of processing device?

There are a number of reasons for being interested in such degeneration phenomena, some of which can be characterized by the key words or phrases: identification, high-gain feedback, almost F mod G invariant subspaces (and almost disturbance decoupling), dynamic observers (and invertability).

1.3. <u>Identification</u>. Suppose we have given some sort of input/output device which is to be modelled "as best as possible" by means of a linear dynamical system (1.1) of dimension n. Now if $S \in GL_n(\mathbb{R})$, then a system $\Sigma = (F,G,H,J)$ and $\Sigma^S = (SFS^{-1},SG,HS^{-1},J)$ have the same input/output operator. Let M be the space of orbits of this action of $GL_n(\mathbb{R})$ on the space L of all n-dimensional systems (with a given number of inputs and outputs). The best we can do on the basis of input/output data alone is to identify the orbit of Σ (and even that is not true if Σ is not completely observable and completely reachable, a fact which can be expected to cause a fair amount of extra trouble). Thus we are trying to identify a point of M

and we can picture identification as finding (or guessing at) a sequence of points in M representing better and better identifications as more and more data come in. From this point of view the question naturally arises. Does a "converging" sequence of points in M necessarily have a limit in M? The answer is no. It is perfectly possible for a sequence of linear dynamical systems (1.1) to have a limiting input/output behaviour which is not the input/output behaviour of any system like (1.1) as the following example shows

$$\Sigma(c):\ \dot{x} = \begin{pmatrix} -c & -c \\ 0 & -c \end{pmatrix} x + \begin{pmatrix} 1 \\ 1 \end{pmatrix} u, \quad y = (c^2, 0)x \tag{1.4}$$

(one input/one output, dimension 2). Let u be a smooth bounded function on $\mathbb{R}$ with compact support in $(0,\infty)$, then if $y_c = V_c u$ a little partial integration shows that $\lim_{c\to\infty} y_c(t) = \frac{d}{dt} u(t)$, uniformly in t on bounded t intervals, and $\frac{d}{dt}$ cannot possibly be the input/output operator of a system (1.1), (e.g. because $\frac{d}{dt}$ is not bounded on smooth bounded functions in $[0,1]$ while all the V_Σ are bounded operators).

The presence of these "holes" is by no means the only difficulty in identification caused by the nontrivial topology and geometry of M. For some more remarks concerning this topic cf. [Haz. 2] (though the point of view I took there is still a good deal too optimistic) and also [BK].

1.5. High-gain Feedback. Consider a system with output feedback loop

$$\dot{x} = Fx + Gu, \quad y = Hx, \quad u = Ly\,. \tag{1.6}$$

What happens when L or certain entries of L go to infinity? For instance in [YKU] it is shown in the case of a large scalar gain factor $L = g$ and under some additional hypothesis the system (1.6) can be transformed into the standard singular perturbation framework

$$\dot{x}_1 = F_{11}x_1 + F_{12}x_2, \quad \mu\dot{x}_2 = F_{21}x_2 + F_{22}x_2, \quad \mu = g^{-1} \tag{1.7}$$

(with $F_{21} = 0$ in the case considered in [YKU], so that there is a separation of slow and fast modes; more precisely there is a fast subsystem which in the setting of [YKU] is asymptotically stable (if μ is small enough) feeding into a slow system). Of course setting $\mu = 0$ in (1.7) yields little information about (1.7) for small μ and the idea is rather to study (1.7) and (1.6) as perturbations of the limit behaviour as μ goes to zero or various coefficients of L go to infinity. In the setting of [YKU] the limit input/output operator is the zero operator, but in general this need not be the case, and one may hope that on the basis of some knowledge about what limit operators can arise it will prove possible to obtain some results on the lines of [YKU] and related papers in more general situations.

For some motivation for studying (very) high-gain feedback cf. [YKU] and some of the references therein, cf. also below in 1.8.

1.8. Almost F mod G Invariant Subspaces and Almost Disturbance Decoupling. An F mod G invariant subspace for $\dot{x} = Fx + Gu$ is a subspace V of the state space such that once one is in it one can stay in it. As is well known (cf. [Won]) these subspaces "solve" the disturbance decoupling problem. An almost F mod G invariant subspace is one such that once one is in it one can stay arbitrarily close to it, and these spaces "solve" an almost disturbance decoupling problem, which turns out to be important especially when the disturbances (partly) come in on the same channels as the inputs (cf. [Wil 1, Wil 2]).

A subspace V of dimension r is almost F mod G invariant if and only if there is for every $\varepsilon > 0$ a feedback matrix K_ε such that $(F+GK_\varepsilon)V$ is within ε of V (in a suitable sense), and if V is almost F mod G invariant but not F mod G invariant, K_ε will not remain finite as $\varepsilon \to 0$. Thus implementing a decoupling by means of an almost F mod G invariant subspace will give rise to a family of systems.

$$\dot{x} = (F+GK_\varepsilon)x + Gu + G'v, \quad y = Hx \tag{1.9}$$

where K_ε does not necessarily remain finite as $\varepsilon \to 0$.

1.10. Dynamic Observers. In [BM1], [BM2] Basile and Marro consider the problem of constructing observers for the state of a system (1.1) when the inputs are unknown. For this it is advantageous to have differential operators (cf. loc. cit) and these, as is suggested by the example (1.4), may be approximated by systems (1.1) (of comparable rank), thus giving us arbitrarily good approximate observers of the form (1.1).

1.11. More General Linear Systems? As we shall see the limit operators as $c \to \infty$ of the input/output operators V_c of a family of systems $\Sigma(c)$ are necessarily of the form $V_\Sigma + L(D)$, where Σ is a system (1.1) (and V_Σ its input/output operator) and where $L(D)$ is a polynomial matrix (with constant coefficients) in the differentiation operator $D = \frac{d}{dt}$. I.e. the possible limit operators are the input/output operators of systems of the type

$$\dot{x} = Fx + Gu, \quad y = Hx + J(D)u \tag{1.12}$$

where $J(s)$ is a matrix of polynomials, arguing that this wider class of systems is in some ways a more natural class to study than the class of systems (1.1), cf. also [Ros 1, Ros 2].

2. STATEMENT OF THE THEOREMS. The first thing to do is to specify in what sense we shall understand the phrase "the family of input/output operators L_c converges to the operator L as $c \to \infty$." And, in turn, this means that we must describe the spaces of functions between which these operators act.

2.1. The Spaces $\mathscr{F}^{(o)}(\mathbb{R}^r)$ and $\mathscr{F}_0(\mathbb{R}^r)$. The elements of $\mathscr{F}^{(o)}(\mathbb{R}^r)$ are all smooth functions $z: \mathbb{R} \to \mathbb{R}^r$ with support in $(0,\infty)$ and of no more than exponential growth. Here the support of a function z is as usual defined as the closure of the set of all $t \in \mathbb{R}$ where $z(t) \neq 0$. Thus $z \in \mathscr{F}^{(o)}(\mathbb{R}^r)$ iff there are an $\varepsilon > 0$, an $M > 0$, and $b \geq 0$ such that $z(t) = 0$ for $t \leq \varepsilon$ and

$$\| e^{-bt} z(t) \| \leq M \quad \text{for all} \quad t \tag{2.2}$$

(Both ε and b (and of course also M) may depend on the function z.) This class of functions includes the smooth functions of slow growth with support in $(0,\infty)$ (cf. [Ze, Chapter IV]), which space in turn contains the subspace $\mathscr{F}_o(\mathbb{R}^r)$ of smooth functions with compact support in $(0,\infty)$.

A sequence of functions $z_c \in \mathscr{F}^{(o)}(\mathbb{R}^2)$ is said to *converge* to $z \in \mathscr{F}^{(o)}(\mathbb{R}^r)$ if there is a b such that

$$\lim_{c\to\infty} \sup_t \| e^{-bt}(z_c(t)-z(t))\| = 0 \tag{2.3}$$

Note that (2.3) in any case implies that the functions $z_c(t)$ converge to $z(t)$ uniformly in t on bounded t intervals.

This defines a topology on $\mathscr{F}^{(o)}(\mathbb{R}^r)$, which is in fact the inductive limit topology defined by the inductive system of normal topological vector spaces

$$\mathscr{F}_b^{(o)}(\mathbb{R}^r),\ i_{b,b'}: \mathscr{F}_b^{(o)}(\mathbb{R}^r) \to \mathscr{F}_{b'}^{(o)}(\mathbb{R}^r),\quad b' \ge b \tag{2.4}$$

where for a given $b \in \mathbb{R}$

$$\mathscr{F}_b^{(o)}(\mathbb{R}^r) = \{z \in \mathscr{F}^{(o)}(\mathbb{R}^r) \mid \sup_t \| e^{-bt}z(t)\| =: \|z\|_b < \infty\} \tag{2.5}$$

with the norm $\|z\|_b$, and where $i_{b,b'}$ is defined by $z(t) \to e^{(b'-b)t}z(t)$.

The space $\mathscr{F}^{(o)}(\mathbb{R}^r)$ tries hard to be complete in the sense of the following lemma.

