593

INTERNAL DESCRIPTION OF MULTILINEAR SYSTEMS

by

JOSEPH GERSON PEARLMAN

A Thesis submitted for the degree
of Doctor of Philosophy

April 1977

Department of Computing and Control
Imperial College of Science and Technology
University of London

To my parents

ABSTRACT

This thesis is a contribution to the field of multilinear system theory, and investigates the state space realization and stability of dynamical systems characterized by multilinear input/output maps.

A summary of research done to date in this field is presented, together with a number of original results. The principal work which has been carried out in recent years has been for the case of bilinear input/output maps, where necessary and sufficient conditions for such a map to be realizable in finite-dimensional state space form have been obtained. A major contribution of this thesis is the determination of necessary and sufficient conditions for a realization of such a map to be observable and quasi-reachable, and of reduction procedures for obtaining a realization which is quasi-reachable and observable from one which is not. Previous thoughts and ideas on constructing realizations direct from the transfer function (notably by Kalman [K1]) are formalized here, and sufficient conditions for stability of the output sequence due to finite length input sequences are demonstrated.

Multi-output bilinear systems are examined separately, as these require relaxation of the idea of observability to that of quasi-observability, and although conditions for quasi-reachability and quasi-observability are obtained, together with a reduction procedure for quasi-reachability, the results are not quite as definitive as those for single output bilinear systems.

Sufficient conditions for stability and state space quasi-reachability of a particular class of multilinear input/output maps are shown, and necessary conditions are obtained in terms of the input-to-state transfer functions for the state space realization of a general multilinear input/output map to be quasi-reachable.

ACKNOWLEDGEMENTS

I should most of all like to thank my supervisor, Dr M. J. Denham, for his help and encouragement during the course of this research.

Further thanks are due to the members of staff of the Department of Computing and Control for their interesting and inspiring lecture courses during my time in the department, and also to the students there, in particular Robin Becker, Jose Farah and Richard Lewis, for stimulating discussions on all aspects of control theory and mathematics.

I am grateful to the Science Research Council for their monetary support during most of the period of my research and to my present employers, Rothamsted Experimental Station, for letting me have the odd half-day off to complete this thesis.

A final thank you goes to Mrs Joan Jelinek, for her expert typing of this thesis, and for her painstaking deciphering of my cramped handwriting.

CONTENTS

NOTATION

The following standard notation is used in this thesis, primarily in Chapter 2:

Z	— the set of negative integers including zero
N	— the set of natural numbers
$R[x]$	— the ring of polynomials in x
$R[[x]]$	— the ring of power series in x
$R[x,y]$	— the ring of polynomials in x and y
$R[[x,y]]$	— the ring of power series in x and y
$R[(x)]$	— the ring of rational power series in x
$R[(x,y)]$	— the ring of rational power series in x and y
$R^{real}[(x,y)]$	— the subring of $R[(x,y)]$ generated by $R[(x)]$, $R[(y)]$ and $R[(xy)]$ (which consists of power series expansions of rational functions with denominator $p_1(x)p_2(y)p(xy)$).

CHAPTER 1. INTRODUCTION

Ever since the time of Newton and Leibniz and the birth of the calculus there has been a continued interest in the theory of differential equations and dynamical systems, and this interest has been vigorously pursued throughout the intervening years from then until now. Great strides have been made, particularly in the area of linear system theory, including linear control systems and the theory of state space realizations. Non-linear system theory has remained somewhat intractable however, and apart from specific examples is usually dealt with by means of approximations. One form of approximation which is applied to non-linear control systems provides a motivation for this thesis — the Volterra series approximation. This is based on the ability to write the solution to the differential equation

$$y^{(n)}(t) = g(y, y', \ldots, y^{(n-1)}, t, u(t))$$

where $y^{(i)} = \frac{d^i y}{dt^i}$

as the infinite series

$$y(t) = \sum_{j=1}^{\infty} \int_0^t \cdots \int_0^t h_j(t; t_1, \ldots, t_j) u(t_1) \ldots u(t_j) dt_1 \ldots dt_j.$$

This is called the Volterra series, and a Volterra series approximation is obtained by truncating this series so that summation is from 1 up to some integer n.

These approximations were studied in detail during the late 50s and early 60s, by a number of researchers at M.I.T., prominent among whom were Wiener [W1], Lee and Schetzen [LS1], Bush [BU1], and George [GE1].

Various techniques were invented in the time invariant case for dealing with these approximations; notable among these are the multiple Laplace transform operators $s_1, \ldots, s_n$ of [GE1], and the determination of the kernels $h_j(t-t_1, \ldots, t-t_j)$ by means of suitable 'white' noise or pseudo-random inputs [LS1], [BO1].

Similar operator techniques were applied by Alper [AL1] to Volterra series representations of discrete-time input/output maps, which have the form

$$y_{k+1} = \sum_{r=1}^{\infty} \sum_{i_1=0}^{k} \sum_{i_r=0}^{k} w_r(k-i_1,\ldots,k-i_r)\ u_{i_1}\ldots u_{i_r} \tag{1.1}$$

and this approach was developed further by Kalman, whose 1968 paper [K1] provides the first study of multilinear machines. One of the purposes of this study was to investigate in depth the properties of the Volterra kernels w_r, and to facilitate this he examined the system governed by the kernel w_r, but with the input channels all distinct. In the bilinear case, for the kernel w_2, this results in the so-called bilinear input/output map

$$y_{k+1} = \sum_{i_1,i_2=0}^{k} w_2(k-i_1,k-i_2)\ u_{i_1}v_{i_2} \tag{1.2}$$

where we no longer have the constraint $u_i = v_i$ $(i=0,\ldots,k)$ as in (1.1).

As with linear input/output maps, various questions can be asked about (1.2), in particular questions concerning state space realizability, and the answer as to what form the realization should take was given in [K1] and by Arbib [A1]. This was taken further by Fornasini and Marchesini [FM1], [FM2] who derived necessary and sufficient conditions for finite state realizability (i.e. conditions for writing (1.2) in state space form, where the dimension of the state space is finite) in terms of the transfer function description of (1.2).

However there still remained several other problems associated with state space descriptions of (1.2); in particular, when is such a realization minimal, controllable, observable? How can one obtain a minimal realization from a non-minimal one? While the ultimate objective, to characterize (1.1) via some 'nice' state space description, has not been fulfilled in this thesis, those questions concerning state space descriptions of (1.2) have now been answered, and the results can perhaps be extended to the analysis of (1.1).

Various other problems have been thrown up by (1.2) such as minimal realizations when there is more than one output, and, rather surprisingly, it has so far proved impossible to provide the same definitive results which were found for the single output case. The principal deficiency is that although the results on reachability and observability are similar to those for single outputs, there has as yet been little success in establishing an isomorphism theorem for minimal realizations in the multi-output case. An interesting analogy here is with observability of single output cascaded linear systems (or dually, controllability of single input cascaded linear systems), where fairly straightforward conditions for observability can be established; these conditions do not hold in the multi-output case and any conditions in this case are far more complicated than for single outputs. This analogy is mentioned again in the Appendix to this thesis.

Of additional interest in the context of Volterra series expansions of non-linear input/output maps are the higher-dimensional analogues of (1.2), i.e.

$$y_{k+1} = \sum_{i_1,\dots,i_n=0}^{k} w_n(k-i_1,\dots,k-i_n)u_{i_1}\dots u_{i_n}. \qquad (1.3)$$

It appears that state space descriptions of such multilinear input/output maps have even deeper structure than those of bilinear input/output maps, and although conditions for reachability can be obtained for certain classes of multilinear maps, the problem of minimal realizations of general multilinear maps still remains unsolved. Indeed, whereas for example it has been shown in [FM1] that all bilinear input/output maps with finite-dimensional state space representation can also be represented by some transfer function $N(z_1,z_2)/p_1(z_1)p_2(z_2)p(z_1z_2)$ where N, p_1, p_2 and p are all polynomials with the indicated arguments, a similar result does not hold in the multilinear case; in particular for $n=3$, the set of transfer functions with finite dimensional state-space representations contains the ring of transfer functions of the form

$$N(z_1,z_2,z_3)/p_1(z_1)p_2(z_2)p_3(z_3)p(z_1z_2z_3),$$

but is in turn strictly contained in the ring of transfer functions of the form

$$N(z_1,z_2,z_3)/p_1(z_1)p_2(z_2)p_3(z_3)p_{12}(z_1z_2)p_{23}(z_2z_3)p_{13}(z_1z_3)p(z_1z_2z_3).$$

Research on the continuous time analogues of (1.1), (1.2) and (1.3) has also been undertaken, but the only significant result has been to establish classes of similarity transformations on the state space representations which preserve the input/output map.

Contents of Thesis and Original Contributions

Chapter 2

This begins with a summary of the work done in [FM1] introducing the reader to bilinear input/output maps, in particular to the necessary and sufficient conditions for a bilinear input/output map to have a finite-dimensional state space representation. Some of the proofs of [FM1] are expanded upon for the sake of clarity, and the errors in those proofs are corrected; in addition, Lemma 2.2.2, which is required for the proof of finite-dimensionality, is proved in an apparently original way and simultaneously provides a matrix representation of the so-called ring of recognizable series for commuting operators. This lemma was originally proved in [F1] for the more general case of non-commutative operators, but that proof did not entail the construction of the matrix representations supplied here. New alternative methods of realizing bilinear input/output maps are presented in §2.4, and these in effect formalize the ideas of Kalman [K1], and set the scene for the results of Chapters 3 and 4; a more general state space representation of bilinear input/output maps than that of [FM1] is presented, and Theorem 2.4.1 demonstrates how to compute the transfer function corresponding to it. Sufficient conditions for stability of the output sequence due to a finite input sequence are obtained in Theorem 2.5.1, and this uses the result of Lemma 2.2.2.

Chapter 3

The whole of this chapter is original, and begins with definitions of quasi-reachability, observability and canonical and (co-)minimal realizations, and a presentation of the class of similarity transformations on the state space representations of bilinear input/output maps which preserve the nature of these maps. The remainder of the chapter is devoted to lengthy proofs of the necessary and sufficient conditions for such state space representations to be quasi-reachable and observable.

Chapter 4

This chapter is also completely original, and demonstrates procedures for reducing a realization which is not quasi-reachable or observable to one which is. In addition it is shown that reduction to quasi-reachable form reduces the dimension of the state space and that reduction to observable form at least does not increase the dimension. It is then apparent that all bilinear input/output maps can be represented by a canonical (i.e. observable and quasi-reachable) realization, and in addition it is shown that all such canonical realizations are isomorphic under the similarity transformations introduced in Chapter 3, and hence are minimal. Some canonical forms for these state space realizations are also presented.

Chapter 5

The results of this chapter are again original, although not as definitive as those of the two preceding chapters. Discussion centres on state space realizations of multi-output bilinear systems, and a specific example is used to illustrate that it may not always be possible to construct a realization which is both observable and quasi-reachable, and it is therefore necessary to introduce the new concept of quasi-observability, analogous to quasi-reachability. It is then possible to construct a realization which is quasi-reachable and quasi-observable, but it has not yet been possible to provide the conditions for such a realization to be minimal, in the sense that any two quasi-reachable and quasi-observable realizations are isomorphic under some class of transformations.

Chapter 6

This chapter begins with two new results on quasi-reachability and stability for a particular class of multilinear input/output maps. The particular class of maps considered are those whose transfer functions have denominators which can be factorized as $p_1(z_1)\ldots p_n(z_n)p(z_1\ldots z_n)$, and Theorems 6.1.1 and 6.2.1 are a generalization of the earlier results on quasi-reachability and stability for bilinear input/output maps. The chapter then continues with a review of work done to date in the field of multilinear input/output maps; the main contributions are contained in three papers by Kalman [K1], Arbib [A1] and Anderson, Arbib and Manes [AAM1], the last of these analysing the problem from a category-theoretic viewpoint. All of these papers indicate how to set up a state space realization, but fail to tackle the problem of reachability, observability, etc., in a satisfactory way. It is however possible, as is shown for the case of trilinear systems, to provide necessary conditions for quasi-reachability by invoking the idea of linear independence of the input-to-state transfer functions and their Kronecker products, and these necessary conditions can be extended to all multilinear systems.

Chapter 7

This last chapter is the conclusion to the thesis, with suggestions for further work. The major field suggested for further research is that of continuous time bilinear state space realizations, and while no results on reachability and observability have yet been achieved, an original result concerning similarity transformations on these realizations is introduced in Theorem 7.1.

Appendix

For ease of reference this draws together two results on linear system theory obtained earlier in the thesis, and an original result concerning reachability conditions for cascaded linear systems is presented.

CHAPTER 2. BILINEAR INPUT/OUTPUT MAPS

In this chapter we introduce the formal definition of a bilinear input/output map and present necessary and sufficient conditions for the existence of a finite dimensional state space realization. In addition we will see that this is equivalent to the canonical, or Nerode, state space being reachable in finite time, which will tie up with our intuitive notion of what the Nerode state space represents. This of course is a result of our being able to view "state" as a partial memory of past inputs.

The necessary and sufficient conditions for realizability which we shall examine in this chapter were derived by Fornasini and Marchesini [FM1]. However some of their proofs are not clear and we shall present them here in greater detail. In addition we shall present an alternative proof of a theorem by Fliess [F1], which will provide us with a result on stability for bilinear input/output maps. We shall mention this again in Chapter 5, when we study a related stability result for multilinear input/output maps.

2.1 Preliminaries and Definitions

We shall work in the field of real numbers, $\mathbb{R}$, but of course the results will hold over all fields, finite or infinite.

Let U, V and Y denote the following spaces:

$$U = \{u \in \mathbb{R}^{Z-} \text{ with compact support}\}$$

$$V = \{v \in \mathbb{R}^{Z-} \text{ with compact support}\}$$

$$Y = \{y \in \mathbb{R}^{\mathbb{N}-\{0\}}\}$$

where

$Z-$ is the set of negative integers including zero

$\mathbb{N}$ is the set of natural numbers.

U and V will then be made up of sequences of the form $(\ldots,0,u_r,\ldots,u_0)$ and $(\ldots,0,v_s,\ldots,v_0)$ respectively, and Y will be made up of sequences $(y_1,y_2,\ldots)$. $U\times V$ is termed the input space and Y the output space. An input/output function $f : U\times V\to Y$, will then map a finite number of inputs from U and V into an infinite output sequence in Y.

The bilinear nature of the map f is described by the following definition.

Definition 2.1.1

A map $f : U\times V\to Y$ is a bilinear discrete-time stationary input/output map if it satisfies the following conditions:

i) bilinearity —

$$f(k_1u_1+k_2u_2,v_1) = k_1f(u_1,v_1) + k_2f(u_2,v_1)$$

$$f(u_1,k_1v_1+k_2v_2) = k_1f(u_1,v_1) + k_2f(u_1,v_2)$$

for all $u_1,u_2\in U$, $v_1,v_2\in V$ and $k_1,k_2\in R$;

ii) stationarity —

the map f is invariant under translation with respect to time in the following sense:

$$f(\sigma_1u,\sigma_2v) = \sigma^*f(u,v) \tag{2.1.1}$$

where σ and σ^* represent delay operators as follows:

$$\sigma_1(\ldots,0,u_i,\ldots,u_0) = (\ldots,0,u_i,\ldots,u_0,0) \tag{2.1.2}$$

$$\sigma_2(\ldots,0,v_j,\ldots,v_0) = (\ldots,0,v_j,\ldots,v_0,0) \tag{2.1.3}$$

$$\sigma^*(y_1,y_2,y_3,\ldots) = (y_2,y_3,\ldots). \tag{2.1.4}$$

It now becomes apparent that we can identify $U\times V$ with $R[z_1]\times R[z_2]$, where $R[z]$ is the ring of polynomials in z, and that we can identify Y with $z^{-1}R[[z^{-1}]]$, the ring of formal power series in the one indeterminate z^{-1}. This we can do via the isomorphisms

$$\psi_1 : U\to R[z_1]$$

$$: (\ldots,0,u_r,\ldots,u_0)\to u_rz_1^r + \ldots + u_0$$

$$\psi_2 : V \to R[z_2]$$
$$: (\ldots,0,v_s,\ldots,v_0) \to v_s z_2^s + \ldots + v_0$$
$$\psi_3 : Y \to z^{-1}R[[z^{-1}]]$$
$$: (y_1,y_2,\ldots) \to z^{-1}y_1 + z^{-2}y_2 + \ldots$$

and it is clear that we can in addition identify σ_1 with z_1, σ_2 with z_2 and σ^* with z, although for the z mapping we must also include the operation of omitting any term involving non-negative powers of z, i.e. $z(z^{-1}y_1 + z^{-2}y_2 + \ldots) \triangleq z^{-1}y_2 + \ldots$.

We shall now find that because of the bilinearity of the map f we shall be able to identify f with a "causal" power series $s(z_1,z_2) \in (z_1z_2)^{-1}R[[z_1^{-1},z_2^{-1}]]$, and then z will be equal to z_1z_2.

Let us consider $f(u,v)$ where $u = (\ldots,0,u_r,\ldots,u_0)$ and $v = (\ldots,0,v_s,\ldots,v_0)$. Then because of bilinearity we have

$$f(u,v) = \sum_{i,j} f\big((\ldots,0,u_i,0,\ldots,0),(\ldots,0,v_j,0,\ldots,0)\big)$$
$$= \sum_{i,j} u_iv_jf(e_i,f_j)$$

where $e_i = (\ldots,0,1,0,\ldots,0)$ and $f_j = (\ldots,0,1,0,\ldots,0)$ with a 1 in the $-i$ and $-j$ positions respectively.

Now let $f(e_i,f_j) = (s_{ij}^1,s_{ij}^2,\ldots)$,

i.e. $s_{ij}^1,s_{ij}^2,\ldots$ is the output sequence due to unit inputs at time $-i$ at the U channel and time $-j$ at the V channel. Hence

$$f(u,v) = \sum_{i,j} u_iv_j(s_{ij}^1,s_{ij}^2,\ldots)$$

and operating on this with the delay operator σ^*, we see from (2.1.4) that

$$\sigma^*f(u,v) = \sum_{i,j} u_iv_j(s_{ij}^2,s_{ij}^3,\ldots) \tag{2.1.5}$$

Now from (2.1.2) and (2.1.3) it is clear that

$$\sigma_1e_i = e_{i+1} \text{ and } \sigma_2f_j = f_{j+1}$$

so utilizing equation (2.1.1) we obtain

$$\sigma^*f(u,v) = f(\sigma_1 u,\sigma_2 v)$$
$$= \sum_{i,j} u_i v_j f(e_{i+1},f_{j+1})$$
$$= \sum_{i,j} u_i v_j (s^1_{i+1,j+1},s^2_{i+1,j+1},\ldots). \qquad (2.1.6)$$

Equating coefficients of $u_i v_j$ in (2.1.5) and (2.1.6) we now obtain

$$(s^1_{i+1,j+1},s^2_{i+1,j+1},\ldots) = (s^2_{ij},s^3_{ij},\ldots),$$

i.e. $s^k_{i+1,j+1} = s^{k+1}_{ij}$ for all $i,j \geq 0$, $k \geq 1$.

By induction we see that

$$s^{k+1}_{ij} = s^1_{i+k+1,j+k+1}$$

so we can write

$$f(e_i,f_j) = (s_{ij},s_{i+1,j+1},s_{i+2,j+2},\ldots)$$

where we have written $s_{ij} = s^1_{ij}$ for convenience.

Intuitively this means that the response at time k+1 due to inputs at times -i and -j is equal to the response at time 1 due to inputs at times -(i+k) and -(j+k). Hence

$$f(u,v) = \sum_{i,j} u_i v_j (s_{ij},s_{i+1,j+1},\ldots)$$

so that the output sequence is dependent solely on the values of the input sequence and on the numbers s_{ij} $(i,j \geq 1)$.

It is now apparent that we can identify f with

$s(z_1,z_2) = (z_1z_2)^{-1} \sum_{i,j} s_{ij} z_1^{-i} z_2^{-j}$ where $f : R[z_1] \times R[z_2] \to (z_1z_2)^{-1} R[[(z_1z_2)^{-1}]]$ is defined by

$$f\big(u(z_1),v(z_2)\big) = (z_1z_2)^{-1} \sum_{i,j} s_{ij} z_1^{-i} z_2^{-j} u(z_1)v(z_2) \odot \sum_{k\geq 1} (z_1z_2)^{-k} \qquad (2.1.7)$$

and the Hadamard product $\odot$ is defined by

$$\sum_{i,j} a_{ij} z_1^{-i} z_2^{-j} \odot \sum_{i,j} b_{ij} z_1^{-i} z_2^{-j} = \sum_{i,j} a_{ij} b_{ij} z_1^{-i} z_2^{-j}.$$

Hence the product (2.1.7) just picks out all terms in $(z_1z_2)^{-k}$ from $s(z_1,z_2)u(z_1)v(z_2)$. As an example, we can see that

$$f(e_r,f_s) \equiv f(z_1^r,z_2^s) = (z_1z_2)^{-1} \sum_{i,j} s_{ij} z_1^{-i} z_2^{-j} z_1^r z_2^s \odot \sum (z_1z_2)^{-k}$$
$$= (z_1z_2)^{-1} \sum_k s_{r+k,s+k} (z_1z_2)^{-k}.$$

2.2 Equivalence Relations and Realizable Series

In this section we introduce the concept of Nerode equivalence classes; the intuitive notion for these rests on the fact that an input sequence i_1 to a system effectively generates a "partial memory" of that input sequence within the system, so that the response of the system to any further inputs is dependent on the original input sequence. If another input sequence i_2 generates the same "partial memory", i.e. the response of the system to any further inputs is the same as that for i_1, then i_1 and i_2 are said to be Nerode equivalent.

A standard example of Nerode equivalence is provided by a linear system of the form $f(z)/p(z)$; if we can only make observations of the inputs and outputs of this system after time 0, then the system can only "partially remember" the input sequence prior to time 0. Writing the input sequence $u(z)$ prior to time 0 as $a(z)p(z) + b(z)$ where $\deg b < \deg p$, the system will in effect "remember" $b(z)$, but not $a(z)$.

For a system like the one above we find that the "partial memory" of any input sequence will be the same as the "partial memory" of some input sequence of length less than deg p, and we refer to this as reachability of the Nerode space in time deg p; similarly for non-linear systems we can think in terms of reachability of the Nerode space in bounded time, and the intuitive feeling at this point is that this implies that there exists a finite dimensional state space realization of the system. This feeling is borne out in the case of bilinear systems as we shall see.

We also define three other equivalence relations and show that taken together they are equivalent to the Nerode equivalence relation. It is then shown that the space of equivalence classes generated by these three relations is finite-dimensional if and only if $s(z_1,z_2)$ is a realizable power series, i.e. s can be written as $N(z_1,z_2)/p_1(z_2)p_2(z_2)p(z_1z_2)$, and this is a necessary and sufficient condition for reachability of the Nerode space in bounded time.

Definition 2.2.1

Two input pairs $(u_1,v_1),(u_2,v_2) \in U \times V$ are Nerode equivalent iff $(u_1 \circ u, v_1 \circ v) = f(u_2 \circ u, v_2 \circ v)$ for all $(u,v) \in U \times V$, where supp u = supp v.

The symbol $\circ$ is the concatenation operator, and is defined by

$$(\ldots 0,u_k',\ldots u_0') \circ (\ldots,0,u_k,\ldots,u_0) = (\ldots 0,u_k',\ldots,u_0',u_k,\ldots,u_0).$$

In polynomial notation, the two input pairs are Nerode equivalent if

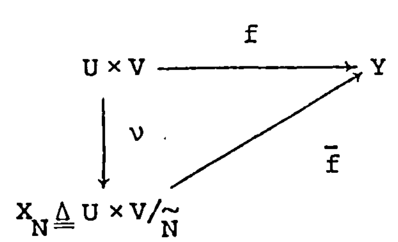
$$f(z_1^{k+1}u_1+u, z_2^{k+1}v_1+v) = f(z_1^{k+1}u_2+u, z_2^{k+1}v_2+v),$$

for all $(u,v) \in U \times V$; $\deg u = \deg v \le k$.

We denote the Nerode equivalence classes by $[u_1,v_1]$ i.e.

$$[u_1,v_1] = \{(u,v) \in U \times V \mid (u,v) \underset{N}{\sim} (u_1,v_1)\}.$$

f can then be factorized as in the following commutative diagram:



where ν is an onto mapping and $\bar{f}$ is (1-1).

$$X_N \triangleq U \times V/\underset{N}{\sim} = \{[u_1,v_1] \mid (u_1,v_1) \in U \times V\}$$

is called the canonical, or Nerode state space.

Definition 2.2.2

X_N is reachable in time m if the mapping ν restricted to input sequences with support less than m is still onto. X_N is said to be reachable in bounded time if it is reachable in time m for some m.

Thus if X_N is reachable in time m, the partial memory which the system has of any input sequence will be the same partial memory that it has of at least one input sequence of length less than or equal to m.

We shall now introduce three more equivalence relations, $\underset{1}{\sim}$, $\underset{2}{\sim}$ and $\underset{3}{\sim}$, which play a major role in what will follow:

$$u_1 \underset{1}{\sim} u_2 \text{ iff } f(u_1 \circ 0^k, 0 \circ v) = f(u_2 \circ 0^k, 0 \circ v)$$
$$\text{for all } k \in N,\ v \in V, \text{ with deg } V < k$$

where $u_1 \circ 0^k$ denotes u_1 followed by k zeros and $0 \circ v$ denotes the zero input sequence followed by v.

$$v_1 \underset{2}{\sim} v_2 \text{ iff } f(0 \circ u, v_1 \circ 0^k) = f(0 \circ u, v_2 \circ 0^k)$$
$$\text{for all } k \in N,\ u \in U, \text{ with deg } u < k$$

$$(u_1, v_1) \underset{3}{\sim} (u_2, v_2) \text{ iff } f(u_1 \circ 0^k, v_1 \circ 0^k) = f(u_2 \circ 0^k, v_2 \circ 0^k)$$
$$\text{for all } k \in N.$$

Remark By stationarity $(u_1, v_1) \underset{3}{\sim} (u_2, v_2)$ iff $f(u_1, v_1) = f(u_2, v_2)$.

The relationship between these equivalence relations and Nerode equivalence is defined by the following lemma.

Lemma 2.2.1

$$(u_1, v_1) \underset{N}{\sim} (u_2, v_2) \text{ iff } u_1 \underset{1}{\sim} v_2,\ v_1 \underset{2}{\sim} v_2,\ (u_1, v_1) \underset{3}{\sim} (u_2, v_2).$$

Proof: Let $(u_1, v_1) \underset{N}{\sim} (u_2, v_2)$, i.e. $f(u_1 \circ u, v_1 \circ v) = f(u_2 \circ u, v_2 \circ v)\ \forall u, v$.

This immediately implies $(u_1, v_1) \underset{3}{\sim} (u_2, v_2)$ by the remark above.

Now $$f(u_1 \circ 0^k, v_1 \circ v) = f(u_1 \circ 0^k, v_1 \circ 0^k) + f(u_1 \circ 0^k, 0 \circ v) \qquad (2.2.1)$$

by bilinearity. Similarly

$$f(u_2 \circ 0^k, v_2 \circ v) = f(u_2 \circ 0^k, v_2 \circ 0^k) + f(u_2 \circ 0^k, 0 \circ v). \qquad (2.2.2)$$

Equating (2.2.1) and (2.2.2), and using $(u_1, v_1) \underset{3}{\sim} (u_2, v_2)$, we obtain

$$f(u_1 \circ 0^k, 0 \circ v) = f(u_2 \circ 0^k, 0 \circ v).$$

Hence $u_1 \underset{1}{\sim} u_2$. We can show in an analogous way that $v_1 \underset{2}{\sim} v_2$.

Conversely, let $u_1 \underset{1}{\sim} u_2$, $v_1 \underset{2}{\sim} v_2$, $(u_1, v_1) \underset{3}{\sim} (u_2, v_2)$.

Then, using bilinearity, we have

$$\begin{aligned} & f(u_1 \circ u, v_1 \circ v) - f(u_2 \circ u, v_2 \circ v) \\ &= f(u_1 \circ 0^k, v_1 \circ 0^k) + f(u_1 \circ 0^k, 0 \circ v) + f(0 \circ u, v_1 \circ 0^k) + f(0 \circ u, 0 \circ v) \\ &- f(u_2 \circ 0^k, v_2 \circ 0^k) + f(u_2 \circ 0^k, 0 \circ v) + f(0 \circ u, v_2 \circ 0^k) + f(0 \circ u, 0 \circ v) \\ &= 0 \end{aligned}$$

where we use $\underset{3}{\sim}$, $\underset{1}{\sim}$ and $\underset{2}{\sim}$ in turn.

Hence $(u_1,v_1) \underset{N}{\sim} (u_2,v_2)$. □

We can now construct the quotient spaces $X_1 = U/\underset{1}{\sim}$ and $X_2 = V/\underset{2}{\sim}$, and these can be endowed with the structure of a linear space. This follows because

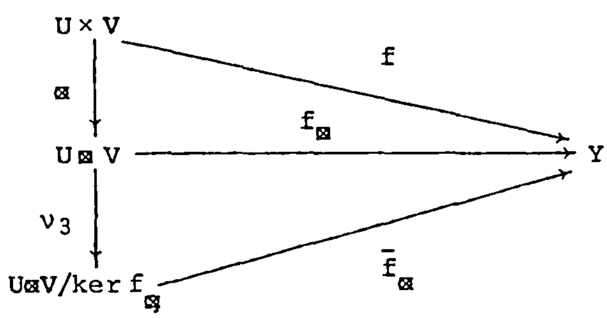
(i) $k[u_1]_{\underset{1}{\sim}} = [ku_1]_{\underset{1}{\sim}}$

(ii) $[u_1]_{\underset{1}{\sim}} + [u_2]_{\underset{1}{\sim}} = [u_1 + u_2]_{\underset{1}{\sim}}$ where $[u]_{\underset{1}{\sim}}$ is the $\underset{1}{\sim}$ equivalence class of u, and similarly for the $\underset{2}{\sim}$ equivalence classes.

However $U \times V/\underset{3}{\sim}$ cannot be endowed with this structure. It is necessary first to embed $U \times V$ in the tensor space $U \otimes V$, where $U \otimes V \cong \mathcal{R}[z_1,z_2]$. We then define the map $f_\otimes : U \otimes V \to Y$, by identifying $f_\otimes$ with $(z_1z_2)^{-1}\sum s_{ij}z_1^{-i}z_2^{-j}$, and its domain with $\mathcal{R}[z_1,z_2]$. That is

$$f_\otimes : \mathcal{R}[z_1,z_2] \to (z_1z_2)^{-1}\mathcal{R}[[(z_1z_2)^{-1}]]$$
$$: w(z_1,z_2) \to (z_1z_2)^{-1}\sum s_{ij}z_1^{-i}z_2^{-j}w(z_1,z_2) \odot \sum (z_1z_2)^{-k}.$$

It is clear that $f_\otimes$ is a linear map and we can therefore write down the commutative diagram:



where ν_3 is onto and $\bar{f}_\otimes$ is (1-1).

It is now immediate that $(u_1,v_1) \underset{3}{\sim} (u_2,v_2)$ iff $f_\otimes(u_1 \otimes v_1) = f_\otimes(u_2 \otimes v_2)$. This is because $f(u,v) = f_\otimes(u \otimes v)$. We can further see that $U \otimes V/\ker f_\otimes \triangleq X_3$ can be endowed with the structure of a linear space.

Before coming to the realizability theorems we shall prove the following technical lemma which will be of use not only for the purposes of studying realizability, but also for discussing stability of bilinear input/output maps and, by an obvious extension, to more general multilinear input/output maps.

Note that this lemma has been proved by Fliess [Fl] for series which are somewhat more general — the so-called recognizable formal series, where z_1 and z_2 do not necessarily commute and where $\mathcal{R}$ is replaced by a semi-ring. However the present statement of the proof is as general as we need it for our purposes, and the proof, though fairly obvious, appears to be original, and also supplies a representation of the matrices and vectors involved.

<u>Lemma 2.2.2</u>

r is a power series in $(z_1z_2)^{-1}\mathcal{R}[(z_1)]\otimes\mathcal{R}[(z_2)]$, i.e. r can be written as $r = M(z_1,z_2)/p_1(z_1)p_2(z_2)$ for some $p_1(x), p_2(x) \in \mathcal{R}[x]$ and $M(z_1,z_2) \in \mathcal{R}[z_1,z_2]$ iff there exists an integer N, vectors $b, c \in \mathcal{R}^N$, matrices $M_1, M_2 \in \mathcal{R}^{N\times N}$ with $M_1M_2 = M_2M_1$, such that $r = (z_1z_2)^{-1}\sum_{i,j\geq 0} c^TM_1^iM_2^jbz_1^{-i}z_2^{-j}$.

Proof: Let

$$
\begin{aligned}
(z_1z_2)r &= \sum_{i,j\geq 0} c^TM_1^iM_2^jbz_1^{-i}z_2^{-j} \\
&= \sum_{i=0}^{\infty}\sum_{j=i}^{\infty} c^TM_1^iM_2^jbz_1^{-i}z_2^{-j} + \sum_{j=0}^{\infty}\sum_{i=j+1}^{\infty} c^TM_1^iM_2^jbz_1^{-i}z_2^{-j} \\
&= \sum_{i=0}^{\infty} c^TM_1^iM_2^iz_1^{-i}z_2^{-i}(I-M_2z_2^{-1})^{-1}b + \sum_{j=0}^{\infty} c^TM_1^{j+1}M_2^jz_1^{-(j+1)}z_2^{-j}(I-M_1z_2^{-1})^{-1}b \\
&= c^T\left(I-M_1M_2(z_1z_2)^{-1}\right)^{-1}(I-M_2z_2^{-1})^{-1}b \\
&\quad + c^TM_1z_1^{-1}\left(I-M_1M_2(z_1z_2)^{-1}\right)^{-1}(I-M_1z_1^{-1})^{-1}b \\
&= c^T\left(I-M_1M_2(z_1z_2)^{-1}\right)^{-1}[(I-M_2z_2^{-1})^{-1} + M_1z_1^{-1}(I-M_1z_1^{-1})^{-1}]b \\
&= c^T\left(I-M_1M_2(z_1z_2)^{-1}\right)^{-1}[I-M_1z_1^{-1} + M_1z_1^{-1}(I-M_2z_2^{-1})] \\
&\qquad (I-M_2z_2^{-1})^{-1}(I-M_1z_1^{-1})^{-1}b \\
&= c^T(I-M_2z_2^{-1})^{-1}(I-M_1z_1^{-1})^{-1}b \\
&= z_1z_2c^T(z_2I-M_2)^{-1}(z_1I-M_1)^{-1}b \\
&\in \mathcal{R}[(z_1)]\otimes\mathcal{R}[(z_2)].
\end{aligned}
$$

Note that all the above equalities are obtained using the fact that M_1 and M_2 commute.

Conversely, let $r \in (z_1 z_2)^{-1} R[(z_1)] \otimes R[(z_2)]$. Then r can be written as

$$r = \frac{1}{p_1(z_1)p_2(z_2)} \sum_{i=0}^{n_1-1} \sum_{j=0}^{n_2-1} \alpha_{ij} z_1^i z_2^j$$

where $p_1(z_1) = z_1^{n_1} + \beta_1 z_1^{n_1-1} + \ldots + \beta_{n_1}$

and $p_2(z_2) = z_2^{n_2} + \gamma_1 z_2^{n_2-1} + \ldots + \gamma_{n_2}$.

Consider now the term $z_1^i/p_1(z_1)$ where $i < n$. This we can write as

$$z_1^i/p_1(z_1) = z_1^{-1} \sum_{r\geq 0} c_{1i}^T A_1^r b_1 z_1^{-r}$$

where $A_1 = \begin{pmatrix} 0 & 1 & & 0 \\ & 0 & \ddots & \\ & & & 1 \\ -\beta_{n_1} & \cdots & & -\beta_1 \end{pmatrix}$ $b_1 = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$ $c_{1i} = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix}$ - i+1th position.

This follows directly from our knowledge of linear systems realization theory.

In a similar way we can write

$$z_2^j/p_2(z_2) = z_2^{-1} \sum_{s\geq 0} c_{2j}^T A_2^s b_2 z_2^{-s}$$

where $A_2 = \begin{pmatrix} 0 & 1 & & \\ & 0 & \ddots & \\ & & & 1 \\ -\gamma_{n_2} & \cdots & & -\gamma_1 \end{pmatrix}$ $b_2 = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$ $c_{2j} = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix}$ - j+1th position.

Hence we can write

$$z_1^i z_2^j/p_1(z_1)p_2(z_2) = (z_1 z_2)^{-1} \sum_{r\geq 0} \sum_{s\geq 0} c_{1i}^T A_1^r b_1 c_{2j}^T A_2^s b_2 z_1^{-r} z_2^{-s}$$

$$= (z_1 z_2)^{-1} \sum_{r\geq 0} \sum_{s\geq 0} c_{1i}^T \otimes c_{2j}^T A_1^r \otimes A_2^s b_1 \otimes b_2 z_1^{-r} z_2^{-s}$$

and it then follows that

$$r = \sum_{i,j} \alpha_{ij} z_1^i z_2^j / p_1(z_1)p_2(z_2)$$

$$= (z_1 z_2)^{-1} \sum_{r\geq 0} \sum_{s\geq 0} c^T A_1^r \otimes A_2^s b_1 \otimes b_2 z_1^{-r} z_2^{-s}$$

where $c = \sum_{i,j} \alpha_{ij} c_{1i} \otimes c_{2j}$.

If we then write $b = b_1 \otimes b_2$, $M_1 = A_1 \otimes I$, $M_2 = I \otimes A_2$, then the lemma is proven, since

$$\begin{aligned} M_1 M_2 &= (A_1 \otimes I)(I \otimes A_2) \\ &= A_1 \otimes A_2 \\ &= (I \otimes A_2)(A_1 \otimes I) \\ &= M_2 M_1. \end{aligned}$$

□

It is obvious from the proof of the above lemma that we can state the following generalization:

<u>Lemma 2.2.3</u>

r is a power series in $(z_1 \ldots z_n)^{-1} R[(z_1)] \otimes \ldots \otimes R[(z_n)]$ iff there exists an integer N, vectors $b, c \in R^N$, matrices $M_1, \ldots, M_n \in R^{N \times N}$ with $M_i M_j = M_j M_i$ for all i,j such that $r = (z_1 \ldots z_n)^{-1} \sum c^T M_1^{i_1} \ldots M_n^{i_n} b z_1^{-i_1} \ldots z_n^{-i_n}$.

Now, following [FM1], we define the power series:

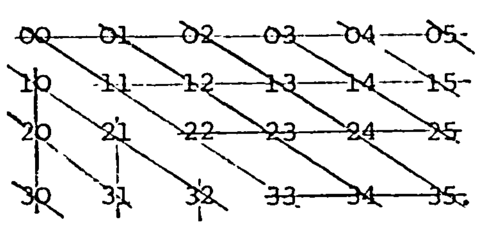
$$r_i(z) = \sum_{k=0}^{\infty} s_{i,i+k} z^{-k} \qquad i = 0,1,\ldots \qquad \text{— row series}$$

$$c_j(z) = \sum_{k=0}^{\infty} s_{j+k,j} z^{-k} \qquad j = 1,2,\ldots \qquad \text{— column series}$$

$$d_{ij}(z) = \sum_{k=0}^{\infty} s_{i+k,j+k} z^{-k} \qquad i,j = 0,1,\ldots \qquad \text{— diagonal series}$$

formed from the general formal power series $s = (z_1 z_2)^{-1} \sum_{i,j} s_{ij} z_1^{-i} z_2^{-j}$.

Diagrammatically, we are doing the summations as shown below:



We also define $R^{real}[(z_1, z_2)]$ as the subring of $R[(z_1, z_2)]$ generated by $R[(z_1)]$, $R[(z_2)]$ and $R[(z_1 z_2)]$, and called the ring of realizable power series.

We next present a result from [FM1], the proof of which seems to be somewhat questionable there, but which is proved correctly here.

Lemma 2.2.4

Let s be a formal power series. Then $s \in \mathcal{R}^{real}[(z_1,z_2)]$ iff the members of the sets $\{r_i\}$, $\{c_j\}$ and $\{d_{ij}\}$ are power series expansions of rational functions in one indeterminate, having a common denominator for each set.

Proof: Let $p_1(z)$, $p_2(z)$, $p(z)$ be the common denominators of the rational functions associated with $\{r_i\}$, $\{c_j\}$ and $\{d_{ij}\}$ respectively.

Now $s = (z_1z_2)^{-1}\sum s_{ij}z_1^{-i}z_2^{-j}$ can be written as

$$s = (z_1z_2)^{-1}\left[\sum_{i=0}^{\infty}\sum_{k=0}^{\infty} s_{i,i+k}(z_1z_2)^{-i}z_2^{-k} + \sum_{j=1}^{\infty}\sum_{k=0}^{\infty} s_{j+k,j}(z_1z_2)^{-j}z_1^{-k}\right]$$

$$\triangleq (z_1z_2)^{-1}(s_1+s_2).$$

Consider then

$$s_1 = \sum_{i=0}^{\infty}\sum_{k=0}^{\infty} s_{i,i+k}(z_1z_2)^{-i}z_2^{-k}$$

$$= \sum_{i=0}^{\infty} a_i(z_2)(z_1z_2)^{-1}/p_2(z_2)$$

for some $a_i(z_2)$ with $\deg a_i \le \deg p_2$

and by interchanging summations we have

$$s_1 = \sum_{k=0}^{\infty} b_k(z_1z_2)z_2^{-k}/p(z_1z_2)$$

for some $b_k(z_1z_2)$ with $\deg b_k \le \deg p$.

