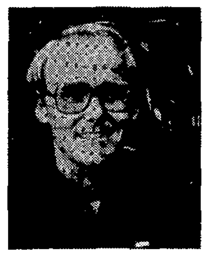
[8] D. E. Knuth, "Matroid partitioning," Report, Stan-CS-73-342, Stanford Univ., 1973.
[9] E. S. Kuh and R. A. Rohrer, "The state variable approach to network analysis," *Proc. IEEE*, vol. 53, pp. 672–686, July 1965.
[10] E. L. Lawler, "Matroid intersection algorithms," *Math. Progr.*, vol. 9, pp. 31–56, 1975.
[11] S. K. Mark and M. N. S. Swamy, "The generalized tree for state variables in linear active networks," *Circ. Th. and Appl.*, vol. 4, pp. 87–92, 1976.
[12] W. Mayeda, *Graph Theory*. New York: Wiley, 1972.
[13] G. J. Minty, "On the axiomatic foundations of the theories of directed linear graphs, electrical networks and network-programming," *J. Math. Mech.*, vol. 15, no. 3, 1966.
[14] T. Ozawa, "Topological conditions for the solvability of active linear networks," *Circuit Theory and Appl.*, vol. 4, pp. 125–136, 1976.
[15] B. Petersen, "Electrical networks and matroids," Ph.D. dissertation, Technical University of Denmark, Lyngby, 1976.
[16] ——, "Circuits in the union of matroids. An algorithmic approach," 1978 (to be submitted).
[17] E. J. Purslow, "Solvability and analysis of linear active networks by use of the state equations," *IEEE Trans. Circuit Theory*, vol. CT-17, Nov. 1970.
[18] A. Recski, "On partitional matroids with applications," *Coll. Math. Soc. J. Bolyai*, vol. 10, Infinite and finite sets, Keszthely (Hungary), 1973, North-Holland-American Elsevier, Amsterdam, New York, 1974, pp. 1169–1179.
[19] ——, "Matroids and independent state variables," presented at *Proc. Second European Conf. Circuit Theory and Design*, Genova, 1976.
[20] S. Seshu and M. B. Reed, *Linear Graphs and Electrical Networks*. Reading, MA: Addison-Wesley, 1961.
[21] J. Tow, "Order of complexity of linear active networks," *Proc. IEE*, vol. 115, pp. 1259–1262, Sept. 1968.
[22] R. J. Wilson, "An introduction to matroid theory," *Amer. Math. Monthly*, vol. 80 no. 5, pp. 500–525, May 1973.



Bjørn Petersen was born in Denmark, in 1948. He received the M.Sc. degree in electrical engineering from the Technical University of Denmark, Lyngby, Denmark, in 1976, and is currently working towards the Ph.D. degree at the Institute of Circuit Theory and Telecommunication, and the Department of Mathematics, Technical University of Denmark.

Until February 1977, he was an Associate Professor at the Department of Mathematics, University of Denmark, Lyngby. He has carried out research in the fields of circuit theory, graph theory, matroid theory, and computational complexity. His fields of interest are linear active network theory and matroid algorithms.

Realization Theory of Discrete-Time Nonlinear Systems: Part I — The Bounded Case

EDUARDO D. SONTAG

Abstract—**A state-space realization theory is presented for a wide class of discrete time input/output behaviors. Although in many ways restricted, this class does include as particular cases those treated in the literature (linear, multilinear, internally bilinear, homogeneous), as well as certain nonanalytic nonlinearities. The theory is conceptually simple, and matrix-theoretic algorithms are straightforward. Finite-realizability of these behaviors by state-affine systems is shown to be equivalent both to the existence of high-order input/output equations and to realizability by more general types of systems.**

Introduction

THIS WORK deals with some aspects of realization theory of deterministic nonlinear discrete-time systems. The realization theory of *linear* systems is by now a successful part of system theory, which has resulted in a deep understanding of behavior and has permitted the application of state-space methods of analysis and synthesis. It may be reasonable to expect, then, that a corresponding theory will eventually derive analogous benefits for nonlinear systems.

For the most part, this paper presents a "linearized" realization theory via systems which are linear in state variables but arbitrarily nonlinear in inputs, *state-affine* systems. While such systems are highly restrictive vis à vis general nonlinear models, they do include those for which a detailed realization theory has been developed, in particular, linear, internally bilinear, and multilinear systems. The importance of S-A representations in the analysis of certain nuclear reactors, heat-transfer processes, and population models, among others, has been made explicit by various authors (see, for instance, [34]); other applications being currently explored are in the areas of image processing and in stochastic filtering. Moreover, in some

Manuscript received September 11, 1977; revised July 6, 1978. This research was supported in part by U. S. Army Research Grant DAH C04-74-G-0153 through the Center for Mathematical System Theory, Gainesville, FL 32611.
The author is with the Department of Mathematics, Rutgers University, New Brunswick, NJ 08903.

0098-4094/79/0500-0342$00.75

cases the canonical realization of a given input/output behavior admits a S-A structure; this is the case with *bounded polynomial responses,* those whose output values at any given instant are arbitrary sums and products of previous input values, subject only to the restriction that there is a *bound*—hence the name—to the exponents to which each single input can be raised in calculating outputs.

Bounded responses were originally defined in a polynomial context, but the present work treats directly a more general case, which has the advantage of including many types of nonanalytic nonlinearitites (piecewise polynomial, in particular).

The first part of this paper deals with an abstract realization theory, while the second presents a concrete matrix realization algorithm which both generalizes and unifies those known in the literature for the various classes of systems. These two parts result in particular in a self-contained realization methodology for S-A systems, and serve also as an introduction to a more general (and strictly nonlinear) realization theory. Part three provides further finiteness results, including a generalization of the linear system fact that finite realizability is equivalent to the existence of a (high-order) input/output difference equation, and studying the relationship between state-affine realizability and more general realizability of a bounded response. The paper closes with some remarks in part four.

I. Abstract Realization Theory

This section develops a realization theory which will give the theoretical basis for the algorithms presented later. The following notational conventions will hold throughout this work.

For any set S and integer $t \geqslant 0$, S^t will be the set of all sequences $w=(w_1,\cdots,w_t)$ of length t, w_i in S; note $S^0=$ set containing only the empty sequence e. The set S^* (resp. S^+) is the union of all S^t, $t \geqslant 0$ (resp. $t \geqslant 1$); $|w|$ denotes the length of w. For notational convenience, a $w=(w_1,\cdots,s_w)$ will be also denoted as $w_1 w_2 \cdots w_t$. (This should not be confused with the product of the w_i, when such a product would be also defined.) The expression a^t will, correspondingly, mean $a \cdots a$ (t times). The *concatenation* of two sequences $v=v_1 \cdots v_s$ and $w=w_1 \cdots w_t$ is $vw:=v_1 \cdots v_s w_1 \cdots w_t$ of length $s+t$. For any function f defined on S^+, the restriction of f to sequences in S^t is f_t, while the maps $v \mapsto f(wv)$ (w in U^+) are denoted by f_w, and the maps $v \mapsto f(vw)$ are denoted by f^w.

An arbitrary field k will be fixed throughout the discussion; "vector space" will mean k-vector space, "linear" will mean k-linear, etc. (Some results will be stated only for $k=$reals or complexes, but most are valid without any restrictions on k.) Recall that *affine manifold*$=$translate of a subspace, and *affine map*$=$linear$+$translation. Unless otherwise stated, U will denote an arbitrary set (of input values) and Y an arbitrary vector space (of output values). When U is a subset of a vector space k^m, the notation $u_{(i)}$ for the vector u in U will mean the ith coordinate of u—the notation u_i being reserved for the ith element of a sequence of vectors. If A,B are sets, $[A,B]$ will denote the set of all maps $f: A \to B$. When B is a vector space, $[A,B]$ will denote the corresponding vector space, with the pointwise operations: $(f+rg)(a):=f(a)+rg(a)$.

A. Response Maps

An "input/output map" sends input sequences into output sequences. When such a map is causal, it can be equivalently specified by its associated "response map," which describes how past inputs affect present outputs; this latter object is defined directly.

Definition 1.1: A *response* is a map $f: U^+ \to Y$. A *strictly causal* response has, for each t, $f_t(u_1,\cdots,u_t)$ independent of u_t; for a *memoryless* response, $f_t(u_1,\cdots,u_t)$ depends *only* on u_t; an *equilibrium response* is one for which there is some $\bar{u}$ in U with $f(\bar{u}w)=f(w)$ for all w in U^+. A response is *polynomial* (resp. *analytic*) iff U is a subset of k^m, $Y=k^p$, and f_t is a polynomial function in mt variables (resp. $k=\mathbb{R}$ or C, U is an analytic manifold, $Y=k^p$, and each f_t is analytic), for all $t \geqslant 1$.

The interpretation of the above is that output values y_t at time t are functions of inputs $u_1,\cdots,u_t$ at times $1,\cdots,t$, "strictly" causal meaning that present inputs cannot affect present outputs. "Equilibrium" is equivalent to "shift invariance" of the corresponding input/output map, where the "equilibrium input value" $\bar{u}$ plays the role of a "zero input."

When U is Euclidean space $\boldsymbol{R}^m$ (or an open set thereof), Y is just $\boldsymbol{R}$, and each f_t is a real analytic map, it is customary to represent f by a "Volterra series"

$$y(t)=\sum_{i=0}^{\infty} L_{ti}(u_1,\cdots,u_t) \tag{1.2}$$

where L_{ti} is a homogeneous degree i polynomial in the coordinates of $(u_1,\cdots,u_t)$. Other representations in this case are the "multidimensional z-transform" [1] and the "regular transfer function" [8]. Each of these alternative representations has its computational advantages depending on the problem to be solved. Since they all give the same information about f, and since the passage between them is well understood, this paper uses the more abstract definition in (1.1), which has the advantage of allowing for nonanalytic nonlinearities in f.

B. Systems

Definition 1.3: A *system* $\Sigma=(X,P,Q,\bar{x})$ is defined by a vector space X, maps $P: X \times U \to X$ and $Q: X \times U \to Y$ and an $\bar{x}$ in X (called, respectively, the *state space*, *next state*, *transition* or *control* map, *output* or *measurement* map, and the *initial* state). The *dimension* of Σ is the (possibly infinite) dimension of X; zero-dimensional systems are called *memoryless*. A *state-output* system has Q independent of inputs U. When there is an $\bar{u}$ in U with

$P(\bar{x},\bar{u})=\bar{x}$, Σ has *equilibrium initial state* (EIS). When P and Q are affine in x, Σ is a state-affine (S-A) system. When P and Q are linear, Σ is *state-linear*. A *polynomial* (resp. *analytic*) system has $X=k^n$, U a subset of some k^m, $Y=k^p$, and P,Q polynomial (resp. $k=\mathbb{R}$ or $\mathbb{C}$, U a manifold, $X=k^n$, $Y=k^p$, P, and Q analytic).

A system as defined above corresponds to a set of difference equations

$$\begin{aligned} x_{t+1}&=P(x_t,u_t)\\ y_t&=Q(x_t,u_t), \quad t=1,2,3,\cdots \end{aligned} \tag{1.4}$$

with the initial condition $x_1=\bar{x}$. State-output systems, in which y_t is a function of x_t alone, are the ones found most frequently in the control literature, and are called "Moore machines" in automata theory. In the state-affine case one has equations

$$\begin{aligned} x_{t+1}&=F(u_t)x_t+G(u_t)\\ y_t&=H(u_t)x_t+I(u_t) \end{aligned} \tag{1.5}$$

with $F(u)$, $G(u)$ linear maps and $G(u)$, $I(u)$ vectors for each u in U. The notations F,G,H,I will be used freely instead of P,Q. The symbolic notation $x'=P(x,u)$, $y=Q(x,u)$ (where the prime indicates a time-shift operator) will be used freely.

The extended transition map $P^*: X\times U^*\to X$ is defined recursively by $P^*(x,e):=x$, $P^*(x,wu):=P(P^*(x,w),u)$; P^t is the restriction to sequences of length t. With a slight abuse of notation, P^* will be denoted also by P. For any $w=vu$ in U^+, where v is in U^* and u in U, one writes $Q^+(x,w)$ (or just $Q(x,w)$): $=Q(P(x,v),u)$, and Q^t for the restriction to $X\times U^t$. The *reachability map* of Σ is $g: U^*\to X$ where $g(w):=P(\bar{x},w)$; the *reachable states* are those in its image. Σ is said to *realize* its response $f_\Sigma: U^+\to Y: w\to Q(\bar{x},w)$.

