

MULTIVARIABLE CONTROL SYSTEMS, FINITE-STATE LINEAR  
SEQUENTIAL MACHINES, AND PROJECTIVE GEOMETRIES:  
SOME EXPLICIT INTERCONNECTIONS

A THESIS

Presented to

The Faculty of the Division of Graduate  
Studies and Research

by

Ghulam Jailani Zalmai

In Partial Fulfillment

of the Requirements for the Degree

Doctor of Philosophy

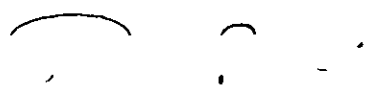
in the School of Industrial and Systems Engineering

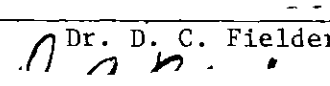
Georgia Institute of Technology

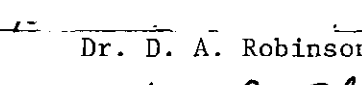
December 1977

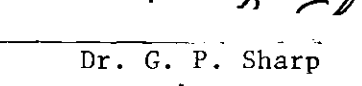
MULTIVARIABLE CONTROL SYSTEMS, FINITE-STATE LINEAR  
SEQUENTIAL MACHINES, AND PROJECTIVE GEOMETRIES:  
SOME EXPLICIT INTERCONNECTIONS

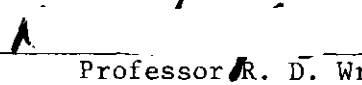
Approved:

  
\_\_\_\_\_  
Dr. P. Zunde, Chairman

  
\_\_\_\_\_  
Dr. D. C. Fielder

  
\_\_\_\_\_  
Dr. D. A. Robinson

  
\_\_\_\_\_  
Dr. G. P. Sharp

  
\_\_\_\_\_  
Professor R. D. Wright

Date approved by Chairman: 1/16/78

## ACKNOWLEDGEMENTS

I wish to express my sincere appreciation to Dr. P. Zunde, my dissertation advisor, for his continual assistance and encouragement throughout the course of this research.

I am grateful to all members of the reading committee:

Dr. D. C. Fielder, Dr. D. A. Robinson, Dr. G. P. Sharp, and Professor R. D. Wright, for their constructive criticisms and valuable suggestions.

I owe the greatest debt of gratitude to Dr. W. W. Hines and Dr. R. N. Lehrer for their material and moral support, and for their genuine willingness to help, throughout my graduate studies at Georgia Tech. It is a pleasure to record my sincere thanks and heartfelt appreciation to them.

## TABLE OF CONTENTS

	Page
ACKNOWLEDGEMENTS . . . . .	ii
GLOSSARY OF SELECTED SYMBOLS AND NOTATION . . . . .	vi
SUMMARY . . . . .	ix
Chapter	
I. INTRODUCTION . . . . .	1
1.1 General Finite-State Sequential Machines	
1.2 Linear Finite-State Sequential Machines	
II. LITERATURE SURVEY AND STATEMENT OF THE RESEARCH PROBLEM . . .	12
2.1 Literature Survey	
2.2 Statement and Relevance of the Research Problem	
III. INTRODUCTION TO LINEAR SEQUENTIAL MACHINES . . . . .	26
3.1 Mathematical Description of Finite-State Sequential Machines	
3.2 Interconvertibility of Mealy and Moore LSMs	
3.3 Input-State and Input-Output Transfer Characteristics of LSMs	
3.4 Indistinguishability, Isomorphism, Minimality, and Similarity in LSMs	
3.5 Input-Output Representation of LSMs	
3.6 Polynomial Representation of LSMs	
IV. STATE REACHABILITY AND STATE CONTROLLABILITY OF LSMS . . . .	52
4.1 State Reachability of LSMs	
4.2 State Unreachability of LSMs	
4.3 State Controllability of LSMs	
V. SOME CONSEQUENCES OF STATE REACHABILITY . . . . .	87
5.1 Canonical LSMs	
5.2 State Reachability and Canonical LSMs	
5.3 State Reachability and Feedback	

VI.	STATE REACHABILITY REVISITED . . . . .	145
6.1	Twenty Four Equivalent Criteria for the State Reachability Property of Single-Input LSMs	
6.2	Equivalence Classes of State Reachable Single-Input LSMs	
6.3	Eighteen Equivalent Criteria for the State Reachability Property of Multi-Input LSMs	
6.4	Equivalence Classes of State Reachable Multi-Input LSMs	
VII.	THE JORDAN CANONICAL FORM AND SELECTIVE STATE REACHABILITY OF LSMs . . . . .	204
7.1	The Jordan Canonical Form for LSMs	
7.2	Selective State Reachability of LSMs	
VIII.	PROJECTIVE-GEOMETRIC STRUCTURES AND LSMs . . . . .	264
8.1	Geometric Definition of State Reachability of LSMs	
8.2	(A, B)-Invariant Flats	
8.3	(A, B)-Invariant Flats and Output Invariance	
8.4	Reachability Flats	
8.5	Eigenvalue Assignability	
8.6	Reachability Flat Algorithm	
8.7	Supremal Reachability Flats	
8.8	Noninteraction in LSMs	
IX.	OUTPUT REACHABILITY AND OUTPUT CONTROLLABILITY OF LSMs . . .	304
9.1	Output Reachability of LSMs	
X.	OBSERVABILITY AND STATE OBSERVER DESIGN FOR LSMs . . . . .	310
10.1	State Observability of LSMs	
10.2	Consequences of the Duality Theorem	
10.3	State Minimization and Observability of LSMs	
10.4	State Observer Design for LSMs	
XI.	CONCLUSIONS AND RECOMMENDATIONS. . . . .	344
APPENDIX	. . . . .	350
	Finite Projective and Affine Spaces and Geometries	
BIBLIOGRAPHY	. . . . .	360
VITA	. . . . .	369

## LIST OF FIGURES

	Page
1.1.1. General Mealy Representation of a Sequential Machine . . .	3
1.1.2. Moore Model of a Sequential Machine . . . . .	4
1.1.3. A Simple State Transition Graph . . . . .	5
1.1.4. General Transition Table . . . . .	6
1.1.5. Transition Table for Machine of Fig. 1.1.3. . . . .	6
1.2.1. State Transition Graph for the LSM of Example 1.2.1. . . .	9
1.2.2. Transition Table for LSM of Example 1.2.1. . . . .	10
1.2.3. Realization Diagram for the LSM of Example 1.2.1. . . . .	11
3.4.1. Isomorphism, Indistinguishability, and Similarity Relations for Minimal and Non-minimal LSMs . . . . .	40
4.1.1. State Transition Graph for the LSM of Example 4.1.1. . . .	54
4.1.2. State Transition Graph for the LSM of Example 4.1.2. . . .	56
4.1.3. State Transition Graph for the LSM of Example 4.1.3. . . .	57
5.1.1. Compound Circuits of a LSM in Canonical Form . . . . .	106
5.1.2. Realization Circuits for LFSRs . . . . .	116
8.1.1. Diagram of Homomorphisms for Theorem 8.1.2. . . . .	270
10.4.1. State Observer Realization Diagram . . . . .	324
10.4.2. Realization Circuit for the LSM of Example 10.4.1. . . . .	329
10.4.3. Realization Circuit for the 2-Dimensional State Observer for the LSM of Example 10.4.11. . . . .	337

## GLOSSARY OF SELECTED SYMBOLS AND NOTATION

$A(X)$	affine geometry generated by the cosets of $X$
$A S$	restriction of $A$ to $S$
$\{A \mid R(B)\}$	$\equiv R(B) + AR(B) + A^2R(B) + \dots + A^{n-1}R(B)$ , reachability flat of the LSM $(A, B, C)$
$A^{-1*}S$	$\equiv \{x : Ax \in S\}$
$F(S)$	$\equiv \{F : P(X) \longrightarrow P(U) : S \in P_{A+BF}(X)\}$
$GF(q)$	Galois field: the set of all residue classes of integers modulo a prime; $q = p^r$ , where $p$ is a prime and $r$ an integer.
$GF(q)^n$	$\equiv GF(q^n)$ , $n$ -dimensional vector space over $GF(q)$
$GF(q)^{m \times n}$	vector space of $m \times n$ matrices over $GF(q)$
$GF(q)[\xi]$	ring of polynomials over $GF(q)$
$GF(q)[\xi]^{\ell \times m}$	vector space of $\ell \times m$ polynomial matrices over $GF(q)$
$GL(n, q)$	group of isomorphisms of $GF(q)^n$
$I_n$	$n \times n$ identity matrix
$I(A, B; X), I(X)$	$\equiv \{S \in P(X) : AS \subseteq S + R(B)\}$ , set of $(A, B)$ -invariant flats of $P(X)$
$I^O(A, B; X), I^O(X)$	dual of $I(X)$
ILSM	Internal Linear Sequential Machine
$K(A, B, n), K$	$\equiv [B, AB, A^2B, \dots, A^{n-1}B]$ , reachability matrix of the LSM $(A, B, C)$
$L(A, C, n), L$	$\equiv [C^T, A^T C^T, (A^T)^2 C^T, \dots, (A^T)^{n-1} C^T]^T$ , observability matrix of the LSM $(A, B, C)$
LFSR	Linear Feedback Shift Register
$\ell g(u)$	Length of $u \in U^*$
LSM	Linear Sequential Machine

$m$	$\equiv \dim U$
$n$	$\equiv \dim X$
$\underline{n}$	$\equiv \{1, 2, \dots, n\}$
$\underline{\underline{n}}$	$\equiv \{0, 1, 2, \dots, n\}$
$N(C)$	null space of $C$
$PGL(n+1, q)$	group of collineations of $P(X)$
$P(X)$	projective geometry generated by the set of all subspaces of $X$
$P(X^0), P^0(X)$	dual of $P(X)$
$P_A(X)$	$\equiv \{S \in P(X) : AS \subseteq S\}$ , set of $A$ -invariant flats of $P(X)$
$R(A, B; X)$	$\equiv \{R \in P(X) : R = \{A + BF \mid R(B) \cap R\}\}$ , set of generalized reachability flats of the LSM $(A, B, C)$
$R(B)$	range of $B$
RFA	Reachability Flat Algorithm
$U$	$\equiv GF(q)^m$ , input space of the LSM $(A, B, C)$
$U^j$	$\equiv \{u(0)u(1) \dots u(j-1) : u(i) \in U, i \in \underline{j-1}\}$
$U^*$	$\equiv \{u(0)u(1) \dots u(\ell-1) : \ell \geq 0, u(i) \in U\}$
$X_n(q), X$	$\equiv GF(q)^n$ , state space of the LSM $(A, B, C)$
$X_n^0(q), X^0$	canonical dual of $X_n(q), X$
$x_i$	$i$ th scalar
$x^i$	$i$ th vector
$(x)^i$	$i$ th power of the scalar $x$
$\langle x^1, x^2, \dots, x^n \rangle$	$\equiv \text{span} \{x^1, x^2, \dots, x^n\}$



$\mathcal{Y}$   $\equiv \text{GF}(q)^r$ , output space of the LSM (A, B, C)

$Z_1 \oplus Z_2 \oplus \dots \oplus Z_s$  block-diagonal matrix with diagonal blocks  
 $Z_i$ ,  $i \in \underline{s}$

## SUMMARY

In this dissertation, deterministic time-invariant finite-state linear sequential machines of the Moore type are treated as discrete-time control systems over the finite field  $GF(q)$ . Adopting a modern multivariable control theory approach and a finite-geometric point of view, various structural aspects of linear machines are investigated. In addition to presenting a mathematically formal account of the central concepts of state reachability and state controllability, and numerous equivalent formulations of these concepts, the relationships among state reachability, structural invariants, canonical forms, and state variable feedback are discussed. The concept of generalized eigenvectors is utilized in the framework of the Jordan canonical form to formulate additional reachability criteria, and introduce and develop in detail the notion of selective state reachability for linear sequential machines. State reachability is further studied in the context of the finite projective geometry and certain classes of flats related to the structural properties of linear machines are identified, some of their applications are demonstrated, and algorithms for their computation are discussed. Finally, the concept of state observability is investigated and its relationship to state reachability is established through a duality theorem. The role of the observability property in the state reconstruction process is illustrated by developing some design procedures for full- and reduced-order Luenberger type state observers for both single- and multi-input linear sequential machines.

## CHAPTER I

## INTRODUCTION

The past decade has witnessed a phenomenal proliferation of mathematical disciplines concerning models for digital cybernetical systems, that is, systems which receive, store, process, and discharge information under the control of a clock pulse. The unifying infrastructure for the majority of these models is a composite mathematical concept called a *finite-state sequential machine or automaton*. A general finite-state sequential machine is an idealized model for a large number of physical devices and phenomena encountered in many fields of science and technology. This important branch of dynamical systems theory has found numerous applications in practically every area of scientific and engineering investigation - from psychology to business administration, and from communication to linguistics. Ideas and techniques originally developed for sequential machines have been found useful in such diverse and seemingly unrelated problems as the investigation of human nervous activity, the analysis of English syntax, and the design of digital computers. Moreover, due to its unifying nature, this class of dynamical systems is undoubtedly one of the most valuable contributors to the growing trend in interdisciplinary cooperation which is becoming exceedingly indispensable for the progress and efficient utilization of today's scientific and technological endeavors and innovations.

### 1.1. General Finite-State Sequential Machines

A *finite-state sequential machine*  $M$  is a quintuple  $M = (X, U, Y, \phi, \eta)$ , where

- $X \equiv \{x^1, x^2, \dots, x^a\}$  is a finite nonempty set of states,
- $U \equiv \{u^1, u^2, \dots, u^b\}$  is a finite nonempty set of inputs,
- $Y \equiv \{y^1, y^2, \dots, y^c\}$  is a finite nonempty set of outputs,
- $\phi$  is a map from  $X \times U$  into  $X$ , called the *next state map*, and
- $\eta$  is a map from  $X \times U$  into  $Y$ , called the *output map*.

Physically  $M$  can be interpreted as a device whose input, output, and internal state at "time" (clock period)  $k$  are denoted by  $u(k)$ ,  $y(k)$ , and  $x(k)$ , respectively. These variables are defined for discrete--and for convenience, integral--values of  $k$  only, and assume values from the finite nonempty sets  $U$ ,  $Y$ , and  $X$ , respectively. Given the state  $x(k)$  and input  $u(k)$  at time  $k$ , the map  $\phi$  specifies the state at time  $k + 1$ , and the map  $\eta$  the output at time  $k$  as follows:

$$x(k + 1) = \phi(x(k), u(k)) \quad (1.1.1)$$

$$y(k) = \eta(x(k), u(k)) \quad (1.1.2)$$

The application of a sequence of  $\ell$  input symbols (or an input sequence of length  $\ell$ ) to  $M$  results in a sequence of states (or state sequence) and a sequence of output symbols (or an output sequence) of the same length. Given an input sequence  $\tilde{u}$ , the state of  $M$  when  $\tilde{u}$  is applied (the initial state of  $M$ ) and the maps  $\phi$  and  $\eta$ , the corresponding state sequence  $\tilde{x}$ , and the output sequence  $\tilde{y}$  can be computed recursively from equations (1.1.1) and (1.1.2).

A general representation of a sequential machine is shown in Fig. 1.1.1.

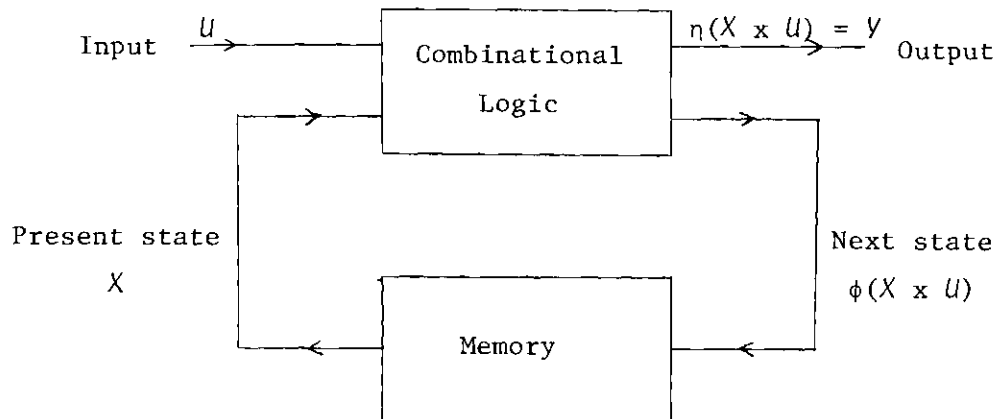


Fig. 1.1.1. General Mealy Representation of a Sequential Machine

A sequential machine of the above description is called a *Mealy machine*, named after G. H. Mealy who studied machines of this general type in [78]. A modification of Mealy model, which is frequently encountered, defines the output map  $\eta$  as restricted to a map of  $X$  into  $Y$ , that is, the output of the machine is dependent only on the state of the machine. This model is called a *Moore machine*, named after E. F. Moore [82] who gave a more abstract formulation and started the formal study of sequential machines which were initially introduced by Huffman [55]. These two sequential machine models provide a means for representing the formal properties of any deterministic machine. It can be shown that these representations can be converted from one to the other with certain trade-offs [12]. Fig. 1.1.1 and Fig. 1.1.2 illustrate the main features of these machine models.

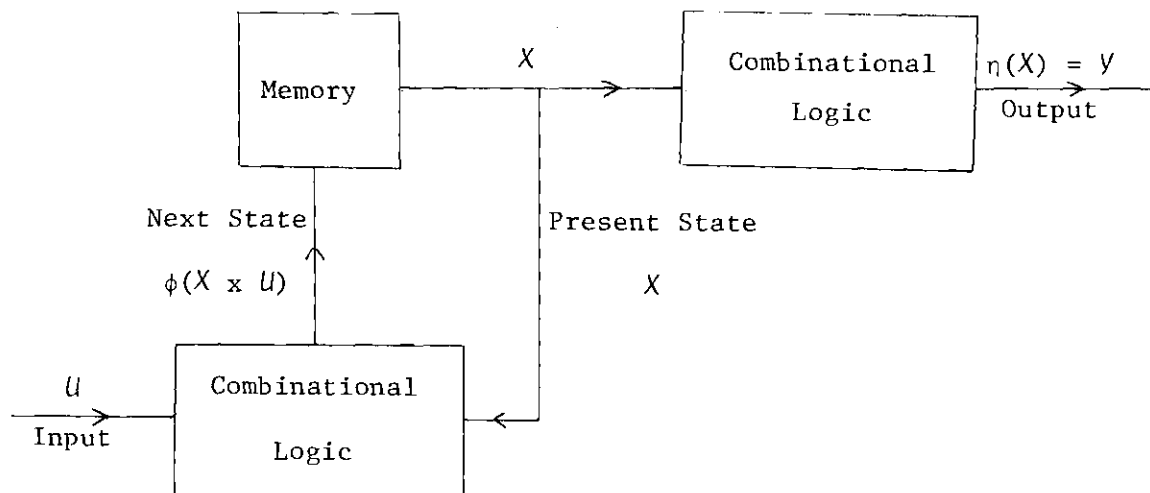


Fig. 1.1.2. Moore Model of a Sequential Machine

The characterizing maps  $\phi$  and  $\eta$  of a machine can be explicitly specified in a number of ways. In certain cases it may be possible to give them as compact mathematical expressions such as

$$\phi(x(k), u(k)) = Ax(k) + \sum_{i=1}^m N_i u_i(k)x(k)$$

$$\eta(x(k), u(k)) = Cx(k)$$

where, for example, the states, inputs, and outputs are  $n$ -,  $m$ -, and  $r$ -vectors, respectively, and  $A$ ,  $N_i$ ,  $i \in \underline{m}$ , and  $C$  are matrices of appropriate dimensions with elements over a certain finite field, say  $GF(q)$  (Galois field), such that the machine operations are compatible with the properties of the ground field.

Two other conventional methods of describing the characterizing maps  $\phi$  and  $\eta$  of  $M$  are by means of *state transition graphs* and *transition tables*. A transition graph is a labeled oriented graph which has one

vertex for each state of  $M$ , and one edge for each state input pair of  $M$ . For example, consider the sequential machine represented by the following state transition graph:

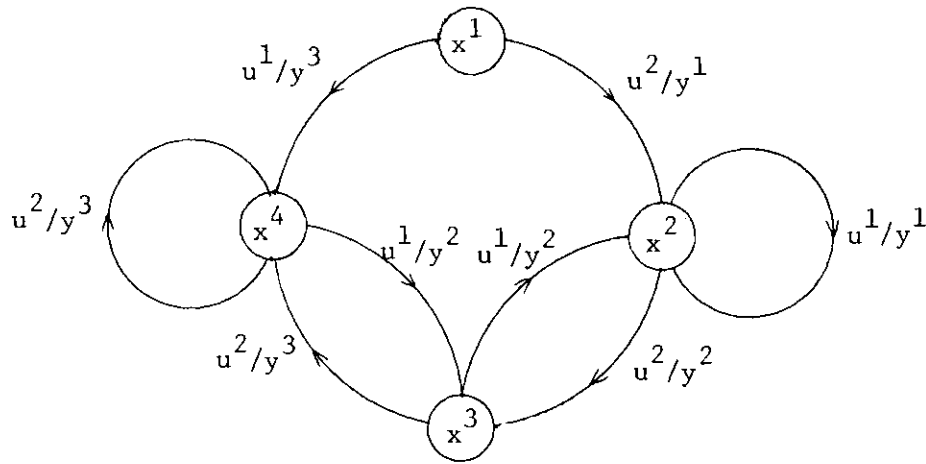


Fig. 1.1.3. A Simple State Transition Graph

In this graph, an edge directed from  $x^i$  to  $x^j$  having the label  $u^a/y^b$  indicates that when the machine is in state  $x^i$ , an input  $u^a$  will produce the current output  $y^b$  and will result in the next state  $x^j$ .

For the sequential machine represented by the state transition graph of Fig. 1.1.3, we have the following characterizing sets and maps:

$$X = \{x^1, x^2, x^3, x^4\}$$

$$U = \{u^1, u^2\}$$

$$Y = \{y^1, y^2, y^3\}$$

$$\phi(x^1, u^1) = x^4$$

$$\phi(x^1, u^2) = x^2$$

$$\phi(x^2, u^1) = x^2$$

$$\phi(x^2, u^2) = x^3$$

$$\phi(x^3, u^1) = x^2$$

$$\eta(x^1, u^1) = y^3$$

$$\eta(x^1, u^2) = y^1$$

$$\eta(x^2, u^1) = y^1$$

$$\eta(x^2, u^2) = y^2$$

$$\eta(x^3, u^1) = y^2$$

$$\phi(x^3, u^2) = x^4$$

$$\phi(x^4, u^1) = x^3$$

$$\phi(x^4, u^2) = x^4$$

$$\eta(x^3, u^2) = y^3$$

$$\eta(x^4, u^1) = y^2$$

$$\eta(x^4, u^2) = y^3$$

The format of a general transition table is shown in Fig. 1.1.4.

		x(k+1)				y(k)			
x(k)	u(k)	u <sup>1</sup>	u <sup>2</sup>	. . .	u <sup>r</sup>	u <sup>1</sup>	u <sup>2</sup>	. . .	u <sup>r</sup>
x <sup>1</sup> (k)				Entries from X				Entries from Y	
x <sup>2</sup> (k)									
.									
.									
x <sup>q</sup> (k)									

Fig. 1.1.4. General Transition Table

As an example, we will show the transition table of the sequential machine represented by state transition graph of Fig. 1.1.3, as follows:

		x(k + 1)		y(k)	
x(k)	u(k)	u <sup>1</sup>	u <sup>2</sup>	u <sup>1</sup>	u <sup>2</sup>
x <sup>1</sup>		x <sup>4</sup>	x <sup>2</sup>	y <sup>3</sup>	y <sup>1</sup>
x <sup>2</sup>		x <sup>2</sup>	x <sup>3</sup>	y <sup>1</sup>	y <sup>2</sup>
x <sup>3</sup>		x <sup>2</sup>	x <sup>4</sup>	y <sup>2</sup>	y <sup>3</sup>
x <sup>4</sup>		x <sup>3</sup>	x <sup>4</sup>	y <sup>2</sup>	y <sup>3</sup>

Fig. 1.1.5. Transition Table for Machine of Fig. 1.1.3



Gill [38] describes a large variety of situations that lend themselves to representation by the basic finite-state sequential model.

### 1.2. Linear Finite-State Sequential Machines

A small but extremely important subclass of sequential machines results when the next state and output maps,  $\phi$  and  $\eta$ , are assumed to be linear, that is, if we assume that there exist, for each  $k$ , four  $\text{GF}(q)$ -homomorphisms

$$\begin{aligned} A(k) &: X \rightarrow X \\ B(k) &: U \rightarrow X \\ C(k) &: X \rightarrow Y \\ D(k) &: U \rightarrow Y \end{aligned} \tag{1.2.1}$$

such that state transitions and outputs are given by the following vector difference equations:

$$\begin{aligned} x(k+1) &= A(k)x(k) + B(k)u(k) \\ y(k) &= C(k)x(k) + D(k)u(k) \end{aligned} \tag{1.2.2}$$

In (1.2.1),  $X$  is the state space,  $U$  is the input space, and  $Y$  is the output space of the linear sequential machine (1.2.2). If we restrict ourselves to finite-dimensional vector spaces over a finite field  $\text{GF}(q)$ , then the homomorphisms (1.2.1) are simply matrices of appropriate dimensions over  $\text{GF}(q)$ , and (1.2.2) is called a *finite-state linear time-varying sequential machine*. Furthermore, if the characterizing matrices of (1.2.2) do not depend on the "time"  $k$ , then the linear sequential machine is said to be *time-invariant* and is described as follows:

$$x(k + 1) = Ax(k) + Bu(k) \quad (1.2.3)$$

$$y(k) = Cx(k) + Du(k)$$

or, for the sake of notational simplicity, as  $(A, B, C, D)$ . If  $GF(q)^j$  denotes the vector space of  $j$ -component vectors and  $GF(q)^{i \times j}$  the vector space of  $i \times j$  matrices over the field  $GF(q)$ , then at time (clock period)  $k$ ,  $x(k) \in GF(q)^n$  is the state,  $u(k) \in GF(q)^m$  is the input, and  $y(k) \in GF(q)^r$  is the output of the machine (1.2.3). Moreover,  $A \in GF(q)^{n \times n}$ ,  $B \in GF(q)^{n \times m}$ ,  $C \in GF(q)^{r \times n}$ , and  $D \in GF(q)^{r \times m}$ .

In analogy with the classification of general sequential machines, the linear sequential machine (1.2.3) whose output depends on both the state and input, is called a *Mealy linear sequential machine*. However, if the output of a linear sequential machine depends only on the state, then it is called a *Moore linear sequential machine*, and has the form

$$x(k + 1) = Ax(k) + Bu(k) \quad (1.2.4)$$

$$y(k) = Cx(k)$$

Example 1.2.1. Consider the following single-input, single-output Moore linear sequential machine over  $GF(2)$ :

$$\begin{pmatrix} x_1(k + 1) \\ x_2(k + 1) \\ x_3(k + 1) \\ x_4(k + 1) \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} x_1(k) \\ x_2(k) \\ x_3(k) \\ x_4(k) \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix} u(k)$$

$$y(k) = x_1(k) + x_4(k)$$

The state set, the input set, the output set, the state graph, and the transition table for this linear machine are shown below.

$$X = \left\langle \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix} \right\rangle$$

$$U = \{0, 1\}$$

$$Y = \{0, 1\}$$

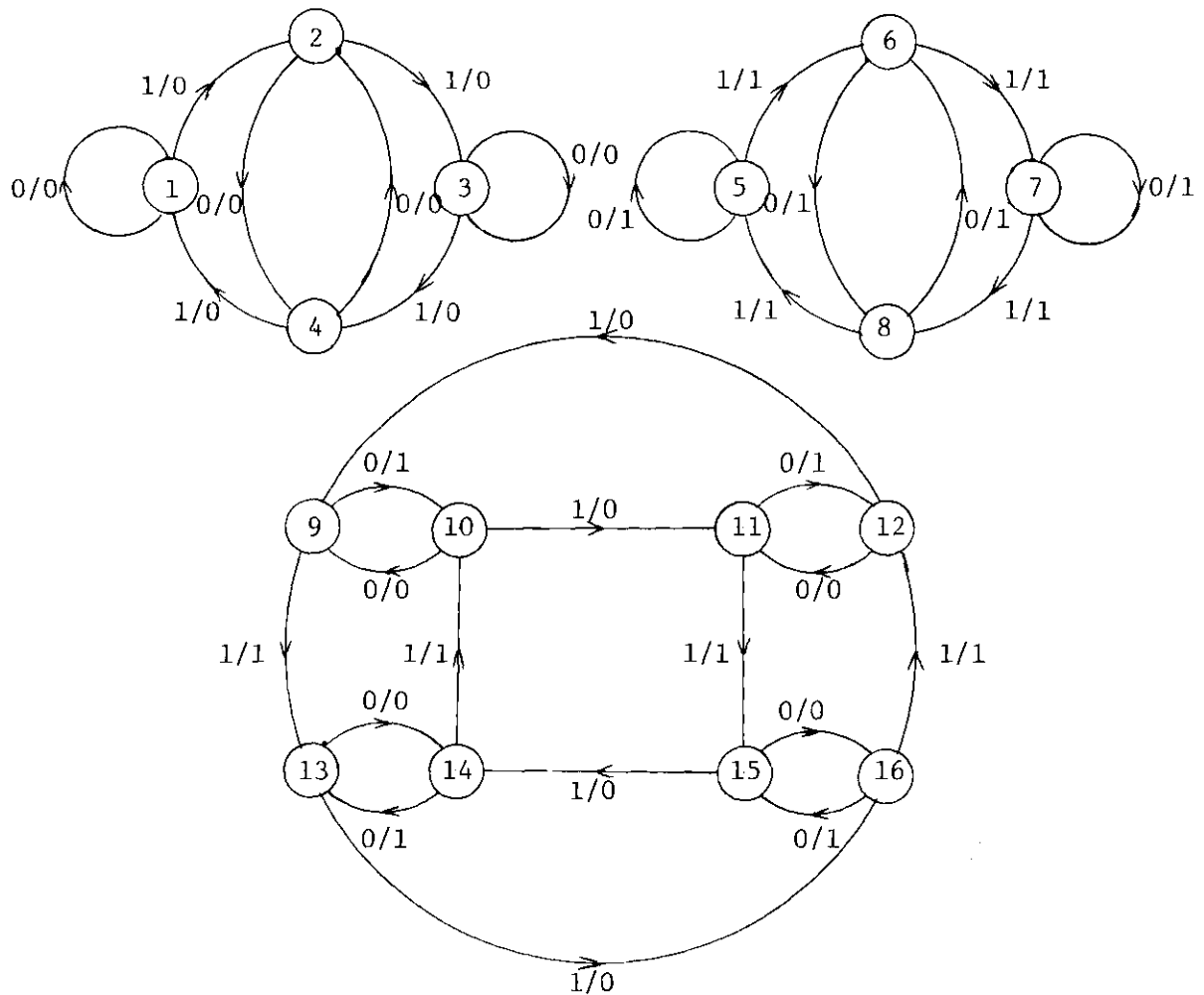


Fig. 1.2.1. State Transition Graph for the LSM of Example 1.2.1

$x(k)$	$u(k)$	$x(k + 1)$		$y(k)$	
		0	1	0	1
$x^1(k) = [0, 0, 0, 0]^T$		$[0, 0, 0, 0]^T$	$[1, 0, 0, 1]^T$	0	0
$x^2(k) = [1, 0, 0, 1]^T$		$[0, 0, 1, 0]^T$	$[1, 0, 1, 1]^T$	0	0
$x^3(k) = [1, 0, 1, 1]^T$		$[1, 0, 1, 1]^T$	$[0, 0, 1, 0]^T$	0	0
$x^4(k) = [0, 0, 1, 0]^T$		$[1, 0, 0, 1]^T$	$[0, 0, 0, 0]^T$	0	0
$x^5(k) = [1, 0, 0, 0]^T$		$[1, 0, 0, 0]^T$	$[0, 0, 0, 1]^T$	1	1
$x^6(k) = [0, 0, 0, 1]^T$		$[1, 0, 1, 0]^T$	$[0, 0, 1, 1]^T$	1	1
$x^7(k) = [0, 0, 1, 1]^T$		$[0, 0, 1, 1]^T$	$[1, 0, 1, 0]^T$	1	1
$x^8(k) = [1, 0, 1, 0]^T$		$[0, 0, 0, 1]^T$	$[1, 0, 0, 0]^T$	1	1
$x^9(k) = [0, 1, 0, 1]^T$		$[0, 1, 1, 0]^T$	$[1, 1, 1, 1]^T$	0	0
$x^{10}(k) = [0, 1, 1, 0]^T$		$[0, 1, 0, 1]^T$	$[1, 1, 0, 0]^T$	1	1
$x^{11}(k) = [1, 1, 0, 0]^T$		$[0, 1, 0, 0]^T$	$[1, 1, 0, 1]^T$	0	0
$x^{12}(k) = [0, 1, 0, 0]^T$		$[1, 1, 0, 0]^T$	$[0, 1, 0, 1]^T$	1	1
$x^{13}(k) = [1, 1, 1, 1]^T$		$[0, 1, 1, 1]^T$	$[1, 1, 1, 0]^T$	1	1
$x^{14}(k) = [0, 1, 1, 1]^T$		$[1, 1, 1, 1]^T$	$[0, 1, 1, 0]^T$	0	0
$x^{15}(k) = [1, 1, 0, 1]^T$		$[1, 1, 1, 0]^T$	$[0, 1, 1, 1]^T$	1	1
$x^{16}(k) = [1, 1, 1, 0]^T$		$[1, 1, 0, 1]^T$	$[0, 1, 0, 0]^T$	0	0

Fig. 1.2.2. Transition Table for LSM of Example 1.2.1

A linear sequential machine can always be realized by using three primitive components over  $GF(p)$ , namely, modulo- $p$  adders, modulo- $p$  scalars, and unit delayers. The number of delayers in an LSM is called the *dimension* of the LSM. Scalars with 0 and 1 signify an open connection and a closed connection, respectively. Thus, an LSM over  $GF(2)$ ,

a binary machine such as the LSM of Example 1.2.1, consists of adders modulo 2 which are commonly known as EXCLUSIVE-OR gates. In general, for any given LSM  $M = (A, B, C, D)$  an electronic circuit can always be constructed which simulates the operation of the machine. Conversely, for any meaningful interconnection of a finite number of primitive components over  $GF(p)$  representing the operation of an LSM, we can always write down the describing state and output equations of the LSM.

A realization circuit for the LSM of Example 1.2.1 is shown in Fig. 1.2.3.

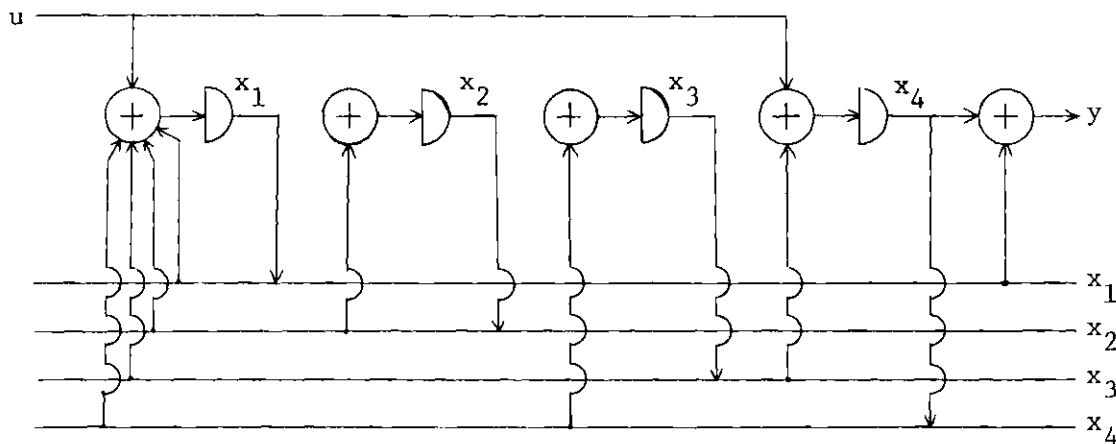


Fig. 1.2.3. Realization Diagram for the LSM of Example 1.2.1

### Summary and Conclusions

For the sake of an overall comparison, in this introductory chapter the general finite-state sequential machine model and the special class of linear sequential machines were briefly described.