Now from Apostol [AP1], we know that if a sum $\sum_{m,n} G(m,n)$ can be written as $\sum_{m,n} G(m,n) = \sum_m A_m = \sum_n B_n$, then the sum does indeed exist. In particular, for the case we are considering, we can readily show that

$$s_1 = \frac{1}{p_2(z_2)p(z_1z_2)} N_1(z_1,z_2)$$

for some $N_1 \in \mathcal{R}[z_1,z_2]$.

Similarly

$$s_2 = \frac{1}{p_1(z_1)p(z_1z_2)} N_2(z_1,z_2)$$

so that

$$s = (z_1z_2)^{-1}(s_1+s_2) = \frac{1}{p_1(z_1)p_2(z_2)z_1z_2p(z_1z_2)} [p_1(z_1)N_1(z_1,z_2) + p_2(z_2)N_2(z_1,z_2)]$$

$$\in R^{real}[(z_1,z_2)].$$

Conversely, suppose $s = \frac{N(z_1,z_2)}{p_1(z_1)p_2(z_2)p(z_1z_2)}$.

By multiplying top and bottom by $(z_1z_2)^k$ for appropriate k it is clear that we can factorize s as

$$s = \frac{f(z_1z_2)}{p(z_1z_2)} \frac{M(z_1,z_2)}{q_1(z_1)q_2(z_2)}$$

where $\deg f \le \deg p$

$\deg_{z_1} M < \deg q_1$, $\deg_{z_2} M < \deg q_2$

where $\deg_{z_i} M$ is the highest power of z_i appearing in M,

and $q_1(z_1) = z_1^k p_1(z_1)$, $q_2(z_2) = z_2^k p_2(z_2)$.

Hence $s = \sum_{k\ge 0} a_k(z_1z_2)^{-k} \sum_{i,j\ge 0} c^T M_1^i M_2^j b z_1^{-(i+1)} z_2^{-(j+1)}$ by Lemma 2.2.2.

It then follows by equating this with $s = \sum s_{mn} z_1^{-(m+1)} z_2^{-(n+1)}$ that

$$s_{mn} = \sum_{k=0}^{\min(m,n)} a_k c^T M_1^{m-k} M_2^{n-k} b.$$

Hence if we consider the column series,

$$\begin{aligned} c_n &= \sum_{r=0}^{\infty} s_{n+r,n} z^{-r} \\ &= \sum_{r=0}^{\infty} \sum_{k=0}^{n} a_k c^T M_1^{n+r-k} M_2^{n-k} b \\ &= \sum_{k=0}^{n} \sum_{r=0}^{\infty} a_k c^T M_1^{n+r-k} M_2^{n-k} b z^{-r} \\ &= \sum_{k=0}^{n} a_k c^T M_1^{n-k} z^{-1} (I-M_1 z^{-1})^{-1} M_2^{n-k} b \\ &= \sum_{k=0}^{n} a_k c^T (M_1M_2)^{n-k} (zI-M_1)^{-1} b \end{aligned}$$

which is a finite sum having denominator $\det(zI-M_1)$ for all $n \ge 1$.

Similarly the row series $(m \ge 0)$ all have common denominator $\det(zI-M_2)$.

$$\text{Now } d_{mn}(z) = \sum_{r=0}^{\infty} s_{m+r,n+r} z^{-r}.$$

$$\begin{aligned}
\text{Hence } d_{mn}(z_1z_2) &= \sum_{r=0}^{\infty} s_{m+r,n+r} z_1^{-r} z_2^{-r} \\
&= z_1^{m+1} z_2^{n+1} s \odot \sum_{k\geq 0} (z_1z_2)^{-k} \\
&= \frac{f(z_1z_2)}{p(z_1z_2)} \frac{z_1^{m+1} z_2^{n+1}}{q_1(z_1)q_2(z_2)} M(z_1,z_2) \odot \sum_{k\geq 0} (z_1z_2)^{-k} \\
&= \frac{f(z_1z_2)}{p(z_1z_2)} z_1^{m+1} z_2^{n+1} \sum_{i,j\geq 0} c^T M_1^i M_2^j b z_1^{-(i+1)} z_2^{-(j+1)} \odot \sum_{k\geq 0} (z_1z_2)^{-k}
\end{aligned}$$

by Lemma 2.2.2.

Because of the term to the right of $\odot$, we can neglect any term to the left of $\odot$ where the powers of z_1 and z_2 are not equal.

Now, assume $m \geq n$. By considering equal powers of z_1 and z_2 (by setting $m-i = n-j$) in the above expression we obtain

$$\begin{aligned}
d_{mn}(z_1z_2) &= \frac{f(z_1z_2)}{p(z_1z_2)} \sum_{j\geq 0} c^T M_1^{j+m-n} M_2^j b z_1^{n-j} z_2^{n-j} \odot \sum_{k\geq 1} (z_1z_2)^{-k} \\
&= \frac{f(z_1z_2)}{p(z_1z_2)} c^T M_1^{m-n} (z_1z_2)^{n+1} (z_1z_2 I - M_1M_2)^{-1} \odot \sum (z_1z_2)^{-k}.
\end{aligned}$$

Hence $d_{mn}(z)$ has denominator $p(z)\det(zI-M_1M_2)$ for all $m \geq n$, and likewise for all $n \geq m$. □

We next show that the $\mathcal{R}$-linear spaces X_3, X_1 and X_2 are all finite dimensional if $s \in \mathcal{R}^{real}[(z_1,z_2)]$, by relating X_3, X_1 and X_2 to the diagonal, column and row series defined above.

We first of all examine $X_3 = U \otimes V/\ker f_{\otimes}$, and we note that

$$\begin{aligned}
f_{\otimes}(z_1^i z_2^j) &= (z_1z_2)^{-1} \sum_{m,n\geq 0} s_{mn} z_1^{-m} z_2^{-n} z_1^i z_2^j \odot \sum_{k\geq 1} (z_1z_2)^{-k} \qquad \text{(by definition of } f_{\otimes}) \\
&= (z_1z_2)^{-1} \sum_{k=0}^{\infty} s_{i+k,j+k} (z_1z_2)^{-k} \\
&= (z_1z_2)^{-1} d_{ij}(z_1z_2).
\end{aligned}$$

Using this relationship, we can establish the following result.

Lemma 2.2.5

X_3 is finite-dimensional iff $\{d_{ij}(z)\}$ are power series expansions of rational functions, having common denominator.

Proof: Since $\{z_1^i z_2^j\}$ is a basis for $R[z_1,z_2]$, it is apparent that

$$\text{Im} f_{\alpha} = \text{span}\{f_{\alpha}(z_1^i z_2^j) \;:\; i,j = 0,1,\ldots\}.$$

Let us assume that X_3 is finite-dimensional. Then $\text{Im} f_{\alpha}$ has a finite basis, so equivalently $\text{span}\{d_{ij}(z)\}$ has a finite basis, say $d_{i_1 j_1}(z),\ldots,d_{i_n j_n}(z)$.

Hence $d_{i_{k+1},j_{k+1}}(z) = \sum_{m=1}^{n} b_{km} d_{i_m j_m}(z), \quad b_{km} \in R, \quad k = 1,\ldots,n.$

But by definition of the diagonal series we have

$$d_{i_{k+1},j_{k+1}}(z) = z(d_{i_k j_k}(z) - s_{i_k j_k}) \qquad k = 1,\ldots,n.$$

Equating these, we obtain

$$(zI - B)\underline{d} = \underline{s}$$

where $B = (b_{km})$ $\underline{d} = [d_{i,j'}(z)\ldots d_{i_n j_n}(z)]^T$, $\underline{s} = [s_{i,j'}\ldots s_{i_n j_n}]^T$ so that all $d_{ij}(z)$ have common denominator $\det(zI-B)$.

Conversely, suppose $d_{ij}(z) = N_{ij}(z)/p(z)$ with deg $p = n$. Then $\dim X_3 = \dim \text{span}\{d_{ij}(z)\} \le m$. □

Before proving analogous results for the row and column series, we shall define the morphisms f_1 and f_2 and relate them to the equivalence relations $\tilde{1}$ and $\tilde{2}$.

We define $f_1 : U \to R[[z^{-1}]]^{1\times\infty}$ by

$$f_1(u) = \left(f(z_1 u,1), f(z_1^2 u,1),\ldots\right).$$

The linear space $R[[z^{-1}]]^{1\times\infty}$ admits the structure of an $R[z_1]$ module if we have the multiplication $z_1(s_1,s_2,s_3,\ldots) = (s_2,s_3,\ldots)$. Hence f_1 is an $R[z_1]$ morphism since

$$\begin{aligned} z_1 f_1(u) &= z_1\left(f(z_1 u,1), f(z_1^2 u,1),\ldots\right) \\ &= \left(f(z_1^2 u,1),\ldots\right) \\ &= f_1(z_1 u). \end{aligned}$$

We next show that $f_1(u_1) = f_1(u_2)$ iff $u_1 \underset{1}{\sim} u_2$. From the definition of f_1 it is clear that $f_1(u_1) = f_1(u_2)$ iff

$$f(z_1^k u_1, 1) = f(z_1^k u_2, 1). \qquad \text{for all } k \geq 1.$$

Hence
$$\begin{aligned} f(z_1^r u_1, a_1 z_2^{r-1} + ,,, + a_r) &= \sum_{i=1}^{r} a_i (z_1 z_2)^{r-i} f(z_1^i u_1, 1) \\ &= \sum_{i=1}^{r} a_i (z_1 z_2)^{r-i} f(z_1^i u_2, 1) \\ &= f(z_1^r u_2, a_1 z_2^{r-1} + \ldots + a_r) \end{aligned}$$

so that $u_1 \underset{1}{\sim} u_2$.

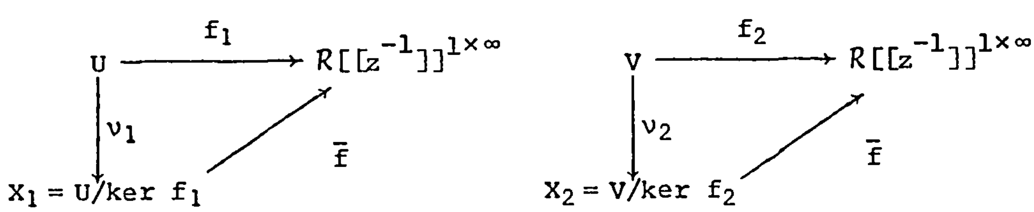
Conversely, let $u_1 \underset{1}{\sim} u_2$. Then $f(z_1^k u_1, 1) = f(z_1^k u_2, 1)$ for all $k \geq 1$, so that $f_1(u_1) = f_1(u_2)$.

In a similar manner we define the $R[z_2]$ morphism $f_2 : V \to R[[z^{-1}]]^{1\times\infty}$ by

$$f_2(v) = \left(f(1, z_2 v), f(1, z_2^2 v), \ldots\right)$$

and we can show that $f_2(v_1) = f_2(v_2)$ iff $v_1 \underset{2}{\sim} v_2$.

Having established these relationships, it is clear that an equivalent definition of the spaces $X_1 = U/\underset{1}{\sim}$ and $X_2 = V/\underset{2}{\sim}$ is $X_1 = U/\ker f_1$ and $X_2 = V/\ker f_2$. We can then obtain the commutative diagrams



where ν_1 is onto and $\bar{f}_1$ is (1-1); X_1 is then naturally endowed with $R[z_1]$ module structure and $\ker f_1$ is a principal ideal in $R[z_1]$. (Similarly for ν_2, $\bar{f}_2$ and X_2.)

We now have sufficient machinery to obtain the following results:

Lemma 2.2.6

X_1 is finite-dimensional iff the column series $c_j(z)$, $j = 0,1,\ldots$ are power series expansions of rational functions having common denominator.

Proof: Let X_1 be finite dimensional. Then $\ker f_1 = (w_1(z_1))$ for some polynomial $w_1(z_1) \in R[z_1]$. Moreover $\dim X_1 = \deg w_1$.

Now let $w_1(z_1) = z_1^n + \alpha_1 z_1^{n-1} + \ldots + \alpha_n \in \ker f$.

Then $0 = f_1\big(w_1(z_1)\big) = w_1(z_1) f_1(1)$

where $f_1(1) = (d_{11}, d_{21}, d_{31}, \ldots)$

Hence $0 = \alpha_n(d_{11}, d_{21}, \ldots) + \alpha_{n-1}(d_{21}, d_{31}, \ldots)$

$+ \ldots + (d_{n+1,1}, d_{n+2,1}, \ldots)$.

This in turn implies that

$$\alpha_n s_{1+k,1+k} + \alpha_{n-1} s_{2+k,1+k} + \ldots + s_{n+1+k,1+k} = 0 \quad \text{for all } k$$

$$\alpha_n s_{2+k,1+k} + \alpha_{n-1} s_{3+k,1+k} + \ldots + s_{n+2+k,1+k} = 0 \quad \text{for all } k$$

. . . etc.

Hence

$$s_{r+k,1+k} = (1\,0\,\ldots\,0)\begin{pmatrix} 0\,1 & & 0 \\ & 0 & 1 \\ -\alpha_n & \cdots\cdots & -\alpha_1 \end{pmatrix}^r \begin{pmatrix} s_{1+k,1+k} \\ \\ s_{n+k,1+k} \end{pmatrix} \triangleq c^T A^r b.$$

Hence the column series are power series expansions of rational functions with common denominator $w_1(z)$. This follows immediately from

$$\sum_{r=1} s_{r+k,1+k} z^{-r} = \sum_{r=1} c^T A^r b z^{-r} = c^T A(zI - A)^{-1} b$$

where $w_1(z) = \det(zI - A)$.

Conversely, suppose that the column series $c_i(z)$ correspond to rational functions $N_i(z)/w_1(z)$, $w_1(z) \neq 0$.

Then

$$f\big(z_1^r w_1(z_1), 1\big) = (z_1 z_2)^{-1} \sum_{i,j} s_{ij} z_1^{-i} z_2^{-j} z_1^r (z_1^n + \alpha_1 z^{n-1} + \ldots + \alpha_n) \circledcirc \sum_{k\geq 1} (z_1 z_2)^{-k}$$

$$= (z_1 z_2)^{-1} \sum_{i,j} s_{ij} z_2^{-j} (z_1^{n+r-i} + \alpha_1 z_1^{n+r-1-i} + \ldots + \alpha_n z_1^{r-i}) \circledcirc \sum_{k\geq 1} (z_1 z_2)^{-k}$$

$$= (z_1 z_2)^{-1} \left[\sum_{\ell=1}^{n} \alpha_\ell \left(\sum_{k=1}^{\infty} s_{r+n-\ell+k,k} (z_1 z_2)^{-k} \right) + \sum_{k=1}^{\infty} s_{r+n+k,k} (z_1 z_2)^{-k} \right].$$

Let us examine the coefficient of $(z_1 z_2)^{-(k+1)}$; this is equal to

$$s_{r+n+k,k} + \sum_{i=1}^{n} \alpha_i s_{r+n-i+k,k}.$$

Now the column series $c_k(z) = \sum_{p=0}^{\infty} s_{p+k,k} z^{-p}$ sums to $N_k(z) w_1(z)$.

Hence the coefficients of the negative powers of z in $w_1(z)c_k(z) = w_1(z)\sum_{p=0}^{\infty} s_{p+k,k}z^{-k}$ are all equal to zero. This immediately implies, on examination of the coefficient of $z^{-(k+1)}$ in this sum, that $s_{r+n+k,k} + \sum_{i=1} \alpha_i s_{r+n-i+k,k} = 0$ for $r \geq 1$.

It is now clear $f(z_1^r w_1(z_1),1) = 0$ for all r, so that $w_1(z_1) \in \ker f_1$, and since $R[z_1]$ is a principal ideal domain, $\ker f_1 = (w_1)$, so that X_1 is finite-dimensional. □

Using similar reasoning, we obtain the following.

Lemma 2.2.7

X_2 is finite dimensional iff the ~~column~~ row series $r_i, i = 0,1,\ldots$, are power series expansions of rational functions having common denominator.

Hence, combining the lemmas that we have just proved, we find that the space $X_1 \oplus X_2 \oplus X_3$ is finite-dimensional if and only if $s(z_1,z_2)$ is a realizable series. In §2.3, we shall demonstrate how to obtain a state space realization of the bilinear input/output map f represented by a realizable series $s(z_1,z_2)$, based on the use of the module-morphisms f_1, f_2 and $f_\otimes$ and their kernels. Before that, however, we shall devote a few further lines to conditions for reachability in bounded time of the Nerode space X_N, details of which may be found in [FM1].

The principal result concerning this is that X_N is reachable in bounded time if and only if $s(z_1,z_2)$ is a realizable series. In other words, the intuitive notion that reachability in bounded time is an equivalent concept to that of being able to write down a finite-dimensional realization for f is confirmed. We shall omit the proof of this result as it is not fundamental to any of the work presented later; however, it is worthwhile giving some indication of the path taken, as this bears some similarity to the proof of quasi-reachability of the state-space realization of Chapter 3, and it also provides an opportunity to point out an error in the proof of [FM 1].

Defining $(w_1(z_1)) = \ker f_1$ and $(w_2(z_2)) = \ker f_2$ as above, it is possible to write any input sequence $\in U \times V \cong R[z_1] \times R[z_2]$ as

$$\left(p_1(z_1)w_1(z_1) + q_1(z_1),\ p_2(z_2)w_2(z_2) + q_2(z_2)\right)$$

where $\deg q_1 < \deg w_1$, $\deg q_2 < \deg w_2$.

Fornasini and Marchesini then construct an algorithm to choose polynomials $g_1(z_1)$ and $g_2(z_2)$ such that (i) $f(g_1w_1,g_2w_2) = f(p_1w_1,p_2w_2)$, (ii) $f(g_1w_1,q_2) = f(p_1w_1,q_2)$ and (iii) $f(q_1,g_2w_2) = f(q_1,p_2w_2)$ with $\deg g_1$ and $\deg g_2$ always less than some specified integer M (dependent on the particular map f). The Nerode equivalent input is then $(g_1w_1 + q_1, g_2w_2 + q_2)$, which can be seen from the equivalence of $f\left(z_1^k(g_1w_1+q_1)+u, z_2^k(g_2w_2+q_2)+v\right)$ and $f\left(z_1^k(p_1w_1+q_1)+u, z_2^k(p_2w_2+q_2)+v\right)$ where deg u and deg $v < k$.

The reader of [FM 1] will find, however, that for the construction given, conditions (ii) and (iii) are not necessarily satisfied, e.g. in the case

$$s = \frac{1}{(z_1^2+az_1+b)(z_2-c)z_2^2}\ .$$

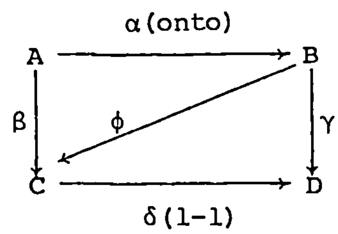
This deficiency can be remedied by replacing the truncation map T^ℓ of Lemma 2.6 of [FM 1] by the truncation map T^L, where $L = \max(\ell, \deg w_1, \deg w_2)$. This will then ensure that conditions (ii) and (iii) hold.

2.3 Finite-Dimensional Realization

In this section we demonstrate how to derive updating equations in $X_1 \oplus X_2 \oplus X_3$ for the case when X_N is reachable in bounded time. The only tool that we shall need will be Zeiger's Lemma [K2], which we state here without proof for the special case of modules.

Lemma 2.3.1

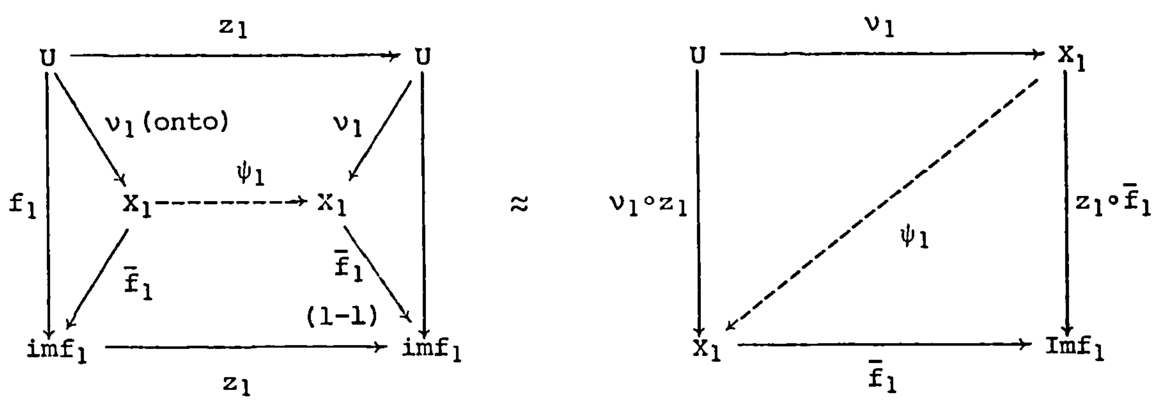
Let A, B, C, D be arbitrary modules. Consider the commutative diagram



where α, β, γ, δ are morphisms, with α onto and δ one-to-one. Then there exists a unique morphism $\phi : B \to C$ such that the diagram remains commutative.

Updating Equations in X_1 and X_2

Cosider the following commutative diagram of $R[z_1]$-module morphisms:



Since ν_1 is onto and $\bar{f}_1$ is (1-1), there exists a unique $R[z_1]$-morphism $\psi_1 : X_1 \to X_1$ such that

$$\nu_1 \circ z_1 = \psi_1 \circ \nu_1 .$$

Hence we can write

$$f_1(z_1 u + u_k) = \bar{f}_1 \circ \nu_1 (z_1 u + u_k) \tag{2.3.1}$$

$$= \bar{f}_1 \circ \big(\nu_1 \circ z_1(u) + u_k \nu_1(1)\big)$$

$$= \bar{f}_1 \circ \big(\psi_1 \circ \nu_1(u) + u_k \nu_1(1)\big) \tag{2.3.2}$$

and since $\bar{f}_1$ is (1-1) we can equate (2.3.1) and (2.3.2) to obtain

$$\nu_1(z_1u+u_k) = \psi_1\circ\nu_1(u) + u_k\nu_1(1).$$

So with respect to a basis in X_1 we can write

$$x^1_{k+1} = A_1x^1_k + b_1u_k \tag{2.3.3}$$

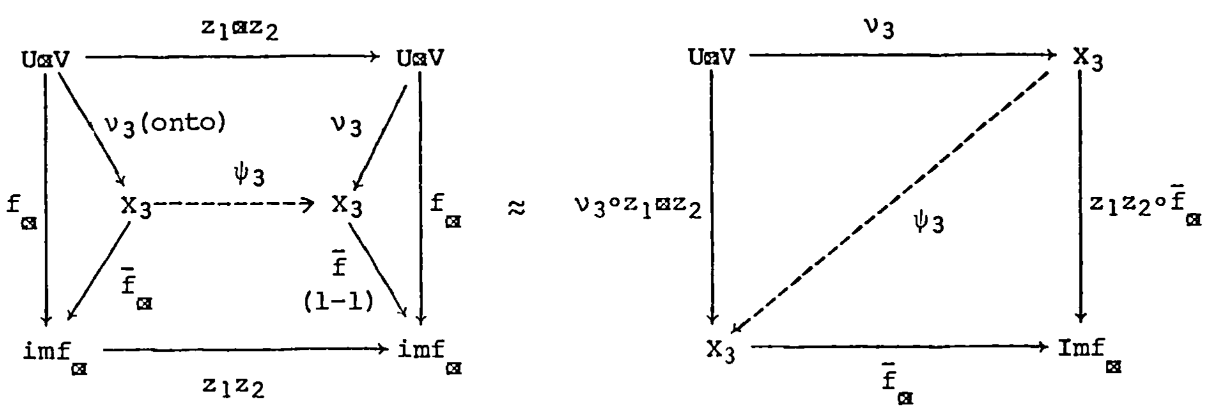
where A_1 and b_1 are representations of ψ_1 and $\nu_1(1)$.

In a similar manner we obtain updating equations for X_2:

$$x^2_{k+1} = A_2x^2_k + b_2v_k. \tag{2.3.4}$$

Updating Equation in X_3

Consider the following commutative diagram of $R[z_1z_2]$- module morphisms:



Since ν_3 is onto and $\bar{f}_⊠$ is one-to-one, a unique $R[z_1z_2]$-morphism $\psi_3 : X_3 \to X_3$ exists, such that

$$\nu_3\circ(z_1⊠z_2) = \psi_3\circ\nu_3.$$

Let us introduce the projection mapping

$$\pi : R[[z^{-1}]]^{1\times\infty} \to R[[z^{-1}]] : (s_1,s_2,\ldots) \to s_1$$

which satisfies the following equations:

$$f_⊠(z_1u⊠1) = \pi\circ f_1(u)$$

$$f_⊠(1⊠z_2v) = \pi\circ f_2(v).$$

Then by bilinearity of f we can write

$$f(z_1u+u_k, z_2v+v_k) = f(z_1u,z_2v) + v_kf(z_1u,1) + u_kf(1,z_2v) + u_kv_kf(1,1)$$

and hence

$$\bar{f}_{\otimes}\circ\nu_3\big((z_1u+u_k)\otimes(z_2v+v_k)\big) = \bar{f}\otimes\circ\nu_3(z_1u\otimes z_2v) + v_k\bar{f}_{\otimes}\circ\nu_3(z_1u\otimes 1) + u_k\bar{f}_{\otimes}\circ\nu_3(1\otimes z_2v) + u_kv_k\bar{f}_{\otimes}\circ\nu_3(1\otimes 1)$$

$$= \bar{f}_{\otimes}\circ\psi_3\circ\nu_3(u\otimes v) + v_k\bar{f}_{\otimes}\circ\tau_1\circ\nu_1(u) + u_k\bar{f}_{\otimes}\circ\tau_2\circ\nu_2(v) + u_kv_k\bar{f}_{\otimes}\circ\nu_3(1\otimes 1)$$

where $\tau_i = \bar{f}_{\otimes}^{-1}\circ\pi_i\circ\bar{f}_i : X_i \to X_3 \quad i = 1,2.$

Since $\bar{f}_{\otimes}$ is one-to one, we can write with respect to a basis in X :

$$x_{k+1} = Ax_k + Q_1x_k^1v_k + Q_2x_k^2u_k + bu_kv_k \tag{2.3.5}$$

$$y_k = h^Tx_k \tag{2.3.6}$$

where A, Q_1, Q_2, b and h^T are representations of ψ_3, τ_1, τ_2, $\nu_3(1\otimes 1)$ and $\bar{f}_{\otimes}$ respectively in the chosen basis.

We shall now show that if a bilinear input/output map can be represented in the above state space form, then it can also be represented by a realizable power series s, which can be directly computed from the system matrices. This we do by evaluating $s = (z_1z_2)^{-1}\sum s_{ij}z_1^{-i}z_2^{-j}$, where s_{ij} is the output at time 0 due to unit inputs at times -i and -j in the U and V channels respectively, as was defined previously in §2.1.

Consider unit inputs at time -(i+k) and -i in channels U and V respectively for k ≥ 1. We then obtain the following:

$$x^1_{-(i+k)+1} = b_1 \qquad \text{by (2.3.3)}$$

$$x^1_{-i} = A_1^{k-1}b_1 \qquad \text{by (2.3.3)}$$

$$x_{-i+1} = Q_1A_1^{k-1}b_1 \qquad \text{by (2.3.5)}$$

$$x_1 = A^iQ_1A_1^{k-1}b_1 \qquad \text{by (2.3.5)}$$

$$y_1 = h^TA^iQ_1A_1^{k-1}b_1 \qquad \text{by (2.3.6).}$$

Hence $s_{i+k,i} = h^TA^iQ_1A_1^{k-1}b_1$. Similarly we obtain $s_{j,j+k} = h^TA^jQ_2A_2^{k-1}b_2$. Finally for unit inputs at time -i in both channels U and V, we obtain $s_{ij} = h^TA^ib$. We can then compute s as

$$s = \sum_{i=0}^{\infty}\sum_{k=1}^{\infty} h^T A^i Q_1 A_1^{k-1} b_1 (z_1 z_2)^{-(i+1)} z_1^{-k} + \sum_{j=0}^{\infty}\sum_{k=1}^{\infty} h^T A^j Q_2 A_2^{k-1} b_2 (z_1 z_2)^{-(j+1)} z_2^{-k} + \sum_{i=0}^{\infty} h^T A^i b (z_1 z_2)^{-(i+1)}$$

$$= h^T (z_1 z_2 I - A)^{-1} \{ Q_1 (z_1 I - A_1)^{-1} b_1 + Q_2 (z_2 I - A_2)^{-1} b_2 + b \}.$$

In order to illustrate the realization procedure described above, we will carry out each step for a simple example.

Let $s = \dfrac{1}{(z_1-a)(z_2-b)(z_1z_2-c)}$

It can then readily be seen that

$$f(1,1) = \frac{1}{(z-ab)(z-c)}$$

$$f(z_1^k,1) = \frac{a^k}{(z-ab)(z-c)}$$

$$f(1,z_2^k) = \frac{b^k}{(z-ab)(z-c)}$$

and we can now compute $f(u,v)$ for any $(u,v) \in U \times V$. Note in particular that $f(z_1,z_2) = \dfrac{z}{(z-ab)(z-c)}$, so that a basis for $\operatorname{im} f_{\otimes}$ is given by $f_{\otimes}(1 \otimes 1)$ and $f_{\otimes}(z_1 \otimes z_2)$, i.e. $\dim X_3 = 2$.

Now

$$f_1(1) = \left(f(z_1,1), f(z_1^2,1), \ldots\right)$$
$$= \left(\frac{a}{(z-ab)(z-c)}, \frac{a^2}{(z-ab)(z-c)}, \ldots\right)$$

and we can readily see that $f_1(z_1) = z_1 f_1(1) = a f_1(1) = f_1(a)$, so that $\ker f_1 = (z_1-a)$, and similarly $\ker f_2 = (z_2-b)$. It is also clear that $[u]_1 = x_k^1 [1]_1$ for all $u \in U$, for some scalar x_k^1 dependent on u, where $[w]_1$ denotes the equivalence class of w under $\tilde{1}$. We can then write

$$[z_1 u + u_k]_1 = x_{k+1}^1 [1]_1 \tag{2.3.7}$$

$$= z_1 [u]_1 + u_k [1]_1$$

$$= a[u]_1 + u_k [1]_1 \quad \text{since } z_1 - a \in \ker f_1$$

$$= a x_k^1 [1]_1 + u_k [1]_1. \tag{2.3.8}$$

Equating (2.3.7) and (2.3.8) we obtain

$$x_{k+1}^1 = a x_k^1 + u_k \tag{2.3.9}$$

and similarly for z_2v+v_k, where we define $[v]_2 = x_k^2[1]_2$, we obtain

$$x_{k+1}^1 = bx_k^2 + v_k. \tag{2.3.10}$$

Now from our choice of basis for im $f_\otimes$ above, we can write

$$f_\otimes(u\otimes v) = \alpha_k f_\otimes(1\otimes 1) + \beta_k f_\otimes(z_1\otimes z_2)$$

for some scalar α_k and β_k dependent on u and v.

Hence

$$\begin{aligned} f_\otimes\big((z_1u+u_k)\otimes(z_2v+v_k)\big) &= f_\otimes(z_1u\otimes z_2v) + v_k f_\otimes(z_1u\otimes 1) + u_k f_\otimes(1\otimes z_2v) + u_k v_k f_\otimes(1\otimes 1) \\ &= z_1\otimes z_2\{\alpha_k f_\otimes(1\otimes 1) + \beta_k f_\otimes(z_1\otimes z_2)\} \\ &\quad + v_k f_\otimes(z_1u\otimes 1) + u_k f_\otimes(1\otimes z_2v) + u_k v_k f_\otimes(1\otimes 1). \end{aligned} \tag{2.3.11}$$

Now

$$z_1\otimes z_2\ f_\otimes(1\otimes 1) = f_\otimes(z_1\otimes z_2)$$

and $$z_1\otimes z_2\ f_\otimes(1\otimes 1) = f_\otimes(z_1^2\otimes z_2^2)$$

and $$\begin{aligned} f_\otimes(z_1^2\otimes z_2^2) &= \frac{z^2}{(z-ab)(z-c)} \odot \sum_{k\geq 1} z^{-k} \\ &= \frac{(ab+c)z - abc}{(z-ab)(z-c)} \\ &= (ab+c)f_\otimes(z_1\otimes z_2) - abcf_\otimes(1\otimes 1). \end{aligned} \tag{2.3.12}$$

Furthermore, we note that

$$\begin{aligned} \big(f_\otimes(z_1u\otimes 1), f_\otimes(z_1^2u\otimes 1), \ldots\big) &= f_1(u_1) \\ &= x_k^1 f_1(1) \text{ since } [u_1]_1 = x_k^1[1]_1 \\ &= x_k^1\big(f_\otimes(z_1\otimes 1), f_\otimes(z_1^2\otimes 1), \ldots\big). \end{aligned}$$

Hence $$\begin{aligned} f_\otimes(z_1u\otimes 1) &= x_k^1 f_\otimes(z_1\otimes 1) \\ &= ax_k^1 f_\otimes(1\otimes 1). \end{aligned} \tag{2.3.13}$$

Similarly we obtain

$$f_\otimes(1\otimes z_2v) = bx_k^2 f_\otimes(1\otimes 1). \tag{2.3.14}$$

Inserting (2.3.12)-(2.3.14) into (2.3.11) we obtain

$$\begin{aligned} \alpha_{k+1} f_\otimes(1\otimes 1) + \beta_{k+1} f_\otimes(z_1\otimes z_2) &\triangleq f_\otimes\big((z_1u+u_k)\otimes(z_2v+v_k)\big) \\ &= \alpha_k f_\otimes(z_1\otimes z_2) + \beta_k[(ab+c)f_\otimes(z_1\otimes z_2) - abcf_\otimes(1\otimes 1)] \\ &\quad + ax_k^1 v_k f_\otimes(1\otimes 1) + bx_k^2 u_k f_\otimes(1\otimes 1) + u_k v_k f_\otimes(1\otimes 1) \end{aligned}$$

and equating the coefficients of $f_{\boxtimes}(1\boxtimes 1)$ and of $f_{\boxtimes}(z_1\boxtimes z_2)$ we obtain

$$\begin{pmatrix}\alpha_{k+1}\\ \beta_{k+1}\end{pmatrix} = \begin{pmatrix}0 & -abc\\ 1 & ab+c\end{pmatrix}\begin{pmatrix}\alpha_k\\ \beta_k\end{pmatrix} + \begin{pmatrix}a\\ 0\end{pmatrix}x_k^1 v_k + \begin{pmatrix}b\\ 0\end{pmatrix}x_k^2 u_k + \begin{pmatrix}1\\ 0\end{pmatrix}u_k v_k. \qquad (2.3.13)$$

Finally we note that f(1,1) has zero output at time + 1, and $f(z_1,z_2)$ has output 1 at time + 1, so that

$$y_k = \begin{pmatrix}0 & 1\end{pmatrix}\begin{pmatrix}\alpha_k\\ \beta_k\end{pmatrix}. \qquad (2.3.14)$$

We can immediately check that the state space realization given by (2.3.7), (2.3.8), (2.3.13) and (2.3.14) is correct by calculating

$$s = \begin{pmatrix}0 & 1\end{pmatrix}\begin{pmatrix}z_1z_2 & abc\\ -1 & z_1z_2-ab-c\end{pmatrix}^{-1}\left\{\begin{pmatrix}a\\ 0\end{pmatrix}(z_1-a)^{-1} + \begin{pmatrix}b\\ 0\end{pmatrix}(z_2-b)^{-1} + \begin{pmatrix}1\\ 0\end{pmatrix}\right\}$$

$$= \frac{1}{(z_1z_2-c)(z_1z_2-ab)}\left(\frac{a}{z_1-a} + \frac{b}{z_2-b} + 1\right)$$

$$= \frac{1}{(z_1-z)(z_2-b)(z_1z_2-c)}.$$

Note that the state space description obtained above is observable (in the sense defined later in Chapter 3), but it is not reachable, since

$$x_k^1 x_k^2 = \alpha_k + ab\beta_k \quad \text{for all } k.$$

2.4 Alternative Methods of Realization

The realization which was produced at the end of the last section is typical of state space realizations of bilinear input/output maps formed by consideration of the equivalence relations $\tilde{1}$, $\tilde{2}$ and $\tilde{3}$, in that it is not reachable. A reasonable method of correcting this deficiency for the example above is to substitute $x_k^1 x_k^2 - ab\beta_k$ for α_k wherever it occurs. The dynamic equation for β_{k+1} can then be expressed as

$$\beta_{k+1} = (x_k^1 x_k^2 - ab\beta_k) + (ab + c)\beta_k$$
$$= x_k^1 x_k^2 + c\beta_k$$

and it is now clear that we are left with a three-state realization describing the map s, which it is fairly easy to see is both reachable and controllable (by the usual definitions of reachability and controllability) provided that a and b are non-zero.

We shall now introduce a state space description of bilinear input/output maps which generalizes the preceding analysis and which formalizes the ideas of Kalman's seminal paper on multilinear systems [Kl]. We shall follow this by a discussion of the advantages of this representation over the one of §2.3, in particular how it is possible to go straight from the transfer function to the state space description, by-passing any consideration of equivalence classes, which even for the simple example above was somewhat tedious. The state space description which is the basis for our later results on reachability, observability and minimal realizations is as follows:

$$x^1_{k+1} = A_1 x^1_k + b_1 u_k \tag{2.4.1}$$

$$x^2_{k+1} = A_2 x^2_k + b_2 v_k \tag{2.4.2}$$

$$x_{k+1} = Ax_k + Cx^1_k \otimes x^2_k + Q_1 x^1_k v_k + Q_2 x^2_k u_k + bu_k v_k \tag{2.4.3}$$

$$y_k = h^T x_k + d^T x^1_k \otimes x^2_k. \tag{2.4.4}$$

Note the inclusion of the term $x^1_k \otimes x^2_k$, where $\otimes$ is the Kronecker product, and since this is bilinear in U and V, it is obvious by induction on x_k that both x_k and y_k are also bilinear in U and V. In Chapter 4 we will show that by the addition of this term it will always be possible to set up a state space description of any bilinear input/output map, with finite dimensional Nerode space, which is both quasi-reachable and observable.

We shall also see that the matrices A_1, A_2 and A have a direct interpretation in terms of the transfer function $s = N(z_1,z_2)/p_1(z_1)p_2(z_2)p(z_1z_2)$; in fact the characteristic polynomials of these matrices will be equal to

$p_1(z)$, $p_2(z)$ and $p(z)$ respectively, to within some factor z^r. To demonstrate this fact we shall prove the following theorem which shows how to compute the transfer function associated with (2.4.1)-(2.4.4) and afterwards give examples of how to set up a suitable state space description.

Theorem 2.4.1

The transfer function associated with equations (2.4.1)-(2.4.4) is given by

$$s = h^T(z_1z_2I-A)^{-1}\{C(z_1I-A_1)^{-1}b_1 \otimes (z_2I-A_2)^{-1}b_2 + Q_1(z_1I-A_1)^{-1}b_1 + Q_2(z_2I-A_2)^{-1}b_2 + b\} + d^T(z_1I-A_1)^{-1}b_1 \otimes (z_2I-A_2)^{-1}b_2. \quad (2.4.5)$$

Proof: We shall set up a state space analogous to that in §2.3, and then employ the formula derived there to calculate s.