Definition 1.6: Σ is *span-reachable* iff X is the smallest affine manifold containing the reachable states, *observable* iff the functions $Q(x,\cdot): U^+\to Y$, x in X, are all distinct, and *span-canonical* iff both span-reachable and observable.

Definition 1.7: An (affine) *system morphism* $T:\Sigma_1\to\Sigma_2$ is an affine map $T: X_1\to X_2$ satisfying $T\bar{x}_1=\bar{x}_2$ and, for all x, u, $T(P_1(x,u))=P_2(T(x),u)$ and $Q_1(x,u)=Q_2(T(x),u)$.

The above definitions are useful in the context of the S-A realization problem, but "*span*-reachability" and "*affine* system morphism" are of course too weak in a purely nonlinear context (see [42], [46]). When systems are normalized (by a translation) to have initial state $\bar{x}=0$, span reachability is of course the same as requiring that X be the smallest *subspace* containing reachable states, and system morphisms become *linear* maps. Since a translation is required when normalizing, it is clear that a definition of morphism should include translations when the category of *all* S-A systems is considered; otherwise, no uniqueness can be expected (as with the internally bilinear counterexamples in [52].

Existance of a morphism $\Sigma_1\to\Sigma_2$ forces equality of responses; the following is a partial converse to this fact.

Theorem 1.8: Let Σ_1 and Σ_2 be S-A systems having the same response map. Assume that Σ_1 is span reachable and Σ_2 is observable. Then there is a unique morphism $T:\Sigma_1\to\Sigma_2$. Furthermore, T is onto if Σ_2 is also span reachable, and one-to-one if Σ_1 is observable.

Proof: By span reachability, any state x in X_1 can be written as $\Sigma r_i g_i(w_i)$, $\Sigma r_i=1$, for some finite set of input sequences w_i and scalars r_i. Write $T(x):=\Sigma r_i g_2(w_i)$. To see that this gives a well-defined map $T: X_1\to X_2$, consider any other expression $x=\Sigma s_i g_i(w_i)$, $\Sigma s_i=1$ (for the same set of w_i's, adding zeroes if necessary to the r_i,s_i). Since $Q_1(\cdot,w)$ is affine, $\Sigma r_i Q_1(g_1(w_i),w)=\Sigma s_i Q_1(g_1(w_i,w))$, i.e., $\Sigma r_i f(w_i w)=\Sigma s_i f(w_i w)$, for any w in U^+. This implies that $Q_2(\Sigma r_i g_2(w_i),w)=Q_2(\Sigma s_i g_2(w_i),w)$ for all w in U^+ and thus, by observability of Σ_2, that $\Sigma r_i g_2(w_i)=\Sigma s_i g_2(w_i)$. Thus T is well defined, and it is clearly affine. Uniqueness is clear, since $T(g_1(w_i))=g_2(w_i)$ is forced by the definition of morphism. The last two statements follow by analogous arguments.

Remark 1.9: If a given S-A system Σ is not span reachable, one may of course restrict P and Q to the affine span $\bar{X}$ of the reachable states (considered as a vector space in itself, once that an arbitrary x in $\bar{X}$ is choosen as origin), resulting in a span-reachable realization of the same response. If Σ is not observable, a dual construction gives an observable realization: it is sufficient to form the quotient of X by the subspace of all states indistinguishable from zero (i.e., all x with $Q(x,\cdot)=Q(0,\cdot)$); P and Q naturally induce maps in the quotient. Thus existence of a single S-A realization of an f already implies existence of a span-canonical one. Such a realization is constructed below, for any f.

Definition 1.10: The *image realization* of the response f has $X:=[U^+,Y]$, $\bar{x}:=f$, $P(b,u):=b_u$, and $Q(b,u):=b(u)$.

The above realization is an S-A, in fact state-linear, system. This is clear from the definition of the vector space structure on X. Moreover, it is observable. Indeed, if $Q(b,w)=b(w)$ is known for all w in U^+, then b is determined uniquely, as an element of X. By Remark (1.9), a span-canonical S-A realization can be obtained by restriction to the affine span L_f of the reachable set $\{f_w, w$ in $U^+\}$. Thus from (1.8) and (1.9) the following theorem results.

Theorem 1.11: Any response f has a span-canonical S-A realization Σ_f, unique up to isomorphism.

Definition 1.12: f is *S-A finitely realizable* iff there exists some finite-dimensional S-A system realizing f; a *minimal* realization is then one whose dimension is smallest among all possible S-A realizations of f.

From (1.8) and (1.11) the next theorem results.

Theorem 1.13: Let f be S-A finitely realizable. Then a S-A realization of f is minimal iff it is span-canonical.

It is important to remark that the above minimality holds *only* with respect to the class of S-A realizations. An extreme case of this is illustrated by the one-dimensional system Σ_0 with $U=Y=k$ and equations $x'=x+u, y=x^s$

(s = positive integer), $\bar{x}=0$. The response f_0 of this system is also realizable by an S-A system, since polynomial nonlinearities can be replaced by equations in the powers of state variables. But it follows from latter results that f_0 admits no S-A realizations of dimension less than s. (A fuller discussion of this issue, along with some more remarks on the example Σ_0, which is completely reachable and observable, can be found in [42].)

A response f is strictly causal (resp. equilibrium) iff its span-canonical realization is state-output (resp. has EIS).

Proposition 1.14: For any response f, the following statements are equivalent: a) f is strictly causal, b) f has a state-output realization, c) any span-reachable S-A realization of f is state-output. Similarly, there is equivalence of: a′) f is an equilibrium response, b′) f has a realization with EIS, c′) any observable realization of f has EIS.

Proof: Both c) implies b) implies a), and c′) implies b′) implies a′) are trivial.

Assume now that f is strictly causal, so $f(wu)=f(wv)$ for all w in U^*, u,v in U. Thus $Q(g(w),u)=Q(g(w),v)$ for all w in U^*. Since Q is affine, $Q(\Sigma r_i g(w_i),u)=Q(\Sigma r_i g(w_i),v)$ for all affine combinations of reachable states, i.e., for all states. So Q is independent of inputs, i.e., a) implies c).

If f is an equilibrium response then, for any Σ realizing f, $Q(\bar{x},w)=f(w)=f(\bar{u}w)=Q(\bar{x},\bar{u}w)=Q(P(\bar{x},\bar{u}),w)$ for all w in U^+. If Σ is observable, $\bar{x}=P(\bar{x},\bar{u})$. So a′) implies c′).

C. Bounded Responses and Finite-Type Systems

For any set $\delta_i, i=0,\cdots,s$ of functions from U into k, any $w=u_1\cdots u_t$ in U^t, and any multiindex $\alpha=\alpha_1\cdots\alpha_t$, (each α_i an integer between 0 and s,) $\delta_\alpha(w)$ denotes the product $\delta_{\alpha_1}(u_1)\delta_{\alpha_2}(u_2)\cdots\delta_{\alpha_t}(u_t)$. The "tensors" on such a set give rise to bounded responses.

Definition 1.15: A response f is *bounded*, of *type* $J=\{\delta_0,\cdots,\delta_s\}$, where the δ_i are functions $U\to k$, iff for each $t\geq 1$ there are (finitely many) vectors a_α in Y with

$$f(w)=\Sigma\delta_\alpha(w)a_\alpha \tag{1.16}$$

for all w in U^t.

Conventions 1.17: Without loss of generality it will be always assumed that δ_0 is the unit constant function: $\delta_0(u)=1$ for all u in U. This will greatly simplify notations. Furthermore, the family of functions J will be assumed *linearly independent* (i.e., if there are scalars r_i with $\Sigma r_i\delta_i(u)=0$ identically on U then all $r_i=0$); if this were not the case, a maximal linearly independent set can be extracted from a given J. For a type $J=\{\delta_0,\cdots,\delta_s\}$, $[J]$ denotes the set of integers $\{0,\cdots,s\}$.

An obvious and trivial example of bounded response is just any map $\delta_1: U\to Y$, inducing a memoryless response $f(u_2\cdots u_t)=\delta_0(u_1)\cdots\delta_0(u_{t-1})\delta_1(u_t)$. More interesting examples follow.

Example 1.18: As explained in the introduction, the terminology "bounded" has its origins in the main (and motivating) example: *polynomial* bounded responses. This case corresponds to U = a subset of k^m, $m\geq 1$, and the δ_i being a set of possible monomials in the m input variables (for instance, J = all monomials of degree $\leq d$, for some d). Those classes of responses for which a satisfactory realization theory is (at least partially) developed are all bounded. They are as follows. a) *Linear* responses: $U=k^m$ and each $f_t: k^{mt}\to Y$ is linear; in particular, these are of type $J_L=\{\delta_0,\cdots,\delta_m\}$ with $\delta_0=1$ and δ_i = ith projection $k^m\to k$ = monomial $u_{(i)}$. b) *Internally biaffine* responses [4], [11], [17], [18], [25]–[27], [50], [52]: $U=k^m$ and no inputs at any given instant appear multiplied among themselves in future outputs; alternatively, these are precisely all those responses of type J_L. c) *Multilinear* responses [2], [3], [20], [21], [28], [29], [32], [38]: $U=k^m$ and, indentifying U^t with $(k^t)^m$ = set of m-vectors of length-t scalar sequences, the map $f_t:(k^t)^m\to Y$ is a m-linear for each t; alternatively, these are responses of type $J_m=\{u_{(1)}^{j_1}\cdots u_{(m)}^{j_m}, j_i=0 \text{ or } 1\}$ and each monomial $\delta_\alpha(w)$ in (1.16) for which $a_\alpha\neq 0$ has total degree exactly m. (More generally, one may consider "vector" r-linear responses with $U=\hat{U}_1\times\cdots\times\hat{U}_r$ and each $\hat{U}_i=k^{m_i}$, $i=1,\cdots,r$, $m_1+\cdots+m_r=m$; these result also in bounded responses.) d) *Homogeneous* (degree-d) responses [8], [22], [42]: $U=k^m$ and each f_t is a homogeneous polynomial of degree d; alternatively, these are the responses of type $J_{h,d}=\{u_{(1)}^{j_1}\cdots u_{(m)}^{j_m}, j_1+\cdots+j_m\leq d)$ such that $\delta_\alpha(w)$ has total degree exactly d whenever $a_\alpha\neq 0$.

It should be carefully noted that b) and c) above give different, incomparable, classes of responses.

Example 1.19: Another rich source of examples is provided by bounded *piecewise linear* responses. A simple example is given by $U=Y=\boldsymbol{R}$ and $J_{\text{step}}:=\{\delta_0=1,\delta_1\}$, with δ_1 = the step function $\delta_1(x)=0$ for $x\leq 0$ and 1 for $x>0$; then a response like f with $f(u_1\cdots u_t)=0$ if all $u_i\leq 0$ and $=2(u_1+\cdots+u_t)$ otherwise, is of type J; indeed, $f(u_1\cdots u_t)$ equals

$$2[\delta_1(u_1)\delta_0(u_2)\cdots\delta_0(u_t)+\delta_0(u_1)\delta_1(u_2)\cdots\delta_0(u_t) + \cdots + \delta_0(u_1)\cdots\delta_1(u_t)].$$

A more interesting example has U = the interval $[0,1]$, $Y=\boldsymbol{R}$, and, for r a positive integer, $J_r=\{\delta_0=1,\delta_1,\cdots,\delta_{r-1}\}$, where δ_i := characteristic function of $[(i-1)/r, i/r)$. The responses of type J_r are given, for each t, by the step functions on $[0,1]^t$, constant on hypercubes with sides of length $1/r$. Even more interesting types of responses appear when adding monomials to the above: piecewise linear and, more generally, piecewise polynomial responses are obtained.

Realizations of bounded responses will have themselves a special structure.

Definition 1.20: The S-A system Σ is of *finite type* $J=\{\delta_0\cdots\delta_s\}$ iff there are affine maps P_i, Q_i with $P(x,u)=\Sigma\delta_i(u)P_i(x)$ and $Q(x,u)=\Sigma\delta_i(u)Q_i(x)$.