## CHAPTER II

### LITERATURE SURVEY AND STATEMENT OF THE RESEARCH PROBLEM

The primary objective of our research is to examine the role of modern linear multivariable control systems theory in the study of linear sequential machines. In keeping with this objective, we will be primarily interested in the survey and assessment of published and otherwise available results concerning the nonautonomous LSM models. Although the class of autonomous LSMs as the prototype model of general LSMs is of paramount importance in its own right, it will not be given any appreciable consideration in our literature survey since the study of this class does not involve any control concepts and essentially belongs to the realm of recurrence sequence theory over  $GF(q)$ . Moreover, certain aspects of it can be treated as special cases of the general nonautonomous model.

#### 2.1. Literature Survey

The first treatment of linear sequential machines was presented by Huffman [57] in 1955. He considered the analysis and synthesis of LSMs comprising unit delays and modulo-2 adders, and briefly considered also modulo-3 elements. Shortly following this initial work, Elspas [32], Friedland [33], and Hartmanis [50] extended Huffman's ideas in several directions and to more general cases.

Elspas [32] considered autonomous LSMs, that is, LSMs with no inputs, and investigated the relation between internal machine logical structure and sequential behavior (cycle structure) for all possible cycle lengths in general, not merely for maximal cycle lengths. Furthermore, the class of internal machines treated by Elspas was not limited to shift registers with feedback, but included arbitrary interconnections of delay elements and the linear logic element. Finally, he generalized the binary situation to that of a multivalued  $p$ -nary logic, where  $p$  is any prime interger, as suggested by Huffman [57]. The results obtained by Elspas included an analysis procedure for autonomous LSMs which could be used to derive the sequential behavior analytically from a knowledge of the logical structure, realizability criteria, a class of canonical realizations, and effective synthesis procedures for finding economical realizations of LSMs.

Simultaneously and independently, Friedland [33] and Hartmanis [50] also generalized Huffman's results to autonomous LSMs comprising unit delays and modulo- $p$  elements. They also investigated some properties of delay polynomials and their application in developing realization procedures for simple LSMs.

Many different aspects of autonomous LSMs have later been studied by other authors [14], [16], [34], [39], [40], [43], [46], [54], [67], [73], [74], [89], [93], [102], [103], [120].

Due to their wide range of applicability, autonomous LSMs have been extensively studied. These LSMs can be regarded as special devices which independently generate sequences of symbols, rather than transform

externally applied sequences. The study of autonomous LSMs, therefore, is also the study of the important class of LSMs employed as "sequence generators" which are extensively utilized in coding and other digital tasks. Furthermore, the study of autonomous LSMs is an essential step in the evaluation of the total response of nonautonomous LSMs.

In an attempt to capitalize on the many theoretical and practical niceties associated with the property of linearity of sequential machines, some efforts have been made to develop some systematic test procedures for determining whether a sequential machine, given in the form of a transition table, can be represented as an LSM. This problem was initially considered by Srinivisan [100] and then expanded upon by several other authors [16], [29], [51], [115].

The first treatment of certain aspects of LSMs from a modern control theory point of view was given by Cohn [23] who investigated the state controllability properties of LSMs and showed that Kalman's controllability theorem for conventional infinite-state systems holds also for the case of LSMs. In fact, this possibility was already recognized by Kalman [60] as he noted the similarity between his theorem and a theorem due to Moore [82]. Kalman's theorem dealt with controllability and observability of linear differential systems, while Moore's theorem was concerned with strongly connected automata and indistinguishable states. "Evidently," wrote Kalman, "the two theorems are concerned with the same abstract facts, each being stated in a different mathematical framework." At any rate, Cohn was the first to formalize the concept of state controllability for LSMs. More



specifically, he showed that the LSM  $M = (A, B, C, D)$  is  $\ell$ -state controllable, that is, there exists an admissible input sequence  $u(0) u(1) \dots u(\ell-1)$  that will drive the LSM from an initial state  $x^1$  to a final state  $x^2$ , if and only if the rank of the  $n \times \ell m$  matrix  $[B, AB, A^2B, \dots, A^{\ell-1}B]$  is equal to  $n$ . He also proved that for LSMs state controllability coincides with strong connectivity. Furthermore, he presented a procedure for generating input sequences for controlling LSMs in minimum time. Later in [25], the same author studied, in the spirit of control theory, some additional properties of LSMs, namely, definiteness, finite memory, information losslessness, and observability which he called diagnosability. Concerning the concept of observability, he proved that the LSM  $M = (A, B, C, D)$  is  $\ell$ -observable, that is, every initial state  $x(0)$  of the LSM can be uniquely determined from the knowledge of the outputs  $y(0), y(1), \dots, y(\ell)$ , if and only if the rank of the  $n \times \ell r$  matrix  $[C^T, A^T C^T, (A^T)^2 C^T, \dots, (A^T)^{\ell-1} C^T]$  is equal to  $n$ .

Although the importance of the concepts of controllability and observability for LSMs were emphasized and some areas of application were indicated by Cohn [23], [25] and Cohn and Even [26], more concrete examples of application as well as theoretical significance of these concepts were actually presented by Massey and Sain [75], [76], and Massey [77]. These authors realizing the fact that the theories of codes, automata, and continuous systems are intimately intertwined, investigated the explicit interconnections and parallelisms existing among these theories. Their results established and clearly characterized some important relationships between the zero-state response,

the zero-input response, controllability and observability of LSMs and the classes of convolutional codes and cyclic codes which are the most important classes of codes that have been found to date, and the problems of burst correction, error detection, and error propagation for these classes of codes. Their exposition provided explicit examples of the resulting benefits accruing to each of these areas from the others, indicating the advantages of an increased exchange of ideas among these disciplines.

We would like to point out the fact that the central theme of the work of these authors is obviously an exemplary reiteration of the urgent need for developing a unified framework encompassing many of the seemingly different disciplines of dynamical systems theory. Pioneering efforts in this direction have already been made by Kalman [61], Arbib [1], [2], and others.

More recently, Tzafestas [105], [106], [107] has investigated some aspects of LSMs from a modern control theory point of view. In [105] he has developed output controllability criteria which are essentially similar to those available for conventional linear systems, and in [106] he has indicated a design procedure for a state observer for LSMs. In an effort to indicate the possibility and desirability of developing a unified and integrated sequential machine control theory, Tzafestas [107] has briefly surveyed some techniques of modern control theory applicable to some aspects of LSMs. More specifically, he has considered the following aspects of LSMs which, due to their matrix-theoretic nature, lend themselves to analysis and synthesis techniques

similar to those originally developed for conventional linear systems: state controllability, output controllability, observability, canonical decomposition, minimization, identification, canonical state space models, interconnections, minimal-time control, state reconstruction, decoupling, and inversion.

Fourier and Laplace transform techniques, among others, provide powerful analytic tools for the study of conventional linear systems. It is conceivable that one might attempt to see if similar operational techniques can be developed for LSMs. Such attempts have been made, resulting into a number of transform methods for LSMs.

Hohn [53] has reported that in 1952, J. G. Tryon had invented a delay operator for the study of synchronous digital machines. The Tryon delay operator differs from that later introduced by Huffman [57], in one essential respect. Tryon's method assumes a characteristic inherent delay not less than zero in each type of logical element, that is, these physical devices do not perform their logical operations instantaneously. For example, if pulses are applied to the input leads of an AND-element at time 6, the output is not necessarily obtained at time 6 but rather at time  $6 + i$ , where  $i \geq 0$  is what is called the *inherent delay* of the AND-element. Tryon's delay operator specifically recognizes this inherent delay and dictates algebraically the location of such pure delay elements as are required to assure proper operation of the machine. Huffman's approach is to assume that all logical elements act instantaneously, all delays being concentrated in suitably located pure delay elements. Huffman's operator is in effect a special

case of Tryon's in which the inherent delays of all logical elements are assumed to be zero. Hohn [53] presented the mathematical development of Tryon's operational method and illustrated its applicability in the analysis and synthesis of synchronous linear and nonlinear machines. Despite its generality and precedence, Tryon's method has not been much used in the area of sequential machines. On the contrary, Huffman's delay transform which is essentially an application of the concept of generating functions to the analysis of sequences of symbols, has been extensively used in the study of the special class of quiescent LSMs, that is, LSMs whose initial state (at  $k = 0$ ) is zero. Quiescent LSM's are widely used as special devices which transform input sequences into output sequences in accordance with some fixed rule, which implies a fixed initial state-0 for convenience. Therefore, the study of quiescent LSMs is also the study of the important class of LSMs employed as "sequence transformers."

Huffman's delay transform (d-transform) is applicable to sequences that are zero for  $k < 0$ . The d-transform  $G(d)$  of a sequence  $\{g(k)\}$  is defined by the following expression

$$G(d) = \sum_{k=0}^{\infty} g(k)d^k$$

Using this operational procedure, one can express the input-output relation of an LSM in terms of polynomials in the indeterminate  $d$ , called *delay polynomials*, and hence represent the LSM in terms of transfer functions similar to those of the classical linear control systems.

Therefore, some analysis and synthesis techniques from classical linear control theory which are based on transfer function methods can be easily modified and adapted to LSMs.

Although the d-transform seems to be very similar to the z-transform of sampled-data systems, there are basic differences between these two transforms. For example, the transform variable  $z$  in the z-transform is a complex variable which can be given many meaningful interpretations in the context of sampled-data theory, while the indeterminate  $d$  in the d-transform is practically devoid of any useful interpretations. Other dissimilarities obviously exist in relation to the questions of convergence properties, transform pair properties, etc.

Another operational technique which has been developed for LSMs is the Laplace-Galois transform introduced by Tsyarkin and Faradzev [104]. With the aid of this transform method, one can introduce and utilize some important classical control concepts such as transfer functions and frequency domain characteristics for sequential machines.

Richalet [95], making use of the theory of Galois fields and formal series, has introduced the fundamentals of an operational calculus for the finite sequence space of finite fields and rings, and has demonstrated its applicability to LSMs.

Richalet introduces his transform technique by associating a formal series  $V(\sigma)$  with an infinite sequence  $\{v(k)\}_{k=0}^{\infty}$  of elements of the field  $GF(q)$  by the following rule

$$\{v(k)\}_{k=0}^{\infty} \xrightarrow{V} \sum_{k=0}^{\infty} \frac{v(k)}{\sigma^{k+1}}$$

and calls the formal series  $V(\sigma) \equiv \sum_{k=0}^{\infty} \frac{v(k)}{\sigma^{k+1}}$ , the discrete Laplace transform modulo  $p$  of the sequence  $\{v(k)\}_{k=0}^{\infty}$ . Then he investigates some properties of the transform pairs analogous to the transform properties of the ordinary Laplace transform, such as initial and final value theorems, multiplication by the transform variable  $\sigma$ , translation, scaling, differentiation, convolution, and inversion. Finally, he briefly demonstrates the relevance of this operational technique to some simple analysis and synthesis problems of LSMs.

Except for Huffman's delay transform, the other operational procedures introduced for the study of LSMs, do not seem to have found application in any appreciable extent.

## 2.2. Statement and Relevance of the Research Problem

In the past, certain important classes of problems such as analysis and design of encoders and decoders, error detection, and error correction in the area of coding theory, computation in finite fields, information and data transmission and storage, have been treated largely by techniques from the domain of automata theory which were totally unrelated to the discipline of modern control systems theory. However, fairly recent preliminary research has revealed the fact that there exist many interconnections and parallelisms between these theories which could be effectively exploited for the purpose of developing a unified framework for these and other related fields. This unification will, on the one hand, make some of the above-mentioned and other classes of problems amenable to treatment by the methods of

control theory and, on the other hand, provide valuable opportunity for deriving additional results and insights from the cross-fertilization of these two systems disciplines.

One of the areas of automata theory which has enjoyed great generality in modeling many physical phenomena in different areas of science and engineering is the class of finite-state sequential machines. A small but extremely important subclass of general sequential machines is the special subclass of finite-state time-invariant linear sequential machines whose mathematical representation is given by equations (1.2.3). Linear sequential machines are of great interest for two important reasons. First, these linear machines constitute a subclass of the class of finite-state machines where powerful theories of finite groups, rings, fields and other algebraic structures, and of linear vector spaces can be exploited to advantage. As such, LSMs constitute a link between the general sequential machine and the general linear machine and offer insight into the operation of both. Furthermore, LSMs provide insight into the methods that may be used to decompose complex machines into an interconnection of smaller machines. Secondly, LSMs have found many applications in computer control circuitry, design of digital control and communication systems, generation of linear codes, synthesis of encoders and decoders, implementation of error detection and correction codes, computation in the ring of polynomials, computation in finite fields, counting and timing, generation of minimum time test sequences, generation of pseudo-random sequences (for use in the implementation of Monte Carlo programs, range measurements in

radar, probabilistic experiments, etc.), and in other aspects of automata theory.

Certain aspects of LSMs have been studied in the context of automata theory and rather superficially in the framework of modern control theory. In this research, treating LSMs as discrete-time finite-state control systems and adopting a modern multivariable control theory approach, we will investigate the possibility of developing a fairly comprehensive structure theory for them. We will select the dual concepts of reachability and observability as the pivotal components of this theory. This is, of course, a natural choice since, as it will be demonstrated in the sequel, these concepts and their extensions prove to be of enormous importance in various analysis and synthesis aspects such as event synchronization and memory address control in digital systems, minimal-time optimal control, decomposition, noninteraction, disturbance decoupling, canonical representation, feedback shift register realizability, state minimization, identification, feedback compensation, state reconstruction, inversion, and so forth.

As it was pointed out in the preceding section, the concepts of controllability and observability for LSMs have been treated in [23], [75], [77], [105], and [107] in a surprisingly superficial manner. In fact, these treatments are so cursory in scope that not even in a single one of them the crucial distinction is made between the properties of reachability and controllability or between observability and reconstructibility. In addition to performing an in-depth investigation of these dual concepts and some of their ramifications in a state space



setting, we will introduce the fundamentals of a projective-geometric approach for the study and characterization of certain structural aspects of LSMs. This new point of view is motivated by a number of factors: first of all, the area of finite projective geometry has been extensively developed, appears to be endowed with rich combinatorial structures, and has found applications in coding theory. On the other hand, LSMs have been widely used in various phases of the coding process. Therefore, it is natural to expect that establishing some connections between certain areas of LSMs and finite projective geometry, and thus closing the underlying triangle of ideas, will contribute to a more constructive conceptual and practical interplay among LSMs, coding theory and finite geometries. Secondly, a geometric treatment can provide a more general and elegant representational framework for developing a structure theory for LSMs. Finally, the increasing prevalence of geometric ideas in the literature of conventional dynamical systems suggests the desirability of similar geometric concepts in the area of automata theory and, in particular, in the area of LSMs.

In more specific terms, the bulk of our research effort will be devoted to the following aspects of LSMs: state reachability, state controllability, canonical forms, state feedback, output reachability, selective state reachability, geometric state reachability, state observability, and state observer design.

As advocated above, our approach will consist of modern control theory in the framework of finite geometries. The motivation for this deviation from the conventional algebraic and combinatorial approaches of automata theory is twofold:

1. Research in the area of sequential machines has been restricted primarily to state assignment and coding, state reduction, decomposition, and design of physically realizable models. We hope that a new look at LSMs through modern control theory will open up new vistas of theoretical and applied research in the field of sequential machines.
2. The results obtained by adopting a control-theoretic approach to investigate LSMs will further contribute to the development of a more unified framework for automata theory and control theory.

### Summary and Conclusions

In this chapter, a fairly comprehensive literature survey was reported. Due to our primary interest in examining the status of linear machine control theory, the survey was mostly restricted to the area of nonautonomous LSMs.

In the course of the literature survey it was readily revealed that although certain aspects of LSMs were investigated in a fragmentary and superficial manner from a modern control theory point of view, no attempt towards developing a coherent linear machine control theory had ever been made. This fact coupled with the enormous importance of LSMs and the anticipation of initiating a constructive interplay between automata theory and control theory seemed to provide ample

justification for embarking upon a systematic investigation of LSMs from the standpoint of modern multivariable control theory. Consequently, a research plan was formulated and a precise statement of the research problem was presented.

## CHAPTER III

### INTRODUCTION TO LINEAR SEQUENTIAL MACHINES

For the purpose of establishing consistent notation and terminology, most of this chapter will be devoted to a brief review of the basic concepts and definitions pertaining to nonautonomous LSMs. For slightly more comprehensive treatments, the references [14], [46], and [49] may be consulted. The discussion of formal polynomials over  $GF(q)$  and formal polynomial representation of LSMs is intended to point out the real possibility for the development of an extensive linear machine theory which will incorporate state space and formal polynomial concepts simultaneously in a unified framework without resorting to any operational transform technique.

#### 3.1. Mathematical Description of Finite State Sequential Machines

Definition 3.1.1. (cf. [60]) A deterministic sequential machine  $M$  is a composite mathematical concept specified by an octuple  $M = (K, X, U, Y, U^*, Y^*, \phi, \eta)$ , where

- (1)  $K$  is the time (clock period) set which is the ordered Abelian group of integers.
- (2)  $X$  is the state set.
- (3)  $U$  is the set of input symbols.

(4)  $U^* \equiv \{u : K \rightarrow U\}$  is the set of all admissible input maps, that is, sequences  $\dots u(-1) u(0) u(1) \dots$ ;  $u(k) \in U$ , and satisfies the following conditions:

- (a) (Nontriviality).  $U^*$  is nonempty.
- (b) (Concatenation of inputs). An input string  $u(k_1) u(k_2) \dots u(k_i)$  is a  $u \in U^*$  restricted to  $\{k_1, k_2, \dots, k_i\} \cap K$ . If  $u, \hat{u} \in U^*$ , then there exists a  $u'' \in U^*$  such that  $u''(k_1)u''(k_2) \dots u''(k_r) = u(k_1)u(k_2) \dots u(k_r)$  and  $u''(k_{r+1})u''(k_{r+2}) \dots u''(k_s) = \hat{u}(k_{r+1})\hat{u}(k_{r+2}) \dots \hat{u}(k_s)$ , where  $k_1 < k_r < k_s, k_i \in K \forall i$ .

(5)  $V$  is the set of output symbols.

(6)  $V^* \equiv \{y : K \rightarrow V\}$  is the set of output maps, that is, sequences  $\dots y(-1) y(0) y(1) \dots$ ;  $y(k) \in V$ .

(7)  $\phi$  is the state transition map  $\phi : K \times K \times X \times U^* \rightarrow X$  whose value is  $x(k) = \phi(k, k_0, x(k_0), u) \in X$  resulting at clock period  $k \in K$  from the initial state  $x(k_0) \in X$  at initial clock period  $k_0 \in K$  under the action of the input sequence  $u \in U^*$ .  $\phi$  has the following properties:

- (a) (Direction of time).  $\phi$  is defined for all  $k \geq k_0$ , but not necessarily for all  $k < k_0$ ;  $k, k_0 \in K$ .
- (b) (Consistency).  $\phi(k, k, x, u) = x \forall k \in K, \forall x \in X$ , and  $\forall u \in U^*$ .

(c) (Group property). For any  $k_1, k_2, k_3 \in K$  such that

$$k_1 < k_2 < k_3, \text{ we have } \phi(k_3, k_1, x, u) =$$

$$\phi(k_3, k_2, \phi(k_2, k_1, x, u), u) \quad \forall x \in X \text{ and } \forall u \in U^*.$$

(d) (Causality).  $u, u' \in U^*, u(k_1) u(k_2) \dots u(k_r) =$

$$u'(k_1) u'(k_2) \dots u'(k_r) \implies \phi(k_1, k_0, x, u) =$$

$$\phi(k_1, k_0, x, u').$$

(8)  $\eta$  is the output (readout) map  $\eta : K \times X \times U^* \longrightarrow Y$  which

defines the output value  $y(k) = \eta(k, x(k), u)$  in state

$x(k) \in X$  at clock period  $k \in K$ . The map  $\{k_0, k_1, \dots, k_r\} \longrightarrow Y$  given by  $\bar{k} \longmapsto \eta(\bar{k}, \phi(\bar{k}, k_0, x, u), u)$ , is an

output string, that is, the restriction  $y(k_0) y(k_1) \dots$

$y(k_r)$  of some  $y \in Y^*$  to  $\{k_0, k_1, \dots, k_r\}$ .

Definition 3.1.2. A sequential machine  $M = (K, X, U, Y, U^*, Y^*,$

$\phi, \eta)$  is *time-invariant* if and only if

(a)  $U^*$  is closed under the shift operator  $\nabla^{k'} : u \longrightarrow u'$

defined by  $u'(k) \equiv u(k + k') \quad \forall k, k' \in K$  and  $\forall u, u' \in U^*$ .

(b)  $\phi(k, k', x, u) = \phi(k + \ell, k' + \ell, x, \nabla^{-\ell} u) \quad \forall \ell \in K$ .

(c) The map  $\eta(k, \cdot, \cdot) : X \times U^* \longrightarrow Y$  is independent of  $k$ .

From the above definition it follows that for time-invariant machines the state transition and output maps assume the following simpler forms:

$$\overset{\vee}{\phi} : K \times X \times U^* \longrightarrow X, (k_1, \tilde{x}, u) \longmapsto \phi(k_1, 0, \tilde{x}, u)$$

$$\overset{\vee}{\eta} : X \times U^* \longrightarrow Y, (\tilde{x}, u) \longmapsto \eta(0, \tilde{x}, u)$$

since we know that for all choices of  $k_0 \in K$ ,  $\phi(k_1, k_0, \tilde{x}, u) =$

$$\overset{\vee}{\phi}(k_1 - k_0, \tilde{x}, \overset{k_0}{\nabla} u) \text{ and } \eta(k_1, \tilde{x}, u) = \overset{\vee}{\eta}(\tilde{x}, u).$$

Definition 3.1.3. A sequential machine  $M = (K, X, U, Y, U^*, V^*, \phi, \eta)$  is *finite-dimensional* if and only if  $X$  is a finite-dimensional linear space;  $M$  is *finite-state* if and only if  $X$  is a finite set.

Definition 3.1.4. A sequential machine  $M = (K, X, U, Y, U^*, V^*, \phi, \eta)$  is *linear* if and only if

- (a)  $X, U, U^*, Y,$  and  $V^*$  are vector spaces (over a given arbitrary field  $F$ ).
- (b) The map  $\phi(k, k_0, \cdot, \cdot) : X \times U^* \longrightarrow X$  is an  $F$ -homomorphism for all  $k, k_0 \in K$ .
- (c) The map  $\eta(k, \cdot, \cdot) : X \times U^* \longrightarrow Y$  is an  $F$ -homomorphism for all  $k \in K$ .

A special class of sequential machines will constitute the central subject of our investigation. The members of this class are assumed to be deterministic, finite-dimensional, finite-state, time-invariant, and linear. To give a precise description of this class of sequential machines, we will formally transliterate the preceding qualifications into the language of Definitions 3.1.1 - 3.1.4 as follows:

$K$  = time (clock period) set = set of integers;

$X$  = state space =  $\text{GF}(q)^n$  = finite vector space of  $n$ -tuples over the Galois field  $\text{GF}(q)$ ;

$U$  = set of input values =  $\text{GF}(q)^m$ ;

$U^*$  = input space = set of arbitrary maps  $u : K \longrightarrow U$ , that is, arbitrary sequences  $\dots u(-1) u(0) u(1) \dots, u(k) \in U$ ;

$Y$  = set of output values =  $\text{GF}(q)^r$ ;

$V^*$  = output space = set of arbitrary maps  $y : K \longrightarrow Y$ ;

$\phi$  = state transition map  $K \times X \times U^* \longrightarrow X$  given by  $(k + 1, k, x, u) \longmapsto \phi(k + 1, k, x, u) = Ax(k) + Bu(k)$ , where  $A$  and  $B$  are  $GF(q)$ -homomorphisms :  $A : X \longrightarrow X$ ,  $B : U \longrightarrow X$ ;  
 $\eta$  = readout map  $X \times U^* \longrightarrow Y$  given by  $(x, u) \longmapsto \eta(x, u) = Cx(k) + Du(k)$ , where  $C$  and  $D$  are  $GF(q)$ -homomorphisms :  $C : X \longrightarrow Y$ ,  $D : U \longrightarrow Y$ .

We will usually not make a distinction between  $(A, B, C, D)$  as a quadruple of  $GF(q)$ -homomorphisms or as a quadruple of matrices over  $GF(q)$  representing these homomorphisms with respect to a given basis of the underlying finite vector space over  $GF(q)$ .

For the purpose of future reference, we will summarize the above conventions in the following definition.

Definition 3.1.5. A deterministic, linear, time-invariant, finite-state,  $n$ -state,  $m$ -input,  $r$ -output sequential machine is a dynamical object whose behavior evolves according to the vector difference equations

$$x(k + 1) = Ax(k) + Bu(k) \quad (3.1.1a)$$

$$y(k) = Cx(k) + Du(k) \quad (3.1.1b)$$

where at clock period  $k$ ,  $x(k) \in GF(q)^n$  is the state,  $u(k) \in GF(q)^m$  is the input, and  $y(k) \in GF(q)^r$  is the output of the machine. Moreover,  $A \in GF(q)^{n \times n}$ ,  $B \in GF(q)^{n \times m}$ ,  $C \in GF(q)^{r \times n}$ , and  $D \in GF(q)^{r \times m}$ .

The defining equations (3.1.1) can be equivalently represented in component form as



$$x_i(k+1) = \sum_{j=1}^n a_{ij} x_j(k) + \sum_{j=1}^m b_{ij} u_j(k), i \in \underline{n} \quad (3.1.2a)$$

$$y_\ell(k) = \sum_{j=1}^n c_{\ell j} x_j(k) + \sum_{j=1}^m d_{\ell j} u_j(k), \ell \in \underline{r} \quad (3.1.2b)$$

where  $a_{ij}, b_{is}, c_{\ell j}, d_{\ell j} \in GF(q)$ ,  $i, j \in \underline{n}$ ,  $s \in \underline{m}$ ,  $\ell \in \underline{r}$ , are elements of the matrices  $A, B, C$ , and  $D$ , respectively.

Since in the sequel we will be concerned exclusively with a machine of the type (3.1.1), for the sake of linguistic and notational simplicity it will be referred to as a linear sequential machine (LSM) - other qualifications being understood and generally not explicitly mentioned - and denoted by  $(A, B, C, D)$ .

### 3.2. Interconvertibility of Mealy and Moore LSMs

In LSM (3.1.1) we observe that the current output depends on both the current state and the current input of the machine. This type of LSM is called a *Mealy machine*. On the other hand, if in (3.1.1) the matrix  $D = 0$ , that is, if the LSM is described by the equations

$$x(k+1) = Ax(k) + Bu(k) \quad (3.2.1a)$$

$$y(k) = Cx(k) \quad (3.2.1b)$$

then it is called a *Moore machine* which is a state-output device whose current output depends only on its current state. Having lost the ability to consult the input in determining the output, it might seem that a Moore LSM is more limited than a Mealy LSM. However, it can be shown [12] that any Mealy machine can be simulated by a state-output

machine of the Moore type and vice versa. This interconvertibility property is true for any general Mealy and Moore machines. To see the conversion procedure for the case of LSMs, consider a Mealy LSM  $M = (A, B, C, D)$ , and let

$$\tilde{x}(k) \equiv \begin{Bmatrix} y(k-1) \\ x(k) \end{Bmatrix}, \quad \tilde{y}(k) \equiv y(k-1), \quad \text{and} \quad \tilde{u}(k) \equiv u(k)$$

Then the LSM  $\tilde{M} = (\tilde{A}, \tilde{B}, \tilde{C})$ , where

$$\tilde{A} \equiv \begin{Bmatrix} 0 & C \\ 0 & A \end{Bmatrix}, \quad \tilde{B} \equiv \begin{Bmatrix} D \\ B \end{Bmatrix}, \quad \text{and} \quad \tilde{C} \equiv [I_r \ 0]$$

is of the Moore type. Comparing the LSMs  $M$  and  $\tilde{M}$ , we notice that  $\tilde{M}$  has more states than  $M$ , and will always be one clock period behind  $M$ . In other words, to each state  $x$  of  $M$  there corresponds a state  $\tilde{x}$  of  $\tilde{M}$  such that the string of outputs that results by feeding a given string of inputs into  $M$  started in state  $x$  and into  $\tilde{M}$  started in state  $\tilde{x}$  will be just the same, except for a unit delay in the output of  $\tilde{M}$ .

In a similar manner, a Moore machine  $\bar{M} = (\bar{A}, \bar{B}, \bar{C})$  can be converted to a Mealy machine  $\hat{M} = (\hat{A}, \hat{B}, \hat{C}, \hat{D})$  by defining  $\hat{x}(k) \equiv \bar{x}(k)$ ,  $\hat{y}(k) \equiv \bar{y}(k+1)$ ,  $\hat{u}(k) \equiv \bar{u}(k)$ ,  $\hat{A} \equiv \bar{A}$ ,  $\hat{B} \equiv \bar{B}$ ,  $\hat{C} \equiv \bar{C}\bar{A}$ , and  $\hat{D} \equiv \bar{C}\bar{B}$ .

From the above observations it is clear that there will be no loss of generality if we consider only Moore LSMs. Therefore, in the sequel we will study exclusively LSMs of the Moore type since our results, if desired, can be readily restated for LSMs of the Mealy type.

In the sequel, we will have occasion to look specifically at a single-input single-output Moore LSM which results from (3.2.1) when  $b \in GF(q)^n$ ,  $u(k) \in GF(q)$ ,  $c \in GF(q)^n$ , and  $y(k) \in GF(q)$ , and has the form

$$\begin{aligned} x(k+1) &= Ax(k) + bu(k) \\ y(k) &= c^T x(k) \end{aligned} \tag{3.2.2}$$

### 3.3. Input-State and Input-Output Transfer Characteristics of LSMs

Given an initial state and an input sequence, the corresponding state and output sequences of an LSM can be computed recursively from equations (3.2.1a) and (3.2.1b), respectively. To see this, let  $x(0)$  denote the initial state of the LSM at  $k = 0$ . Then applying equation (3.2.1a) recursively, we obtain

$$\begin{aligned} x(1) &= Ax(0) + Bu(0) \\ x(2) &= A^2x(0) + ABu(0) + Bu(1) \\ &\vdots \\ x(k) &= A^kx(0) + \sum_{j=0}^{k-1} A^{k-j-1} Bu(j) \end{aligned} \tag{3.3.1}$$

In view of equation (3.2.1b), the output is given by

$$y(k) = CA^kx(0) + \sum_{j=0}^{k-1} CA^{k-j-1} Bu(j) \tag{3.3.2}$$

Equation (3.3.2) is a general expression for the response of the LSM  $M = (A, B, C)$ , and is composed of two distinct parts: the term  $CA^kx(0)$

is the *autonomous (zero-input) response* and the convolution sum  $\sum_{j=0}^{k-1} CA^{k-j-1} Bu(j)$  is the *signal (zero-state) response* of the LSM.

From the form of the signal response it is clear that the convolution factor  $CA^{k-1}B$  is the *weighting sequence* (Kronecker delta response) of the LSM.

In conjunction with the state transition and output maps, equations (3.3.1) and (3.3.2) can be equivalently written as follows:

$$\phi(x(0), u(0) \ u(1) \ . \ . \ . \ u(\ell-1)) = [A^\ell \ A^{\ell-1}B \ . \ . \ . \ AB \ B] \begin{pmatrix} x(0) \\ u(0) \\ u(1) \\ \vdots \\ u(\ell-1) \end{pmatrix} \quad (3.3.3)$$

$$\eta(x(0), u(0) \ u(1) \ . \ . \ . \ u(\ell-1)) = \begin{pmatrix} C & 0 & 0 & . & . & . & 0 \\ CA & CB & 0 & . & . & . & 0 \\ \vdots & \vdots & \vdots & & & & \vdots \\ CA^{\ell-2} & CA^{\ell-3}B & CA^{\ell-4}B & . & . & . & 0 \\ CA^{\ell-1} & CA^{\ell-2}B & CA^{\ell-3}B & . & . & . & CB \end{pmatrix} \begin{pmatrix} x(0) \\ u(0) \\ \vdots \\ u(\ell-2) \\ u(\ell-1) \end{pmatrix} \quad (3.3.4)$$

Some additional relationships among input, state, and output of an LSM are given in the following theorem. The verification of these relationships is straightforward and hence omitted.

**Theorem 3.3.1.** For each  $x, x' \in X$ ,  $u \in U^*$ ,  $d \in GF(q)$

- (a)  $\phi(x + dx', u) = \phi(x, u) + d\phi(x', 0^{\lg(u)})$
- (b)  $\eta(x + dx', u) = \eta(x, u) + d\eta(x', 0^{\lg(u)})$
- (c)  $\eta(x, u) = \eta(x, 0^{\lg(u)}) + \eta(0, u)$
- (d) Let  $u, u' \in U^*$  such that  $\lg(u) = \lg(u')$ . Then
 
$$\phi(x, u) = \phi(x, u') \iff \phi(0, u) = \phi(0, u')$$

- (e)  $\phi(x, u) = \phi(x', u) \iff \phi(x, 0^{\lg(u)}) = \phi(x', 0^{\lg(u)})$   
 (f)  $\eta(x, u) = \eta(x', u) \iff \eta(x, 0^{\lg(u)}) = \eta(x', 0^{\lg(u)})$ .

### 3.4. Indistinguishability, Isomorphism, Minimality, and Similarity in LSMs

Definition 3.4.1. Let  $M$  and  $\tilde{M}$  be LSMs. The states  $x$  of  $M$  and  $\tilde{x}$  of  $\tilde{M}$  are said to be  $\ell$ -indistinguishable, and written  $x \overset{\ell}{\sim} \tilde{x}$ , if and only if  $\eta(x, w) = \tilde{\eta}(\tilde{x}, w)$  for all input sequences of length at most  $\ell$ , where  $\eta$  and  $\tilde{\eta}$  are the output maps of  $M$  and  $\tilde{M}$ , respectively. The states  $x$  and  $\tilde{x}$  are said to be indistinguishable, and written  $x \sim \tilde{x}$ , if and only if  $\eta(x, w) = \tilde{\eta}(\tilde{x}, w)$  for all input sequences  $w$ ; otherwise  $x$  and  $\tilde{x}$  are said to be distinguishable.  $M$  and  $\tilde{M}$  may refer to the same LSM.

From the above definition it is clear that for all  $\ell \leq \ell_1$ ,  $x \overset{\ell_1}{\sim} \tilde{x}$  implies that  $x \overset{\ell}{\sim} \tilde{x}$ . The relation  $\sim$  is clearly an equivalence relation on  $X \cup \tilde{X}$ , where  $X$  and  $\tilde{X}$  are the state sets of the LSMs  $M$  and  $\tilde{M}$ , respectively.

Let  $x^1$  and  $x^2$  be two arbitrary states of the LSM  $M$ . Then from part (c) of Theorem 3.3.1 it follows that

$$\eta(x^1, u(0) u(1) \dots u(\ell-1)) = \eta(x^1, 0^\ell) + \eta(0, u(0) u(1) \dots u(\ell-1))$$

and

$$\eta(x^2, u(0) u(1) \dots u(\ell-1)) = \eta(x^2, 0^\ell) + \eta(0, u(0) u(1) \dots u(\ell-1))$$

which clearly show that  $x^1 \overset{\ell}{\sim} x^2$  if and only if

$$\eta(x^1, 0^\ell) = \eta(x^2, 0^\ell) \tag{3.4.1}$$

Therefore, we have the following result.

Theorem 3.4.1. Two states of a given LSM are  $\ell$ -indistinguishable if and only if they yield the same response to all zero input sequences of length at most  $\ell$ .

Using the expression (3.3.4), (3.4.1) becomes

$$Lx^1 = Lx^2 \quad (3.4.2)$$

where

$$L \equiv \begin{pmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^{\ell-1} \end{pmatrix} \quad (3.4.3)$$

From (3.4.2) it follows that

$$CA^{j-1}(x^1 - x^2) = 0_Y = CA^{j-1}0_X, \quad j \in \underline{\ell}$$

which implies that two states  $x^1$  and  $x^2$  of an LSM  $M$  are distinguishable if and only if their difference  $(x^1 - x^2)$  is distinguishable from  $0_X$ .

Clearly (3.4.2) is equivalent to  $L(x^1 - x^2) = 0$  which implies that

$$x^1 - x^2 \in N(L) \quad (3.4.4)$$

Therefore, we have proved the following indistinguishability criterion.