First we shall compute the transition map of $x_k^1 \otimes x_k^2$ from (2.4.1) and (2.4.2), and combining this with (2.4.3) we obtain the composite state transition map

$$\begin{pmatrix} x_{k+1}^1 \otimes x_{k+1}^2 \\ x_{k+1} \end{pmatrix} = \begin{pmatrix} A_1\otimes A_2 & 0 \\ C & A \end{pmatrix} \begin{pmatrix} x_k^1 \otimes x_k^2 \\ x_k \end{pmatrix} + \begin{pmatrix} A_1\otimes b_2 \\ Q_1 \end{pmatrix} x_k^1 v_k + \begin{pmatrix} b_1\otimes A_2 \\ Q_2 \end{pmatrix} x_k^2 u_k + \begin{pmatrix} b_1\otimes b_2 \\ b \end{pmatrix} u_k v_k.$$

This equation together with (2.4.1), (2.4.2) and (2.4.4) are of the same form as the state space description of §2.3, so that the transfer function $s(z_1,z_2)$ is computed as

$$s = [d^T h^T] \begin{bmatrix} z_1z_2I-A_1\otimes A_2 & 0 \\ -C & z_1z_2I-A \end{bmatrix}^{-1} \left(\begin{bmatrix} A_1\otimes b_2 \\ Q_1 \end{bmatrix} (z_1I-A_1)^{-1}b_1 + \begin{bmatrix} b_1\otimes A_2 \\ Q_2 \end{bmatrix} (z_2I-A_2)^{-1}b_2 + \begin{bmatrix} b_1\otimes b_2 \\ b \end{bmatrix} \right)$$

$$= [d^T h^T] \begin{bmatrix} (z_1z_2I-A_1\otimes A_2)^{-1} & 0 \\ (z_1z_2I-A)^{-1}\, C(z_1z_2I-A_1\otimes A_2)^{-1} & (z_1z_2I-A)^{-1} \end{bmatrix} \left(\begin{bmatrix} A_1\otimes b_2 \\ Q_1 \end{bmatrix} (z_1I-A_1)^{-1}b_1 + \begin{bmatrix} b_1\otimes b_2 \\ b \end{bmatrix} + \begin{bmatrix} b_1\otimes A_2 \\ Q_2 \end{bmatrix} (z_2I-A_2)^{-1}b_2 \right). \quad (2.4.6)$$

$$\text{Now } (z_1z_2I-A_1\otimes A_2)^{-1}\{A_1\otimes b_2(z_1I-A_1)^{-1}b_1 + b_1\otimes b_2 + b_1\otimes A_2(z_2I-A_2)^{-1}b_2\}$$
$$= (z_1z_2I-A_1\otimes A_2)^{-1}\{A_1\otimes(z_2I-A_2) + (z_1I-A_1)\otimes(z_2I-A_2) + (z_1I-A_1)\otimes A_2\}$$
$$(z_1I-A_1)^{-1}b_1\otimes(z_2I-A_2)^{-1}b_2$$
$$= (z_1z_2I-A_1\otimes A_2)^{-1}\{A_1\otimes z_2I - A_1\otimes A_2 + z_1z_2I - A_1\otimes z_2I - z_1I\otimes A_2 + A_1\otimes A_2$$
$$+ z_1I\otimes A_2 - A_1\otimes A_2\}(z_1I-A_1)^{-1}b_1\otimes(z_2I-A_2)^{-1}b_2$$
$$= (z_1I-A_1)^{-1}b_1\otimes(z_2I-A_2)^{-1}b_2.$$

Hence from (2.4.6) we see that

$$s = d^T(z_1I-A_1)^{-1}b_1\otimes(z_2I-A_2)^{-1}b_2 + h^T(z_1z_2I-A)^{-1}C(z_1I-A_1)^{-1}b_1\otimes(z_2I-A_2)^{-1}b_2$$
$$+ h^T(z_1z_2I-A)^{-1}\{Q_1(z_1I-A_1)^{-1}b_1 + Q_2(z_2I-A_2)^{-1}b_2 + b\}. \qquad \square$$

By comparing the expression (2.4.5) for s with the state space equations (2.4.1)-(2.4.4), it becomes apparent how to set up a suitable state space description. Consider the example of section 2.3:

$$s = \frac{1}{(z_1-a)(z_2-b)(z_1z_2-c)}.$$

By associating the A_1 matrix with z_1-a and the A_2 matrix with z_2-b and regarding the bilinear output $1/(z_1-a)(z_2-b)$ as the input to the linear system with transfer function $1/z-c$, it is possible to write down the simple state space description as

$$x^1_{k+1} = ax^1_k + u_k$$
$$x^2_{k+1} = bx^2_k + v_k$$
$$x_{k+1} = cx_k + x^1_kx^2_k$$
$$y_k = x_k.$$

Returning to more general $s \in R^{real}[[z^{-1}, z_2^{-1}]]$, we shall consider two cases; this is done for convenience, rather than because the realizations corresponding to these two cases differ in any significant manner.

Case 1: $s = N(z_1,z_2)/p_1(z_1)p_2(z_2)$

where $N(z_1,z_2) = \sum_{j=0}^{n-1}\sum_{i=0}^{m-1} g_{ij}z_1^iz_2^j$

$$p_1(z_1) = z_1^m + a_1 z_1^{m-1} + \dots + a_m$$
$$p_2(z_2) = z_2^n + b_1 z_2^{n-1} + \dots + b_n$$

A possible state space description is

$$x_{k+1}^1 = \begin{bmatrix} 0 & 1 & & 0 \\ & 0 & \ddots & \\ & & & 1 \\ -a_m & \dots & & -a_1 \end{bmatrix} x_k^1 + \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} u_k$$

$$x_{k+1}^2 = \begin{bmatrix} 0 & 1 & & \\ & & \ddots & \\ & & & 1 \\ -b_n & & & -b_1 \end{bmatrix} x_k^2 + \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} v_k$$

$$y_k = [g_{01} \dots g_{0,n-1} \dots g_{m-1,1} \dots g_{m-1,n-1}] x_k^1 \otimes x_k^2.$$

That this does realize s is an immediate consequence of the fact that the transfer function vector from u_k to x_k^1 is $[1 z_1 \dots z_1^{m-1}]^T / p_1(z_1)$, and likewise for the transfer function vector from v_k to x_k^2.

<u>Case 2</u>: $s = N(z_1 z_2)/p_1(z_1)p_2(z_2)p(z_1 z_2)$ where $\deg p \geq 1$.

By multiplying numerator and denominator of the expression for s by $(z_1 z_2)^k$ for appropriate choice of k we can factorize s as

$$s = \frac{f(z_1 z_2)}{p(z_1 z_2)} \quad \frac{M(z_1, z_2)}{q_1(z_1) q_2(z_2)} \tag{2.4.7}$$

where $\deg_{z_1} M \leq \deg q_1$, $\deg_{z_2} M \leq \deg q_2$, $\deg f < \deg p$ and $q_1(z_1) = z_1^k p_1(z_1)$, $q_2(z_2) = z_2^k p_2(z_2)$.

We can now view $M(z_1, z_2)/q_1(z_1)q_2(z_2)$ as the input to a system with transfer function $f(z)/p(z)$, so writing this as

$$\frac{M(z_1, z_2)}{q_1(z_1) q_2(z_2)} = \frac{M_1(z_1, z_2)}{q_1(z_1) q_2(z_2)} + \frac{M_2(z_1)}{q_1(z_1)} + \frac{M_3(z_2)}{q_2(z_2)} + m_4,$$

where we now require $\deg_{z_1} M_1 < \deg q_1$, $\deg_{z_2} M_1 < \deg q_2$,

$\deg M_2 < \deg q_1$, $\deg M_3 < \deg q_2$

(which can always be satisfied). We can employ Theorem 2.4.1 to enable us to write down a state space description for s as follows:

$$x^1_{k+1} = \begin{pmatrix} 0 & 1 & & \\ & & \ddots & \\ & & & 1 \\ -a_{n_1} & \cdots & & -a_1 \end{pmatrix} x^1_k + \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} u_k$$

$$x^2_{k+1} = \begin{pmatrix} 0 & 1 & & \\ & & \ddots & \\ & & & 1 \\ -b_{n_2} & \cdots & & -b_1 \end{pmatrix} x^2_k + \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} v_k$$

$$x_{k+1} = \begin{pmatrix} 0 & 1 & & \\ & & \ddots & \\ & & & 1 \\ -g_n & \cdots & & -g_1 \end{pmatrix} x_k + \begin{pmatrix} 0 \\ \vdots \\ 0 \\ c_1^T \end{pmatrix} x^1_k \otimes x^2_k + \begin{pmatrix} 0 \\ \vdots \\ 0 \\ c_2^T \end{pmatrix} x^1_k v_k + \begin{pmatrix} 0 \\ \vdots \\ 0 \\ c_3^T \end{pmatrix} x^2_k u_k + \begin{pmatrix} 0 \\ \vdots \\ 0 \\ m_4 \end{pmatrix} u_k v_k$$

$$y_k = [d_n \ldots . d_1] x_k$$

where the numbers in the above matrices are given by

$$q_1(z_1) = z_1^{n_1} + a_1 z_1^{n_1-1} + \ldots + a_{n_1}$$

$$q_2(z_2) = z_2^{n_2} + b_1 z_2^{n_2-1} + \ldots + b_{n_2}$$

$$p(z) = z^n + g_1 z^{n-1} + \ldots + g_n$$

$$f(z) = d_1 z^{n-1} + \ldots + d_n$$

$$M_1(z_1,z_2) = c_1^T \begin{pmatrix} 1 \\ z_1 \\ \vdots \\ z_1^{n_1-1} \end{pmatrix} \otimes \begin{pmatrix} 1 \\ z_2 \\ \vdots \\ z_2^{n_2-1} \end{pmatrix}$$

$$M_2(z_1) = [1 \; z_1 \ldots . z_1^{n_1-1}] c_2$$

$$M_3(z_2) = [1 \; z_2 \ldots . z_2^{n_2-1}] c_3 .$$

This particular realization will not in general be a canonical realization as we shall define it later. However we only desire at this stage to demonstrate that we can in general set up a state space description involving the term $x^1_k \otimes x^2_k$, and that this only requires linear system realization procedures, assuming that the transfer function of the bilinear system is known.

This is a considerable simplification of the realization procedures of [FM1] since it essentially only involves the construction of substates corresponding to $p_1(z_1)$, $p_2(z_2)$ and $p(z_1 z_2)$ respectively. The procedures of [FM1] require the construction of substates corresponding to the $p_1(z_1)p_2(z_2)$ interaction, as evidenced by (2.3.13).

2.5 Input/output Stability

In this section we will present some new sufficient conditions for the output sequence $\{y_k : k \geq 1\}$ due to finite input sequences $\{u_i : i \leq 0\}$, $\{v_j : j \leq 0\}$, to be stable in the ℓ_1-norm, or ℓ_1-stable.

As above, let $s = N(z_1,z_2)/p_1(z_1)p_2(z_2)p(z_1z_2)$; then $N(z_1,z_2)$ can be completely factorized with respect to polynomials in z_1z_2 as $N(z_1,z_2) = M(z_1,z_2)f(z_1z_2)$ i.e.

$$s = \frac{f(z_1z_2)}{p(z_1z_2)} \frac{M(z_1,z_2)}{p_1(z_1)p_2(z_2)} .$$

If $\deg f \leq \deg p$, $\deg_{z_1} M < \deg p_1$ and $\deg_{z_2} M < \deg p_2$ we leave the above expression for s as it is.

If either or both of $\deg_{z_1} M \geq \deg p_1$ and $\deg_{z_2} M \geq \deg p_2$ hold we shall express s as

$$s = \frac{M(z_1,z_2)}{p_1(z_1)z_1^m p_2(z_2)z_2^m} \frac{(z_1z_2)^m f(z_1z_2)}{p(z_1z_2)}$$

where $m = \max\{\deg_{z_1} M - \deg p_1,\ \deg_{z_2} M - \deg p_2\} + 1$.

If $\deg f > \deg p$, we shall express s as

$$s = \frac{M(z_1,z_2)z_1^r z_2^r}{p_1(z_1)p_2(z_2)} \frac{f(z_1z_2)}{(z_1z_2)^r p(z_1z_2)}$$

where $r = \deg f - \deg p$.

In any event, we can rewrite s as

$$s = \frac{R(z_1,z_2)}{q_1(z_1)q_2(z_2)} \frac{g(z_1z_2)}{q(z_1z_2)} \tag{2.5.1}$$

where $\deg_{z_1} R < \deg q_1$, $\deg_{z_2} R < \deg q_2$, $\deg g \leq \deg q$.

Note the similarity between this expression and expression (2.4.7). The difference is that in this case the numerator of s is completely factorized with respect to any polynomials in the term z_1z_2, but this is not necessarily so with (2.4.7).

Having set up this useful expression for s, we can now state the following

Theorem 2.5.1

If either of the following conditions hold, then the output sequence due to a finite length input sequence from $U \times V$ is ℓ_1-stable:

(i) all zeros of p(z) and all terms of the form $\{\alpha_i\beta_j\}$, where $\{\alpha_i\}$ and $\{\beta_j\}$ are the zeros of $p_1(z_1)$ and $p_2(z_2)$ respectively, lie within the unit circle;

(ii) all zeros of p(z) lie within the unit circle, and all terms $\{\alpha_i\beta_j\}$ not lying within the unit circle are zeros of g(z).

Proof: The output sequence due to inputs (z_1^i, z_2^j) for $i \geq j$ is given by

$$y(z_1z_2) = z_1^i z_2^j \frac{R(z_1,z_2)}{q_1(z_1)q_2(z_2)} \times \frac{g(z_1z_2)}{q(z_1z_2)} \odot \sum_{k\geq 1} (z_1z_2)^{-k}$$

$$= z_1^i z_2^j \sum_{r,s\ 1} c^T A_1^r \otimes A_2^s b_1 \otimes b_2 z_1^{-(r+1)} z_2^{-(s+1)} \frac{g(z_1z_2)}{q(z_1z_2)} \odot \sum_{k\geq 1} (z_1z_2)^{-k}$$

by Lemma 2.2.2.

So equating powers of z_1 and z_2 i.e. by setting $i - r = j - s$, we obtain

$$y(z_1z_2) = \sum_{s\geq 1} c^T A_1^{i-j+s} \otimes A_2^s (z_1z_2)^{j-s-1} b_1 \otimes b_2 \frac{g(z_1z_2)}{q(z_1z_2)} \odot \sum_{k\geq 1} (z_1z_2)^{-k}$$

$$= c^T (z_1z_2I - A_1 \otimes A_2)^{-1} A_1^{i-j} b_1 \otimes b_2 \frac{(z_1z_2)^j g(z_1z_2)}{q(z_1z_2)} \odot \sum (z_1z_2)^{-k}$$

Simlarly, for $i \leq j$, we obtain an output sequence

$$y(z_1z_2) = c^T (z_1z_2I - A_1 \otimes A_2)^{-1} b_1 \otimes A_2^{j-i} b_2 \frac{(z_1z_2)^i g(z_1z_2)}{q(z_1z_2)} \odot \sum (z_1z_2)^{-k} \quad (2.5.2)$$

We can now immediately see from our knowledge of linear systems that if $\det(z_1z_2I - A_1 \otimes A_2)$ and $q(z_1z_2)$ both have zeros within the unit circle, then the output sequence due to any finite input sequence from $U \times V$ will be ℓ_1-stable. Now q(z) has zeros which are either zero or else zeros of

$p(z)$, and $\det(zI - A_1 \otimes A_2)$ has zeros which are either zero or else of the form $\{\alpha_i \beta_j\}$ where $\{\alpha_i\}$ are the zeros of $p_1(z_1)$ and $\{\beta_j\}$ are the zeros of $p_2(z_2)$. Hence (i) is a sufficient condition for ℓ_1-stability.

Likewise, we can see from (2.5.2) that if $g(z)$ cancels all zeros of $\det(zI - A_1 \otimes A_2)$ which lie on or outside the unit circle, then the output sequence $y(z_1 z_2)$ is ℓ_1-stable. Hence (ii) is a sufficient condition for ℓ_1-stability.

CHAPTER 3. CANONICAL REALIZATIONS OF BILINEAR INPUT/OUTPUT MAPS

In this chapter we analyse state space representations of bilinear input/output maps in greater depth. The motivation for this is that state space representations will in general be neither controllable nor observable, and it may be helpful for the purpose of identification of parameters to be able to construct a realization possessing the properties of controllability and observability.

For the case of linear discrete-time systems it is common to talk about state space reachability rather than state space controllability (where controllability refers to zero state controllability) since a zero-eigenvalue mode which is unaffected by inputs will certainly attain zero value in finite time. For this reason reachability rather than controllability is considered here as well, but as we shall see, it is necessary to relax the concept of reachability to that of quasi-reachability, and in §3.2 necessary and sufficient conditions are obtained for a state space realization of a bilinear input/output map to be quasi-reachable.

Observability too has to be treated in a slightly different manner from that of linear systems, and the idea of a realization being observable if its initial state can be determined with the help of a finite number of "experiments" has to be invoked. Necessary and sufficient conditions are obtained in §3.3 for a state space realization of a bilinear input/output map to be observable.

In §3.1, formal definitions of these and other concepts are introduced, as are the similarity transformations on the state space which produce equivalent realizations of bilinear input/output maps. In Chapter 4 it will be shown that any two minimal realizations (Definition 3.1.4) of a bilinear map f are isomorphic under these transformations.

3.1 Preliminaries

Definition 3.1.1

A state space realization of an input/output map is quasi-reachable if the closure of the set of states reachable from the zero state is the whole space.

Definition 3.1.2

A state space realization of an input/output map is observable if no two states are equivalent. We say that two states x_1 and x_2 are equivalent if $f(x_1,w) = f(x_2,w)$ for all w where $f: X \times W \to Y$ represents the map from an initial state $x \in X$ and an input sequence $w \in W$ to the output space Y.

Definition 3.1.3

A state space realization is canonical if it is both quasi-reachable and observable.

Definition 3.1.4 [AM1]

A state space realization M of an input/output map f is (co)minimal if it is observable, and if for every other observable realization M' of f, there exists a unique mapping $\emptyset: M \to M'$.

We now reintroduce the state space realization first mentioned in Chapter 2:

$$x^1_{k+1} = A_1 x^1_k + b_1 u_k \tag{3.1.1}$$

$$x^2_{k+1} = A_2 x^2_k + b_2 v_k \tag{3.1.2}$$

$$x_{k+1} = Ax_k + Cx^1_k \otimes x^2_k + Q_1 x^1_k v_k + Q_2 x^2_k u_k + bu_k v_k \tag{3.1.3}$$

$$y_k = h^T x_k + d^T x^1_k \otimes x^2_k \tag{3.1.4}$$

where $x^1_k \in R^{n_1}$, $x^2_k \in R^{n_2}$, $x_k \in R^n$, and the system matrices have dimension consistent with these.

Before going on to discuss reachability and observability in §3.2

and §3.3, it is of interest to discover the class of transformations on this state space which preserves the behaviour of the system, and it is this which provides the setting for the reduction procedures of Chapter 4.

From our knowledge of linear system theory, it is immediately obvious that there exist three particular classes of similarity transformations which preserve the behaviour of the system, namely $x_k^1 \to T_1 x_k$, $x_k^2 \to T_2 x_k^2$ and $x_k \to T x_k$, where T_1, T_2 and T are non-singular square matrices, with the associated transformations

$$A_1 \to T_1 A_1 T_1^{-1} \quad b_1 \to T_1 b_1 \quad A_2 \to T_2 A_2 T_2^{-1} \quad b_2 \to T_2 b_2$$
$$A \to TAT^{-1} \quad C \to TCT_1^{-1} \otimes T_2^{-1} \quad Q_1 \to TQ_1 T_1^{-1} \quad Q_2 \to TQ_2 T_2^{-1} \quad b \to Tb$$
$$h^T \to h^T T^{-1} \quad d^T \to d^T T_1^{-1} \otimes T_2^{-1}.$$

However there is one further similarity transformation which is not so clearly apparent:

Proposition 3.1.1

Let (3.1.1)-(3.1.4) be a realization of the bilinear input/output map $f : U \times V \to Y$. Then for any $W \in R^{n \times n_1 n_2}$, (3.1.1)-(3.1.4) is also a realization of f under the transformation

$$C \to W(A_1 \otimes A_2) + C - AW$$
$$Q_1 \to Q_1 + W(A_1 \otimes b_2) \qquad Q_2 \to Q_2 + W(b_1 \otimes A_2)$$
$$b \to b + W(b_1 \otimes b_2) \qquad d^T \to d^T - h^T W.$$

Proof: We calculate the transfer function $s_T(z_1,z_2)$ of the transformed system according to the methods of Chapter 2, and show that it is equal to the original transfer function $s(z_1,z_2)$.

The transfer function $s_T(z_1,z_2)$ is given by

$$\begin{aligned} s_T(z_1,z_2) = h^T(z_1 z_2 I - A)^{-1}\{&[W(A_1 \otimes A_2) + C - AW](z_1 I - A_1)^{-1} b_1 \otimes (z_2 I - A_2)^{-1} b_2 \\ &+ [Q_1 + W(A_1 \otimes b_2)](z_1 I - A_1)^{-1} b_1 \\ &+ [Q_2 + W(b_1 \otimes A_2)](z_2 I - A_2)^{-1} b_2 + b + W(b_1 \otimes b_2)\} \\ + (d^T - h^T W)(z_1 I - A_1)^{-1} & b_1 \otimes (z_2 I - A_2)^{-1} b_2 \end{aligned}$$

$$= h^T(z_1z_2I-A)^{-1}\{C(z_1I-A_1)^{-1}b_1 \otimes (z_2I-A_2)^{-1}b_2 + Q_1(z_1I-A_1)^{-1}b_1 + Q_2(z_2I-A_2)^{-1}b_2+b\}$$
$$+ d^T(z_1I-A_1)^{-1}b_1 \otimes (z_2I-A_2)^{-1}b_2$$
$$+ h^T(z_1z_2I-A)^{-1}\{[W(A_1 \otimes A_2) - AW](z_1I-A_1)^{-1}b_1 \otimes (z_2I-A_2)^{-1}b_2$$
$$+ W(A_1 \otimes b_2)(z_1I-A_1)^{-1}b_1 + W(b_1 \otimes A_2)(z_2I-A_2)^{-1}b_2 + W(b_1 \otimes b_2)\}$$
$$- h^TW(z_1I-A_1)^{-1}b_1 \otimes (z_2I-A_2)^{-1}.$$

Then using the identity $(zI-F)^{-1}F = -I + z(zI-F)^{-1} = F(zI-F)^{-1}$ and the expression for $s(z_1,z_2)$ given in (2.4.5) we obtain

$$s_T(z_1,z_2) = s(z_1,z_2) + h^T(z_1z_2I-A)^{-1}\{W(-I+z_1(z_1I-A_1)^{-1})b_1 \otimes (-I+z_2(z_2I-A_2)^{-1})b_2$$
$$+ W(-I+z_1(z_1I-A_1)^{-1})b_1 \otimes b_2$$
$$+ Wb_1 \otimes (-I+z_2(z_2I-A_2)^{-1})b_2 + Wb_1 \otimes b_2$$
$$+ h^T(I-z_1z_2(z_1z_2I-A)^{-1})W(z_1I-A_1)^{-1}b_1 \otimes (z_2I-A_2)^{-1}b_2$$
$$- h^TW(z_1I-A_1)^{-1}b_1 \otimes (z_2I-A_2)^{-1}b_2$$
$$= s(z_1,z_2).$$

□

Remark: This transformation is equivalent to the similarity transformation $\begin{pmatrix} I \otimes I & O \\ W & I \end{pmatrix}$ applied to the linear system defined by

$$[d^T\ h^T]\ ,\ \begin{pmatrix} A_1 \otimes A_2 & O \\ C & A \end{pmatrix}\ ,\ \begin{pmatrix} A_1 \otimes b_2 & b_1 \otimes A_2 & b_1 \otimes b_2 \\ Q_1 & Q_2 & b \end{pmatrix}$$

In Chapter 4 we shall see that these four classes of similarity transformation define the isomorphism between two minimal realizations of a bilinear input/output map.

3.2 Reachability of the State Space

In Chapter 2, we mentioned the intuitive idea of reachability in bounded time and in [FM1] it was shown that this is equivalent to the existence of a finite-dimensional state-space realization. Here we bring in some intuitive ideas of state space reachability and demonstrate that within certain restrictions they do indeed hold.

We first of all digress for a moment to discuss linear systems. By a well-known theorem we know that the system

$$x_{k+1} = Fx_k + gu_k$$

is not reachable iff there exists a row vector a^T such that $a^T g = 0$ and $a^T F = \lambda a^T$ for some $\lambda \in \mathcal{C}$. In other words, $a^T x_{k+1} = \lambda a^T x_k$, and given a zero initial state, the state space evolves on the hyperplane $a^T x = 0$.

With bilinear systems, using a certain amount of intuitive reasoning, we may expect the state space to evolve on some hyper surface $p^T x + q^T x^1 \otimes x^2 = 0$ if the state space realization (3.1.1)-(3.1.3) is not reachable. To be more precise, we expect that

$$p^T x_{k+1} + q^T x^1_{k+1} \otimes x^2_{k+1} = \lambda (p^T x_k + q^T x^1_k \otimes x^2_k)$$

identically, for some $\lambda \in \mathcal{C}$.

In fact, we shall see that, subject to certain assumptions detailed in Theorem 3.2.1, this condition is both necessary and sufficient for non-reachability.

Before we come to the main body of this section, we recall the following definitions from linear system theory. Let $F \in \mathcal{R}^{n \times n}$, $H \in \mathcal{R}^{r \times n}$, $G \in \mathcal{R}^{n \times m}$. Then

(i) (F,G) is a reachable pair iff $\text{rank}[G\ FG\ \dots\ F^{n-1}G] = n$, and

(ii) (H,F) is an observable pair iff $\text{rank}\,[H^T F^T H^T \dots (F^{n-1})^T H^T] = n$.

We shall now make the following two assumptions concerning the state-space description (3.1.1)-(3.1.4):

(A1) (A_1, b_1) and (A_2, b_2) are reachable pairs

(A2) (h^T, A) is an observable pair.

If either of these assumptions does not hold, we know from the well-known linear system theory results of Kalman [K2] how to reduce

(3.1.3)-(3.1.4) to suit our requirements. In particular, assumption (A2) tells us that if we diagonalize the matrix A into Jordan form, then there is only one Jordan block corresponding to each distinct eigenvalue of A, and hence just one Jordan block corresponding to zero eigenvalues of A.

We now state the following technical lemma concerning the transfer functions of the system.

Lemma 3.2.1

Let $\begin{bmatrix} A_1 \otimes A_2 & 0 \\ C & A \end{bmatrix}, \begin{bmatrix} A_1 \otimes b_2 & b_1 \otimes A_2 & b_1 \otimes b_2 \\ Q_1 & Q_2 & b \end{bmatrix}$ be a reachable pair.

Then the components of $x^1(z_1) \otimes x^2(z_2)$ and $x(z_1,z_2)$ are linearly independent, where

$$x^1(z_1) = (z_1 I - A_1)^{-1} b_1 \quad (3.2.1)$$

$$x^2(z_2) = (z_2 I - A_2)^{-1} b_2 \quad (3.2.2)$$

$$x(z_1,z_2) = (z_1 z_2 I - A)^{-1} [C(z_1 I - A_1)^{-1} b_1 \otimes (z_2 I - A_2)^{-1} b_2 + Q_1 (z_1 I - A_1)^{-1} b_1 + Q_2 (z_2 I - A_2)^{-1} b_2 + b] \quad (3.2.3)$$

are the transfer functions of x_k^1, x_k^2 and x_k respectively.

Proof: Suppose there exist row vectors p^T and q^T such that

$$p^T x^1(z_1) \otimes x^2(z_2) + q^T x(z_1,z_2) = 0.$$

Expanding $x(z_1,z_2)$ in powers of $z_1^{-i} z_2^{-j}$ we obtain

$$x(z_1,z_2) = \sum_{k \geq 0} (z_1 z_2)^{-(k+1)} A^k [C \sum_{i,j \geq 0} (A_1^i \otimes A_2^j)(b_1 \otimes b_2) z_1^{-(i+1)} z_2^{-(j+1)} + Q_1 \sum_{i \geq 0} A_1^i b_1 z_1^{-(i+1)} + Q_2 \sum A_2^j b_2 z_2^{-(j+1)} + b].$$

The coefficient of $(z_1 z_2)^{-(r+1)}$ is then

$$\begin{bmatrix} 0 & I \end{bmatrix} \begin{bmatrix} A_1 \otimes A_2 & 0 \\ C & A \end{bmatrix}^r \begin{bmatrix} b_1 \otimes b_2 \\ b \end{bmatrix} \quad r = 0,1,\ldots$$

The coefficient of $(z_1 z_2)^{-(r+1)} z_1^{-(s+1)}$ is

$$\begin{bmatrix} 0 & I \end{bmatrix} \begin{bmatrix} A_1 \otimes A_2 & 0 \\ C & A \end{bmatrix}^r \begin{bmatrix} A_1^{s+1} b_1 \otimes b_2 \\ Q_1 A_1^s b_1 \end{bmatrix} \quad r,s = 0,1,\ldots$$

The coefficient of $(z_1z_2)^{-(r+1)}z_2^{-(s+1)}$ is

$$[0\ I]\begin{bmatrix} A_1\otimes A_2 & 0 \\ C & A\end{bmatrix}^r \begin{bmatrix} b_1\otimes A_2^{s+1}b_2 \\ Q_2A_2^s b_2\end{bmatrix} \quad r,s = 0,1,\ldots$$

Similarly the coefficient of $(z_1z_2)^{-(r+1)}$ in $x_1(z_1)\otimes x^2(z_2)$ is

$$(A_1\otimes A_2)^r b_1\otimes b_2 \quad r = 0,1,\ldots$$

the coefficient of $(z_1z_2)^{-(r+1)}z_1^{-(s+1)}$ is

$$(A_1\otimes A_2)^r A_1^{s+1}b_1\otimes b_2 \quad r,s = 0,1,\ldots$$

the coefficient of $(z_1z_2)^{-(r+1)}z_2^{-(s+1)}$ is

$$(A_1\otimes A_2)^r b_1\otimes A_2^{s+1}b_2 \quad r,s = 0,1,\ldots .$$

So $p^Tx^1\otimes x^2 + q^Tx = 0$ implies that

$$[p^Tq^T]\begin{bmatrix} A_1\otimes A_2 & 0 \\ C & A\end{bmatrix}^r \begin{bmatrix} b_1\otimes b_2 \\ b\end{bmatrix} = 0 \quad r = 0,1,\ldots \tag{3.2.4}$$

$$[p^Tq^T]\begin{bmatrix} A_1\otimes A_2 & 0 \\ C & A\end{bmatrix}^r \begin{bmatrix} A_1^{s+1}b_1\otimes b_2 \\ Q_1A_1^s b_1\end{bmatrix} = [p^Tq^T]\begin{bmatrix} A_1\otimes A_2 & 0 \\ C & A\end{bmatrix}^r \begin{bmatrix} A_1\otimes b_2 \\ Q_1\end{bmatrix} A_1^s b_1 = 0 \quad r,s = 0,1,\ldots \tag{3.2.5}$$

$$[p^Tq^T]\begin{bmatrix} A_1\otimes A_2 & 0 \\ C & A\end{bmatrix}^r \begin{bmatrix} b_1\otimes A_2^{s+1}b_2 \\ Q_2A_2^s b_2\end{bmatrix} = [p^Tq^T]\begin{bmatrix} A_1\otimes A_2 & 0 \\ C & A\end{bmatrix}^r \begin{bmatrix} b_1\otimes A_2 \\ Q_2\end{bmatrix} A_2^s b_2 = 0 \quad r,s = 0,1,\ldots \tag{3.2.6}$$

But (A_1,b_1) and (A_2,b_2) are reachable pairs, and hence (3.2.5) and (3.2.6) reduce to

$$[p^Tq^T]\begin{bmatrix} A_1\otimes A_2 & 0 \\ C & A\end{bmatrix}^r \begin{bmatrix} A_1\otimes b_2 \\ Q_1\end{bmatrix} = 0 \quad r = 0,1,\ldots$$

$$[p^Tq^T]\begin{bmatrix} A_1\otimes A_2 & 0 \\ C & A\end{bmatrix}^r \begin{bmatrix} b_1\otimes A_2 \\ Q_2\end{bmatrix} = 0 \quad r = 0,1,\ldots$$

which together with (3.2.4) provide a contradiction to

$$\begin{bmatrix} A_1\otimes A_2 & 0 \\ C & A\end{bmatrix}, \begin{bmatrix} A_1\otimes b_2 & b_1\otimes A_2 & b_1\otimes b_2 \\ Q_1 & Q_2 & b\end{bmatrix}$$

being a reachable pair. $\square$

We now state the main result of this chapter:

Theorem 3.2.1

The system (3.1.1)-(3.1.4), with assumptions (A1) and (A2), is quasi-reachable iff

$$\begin{bmatrix} A_1 \otimes A_2 & 0 \\ C & A \end{bmatrix}, \begin{bmatrix} A_1 \otimes b_2 & b_1 \otimes A_2 & b_1 \otimes b_2 \\ Q_1 & Q_2 & b \end{bmatrix} \triangleq (F,G)$$

is a reachable pair.

Proof: Suppose that (F,G) is not a reachable pair. Then there exist row vectors p and q such that

$$[p^T \; q^T] \begin{bmatrix} A_1 \otimes A_2 & 0 \\ C & A \end{bmatrix} = \lambda [p^T \; q^T] \text{ and } [p^T \; q^T] \begin{bmatrix} A_1 \otimes b_2 & b_1 \otimes A_2 & b_1 \otimes b_2 \\ Q_1 & Q_2 & b \end{bmatrix} = 0$$

and by expanding x_{k+1} and $x^1_{k+1} \otimes x^2_{k+1}$ in terms of x_k, x^1_k, x^2_k, u_k and v_k it is clear that $p^T x_{k+1} + q^T x^1_{k+1} \otimes x^2_{k+1} = \lambda (p^T x_k + q^T x^1_k \otimes x^2_k)$.

Hence, given a zero initial state, i.e. $x_o = 0$, $x^1_o = 0$, $x^2_o = 0$, we see that the state space evolves on the hypersurface $p^T x + q^T x^1 \otimes x^2 = 0$ for all time, so that the system is certainly not quasi-reachable.

Conversely, suppose that (F,G) is a reachable pair. We shall now proceed to show quasi-reachability of the state space using a similar approach to that of [K2]. This we do by specifying a desired state, and then constructing input sequences from $U \times V$ which reach this desired state at time +1. Note that the state x_k at time +1 is given by the vector coefficient of $(z_1 z_2)^{-1}$ in the expansion of $x(z_1,z_2)u(z_1)v(z_2)$, because x_k is a bilinear function of U and V. We are of course assuming that at some time -J, where J is greater than the length of the input sequence that we shall construct, we have x^1_{-J}, x^2_{-J} and x_{-J} all zero.

Now let $\psi_1(z)$ and $\psi_2(z)$ be the characteristic polynomials of A_1 and A_2 respectively. Then, given desired states x^1_1 and x^2_1 we know from linear system theory that there exist unique input sequences

$q_1(z_1)$ and $q_2(z_2)$ with deg q_i < deg ψ_i (i = 1,2) such that the input sequences $p_1(z_1)\psi_1(z_1) + q_1(z_1)$ and $p_2(z_2)\psi_2(z_2) + q_2(z_2)$ applied to (3.1.1) and (3.1.2) respectively reach x_1^1 and x_1^2 for all $p_1(z_1)$ and $p_2(z_2)$.

Hence, once the desired state $\begin{bmatrix} x_1^1 \\ x_1^2 \\ x_1 \end{bmatrix}$ is specifed, the reachability problem becomes one of constructing polynomials $p_1(z_1)$ and $p_2(z_2)$ which enable us to reach the state x_1 via (3.1.1)-(3.1.3).

The construction of these polynomials is fairly long and detailed, so we shall first outline the two major remaining stages of the proof:

1. Using a suitable choice of matrix T, we apply a similarity transformation to equation (3.1.3) in such a way that

$$TAT^{-1} = \begin{bmatrix} J_1 & 0 \\ 0 & Jo \end{bmatrix} \tag{3.2.7}$$

where $Jo = \begin{bmatrix} 0 & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & 0 \end{bmatrix} \in R^{mxm}$ and J_1 is non-singular.

We then show in Lemma 3.2.2 that the subsystem corresponding to Jo, together with equations (3.1.1) and (3.1.2) is quasi-reachable and we show how to construct the input sequences necessary to achieve the desired state.

2. We then construct a further input sequence with the aid of another technical lemma, which enables us to reach the remaining desired components of x_1. In fact it becomes clear that if A has no zero eigenvalues then the state space is not only quasi-reachable but completely reachable as well.

Let us now consider the special case A = Jo, where $Jo \in R^{mxm}$.

Lemma 3.2.2

The state space realization (3.1.1)-(3.1.4) is quasi reachable iff

(F,G), as defined in Theorem 3.2.1, is a reachable pair, where A = Jo.

Proof: The transfer function $x(z_1,z_2)$ is calculated as

$$(z_1z_2I - Jo)x(z_1z_2) = [C(z_1I - A_1)^{-1}a(z_2I - A_2)^{-1}b_2 + Q_1(z_1I - A_1)^{-1}b_1 + Q_2(z_2I - A_2)^{-1}b_2 + b] \quad (3.2.8)$$

The RHS of (3.2.8) can be written as the vector

$$\frac{1}{\psi_1(z_1)\psi_2(z_2)}\begin{bmatrix} R_1(z_1,z_2) \\ \vdots \\ R_m(z_1,z_2) \end{bmatrix}$$

so that

$$x(z_1,z_2) \underline{\Delta} \begin{bmatrix} x_1(z_1,z_2) \\ \vdots \\ x_m(z_1,z_2) \end{bmatrix} = \frac{1}{(z_1z_2)^m}\begin{bmatrix} (z_1z_2)^{m-1} \ldots\ldots 1 \\ \vdots \\ 0 \qquad (z_1z_2)^{m-1} \end{bmatrix}\begin{bmatrix} R_1(z_1,z_2) \\ \\ R_m(z_1,z_2) \end{bmatrix}\frac{1}{\psi_1(z_1)\psi_2(z_2)} \quad (3.2.9)$$

Let us now examine the state sequence from time + 1 onwards due to the input sequence

$$[(\alpha_0+\alpha_1z_1+\ldots+\alpha_sz_1^s)\psi_1(z_1) + q_1(z_1)][(\beta_0+\beta_1z_1+\ldots+\beta_tz_2^t)\psi_2(z_2) + q_2(z_2)]$$
$$\underline{\Delta}\ [\alpha(z_1)\psi_1+q_1][\beta(z_2)\psi_2+q_2]\ .$$

We shall label this state sequence $y(z_1z_2) = [y_1(z_1z_2)\ldots y_m(z_1z_2)]^T$ so that

$$y_{m-r}(z_1z_2) = x_{m-r}(z_1,z_2)[\alpha(z_1)\psi_1+q_1][\beta(z_2)\psi_2+q_2]\odot\sum(z_1z_2)^{-k} \qquad (r = 0,\ldots,m-1)$$

$$= \frac{1}{\psi_1(z_1)\psi_2(z_2)}\left[\frac{R_{m-r}}{z_1z_2} + \ldots + \frac{R_m}{(z_1z_2)^{r+1}}\right][\alpha(z_1)\psi_1+q_1][\beta(z_2)\psi_2+q_2] \odot \sum(z_1z_2)^{-k} \qquad (r = 0,\ldots,m-1)$$

On examination of the terms to the left of $\odot$ which involve α_i and β_j for i and j greater than r, it is clear that these do not contribute to $y_{m-r}(z_1z_2)$, since $\alpha_iz_1^i\psi_1(z_1)$ cancels out all z_1 terms in the denominator for $i > r$, and $\beta_jz_2^j\psi_2(z_2)$ cancels out all z_2 terms in the denominator for $j > r$. Hence

$$y_{m-r}(z_1z_2) = \frac{1}{\psi_1(z_1)\psi_2(z_2)}\left[\frac{R_{m-r}}{z_1z_2} + \dots + \frac{R_m}{(z_1z_2)^{r+1}}\right] \times$$

$$[(\alpha_o + \dots + \alpha_r z_1^r)\psi_1 + q_1][(\beta_o + \dots + \beta_r z_2^r)\psi_2 + q_2] \circledcirc \sum(z_1z_2)^{-k}$$

$$(r = 0,\dots,m-1)$$

In addition, it is clear that all terms involving multiplication of α_r and $R_{m-r},\dots,R_{m-1}$ make no contribution to $y_{m-r}(z_1z_2)$ since once again we have a cancellation of all z_1 terms in the denominator. The same goes for multiplication of β_r with $R_{m-r},\dots,R_{m-1}$, so using the bilinearity principle we can now write

$$y_{m-r}(z_1z_2) = y^q_{m-r}(z_1z_2) + \frac{R_m}{(z_1z_2)^{r+1}\psi_1(z_1)\psi_2(z_2)}$$

$$\times\Big[\alpha_r\beta_r(z_1z_2)^r\psi_1\psi_2$$

$$+\alpha_r z_1^r\psi_1[(\beta_o + \dots + \beta_{r-1}z_2^{r-1})\psi_2 + q_2]$$

$$+\beta_r z_2^r\psi_2[(\alpha_o + \dots + \alpha_{r-1}z_1^{r-1})\psi_1 + q_1]\Big]$$

$$\circledcirc \sum(z_1z_2)^{-k}$$

$$(r = 0,\dots,m-1)$$

where $y^q_{m-r}(z_1z_2)$ is just $y_{m-r}(z_1z_2)$ for $\alpha_r = \beta_r = 0$.