In terms of linear maps and translations, in the above case there are linear maps $F_i: X\to X$, $H_i: X\to Y$, and vectors G_i in X, I_i in Y, such that system equations become

$$x' = \sum_{i=0}^{s} \delta_i(u) F_i x + \sum_{i=0}^{s} \delta_i(u) G_i$$
$$y = \sum_{i=0}^{s} \delta_i(u) H_i x + \sum_{i=0}^{s} \delta_i(u) I_i.$$

Example 1.21: a) A *linear* system has type J_L (cf., (1.18a)) and equations

$$x' = \delta_0(u) F x + \sum \delta_i(u) G_i$$
$$y = \delta_0(u) H x + \sum \delta_i(u) I_i$$

with $\bar{x} = 0$; usually one assumes all $I_i = 0$. b) An *internally biaffine* system is any system of type J_L; (these systems are sometimes called "internally bilinear" or just "bilinear" in the references in (1.18b), especially in the EIS case; the latter terminology is misleading, since it is also used for the next example). c) A *standard multilinear system* is best described by a "wiring diagram" as in the references in (1.18c); in the bilinear ($m=2$) case, these are state-affine systems with $U = k^2$ and of type $J = \{1, u_{(1)}, u_{(2)}, u_{(1)}u_{(2)}\}$ whose equations can be decomposed into three blocks:

$$x_1' = A_{11}x_1 + u_{(1)}B_1$$
$$x_2' = A_{22}x_2 + u_{(2)}B_2$$
$$x_3' = A_{33}x_3 + u_{(2)}A_{31}x_1 + u_{(1)}A_{32}x_2 + u_{(1)}u_{(2)}B_3$$
$$y = Cx_3$$

where the A_{ij}, B_i, C are matrices and vectors of appropriate sizes. These systems are clearly of finite type.

Example 1.22: Piecewise linear and step-function types are useful in modeling "logical" controls. For instance, a system with transitions $x' = Ax$ if u is positive but $x' = Bx$ otherwise, is of type $J_{\text{step}}: x' = \delta_1(u)(A-B)x + \delta_0(u)Bx$.

Remark 1.23: Every *finite-dimensional* S-A system is obviously of finite type, when Y is finite dimensional. Indeed, it is only necessary to choose bases for X and Y; if dim $X = n$, dim $Y = p$ this results in $n^2 + nm + np + p$ functions $\delta_i(u)$ as entries of the corresponding matrices F, G, H, I. Adding $\delta_0 = 1$ if necessary, the given system is clearly of finite type.

The connection between bounded responses and finite type systems is given in the following theorem.

Theorem 1.24: A response f is bounded, of type J, iff its span-canonical realization Σ_f is of finite type J.

Proof: If f has *any* realization Σ of type J, the definition of f_Σ shows that f has type J. The converse statement will be immediate from the construction to be given in (2.9).

D. Some Useful Formulas

Explicit formulas can be given in the finite-type case for reachability maps, responses, morphisms, etc; these will be used later.

Remark 1.25: A morphism between two finite-type systems $\Sigma, \hat{\Sigma}$ can be described somewhat more explicitely than in (1.7). Adding zero matrices and vectors if necessary, it may be assumed that both systems are of the same type J. It is easy to verify then that an affine map $T = Ax + b$ mapping X into $\hat{X}$ induces a morphism from Σ into $\hat{\Sigma}$ iff $T\bar{x} = \hat{\bar{x}}$, $AF_i = \hat{F}_i A$, $AG_i = \hat{G}_i + \hat{F}_i b$ for $i \neq 0$, and $AG_0 = \hat{G}_0 + \hat{F}_0 b - b$, $\hat{H}_i A = H_i$, and $\hat{I}_i + \hat{H}_i b = I_i$.

Notation 1.26: For each multiindex $\alpha = \alpha_1 \cdots \alpha_t$ in $[J]^*$ (J = type of Σ), G_α in X is defined by

$$G_\alpha := F_{\alpha_t} \cdots F_{\alpha_2} G_{\alpha_1}.$$

The vectors W_α, α in $[J]^*$, are defined by $W_\alpha := G_\alpha$ if $\alpha_1 \neq 0$, α in $[J]^+$, $W_e := 0$, and $W_{0\alpha} := G_{0\alpha} + W_\alpha$ otherwise. Finally, the vectors V_α, $\alpha = \alpha_t \cdots \alpha_1$ in $[J]^+$ are defined as

$$V_\alpha := F_{\alpha_t} \cdots F_{\alpha_1}\bar{x} + W_\alpha.$$

A straightforward induction gives

Lemma 1.27: With the above notations,

$$g(u_1 \cdots u_t) = \sum \delta_\alpha(u_1 \cdots u_t) V_\alpha$$

and

$$f_\Sigma(u_1 \cdots u_t) = \sum \delta_\alpha(u_1 \cdots u_t) H_{\alpha_t} V_{\alpha_1 \cdots \alpha_{t-1}} + I(u_t) \tag{1.28}$$

both sums over all α in $[J]^t$.

The following algebraic facts will be needed.

Lemma 1.29: a) If $\{\delta_i, i \text{ in } [J]\}$ is a linearly independent set of functions $U \to k$ then, for each $t \geq 1$, the set $\{\delta_\alpha, \alpha \text{ in } [J]^t\}$ of functions $U^t \to k$ is also linearly independent.

b) Let $\delta_1, \cdots, \delta_r$ be a family of linearly independent functions $U \to k$ and let X be a vector space, $b_1, \cdots, b_r$ vectors in X. Consider $X_1 :=$ subspace generated by the b_i, and $X_2 :=$ subspace generated by the vectors $\{\sum \delta_i(u) b_i, u \text{ in } U\}$. Then $X_1 = X_2$.

Proof: a) For $t = 1$ this is true by hypothesis. If true for t but not for $t+1$, there is some relation $\sum r_\alpha \delta_\alpha = 0$, the α in $[J]^{t+1}$. Then, for each w in U^t and u in U:

$$\sum a_i(w) \delta_i(u) = 0, \qquad \text{for all } u \text{ in } U$$

with the a_i in the subspace generated by the δ_α, α in $[J]^t$. Independence of the δ_i forces $a_i(w) = 0$ for all i and all w in U^t, so by induction all $r_\alpha = 0$.

b) Clearly X_2 is included in X_1. Let $L: X \to k$ be any linear functional such that $L(X_2) = 0$. Then, for any u in U,

$$0 = L\Big(\sum \delta_i(u) b_i\Big) = \sum \delta_i(u) L(b_i).$$

By linear independence of the δ_i, all $L(b_i) = 0$. Thus $L(X_1)$ is also zero. So X_1 is included in X_2.

The following lemma permits checking span reachability. A system with $\bar{x} \neq 0$ can always be transformed by a translation into one with $\bar{x} = 0$. The original system is span-reachable iff the second one is. The advantage of this normalization is that since $\bar{x} = 0$ is now always in the affine span of the reachable set, this span is a subspace and span-reachability means $X =$ subspace generated by $g(U^*)$.

Proposition 1.30: With $\bar{x}=0$, an S-A system is span-reachable iff

$$\text{span}\{G_\alpha, \alpha \text{ in } [J]^+\} = \text{span}\{V_\alpha, \alpha \text{ in } [J]^+\} = X.$$

Proof: Consider the affine manifolds

$$X_t := \text{affine span of } \{g(w), |w| \leq t\}.$$

Here these are just the subspaces generated by the $g(w), |w| < t$, since $g(\bar{u}\cdots\bar{u}) = \bar{x} = 0$ is in every generating set. But the subspace generated by X_t is the sum of the $\hat{X}_i$, $i=0,\cdots,t$, where $\hat{X}_i :=$ subspace generated by $g(U^i)$. By (1.27) and (1.29), each of the $\hat{X}_i$ is generated by the V_α, α in $[J]^i$. Since here $V_\alpha = W_\alpha$ and by definition $V_\alpha = G_\alpha$ if $\alpha_1 \neq 0$ and $V_{0\alpha} = G_{0\alpha} + V_\alpha$, the result follows.

Since equality $X_t = X_{t+1}$ clearly implies $X_{t+1} = X_{t+2} = \cdots$, a dimensionality argument gives the following corollary.

Corollary 1.31: An n-dimensional system is span-reachable iff $X_n = X$. If $\bar{x}=0$, this is equivalent to $X=$span of all G_α, α in J^t, $t \leq n$.

Since a S-A system is observable iff there are not states x with $Q(x,\cdot) = Q(0,\cdot)$, observability becomes also a linear condition. Denoting for α in $[J]^+$:

$$H_\alpha := H_{\alpha_t} F_{\alpha_{t-1}} \cdots F_{\alpha_1} : X \to Y$$

an argument analogous to the preceeding ones gives.

Proposition 1.32: A S-A system is observable iff the intersection of the kernels of the H_α, α in $[J]^+$, is zero. When the system is n-dimensional, it is sufficient to form the intersections of the ker H_α, α in $[J]^t$, $t \leq n$.

Notational Conventions 1.33: A few normalizations will simplify notations considerably. First, it may be assumed that, if a response f, or a system Σ, is of type $J = \{\delta_0,\cdots,\delta_s\}$, with U a subset of k^m, then $\delta_i(0)=0$ for all $i=1,\cdots,s$: if this were not the case, it is enough to replace δ_i by $\delta_i' := \delta_i - r\delta_0$, where $r = \delta_i(0)$; the new J is again linearly independent and the class of responses, and systems, of type J is unchanged. (In the polynomial case, the δ_i are nonconstant monomials in the input variables for $i \neq 0$, so $\delta_i(0)=0$ is always satisfied.) When f is an equilibrium response, or Σ has EIS, a coordinate translation can be effected in the input space U resulting in $\bar{u}=0$. Similarly, a translation in X gives $\bar{x}=0$. These normalizations will be assumed when dealing then with the equilibrium case.

The set of all those multiindexes α in $[J]^+$ for which $\alpha_1 \neq 0$ will be denoted by $[J]_+$. Then, f is an equilibrium response iff $a_{0\alpha} = a_\alpha$ for all α in $[J]^+$, and Σ has EIS iff $G_0 = 0$. Thus (1.30) simplifies to the following lemma.

Lemma 1.34: An EIS system is span-reachable iff $X=$ span $\{G_\alpha, \alpha \text{ in } [J]_+\}$.

Remark 1.35: The given span reachability and observability conditions reduce to the well-known ones for linear and internally biaffine systems. As in those examples, these conditions result for EIS systems in a "duality" between span reachability and observability. It is as yet unclear whether these dualities have any system-theoretic meaning in the S-A case.

Numerical Examples 1.36: Consider a system Σ_1 with $U=\mathbf{R}$, $X=\mathbf{R}^3$, $\bar{x}=0$, and transitions

$$x_1' = x_2 \quad \text{if} \quad u \leq 0, \quad x_1+x_2+1 \quad \text{otherwise}$$
$$x_2' = x_3 \quad \text{if} \quad u \leq 0, \quad x_1+x_3 \quad \text{otherwise}$$
$$x_3' = x_2 - x_3.$$

This system is of type J_{step} (cf., (1.19)), where

$$G_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \quad F_0 = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & -1 \end{bmatrix} \quad F_1 = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Since $G_1, G_{11} = (1 \;\; 1 \;\; 0)'$ and $G_{110} = (1 \;\; 0 \;\; 1)'$ span X, Σ_1 is span reachable. Similarly, consider Σ_2 with same $U, X, \bar{x}$, and transitions

$$x_1' = u^2x_1 + ux_2 + u^2$$
$$x_2' = u^2x_1 + ux_3$$
$$x_3' = ux_2 - ux_3.$$

This system is of type $\{\delta_0, \delta_1, \delta_2\}$ with $\delta_1(u) = u$ and $\delta_2(u) = u^2$, and with $G_1 = 0$, $G_2 =$ the above G_1, $F_0 = 0$, $F_i =$ the above F_{i-1}, $i=1,2$. So G_2, G_{22}, and G_{221} span X, and Σ_2 is also span-reachable. Finally, consider Σ_3 with U, X as above, $\bar{x} = (1 \;\; 2 \;\; 0)'$ and

$$x_1' = u^2x_1 + ux_2 - 2u - 1$$
$$x_2' = u^2x_1 + ux_3 - u^2 - 2$$
$$x_3' = ux_2 - ux_3 - 2u.$$

Normalizing to $\bar{x}=0$ by $x \to x - \bar{x}$, Σ_2 is obtained. So Σ_3 is span reachable too.

II. Matrix Minimal Realization Algorithms

This part introduces a general matrix mathod for minimal S-A realization of bounded responses. A reduction is first carried out, which allows restricting attention to the case of equilibrium responses without any loss of generality. Notations introduced in previous sections are freely utilized.