Theorem 3.4.2. Two states of a given LSM  $M = (A, B, C)$  are  $\ell$ -indistinguishable if and only if their difference is in the null space of the linear map  $L$  given by (3.4.3).

From (3.4.4) it follows that the set  $E_0$  of all states which are indistinguishable from the zero state is the null space of  $L$ , that is,  $E_0 \equiv \{x \in X : x \sim 0\} = N(L) \subseteq X$ . The set  $E_0$  is clearly a subgroup of the additive Abelian group  $X$  and induces a coset partition on  $X$  which can be given the following characterization:

Theorem 3.4.3. The cosets of the additive Abelian group  $X$  induced by the subgroup  $E_0 \equiv \{x \in X : x \sim 0\}$  are the equivalence classes of  $X$ .

Proof. Two states  $x^1, x^2 \in X$  belong to the same coset if and only if  $x^1 - x^2 \in E_0$ , hence if and only if  $x^1 - x^2 \in N(L)$ , hence if and only if  $L(x^1 - x^2) = 0$ , hence if and only if  $Lx^1 = Lx^2$ , hence by (3.4.4), if and only if  $x^1 \sim x^2$ .

Theorem 3.4.4. Let  $M$  and  $\tilde{M}$  be LSMs. If states  $x$  of  $M$  and  $\tilde{x}$  of  $\tilde{M}$  are indistinguishable, then  $\phi(x, u) = \tilde{\phi}(\tilde{x}, u)$  for all  $u \in U^*$ .

Proof. Let  $u \in U^*$ . Then  $\eta(\phi(x, u'), u) = \eta(x, u'u) = \tilde{\eta}(x, u'u) = \eta(\tilde{\phi}(x, u'), u)$ . Hence  $\phi(x, u) = \tilde{\phi}(\tilde{x}, u)$ .

Definition 3.4.2. Let  $M = (A, B, C)$  and  $\tilde{M} = (\tilde{A}, \tilde{B}, \tilde{C})$  be LSMs with state spaces  $X$  and  $\tilde{X}$ , respectively. A map  $\alpha : X \rightarrow \tilde{X}$  is said to be a *homomorphism* from  $M$  into  $\tilde{M}$  if  $\alpha(\phi(x, u)) = \tilde{\phi}(\alpha(x), u)$  and  $\alpha(\eta(x, u)) = \tilde{\eta}(\alpha(x), u)$  for all  $(x, u) \in X \times U^*$ . If such a map exists, then  $\tilde{M}$  is said to be a *homomorphic image* of  $M$ . Furthermore,  $M$  is said to be *isomorphic* to  $\tilde{M}$  if there exists an isomorphism of  $M$  onto  $\tilde{M}$ , that is, if a one-to-one relationship can be established between  $X$  and  $\tilde{X}$  in the following manner: If a state  $x$  of  $M$  corresponds to a state  $\tilde{x}$  of  $\tilde{M}$ , then for every input  $u$ ,  $Cx = \tilde{C}\tilde{x}$  and the state  $\hat{x} = Ax + Bu$  in  $M$  corresponds to the state  $\hat{\tilde{x}} = \tilde{A}\tilde{x} + \tilde{B}u$  in  $\tilde{M}$ .

Thus, if  $M$  and  $\tilde{M}$  are isomorphic, their state-output graphs are identical except, possibly, for vertex labeling. Clearly isomorphic LSMs are indistinguishable but not conversely.

Definition 3.4.3. An LSM  $M$  is said to be *minimal* if and only if  $x^1 \sim x^2 \implies x^1 = x^2 \ \forall x^1, x^2 \in X$ .

From the definitions of indistinguishability, isomorphism, and minimality, the following results are immediate.

Theorem 3.4.5. Let the LSM  $M$  be indistinguishable from a minimal LSM  $\tilde{M}$  of dimension  $r$ . Then no LSM indistinguishable from  $M$  has dimension smaller than  $r$ .

Theorem 3.4.6. If  $M$  and  $\tilde{M}$  are indistinguishable and minimal LSMs, then they are isomorphic.

Definition 3.4.4. The LSM  $M = (A, B, C)$  is said to be *similar* to the LSM  $\tilde{M} = (\tilde{A}, \tilde{B}, \tilde{C})$  if there exists a nonsingular matrix  $P$  such that  $\tilde{A} = PAP^{-1}$ ,  $\tilde{B} = PB$ , and  $\tilde{C} = CP^{-1}$ .

Theorem 3.4.7. If the LSM  $M = (A, B, C)$  is similar to the LSM  $\tilde{M} = (\tilde{A}, \tilde{B}, \tilde{C})$ , then

- (a)  $x \sim Px \ \forall x \in X$
- (b)  $M$  is isomorphic to  $\tilde{M}$
- (c)  $M \sim \tilde{M}$

Proof. To show part (a), let  $u(0) u(1) \dots u(\ell-1) \in U^*$  and  $x \in X$ . Then

$$\begin{aligned} \tilde{\eta}(Px, u(0)u(1) \dots u(\ell-1)) &= \tilde{C}\tilde{A}^\ell Px + \sum_{j=0}^{\ell-1} \tilde{C}\tilde{A}^{\ell-j-1} \tilde{B}u(j) \\ &= CP^{-1}(PAP^{-1})^\ell Px + \sum_{j=0}^{\ell-1} CP^{-1}(PAP^{-1})^{\ell-j-1} PBu(j) \end{aligned}$$



Since  $(PAP^{-1})^n = PA^n P^{-1}$  for each positive integer  $n$ , we have

$$\begin{aligned}\tilde{\eta}(Px, u(0)u(1) \dots u(\ell-1)) &= CA^\ell x + \sum_{j=0}^{\ell-1} CA^{\ell-j-1} Bu(j) \\ &= \eta(x, u(0)u(1) \dots u(\ell-1))\end{aligned}$$

Thus  $x \sim Px$ .

To show (b), let  $\alpha(x) \equiv Px$ . Then  $\alpha : X \longrightarrow X$  is one-to-one and onto since  $P$  is invertible. By (a)

$$\begin{aligned}\eta(x, u) &= \tilde{\eta}(\alpha(x), u) \quad \forall u \in U^* \\ \alpha(\phi(x, u')) &= \alpha(Ax + Bu') = P(Ax + Bu') \\ &= (PAP^{-1})Px + PBu' \\ &= \tilde{A}\alpha(x) + \tilde{B}u' \\ &= \tilde{\phi}(\alpha(x), u')\end{aligned}$$

Thus  $M$  is isomorphic to  $\tilde{M}$  and hence  $M \sim \tilde{M}$ .

The relationships among isomorphism, indistinguishability, and similarity, as applied to minimal and nonminimal LSMs, are summarized in the following implication diagram:

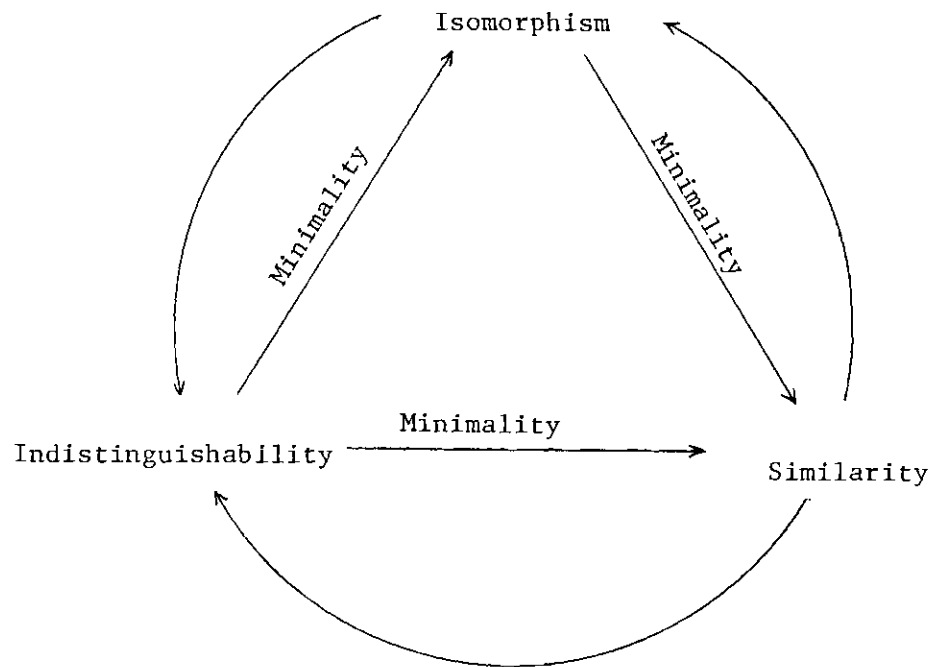


Fig. 3.4.1. Isomorphism, indistinguishability, and similarity relations for minimal and non-minimal LSMs.

### 3.5. Input-Output Representation of LSMs

The description of an LSM given by equations (3.1.1) is essentially an internal description in the sense that the operational structure of the LSM in terms of the evolution of the state set is completely specified, and the output of the LSM is generated indirectly via a transformation of the state. However, there are many situations in which the internal structure of the LSM is not available and hence the only access to the machine is by means of the input terminals and output terminals. In such cases, the input-output behavior of the machine can be abstracted from a collection of input-output pairs obtained by

feeding strings of inputs to the "black box" and observing the corresponding strings of outputs. It is clear that even then the description of the machine is not entirely free of the state since the response of the machine to a string of inputs depends on the state of the machine at the beginning of the period of observation and, in the case of time-varying machines, on the time at which the observations begin.

In order to give a precise input-output description of a machine, we need to introduce a response map to serve as the link between the inputs and the corresponding outputs of the machine. Initially, we will consider the general case. Let

$$\rho_{k_0, x^0} : K \times U^* \longrightarrow Y, (k, u) \longmapsto \eta(k, \phi(k, k_0, x^0, u)) \quad (3.5.1)$$

be the response map of the machine  $M = (K, X, U, Y, U^*, Y^*, \phi, \eta)$ .

According to this input-output correspondence, the machine is started in state  $x^0$  at clock period  $k_0$ , an admissible input sequence is applied to obtain state  $\phi(k, k_0, x^0, u)$  at clock period  $k$ , and then the map  $\eta$  is applied to determine the corresponding output at clock period  $k$ .

Since for time-invariant machines (Definition 3.1.2.)

$$\phi : K \times X \times U^* \longrightarrow X, (k, x^0, u) \longmapsto \phi(k, 0, x^0, u)$$

$$\eta : X \times U^* \longrightarrow Y, (x^0, u) \longmapsto \eta(0, x^0, u)$$

and

$$\rho_{k_0+l, x^0}(k+l, \nabla^{-l}u) = \rho_{k_0, x^0}(k, u) \quad (3.5.2)$$

where the shift operator  $\nabla^{k'} : u \longrightarrow u'$  is defined by  $u'(k) \equiv u(k + k')$   $\forall k, k' \in K$  and  $\forall u, u' \in U^*$ , (3.5.1) can be simply written as

$$\rho_{x^0} : K \times U^* \longrightarrow Y, (k, u) \longmapsto \eta(\phi(k, x^0, u)) \quad (3.5.3)$$

without explicitly indicating  $k_0$ . That is, the initial clock period can always be taken to be zero. Of course,  $\rho_{k_0, x^0}$  is recoverable from  $\rho_{x^0}$  for any  $x^0 \in X$  by the equation

$$\rho_{k_0, x^0}(k, u) = \rho_{x^0}(k - k_0, \nabla^{k_0} u)$$

obtained from (3.5.2) by setting  $\ell = -k_0$ .

Let  $0_U \in U$  denote the zero input and  $0_Y \in Y$  the zero output, and define the *zero input map*  $0_{U^*} \in U^*$  by the equation  $0_{U^*}(k) = 0_U \forall k \in K$ . Then a state  $x^0 \in X$  is called a *zero state* of the machine  $M$  whenever  $\rho_{k_0, x^0}(k, 0_{U^*}) = 0_Y \forall k \geq k_0; k, k_0 \in K$ . Since for a machine  $M$  it is possible to have many distinguishable zero states different from the additive zero  $0_X$  of the state space, we will assume that  $M$  has at least one zero state and denote it by  $x^\theta$ . If  $X$  has an additive zero  $0$ , then the maps  $\phi$  and  $\eta$  can be set up so as to allow the choice  $x^\theta = 0$ .

If we fix upon  $0_U, 0_{U^*}, 0_Y$ , and  $x^\theta$ , the map

$$\rho_{k_0, x^\theta} : K \times U^* \longrightarrow Y \quad (3.5.4)$$

is called the *zero-state response* of  $M$  started at  $k_0$ , and the map

$$\rho_{k_0} : K \times X \longrightarrow Y, (k, x^0) \longmapsto \rho_{k_0, x^0}(k, 0_{U^*}) \quad (3.5.5)$$

is called the *zero-input response* of  $M$  started at  $k_0$ .

In fact the above particular response maps can be regarded as special cases of an overall response map of  $M$  defined by

$$\rho : K \times K \times X \times U^* \longrightarrow Y, (k, k_0, x^0, u) \longmapsto \eta(k, \phi(k, k_0, x^0, u))$$

since (3.5.1), (3.5.4), and (3.5.5) may be written as

$$\rho_{k_0, x^0} = \rho(\cdot, k_0, x^0, \cdot) : K \times U^* \longrightarrow Y$$

$$\rho_{k_0, x^\theta} = \rho(\cdot, k_0, x^\theta, \cdot) : K \times U^* \longrightarrow Y$$

and

$$\rho_{k_0} = \rho(\cdot, k_0, \cdot, 0_{U^*}) : K \times X \longrightarrow Y$$

respectively. Moreover, if we specify two clock periods  $k_0$  and  $k_1$  in  $K$ , then we see that

$$\rho(k_1, k_0, \cdot, \cdot) = \eta(k_1, \cdot) \circ \phi(k_1, k_0, \cdot, \cdot)$$

Now if we consider an LSM  $(A, B, C)$ , then from the preceding discussion and Section 3.3 it follows that

$$y(\ell) = \rho_{x^0}^0(u(0)u(1) \dots u(\ell-1)) = CA^\ell x^0 + \sum_{j=0}^{\ell-1} CA^{\ell-j-1} Bu(j)$$

where  $\rho_{\mathbf{x}0}(0^\ell) = \mathbf{CA}^\ell \mathbf{x}^0$  is the zero-input response and  $\rho_{0\chi}(u(0)u(1) \dots u(\ell-1)) = \sum_{j=0}^{\ell-1} \mathbf{CA}^{\ell-j-1} \mathbf{B}u(j)$  is the zero-state response of the LSM. That is,

$$\rho_{\mathbf{x}0}(u(0)u(1) \dots u(\ell-1)) = \rho_{\mathbf{x}0}(0^\ell) + \rho_{0\chi}(u(0)u(1) \dots u(\ell-1))$$

Consider the special unit-pulse sequence defined by

$$\delta(k) \equiv 1, k = 0$$

$$\equiv 0, \text{ elsewhere}$$

It is easily seen that any arbitrary sequence  $u(0)u(1) \dots u(\ell-1)$  can be expressed as a weighted sum of  $\delta(k)$ , that is,

$$u(k) = \sum_{j=0}^k \delta(k-j) u(j), k = 0, 1, \dots, \ell-1 \quad (3.5.6)$$

In view of the properties of linearity and time-invariance of an LSM, the zero-state response to the input sequence given by (3.5.6) may be written as

$$\rho_{0\chi}(u(0)u(1) \dots u(\ell-1)) = \sum_{j=0}^{\ell-1} g(\ell-j) u(j) \quad (3.5.7)$$

where  $g(\ell)$  is the unit-pulse response of the LSM, that is,

$$g(\ell) \equiv \rho_{0\chi}(100 \dots 0) = \begin{cases} \mathbf{CA}^{\ell-1} \mathbf{B}, & \ell > 0 \\ 0 & \ell = 0 \end{cases} \quad (3.5.8)$$

This result implies that the unit-pulse response completely specifies the input-output behavior of an LSM started at state  $0_\chi$  since knowing

$g$ , the response of the LSM to any arbitrary input sequence is just the convolution sum given by (3.5.7). Therefore, the essential input-output properties of an LSM with zero initial state are captured in the special response sequence  $g$  which can be regarded as the input-output operational model of the LSM.

In the next section we will see that the simple notion of the unit-pulse response sequence makes it possible to describe the input-output behavior of an LSM in terms of polynomials over  $GF(q)$ . This possibility seems to open up new potential avenues of research in various aspects of LSMs.

### 3.6. Polynomial Representation of LSMs

If a polynomial  $f(\xi)$  in the indeterminate  $\xi$  is regarded as a special algebraic object, then, in conjunction with formal power series and realizable rational functions over  $GF(q)$ , it is possible to describe the input-output behavior of an LSM in terms of polynomial matrices, that is, matrices with polynomial elements. We will show that using this particular external description of LSMs, a link can be established between the state variable and the input-output representations without the use of any transform techniques. First, we will review the polynomial representation of sequences over  $GF(q)$ .

Consider the set  $S$  of all infinite sequences over  $GF(q)$

$$S \equiv \{ \{s_1, s_2, \dots, s_r, \dots\} : s_i \in GF(q) \} \quad (3.6.1)$$

where only a finite number of the entries is nonzero, and let  $f, g \in S$ . Then  $f$  and  $g$  can be expressed as

$$f = \{a_0, a_1, \dots, a_m, \dots\}, a_i \in GF(q)$$

$$g = \{b_0, b_1, \dots, b_n, \dots\}, b_i \in GF(q)$$

$f = g$  if and only if  $a_i = b_i$ ,  $i = 0, 1, 2, \dots$ . If we define addition of two sequences  $f$  and  $g$  as

$$f + g = \{a_0 + b_0, a_1 + b_1, \dots\} \quad (3.6.2)$$

multiplication of a sequence  $f$  by a scalar  $c$  as

$$cf = \{ca_0, ca_1, \dots, ca_m, \dots\} \quad (3.6.3)$$

and multiplication of two sequences  $f$  and  $g$  as

$$fg = \{c_0, c_1, \dots\} \quad (3.6.4)$$

where  $c_i = \sum_{j=0}^i a_j b_{i-j}$ ,  $i = 0, 1, 2, \dots$ , then it follows that the set  $S$  becomes a commutative ring.

Let  $\xi \equiv \{0, 1, 0, \dots\}$ , that is, the second term is 1, and all others are zero. Then from (3.6.4) it follows that  $(\xi)^2 = \{0, 0, 1, 0, \dots\}$ , and by an induction argument we get  $(\xi)^\ell = \{0, 0, \dots, 0, 1, 0, \dots\}$ , where the first  $\ell$  terms of  $(\xi)^\ell$  are 0, the  $(\ell+1)$ th is 1, and all later ones are 0.

Now consider any element  $f = \{a_0, a_1, \dots, a_m, 0, \dots\}$  of the ring  $S$ . Then in view of the definition of  $(\xi)^\ell$ , we can express  $f$  as



$$f = a_0\{1, 0, \dots\} + a_1\{0, 1, 0, \dots\} + \dots + a_m\{0, 0, \dots, 0, 1, 0, \dots\}$$

$$f = a_0 + a_1\xi + a_2(\xi)^2 + \dots + a_m(\xi)^m \quad (3.6.5)$$

This algebraic expression is called a *formal polynomial* in the indeterminate  $\xi$  and the set  $S = GF(q)[\xi]$  of all such formal polynomials, the *ring of formal polynomials* over  $GF(q)$ . It is clear that the units of this ring are polynomials of zero degree, where the degree of a polynomial  $f$  is defined to be the index of the leading nonzero coefficient, its only divisor of zero is 0, and its primes are polynomials irreducible in  $GF(q)[\xi]$ , where a nonconstant polynomial  $f$  is said to be *irreducible* if there do not exist polynomials  $f_1, f_2$  in  $GF(q)[\xi]$  such that  $f = f_1f_2$ .

Another algebraic object related to the above description of a polynomial is a *formal power series* over  $GF(q)$  in the indeterminate  $\xi$ , which is an infinite sequence

$$g \equiv \{a_0, a_1, a_2, \dots\}, a_i \in GF(q) \quad (3.6.6)$$

In view of the definition of  $(\xi)^k$ , it follows that (3.6.6) can be equivalently expressed as

$$g = a_0 + a_1\xi + a_2(\xi)^2 + \dots$$

If we define addition and multiplication for infinite sequences of the form (3.6.6) as in (3.6.2), (3.6.3) and (3.6.4), then it is easily seen that the set of all formal power series forms a ring, denoted by  $GF(q)[[\xi]]$ , which contains the ring of polynomials  $GF(q)[\xi]$  over  $GF(q)$  as a subring.

The set of realizable rational functions with elements of the form  $f = \frac{g}{h}$ ,  $\deg g \leq \deg h$ ;  $g, h \in S$ , or

$$f = c_0 + c_1(\xi)^{-1} + c_2(\xi)^{-2} + \dots + c_i \in GF(q)$$

also forms a ring and is denoted by  $GF(q)[(\xi)^{-1}]$ . The units of this ring are elements of order zero, its only divisor of zero is 0, and its only prime element is  $(\xi)^{-1}$ . Clearly  $GF(q)[(\xi)^{-1}]$  contains  $GF(q)[\xi]$  as a subring.

For our purposes it is extremely important to realize that a formal polynomial is just an algebraic object completely equivalent to a finite sequence of elements of  $GF(q)$ , and it is not a function of a complex variable. The uninterpreted symbol  $\xi$  plays the role of a "position marker." In fact,  $\xi$  can be interpreted as a linear mapping describing the dynamics in the context of LSMs. Now we are in a position to discuss the polynomial representation of LSMs. In Section 3.5 it was shown that the unit-pulse response given by

$$g(\ell) \equiv \rho_{0_X} (100 \dots 0) = \begin{cases} CA^{\ell-1}B, & \ell = 1, 2, \dots \\ 0, & \ell = 0 \end{cases}$$

completely specifies the input-output behavior of the LSM  $(A, B, C)$  started at  $0_X$ . Recalling the definition of formal polynomials, we can express  $g(\ell)$  as an element of  $GF(q)[\xi]$  as

$$\begin{aligned} g(\ell) &= CB\xi + CAB(\xi)^2 + CA^2B(\xi)^3 + \dots \\ &= C[\xi I_n + A(\xi)^2 + A^2(\xi)^3 + \dots]B \\ &= C\xi(I_n - \xi A)^{-1}B \end{aligned}$$

and as an element of  $GF(q)[(\xi)^{-1}]$  as

$$\begin{aligned}
 g(\ell) &= CB(\xi)^{-1} + CAB(\xi)^{-2} + CA^2B(\xi)^{-3} + \dots \\
 &= C[I_n(\xi)^{-1} + A(\xi)^{-2} + A^2(\xi)^{-3} + \dots]B \\
 &= C(\xi I_n - A)^{-1}B
 \end{aligned} \tag{3.6.7}$$

The expression (3.6.7) very clearly indicates the intimate relationship between the state variable and polynomial representations of LSMs and points out the possibility of merging these two approaches whose constructive interplay can result into a general and unified theory for LSMs. The expression (3.6.7) unmistakably resembles the transfer function matrix that represents the dynamical behavior of conventional linear systems in the frequency domain. However, it should be emphasized that in deriving (3.6.7) no operational transform technique was employed and it does not involve any complex variables, in direct contrast to the transfer function matrix of conventional linear systems which is obtained by using the Laplace transform in the continuous case, and the Z-transform in the discrete case and hence involves functions of a complex variable.

Considering (3.6.7) as a mapping

$$C(\xi I_n - A)^{-1}B : GF(q)[(\xi)^{-1}]^m \times 1 \longrightarrow GF(q)[(\xi)^{-1}]^r \times 1$$

the input-state and input-output pairs of the LSM (A, B, C) can be related as follows:

$$x(\xi) = (\xi I_n - A)^{-1} Bu(\xi)$$

and

$$\begin{aligned} y(\xi) &= C(\xi I_n - A)^{-1} Bu(\xi) \\ &= \frac{1}{\Delta(\xi)} \left\{ [(\xi)^{n-1} + a_{n-1}(\xi)^{n-2} + \dots + a_1] CBu(\xi) \right. \\ &\quad + [(\xi)^{n-2} + a_{n-1}(\xi)^{n-3} + \dots + a_2] CABu(\xi) \\ &\quad + \dots + (\xi + a_{n-1}) CA^{n-2} Bu(\xi) \\ &\quad \left. + CA^{n-1} Bu(\xi) \right\} \end{aligned}$$

where

$$\Delta(\xi) \equiv \det(\xi I_n - A) = (\xi)^n + a_{n-1}(\xi)^{n-1} + \dots + a_1\xi + a_0$$

is the characteristic polynomial of  $A$ .

From the above brief discussion it is evident that an extensive linear machine theory based on the theory of formal polynomial matrices over  $GF(q)[\xi]$  can be developed, without employing any operational transform techniques, which will parallel, in many respects, the works of Rosenbrock [97], Wolovich [110], and others in conventional linear time-invariant dynamical systems.

### Summary and Conclusions

We summarized in this chapter some essential definitions and properties of nonautonomous LSMs which will be needed in the sequel. Slightly more detailed discussions can be found in [14], [46], and [49].

In Section 3.6 a connection was established between the state variable and the input-output representations of LSMs by the use of formal polynomials over  $GF(q)$ . It was pointed out that this approach can be utilized to develop an extensive linear machine theory which will parallel, in many respects, the works of Rosenbrock [97], Wolovich [110], and others in conventional linear time-invariant dynamical systems. The idea underlying the formal polynomial representation was originally used by Kalman [61] in his module-theoretic investigation of linear systems.

## CHAPTER IV

## STATE REACHABILITY AND STATE CONTROLLABILITY OF LSMs

Our primary purpose in this chapter is to present a mathematically formal account of the pivotal concepts of state reachability and state controllability for LSMs. In order to avoid excessive clutter at the outset, we will keep the degree of detail to a minimum by relegating to the subsequent chapters the discussion of implications, extensions, consequences, and alternative formulations of the criteria developed in the present chapter.

4.1. State Reachability of LSMs

Definition 4.1.1. A state  $\tilde{x} \neq 0_X$  of the LSM  $M = (A, B, C)$  is said to be *reachable* from the state  $x^0 \in X$  if there exists an input sequence  $u \in U^*$  such that  $\phi(x^0, u) = \tilde{x}$ ; if  $\lg(u) = \ell$ ,  $\tilde{x}$  is said to be  *$\ell$ -reachable* from  $x^0$ ;  $M$  is said to be  *$\ell$ -state reachable* if every state of  $M$  is  $\ell$ -reachable from  $x^0$  for at least one particular  $\ell$ . The smallest such integer  $\ell_r$  is called the *state reachability index* of the LSM.

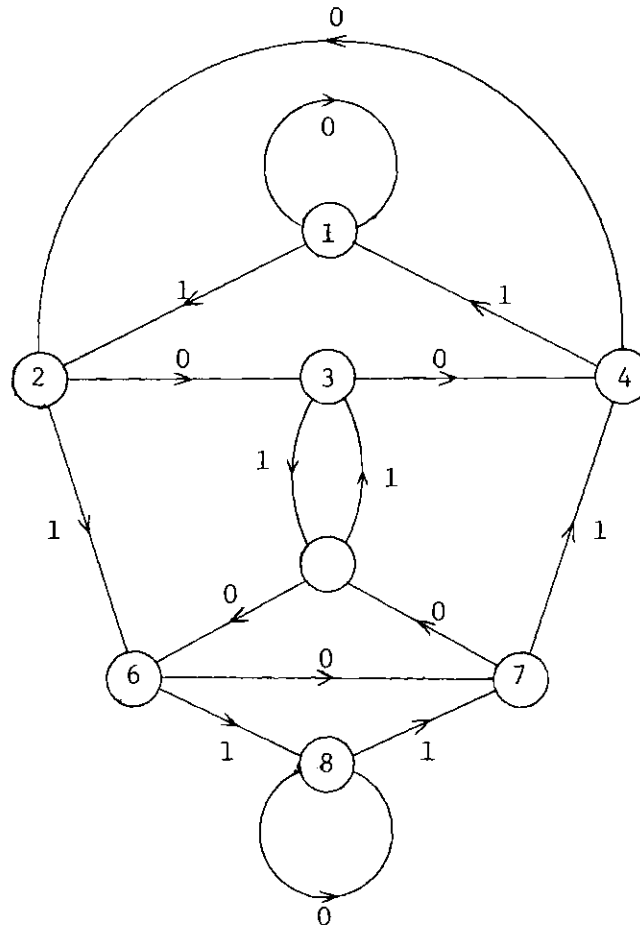
The above definition of state reachability may be simply rephrased as follows:

Definition 4.1.2. The LSM  $M = (A, B, C)$  is state reachable from the initial state  $x^0 \in X$  if and only if the state transition map  $\phi(x^0, \cdot) : U^* \longrightarrow X$  is an epimorphism.

Example 4.1.1. To illustrate the concept of state reachability, consider the state equation of a single-input LSM over GF(2) given by

$$\begin{bmatrix} x_1(k+1) \\ x_2(k+1) \\ x_3(k+1) \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_1(k) \\ x_2(k) \\ x_3(k) \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} u(k)$$

From the state transition graph shown in Fig. 4.1.1, it is clear that this LSM is 3-state reachable since starting in any one of the eight states, any other state of the LSM can be reached in at most three transition steps. For example, starting in state  $5 \equiv [1, 0, 1]^T$ , state  $1 \equiv [0, 0, 0]^T$  can be reached by applying the input sequence 101, state  $2 \equiv [1, 0, 0]^T$  can be reached by applying the input sequence 100, state  $3 \equiv [0, 1, 0]^T$  can be reached by applying the single input symbol 1, and so forth.



$$1 \equiv x^1 = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, 2 \equiv x^2 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, 3 \equiv x^3 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, 4 \equiv x^4 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix},$$

$$5 \equiv x^5 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, 6 \equiv x^6 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, 7 \equiv x^7 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, 8 \equiv x^8 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

Fig. 4.1.1. State Transition graph for the LSM of Example 4.1.1



In Definition 4.1.1 the concept of state reachability is defined for any arbitrary final state  $\tilde{x} \in X$ . However, if  $\tilde{x} = 0_X$ , then it is called *state controllability* which is made more precise in the following definition.

Definition 4.1.3. A state  $\hat{x} \in X$  of the LSM  $M = (A, B, C)$  is said to be  $\ell$ -controllable if it can be driven from any initial state  $x^0 \in X$  to the zero state  $0_X$  in exactly  $\ell$  time steps (clock periods), that is, if there exists an input sequence  $u \in U^*$ ,  $\ell g(u) = \ell$ , such that  $\phi(x^0, u) = 0_X$ ;  $M$  is said to be  $\ell$ -state controllable if every state of  $M$  is  $\ell$ -controllable for at least one particular  $\ell$ . The smallest such integer  $\ell_c$  is called the *state controllability index* of  $M$ .

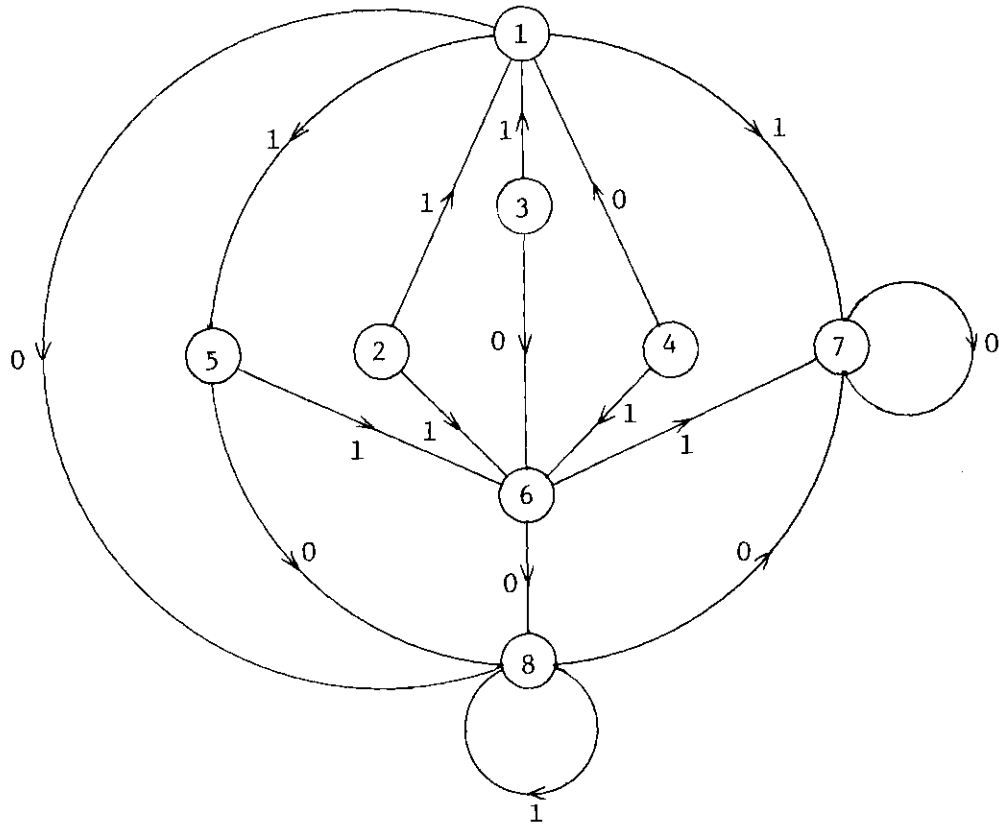
It is clear that the LSM of Example 4.1.1 is 3-state controllable since from its state transition graph, shown in Fig. 4.1.1., it is easily seen that any one of the seven states 2, 3, . . . , 8, can be driven to the zero state  $0_X \equiv [0, 0, 0]^T$  in at most three steps.

In the following example we consider an LSM which is state controllable but not state reachable.

Example 4.1.2. Consider the following state equation of a single-input LSM over  $GF(2)$ :

$$\begin{bmatrix} x_1(k+1) \\ x_2(k+1) \\ x_3(k+1) \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1(k) \\ x_2(k) \\ x_3(k) \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} u(k)$$

The state transition graph of this LSM is shown in Fig. 4.1.2. Clearly this LSM is 2-state controllable but not state reachable.



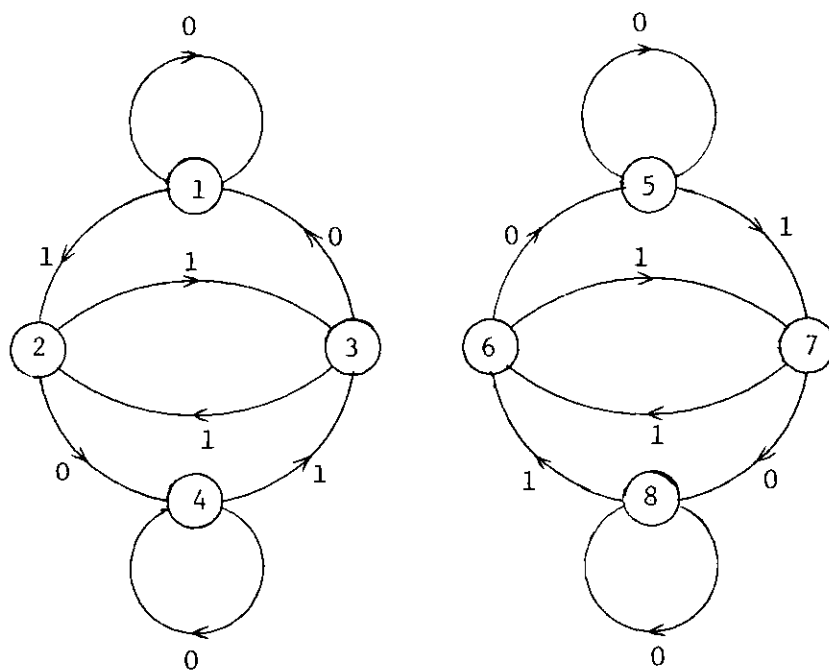
$$1 \equiv x^1 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad 2 \equiv x^2 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \quad 3 \equiv x^3 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad 4 \equiv x^4 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix},$$

$$5 \equiv x^5 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \quad 6 \equiv x^6 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad 7 \equiv x^7 = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \quad 8 \equiv x^8 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

Fig. 4.1.2. State Transition Graph for the LSM of Example 4.1.2

Example 4.1.3. Obviously there exist LSMs which are neither state reachable nor state controllable. The state equation of such an LSM is given below and its state transition graph is shown in Fig. 4.1.3.

$$\begin{bmatrix} x_1(k+1) \\ x_2(k+1) \\ x_3(k+1) \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1(k) \\ x_2(k) \\ x_3(k) \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} u(k)$$



$$1 \equiv x^1 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \quad 2 \equiv x^2 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad 3 \equiv x^3 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \quad 4 \equiv x^4 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix},$$

$$5 \equiv x^5 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \quad 6 \equiv x^6 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \quad 7 \equiv x^7 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \quad 8 \equiv x^8 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

Fig. 4.1.3. State Transition Graph for the LSM of Example 4.1.3

As indicated above, we are explicitly distinguishing between the related concepts of state reachability and controllability since these concepts in contradistinction to the case of conventional continuous-time systems, are not identical for LSMs. Later we will see that it is "easier" for an LSM to be state controllable than it is to be state reachable, by showing that state reachability implies state controllability but not conversely. This fact is illustrated by the state transition graphs of Fig. 4.1.1 and Fig. 4.1.2 of the LSMs of Example 4.1.1 and Example 4.1.2, respectively.