We can simplify this to

$$y_{m-r}(z_1z_2) = y^q_{m-r}(z_1z_2) + \frac{R_m}{z_1z_2}\left[\alpha_r\beta_r + \frac{\alpha_r(\beta_o + \dots + \beta_{r-1}z_2^{r-1})}{z_2^r} + \frac{\alpha_r q_2(z_2)}{z_2^r\psi_2(z_2)}\right.$$

$$\left. + \frac{\beta_r(\alpha_o + \dots + \alpha_{r-1}z_1^{r-1})}{z_1^r} + \frac{\beta_r q_1(z_1)}{z_1^r\psi_1(z_1)}\right] \circledcirc \sum(z_1z_2)^{-k} \qquad (3.2.10)$$

$$(r = 0,\dots,m-1)$$

We immediately notice that any terms of $R_m(z_1,z_2)$ with a factor of z_1z_2 make no contribution, since this factor cancels with the z_1z_2 term in the denominator outside the square brackets of (3.2.10), and all terms inside the square brackets have denominator with terms either in z_1 or in z_2, but not involving both z_1 and z_2. Hence the only terms of $R_m(z_1,z_2)$ which contribute to $y_{m-r}(z_1z_2)$ are those of the form

$$a_1z_1 + \dots + a_{n_1}z_1^{n_1} + b_1z_2 + \dots + b_{n_2}z_2^{n_2} + c \triangleq a(z_1) + b(z_2) + c.$$

Let us now write

$$\frac{a(z_1)q_1(z_1)}{\psi_1(z_1)} = f_{n_1-1}z_1^{n_1-1} + \ldots + f_1 z_1 + f_o + \text{terms in } z_1^{-1} \qquad (3.2.11)$$

and

$$\frac{b(z_2)q_2(z_2)}{\psi_2(z_2)} = g_{n_2-1}z_2^{n_2-1} + \ldots + g_1 z_2 + g_o + \text{terms in } z_2^{-1}. \qquad (3.2.12)$$

It then follows from (3.2.10) that

$$y_{m-r}(z_1z_2) = y^q_{m-r}(z_1z_2) + (z_1z_2)^{-1}[\alpha_r\beta_r c + \alpha_r(\beta_o b_r + \beta_1 b_{r-1} + \ldots + \beta_{r-1}b_1) + \alpha_r g_r + \beta_r(\alpha_o a_r + \alpha_1 a_{r-1} + \ldots + \alpha_{r-1}a_0) + \beta_r f_r] \qquad (3.2.13)$$

$$(r = 0,\ldots,m-1)$$

As we remarked earlier, the term of interest to us is the coefficient of $(z_1z_2)^{-1}$, and it is clear from (3.2.13) that if c is non-zero, by suitable choice of α_r and β_r we can achieve any desired value of this coefficient. Hence if c is non-zero, not only do we have quasi-reachability of (3.1.1)-(3.1.4), but complete reachability.

Alternatively, suppose $c = 0$. We then have two cases to consider:

(1) either $a(z_1) = 0$ or $b(z_2) = 0$,

or (2) neither $a(z_1)$ nor $b(z_2)$ are identically zero.

Note that if both $a(z_1)$ and $b(z_2)$ (as well as c) are zero, then we can write $R_m(z_1,z_2)$ as $z_1z_2R'_m(z_1,z_2)$, in which case we see from (3.2.9) that

$$x_m(z_1,z_2) = \frac{1}{z_1z_2}R_m(z_1,z_2)\frac{1}{\psi_1(z_1)\psi_2(z_2)} = \frac{R'_m(z_1,z_2)}{\psi_1(z_1)\psi_2(z_2)}$$

which is linearly dependent on the components of $x^1(z_1)\otimes x^2(z_2)$, in contradiction of Lemma 3.2.1.

Let us first assume, then, that $a(z_1) \neq 0$ and $b(z_2) \neq 0$, and suppose that

$$a_i = 0 \ (i=1,\ldots,s_1) \text{ and } a_{s_1+1} \neq 0$$

$$\text{and } b_j = 0 \ (j=1,\ldots,s_2) \text{ and } b_{s_2+1} \neq 0$$

for some $s_1 \leq m_1$, $s_2 \leq m_2$. Then from (3.2.13) we have

$$\bar{y}_{m-r} = \bar{y}^q_{m-r} + \alpha_r g'_r + \beta_r f'_r \quad (r = 0,\ldots,m-1) \tag{3.2.14}$$

where $\bar{y}_{m-r}$ and $\bar{y}^q_{m-r}$ are the coefficients of $(z_1z_2)^{-1}$ in $y_{m-r}(z_1z_2)$ and $y^q_{m-r}(z_1z_2)$ respectively and

$$g'_r = \beta_o b_r + \beta_1 b_{r-1} + \ldots + \beta_{r-1} b_1 + g_r \quad r = 0,\ldots,m-1 \tag{3.2.15}$$

$$f'_r = \alpha_o a_r + \alpha_1 a_{r-1} + \ldots + \alpha_{r-1} a_1 + f_r \quad r = 0,\ldots,m-1 \ . \tag{3.2.16}$$

From (3.2.14) we see that a sufficient condition for reachability is that g'_r and f'_r are non-zero for $r=0,\ldots,m-1$. However from (3.2.15) and (3.2.16) it is clear that $g'_r = g_r (r=0,\ldots,s_2)$ and $f'_r = f_r (r=0,\ldots,s_1)$, so if $g_r = 0$ and $f_r = 0$ for any of these values of r, we have $\bar{y}_{m-r} = \bar{y}^q_{m-r}$, so the state space is not reachable. We therefore constrain $g_r(r=0,\ldots,s_2)$ and $f_r(r=0,\ldots,s_1)$ to be non-zero; from (3.2.11) and (3.2.12) we see that this is just a restriction on the coefficients of $q_2(z_2)$ and $q_1(z_1)$ not to lie within a certain union of hyperplanes of R^{n_2} and R^{n_1} respectively, and this in turn is a restriction on x^2_1 and x^1_1 not to lie within a certain union of hyperplanes also in R^{n_2} and R^{n_1}. If we can now show that we can attain any value of x_1 provided that x^1_1 and x^2_1 do not lie in the hyperplanes characterized by the above discussion, it then follows that the closure of the reachable set of (3.1.1)-(3.1.4) is the whole space $R^{n_1+n_2+n}$, so that the system is quasi-reachable. To do this, we just need to ensure that $g'_k(k>s_2)$ and $f'_k(k>s_1)$ are non-zero. Now from (3.2.15) and (3.2.16) we see that

$$g'_{s_1+r+1} = \beta_o b_{s_2+r+1} + \ldots + \beta_{r-1} b_{s_2+2} + \beta_r b_{s_2+1} + g_{s_2+r+1}$$

$$(r = 0,\ldots,m-s_2-1)$$

$$f'_{s_1+r+1} = \alpha_o a_{s_1+r+1} + \ldots + \alpha_{r-1} a_{s_1+2} + \alpha_r a_{s_1+1} + f_{s_1+r+1}$$

$$(r = 0,\ldots,m-s_1-1)$$

so our methodology is to choose α_r and β_r in such a way that f'_{s_1+r+1} and g'_{s_2+r+1} are non-zero, and at the same time ensure that the desired value of $\bar{y}_{m-r}$ in (3.2.14) is achieved. It is readily seen that we can attain these objectives, so quasi-reachability is proven.

Now consider $b(z_2) = 0$. From (3.2.12) we see that $g_r = 0$ for all r, so the coefficient of α_r in (3.2.14) vanishes. Then, as above, we restrict f_i $(i \leq s_1)$ to be non-zero; we then choose α_r $(r = 0, m-s_1-1)$ arbitrarily, since it makes no contribution to $\bar{y}_{m-r}$, to ensure that f'_{s_1+r+1} is non-zero. Finally we choose α_r so that the desired value of $\bar{y}_{m-r}$ is attained. Hence we have quasi-reachability.

We follow an analogous argument for the case $a(z_1) = 0$. □

Let us now return to the proof of Theorem 3.2.1; we have seen from Lemma 3.2.2 that we can achieve all states associated with the subsystem of x_k corresponding to zero eigenvalues, and it is clear from the proof of this lemma that all inputs of the form $z_1^{t_1}\psi_1(z_1)$ and $z_2^{t_2}\psi_2(z_2)$ have no influence on that subsystem or on the x_k^1 and x_k^2 states for $t_1, t_2 > m-1$.

If we now calculate the transfer function of the remaining x_k states, it is clear that this will be of the form

$$\hat{x}(z_1,z_2) = \frac{S(z_1,z_2)}{\phi(z_1z_2)\psi_1(z_1)\psi_2(z_2)} \in R^{n-m}[(z_1,z_2)] \qquad (3.2.17)$$

where, and from (3.2.7), $\phi(z)$ is the characteristic polynomial of $J_1, \phi(z) = z^{n-m} + \ldots + \phi_1 z + \phi_0$.

We can write this as

$$\hat{x}(z_1,z_2) = \frac{S(z_1,z_2)(z_1z_2)^m}{\phi(z_1z_2)\bar{\psi}_1(z_1)\bar{\psi}_2(z_2)} \qquad (3.2.18)$$

where $\bar{\psi}_1 = z_1^m\psi_1$ and $\bar{\psi}_2 = z_2^m\psi_2$.

Now consider constructing an input sequence of the form

$$(p_1(z_1)\bar{\psi}_1(z_1) + \bar{q}_1(z_1)) \times (p_2(z_2)\bar{\psi}_2(z_2) + \bar{q}_2(z_2))$$

where $\bar{q}_1 = \alpha(z_1)\psi_1 + q_1$ and $\bar{q}_2 = \beta(z_2)\psi_2 + q_2$.

Let the vector bilinear map (3.2.17) be represented by

$$g:\ R[z_1] \times R[z_2] \longrightarrow R^{n-m}[[(z_1z_2)^{-1}]]$$

$$:\ (u(z_1),v(z_2)) \longrightarrow \hat{x}(z_1,z_2)u(z_1)v(z_2) \odot \sum (z_1z_2)^{-k}.$$

Then

$$g(p_1\bar{\Psi}_1+\bar{q}_1, p_2\bar{\Psi}_2+\bar{q}_2) = g(p_1\bar{\Psi}_1, p_2\bar{\Psi}_2) + g(p_1\bar{\Psi}_1, \bar{q}_2) + g(\bar{q}_1, p_2\bar{\Psi}_2) + g(\bar{q}_1, \bar{q}_2). \quad (3.2.19)$$

We shall now set up p_1 and p_2 in such a way that the two middle terms of (3.2.19) become zero, and then concentrate on demonstrating reachability via $g(p_1\bar{\Psi}_1, p_2\bar{\Psi}_2)$. The reason for doing this, as we shall see very shortly, is to enable us to select the coefficients of $p_2(z_2)$ by solving a series of linear equations.

Consider then

$$g(p_1\bar{\Psi}_1, \bar{q}_2) = \frac{(z_1z_2)^m S(z_1z_2)}{\phi\bar{\Psi}_1\bar{\Psi}_2},\ p_1\bar{\Psi}_1\bar{q}_2 \odot \sum (z_1z_2)^{-k}$$

$$= \frac{(z_1z_2)^m S(z_1z_2)}{\phi\bar{\Psi}_2}\ p_1\bar{q}_2 \odot\ (z_1z_2)^{-k}.$$

Now $S(z_1,z_2) \triangleq [S_1,\ldots,S_{n-m}]^T$ is made up of terms $\varepsilon_{ij}z_1^iz_2^j$.

Let us now define

$$m_1 = \max\{j-i \mid \varepsilon_{ij} \neq 0,\ \varepsilon_{ij}z_1^iz_2^j \text{ occurs in one of } S_1,\ldots,S_{n-m}\}$$

and let

$$p_1(z_1) = z_1^{m_1}\bar{p}_1(z_1)$$

for some $\bar{p}_1$ to be constructed later.

$$\text{Then } \frac{(z_1z_2)^m Sp_1\bar{q}_2}{\phi\bar{\Psi}_2} = \frac{(z_1z_2)^m S\ z_1^{m_1}\bar{p}_1\ \sum\gamma_i z_2^{-i}}{\phi(z_1z_2)} \quad (3.2.20)$$

where $\sum\limits_{i\geq 1}\gamma_i z_2^{-i}$ is the expansion of $\frac{\bar{q}_2(z_2)}{\bar{\Psi}_2(z_2)}$ in negative powers of z_2.

It is now clear that because of our choice of m_1, all terms in the numerator are of the form $a_{k\ell}z_1^kz_2^\ell$ with $k > \ell$ for all $a_{k\ell} \neq 0$. Hence the expansion of (3.2.20) in negative powers of z_1 and z_2 contains no terms of the form $b_{kk}(z_1z_2)^{-k}$ with non-zero b_{kk}, so that $g(p_1\bar{\Psi}_1, \bar{q}_2) = 0$.

In a similar manner we can choose m_2 to ensure that $g(\bar{q}_1, p_2\bar{\psi}_2) = 0$ where $p_2(z_2) = z_2^{m_2}\bar{p}_2(z_2)$.

Let us now consider

$$g(p_1\bar{\psi}_1, p_2\bar{\psi}_2) = \frac{p_1\bar{\psi}_1 p_2\bar{\psi}_2 S(z_1z_2)^m}{\phi\bar{\psi}_1\bar{\psi}_2} \circledcirc \sum(z_1z_2)^{-k}$$

$$= \bar{p}_1\bar{p}_2 \frac{z_1^{m_1}z_2^{m_2}S(z_1z_2)^m}{\phi} \circledcirc \sum(z_1z_2)^{-k} \qquad (3.2.21)$$

and let us write

$$(z_1z_2)^m z_1^{m_1}z_2^{m_2}S = N(z_1,z_2) + \phi(z_1z_2)M(z_1,z_2) \qquad (3.2.22)$$

where $N(z_1,z_2)$ contains no term with a factor $(z_1z_2)^{n-m}$.

We assert that the components of $N(z_1,z_2)$ are linearly independent. For, suppose the contrary; then there exists c^T such that $c^T N = 0$. Hence $(z_1z_2)^m z_1^{m_1}z_2^{m_2}c^T S = \phi c^T M$, by (3.2.22), so that ϕ divides $c^T S$, since ϕ has no zero roots. Then by (3.2.17)

$$c^T\hat{x}(z_1,z_2) = \frac{c^T S(z_1,z_2)}{\phi\psi_1\psi_2} = \frac{k(z_1,z_2)}{\psi_1\psi_2}, \text{ say.}$$

But this is linearly dependent on the components of $x^1 \otimes x^2$, which is a contradiction of Lemma 3.2.2.

Now, substituting (3.2.22) into (3.2.21), we obtain

$$g(p_1\bar{\psi}_1, p_2\bar{\psi}_2) = \bar{p}_1\bar{p}_2(M + \frac{N}{\phi}) \circledcirc \sum(z_1z_2)^{-k}$$

$$= \bar{p}_1\bar{p}_2 \frac{N}{\phi(z_1z_2)} \circledcirc \sum(z_1z_2)^{-k} \qquad (3.2.23)$$

since M is a polynomial in z_1 and z_2.

Now the terms in $N(z_1,z_2)$ will be members of the sets

$$B = \{(z_1z_2)^k z_1^j : k = 0,\ldots,\bar{n}-1;\ j = 0,\ldots,\ell_1\}$$

and

$$C = \{(z_1z_2)^k z_2^i : k = 0,\ldots,\bar{n}-1;\ i = 1,\ldots,\ell_2\}$$

for some ℓ_1 and ℓ_2, and $\bar{n} = n-m$.

Let us arrange these terms in the following way:

$$\begin{array}{ccccc} z_2^{\ell_2} \;\cdots\cdots\cdots & z_2 & 1 & z_1 & z_1^{\ell_1} \\ (z_1z_2)z_2^{\ell_2} \;\cdots\cdots & (z_1z_2)z_2 & z_1z_2 & (z_1z_2)z_1 & (z_1z_2)z_1^{\ell_1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ (z_1z_2)^{\bar{n}-1}z_2^{\ell_2} \cdots\cdots & (z_1z_2)^{\bar{n}-1}z_2 & (z_1z_2)^{\bar{n}-1} & (z_1z_2)^{\bar{n}-1}z_1 & (z_1z_2)^{\bar{n}-1}z_1^{\ell_1} \end{array}$$

We label the columns of this table as $e_{-\ell_2},\ldots,e_{-1},e_o,e_1,\ldots e_{\ell_1}$, and it is clear that if we can show that the transfer function vector

$$w(z_1,z_2) = \frac{1}{\phi(z_1z_2)}\begin{bmatrix} e_{-\ell_2} \\ \vdots \\ e_o \\ \vdots \\ e_{\ell_1} \end{bmatrix} \in R^{\bar{n}(\ell_1+\ell_2+1)}[(z_1,z_2)] \qquad (3.2.24)$$

is "reachable", then it follows that $N(z_1z_2)/\phi(z_1z_2)$ is also "reachable" and hence that the system (3.1.1)-(3.1.4) is quasi-reachable. By "reachability" of $w(z_1,z_2)$ in this context, we mean that for all specified $y \in R^{\bar{n}(\ell_1+\ell_2+1)}$, there exists $\bar{p}_1(z_1) \in R[z_1]$ and $\bar{p}_2(z_2) \in R[z_2]$ such that the vector coefficient of $(z_1z_2)^{-1}$ in $w(z_1,z_2)\bar{p}_1(z_1)\bar{p}_2(z_2)$ is equal to y.

Note that although $N(z_1,z_2)/\phi(z_1z_2)$ and $w(z_1,z_2)$ are not necessarily strictly causal (that is, there may be higher powers of z_1 in the numerator than in the denominator), we counter this by only allowing inputs to be inserted before time 0, and observe the outputs at time + 1. Furthermore, it is obvious from the earlier development how we arrived at $N(z_1,z_2)/\phi(z_1z_2)$ and there is nothing spurious about the way we use it in (3.2.23) as though it were a transfer function with inputs $\bar{p}_1(z_1)$ and $\bar{p}_2(z_2)$.

Before constructing our input sequence, we prove the following

<u>Lemma 3.2.3</u>:

Let (A,b) be a controllable pair. Then for all $\ell > 0$, there exists an integer $N > \ell$ such that $(A^N, A^k b)$ is a controllable pair for all k iff A is non-singular.

Proof: Let us write (A,b) in the following canonical form [K2]:

$$A = \begin{bmatrix} J_{\lambda_1} & & 0 \\ & \ddots & \\ 0 & & J_{\lambda_n} \end{bmatrix} \qquad b = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}$$

where $J_{\lambda_i} \in R^{n_i \times n_i}$ is the Jordan block

$$\begin{bmatrix} \lambda_i & 1 & & 0 \\ & \ddots & \ddots & \\ & & & 1 \\ 0 & & & \lambda_i \end{bmatrix}$$

and $b_i = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} \in R^{n_i}$.

Suppose A is singular; then $\lambda_j = 0$ for some j. If λ_j has multiplicity m, then $J^m_{\lambda_j} b_j = 0$, so that $(A^N, A^m b)$ is not controllable for any N.

Conversely, suppose A is non-singular. Then it is clear that $(A^N, A^k b)$ is controllable iff (A^N, b) is controllable.

But (A^N, b) is uncontrollable iff $\lambda_j^N = \lambda_k^N$ for some j, k. Let us write $\lambda_j = r_j e^{i\theta j}$, $\lambda_k = r_k e^{i\theta k}$, where $r_j, r_k > 0$. Then $\lambda_j^N = \lambda_k^N$ implies

$$r_j^N e^{iN\theta j} = r_k^N e^{iN\theta k}$$

so $\quad r_j = r_k$ and $e^{iN\theta j} = e^{iN\theta k}$

or $\quad N(\theta_j - \theta_k) = 2n\pi$ for some integer n. (3.2.25)

Let N_{jk} be the minimum value of N for which this occurs. Then any other N satisfying (3.2.25) is an integer multiple of N_{jk}. Now choose $\bar{N} > \ell$ coprime to $\{N_{jk}: j,k = 1,\ldots,n;\ j \neq k\}$.

Then $(A^{\bar{N}}, b)$ is a controllable pair. □

We now return once again to the proof of Theorem 4.2. Let (c^T, A, b) be a minimal realization of $1/\phi(z_1 z_2)$. We choose N to satisfy the conditions of the above lemma for $\ell = \ell_1 + \ell_2$ and define $\bar{p}_1(z_1)$ and $\bar{p}_2(z_2)$ as follows:

$$\bar{p}_1(z_1) = z_1^{\ell_2}(1 + z_1^N + z_1^{2N} + \ldots + z_1^{(\bar{n}-1)N}) \qquad (3.2.26)$$

$$\bar{p}_2(z_2) = \alpha_{-\ell_2,1}+\alpha_{-\ell_2+1,1}z_2+\ldots+\alpha_{01}z_2^{\ell_2}+\alpha_{11}z_2^{\ell_2+1}+\ldots+\alpha_{\ell_1,1}z_2^{\ell_1+\ell_2}$$
$$+ z_1^N(\alpha_{-\ell_2,2}+\alpha_{-\ell_2+1,2}z_2+\ldots+\alpha_{02}z_2^{\ell_2}+\alpha_{12}z_2^{\ell_2+1}+\ldots+\alpha_{\ell_1,2}z_2^{\ell_1+\ell_2})$$
$$+ \ldots$$
$$+ z_1^{N(\bar{n}-1)}(\alpha_{-\ell_2,\bar{n}}+\ldots+\alpha_{0\bar{n}}z_2^{\ell_2}+\ldots+\alpha_{\ell_1,\bar{n}}z_2^{\ell_1+\ell_2}) \qquad (3.2.27)$$

Again, a diagram may prove useful to visualize the polynomials:

$\bar{p}_1$: 0.....010....00.....00....010.....00........0...00....010....0

$\bar{p}_2$: x.....xxx....x0.....0x....xxx.....x0........0...0x....xxx....x

Note that by our choice of $N > \ell_1+\ell_2$ there is no overlapping of terms in (3.2.27). We shall now find that for all i, the inputs $\alpha_{ij}(j=1,\ldots\bar{n})$ only affect the outputs of the transfer functions of (3.2.24) whose numerators lie in the column e_i.

For we see that the output from the transfer function with numerator $(z_1z_2)^s z_1^j$ is

$$\frac{(z_1z_2)^s z_1^j}{\phi(z_1z_2)}\bar{p}_1(z_1)\bar{p}_2(z_2) \odot \Sigma(z_1z_2)^{-k} \qquad (s = 0,\ldots,\bar{n}-1)$$
$$= \frac{(z_1z_2)^{s+\ell_2+j}(\alpha_{j,1}+(z_1z_2)^N\alpha_{j,2}+\ldots+(z_1z_2)^{N(\bar{n}-1)}\alpha_{j,\bar{n}})}{\phi(z_1z_2)} \qquad (3.2.28)$$

by inspection.

Similarly the output from the transfer function with numerator $(z_1z_2)^s z_2^i$ is

$$\frac{(z_1z_2)^{s+\ell_2}(\alpha_{-i,1}+(z_1z_2)^N\alpha_{-i,2}+\ldots+(z_1z_2)^{N(\bar{n}-1)}\alpha_{-i,\bar{n}})}{\phi(z_1z_2)}$$
$$(s = 0,\ldots,\bar{n}-1)$$

Let us now label the output at time 1 from the transfer function $e_j/\phi(z_1z_2)$ $(j \geq 0)$ due to the input sequence (3.2.26) and (3.2.27) by $[y_{j1}\ldots y_{js}\ldots y_{j\bar{n}}]^T$. It then follows from (3.2.28) that

$$\begin{pmatrix} y_{j1} \\ \vdots \\ y_{js} \\ \vdots \\ y_{j\bar{n}} \end{pmatrix} = \begin{pmatrix} \sum_{k=1}^{\bar{n}} c^T A^{\ell_2+j} A^{kN} b\alpha_{j,k} \\ \vdots \\ \sum_{k=1}^{\bar{n}} c^T A^{s+\ell_2+j} A^{kN} b\alpha_{j,k} \\ \vdots \\ \sum_{k=1}^{n} c^T A^{\bar{n}-1+\ell_2+j} A^{kN} b\alpha_{j,k} \end{pmatrix},$$

$$= \begin{pmatrix} c^T \\ \cdot \\ c^T A \\ \cdot \\ c^T A^{\bar{n}-1} \end{pmatrix} [A^{\ell_2+j}b \;\; A^N A^{\ell_2+j}b \ldots A^{N(\bar{n}-1)} A^{\ell_2+j}b] \begin{pmatrix} \alpha_{j,1} \\ \vdots \\ \alpha_{j,\bar{n}} \end{pmatrix} \quad (3.2.29)$$

where (c^T,A,b) is a minimal realization of $\frac{1}{\phi(z_1z_2)}$ as defined above.

Now the first two matrices of (3.2.29) are invertible since (c^T,A) is an observable pair and $(A^N,A^{\ell_2+j}b)$ is a controllable pair by Lemma 3.2.3, so given specified values $y_{j,k}$ $(k = 1,\ldots,\bar{n})$ we can obtain unique $\alpha_{j,1},\ldots,\alpha_{j,\bar{n}}$ which reach $y_{j,k}$.

A similar situation holds for the outputs at time 1 from the transfer functions with numerator $(z_1z_2)^s z_2^i$ $(s = 0,\ldots,\bar{n}-1)$, so it is clear that the transfer function vector (3.2.24) can reach any desired output, so that our theorem is proved. □

In linear system theory, we usually ask not only about reachability, but about controllability as well; if a system is both reachable from and controllable to the origin, it follows that the system is completely controllable, i.e. we attain any one state in finite time starting from any other.

In the same way, subject to the quasi-reachability constraint, we can prove a similar theorem for bilinear systems:

Theorem 3.2.2

If the conditions of Theorem 3.2.1 hold, then every state of the system (3.1.1)-(3.1.4) is controllable to the origin.

Proof: We first show that from an initial state which is not reachable from the origin, we can attain a state which is reachable from the origin. We then show that any reachable state is controllable to the origin.

By Lemma 3.2.2, a non-reachable state $\begin{pmatrix} x_o^1 \\ x_o^2 \\ x_o \end{pmatrix}$ is one which is characterized by x_o^1 and x_o^2 lying on some finite union of U_1, U_2 of hyperplanes in R^{n_1} and R^{n_2} respectively, and the substate $\bar{x}_o \in R^m$ of x_o, corresponding to zero eigenvalues of A, being incompatible with these. It is of course clear, from Lemma 3.2.2, that there do exist reachable states of (3.1.1)-(3.1.4) for any $x_o^1 \in R^{n_1}$ $x_o^2 \in R^{n_2}$.

Let us now partition x_k as $\begin{pmatrix} \bar{x}_k \\ \hat{x}_k \end{pmatrix}$, using the transformation (3.2.7), where $\bar{x}_k, \hat{x}_k$ are the subsystems corresponding to J_0 and J_1 respectively. Then if there are no inputs from UxV for the next m stages, it is clear from (3.1.1)-(3.1.3) that the state at time m is given by

$$x_m^1 = A_1^m x_o^1 \qquad x_m^2 = A_2^m x_o^2$$

$$\begin{pmatrix} x_m^1 \otimes x_m^2 \\ \bar{x}_m \\ \hat{x}_m \end{pmatrix} = \begin{pmatrix} A_1 \otimes A_2 & 0 & 0 \\ C_1 & J_0 & 0 \\ C_2 & 0 & J_1 \end{pmatrix}^m \begin{pmatrix} x_o^1 \otimes x_o^2 \\ \bar{x}_o \\ \hat{x}_o \end{pmatrix} \tag{3.2.30}$$

Now $J_0^m = 0$, so (3.2.30) can be rewritten as

$$\begin{pmatrix} x_m^1 \otimes x_m^2 \\ \bar{x}_m \\ x_m \end{pmatrix} = \begin{pmatrix} A_1^m \otimes A_2^m & 0 & 0 \\ \mathcal{D}_1 & 0 & 0 \\ \mathcal{D}_2 & 0 & J_1^m \end{pmatrix} \begin{pmatrix} x_o^1 \otimes x_o^2 \\ \bar{x}_o \\ x_o \end{pmatrix}$$

In other words $\bar{x}_o$ makes no contribution to the state at time m, and the state $\begin{pmatrix} x_m^1 \\ x_m^2 \\ \bar{x}_m \\ \hat{x}_m \end{pmatrix}$ could equivalently have been reached from a state

$\begin{pmatrix} x_o^1 \\ x_o^2 \\ \bar{x}_o \\ \hat{x}_o \end{pmatrix}$ which was reachable from the origin. Hence $\begin{pmatrix} x_m^1 \\ x_m^2 \\ \bar{x}_m \\ \hat{x}_m \end{pmatrix}$ is reachable from the origin.

We now show that any reachable state is controllable to the origin. If it is reachable, then it is attained by an input $(p_1\psi_1+q_1, p_2\psi_2+q_2)$ for some $p_1, q_1 \in R[z_1]$, $p_2, q_2 \in R[z_2]$.

We shall construct the input sequence which sends the state to zero by the following concatenation:

$$((p_1\psi_1+q_1)\circ r_1\circ s_1\psi_1\circ 0^m), ((p_2\psi_2+q_2)\circ r_2\circ s_2\psi_2\circ 0^m)$$

with (r_1, r_2) and (s_1, s_2) determined sequentially.

Step 1: Multiply $p_1\psi_1+q_1$ and $p_2\psi_2+q_2$ by z_1^k and z_2^k respectively, where $k = \max(\deg\psi_1, \deg\psi_2)$.

Now choose r_1 and r_2 such that

$$z_i^k q_i + r_i \equiv 0 \pmod{\psi_i} \qquad i = 1,2$$

$$\text{with } \deg r_i = \deg \psi_i$$

and define $p_{ii}\psi_i = z_i^k(p_i\psi_i+q_i) + r_i \qquad i = 1,2$

Step 2: Choose the integer N as in Lemma 3.2.3, and in addition the integer $M > N(n-1) + \ell_1 + \ell_2 + m$ (ℓ_1 and ℓ_2 defined as in Theorem 3.2.1) in such a way that

$$g(z_1^M p_{11}\psi_1, z_2^m s_2\psi_2) = 0 = g(z_1^m s_1\psi_1, z_2^M p_{22}\psi_2)$$

where g represents the $\hat{x}(z_1,z_2)$ transfer function (3.2.17), so that we obtain $g(z_1^M p_{11}\psi_1 + z_1^m s_1\psi_1, z_2^M p_{22}\psi_2 + z_2^m s_2\psi_2)$

$$= g(z_1^M p_{11}\psi_1, z_2^M p_{22}\psi_2) + g(z_1^m s_1\psi_1, z_2^m s_2\psi_2)$$

We can then choose s_1 and s_2 appropriately, as explained in Theorem 3.2.1. In addition the factors z_1^m and z_2^m ensure that the subsystem of x_k corresponding to zero eigenvalues becomes zero. □

3.3 Observability of the State Space

In Definition 3.1.2 we said that two initial states were distinguishable if there exist finite length input sequences producing different outputs for each initial state.

For linear systems, because we have no coupling of initial states with inputs other than with respect to addition, it is possible to distinguish initial states by observing a finite number of outputs due to one input sequence, and the actual input sequence itself is immaterial. For bilinear systems we will in general need a number of "experiments" — that is, several distinct input sequences all starting at the same initial state — to distinguish initial states. This is because we have a multiplicative coupling between inputs and initial states. We demonstrate this by the following

Lemma 3.3.1

Let $f: X_1 x X_2 x X x U x V \to Y$ represent the map from initial states $x_o^1 \in X_1$, $x_o^2 \in X_2$, $x_o \in X$ and input sequences $u \in U$, $v \in V$ to the output Y as specified by equations (3.1.1)-(3.1.4). Then

$$f(x_o^1,x_o^2,x_o;u,v) = f(x_o^1,0,0;0,v) + f(0,x_o^2,0;u,0) + f(x_o^1,x_o^2,0;0,0) + f(0,0,x_o;0,0) + f(0,0,0;u,v) \qquad (3.3.1)$$

Proof: By (3.1.3) and (3.1.4) we see that f is linear in x_o, so that

$$f(x_o^1,x_o^2,x_o;u,v) = f(0,0,x_o;0,0) + f(x_o^1,x_o^2,0;u,v). \qquad (3.3.2)$$

Now, x_o^1 represents a linear sum of past inputs from U and x_o^2 represents a linear sum of past inputs from V, so that f is bilinear with respect to (x_o^1,u) and (x_o^2,v). Hence

$$f(x_o^1,x_o^2,0;u,v) = f(x_o^1,0,0;0,v) + f(0,x_o^2,0;u,0) + f(x_o^1,x_o^2,0;0,0) + f(0,0,0;u,v). \qquad (3.3.3)$$

Combining (3.3.2) and (3.3.3) we obtain (3.3.1). □

An immediate consequence of this lemma is that two initial states are equivalent iff the first four terms of (3.3.1), for each initial state, are equal for all input sequences in UxV.

Let us consider $f(x_o^1,x_o^2,0;0,0)$. From (3.1.1) and (3.1.2) we see that

$$x_{k+1}^1 = A_1x_k^1 \text{ and } x_{k+1}^2 = A_2x_k^2$$

so that we can write

$$x_{k+1}^1 \otimes x_{k+1}^2 = A_1 \otimes A_2 x_k^1 \otimes x_k^2$$

and together with (3.1.3) this gives us

$$\begin{pmatrix} x_{k+1}^1 \otimes x_{k+1}^2 \\ x_{k+1} \end{pmatrix} = \begin{pmatrix} A_1 \otimes A_2 & 0 \\ C & A \end{pmatrix} \begin{pmatrix} x_k^1 \otimes x_k^2 \\ x_k \end{pmatrix}$$

so that $y_k = [d^T h^T] \begin{pmatrix} A_1 \otimes A_2 & 0 \\ C & A \end{pmatrix}^k \begin{pmatrix} x_o^1 \otimes x_o^2 \\ 0 \end{pmatrix}$ (3.3.4)

Next, $f(0,0,x_o;0,0)$ immediately gives us

$$y_k = h^T A^k x_o \qquad (3.3.5)$$

by inspection of (3.1.3) and (3.1.4), and this is the reason for our original requirement that (h^T,A) be an observable pair.

In fact, since $f(x_o^1,x_o^2,x_o;0,0) = f(x_o^1,x_o^2,0;0,0) + f(0,0,x_o;0,0)$ (3.3.6)

we can combine (3.3.4) and (3.3.5) to obtain

$$y_k = [d^T h^T] \begin{pmatrix} A_1 \otimes A_2 & 0 \\ C & A \end{pmatrix}^k \begin{pmatrix} x_o^1 \otimes x_o^2 \\ x_o \end{pmatrix}.$$

Now $f(x_o^1,0,0;0,v) = \sum_{i=0}^{r} f(x_o^1,0,0;0,\tilde{v}_i)$

where $\tilde{v}_i = (0\ldots0,v_i,0,\ldots,0)$

by bilinearity, where v is the input sequence $(v_o,v_1,\ldots,v_r)$. Note that we are considering inputs v_k at times $k \geq 0$.

Consider then $f(x_o^1,0,0;0,\tilde{v}_i)$. At time i, we have $x_i^1 = A_1^i x_o^1, x_i^2 = 0$, $x_i = 0$, by examination of (3.1.1)-(3.1.3). At time i+1, we have

$$x_{i+1}^1 = A_1^{i+1}x_o^1,\ x_{i+1}^2 = b_2v_i,\ x_{i+1} = Q_1A_1^i x_o^1 v_i$$

and since all further $v_k (k > i)$ are zero, we have

$$\begin{pmatrix} x^1_{k+i+1} \otimes x^2_{k+1+i} \\ x_{k+1+i} \end{pmatrix} = \begin{pmatrix} A_1 \otimes A_2 & 0 \\ C & A \end{pmatrix}^k \begin{pmatrix} A_1^{i+1} x_o \otimes b_2 v_i \\ Q_1 A_1^i x_o^1 v_i \end{pmatrix}$$

so that $$y_{k+i+1} = [d^T h^T] \begin{pmatrix} A_1 \otimes A_2 & 0 \\ C & A \end{pmatrix}^k \begin{pmatrix} A_1 \otimes b_2 \\ Q_1 \end{pmatrix} A_1^i x_o^1 v_i \qquad (3.3.7)$$

by removing the term $A_1^i x_o^1 v_i$ to the right of the brackets.

Similarly, $f(0, x_o^2, 0; \tilde{u}_j, 0)$ gives a sequence of outputs

$$y_{k+j+1} = [d^T h^T] \begin{pmatrix} A_1 \otimes A_2 & 0 \\ C & A \end{pmatrix}^k \begin{pmatrix} b_1 \otimes A_2 \\ Q_2 \end{pmatrix} A_2^j x_o^2 u_j . \qquad (3.3.8)$$

Remark 3.3.1 We note that the identity (3.3.1) tells us that we can actually "observe" the output sequence (3.3.7) by first performing an "experiment" with no inputs, and then, starting at the same initial state, perform another experiment with all inputs zero except v_i. Similarly with the output sequence (3.3.8).

We now present the main theorem on observability, which also demonstrates the sufficiency of (h^T, A) being an observable pair.

Theorem 3.3.1

The system (3.1.1)-(3.1.4) is observable iff

(i) (h^T, A) is observable

(ii) $$\left([d^T h^T] , \begin{pmatrix} A_1 \otimes A_2 & 0 \\ C & A \end{pmatrix} , \begin{pmatrix} A_1 \otimes b_2 \\ Q_1 \end{pmatrix} , A_1 \right) \text{ is biobservable} \qquad (3.3.9)$$

(iii) $$\left([d^T h^T] , \begin{pmatrix} A_1 \otimes A_2 & 0 \\ C & A \end{pmatrix} , \begin{pmatrix} b_1 \otimes A_2 \\ Q_2 \end{pmatrix} , A_2 \right) \text{ is biobservable} \qquad (3.3.10)$$

where (a^T, M, L, T) is biobservable iff $a^T M^i L T^j y = 0$ for all i,j implies $y = 0$.

Remark 3.3.2 To check biobservability, we calculate the observability subspace H generated by (a^T, M). Then, letting H be a matrix whose row vectors are a basis for H, it is clear that (a^T, M, L, T) is biobservable iff (HL, T) is an observable pair.

Proof: We have already seen that condition (i) is necessary, for if the initial state is $(0,0,x_o)$, its contribution to the output is $h^T A^k x_o$.

Consider now the initial state $(x_o^1,0,0)$; then $x_o^1 \otimes 0 = 0$, so the contribution from $f(x_o^1,0,0;0,0)$ is zero, as we see from (3.3.4). So the only contribution which x_o^1 makes is via $f(x_o^1,0,0;0,v)$. Hence (ii) is necessary.

Similarly, by considering the initial state $(0,x_o^2,0)$, we see that condition (iii) is necessary.

To show sufficiency, we note by Remark 3.3.1 that we can always "observe" $f(x_o^1,0,0;0,v_i)$, so that condition (ii) is sufficient. Likewise, we see the sufficiency of condition (iii). Finally, since we already have x_o^1 and x_o^2 observable, we see from (3.3.5) and (3.3.6) that (i) is sufficient for x_o to be observable. □

This theorem, together with Theorem 3.2.1, provides us with necessary and sufficient conditions for a state space realization of a bilinear input/output map to be quasi-reachable and observable. In the next chapter we shall demonstrate how to obtain a realization with the properties of quasi-reachability and observability from a realization which does not possess them, and since we have seen in Chapter 2 that some state space realization can always be constructed, it will then follow that a quasi-reachable and observable realization always exists.

CHAPTER 4

REDUCTION PROCEDURES AND CANONICAL FORMS FOR BILINEAR INPUT/OUTPUT MAPS

We have seen in Chapter 2 that it is possible to construct a state space realization of any bilinear input/output map, and in Chapter 3 we have demonstrated necessary and sufficient conditions for such a realization to be quasi-reachable and observable or canonical (Definition 3.1.3). In this chapter we shall see how to reduce any realization to a canonical one, and in addition we shall find that the term reduction is well-chosen, since in the case of reduction to quasi-reachable realization, the dimension of the state space is reduced, and in the case of reduction to observable realization, the dimension of the state space is at least not increased. Note that the dimension of the state space may well stay the same, as in Example 1 below, on reduction to observable state space form.

We shall deal with reduction to observable state space form in §4.1 and in §4.2 we demonstrate reduction to quasi-reachable form. We choose this order of doing things rather than the conventional reduction to reachable form followed by reduction to observable form, basically because it is simpler; the fact that we are dealing with quasi-reachability rather than complete reachability means that it is more convenient to deal with this factor second.

In §4.3 we show that a realization is canonical if and only if it is co-minimal (Definition 3.1.4), and that all co-minimal realizations are isomorphic under the transformations defined in Chapter 3. (Henceforth we shall omit the prefix co- before minimal, although by convention a minimal realization has as its definition the analogue of Definition 3.1.4, where observable is replaced by reachable.)

In §4.4 we present two canonical forms for realizations of bilinear input/output maps.

4.1 Reduction to Observable Realization

When we talk about an unobservable state in linear system theory, we mean a particular mode of the state space which can be partitioned off from the other states and which neither contributes to the output nor to any of the other states. Naturally, with bilinear systems we encounter the same phenomenon; however as the following example shows, this is not the only kind of unobservable state:

Example 1:

$$s = \frac{1}{z_1(z_2-a)} . \tag{4.1.1}$$

An obvious choice of state space representation is

$$x^1_{k+1} = u_k \quad x^2_{k+1} = ax^2_k + v_k \quad y_k = x^1_k x^2_k . \tag{4.1.2}$$

However, if we check the conditions of Theorem 3.3.1, we see that this is unobservable; more straightforwardly, we see that if $x^2_o = 0$, then the value of x^1_o has no effect on the output. Note, though, that we could perhaps call this state space description quasi-observable, since if $x^2_o \neq 0$, we can observe the effect of x^1_o as well. This idea of quasi-observability will arise with multi-output bilinear maps. If we now regard the transfer function (4.1.1) as

$$s = \frac{z_2}{z_1 z_2 (z_2-a)}$$

a natural choice of state space description is

$$x^2_{k+1} = ax^2_k + v_k \quad x_{k+1} = u_k(ax^2_k + v_k) \quad y_k = x_k \tag{4.1.3}$$

and we can check that this is indeed observable, although only quasi-reachable.

The reduction procedure that we detail here will tell us how to switch from (4.1.2) to (4.1.3), and in addition we shall see that the word reduction is not inappropriate — at worst the dimension of the state space will remain the same after reduction, as in the example above. Otherwise, the dimension of the state space will indeed be reduced.

Let us now turn to the reduction procedure itself. We shall assume that (h^T, A) is observable; if not, we reduce the state space in the usual way.