2.6. LEMMA. *Let* $\eta > 0$ *and let* $z_c \in \mathscr{F}^{(o)}(\mathbb{R}^r)$ *be a sequence of functions with support in* $[\eta,\infty)$ *for all* c. *Suppose that there is a* $b \in \mathbb{R}$ *such that for all* $\varepsilon > 0$ *there is a* c_o *such that*

$$\sup_t \| e^{-bt}(z_c(t)-z_{c'}(t)\| < \varepsilon \quad \text{for all } c,c' \ge c_o \tag{2.7}$$

Then the z_c *converge to a function* $z \in \mathscr{F}^{(o)}(\mathbb{R}^r)$ *with support in* $[\eta,\infty)$ *as* $c \to \infty$ *(where the convergence is in the sense of* (2.3)*).*

Proof. Let $z(t)$ be the pointwise limit of $z_c(t)$ as $t \to \infty$ (which clearly exists by (2.7)). Then supp $z(t) \subset [\eta,\infty)$ and $z_c(t)$ converges to $z(t)$ uniformly on bounded t intervals (again by (2.7)). It follows that $z(t)$ is smooth. Take $\varepsilon = 1$ and let c_1 be such that that (2.7) holds for this ε with $c_0 = c_1$. Let $z_{c_1}(t) \in \mathscr{F}_{b_1}^{(o)}(\mathbb{R}^r)$. We can assume $b_1 \geq b$. Then, using $b_1 \geq b$,

$$e^{-b_1 t}\| z(t)\| \leq e^{-b_1 t}\| z_{c_1}(t)\| + e^{-bt}\| z_{c_1}(t)-z_{c'}(t)\| + e^{-bt}\| z_{c'}(t)-z(t)\|$$

Choosing c' depending on t such that $\| z_{c'}(t)-z(t)\| < 1$ it follows that $z(t) \in \mathscr{F}_{b_1}^{(o)}(\mathbb{R}^r) \subset \mathscr{F}^{(o)}(\mathbb{R}^r)$, proving the lemma.

Just what $b \in \mathbb{R}$ is used in (2.3) is largely irrelevant. Firstly, if (2.3) holds for a given b then it still holds with b replaced by $b' \geq b$. Secondly, if (2.3) holds and $z \in \mathscr{F}_{b'}^{(o)}(\mathbb{R}^r)$ then $z_c \in \mathscr{F}_{b''}^{(o)}(\mathbb{R}^r)$ for all large enough c where $b'' = \max(b,b')$. The converse of this: "if $z_c(t) \in \mathscr{F}_{b'}^{(o)}(\mathbb{R}^r)$ for all large enough c then $z(t) \in \mathscr{F}_{b''}^{(o)}(\mathbb{R}^r)$ with $b'' = \max(b,b')$" follows as in the lemma. Thirdly, and lastly, it does not really matter if one uses "too big a b" in (2.3). Indeed, $z(t)$ as the pointwise limit of the $z_c(t)$ is of course independent of b. What (2.3) does is to require a certain mild uniformity about the way the limit is approached. (It is, incidentally, perfectly possible for a sequence of functions $z_c(t) \in \mathscr{F}_b^{(o)}(\mathbb{R}^r)$ to converge to zero when considered as elements of $\mathscr{F}_{b'}^{(o)}(\mathbb{R}^r)$ for $b' > b$ while not converging when considered as a sequence in $\mathscr{F}_b^{(o)}(\mathbb{R}^r)$; take for example $z_c(t) = 0$ for $t \leq c$, $z_c(t) = e^{bt} - e^{bc}$ for $t \geq c$, suitably smoothed.)

2.8. The Spaces $\mathscr{F}(\mathbb{R}^r)$. For the purposes below the spaces $\mathscr{F}^{(o)}(\mathbb{R}^r)$ are still too big to be suitable as input spaces (essentially because we shall want differentiation to be a continuous operator). On the other hand $\mathscr{F}_o(\mathbb{R}^r)$, while eminently suitable as an input function space is not large enough to accommodate output functions. As we shall need to be able to use the outputs of one dynamical system as the inputs of another, we need an intermediate space. A suitable one is

$$\mathscr{F}(\mathbb{R}^r) = \{z \in \mathscr{F}^{(o)}(\mathbb{R}^r) \mid z^{(k)} \in \mathscr{F}^{(o)}(\mathbb{R}^r) \text{ for all } k = 0,1,2,\dots\} \tag{2.9}$$

where $z^{(k)}$ denotes the k-th derivative of z. We give $\mathscr{F}(\mathbb{R}^r)$ the topology determined by $z_c \to z$ as $c \to \infty$ iff $z_c^{(k)} \to z^{(k)}$ for all $k = 0,1,2,\dots$ in $\mathscr{F}^{(o)}(\mathbb{R}^r)$. Thus the family z_c converges to z as $c \to \infty$ iff there are real numbers $b_o, b_1, \dots$ such that for all k

$$\lim_{c\to\infty} \sup_t e^{-b_k t} \| z_c^{(k)}(t) = z^{(k)}(t) \| = 0$$

When dealing with systems of dimension $\leq n$ only, one can also work with $\mathscr{F}^{(n)}(\mathbb{R}^r) = \{z \in \mathscr{F}^{(o)}(\mathbb{R}) \mid z^{(k)} \in \mathscr{F}^{(o)}(\mathbb{R})$, $k = 0,.,\dots,n+1\}$.

2.10. Convergence of Input/Output Operators. Now let $\Sigma = (F,G,H,J)$ be a linear dynamical system with direct feed-through term

$$\dot{x} = Fx + Gu, \quad y = Hx + Ju \qquad x \in \mathbb{R}^n, \quad y \in \mathbb{R}^p, \quad u \in \mathbb{R}^m \tag{2.11}$$

where F, G, H, J are real matrices of the appropriate dimensions (independent of t). Then the associated input/output operator is defined by

$$V_\Sigma : u(t) \mapsto y(t) = Ju(t) + \int_0^t He^{F(t-\tau)}GU(\tau)d\tau \tag{2.12}$$

Let $\mathscr{U} = \mathscr{F}(\mathbb{R}^m)$, $\mathscr{Y} = \mathscr{F}(\mathbb{R}^p)$, $\mathscr{U}_o = \mathscr{F}_o(\mathbb{R}^m)$, $\mathscr{Y}_o = \mathscr{F}_o(\mathbb{R}^p)$. Then V_Σ is a continuous linear operator $\mathscr{U} \to \mathscr{Y}$. Indeed if $u \in U$ is such that $\|u\|_b < \infty$ and if $b' > \max\{\mathrm{Re}\lambda, 0\}$ where λ runs through the eignevalues of F then $\| V\Sigma(u) \|_{b+b'} < \infty$.

Thus for every $b \geq 0$ there is a $b' \geq 0$, usually necessarily larger than b, such that V_Σ maps $\mathscr{F}_b^{(o)}(\mathbb{R}^m)$ into $\mathscr{F}_{b'}^{(o)}(\mathbb{R}^p)$, with b' depending on Σ. Thus, when dealing with families of systems one is practically forced to use the union of the all the $\mathscr{F}_b^{(o)}(\mathbb{R}^p)$, i.e. $\mathscr{F}^{(o)}(\mathbb{R}^p)$, and if one would like differential operators to be continuous one is almost obliged to work with $\mathscr{F}(\mathbb{R}^p)$ and $\mathscr{F}(\mathbb{R}^m)$. From now on we fix the dimensions m,n,p of the systems (2.11) which we are considering. Let L denote the space of all systems (2.11). I.e. L is the space of all real quadruples of matrices (F,G,H,J) of the dimensions $n \times n$, $n \times m$, $p \times n$, $p \times m$ respectively.

We shall use L^{co}, L^{cr}, $L^{co,cr}$ to denote the subspaces of completely observable, (abbreviated co), resp. completely reachable (cr), resp. completely reachable and completely observable systems.

We now define

2.13. DEFINITION. *The family of systems* $\Sigma(c) \subset L$ *converges in input/output behaviour to an operator* V *iff for all* $u \in U$ *the functions* $V_{\Sigma(c)}u$ *converge to* Vu *in* $\mathscr{Y}$ *as* $c \to \infty$.

Let $\mathrm{supp}(u) \subset [\eta,\infty)$ (such an η necessarily exists because $\mathrm{supp}(u) \subset (0,\infty)$ and $\mathrm{supp}(u)$ is closed by definition). Then $\mathrm{supp}\, V_{\Sigma(c)}(u) \subset [\eta,\infty)$. It follows by lemma 2.6 that one can decide whether the family $(\Sigma(c))$ converges without mentioning (or knowing) the limit operator V. The family $(\Sigma(c))_c$ converges in input/output behaviour iff there are for every $u \in U$ a sequence of numbers $b_0,b_1,b_2,\ldots$ such that for every $\varepsilon > 0$, $k = 0,.,2,\ldots$ there is a $c(\varepsilon,k)$ such that

$$\sup_t\{e^{-b_k t}\|(D^k V_{\Sigma(c)}u)(t) - (D^k V_{\Sigma(c')}u)(t)\|\} < \varepsilon$$

$$\text{if} \quad c,c' \geq c(\varepsilon,k) \tag{2.14}$$

where D is the differentiation operator $D = \frac{d}{dt}$. Thus if $(\Sigma(c))$ converges in input/output behaviour (in the sense that (2.14) holds) then there is a well-defined limit operator V. (This uses of course (cf. (2.14)) that D is a continuous operator $\mathcal{U} \to \mathcal{U}$). Whether this limit operator V is continuous is unclear at this stage. (It is though, as will be shown below in section 5).