There is also another "connectedness" concept, originally belonging to graph theory, that is sometimes used to characterize certain structural properties of machines. This is the concept of strongly state connectedness which, for the case of LSMs, turns out to be equivalent to state reachability, and is defined below.

Definition 4.1.4. A machine  $M$  is said to *strongly state connected* if for each pair of states  $\hat{x}$  and  $\tilde{x}$  of  $M$ , there exists an input sequence  $u \in U^*$  such that  $\phi(\hat{x}, u) = \tilde{x}$ .

The LSM of Example 4.1.1 is obviously strongly state connected while that of Example 4.1.3 is not.

For the purpose of investigating the reachability and controllability aspects of LSMs, we will need some additional notation and a few auxiliary results which will be presented next.

Let

$$\begin{aligned}\chi^{(x^0)} &\equiv \{\hat{x} : \text{there exists a } u \in \mathcal{U}^*, \text{ such that } \phi(x^0, u) = \hat{x}\} \\ &= \phi(x^0, \mathcal{U}^*), \text{ the range of the map } \phi(x^0, \cdot)\end{aligned}\quad (4.1.1)$$

$$\begin{aligned}\chi_j^{(x^0)} &\equiv \{\hat{x} : \text{there exists } u(0)u(1) \dots u(j-1), u(\cdot) \in \mathcal{U}, \text{ such that} \\ &\quad \phi(x^0, u(0)u(1) \dots u(j-1)) = \hat{x}\}\end{aligned}\quad (4.1.2)$$

$$\begin{aligned}\chi_\ell^{(x^0)} &\equiv \{\hat{x} : \text{there exists } u(0)u(1) \dots u(j-1), 0 \leq j \leq \ell, \text{ such that} \\ &\quad \phi(x^0, u(0)u(1) \dots u(j-1)) = \hat{x}\} \\ &= \bigcup_{j=0}^{\ell} \chi_j^{(x^0)} = \bigcup_{j=0}^{\ell} \phi(x^0, \mathcal{U}^j)\end{aligned}\quad (4.1.3)$$

$$\text{Lemma 4.1.1. } \chi^{(x^0)} = \bigcup_{j=0}^{\infty} \chi_j^{(x^0)} = \bigcup_{j=0}^{\infty} \chi_j^{(x^0)}$$

Proof. Immediate from the definitions of  $\chi^{(x^0)}$ ,  $\chi_j^{(x^0)}$ , and  $\chi_j^{(x^0)}$ .  $\square$

$$\text{Lemma 4.1.2. } \chi_0^{(x^0)} \subseteq \chi_1^{(x^0)} \subseteq \dots \subseteq \chi^{(x^0)}$$

Proof. Immediate from the definition of  $\chi_j^{(x^0)}$ .  $\square$

In order to illustrate the above notation, consider the LSM of

Example 4.1.1. Taking  $x^0 = x^5 \equiv [1, 0, 1]^T$ , we have

$$\chi^{(x^5)} = \left\langle \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \right\rangle$$

$$e_{\chi_0^{(x^5)}} = \left\langle \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \right\rangle$$

$$e_{\chi_1^{(x^5)}} = \left\langle \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \right\rangle$$

$$e_{\chi_2^{(x^5)}} = \left\langle \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right\rangle$$

$$e_{\chi_3^{(x^5)}} = \left\langle \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right\rangle$$

$$\chi_0^{(x^5)} = \left\langle \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \right\rangle$$

$$\chi_1^{(x^5)} = \left\langle \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \right\rangle$$

$$\chi_2^{(x^5)} = \left\langle \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right\rangle$$

$$\chi_3^{(x^5)} = \left\langle \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right\rangle$$

Lemma 4.1.3. If there exists an integer  $r$  such that  $\chi_{r'}^{(x^0)} = \chi_r^{(x^0)}$  for all  $r' \geq r$ , then

$$\chi^{(x^0)} = \bigcup_{j=0}^{\infty} \chi_j^{(x^0)} = \chi_r^{(x^0)} = \bigcup_{j=0}^r \chi_j^{(x^0)}$$

Proof. Let  $x \in \bigcup_{j=0}^{\infty} \chi_j^{(x^0)}$ . Then  $x \in \chi_s^{(x^0)}$  for some integer  $s$ .

If  $s \leq r$ , then  $\chi_s^{(x^0)} \subseteq \bigcup_{j=0}^r \chi_j^{(x^0)}$  so  $x \in \bigcup_{j=0}^r \chi_j^{(x^0)}$ ; and if  $s > r$ , then

$$\chi_s^{(x^0)} = \chi_r^{(x^0)} \subseteq \bigcup_{j=0}^r \chi_j^{(x^0)} \text{ which shows that } x \in \bigcup_{j=0}^r \chi_j^{(x^0)}. \text{ Hence}$$

$$\chi^{(x^0)} \subseteq \chi_r^{(x^0)}. \text{ The reverse inclusion is obvious. } \square$$

Lemma 4.1.4. Let  $M = (A, B, C)$  be an LSM with a state  $x^0 \in X$  for which there exists an integer  $j$  such that  $\chi_j^{(x^0)} = \chi_{j+1}^{(x^0)}$ . Then  $\chi_j^{(x^0)} = \chi^{(x^0)}$ .

Proof. By Lemma 4.1.3 we need only prove that  $\chi_r^{(x^0)} = \chi_{r'}^{(x^0)}$  for all  $r' > r$ . We will accomplish this by induction on  $j$  for  $r' = r + j$ . For  $r' = r + 1$ ,  $\chi_{r+1}^{(x^0)} = \chi_r^{(x^0)}$  by hypothesis. Next suppose that  $\chi_r^{(x^0)} = \chi_{r'}^{(x^0)}$  for  $r' = \ell$ . Now we want to show that  $\chi_r^{(x^0)} = \chi_{\ell+1}^{(x^0)}$ .

It is clear that

$$\chi_{\ell+1}^{(x^0)} = \phi(x^0, u^{\ell+1}) = \phi(x^0, u^{r+1} u^{\ell-r})$$

Thus

$$\chi_{\ell+1}^{(x^0)} = \phi(\chi_{r+1}^{(x^0)}, u^{\ell-r})$$

Also,  $\phi(x^0, u^1 u^2) = \phi(\phi(x^0, u^1), u^2)$  for all  $u^1 \in U^{r+1}$  and  $u^2 \in U^{\ell-r}$  so that

$$\begin{aligned} x_{\ell+1}^{(x^0)} &= \phi(x_r^{(x^0)}, u^{\ell-r}) \text{ by hypothesis} \\ &= x_{\ell}^{(x^0)} \\ &= x_r^{(x^0)} \text{ by induction hypothesis } \square \end{aligned}$$

The results contained in Lemma 4.1.1 - Lemma 4.1.4 are applicable to general sequential machines and are not limited to LSMs. However, we will make use of these general results for the study of state reachability properties of LSMs.

Theorem 4.1.1. For the LSM  $M = (A, B, C)$  the set of states reachable from the zero state  $0_X$  in at most  $\ell$  clock periods is the range  $R(K)$  of the linear map

$$K(A, B, \ell) \equiv [A^{\ell-1}B, A^{\ell-2}B, \dots, AB, B] : U^{\ell} \longrightarrow X$$

That is,

$$x_{\ell}^{(0_X)} = R([A^{\ell-1}B, A^{\ell-2}B, \dots, AB, B])$$

Proof. First, we observe that a zero input leaves the zero state  $0_X$  unchanged since  $0_X = A0_X + B0_X$ . This means that if a state  $\hat{x}$  can be reached from the zero state by applying an input sequence  $u(0)u(1) \dots u(j-1)$  of length  $j < \ell$ , then  $\hat{x}$  can also be reached from  $0_X$  by first applying the input sequence  $0^{\ell-j}$ , that is, a string of  $\ell-j$



successive inputs, each with value 0, and then applying  $u(0)u(1) \dots u(j-1)$ . Thus, for all  $u(0)u(1) \dots u(j-1) \in U^j$  and all  $\ell > j$ , the LSM  $(A, B, C)$  satisfies

$$\phi(0_X, u(0)u(1) \dots u(j-1)) = \phi(0_X, 0^{\ell-j} u(0)u(1) \dots u(j-1))$$

That is, for the LSM  $(A, B, C)$ ,  $X_\ell^{(0_X)} = X_\ell^{e(0_X)}$ .

Therefore,  $\hat{x} \in X_\ell^{(0_X)}$  if and only if there exists an input sequence  $u(0)u(1) \dots u(\ell-1)$  of length exactly  $\ell$  such that

$$\begin{aligned} \hat{x} &= A^\ell 0_X + \sum_{j=0}^{\ell-1} A^{\ell-j-1} B u(j) \\ &= [A^{\ell-1}B, A^{\ell-2}B, \dots, AB, B] \begin{bmatrix} u(0) \\ u(1) \\ \vdots \\ u(\ell-1) \end{bmatrix} \end{aligned}$$

Clearly  $\hat{x} \in X_\ell^{(0_X)} = R([A^{\ell-1}B, A^{\ell-2}B, \dots, AB, B]) \equiv R(K)$ .  $\square$

The state reachability subspace  $R(K) \subseteq X$  may be expressed in the equivalent form

$$R(K) = R(B) + AR(B) + A^2R(B) + \dots + A^{\ell-1}R(B) \quad (4.1.4)$$

where  $AR(B) \equiv \{Ax : x \in R(B)\}$ . From this expression it follows that  $R(K)$  is the smallest subspace of the state space  $X$  that contains the range  $R(B)$  of the input matrix  $B$ . Another property that contributes to the significance of  $R(K)$  in many structural aspects of the LSM  $(A, B, C)$  is its invariance under  $A$ . To see that  $R(K)$  is  $A$ -invariant, let  $z \in R(K)$ . Then  $z = Ax$  for some  $x \in R(K)$ . But  $x \in R(K)$  means that

$x$  can be expressed as a linear combination of the columns of  $B, AB, \dots, A^{\ell-1}B$ , so that  $Ax$  can be expressed as a linear combination of the columns of  $AB, A^2B, \dots, A^{\ell}B$ . Since by the Cayley-Hamilton Theorem,  $A^{\ell}$  can be expressed as a linear combination of the matrices  $I_n, A, A^2, \dots, A^{\ell-1}$ , then  $A^{\ell}B$  can be expressed as a linear combination of  $B, AB, \dots, A^{\ell-1}B$ . Therefore,  $Ax$  can be expressed as a linear combination of the columns of  $B, AB, \dots, A^{\ell-1}B$ . That is,  $z = Ax \in \mathcal{R}(K)$ . Thus  $A\mathcal{R}(K) \subseteq \mathcal{R}(K)$ . We have thus proved the following theorem.

**Theorem 4.1.2.** The state reachability subspace  $\mathcal{R}(K) \subseteq X$  of an LSM  $(A, B, C)$  is the smallest  $A$ -invariant subspace that contains the range  $\mathcal{R}(B)$  of the input matrix  $B$ .

**Theorem 4.1.3.** Every state of the  $n$ -dimensional LSM  $(A, B, C)$  reachable from the zero state  $0_X$  can be reached in at most  $n$  clock periods, that is,  $X_n^{(0_X)} = X^{(0_X)}$ .

**Proof.** It is clear that for each integer  $r$ ,  $X_r^{(0_X)} \subseteq X_{r+1}^{(0_X)} \subseteq X$ . Thus, if for any  $r$ ,  $X_{r+1}^{(0_X)} \neq X_r^{(0_X)}$ , we must have  $\dim X_{r+1}^{(0_X)} \geq 1 + \dim X_r^{(0_X)}$ . Now consider the chain of subsets

$$X_0^{(0_X)} \subseteq X_1^{(0_X)} \subseteq \dots \subseteq X_n^{(0_X)} \subseteq X^{(0_X)}$$

If  $X_n^{(0_X)} = X_{n-1}^{(0_X)}$ , then  $X^{(0_X)} = X_{n-1}^{(0_X)}$  by Lemma 4.1.1, and certainly

$X_n^{(0_X)} = X^{(0_X)}$ . If  $X_n^{(0_X)} \neq X_{n-1}^{(0_X)}$ , then we must have  $X_0^{(0_X)} \neq X_1^{(0_X)} \neq$

$X_2^{(0_X)} \neq \dots \neq X_{n-1}^{(0_X)} \neq X_n^{(0_X)}$ . But the dimension of  $X_j^{(0_X)}$  increases by

at least one at each step of this chain. Thus  $\dim X_n^{(0_X)} \geq n$ . But

$X_n^{(0_X)}$  is a subspace of  $X$ , which itself is of dimension  $n$ . Hence

$X_n^{(0_X)} = X$  and since  $X_n^{(0_X)} \subseteq X^{(0_X)} \subseteq X$ , we have  $X_n^{(0_X)} = X^{(0_X)}$ .  $\square$

Theorem 4.1.4. The  $n$ -dimensional LSM  $(A, B, C)$  is state reachable from the zero state  $0_X$  if and only if

$$\text{rank } [A^{n-1}B, A^{n-2}B, \dots, AB, B] = n$$

Proof. The LSM  $(A, B, C)$  is state reachable if and only if  $X = X^{(0_X)}$  or, since  $X^{(0_X)} = X_n^{(0_X)}$  by Theorem 4.1.3, if and only if

$$\dim X = \dim X_n^{(0_X)}$$

$$n = \dim R([A^{n-1}B, A^{n-2}B, \dots, AB, B])$$

$$n = \text{rank } [A^{n-1}B, A^{n-2}B, \dots, AB, B] \quad \square$$

Corollary 4.1.1. A single-input LSM  $(A, b, c^T)$  is state reachable if and only if the matrix  $[A^{n-1}b, A^{n-2}b, \dots, Ab, b] \in GF(q)^{n \times n}$  is nonsingular.

Corollary 4.1.2. If the characteristic matrix  $A$  of the LSM  $(A, B, C)$  is diagonal with distinct elements, then the LSM is state reachable if and only if the input matrix  $B$  has no zero rows.

Corollary 4.1.3. The LSM  $(A, B, C)$  is state reachable if and only if the matrix  $KK^T = \sum_{j=0}^{n-1} A^{n-j-1} BB^T (A^T)^{n-j-1} \in GF(q)^{n \times n}$  is nonsingular.

Proof. It follows from the fact that for any  $n \times \ell$  matrix  $Q$ ,  $\text{rank } Q = n \iff \text{rank } QQ^T = n$ , and Theorem 4.1.4.  $\square$

Corollary 4.1.4. The LSM  $(A, B, C)$  is  $\ell$ -state reachable,  $\ell \leq n$ , if and only if  $\text{rank } [A^{\ell-1}B, A^{\ell-2}B, \dots, AB, B] = n$ .

Proof. By definition, the LSM  $(A, B, C)$  is  $\ell$ -state reachable if and only if for every pair of states  $\tilde{x}$  and  $\hat{x}$  there exists an input sequence  $u \in U^\ell$  such that  $\phi(\hat{x}, u) = \tilde{x}$ , or

$$\hat{x} = A^{\ell-1}\tilde{x} + \sum_{j=0}^{\ell-1} A^{\ell-j-1} B u(j)$$

$$\hat{x} - A^{\ell-1}\tilde{x} = [A^{\ell-1}B, A^{\ell-2}B, \dots, AB, B] \begin{pmatrix} u(0) \\ u(1) \\ \vdots \\ u(\ell-1) \end{pmatrix}$$

$$\hat{x} - A^{\ell-1}\tilde{x} = Ku$$

Since  $\hat{x} - A^{\ell-1}\tilde{x}$  can be an arbitrary vector  $v \in GF(q)^n$ ,  $\ell$ -state reachability reduces to the condition that the vector equation  $Ku = v$  be solvable for all  $v$ . Therefore, this condition implies that the LSM is  $\ell$ -state reachable if and only if  $\text{rank } K = n$ .  $\square$

Corollary 4.1.5. Let the minimal polynomial of  $A$  be of degree  $r \leq n$ . Then the LSM  $(A, B, C)$  is  $\ell$ -state reachable for some  $\ell$ , if and only if it is  $r$ -state reachable.

Proof. If the minimal polynomial of  $A$  is  $f_m(\lambda) = \sum_{i=0}^r a_i(\lambda)^i$ , then by the Cayley-Hamilton Theorem

$$\sum_{i=0}^r a_i A^i = 0$$

$$\sum_{i=0}^r a_i A^i B = 0$$

$$A^r B = - \sum_{i=0}^{r-1} a_i A^i B$$

and hence the columns of  $A^s B$  are linearly dependent on the columns of  $[A^{r-1} B, A^{r-2} B, \dots, AB, B]$  for all  $s \geq r$ . Thus, if  $K(A, B, s)$  has rank  $n$  for any  $s$ , then  $K(A, B, r)$  must also have rank  $n$ .  $\square$

Corollary 4.1.6. The LSM  $(A, B, C)$  is  $\ell$ -state reachable for some  $\ell \leq n$ , if and only if it is  $n$ -state reachable.

Proof. It follows from the fact that the degree of the minimal polynomial of  $A$  does not exceed  $n$ .  $\square$

Using Theorem 4.1.4, we can characterize the reachability index  $\ell_r$  (Definition 4.1.1) of an LSM  $(A, B, C)$  as follows:

$$\ell_r = \min\{j : 1 \leq j \leq n, \text{rank}[B, AB, A^2 B, \dots, A^{j-1} B] = n\}$$

In view of (4.1.4), this integer can be equivalently described as

$$\ell_r = \min\{j : 1 \leq j \leq n, R(B) + AR(B) + A^2 R(B) + \dots + A^{j-1} R(B) = X\}$$

Theorem 4.1.4 provides a straightforward computational procedure for checking the state reachability property of an LSM. However, this criterion requires the calculation of the entire reachability matrix  $K(A, B, n)$  which may not actually be needed. In many instances, we need not calculate  $K(A, B, n)$  but only a matrix with a smaller number of columns. This claim follows directly from Lemma 4.1.4 and Theorem 4.1.1. Therefore, restating Lemma 4.1.4 in terms of state reachability matrices, we obtain the following result.

Theorem 4.1.5. If  $j$  is the least integer such that  $\text{rank } K(A, B, j) = \text{rank } K(A, B, j+1)$ , then  $\text{rank } K(A, B, \ell) = \text{rank } K(A, B, j)$  for all integers  $\ell > j$ , and  $j \leq \min\{n-r, \bar{n}-1\}$ , where  $r$  is the rank of  $B$  and  $\bar{n}$  is the degree of the minimal polynomial of  $A$ .

Corollary 4.1.7. (Simplified Reachability Criterion) If  $\text{rank } B = r$ , then the  $n$ -dimensional LSM  $(A, B, C)$  is state reachable if and only if  $\text{rank } K(A, B, n-r+1) = n$ .

Corollary 4.1.8. If  $\text{rank } B = r$ , then the LSM  $(A, B, C)$  is state reachable if and only if the matrix  $K(A, B, n-r+1) K^T(A, B, n-r+1) \in \text{GF}(q)^{n \times n}$  is nonsingular.

Proof. It follows from Corollary 4.1.3 and Theorem 4.1.5.  $\square$

Corollary 4.1.9. If  $\text{rank } B = 1$  and the LSM  $(A, B, C)$  is state reachable, then  $f_c = f_m$ , where  $f_c$  and  $f_m$  are the characteristic and minimal polynomials of  $A$ .

Example 4.1.4. In order to illustrate the application of the simplified reachability criterion of Corollary 4.1.7, consider the following state equation of a three-dimensional LSM over  $\text{GF}(3)$ :

$$\begin{bmatrix} x_1(k+1) \\ x_2(k+1) \\ x_3(k+1) \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 2 & 0 \end{bmatrix} \begin{bmatrix} x_1(k) \\ x_2(k) \\ x_3(k) \end{bmatrix} + \begin{bmatrix} 1 & 2 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} u_1(k) \\ u_2(k) \end{bmatrix}$$

Since  $\text{rank } B = 2$ , we need to check  $K(A, B, 2) = [AB, B]$

$$\text{rank } [AB, B] = \text{rank} \begin{pmatrix} 2 & 2 & | & 1 & 2 \\ 1 & 1 & | & 0 & 1 \\ 0 & 2 & | & 1 & 0 \end{pmatrix} = 3$$

Hence the given LSM is state reachable.

#### 4.2. State Unreachability of LSMs

In Theorem 4.1.2 it was shown that the state reachability subspace  $R(K) \equiv R([B, AB, \dots, A^{n-1}B])$  of the LSM  $(A, B)$  is the smallest  $A$ -invariant subspace of the state space  $X$ , that contains the range  $R(B)$  of the control matrix  $B$ . Thus if the LSM is state reachable, then  $R(K) = X$ , that is,  $R(K)^\perp = \{0\}$ . However, if the LSM is not state reachable, then  $R(K)$  is a proper subspace of  $X$ , that is,  $\text{rank } K < n$ , and hence  $R(K)^\perp \neq \{0\}$ . In this case, then, one would be tempted to expect that, in analogy with the similar situation in conventional linear dynamical systems, the state space  $X$  can be expressed as a direct sum of  $R(K)$  and  $R(K)^\perp$ , resulting into the so-called unreachable canonical form which essentially separates the reachable and unreachable portions of the LSM. However, due to the peculiarities of the ground field  $GF(q)$ , this direct sum decomposition is, in general, no longer possible for LSMs. At this point we will momentarily digress to briefly elaborate on this particular property of LSMs. We begin with the definition of a bilinear form.

A *bilinear form*  $f$  on a vector space  $V$  over an arbitrary field  $F$  is a function  $f : V \times V \longrightarrow F$  such that

$$f(rv + sv', v'') = rf(v, v'') + sf(v', v'')$$

$$f(v, rv' + sv'') = rf(v, v') + sf(v, v'')$$

for all  $v, v', v'' \in V$  and for all  $r, s \in F$ . If, in addition,  $f(v, v') = f(v', v)$  for all  $v, v' \in V$ , then  $f$  is said to be *symmetric*.

From the above definition it is clear that the usual inner product is a symmetric bilinear form.

The matrix of a bilinear form  $f$  relative to a fixed basis  $\{v^1, v^2, \dots, v^n\}$  of  $V$  is the  $n \times n$  matrix  $F$  over  $F$  with entries  $f_{ij} \equiv f(v^i, v^j)$ ;  $i, j \in \underline{n}$ . This matrix completely determines the form  $f$ . To see this, we express the vectors  $v, v' \in V$  in terms of the given basis elements of  $V$  as

$$v = \sum_{i=1}^n a_i v^i, \quad v' = \sum_{j=1}^n b_j v^j$$

where  $a_i, b_j \in F$ ;  $i, j \in \underline{n}$ . Then the bilinearity of  $f$  shows that  $f$  is determined by  $F$  as

$$f(v, v') = \sum_{i=1}^n \sum_{j=1}^n a_i f(v^i, v^j) b_j$$

$$= \sum_{i=1}^n \sum_{j=1}^n a_i f_{ij} b_j$$

$$= [a_1, a_2, \dots, a_n] \begin{pmatrix} f_{11} & f_{12} & \dots & f_{1n} \\ f_{21} & f_{22} & \dots & f_{2n} \\ \vdots & \vdots & & \vdots \\ f_{n1} & f_{n2} & \dots & f_{nn} \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$$

$$\equiv a^T F b$$



where  $a, b \in F^n$  and  $F \in F^{n \times n}$ .

Since two  $n \times n$  matrices represent the same bilinear form if and only if they are equivalent, and equivalent matrices have the same rank, the rank of a bilinear form is defined to be the rank of any one (and hence of every one) of the matrices of a bilinear form. In view of this fact, the bilinear form  $f$  on  $V$  is said to be *nondegenerate* if the rank of the matrix of  $f$  is equal to the dimension of  $V$ , otherwise  $f$  is called *degenerate*.

It is well known that for the case of a vector space  $W$  over the field of real numbers,  $W$  can always be expressed as a direct sum of the subspaces  $W_1$  and  $W_1^\perp$  with respect to the usual inner product bilinear form. On the contrary, this decomposition is not, in general, possible for a vector space  $X$  over  $GF(q)$ . For example, let

$$X \equiv GF(2)^5$$

and

$$X \supset X_1 \equiv \left\langle \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \\ 0 \end{pmatrix} \right\rangle$$

Then

$$X_1^\perp = \left\langle \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right\rangle$$

Although  $\dim X = \dim X_1 + \dim X_1^\perp$ , we see that  $X \neq X_1 \oplus X_1^\perp$ . However, if we choose

$$X \supset X_2 \equiv \left\langle \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \right\rangle$$

then

$$X_2^\perp = \left\langle \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \right\rangle$$

and  $X = X_2 \oplus X_2^\perp$ . It can be easily checked that the usual inner product bilinear form is degenerate on the subspace  $X_1$  in the first case and nondegenerate on  $X$  and  $X_2$  in the second case. In general, nondegeneracy of a symmetric bilinear form on both  $V$  and  $V_1 \subset V$  turns out to be a sufficient condition for the existence of the direct sum decomposition  $V = V_1 \oplus V_1^\perp$ . We will prove this result for the special case of an inner product bilinear form.

Lemma 4.2.1. Let  $V$  be a finite-dimensional inner product space over an arbitrary field  $F$ , and let the inner product bilinear form  $f$  on  $V$  be nondegenerate on the subspace  $V_1$  of  $V$ . Then  $V$  can be expressed as  $V = V_1 \oplus V_1^\perp$ .

Proof. Let  $\{v^1, v^2, \dots, v^m\}$  be a basis of  $V_1$ . If  $x \in V$  and  $y \in V_1^\perp$ , we want to show that there exist  $a_i \in F$ ,  $i \in \underline{m}$ , such that

$$x = \sum_{i=1}^m a_i v^i + y$$

Let  $f(v^i, v^j) \equiv c_{ij}$ ;  $i, j \in \underline{m}$ . Since  $f(y, v^i) = 0$ , we have

$$\sum_{i=1}^m c_{ij} a_i = f(x, v^j), \quad j \in \underline{m}$$

This system of  $m$  equations for the  $m$  unknowns  $a_i$ ,  $i \in \underline{m}$ , has a (unique) solution since by hypothesis the matrix of coefficients is nonsingular.  $\square$

Now we return to the discussion of state unreachability of LSMs.

Theorem 4.2.1. Suppose that the LSM  $(A, B)$  is state unreach-  
able, and let the inner product bilinear form on  $X$  be nondegenerate on  
the subspace  $R(K)$ . Then there exists an isomorphism  $P : X \rightarrow X$ ,  
 $x \mapsto Px$ ,  $P \in GL(n, q)$ , such that the isomorphic LSM  $(PAP^{-1}, PB) \equiv$   
 $(\tilde{A}, \tilde{B})$  has the form

$$\left( \begin{bmatrix} \tilde{A}_{11} & \tilde{A}_{12} \\ 0 & A_{21} \end{bmatrix}, \begin{bmatrix} \tilde{B}_1 \\ 0 \end{bmatrix} \right) \quad (4.2.1)$$

or equivalently,

$$\begin{aligned} \tilde{x}^I(k+1) &= \tilde{A}_{11} \tilde{x}^I(k) + \tilde{A}_{12} \tilde{x}^{II}(k) + \tilde{B}_1 u(k) \\ \tilde{x}^{II}(k+1) &= \tilde{A}_{22} \tilde{x}^{II}(k) \end{aligned} \quad (4.2.2)$$

where  $\tilde{x}^I \in GF(q)^r$ ,  $\tilde{x}^{II} \in GF(q)^{n-r}$ ,  $\tilde{A}_{11} \in GF(q)^{r \times r}$ ,  $A_{12} \in GF(q)^{r \times (n-r)}$ ,  $\tilde{A}_{22} \in GF(q)^{(n-r) \times (n-r)}$ , and  $\tilde{B}_1 \in GF(q)^r \times m$ .

Proof. Since, by hypothesis, the inner product bilinear form is nondegenerate on  $R(K)$  and  $\dim R(K) < n = \dim X$ , in view of Lemma 4.2.1, the state space can be expressed as  $X = R(K) \oplus R(K)^\perp$ . Let  $\{v^1, v^2, \dots, v^r\}$ ,  $\{v^{r+1}, v^{r+2}, \dots, v^n\}$ , and  $\{v^1, v^2, \dots, v^r, v^{r+1}, \dots, v^n\}$  be bases for  $R(K)$ ,  $R(K)^\perp$ , and  $X$ , respectively. Then any  $x \in R(K)$  can be uniquely represented as a linear combination of  $\{v^i, i \in \underline{r}\}$ . Since  $R(K)$  is  $A$ -invariant,  $x^i \in R(K)$  implies that  $Ax^i \in R(K)$ ,  $i \in \underline{r}$ , and hence

$$\begin{aligned} Ax^i &= a_{1i} v^1 + a_{2i} v^2 + \dots + a_{ri} v^r \\ &= a_{1i} v^1 + a_{2i} v^2 + \dots + a_{ri} v^r + 0v^{r+1} + 0v^{r+2} + \dots + 0v^n, \end{aligned}$$

$$i \in \underline{r} \tag{4.2.3}$$

for appropriate  $a_{hi} \in GF(q)$ . From (4.2.3) it follows that the matrix representation of  $A$  with respect to the basis  $\{v^i, i \in \underline{n}\}$  is precisely of the form given in (4.2.1). Since  $R(B) \subseteq R(K)$ , that is, for  $j \in \underline{m}$ , the  $j$ th column of  $B$  is in  $R(K)$ , its representation with respect to the basis  $\{v^i, i \in \underline{n}\}$  must have the last  $n-r$  entries equal to zero. Thus  $B$  has the form indicated in (4.2.1). Clearly  $p^{-1}$  is the matrix formed by the vectors  $\{v^i, i \in \underline{n}\}$ .  $\square$

Theorem 4.2.1 can be used to separate the reachable part of an unreachable LSM. That is, given any unreachable LSM, this result can be applied to determine a submachine of smaller dimension which is state reachable. This is readily seen, by directly computing the reachability matrix  $\tilde{K}$  of the LSM represented by (4.2.1), as follows:

$$\begin{aligned}\tilde{K} &\equiv [\tilde{B}, \tilde{A}\tilde{B}, \tilde{A}^2\tilde{B}, \dots, \tilde{A}^{n-1}\tilde{B}] \\ &= \begin{pmatrix} \tilde{B}_1 & \tilde{A}_{11}\tilde{B}_1 & \tilde{A}_{11}^2\tilde{B}_1 & \dots & \tilde{A}_{11}^{n-1}\tilde{B}_1 \\ 0 & 0 & 0 & \dots & 0 \end{pmatrix} \quad (4.2.4)\end{aligned}$$

Since  $\tilde{K} = PK$ , it is clear that  $\text{rank } \tilde{K} = \text{rank } PK = \text{rank } K = r$ . Thus the first  $r$  rows of (4.2.4), which constitute the state reachability matrix for the submachine  $(\tilde{A}_{11}, \tilde{B}_1)$ , are linearly independent, and hence  $(\tilde{A}_{11}, \tilde{B}_1)$  is state reachable.

From the representation (4.2.2) where  $\tilde{x} = \begin{pmatrix} \tilde{x}^I \\ \tilde{x}^{II} \end{pmatrix} + \begin{pmatrix} 0 \\ \tilde{x}^{II} \end{pmatrix}$ ,  $\begin{pmatrix} \tilde{x}^I \\ 0 \end{pmatrix} \in R(K)$ , and  $\begin{pmatrix} 0 \\ \tilde{x}^{II} \end{pmatrix} \in R(K)^\perp$ , it is clear that the state subvector  $\tilde{x}^{II}$  is not affected directly by the input or indirectly through  $\tilde{x}^I$ , and hence is disregarded in the reduced submachine  $(\tilde{A}_{11}, \tilde{B}_1)$ . Furthermore, from the second equation of (4.2.2) we observe that if  $\tilde{x}^{II}(0) = 0$ , then  $\tilde{x}^{II}(k) = 0$  for all subsequent clock periods. Therefore, the submachine  $(\tilde{A}_{11}, \tilde{B}_1)$  is zero-state equivalent to the LSM (4.2.2) and, consequently, to the original LSM  $(A, B)$ .

The following example will illustrate the above ideas.

Example 4.2.1. Consider the following LSM over GF(2):

$$\begin{pmatrix} x_1(k+1) \\ x_2(k+1) \\ x_3(k+1) \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1(k) \\ x_2(k) \\ x_3(k) \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} u(k)$$

The state reachability matrix of this LSM is

$$K \equiv [B, AB, A^2B] = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

which has rank 2. The matrix  $P^{-1}$  can be chosen as

$$P^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

where the first two columns are the linearly independent columns of  $K$  and the third column is chosen arbitrarily but with the provision that  $P^{-1}$  is nonsingular. The isomorphic LSM  $(\tilde{A}, \tilde{B}) \equiv (PAP^{-1}, PB)$  has the required form

$$\begin{pmatrix} \tilde{x}_1(k+1) \\ \tilde{x}_2(k+1) \\ \hline \tilde{x}_3(k+1) \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 1 & 1 \\ \hline 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \tilde{x}_1(k) \\ \tilde{x}_2(k) \\ \hline \tilde{x}_3(k) \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \\ \hline 0 \end{pmatrix} u(k)$$

The remainder of this section will be devoted to a rederivation of the isomorphic LSM (4.2.1) by using a different approach (cf.[43]).

A careful examination of the general expression for solution of the state equations of the LSM (A, B) given by

$$\begin{aligned}
 x(n) &= A^n x(0) + \sum_{j=0}^{n-1} A^{n-j-1} B u(j) \\
 &= A^n x(0) + [A^{n-1}B, A^{n-2}B, \dots, AB, B] \begin{bmatrix} u(0) \\ u(1) \\ \vdots \\ u(n-1) \end{bmatrix} \quad (4.2.5)
 \end{aligned}$$

in conjunction with the rank deficiency of the state reachability matrix  $K \equiv [A^{n-1}B, A^{n-2}B, \dots, AB, B]$ , will reveal the existence of certain fundamental relationships between the reachable and unreachable subvectors of the state  $x(k)$ . In order to investigate some of these relationships, we need to rewrite the state equation of the LSM (A, B) in the following partitioned form:

$$\begin{bmatrix} x^I(k+1) \\ x^{II}(k+1) \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} x^I(k) \\ x^{II}(k) \end{bmatrix} + \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} u(k) \quad (4.2.6)$$

where  $x^I(k) \in GF(q)^r$ ,  $x^{II}(k) \in GF(q)^{(n-r)}$ ,  $A_{11} \in GF(q)^{r \times r}$ ,  $A_{12} \in GF(q)^{r \times (n-r)}$ ,  $A_{21} \in GF(q)^{(n-r) \times r}$ ,  $A_{22} \in GF(q)^{(n-r) \times (n-r)}$ ,  $B_1 \in GF(q)^{r \times m}$ , and  $B_2 \in GF(q)^{(n-r) \times m}$ . Now if  $\text{rank } K = r < n$ , then  $n - r$  rows of  $K$  can be expressed as linear combinations of the remaining  $r$  linearly independent rows, and therefore, in view of (4.2.5), the following relationship exists between the zero-state solution subvectors  $x_{os}^I$  and  $x_{os}^{II}$ :

$$x_{os}^{II}(k) = Tx_{os}^I(k) \quad (4.2.7)$$

where  $T \in GF(q)^{(n-r) \times r}$ .