Now, let H be a matrix whose rows are a basis for the observability subspace $\mathcal{H}$ of

$$\left([d^T h^T] \ , \begin{bmatrix} A_1 \boxtimes A_2 & O \\ C & A \end{bmatrix} \right) .$$

It then follows from Remark 3.3.2 that the biobservability subspaces corresponding to A_1 and A_2 are the observability subspaces of

$$\left(H \begin{bmatrix} A_1 \boxtimes b_2 \\ Q_1 \end{bmatrix} , A_1 \right) \text{ and } \left(H \begin{bmatrix} b_1 \boxtimes A_2 \\ Q_2 \end{bmatrix} , A_2 \right) \text{ respectively.}$$

Let T_1 and T_2 be matrices whose rows are a basis for these biobservability subspaces. In particular, this implies that

$$T_1A_1 = S_1T_1 \quad \text{and} \quad T_2A_2 = S_2T_2 \tag{4.1.4}$$

for some S_1 and S_2.

We shall now write the basis matrix H of $\mathcal{H}$ as

$$H = \begin{pmatrix} U & O \\ V & O \\ W & I_n \end{pmatrix} \quad \text{where } U \subset T_1 \boxtimes T_2 \tag{4.1.5}$$

and V is linearly independent of $T_1 \boxtimes T_2$.

This we can do since $\begin{pmatrix} A_1 \boxtimes A_2 & O \\ C & A \end{pmatrix}$ is lower block triangular and (h^T, A) is an observable pair.

Now because of the invariant subspace property of $\mathcal{H}$, we can write

$$\begin{pmatrix} U & O \\ V & O \\ W & I \end{pmatrix} \begin{pmatrix} A_1 \boxtimes A_2 & O \\ C & A \end{pmatrix} = \begin{pmatrix} L & O & O \\ L_1 & K_1 & O \\ L_2 & K_2 & A \end{pmatrix} \begin{pmatrix} U & O \\ V & O \\ W & I \end{pmatrix} \tag{4.1.6}$$

for some matrices L, L_1, L_2, K_1 and K_2. The only identity that we obtain from this matrix equality which is not immediately obvious is $U(A_1 \boxtimes A_2) = LU$, but this follows from the fact that $U \subset T_1 \boxtimes T_2$, which is an invariant subspace of $A_1 \boxtimes A_2$.

Further, because $[d^T h^T]$ is contained in H, we can write

$$[d^T h^T] = [k_1^T k^T h^T] \begin{pmatrix} U & O \\ V & O \\ W & I \end{pmatrix} \tag{4.1.7}$$

for some k_1^T, k^T.

It is now immediate, from (4.1.6) and (4.1.7) and from the fact that H has full row rank, that

$$\left([k_1^T\ k^T\ h^T],\ \begin{bmatrix} L & 0 & 0 \\ L_1 & K_1 & 0 \\ L_2 & K_2 & A \end{bmatrix} \right)$$

is an observable pair, and in particular we see that

$$\left([k^T\ h^T],\ \begin{bmatrix} K_1 & 0 \\ K_2 & A \end{bmatrix} \right) \triangleq (\hat{h}^T, \hat{A})$$

is an observable pair. This last remark follows from $\begin{pmatrix} L & 0 & 0 \\ L_1 & K_1 & 0 \\ L_2 & K_2 & A \end{pmatrix}$ being lower block triangular.

We also note that by (3.3.9) we have

$$\begin{pmatrix} U & 0 \\ V & 0 \\ W & I \end{pmatrix} \begin{pmatrix} A_1 \otimes b_2 \\ Q_1 \end{pmatrix} = \begin{pmatrix} U(A_1 \otimes b_2) \\ V(A_1 \otimes b_2) \\ W(A_1 \otimes b_2) + Q_1 \end{pmatrix} \subset T_1 \qquad (4.1.8)$$

and similarly by (3.3.10)

$$\begin{pmatrix} U & 0 \\ V & 0 \\ W & I \end{pmatrix} \begin{pmatrix} b_1 \otimes A_2 \\ Q_2 \end{pmatrix} = \begin{pmatrix} U(b_1 \otimes A_2) \\ V(b_1 \otimes A_2) \\ W(b_1 \otimes A_2) + Q_2 \end{pmatrix} \subset T_2 \qquad (4.1.9)$$

We shall use these facts (4.1.8)-(4.1.9) in the reduction procedure.

To actually perform the reduction to observable form we shall first of all add on some dummy states to the substate x_k; the number of dummy states will be equal to the rank of the matrix V. We shall then transform equations (3.1.1)-(3.1.4) using the transformations from Proposition 3.1.1. Finally, we shall eliminate those states in the null-spaces of T_1 and T_2 in the same way as we do for linear systems.

Step 1. Addition of Dummy States

We ~~rewrite~~ augment equations (3.1.3) and (3.1.4) as follows:

$$\hat{x}_{k+1} = \begin{bmatrix} K_1 & 0 \\ K_2 & A \end{bmatrix} \hat{x}_k + \begin{bmatrix} 0 \\ C \end{bmatrix} x_k^1 \otimes x_k^2 + \begin{bmatrix} 0 \\ Q_1 \end{bmatrix} x_k^1 v_k + \begin{bmatrix} 0 \\ Q_2 \end{bmatrix} x_k^2 u_k + \begin{bmatrix} 0 \\ b \end{bmatrix} u_k v_k \qquad (4.1.10)$$

$$y_k = d^T x_k^1 \otimes x_k^2 + [k^T\ h^T]\hat{x}_k. \qquad (4.1.11)$$

Note that the upper subsystem of (4.1.10) plays no clear role at the moment, since if it starts at the zero state, it will remain zero for all time; however the reason for its addition will become apparent in due course.

Let us also bear in mind that during the remaining steps of the reduction procedure, the pair $\left([k^T\ h^T], \begin{bmatrix} K_1 & O \\ K_2 & A \end{bmatrix}\right)$ will remain unchanged, so the fact that they are an observable pair is crucial.

In addition we remark that the calculation of k^T, K_1 and K_2 is done in the usual way, i.e. we append the matrix $[U_1\ O]$ to H, where the rows of U_1 are linearly independent of those of U and V, and then perform a similarity transformation on the pair $\left([d^T h^T], \begin{bmatrix} A_1 \otimes A_2 & O \\ C & A \end{bmatrix}\right)$, extracting the required values of k^T, K_1 and K_2 from the positions indicated by (4.1.6) and (4.1.7).

Step 2. Transformation of System Equations

We now transform equations (4.1.10) and (4.1.11) as prescribed by Proposition 3.1.1, using the matrix $\begin{bmatrix} V \\ W \end{bmatrix}$:

$$\begin{pmatrix} O \\ C \end{pmatrix} \rightarrow \begin{pmatrix} V \\ W \end{pmatrix} A_1 \otimes A_2 + \begin{pmatrix} O \\ C \end{pmatrix} - \begin{pmatrix} K_1 & O \\ K_2 & A \end{pmatrix} \begin{pmatrix} V \\ W \end{pmatrix} = \begin{pmatrix} V(A_1 \otimes A_2) - K_1 V \\ C + W(A_1 \otimes A_2) - K_2 V - AW \end{pmatrix} \triangleq \hat{C} \tag{4.1.12}$$

$$\begin{pmatrix} O \\ Q_1 \end{pmatrix} \rightarrow \begin{pmatrix} O \\ Q_1 \end{pmatrix} + \begin{pmatrix} V \\ W \end{pmatrix} A_1 \otimes b_2 = \begin{pmatrix} V(A_1 \otimes b_2) \\ Q_1 + W(A_1 \otimes b_2) \end{pmatrix} \triangleq \hat{Q}_1 \tag{4.1.13}$$

$$\begin{pmatrix} O \\ Q_2 \end{pmatrix} \rightarrow \begin{pmatrix} O \\ Q_2 \end{pmatrix} + \begin{pmatrix} V \\ W \end{pmatrix} b_1 \otimes A_2 = \begin{pmatrix} V(b_1 \otimes A_2) \\ Q_2 + W(b_1 \otimes A_2) \end{pmatrix} \triangleq \hat{Q}_2 \tag{4.1.14}$$

$$\begin{pmatrix} O \\ b \end{pmatrix} \rightarrow \begin{pmatrix} O \\ b \end{pmatrix} + \begin{pmatrix} V \\ W \end{pmatrix} b_1 \otimes b_2 = \begin{pmatrix} V(b_1 \otimes b_2) \\ b + W(b_1 \otimes b_2) \end{pmatrix} \triangleq \hat{b} \tag{4.1.15}$$

$$d^T \rightarrow d^T - [k^T\ h^T] \begin{pmatrix} V \\ W \end{pmatrix} = d^T - k^T V - h^T W \triangleq \hat{d}^T \tag{4.1.16}$$

Now from (4.1.6) we see that

$$V(A_1 \otimes A_2) - K_1 V = L_1 U \subset T_1 \otimes T_2 \tag{4.1.17}$$

and $\qquad W(A_1 \boxtimes A_2) + C - K_2V - AW = L_2U \quad \subset T_1 \boxtimes T_2,$

i.e. $\hat{C} \subset T_1 \boxtimes T_2$, and from (4.1.7) we see that

$$d^T - k^TV - h^TW = k_1^TU$$

i.e. $\hat{d}^T \subset T_1 \boxtimes T_2$.

So (4.1.10) and (4.1.11) are now transformed to

$$\hat{x}_{k+1} = \begin{pmatrix} K_1 & O \\ K_2 & A \end{pmatrix}\hat{x}_k + \begin{pmatrix} L_1 & U \\ L_2 & U \end{pmatrix} x_k^1 \boxtimes x_k^2 + \begin{pmatrix} V(A_1 \boxtimes b_2) \\ Q_1 + W(A_1 \boxtimes b_2) \end{pmatrix} x_k^1 v_k$$

$$+ \begin{pmatrix} V(b_1 \boxtimes A_2) \\ Q_1 + W(b_1 \boxtimes A_2) \end{pmatrix} x_k^2 u_k + \begin{pmatrix} V(b_1 \boxtimes b_2) \\ b + W(b_1 \boxtimes b_2) \end{pmatrix} u_k v_k$$

$$y_k = k_1^TUx_k^1 \boxtimes x_k^2 + [k^T h^T]\hat{x}_k$$

<u>Remark 4.1.1</u> The upper subsystem of $\hat{x}_k$ satisfies the system equation for $Vx_k^1 \boxtimes x_k^2$, since

$$Vx_{k+1}^1 \boxtimes x_{k+1}^2 = V(A_1x_k^1 + b_1u_k) \boxtimes (A_2x_k^2 + b_2v_k)$$

$$= K_1(Vx_k^1 \boxtimes x_k^2) + L_1Ux_k^1 \boxtimes x_k^2 + V(A_1 \boxtimes b_2)x_k^1 v_k + V(b_1 \boxtimes A_2)x_k^2 u_k + V(b_1 \boxtimes b_2)u_k v_k$$

using identity (4.1.17).

It is now apparent, therefore, that one of the intentions of the reduction procedure is to set up a new x_k substate to replace those substates of x_k^1 and x_k^2 which are unobservable separately, but which are observable as substates of $x_k^1 \boxtimes x_k^2$.

<u>Step 3. Elimination of Unobservable States</u>

From (4.1.8) and (4.1.9) we see that $\hat{Q}_1 \subset T_1$ and $\hat{Q}_2 \subset T_2$. Hence by choosing X_1 and X_2 to make $\begin{pmatrix} T_1 \\ X_1 \end{pmatrix}$ and $\begin{pmatrix} T_2 \\ X_2 \end{pmatrix}$ full rank, and calculating

$$[V_1W_1] = \begin{pmatrix} T_1 \\ X_1 \end{pmatrix}^{-1} \text{ and } [V_2W_2] = \begin{pmatrix} T_2 \\ X_2 \end{pmatrix}^{-1}$$

we can employ the usual linear system reduction procedure via similarity transformations on x_k^1 and x_k^2 to obtain

$$\hat{x}^1_{k+1} = T_1A_1V_1\hat{x}^1_k + T_1b_1u_k$$
$$\hat{x}^2_{k+1} = T_2A_2V_2\hat{x}_k + T_2b_2v_k$$
$$\hat{x}_{k+1} = \hat{A}\hat{x}_k + \hat{C}V_1 \otimes V_2\hat{x}^1_k \otimes \hat{x}^2_k + \hat{Q}_1V_1\hat{x}^1_k v_k + \hat{Q}_2V_2\hat{x}^2_k u_k + \hat{b}u_k v_k$$
$$y_k = \hat{d}^T V_1 \otimes V_2 \hat{x}^1_k \otimes \hat{x}^2_k + \hat{h}^T \hat{x}_k$$

where we use the fact that $UW_1 \otimes W_2 \subset (T_1 \otimes T_2)(W_1 \otimes W_2) = T_1W_1 \otimes T_2W_2 = 0$.

We now wish to show that the term "reduction" does indeed apply; we shall see shortly that we can immediately eliminate all the modes of x^1_k and x^2_k contained in ker T_1 and ker T_2 respectively which are associated with non-zero eigenvalues of A_1 and A_2. However some of the modes associated with zero eigenvalues will reappear in some sense in V, and are converted into x_k states. Intuitively, we can view this as the transfer function $\frac{1}{z_1z_2}$ giving rise to an x_k state or equivalently the transfer function $\frac{1}{z_1} \cdot \frac{1}{z_2}$ producing an x^1_k and an x^2_k state.

Let us consider the eigenvectors and generalized eigenvectors of A_1; then it is well-known that the null-space of T_1, i.e. the unobservable subspace of A_1, has as a basis a subset of these eigenvectors. If we then take the subset of those eigenvectors which correspond to non-zero eigenvalues, which we label as $[y_1 \ldots y_k] \triangleq Y_1$, it is clear that $T_1\mathcal{Y}_1 = 0$ and $A_1\mathcal{Y}_1 = \mathcal{Y}_1$, where $\mathcal{Y}_1$ is the subspace generated by Y_1. It then follows that there exists some non-singular matrix M_1 such that

$$Y_1 = A_1Y_1M_1. \qquad (4.1.18)$$

Lemma 4.1.1

Let U, V, W, A_2 and b_2 be defined as above; let (A_2,b_2) be a controllable pair. Then for all x^2_o there exists a matrix Z_1 such that

$$\begin{pmatrix} U & 0 \\ V & 0 \\ W & I \end{pmatrix} \begin{pmatrix} Y_1 \otimes x^2_o \\ Z_1 \end{pmatrix} = 0. \qquad (4.1.19)$$

<u>Remark</u>: This tells us that all initial states $\begin{pmatrix} x_o^1 \\ x_o^2 \\ x_o \end{pmatrix}$ where $x_o^1 \in \mathcal{Y}_1$ are indistiguishable from the initial state $\begin{pmatrix} 0 \\ x_o^2 \\ x_o - Z_1 y \end{pmatrix}$ where $x_o^1 = Y_1 y$.

This follows from

$$\begin{pmatrix} U & 0 \\ V & 0 \\ W & I \end{pmatrix} \begin{pmatrix} Y_1 y \boxtimes x_o^2 \\ x_o \end{pmatrix} + \begin{pmatrix} U & 0 \\ V & 0 \\ W & I \end{pmatrix} \begin{pmatrix} 0 \boxtimes x_o^2 \\ Z_1 y - x_o \end{pmatrix} = \begin{pmatrix} U & 0 \\ V & 0 \\ W & I \end{pmatrix} \begin{pmatrix} Y_1 y \boxtimes x_o^2 \\ Z_1 y \end{pmatrix} = 0$$

Proof of Lemma:

Since $\begin{pmatrix} U & 0 \\ V & 0 \\ W & I \end{pmatrix}$ is the observability subspace generated by $[d^T h^T]$ and $\begin{pmatrix} A_1 \boxtimes A_2 & 0 \\ C & A \end{pmatrix}$, (4.1.19) is equivalent to

$$[d^T h^T] \begin{pmatrix} A_1 \boxtimes A_2 & 0 \\ C & A \end{pmatrix}^i \begin{pmatrix} Y_1 \boxtimes x \\ Z_1 \end{pmatrix} = 0 \quad \text{for all } i.$$

Now Theorem 3.3.1 tells us that if $Y_1 \subset \ker T_1$ then

$$[d^T h^T] \begin{pmatrix} A_1 \boxtimes A_2 & 0 \\ C & A \end{pmatrix}^i \begin{pmatrix} A_1 \boxtimes b_2 \\ Q_1 \end{pmatrix} A_1^j Y_1 = 0 \quad \text{for all } i,j.$$

Setting $j = 0$, postmultiplying by M_1 and substituting from (4.1.18) we obtain

$$[d^T h^T] \begin{pmatrix} A_1 \boxtimes A_2 & 0 \\ C & A \end{pmatrix}^i \begin{pmatrix} Y_1 \boxtimes b_2 \\ Q_1 Y_1 M_1 \end{pmatrix} = 0 \quad \text{for all } i. \qquad (4.1.20)$$

Now expanding (4.1.20) we obtain

$$[d^T h^T] \begin{pmatrix} A_1 \boxtimes A_2 & 0 \\ C & A \end{pmatrix}^{i-1} \begin{pmatrix} A_1 Y_1 \boxtimes A_2 b_2 \\ C Y_1 \boxtimes b_2 + A Q_1 Y_1 M_1 \end{pmatrix} = 0 \text{ for all } i.$$

As before we postmultiply by M_1 and substitute from (4.1.19) to obtain

$$[d^T h^T] \begin{pmatrix} A_1 \boxtimes A_2 & 0 \\ C & A \end{pmatrix}^i \begin{pmatrix} Y_1 \boxtimes A_2 b_2 \\ K_1 \end{pmatrix} = 0$$

where $K_1 = C Y_1 M_1 \boxtimes b_2 + A Q_1 Y_1 M_1^2$.

In a similar way we obtain

$$[d^T h^T] \begin{pmatrix} A_1 \boxtimes A_2 & 0 \\ C & A \end{pmatrix}^i \begin{pmatrix} Y_1 \boxtimes A_2^r b_2 \\ K_r \end{pmatrix} = 0$$

for all i and r, where $K_{r+1} = CY_1M_1 \boxtimes A_2^r b_2 + AK_r M$.

Finally, since (A_2, b_2) is a controllable pair we can write

$$x_o^2 = \sum_{i=0}^{n_2-1} \alpha_j A_2^j b_2 \quad \text{for some } \alpha_o, \ldots, \alpha_{n_2-1}$$

and it then follows that

$$[d^T h^T] \begin{pmatrix} A_1 \boxtimes A_2 & 0 \\ C & A \end{pmatrix}^i \begin{pmatrix} Y_1 \boxtimes x_o^2 \\ \sum_{j=0}^{n_2-1} \alpha_j K_j \end{pmatrix} = 0$$

So our lemma is proved with $Z_1 = \sum \alpha_j K_j$.

We obtain a similar result for Y_2 where $T_2 \mathcal{Y}_2 = 0$ and $A_2 \mathcal{Y}_2 = \mathcal{Y}_2$, so we know that $\mathcal{Y}_1$ and $\mathcal{Y}_2$ can be discarded from the state space description.

Let us now examine the modes associated with zero eigenvalues. Some of them may end up being completely discarded as with $\mathcal{Y}_1$ and $\mathcal{Y}_2$; in general, however, some will be transferred through the state space.

We assume that (A_1, b_1) is a controllable pair. Hence A_1 is cyclic and has just one Jordan block of zero eigenvalues. Let us suppose that there are $\ell_1 + r_1$ of these zero eigenvalues; then there exists a vector x_1 such that $A_1^k x_1$ is non-zero for $k < \ell_1 + r_1$ and is zero for $k = \ell_1 + r_1$. We also suppose that $T_1 A_1^{\ell_1 - 1} x_1$ is non-zero, but

$$T_1 A_1^{\ell_1} x_1 = 0. \tag{4.1.21}$$

Then $T_1 A_1^{\ell_1+1} x_1 = S_1 T_1 A_1^{\ell_1} x_1 = 0$ by (4.1.4) and similarly

$$T_1 A_1^{\ell_1+j} x_1 = S_1^j T_1 A^{\ell_1} x_1 = 0 \quad \text{for all } j \geq 0. \tag{4.1.22}$$

Hence a basis for ker T_1 is given by $Y_1, A^{\ell_1} x_1, \ldots, A^{\ell_1 + r_1 - 1} x_1$.

Now it is clear that there exists c_1^T such that $c_1^T A_1^{\ell_1 + r_1 - 1} x_1 \neq 0$ and $c_1^T Y_1 = 0$, and it then follows that $c_1^T, c_1^T A_1, \ldots, c_1^T A_1^{r_1 - 1}$ and T_1 are linearly independent; for otherwise there would exist non-zero $\alpha_1,\ldots,\alpha_{r_1}$ and β^T such that

$$m^T = \sum_{i=1}^{r_1} \alpha_1 c_1^T A_1^{i-1} + \beta^T T_1 = 0.$$

Multiplying on the right by $A_1^{\ell_1+r_1-1}x_1 \subset \ker T_1$, we obtain

$$m^T A_1^{\ell_1+r_1-1} = \alpha_1 c_1^T A_1^{\ell_1+r_1-1} = 0$$

so that $\alpha_1 = 0$. Similarly multiplying by $A_1^{\ell_1+r_1-2}x_1 \subset \ker T_1$, we obtain

$$m^T A_1^{\ell_1+r_1-2} = \alpha_2 c_1^T A_1^{\ell_1+r_1-1} = 0$$

so that $\alpha_2 = 0$. Similarly we find that $\alpha_3 = \ldots = \alpha_{r_1} = 0$.

It is now clear that $c_1^T, c_1^T A_1, \ldots, c_1^T A_1^{r_1-1}$ and T_1 are a basis for the annihilator of Y_1, since by (4.1.18) we have that

$$c_1^T A_1^k Y_1 = c_1^T Y_1 M_1^{-k} = 0 \quad \text{for all } k.$$

Similarly for T_2 and A_2, and the corresponding ℓ_2 and r_2 there exists c_2^T such that $c_2^T A_2^{\ell_2+r_2-1}x_2 \neq 0$ and $c_2^T Y_2 = 0$, and it then follows that $c_2^T,\ldots,c_2^T A_2^{r_2-1}$ and T_2 are a basis for the annihilator of Y_2.

Now we proved in Lemma 4.2.1 that

$$\begin{pmatrix} U \\ V \end{pmatrix} Y_1 \otimes x_o^2 = 0 \qquad \text{for all } x_o^2$$

and

$$\begin{pmatrix} U \\ V \end{pmatrix} x_o^1 \otimes Y_2 = 0 \qquad \text{for all } x_o^1.$$

It then follows that $\begin{pmatrix} U \\ V \end{pmatrix}$ must be spanned by $\{e_i \otimes f_j\}$ where $\{e_i\}$ and $\{f_j\}$ are bases for the annihilators of Y_1 and Y_2 respectively. Hence by the preceding discussion we can immediately see that $\begin{pmatrix} U \\ V \end{pmatrix}$ is spanned by

$$\{T_1 \cup c_1^T,\ldots,c_1^T A_1^{r_1-1}\} \otimes \{T_2 \cup c_2^T,\ldots c_2^T A_2^{r_2-1}\}. \qquad (4.1.23)$$

Now, let $[v^T\ 0]$ be the first row of $[V\ 0]$ that we obtain in (4.1.5) after expansion of $[d^T\ h^T] \begin{pmatrix} A_1 \otimes A_2 & 0 \\ C & A \end{pmatrix}^{n+1}$.

Then by (4.1.23) it is clear that we can write v^T as

$$v^T = w^T T_1 \otimes T_2 + \sum_{j=0}^{r_2-1} \gamma_j^T T_1 \otimes c_2^T A_2^j + \sum_{i=0}^{r_1-1} c_1^T A_1^i \otimes \delta_i^T T_2 + \sum_{i=0}^{r_1-1} \sum_{j=0}^{r_2-1} \alpha_{ij} c_1^T A_1^i \otimes c_2^T A_2^j \qquad (4.1.24)$$

for some w^T, $\{\gamma_j^T\}$, $\{\delta_i^T\}$ and $\{\alpha_{ij}\}$.

The remaining rows of V and U are then calculated from $v^T(A_1 \otimes A_2)^k$, $k > 0$. Let $r = \max(r_1, r_2)$. Then using the facts that

i) $T_1 A_1^r \subset T_1$ and $T_2 A_2^r \subset T_2$ since T_1 and T_2 are bases for invariant subspaces of A_1 and A_2 respectively,

ii) $c_1^T A_1^{r_1} \subset T_1$ since $c_1^T A_1^{r_1} Y_1 = c_1^T Y_1 M_1^{-r_1} = 0$ and $c_1^T A_1^{r_1}(A_1^k x_1) = 0$, $k = \ell_1, \ldots, \ell_1 + r_1 - 1$ i.e. $c_1^T A_1^{r_1}$ annihilates the null-space of T_1 and similarly

iii) $c_2^T A_2^{r_2} \subset T_2$,

it is clear from (4.1.24) that

$$v^T(A_1 \otimes A_2)^r \subset T_1 \otimes T_2.$$

Hence $v^T(A_1 \otimes A_2)^r \subset U$, so that

$$\text{rank } V \le r = \max(r_1, r_2).$$

Now in Remark 4.1.1 we saw that the number of additional substates added to x_k was equal to rank V, and from Step 3 of the reduction procedure it is clear that the dimensions of x_k^1 and x_k^2 are reduced by at least r_1 and r_2 respectively. Hence the dimension of the whole state space is reduced by at least

$$r_1 + r_2 - \max(r_1, r_2) \ge 0.$$

Remark 4.1.2

The fact that our examination of V can be based on its first row v^T when dealing with a single output is the departure point when we turn to multioutput bilinear systems. As we shall see in the example at the beginning of Chapter 5, any attempt to set up a realization which is both completely observable and quasi-reachable may well break down because the above reduction procedure to observable form actually increases the dimension of the state space rather than decreases it. We are then left with a realization which is no longer quasi-reachable.

4.2 Reduction to Quasi-Reachable Realization

Reduction in the case of a realization with uncontrollable states is much simpler than reduction for unobservable states. Before desscribing the procedure, we prove the following

Lemma 4.2.1

If (A_1,b_1) and (A_2,b_2) are controllable pairs, then $(A_1\otimes A_2,\ [A_1\otimes b_2 \vdots b_1\otimes A_2 \vdots b_1\otimes b_2])$ is a controllable pair.

Proof: Suppose otherwise. Then there exists $v \in R^{n_1n_2}$ such that

$$v^T A_1\otimes A_2 = \lambda v^T \qquad \text{for some } \lambda \in C \tag{4.2.1}$$

and $$v^T[A_1\otimes b_2 \ b_1\otimes A_2 \ b_1\otimes A_2] = 0.$$

In particular $v^T A_1\otimes b_2 = 0$ implies $v^T A_1^k b_1\otimes b_2 = 0$ for all k

and $$v^T b_1\otimes A_2 = 0 \text{ implies } v^T b_1\otimes A_2^k b_2 = 0 \text{ for all k.}$$

Then from (4.2.1) we have

$$v^T A_1^{j+k} b_1\otimes A_2^j b_2 = \lambda^j v^T A_1^k b_1\otimes b_2 = 0 \quad \text{for all } j,k$$

and similarly $$v^T A_1^j b_1\otimes A_2^{j+k} b_2 = \lambda^j b_1\otimes A_2^k b_2 = 0 \quad \text{for all } j,k$$

and further $$v^T A_1^j b_1\otimes A_2^j b_2 = \lambda^j v^T b_1\otimes b_2 = 0 \quad \text{for all } j.$$

Now (A_1,b_1) and (A_2,b_2) controllable implies that $\{A_1^i b_1\otimes A_2^j b_2 : i = 0,\dots,n_1-1;\ j = 0,\dots,n_2-1\}$ is a basis for $R^{n_1n_2}$, so $v^T A_1^i b_1\otimes A_2^j b_2 = 0$ for all i,j implies $v^T = 0$. □

Let us suppose that the system (3.1.1)-(3.1.3) is not quasi-reachable. By Theorem 3.2.1, this means that the controllability matrix of

$$\begin{pmatrix} A_1 \otimes A_2 & O \\ C & A \end{pmatrix} \text{ and } \begin{pmatrix} A_1 \otimes b_2 & b_1 \otimes A_2 & b_1 \otimes b_2 \\ Q_1 & Q_2 & b \end{pmatrix} \triangleq \begin{pmatrix} B_1 \\ B_2 \end{pmatrix}$$

does not have full rank.

With the aid of Lemma 4.2.1 above we can normalize the controllability matrix to $\begin{pmatrix} I & O \\ L & L_1 \end{pmatrix} \triangleq R$ where I is the identity matrix of $R^{n_1 n_2}$. This follows because $\begin{pmatrix} A_1 \otimes A_2 & O \\ C & A \end{pmatrix}$ is a lower triangular matrix. Now, because of the invariant subspace property of R we can write

$$\begin{pmatrix} A_1 \otimes A_2 & O \\ C & A \end{pmatrix} \begin{pmatrix} I & O \\ L & L_1 \end{pmatrix} = \begin{pmatrix} I & O \\ L & L_1 \end{pmatrix} \begin{pmatrix} A_1 \otimes A_2 & O \\ E_1 & E_2 \end{pmatrix} \quad (4.2.2)$$

for some E_1, E_2, and because $\begin{pmatrix} B_1 \\ B_2 \end{pmatrix}$ is contained in R we have

$$\begin{pmatrix} B_1 \\ B_2 \end{pmatrix} = \begin{pmatrix} I & O \\ L & L_1 \end{pmatrix} \begin{pmatrix} B_1 \\ E \end{pmatrix} \quad (4.2.3)$$

for some matrix E.

We now append the matrix $\begin{pmatrix} O \\ L_2 \end{pmatrix}$ to R, where L_2 is linearly independent of L_1, and calculate

$$\begin{pmatrix} I & O & O \\ L & L_1 & L_2 \end{pmatrix}^{-1} = \begin{pmatrix} I & O \\ -N_1 L & N_1 \\ -N_2 L & N_2 \end{pmatrix}, \text{ say.}$$

Then using (4.2.2) we calculate

$$\begin{pmatrix} I & O \\ -N_1 L & N_1 \\ -N_2 L & N_2 \end{pmatrix} \begin{pmatrix} A_1 \otimes A_2 & O \\ C & A \end{pmatrix} \begin{pmatrix} I & O & O \\ L & L_1 & L_2 \end{pmatrix} = \begin{pmatrix} A_1 \otimes A_2 & O & O \\ N_1 [C + AL - L(A_1 \otimes A_2)] & N_1 A L_1 & N_1 A L_2 \\ O & O & N_2 A L_2 \end{pmatrix}$$

and

$$\begin{pmatrix} I & O \\ -N_1 L & N_1 \\ -N_2 L & N_2 \end{pmatrix} \begin{pmatrix} B_1 \\ B_2 \end{pmatrix} = \begin{pmatrix} B_1 \\ N_1 (B_2 - L B_1) \\ O \end{pmatrix}$$

and this last identity follows from (4.2.3).

We immediately see that

$$\left[\begin{pmatrix} A_1 \otimes A_2 & O \\ N_1 [C + AL - L(A_1 \otimes A_2)] & N_1 A L_2 \end{pmatrix}, \begin{pmatrix} B_1 \\ N_1 (B_2 - L B_1) \end{pmatrix} \right]$$

is a controllable pair.

It is now obvious what we need to do to reduce the state space to controllable form; we first transform (3.1.1)-(3.1.4) using the transformation defined in Proposition 3.1.1, with $W = -L$; we then just employ the ordinary linear system type similarity transformation $x_k \to \begin{pmatrix} N_1 \\ N_2 \end{pmatrix} x_k$, so that our reduced system equations (3.1.3) and (3.1.4) can be rewritten as

$$\hat{x}_{k+1} = N_1 A L_1 \hat{x}_k + N_1[C + AL - L(A_1 \otimes A_2)] x_k^1 \otimes x_k^2$$
$$+ N_1[Q_1 - L(A_1 \otimes b_2)] x_k^1 v_k + N_1[Q_2 - L(b_1 \otimes A_2)] x_k^2 u_k + N_1[b - L(b_1 \otimes b_2)] u_k v_k$$
$$y_k = (d^T + h^T L) x_k^1 \otimes x_k^2 + h^T L\, \hat{x}_k.$$

Remark 4.2.1

It is readily seen that the transformations used here all preserve observability.

We now return to our previous example $s = \frac{1}{z_1(z_2 - a)}$ with state space description

$$x_{k+1}^1 = u_k \qquad x_{k+1}^2 = a x_k^2 + v_k \qquad y_k = x_k^1 x_k^2.$$

We have $d^T = [1] \qquad A_1 \otimes A_2 = [0]$.

So the observability subspace H of $(d^T, A_1 \otimes A_2)$ is $[1]$. Now $A_1 \otimes b_2 = [0]$ and $b_1 \otimes A_2 = [a]$, so that $T_1 = 0$ and $T_2 = [1]$ and the system is not observable.

It is clear that $H = [1] \not\subset T_1 \otimes T_2 = [0] \otimes [1] = [0]$; hence $V = [1]$ and using (4.1.6) and (4.1.7) we obtain

$$K_1 = [0]$$
$$K_2 = \emptyset \text{ since } A = \emptyset$$

and $\qquad k^T = [1].$

Finally we choose $X_1 = [1]$ and $V_2 = T_2^{-1} = [1]$, $V_1 = 0$.

Then $\qquad \hat{x}_{k+1}^2 = a x_k^2 + v_k$

$$\hat{x}_{k+1} = K_1 x_k + V(b_1 \otimes A_2) V_2 \hat{x}_k^2 u_k + V(b_1 \otimes b_2) u_k v_k$$
$$= 0 x_k + [1][a][1] \hat{x}_k^2 u_k + [1][1] u_k v_k$$
$$= a \hat{x}_k^2 u_k + u_k v_k$$
$$y_k = k^T x_k = x_k.$$

Remark 4.2.2

Reduction to quasi-reachable form is in effect just pole-zero cancellation, similar to that encountered in linear systems. For example the transfer function

$$s = \frac{1}{(z_1^2+az_1+b)(z_2^2+cz_2+d)(z_1z_2+e)}$$

can be rewritten as

$$s = \frac{z_1z_2+f}{(z_1^2+az_1+b)(z_2^2+cz_2+d)} \cdot \frac{1}{(z_1z_2+e)(z_1z_2+f)}$$

Then by considering this as a linear system $1/(z+e)(z+f)$ with input from a bilinear map $(z_1z_2+f)/(z_1^2+az_1+b)(z_2^2+cz_2+d)$, a state space realization

$$x_{k+1}^1 = \begin{bmatrix} 0 & 1 \\ -b & -a \end{bmatrix} x_k^1 + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u_k \qquad x_{k+1}^2 = \begin{bmatrix} 0 & 1 \\ -d & -c \end{bmatrix} x_k^1 + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u_k$$

$$x_{k+1} = \begin{bmatrix} 0 & 1 \\ -ef & -(e+f) \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 & 0 \\ f & 0 & 0 & 1 \end{bmatrix} x_k^1 \otimes x_k^2$$

$$y_k = [1 \ \ 0]\, x_k$$

can readily be constructed. It is now easy to check that the quasi-reachability conditions of Theorem 3.2.1 are not satisfied.

4.3 Minimal Realizations

In this section we will show that a realization is minimal if and only if it is canonical. This we do by demonstrating that all canonical (i.e. quasi-reachable and observable) realizations of a bilinear input/output map f are isomorphic to one another in the sense that any two are related by the four types of transformations detailed in §3.1. Having done this it will then be apparent that there is a unique mapping from any observable realization of f to any specified canonical realization; hence, according to Definition 3.1.4, a canonical realization will also be minimal. Finally, to show that a minimal realization is canonical we use the method of contradiction; suppose that a minimal realization M is not quasi-reachable. Then given any observable realization M', there is not in general a unique mapping $\phi : M \to M'$, a contradiction of the fact that M is minimal. Hence M must be quasi-reachable.

We shall now build up to Theorem 4.3.1, which states that all canonical realizations are isomorphic to one another, by means of Lemmas 4.3.1-4.3.3. We do not make any assumptions regarding the dimensions (n_1,n_2,n) and $(\hat{n}_1,\hat{n}_2,\hat{n})$ of the substates (x_k^1,x_k^2,x_k) and $(\hat{x}_k^1,\hat{x}_k^2,\hat{x}_k)$ of any two canonical realizations M and $\hat{M}$, but we shall find that the dimensions of corresponding substates are equal.

In Lemma 4.3.1 we shall require the following results from Chapter 3 regarding canonical realizations M and $\hat{M}$. We only state the results for M, but they will be identical for $\hat{M}$:

(R1) (A_1,b_1) and (A_2,b_2) are reachable pairs

(R2) $\left([d^T\ h^T],\ \begin{pmatrix} A_1 \otimes A_2 & O \\ C & A \end{pmatrix}, \begin{pmatrix} A_1 \otimes b_2 \\ Q_1 \end{pmatrix},\ A_1 \right)$ and

$\left([d^T\ h^T],\ \begin{pmatrix} A_1 \otimes A_2 & O \\ C & A \end{pmatrix}, \begin{pmatrix} b_1 \otimes A_2 \\ Q_2 \end{pmatrix},\ A_2 \right)$ are biobservable pairs.

Lemma 4.3.1

Let $M = (A_1, b_1, A_2, b_2, A, C, Q_1, Q_2, b, h^T, d^T)$ and $\hat{M} = (\hat{A}_1, \hat{b}_1, \hat{A}_2, \hat{b}_2, \hat{A}, \hat{C}, \hat{Q}_1, \hat{Q}_2, \hat{b}, \hat{h}^T, \hat{d}^T)$ be canonical realizations of a bilinear input/output map f. Then there exist non-singular matrices $T_1 \in R^{n_1 \times n_1}$, $T_2 \in R^{n_2 \times n_2}$ such that

$$\hat{A}_1 = T_1 A_1 T_1^{-1}, \quad \hat{b}_1 = T_1 b_1, \quad \hat{A}_2 = T_2 A_2 T_2^{-1}, \quad \hat{b}_2 = T_2 b_2.$$

Proof: If M and $\hat{M}$ are realizations of f, then the following equalities hold on expansion of the terms in (2.4.5):

$$s_{i+j+1,i} = [d^T\ h^T]\begin{pmatrix} A_1 \otimes A_2 & 0 \\ C & A \end{pmatrix}^i \begin{pmatrix} A_1 \otimes b_2 \\ Q_1 \end{pmatrix} A_1^j b_1 = [\hat{d}^T\ \hat{h}^T]\begin{pmatrix} \hat{A}_1 \otimes \hat{A}_2 & 0 \\ \hat{C} & \hat{A} \end{pmatrix}^i \begin{pmatrix} \hat{A}_1 \otimes \hat{b}_2 \\ \hat{Q}_1 \end{pmatrix} \hat{A}_1^j \hat{b}_1 \quad (4.3.1)$$

$$s_{i,i+j+1} = [d^T\ h^T]\begin{pmatrix} A_1 \otimes A_2 & 0 \\ C & A \end{pmatrix}^i \begin{pmatrix} b_1 \otimes A_2 \\ Q_2 \end{pmatrix} A_2^j b_2 = [\hat{d}^T\ \hat{h}^T]\begin{pmatrix} \hat{A}_1 \otimes \hat{A}_2 & 0 \\ \hat{C} & \hat{A} \end{pmatrix}^i \begin{pmatrix} \hat{b}_1 \otimes \hat{A}_2 \\ \hat{Q}_2 \end{pmatrix} \hat{A}_2^j \hat{b}_2 \quad (4.3.2)$$

$$s_{i,i} = [d^T\ h^T]\begin{pmatrix} A_1 \otimes A_2 & 0 \\ C & A \end{pmatrix}^i \begin{pmatrix} b_1 \otimes b_2 \\ b \end{pmatrix} = [\hat{d}^T\ \hat{h}^T]\begin{pmatrix} \hat{A}_1 \otimes \hat{A}_2 & 0 \\ \hat{C} & \hat{A} \end{pmatrix}^i \begin{pmatrix} \hat{b}_1 \otimes \hat{b}_2 \\ \hat{b} \end{pmatrix} \quad (4.3.3)$$

where $s = (z_1 z_2)^{-1} \sum_{i,j \geq 1} s_{ij} z_1^{-i} z_2^{-j}$ is the transfer function representation of f.

Now let $H = \begin{pmatrix} p^T \\ p^T F \\ \vdots \\ p^T F^{k-1} \end{pmatrix}$ and $\hat{H} = \begin{pmatrix} \hat{p}^T \\ \hat{p}^T \hat{F} \\ \vdots \\ \hat{p}^T \hat{F}^{k-1} \end{pmatrix}$

span the observability subspaces generated by

$$(p^T, F) \triangleq \left([d^T h^T], \begin{bmatrix} A_1 \otimes A_2 & 0 \\ C & A \end{bmatrix}\right) \text{ and } (\hat{p}^T, \hat{F}) \triangleq \left([\hat{d}^T \hat{h}^T], \begin{bmatrix} \hat{A}_1 \otimes \hat{A}_2 & 0 \\ \hat{C} & \hat{A} \end{bmatrix}\right).$$

Equality (4.3.1) then implies that

$$H \begin{pmatrix} A_1 \otimes b_2 \\ Q_1 \end{pmatrix} A_1^j b_1 = \hat{H} \begin{pmatrix} \hat{A}_1 \otimes \hat{b}_2 \\ \hat{Q}_2 \end{pmatrix} \hat{A}_1^j \hat{b}_1 \text{ for all } j \geq 0.$$

Using results (R1) and (R2) above and the theory of Hankel matrices [K2], we can now deduce that there exists an invertible matrix T_1 such that

$$\hat{A}_1 = T_1 A_1 T_1^{-1} \quad \text{and} \quad \hat{b}_1 = T_1 b_1.$$

In a similar way, equality (4.3.2) implies that there exists an invertible matrix T_2 such that

$$\hat{A}_2 = T_2A_2T_2^{-1} \quad \text{and} \quad \hat{b}_2 = T_2b_2.$$

Using this result, it is now clear that to establish a relationship between two canonical realizations of a bilinear input/output map, it is sufficient to study realizations of the form $M = (A_1,b_1,A_2,b_2,A,C,Q_1,Q_2,b,h^T,d^T)$ and $\hat{M} = (A_1,b_1,A_2,b_2,\hat{A},\hat{C},\hat{Q}_1,\hat{Q}_2,\hat{b},\hat{h}^T,\hat{d}^T)$.