2.15. Differential Operators. Let $\mathcal{U}$ and $\mathcal{Y}$ be as above. Then a (matrix) differential opeartor (in this paper) is an operator of the form

$$V(D)\colon u(t) \mapsto y(t), \quad y_j(t) = \sum_{i=1}^{m} v_{ji}(D)u_i(t)$$

where $v_{ji}(D)$ is a polynomial with constant real coefficients in $D = d/dt$. Every polynomial $V(s)$ (of size $p \times m$) thus defines a continuous linear operator $\mathcal{U} \to \mathcal{Y}$.

2.16. The Scalar Case. If $m = 1 = p$, i.e. if we are dealing with one input and one output the main theorem of this paper says that

2.17. THEOREM. *Let* $(\Sigma(c))$ *be a family of one input/one output linear dynamical systems* (2.11) *of dimension* $\leq n$ *converging in input/output behaviour to the operator* $V\colon \mathcal{U} \to \mathcal{Y}$. *There there exist a system* Σ *and a polynomial* $L(s)$ *such that* $V = V_\Sigma + L(D)$, *where moreover* $\dim(\Sigma) + \text{degree } L(s) \leq n$. *It follows in particular that the limit operator* V *is continuous. Inversely, if* V *is an operator of the form* $V = V_\Sigma + L(D)$ *where* $L(s)$ *is a polynomial of degree* $\leq n - \dim(\Sigma)$, *then there exists a family* $(\Sigma(c)) \subset L^{co,cr}$ *such that* $\Sigma(c)$ *converges in input/output behaviour to* V.

In case one wants to restrict oneself to systems (2.11) with $J = 0$ the theorem remains essentially the same except that the essential inequality $\dim(\Sigma) + \text{degree } (L(s)) \leq n$ gets replaced by $\dim(\Sigma) + \text{degree } (L(s)) \leq n-1$ (where by definition degree $(0) = -1$). This is stated and proved (more or less) in [Haz 1] and the proof readily adapts to a proof of the present theorem. In section 5 below a different proof of theorem 2.17 is given which also covers the multivariable case.

2.18. Degree of a Matrix Polynomial (Differential Operator). Obviously if $\Sigma(c)$ is a family of systems of dimension $\leq n$ which converges to the $p \times m$ matrix differential operator $L(D)$

then all the entries of $L(s)$ have degree $\leq n$ (by the result in the scalar case). One might think that inversely every such operator arises as a limit of systems $\leq n$. This, however, is not the case as the example.

$$L(D) = \begin{pmatrix} D & 0 \\ 0 & D \end{pmatrix} \tag{2.19}$$

shows. One shows readily by explicit calculation that the operator (2.19) cannot arise as a limit of one-dimensional systems. A more sensitive definition of "degree" is needed.

2.19. DEFINITION. *Let* $L(s)$ *be a matrix polynomial. Then we define*

$$\deg L(s) = \max_{m} (\text{degree}(M)) \tag{2.20}$$

where M *runs over all the minors of* L. *This agrees with the MacMillan degree of a polynomial matrix, (lemma* 4.10*, or cf.* [AV]*, section* 3.6. *properties* 5 *and* 10*)*.

2.21. <u>The Multivariable Case</u>. In the case of more inputs more outputs the main theorem now is precisely analogous to theorem 2.17. I.e.

2.22. THEOREM. *Let* $\Sigma(c)$ *be a family of* n *dimensional systems with* m *intputs and* p *outputs. Suppose that* $\Sigma(c)$ *converges in input/output behaviour to the operator* $V: \mathcal{U} \to \mathcal{Y}$ *as* $c \to \infty$. *Then there exist a system* Σ *and a* $p \times m$ *matrix polynomial* $L(s)$ *such that* $V = V_\Sigma + L(D)$ *(so that* V *is continuous) and moreover* $\dim(\Sigma) + \text{degree } L(s) \leq n$. *Inversely if* V *is an operator of the form* $V_\Sigma + L(D)$ *with* $\dim(\Sigma) + \text{degree } L(s) \leq n$, *then there exists a family of completely observable and completely reachable systems* $\Sigma(c)$ *of dimension* $\leq n$ *which converges in input/output behaviour to* V *as* $c \to \infty$.

The proof of the first half of the theorem uses the continuity (in this case) of the Laplace transform and the upper semicontinuity of the MacMillan degree (theorem 4.16) and thus gives us (besides lemma 4.10) yet another characterization of the MacMillan degree of a matrix of rational functions.

2.23. THEOREM. *Let* $L(s)$ *be a matrix of rational functions. Then the MacMillan degree of* $L(s)$ *is* $\le n$ *iff there exists a sequence* $L_c(s)$ *of proper rational function matrices of degree* n *such that* $L_c(s)$ *converges to* $L(s)$ *for* $c \to \infty$ *pointwise in* s *for infinitely many values of* s. *Moreover one can see to it that the poles of* $L_c(s)$ *fall into two sets one equal (together with multiplicities) to the set of poles* $\neq \infty$ *of* $L(s)$ *while remaining poles of* $L_c(s)$ *all go to* $-\infty$ *as* $c \to \infty$.

It is not true, however, that one can always obtain $L(s)$ as a limit of the $L_c(s)$ in the sense of the mappings on the Riemann sphere that these matrices of rational functions define. This in fact only happens when $L(s)$ is itself proper.

To prove Theorem 2.23 without the extra requirement that the remaining poles of $L_c(s)$ go to $-\infty$ as c goes to ∞ is quite easy (Proposition 4.18). The extra requirement complicates things considerably and I know of no direct proof except for certain special, albeit generic, cases. (Like "the matrix of coefficients of maximal powers of s in each row is of maximal rank"). Another corrollary of the proof of the second half of Theorem 2.22 is

2.24. COROLLARY. *Let* $L(s)$ *be a polynomial matrix of size* $p \times m$. *Then* $L(s)$ *has degree* $\le n$ *if and only if it can be obtained from the zero matrix by means of the operations.*

(i) *addition of a matrix of constants*

(ii) *multiplication on the left by a nonsingular polynomial* $p \times p$ *matrix of degree* 1

(iii) *multiplication on the right by a nonsingular* $m \times m$ *matrix of constants*

where one uses at most n *times an operation of type* (ii).

There is of course an analogous statement with right instead of left in (ii) and left instead of right in (iii), and also an analogous statement where in both (ii) and (iii) multiplications on both sides are allowed.

3. ON LIMITS OF RATIONAL FUNCTIONS. The degree of a rational function $T(s) = q(s)^{-1}p(s)$, $p(s), q(s) \in k[s]$ with no common factors is equal to $\delta(T) = \max(\delta(p), \delta(q))$ where the degree of a polynomial is defined as usual. We shall need the following intuitively obvious fact.

3.1. PROPOSITION. *Let* $T_c(s)$ *be a sequence of rational functions of degree* $\le n$. *Suppose that* $\lim_{c\to\infty} T_c(s)$ *exists (and is finite) for infinitely many* s. *Then there exists a rational function* $T(s)$ *of degree* $\le n$ *such that* $\lim_{c\to\infty} T_c(s) = T(s)$ *for all but finitely many* s *(and if the* $T_c(s)$ *and* $T(s)$ *are interpreted as functions* $\mathbb{C} \to \mathbb{P}^1(\mathbb{C})$ *then* $T_c(s)$ *converges to* $T(s)$ *in the compact open topology).*

Proof. Write

$$T_c(s) = \frac{p_c(s)}{q_c(s)} = \frac{a_n(c)s^n + a_{n-1}(c)s^{n-1} + \dots + a_1(c)s + a_0(c)}{b_n(c)s^n + b_{n-1}(c)s^{n=1} + \dots + b_1(c)s + b_0(c)} \tag{3.2}$$

and associate to $T_c(s)$ the point $\psi(c) \in \mathbb{P}^{2n+1}(\mathbb{C})$ with the homogeneous coordinates $(a_n, \dots, a_0, b_n, \dots, b_0)$. Note that this is well defined because the coefficients of $p_c(s)$ and $q_c(s)$ are well defined up to a common scalar factor. (This map is not continuous if the space of all rational functions of degree $\le n$ is given the compact open topology of maps $\mathbb{C} \to \mathbb{P}^1(\mathbb{C})$; but it is continuous on the open subspace of function of degree n, and on the subspaces of functions of fixed degree i).

Let $M \subset \mathbb{P}^{2n+1}(\mathbb{C})$ be the subspace of all points $(x_n, \dots, x_0, y_n, \dots, y_0) \in \mathbb{P}^{2n+1}(\mathbb{C})$ such that at least one y_i is unequal to zero. Because $\mathbb{P}^{2n+1}(\mathbb{C})$ is compact the sequence $\{\psi(c)\}$ has limit points.