A. Reduction to the Equilibrium Case

For any response f of type $J = (\delta_0,\cdots,\delta_s\}$, a new response $\hat{f}$ will be derived from f as follows. The new input set $\hat{U}$ will be $U \times k$, (to be interpreted as "adding a new input channel") and $\hat{f}$ will be of type $\hat{J} = \{\hat{\delta}_0,\cdots,\hat{\delta}_s,\hat{\delta}_{s+1}\}$, where $\hat{\delta}_i(u,v) := \delta_i(u)$ for all (u,v) in $U \times k$, $i=0,\cdots,s$, and $\hat{\delta}_{s+1}(u,v) := v$. Clearly, $\hat{J}$ is linearly independent since J was. Using the notation in (1.16), for α in $[\hat{J}]^t$:

$$\hat{a}_{0\alpha} := \hat{a}_\alpha, \qquad \text{for all} \quad \alpha \quad \text{in} \quad [\hat{J}]^+$$
$$\hat{a}_{(s+1)\alpha} := a_{0\alpha} - a_\alpha, \qquad \text{if } \alpha \text{ in } [J]^+$$
$$\hat{a}_0 := a_\alpha, \qquad \text{if } \alpha \text{ in } [J]^+ \text{ with } \alpha_1 \neq 0$$
$$\hat{a}_\alpha := 0, \qquad \text{if } \alpha_1 \neq 0 \text{ but } \alpha_i = s+1 \text{ for some } i = 2,\cdots,t. \tag{2.1}$$

For any $w=u_1\cdots u_t$ in U^t, $\hat{w}$ in $\hat{U}^t$ is the sequence with $\hat{u}_i:=(u_i,1)'$. A straightforward calculation shows the following lemma.

Lemma 2.2: $f(w)=\hat{f}(\hat{w})$ for all w in U^+.

For any system Σ of type J having $\bar{x}=0$, a system $\hat{\Sigma}$ of type $\hat{J}$ and input set $\hat{U}$ can be derived as follows (with $\hat{\bar{x}}:=0$, $\hat{X}:=X$):

$$\hat{F}_i[\text{resp. } \hat{H}_i,\hat{I}_i]:=F_i[\text{resp. } H_i,I_i], \quad \text{if } i=0,\cdots,s$$
$$:=\text{zero}, \quad \text{if } i=s+1;$$
$$\hat{G}_i:=G_i \quad i=1,\cdots,s \quad \hat{G}_0:=0 \quad \hat{G}_{s+1}:=G_0 \tag{2.3}$$

then $\hat{\Sigma}$ has EIS by definition. Since $\hat{G}_\alpha=G_\alpha$ whenever α is in $[J]^+$ with $\alpha_1\neq 0$, $s+1$, $\hat{G}_{(s+1)\alpha}=G_{0\alpha}$ if α is in $[J]^+$, and $\hat{G}_\alpha=0$ otherwise, the following lemma follows from (1.30).

Lemma 2.4: Σ is span-canonical iff $\hat{\Sigma}$ is.

This result is in fact also a consequence of the following lemma.

Lemma 2.5: f has an S-A realization of dimension n and type J iff $\hat{f}$ has an EIS S-A realization of dimension n and type $\hat{J}$.

Proof: Given Σ realizing f, $\hat{\Sigma}$ realizes $\hat{f}$. Indeed, using the notations in (1.26), $\hat{V}_\alpha=\hat{G}_\alpha=G_\alpha=V_\alpha$ if $\alpha_1\neq 0$, $s+1$, $\hat{V}_{0\alpha}=\hat{G}_{0\alpha}+\hat{V}_\alpha=\hat{V}_\alpha$ (since all $\hat{G}_{0\alpha}=0$), and $\hat{V}_{(s+1)\alpha}=G_{0\alpha}=V_{0\alpha}-V_\alpha$; this implies the formulas in (2.1).

Conversely, if Σ^* realizes $\hat{f}$, then a system Σ can be defined setting $v=1$, i.e., $F_i:=F_i^*$ for $i=1,\cdots,s$, $F_0:=F_0^*+F_{s+1}$, similarly for G,H,I. By (2.2), Σ realizes f.

The conclusion from the above results is that *it is enough to develop a minimal realization algorithm for equilibrium responses*. If a given f is not such a response, then $\hat{f}$ can be constructed and a minimal realization of f can be obtained from one for $\hat{f}$. (System-theoretically, the addition of an input channel corresponds to adding a "reset control"; in fact, $\hat{U}=U\times\{0,1\}$ would have worked equally well.)

Further Simplification 2.6: If Σ has EIS $\bar{x}=0$, clearly $f_\Sigma(u)=I(u)$ for all u in U. Given an equilibrium response f, the response f' with

$$f'(u_1\cdots u_t):=f(u_1\cdots u_t)-f(u_t)$$

satisfies $f'(u)=0$. A realization Σ' of f' will thus have $I'(u)=0$, and will give rise to a realization Σ of f by adding $I(u):=f(u)$. *It will therefore be assumed that $f(u)=0$ for all u in U.*

B. The Algorithm

A bounded equilibrium response of type $J=\{\delta_0,\cdots,\delta_s\}$ will be fixed. For the explicit matrix calculations it will be assumed that $Y=k^p$, some $p\geq 1$.

Definition 2.7: The *behavior* (or *generalized Hankel*) *matrix* $B(f)$ of f is the doubly infinite block matrix constructed as follows. Rows of $B(f)$ are indexed by $[J]^+$ and columns are indexed by $[J]_+$. The (β,α)th block entry of $B(f)$, β in $[J]^+$, α in $[J]_+$, is the vector $a_{\alpha\beta}$. The submatrix obtained by restricting to those rows with indexes $|\beta|\leq r$ and columns $|\alpha|\leq t$ is denoted by $B(f)_{rt}$. The order of the rows and columns of $B(f)$ is irrelevant for the validity of the algorithm to be described; in fact different orderings may be useful for various purposes, as described later. It will be convenient however to assume that they are ordered lexicographically, i.e., (block) rows are ordered as $0,1,2,3,\cdots,s,00,01,\cdots,0s,10,\cdots,ss,000,\cdots$, etc., and columns are ordered as $1,2,\cdots,s,10,11,\cdots,1s,20,\cdots,ss,100,101,\cdots$, etc.

Realization Algorithm 2.8: Assume that f is an S-A finitely realizable response or, equivalently, that $B(f)$ has a finite rank n. Then $B(f)_{nn}$ already has rank n. Pick a nonsingular n by n submatrix Φ of $B(f)_{nn}$. Let $B_{\alpha_1},\cdots,B_{\alpha_n}$ be the columns, and $B^{\beta_1,i_1},\cdots,B^{\beta_n,i_n}$ the rows, defining Φ ($B^{\beta,i}$ means the ith entry in the block row of index β). Let Φ_i be the submatrix of $B_{n,n+1}$ defined by the same rows but with columns $B_{\alpha_1 i},\cdots,B_{\alpha_n i}$, $i=0,\cdots,s$. Then

$$G_i:=\Phi^{-1}(a_{i\beta_1},\cdots,a_{i\beta_n})', \quad i=1,\cdots,s$$
$$F_i:=\Phi^{-1}\Phi_i, \quad i=0,\cdots,s$$
$$H_i:=(a_{\alpha_1 i},\cdots,a_{\alpha_n i}), \quad i=0,\cdots,s \tag{2.9}$$

define a minimal S-A realization of f (of dimension n).

Proof of Correctness: In general, a span-canonical realization Σ of any bounded f can be constructed as follows. The state-space X is the linear span of the columns of $B(f)$ (seen as infinite vectors), G_i is the ith column, $i=1,\cdots,s$, F_i is the linear map induced by the column shifts $B_\alpha\to B_{\alpha i}$, $i=0,\cdots,s$, and H_i is the map $X\to Y$ induced by restricting a column to its ith row block, $i=0,\cdots,s$. The maps F_i are well defined, since any relation among columns $\Sigma r_\alpha B_\alpha=0$ (finite sum), r_α in k, implies $\Sigma r_\alpha B_{\alpha i}=0$ for all i in J. Indeed, the jth (block) row of $\Sigma r_\alpha B_{\alpha i}$ is $\Sigma r_\alpha a_{\alpha ij}$, which is also the ijth row of $\Sigma r_\alpha B_\alpha$, hence zero. Let $\bar{x}=0$.

Clearly, $G_\alpha=B_\alpha$ for every α in $[J]_+$, so Σ is span-reachable. For any $x=\Sigma r_\alpha B_\alpha$ in the intersection of all the ker H_β, β in $[J]^+$,

$$0=H_\beta x=\sum r_\alpha H_\beta B_\alpha=\sum r_\alpha a_{\alpha\beta}$$

is the βth row of x. Thus $x=0$, and Σ is observable.

It only remains to prove that Σ is indeed a realization of f. Pick $t\geq 1$, w in U^t. By (1.27), $f_\Sigma(w)=\Sigma\delta_\alpha(w)H_{\alpha_t}V_{\alpha_1\cdots\alpha_{t-1}}=\Sigma\delta_\alpha(w)H_{\alpha_t}G_{\alpha_1\cdots\alpha_{t-1}}=\Sigma\delta_\alpha(w)a_\alpha=f(w)$, as wanted.

Thus Σ is a span-canonical realization of f, and the latter is S-A finitely realizable iff Σ is finite dimensional. Further, Σ is a minimal S-A realization of f, and its dimension n is the dimension of X, i.e., the rank of $B(f)$. By (1.31), (1.32), the rank of $B(f)_{nn}$ is also n, so the latter has a nonsingular Φ as claimed. The explicit formulas are just the expressions of the F_i,G_i,H_i using the columns defining Φ as a basis of X.

Remarks 2.10: a) The above is a minor variation of the algorithm given by [41] for linear systems, and for representing power series in noncommutative variables given by [15]; the method of proof is analogous to the one given by [39] for linear systems over rings. Alternatively, the algorithm for linear systems given by HO (see [30, ch. 10]) could be also easily generalized—here one *factors* B_{nn} rather than finding a full submatrix. It is as yet unclear which of the alternatives provides a "better" algorithm. A theory of *partial* realizations can be also derived from the above constructions, again by generalizing the linear case. In fact, this has already been done by [25] for the (EIS) internally biaffine case. It is interesting to remark that, since Σ in (2.9) is span canonical, it follows that the column space X is isomorphic to L_f (by (1.10), (1.11)). Generalizations of behavior-like matrices to the non-bounded case, (such that finite rank be equivalent to some sort of realizability), can be found in [42], [46] where a Jacobian matrix of differentials of f replaces $B(f)$.

b) When restricting attention to a particular class C of responses, it is to be expected that many, if not most, columns of $B(f)$ will be zero independently of the particular f. Thus, for computational purposes one orders columns in suitable ways, dropping those which are zero (examples below).

c) The application of the above algorithm to noisy data is a problem which should be more carefully studied, since (2.9) is numerically unstable as given. It is, however, possible to modify the algorithm such as to achieve stability, at the cost of an increase in its computational complexity; for the HO version of (2.9) (cf., Remark a)), this modification follows along the lines of the work (in the linear case) of [13].

d) A serious practical question concerns the coefficients a_α themselves (B_{nn} involves those a_α with $|\alpha| \leqslant 2n$). In the polynomial case, a least squares fitting of the polynomial f_{2t} may be needed, given a bound on the dimension t of a S-A realization of f; alternatively, f_t could be obtained from the corresponding Volterra or other series, if such data is available. It is interesting to remark that f_{2n} is itself determined by a restriction f_{2r}, where r is the (usually much smaller) dimension of the reachable set of Σ_f ([48], theorem 6.1) in the EIS case; it would be interesting to have an algorithm which makes efficient use of this fact.

Besides its generality, the above algorithm (and its variation mentioned in a) above) serves to unify the literature.

Example 2.11: For linear responses (1.18a), the only nonzero columns of $B(f)$ are those with indexes $i0^j$, with $i=1,\cdots,m$ and $j \geqslant 0$; the only nonzero rows are those with indexes 0^i, $i \geqslant 1$. Ordering the columns as $1,2,\cdots,m,10,20,\cdots,m0,100,\cdots,$ and the rows as $0,00,000,\cdots,B(f)$ is precisely the Hankel matrix (A_{i+j}), where the A_r are the "impulse response" or "Markov" parameters $A_r=(a_{10^r},\cdots,a_{m0^r})'$. Then (2.9) reduces to the algorithm of [41].