Returning then to the expressions (4.3.1)-(4.3.3) for $s_{i+j+1,i}$, $s_{i,i+j+1}$, s_{ii}, where we now assume that $\hat{A}_1 = A_1$, $\hat{b}_1 = b_1$, $\hat{A}_2 = A_2$, $\hat{b}_2 = b_2$, it is clear that since (A_1,b_1) and (A_2,b_2) are reachable pairs, we can rewrite (4.3.1) and (4.3.2) as

$$[d^Th^T]\begin{pmatrix} A_1\otimes A_2 & 0 \\ C & A \end{pmatrix}^i \begin{pmatrix} A_1\otimes b_2 \\ Q_1 \end{pmatrix} = [\hat{d}^T\hat{h}^T]\begin{pmatrix} A_1\otimes A_2 & 0 \\ \hat{C} & \hat{A} \end{pmatrix}^i \begin{pmatrix} A_1\otimes b_2 \\ \hat{Q}_1 \end{pmatrix} \qquad (i = 0,1,\ldots)$$

$$[d^Th^T]\begin{pmatrix} A_1\otimes A_2 & 0 \\ C & A \end{pmatrix}^i \begin{pmatrix} b_1\otimes A_2 \\ Q_2 \end{pmatrix} = [\hat{d}^T\hat{h}^T]\begin{pmatrix} A_1\otimes A_2 & 0 \\ \hat{C} & \hat{A} \end{pmatrix}^i \begin{pmatrix} b_1\otimes A_2 \\ \hat{Q}_2 \end{pmatrix} \qquad (i = 0,1,\ldots)$$

and combining these with (4.3.3), we obtain the following equality:

$$[d^Th^T]\begin{pmatrix} A_1\otimes A_2 & 0 \\ C & A \end{pmatrix}^i \begin{pmatrix} A_1\otimes b_2 & b_1\otimes A_2 & b_1\otimes b_2 \\ Q_1 & Q_2 & b \end{pmatrix} = [\hat{d}^T\hat{h}^T]\begin{pmatrix} A_1\otimes A_2 & 0 \\ \hat{C} & \hat{A} \end{pmatrix}^i \begin{pmatrix} A_1\otimes b_2 & b_1\otimes A_2 & b_1\otimes b_2 \\ \hat{Q}_1 & \hat{Q}_2 & \hat{Q} \end{pmatrix}$$

which we shall rewrite for convenience as

$$[d^Th^T]\begin{pmatrix} F & 0 \\ C & A \end{pmatrix}^i \begin{pmatrix} G \\ B \end{pmatrix} = [\hat{d}^T\hat{h}^T]\begin{pmatrix} F & 0 \\ \hat{C} & \hat{A} \end{pmatrix}^i \begin{pmatrix} G \\ \hat{B} \end{pmatrix} \quad (i = 0,1,\ldots). \qquad (4.3.4)$$

Note that (h^T,A) and $(\hat{h}^T,\hat{A})$ are observable pairs and

$$\left(\begin{bmatrix} F & 0 \\ C & A \end{bmatrix}, \begin{bmatrix} G \\ B \end{bmatrix}\right) \text{ and } \left(\begin{bmatrix} F & 0 \\ \hat{C} & \hat{A} \end{bmatrix}, \begin{bmatrix} G \\ \hat{B} \end{bmatrix}\right) \text{ are reachable pairs.}$$

Before we show the relationship between the matrices of M and $\hat{M}$ we shall prove the following lemma:

<u>Lemma 4.3.2</u>

Let $\left(\begin{bmatrix} K & O \\ L & A \end{bmatrix}, \begin{bmatrix} M \\ N \end{bmatrix}\right)$ and $\left(\begin{bmatrix} K & O \\ \hat{L} & \hat{A} \end{bmatrix}, \begin{bmatrix} M \\ \hat{N} \end{bmatrix}\right)$ be reachable pairs which are related by a similarity transformation. Then this similarity transformation is of the form $\begin{pmatrix} I & O \\ W & T \end{pmatrix}$, where T is an invertible matrix.

Proof: By the above assumptions there exist matrices T, U, V and W such that

$$\begin{pmatrix} V & U \\ W & T \end{pmatrix}\begin{pmatrix} K & O \\ L & A \end{pmatrix} = \begin{pmatrix} K & O \\ \hat{L} & \hat{A} \end{pmatrix}\begin{pmatrix} V & U \\ W & T \end{pmatrix} \qquad (4.3.5)$$

$$\text{and} \quad \begin{pmatrix} V & U \\ W & T \end{pmatrix}\begin{pmatrix} M \\ N \end{pmatrix} = \begin{pmatrix} M \\ \hat{N} \end{pmatrix}. \qquad (4.3.6)$$

From (4.3.6) we obtain

$$VM + UN = M \quad \text{and hence}$$

$$[V-I \quad U]\begin{pmatrix} M \\ N \end{pmatrix} = 0. \qquad (4.3.7)$$

From (4.3.5) we obtain

$$VK + UL = KV$$

and adding -K to each side we obtain

$$(V-I)K + UL = K(V-I). \qquad (4.3.8)$$

Also from (4.3.5) we obtain

$$UA = KU, \qquad (4.3.9)$$

so combining (4.3.8) and (4.3.9) we get

$$[V-I \ U]\begin{pmatrix} K & O \\ L & A \end{pmatrix} = K[V-I \ U]. \qquad (4.3.10)$$

Combining (4.3.7) and (4.3.10) we see that

$$[V-I \ U]\left(\begin{bmatrix} M \\ N \end{bmatrix}\begin{bmatrix} K & O \\ L & A \end{bmatrix}\begin{bmatrix} M \\ N \end{bmatrix}\cdots\cdots\begin{bmatrix} K & O \\ L & A \end{bmatrix}^n\begin{bmatrix} M \\ N \end{bmatrix}\right) = 0$$

which can only hold if [V-I U] = 0 because of the controllability assumption. Hence V = I and U = 0. □

We shall now prove a similar result to this lemma, involving the matrices in the expression (4.3.4).

Lemma 4.3.3

Let $$[d^T h^T]\begin{pmatrix} F & O \\ C & A \end{pmatrix}^i \begin{pmatrix} G \\ B \end{pmatrix} = [\hat{d}^T \hat{h}^T]\begin{pmatrix} F & O \\ \hat{C} & \hat{A} \end{pmatrix}^i \begin{pmatrix} G \\ \hat{B} \end{pmatrix} \text{ for all } i, \qquad (4.3.11)$$

where (h^T, A) and $(\hat{h}^T, \hat{A})$ are observable pairs and $\left(\begin{bmatrix} F & O \\ C & A \end{bmatrix}, \begin{bmatrix} G \\ B \end{bmatrix} \right)$ and $\left(\begin{bmatrix} F & O \\ \hat{C} & \hat{A} \end{bmatrix}, \begin{bmatrix} G \\ \hat{B} \end{bmatrix} \right)$ are reachable pairs. Then there exists a similarity transform relating these matrices which is of the form $\begin{pmatrix} I & O \\ Y & T \end{pmatrix}$ where T is invertible.

Proof: Let us first note that the observability assumptions imply that A and $\hat{A}$ are cyclic.

Without loss of generality let $n = \dim A \leq \dim \hat{A}$, and suppose that the characteristic equation of A is given by

$$A^n + a_1 A^{n-1} + \ldots + a_n I = 0. \qquad (4.3.12)$$

Consider the following equality which can be derived from the assumption (4.3.11):

$$[d^T h^T]\left(\begin{bmatrix} F & O \\ C & A \end{bmatrix}^n + a_1 \begin{bmatrix} F & O \\ C & A \end{bmatrix}^{n-1} + \ldots + a_n \begin{bmatrix} I & O \\ O & I \end{bmatrix} \right) \begin{bmatrix} F & O \\ C & A \end{bmatrix}^i \begin{bmatrix} G \\ B \end{bmatrix} =$$
$$[\hat{d}^T \hat{h}^T]\left(\begin{bmatrix} F & O \\ \hat{C} & \hat{A} \end{bmatrix}^n + a_1 \begin{bmatrix} F & O \\ \hat{C} & \hat{A} \end{bmatrix}^{n-1} + \ldots + a_n \begin{bmatrix} I & O \\ O & I \end{bmatrix} \right) \begin{bmatrix} F & O \\ \hat{C} & \hat{A} \end{bmatrix}^i \begin{bmatrix} G \\ \hat{B} \end{bmatrix}$$
$$(i = 0,1,2,\ldots)$$

It follows from (4.3.12) that the left-hand side of this expression reduces to $[u^T \; O]\begin{pmatrix} F & O \\ C & A \end{pmatrix}^i \begin{pmatrix} G \\ B \end{pmatrix}$ where u^T is defined in the obvious way, and the right-hand side can be expressed as $[\hat{u}^T \hat{w}^T]\begin{pmatrix} F & O \\ \hat{C} & \hat{A} \end{pmatrix}^i \begin{pmatrix} G \\ \hat{B} \end{pmatrix}$.

Hence

$$[\hat{u}^T \hat{w}^T]\begin{pmatrix} F & O \\ \hat{C} & \hat{A} \end{pmatrix}^i \begin{pmatrix} G \\ \hat{B} \end{pmatrix} = [u^T \; O]\begin{pmatrix} F & O \\ C & A \end{pmatrix}^i \begin{pmatrix} G \\ B \end{pmatrix}$$
$$= u^T F^i G$$
$$= [u^T \; O]\begin{pmatrix} F & O \\ \hat{C} & \hat{A} \end{pmatrix}^i \begin{pmatrix} G \\ \hat{B} \end{pmatrix}.$$

This implies that

$$[\hat{u}^T - u^T \; \hat{w}^T]\begin{pmatrix} F & 0 \\ \hat{C} & \hat{A} \end{pmatrix}^i \begin{pmatrix} G \\ \hat{B} \end{pmatrix} = 0 \quad \text{for all } i.$$

Hence $[\hat{u}^T - u^T \; \hat{w}^T] = 0$ because of the reachability assumption.

An immediate consequence of this is that $\hat{A}$ also satisfies (4.3.12), and since $\hat{A}$ is cyclic, we must have $\dim \hat{A} = \dim A$.

It now follows from the fact that $\hat{u} = u$ and $\hat{w} = 0$, that we can write a basis for the observability subspaces H and $\hat{H}$ of $\left([d^T h^T], \begin{bmatrix} F & 0 \\ C & A \end{bmatrix}\right)$ and $\left([\hat{d}^T \hat{h}^T], \begin{bmatrix} F & 0 \\ \hat{C} & \hat{A} \end{bmatrix}\right)$ as the rows of $\begin{pmatrix} U & 0 \\ P & Q \end{pmatrix} \triangleq H$ and $\begin{pmatrix} U & 0 \\ \hat{P} & \hat{Q} \end{pmatrix} \triangleq \hat{H}$ respectively, where $U = \begin{pmatrix} u^T F^{k-1} \\ \vdots \\ u^T F \\ u^T \end{pmatrix}$ where $k \le \dim F$.

Further, we know that Q and $\hat{Q}$ are full rank because (h^T, A) and $(\hat{h}^T, \hat{A})$ are observable pairs, so by rearranging the rows of H and $\hat{H}$, we can write down a basis for the observability subspaces as $\begin{pmatrix} U & 0 \\ V & I \end{pmatrix}$ and $\begin{pmatrix} U & 0 \\ \hat{V} & I \end{pmatrix}$.

Because of the invariant subspace property of H and $\hat{H}$ we can now write

$$\begin{pmatrix} U & 0 \\ V & I \end{pmatrix}\begin{pmatrix} F & 0 \\ C & A \end{pmatrix} = \begin{pmatrix} K & 0 \\ L & A \end{pmatrix}\begin{pmatrix} U & 0 \\ V & I \end{pmatrix} \tag{4.3.13}$$

for appropriate K and L, and

$$\begin{pmatrix} U & 0 \\ \hat{V} & I \end{pmatrix}\begin{pmatrix} F & 0 \\ \hat{C} & \hat{A} \end{pmatrix} = \begin{pmatrix} K & 0 \\ \hat{L} & \hat{A} \end{pmatrix}\begin{pmatrix} U & 0 \\ \hat{V} & I \end{pmatrix} \tag{4.3.14}$$

for appropriate $\hat{L}$. Note that K is the same in both (4.4.13) and (4.4.14). Furthermore, since $[d^T h^T] \subset H$ and $[\hat{d}^T \hat{h}^T] \subset \hat{H}$, we can write

$$[d^T h^T] = [k^T h^T]\begin{pmatrix} U & 0 \\ V & I \end{pmatrix} \tag{4.3.15}$$

$$\text{and } [\hat{d}^T \hat{h}^T] = [\hat{k}^T \hat{h}^T]\begin{pmatrix} U & 0 \\ \hat{V} & I \end{pmatrix} \tag{4.3.16}$$

for appropriate k^T and $\hat{k}^T$.

We can now write (4.3.11) in terms of minimal representations, namely

$$[k^T h^T]\begin{pmatrix} K & O \\ L & A \end{pmatrix}^i \begin{pmatrix} UG \\ VG+B \end{pmatrix} = [\hat{k}^T \hat{h}^T]\begin{pmatrix} K & O \\ \hat{L} & \hat{A} \end{pmatrix}^i \begin{pmatrix} UG \\ VG+\hat{B} \end{pmatrix} \quad (i = 0,1,\ldots)$$

By Lemma 4.3.2 we know that there exists a similarity transformation relating the two sets of matrices as follows:

$$\begin{pmatrix} I & O \\ W & T \end{pmatrix}\begin{pmatrix} K & O \\ L & A \end{pmatrix} = \begin{pmatrix} K & O \\ \hat{L} & \hat{A} \end{pmatrix}\begin{pmatrix} I & O \\ W & T \end{pmatrix}$$

so in particular $$\hat{A} = TAT^{-1} \tag{4.3.17}$$

$$WK + TL = \hat{L} + \hat{A}W \tag{4.3.18}$$

$$\begin{pmatrix} I & O \\ W & T \end{pmatrix}\begin{pmatrix} UG \\ VG+B \end{pmatrix} = \begin{pmatrix} UG \\ \hat{V}G+\hat{B} \end{pmatrix}$$

so in particular

$$WUG + TVG + TB = \hat{V}G + \hat{B} \tag{4.3.19}$$

and $$[k^T h^T] = [\hat{k}^T \hat{h}^T]\begin{pmatrix} I & O \\ W & T \end{pmatrix}$$

so in particular $$\hat{h}^T = h^T T^{-1} \tag{4.3.20}$$

and $$k^T = \hat{k}^T + \hat{h}^T W. \tag{4.3.21}$$

We shall now show that (4.3.13)-(4.3.20) together imply that the similarity transformation relating the two sets of matrices in (4.3.11) is $\begin{pmatrix} I & O \\ TV+WU-\hat{V} & T \end{pmatrix}$.

(i) Rearranging (4.3.19) we obtain

$$\hat{B} = TB + (TV+WU-\hat{V})G. \tag{4.3.22}$$

(ii) From (4.3.16) we have

$$\hat{d}^T = \hat{k}^T U + \hat{h}^T \hat{V}.$$

Substituting for $\hat{k}^T$ from (4.3.21) we obtain

$$\hat{d}^T = k^T U - \hat{h}^T WU + \hat{h}^T \hat{V}$$

and substituting for $k^T U$ from (4.3.15) and h^T from (4.3.20) we obtain

$$\hat{d}^T = d^T - \hat{h}^T V - h^T WU + \hat{h}^T \hat{V}$$

$$= d^T - h^T T^{-1}(TV+WU-\hat{V}). \tag{4.3.23}$$

(iii) From (4.3.14) we have

$$\hat{C} = \hat{L}U + \hat{A}\hat{V} - \hat{V}F$$

$$= (WK+TL-\hat{A}W)U + \hat{A}\hat{V} - \hat{V}F \text{ by (4.3.18)}.$$

Substituting for LU and KU from (4.3.13) we obtain

$$\hat{C} = WUF + T(VF+C-AV) - \hat{A}WU + \hat{A}\hat{V} - \hat{V}F$$

$$= (TV+WU-\hat{V})F + TC - TAT^{-1}(TV+WU-\hat{V}) \qquad (4.3.24)$$

on substituting for $\hat{A}$ from (4.3.17).

It is now clear that the relationships (4.3.17), (4.3.20), (4.3.22) (4.3.23) and (4.3.24) together give the required result. □

We are now in a position to put together the results of Lemmas 4.3.1 and 4.3.3 to provide the main result of this section.

Theorem 4.3.1

Let $M = (A_1,b_1,A_2,b_2,A,C,Q_1,Q_2,b,h^T,d^T)$ and $\hat{M} = (\hat{A}_1,\hat{b}_1,\hat{A}_2,\hat{b}_2,\hat{A},\hat{C},\hat{Q}_1,\hat{Q}_2,\hat{b},\hat{h}^T\hat{d}^T)$ be canonical realizations of a bilinear input/output map. Then there exist unique invertible matrices T_1, T_2 and T and a unique matrix Y such that the following relationships hold:

$$\hat{A}_1 = T_1A_1T_1^{-1} \quad \hat{b}_1 = T_1b_1 \quad \hat{A}_2 = T_2A_2T_2^{-1} \quad \hat{b}_2 = T_2b_2$$

$$\hat{A} = TAT^{-1} \qquad \hat{h}^T = h^TT^{-1}$$

$$\hat{C} = TCT_1^{-1}\otimes T_2^{-1} + YT_1\otimes T_2(A_1\otimes A_2)T_1^{-1}\otimes T_2^{-1} - TAT^{-1}Y$$

$$\hat{Q}_1 = TQ_1T_1^{-1} + YT_1\otimes T_2(A_1\otimes b_2)T_1^{-1} \qquad \hat{Q}_2 = TQ_2T_2^{-1} + YT_1\otimes T_2(b_1\otimes A_2)T_2^{-1}$$

$$\hat{b} = Tb + YT_1\otimes T_2(b_1\otimes b_2) \qquad \hat{d}^T = d^TT_1^{-1}\otimes T_2^{-1} - h^TT^{-1}Y.$$

Proof: The existence of T_1, T_2, T and Y follows from Lemmas 4.3.1 and 4.3.2. Uniqueness of T_1, T_2 and T follows immediately from the facts that (A_1,b_1) and (A_1,b_1) are reachable pairs and (h^T,A) is an observable pair.

To show uniqueness of Y, suppose that Y_1 also satisfies the above equalities. In particular, we obtain from the equalities for $\hat{C}$ and $\hat{d}^T$ the following:

$$\hat{C} - TCT_1^{-1}\otimes T_2^{-1} = Y\hat{A}_1\otimes\hat{A}_2 = \hat{A}Y = Y_1\hat{A}_1\otimes\hat{A}_2 - \hat{A}Y_1$$

$$\text{i.e. } (Y-Y_1)\hat{A}_1\otimes\hat{A}_2 = \hat{A}(Y-Y_1) \tag{4.3.25}$$

and $$\hat{d}^T - d^T T_1^{-1}\otimes T_2^{-1} = -\hat{h}^T Y = -\hat{h}^T Y$$

$$\text{i.e. } h^T(Y-Y_1) = 0. \tag{4.3.26}$$

Using (4.3.25) and (4.3.26) we obtain

$$\begin{pmatrix} \hat{h}^T \\ \vdots \\ \hat{h}^T\hat{A}^{n-1} \end{pmatrix}(Y-Y_1) = 0$$

so that (h^T,A) observable implies $Y = Y_1$. □

We are now in a position to obtain a result connecting minimal and canonical realizations:

<u>Theorem 4.3.2</u>

A realization of a bilinear input/output map is minimal iff it is canonical.

Proof: Referring back to §4.2, we see that reduction to quasi-reachable form from an observable realization is equivalent to linear system reduction to reachable form of the pair $\begin{pmatrix} A_1\otimes A_2 & 0 \\ C & A \end{pmatrix}, \begin{pmatrix} A_1\otimes b_2 & b_1\otimes A_2 & b_1\otimes b_2 \\ Q_1 & Q_2 & b \end{pmatrix}$.

From this we can deduce, with the aid of Theorem 4.3.1, that there exists a unique mapping taking any observable realization of a bilinear input/output map f to any specified canonical realization of f. Hence, according to Definition 3.1.4, a canonical realization is a minimal realization.

Conversely, let M be a minimal realization of f; then by Definition 3.1.4, it must be observable. Suppose, however, that M is not a canonical realization; it follows, then, that M is not quasi-reachable. If this is the case, there will not in general be a unique mapping from any observable realization of f to M, a contradiction of M being minimal. Hence we deduce that M is canonical. □

Remark:

It is now apparent from this result and from the reduction procedures that produce observable and quasi-reachable realizations that a minimal realization is one with the smallest number of states necessary to describe the input/output map in state space form.

4.4 Canonical Forms

We present a few definitions from [D1] before getting down to the main business of presenting realizations of bilinear input/output maps which are unique with respect to their structure and which contain as many fixed zeros and ones as possible.

Definition 4.4.1

Let E be an equivalence relation on the set S. A set of canonical forms for S under E is a subset C of S such that for all $s \in S$, there exists a unique $c \in C$ such that sEc. Let $\emptyset: S \to C$ be defined by $\emptyset(s) = c$. Clearly $\text{Im}\emptyset = C \cong S/E$.

Definition 4.4.2

A function $f: S \to V$ is an invariant (for S under E) if for all $s_1, s_2 \in S$, $s_1 E s_2$ implies $f(s_1) = f(s_2)$. f is a complete invariant if $f(s_1) = f(s_2) \leftrightarrow s_1 E s_2$. f is an independent invariant if Im f = V.

Clearly $\emptyset: S \to C$ is a complete independent invariant, and conversely, a complete independent invariant $f: S \to V \subset S$ generates canonical forms (V is a set of canonical forms).

For bilinear dynamical systems, we shall say that $M_1 E M_2$ if the transfer functions obtained from them are equal. For single input/single output linear systems it is well known that the following two realizations are canonical, each having invariants $\{a_1, \ldots, a_n, c_1, \ldots, c_n\}$ and

$\{a_1,\ldots,a_n,b_1,\ldots,b_n\}$ which uniquely specify the transfer function of the system

(i) $c^T = [c_1 \ldots c_n]$ $A = \begin{bmatrix} 0 & 1 & & 0 \\ & & \ddots & \\ & 0 & & 1 \\ -a_n & \cdots & \cdots & -a_1 \end{bmatrix}$ $b = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}$

(ii) $c^T = [0 \ldots 0\ 1]$ $A = \begin{bmatrix} 0 & & & a_n \\ 1 & & 0 & \vdots \\ & \ddots & & \vdots \\ 0 & & 1 & -a_1 \end{bmatrix}$ $b = \begin{bmatrix} b_1 \\ \vdots \\ \vdots \\ b_n \end{bmatrix}$

Given a realization $M = (A_1,b_1,A_2,b_2,A,C,Q_1,Q_2,b,d^T,h^T)$ of a bilinear input/output map we know that (A_1,b_1) and (A_2,b_2) are reachable pairs, so that we can set up these pairs like A and b in (i), and in addition (h^T,A) is an observable pair so we can set this up like (c^T,A) in (ii).

We can now ask whether there is any representation of the remaining matrices specifying M which has a similar well-defined form and is also unique. The answer to this is in the affirmative, and we can derive several different canonical forms.

Before specifying these canonical forms we remark that having defined the form of the matrices A_1,b_1,A_2,b_2,A and h^T, we are no longer permitted to transform M via the similarity transformations T_1, T_2 and T, defined in Theorem 4.4.4, and the only freedom allowed us with regard to changing parameters is therefore via the matrix $Y \in R^{n \times n_1 n_2}$.

Canonical Form 1:

We assume that n_1,n_2 and n are all greater than zero. If any of these are zero, then clearly $Y = 0$.

The canonical form presented here will be specified by $d^T = [0 \ldots 0]$ and all rows of C except the first are zero. We shall show that not only does this canonical form exist, but that it is unique. An immediate corollary of this will be that a complete set of independent invariants for an input/output map f will be given by $\{\underline{a}_1, \underline{a}_2, \underline{a}, \underline{c}, Q_1, Q_2, b\}$ where $\underline{a}_1, \underline{a}_2$ are the bottom rows of A_1, A_2 respectively, $\underline{a}$ is the last column of A and $\underline{c}$ is the first row of C.

From Theorem 4.4.4 we have

$$\hat{d}^T = d^T - h^T Y \quad \text{and} \quad \hat{C} = C + YA_1 \otimes A_2 - AY.$$

We can obtain $\hat{d}^T = 0$ by setting the last row of Y equal to d^T, for then $d^T = d^T - [0.....0\ 1]\begin{pmatrix} Y_1 \\ d^T \end{pmatrix} = 0$ where $Y = \begin{pmatrix} Y_1 \\ \hline d^T \end{pmatrix}$.

If we now define $\hat{C} = \begin{pmatrix} \hat{c}_1^T \\ \vdots \\ \hat{c}_n^T \end{pmatrix}$, $C = \begin{pmatrix} c_1^T \\ \vdots \\ c_n^T \end{pmatrix}$, $Y = \begin{pmatrix} y_1^T \\ \vdots \\ y_n^T \end{pmatrix}$ where $y_n^T = d^T$

we obtain
$$\begin{pmatrix} \hat{c}_1^T \\ \vdots \\ \hat{c}_n^T \end{pmatrix} = \begin{pmatrix} c_1^T \\ \vdots \\ c_n^T \end{pmatrix} + \begin{pmatrix} y_1^T \\ \vdots \\ y_n^T \end{pmatrix} A_1 \otimes A_2 - \begin{pmatrix} 0 & & & -a_1 \\ 1 & \ddots & 0 & \vdots \\ & \ddots & \ddots & \vdots \\ 0 & & 1 & -a_n \end{pmatrix} \begin{pmatrix} y_1^T \\ \vdots \\ y_n^T \end{pmatrix}$$

and hence
$$\hat{c}_k^T = c_k^T + y_k^T A_1 \otimes A_2 - y_{k-1}^T + a_k y_n^T \quad (k = 2,\ldots,n)$$
$$y_n^T = d^T.$$

To obtain $\hat{c}_k^T = 0 \quad k = 2,\ldots,n$ we choose y_k^T sequentially in the manner

$$\begin{aligned} y_{k-1}^T &= c_k^T + y_k^T A_1 \otimes A_2 + a_k y_n^T \\ &= c_k^T + y_k^T A_1 \otimes A_2 + a_k d^T. \end{aligned}$$

This gives unique values for $\{y_k\}$, and hence

$$\hat{c}_1^T = c_1^T + y_1^T A_1 \otimes A_2 + a_1 d^T$$

is uniquely specified. Finally, since it is now apparent that Y is uniquely defined, we must have $\hat{Q}_1 = Q_1 + YA_1 \otimes b_2$, $\hat{Q}_2 = Q_2 + Y(b_1 \otimes A_2)$ and $\hat{b} = b + Y(b_1 \otimes b_2)$ uniquely defined.

Note that using the nn_1n_2 elements of Y we have specified a canonical form with nn_1n_2 zeros.

Canonical Form 2

The canonical form presented here will be specified by $b = [0 \ldots 0]^T$, all columns except the first of Q_1 and Q_2 are zero, and C is structured as follows:

$$C = \left[\begin{array}{cccccccc} \times\cdots\cdots\times & \times & & \times & & & \times & \\ \vdots \qquad\quad \vdots & \vdots & 0 & \vdots & 0 & \cdots\cdots & \vdots & 0 \\ \times\cdots\cdots\times & \times & & \times & & & \times & \end{array}\right]$$

$$\underbrace{\longleftrightarrow}_{n_1 \text{ columns}} \quad n_1+1 \quad 2n_1+1 \qquad (n_2-1)n_1+1$$

Again, we shall show that this canonical form exists and that it is unique. We shall then find that a complete set of independent invariants is given by $\{\underline{a}_1, \underline{a}_2, \underline{a}, \underline{q}_1, \underline{q}_2, C', d^T\}$, where $\underline{a}_1, \underline{a}_2$ are the bottom rows of A_1, A_2 respectively, $\underline{a}$ is the last column of A, $\underline{q}_1$ and $\underline{q}_2$ are the first columns of Q_1 and Q_2 respectively and C' is the $n \times (n_1+n_2-1)$ matrix made up of the non-zero columns of C.

From Theorem 4.4.4 we have

$$\hat{b} = b + Y(b_1 \otimes b_2) \qquad \hat{Q}_1 = Q_1 + Y(A_1 \otimes b_2) \qquad \hat{Q}_2 = Q_2 + Y(b_1 \otimes A_2)$$

$$\hat{C} = C + Y(A_1 \otimes A_2) - AY.$$

Defining $Y = [y_{11} \ldots y_{n_1,1}\, y_{12} \ldots y_{n_1,2} \ldots y_{1n_2} \ldots y_{n_1n_2}]$

and letting $\qquad y_{n_1n_2} = -b$

it is clear that

$$\hat{b} = b + Y(b_1 \otimes b_2) = b + [y_{11} \ldots y_{n_1n_2}]\begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Defining $\hat{Q}_1 = [\hat{q}_{11}\hat{q}_{12} \ldots \hat{q}_{1n_1}] \quad Q_1 = [q_{11}q_{12} \ldots q_{1n_1}]$,

we can write

$$[\hat{q}_{11} \ldots \hat{q}_{1n_1}] = [q_{11} \ldots q_{1n_1}] + [y_{11} \ldots y_{n_1n_2}]\begin{pmatrix} & 0 & \\ & 0 & \\ & & \\ 0\,1 & & 0 \\ 0 & & 1 \\ -a_1 & \cdots & -a_{n_1} \end{pmatrix}$$

and hence

$$\hat{q}_{1k} = q_{1k} - a_k^1 y_{n_1n_2} + y_{k-1,n_2} \qquad k = 2,\ldots,n,$$

$$y_{n_1n_2} = -b.$$

To obtain $\hat{q}_{1k} = 0 \quad k = 2,\ldots,n,$

we choose y_{kn_2} sequentially as

$$y_{k-1,n_2} = a_k^1 y_{n_1n_2} - q_{1k}$$
$$= -a_k^1 b - q_{1k}.$$

Defining $\hat{Q}_1$ and Q_2 similarly, it is clear that we can obtain $\hat{q}_{2k} = 0 \quad k = 2,\ldots,n_2$ by choosing y_{n_1k} sequentially as

$$y_{n_1,k-1} = -a_k^2 b - q_{2k}.$$

Note that $\quad \hat{q}_{11} = q_{11} - a_1^1 b$

and $\quad \hat{q}_{21} = q_{21} - a_1^2 b$

are uniquely specified.

Finally, defining $\hat{C} = [\hat{c}_{11}\ldots\hat{c}_{n_1,1}\ldots\hat{c}_{1,n_2}\ldots\hat{c}_{n_1n_2}]$

and $C = [c_{11}\ldots c_{n_1,1}\ldots c_{1,n_2}\ldots c_{n_1n_2}]$

we can write

$$[\hat{c}_{11}\ldots\hat{c}_{n_1n_2}] = [c_{11}\ldots c_{n_1n_2}] + A[y_{11}\ldots y_{n_1n_2}]$$

$$-[y_{11}\ldots y_{n_1n_2}]\begin{bmatrix} 0 & \begin{matrix} 0\;1 & & 0 \\ & \ddots & \\ 0 & & 1 \\ -a_1^1 & \cdots & -a_{n_1}^1 \end{matrix} & & 0 \\ \vdots & & \ddots & \\ \vdots & 0 & & \begin{matrix} 0\;1 & & 0 \\ & \ddots & \\ & & 1 \\ -a_1^1 & \cdots & -a_{n_1}^1 \end{matrix} \\ 0 & & & \\ -a_1^2A_1 & \cdots & \cdots & -a_{n_2}^2A_1 \end{bmatrix}$$

where

$$A_1 = \begin{pmatrix} 0 & 1 & & 0 \\ & \ddots & \ddots & \\ & 0 & \ddots & 1 \\ -a_1^1 & \cdots & \cdots & -a_{n_1}^1 \end{pmatrix}$$

and hence

$$\hat{c}_{jk} = c_{jk} + Ay_{jk} - y_{j-1,k-1} + a_k^1 y_{j-1,n_1} + a_j^2 \bar{Y} A_1 \qquad j = 2,\ldots,n_1$$

$$\text{where } \bar{Y} \triangleq [y_{1n_2} \cdots y_{n_1 n_2}]. \qquad k = 2,\ldots,n_2$$

Hence we can set $\{\hat{c}_{jk} : j = 2,\ldots,n_1; \ k = 2,\ldots,n_2\}$ to zero by appropriate choice of $\{y_{j-1,k-1} : j = 2,\ldots,n_1, \ k = ,\ldots,n_2\}$, and it is clear that the remaining columns of $\hat{C}$ are uniquely defined.

Note again that the number of zeros we have inserted into $\hat{C},\Omega_1,\Omega_2$ and b is given by $n(n_1-1)(n_2-1) + n(n_1-1) + n(n_2-1) + n = nn_1n_2$.

CHAPTER 5. MULTI-OUTPUT BILINEAR SYSTEMS

5.1 Preamble

Before discussing canonical and minimal realizations of multioutput bilinear systems, we shall examine the following example of a two-output bilinear input/output map in the context of observability and quasi-reachability of single output maps.

Let the map be represented by the transfer function $s \in R^{real}[(z_1, z_2)]$.

$$s = \frac{1}{z_1(z_2^2+az_2+b)} \begin{pmatrix} 1 \\ z_2 \end{pmatrix} \tag{5.1.1}$$

Following on from our discussion of state-space descriptions of bilinear input/output maps in Chapter 2, an obvious choice of state-space realization in this case is

$$x^1_{k+1} = u_k \qquad x^2_{k+1} = \begin{pmatrix} 0 & 1 \\ -b & -a \end{pmatrix} x^2_k + \begin{pmatrix} 0 \\ 1 \end{pmatrix} v_k$$
$$y_k = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} x^1_k \otimes x^2_k. \tag{5.1.2}$$

Using the notation of Chapter 4, we see that

$$H = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad T_1 = 0 \qquad T_2 = \begin{pmatrix} 0 & 1 \\ -b & -a \end{pmatrix}$$

and this implies that the realization (5.1.2) is not observable. In particular, if the initial substate $x^2_0 = 0$, it is impossible to determine the value of the initial substate x^1_0 by any sequence of experiments. So, to construct an observable realization, we employ the reduction procedure of §4.2 to obtain the following:

$$x^2_{k+1} = \begin{pmatrix} 0 & 1 \\ -b & -a \end{pmatrix} x^2_k + \begin{pmatrix} 0 \\ 1 \end{pmatrix} v_k$$
$$x_{k+1} = \begin{pmatrix} 0 & 1 \\ -b & -a \end{pmatrix} x^2_k u_k + \begin{pmatrix} 0 \\ 1 \end{pmatrix} v_k u_k \tag{5.1.3}$$
$$y_k = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} x_k.$$

It is now clear that not only is this a state space with higher dimension than the one we started off with, but it is also not quasi-reachable since $x_{k+1} = u_k x_{k+1}^2$. The reason that we obtained a higher-dimensional state space can be explained by writing H in the form described in §4.2, i.e.

$$H = \begin{pmatrix} U & O \\ V & O \\ W & In \end{pmatrix}$$

and in our case here we have n = O and V = I, so that H = V. Whereas in the single output case V could be written in the form

$$V = \begin{pmatrix} v^T A_1^k \otimes A_2^k \\ \vdots \\ v^T A_1 \otimes A_2 \\ v^T \end{pmatrix}$$

it is impossible to do this for (5.1.2) for any v^T, since

$$A_1 \otimes A_2 = [0] \otimes \begin{pmatrix} 0 & 1 \\ -b & -a \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

so that the proof showing that the state space dimension does not increase breaks down at this point.

Turning our attention now to the reason for (5.1.3) not being quasi-reachable, we can calculate the transfer function s in a natural way from (5.1.3) as

$$s = \frac{1}{z_1 z_2} \begin{pmatrix} \dfrac{z_2}{z_2^2+az_2+b} \\ \dfrac{z_2^2}{z_2^2+az_2+b} \end{pmatrix}$$

and our proof of Lemma 3.2.2 indicates that because both elements of the transfer function vector have numerators containing terms in $z_2, z_2^2, \ldots$ but no terms in $1, z_1, z_1^2, \ldots$, the state space is not quasi-reachable.

It follows also from Lemma 3.2.2 that if the A matrix has more than two Jordan blocks corresponding to zero eigenvalues, or equivalently if the dimension of the null-space of A is greater than two, then the state space is certainly not quasi-reachable. However, suppose that A has two Jordan blocks of zero eigenvalues, and the input to state transfer functions corresponding to these are written as

$$\begin{pmatrix} x_1(z_1,z_2) \\ \vdots \\ x_{m_1}(z_1,z_2) \end{pmatrix} = \frac{1}{(z_1z_2)^{m_1}} \begin{pmatrix} (z_1z_2)^{m_1-1} \ldots\ldots & 1 \\ & \ddots \\ 0 & & (z_1z_2)^{m_1-1} \end{pmatrix} \begin{pmatrix} R_1 \\ \\ R_{m_1} \end{pmatrix} \frac{1}{\psi_1(z_1)\psi_2(z_2)}$$

and

$$\begin{pmatrix} x_{m_1+1}(z_1,z_2) \\ \\ x_{m_2}(z_1,z_2) \end{pmatrix} = \frac{1}{(z_1z_2)^{m_2}} \begin{pmatrix} (z_1z_2)^{m_2-1} \ldots\ldots & 1 \\ \vdots & \ddots \\ 0 & & (z_1z_2)^{m_2-1} \end{pmatrix} \begin{pmatrix} R_{m_1+1} \\ \vdots \\ R_{m_1+m_2} \end{pmatrix} \frac{1}{\psi_1(z_1)\psi_2(z_2)}$$

where $R_1,\ldots,R_{m_1+m_2} \in \mathcal{R}[z_1,z_2]$.

Then the state space corresponding to these will be quasi-reachable if the terms of R_{m_1} and $R_{m_1+m_2}$ which remain after discarding those divisible by z_1z_2 are not all divisible by z_1 or all divisible by z_2. This is an immediate consequence of the detailed discussion in the proof of Lemma 3.2.2, where we discovered that the crucial terms in the study of reachability for zero-eigenvalue modes were those terms in the numerator which were not divisible by z_1z_2.

We now have an indication of how to test whether a state space realization of multi-output bilinear map is quasi-reachable or not, and this is formalized in the next section. Complementing this, in §5.3 we provide an algorithm for reducing a realization which is not quasi-reachable to one which is.

There is another problem that remains, however, and that is the question of observability, and in order to cope with a realization such as (5.1.2) it is necessary, in §5.4, to formulate the concept of quasi-observability. As we shall see, this is not quite enough to generate a theory of minimal realizations (Definition 3.1.4), and this remains an open problem.

5.2 Quasi-Reachability

We have seen in Chapter 2 how to construct a realization of a bilinear input/output map. We can therefore assume that for the multi-output case it is in general possible to obtain a realization of the form

$$\begin{aligned} x^1_{k+1} &= A_1 x^1_k + b_1 u_k \\ x^2_{k+1} &= A_2 x^2_k + b_2 v_k \\ x_{k+1} &= A x_k + C x^1_k \otimes x^2_k + Q_1 x^1_k v_k + Q_2 x^2_k u_k + b u_k v_k \\ y_k &= H x_k + D x^1_k \otimes x^2_k . \end{aligned}$$

We shall assume that (A_1,b_1) and (A_2,b_2) are reachable pairs, and that the pair $\left(\begin{bmatrix} A_1\otimes A_2 & 0 \\ C & A \end{bmatrix}, \begin{bmatrix} A_1\otimes b_2 & b_1\otimes A_2 & b_1\otimes b_2 \\ Q_1 & Q_2 & b \end{bmatrix}\right)$ is also reachable and (H,A) is an observable pair. If any of these conditions do not hold, then we know from standard linear system theory and from the results of Chapter 4 how to remedy this. Note that these conditions do not imply quasi-reachability of the state space, but only that the components of the transfer functions $x(z_1,z_2)$ and $x^1(z_1)\otimes x^2(z_2)$ are linearly independent, and quasi-reachability will certainly break down if dim ker $A > 2$.

Before determining conditions for quasi-reachability in the case dim ker $A = 2$, we shall give an example of how quasi-reachability works in this case. To show quasi-reachability for general cases of dim ker $A = 2$, the argument proceeds in a similar way to that of Lemma 3.2.2, and we do not include it here.