3.3 LEMMA. *If* $\lim_{c\to\infty} T_c(s)$ *exists for infinitely many* s *then all limit points of the sequence* $\{\psi(c)\}$ *are in* M.

Proof. Suppose that $\lim_{c\to\infty} T_c(s) = T(s) \in \mathbb{C}$, and suppose that $\{\psi(c)\}$ has a limit point in $\mathbb{P}^{2n+1}(\mathbb{C})\backslash M$. Let this limit point be $x = (a_n,\ldots,a_{i+1},1,0,\ldots,0)$. Taking a subsequence we can assume that $\{\psi(c)\}$ converges to x. For large enough c we then have $a_i(c) \neq 0$ and multiplying both $p_c(s)$ and $q_c(s)$ with $a_i(c)^{-1}$ we can assume that $a_i(c) = 1$ for all c. We then have for all c

$$a_n(c)s^n + \ldots + a_{i+1}(c)s^{i+1} + s^i + a_{i-1}(c)s^{i-1} + \ldots + a_0(c)$$
$$= T_c(s)(b_n(c)s^n + \ldots + b_0(c)) \qquad (3.4)$$

with

$$\lim_{c\to\infty} b_j(c) = 0, \quad j = 0,\ldots,n$$
$$\lim_{c\to\infty} a_j(c) = 0, \quad j = 0,\ldots,i-1 \qquad (3.5)$$
$$\lim_{c\to\infty} a_j(c) = a_j, \quad j = i+1,\ldots,n$$

Taking the limit as $c\to\infty$ in (3.4) and using the relations (3.5) one finds because $\lim_{c\to\infty} T_c(s) = T(s) \neq \infty$

$$a_n s^n + \ldots + a_{i+1}s^{i+1} + s^i = 0 \qquad (3.6)$$

and there are only finitely many s for which this can hold. Thus there can be no limit points of $\{\psi(c)\}$ in $\mathbb{P}^{2n+1}\backslash M$ if $\lim_{c\to\infty} T_c(s)$ exists (and is finite) for infinitely many s.

The proof of proposition now continues as follows. Let $x \in M \subset \mathbb{P}^{2n+1}(\mathbb{C})$, $x = (x_n,\ldots,x_0,y_n,\ldots,y_0)$. Because at least one of the $y_i \neq 0$ the expression

$$T_x(s) = \frac{x_n s^n + \ldots + x_1 s + x_0}{y_n s^n + \ldots + y_1 s + y_0} \qquad (3.7)$$

is well-defined for all but finitely many s. Now let $x \in M$ be a limit point of $\{\psi(c)\}$. Let i be the largest index such that $y_i \neq 0$. Multiplying all coordinates with y_i^{-1} if necessary, we can assume $y_i = 1$. Take a subsequence of $\{\psi(c)\}$ which converges to x. For large enough c we then have

$b_i(c) \neq 0$. Multiplying both $p_c(s)$ and $q_c(s)$ with $b_i(c)^{-1}$ we then obtain sequence of rational functions.

$$T_c(s) = \frac{a_n(c)^s s^n + \ldots + a_1(c)s + a_0(c)}{b_n(c)s^n + \ldots + s^i + \ldots + b_1(c)s + b_0(c)} \tag{3.8}$$

such that as $c \to \infty$.

$$a_j(c) \to x_j, \quad b_j(c) \to y_j, \quad j = 0,1,\ldots,n \tag{3.9}$$

It follows that $\lim_{c\to\infty} T_c(s) = T_x(s)$ for all but finitely many s, where the limit is a priori over the subsequence. In turn this says that $\lim_{c\to\infty} T_c(s) = T_x(s)$ for all but finitely many s of the infinitely many s for which $\lim T_c(s)$ was assumed to exist.

This holds for all limit points of $\{\psi(c)\}$, hence if x' is a second limit point of $\{\psi(c)\}$ then $T_x(s) = T_{x'}(s)$ for infinitely many s so that $T_x(s) = T_{x'}(s)$ if both x,x' are limit points of $\{\psi(c)\}$, and this in turn says that $\lim_{c\to\infty} T_c(s) = T_x(s)$ for all but finitely many s, where now we are dealing with the original sequence $\{T_c(S)\}$. This concludes the proof of the proposition (except for the last statement between brackets which is easy because by the above the convergence $T_c(s) \to T_x(s)$ really means that the coefficients, suitably normalized, converge).

3.10. COROLLARY. (of the proof) *Let* $T_c(s) \to T(s)$ *as* $c \to \infty$ *and let* $T_c(s) = q_c(s)^{-1}p_c(s)$, $T(s) = q(s)^{-1}p(s)$ *with no common factors. Suppose that degree* $p_c(s) \leq n'$ *for all* c. *Then degree* $p(s) \leq n'$.

This follows immediately because (using the notations of the proof) after a suitable normalization and for c large enough the coefficients of $p_c(s)$ converge to the coefficients of $p_x(s)$ where $p_x(s)$ is the numerator of (3.7), and because $q(s)^{-1}p(s) = T(s) = T_x(s) = q_x(s)^{-1}p_x(s)$ where $q_x(s)$ is the denominator of (3.7). So degree $p_x(s) \leq$ degree $p_c(s)$ for all large enough c. (Of course $p_x(s)$ and $q_x(s)$ may have common factors so that degree $p(s)$ may be smaller than $\liminf_{c\to\infty}(\text{degree}(p_c(s))))$.

4. ON THE DEGREE OF RATIONAL MATRICES. Recall that the MacMillan degree $\delta(T)$ of a matrix of rational functions $T(s)$ can be defined in a variety of ways ([Ka], [AV, section 3.6], [Ros, section 3.4]). First let $T(s)$ be proper, i.e. $\lim_{s\to\infty} T(s)$ exists, then $\delta(T) = \nu(T)$, which is by definition the minimal dimension of a realization (F,G,H,J) of $T(s)$. If $T(s)$ is not proper write

$$T(s) = T_-(s) + T_1 s + T_2 s^2 + \dots + T_r s^r, \qquad V(s) = T_1 s^{-1} + \dots + T_r s^{-r} \tag{4.1}$$

where $T_-(s)$ is the proper part of $T(s)$. Then $V(s)$ is also proper (in fact strictly proper, meaning that $\lim_{s\to\infty} V(s) = 0$) and we define

$$\delta(T) = \nu(T_-) + \nu(V) \quad . \tag{4.2}$$

This definition shows that if $T(s) = T_-(s) + T_+(s)$, where $T_-(s)$ is proper and $T_+(s)$ is polynomial then

$$\delta(T) = \delta(T_+) + \delta(T_-) \quad . \tag{4.3}$$

(It does not matter how the "constant part" of $T(s)$ is split up between T_- and T_+).

Another way to obtain $\delta(T)$ goes as follows (cf. [Kal]). Let $T(s)$ be a $p\times m$ matrix of rational functions. For each $m\times p$ matrix of constants K write

$$\det(I_m + KT(s)) = b_K(s)^{-1} a_K(s) \tag{4.4}$$

where I_m is the $m\times m$ identity matrix and $a_K(s)$, $b_K(s)$ are polynomials without common factors. Let

$$\delta_K(T) = \text{degree}(a_K(s)) \tag{4.5}$$

Then one has the proposition (cf. [Kal])

$$\delta(T) = \max_K \; \delta_K(T) \tag{4.6}$$

We shall need a few elementary properties of $\delta(T)$. If A and B are matrices of constants such that $AT(s)B$ is defined then (cf. [AV, (3.6.6)]

$$\delta(ATB) \le \delta(T) \tag{4.7}$$

(which is also immediately obvious from definition 4.2).

Now let $T'(s)$ be obtained from $T(s)$ by augmenting $T(s)$ with some rows and columns of constants. Then

$$\delta(T') = \delta(T) \tag{4.8}$$

This is seen as follows. Let $T(s)$ and $V(s)$ be as in (3.1) and let $T'_-(s)$ and $V'(s)$ be the analogous matrices for $T'(s)$. Then if (F,G,H,J) realizes $T_-(s)$ a realization for $T'_-(s)$ is obtained by adding some zero columns to G, some zero rows to H and by augmenting J with the same rows and columns of constants as were used to obtain $T'(s)$ from $T(s)$. Similarly a realization (F_1,G_1,H_1,J_1) for $V(s)$ can be changed in a realization of the same dimension for $V'(s)$ by augmenting G_1 with zero columns, H_1 with zero rows and J_1 with both zero rows and zero columns. This shows that $\delta(T') \le \delta(T)$. The opposite inequality follows from (4.7) because $T(s)$ is a submatrix of $T'(s)$.

A third result we need is: Let $T(s)$ be square such that $\det(T(s)) \not\equiv 0$. Then (cf. e.g. [Rose, theorem 7.2, p. 135)]

$$\delta(T^{-1}) = \delta(T) \tag{4.9}$$

As an application of (4.8) and (4.9) we show (using a few tricks which will also be useful further on).

4.10 LEMMA. *Let* $T(s)$ *be a matrix of polynomials. Then*

$$\delta(T) = \max_{M(s)} \; \{\mathrm{degree}(\det(M(s))\} \tag{4.11}$$

where $M(s)$ runs through all square submatrices of $T(s)$.