Example 2.12: For internally biaffine responses (1.18b), $B(f)$ is essentially the "generalized Hankel matrix" given independently by [25] and [16]. In particular, the matrix given by Isidori [25] is $B(f)$ with columns ordered lexiocographically but with rows ordered as follows: first, all rows of indexes in $[J]$, then all those with indexes in $[J]^2$, $[J]^3$, etc.; for each index length r, the rows are ordered as $\alpha_1,\cdots,\alpha_t$, $t=2^r$, where, if i has the binary expansion $\Sigma b_j 2^{j-1}$ $(j=0,\cdots,r-1)$, then, α_i is the binary expression of the integer $2^{r-1}(\Sigma b_j 2^{-j})$. (The above expression for the α_i can be inductively proved from Isidori's algorithm.)

Example 2.13: For m-linear responses (1.18c), the algorithm coincides with the one given by [29]. For simplicity, let $m=2$. Thus f is of type $J=\{\delta_0=1,\delta_1=u_{(1)},\delta_2=u_{(2)},\delta_3=u_{(1)}u_{(2)}\}$. The nonzero columns of $B(f)$ can be partitioned into three sets: first, all those with indexes 10^i; then all 20^j; finally all $0^i20^j10^r$, $0^i10^j20^r$ and 0^i30^j. The nonzero rows are partitioned into three sets: all 0^i, all 0^i10^j, and all 0^i20^j. Then $B(f)$ has the block structure

$$\begin{bmatrix} 0 & 0 & B_{12} \\ 0 & B_2 & 0 \\ B_1 & 0 & 0 \end{bmatrix} \tag{2.14}$$

where the indicated submatrices are those obtained in the above reference. Since F_1 and F_2 shift columns in the first two blocks into the third, it follows that in a suitable basis the equations of Σ_f have the standard bilinear form in (1.21c).

C. State-Linear Models

The response of any S-A system Σ can be also realized by a state-*linear* system (P,Q linear). It is enough for this to enlarge the state-space to $\hat{X}:=X\times k$ and to add a constant equation $z(t+1)=z(t)$: defining $\hat{P}((x,z)',u):=(F(u)x+zG(u),z)'$ and $\hat{Q}((x,z)',u)=H(u)x+zI(u)$ with $\hat{x}:=(\bar{x},1)$ gives a state-linear $\hat{\Sigma}$ realizing f_Σ. Moreover, Σ is finite dimensional or of finite type iff $\hat{\Sigma}$ is. The realization theory of state-linear systems can be of course developed independently from that of S-A systems. Defining L-span reachable to mean that X is the smallest *subspace* (rather than affine manifold) generated by reachable states and defining L-span canonical: $=L$-span reachable + observable, L-morphism: $=$ *linear* T as in (1.7), the proofs in the previous sections can be repeated with minor modifications to yield the following theorem.

Theorem 2.15: Any response f has a span-canonical state-linear realization, unique up to isomorphism. If f is (S-A) finitely realizable, a state-linear realization is minimal (among all state-*linear* realizations) iff it is L-span canonical.

Remark 2.16: A Hankel of Behavior matrix $B_L(f)$ can be defined in a way suitable for constructing minimal state-linear realizations, by letting (block) rows be induced by $[J]^+$, columns by $[J]^*$, and writing $a_{\alpha\beta}$ in the (β,α)-the position (thus $B(f)$ is a submatrix of $B_L(f)$ when f has EIS). A minimal realization is now obtained letting $X=$ column space of $B_L(f)$, $\bar{x}:=$ column of index $e=$ empty word, and $F_i=$ shifts, $H_i=$ restrictions, as before. Thus the rank of $B_L(f)$ gives the dimension of a minimal state-linear realization. (Since the affine span of reachable states does not in general contain the origin—the normalization $\bar{x}:=0$ destroys the state-linear form—this dimension is in general one more than the dimension of Σ_f, the minimal S-A realization.)

Definition 2.17: The *exponent series* ϕ_f of a bounded response f of type $J=\{\delta_0,\cdots,\delta_s\}$ is a power series in noncommutative variables $z_0,\cdots,z_s$, defined by taking a_α, $\alpha=\alpha_1\cdots\alpha_t$ as the coefficient of the monomial $z_{\alpha_1}\cdots z_{\alpha_t}$.

For instance, the first response in (1.19) has series

$$2(z_1+z_1z_0+z_0z_1+z_1z_0^2+z_0z_1z_0+z_0^2z_1+\cdots).$$

Recall that a power series is *rational* iff it can be expressed in terms of polynomials using a finite number of sums, products and inversions.

Theorem 2.18: A bounded response f is S-A (or state-linear) finitely realizable iff ϕ_f is rational.

Proof: The Hankel matrix $B_L(f)$ is precisely the Hankel matrix of ϕ_f, as defined in [17]. Rationality is then equivalent to finite rank of $B_L(f)$, by the results there.

For example, the system of type $J=\{1,u,u^2\}$:

$$x_1'=x_1+2x_1u-x_2u^2$$

$$x_2'=x_1u^2+x_2$$

$$y=x_1$$

has a response f with rational exponent series

$$\left(1-z_0-2z_1+(1-z_0)^{-1}z_2\right)^{-1}.$$

This is calculated easily using the methods in [14]. In general, automata-theoretic techniques (the "regular calculus") result in a calculus for state-linear realizations. An application is given in (3.10) below.

III. Other Finiteness Results

A. Compositions of Systems; Finite Responses

Realizations of finite Volterra series, and of their non-analytic analogues, require introducing some new notions.

Definition 3.1: The *series composition* or *cascade* of two systems Σ_1, Σ_2 with $Y_1=U_2$ is the system $\Sigma:=\Sigma_1\cdot\Sigma_2$ with $X:=X_1\times X_2$, $U:=U_1$, $Y:=Y_2$, $\bar{x}:=(\bar{x}_1,\bar{x}_2)$ and

$$P((x_1,x_2),u):=(P_1(x_1,u),P_2(x_2,Q_1(x_1,u)))$$
$$Q((x_1,x_2),u):=Q_2(x_2,Q_1(x_1,u)). \tag{3.2}$$

It is easy to verify that series composition defines an associative operation on systems, the input/output map of $\Sigma_1\cdot\Sigma_2$ being the composition of the corresponding input/output maps.

Definition 3.3: Σ is a *cascade of linear* (resp. S-A, *internally biaffine*) systems iff $\Sigma=\Sigma_0\cdot\Sigma_1\cdots\Sigma_r$, where Σ_0 is memory-free and, for each $i=1,\cdots,r$, Σ_i has $P_i(x,u)$ linear in both x and u (resp. affine in x, affine in x and u). The cascade has *polynomial interconnections* iff the maps Q_i are polynomial (and each U_i, Y_i is a finite-dimensional vector space), and Σ_0 is polynomial.

Definition 3.4: A bounded response f of type $J=\{\delta_0=1,\delta_1,\cdots,\delta_s\}$ is *finite* iff there exists some integer K, which depends only on f, such that a_0 is zero whenever more than d of the entries α_i are nonzero.

For instance (cf. (1.18)), linear and multilinear responses are always finite, while internally biaffine ones in general are not. The main result here is as follows.

Theorem 3.5: The span-canonical S-A realization of a finite response f is (isomorphic to) a cascade of linear systems. Conversely, the response of a cascade of linear systems with polynomial interconnections is finite.

Proof: Since the construction in (2.1) and (2.2) preserves finiteness, it is enough to treat the EIS case. In the proof of algorithm (2.8) let X_j be the subspace of X generated by all those columns B_α, α in $[J]_+$, for which j or more of the α_i are nonzero. Then the shift F_0 maps X_j into itself, while the shift F_i maps X_j into X_{j+1}, for $i=1,\cdots,s$. If K is as in (2.2), then $X_{K+1}=0$. A basis $v_1,\cdots,v_n$ of X can be chosen by letting $v_r,\cdots,v_n$ be a basis of X_K, completing this to a basis $v_t,\cdots,v_r,\cdots,v_n$ of X_{K-1}, etc. This results in transition equations which can be arranged in blocks as follows:

$$x_1'=A_1x_1+B_1(u)$$

$$x_2'=A_2x_2+B_2(x_1,u)$$

$$\vdots$$

$$x_K'=A_Kx_K+B_K(x_1,x_2,\cdots,x_{K-1},u)$$

$$y=\sum C_i(u)x_i+I(u) \tag{3.6}$$

where the B_i are affine in the x_j and of type J in the u. Defining Σ_0 as the zero-dimensional map $B_1(u)$, Σ_1 as $x_1'=A_1x_1+u$, $y=(x_1,B_2(x_1,u))$, and in general Σ_i as $x_i'=A_ix_i+G_i(u)$, $y_i=(x_1,\cdots,x_i,B_{i+1}(x_1,\cdots,x_i,u))$ with $G_i=$ projection in last block entry of input, permits expressing (3.6) as in (3.3).

The converse is an easy induction.

In the polynomial case a degree $d(\alpha)$ is well defined for each δ_α, α in $[J]^+$, and a response of type J is bounded if and only if there is some d such that all monomials δ_α of degree greater than d appear with zero coefficient in f. In terms of the Volterra representation (2.1), this amounts to $L_{ti}=0$ for $i>d$. Thus one says that "f has a *finite Volterra series*;" this is in fact the motivation for the terminology. (It should be noted that "finite" is *relative* to the type of f; for example, every memory-free response is finite as a response of type $f(u)$, so for instance $y_t=\exp(u_t)$ is finite in this sense although it gives rise to an "infinite Volterra series" in the classical sense.) In the polynomial case, then, the proof of (3.5) can be repeated but letting now X_j be instead the span of the columns B_α with $d(\alpha)=j$. Then F_α maps X_j into $X_{j+d(\alpha)}$, and the equations (3.6) can be simplified even further. For example, if $m=1$ and $J=\{1,u,u^2,\cdots,u^d\}$, the span-canonical realization of a finite f of type J can be written in the following block form:

$$x_1' = F_0^{1,1}x_1 + uG_{1,1}$$
$$\vdots$$
$$x_t' = \sum_{j=1}^{t} F_0^{t,j}x_j + u\left[\sum_{j=1}^{t-1} F_1^{t,j}x_j\right] + \cdots$$
$$+ u^{t-1}F_{t-1}^{t,1}x_1 + uG_{t,1} + \cdots + u^tG_{t,t}$$
$$\vdots$$
$$y = \sum_{j=1}^{d} H_{0j}x_j + u\sum_{j=1}^{d-1} H_{1j}x_j + \cdots + u^{d-1}H_{d-1,1}x_1. \tag{3.7}$$

When f is, further, homogeneous of degree d (cf. (1.18d)), the above equations have

$$G_{ij}=0, \quad \text{if } i\neq j \qquad F_i^{t,j}=0, \quad \text{if } i+j\neq t$$
$$\text{and } H_{ij}=0, \quad \text{if } i+j\neq d. \tag{3.8}$$

Definition 3.9: The dth *truncation* of the polynomial response f is the finite response $f^{(d)}$ obtained by dropping all terms a_α with $d(\alpha)\geqslant d$.

Proposition 3.10: If f is an S-A finitely realizable bounded polynomial response, then $f^{(d)}$ is also S-A finitely realizable, for each d.

Proof: By (2.18), ϕ_f is rational. Let S_d be the power series in $z_0,\cdots,z_s$ which has $a_\alpha=1$ when $d(\alpha)\leqslant d$ and $a_\alpha=0$ otherwise; S_d is easily seen to be rational. For example, if $J=\{\delta_0=1,\ \delta_1=u\}$ then S_2 is

$$(1-z_0)^{-1}+(1-z_0)^{-1}z_1(1-z_0)^{-1}$$
$$+(1-z_0)^{-1}z_1(1-z_0)^{-1}z_1(1-z_0)^{-1}.$$

But the exponent series of $f^{(d)}$ is the coefficientwise or "Hadamard" product of ϕ_f and S_d. By the results of [15] it is itself rational, being a Hadamard product of rational series. Thus $f^{(d)}$ is finitely realizable.

Corollary 3.11: If $k=\boldsymbol{R}$ or $\boldsymbol{C}$ and f is realizable by an analytic system with EIS, then $f^{(d)}$ is S-A finitely realizable.

Proof: Let Σ be an analytic realization of f, and assume without loss of generality that $\bar{x}=0$, $\bar{u}=0$. Introducing if necessary new state variables for the monomials in state variables with total degree $\leqslant d$, it may be assumed that all terms in the power series expansions of P,Q are either linear in x (or constant) or of total degree greater than d. If $\hat{P},\hat{Q}$ are obtained from P and Q by dropping all terms of total degree greater than d, an S-A system $\hat{\Sigma}$ is obtained whose (necessarily bounded) response $\hat{f}$ coincides with f on all terms of degree $\leqslant d$ (the EIS hypothesis is crucial here, since it implies that P has no constant term and hence that no term of degree more than d in state variables can contribute a term of degree $\leqslant d$ when iterating P). Then $f^{(d)}=\hat{f}^{(d)}$ is S-A realizable by (3.10) applied to $\hat{f}$.