Example: We shall just consider the transfer functions involving zero-eigenvalue modes of A, and assume that a reachable state space has been set up for the linear sub-systems involved. For example, let

$$s_1 = (z_1^2 + z_2)/z_1 z_2 \psi_1(z_1)\psi_2(z_2)$$
$$s_2 = (z_1 + kz_2^2)/z_1 z_2 \psi_1(z_1)\psi_2(z_2).$$

From Lemma 3.12 we know that the only inputs affecting the output at time + 1 of s_1 and s_2 are those of the form

$$(\alpha\psi_1(z_1) + q_1(z_1), \beta\psi_2(z_2) + q_2(z_2)) \in U \times V,$$

where α and β are scalars.

The output is then given by

$$y_1 = \frac{1}{z_1 z_2}\left\{\frac{\alpha q_2(z_2) z_2}{\psi_2(z_2)} + \frac{\beta q_1(z_1) z_1^2}{\psi_1(z_1)}\right\} \odot \sum (z_1 z_2)^{-k} + y_1^q$$

$$y_2 = \frac{1}{z_1 z_2}\left\{\frac{k\alpha q_2(z_2) z_2}{\psi_2(z_2)} + \frac{\beta q_1(z_1) z_1}{\psi_1(z_1)}\right\} \odot \sum (z_1 z_2)^{-k} + y_2^q$$

(following the notation of Chapter 3).

If we now write out the series expansions of q_1/ψ_1 and q_2/ψ_2 as

$$q_1/\psi_1 = f_1 z_1^{-1} + f_2 z_1^{-2} + \ldots$$
$$q_2/\psi_2 = g_1 z_2^{-1} + g_2 z_2^{-2} + \ldots$$

the output at time + 1 is given by

$$y_{10} = \alpha g_1 + \beta f_2 + \bar{y}_1^q$$
$$y_{20} = \alpha k g_2 + \beta f_1 + \bar{y}_2^q$$

and these two simultaneous equations can be solved for α and β provided that $f_1 g_1 - k f_2 g_2 \neq 0$. This of course is just a restriction on the components of $q_1(z_1)$ and $q_2(z_2)$ not to lie on some manifold $\in R^{n_1+n_2+2}$. When this condition is satisfied, any given value of y_{10} and y_{20} can be attained, so that the whole system is quasi-reachable.

We shall now determine necessary and sufficient conditions for the state space realization of a bilinear input/output map to be quasi-reachable when dim ker A = 2. We shall assume that A has been transformed by a similarity transformation $A \to TAT^{-1}$ in such a way that it can be written as

$$A = \left(\begin{array}{c|cc} A' & \multicolumn{2}{c}{A''} \\ \hline 0\ldots\ldots0 & 0 & 0 \\ 0\ldots\ldots0 & 0 & 0 \end{array}\right)$$

The transfer function of the zero-eigenvalue modes will then be given by

$$s = \frac{1}{z_1z_2}[C''(z_1I-A_1)^{-1}b_1\boxtimes(z_2I-A_2)^{-1}b_2 + Q_1''(z_1I-A_1)^{-1}b_1 + Q_2''(z_2I-A_2)^{-1}b_2 + b'']$$

where C'', Q_1'', Q_2'' and b'' respresent the bottom two rows of C, Q_1, Q_2 and b.

Expanding $(z_1I-A_1)^{-1}$ and $(z_2I-A_2)^{-1}$ as

$$(z_1I-A_1)^{-1} = z_1^{n_1-1}I + z_1^{n_1-2}(A_1+\alpha_1I) + \ldots + (A_1^{n_1-1} + \ldots + \alpha_{n_1-1}I)$$

$$(z_2I-A_2)^{-1} = z_2^{n_2-1}I + z_2^{n_2-2}(A_2+\beta_1I) + \ldots + (A_2^{n_2-1} + \ldots + \beta_{n_2-1}I)$$

where $A_1^{n_1} + \alpha_1A_1^{n_1-1} + \ldots + \alpha_{n_1-1}A_1 + \alpha_{n_1}I = 0$

and $A_2^{n_2} + \beta_1A_2^{n_2-1} + \ldots + \beta_{n_2-1}A_2 + \beta_{n_2}I = 0$ represent the characteristic polynomials of A_1 and A_2, we obtain

$$s = \frac{1}{z_1z_2\psi_1\psi_2}\Big\{C''[z_1^{n_1-1}I+..+(A_1^{n_1-1}+..+\alpha_{n_1-1}I)]b_1\boxtimes[z_2^{n_2-1}I+..+(A_2^{n_2-1}+..+\beta_{n_2-1}I)]b_2$$
$$+ Q_1''[z_1^{n_1-1}I+..+(A_1^{n_1-1}+..+\alpha_{n_1-1}I)]b_1\psi_2(z_2)$$
$$+ Q_2''[z_2^{n_2-1}I+..+(A_2^{n_2-1}+..+\beta_{n_2-1}I)]b_2\psi_1(z_1) + \psi_1(z_1)\psi_2(z_2)b''\Big\}.$$

We have seen by Lemma 3.2.2 that the system is not quasi-reachable if and only if either all the coefficients of $1, z_1, z_1^2, \ldots$ inside the brackets $\{\;\}$ are zero or else all the coefficients of $1, z_2, z_2^2, \ldots$ are zero.

Suppose the coefficients of $1, z_1, z_1^2, \ldots$ are all zero; this is equivalent to

$$C''[z_1^{n_1-1}I+\ldots+(A_1^{n_1-1}+\ldots+\alpha_{n_1-1}I)]b_1\boxtimes[A_2^{n_2-1}+\ldots+\beta_{n_2-1}I]b_2$$
$$+ Q_1''[z_1^{n_1-1}I+\ldots+(A_1^{n_1-1}+\ldots+\alpha_{n_1-1}I)]b_1\beta_{n_2}$$
$$+ Q_2''[A_2^{n_2-1}+\ldots+\beta_{n_2-1}I]b_2(z_1^{n_1}+\ldots+\alpha_{n_1}) + \beta_{n_2}(z_1^{n_1}+\ldots+\alpha_{n_1})b'' = 0.$$

Taking the coefficients of $z_1^{n_1}, \ldots, 1$, in turn, we obtain

$z_1^{n_1}$: $\quad Q_2''(A_2^{n_2-1}+\ldots+\beta_{n_2-1}I)b_2 + \beta_{n_2}b'' = 0$

$z_1^{n_1-1}$: $\quad C''b_1\boxtimes(A_2^{n_2-1}+\ldots+\beta_{n_2-1}I)b_2+\beta_{n_2}Q_1''b_1+\alpha_1Q_2''(A_2^{n_2-1}+\ldots+\beta_{n_2-1}I)b_2+\alpha_1\beta_{n_2}b'' = 0$

1: $\quad C''(A_1^{n_1-1}+\ldots+\alpha_{n_1-1}I)b_1\boxtimes(A_1^{n_1-1}+\ldots+\beta_{n_2-1}I)b_2$

$$+ \beta_{n_2}Q_1''(A_1^{n_1-1}+\ldots+\alpha_{n_1-1}I)b_1 + \alpha_{n_1}Q_2''(A_2^{n_2-1}+\ldots+\beta_{n_2-1}I)b_2 + \alpha_{n_1}\beta_{n_2}b'' = 0.$$

By subtracting $\alpha_i[Q_2''(A_2^{n_2-1}+\ldots+\beta_{n_2-1}I)b_2 + \beta_{n_2}b''] = 0$ from each of these equalities in turn we obtain

$$C''b_1 \otimes (A_2^{n_2-1}+\ldots+\beta_{n_2-1}I)b_2 + \beta_{n_2}Q_1''b_1 = 0$$

$$\cdot\;\cdot\;\cdot\;\cdot\;\cdot$$

$$C''(A_1^{n_1-1}+\ldots+\alpha_{n_1-1}I)b_1 \otimes (A_2^{n_2-1}+\ldots+\beta_{n_2-1}I)b_2 + \beta_{n_2}Q_1''(A_1^{n_1-1}+\ldots+\alpha_{n_1-1}I)b_1 = 0.$$

Now $b_1, (A_1+\alpha_1 I)b_1, \ldots, (A_1^{n_1-1}+\ldots+\alpha_{n_1-1}I)b_1$ are linearly independent and span R^{n_1}. Hence this series of equalities reduces to

$$\begin{aligned} &C''I \otimes (A_2^{n_2-1}+\ldots+\beta_{n_2-1}I)b_2 + \beta_{n_2}Q_1'' = 0 \\ \text{and}\quad &Q_2''(A_2^{n_2-1}+\ldots+\beta_{n_2-1}I)b_2 + \beta_{n_2}b'' = 0 \end{aligned} \qquad (5.2.1)$$

which are the necessary and sufficient conditions for the coefficients of $1, z_1, \ldots, z_1^{n_1}$ to vanish. Note that for $\beta_{n_2} \neq 0$, this is equivalent to

$$\begin{aligned} &C''I \otimes A_2^{-1}b_2 = Q_1'' \\ &Q_2''A_2^{-1}b_2 = b''. \end{aligned} \qquad (5.2.1)'$$

In a similar way we obtain the necessary and sufficient conditions for the coefficients of $1, z_2, \ldots, z_2^{n_2}$ to vanish as

$$\begin{aligned} &C''(A_1^{n_1-1}+\ldots+\alpha_{n_1-1}I)b_1 \otimes I + \alpha_{n_1}Q_2'' = 0 \\ \text{and}\quad &Q_1''(A_1^{n_1-1}+\ldots+\alpha_{n_1-1}I)b_1 + \alpha_{n_1}b'' = 0. \end{aligned} \qquad (5.2.2)$$

Hence, provided that neither (5.2.1) nor (5.2.2) are satisfied, and that the reachability conditions above hold, we know that the system is quasi-reachable. Note that we no longer require A to be a cyclic matrix. In fact, our proof of reachability in Chapter 3 was general enough to guarantee that provided that the components of $x(z_1,z_2)$ and $x^1(z_1) \otimes x^2(z_2)$ were linearly independent, then the whole system is quasi-reachable (if we include the special condition for dim ker A = 2). So for example the system

$$x_{k+1}^1 = ax_k^1 + u_k \qquad x_{k+1}^2 = bx_k^2 + v_k$$

$$x_{k+1} = \begin{pmatrix} c & 0 \\ 0 & c \end{pmatrix} x_k + \begin{pmatrix} 1 \\ 0 \end{pmatrix} x_k^1 x_k^2 + \begin{pmatrix} 0 \\ 1 \end{pmatrix} x_k^1 v_k$$

is reachable for $c \neq 0$ and quasi-reachable for $c = 0$.

5.3 Reduction to Quasi-Reachable Realization

We shall consider three possible alternatives as to why a multi-output bilinear state space realization is not quasi-reachable, and show how to reduce the state space in each case. This will then provide us with an interative procedure (with a finite number of iterations) for reducing any state space realization of a multi-output bilinear system to quasi-reachable form. The three possible alternatives are:

(1) dim ker $A \geq 2$ and (5.2.1) holds but not (5.2.2)

(2) dim ker $A \geq 2$ and (5.2.2) holds but not (5.2.1)

(3) dim der $A > 2$ and neither (5.2.1) nor (5.2.2) hold.

Note that we need not consider the case where both (5.2.1) and (5.2.2) hold, for then the numerator of the transfer function of the zero-eigenvalue modes would be divisible by z_1z_2. Thus we would have a cancellation of z_1z_2 in both numerator and denominator, and the transfer functions would then be linearly dependent on the components of $x^1(z_1)\otimes x^2(z_2)$. This case has then been covered by the reduction procedure of Chapter 4.

It is also clear that the reduction procedure for case (2) will be completely analogous to that of case (1), so we shall only deal with the latter.

In all three cases we shall assume that the x_k subsystem is written as

$$x_{k+1} = \begin{pmatrix} A' & A'' \\ O & Om \end{pmatrix} x_k + \begin{pmatrix} C' \\ C'' \end{pmatrix} x_k^1 \otimes x_k^2 + \begin{pmatrix} Q_1' \\ Q_1'' \end{pmatrix} x_k^1 v_k + \begin{pmatrix} Q_2' \\ Q_2'' \end{pmatrix} x_k^2 u_k + \begin{pmatrix} b' \\ b'' \end{pmatrix} u_k v_k \quad (5.3.1)$$

with $y_k = [H' \; H'']x_k + Dx_k^1 \otimes x_k^2$ where m = dim ker A and the matrices (A_1, b_1) and (A_2, b_2) are in reachable canonical form.

<u>Case (1)</u> Here we have no terms in $1, z_1, \ldots, z_1^{n_1}$ in the numerators of the transfer functions, so that they must all be divisible by z_2. Therefore we can write

$$s(z_1,z_2) = z_2R/z_1z_2\psi_1(z_1)\psi_2(z_2)$$
$$= R/z_1\psi_1(z_1)\psi_2(z_2) \qquad \text{for } R \in \mathcal{R}[z_1,z_2].$$

It is now clear that the sensible path to take to obtain a quasi-reachable realization is to increase the dimension of the x_k^1 subsystem by 1, and to dispense with all the zero-eigenvalue modes of x_k. Let us label these modes by $\bar{x}_k$ so that x_k may be written as $x_k = \begin{pmatrix} \hat{x}_k \\ \bar{x}_k \end{pmatrix}$; then the system equation for these is given by

$$\bar{x}_{k+1} = C''x_k^1 \otimes x_k^2 + Q_1''x_k^1v_k + Q_2''x_k^2u_k + b''u_kv_k. \tag{5.3.2}$$

Consider the case $\det A_2 = \beta_{n_2} \neq 0$. Substituting into this equation from (5.2.1)', we obtain

$$\begin{aligned}\bar{x}_{k+1} &= C''x_k^1 \otimes x_k^2 + C''x_k^1 \otimes A_2^{-1}b_2v_k + Q_2''x_k^2u_k + Q_2''A_2^{-1}b_2u_kv_k \\ &= C''x_k^1 \otimes A_2^{-1}x_{k+1}^2 + Q_2''A_2^{-1}x_{k+1}^2u_k.\end{aligned} \tag{5.3.3}$$

If we now adjoin a new state $\bar{x}_k^1$ to x_k^1 as follows,

$$\hat{x}_{k+1}^1 = \begin{pmatrix} 0 & 1 & & & 0 \\ \vdots & & \ddots & & \\ \vdots & & & \ddots & \\ \vdots & 0 & & & 1 \\ 0 & -\alpha_{n_1} & \cdots & & -\alpha_1 \end{pmatrix}\hat{x}_k^1 + \begin{pmatrix} 0 \\ \vdots \\ \vdots \\ 0 \\ 1 \end{pmatrix}u_k$$

such that $\hat{x}_k^1 = \begin{pmatrix} \bar{x}_k^1 \\ x_k^1 \end{pmatrix}$ for $\bar{x}_k^1 \in \mathcal{R}^1$,

then $$x_k^1 = \begin{pmatrix} 1 & & & 0 \\ & \ddots & 0 & \vdots \\ & & \ddots & \vdots \\ 0 & & 1 & 0 \end{pmatrix}\hat{x}_{k+1}^1 \triangleq J_1\hat{x}_{k+1}^1$$

and $$\begin{aligned}u_k &= \hat{x}_{k+1,n_1+1}^1 + \alpha_{n_1}x_{k,1}^1 + \cdots + \alpha_1x_{k,n_1}^1 \\ &= [\alpha_{n_1}\cdots\cdots\alpha_1 \; 1]\hat{x}_{k+1}^1 \triangleq p_1^T\hat{x}_{k+1}^1.\end{aligned}$$

Hence

$$\bar{x}_k = (C''J_1 \otimes A_2^{-1} + Q_2''p^T \otimes A_2^{-1})\hat{x}_k^1 \otimes x_k^2 \triangleq \bar{C}\hat{x}_k^1 \otimes x_k^2.$$

We can then write the equations for $\hat{x}_k$ and y_k as

$$\begin{aligned}\hat{x}_{k+1} &= A'\hat{x}_k + (A''\bar{C} + C'J_2 \otimes I)\hat{x}_k^1 \otimes x_k^2 + Q_1'J_2\hat{x}_k^1 + Q_2'x_k^2u_k + b'u_kv_k \\ y_k &= H'\hat{x}_k + (H''\bar{C} + DJ_2 \otimes I)\hat{x}_k^1 \otimes x_k^2\end{aligned}$$

where $J_2 = \begin{pmatrix} 0 & \\ \vdots & I_{n_1} \\ 0 & \end{pmatrix}$.

In the case $\beta_{n_2} = 0$, from (5.2.1) we obtain the identities

$$C'' I \otimes (A_2^{n_2-1}+\ldots+\beta_{n_2-1} I) b_2 = 0 \tag{5.3.4}$$

$$\text{and } Q_2''(A_2^{n_2-1}+\ldots+\beta_{n_2-1} I) b_2 = 0. \tag{5.3.5}$$

For (A_2,b_2) in reachable canonical form, it is easy to show that

$$(A_2^{n_2-1}+\ldots+\beta_{n_2-1} I) b_2 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \in R^{n_2}$$

so that (5.3.4) and (5.3.5) imply that C'' and Q_2'' are of the form

$$Q_2'' = \begin{bmatrix} 0 & \\ \vdots & \times \\ 0 & \end{bmatrix} \qquad C'' = \begin{bmatrix} 0 \ldots 0 & \\ \vdots \quad \vdots & \times \\ 0 \ldots 0 & \end{bmatrix}.$$

This in turn implies that equation (5.3.2) is independent of $x_{k,1}^2$, and it is then possible to express v_k and $x_{k,2}^2,\ldots,x_{k,n_2}^2$ as linear functions of x_{k+1}^2, to obtain an equation analogous to (5.3.3). From this point on, the construction is the same as for $\beta_{n_2} \neq 0$.

<u>Case (2)</u> As we remarked above, this is treated in an analogous way to Case (1).

<u>Case (3)</u> In this case all we do is set up new states in x_k^1 and x_k^2 and get rid of states corresponding to zero eigenvalues in the x_k subsystem.

We adjoin the new states as follows:

$$\hat{x}_{k+1}^1 = \begin{pmatrix} 0 & 1 & & 0 \\ & & \ddots & \\ & 0 & & 1 \\ 0 & -\alpha_{n_1} & \ldots & -\alpha_1 \end{pmatrix} \hat{x}_k^1 + \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} u_k$$

$$\hat{x}_{k+1}^2 = \begin{pmatrix} 0 & 1 & & \\ & & \ddots & \\ & & & 1 \\ 0 & -\beta_{n_2} & \ldots & -\beta_1 \end{pmatrix} \hat{x}_k^2 + \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} v_k$$

so that $$x_k^1 = \begin{pmatrix} I_{n_1} & \begin{matrix} 0 \\ \vdots \\ 0 \end{matrix} \end{pmatrix} \hat{x}_{k+1}^1 \triangleq J_1 \hat{x}_{k+1}^1$$

$$= \begin{pmatrix} \begin{matrix} 0 \\ \vdots \\ 0 \end{matrix} & I_{n_1} \end{pmatrix} \hat{x}_k^1 \triangleq J_2 \hat{x}_k^1$$

and $$x_k^2 = \begin{pmatrix} I_{n_2} & \begin{matrix} 0 \\ \vdots \\ 0 \end{matrix} \end{pmatrix} \hat{x}_{k+1}^2 \triangleq K_1 \hat{x}_{k+1}^2$$

$$= \begin{pmatrix} \begin{matrix} 0 \\ \vdots \\ 0 \end{matrix} & I_{n_2} \end{pmatrix} \hat{x}_k^2 \triangleq K_2 \hat{x}_k^2$$

so that we obtain

$$\begin{aligned} x_{k+1} &= C'' x_k^1 \otimes x_k^2 + Q_1'' x_k^1 v_k + Q_2'' u_k + b'' u_k v_k \\ &= C'' J_1 \otimes J_2 \hat{x}_{k+1}^1 \otimes \hat{x}_{k+1}^2 + Q_1'' \hat{x}_{k+1}^1 (\hat{x}_{k+1,n_2+1}^2 + \beta_1 x_{k,n_2}^2 + \ldots + \beta_{n_2} x_{k,1}^2) \\ &\quad + Q_2'' \hat{x}_{k+1}^2 (\hat{x}_{k+1,n_1+1}^1 + \alpha_1 x_{k,n_1} + \ldots + \alpha_{n_1} x_{k,1}^1) \\ &\quad + b (\hat{x}_{k+1,n_1+1} + \alpha_1 x_{k,n_1}^1 + \ldots + \alpha_{n_1} x_{k,1}^1)(\hat{x}_{k+1,n_2+1}^2 + \ldots + \beta_{n_2} x_{k,1}^2) \\ &= (C'' J_1 \otimes J_2 + Q_1'' I \otimes p_2^T + Q_2'' p_1^T \otimes J + b p_1^T \otimes p_2^T) \\ &\qquad \text{where } p_2^T = [\beta_{n_2} \ldots \beta_1 \ 1] \\ &\triangleq \bar{C} \hat{x}_{k+1}^1 \otimes \hat{x}_{k+1}^2 . \end{aligned}$$

We can then write the equations for $\hat{x}_{k+1}$ and y_k as

$$\hat{x}_{k+1} = A' \hat{x}_k + (A'' \bar{C} + C' J_2 \otimes K_2) \hat{x}_k^1 \otimes \hat{x}_k^2 + Q_1' J_2 \hat{x}_k^1 v_k + Q_2' K_2 \hat{x}_k^2 u_k + b' u_k v_k$$

$$y_k = H' \hat{x}_k + (H'' \bar{C} + D J_2 \otimes K_2) \hat{x}_k^1 \otimes \hat{x}_k^2 .$$

Having gone through one of the reduction procedures (1), (2) or (3), we examine dim ker A', and if this is greater than or equal to 2, we repeat the above tests and if necessary reduce the system further. Since the dimension of the A matrix is constantly being reduced as a result of these procedures, after a finite number of iterations we must reach a stage when the whole system is quasi-reachable.

5.4 Quasi-Observability and Canonical Realizations

It has not proved possible as yet to provide a really good definition of observability or quasi-observability which will lead to a realization which is minimal in some sense and also canonical. The best that we have done so far is as follows:

Definition 5.4.1

A state-space realization is quasi-observable if the closure of the observable set is the whole space.

Referring back to the example (5.2) at the beginning of the chapter, we can readily see that provided $x_0^2 \in R^2$ is not equal to zero, it is possible to observe the value of $x_0^1 \in R^1$.

With the aid of the work done on observability conditions in Chapter 3, we can also formulate some ideas of observability for multi-output systems. Let us write a basis for the observability subspace of $\left([D\ H], \begin{bmatrix} A_1 \otimes A_2 & 0 \\ C & A \end{bmatrix} \right)$ as $\bar{H} = \begin{pmatrix} U & 0 \\ V & 0 \\ W & I \end{pmatrix}$

where $U \subset T_1 \otimes T_2$ where T_1, T_2 are the observability subspaces of $\left(\bar{H}\begin{bmatrix} A_1 \otimes b_2 \\ Q_1 \end{bmatrix}, A_1 \right)$ and $\left(\bar{H}\begin{bmatrix} b_1 \otimes A_2 \\ Q_2 \end{bmatrix}, A_2 \right)$ respectively and the rows of V are linearly independent of the rows of $T_1 \otimes T_2$.

If we were dealing with a single output system, then as we have shown in Chapter 4, we would have $\dim V \leq \dim \ker T_1 + \dim \ker T_2$. However this is no longer the case in general for multi-output systems, so we are led to the idea of a system being quasi-observable if (H,A) is an observable pair, and if $\dim V \geq \dim \ker T_1 + \dim \ker T_2$. We also assume in this case that all non-observable modes of x_k^1 and x_k^2 associated with non-zero eigenvalues have been eliminated, i.e. there exist no

$x_0^1 \in \ker T_1$ or $x_0^2 \in \ker T_2$ such that $V(x_0^1 \otimes I) = 0$ or $V(I \otimes x_0^2) = 0$. To eliminate any of these modes we follow the reduction procedure to observable realization of Chapter 4, followed by the conversion to quasi-reachable form as detailed in §5.2.

However this is still unsatisfactory with regard to obtaining canonical realizations which are related by a unique mapping. An example which has two realizations which are both quasi-reachable and quasi-observable illustrates this point:

$$s = \left[\frac{1}{z_1 z_2} \quad \frac{1}{z_1(z_2^2+az_2+b)} \quad \frac{z_2}{z_1(z_2^2+az_2+b)} \right]^T \qquad (b \neq 0)$$

Realization 1:
$$x_{k+1}^1 = u_k \qquad x_{k+1}^2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -b & -a \end{pmatrix} x_k^2 + \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} v_k$$

$$y_k = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} x_k^1 \otimes x_k^2$$

Realization 2:
$$x_{k+1}^1 = u_k \qquad x_{k+1}^2 = \begin{pmatrix} 0 & 1 \\ -b & -a \end{pmatrix} x_k^2 + \begin{pmatrix} 0 \\ 1 \end{pmatrix} v_k$$

$$x_{k+1} = u_k v_k$$
$$y_k = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} x_k + \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} x_k^1 \otimes x_k^2.$$

In fact, it has so far proved impossible to find a really satisfactory definition of canonical realizations for multi-output bilinear systems which ensures that they are minimal realizations too. The only way that minimality can be brought in directly is that a canonical realization is one with the smallest state dimension.

CHAPTER 6. MULTILINEAR SYSTEMS

This chapter opens with the study of a particular class of multilinear systems, which bear a great deal of resemblance to bilinear systems and concludes with a summary of some of the ideas contained in previous work on multilinear system theory.

The particular class of multilinear systems that is studied in §6.1 and §6.2 is based on the class of multilinear input/output maps whose denominators (in the case of n-linear maps) can be factorized as $p_1(z_1)\dots\dots p_n(z_n)p(z_1\dots\dots z_n)$. In §6.1, necessary and sufficient conditions are presented for a particular class of state space realizations of these maps to be quasi-reachable, and the proof is a natural extension of the proof for quasi-reachability of state space realizations of bilinear input/output maps. In §6.2, a stability result similar to that of §2.5 for bilinear maps is obtained for this class of multilinear maps, and this provides sufficient conditions for the boundedness of the output sequence due to a finite length input sequence.

In §6.3, multilinear input/output maps are characterized in a more formal way, analogously to the treatment of bilinear input/output maps in §2.1, and in addition some of the notions of the category-theory approach of [AAM1] to multilinear or multidecomposable systems are introduced. The main purpose of this approach is to provide the right sort of input and output spaces in which to work, but as yet it has not produced a theory of minimal realizations (except in the linear or decomposable case [AM1]).

6.1 Quasi-Reachability of a Class of Multilinear Systems

We shall consider the following specialized class of multilinear systems:

$$x^j_{k+1} = A_j x^j_k + b_j u^j_k$$

$$x_{k+1} = Ax_k + Cx^1_k \otimes \cdots \otimes x^r_k + \sum_{j=1}^{r} \sum_{i_j=0}^{1} Q_{i_1 \cdots i_r} v^1_{i_1} \otimes \cdots \otimes v^r_{i_r} + bu^1_k \cdots u^r_k \quad (6.1.1)$$

(i_j not all equal)

where $v^j_{i_j} = \begin{cases} x^j_k & \text{if } i_j = 0 \\ u^j_k & \text{if } i_j = 1 \end{cases}$

and dim ker A = 0 or 1.

This is specialized in the sense that the transfer function of the x_k state has denominator of the form $p_1(z_1)\ldots p_r(z_r)p(z_1\ldots z_r)$, with no polynomials in the denominator of the form $p_{12}(z_1z_2)$, $p_{345}(z_3z_4z_5)$.

In much the same way as we provided necessary and sufficient conditions for quasi-reachability of bilinear systems, we can express conditions for quasi-reachability of this multilinear system as follows:

Theorem 6.1.1

The system (6.1.1) is quasi-reachable iff the following conditions hold:

(i) $(A_1,b_1),\ldots,(A_r,b_r)$ are all controllable pairs.

(ii) $\begin{pmatrix} A_1 \otimes \cdots \otimes A_r & 0 \\ C & A \end{pmatrix}, \begin{pmatrix} A_1 \otimes b_2 \otimes \cdots \otimes b_r \cdots\cdots b_1 \otimes A_2 \otimes \cdots \otimes A_r & b_1 \otimes \cdots \otimes b_r \\ Q_{10\cdots 0} \cdots\cdots\cdots\cdots Q_{01\cdots 1} & b \end{pmatrix}$

$\triangleq$ (F,G) is a reachable pair.

Proof: Clearly these conditions are necessary for quasi-reachability, since if any $(A_i.b_i)$ were not a controllable pair then the substate x^i_k would not be reachable, and if (F,G) were not a controllable pair, then there would exist vectors $p \in R^{n_1\cdots n_r}$, $q \in R^n$ such that

$$[p^T \; q^T]F = \lambda[p^T \; q^T] \text{ for some } \lambda \in C \text{ and } [p^T \; q^T]G = 0$$

and then we would have

$$p^T x^1_{k+1} \otimes \cdots \otimes x^r_{k+1} + q^T x_{k+1} = \lambda(p^T x^1_k \otimes \cdots \otimes x^r_k + q^T x_k).$$

To show sufficiency, we first note that the components of the transfer functions

$x^1(z_1)\otimes\ldots\otimes x^r(z_r)$ and $x(z_1,\ldots,z_r)$ are linearly independent. This follows from a similar argument to that of Lemma 3.2.1.

Following the now established procedure of Chapter 3, given any desired $x_1^1,\ldots,x_1^r$, we can construct input sequences $p_1(z_1)\psi_1(z_1)+q_1(z_1),\ldots,p_r(z_r)\psi_r(z_r)+q_r(z_r)$ which reach these substates, for unique $q_i(z_i)$ where $\deg q_i < \deg \psi_i$, and for any $p_i(z_i) \in R[z_i]$, where $\psi_i(z)$ is the minimal polynomial of A_i.

Now with the aid of a similarity transformation on x_k, we can write A as $A = \text{diag}\,(J_1,J_0)$, where J_1 and J_0 are square matrices with $\ker J_1 = \phi$, and

$$J_0 = \begin{pmatrix} 0 & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & 0 \end{pmatrix} \in R^{m\times m}.$$

It is now clear that if x_k is partitioned as $[\hat{x}_k^T\ \bar{x}_k^T]$ with $\hat{x}_k$ and $\bar{x}_k$ associated with J_1, J_0 respectively, then only the first r terms of $p_i(z_i)$ $(i=1,\ldots,r) \triangleq \alpha_i(z_i)$ will affect the substate $\bar{x}_1$, and in a similar way to that for bilinear systems, we find that $\bar{x}_1$ can almost always be reached. To be more precise, any given value of $\bar{x}_1$ can be reached provided that the coefficients of $q_i(z_i)$ or equivalently the elements of x_1^i $(i=1,\ldots,r)$ do not lie on a certain finite union of hyperplanes.

We can now write the transfer function for $\hat{x}(z_1,\ldots,z_r) \in R^{n-m}[(z_1,\ldots,z_r)]$ as

$$\hat{x}(z_1,\ldots,z_r) = \frac{s(z_1,\ldots,z_r)(z_1\cdot\ldots\cdot z_r)^m}{\phi(z_1\ldots z_r)\bar{\psi}_1(z_1)\cdot\ldots\cdot\bar{\psi}_r(z_r)}$$

where $\bar{\psi}_i(z_i) = z_i^m\psi_i(z_i)$

and our problem is now to construct input sequences of the form $p_i(z_i)\bar{\psi}_i(z_i) + \bar{q}_i(z_i)$ which enable us to reach $\bar{x}_1$ (where $\bar{q}_i(z_i) = \alpha_i(z_i)\psi_i(z_i) + q_i(z_i)$).

We shall now let $p_i(z_i) = z_i^{m_i}\bar{p}_i(z_i)$, for some $\bar{p}_i$ to be determined later, where m_i $(i=1,\ldots,r)$ is given by

$m_i = \max_k \{j_k - j_i \;\; \varepsilon_{j_1\ldots j_r} \neq 0, \varepsilon_{j_1\ldots j_r} z_1^{j_1}\ldots z_r^{j_r}$ occurs in one of $s_1,\ldots,s_{n-m}\}$.

It is then clear that

$$\hat{x}(z_1,\ldots,z_r)w_1(z_1)\ldots w_r(z_r) \odot \sum_{k\geq 0}(z_1\ldots z_r)^{-k} = 0$$

where $w_i(z_i) = p_i(z_i)\bar{\psi}_i(z_i)$ or $q_i(z_i)$ except in the case where $w_i(z_i) = q_i(z_i)$ for all i or else $w_i(z_i) = p_i(z_i)\bar{\psi}_i(z_i)$ for all i.

If we can now show that

$$\hat{x}(z_1,\ldots,z_r)\prod_{i=1}^{r} p_i(z_i)\bar{\psi}_i(z_i) \odot (z_1\ldots z_r)^{-1}$$

can attain any value of $x_1 \in R^{n-m}$, then the proof is complete.

We first write the numerator of x as

$$(z_1\ldots z_r)^m z_1^{m_1}\ldots z_r^{m_r}S = N(z_1,\ldots,z_r) + \phi(z_1\ldots z_r)M(z_1,\ldots,z_r)$$

where N contains no term with a factor $(z_1\ldots z_r)^{n-m}$. By a similar argument to that of §3.1, we find that the components of N are linearly independent, and that

$$\hat{x}p_1(z_1)\bar{\psi}_1\ldots p_r(z_r)\bar{\psi}_r \odot \sum(z_1\ldots z_r)^{-k}$$
$$= \frac{(z_1\ldots z_r)^m z_1^{m_1}\ldots z_r^{m_r}S}{\phi(z\ldots z_r)}\bar{p}_1(z_1)\ldots\bar{p}_r(z_r) \odot \sum(z_1\ldots z_r)^{-k}$$
$$= \frac{N(z_1,\ldots,z_r)}{\phi(z_1\ldots z_r)}\bar{p}_1(z_1)\ldots\bar{p}_r(z_r) \odot \sum(z_1\ldots z_r)^{-k}.$$

We shall choose the $\bar{p}_i$ in such a way that $\bar{p}_1(z_1),\ldots,\bar{p}_{r-1}(z_{r-1})$ have coefficients either 0 or 1, while the coefficients of $\bar{p}_r(z_r)$ are chosen to solve a set of linear equations.

The terms of $N(z_1,\ldots,z_r)$ can be written as members of the sets

$$A = \{z_1^{i_1}\ldots z_{r-1}^{i_{r-1}}(z_1\ldots z_{r-1})^k(z_1\ldots z_r)^\ell\}$$

and $B = \{z_1^{i_1}\ldots z_{r-1}^{i_{r-1}}z_r^k(z_1\ldots z_r)^\ell\}$

where $i_1,\ldots,i_{r-1}$ range from zero to $\ell_1,\ldots,\ell_{r-1}$ respectively with the proviso that at least one of $i_1,\ldots,i_{r-1}$ is zero, k ranges from zero to K and ℓ ranges from zero to $\bar{n}-1$ (where $\bar{n} = n-m$).

The input sequences will be structured in a way that is very similar to that for bilinear systems. First of all we shall divide the input sequences into $\bar{n}$ sections, and as with bilinear systems we shall choose an integer N greater than an integer M to be specified shortly such that (A^N, b) is a reachable pair, where (c^T, A, b) is a minimal realization of $1/\phi(z)$. The inputs $\bar{p}_i(z_i)$ will be of the form

$$\bar{p}_i(z_i) = u_i^0(z_i) + z_i^N u_i^1(z_i) + \dots + z_i^{N(\bar{n}-1)} u_i^{\bar{n}-1}(z_i) \quad (i = 1,\dots,r)$$

$$\text{where } \deg u_i^j(z_i) \le M.$$

All the $u_i^j(z_i)$ for each i will have the same structure, and each of these will be divided up into sections each of length $\ge 2K+1$. Each of these sections will be characterized by $z_1^{i_1} \dots z_{r-1}^{i_{r-1}}$; let $I_1 = \max(i_1,\dots,i_{r-1})$ for the first choice of $i_1,\dots,i_{r-1}$. Then the inputs corresponding to this which we shall choose will be $z_1^{K+I_1-i_1},\dots,z_{r-1}^{K+I_1-i_{r-1}}$ and $z_r^{I_1}(\alpha_{-K} + z_r\alpha_{-K+1} + \dots + z_r^K\alpha_0 + \dots + z_r^{2K}\alpha_k)$.

If this were the only set of inputs to the system, it is clear that the only effects would be on the transfer functions

$$\frac{z_1^{i_1}\dots z_{r-1}^{i_{r-1}}}{\phi(z_1\dots z_r)} \begin{pmatrix} z_r^K \\ \vdots \\ z_r \\ 1 \\ (z_1\dots z_{r-1}) \\ \vdots \\ (z_1\dots z_{r-1})^K \end{pmatrix} (z_1\dots z_r)^{\ell} \qquad (\ell = 0,\dots,\bar{n}-1)$$

and the outputs of these will then be equal to

$$\frac{z^{K+I_1+\ell}}{\phi(z)} \begin{pmatrix} \alpha_{-K} \\ \vdots \\ -\alpha_{-1} \\ \alpha_0 \\ \alpha_1 z \\ \vdots \\ \alpha_K z^K \end{pmatrix} \qquad (\ell = 0,\dots,\bar{n}-1)$$

The next section of the input sequence will be chosen in a similar fashion, except that we have to multiply it by $z_i^{I_1+2K+1}$ $(i = 1,\ldots,r)$ and we continue in this way until all the $u_i^0(z_i)$ $(i = 1,\ldots,r)$ are completely characterized, although the values $\alpha_{-K},\ldots,\alpha_K$ etc. are yet to be chosen. It is easy to check that the various sections of the input sequences do not interact with one another through the transfer functions.

Before completing this proof we give an example of what the inputs look like for the case $r = 3$.

Suppose the numerator terms are grouped together as $\{z_3,1,z_1z_2\}$, $\{z_1z_3,z_1,z_1^2z_2\}$, $\{z_2z_3,z_2,z_1z_2^2\}$. The input sequences will be as follows, in ascending powers of z_i:

u_1	0 1 0	1 0 0	0 1 0
u_2	0 1 0	0 1 0	1 0 0
u_3	a_1 a_2 a_3	a_4 a_5 a_6	a_7 a_8 a_9

The outputs will then be given by $(\{a_1z,a_2z,a_3z^2\},\{a_4z^4,a_5z^4,a_6z^5\},$ $\{a_7z^7,a_8z^7,a_9z^8\})/\phi(z)$.

To return to the proof, we now choose $N > M = I_1 + \ldots + I_s + s(2K+1)$, where s is the number of different combinations of $\{i_1,\ldots,i_r\}$, in such a way that N satisfies the conditions of Lemma 3.2.3. It is then clear that the various unknowns, α_i, etc., can be chosen to ensure that the whole vector transfer function is 'reachable', in the sense that any output at time + 1 can be attained. This follows from exactly the same arguments as those in Theorem 3.2.1.

Hence we can construct an input sequence which reaches any state which does not lie on a certain finite union of hyperplanes, so that the system (5.1.1) is quasi-reachable. □

Note that the A-matrix was not required to be cyclic other than with respect to its zero eigenvalues.

6.2 Input/Output Stability of a Class of Multilinear Systems

Closely associated with the state space realization of §6.1 is the transfer function

$$s = N(z_1,\ldots,z_r)/p_1(z_1)\ldots p_r(z_r)p(z_1\ldots z_r) \qquad (6.2.1)$$

which, using analogous arguments to those of §2.5 for the special case of bilinear systems, can be realized in the form of (6.1.1) with the observation

$$y_k = d^T x_k^1 \otimes \ldots \otimes x_k^r + h^T x_k$$

and although this realization may not in general be observable it can always be reduced to quasi-reachable form, using similar methods to those of §4.2.

We shall now produce a similar result to that of Theorem 2.3 concerning stability of the output sequence of (6.2.1) due to a finite length input sequence. Let us factorize the numerator of s as

$$N(z_1,\ldots,z_r) = M(z_1,\ldots,z_r)f(z_1\ldots z_r).$$

Depending on whether $\deg_{z_i} M \geq \deg p_i$ or $\deg f > \deg p$, we multiply top and bottom of (6.2.1) by $(z_1\ldots z_r)^s$ to obtain (in analogy with §2.5)

$$s = \frac{R(z_1,\ldots,z_r)}{q_1(z_1)\ldots q_r(z_r)} \times \frac{g(z_1\ldots z_r)}{q(z_1\ldots z_r)}$$

where $\deg z_i R < \deg q_i \quad (i = 1,\ldots,r)$

$\deg g \leq \deg q.$

With the preliminaries over, we can now state the following sufficient conditions for output stability.

Theorem 6.2.1

If either of the following conditions hold, then the output sequence due to a finite length input sequence from $U_1 \times \ldots \times U_r$ is ℓ_1-stable:

(i) all zeros of $p(z)$ and all terms of the form $\{\alpha_{i_1}^1 \ldots \alpha_{i_r}^r\}$ (where $\{\alpha_{i_1}^1\},\ldots,\{\alpha_{i_r}^r\}$ are zeros of $p_1(z),\ldots,p_r(z)$ respectively) lie within the unit circle;

(ii) all zeros of $p(z)$ lie within the unit circle and all terms $\{\alpha^1_{i_1} \dots \alpha^r_{i_r}\}$ not lying within the unit circle are zeros of $f(z)$.