Similarly, if we assume that  $\text{rank } K = r < n$ , and consider the expression for the general solution of the state equation

$$\begin{pmatrix} x^I(k) \\ x^{II}(k) \end{pmatrix} = \begin{pmatrix} x_{os}^I(k) \\ x_{os}^{II}(k) \end{pmatrix} + \begin{pmatrix} x_{oi}^I(k) \\ x_{oi}^{II}(k) \end{pmatrix} \quad (4.2.8)$$

then the following relationship holds:

$$x^{II}(k) = Tx^I(k) + z(x(0), k) \quad (4.2.9)$$

where  $T$  is defined by (4.2.7) and  $z \in GF(q)^{n-r}$  such that  $z(0, k) = 0$ .

Relation (4.2.9) follows immediately from the decomposed expression for  $x^{II}(k)$  given by (4.2.8), and (4.2.7) as follows:

$$\begin{aligned} x^{II}(k) &= x_{os}^{II}(k) + x_{oi}^{II}(k) \\ &= Tx_{os}^I(k) + Tx_{oi}^I(k) - Tx_{oi}^I(k) + x_{oi}^{II}(k) \\ &= Tx^I(k) + z(x(0), k) \end{aligned}$$

where

$$z(x(0), k) \equiv x_{oi}^{II}(k) - Tx_{oi}^I(k). \quad (4.2.10)$$



Using the above results, we will introduce a special isomorphism that will transform a state unreachable LSM (A, B) to an isomorphic LSM ( $\tilde{A}$ ,  $\tilde{B}$ ) having the unreachable canonical form given by (4.2.1).

Theorem 4.2.2. The isomorphism

$$P^{-1} \equiv \begin{bmatrix} I_r & 0 \\ T & I_{n-r} \end{bmatrix}^{-1} : X \longrightarrow X, \quad (4.2.11)$$

$$\begin{bmatrix} x^I(k) \\ x^{II}(k) \end{bmatrix} \xrightarrow{P^{-1}} \begin{bmatrix} I_r & 0 \\ T & I_{n-r} \end{bmatrix}^{-1} \begin{bmatrix} x^I(k) \\ x^{II}(k) \end{bmatrix}$$

transforms the state unreachable LSM (A, B) to the unreachable canonical form given by (4.2.1).

Proof. In view of the relation (4.2.9), the general solution of the state equation of the LSM (A, B) can be written as

$$\begin{bmatrix} x^I(k) \\ x^{II}(k) \end{bmatrix} = \begin{bmatrix} x^I(k) \\ Tx^I(k) + z(x(0), k) \end{bmatrix} = \begin{bmatrix} I_r & 0 \\ T & I_{n-r} \end{bmatrix} \begin{bmatrix} x^I(k) \\ z(x(0), k) \end{bmatrix} \quad (4.2.12)$$

which must satisfy the equation (4.2.6), that is,

$$\begin{bmatrix} I_r & 0 \\ T & I_{n-r} \end{bmatrix} \begin{bmatrix} x^I(k+1) \\ z(x(0), k+1) \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} I_r & 0 \\ T & I_{n-r} \end{bmatrix} \begin{bmatrix} x^I(k) \\ z(x(0), k) \end{bmatrix} \quad (4.2.13)$$

Since

$$\begin{pmatrix} I_r & 0 \\ T & I_{n-r} \end{pmatrix}^{-1} = \begin{pmatrix} I_r & 0 \\ -T & I_{n-r} \end{pmatrix}$$

equation (4.2.13) reduces to

$$\begin{pmatrix} x^I(k+1) \\ z(x(0), k+1) \end{pmatrix} = \begin{pmatrix} A_{11} + A_{12}T & A_{12} \\ -TA_{11} - TA_{12}T + A_{21} + A_{22}T & A_{22} - TA_{12} \end{pmatrix} \begin{pmatrix} x^I(k) \\ z(x(0), k) \end{pmatrix} + \begin{pmatrix} B_1 \\ B_2 - TB_1 \end{pmatrix} u(k) \quad (4.2.14)$$

Assuming that  $\text{rank } K = r < n$ , it follows that in (4.2.14),  $B_2 - TB_1 = 0$  since according to (4.2.7) the matrix  $T$  expresses the linear dependence of the  $n - r$  rows of  $K$  on the remaining  $r$  linearly independent rows. Therefore, if we show that in (4.2.14),

$$\tilde{A}_{21} \equiv -TA_{11} - TA_{12}T + A_{21} + A_{22}T = 0 \quad (4.2.15)$$

then the proof is complete by taking  $\tilde{x}^I(k) \equiv x^I(k)$ ,  $\tilde{x}^{II}(k) \equiv z(x(0), k)$ ,  $\tilde{A}_{11} \equiv A_{11} + A_{12}T$ ,  $\tilde{A}_{12} \equiv A_{12}$ ,  $\tilde{A}_{22} \equiv A_{22} - TA_{12}$ , and  $\tilde{B}_1 \equiv B_1$ . To show (4.2.15), we will consider the augmented reachability matrix  $\bar{K} \equiv$

$[K \mid A^n B]$  which has the same rank as  $K$  since, by the Cayley-Hamilton Theorem,  $A^n = \sum_{i=1}^{n-1} a_i A^i$ . Using (4.2.6),  $\bar{K}$  can be written as

$$\bar{K} = \begin{bmatrix} B_1 & (A_{11} + A_{12}T)B_1 & (A_{11} + A_{12}T)^2 B_1 & \dots & (A_{11} + A_{12}T)^n B_1 \\ TB_1 & T(A_{11} + A_{12}T)B_1 & T(A_{11} + A_{12}T)^2 B_1 & \dots & T(A_{11} + A_{12}T)^n B_1 \end{bmatrix} \quad (4.2.16)$$

In forming (4.2.16), we have used the linear dependence relations

$B_2 = TB_1$  and  $A_{21}B_1 + A_{22}B_2 = T(A_{11}B_1 + A_{12}B_2)$ . From (4.2.16) it is seen that the general expression for the  $n - r$  rows of the  $i$ th column block has the form

$$T(A_{11} + A_{12}T)^i B_1 = 0, \quad i \in \underline{n} \quad (4.2.17)$$

However, in view of the generation scheme used in (4.2.16), the last  $n - r$  rows of the  $i$ th column block can be obtained from the  $(i - 1)$ th column block as follows:

$$[A_{21} \quad A_{22}] \begin{bmatrix} (A_{11} + A_{12}T)^{i-1} B_1 \\ T(A_{11} + A_{12}T)^{i-1} B_1 \end{bmatrix} = (A_{21} + A_{22}T)(A_{11} + A_{12}T)^{i-1} B_1, \quad i \in \underline{n} \quad (4.2.18)$$

Equating (4.2.17) and (4.2.18), we obtain the following relation:

$$(A_{21} + A_{22}T - TA_{11} - TA_{12}T)(A_{11} + A_{12}T)^{i-1} B_1 = 0, \quad i \in \underline{n} \quad (4.2.19)$$

Since the matrix

$$\tilde{K} \equiv [B_1(A_{11} + A_{12}T)B_1 \quad (A_{11} + A_{12}T)^2 B_1 \quad \dots \quad (A_{11} + A_{12}T)^{n-1} B_1]$$

has rank  $r$ , its columns span an  $r$ -dimensional subspace of  $X$  which contains the rows of the matrix  $\tilde{A}_{21} \equiv A_{21} + A_{22}T - TA_{11} - TA_{12}T$ . But the

rows of  $\tilde{A}_{21}$  are contained in  $R(\tilde{K})$  and, according to (4.2.19), are also orthogonal to  $R(\tilde{K})$  which is possible if and only if  $\tilde{A}_{21} = 0$ .  $\square$

In the above analysis, we observe that the state subvector  $x^I(k) \in GF(q)^r$  and the submatrix  $B_1$  remain invariant under the action of the special isomorphism (4.2.11), that is,  $\tilde{x}^I(k) = x^I(k)$  and  $\tilde{B}_1 = B_1$ . Similarly, we see that the transformed state subvector  $\tilde{x}^{II}(k)$  is directly related to the zero-input solution of the state equation of the original LSM  $(A, B)$  since  $\tilde{x}^{II}(k) = x_{oi}^{II}(k) - Tx_{oi}^I(k)$ .

#### 4.3. State Controllability of LSMs

Earlier we defined a state controllable LSM  $(A, B)$  as one which can be driven from any initial state  $\hat{x}$  to the *zero state*  $0_X$  in a finite number of clock periods. In other words, an LSM  $(A, B)$  is state controllable if and only if there exists an input sequence  $u(0)u(1) \dots u(\ell-1) \in U^*$  such that

$$0_X = A^\ell \hat{x} + \sum_{j=0}^{\ell-1} A^{\ell-j-1} B u(j) \quad (4.3.1)$$

Rewriting (4.3.1) as

$$A^\ell \hat{x} = \sum_{j=0}^{\ell-1} A^{\ell-j-1} B [-u(j)]$$

and letting  $-u(j) \equiv \bar{u}(j)$ ,  $j \in \underline{\ell-1}$ , we see that  $\ell$ -controllability of the initial state  $\hat{x}$  reduces to the  $\ell$ -reachability of the state  $A^\ell \hat{x}$  with the input sequence  $\bar{u}(0)\bar{u}(1) \dots \bar{u}(\ell-1)$ , which implies that  $\hat{x}$  is  $\ell$ -controllable if and only if

$$A^{\ell\wedge} x \in R([A^{\ell-1}B, A^{\ell-2}B, \dots, AB, B]) \quad (4.3.2)$$

Since  $R(A^\ell) = A^\ell X \equiv \{A^{\ell\wedge} x : x \in X\}$ , in view of (4.3.2), we can say that the LSM  $(A, B)$  is  $\ell$ -state controllable if and only if there exists an integer  $\ell$  such that

$$R(A^\ell) \subseteq R([A^{\ell-1}B, A^{\ell-2}B, \dots, AB, B])$$

Lemma 4.3.1. Let  $A : X \longrightarrow X$  be a linear map. Then  $R(A^n) \subseteq R([A^{n-1}B, A^{n-2}B, \dots, AB, B])$  if and only if there exists an integer  $\ell$  such that  $R(A^\ell) \subseteq R([A^{\ell-1}B, A^{\ell-2}B, \dots, AB, B])$ .

Proof. It follows immediately from Corollary 4.1.6.  $\square$

Lemma 4.3.2. Let  $A : X \longrightarrow X$  be a linear map. Then  $R(A^n) \subseteq R(A^i)$  for all integers  $i < n$ , and  $R(A^n) = R(A^i)$  when  $i \geq n$ .

Proof. Since  $\dim X = n$  and  $A$  is linear, we have the following chain of subspaces in  $X$ :

$$\{0_X\} \subseteq \dots \subseteq R(A^{i+1}) \subseteq R(A^i) \subseteq \dots \subseteq R(A^2) \subseteq R(A) \subseteq X$$

where  $R(A^i) = A^i X$ , the range of the linear map  $A^i$ . Therefore,  $1 + \dim R(A^{i+1}) \leq \dim R(A^i)$  whenever  $R(A^{i+1}) \subset R(A^i)$ , that is, whenever the inclusion is proper. This conclusion implies that there must be an integer  $\ell$ ,  $0 \leq \ell \leq n$ , such that  $\dim R(A^i) = \dim R(A^\ell)$  for all  $i \geq \ell$ . In view of the fact that for any subspaces  $V, W \subseteq X$ ,  $\dim V = \dim W$  implies that  $V = W$ , and since  $R(A^i) \subseteq R(A^\ell)$  for all  $i \geq \ell$ , then  $R(A^i) = R(A^\ell)$  for all  $i \geq \ell$  and the range  $R(A^i)$  of  $A^i$  stops decreasing at some step  $i = \ell$ ,  $i \leq n$ . Now the result follows from Lemma 4.3.1.  $\square$

Theorem 4.3.1. The LSM  $(A, B)$  is state controllable if and only if

$$R(A^n) \subseteq R([A^{n-1}B, A^{n-2}B, \dots, AB, B])$$

Proof. It follows from Lemma 4.3.2.  $\square$

Corollary 4.3.1. If the LSM  $(A, B)$  is state reachable, then it is state controllable.

Proof. If the LSM  $(A, B)$  is state reachable, then by Theorem 4.1.4,  $R([A^{n-1}B, A^{n-2}B, \dots, AB, B]) = X$  and hence  $R(A^n) \subseteq R([A^{n-1}B, A^{n-2}B, \dots, AB, B])$  which, in view of Theorem 4.3.1, implies that  $(A, B)$  is state controllable.  $\square$

Corollary 4.3.2. If the LSM  $(A, B)$  is state controllable and  $A : X \rightarrow X$  is an isomorphism, then  $(A, B)$  is state reachable.

Proof. If  $A$  is an isomorphism, then  $R(A^n) = R(A) = X$ . Therefore, if  $(A, B)$  is state controllable, then by Theorem 4.3.1,  $R(A^n) = X \subseteq R([A^{n-1}B, A^{n-2}B, \dots, AB, B])$ . By Theorem 4.1.4,  $\dim R([A^{n-1}B, A^{n-2}B, \dots, AB, B]) = n$  which implies that  $(A, B)$  is state reachable.  $\square$

Corollary 4.3.3. Let  $S_r$  denote the subspace of reachable states and  $S_c$  the subspace of controllable states of the LSM  $(A, B)$ . Then  $S_c \supseteq S_r$ .

Corollary 4.3.4. If the characteristic matrix  $A$  of the LSM  $(A, B)$  is  $n$ -nilpotent, that is, if  $A^n = 0$ , then the LSM is state controllable.

Proof. If  $A^n = 0$ , then obviously  $R(A^n) = \{0_X\} \subset R([A^{n-1}B, A^{n-2}B, \dots, AB, B])$  and hence by Theorem 4.3.1,  $(A, B)$  is state controllable.  $\square$

It should be pointed out that an LSM can be state reachable and hence state controllable with the characteristic matrix  $A$  being neither invertible nor nilpotent. For example, the LSM

$$\left( \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right)$$

over  $GF(2)$ , is state reachable and hence state controllable, while  $A$  is neither invertible nor 2-nilpotent.

It is interesting to note that Corollary 4.3.1 and Corollary 4.3.2 point out a special feature of an LSM in that state reachability and state controllability are not necessarily the same. In marked distinction to LSMs, the concepts of state reachability and controllability are equivalent for conventional continuous-time systems regardless of the invertibility of the characteristic matrix  $A$ , because for continuous-time systems the state transition map is always invertible and the state trajectory may be solved both forward and backward in time.

### Summary and Conclusions

In this chapter a concise and formal exposition of the concepts of state reachability and state controllability for LSMs, in a state space setting, was presented and the relationship between these two distinct concepts was clearly demonstrated. In addition, the unreachability property of LSMs and some of its consequences were investigated.

In this connection, a peculiarity of the ground field which, in contrast to the case of conventional linear systems, precludes arbitrary decomposition of the state space of LSMs was pointed out. Consequently, it was shown that in order to be able to obtain a direct sum decomposition of the state space in terms of suitable constituent subspaces for the purpose of deriving the so-called unreachability canonical form for LSMs, in general, an additional assumption concerning the non-degeneracy of the usual inner product bilinear form must be made. A different derivation of the unreachability canonical form based on the linear dependence property of the rows of the state reachability matrix of unreachable LSMs was also discussed in detail (cf. [1], [2], [20], [21], [23], [37], [46], [48], [49], [58], [61], [66], [88], [105], [107], [109], [114], [118]).

As reported earlier, state controllability of LSMs was investigated in [23], [75], [105], and [107] in a rather superficial manner. The crucial distinction between the concepts of state reachability and state controllability of LSMs has not been made in any of these investigations. Consequently, this confusion has given rise to certain minor and major flaws in some of the conclusions of these studies.



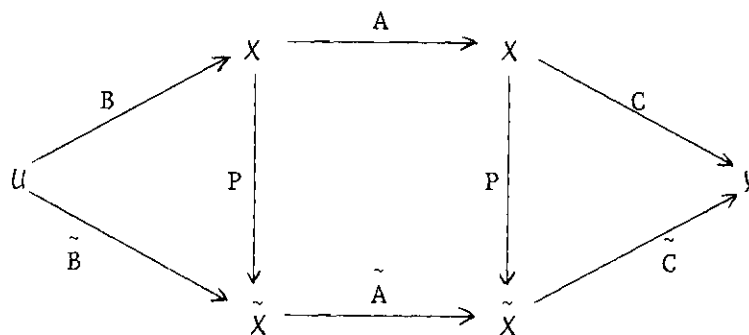
## CHAPTER V

## SOME CONSEQUENCES OF STATE REACHABILITY

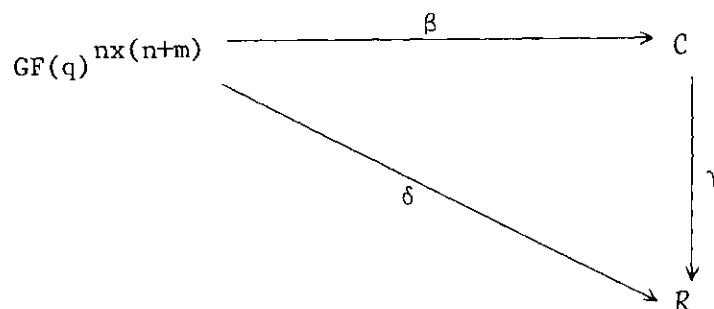
In this chapter we will discuss some implications of the property of state reachability in relation to canonical forms and state feedback in LSMs. In the area of LSMs the only canonical form that has been extensively used is the classical companion form associated with an autonomous LSM which is called a *linear feedback shift register*, and is widely used for sequence generation and coding. Canonical forms for general nonautonomous LSMs should find applications in various areas such as feedback design, observer design, data filtering, realization, and identification. In this and subsequent chapters, we will introduce a number of canonical and quasi-canonical representations for LSMs and point out some of their applications. We will conclude this chapter with a discussion of the interplay among state reachability, state feedback, and canonical forms.

5.1. Canonical LSMs

For the purpose of studying those properties of the LSM  $M = (A, B, C)$  which remain invariant under the action of an isomorphism  $P : X \longrightarrow \tilde{X}$ ,  $x(k) \longmapsto Px(k) \equiv \tilde{x}(k)$ ,  $P \in GF(n, q)$ , effective use can be made of the concept of isomorphic LSMs. We recall that two LSMs  $M = (A, B, C)$  and  $\tilde{M} = (\tilde{A}, \tilde{B}, \tilde{C})$  are said to be isomorphic if and only if there exists an isomorphism  $P : X \longrightarrow \tilde{X}$  such that the following diagram of homomorphisms is commutative:



That is,  $\tilde{M} = (\tilde{A}, \tilde{B}, \tilde{C}) = (PAP^{-1}, PB, CP^{-1})$ . The fact that the isomorphism  $P$  is arbitrary poses the obvious problem of choosing this  $P$  so as to simplify, as much as possible, the study of the LSM of interest. This problem is at least partially solved by introducing some "canonical" forms. The basic idea of this approach, therefore, is to replace  $M$  by  $\tilde{M}$  such that the characterizing matrices of  $\tilde{M}$  have certain desirable structures and properties. In order to explain this process of simplification more precisely, consider the following diagram of homomorphisms:



where  $GF(q)^{n \times (n+m)}$  is the set of all LSMs  $(A, B)$ ,  $C$  is a set of canonical forms, and  $R$  is a suitable set. The roles of the homomorphisms  $\beta$ ,  $\gamma$ , and  $\delta$  are clear: for every LSM  $(A, B) \in GF(q)^{n \times (n+m)}$ ,

$\beta$  determines a corresponding canonical form in  $\mathcal{C}$ , any property of the LSM  $(A, B)$  can be described by  $\delta$ , and  $\gamma$  is the "simpler" homomorphism that replaces  $\delta$ .

In fact, this scheme by which canonical forms are introduced is a special case of a problem of universality of universal algebra [72] which for our problem can be precisely formulated as follows: Let  $\mathcal{W}(R)$  denote the set of maps  $\delta : GF(q)^{n \times (n+m)} \rightarrow R$ ,  $(A, B) \mapsto \delta((A, B))$ , such that for every isomorphism represented by  $P \in GF(q)^{n \times n}$ , we have  $\delta((PAP^{-1}, PB)) = \delta((A, B))$ . Then our problem of universality can be stated as follows: Find a set  $G$  and a homomorphism  $\beta \in \mathcal{W}(G)$  such that the following property is true: For every set  $R$  and every homomorphism  $\delta \in \mathcal{W}(R)$  there exists exactly one homomorphism  $\gamma : \mathcal{C} \rightarrow R$  such that  $\delta = \gamma \circ \beta$ .

In order to be able to discuss in more concrete terms the important concepts of structural invariants and canonical forms for LSMs, below we will provide precise definitions of these concepts, and then specialize them in terms of certain classes of LSMs. We will clearly state the set of conditions which will serve as unambiguous qualifications for an LSM representation to be termed "canonical" which has frequently been a source of confusion in the literature.

Let  $E$  be an equivalence relation on a set  $S$ . If  $T$  is another set, a map  $f : S \rightarrow T$  is called an *invariant* of  $E$  if  $aEb$  implies that  $f(a) = f(b)$ ; it is called a *complete invariant* for  $E$  when  $aEb$  if and only if  $f(a) = f(b)$ . A list  $\{f_i, i \in \underline{\ell}\}$  of maps  $f_i : S \rightarrow T$  is called a *complete set of invariants* for  $E$  if each  $f_i$  is an invariant for  $E$  and

the assignment  $a \mapsto \{f_1(a), f_2(a), \dots, f_\ell(a)\}$  is a complete invariant  $S \rightarrow T_1 \times T_2 \times \dots \times T_\ell$ . Clearly there always exists a complete invariant for  $E$ , namely the projection  $g: S \rightarrow S/E$ , where  $S/E$  is the quotient space of  $S$  by  $E$  [72].

A set of *canonical forms* for an equivalence relation  $E$  on a set  $S$  is a subset  $C$  of  $S$  such that there is to each  $s \in S$  exactly one  $c \in C$  with  $sEc$ . This amounts to requiring that the projection  $g: S \rightarrow S/E$ , restricted to  $C$ , be an isomorphism.

In order to explain the notion of invariants and canonical representations for LSMs, suppose that the elements  $P$  of the group  $GF(n, q)$  act on the set of LSMs  $(A, B)$  according to the rule

$$(A, B) \mapsto (PAP^{-1}, PB) \equiv (\tilde{A}, \tilde{B}) \quad (5.1.1)$$

which is, of course, equivalent to a change of basis in the state space  $X$ . The  $GF(n, q)$ -orbit (equivalence class) of  $(A, B)$ , denoted by  $GF(A, B)$ , is the set of all LSMs  $(PAP^{-1}, PB)$ . Now if we can find an explicit and computable set of integers, polynomials, etc. dependent on  $(A, B)$  which (i) are preserved under the action of  $GF(n, q)$ , and (ii) allow us to decide whether the LSMs  $(A, B)$  and  $(\tilde{A}, \tilde{B})$  belong to  $GF(A, B)$ , then these integers, polynomials, etc. will constitute a complete set of invariants for  $GF(A, B)$ .

An element  $(\tilde{A}_c, \tilde{B}_c) \in GF(A, B)$  will qualify as a canonical form if it satisfies the following two requirements: (i) every LSM  $(A, B)$  of order  $n$  can be carried to  $(\tilde{A}_c, \tilde{B}_c)$  by the rule (5.1.1), and (ii)  $(\tilde{A}_c, \tilde{B}_c)$  can be completely described by a complete set of invariants for  $GF(A, B)$ .

Next, we will introduce and discuss an important set of complete structural invariants and its role in the invariant description of state reachable LSMs in terms of a very useful canonical form. In more precise terms, we will identify a map  $\beta : \{\text{state reachable LSM } (A, B) \text{ of order } n\} \longrightarrow \{\text{list of positive integers}\}$ ,  $(A, B) \longmapsto \{n_1, n_2, \dots, n_m\}$ , and show that  $\beta$  is indeed a complete orbital invariant of  $GF(n, q)$  which provides a complete parametric description of the resulting canonical LSM.

Suppose that the LSM  $M = (A, B)$  is state reachable, and let

$$\text{rank } B \equiv \text{rank } [b^1, b^2, \dots, b^m] = m \quad (5.1.2)$$

Condition (5.1.2) is imposed merely for the sake of notational simplicity, and is not an absolute requirement for the ensuing discussion. Consider the columns of the state reachability matrix  $K \equiv [B, AB, \dots, A^{n-1}B]$  of  $M = (A, B)$  in the following order:

$$b^1, b^2, \dots, b^m; Ab^1, Ab^2, \dots, Ab^m; A^2b^1, A^2b^2, \dots, A^2b^m; \dots (5.1.3)$$

Let

$$A(r, s) \equiv \{A^{\ell} b^i : \ell m + i < r m + s; i, \ell, r, s \text{ are positive integers}\} (5.1.4)$$

That is,  $A(r, s)$  is the set of all vectors  $A^{\ell} b^i$  which occur before  $A^r b^s$  in (5.1.3). Let  $n_i, i \in \underline{m}$ , denote the smallest positive integer such that  $A^{n_i-1} b^i \notin \langle A(n_i, i) \rangle$  and  $A^{n_i} b^i \in \langle A(n_i, i) \rangle$ . Therefore,  $n_i, i \in \underline{m}$ , can be characterized by the condition

$$A^{\ell} b^i \in \langle A(\ell, i) \rangle \iff \ell \geq n_i \quad (5.1.5)$$

Kalman [62] has shown that the numbers  $n_i$  are identical with a certain type of classical Kronecker invariants, namely the minimal column indices, associated with the singular pencil of matrices  $[\xi I_n - A, B]$ . The numbers  $n_i$  have also been referred to as controllability indices in the literature of linear systems. Therefore, we will refer to  $n_i$  as the *ith Kronecker invariant* and also the *ith reachability index* of the LSM  $(A, B)$ . Furthermore, let

$$R \equiv \{A^{\ell} b^j : \ell < n_j\} \quad (5.1.6)$$

Theorem 5.1.1.

$$A^{\ell} b^i \in \langle A(\ell, i) \cap R \rangle \quad \forall \ell \geq n_i, i \in \underline{m} \quad (5.1.7)$$

Proof. Consider the first vector from (5.1.3) which is not in  $R$ . This vector is necessarily of the form  $A^{n_t} b^t$ , where  $n_t m + t \leq n_{t, m} + t'$ , for all  $t' \in \underline{m}$ . But in view of (5.1.5), we have  $A^{n_t} b^t \in \langle A(n_t, t) \rangle$ . Since  $A^{n_t} b^t$  is the first vector in (5.1.3) which is not in  $R$ , it follows that  $\langle A(n_t, t) \rangle \subseteq R$ , implying that (5.1.7) holds for  $\ell = n_t$  and  $i = t$ . Now suppose that (5.1.7) is true for every vector  $A^{\ell} b^i$  such that  $\ell m + i < r m + s$  and  $\ell \geq n_i$ . Furthermore, assume that  $r \geq n_s$ . Then  $\langle A(\ell, i) \rangle \subseteq \langle A(r, s) \rangle$  and by the induction hypothesis

$$A^{\ell} b^i \in \langle A(r, s) \cap R \rangle \quad (5.1.8)$$

for all  $A^{\ell} b^i \in A(r, s)$  such that  $A^{\ell} b^i \notin R$ . Obviously (5.1.7) holds if  $A^{\ell} b^i \in A(r, s)$  such that  $A^{\ell} b^i \in R$ . Therefore, (5.1.8) is true for all  $A^{\ell} b^i \in A(r, s)$ . Hence

$$\langle A(r, s) \rangle \subseteq \langle A(r, s) \cap R \rangle \quad (5.1.9)$$

Since by assumption  $r \geq n_s$ , from (5.1.5) we have  $A^{r,s} \in \langle A(r, s) \rangle$ .

Thus (5.1.7) is true for  $\ell = r$  and  $i = s$ . Thus the theorem is proved by induction.  $\square$

Theorem 5.1.2. The elements of the set  $R$ , given by (5.1.6), are linearly independent.

Proof. Suppose that the elements of  $R$  are linearly dependent. Then there exist scalars  $e_{j\ell} \in \text{GF}(q)$ , not all zero, such that

$$\sum_{j=1}^m \sum_{\ell=1}^{n_j-1} e_{j\ell} A_{\ell}^{r,s} = 0 \quad (5.1.10)$$

Let  $A^{r,s}$  be the last vector from (5.1.3) such that  $r \leq n_s - 1$  and  $e_{sr} \neq 0$ . Then from (5.1.10) it follows that  $A^{r,s} \in \langle A(r, s) \rangle$ , implying that  $r > n_s$  which is a contradiction.  $\square$

Theorem 5.1.3. The integers  $n_i$ ,  $i \in \underline{m}$ , defined by (5.1.5), satisfy the relation

$$n_1 + n_2 + \dots + n_m = n \quad (5.1.11)$$

Proof. By Theorem 5.1.1 every vector of (5.1.3) is a linear combination of the elements of the set  $R$  given by (5.1.6). On the other hand, since the LSM  $M = (A, B)$  is reachable, there are exactly  $n$  linearly independent vectors in (5.1.3). Therefore, from Theorem 5.1.2 which shows that the cardinality of  $R$  is equal to  $n$ , and the definition of  $R$ , given by (5.1.6), we obtain (5.1.11).  $\square$

Corollary 5.1.1. There exists one set of ordered numbers

$a_{ij\ell} \in \text{GF}(q)$ , defined for  $i \in \underline{m}$ ,  $j \in \underline{i-1}$ ,  $\ell \in \min\{n_i, n_j-1\}$   
and for  $i \in \underline{m}$ ,  $j = i, i+1, \dots, m$ ;  $\ell \in \min\{n_i, n_j-1\}$ , such that

$$\begin{aligned} A_{b^i}^{n_i} &= \sum_{j=1}^{i-1} \sum_{\ell=0}^{\min\{n_i, n_j-1\}} a_{ij\ell} A_{b^j}^{\ell} \\ &+ \sum_{j=i}^m \sum_{\ell=0}^{\min\{n_i, n_j\}-1} a_{ij\ell} A_{b^j}^{\ell}, \quad i \in \underline{m} \end{aligned} \quad (5.1.12)$$

where  $n_i$ ,  $i \in \underline{m}$ , are defined by (5.1.5).

Theorem 5.1.4. The numbers  $n_i$ , defined by (5.1.5), and  $a_{ij\ell}$ , defined in Corollary 5.1.1, remain invariant under the action of the group  $\text{GF}(n, q)$  on the set of LSMs  $M = (A, B)$  according to the rule  $(A, B) \mapsto (PAP^{-1}, PB)$ .

Proof. This is obvious since  $n_i$  and  $a_{ij\ell}$  are defined only in terms of the vectors  $A_{b^j}^{\ell}$  of (5.1.3) and consequently if  $(A, B)$  is changed to  $(PAP^{-1}, PB)$ , then the vectors  $A_{b^j}^{\ell}$  will be changed to  $\tilde{A}_{b^j}^{\ell} = (PAP^{-1})^{\ell} P b^j = P A_{b^j}^{\ell} = P A_{b^j}^{\ell}$ . Clearly premultiplication of the vectors in (5.1.3) by a nonsingular matrix  $P$  does not change the numbers  $n_i$  and  $a_{ij\ell}$ .  $\square$

Theorem 5.1.5. The set of invariants  $\{n_i, a_{ij\ell}\}$  is complete. In other words, if for two LSMs  $(A, B)$  and  $(\tilde{A}, \tilde{B})$  of the same dimension the invariants  $n_i$  and  $a_{ij\ell}$  coincide, then there exists an isomorphism  $P : X \rightarrow X$ ,  $P \in \text{GF}(n, q)$ , such that  $(\tilde{A}, \tilde{B}) = (PAP^{-1}, PB)$ .



Proof. Let the invariants  $\{n_i, a_{ij\ell}\}$  and  $\{\tilde{n}_i, \tilde{a}_{ij\ell}\}$  of the LSMS  $(A, B)$  and  $(\tilde{A}, \tilde{B})$ , respectively, coincide, and consider the following matrices  $Q, \tilde{Q} \in GF(q)^{n \times n}$ :

$$Q \equiv [b^1, Ab^1, \dots, A^{n_1-1}b^1; b^2, Ab^2, \dots, A^{n_2-1}b^2; \dots, b^m, Ab^m, \dots, A^{n_m-1}b^m] \quad (5.1.13)$$

$$\tilde{Q} \equiv [\tilde{b}^1, \tilde{A}\tilde{b}^1, \dots, \tilde{A}^{\tilde{n}_1-1}\tilde{b}^1; \tilde{b}^2, \tilde{A}\tilde{b}^2, \dots, \tilde{A}^{\tilde{n}_2-1}\tilde{b}^2; \dots, \tilde{b}^m, \tilde{A}\tilde{b}^m, \dots, \tilde{A}^{\tilde{n}_m-1}\tilde{b}^m] \quad (5.1.14)$$

By Theorem 5.1.2 and Theorem 5.1.3, these matrices are nonsingular.

Therefore, we can define

$$P \equiv \tilde{Q}Q^{-1} \quad (5.1.15)$$

Thus we have

$$\tilde{Q} = PQ \quad (5.1.16)$$

In view of (5.1.13), (5.1.14), and the assumption that  $n_i = \tilde{n}_i$  and  $a_{ij\ell} = \tilde{a}_{ij\ell}$ , (5.1.16) is equivalent to

$$\tilde{A}^{\ell} \tilde{b}^j = P A^{\ell} b^j, \quad j \in \underline{m}, \quad \ell \in \underline{n_j-1} \quad (5.1.17)$$

From (5.1.12) and from the similar equation written for  $(\tilde{A}, \tilde{B})$  we obtain the relations

$$\tilde{A}^{n_j} \tilde{b}^j = P A^{n_j} b^j, \quad j \in \underline{m} \quad (5.1.18)$$

From (5.1.17), for  $\ell = 0$ , we obtain  $\tilde{B} = PB$ . From (5.1.17) and (5.1.18) it follows that

$$\tilde{\tilde{A}}Q = PAQ \quad (5.1.19)$$

Since from (5.1.16)  $Q = P^{-1}\tilde{Q}$ , (5.1.19) gives  $\tilde{A} = PAP^{-1}$ .  $\square$

Combining the above results, we conclude that for a reachable LSM  $M = (A, B)$  with  $\text{rank } B = m$ , the numbers  $\{n_i, a_{ij\ell}\}$ , where  $n_i$  are defined by (5.1.5), and  $a_{ij\ell}$  are specified in Corollary 5.1.1, constitute a complete set of invariants with respect to the action of the group  $GF(n, q)$  on  $(A, B)$  according to the rule  $(A, B) \mapsto (PAP^{-1}, PB)$ .