Remark 3.12: Corollary (3.11) is false without the EIS assumption. For instance, the finite ($d=1$) response with $f(u)=1$ and $f(u_1\cdots u_t)=2^{2^{(t-2)}}u_{t-1}$ is realizable by the (analytic but not S-A) system

$$x_1'=x_2u$$
$$x_2'=x_2^2$$
$$y=x_1 \tag{3.13}$$

with $\bar{x}=(1\ \ 2)'$. But $f^{(1)}=f$ has exponent series $\phi_f=(\Sigma 2^{2^t}z_0^t)z_1$, which is not rational. Thus f admits no finite-dimensional S-A realization. (Note that (3.13) could be interpreted, however, as a "time varying" S-A system, with the x_2-coordinate corresponding to time evolution.)

Although cascades of linear systems with polynomial interconnections give rise to reasonably well-behaved maps, it should be remarked that a simple cascade of a linear and an internally biaffine system, with linear interconnections, already may give rise to rather complicated responses. For example, if Σ_1 and Σ_2 are systems with $X=U=Y:=k$, $\bar{x}:=0$, and Σ_1 has $x'=x+u, y=x$, Σ_2 has $x'=xu+x+u$, $y=x$, the response of $\Sigma_1\cdot\Sigma_2$ does not admit any polynomial realization with desirable observability properties (see [46, example (18.1)]). In general, cascades of polynomial S-A systems with polynomial interconnections give rise to responses for which the degree of f_t in each fixed u_i, $1\leqslant i\leqslant t$, grows polynomially in t. (So, for example, the response of $x'=x^2u$, $\bar{x}=1$, $y=x$, cannot be realized in that way.) Cascades of S-A systems can be characterized slightly more abstractly introducing some further concepts.

Definition 3.14: The *graph* $G(\Sigma)$ of a system Σ with $X=k^n$ and equations $x_i=P_i(x,u)$, $y=Q(x,u)$, has n nodes $\{1,\cdots,n\}$ and an arc from i into j iff P_j is a nonconstant function of x_i. The system Σ has *no nonlinear feedback* iff the following condition holds for each $i=1,\cdots,n$: if $j_1,\cdots,j_r$ are the nodes which can be reached from i by a path in $G(\Sigma)$, then P_i is (jointly) affine in $x_{j_1},\cdots,x_{j_r}$.

Proposition 3.15: The following statements are equivalent for a response f: a) f is realizable by a cascade of (finite-dimensional) state-affine systems, b) f is realizable by a cascade of (finite-dimensional) internally biaffine systems, and c) f is realizable by a system with no nonlinear feedback.

Proof: It is trivial to prove that b) implies a) implies c). To prove that a) implies b), note that any type-J state affine $x'=\Sigma\delta_i(u)P_i(x)$, $y=\Sigma\delta_i(u)Q_i(x)$ can be written as a cascade of a memory-free system $Q(0,u)=(\delta_0(u),\cdots,\delta_s(u))$ and the internally biaffine system $x'=\Sigma u_{(i)}P_i(x)$, $y=\Sigma u_{(i)}Q_i(x)$. To prove that c) implies a), assume that f is realizable by Σ as in (2.31) and define an equivalence relation on $\{1,\cdots,n\}$ by: i equivalent to j iff both i is reachable in $G(\Sigma)$ from j and vice versa. Let "$\leqslant$" be the partial order induced on equivalence classes by: $C_i\leqslant C_j$ iff nodes in C_j can be reached from the ones in C_i. Extend $\leqslant$ to a total order, and relabeling classes if necessary, suppose $C_1\leqslant C_2\leqslant\cdots$. Letting X_i: = subset of variables with indexes in C_i, it follows that 1) P_i is constant in the variables in $X_j, j>i$, and 2) P_i is affine on the variables in X_i. A permutation of the variables gives then the required decomposition.

B. Non-S-A Realizability of Bounded Responses

It is conceivable that an analytic bounded response f not be S-A finitely realizable but that there may exist a more general, for instance analytic, system Σ realizing f. In fact, this may happen even for a linear response, as in (3.12). It is clear from (3.11) however, that for a polynomial finite f, Σ analytic *and EIS*, f is S-A finitely realizable. Although the method of proof utilized in (3.11) does not generalize to the bounded *non*finite case, the result will still be true, as proved below.

Definition 3.16: For each w in U^+ the corresponding *observable* f^w is the map $f^w(v):=f(vw)$. The *observation space* L^f of f is the affine span of $\{f^w, w \text{ in } U^*\}$ in $[U^+,Y]$.

The observation space is an object dual to the span-canonical state space L_f generated by the f_w (recall $f_w(v)=f(wv)$). More precisely, if $\hat{L}_f$ (resp. $\hat{L}^f$) is the subspace of $[U^+,Y]$ generated by the f_w (resp. f^w) then the nondegenerate bilinear map

$$\hat{L}_f\times\hat{L}_f\to Y \tag{3.17}$$

defined on generators by

$$(f_w,f^v)\mapsto f(wv)$$

induces one-to-one linear maps $\hat{L}_f\to(\hat{L}^f)'$ and $\hat{L}^f\to(\hat{L}_f)'$ (where M'=space of all linear $M\to Y$). Thus if Y is finite dimensional (so that M' is finite dimensional for any finite dimensional M), $\hat{L}_f$ is finite dimensional iff $\hat{L}^f$ is. Since L_f (resp. L^f) has finite dimension iff $\hat{L}_f$ (resp. $\hat{L}^f$) does (e.g., $\dim L_f\leqslant\dim\hat{L}_f\leqslant\dim L_f+1$), the following lemma is concluded.

Lemma 3.18: The response f, with $Y=k^p$, is S-A finitely realizable iff L^f, or equivalently $\hat{L}^f$, is finite dimensional.

The *restriction* $\hat{L}^f_t$, $t=1,2,\cdots$ is the subspace of $[U^t,Y]$ generated by $\{f^w_t, w \text{ in } U^*\}$. In other words if $R_t:[U^+,Y]\to[U^t,Y]$ is the restriction operator $b\mapsto b_t$, then $\hat{L}^f_t=R_t(\hat{L}^f)$. When f is a bounded response of type $J=\{\delta_0,\cdots,\delta_s\}$ all $\hat{L}^f_t$ are finite dimensional, since $\hat{L}^f_t$ is included in the space generated by all δ_α, α in $[J]^t$. Thus if some R_t is one-to-one on $\hat{L}^f$, the latter is finite dimensional. The kernels K_t of the $R_t|\hat{L}^f$ give subspaces $K_1\supseteq K_2\supseteq\cdots$ the intersection of all of which is zero.

Lemma 3.19: L^f is finite dimensional iff there is some t with $K_t=K_{t+1}$.

Proof: If L^f is finite dimensional, the chain of subspaces $K_1,K_2,\cdots$ must be finite.

Conversely, it is enough to prove that $K_t=K_{t+1}$ implies $K_{t+1}=K_{t+2}$; an inductive argument will then give that all K_i, $i>t$ are equal to K_t and hence K_t is their intersection, i.e., zero, so R_t is one-to-one on $\hat{L}^f$. Let then $\Sigma r_jf^{w_j}$ be in K_{t+1}. Thus $\Sigma r_jf^{w_j}(v)=0$ for all v in U^{t+1}. So $\Sigma r_jf^{uw_j}(w)=\Sigma r_jf^{w_j}(uw)=0$ for all u in U and w in U^+. Thus $\Sigma r_jf^{uw_j}$ is in K_t, and hence also in K_{t+1}, for all u in U. So $\Sigma r_jf^{w_j}(u\hat{w})=\Sigma r_jf^{uw_j}(\hat{w})=0$ for all u in U and w in U^{t+1}, i.e., $\Sigma r_jf^{w_j}$ is also in K_{t+2}, as wanted.

Theorem 3.20: Assume that f is a bounded response having an EIS realization Σ such that either a) Σ is analytic and U is connected, or b) Σ is polynomial and U is an irreducible algebraic set (i.e., a subset of k^m, defined by polynomial equations, which cannot be written as a union of two proper such sets; for instance $k=\boldsymbol{R}$, $U=\boldsymbol{R}^m$); then f is S-A finitely realizable.

Proof: Both cases will be proved by using Lemma (3.19). To prove a), consider the t-step reachability maps $g_t:U^t\to X$. The *rank* r_t of g_t will mean the maximal possible value of the differential dg_t of g at all possible points w in U^t. Since all $g(U^t)\subseteq g(U^{t+1})$, it follows that $r_t\leqslant r_{t+1}$ for all t. Thus there is some $t\leqslant n=\dim\Sigma$ with $r_t=r_{t+1}=r$. Let w in U^t be such that rank $dg_t(w)=r$. Since g_t is the composition of g_{t+1} and the map $U^t\to U^{t+1}:v\mapsto\bar{u}v$ ($\bar{u}$=equilibrium input), it follows that rank $dg_{t+1}(\bar{u}w)=r$. It follows from the rank theorem (see, e.g., [12, ch. 8]) that there is an open neighborhood $V=V_1\times V_2$ of $(\bar{u},w)$ in U^{t+1}, and an r-dimensional submanifold M of X such that $g_{t+1}(V)=M$. Since $g_t(V_2)\subseteq M$ and rank $dg_t(w)=r$, by the inverse function theorem V_2 contains an open neighborhood V_3 of w with $g_3(V_3)$ open in M. It follows that $V_4:=g_{t+1}^{-1}(V_3))$ is an open subset of V, and hence of U^{t+1}.

Now let $\Sigma r_if^{w_i}$ be in K_t. Then, for each v in $V_4, g_{t+1}(v)=g_t(\hat{w})$ for some $\hat{w}$ in U^t; thus $\Sigma r_if^{w_i}(v)=\Sigma r_iQ(g_{t+1}(v),w_i)=\Sigma r_iQ(g_t(\hat{w}),w_i)=\Sigma r_if^{w_i}(\hat{w})=0$. This means that $\Sigma r_if^{w_i}$ is zero in the open subset V_4 of the connected set U^{t+1}. Since this map is analytic, it must be zero on all of U^{t+1}, as wanted.

In the polynomial case, it follows from [48, prop. 5.4] that $g_n(U^n)$ and $g_{n+1}(U^{n+1})$ have the same closure in the Zariski topology of X. Thus $\Sigma r_if^{w_i}_t=0$ implies that $\Sigma r_iQ(\cdot,w_i)$ is zero on $g_n(U^n)$ and hence, by continuity, on its closure, which includes $g_{n+1}(U^{n+1})$. Thus $\Sigma r_if^{w_i}_{t+1}=0$, as before.

It is clear that a stronger version of a) above is valid: if Σ is only assumed to have P,Q differentiable, but f is assumed analytic, then f is again S-A finitely realizable.

C. Input/Output Equations

A fundamental result in linear system theory states that a linear response is finitely realizable iff its transfer matrix is rational. In other words, finite realizability is equivalent to inputs and outputs being related by a (high-order) difference equation. This result can be generalized to bounded maps. For simplicity, only polynomial maps will be treated, and $Y=k$ will be assumed. The input set U will be taken to be an irreducible algebraic set (cf. 3.20 b).

Definition 3.21: An (output-affine) *difference equation* (of order r) for the response f is an equation

$$E(y_t,\cdots,y_{t-r},u_t,\cdots,u_{t-r})$$
$$=\sum_{i=0}^{r} b_i(u_t,\ldots,u_{t-r})y_{t-i}+b_{r+1}(u_t,\cdots,u_{t-r})=0 \quad (3.22)$$

with the b_i polynomial in $u_t,\cdots,u_{t-r}$, $b_0\neq 0$, which is understood to hold for all input/output pairs $(u(\ ),y(\ ))$ with $y_t=f(u_1,\cdots,u_t)$, $t>r$. The equation is *output linear* iff $b_{r+1}=0$.