Proof: Using the representation lemma for recognizable sequences, Lemma 2.2.2, it is easy to see that an input of the form $z_1^{i_1} \dots z_r^{i_r}$ will produce an output sequence given by

$$y(z) = c^T(zI-A_1 \otimes \dots \otimes A_r)^{-1} A_1^{i_1-I} b_1 \otimes \dots \otimes A_r^{i_r-I} b_r \frac{z^I g(z)}{q(z)} \odot \sum_{k\geq 1} z^{-r} \qquad (6.2.2)$$

where $I = \min(i_1, \dots, i_r)$.

It is now immediately obvious that because (1) the eigenvalues of A_i $(i = 1,\dots,r)$ are the same as the zeros of $p_i(z_i)$ with perhaps the addition of a few zero eigenvalues, (2) the zeros of $q(z)$ are the same as those of $p(z)$ with the possible addition of a few zeros, the conditions (i) or (ii) are sufficient for the output sequence given by $y(z)$ to be ℓ_1-stable.

Finally, because a finite input sequence leads to the addition of only a finite number of terms of form (6.2.2), the theorem is proven. □

6.3 Characterization of Multilinear Systems

We shall define the input space and output spaces in a similar way to that of Chapter 2, where in this case we deal with $m \geq 2$ input channels.

$$U_i = \{u \in R^{Z-} \text{ with compact support}\} \quad i = 1,\dots,m$$
$$Y = \{u \in R^{N-\{0\}}\}.$$

Then we say that a map $f : U_1 \times \dots \times U_m$ is a multi-linear discrete-time input/output map if it satisfies the conditions:

(i) Multilinearity:

For all fixed $u_i \in U_i$ for $i \neq j$, the map

$$f(u_1,\dots,u_{j-1}, \cdot, u_{j+1},\dots,u_m) : U_j \to Y$$

is linear $(j = 1,\dots,m)$.

(ii) Stationarity:

$$f(\sigma u_1,\ldots,\sigma u_m) = \sigma^* f(u_1,\ldots,u_m)$$

where σ and σ^* are shift operators.

As with bilinear systems, we can easily show that after setting up the isomorphisms $U_i \cong R[z_i]$, $Y \cong R[[z^{-1}]]$, the map f will then be isomorphic to the causal power series $s = (z_1\ldots z_m)^{-1}\sum s_{i_1\ldots i_m} z_1^{-i_1}\ldots z_m^{-i_m}$, with output $y \in R[[(z_1\ldots z_m)^{-1}]]$ given by

$$y = s\, u_1(z_1)\ldots u_m(z_m) \odot \sum_{k\geq 1} (z_1\ldots z_m)^{-k}$$

for inputs $u_i(z_i) \in R[z_i]$.

Again, in analogy with bilinear systems, we can define a series of equivalence relations which, when taken together, are equivalent to Nerode equivalence. An example of this is given by the case $m = 3$; we define the equivalence relations $\tilde{1}$, $\tilde{2}$, $\tilde{3}$, $\widetilde{12}$, $\widetilde{13}$, $\widetilde{23}$, $\widetilde{123}$, as follows (where, for convenience, we consider input spaces U, V and W):

$$u_1 \underset{1}{\sim} u_2 \quad \text{iff } f(z_1^k u_1,\phi,\psi) = f(z_1^k u_2,\phi,\psi)$$

for all k and for all ϕ,ψ with deg $\phi,\psi < k$

with similar definitions for $\underset{2}{\sim}$ and $\underset{3}{\sim}$.

$$(v_1,w_1) \underset{23}{\sim} (v_2,w_2) \text{ iff } f(\theta,z_2^k v_1,z_3^k w_1) = f(\theta,z_2^k v_2,z_3^k w_2)$$

for all k and for all θ with deg θ < 1·

with similar definitions for $\underset{12}{\sim}$ and $\underset{13}{\sim}$.

$$(u_1,v_1,w_1) \underset{123}{\sim} (u_2,v_2,w_2) \text{ iff } f(u_1,v_1,w_1) = f(u_2,v_2,w_2).$$

It is then easy to show that $(u_1,v_1,w_1) \underset{N}{\sim} (u_2,v_2,w_2)$ iff $u_1 \underset{1}{\sim} u_2$, $v_1 \underset{2}{\sim} v_2$, $w_1 \underset{3}{\sim} w_2$, $(u_1,v_1) \underset{12}{\sim} (u_2,v_2)$, $(u_1,w_1) \underset{13}{\sim} (u_2,w_2)$, $(v_1,w_1) \underset{23}{\sim} (v_2,w_2)$ and $(u_1,v_1,w_1) \underset{123}{\sim} (u_2,v_2,w_2)$.

Further, by analogy with bilinear systems, it is clear that the quotient spaces $X_1 = U/\underset{1}{\sim}$, $X_2 = V/\underset{2}{\sim}$, and $X_3 = W/\underset{3}{\sim}$ may be endowed with the structure of a linear space. Then, by embedding U×V in the tensor space U⊗V, we can show that there exists a linear map $f_{1\otimes 2}$ inducing an equivalence relation in $U\times V$, and then $(u_1,v_1) \underset{12}{\sim} (u_2,v_2)$ iff $u_1 \otimes v_1 \equiv u_2 \otimes v_2$ (mod $f_{1\otimes 2}$). We can do the same with $U\times W$ and $V\times W$ to obtain linear spaces $X_{12} = U\otimes V/\ker f_{1\otimes 2}$, $X_{13} = U\otimes W/\ker f_{1\otimes 3}$ and $X_{23} = V\otimes W/\ker f_{2\otimes 3}$ into which we can naturally embed the equivalence classes under $\underset{12}{\sim}$, $\underset{13}{\sim}$ and $\underset{23}{\sim}$ respectively. Finally by embedding $U\times V\times W$ in $U\otimes V\otimes W$, we can show that there exists a linear map $f_{\otimes}$ inducing an equivalence relation in $U\times V\times W$ and then $(u_1,v_1,w_1) \underset{123}{\sim} (u_2,v_2,w_2)$ iff $u_1\otimes v_1\otimes w_1 = u_2\otimes v_2\otimes w_2$ (mod $f_{\otimes}$). The equivalence classes under $\underset{123}{\sim}$ are then naturally embedded in the linear space $X_{123} = U\otimes V\otimes W/\ker f_{\otimes}$.

X_i $(i=1,2,3)$ will then be an $R[z_i]$-module, X_{ij} $(j>i)$ will be an $R[z_iz_j]$-module and X_{123} will be an $R[z_1z_2z_3]$-module. Using Zeiger's Lemma, we will be able to set up a state space realization as follows:

$$
\begin{aligned}
x^i_{k+1} &= A_i x^i_k + b_i u^i_k \\
x^{ij}_{k+1} &= A_{ij} x^{ij}_k + Q^{ij}_1 x^i_k u^j_k + Q^{ij}_2 x^j_k u^i_k + b_{ij} u^i_k u^j_k \quad (j>i) \\
x_{k+1} &= Ax_k + Q_1 x^{23}_k u^1_k + Q_2 x^{13}_k u^2_k + Q_3 x^{12}_k u^3_k + b u^1_k u^2_k u^3_k \\
y_k &= Hx_k .
\end{aligned}
\tag{6.3.1}
$$

In general, for $m \geq 1$, we can characterize this type of realization as follows, and we quote directly from [AAM1], where it is described as an m-line system:

<u>Definition 6.3.1</u>

An m-line system M with input objects $U_1,\ldots,U_m$ and output object Y is defined by induction on m as follows:

For $k=1$: M is a linear system $M = (X,F,U_1,G,Y,H)$

$G : U_1 \to X$, $F : X \to X$, $H : X \to Y$.

For $k>1$: M is specified by

(i) a state transition map $F : X\to X$ and output map $H : X\to Y$;

(ii) for each proper non-empty subset α of $\{1,\ldots,k\}$ an $|\alpha|$ line system M_α with input objects $\{V_i : i\in\alpha\}$ and output object Y_α;

(iii) for each proper non-empty α, a morphism $J_\alpha : Y_\alpha \otimes I^{\bar{\alpha}} \to X$ where $I^{\bar{\alpha}} = \otimes\{I_j \mid j \notin \alpha\}$;

(iv) a morphism $J_\phi : U_1 \otimes \ldots \otimes u_m \to X$.

As we can see, this definition agrees with the state space realizations (2.3.5) and (6.3.1) obtained for bilinear and trilinear systems. However, as we have observed with bilinear systems, this realization will in general not be reachable or even quasi-reachable. The conditions for observability will be the straightforward ones we had for bilinear systems; in the case of (6.3.1) these will be (H,A), (Q_1,A_{23}), (Q_2,A_{13}), (Q_3,A_{12}), $([Q_1^{12}Q_1^{12}],A_1)$, $([Q_1^{23}Q_2^{12}],A_2)$, $([Q_2^{13}Q_2^{23}],A_3)$ must all be observable pairs.

As has been mentioned by Kalman [K1], the state space as defined by Definition 6.3.1, will lie on some algebraic variety, and it will therefore be possible to reduce the system in such a way as to include multiplication of states. However it has not yet proved possible to demonstrate necessary and sufficient conditions for quasi-reachability. We would again expect, as with bilinear state space descriptions, to have quasi-reachability if the various tensor products of transfer functions are linearly independent, so for instance in the trilinear case we will require the components of the following set of vectors to be linearly independent:

(i) $x^i(z_i)$; $i = 1,2,3$

(ii) $x^i(z_i) \otimes x^j(z_j)$, $x^{ij}(z_i,z_j)$; $j > i$

(iii) $x^1(z_1) \otimes x^2(z_2) \otimes x^3(z_3)$, $x^1(z_1) \otimes x^{23}(z_2,z_3)$, $x^2(z_2) \otimes x^{13}(z_1,z_2)$, $x^3(z_3) \otimes x^{12}(z_1,z_2)$, $x(z_1,z_2,z_3)$,

and these will clearly be necessary conditions for quasi-reachability.

Finally, it is perhaps necessary to comment on the use of category theory in the analysis of multilinear systems. In their first paper on decomposable systems [AM1], Arbib and Manes demonstrated how to set up canonical realizations of systems of the form

$$x_{k+1} = a(x_k) \circ b(u_k) \qquad u_k \in U$$
$$y_k = c(x_k) \qquad y_k \in Y$$

where $\circ$ indicates some operation particular to the system, e.g. addition for modules, or multiplication for groups (when a, b, and c would be homomorphisms).

In their paper they showed that the input spaces $U^{\S}$ suitable for analysing this system would have to be in the same category as U, in fact a countable copower of it, so that if U is a module, then $U^{\S}$ would have to be a module, and if U were a group, then in the same way $U^{\S}$ would have to be a group. Analogously, they showed that the output space $Y_{\S}$ would have to be in the same category as Y, in fact a countable power of it.

However, when it comes to multidecomposable systems, any algebraic entity which is not at least a ring seems to be unsuited to this form of analysis, although Abelian groups might possibly fit into the scheme better.

More suitable is the category theory approach adopted by Goguen [G1], who treats discrete-time machines in closed monoidal categories. Rather than present various definitions concerning categories, we shall outline the application of his work to affine maps, and we comment that this might provide an extension of the results of multilinear maps to those of multiaffine maps.

Goguen began within the framework of a particular category $\underline{C}$, e.g. the category of groups, or of vector spaces, or of affine spaces, or of sets, together with the various mappings (e.g. in the above cases we would be considering homomorphisms, linear maps, affine maps, set maps) within which there existed a monoidal structure (defined by $\otimes$), i.e. if $A,B \in \underline{C}$, then $A\otimes B \in \underline{C}$, $(A\otimes B)\otimes C \cong A\otimes(B\otimes C)$. He then assumed that this monoidal category was closed, i.e. for all mappings $f : A \boxtimes B \to C$ (for $A,B,C \in \underline{C}$), there existed an entity $[B,C] \in \underline{C}$, such that there is a natural isomorphism between $f : A\boxtimes B \to C$ and $f' : A \to [B,C]$.

For the case $\underline{C}$ = category of sets, $\boxtimes$ is the Cartesian product, and if we have

$$f : A\times B \to C : (a,b) \to f(a,b)$$

then f' is given by

$$f' : A \to [B,C] : a \to f(a,\ )$$

and it is clear that [B,C] is the collection of set mappings from B to C, which is itself a set.

Note that for $\underline{C}$ = category of vector spaces with $\boxtimes$ the Cartesian product, there exists no such suitable entity, since we have

$$f : A\times B \to C : (a,b) \quad f_1(a) + f_2(b)$$

and if we define f' by

$$f' : A \to [B,C] : a \to f_1(a) + f_2(\ )$$

then this is clearly not a linear space, so $[B,C] \notin \underline{C}$.

For vector spaces it is easy to see that $\boxtimes$ = tensor product produces a closed category.

In the case that interests us here, we consider $\underline{C}$ = category of affine spaces with Ⓐ = affine tensor product, i.e. Ⓐ $: A\times B \to A\boxtimes B + A + B$, where $\boxtimes$ is now defined as the usual tensor product.

Then the affine map f applied to $A\boxtimes B$ is given by

$$\begin{aligned} f &: A\boxtimes B \to C \\ &: (a,b) \to f_1(a\boxtimes b) + f_2(a) + f_3(b) + c \end{aligned} \qquad (6.3.2)$$

(where f_1, f_2 and f_3 are linear).

It follows that

$$\begin{aligned} f' &: A \to [B,C] \\ &: a \to f_1(a\boxtimes\cdot) + f_2(a) + f_3(\cdot) + c \end{aligned}$$

is affine, so [B,C] is an affine space, and hence is in the same category as A, B and C.

Goguen then showed that given an input/output function $f: U^* \to Y$, where $U^* = \bigcup_{k\geq 1} U \otimes \ldots \otimes U$, the countable copower of U, a suitable state space is given by $[U^*,Y]$, and a minimal realization will be provided by the reachable set of $[U^*,Y]$. In the case of linear systems we have f_1 and c (as in (6.3.2)) both zero, but otherwise we have a well-defined input/output map. If we write the impulse response of this linear system as $(s_1,s_2,s_3,\ldots.)$, then the elements of $[U^*,Y]$ which will be of interest to us will be

$$g_1 \triangleq s_1(\ ) + s_1(\ ) + s_1(\ ) + \ldots : U+ \to Y$$

$$g_2 \triangleq s_1a_1 + s_2(\ ) + s_3(\ ) + s_4(\ ) + \ldots : U^* \to Y$$

$$g_3 \triangleq s_1a_1 + s_2a_2 + s_3(\ ) + s_4(\ ) + s_5(\ ) + \ldots : U^* \to Y,$$

etc. Clearly then, if the Hankel matrix formed from $(s_1,s_2,s_3,\ldots.)$ has dimension n, then the number of linearly independent g_i will be equal to $n+1$, so that the dimension of the affine state space will equal $n+1$.

In the case of f_1 and c not equal to zero, a state space description (assuming one exists) would be of the form

$$\begin{aligned} x_{k+1} &= Ax_k + u_kFx_k + bu_k \\ y_k &= Cx_k \end{aligned} \tag{6.3.3}$$

which is termed an affine system in [G1]. (Note that Isidori [I1] and others refer to (6.3.3) as a bilinear system, since the R.H.S. of the transition equation of (6.3.3) is linear in each of u_k and x_k separately.)

A possibility now is to extend this approach of Goguen to biaffine and multiaffine systems, which would provide greater generality than multilinear systems, and would probably be more relevant than the multi-decomposable approach of [AAM1].

CHAPTER 7. CONCLUSION

The main accomplishment of this thesis has been to give a thorough account of the theory of state space realizations of bilinear input/output maps, providing the solution to a number of previously unsolved problems. The contributions to this realization theory have involved a formalization in Chapter 2 of the ideas of Kalman [Kl] regarding the actual setting up of a state space realization directly from the transfer function, thus by-passing the elaborate constructions of Fornasini and Marchesini [FM1]; the derivation of necessary and sufficient conditions in Chapter 3 for a state space realization to be observable and quasi-reachable; reduction procedures for obtaining canonical realizations from realizations which are not observable or quasi-reachable, and furthermore, in Chapter 4, an isomorphism theorem showing that any two such canonical realizations are isomorphic under a well-defined class of transformations.

Quasi-reachability results have also been obtained in Chapter 5 for the case of multi-output bilinear systems, and the concept of quasi-observability was introduced to cover the cases when observability was too strong a requirement. However it was not possible to obtain such definitive results as for the single output case studied in Chapters 3 and 4, and in particular no isomorphism theorem has been obtained for minimal realizations.

Sufficiency conditions were obtained in Chapter 2 on the transfer function of a bilinear input/output map which ensure that the output sequence from this map, due to a finite length input sequence, tends to zero. In Chapter 6, analogous conditions on a particular class of multi-linear transfer functions were obtained, assuring a similar stability result for the corresponding input/output map. Sufficient conditions were also derived in Chapter 6 for a particular form of state space realization of this class of multilinear maps to be quasi-reachable, these conditions again being analogous to those of Chapter 3 on bilinear state space realizations. In addition it was shown in Chapter 6 how to obtain necessary conditions for quasi-reachability for general multilinear state space realizations, but the question of sufficiency still remains open.

Apart from this investigation of conditions for quasi-reachability and observability of realizations of discrete-time multilinear input/output maps, an obvious area for future work on multilinear system theory is the realization of continuous-time input/output maps. In the bilinear case, such a realization may be written as

$$\dot{x}_1 = A_1x_1 + b_1u \tag{7.1}$$

$$\dot{x}_2 = A_2x_2 + b_2v \tag{7.2}$$

$$\dot{x} = Ax + Cx_1 \otimes x_2 + Q_1x_1v + Q_2x_2u + buv \tag{7.3}$$

$$y = h^Tx + d^Tx_1 \otimes x_2. \tag{7.4}$$

Using the intuitive approach of Chapter 3, we expect that this representation will not be reachable if there exist vectors p and q such that $p^Tx_1 \otimes x_2 + q^Tx$ evolves independently of u and v. In particular we expect that

$$\frac{d}{dt}(p^Tx_1 \otimes x_2 + q^Tx) = \lambda(p^Tx_1 \otimes x_2 + q^Tx) \quad \text{for some } \lambda \in \mathcal{C}.$$

On expansion of the left-hand side of this expression we find that this property is equivalent to the pair

$$\left(\begin{bmatrix} A_1 \otimes I + I \otimes A_2 & O \\ C & A \end{bmatrix}, \begin{bmatrix} b_1 \otimes I & I \otimes b_2 & O \\ Q_1 & Q_2 & b \end{bmatrix} \right) \tag{7.5}$$

not being reachable. Although this condition is sufficient for non-reachability of (7.1)-(7.3) it has not yet been shown to be necessary. However we can still obtain some information from (7.5) concerning similarity transformations, namely that the similarity transformation $\begin{bmatrix} I & O \\ W & I \end{bmatrix}$

applied to the pair (7.5) and $[d^T h^T]$ yields a system which is equivalent to that of (7.1)-(7.4). We state this result formally, as follows:

Theorem 7.1

Let (7.1)-(7.4) be a realization of a continuous-time bilinear input/output map $f: U\times V \to Y$, where $x_1(t) \in R^{n_1}$, $x_2(t) \in R^{n_2}$, $x(t) \in R^n$. Then for any $W \in R^{n\times n_1 n_2}$, (7.1)-(7.4) is also a realization of f under the transformation

$$\begin{aligned} C &\to W(A_1 \otimes I + I \otimes A_2) + C - AW \\ Q_1 &\to Q_1 + W(I\otimes b_2) \\ Q_2 &\to Q_2 + W(b_1 \otimes I) \\ d^T &\to d^T - h^T W. \end{aligned} \quad (7.6)$$

Proof: From (7.1)-(7.4) we can immediately write down the expression for y(t) as

$$\begin{aligned} y(t) = h^T[\int_0^t e^{A(t-\tau)}\{C\int_0^t e^{A_1(\tau-\tau_1)} b_1 u(\tau_1)d\tau_1 \otimes \int_0^t e^{A_2(\tau-\tau_2)} b_2 v(\tau_2)d\tau_2 \\ + Q_1\int_0^t e^{A_1(\tau-\tau_1)} b_1 u(\tau_1)d\tau_1 v(\tau) + Q_2\int_0^t e^{A_2(\tau-\tau_2)} b_2 v(\tau_2)d\tau_2 u(\tau) \\ + bu(\tau)v(\tau)\}d\tau] \\ + d^T[\int_0^t e^{A_1(\tau-\tau_1)} b_1 u(\tau_1)d\tau_1 \otimes \int_0^t e^{A_2(\tau-\tau_2)} b_2 v(\tau_2)d\tau_2] \end{aligned}$$

Now by inspection we can see that the difference $\tilde{y}(t)$ between this value and that of the output of the transformed system is equal to

$$\begin{aligned} \tilde{y}(t) = h^T[\int_0^t e^{A(t-\tau)} W\{\int_0^\tau A_2 e^{A_1(\tau-\tau_1)} b_1 u(\tau_1)d\tau_1 \otimes \int_0^\tau e^{A_2(\tau-\tau_2)} b_2 v(\tau_2)d\tau_2 \\ + \int_0^\tau e^{A_1(\tau-\tau_1)} b_1 u(\tau_1)d\tau_1 \otimes \int_0^\tau A_2 e^{A_2(\tau-\tau_2)} b_2 v(\tau_2)d\tau_2 \\ + \int_0^\tau e^{A_1(\tau-\tau_1)} b_1 u(\tau_1)d\tau_1 \otimes b_2 v(\tau) + b_1 u(\tau)\otimes \int_0^\tau e^{A_2(\tau-\tau_2)} b_2 v(\tau_2)d\tau_2\}d\tau] \\ - *h^T\int_0^t e^{A(t-\tau)} AW\{\int_0^\tau e^{A_1(\tau-\tau_1)} b_1 u(\tau_1)d\tau_1 \otimes \int_0^\tau e^{A_2(\tau-\tau_2)} b_2 v(\tau_2)d\tau_2\}d\tau* \\ - h^T W\int_0^t e^{A_1(t-\tau_1)} b_1 u(\tau_1)d\tau_1 \otimes \int_0^t e^{A_2(t-\tau_2)} b_2 v(\tau_2)d\tau_2. \end{aligned}$$

If we now integrate by parts the term enclosed by asterisks (*), and use the fact that $\frac{d}{d\tau} e^{A(t-\tau)} = -e^{A(t-\tau)}A$, we obtain

$$h^T[-e^{A(t-\tau)}W\{\int_0^\tau e^{A_1(\tau-\tau_1)}b_1u(\tau_1)d\tau_1 \otimes \int_0^\tau e^{A_2(\tau-\tau_2)}b_2v(\tau_2)d\tau_2\}]_{\tau=0}^{\tau=t}$$

$$+ h^T\int_0^\tau e^{A(t-\tau)}W\frac{d}{d\tau}\{\int_0^\tau e^{A_1(\tau-\tau_1)}b_1\dot{u}(\tau_1)d\tau_1 \otimes \int_0^\tau e^{A_2(\tau-\tau_2)}b_2v(\tau_2)d\tau_2\}d\tau.$$

Evaluating these terms we find that they are equal to the remaining terms of $\tilde{y}(t)$. Hence $\tilde{y}(t) = 0$ and the theorem is proved. □

Although this is only a preliminary result, further work will hopefully show that the conjecture, that the system (7.1)-(7.3) is controllable iff $(A_1,b_1),(A_2,b_2)$ and the pair (7.5) are controllable, does hold true.

One question that must be asked at this point is whether the multilinear approach to non-linear dynamical systems is likely to bear any fruit, but unfortunately it is still difficult to give a definite answer. Even in the simplest single input non-linear case, when the input/output map is identical to its own second-order Volterra kernel, W_2, quasi-reachability can easily break down. The following two examples illustrate this:

1) Let the state space description derived by considering W_2 as bilinear input/output map be given by

$$x_{k+1}^1 = A_1x_k^1 + b_1u_k$$
$$x_{k+1}^2 = A_2x_k^2 + b_2u_k$$
$$y_k = d^Tx_k^1 \otimes x_k^2$$

where the two separate input channels are now regarded as identical. Then it is obvious that the state space is not reachable if A_1 and A_2 have common eigenvalues.

2) Let the state space description be given by

$$x_{k+1} = ax_k + bu_k^2 \qquad (a,b > 0)$$
$$y_k = x_k.$$

It is clear that if $x_0 = 0$, then $x_k \geq 0$ for all k.

It is also conceivable that in the general non-linear case the Volterra approximation could well produce a larger state space than one derived straight from the input/output map itself.

However, multilinear system theory is undoubtedly of use when it comes to modelling a system with more than one input channel, when it is known that the inputs from separate channels interact multiplicatively to produce an output.

Various other approaches to non-linear system theory, besides the multilinear approach and the classical methods of examining concepts such as stability by means of approximations and norm inequalities, have been made in recent years. Fliess [F2] looks at Volterra series approximations with the aid of non-commutative formal series, and Sontag [S1],[S2] discusses discrete time polynomial systems, which are systems for which the state transitions are polynomial functions of the inputs and state variables. This supplements the work of such people as Isidori [I1] and Fliess [F2],[F3], who have studied the so-called bilinear system of the form

$$x_{k+1} = Ax_k + u_k F x_k + bu_k \qquad (7.7)$$
$$(x_k \in R^n).$$

This system also falls naturally into the class of affine systems discussed in Chapter 6, and has been looked at by Goguen [G1] in this category-theoretic context.

At this point it is worthwhile stating a conjecture concerning the reachable set of (7.7), which arose after reading [G1]:

Assume that there is no transformation $x_k \to Tx_k$ on (7.7), such that a substate $\hat{x}_k$ of Tx_k can be partitioned off as

$$\hat{x}_{k+1} = (\hat{A}+u_k\hat{F})\hat{x}_k$$

(or equivalently, assume that the reachable set of (7.7) is not contained within a subspace of R^n); then the reachable set, S, of (7.7) is given by

$$S = \{Ax + uFx + bu : x \in R^n, u \in R\}.$$

This is a trivial result for F = 0, since the subspace assumption holds iff (A,b) is a reachable pair, in which case it is obvious that $\{AX + bu : x \in R^n, u \in R\} = R^n$.

All in all, however, non-linear system theory, with its related aspects of stability and controllability, etc., is still very much an unresolved topic, and a great deal more research is required to bring the state of the art anywhere near that of linear system theory.

Nevertheless,nonlinear systems in general are still amenable to study by less exact methods.In particular,global input/output stability properties of the system (7.1)-(7.3) and of the continuous-time analogues of (7.7) are particularly suited to the off-axis circle criterion of Cho and Narendra [X1].The more recent application of circle theorems by Shankar and Atherton [X2] to nonlinear multivariable systems is also significant,as is the less recent but important theory of Liapounov functions(see e.g.[WI1]).

APPENDIX. LINEAR SYSTEM THEORY RESULTS

During the course of this thesis, two interesting results have been proved in linear system theory, and it seems convenient to restate them here, together with an independent theorem on cascaded linear systems which was proved in an early attempt at attacking the quasi-reachability result of Chapter 3.

The two earlier results are as follows:

1) Lemma 3.2.3

Let (A,b) be a reachable pair. Then for all $\ell > 0$, there exists an integer $N > \ell$ such that $(A^N, A^k b)$ is a reachable pair iff A is non-singular.

The interpretation of this result is that if a discrete-time system is constrained in such a way that all inputs must be separated by at least ℓ time intervals, with the initial input only permitted after time k, then invertibility of A guarantees that the system is still reachable.

2) Lemma 4.3.3

Let $[d^T h^T]\begin{bmatrix} F & O \\ C & A \end{bmatrix}^i \begin{bmatrix} G \\ B \end{bmatrix} = [\hat{d}^T \hat{h}^T]\begin{bmatrix} F & O \\ \hat{C} & \hat{A} \end{bmatrix}^i \begin{bmatrix} G \\ B \end{bmatrix}$ for all i,

where (h^T, A) and $(\hat{h}^T, \hat{A})$ are observable pairs

and $\left(\begin{bmatrix} F & O \\ C & A \end{bmatrix}, \begin{bmatrix} G \\ B \end{bmatrix} \right)$ and $\left(\begin{bmatrix} F & O \\ \hat{C} & \hat{A} \end{bmatrix}, \begin{bmatrix} G \\ \hat{B} \end{bmatrix} \right)$ are reachable pairs;

then there exists a similarity transformation relating the system matrices, which is of the form $\begin{bmatrix} I & O \\ Y & T \end{bmatrix}$, with T invertible.

The interpretation of this result is that given a cascaded linear system, which is known to be reachable, although not necessarily observable, and with the requirement that the (F,G) subsystem be included in the state space realization, then any two realizations of this system will be isomorphic, despite the fact that they may not be completely observable.

The new result that we present here concerns the cascaded linear system represented by

$$x_{k+1} = \begin{bmatrix} F & K \\ O & A \end{bmatrix} x_k + \begin{bmatrix} g \\ b \end{bmatrix} u_k \qquad \text{(A.1)}$$

where $g \in R^{n_1}$, $b \in R^{n_2}$, A and F are square matrices and (A,b) and (F,[K g]) are reachable pairs.

A long-standing problem has been to provide necessary and sufficient conditions for (A.1) to be reachable, without having to check whether the Kalman controllability matrix has rank n_1+n_2. Instead it has been hoped that a check for reachability will be provided by examining whether some other matrix has rank $m < n_1+n_2$.

Equation (A.1) has been studied by various people; in particular we mention Chen and Desoer [CD1], Chen [C1] and Davison and Wang [DW1], who have all made valid contributions to the multiple input case.

Here we present necessary and sufficient conditions for reachability of (A.1) which only require the examination of a matrix to be defined below, as to whether it has rank n_1 or less. Unfortunately these conditions are only valid for the single input case, and it is not clear how the approach taken here might be extended to take in the more general multiple input case.

Theorem A.1

The linear system (A.1) with (A,b) and (F,[K g]) reachable pairs is itself reachable iff

$$\text{rank } M(F) = n_1$$

where $M(z) = \text{adj}(zI-F)\,[K\,\text{adj}(zI-A)b + (\det(zI-A))g] \triangleq M_0+M_1z+\ldots+M_pz^p$

(so that $M(F) = M_0 \otimes I + M_1 \otimes F + \ldots + M_p \otimes F^p$)

Proof: We form the transfer function corresponding to (A.1) to obtain

$$x_1(z) = (zI-F)^{-1}[K(zI-A)^{-1}b+g]$$

$$x_2(z) = (zI-A)^{-1}b.$$

Rewriting this more conveniently we obtain

$$\begin{bmatrix} x_1(z) \\ x_2(z) \end{bmatrix} = \frac{1}{\psi_A(z)\psi_F(z)} \begin{bmatrix} W_F(z)(KX_A(z)+\psi_A(z)g) \\ X_A(z)\psi_F(z) \end{bmatrix}$$

$$\triangleq \frac{1}{\psi_A(z)\psi_F(z)} \begin{bmatrix} y_1(z) \\ y_2(z) \end{bmatrix} \qquad (A.2)$$

where $X_A(z) = \text{adj}(zI-A)b$

$W_F(z) = \text{adj}(zI-F)$

$\psi_A(z) = \det(zI-A)$

$\psi_F(z) = \det(zI-F)$.

Now (A.1) is not reachable iff $\exists\, a_1 \in R^{n_1}, a_2 \in R^{n_2}$, with $a_1 \neq 0$ such that

$$a_1^T x_1(z) + a_2^T x_2(z) = 0 \text{ identically.} \qquad (A.3)$$

Now the components of the numerator $y_2(z)$ are all contained in the ideal $(\psi_F(z))$, so that (A.3) holds

iff there exists a vector $a_1 \in R^{n_1}$ s.t. $a_1^T y_2(z) \in \psi_F(z)$

i.e. iff there exists an a_1 s.t. $a_1^T W_F(z)[K\ g]\begin{bmatrix} X_A(z) \\ \psi_A(z) \end{bmatrix} = k(z)\psi_F(z)$ (A.4)

for some $k(z) \in R[z]$.

Note that $\psi_A(z)$ and the components of $X_A(z)$ are linearly independent, because of (A,b) being a reachable pair.

Let us now write $X_F(z) = W_F(z)[K\ g]$

$$= \begin{bmatrix} x_{11}(z) & \cdots & x_{1n}(z) \\ \vdots & & \vdots \\ x_{n_1,1}(z) & \cdots & x_{n_1,n}(z) \end{bmatrix} \text{ where } n = n_2+1$$

and $\begin{bmatrix} X_A(z) \\ \psi_A(z) \end{bmatrix} = \begin{bmatrix} W_1(z) \\ \vdots \\ W_n(z) \end{bmatrix}$

Now $(F,[K\ g])$ is a reachable pair iff all minors of $X_F(z)/\psi_F(z)$ have common denominator equal to $\psi_F(z)$.

Hence $x_{ij}(z)x_{k\ell}(z) - x_{i\ell}(z)x_{kj}(z) = a_{ijk\ell}(z)\psi_F(z)$ for some $a_{ijk\ell}(z) \in \mathcal{R}[z]$.

We now have by (A.4)

$$\begin{aligned} &c_1[x_{11}(z)w_1(z)+\ldots+x_{1n}w_n(z)] \\ &+ c_2[x_{21}(z)w_1(z)+\ldots+x_{2n}(z)w_n(z)] \\ &+ \ldots + c_{n_1}[x_{n_1,1}(z)w_1(z)+\ldots+x_{n_1,n}w_n(z)] = k(z)\psi_F(z) \end{aligned} \qquad (A.5)$$

where $a_1^T = [c_1 \ldots c_n]$.

Multiplying (A.5) by $x_{11}(z)$ and substituting

$$x_{11}(z)x_{k\ell}(z) = x_{1\ell}(z)x_{k1}(z) + a_{11k\ell}(z)\psi_F(z)$$

we obtain

$$\begin{aligned} &c_1x_{11}(z)[x_{11}(z)w_1(z)+\ldots+x_{1n}(z)w_n(z)] \\ &+ c_2x_{21}(z)[x_{11}(z)w_1(z)+\ldots+x_{1n}(z)w_n(z)] \\ &+ \ldots + c_{n_1}x_{n_1,1}(z)[x_{11}(z)w_1(z)+\ldots+x_{1n}(z)w_n(z)] = b_{11}(z)\psi_F(z) \end{aligned}$$

where $b_{11}(z) = \sum_{k,\ell} a_{11k\ell}(z)w_\ell(z) + k(z)$.

Rearranging, we obtain

$$[c_1x_{11}(z)+\ldots+c_{n_1}x_{n_1,1}(z)][x_{11}(z)w_1(z)+\ldots+x_{1n}(z)w_n(z)] = b_{11}(z)\psi_F(z).$$

In a similar manner, multiplying (A.5) by $x_{1j}(z)$ $(j = 2,\ldots,n)$, we obtain

$$[c_1x_{1j}(z)+\ldots+c_{n_1}x_{n_1,j}(z)][x_{11}(z)w_1(z)+\ldots+x_{1n}(z)w_n(z)] = b_{1j}(z)\psi_F(z).$$

Now $(F,[K\ g])$ is a reachable pair, so that $a_1^TX_F(z) = 0$ implies $a_1 = 0$. By hypothesis, we have $a_2 \neq 0$, so that at least one of $c_1x_{ij}(z)+\ldots+c_{n_1}x_{n_1,j}(z) \neq 0$, but this last term has degree less than $\deg \psi_F(z)$, so that $x_{11}(z)w_1(z)+\ldots+x_{1n}(z)w_n(z)$ shares a common polynomial factor with F.

In a similar way, by multiplying (A.5) by $x_{ij}(z)$, we discover that

$$[c_1x_{1j}(z)+\ldots+c_{n_1}x_{n_1,j}(z)][x_{i1}(z)w_1(z)+\ldots+x_{in}(z)w_n(z)] = b_{ij}(z)\psi_F(z)$$

for all $i = 1,\ldots,n_1;\ j = 1,\ldots,n$.

Without loss of generality, let $c_1x_{1k}(z) + \ldots + c_{n_1}x_{n_1,k}(z) \neq 0$. It is then clear that all the polynomials $x_{i1}(z)w_1(z)+\ldots+x_{in}(z)w_n(z)$ share the same common polynomial factor of $\psi_F(z)$.

Utilizing Theorem 1 of Barnett [B1], we see that a necessary and sufficient condition for this to hold is that

$$\text{rank}\begin{pmatrix} x_{11}(F)w_1(F)+\ldots+x_{1n}(F)w_n(F) \\ x_{n_1,1}(F)w_1(F)+\ldots+x_{n_1,n}(F)w_n(F) \end{pmatrix} < n_1$$

which is precisely the condition stated. □

We can write a dual result for observability as follows using similar notation to that of Lemma 4.3.3:

The linear system

$$x_{k+1} = \begin{pmatrix} F & O \\ C & A \end{pmatrix} x_k + \begin{pmatrix} G \\ B \end{pmatrix} u_k$$

$$y_k = [d^T \; h^T]x_k$$

where (h^T, A) and $\left(\begin{bmatrix} d^T \\ C \end{bmatrix}, F \right)$ are observable pairs, it itself observable iff

$$\text{rank } M(F) = n_1$$

where $M(z) = [d^T(\det(zI-A)) + h^T \text{adj}(zI-A)C]\text{adj}(zI-F)$

and $F \in R^{n_1 \times n_1}$.

Of interest here is that both this result and Lemma 4.3.3 only hold for single output systems, which throws up a further analogy between bilinear input/output maps and cascaded linear systems. (The first analogy is that a cascaded linear system results from constraining a bilinear state space realization to sustain the substate x_k^1 at a constant level using a constant value for the input u_k.)

REFERENCES

[A1] M.A. Arbib, A characterization of multilinear systems, IEEE Transactions on Automatic Control, December 1969

[AAM1] B.D.O. Anderson, M.A. Arbib and E.G. Manes, Foundations of system theory: multidecomposable systems, Journal of the Franklin Institute, June 1976

[AL1] P. Alper, Some Aspects of the Volterra Series, Electronics Laboratory, Technological University, Delft, April 1964

[AM1] M.A. Arbib and E.G. Manes, Foundations of system theory: decomposable machines, Automatica, Vol.10, 1974

[AP1] T. Apostol, Mathematical Analysis, Addison-Wesley, 1974

[B1] S. Barnett, Matrices, polynomials and linear time-invariant systems, IEEE Transactions on Automatic Control, February 1973

[BO1] H.A. Barker and S.N. Obidegwu, Combined cross-correlation method for the measurement of second order Volterra kernels, Proceedings IEE, January 1973

[BU1] A.M. Bush, Some Techniques for the Synthesis of Nonlinear Systems, Technical Report 441, Research Lab. Elect., M.I.T., 1966

[C1] C.T. Chen, Introduction to Linear System Theory, Holt, Rinehart and Winston, 1970

[CD1] C.T. Chen and C.A. Desoer, Controllability of and observability of composite systems, IEEE Transactions on Automatic Control, August 1967

[D1] M.J. Denham, Canonical forms for the identification of multivariable linear systems, IEEE Transactions on Automatic Control, December 1974

[DW1] E.J. Davison and S.H. Wang, New results on the controllability and observability of general composite systems, IEEE Transactions on Automatic Control, February 1975

[F1] M. Fliess, Matrices de Hankel, Journal de Mathematiques, Pures et Appliquées, Vol.53, 1974

[F2] M. Fliess, Un outil algebrique: les séries formelles non commutatives, Conference on Algebraic System Theory, Udine 1975

[F3] M. Fliess, Sur les systèmes dynamiques bilinéaires qui sont linéaires, Comptes Rendues Acad. Sc. Paris, Série A, April 1974

[FM1] E. Fornasini and G. Marchesini, Algebraic realization theory of bilinear discrete time input/output maps, Journal of the Franklin Institute, January 1976

[FM2] E. Fornasini and G. Marchesini, On the internal structure of bilinear discrete time input/output maps, in Geometric Methods in System Theory, ed. R. Brockett and D.Q. Mayne, Reidel, Dordrecht, 1973

[G1] J.A. Goguen, Discrete-time machines in closed monoidal categories I, Journal of Computer and System Sciences, February 1975

[GE1] D.A. George, Continuous Non-Linear Systems, Technical Report 355, Research Lab. Elec., M.I.T., 1959

[I1] A. Isidori, Direct construction of minimal bilinear realizations from non-linear input/output maps, IEEE Transactions on Automatic Control, December 1973

[K1] R.E. Kalman, Pattern recognition properties of multilinear machines, IFAC Symposium on Technical and Bilogical Problems of Control, Yerevan, Armenian SSR, 1968

[K2] R.E. Kalman, P.L. Falb and M.A. Arbib, Topics in Mathematical System Theory (Chapter 10), McGraw-Hill, 1969

[LS1] Y.W. Lee and M. Schetzen, Measurement of the kernels of a nonlinear system by Crosscorrelation, M.I.T. Research Lab. Elec., Quarterly Progress Report 60, 1963

[S1] E.D. Sontag and Y. Rouchaleau, On discrete time polynomial systems, Conference on Algebraic System Theory, Udine, 1975

[S2] E.D. Sontag, On realizations of discrete-time nonlinear systems, Notices of the American Mathematical Society, June 1976

[W1] N. Wiener, Non-Linear Problems in Random Theory, M.I.T. Press, 1958

[WI1] J.L.Willems,Stability Theory of Dynamical Systems, Nelson,1970

[X1] Y.S.Cho and K.S.Narendra,An off-axis circle criterion for the stability of feedback systems with a monotonic nonlinearity,IEEE Transactions on Automatic Control,1968.

[X2] S.Shakar and D.P.Atherton,Graphical stability analysis of nonlinear multivariable control systems,International Journal of Control,March,1977.