Proof. Define $\delta'(T)$ as being equal to the right hand side of (4.11). Then we have to prove that $\delta(T) = \delta'(T)$. Then the analogues of (4.7) and (4.8) also hold for δ', i.e.

$$\delta'(ATB) \leq \delta'(T), \quad \delta'(T') = \delta'(T) \tag{4.12}$$

To see this recall that a minor of a product of matrices is a sum of products of minors (of the same size) of the factors (cf. e.g. [Ros1], Thm. 1.3, p. 5) and that a minor of a matrix T' obtained by adding a row of constants or column of constants to T is either a minor of T or a sum of minors (of one size smaller) of T with constant coefficients. This proves (4.12).

It follows that if A and B are invertible then $\delta'(ATB) = \delta'(T)$. So by taking A and B to be suitable permutation matrices we can assume that T is of the form

$$T = \begin{pmatrix} T_{11} & T_{12} \\ & \\ T_{21} & T_{22} \end{pmatrix}$$

with $\deg(\det(T_{11}) = \delta'(T)$. Let the dimensions of T_{11}, T_{12}, T_{21}, T_{22} be respectively $r \times r$, $r \times (m-r)$, $(p-r) \times r$, $(p-r) \times (m-r)$. Let $T'(s)$ be the matrix

$$T'(s) = \begin{pmatrix} T_{11} & T_{12} & 0 \\ T_{21} & T_{22} & I \\ 0 & I' & 0 \end{pmatrix}$$

where I is the $(p-r) \times (p-r)$ unit matrix and I' the $(m-r) \times (m-r)$ unit matrix. Then by (4.12)

$$\delta'(T') = \delta'(T) \tag{4.13}$$

Also $\det(T') = \det(T_{11})$ so that degree $\det(T') \geq \text{degree}(M)$ for all minors M of T'. It follows that $T'(s)^{-1}$ is proper so that

$$\delta(T'(s)^{-1}) = \nu(T'(s)^{-1}) \tag{4.14}$$

At this stage we need one more property of the degree function which is essentially proved in [Ros1], cf. Thm. 4.3 on p. 115, cf. also [MH, section 2]. Viz.

4.15. LEMMA. *Let* $T(s)$ *be a* $p \times m$ *proper matrix of rational functions. Then there are polynomial matrices* $N(s)$, $D(s)$, *of sizes* $p \times m$, $m \times m$ *such that*

(i) $T(s) = N(s)D(s)^{-1}$

(ii) $N(s)$ *and* $D(s)$ *are right coprime, which means that there are polynomial matrices* $X(s)$, $Y(s)$ *such that* $X(s)N(s) + Y(s)D(s) = I_m$.

Moreover $N(s)$ *and* $D(s)$ *are unique up to a common unimodular right factor and* $\nu(T(s)) = \deg(\det D(s))$.

(The last statement of the lemma is more usually stated for strictly proper $T(s)$, i.e., matrices of rational functions $T(s)$ such that $\lim_{s\to\infty} T(s) = 0$; the slight extension is immediate; indeed if $T(s)$ is proper and $T(s) = J + \bar{T}(s)$, with $\bar{T}(s)$ strictly proper, $\bar{T}(s) = \bar{N}(s)\bar{D}(s)^{-1}$. Then $T(s) = N(s)D(s)^{-1}$ with $N(s) = J\bar{D}(s) + \bar{N}(s)$, $D(s) = \bar{D}(s)$, and if $\bar{X}(s)\bar{N}(s) + \bar{Y}(s)\bar{D}(s) = I_m$, then $X(s)N(s) + Y(s)D(s) = I_m$, with $X(s) = \bar{X}(s)$, $Y(s) = \bar{Y}(s) - \bar{X}(s)J$.)

Continuing with the proof of lemma 4.10. Applying lemma 4.15 to $T'(s)$ we find

$$\nu(T'(s)^{-1}) = \mathrm{degree}(\det(T'(s))) \tag{4.16}$$

So combining (4.8), (4.9), (4.12)-(4.14), (4.15) we have

$$\begin{aligned}\delta(T) &= \delta(T') = \delta((T')^{-1}) = \nu((T')^{-1}) \\ &= \mathrm{degree}(\det(T')) = \mathrm{degree}(\det(T_{11})) \\ &= \delta'(T)\end{aligned}$$

which concludes the proof of lemma 4.10.

4.17. THEOREM. (upper continuity of $\delta(T)$). *Let* $T_c(s)$ *be a sequence of matrices of rational functions of* s. *Suppose that the sequence converges to matrix of rational functions*

$T(s)$ *as* $c \to \infty$ *and suppose that* $\delta(T_c(s)) \le n$ *for all large enough* c. *Then* $\delta(T) \le n$.

Here a sequence of matrices of rational functions is said to converge iff the sequences of entries converge in the sense of section 3 above; i.e. $T_c(s)$ converges as $c \to \infty$ iff $\lim_{c\to\infty} T_c(s)$ exists for infinitely many s and then the limit is necessarily a matrix of rational functions $T(s)$ and $\lim_{c\to\infty} T_c(s) = T(s)$ for all but finitely many s.

The proof of the theorem is easy. We have for each $m \times n$ matrix of constants K that

$$\lim_{c\to\infty} \det(I_m + KT_c(s)) = \det(I_m + KT(s))$$

Hence using proposition 3.1 (which among other things contains the scalar case of theorem (4.16)), or rather using corollary 3.10, and using the second definition of the degree of a rational matrix discussed above (cf. (4.4)-(4.6), we have for large enough c (which may depend on K)

$$\delta_K(T) = \text{degree}(a_K(s)) \le \text{degree}(a_{K,c}(s)) = \delta_K(T_c) \le n$$

where

$$\frac{a_K(s)}{b_K(s)} = \det(I_m + KT(s)), \quad \frac{a_{K,c}(s)}{b_{K,c}(s)} = \det(I_m + KT_c(s))$$

(without common factors). It follows that $\delta(T) = \max_K\{\delta_K(T)\} \le n$.

It is now not difficult to prove Theorem 2.23 without the extra requirement that the poles of $L_c(s)$ unequal to the finite poles of $L(s)$ go to $-\infty$ as $c \to \infty$. Indeed the upper semicontinuity property of theorem 4.17 takes care of the "if" part. So let $L(s)$ be of degree n. Write $L(s) = A(s) + T(s)$, where $T(s)$ is proper and $A(s)$ is polynomial. Then $\delta(L) = \delta(T) + \delta(A)$. So if $A(s) = \lim_{n\to\infty} T_n(s)$, with $T_n(s)$ proper and $\delta(T_n(s)) \le \delta(A(s))$ we will be done.

4.18. PROPOSITION. *Let* $A(s)$ *be a polynomial matrix of degree* δ. *Then there exist a sequence of proper rational matrices* $T_n(s)$ *of degree* $\leq \delta$ *such that* $\lim_{n\to\infty} T_n(s) = A(s)$.

Proof. By multiplying $A(s)$ on the left and on the right with suitable invertible matrices we can assume that A is of the form

$$A = \begin{pmatrix} A_{11} & A_{12} \\ & \\ A_{21} & A_{22} \end{pmatrix}$$

with $\deg(\det(A_{11})) = \delta$. As above let

$$A' = \begin{pmatrix} A_{11} & A_{12} & 0 \\ A_{21} & A_{22} & I \\ 0 & I & 0 \end{pmatrix}$$

Then $\delta = \delta(A') = \text{degree} \det(A')$. Now let

$$T_n'(s) = nA'(nI + A')^{-1}$$

(Note that $(nI + A'(s))^{-1}$ exists if we assume, as we can, that $\delta > 0$). Then clearly for a fixed s, $\lim_{n\to\infty} T_n'(s) = A'(s)$. We claim that $T_n'(s)$ is proper for all but finitely many n. Indeed for a fixed n

$$\begin{aligned} T_n' = nA'(nI + A')^{-1} &= nA'(A')^{-1}(nA'^{-1} + I)^{-1} \\ &= ((A')^{-1} + n^{-1}I)^{-1} . \end{aligned} \tag{4.19}$$

Now because $\delta(A') = \deg(\det(A'))$ we know that $(A')^{-1}$ is proper. Let $J = \lim_{s\to\infty} (A')^{-1}$. Then if $-n^{-1}$ is not an eigenvalue of J it follows from (4.19) that $\lim_{s\to\infty} T_n'(s)$ exists, proving that $T_n'(s)$ is proper for all but finitely many n.

Finally, by lemma (4.15), if $T_n'(s)$ is proper,

$$\nu(T_n'(s)) \leq \deg(\det(nI + A')) \tag{4.20}$$

Now $\det(nI + A')$ is a polynomial in s whose coefficients are sums of minors of A'. Hence $\deg(\det(nI + A')) \le \max_M \deg(M) = \delta(A') = \delta$ where M runs through the minors of A'.

Now let $T_n(s)$ be obtained from $T_n'(s)$ by removing the appropriate columns and rows. Then $\lim_{n\to\infty} T_n(s) = A(s)$, $T_n(s)$ is proper if $T_n'(s)$ is proper and $\delta(T_n) \le \delta(T_n')$ proving proposition 4.18.