Numerical Example 3.23: Let $U=k$ and f the response of the system

$$x'=x+u, \quad \bar{x}=0$$
$$y=x^2.$$

For any state $x_1=x$ and any input sequence $u_1u_2u_3$ the resulting outputs are $y_1=x^2$, $y_2=(x+u_1)^2$, $y_3=(x+u_1+u_2)^2$. This gives a relation between u_1,u_2,y_1,y_2,y_3 in which no x appears. Indeed,

$$2u_1u_2x=u_2y_2-u_2y_1-u_1^2u_2$$

and

$$y_3=y_2+2x_1u_2+2u_1u_2+u_2^2$$

imply the second order difference equation

$$-u_2y_3+2u_2y_2-u_2y_1-u_1^2u_2+2u_1u_2^2+u_2^3-u_2=0. \quad (3.24)$$

It should be observed that for the f in this example (as in general for nonlinear responses) there exists no possible "regression" $y_t=R(u_{t-r},\cdots,u_t,y_{t-r},\cdots,y_{t-1})$: if there would exist such an R, application of $w=u_1\cdots u_{r+1}$ with $u_1:=1$, $u_i:=0$ for $i=2,\cdots,r$, $u_{r+1}:=1$, and application of $w'=u_1'u_2\cdots u_{r+1}$ with $u_1':=-1$, results in a contradiction in calculating y_{r+1}. Furthermore, note that (3.24) does not uniquely determine f, since the response obtained by beginning in any other initial state also satisfies (3.24).

Theorem 3.25: A polynomial response satisfies an output-affine equation iff it is S-A finitely realizable. If this is the case, f satisfies also an output-linear equation.

("Only if"): For an n-dimensional S-A system Σ there are functions $Q_i^t: U^t\to Y$, $i=0,\cdots,n$, $t=1,2,\cdots$ such that

$$Q^t(x,w)=Q_0^t(w)+Q_1^t(w)x_1+\cdots+Q_n^t(w)x_n \quad (3.26)$$

for all x in X and w in U^t. By abuse of notation, $Q_i^t(u_1\cdots u_s)$ will mean $Q_i^t(u_1\cdots u_t)$ if $t<s$. When Σ is polynomial, the Q_i^t are all polynomial. Denoting $Q_t:=(Q_0^t,\cdots,Q_n^t)$, the $n+2$ vectors $Q_1,\cdots,Q_{n+2}$ all belong to the $(n+1)$-dimensional vector space K^{n+1} over the field K of rational functions on U^{n+2} (irreducibility of U implies that K is well defined). Thus there exists a relation among these vectors, or componentwise:

$$\sum_{j=0}^{n+2} b_j(w)Q_i^{n+2-j}=0, \quad i=0,\cdots,n \quad (3.27)$$

with the $b_i(w)$ rational and not all $b_i=0$. Without loss of generality, $b_0\neq 0$. Multiplying by a common denominator, the $b_i(w)$ may be assumed polynomials. From (3.26) and (3.27),

$$\sum_{j=0}^{n+2} b_j(w)Q^{n+2-j}(x,w)=0 \quad (3.28)$$

for all x in X and w in U^{t+2}. Let $r:=n+2$. Then if $u_1\cdots u_t$ is any input sequence with $t>r$, $y_j=Q^j(\bar{x},u_1\cdots u_t)=Q^j(P(\bar{x},u_1\cdots u_{t-r-1}),u_{t-1}\cdots u_t)$ for $j=t-r,\cdots,r$. So (3.28) becomes $\sum b_j(u_{t-r}\cdots u_t)y_{t-j}=0$, an output-linear equation as wanted.

("If"): Assume that f satisfies an output-affine equation. In terms of the observables f^w, (3.22) translates into the statement

$$b_0(w)f^w=\sum_{i=0}^{r-1} b_{r-i}(w)f^{u_1\cdots u_i}+b_{r+1}(w) \quad (3.29)$$

for all $w=u_1\cdots u_r$ in U^r. Let V be the subset of all w in U^r with $b_0(w)\neq 0$ and S the subspace of $[U^+,Y]$ generated by all $\{f^w,|w|<r\}$ and the constant function 1. The space S is finite dimensional. To see this, it is sufficient to prove that each subspace S_t generated by those f^w with w in U^t is finite dimensional, for all t. Indeed, for each fixed t and v in U^+, w in U^t, $f^w(v)=f(vw)$ is a polynomial in the variables w (because, by induction on (3.22), $f_t(u_1\cdots u_t)$ has a degree in u_{t-j} bounded independently of j,) so there are functions $d_\alpha: U^+\to Y$ with

$$f^w=\sum \delta_\alpha(w)d_\alpha \quad \text{(finite sum)} \quad (3.30)$$

for linearly independent monomials δ_α. Then the d_α generate S_t.

To complete the proof, it will be enough to show that S_r is included in S, since then an inductive argument gives that $S_{r+1}, S_{r+2},\cdots$ are all in S, hence that L^f in included in S and is therefore finite dimensional. But (3.29) says that f^w is in S whenever v is in V. Thus it will be enough to prove that the span of the $\{f^w,w \text{ in } V\}$ is all of S_r. Let d_α,δ_α be as in (3.30), for $t=r$. Then the span of the $\{f^w,w$ in $V\}$ already includes all the d_α: if this were not the case, there is an α_0 and a linear $T:[U^+,Y]\to k$ with $T(d_{\alpha_0})\neq 0$ but $T(f^w)=0$ for all w in V. So

$$0=T\left(\sum \delta_\alpha(w)d_\alpha\right)=\sum \delta_\alpha(w)T(d_\alpha) \quad (3.31)$$

for all w in the open dense set V. By continuity (Weyl's principle), (3.31) holds for all w in U^r. But this contradicts the linear independence of the δ_α, since $T(d_{\alpha_0})\neq 0$.

Numerical Example 3.23: Applying the above procedure in (3.23) results in an output-linear equation

$$b_0y_4+b_1y_3+b_2y_2=0$$

where $b_0=(u_1+y_2)^2-(u_1+u_2)$, $b_1=u_1(u_1+u_2+u_3)-(u_1+u_2+u_3)^2$, and $b_2=b_1-b_0$.

Remark 3.33: It is easy to generalize (3.25) to the case $Y=k^p$. Then finitely realizability becomes equivalent to the existence of an equation as in (3.22) with all b_i being now p by p matrices, $i=0,\cdots,r$, and b_{r+1} a p-vector, det $b_0 \neq 0$. It is also possible to give generalizations to the analytic case (U connected, K = field of meromorphic functions) and to the case of algebraic input/output difference equations which are not output-affine (see [42, section 16], [46]).

IV. Discussion

The abstract realization theory in the first part of this paper, including the construction of the span-canonical state-space L_f, is mostly a rather immediate adaptation of standard automata-theoretic constructions, as found for instance in Padulo and Arbib [36, ch. 8]. Although the terminology "state-affine system" was apparently first used in Sontag and Rouchaleau [48], these models had already appeared in the literature under the name "variable structure systems." The notion of bounded response and the terminology "span-canonical" were introduced in [43] and [44] but the latter concept—and its importance for uniqueness—was well known (see Brockett's paper [4]). The uniqueness part of (1.11) was proved in [48, corollary 8.4] as a corollary of a more general uniqueness result (statements are made there in the general context of EIS, but this hypothesis is not needed in proving them (see [42] and [45]).

Mathematically, the state-linear realization problem has the same structure as the question of representing power series in noncommuting variables. This notion was introduced by Schutzenberger [40] as a generalization of automaton-theoretic ideas, and has been rediscovered since by many authors, notably in the context of stochastic automata. Representations are called *sequential systems* by Turakainen [53], *generalized linear automata* by Muchnik [35], and *automata with multiplicities* by Eilenberg [14]. The Hankel matrix results for noncommutative power series were obtained by Fliess [15] and [17] as a generalization of well-known linear-system results. As an application of his own results on representations, Muchnik [35] seems to mention in passing internally bilinear systems, but does not develop his ideas further. Independently, Fliess [16], [18], and [19] discovered, and carried out in detail, this application to internally bilinear systems. Simultaneous with this, Isidori [25] arrived at a similar algorithm for internally bilinear systems. The matrix procedure given here should be regarded as a generalization of the last two references. The reduction to the equilibrium case allowed reducing the complexity of this generalization, but it is clear that a matrix procedure could be also given directly; the formulas then resemble those given by Tarn and Nonoyama [52] for internally bi-affine systems.

The results in Section III (except (3.10), (3.11)) were announced in Sontag [43] and [44] and proved in Sontag [42]. A (weaker) analytic case of Theorem (3.5), and Corollary (3.11), can be obtained also as immediate consequences of the work done independently by Gilbert [22], [24], who uses a totally different approach, based on variations of f, which results in a non-S-A cascade. The statement of (3.11) was suggested by an analogue in continuous-time proved by Brockett [5]; the proof in continuous-time is based upon factorizations (of Volterra kernels) whose existance is due to reversibility properties of *differential* (as opposed to difference) equations.

Various questions can be raised regarding the suitability of a S-A realization theory, even in the bounded case. Although for equilibrium responses, bounded implies (3.20) S-A realizability, it was already remarked that lower dimensional representations will in general result when more general classes of systems are considered; a tradeoff between dimensionality and complexity of transition and output maps is often involved. S-A realizations have an obvious advantage from an analysis viewpoint—the more general polynomial theory in [42], for example, lacks the straightforward algorithms found here. From a control-theoretic viewpoint, on the other hand, S-A realizations do not have desirable controllability properties. It is interesting to speculate on the impact of microprocessor technology, which may very well render attractive the concept of a high-order realization where each component (state-variable) computes a relatively simple function, as with S-A systems.

Regarding alternative system configurations, it should be pointed out that the "naive" definitions of analytic and polynomial systems should be in general refined in order to account for state-space constraints (e.g., X = manifold, algebraic variety, etc.). An interesting question regarding such realizations is that of constructing minimal ones (with respect to those classes). Although an abstract treatment of this problem can be given, only in a few very special cases are there *constructions* of minimal polynomial realizations: the linear case, the bilinear case (Kalman [28], [29] and Pearlman [38]), and the degree two homogeneous case (Gilbert [23]). (By contrast, in the *continuous*-time, finite Volterra series case, the recent work of Crouch [10]—proving that the canonical realizations of Sussman [51] are Euclidean—may eventually provide effective minimal realization procedures.)

There are many open problems suggested by this work. No topological considerations have been made except at various technical points. Certain statements can be proved easily (e.g., if U is a topological space, $K = \boldsymbol{R}$ or $\boldsymbol{C}$, and each f_t is continuous, then the image realization has P, Q separately continuous), but deeper issues still await a careful study. Somewhat related are questions of approximation. It follows from the Stone–Weirstrass theorem that, if each f_t is continuous and U is compact, there are S-A finitely realizable responses arbitrarily approximating f on finite-time intervals, in fact, a linear system cascaded with a memory-free one suffice for this purpose (this is, via the bilinearization of Brockett [4], the gist of the continuous-time results of Fliess [18] and Sussmann [50]). For the more important case of infinite-interval approximations, however, it appears that results like (3.11) will be more relevant. We conjecture (see [46], for instance, that truncations of a finitely realizable response f can be themselves "realized approximately" (in some precise sense) by systems of dimension $n+1$, where n is the dimension of the canonical realization of f.

It is interesting to note that in the case $k=$finite field *any* response with $U=k^m$ is polynomial, and in fact bounded. This brings up the possibility of applying the results above to the state-assignment problem for automata; the research here remains to be done. A possible generalization deals with $k=$ring (e.g., the integers); some preliminary results are given in Fliess [17] and Sontag and Rouchaleau [49] which are applicable to the internally bilinear case.

The results on realization may be extended to the consideration of S-A systems in which P, Q depend explicitly of time; with was done by Fliess [18] for internally bilinear systems. A deeper understanding will have to make explicit consideration of the allowed time-variations.

Another set of open problems deals with "real-time" identification and realization, when no "resetting" to initial states is needed; some work in this direction can be found in Sontag [47].

Applications of results and methods of (a previous version of) this paper can be found in Marcus [33] and Kamen [31].

Acknowledgment

Most of the material in this paper was part of the author's doctoral dissertation under Prof. R. E. Kalman, who arranged for long-term support making it possible, and who directed the author's attention to the work of M. Fliess and others. Many thanks are also due to the referees of this paper, who suggested the detailed exposition of the nonstrictly causal and/or nonequilibrium cases, as opposed to just mentioning these as possible generalizations of the basic framework, and who also suggested proving in detail (3.20) and (3.25), independently of the material in [42].