Next we want to relate the reachability indices  $n_i$ ,  $i \in \underline{m}$ , to a particular quasi-canonical form of state reachable LSMs. The assumption of reachability implies that there are precisely  $n$  linearly independent vectors in (5.1.3). Let us choose these  $n$  vectors, which will form a basis for  $X$ , in the following order:  $b^1, b^2, \dots, b^m$ ;  $Ab^1, Ab^2, \dots, Ab^m$ ;  $A^2b^1, A^2b^2, \dots, A^2b^m$ ;  $\dots$ ;  $A^{n-1}b^1, A^{n-1}b^2, \dots, A^{n-1}b^m$ . If a vector, say  $Ab^3$ , is skipped because of linear dependence on its predecessors, that is, on the vectors  $b^1, b^2, \dots, b^m$ ;  $Ab^1, Ab^2$ , then all vectors of the form  $A^j b^3$ ,  $j \geq 2$ , can also be skipped because by the Cayley-Hamilton Theorem they are also dependent on the previous vectors. After the  $n$  linearly independent vectors are chosen in this order, we rearrange them as follows:

$$b^1, Ab^1, \dots, A^{n_1-1}b^1; b^2, Ab^2, \dots, A^{n_2-1}b^2; \dots; b^m, Ab^m, \dots, A^{n_m-1}b^m \quad (5.1.20)$$

where  $n_1 + n_2 + \dots + n_m = n$ . The  $m$  vectors  $A^{n_i}b^i$ ,  $i \in \underline{m}$ , of (5.1.3) can be expressed in terms of the basis vectors (5.1.20) as follows:

$$A^{n_i}b^i = \sum_{j=1}^m \sum_{\ell=0}^{n_j-1} c_{ij\ell} A^{\ell}b^j, \quad i \in \underline{m} \quad (5.1.21)$$

for appropriate scalars  $c_{ij\ell} \in \text{GF}(q)$ . By direct computation, it can be easily shown that  $(A, B)$  has the following representation  $(\tilde{A}, \tilde{B})$  with respect to the basis (5.1.20):

$$\tilde{A} \equiv \left( \begin{array}{cccccc|cccc|cccc|cccc} 0 & 0 & . & . & . & 0 & * & & & & & & & & & * & & & * \\ 1 & 0 & . & . & . & 0 & * & & & & & & & & & * & & & * \\ 0 & 1 & . & . & . & 0 & * & & & & & & & & & * & & & * \\ \vdots & \vdots & & & & \vdots & \vdots & & & & & & & & & \vdots & & & \vdots \\ 0 & 0 & . & . & . & 1 & * & & & & & & & & & * & & & * \\ \hline & & & & & & * & 0 & 0 & . & . & . & 0 & * & & & * & & & \\ & & & & & & * & 1 & 0 & . & . & . & 0 & * & & & * & & & \\ & & & & & & * & 0 & 1 & . & . & . & 0 & * & & & * & & & \\ & & & & & & \vdots & \vdots & \vdots & & & \vdots & \vdots & & & \vdots & & & \\ & & & & & & * & 0 & 0 & . & . & . & 1 & * & & & * & & & \\ \hline & & & & & & \vdots & & & & & & & & & \vdots & & & \\ & & & & & & \vdots & & & & & & & & & \vdots & & & \\ & & & & & & * & & & & & & & 0 & 0 & . & . & . & 0 & * \\ & & & & & & * & & & & & & & 1 & 0 & . & . & . & 0 & * \\ & & & & & & * & & & & & & & 0 & 1 & . & . & . & 0 & * \\ & & & & & & \vdots & & & & & & & \vdots & \vdots & & \vdots & \vdots & \\ & & & & & & \vdots & & & & & & & \vdots & \vdots & & \vdots & \vdots & \\ & & & & & & * & & & & & & & 0 & 0 & . & . & . & 1 & * \end{array} \right).$$

(5.1.22)

$$\tilde{B} \equiv \begin{pmatrix} 1 & 0 & . & . & . & 0 & 0 \\ 0 & 0 & . & . & . & 0 & 0 \\ \vdots & \vdots & & & & \vdots & \vdots \\ 0 & 0 & . & . & . & 0 & 0 \\ \hline 0 & 1 & . & . & . & 0 & 0 \\ 0 & 0 & . & . & . & 0 & 0 \\ \vdots & \vdots & & & & \vdots & \vdots \\ 0 & 0 & . & . & . & 0 & 0 \\ \hline \vdots \\ \hline 0 & 0 & . & . & . & 0 & 1 \\ 0 & 0 & . & . & . & 0 & 0 \\ \vdots & \vdots & & & & \vdots & \vdots \\ 0 & 0 & . & . & . & 0 & 0 \end{pmatrix} \quad (5.1.23)$$

The entries of  $\tilde{A}$  marked \* are given by the coefficients  $c_{ijl}$  of the linear combination (5.1.21).

The above representation  $(\tilde{A}, \tilde{B})$  is called the Luenberger "canonical" form [71] in the area of linear systems theory. Clearly the LSM  $(\tilde{A}, \tilde{B})$  is not in canonical form as it obviously does not satisfy the requirements of a canonical form discussed earlier in this section. A more appropriate adjective to describe  $(\tilde{A}, \tilde{B})$  would be "quasi-canonical." This type of confusion concerning the term "canonical" is rather common in the literature of linear systems theory.

However, the above quasi-canonical form can be easily transformed to a canonical form via state feedback homomorphisms of the type  $(\tilde{A}, \tilde{B}) \mapsto (\tilde{A} + \tilde{B}F, \tilde{B}G)$ , where  $F$  and  $G$  are appropriate matrices, which will make zero all the \* entries of the matrix  $\tilde{A}$ . We will explain this

procedure by first considering a transposed version of  $\tilde{A}$  through a new basis. As a result of this method, we will derive Brunovsky's canonical form [15] for LSMs.

Let  $v^{ijT}$ ,  $i \in \underline{m}$ ,  $j \in \underline{n_i}$ , denote the rows of the matrix  $P^{-1}$  formed by the basis vectors (5.1.20), and write  $P^{-1}$  in terms of its row vectors as

$$P^{-1} = \begin{pmatrix} v^{11T} \\ v^{12T} \\ \vdots \\ v^{1n_1T} \\ v^{21T} \\ v^{22T} \\ \vdots \\ v^{2n_2T} \\ \vdots \\ v^{m1T} \\ v^{m2T} \\ \vdots \\ v^{mn_mT} \end{pmatrix}$$

Now let  $v_\ell = n_1 + n_2 + \dots + n_\ell$ ,  $\ell \in \underline{m}$ . Using the  $v_\ell$ th rows of  $P^{-1}$  we form the following matrix:

$$P_0 \equiv \begin{pmatrix} 1n_1^T \\ v \\ 1n_1^T \\ v \quad A \\ \cdot \\ \cdot \\ \cdot \\ 1n_1^T \quad n_1^{-1} \\ v \quad A \quad n_1^{-1} \\ 2n_2^T \\ v \\ 2n_2^T \\ v \quad A \\ \cdot \\ \cdot \\ \cdot \\ 2n_2^T \quad n_2^{-1} \\ v \quad A \quad n_2^{-1} \\ \cdot \\ \cdot \\ \cdot \\ mn_m^T \\ v \\ mn_m^T \\ v \quad A \\ \cdot \\ \cdot \\ \cdot \\ mn_m^T \quad n_m^{-1} \\ v \quad A \quad n_m^{-1} \end{pmatrix} \quad (5.1.24)$$

We want to show that the rows of  $P_0$  are linearly independent and hence constitute a basis for  $X$ . Suppose that the rows of  $P_0$  are linearly dependent. Then there exist scalars  $c_{ij} \in GF(q)$ ,  $i \in \underline{m}$ ,  $j \in \underline{n_i}$ , not all zero, such that

$$\sum_{i=1}^m \sum_{j=1}^{n_i} c_{ij} v^{in_i T} A^{j-1} = 0 \quad (5.1.25)$$

Taking the inner product of both sides of (5.1.25) with  $b^r$  yields

$$c_{rn_r} = 0 \quad (5.1.26)$$

since by the definition of  $v^{in_i T}$  each term in the inner product is zero except the one involving  $v^{rn_r T} A^{r-1} b^r$  which is unity. In view of (5.1.26), (5.1.25) can be written equivalently as

$$\sum_{i=1}^m \sum_{j=1}^{n_i-1} c_{ij} v^{in_i T} A^{j-1} = 0$$

Taking the inner product of both sides of this equation with  $Ab^r$  produces  $c_{r, n_r-1} = 0$ . Continuing in this manner, by induction it is seen that each  $c_{ij} = 0$ ,  $i \in \underline{m}$ ,  $j \in \underline{n_i}$ , which is a contradiction. Therefore, the rows of  $P_o$  are linearly independent. Now by direct computation, it is easy to see that the isomorphic LSM  $\tilde{M} = (\tilde{A}, \tilde{B}) \equiv (P_o A P_o^{-1}, P_o B)$  has the following form:





$$\tilde{B} = \begin{pmatrix} 0 & 0 & . & . & . & 0 & 0 \\ 0 & 0 & . & . & . & 0 & 0 \\ . & . & & & & . & . \\ . & . & & & & . & . \\ . & . & & & & . & . \\ 0 & 0 & . & . & . & 0 & 0 \\ 1 & 0 & . & . & . & 0 & 0 \\ \hline 0 & 0 & . & . & . & 0 & 0 \\ 0 & 0 & . & . & . & 0 & 0 \\ . & . & & & & . & . \\ . & . & & & & . & . \\ . & . & & & & . & . \\ 0 & 0 & . & . & . & 0 & 0 \\ 0 & 1 & . & . & . & 0 & 0 \\ \hline . & & & & & & \\ . & & & & & & \\ . & & & & & & \\ \hline 0 & 0 & . & . & . & 0 & 0 \\ 0 & 0 & . & . & . & 0 & 0 \\ . & . & & & & . & . \\ . & . & & & & . & . \\ . & . & & & & . & . \\ 0 & 0 & . & . & . & 0 & 1 \end{pmatrix} \quad (5.1.27b)$$

where the \* entries of the matrix  $\tilde{A}$  are given by the coefficients  $d_{ij\ell}$  of linear combinations of the type

$${}^i n_i^T A^i = \sum_{j=1}^m \sum_{\ell=1}^{n_j} d_{ij\ell} {}^j n_j^T A^{\ell-1}, \quad i \in \underline{m} \quad (5.1.28)$$

It is obvious that the basis (5.1.24) and the constants  $d_{ij\ell}$  are not unique and consequently  $(\tilde{A}, \tilde{B})$ , given by (5.1.27), does not constitute a completely invariant description of the original LSM  $(A, B)$ . However, we can readily see that the constants  $d_{ij\ell}$  in  $\tilde{A}$  can be made zero by applying a feedback law of the form  $u(k) = Fx(k) + w(k)$  to the LSM  $(\tilde{A}, \tilde{B})$ , transforming it to  $(\tilde{A} + \tilde{B}F, \tilde{B})$ . We are keeping  $\tilde{B}$  unaltered since it is already in the desired form. In order to pick a matrix  $F \in GF(q)^{m \times n}$  to do the job, let  $f_{st}$ ,  $s \in \underline{m}$ ,  $t \in \underline{n}$ , denote the entries of  $F$  and observe that the product matrix  $\tilde{B}F$  has the following form:

$$\tilde{B}F \equiv \begin{pmatrix} 0 & 0 & 0 & \dots & 0 & 0 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots & & \vdots & \vdots & & \vdots \\ \vdots & \vdots & \vdots & & \vdots & \vdots & & \vdots & \vdots & & \vdots \\ \vdots & \vdots & \vdots & & \vdots & \vdots & & \vdots & \vdots & & \vdots \\ n_1\text{th row} & f_{11} & f_{12} & f_{13} & \dots & f_{1n_1} & f_{1,n_1+1} & \dots & f_{1n_2} & f_{1,n_2+1} & \dots & f_{1n_m} \\ 0 & 0 & 0 & \dots & 0 & 0 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots & & \vdots & \vdots & & \vdots \\ \vdots & \vdots & \vdots & & \vdots & \vdots & & \vdots & \vdots & & \vdots \\ \vdots & \vdots & \vdots & & \vdots & \vdots & & \vdots & \vdots & & \vdots \\ \tilde{B}F \equiv & 0 & 0 & 0 & \dots & 0 & 0 & \dots & 0 & 0 & \dots & 0 \\ & 0 & 0 & 0 & \dots & 0 & 0 & \dots & 0 & 0 & \dots & 0 \\ (n_1+n_2)\text{th row} & f_{21} & f_{22} & f_{23} & \dots & f_{2n_1} & f_{2,n_1+1} & \dots & f_{2n_2} & f_{2,n_2+1} & \dots & f_{2n_m} \\ & 0 & 0 & 0 & \dots & 0 & 0 & \dots & 0 & 0 & \dots & 0 \\ & & & & & & \vdots & & & & & \\ & & & & & & \vdots & & & & & \\ & & & & & & \vdots & & & & & \\ & 0 & 0 & 0 & \dots & 0 & 0 & \dots & 0 & 0 & \dots & 0 \\ nth row & f_{m1} & f_{m2} & f_{m3} & \dots & f_{mn_1} & f_{m,n_1+1} & \dots & f_{mn_2} & f_{m,n_2+1} & \dots & f_{mn_m} \end{pmatrix}$$

(5.1.29)

Therefore, if we choose

$$\begin{aligned}
 & [f_{i1}, f_{i2}, \dots, f_{in_1}; f_{i, n_1+1}, \dots, f_{in_2}; \dots, f_{in_m}] \\
 & = -[d_{i11}, d_{i12}, \dots, d_{i1n_1}; d_{i21}, d_{i22}, \dots, d_{i2n_2}; \dots, d_{imn_m}], \\
 & \quad i \in \underline{m} \tag{5.1.30}
 \end{aligned}$$

then the resulting matrix  $\tilde{A} + \tilde{B}F$  will have precisely the form of  $\tilde{A}$  with the \* entries equal to zero. Since the above procedure can always be effected under the state reachability assumption, in effect it provides a new derivation of an important canonical form initially introduced by Brunovsky [15] for continuous-time dynamical systems.

We summarize the above result as a theorem.

Theorem 5.1.6. If the LSM  $M = (A, B)$  is reachable, then it has the following canonical form  $\tilde{M} = (\tilde{A}_c, \tilde{B}_c)$ :

$$\begin{aligned}
 \tilde{A}_c & \equiv \tilde{A}_1 \oplus \tilde{A}_2 \oplus \dots \oplus \tilde{A}_{n_m} \\
 \tilde{B}_c & \equiv \tilde{b}_1 \oplus \tilde{b}_2 \oplus \dots \oplus \tilde{b}_{n_m}
 \end{aligned} \tag{5.1.31}$$

where  $\tilde{A}_i \in GF(2)^{n_i \times n_i}$ ,  $\tilde{b}_i \in GF(2)^{n_i}$ ,  $i \in \underline{m}$

$$\tilde{A}_i \equiv \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ 0 & 0 & 0 & \dots & 0 \end{pmatrix}; \quad \tilde{b}_i \equiv \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} \tag{5.1.32}$$

Therefore, any state reachable LSM  $(A, B)$  over  $GF(q)$  is equivalent to the following  $m$  completely decoupled submachines over  $GF(2)$ , each of order  $n_i$ ,  $i \in \underline{m}$ :

$$\begin{aligned}
 x_1^{(i)}(k+1) &= x_2^{(i)}(k) \\
 x_2^{(i)}(k+1) &= x_3^{(i)}(k) \\
 &\vdots \\
 x_{n_i-1}^{(i)}(k+1) &= x_{n_i}^{(i)}(k) \\
 x_{n_i}^{(i)}(k+1) &= w_i(k), \quad i \in \underline{m}
 \end{aligned} \tag{5.1.33}$$

The submachines (5.1.33) can be realized by a parallel array of circuits of the form

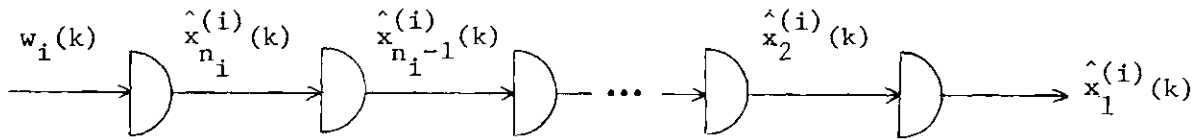


Fig. 5.1.1. Compound Circuits of an LSM in Canonical Form.

The representation (5.1.31) is truly canonical since it can be described completely only by the reachability indices  $n_i$ ,  $i \in \underline{m}$ .

The numbers  $n_i$ ,  $i \in \underline{m}$ , as defined in (5.1.5) do not satisfy any order relations. However, these numbers can be redefined such that they can be ordered [15]. To see this, let

$$r_0 \equiv \text{rank } B$$

$$r_j \equiv \text{rank}[B, AB, A^2B, \dots, A^jB] - \text{rank}[B, AB, A^2B, \dots, A^{j-1}B], \quad j \in \underline{n-1}$$

Since  $\text{rank } B \leq m$ , it is clear that for any state reachable LSM  $(A, B)$

we have  $0 \leq r_j \leq m$ ,  $j \in \underline{n-1}$ , and  $r_0 + r_1 + \dots + r_{n-1} = n$ .

Alternatively, if we define

$$L_j \equiv R(B) + AR(B) + A^2R(B) + \dots + A^jR(B), \quad j \in \underline{n-1}$$

and let  $\Pi_j : L_j \longrightarrow L_{j-1}^\perp$  be the orthogonal projection, then the integers  $r_j$ ,  $j \in \underline{n-1}$ , can be equivalently characterized as follows:

$$r_0 \equiv \text{rank } B$$

$$r_j \equiv \dim L_{j-1}^\perp = \text{rank}[\Pi_j(A^jB)] = \dim(L_j/L_{j-1}), \quad j \in \underline{n-1}$$

where  $L_j/L_{j-1}$  is the quotient space of  $L_j$  by  $L_{j-1}$ . Since if the vector  $A^j b^i$  can be expressed as a linear combination of the set  $\{A^s b^s, s \in \underline{m}\}$ , so can  $A^{j+1} b^i$ , it follows that  $r_0 \geq r_1 \geq \dots \geq r_{n-1}$ , and that we can choose a basis  $B$  of  $X$  from the columns of  $[B, AB, A^2B, \dots, A^{n-1}B]$  such that  $A^j b^i \notin B \implies A^{j+1} b^i \notin B$ . Now if we associate a number  $n'_i$  with every vector  $b^i$  such that  $A^j b^i \in B$ ,  $j \in \underline{n_i-1}$ , but  $A^{n'_i} b^i \notin B$ , then in view of (5.1.31), the numbers  $n'_i$ ,  $i \in \underline{m}$ , are precisely the reachability indices of the LSM  $(A, B)$ . Performing an input coordinate transformation, if necessary, it is possible to have  $n'_1 \geq n'_2 \geq \dots \geq n'_m$ . Consequently  $n'_1$  is the reachability index (Definition 4.1.1) of

the LSM  $(A, B)$ . Furthermore, from the above discussion it follows that  $n'_i, i \in \underline{m}$ , can be uniquely characterized in terms of the numbers  $r_j, j \in \underline{n-1}$ , by defining  $n'_i \equiv$  number of integers in the set  $\{r_0, r_1, \dots, r_{n-1}\}$  which are  $\geq i, i \in \underline{m}$ . Thus, in view of the fact that  $r_0 + r_1 + \dots + r_{n-1} = n$ , we have  $n'_1 + n'_2 + \dots + n'_m = n$ .

The invariants  $n_i$  along with  $a_{ij\ell}$  defined in Corollary 5.1.1, and their relation to canonical forms of conventional linear systems have been studied in detail by Popov [90] through universal algebraic methods. For linear systems, invariants and canonical forms have also been investigated in terms of certain polynomial matrices [19], [52], [96], [97]. Using the polynomial representation of LSMs discussed in Section 3.6, and giving due consideration to the properties of polynomials and polynomial matrices over  $GF(q)[\xi]$ , many results of these investigations can be similarly developed for LSMs.

In the above discussion, the canonical representation (5.1.31) was obtained from the given LSM under two consecutive transformations, namely state coordinate transformation and state feedback transformation. These transformations are special members of a relatively more general transformation group which can be utilized for the study of isomorphic LSMs. Here we will briefly discuss this particular group of transformations.

Let

$$M \equiv \{\text{LSM } M = (A, B, C) : A : X \longrightarrow X, B : U \longrightarrow X, C : X \longrightarrow Y\} \quad (5.1.34)$$

and

$$G_1 \equiv \{(P, F, G) : P : X \longrightarrow X, P \in GF(n, q),$$

$$F : X \longrightarrow U, F \in GF(q)^{m \times n}, \text{ and}$$

$$G : U \longrightarrow U, G \in GL(m, q)\} \quad (5.1.35)$$

Suppose that the elements of the class  $G_1$  act on the elements of the set  $M$  of all LSMs of dimension  $n$  according to the rule

$$(A, B, C) \longmapsto (P(A + BF)P^{-1}, PBG, CP^{-1}) \quad (5.1.36)$$

That is,  $G_1$  is the set of state coordinate, input coordinate, and state feedback transformations. It can be easily shown that the rule (5.1.36) assigns to  $G_1$  the structure of a transformation group with

$$\text{identity: } (I_n, 0, I_m)$$

$$\text{inverse: } (P, -G^{-1}GP^{-1}, G^{-1})$$

and

$$\text{composition rule: } (P_2, F_2, G_2) \circ (P_1, F_1, G_1) = (P_1^{-1}P_2, F_1 + G_1F_2P_1, G_1G_2)$$

Thus an equivalence relation on  $M$  with respect to  $G_1$  may be defined as follows: Two LSMs  $(A_1, B_1, C_1), (A_2, B_2, C_2) \in M$  are said to be  $G_1$ -equivalent if and only if there exists a triple  $(P, F, G) \in G_1$  such that  $(A_1, B_1, C_1) \longmapsto (P(A_2 + B_2F)P^{-1}, PB_2G, C_2P^{-1})$ . The following two transformation groups, which are special cases of  $G_1$ , are also of interest in the study of isomorphic LSMs:

$$G_2 = \{(P, 0, I_n) : P : X \longrightarrow X, P \in GL(n, q)\} \quad (5.1.37)$$

$$G_3 = \{(P, 0, G) : P : X \longrightarrow X, P \in GL(n, q),$$

$$G : U \longrightarrow U, G \in GL(m, q)\} \quad (5.1.38)$$

Clearly the elements of  $G_2$  and  $G_3$  act on the elements of  $M$  according to the rules

$$(A, B, C) \longmapsto (PAP^{-1}, PB, CP^{-1})$$

and

$$(A, B, C) \longmapsto (PAP^{-1}, PBG, CP^{-1})$$

That is,  $G_2$  is the group of state coordinate transformations and  $G_3$  is the group of state coordinate and input coordinate transformations. The equivalence relations  $G_2$ -equivalence and  $G_3$ -equivalence can be defined similar to  $G_1$ -equivalence.

Later in the sequel, we will have occasion to examine further properties of  $G_1$ -equivalence. However, in the remainder of this section we will restrict our attention to some canonical representations under the transformation group  $G_2$ .

One of the important properties of LSMs which remains invariant under state isomorphism is state reachability as shown in the following theorem.



Theorem 5.1.7. Let  $P : X \longrightarrow X$ ,  $x \longmapsto Px \equiv \tilde{x}$ ,  $P \in GL(n, q)$ , be an isomorphism. Then the LSM  $M = (A, B)$  is state reachable if and only if the isomorphic LSM  $\tilde{M} = (\tilde{A}, \tilde{B}) \equiv (PAP^{-1}, PB)$  is state reachable.

Proof. By Theorem 4.1.4,  $\tilde{M}$  is state reachable if and only if  $\text{rank}[\tilde{B}, \tilde{A}\tilde{B}, \dots, \tilde{A}^{n-1}\tilde{B}] = n$ . But

$$\begin{aligned} [\tilde{B}, \tilde{A}\tilde{B}, \dots, \tilde{A}^{n-1}\tilde{B}] &= [PB, PAP^{-1}PB, \dots, PA^{n-1}P^{-1}PB] \\ &= P[B, AB, \dots, A^{n-1}B] \end{aligned}$$

Since  $P$  is nonsingular, it follows that

$$\text{rank}[\tilde{B}, \tilde{A}\tilde{B}, \dots, \tilde{A}^{n-1}\tilde{B}] = \text{rank}[B, AB, \dots, A^{n-1}B] = n. \quad \square$$

The state reachability property is also invariant under input coordinate transformation and hence under transformations of the form  $(A, B) \longmapsto (PAP^{-1}, PBG)$ , where  $G : U \longrightarrow U$ ,  $G \in GL(m, q)$ . This is obvious from the proof of Theorem 5.1.7 since

$$\begin{aligned} \text{rank}[\tilde{B}, \tilde{A}\tilde{B}, \dots, \tilde{A}^{n-1}\tilde{B}] &= \text{rank } P[B, AB, \dots, A^{n-1}B]G \\ &= \text{rank}[B, AB, \dots, A^{n-1}B] = n \end{aligned}$$

A more geometric proof of the above result is given below.

$$\begin{aligned} \{\tilde{A} \mid R(\tilde{B})\} &= \{PAP^{-1} \mid R(PBG)\} = \sum_{j=1}^m (PAP^{-1})^{j-1} R(PBG) \\ &= \sum_{j=1}^n PA^{j-1}P^{-1}PR(BG) \\ &= P \sum_{j=1}^n A^{j-1}R(B) = PX = X \end{aligned}$$

since  $M = (A, B)$  is reachable, that is, since

$$R([B, AB, \dots, A^{n-1}B]) \equiv \sum_{j=1}^n A^{j-1}R(B) \equiv \{A \mid R(B)\} = X$$

This result can be utilized to study other aspects of machine state reachability and further related properties of LSMs in more convenient forms. For example, we may want to employ some special isomorphism to transform a given LSM to some simpler canonical form such as the Jordan canonical form or Luenberger quasi-canonical form where we might obtain much simpler criteria for checking reachability or, in conjunction with the capabilities provided by state and/or output feedback, we may be able to design efficient state estimators or investigate noninteraction properties. In the sequel we will introduce several canonical forms for LSMs and study their structures and properties. However, first we would like to briefly survey the existing canonical forms and their use in the area of linear sequential machines.

We know that the characteristic matrix  $A$  of an LSM  $(A, B)$  plays a key role in determining its operational structure because it describes the interconnections among the storage devices (delayers). Therefore, it is natural to search for a "modified" characteristic matrix  $\tilde{A}$  that will have certain desirable attributes and yet be "similar," in some sense, to  $A$ . This is, of course, the old problem of similarity transformations in linear algebra, since these transformations change the form of a matrix  $A$  but not its characteristic polynomial which embodies most of the essential properties of  $A$ . This idea has been used in the

area of LSMs for the purpose of designing suitable data storage interconnection layouts, traditionally called *linear feedback shift registers*. Generally speaking, a shift register is a device by means of which digital information can be stored temporarily while--during the processing operation--the information is transformed from one memory element to the next under the control of a clock pulse. In connection with LSMs, a feedback shift register is a circuit configuration that realizes the companion form of the characteristic matrix of an internal LSM (ILSM). That is, a linear feedback shift register (LFSR) is an ILSM  $x(k+1) = Ax(k)$  having any one of the following forms:

$$\begin{pmatrix} x_1(k+1) \\ x_2(k+1) \\ \vdots \\ x_n(k+1) \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ a_0 & a_1 & a_2 & \dots & a_{n-1} \end{pmatrix} \begin{pmatrix} x_1(k) \\ x_2(k) \\ \vdots \\ x_n(k) \end{pmatrix} \quad (5.1.39a)$$

$$\begin{pmatrix} x_1(k+1) \\ x_2(k+1) \\ \vdots \\ x_n(k+1) \end{pmatrix} = \begin{pmatrix} a_{n-1} & a_{n-2} & a_{n-3} & \dots & a_1 & a_0 \\ 1 & 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & 0 \end{pmatrix} \begin{pmatrix} x_1(k) \\ x_2(k) \\ \vdots \\ x_n(k) \end{pmatrix} \quad (5.1.39b)$$

$$\begin{pmatrix} x_1(k+1) \\ x_2(k+1) \\ \vdots \\ \vdots \\ x_n(k+1) \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & \dots & 0 & a_0 \\ 1 & 0 & 0 & \dots & 0 & a_1 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & a_{n-1} \end{pmatrix} \begin{pmatrix} x_1(k) \\ x_2(k) \\ \vdots \\ \vdots \\ x_n(k) \end{pmatrix} \quad (5.1.39c)$$

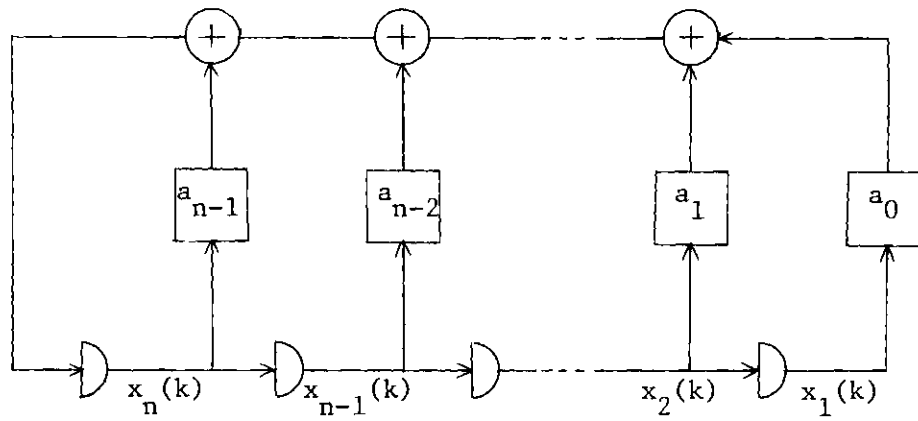
$$\begin{pmatrix} x_1(k+1) \\ x_2(k+1) \\ \vdots \\ \vdots \\ x_n(k+1) \end{pmatrix} = \begin{pmatrix} a_{n-1} & 1 & 0 & \dots & 0 \\ a_{n-2} & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ \vdots & \vdots & \vdots & & \vdots \\ a_1 & 0 & 0 & \dots & 1 \\ a_0 & 0 & 0 & \dots & 0 \end{pmatrix} \begin{pmatrix} x_1(k) \\ x_2(k) \\ \vdots \\ \vdots \\ x_n(k) \end{pmatrix} \quad (5.1.39d)$$

It is easily seen that the LFSRs (5.1.39) have the same characteristic polynomial given by

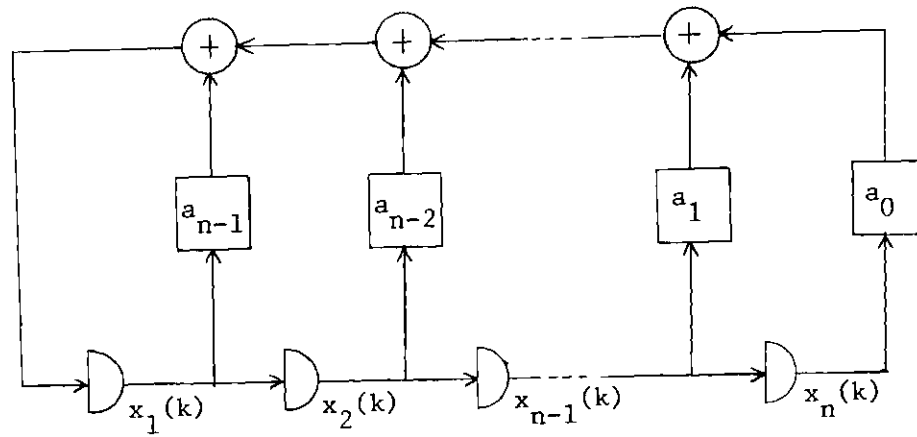
$$f_c(\lambda) = (\lambda)^n - a_{n-1}(\lambda)^{n-1} - a_{n-2}(\lambda)^{n-2} - \dots - a_1\lambda - a_0$$

That is, the companion matrices of these shift registers are similar to one another and hence any one of them can be transformed to any other by some suitable similarity transformation.

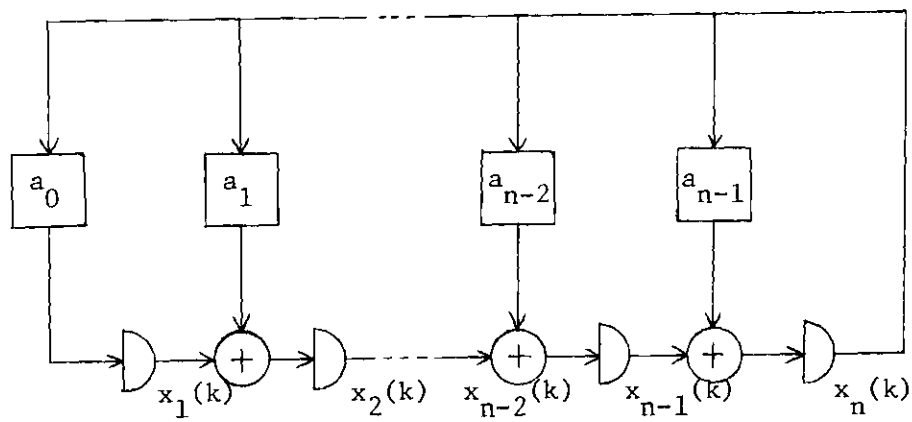
Realization circuits for the LFSRs (5.1.39) are shown in Fig. 5.1.2.



(a)



(b)



(c)

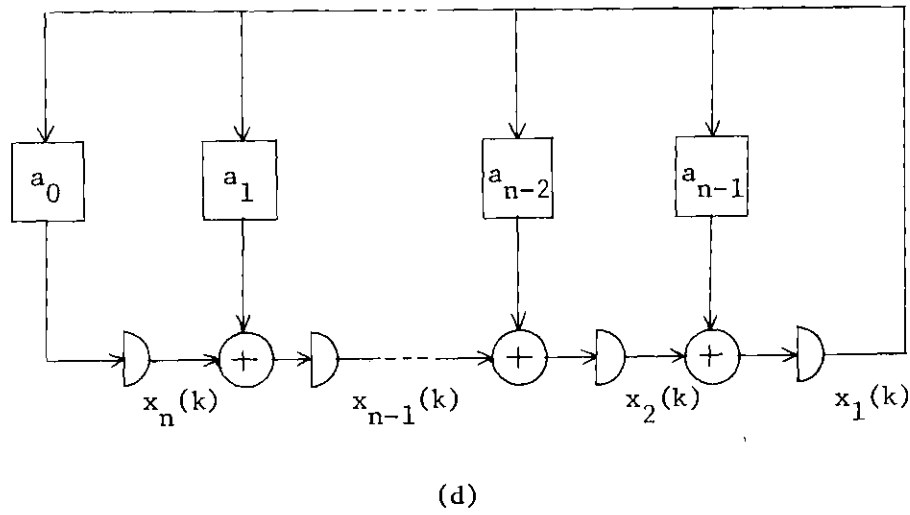


Fig. 5.1.2. Realization Circuits for LFSRs

The existence of the canonical forms (5.1.39) is guaranteed by a classical result of linear algebra which we will briefly discuss in terms of ILSMs.

Suppose that the endomorphism  $A : X \rightarrow X$  is cyclic, that is, the minimal polynomial  $f_m(\lambda)$  of  $A$  is equal to its characteristic polynomial  $f_c(\lambda)$  (the matrix  $A$  is nonderogatory) or equivalently, there exists a  $z \in X$  such that the vectors  $z, Az, A^2z, \dots, A^{n-1}z$  form a basis for  $X$ . The vector  $z$  is called a (cyclic) *generator* for  $X$  (relative to  $A$ ). The set of all generators coincide with the set of vectors  $g(A)z$ , where  $g(\lambda) \in \text{GF}(q)[\lambda]$  is coprime with  $f_m(\lambda)$ .

Suppose that  $A$  is cyclic with generator  $z$ , and let

$$f_m(\lambda) = f_c(\lambda) = (\lambda)^n - a_{n-1}(\lambda)^{n-1} - a_{n-2}(\lambda)^{n-2} - \dots - a_1\lambda - a_0 \quad (5.1.40)$$

Defining the auxiliary polynomials

$$\begin{aligned}
 f^{(0)}(\lambda) &\equiv f_m(\lambda) \\
 f^{(1)}(\lambda) &\equiv (\lambda)^{n-1} - a_1 - a^2\lambda - \dots - a_{n-1}(\lambda)^{n-2} \\
 f^{(n-1)}(\lambda) &\equiv \lambda - a_{n-1} \\
 f^{(n)}(\lambda) &\equiv 1
 \end{aligned} \tag{5.1.41}$$

it is easy to see that  $f^{(i)}(\lambda)$ ,  $i \in \underline{n}$ , satisfy the recursion relation

$$\lambda f^{(i)}(\lambda) = f^{(i-1)}(\lambda) + a_{i-1} f^{(n)}(\lambda), \quad i \in \underline{n} \tag{5.1.42}$$

It is clear that the set of vectors

$$e^i \equiv f^{(i)}(A)z, \quad i \in \underline{n} \tag{5.1.43}$$

where  $e^0 \equiv 0$ , forms a basis for  $X$ . From (5.1.42) and (5.1.43) we obtain the following relations:

$$Ae^i = e^{i-1} + a_{i-1}e^n, \quad i \in \underline{n} \tag{5.1.44}$$

From (5.1.44) it follows that with respect to the basis (5.1.43), the isomorphic ILSM  $\tilde{x}(k+1) = \tilde{\tilde{A}}\tilde{x}(k)$  has the companion form representation given by (5.1.39a). That is, if  $A$  is cyclic, there exists an isomorphism  $P : X \rightarrow X$  such that the isomorphic ILSM  $\tilde{x}(k+1) = \tilde{\tilde{A}}\tilde{x}(k) \equiv (PAP^{-1})\tilde{x}(k)$  has the form (5.1.39a).