5. PROOF OF THE MAIN THEOREM.

5.1. First Half of the Proof of Theorem 2.22. Let $\Sigma(c) \subset L$ be a family of systems of dimension n and suppose they converge in input/output behaviour. This means (cf. 2.10) that for every $u \in U$ the sequence of functions

$$(V_{\Sigma(c)}u) \subset \mathcal{Y} \tag{5.2}$$

converges. In turn this means (as in the proof of lemma 2.6) that there is a b such that for all sufficiently large c

$$V_{\Sigma(c)}u \in \mathcal{F}_b^{(o)}(\mathbb{R}^p) \tag{5.3}$$

If $z \in \mathcal{F}_b^{(o)}(\mathbb{R}^p)$, then $\sup_t \|e^{-bt}z(t)\| < \infty$ so that

$$\int_0^\infty \|e^{-(b+1)t}z(t)\|dt < \infty$$

which implies (cf. [Doe] or [Zem]) that $z(t)$ is Laplace transformable and that $(\pounds z)(s)$ is defined for $\mathrm{Re}(s) \ge b+1$.

Applying this to the $V_{\Sigma(c)}u$ we see that their Laplace transforms are well defined for $s \ge b+1$. This gives us a sequence of functions

$$Y_c(s) = T_c(s)U(s) \tag{5.4}$$

where $Y_c(s)$ is the Laplace transform of $V_{\Sigma(c)}u$, $T_c(s)$ is the transfer function of $\Sigma(c)$ and $U(s)$ is the Laplace transform of $u(t)$.

The Laplace transform $\pounds$ is continuous when considered as an operator on the normed space $\mathcal{F}_{b+1}'(\mathbb{R}^p)$ consisting of all locally integrable functions such that

$$\int_0^\infty \|e^{-(b+1)t} z(t)\| dt < \infty \tag{5.5}$$

equipped with the norm defined by the integral (5.5), cf. [Doe, Kap. III, §8]. As $\mathscr{F}_b^{(o)}(\mathbb{R}^p) \subset \mathscr{F}'_{b+1}(\mathbb{R}^p)$ is a continuous embedding it follows that the sequence (5.4) converges for $\operatorname{Re}(s) \geq b+1$ as $c \to \infty$. Choosing various $u \in U$ judiciously this implies that the family of rational matrix functions $T_c(s)$ converges for infintiely many values of s. According to section 4 above this means that there is a rational matrix function $T(s)$ such that

$$\lim_{c\to\infty} T_c(s) = T(s) \tag{5.6}$$

and moreover $\delta(T) \leq n$ by the upper semicontinuity theorem 4.17. Write

$$T(s) = T'(s) + L(s) \tag{5.7}$$

where $T'(s)$ is proper and where $L(s)$ is polynomial. Let Σ be a co and cr realization of $T'(s)$. Consider the operator

$$V = V_\Sigma + L(D) \tag{5.8}$$

Applying this operator to a $u \in U$ and taking the Laplace transform of the result (which can be done because $Vu \in \mathscr{Y}$ and all functions in $\mathscr{Y}$ are Laplace transformable) we find (for $\operatorname{Re}(s) \geq b'+1$, for some $b' \geq b$)

$$(\pounds Vu)(s) = T'(s)U(s) + L(s)U(s) = \lim_{c\to\infty} T_c(s)U(s)$$
$$= \lim_{c\to\infty} Y_c(s) = (\pounds(\lim_{c\to\infty} y_c))(s)$$

where $y_c = V_{\Sigma(c)}u$, and where we have again used the same continuity property of the Laplace transform. The Laplace transform begin injective on the space of functions under consideration it follows that

$$Vu = \lim_{c\to\infty} V_{\Sigma(c)}u$$

for all $u \in U$. Thus the limit operator is indeed of the form $V = V_\Sigma + L(d)$ with $\dim(\Sigma) + \operatorname{degree} L(s) = \delta(T) \leq n$, which finishes the proof of the first half of theorem 2.22.

To prove the second half we need some lemmas. If A is any matrix we use the following notation for its various minors:

$$A \begin{matrix} i_1,\dots,i_r \\ j_1,\dots,j_r \end{matrix}$$

denotes the determinant of the submatrix of A obtained by removing all rows except those with the indices $i_1,\dots,i_r$ and all columns except those with the indices $j_1,\dots,j_r$. Recall that the minors of a product matrix are given by

$$(AB) \begin{matrix} i_1,\dots,i_r \\ j_1,\dots,j_r \end{matrix} = \sum_{k_1,\dots,k_r} A \begin{matrix} i_1,\dots,i_r \\ k_1,\dots,k_r \end{matrix} \; B \begin{matrix} k_1,\dots,k_r \\ j_1,\dots,j_r \end{matrix} \qquad (5.9)$$

5.10. LEMMA. *Let* $L(s)$ *be a polynomial matrix of size* $p \times m$. *Suppose that for a certain* $1 \le r \le \min(p,m)$

$$\deg L(s) \begin{matrix} 1,\dots,r \\ 1,\dots,r \end{matrix} \ge \deg L(s) \begin{matrix} 2,\dots,r,j \\ 1,\dots,r \end{matrix} \;, \; j = r+1,\dots,p \qquad (5.11)$$

Then there exists an invertible $p \times p$ *matrix of constants* A *such that*

$$\deg (AL(s)) \begin{matrix} 1,\dots,r \\ 1,\dots,r \end{matrix} > \deg (AL(s)) \begin{matrix} 2,\dots,r,j \\ 1,\dots,r \end{matrix} \;, \; j = r+1,\dots,p \qquad (5.12)$$

Proof. Let $E_j(c) = E$, $j \in \{r+1,\dots,p\}$ be the matrix with 1's on the diagonal, a c in spot (j,1) and zero's elsewhere. Then as is easily checked

$$E \begin{matrix} 1,\dots,r \\ i_1,\dots,i_r \end{matrix} = \begin{cases} 1 & \text{if } \{i_1,\dots,i_r\} = \{1,\dots,r\} \\ 0 & \text{otherwise} \end{cases}$$

and for $k \ne j$, $k \in \{r+1,\dots,p\}$

$$E \begin{matrix} 2,\dots,r,k \\ i_1,\dots,i_r \end{matrix} = \begin{cases} 1 & \text{if } \{i_1,\dots,i_r\} = \{2,\dots,r,k\} \\ 0 & \text{otherwise} \end{cases}$$

while

$$E^{2,\dots,r,j}_{i_1,\dots,i_r} = \begin{cases} (-1)^r c & \text{if } \{i_1,\dots,i_r\} = \{1,\dots,r\} \\ 1 & \text{if } \{i_1,\dots,i_r\} = \{2,\dots,r,j\} \\ 0 & \text{otherwise} \end{cases}$$

It now follows from the minor product rule (5.9) that

$$\text{(EL)} \quad {}^{2,\dots,r,k}_{1,\dots,r} = \begin{cases} L^{1,\dots,r}_{1,\dots,r} & \text{if } k = 1 \\ L^{2,\dots,r,k}_{1,\dots,r} & \text{if } k \in \{r+1,\dots,p\}\backslash\{j\} \\ L^{2,\dots,r,j}_{1,\dots,r} + (-1)^r c \; L^{1,\dots,r}_{1,\dots,r} & \text{if } k = j \end{cases}$$

It follows that (5.12) holds if we take for A a suitable product of matrices $E_j(c)$.

5.13. LEMMA. *Let* $L(s)$ *be a polynomial* $p \times m$ *matrix without constant terms of degree* n. *Suppose that for a certain* r *all minors of size* $< r$ *have degree* $< n$ *and that*

$$\deg(L^{1,\dots,r}_{1,\dots,r}) = n > \deg(L^{2,\dots,r,j}_{1,\dots,r}), \; j = r+1,\dots,p \qquad (5.14)$$

Let $d(s)$ *be the diagonal matrix with diagonal entries* $(s,1,\dots,1)$ *and let* $L'(s) = d(s)^{-1}L(s)$. *Then* $L'(s)$ *is polynomial (because the first row of* $L(s)$ *has no constant terms) and* $\deg(L'(s)) = n-1$.

Proof. Because $\deg(d(s)) = 1$ and $\deg L(s) \le \deg(d(s)) + \deg(L'(s))$ we must have $\deg(L'(s)) \ge n-1$. It remains to show that $\deg(L'(s)) \le n-1$. Let $\bar{L}(s)$ be the square matrix

$$\bar{L} = \begin{pmatrix} L_{11} & L_{12} & 0 \\ L_{21} & L_{22} & I \\ 0 & I & 0 \end{pmatrix}$$

where

$$L = \begin{pmatrix} L_{11} & L_{12} \\ L_{21} & L_{22} \end{pmatrix} ,$$

where L_{11} is the top-left $r \times r$ submatrix of $\bar{L}$, and where the I's are the appropriate unit matrices. Then

$$\deg(L) = \deg(\bar{L}) = \deg(\det(L_{11}) = \deg(\det(\bar{L})) = n \qquad (5.15)$$

which implies that $\bar{L}^{-1}$ is proper. We claim that the first column of $\bar{L}^{-1}$ consists of strictly proper rational functions. Indeed the entries of the first column are the functions

$$\det(\bar{L})^{-1}\bar{L}_j^1, \quad j = 1,\dots,m+p-r \qquad (5.16)$$

Now, if $j = 1,\dots,r$, $\bar{L}_j^1$ is the determinant of a $(r-1) \times (r-1)$ submatrix of L_{11} and hence $\deg(\bar{L}_j^1) < n$ by hypothesis. If $j = r+1,\dots,m$ then $\bar{L}_j^1 = 0$ and finally if $j = m+k$, $k = 1,\dots,p-r$ then

$$\bar{L}_j^1 = L\begin{matrix} 2,\dots,r,r+k \\ 1,\dots,r \end{matrix} , \quad j = m+k$$

which by hypothesis is of degree $< n = \deg(\det(L)^{-1}$. This proves the claim.