References

[1] P. Alper, "A consideration of the discrete Volterra series," *IEEE Trans. Automat. Contr.*, vol. AC-10, pp. 322–327, 1965.
[2] B. D. O. Anderson, M. A. Arbib, and E. G. Manes, "Realization of multilinear and multidecomposable machines," in *Lecture Notes in Computer Science*, vol. 25, Ed. E. Manes. Berlin: Springer, 1975.
[3] M. A. Arbib, "A characterization of multilinear systems," *IEEE Trans. Automat. Contr.*, vol. AC-14, pp. 699–702, 1969.
[4] R. W. Brockett, "On the algebraic structure of bilinear systems," in *Theory and Applications of Variable Structure Systems*, Eds. R. Mohler and A. Ruberti. New York: Academic, 1972.
[5] ——, "Volterra series and geometric control theory," *Proc. IFAC Int. Congress*, Boston, MA, 1975. (Appeared in revised form in *Automatica*, vol. 12, pp. 167–176, 1976.)
[6] ——, "Convergence of Volterra series on infinite intervals and bilinear approximations," Harvard Univ. preprint, 1976.
[7] J. W. Carlyle and A. Paz, "Realizations by stochastic finite automata," *J. Comp. Syst. Sci.*, vol. 5, pp. 26–40, 1971.
[8] S. J. Clancy and W. J. Rugh, "On the realization problem for stationary homogeneous discrete-time systems," *Automatica*, 1978.
[9] ——, "The regular transfer function and bilinear and state-affine realizations for stationary, homogeneous, discrete-time systems," in *Proc. 1978 Conf. Inf. Sci. and Syst.*, The John Hopkins Univ., pp. 167–172, 1978.
[10] P. E. Crouch, "Finite Volterra series," Ph.D. dissertation, Harvard Univ., Cambridge, MA, 1977.
[11] P. D'Alessandro, A. Isidori, and A. Ruberti, "Realization and structure theory of bilinear dynamical systems," *SIAM J. Contr.*, vol. 12, pp. 517–535, 1974.
[12] J. Dieudonnee, *Foundations of Modern Analysis*. New York: Academic, 1969.
[13] L. S. de Jong, "Numerical aspects of realization algorithms in linear system theory," Ph.D. dissertation, Technological Univ., Eindhoven, The Netherlands, 1975. (Also in *SIAM J. Contr.*, vol. 16, pp. 646–665, 1978.)
[14] S. Eilenberg, *Automata, Languages, and Machines*, vol. A. New York: Academic, 1974.
[15] M. Fliess, "Sur certaines familles de séries formelles," These de Doctorat d'Etat, Université Paris VII, 1972.
[16] ——, "Sur la realization des systemes dynamiques bilineaires," *C. R. Acad. Sci. Paris*, Series A, 277, pp. 243–247, 1973.
[17] ——, "Matrices de Hankel," *J. Math. Pures Appl.*, vol. 53, pp. 197–224, 1974.
[18] ——, "Un outil algebrique: Les series formelles noncommutatives," Preprints of *CNR-CISM Symp., Algebraic System Theory*, Udine, Italy. (In *Lecture Notes in Economics and Mathematical Systems*, Eds. G. Marchesini and S. Mitter. Berlin: Springer, 1975, p. 131.)
[19] ——, "Un codage non commutatif pour certains systemès échantillonnes non lineaires," *Inf. Contr.*, 1978.
[20] E. Fornasini and G. Marchesini, "On the internal structure of bilinear input/output maps," in *Geometric Methods in System Theory*, Eds. D. Q. Maine and R. W. Brockett. Boston, MA: Reidel, 1973.
[21] ——, "Algebraic realization theory of bilinear discrete-time input/output maps," *J. Franklin Inst.*, vol. 301, pp. 143–159, 1976.
[22] E. Gilbert, "Functional Expansions for the response of nonlinear differential systems," *IEEE Trans. Automat. Contr.*, vol. AC-20, pp. 909–921, 1977.
[23] ——, "Minimal realizations for nonlinear I-O maps: The continuous-time, two-power case," presented at *Proc. 1976 Conf. Inf. Sci. Syst.*, The John Hopkins Univ., 1977.
[24] ——, "Bilinear and 2-power input-output maps: Finite-dimensional realizations and the role of functional series," *IEEE Trans. Automat. Contr.*, vol. AC-23, pp. 418–425, June 1978.
[25] A. Isidori, "Direct construction of minimal bilinear realizations from nonlinear input/output maps," *IEEE Trans. Automat. Contr.*, vol. AC-18, pp. 626–631, 1973.
[26] ——, "New results on the abstract realization theory of nonlinear input/output functions," *Ricerche di Automatica, vol.* 5, *pp.* 1–10, 1974.
[27] A. Isidori and A. Rugerti, "Realization theory of bilinear systems," in *Geometric Methods in System Theory*, Eds. D. O. Mayne and R. W. Brockett. Boston, MA: D. Reidel, 1973.
[28] R. E. Kalman, "Pattern recognition properties of multilinear machines," presented at *IFAC Symp. Yerevan, Armenian SSR*, 1968.
[29] ——, "On the realization of multilinear response maps," in *Ricerehe di Automatica*. 1979.
[30] R. E. Kalman, P. L. Falb, and M. A. Arbib, *Topics in Mathematical System Theory*. New York: McGraw-Hill, 1969.
[31] E. W. Kamen, "On the relation between bilinear maps and linear 2-D maps," 1978.
[32] G. Marchesini and G. Picci, "Some results on the abstract realization theory of multilinear systems," in *Theory and Applications of Variable Structure System*, Eds. R. Mohler and A. Ruberti. New York: Academic, 1972.
[33] S. I. Marcus, "Discrete time optimal nonlinear estimation," *IEEE Trans. Automat. Contr.*, to be published.
[34] R. R. Mohler, *Bilinear Control Processes*. New York: Academic, 1973.
[35] A. A. Muchnik, "General linear automata," (in Russian) in *Problemy Kibernetiki*. Moscow: Nauka, 1970. (Trans. in *Systems Theory Research*, vol. 23, Ed. A. A. Lyapunov. Consultants Bureau, 1973.)
[36] L. Padulo and M. A. Arbib, *System Theory*. Washington, DC: Hemisphere, 1974.
[37] A. Paz, *Introduction to Probabilistic Automata*. New York: Academic, 1971.
[38] J. G. Pearlman, "Internal properties of multilinear systems," Ph.D. dissertation, Imperial College, London, 1977.
[39] Y. Rouchaleau, "Linear, discrete time, finite dimensional dynamical systems over some classes of commutative rings," Ph.D. dissertation, Stanford Univ., Palo Alto, CA, 1972.

[40] M. P. Schutzenberger, "On the definition of a family of automata," *Inf. Contr.*, vol. 4, pp. 245–270, 1961.

[41] L. M. Silverman, "Realization of linear dynamical systems," *IEEE Trans. Automat. Contr.*, vol. AC-16, pp. 554–567, 1971.

[42] E. D. Sontag, "On the internal realization of polynomial response maps," Ph.D. dissertation, Univ of Florida, Gainesville, FL, 1976.

[43] ——, "On the internal realization of nonlinear behaviors," presented at *Int. Symp. Dynamical Syst.*, Univ. of Florida, Gainesville, FL, in *Dynamical Systems*, Eds. A. Bednarek and L. Cesari. New York: Academic, 1977, pp. 493–497.

[44] ——, "On realizations of discrete-time nonlinear systems," *Notices* A.M.S., vol. 23, p. A-419, 1976.

[45] ——, "Algebraic-geometric methods in realization of discrete-time nonlinear systems," *Proc. 1978 Conf. Inf. Sci. Syst.*, (The John Hopkins University, Baltimore, MD), pp. 159–161, 1978.

[46] ——, *Polynomial Response Maps*. Berlin: Springer, 1979.

[47] ——, "On the observability of polynomial systems, I: Finite-time problems," *SIAM J. Contr., Opt.*, vol. 17, pp. 133–151, 1979.

[48] E. D. Sontag and Y. Rouchaleau, "On discrete-time polynomial systems," Preprints of *CNR-CISM Symp. Algebraic Syst. Theory*, Udine, Italy, 1975. (Appeared in revised form in *J. Nonlinear Analysis, Methods, Theory and Applications*, vol. 1, pp. 55–64, 1976.)

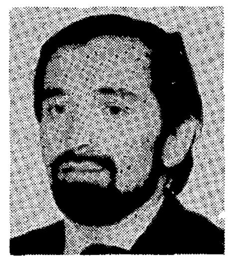
[49] "Sur les anneaux de Fatou forts," *C. R. Acad. Sci. Paris*, vol. 284 A, pp. 331–333, 1979.

[50] H. J. Sussmann, "Semigroup representations, bilinear approximation of input-output maps, and generalized inputs," Preprints of *CNR-CISM Symp. Algebraic System Theory*, Udine, Italy. (Appeared in *Lecture Notes in Economics and Mathematical Systems*, Eds. G. Marchesini and S. Mitter. Berlin: Springer, 1976.)

[51] ——, "Existence and uniqueness of minimal realizations of nonlinear systems," *Math. Syst. Theory*, vol. 10, pp. 263–284, 1976.

[52] T. J. Tarn and S. Nonoyama, "Realization of discrete-time bilinear systems," *Proc. IEEE Conf. Decision and Control* (Clearwater, FL), pp. 125–133, 1976.

[53] P. Turakainen, "On the minimization of linear space automata," *Ann. Acad. Sci. Fen.*, vol. 506-A, pp. 1–14, 1972.



Eduardo D. Sontag was born in Buenos Aires, Argentina, on April 16, 1951. He received the Licenciado degree in mathematics from the University of Buenos Aires, and the Ph.D. degree in mathematics from the Center for Mathematical System Theory at the University of Florida, Gainesville.

From 1976 to 1977, he held a postdoctoral fellowship at the University of Florida, Gainesville. In 1977, he was appointed Assistant Professor of mathematics at Rutgers—The State University, New Brunswick, NJ. He has also worked in nonacademic computer-related positions. His major research interests lie in mathematical system theory, applied algebra, and theoretical computer science. He is the author of various papers in system theory and algebra, and has also written survey papers and a book on artificial intelligence.

Dr. Sontag is a member of Phi Beta Kappa and of the American Mathematical Society.

Letters to the Editor

Planar Reflex Isomorphs of Terakado's Mirror-Antisymmetrical Constant-Resistance Networks

BASIL R. MYERS

***Abstract*—The self-dual one-terminal-pair constant-resistance networks obtained by Terakado's mirror-antisymmetrical disection of two-dimensional regions of resistance are shown to belong to the planar reflex subclass of the class of self-dual maps pioneered by Smith and Tutte.**

I. Introduction

Terakado [1] recently showed how to construct constant-resistance self-dual one-terminal-pair networks by astute "mirror-antisymmetrical" and "four-fold rotationally antisymmetrical" disections of two-dimensional resistance regions. His technique yielded some new configurations [1, figs. 5–7] as well as some well-known classical networks [2, figs. 2–4]. The new configurations are illustrative of a new and large class of constant-resistance networks which can now be synthesized, thus significantly extending the work of earlier contributions (such as [2], [3]).

Manuscript received July 6, 1978.

The author is with the Naval Postgraduate School, Monterey, CA 93940 on leave from the University of Maine at Orono, Orono, ME 04473.

Of the new configurations, we have no further comment on those which he obtained by a four-fold rotationally antisymmetrical disection (typified by the network of [1, fig. 5]). Regarding those which he obtained by mirror antisymmetrical disections [1, figs. 6 and 7], however, we thought it worthwhile to report that we have observed that this particular subclass of networks actually belongs to the planar reflex subclass of the class of self-dual structures revealed almost three decades ago by Smith and Tutte in a classic article on self-dual maps [4], subsequently investigated further by Benedict and Roe [5], [6], as follows.

II. Planar Reflexes

Tutte has commented [7] that his approach in investigating self-dual structures was to assume that a self-dual network could be transformed into its dual by one of the symmetries of the sphere. We borrow verbatim from Smith and Tutte's [4] illustration of the particular symmetry which is a planar reflex. Fig. 1 (which is [4, fig. 2, p. 195]) shows a planar reflex as seen from above the "north pole" of the sphere. The solid lines in the figure represent the part of the planar reflex in the northern hemisphere, the broken lines the part in the southern hemisphere. We observe that a planar reflex is thus a self-duality-preserving reflection about the equatorial plane of the sphere.

A simple example of a planar reflex is the wheel graph of order 4, Fig. 2(a), which is the graph of the familiar Wheatstone

0098-4094/79/0500-0356$00.75 ©1979 IEEE