This result can be extended to the general case of a noncyclic (derogatory) ILSM in which case the isomorphic ILSM has the form

$$\begin{pmatrix} \tilde{x}^1(k+1) \\ \tilde{x}^2(k+2) \\ \vdots \\ \tilde{x}^r(k+1) \end{pmatrix} = \begin{pmatrix} \tilde{A}_1 & & & \\ & \tilde{A}_2 & & \\ & & \ddots & \\ & & & \tilde{A}_r \end{pmatrix} \begin{pmatrix} \tilde{x}^1(k) \\ \tilde{x}^2(k) \\ \vdots \\ \tilde{x}^r(k) \end{pmatrix} \quad (5.1.45)$$

where  $\tilde{x}^i(k) \in \text{GF}(q)^{s_i}$  and  $\tilde{A}_i \in \text{GF}(q)^{s_i \times s_i}$ ,  $i \in \underline{r}$ , are the companion matrices associated with the elementary divisors  $[f_i(\lambda)]^{e_i}$ ,  $i \in \underline{r}$ , of  $A$ , that is,

$$\tilde{A}_i \equiv \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ a_{i0} & a_{i1} & a_{i2} & \dots & a_{i,s_i-1} \end{pmatrix}, \quad i \in \underline{r} \quad (5.1.46)$$

$$f_i(\lambda) = (\lambda)^{s_i} - a_{i,s_i-1}(\lambda)^{s_i-1} - \dots - a_{i1}\lambda - a_{i0}, \quad i \in \underline{r} \quad (5.1.47)$$

$$f_c(\lambda) = [f_1(\lambda)]^{e_1} [f_2(\lambda)]^{e_2} \dots [f_r(\lambda)]^{e_r} \quad (5.1.48)$$

The matrix  $\tilde{A} \equiv \tilde{A}_1 \oplus \tilde{A}_2 \oplus \dots \oplus \tilde{A}_r$  in (5.1.45) is called the *rational canonical form* of  $A$ . Systematic procedures exist for computing



the rational canonical form. Variations of this form in terms of invariant factors of A and the Smith canonical form also exist, but will not be considered here.

The preceding result shows that every arbitrary ILSM is isomorphic to an ILSM that is composed entirely of uncoupled LFSRs. Therefore, the problem of analyzing the state behavior of arbitrary ILSMs reduces to investigating LFSRs which have much simpler structure and lead to efficient and economical synthesis and particularly simple physical realization.

The classical canonical form has also been used for the purpose of investigating certain aspects of LSMs [49]. If an ILSM is cyclic and its minimal polynomial can be expressed as  $f_m(\lambda) = [f(\lambda)]^e$ , where  $f(\lambda)$  is an irreducible polynomial, then the isomorphic ILSM has the following *hypercompanion* form:

$$\begin{bmatrix} \tilde{x}^1(k+1) \\ \tilde{x}^2(k+1) \\ \vdots \\ \tilde{x}^{e-1}(k+1) \\ \tilde{x}^e(k+1) \end{bmatrix} = \begin{bmatrix} \tilde{Q} & \tilde{R} & & & \\ & \tilde{Q} & \tilde{R} & & \\ & & \ddots & \ddots & \\ & & & \tilde{Q} & \tilde{R} \\ & & & & \tilde{Q} \end{bmatrix} \begin{bmatrix} \tilde{x}^1(k) \\ \tilde{x}^2(k) \\ \vdots \\ \tilde{x}^{e-1}(k) \\ \tilde{x}^e(k) \end{bmatrix} \quad (5.1.49)$$

where  $\tilde{Q} \in GF(q)^{\ell \times \ell}$  is the companion matrix of  $f(\lambda)$ , and  $\tilde{R} \in GF(2)^{\ell \times \ell}$ ,  $\ell = \deg f(\lambda)$ , has the following form:

$$\tilde{R} \equiv \begin{pmatrix} 0 & 0 & . & . & . & 0 \\ 0 & 0 & . & . & . & 0 \\ . & . & & & & . \\ . & . & & & & . \\ . & . & & & & . \\ 0 & 0 & . & . & . & 0 \\ 1 & 0 & . & . & . & 0 \end{pmatrix} \quad (5.1.50)$$

On the other hand, if  $A$  is cyclic and its minimal polynomial can be expressed as (5.1.48), then the isomorphic ILSM has the form

$$\begin{pmatrix} \tilde{x}^1(k+1) \\ \tilde{x}^2(k+1) \\ . \\ . \\ \tilde{x}^r(k+1) \end{pmatrix} = \begin{pmatrix} \tilde{A}_1 & & & \\ & \tilde{A}_2 & & \\ & & \ddots & \\ & & & \tilde{A}_r \end{pmatrix} \begin{pmatrix} \tilde{x}^1(k) \\ \tilde{x}^2(k) \\ . \\ . \\ \tilde{x}^r(k) \end{pmatrix} \quad (5.1.51)$$

where

$$\tilde{A}_i \equiv \begin{pmatrix} \tilde{Q}_i & \tilde{R}_i & & \\ & \tilde{Q}_i & \tilde{R}_i & \\ & & \ddots & \\ & & & \tilde{Q}_i & \tilde{R}_i \\ & & & & \tilde{Q}_i \end{pmatrix}, \quad i \in \underline{r} \quad (5.1.52)$$

$\tilde{Q}_i \in GF(q)^{s_i \times s_i}$  appears  $e_i$  times and is the companion matrix of  $f_i(\lambda)$  having the form (5.1.46),  $\tilde{R}_i \in GF(q)^{s_i \times s_i}$  has the form (5.1.50), and  $s_i = \deg f_i(\lambda)$ ,  $i \in \underline{r}$ .

A realization circuit for the internal submachine  $\tilde{x}^i(k+1) = \tilde{A}_i \tilde{x}^i(k)$  would consist of an assemblage of weakly coupled LFSRs. Clearly (5.1.51) consists of  $r$  uncoupled assemblages of this type.

The above discussion just about exhausts the number of canonical forms that have been used in the area of LSMs. It is evident that the feedback shift register constitutes the building block in all of these canonical structures in the sense that in each case the isomorphic ILSM is an assemblage of uncoupled or weakly coupled LFSRs. As pointed out earlier, some of the justifications for choosing LFSRs as the central and elemental components have to do with their structural simplicity, economy and efficiency in synthesis, design, and physical realization aspects. However, there exist other canonical forms such as the Jordan canonical form, Luenberger quasi-canonical form, Brunovsky canonical form, and so forth, which may be used to construct LFSRs. It is conceivable that some of these LFSRs might prove to be functionally superior in certain digital tasks to those based on the companion forms. From the related literature it appears that no comparative investigation has ever been performed in this area of ILSMs. We also observe that in the above discussion of canonical ILSMs, the input element is conspicuously missing which leads us to the conclusion that control-theoretic concepts have not been utilized in conjunction with the canonical structures.

In the sequel we will introduce some additional canonical forms, in the presence of the input element, whose structure and realization will be based on the concept of state reachability.

### 5.2. State Reachability and Canonical LSMs

In this section, we will first look at some of the companion forms from a different point of view and then discuss the Luenberger quasi-canonical forms whose structures are based on the reachability property of LSMs. The Jordan canonical form will be considered later in the sequel.

First, we consider the single-input LSM  $M_1 = (A, b)$  and assume that under the isomorphism  $P : X \rightarrow X$ , it has been transformed to the isomorphic LSM  $\tilde{M}_1 = (\tilde{A}, \tilde{b}) \equiv (P^{-1}AP, P^{-1}b)$ . If  $K_1$  and  $\tilde{K}_1$  denote the reachability matrices of  $M_1$  and  $\tilde{M}_1$ , respectively, then

$$\tilde{K}_1 \equiv [\tilde{b}, \tilde{A}\tilde{b}, \dots, \tilde{A}^{n-1}\tilde{b}] = P^{-1}[b, Ab, \dots, A^{n-1}b] = P^{-1}K_1 \quad (5.2.1)$$

If we assume that  $M_1$  is reachable, then by Corollary 4.1.1,  $K_1$  and in view of (5.2.1),  $\tilde{K}_1$  are nonsingular. Hence  $P$  can be written as

$$P = K_1 \tilde{K}_1^{-1} \quad (5.2.2)$$

Now using this particular isomorphism, we can show that  $\tilde{M}_1 = (\tilde{A}, \tilde{b})$  takes the following canonical form:

$$\begin{pmatrix} \tilde{x}_1(k+1) \\ \tilde{x}_2(k+1) \\ \vdots \\ \tilde{x}_{n-1}(k+1) \\ \tilde{x}_n(k+1) \end{pmatrix} = \begin{pmatrix} 0 & 1 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \\ a_0 & a_1 & \dots & a_{n-1} \end{pmatrix} \begin{pmatrix} \tilde{x}_1(k) \\ \tilde{x}_2(k) \\ \vdots \\ \tilde{x}_{n-1}(k) \\ \tilde{x}_n(k) \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} u(k) \quad (5.2.3)$$

where  $f_c(\lambda) = (\lambda)^n - a_{n-1}(\lambda)^{n-1} - \dots - a_1\lambda - a_0$  is the characteristic polynomial of A.

Assuming that  $\tilde{M}_1$  has the form given by (5.2.3), we can directly calculate  $\tilde{K}_1$  and obtain

$$\tilde{K}_1 \equiv [\tilde{b}, \tilde{A}\tilde{b}, \dots, \tilde{A}^{n-1}\tilde{b}] = \begin{pmatrix} 0 & 0 & 0 & \dots & 1 \\ 0 & 0 & 0 & \dots & e_1 \\ 0 & 0 & 0 & \dots & e_2 \\ \vdots & \vdots & \vdots & & \vdots \\ \vdots & \vdots & \vdots & & \vdots \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 1 & \dots & e_{n-3} \\ 0 & 1 & e_1 & \dots & e_{n-2} \\ 1 & e_1 & e_2 & \dots & e_{n-1} \end{pmatrix} \quad (5.2.4)$$

where

$$e_j \equiv \sum_{i=0}^{j-1} a_{n-i-1} e_{j-i-1}, \quad j \in \underline{n-1}; \quad e_0 \equiv 1$$

and

$$\tilde{K}_1^{-1} = \begin{pmatrix} -a_1 & -a_2 & \dots & -a_{n-1} & 1 \\ -a_2 & -a_3 & & & 1 \\ \vdots & \vdots & & & \vdots \\ \vdots & \vdots & & & \vdots \\ \vdots & \vdots & & & \vdots \\ -a_{n-1} & 1 & & & \\ 1 & & & & \end{pmatrix} \quad (5.2.5)$$

which can be verified by direct multiplication,  $\tilde{K}_1 \tilde{K}_1^{-1} = I_n$ . Therefore, the LSMs  $M_1$  and  $\tilde{M}_1$  are related to each other by the relation  $x \mapsto Px$ , where  $P = K_1 \tilde{K}_1^{-1}$ . If we let the vectors  $v_i$ ,  $i \in \underline{n}$ , denote the columns of  $P$ , we get

$$\begin{aligned}
 v^n &= b \\
 v^{n-1} &= Ab - a_{n-1}b \\
 v^{n-2} &= A^2b - a_{n-1}Ab - a_{n-2}b \\
 &\vdots \\
 v^1 &= A^{n-1}b - a_{n-1}A^{n-2}b - \dots - a_1b
 \end{aligned} \tag{5.2.6}$$

From (5.2.6) it is seen that the following recursive relation holds:

$$Av^i = v^{i-1} + a_{i-1}v^n, \quad i = 2, 3, \dots, n \tag{5.2.7}$$

Using the above analysis, we are now in a position to show that  $\tilde{M}_1$  indeed has the form given by (5.2.3). To this end, let us consider the matrix  $P^{-1}AP = \tilde{K}_1 \tilde{K}_1^{-1} A K_1 \tilde{K}_1^{-1}$ , and denote the rows of  $P^{-1}$  by  $w^{iT}$ ,  $i \in \underline{n}$ . Then for  $i \in \underline{n}$  and  $j \in \underline{n-1}$ , the  $(i, j)$  entry  $(P^{-1}AP)_{ij}$  of  $P^{-1}AP$  is given by

$$(P^{-1}AP)_{ij} = w^{iT}(Av^j) = w^{iT}(v^{j-1} + a_{j-1}v^n)$$

Therefore,

$$(P^{-1}AP)_{ij} = \begin{cases} 1 & \text{if } i = j-1 \\ a_{j-1} & \text{if } i = n \\ 0 & \text{otherwise} \end{cases}$$

This shows that the last  $n-1$  columns of  $P^{-1}AP$  have precisely the form claimed in (5.2.3). To determine the first column, we observe from (5.2.6) that

$$Av^1 = (A^n - a_{n-1}A^{n-1} - \dots - a_1A)b = a_0b = a_0v^n$$

since according to the Cayley-Hamilton Theorem,  $f_c(A) = 0$ . Thus we have for  $i \in \underline{n}$

$$\begin{aligned} (P^{-1}AP)_{i1} &= w^{iT}(Av^1) = a_0w^{iT}v^n \\ &= \begin{cases} a_0 & \text{if } i = n \\ 0 & \text{otherwise} \end{cases} \end{aligned}$$

Similarly, it can be shown that  $\tilde{b} \equiv P^{-1}b$  has the required form.

Another way of showing that the isomorphic LSM  $\tilde{M}_1$  has the canonical form given by (5.2.3) is to consider the columns  $v^i$ ,  $i \in \underline{n}$ , of  $P = K_1\tilde{K}_1^{-1}$ , given by (5.2.6), as a new basis for  $X$ , and then determine the representation of the original LSM  $M_1 = (A, b)$  with respect to this new basis. We would like to briefly discuss this approach since it can be generalized to the case of multi-input LSMs. In order to accomplish this, observe that from (5.2.6) and (5.2.7) we have

$$Av^1 = (A^n - a_{n-1}A^{n-1} - \dots - a_1A - a_0I_n)b + a_0b$$

$$= a_0b = a_0v^n = [v^1, v^2, \dots, v^n] \begin{pmatrix} 0 \\ 0 \\ \vdots \\ \vdots \\ 0 \\ a_0 \end{pmatrix}$$

$$Av^2 = v^1 + a_1v^n = [v^1, v^2, \dots, v^n] \begin{pmatrix} 1 \\ 0 \\ \vdots \\ \vdots \\ 0 \\ a_1 \end{pmatrix}$$

⋮

$$Av^n = v^{n-1} + a_{n-1}v^n = [v^1, v^2, \dots, v^n] \begin{pmatrix} 0 \\ 0 \\ \vdots \\ \vdots \\ 1 \\ a_{n-1} \end{pmatrix}$$

and



$$b = v^n = [v^1, v^2, \dots, v^n] \begin{pmatrix} 0 \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \\ 1 \end{pmatrix}$$

This shows that the LSM  $M_1 = (A, b)$  has the desired representation with respect to the new basis  $\{v^1, v^2, \dots, v^n\}$ .

However, if instead of  $P = K_1 \tilde{K}^{-1}$  we choose  $P = K_1$ , then it is clear that

$$b = [b, Ab, \dots, A^{n-1}b] \begin{pmatrix} 1 \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{pmatrix} \Rightarrow P^{-1}b = \begin{pmatrix} 1 \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{pmatrix}$$

Similarly, it is easy to see that

$$A(A^i b) = P e^{i+2}, \quad i \in \underline{n-2}$$

where  $e^j$  is a vector with 1 in the  $j$ th position and zeros everywhere else, and

$$A(A^{n-1}b) = P \begin{pmatrix} a_0 \\ a_1 \\ \cdot \\ \cdot \\ \cdot \\ a_{n-1} \end{pmatrix}$$

by the Cayley-Hamilton Theorem. Therefore, the LSM  $M_1 = (A, b)$  has the following isomorphic representation with respect to the basis  $\{b, Ab, A^2b, \dots, A^{n-1}b\}$ :

$$\begin{bmatrix} \tilde{x}_1(k+1) \\ \tilde{x}_2(k+1) \\ \tilde{x}_3(k+1) \\ \vdots \\ \tilde{x}_n(k+1) \end{bmatrix} = \begin{bmatrix} 0 & 0 & \dots & a_0 \\ 1 & 0 & \dots & a_1 \\ 0 & 1 & \dots & a_2 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_{n-1} \end{bmatrix} \begin{bmatrix} \tilde{x}_1(k) \\ \tilde{x}_2(k) \\ \tilde{x}_3(k) \\ \vdots \\ \tilde{x}_n(k) \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} u(k) \quad (5.2.8)$$

We summarize the above results in the following theorem.

Theorem 5.2.1. Suppose that the single-input LSM  $M_1 = (A, b)$  is state reachable with reachability matrix  $K_1 \equiv [b, Ab, \dots, A^{n-1}b]$  and characteristic polynomial  $f_c(\lambda) = (\lambda)^n - a_{n-1}(\lambda)^{n-1} - \dots - a_1\lambda - a_0$ . Then under the isomorphism  $P^{-1} \equiv (K_1 \tilde{K}^{-1})^{-1} : X \rightarrow X, x \mapsto P^{-1}x \equiv \tilde{x}$ , where  $\tilde{K}_1 \equiv [\tilde{b}, \tilde{A}\tilde{b}, \dots, \tilde{A}^{n-1}\tilde{b}]$ ,  $\tilde{A} \equiv P^{-1}AP$ , and  $\tilde{b} \equiv P^{-1}b$ , the isomorphic LSM  $\tilde{M}_1 = (\tilde{A}, \tilde{b})$  has the canonical form given by (5.2.3); and under the isomorphism  $P = K_1$ , the isomorphic LSM  $\tilde{M}_1 = (\tilde{A}, \tilde{b})$  has the canonical form given by (5.2.8). If the original LSM  $M_1$  is not state reachable, then no such isomorphisms exist.

In the remainder of this section we will consider the case of multi-input LSMs  $M = (A, B)$ . In deriving the canonical forms (5.2.3) and (5.2.8) for the single-input LSM  $M_1 = (A, b)$ , we chose the state isomorphism  $P$  to be  $K_1 \tilde{K}^{-1}$  and  $K_1$ , respectively. Similar state isomorphisms may be used for the purpose of identifying certain multivariable

machine canonical forms. However, since the reachability matrix  $K \equiv [B, AB, \dots, A^{n-1}B]$  of the multivariable LSM  $M = (A, B)$  is not square, the relation  $\tilde{K} = P^{-1}K$ , the equivalent of  $\tilde{K}_1 = P^{-1}K_1$  for  $M_1 = (A, b)$  given by (5.2.1), can no longer be used. But since  $\tilde{K}K^T = P^{-1}KK^T$ ,  $P = KK^T(\tilde{K}K^T)^{-1}$  can be used as a valid state isomorphism. However, we will not pursue this particular approach but instead utilize the rank condition of the matrix  $K$  under the assumption of state reachability of  $M = (A, B)$  and choose different bases from the columns of  $K$  which will result into different quasi-canonical forms for  $M$ . These will be called the Luenberger quasi-canonical forms for LSMs. This approach is somewhat similar to the one used for the case of  $M_1 = (A, b)$  by choosing  $P = K_1$ .

Let  $b^i$ ,  $i \in \underline{m}$ , denote the columns of the matrix  $B$  and rewrite  $K$  as

$$K = [b^1, b^2, \dots, b^m, Ab^1, Ab^2, \dots, Ab^m, \dots, A^{n-1}b^1, A^{n-1}b^2, \dots, A^{n-1}b^m] \quad (5.2.9)$$

Furthermore, assume that the LSM  $M = (A, B)$  is state reachable. Therefore, there are  $n$  linearly independent columns in  $K$ . Since  $K$  has a total of  $mn$  columns,  $n$  linearly independent columns can be selected in many different ways, giving rise to different quasi-canonical forms. Two such quasi-canonical forms were discussed in detail in Section 5.1, and are given by (5.1.22) – (5.1.23) and (5.1.27). Here, we will employ a different scheme for choosing the  $n$  linearly independent columns of  $K$

and consequently derive another quasi-canonical form for LSMs. In contrast to the discussion of canonical forms in Section 5.1, here no assumption is made about the rank of the matrix  $B$ .

For the purpose of selecting a basis for  $X$  from the columns of  $K$ , we start with the vector  $b^1$  (the first column of  $B$ ) and proceed to  $Ab^1, A^2b^1, \dots$ , until either  $A^{n_1-1}b^1$  is chosen in which case the machine is reachable by the first input alone, or until a dependency arises, that is, until  $A^{n_1}b^1$  can be expressed as a linear combination of  $b^1, Ab^1, \dots, A^{n_1-1}b^1$ . If more independent vectors are required, we select  $b^2$  (the second column of  $B$ ),  $Ab^2, A^2b^2, \dots$ , until a dependency arises, that is, until  $A^{n_2}b^2$  can be expressed as a linear combination of  $b^2, Ab^2, \dots, A^{n_2-1}b^2$ . If  $n_1 + n_2 < n$ , we proceed to  $b^3, Ab^3, \dots, A^{n_3-1}b^3$ , and so forth, until  $n$  linearly independent vectors are obtained. Assume that this procedure yields the following set of vectors:

$$\{b^1, Ab^1, \dots, A^{n_1-1}b^1; b^2, Ab^2, \dots, A^{n_2-1}b^2; \dots; b^s, Ab^s, \dots, A^{n_s-1}b^s\} \quad (5.2.10)$$

Now if the state isomorphism  $P : X \rightarrow X, x \mapsto P^{-1}x \equiv \tilde{x}$ , is chosen to consist of the vectors (5.2.10), then it can be easily verified that the isomorphic LSM  $\tilde{M} = (\tilde{A}, \tilde{B}) \equiv (P^{-1}AP, P^{-1}B)$  has the form given by (5.2.11) in which the  $*$  entries represent possibly nonzero elements.



$$\tilde{B} \equiv \left( \begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & 0 & \\ 0 & 0 & 0 & 0 & 0 & \\ 0 & 0 & 0 & 0 & 0 & \\ \vdots & \vdots & \vdots & \vdots & \vdots & \\ \vdots & \vdots & \vdots & \vdots & \vdots & \\ \vdots & \vdots & \vdots & \vdots & \vdots & \\ \hline 0 & 0 & 0 & 0 & 0 & \\ \hline 0 & 1 & 0 & 0 & 0 & \\ 0 & 0 & 0 & 0 & 0 & \\ 0 & 0 & 0 & 0 & 0 & \\ \vdots & \vdots & \vdots & \vdots & \vdots & \\ \vdots & \vdots & \vdots & \vdots & \vdots & \\ \vdots & \vdots & \vdots & \vdots & \vdots & \\ \hline 0 & 0 & 0 & 0 & 0 & \\ \hline & \vdots & & \vdots & & \\ & \vdots & & \vdots & & \\ & \vdots & & \vdots & & \\ \hline 0 & 0 & 0 & 0 & 1 & \\ 0 & 0 & 0 & 0 & 0 & \\ \vdots & \vdots & \vdots & \vdots & \vdots & \\ \vdots & \vdots & \vdots & \vdots & \vdots & \\ \vdots & \vdots & \vdots & \vdots & \vdots & \\ \hline 0 & 0 & 0 & 0 & 0 & \\ 0 & 0 & 0 & & & \end{array} \right) \begin{array}{l} \\ \\ \\ \\ \\ \\ \\ \\ \\ \tilde{b}^{s+1} \quad \tilde{b}^{s+2} \quad \dots \quad \tilde{b}^m \\ \\ \\ \\ \\ \\ \\ \\ \\ \end{array} \quad (5.2.11b)$$

### 5.3. State Reachability and Feedback

Linear state variable feedback is of fundamental importance in many aspects of the synthesis or design of compensation schemes for LSMs. The important role of feedback in certain areas of linear machine design and behavior will be discussed in more detail in the

sequel. However, in this section we will restrict our attention to a brief presentation of some interrelationships among state reachability, state feedback, and "pole shifting."

Theorem 5.3.1. Consider the LSM  $M = (A, B)$ . Introducing the memoryless state feedback law  $u(k) = Fx(k) + Gw(k)$ , where  $F : X \rightarrow U$  is a state feedback map,  $G : U \rightarrow U$  is an isomorphism, and  $w(k) \in GF(q)^m$  is a new external input, the LSM  $M$  is transformed to the LSM  $\bar{M} \equiv (A + BF, BG)$ . Then  $M$  is state reachable if and only if  $\bar{M}$  is state reachable. That is, the property of state reachability for LSMs is invariant under state feedback transformation.

Proof. Let  $K$  and  $\bar{K}$  denote the reachability matrices of  $M$  and  $\bar{M}$ , respectively. That is,  $K \equiv [B, AB, \dots, A^{n-1}B]$  and  $\bar{K} \equiv [BG, (A + BF)BG, \dots, (A + BF)^{n-1}BG]$ . Suppose that  $M$  is state reachable but  $\bar{M}$  is not. Then  $\text{rank } K = n$  and  $\text{rank } \bar{K} < n$ . Therefore, there exists a nonzero vector  $v \in GF(q)^n$  such that  $v^T \bar{K} = 0$  which implies that

$$v^T (A + BF)^i BG = 0, \quad i \in \underline{n-1} \quad (5.3.1)$$

Since  $G$  is nonsingular, (5.3.1) reduces to

$$v^T (A + BF)^i B = 0, \quad i \in \underline{n-1} \quad (5.3.2)$$

From (5.3.2) it follows that  $v^T A^i B = 0, i \in \underline{n-1}$ , that is,  $v^T [B, AB, \dots, A^{n-1}B] = 0$ , which is a contradiction since the rows of  $K$  are linearly independent. Conversely, assume that  $M$  is not state reachable but  $\bar{M}$  is. Then there exists a nonzero vector  $v \in GF(q)^n$  such that  $v^T \bar{K} = 0$ . This implies that  $v^T A^i B = 0, i \in \underline{n-1}$ , which is equivalent

to  $v^T A^i B G = 0$ ,  $i \in \underline{n-1}$ . From this last set of relations it follows that  $v^T (A + BF)^i B G = 0$ ,  $i \in \underline{n-1}$ , that is,  $v^T K = 0$ , which is a contradiction.  $\square$

If we use the equivalent expression

$$\{A \mid R(B)\} \equiv R(B) + AR(B) + A^2 R(B) + \dots + A^{n-1} R(B)$$

for  $R(K)$ , then an alternative and more elegant proof for the above theorem can be given as follows. Letting  $\hat{A} \equiv A + BF$  and noting that for any  $W \subseteq X$ ,  $(A + BF)W \subseteq AW + R(B)$ , and  $R(BG) = R(B)$ , we have

$$\begin{aligned} \{A + BF \mid R(B)\} &= \{\hat{A} \mid R(B)\} = R(B) + \hat{A} R(B) + \dots + \hat{A}^{n-1} R(B) \\ &= R(B) + \hat{A} R(B) + \hat{A} R(B) + \hat{A} (\dots (\hat{A} R(B) + \hat{A} R(B)) \dots) \\ &\subseteq R(B) + AR(B) + A^2 R(B) + \dots + A^{n-1} R(B) \\ &\equiv \{A \mid R(B)\} \end{aligned}$$

Since the above inclusion holds for all  $A$ ,  $B$ , and  $F$ , if we replace  $F$  by  $-F$  and then  $A$  by  $A + BF$ , we obtain the reverse inclusion  $\{A \mid R(B)\} \subseteq \{A + BF \mid R(B)\}$ . Hence  $\{A + BF \mid R(B)\} = \{A \mid R(B)\}$ .  $\square$

**Theorem 5.3.2.** If the single-input LSM  $M_1 = (A, b)$  is state reachable, then there exists a vector  $v \in GF(q)^n$  such that the characteristic polynomial of the LSM  $(A + bv^T, b)$  has an arbitrary preassigned form.

**Proof.** Since  $M_1$  is state reachable, by Theorem 5.2.1, there exists an isomorphism  $P : X \rightarrow X$  such that the isomorphic LSM  $\tilde{M}_1 = (\tilde{A}, \tilde{b}) \equiv (P^{-1}AP, P^{-1}b)$  has the form



$$\left( \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ a_0 & a_1 & a_2 & \dots & a_{n-1} \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ \vdots \\ \vdots \\ 0 \\ 1 \end{bmatrix} \right)$$

where  $f_c(\lambda) = (\lambda)^n - a_{n-1}(\lambda)^{n-1} - a_{n-2}(\lambda)^{n-2} - \dots - a_1\lambda - a_0$  is the characteristic polynomial of  $A$ . If  $\tilde{v}^T \equiv (\tilde{v}_1, \tilde{v}_2, \dots, \tilde{v}_n)$  is an arbitrary vector, then

$$\tilde{A} + b\tilde{v}^T = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ \tilde{d}_0 & \tilde{d}_1 & \tilde{d}_2 & \dots & \tilde{d}_{n-1} \end{bmatrix}$$

with characteristic polynomial  $\tilde{f}_c(\lambda) = (\lambda)^n - \tilde{d}_{n-1}(\lambda)^{n-1} - \dots - \tilde{d}_1\lambda - \tilde{d}_0$ , where  $\tilde{d}_i \equiv a_i + \tilde{v}_{i+1}$ ,  $i \in \underline{n-1}$ . Now it is obvious that we can choose  $a_i + \tilde{v}_{i+1}$ ,  $i \in \underline{n-1}$ , to match the coefficients of the preassigned characteristic polynomial.  $\square$

Theorem 5.3.3. If  $M = (A, B)$  is a state reachable LSM, then there exists a feedback homomorphism  $F : X \longrightarrow U$  and a vector  $b \in GF(q)^n$  such that the LSM  $\bar{M} = (A + BF, b)$  is reachable and  $b \in R(B)$ .

Proof. Since  $M$  is reachable, the reachability matrix  $K \equiv [B, AB, \dots, A^{n-1}B]$  has rank  $n$ . Consequently, there are  $n$  linearly independent column vectors in  $K$ . We choose these linearly independent column vectors according to the scheme that led to (5.2.10), and let

$$T \equiv [b^1, Ab^1, \dots, A^{n_1-1}b^1; b^2, Ab^2, \dots, A^{n_2-1}b^2; \dots; b^t, Ab^t, \dots, A^{n_t-1}b^t]$$

where  $n_i, i \in \underline{t}$ , are the reachability indices and thus satisfy the relation  $n_1 + n_2 + \dots + n_t = n$ . Furthermore, we define

$$S \equiv [s^1, s^2, \dots, s^n] \in GF(q)^{m \times n}$$

as follows:

$$S = [0, 0, \dots, 0, \overset{\uparrow}{e^2}, 0, \dots, 0, \overset{\uparrow}{e^3}, 0, \dots, 0, \overset{\uparrow}{e^t}, 0, \dots, 0, 0]$$

$n_1$ th column       $(n_1+n_2)$ th column       $(n_1+n_2+\dots+n_{t-1})$ th column       $n$ th column

or more compactly,

$$s^r_j = e^{j+1} \quad \text{if} \quad r_j = \sum_{i=1}^j n_i, \quad j \in \underline{t-1}$$

$$s^j = 0 \quad \text{otherwise}$$

where  $e^i$  is the  $i$ th standard basis vector of  $GF(q)^m$ . Now we will show that choosing the feedback matrix  $F$  as

$$F \equiv ST^{-1} \tag{5.3.3}$$

satisfies the theorem. Rewriting (5.3.3) as  $FT = S$ , it is clear that

$$FA^{j-1}b^j = e^{j+1}, j \in \underline{t-1}$$

$$FA^i b^j = 0, \text{ for all other powers of } A$$

Using these relationships, we can determine the columns of the reachability matrix  $\bar{K}$  of the LSM  $\bar{M} = (A + BF, b)$  as follows:

$$b^1 = b^1$$

$$(A + BF)b^1 = Ab^1$$

$$(A + BF)^2 b^1 = (A + BF)Ab^1 = A^2 b^1$$

$$\vdots$$

$$(A + BF)^{n-1} b^1 = (A + BF)^{n-2} b^1 = A^{n-1} b^1$$

$$(A + BF)^n b^1 = (A + BF)^{n-1} b^1 = A^n b^1 + Be^2 = b^2 + ***$$

$$(A + BF)^{n+1} b^1 = (A + BF)(b^2 + A^n b^1) = Ab^2 + ***$$

$$\vdots$$

$$(A + BF)^{n-1} b^1 = (A + BF)(A^{t-2} b^t + ***) = A^{t-1} b^t + ***$$

where \*\*\* denotes the linear combination of the preceding vectors.

From the above expressions it is clear that the columns of  $\bar{K}$ , that is, the vectors

$$b^1, (A + BF)b^1, (A + BF)^2 b^1, \dots, (A + BF)^{n-1} b^1$$

are linearly independent. Thus  $\text{rank } \bar{K} = n$  and hence  $\bar{M} = (A + BF, b)$  is state reachable. Clearly  $b^1 \in R(B)$ .  $\square$

With the aid of the above theorem, we can extend Theorem 5.3.2 to the multivariable case as shown in the following theorem.

Theorem 5.3.4. If the LSM  $M = (A, B)$  is state reachable, then there exists a feedback homomorphism  $F : X \longrightarrow U$  such that the characteristic polynomial of  $A + BF$  has an arbitrary preassigned form.

Proof. Introducing the feedback law  $u(k) = v(k) + F'x(k)$ ,  $M$  becomes  $\bar{M} \equiv (A + BF', B)$ . Since  $M$  is state reachable, by Theorem 5.3.3 there exists an  $F'$  such that the LSM  $(A + BF', b^1)$ , where  $b^1$  is the first column of  $B$ , is state reachable. Introducing another state feedback law  $v(k) = w(k) + F''x(k)$  with  $F''$  having the form

$$F'' \equiv \begin{pmatrix} f''_1 & f''_2 & \cdot & \cdot & \cdot & f''_n \\ 0 & 0 & \cdot & \cdot & \cdot & 0 \\ \cdot & \cdot & & & & \cdot \\ \cdot & \cdot & & & & \cdot \\ \cdot & \cdot & & & & \cdot \\ 0 & 0 & \cdot & \cdot & \cdot & 0 \end{pmatrix}$$

$\bar{M}$  becomes

$$\begin{aligned} x(k+1) &= (\bar{A} + BF'')x(k) + Bw(k) \\ &= (\bar{A} + b^1 f'')x(k) + Bw(k) \end{aligned}$$

where  $\bar{A} \equiv A + BF'$  and  $f'' = [f''_1, f''_2, \cdot, \cdot, \cdot, f''_n]$ . Since the LSM  $(\bar{A}, b^1)$  is state reachable, by Theorem 5.3.2, the characteristic polynomial of  $\bar{A} + b^1 f''$  has an arbitrary preassigned form. Now if we combine the feedback forms  $u$  and  $v$ , introduced above, as  $u = v + (F' + F'')x \equiv v + Fx$ , then the theorem is proved.  $\square$

The necessary condition of Theorem 5.3.4 also turns out to be sufficient. Proofs of sufficiency will be given in Theorem 6.3.1 and Theorem 7.1.1.

The following lemma plays a central role in our further discussion of the interrelationships between state reachability and feedback.

Lemma 5.3.1. Let  $A \in GF(q)^{n \times n}$ ,  $B \in GF(q)^{n \times m}$ , and  $F \in GF(q)^{m \times n}$ .

Then the following matrix identity holds:

$$\begin{aligned}
 (A + BF)^j &= A^j + BF(A + BF)^{j-1} + ABF(A + BF)^{j-2} \\
 &\quad + \dots + A^{j-2}BF(A + BF) + A^{j-1}BF, \quad j = 1, 2, \dots \\
 &= A^j + [B, AB, \dots, A^{j-1}] \begin{pmatrix} F(A + BF)^{j-1} \\ F(A + BF)^{j-2} \\ \vdots \\ F(A + BF) \\ F \end{pmatrix}
 \end{aligned}
 \tag{5.3.4}$$

Proof. For  $j = 1$  and  $j = 2$ , (5.3.4) holds since

$$\begin{aligned}
 (A + BF)^1 &= A^1 + BF(A + BF)^0 = A + BF \\
 (A + BF)^2 &= A^2 + BF(A + BF) + ABF(A + BF)^0 \\
 &= A^2 + BF(A + BF) + ABF
 \end{aligned}$$

Now suppose that the identity holds for  $j-1$ , that is,

$$\begin{aligned}
 (A + BF)^{j-1} &= A^{j-1} + BF(A + BF)^{j-2} + ABF(A + BF)^{j-3} \\
 &\quad + \dots + A^{j-3}BF(A + BF) + A^{j-2}BF
 \end{aligned}
 \tag{5.3.5}$$

We will show that it also holds for  $j$ . Postmultiplying (5.3.5) by  $A + BF$ , we obtain

$$\begin{aligned}
 (A + BF)^j &= A^{j-1}(A + BF) + BF(A + BF)^{j-1} + ABF(A + BF)^{j-2} \\
 &\quad + \dots + A^{j-3}BF(A + BF)^2 + A^{j-2}BF(A + BF) \\
 &= A^j + BF(A + BF)^{j-1} + ABF(A + BF)^{j-2} \\
 &\quad + \dots + A^{j-2}BF(A + BF) + A^{j-1}BF
 \end{aligned}$$

Therefore, the result is proved by induction.  $\square$

Theorem 5.3.5. Let  $M = (A, B)$  be a nonsingular LSM. Then  $M$  is  $j$ -state reachable if and only if the matrix equation

$$(A + BF)^j = 0 \quad (5.3.6)$$

has a unique solution with respect to the feedback matrix  $F \in GF(q)^{m \times n}$ .