Now let $d'(s)$ be the $(m+p-r) \times (m+p-r)$ diagonal matrix with entries $(s,1,\dots,1)$, and let $\bar{L}' = d'(s)^{-1}\bar{L}$. Then L' is the $p \times m$ top left submatrix of $\bar{L}'$ and hence

$$\text{degree}(L') \le \text{degree}(\bar{L}') \qquad (5.17)$$

On the other hand $(\bar{L}')^{-1} = (\bar{L})^{-1}d'(s)$ is still proper because the first column of $\bar{L}^{-1}$ consists of strictly proper rational functions. Hence (cf. lemma 4.15)

$$\begin{aligned} \deg(\bar{L}') &= \deg((\bar{L}')^{-1}) \le \deg(\det(\bar{L}')) \\ &= \deg(\det(d'(s))^{-1}\det(\bar{L})) \\ &= \deg(s^{-1}\det(L_{11})) = n-1 \end{aligned} \qquad (5.18)$$

because L_{11} has no constants. Combining (5.18) and (5.17) we see that indeed $\deg(L') \le n-1$, proving the lemma. (NB it is not true as a rule that $(L')^{-1}$ is proper.)

Note that lemma 5.13 and 5.10 combine to give a proof of corollary 2.24.

5.19. PROPOSITION. *Let* $L(s)$ *be a polynomial matrix of degree* n. *Then there exists a family of* n*-dimensional systems* $\Sigma(c)$ *such that* $\Sigma(c)$ *converges in input/output behaviour to* $L(D) : \mathcal{U} \to \mathcal{Y}$ *as* $c \to \infty$ *and such that moreover the poles of (the transfer functions of) the* $\Sigma(c)$ *all go to* $-\infty$ *as* $c \to \infty$.

Proof. This is proved by induction, the case $n = 0$ being trivial because $L(s)$ has degree zero iff it is a matrix of constants. The first thing to do next is to obtain the scalar operator $D : \mathcal{F}(\mathbb{R}) \to \mathcal{F}(\mathbb{R})$ as a limit of input/output operators of one dimensional systems. To this end let $\Sigma(c)$, $c = 1,2,\ldots$ (or $c \in \mathbb{R}$) be the family of systems

$$\Sigma(c) = (F_c, G_c, H_c, J_c), \quad J_c = c, \qquad F_c = -c, \quad H_c = c, \quad G_c = -c \tag{5.20}$$

The associated input/output operator of $\Sigma(c)$ is $V_c : \mathcal{F}(\mathbb{R}) \to \mathcal{F}(\mathbb{R})$

$$V_c : u(t) \mapsto y_c(t) = cu(t) + \int_0^t -c^2 e^{-c(t-\tau)} u(\tau)d\tau \tag{5.21}$$

By partial integration (twice) we see that

$$y_c(t) = u^{(1)}(t) - \int_0^t e^{-c(t-\tau)} u^{(2)}(\tau)d\tau \tag{5.22}$$

Let b be such that $u^{(2)} \in \mathcal{F}_b^{(o)}(\mathbb{R})$ (i.e. $\sup_t e^{-bt}|u^{(2)}(t)| < \infty$). Then if $M = \|u^{(2)}\|_b$, we have

$$\left|\int_0^t e^{-c(t-\tau)} u^{(2)}(\tau)d\tau\right| \le \int_0^t e^{-c(t-\tau)} e^{b\tau} M \le (b+c)^{-1} M e^{bt} \tag{5.23}$$

and it follows that the $y_c(t)$ converge to $u^{(1)}(t)$ in $\mathcal{F}(\mathbb{R})$.

More precisely if b is such $u^{(1)}$, $u^{(2)}$ are both in $\mathscr{F}_b^{(o)}(\mathbb{R})$ then $y_c(t) \in \mathscr{F}_b^{(o)}(\mathbb{R})$ and the $y_c(t)$ converge to $u^{(1)}(t)$ in $\mathscr{F}_b^{(o)}(\mathbb{R})$.

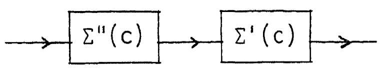
Now suppose with induction that the proposition has been proved for all polynomial matrices of degree $\leq n-1$.

Let $L(s)$ be a polynomial matrix of degree n. First note that if P,Q are invertible matrices of constants then $L(D)$ is the limit of a family as in the statement of the theorem if and only if $PL(D)Q$ is. Also adding a matrix of constants makes no difference. Removing the constants and multiplying $L(s)$ on the left and on the right with suitable invertible matrices of constants we can therefore assume that for a certain minimal $r \in \mathbb{N}$ the top left $r \times r$ minor of $L(s)$ is of degree n. Using lemma 5.10 and lemma 5.13 we see that after a further multiplication on the left by an invertible matrix of constants $L(s)$ factorizes as

$$L(s) = \begin{pmatrix} s & & & 0 \\ & 1 & & \\ & & \ddots & \\ 0 & & & 1 \end{pmatrix} L'(s)$$

with $L'(s)$ polynomial of degree $n-1$. By induction we have that there exists a family of $(n-1)$-dimensional systems $\Sigma'(c) = (F_c', G_c', H_c', J_c')$ such that the poles of $\Sigma'(c)$ go to $-\infty$ as $c \to \infty$ (if $n-1 > 0$, if $n=1, L'(s)$ is constant and one takes $\Sigma'(c) = (0,0,0,L')$) and such that $V_{\Sigma'(c)}$ converges in input/output behaviour to $L'(s)$.

Now let $\Sigma(c)$ be the composed system



(5.24)

where $\Sigma''(c)$ is the m input/m output one dimensional system given by the matrices

$$F''_c = -c, \quad G''_c = (-c,0,\dots,0),$$

$$H''_c = \begin{pmatrix} c \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad J''_c = \begin{pmatrix} c & 0 & \dots & 0 \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{pmatrix}$$

I.e. if $\Sigma'(c) = (F'_c, G'_c, H'_c, J'_c)$ then $\Sigma(c)$ is given by the

$$F_c = \begin{pmatrix} F''_c & 0 \\ G'_c H''_c & F'_c \end{pmatrix}, \quad G_c = \begin{pmatrix} G''_c \\ G'_c J''_c \end{pmatrix},$$

$$H_c = \begin{pmatrix} J'_c H''_c & H'_c \end{pmatrix}, \quad J_c = J'_c J''_c \tag{5.25}$$

(if $n > 1$; if $n = 1$, $F_c = -c$, $G_c = (-c,0,\dots,0)$, $H_c = L'H''_c$, $J_c = L'H''_c$). Then the $\Sigma(c)$ converge in input/output behaviour to $L(D)$. Moreover (as follows from (5.25)) the poles of $\Sigma(c)$ go to $-\infty$ as $c \to \infty$ if $n > 1$. This proves the proposition.

We can be somewhat more precise about how well the $\Sigma(c)$ converge in input/output behaviour to $L(D)$. Indeed one has

5.26. COROLLARY. *Let* $L(D)$ *and* $(\Sigma(c))_c$ *be as above in the proof of proposition* 5.19. *Let* $b \geq 0$ *be such that* $u, u^{(1)}, \dots, u^{(n+1)} \in \mathscr{F}_b^{(o)}(\mathbb{R}^m)$. *Then there is a constant* M *such that*

$$\|V_{\Sigma(c)}u - L(D)u\| \leq c^{-1}Me^{bt} \tag{5.27}$$

In particular if $u \in \mathscr{U}$ *is of compact support or, more generally if* $u, u^{(1)}, \dots, u^{(n+1)}$ *are all bounded, we can take* $b = 0$ *and for such input functions* u, $V_{\Sigma(c)}u$ *converges uniformly in* t *to* $L(D)u$.

This follows readily by induction from the proof of proposition 5.19 above, (5.22), and the estimate (5.23), because $L'(D)u$ is a vector of linear combination of the $u, u^{(1)}, \dots, u^{(n-1)}$.

5.28. Proof of the Second Half of Theorem 2.22. Now let $V : \mathcal{U} \to \mathcal{Y}$ be an operator of the form $V = L(D) + V_\Sigma$ with $\dim(\Sigma) + \deg(L(s)) \le n$. Let $\Sigma(c)$ be a sequence of $\deg(L(s))$-dimensional systems converging to $L(D)$ in input/output behaviour as in proposition 5.19. Then if $\Sigma'(c)$ is the sum system of $\Sigma(c)$ and Σ, the family $\Sigma'(c)$ converges in input/output behaviour to V. More precisely if $\Sigma = (F,G,H,J)$, $\Sigma(c) = (F_c,G_c,H_c,J_c)$ then $\Sigma'(c)$ is given by the matrices

$$F'_c = \begin{pmatrix} F & 0 \\ 0 & F_c \end{pmatrix}, \quad G'_c = \begin{pmatrix} C \\ G_c \end{pmatrix}, \quad H'_c = \begin{pmatrix} H & H_c \end{pmatrix}, \quad J'_c = J + J_c$$

Because the co and cr systems are open and dense in L we can perturb each $\Sigma'(c)$ slightly to a $\Sigma''(c)$ which is co,cr such that $\Sigma''(c)$ still converges to V in input/output behaviour as $c \to \infty$, and such that the behaviour of the poles of the $\Sigma''(c)$ as $c \to \infty$ is like that of the $\Sigma'(c)$ as $c \to \infty$. This finishes the proof of theorem 2.22.

5.29. REMARK. One has of course in the setting of 5.28 abov also an estimate like (5.27) for $\|V_{\Sigma'(c)}u - Vu\|$.

5.30. REMARK. If $\Sigma(c)$ is e.g. the family of (5.20) above, the Markov parameters of the family $J_c, H_cG_c, H_cF_cG_c, H_cF_c^2G_c, \ldots$ definitely do not converge as $c \to \infty$.

One can, of course, examine what the possible limits are of families of systems $\Sigma(c)$ of dimension n which converge in input/output operators and such that moreover the Markov parameters converge as well (or more generally such that the Markov parameters remain bounded) as $c \to \infty$. The answer is simple: tł limit operator is then necessarily of the form V_Σ where Σ iś a possibly lower dimensional system. Inversely every V_Σ with $\dim(\Sigma) \le n$ can arise a limit of input/output operators of co and cr systems of dimension n, cf. [Haz 2].

5.31. Approximation by systems with $J = 0$. Let $T(s)$ be ӓ matrix of rational functions. Write

$$T(s) = T_-(s) + L(s) \qquad (5.3ȝ$$

with $T_-(s)$ strictly proper and $L(s)$ polynomial. Define

$$\begin{aligned} n_r(T) &= \text{dim of the } \mathbb{R}\text{-vectorspace spanned by the rows of } L(s) \\ n_c(T) &= \text{dim of the } \mathbb{R}\text{-vectorspace spanned by the columns of } L(s) \\ q(T) &= \min\{n_r(T), n_c(T)\} . \end{aligned} \tag{5.33}$$

E.g. if $T(s) = L(s) = \begin{pmatrix} s^2 & s & s^3 \\ 1 & s & 1 \end{pmatrix}$, then $n_r(T) = 2$, $n_c(T) = 3$ and if $T(s) = L(s) = \begin{pmatrix} s^2 & s^2 \\ s & s \end{pmatrix}$, $n_r(T) = 2$, $n_c(T) = 1$. Let Σ realize $T_-(s)$. Then the operator $V_\Sigma + L(D)$ is the limit in input/output behaviour of a family of $(\deg(T(s)) + q(T(s))$ - dimensional systems.

This can be seen as follows. Because $T_-(s)$ is strictly proper it suffices to see that $L(D)$ can be obtained as the limit of the input/output operators of a family of $\deg(L(s)) + q(L(s))$ dimensional systems. Assume for definitiveness that $q(T) = n_c(T)$. Then we can factorize $L(s)$ as

$$L(s) = (L'(s) \quad 0)Q$$

where Q is a square invertible matrix of constants and $L'(s)$ has $q(T)$ columns. It now clearly suffices to obtain $L'(D)$ as a limit of $\deg(L) + q(L)$ dimensional systems. To this end let $\Sigma(c)$ be a family of systems converging to $L(D)$ of dimension $\deg(L)$ and let $\Sigma'(c)$ be a $q = q(L)$-dimensional family of systems with $J_c = 0$ for all c with limit input/output operator equal to I, the $q \times q$ identity matrix. Such a family is e.g. given by the matrices

$$F_c = \begin{pmatrix} -c & 0 \\ 0 & -c \end{pmatrix}, \quad G_c = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad H_c = \begin{pmatrix} c & 0 \\ 0 & c \end{pmatrix}, \quad J_c = 0 .$$

Let $T_c'(s)$ be the transfer function matrix of $\Sigma'(c)$ and $T_c(s)$

that of $\Sigma(c)$. Then the $(q+\deg(L))$-dimensional system $\Sigma''(c)$ obtained by applying first $\Sigma'(c)$ and then $\Sigma(c)$ has transfer function matrix $T_c(s)T_c'(s)$, which is strictly proper, and the $\Sigma''(c)$ converge in input/output behaviour to $L'(s)$.

This result is optimal if $p=1$ or $m=1$, but, though definitely generically best possible (meaning that for almost all $T(s)$ with given $q(t) = q$, $\deg(T) + q$ is the best one can do), it is not best possible for every particular $T(s)$. E.g. the factorization

$$L(s) = \begin{pmatrix} s^2 & s^3 \\ s & s^2 \end{pmatrix} = \begin{pmatrix} s & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} s & s^2 \\ 0 & 0 \end{pmatrix}$$

shows that this $L(s)$ can be obtained as the input/output limit of a family of four dimensional systems with $J = 0$, although $\deg(L) = 3$ and $q(L) = 2$.

REFERENCES

[AV] Anderson, B. D. O., and S. Vongpanitlerd, *Network Analysis and Synthesis*, Prentice Hall, 1973.

[BK] Brockett, R. W., and P. S. Krishnaprasad, *A scaling theory for linear systems*, to appear Trans. IEEE AC.

[BM1] Basile, G., and G. Marro, *On the synthesis of unknown-input observers and inverse systems by recursive algorithms*, Report AM-68-8, UC at Berkeley, College of Engineering, 1968.

[BM2] Basile, and G. Marro, *A new characterization of some structural properties of linear systems: unknown-input observability, invertability and functional controllability*, Int. J. Control 17 (1973), 931-943.

[Doe] Doetsch, G., *Theorie und Anwendung der Laplace Transformation*, Springer, 1937 (Dover reprint, 1943).

[Haz1] Hazewinkel, M., *Degenerating families of linear dynamical systems I*, Proc. 1977 IEEE Conf. on Decision and Control (New Orleans, Dec. 1977), 258-264.

[Haz2] Hazewinkel, M., *On identification and the geometry of the space of linear systems*, Proc. Bonn Jan. 1979 Workshop on Stochastic Control and Stochastic Differential Equations, Lect. Notes Control and Information Sciences, 16, Springer, 1979, 401-415.

[Haz3] Hazewinkel, M., *(Fine) moduli spaces for linear systems. What are they and what are they good for?* in Proc. NATO-AMS Summer Inst. on Algebraic and Geometric Methods in Linear Systems Theory, Reidel Publishing Co., to appear.

[Haz4] Hazewinkel, M., *Moduli and canonical forms for linear dynamical systems. II: the topological case,* Math. Systems Theory 10 (1977), 363-385.

[HP] Hazewinkel, M., and A. -M. Perdon, *On the theory of families of linear dynamical systems,* Proc. MTNS 4 (Delft: July 1979), 155-161.

[Kal] Kalman, R. E., *Irreducible realizations and the degree of a rational matrix,* J. Soc. Industrial and Appl. Math. 13, 2(1965), 520-544.

[MH] Martin, C., and R. Hermann, *Applications of algebraic geometry to systems theory: the McMillan degree and Kronecker indices of transfer functions as topological and holomorphic system invariants,* SIAM J. Control and Optimization 16 (1978), 743-755.

[Ros1] Rosenbrock, H. H., *State Space and Multivariable Theory,* Nelson, 1970.

[Ros2] Rosenbrock, H. H., *Systems and polynomial matrices,* Proc. NATO-AMS Summer Inst. on Algebraic and Geometric Methods in Linear Systems Theory (Harvard Univ., June 1979), Reidel Publishing Company, to appear.

[Wil1] Willems, J. C., *Almost A(modB)-invariant subspaces,* to appear in Astérisque (Soc. Math. de France; Journées sur l'analyse des systèmes, Bordeaux, 1978).

[Wil2] Willems, J. C., *Almost A(modB)-invariant subspaces: an approach to high-gain feedback design,* in preparation.

[Won] Wonham, W. M., *Linear multivariable control,* Lect. Notes Economics and Math. System Theory 101, 1974, Springer.

[YKU] Young, K. -K. D., P. V. Kokotovic, and V. I. Utkin, *A singular perturbation analysis of high-gain feedback systems,* IEEE Trans. AC 22 (1977), 931-938.

[Zem] Zemanian, A. H., *Distribution Theory and Transform Analysis,* McGraw-Hill, 1965.

ERASMUM UNIVERSITY OF ROTTERDAM
P.O. BOX 1738
3000 DR ROTTERDAM
THE NETHERLANDS