Proof. Introducing the feedback law

$$u(k) = Fx(k) \quad (5.3.7)$$

$M$  becomes

$$x(k+1) = (A + BF)x(k) \quad (5.3.8)$$

Starting from an arbitrary initial state  $x(0)$  and applying equation (5.3.8)  $j$  times, we obtain the relationship

$$x(j) = (A + BF)^j x(0) \quad (5.3.9)$$

Now postmultiplying both sides of the matrix identity (5.3.4) by  $x(0)$ , we get

$$\begin{aligned} (A + BF)^j x(0) &= A^j x(0) + BF(A + BF)^{j-1} x(0) + ABF(A + BF)^{j-2} x(0) \\ &+ \dots + A^{j-2} BF(A + BF)x(0) + A^{j-1} BFx(0) \end{aligned} \quad (5.3.10)$$

In view of the equations (5.3.6) and (5.3.9), equation (5.3.10) reduces to the following expression:

$$A^j x(0) + BFx(j-1) + ABFx(j-2) + \dots + A^{j-2} BFx(1) + A^{j-1} BFx(0) = 0 \quad (5.3.11)$$

Premultiplying equation (5.3.11) by  $FA^{-j}$  yields

$$\begin{aligned} Fx(0) + FA^{-j} BFx(j-1) + FA^{-j+1} BFx(j-2) \\ + \dots + FA^{-2} BFx(1) + FA^{-1} BFx(0) = 0 \end{aligned} \quad (5.3.12)$$

Substituting  $u(k) = Fx(k)$ ,  $k \in \underline{j-1}$ , equation (5.3.12) reduces to the following equation:

$$-u(0) = FA^{-j} Bu(j-1) + FA^{-j+1} Bu(j-2) + \dots + FA^{-2} Bu(1) + FA^{-1} Bu(0) \quad (5.3.13)$$

Since equation (5.3.13) is an identity for any  $u(k)$ ,  $k \in \underline{j-1}$ , it follows that

$$\begin{aligned} FA^{-1} B &= -I_m \\ FA^{-i} B &= 0, \quad i = 2, 3, \dots, j \end{aligned} \quad (5.3.14)$$

It is clear that equation (5.3.6) is equivalent to equations (5.3.14). The set of equations (5.3.14) will have a unique solution with respect to  $F$  if and only if it contains exactly  $mn$  linearly independent equations or equivalently, if and only if the matrix

$$[A^{-j}B, A^{-j+1}B, \dots, A^{-1}B] = A^{-j}[B, AB, \dots, A^{j-1}B] \equiv A^{-j}K$$

has exactly  $n$  linearly independent columns. But since  $A$  is nonsingular, the matrix  $K$ , which is the state reachability matrix of  $M$ , will have  $n$  linearly independent columns if and only if  $\text{rank } K = n$ , that is, if and only if  $M$  is  $j$ -state reachable.  $\square$

Corollary 5.3.1. A nonsingular LSM  $M = (A, B)$  is state reachable if and only if the set of equations (5.3.14) has a unique solution with respect to the matrix  $F$ .

The above result also contains a solution of the minimum-time feedback control problem of linear machines which can be stated as follows: Given a nonsingular LSM  $M = (A, B)$  with state reachability index  $\ell$ , determine the matrix  $F \in GF(q)^{m \times n}$  in the linear feedback control law (5.3.7) such that  $M$  is driven from any arbitrary initial state to the zero state  $0_X$  in a minimum number of clock periods. To see that a solution to this optimal control problem is provided by Theorem 5.3.5, notice that in order to have  $x(\ell) = 0$  for any arbitrary initial state  $x(0)$ , from equation (5.3.9) it follows that we must have  $(A + BF)^\ell = 0$  which is equation (5.3.6). In the proof of Theorem 5.3.5, it was shown that the existence of a unique solution of  $(A + BF)^\ell = 0$  with



respect to the feedback matrix  $F$  is equivalent to the condition of  $\ell$ -state reachability of  $M$ . We summarize this observation in the following theorem.

Theorem 5.3.6. The matrix  $F \in GF(q)^{m \times n}$  in the linear feedback control law  $u(k) = Fx(k)$  for which the LSM  $(A, B)$ , started at any arbitrary initial state, is driven to the zero state in a minimum number of clock periods, is given by the solution of equation (5.3.14).

Corollary 5.3.2. The state controllability index of the LSM  $(A, B)$  is equal to the smallest integer  $\ell$  for which the equation  $(A + BF)^\ell = 0$  has a unique solution with respect to  $F \in GF(q)^{m \times n}$ .

Corollary 5.3.3. The state controllability index of the LSM  $(A, B)$  is equal to the smallest integer  $\ell$  for which the set of equations (5.3.14) has a unique solution with respect to  $F \in GF(q)^{m \times n}$ .

### Summary and Conclusions

The primary focus of this chapter was on the concepts of isomorphic LSMs, canonical representations, and state feedback in conjunction with the property of state reachability of LSMs.

After formalizing the notions of isomorphic LSMs, canonical forms, and invariants of equivalence relations, first a special set of invariants, called reachability indices, associated with an LSM  $M = (A, B)$ , was thoroughly characterized and then its use in the invariant description of state reachable LSMs was illustrated. In the course of this illustration, a new derivation of Brunovsky's canonical

form [15] for LSMs was presented. The results pertaining to the characterization of the reachability indices are essentially specializations to the case of LSMs of the results due to Popov [90].

It was observed that the only canonical forms used only in the area of autonomous LSMs are the companion and hypercompanion forms which are popular in the area of classical linear algebra. This observation led to the conclusion that canonical forms have never been considered for nonautonomous LSMs and hence control-theoretic concepts have not been utilized for their canonical representation. In this chapter canonical forms for nonautonomous LSMs were investigated. Making use of the property of state reachability, various canonical and quasi-canonical forms, including the companion forms, for both single-input and multi-input LSMs were presented (cf. [17], [21], [58], [61], [71], [88], [107], [110], [114]).

Finally, the effect of state feedback on reachability, eigenvalue assignability, and time-optimal control of LSMs was closely examined (cf. [18], [21], [58], [61], [79], [88], [107], [109], [110], [114]).

## CHAPTER VI

## STATE REACHABILITY REVISITED

In this chapter we will introduce a large number of equivalent state reachability criteria for both single-input and multi-input LSMs. Single-input LSMs will be treated separately rather than special cases of multivariable LSMs because they do constitute an important class of LSMs in their own right, and also due to the fact that there are some reachability criteria for single-input LSMs that cannot be extended to the multivariable case.

In Section 5.1, we saw that the property of state reachability could lead to numerous canonical and quasi-canonical forms for both classes of single-input and multi-input LSMs. Clearly the existence of these isomorphic forms can also be stated as necessary and sufficient conditions for state reachability. However, the number of these forms is prohibitively large. Therefore, by way of illustration, we will include in this presentation only two reachability criteria in terms of the existence of canonical forms for single-input LSMs and only one such criterion for multi-input LSMs. Furthermore, we will re-examine, in this chapter, some groups of transformations in relation to the class of state reachable LSMs. These groups were originally introduced in Section 5.1.

6.1. Twenty Four Equivalent Criteria for the  
State Reachability Property of  
Single-Input LSMs

Theorem 6.1.1. For the single-input LSM  $M_1 = (A, b)$ , the following statements are equivalent:

1<sup>o</sup>. The LSM  $M_1 = (A, b)$  is state reachable.

2<sup>o</sup>. There does not exist any isomorphism  $P : X \rightarrow X$ ,  $P \in GF(n, q)$ , such that the isomorphic LSM  $\tilde{M}_1 = (\tilde{A}, \tilde{b}) \equiv (PAP^{-1}, Pb)$  will have the form

$$\left( \begin{bmatrix} \tilde{A}_{11} & \tilde{A}_{12} \\ 0 & \tilde{A}_{22} \end{bmatrix}, \begin{bmatrix} \tilde{b}^1 \\ 0 \end{bmatrix} \right) \quad (6.1.1)$$

where  $\tilde{A}_{11} \in GF(q)^{r \times r}$ ,  $\tilde{A}_{12} \in GF(q)^{r \times (n-r)}$ ,  $\tilde{A}_{22} \in GF(q)^{(n-r) \times (n-r)}$ , and  $\tilde{b}^1 \in GF(q)^r$ , with  $r < n$ .

Or

There does not exist any LSM  $\tilde{M}_1 = (PAP^{-1}, Pb)$ , isomorphic to  $M_1 = (A, b)$ , which will have the form:

$$\tilde{x}^I(k+1) = \tilde{A}_{11}\tilde{x}^I(k) + \tilde{A}_{12}\tilde{x}^{II}(k) + \tilde{b}^1 u(k) \quad (6.1.2a)$$

$$\tilde{x}^{II}(k+1) = \tilde{A}_{22}\tilde{x}^{II}(k) \quad (6.1.2b)$$

where  $\tilde{x}^I(k) \in GF(q)^r$ ,  $\tilde{x}^{II}(k) \in GF(q)^{n-r}$ ,  $\tilde{A}_{11} \in GF(q)^{r \times r}$ ,  $\tilde{A}_{12} \in GF(q)^{r \times (n-r)}$ ,  $\tilde{A}_{22} \in GF(q)^{(n-r) \times (n-r)}$ , and  $\tilde{b}^1 \in GF(q)^r$ , with  $r < n$ .

3°. The vector  $b$  does not belong to any  $A$ -invariant subspace of dimension smaller than  $n$ .

4°. There exists no nonzero eigenvector  $v$  of the matrix  $A^T$  orthogonal to the vector  $b$ . That is, there exists no vector  $v$  which will simultaneously satisfy the following conditions:

$$v^T(\lambda I_n - A) = 0, v^T b = 0, v \neq 0 \quad (6.1.3)$$

5°. A subspace of  $X$  which is orthogonal to the vector  $b$  does not contain an  $A^T$ -invariant subspace.

6°. The following matrix  $K_1 \in GF(q)^{n \times n}$  has rank  $n$ :

$$K_1 \equiv [b, Ab, A^2b, \dots, A^{n-1}b] \quad (\det K_1 \neq 0) \quad (6.1.4)$$

Or

The linear map  $[b, Ab, A^2b, \dots, A^{n-1}b] : U^* \rightarrow X$  is an epimorphism.

Or

$GF(q)^n$  is cyclic with respect to  $A$ , having generator  $b$ .

Or

There exists no polynomial  $f(\xi) \in GF(q)[\xi]$  of degree less than  $n$  such that  $f(A)b = 0$ .

7°. The rank of the matrix  $K_1 K_1^T = \sum_{j=0}^{n-1} A^{n-j-1} b b^T (A^T)^{n-j-1}$ , where  $K_1$  is defined by (6.1.4), is equal to  $n$ .

8°. The matrix equation

$$(A + bf^T)^n = 0 \quad (6.1.5)$$

has a unique solution with respect to the vector  $f \in GF(q)^n$ , and  $A$  is nonsingular.

9°. The set of linear equations

$$\begin{aligned} f^T A^{-1} b &= -1 \\ f^T A^{-1} b &= 0, \quad i = 2, 3, \dots, n \end{aligned} \quad (6.1.6)$$

has a unique solution with respect to the elements of the vector  $f$ .

10°. The following matrix  $E_1 \in GF(q)^{n^2 \times n^2}$  has rank  $n^2$ :

$$E_1 \equiv \begin{pmatrix} I_n & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & b \\ -A & I_n & 0 & \dots & 0 & 0 & 0 & \dots & 0 & b & 0 \\ 0 & -A & I_n & \dots & 0 & 0 & 0 & \dots & b & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & I_n & 0 & b & \dots & 0 & 0 & 0 \\ 0 & 0 & 0 & \dots & -A & b & 0 & \dots & 0 & 0 & 0 \end{pmatrix} \quad (6.1.7)$$

11°. Given any polynomial vector  $z(\xi) \in GF(q)[\xi]^{n \times 1}$  with elements of degree  $n-1$  or less, there exist a polynomial vector  $x(\xi) \in GF(q)[\xi]^{n \times 1}$  with elements of degree  $n-2$  or less, and a polynomial  $y(\xi) \in GF(q)[\xi]$  of degree  $n-1$  or less, such that

$$(\xi I_n - A)x(\xi) + by(\xi) = z(\xi) \quad (6.1.8)$$

12<sup>o</sup>. There exist a polynomial matrix  $X(\xi) \in GF(q)[\xi]^{n \times n}$  with elements of degree  $n-2$  or less, and a polynomial vector  $y(\xi) \in GF(q)[\xi]^{n \times 1}$  with elements of degree  $n-1$  or less, such that

$$(\xi I_n - A)X(\xi) + by^T(\xi) = I_n \quad (6.1.9)$$

13<sup>o</sup>. The matrix  $[\xi I_n - A, b] \in GF(q)[\xi]^{n \times (n+1)}$  has rank  $n$  for all  $\xi$ .

14<sup>o</sup>. The polynomial matrices  $\xi I_n - A, b$  are coprime, that is,  $[\xi I_n - A, b]$  has the Smith canonical form  $[I_n, 0]$ .

15<sup>o</sup>. The matrix equations

$$PA - AP = 0 \quad (6.1.10)$$

$$Pb = d$$

admit a unique matrix solution  $P \in GF(q)^{n \times n}$  for every vector  $d \in GF(q)^n$ .

16<sup>o</sup>. For any other LSM  $\tilde{M}_1 = (\tilde{A}, \tilde{b})$  which satisfies the conditions

$$\det \tilde{K}_1 = \det[\tilde{b}, \tilde{A}\tilde{b}, \tilde{A}^2\tilde{b}, \dots, \tilde{A}^{n-1}\tilde{b}] \neq 0 \quad (6.1.11)$$

and

$$\det(\lambda I_n - \tilde{A}) = \det(\lambda I_n - A) \quad (6.1.12)$$

there exists a nonsingular matrix  $P \in GF(q)^{n \times n}$  such that

$$\tilde{A} = PAP^{-1}; \tilde{b} = Pb \quad (6.1.13)$$

17<sup>o</sup>. There exists an isomorphism  $P : X \longrightarrow X$ ,  $P \in GF(n, q)$ , such that the isomorphic LSM  $(\tilde{A}, \tilde{b}) \equiv (PAP^{-1}, Pb)$  has the following form:

$$\tilde{A} \equiv \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ a_1 & a_2 & a_3 & \dots & a_n \end{pmatrix}; \tilde{b} \equiv \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} \quad (6.1.14)$$

where  $a_i$ ,  $i \in \underline{n}$ , are the coefficients of the characteristic polynomial  $f_c(\lambda) = (\lambda)^n - \sum_{i=1}^n a_i(\lambda)^{i-1}$  of the matrix  $A$ .

Or

There exists an LSM  $\tilde{M}_1 = (\tilde{A}, \tilde{b}) \equiv (PAP^{-1}, Pb)$  isomorphic to  $M_1 = (A, b)$ , whose state equations have the form

$$\tilde{x}_1(k+1) = \tilde{x}_2(k)$$

$$\tilde{x}_2(k+1) = \tilde{x}_3(k)$$

$$\vdots$$

$$\tilde{x}_{n-1}(k+1) = \tilde{x}_n(k)$$

$$\tilde{x}_n(k+1) = a_1 \tilde{x}_1(k) + a_2 \tilde{x}_2(k) + \dots + a_n \tilde{x}_n(k) + \tilde{b}u(k)$$

18<sup>o</sup>. There exists an isomorphism  $P : X \longrightarrow X$ ,  $P \in GF(n, q)$ , such that the isomorphic LSM  $\tilde{M}_1 = (\tilde{A}, \tilde{b}) \equiv (PAP^{-1}, Pb)$  has the Jordan-Lur'e-Lefschetz form



$$\tilde{A} \equiv \begin{pmatrix} \tilde{A}_1 & & & \\ & \tilde{A}_2 & & \\ & & \ddots & \\ & & & \tilde{A}_v \end{pmatrix}; \tilde{b} \equiv \begin{pmatrix} \tilde{b}^1 \\ \tilde{b}^2 \\ \vdots \\ \tilde{b}^v \end{pmatrix} \quad (6.1.15)$$

where  $\tilde{A}_i \in GF(q)^{n_i \times n_i}$ ,  $i \in \underline{v}$ , are Jordan blocks of the form

$$\tilde{A}_i \equiv \begin{pmatrix} \lambda_i & 1 & & & \\ & \lambda_i & 1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ & & & & \lambda_i \end{pmatrix}, i \in \underline{v} \quad (6.1.16)$$

$\tilde{b}^i \in GF(q)^{n_i}$ ,  $i \in \underline{v}$ , have the form

$$\tilde{b}^i \equiv \begin{pmatrix} 0 \\ 0 \\ \vdots \\ \vdots \\ 0 \\ 1 \end{pmatrix}, i \in \underline{v} \quad (6.1.17)$$

and the numbers  $\lambda_i$  and  $n_i$  are obtained from the expression of the characteristic polynomial

$$\det(\lambda I_n - A) = \prod_{i=1}^v (\lambda - \lambda_i)^{n_i}, \quad (\lambda_r = \lambda_s \text{ whenever } r \neq s) \quad (6.1.18)$$

of the matrix A.

Or

There exists an isomorphism  $P : X \rightarrow X$ ,  $P \in \text{GF}(n, q)$ , such that the isomorphic LSM  $\tilde{M}_1 = (\tilde{A}, \tilde{b}) \equiv (PAP^{-1}, Pb)$  has the form

$$\begin{aligned} \tilde{x}_{i1}(k+1) &= \lambda_i \tilde{x}_{i1}(k) + \tilde{x}_{i2}(k) \\ \tilde{x}_{i2}(k+1) &= \lambda_i \tilde{x}_{i2}(k) + \tilde{x}_{i3}(k) \\ &\vdots \\ \tilde{x}_{i, n_i-1}(k+1) &= \lambda_i \tilde{x}_{i, n_i-1}(k) + \tilde{x}_{in_i}(k) \\ \tilde{x}_{in_i}(k+1) &= \lambda_i \tilde{x}_{in_i}(k) + u(k); \quad i \in \underline{v} \end{aligned} \quad (6.1.19)$$

19<sup>0</sup>. For every polynomial of the form

$$f_1(\lambda) = (\lambda)^n - \sum_{i=1}^n b_i(\lambda)^{i-1} \quad (6.1.20)$$

there exists a vector  $v^0 \in \text{GF}(q)^n$  such that

$$\det[\lambda I_n - (A + bv^{0T})] = f_1(\lambda) \quad (6.1.21)$$

That is, the characteristic polynomial of the matrix  $A + bv^{0T}$  is equal to the given polynomial  $f_1(\lambda)$ .

20°. There exists a nonsingular matrix  $P \in GF(q)^{n \times n}$  such that

$$(\lambda I_n - A)^{-1}b = \frac{P^{-1}w(\lambda)}{\det(\lambda I_n - A)} \quad (6.1.22)$$

where

$$w^T(\lambda) \equiv (1, \lambda, \dots, \lambda^{n-1}). \quad (6.1.23)$$

21°. For any polynomial of the form

$$f_2(\lambda) = d_n(\lambda)^{n-1} + d_{n-1}(\lambda)^{n-2} + \dots + d_1 \quad (6.1.24)$$

there exists a vector  $v \in GF(q)^n$  such that

$$v^T(\lambda I_n - A)^{-1}b = \frac{f_2(\lambda)}{\det(\lambda I_n - A)} \quad (6.1.25)$$

22°. There exists a vector  $v \in GF(q)^n$  for which the expression

$$v^T(\lambda I_n - A)^{-1}b \quad (6.1.26)$$

is irreducible, that is,

$$v^T(\lambda I_n - A)^{-1}b = \frac{f(\lambda)}{g(\lambda)} \quad (6.1.27)$$

where  $f(\lambda)$  and  $g(\lambda)$  are coprime polynomials.

23°. The following matrix  $R \in GF(q)^{(\mu+\nu) \times (\mu+\nu)}$  has rank  $(\mu+\nu)$ :

$$R \equiv \begin{pmatrix} a_v & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & b_\mu \\ a_{v-1} & a_v & \dots & 0 & 0 & 0 & 0 & \dots & b_\mu & b_{\mu-1} \\ a_{v-2} & a_{v-1} & \dots & 0 & 0 & 0 & 0 & \dots & b_{\mu-1} & b_{\mu-2} \\ \vdots & \vdots & & \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ \vdots & \vdots & & \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ \vdots & \vdots & & \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \dots & a_0 & a_1 & b_1 & b_0 & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & a_0 & b_0 & 0 & \dots & 0 & 0 \end{pmatrix}$$

where  $a_i$ ,  $i \in \underline{v}$ , and  $b_j$ ,  $j \in \underline{\mu}$ , are the coefficients of the polynomials  $f(\lambda)$  and  $g(\lambda)$ , respectively, in (6.1.27).

24<sup>o</sup>. There does not exist any nonzero vector  $v \in GF(q)^n$  such that the expression

$$v^T (\lambda I_n - A)^{-1} b \quad (6.1.28)$$

is identically equal to zero.

25<sup>o</sup>. For every set of distinct scalars  $\lambda_i$ ,  $i \in \underline{n}$ , different from the eigenvalues of the matrix  $A$ , the vectors

$$(\lambda_i I_n - A)^{-1} b, \quad i \in \underline{n} \quad (6.1.29)$$

are linearly independent.

Proof. We want to show that the following chain of sequential implications is closed:

$$\begin{aligned}
1^0 &\Rightarrow 2^0 \Rightarrow 3^0 \Rightarrow 4^0 \Rightarrow 5^0 \Rightarrow 6^0 \Rightarrow 7^0 \Rightarrow 8^0 \Rightarrow 9^0 \Rightarrow 10^0 \Rightarrow \\
11^0 &\Rightarrow 12^0 \Rightarrow 13^0 \Rightarrow 14^0 \Rightarrow 15^0 \Rightarrow 16^0 \Rightarrow 17^0 \Rightarrow 18^0 \Rightarrow 19^0 \Rightarrow \\
20^0 &\Rightarrow 21^0 \Rightarrow 22^0 \Rightarrow 23^0 \Rightarrow 24^0 \Rightarrow 25^0 \Rightarrow 1^0
\end{aligned}$$

$1^0 \Rightarrow 2^0$ : Suppose that property  $2^0$  is not satisfied. Then there exists an isomorphism  $P : X \rightarrow X$  such that the isomorphic LSM  $\tilde{M}_1 = (\tilde{A}, \tilde{b}) \equiv (PAP^{-1}, Pb)$  is described by the equations (6.1.2). From equation (6.1.2b) it is clear that if  $\tilde{x}^{II}(k') = 0$  for any clock period  $k'$ , then  $\tilde{x}^{II}(k'') = 0$  for any other clock period  $k''$ . Therefore,  $\tilde{M}_1 = (\tilde{A}, \tilde{b})$  and consequently  $M_1 = (a, b)$  cannot be reachable. Hence  $1^0 \Rightarrow 2^0$ .

$2^0 \Rightarrow 3^0$ : Suppose that property  $3^0$  is not satisfied. Then the vector  $b$  belongs to an  $A$ -invariant subspace  $S \subseteq X$  of dimension  $r < n$ . Let

$$\{s^1, s^2, \dots, s^r, s^{r+1}, \dots, s^n\} \quad (6.1.30)$$

be a basis for  $X$  such that  $\{s^1, s^2, \dots, s^r\}$  forms a basis for  $S$ .

Then any vector  $x \in S$  can be uniquely expressed as

$$x = a_1 s^1 + a_2 s^2 + \dots + a_r s^r \quad (6.1.31)$$

for appropriate  $a_i \in GF(q)$ . Since  $S$  is  $A$ -invariant,  $Ax \in S$  and in view of (6.1.31),

$$\begin{aligned}
Ax &= b_1 A s^1 + b_2 A s^2 + \dots + b_r A s^r \\
&= b_1 A s^1 + b_2 A s^2 + \dots + b_r A s^r + 0 s^{r+1} + \dots + 0 s^n
\end{aligned} \quad (6.3.32)$$

for appropriate  $b_i \in GF(q)$ . From (6.1.31) it follows that the matrix representations  $\tilde{A}$  and  $\tilde{b}$  of  $A$  and  $b$  with respect to the basis (6.1.30) have the forms given by (6.1.1), and hence property  $2^0$  is not satisfied. Therefore,  $2^0 \Rightarrow 3^0$ . Clearly  $P^{-1} = [s^1, s^2, \dots, s^n]$ .

$3^0 \Rightarrow 4^0$ : Suppose that property  $4^0$  is not satisfied. Then there exists a nonzero vector  $v \in GF(q)^n$  such that

$$v^T A = \lambda v^T \quad (6.1.33)$$

$$v^T b = 0 \quad (6.1.34)$$

Assume that a subspace  $W \subseteq X$  is orthogonal to the vector  $b$  so that  $v^T w = 0$  for all  $w \in W$ . From equation (6.1.33) it follows that  $v^T A w = \lambda v^T w = 0$  which implies that  $W$  is an  $A$ -invariant subspace. Since by hypothesis  $v \neq 0$ ,  $\dim W < n$ . Furthermore, equation (6.1.34) shows that  $b \in W$ . That is, there exists an  $A$ -invariant subspace of dimension smaller than  $n$  containing  $b$ . This conclusion obviously contradicts property  $3^0$  and hence  $3^0 \Rightarrow 4^0$ .

$4^0 \Rightarrow 5^0$ : Suppose that property  $5^0$  is not satisfied. Then there exists a subspace  $V \subseteq X$  such that  $v^T b = 0$  for all  $v \in V$ . Furthermore, there exists a subspace  $W \subseteq V$  which is  $A^T$ -invariant. Therefore, for any  $v \in W$  we have  $A^T v \in W$ . We need to show that  $W$  contains an eigenvector of  $A^T$ . Let  $\dim W = \ell$ . Then given any  $v \in W$ , there exists an integer  $\nu$ ,  $1 \leq \nu \leq \ell$ , such that the vectors  $v, A^T v, (A^T)^2 v, \dots, (A^T)^{\nu-1} v$  are linearly independent, but for some  $a_i \in GF(q)$ ,  $i \in \underline{\nu-1}$ ,

$$a_0 v + a_1 A^T v + a_2 (A^T)^2 v + \dots + a_{\nu-1} (A^T)^{\nu-1} v + (A^T)^\nu v = 0 \quad (6.1.35)$$

Let  $\lambda$  be a solution of the following equation:

$$b_0 + b_1\lambda + \dots + b_{v-1}(\lambda)^{v-1} + (\lambda)^v = 0 \quad (6.1.36)$$

Then the  $v+1$  equations

$$\begin{aligned} \lambda c_0 &= b_0 \\ -c_0 + \lambda c_1 &= b_1 \\ &\vdots \\ -c_{v-2} + \lambda c_{v-1} &= b_{v-1} \\ -c_{v-1} &= 1 \end{aligned} \quad (6.1.37)$$

determine a unique solution for the  $v$  quantities  $c_0, c_1, \dots, c_{v-1}$ . It is clear that equation (6.1.36) is the condition that the first of equations (6.1.37) should be satisfied. Substituting for  $b_i$ ,  $i \in \underline{v-1}$ , from (6.1.37) into (6.1.35) and rearranging, we obtain

$$(\lambda I_n - A^T)(c_0 v + c_1 A^T v + c_2 (A^T)^2 v + \dots + c_{v-1} (A^T)^{v-1} v) = 0 \quad (6.1.38)$$

Letting

$$c_0 v + c_1 A^T v + c_2 (A^T)^2 v + \dots + c_{v-1} (A^T)^{v-1} v \equiv w \quad (6.1.39)$$

equation (6.1.38) reduces to

$$A^T w = \lambda w$$

showing that  $w$  is an eigenvector of  $A^T$ . Clearly  $w$  is nonzero and lies in the subspace  $\mathcal{W}$ , which contradicts property  $4^0$ . Hence  $4^0 \implies 5^0$ .

$5^0 \implies 6^0$ : If  $0 \neq v \in \text{GF}(q)^n$ , then obviously  $v^T b = 0$  or  $v^T b \neq 0$ . If  $v^T b = 0$  and property  $5^0$  is satisfied, then not every one of the vectors

$$v, A^T v, (A^T)^2 v, \dots, (A^T)^{n-1} v$$

is orthogonal to  $b$ . Therefore,

$$v^T (b \quad Ab \quad A^2 b \quad \dots \quad A^{n-1} b) \neq 0$$

and hence

$$\text{rank}(b \quad Ab \quad A^2 b \quad \dots \quad A^{n-1} b) = n$$

Therefore,  $5^0 \implies 6^0$ .

$6^0 \iff 7^0$ : This follows immediately from the fact that the rank of a matrix does not change after premultiplying or postmultiplying a nonsingular matrix.

$7^0 \iff 8^0 \iff 9^0$ : Introducing the feedback law

$$u(k) = f^T x(k), \quad f \in \text{GF}(q)^n \quad (6.1.40)$$

the LSM  $(A, b)$  becomes

$$x(k+1) = (A + bf^T)x(k) \quad (6.1.41)$$



Starting from an arbitrary initial state  $x(0)$  and applying equation (6.1.41)  $n$  times, we obtain the relationship

$$x(n) = (A + bf^T)^n x(0) \quad (6.1.42)$$

For the state  $x(n)$  to be zero, we must have  $(A + bf^T)^n = 0$ , that is, the matrix  $(A + bf^T)$  must be  $n$ -nilpotent. Now postmultiplying both sides of the matrix identity (see Lemma 5.3.1)

$$\begin{aligned} (A + bf^T)^n &= A^n + bf^T(A + bf^T)^{n-1} + Abf^T(A + bf^T)^{n-2} \\ &+ \dots + A^{n-3}bf^T(A + bf^T)^2 + A^{n-2}bf^T(A + bf^T) \\ &+ A^{n-1}bf^T \end{aligned} \quad (6.1.43)$$

by  $x(0)$ , we get

$$\begin{aligned} (A + bf^T)^n x(0) &= A^n x(0) + bf^T(A + bf^T)^{n-1} x(0) + Abf^T(A + bf^T)^{n-2} x(0) \\ &+ \dots + A^{n-2}bf^T(A + bf^T)x(0) + A^{n-1}bf^T x(0) \end{aligned} \quad (6.1.44)$$

In view of  $(A + bf^T)^n = 0$  and equation (6.1.42), equation (6.1.44) reduces to

$$A^n x(0) + bf^T x(n-1) + abf^T x(n-2) + \dots + A^{n-2}bf^T x(1) + A^{n-1}bf^T x(0) = 0 \quad (6.1.45)$$

Premultiplying equation (6.1.45) by  $f^T A^{-n}$  yields

$$\begin{aligned} f^T x(0) + f^T A^{-n} bf^T x(n-1) + f^T A^{-n+1} bf^T x(n-2) \\ + \dots + f^T A^{-2} bf^T x(1) + f^T A^{-1} bf^T x(0) = 0 \end{aligned} \quad (6.1.46)$$

Substituting  $u(k) = f^T x(k)$ ,  $k = 0, 1, \dots, n-1$ , equation (6.1.46) gives

$$\begin{aligned} -u(0) &= f^T A^{-n} b u(n-1) + f^T A^{-n+1} b u(n-2) \\ &+ \dots + f^T A^{-2} b u(1) + f^T A^{-1} b u(0) \end{aligned} \quad (6.1.47)$$

Since equation (6.1.47) is an identity for any  $u(k)$ ,  $k \in \underline{n-1}$ , it follows that

$$\begin{aligned} f^T A^{-1} b &= -1 \\ f^T A^{-i} b &= 0, \quad i = 2, 3, \dots, n \end{aligned} \quad (6.1.48)$$

That is, the matrix equation  $(A + bf^T)^n = 0$  is equivalent to the set of scalar equations (6.1.48). Equations (6.1.48) will admit a unique solution with respect to the components of the vector  $f$  if and only if they are linearly independent, that is, if and only if the rank of the matrix

$$(A^{-n}b, A^{-n+1}b, \dots, A^{-1}b) = A^{-n}(b, Ab, A^2b, \dots, A^{n-1}b) \equiv A^{-n}K_1$$

is  $n$ . But since  $A$  is nonsingular, the matrix  $A^{-n}K_1$  will have rank  $n$  if and only if  $\text{rank } K_1 = n$ . Hence  $7^0 \iff 8^0 \iff 9^0$ .

$$9^0 \implies 10^0: \quad \text{Since } 9^0 \implies 6^0, \text{ we want to show that } 6^0 \implies 10^0.$$

In the matrix  $E_1$ , given by (6.1.7), if we add  $A$  times the first (block) row to the second (block) row, then add  $A$  times the second (block) row to the third (block) row, and so on, until  $E_1$  is reduced to the form

$$\tilde{E}_1 \equiv \begin{pmatrix} I_n & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & b \\ 0 & I_n & \dots & 0 & 0 & 0 & \dots & 0 & b & Ab \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & b & Ab & A^2b \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & I_n & 0 & b & \dots & A^{n-4}b & A^{n-3}b & A^{n-2}b \\ 0 & 0 & \dots & 0 & b & Ab & \dots & A^{n-3}b & A^{n-2}b & A^{n-1}b \end{pmatrix} \quad (6.1.49)$$

then  $\tilde{E}_1$  has the same rank as  $E_1$ , and the rank of  $\tilde{E}_1$  is  $n^2$  if and only if the rank of  $(b \quad Ab \quad A^2b \quad \dots \quad A^{n-1}b)$  is  $n$ . Thus  $6^o \iff 10^o$ .

$10^o \implies 11^o$ : In the equation  $(\xi I_n - A)x(\xi) + by(\xi) = z(\xi)$ , writing  $x(\xi)$ ,  $y(\xi)$ , and  $z(\xi)$  as

$$x(\xi) = x^0 + x^1\xi + x^2(\xi)^2 + \dots + x^{n-2}(\xi)^{n-2}$$

$$y(\xi) = y_0 + y_1\xi + y_2(\xi)^2 + \dots + y_{n-1}(\xi)^{n-1}$$

$$z(\xi) = z^0 + z^1\xi + z^2(\xi)^2 + \dots + z^{n-1}(\xi)^{n-1}$$

multiplying out the products, and equating coefficients of like powers of  $\xi$ , we obtain the following set of equations:

$$\begin{array}{rcl} x^{n-1} & +by_{n-1} & = z^{n-1} \\ -Ax^{n-1} + x^{n-2} & +by_{n-2} & = z^{n-2} \\ \cdot & \cdot & \\ \cdot & \cdot & \\ -Ax^1 + x^0 & +by_1 & = z^1 \\ -Ax^0 & +by_0 & = z^0 \end{array}$$

The coefficient matrix of this set of equations is precisely the matrix  $E_1$  given by (6.1.7), and hence this set of equations will have a unique solution for  $z^0, z^1, \dots, z^{n-1}$ , if and only if the matrix  $E$  has rank  $n^2$ .

$11^0 \Rightarrow 12^0$ : Let  $e^i$  denote the  $i$ th column of the identity matrix  $I_n$ , and in equation (6.1.8), let  $z(\xi) = e^i$ ,  $i \in \underline{n}$ . Furthermore, if we denote by  $(x^{(i)}(\xi), y^{(i)}(\xi))$  the corresponding solutions, then the matrix  $X(\xi)$  having  $x^{(i)}(\xi)$  as its columns and the vector  $y(\xi)$  having  $y^{(i)}(\xi)$  as its components will satisfy equation (6.1.9). Therefore,  $11^0 \Rightarrow 12^0$ .

$12^0 \Rightarrow 13^0$ : Rewriting equation (6.1.9) as

$$[(\xi I_n - A), b] \begin{pmatrix} X(\xi) \\ y^T(\xi) \end{pmatrix} = I_n$$

it is clear that  $\text{rank} [(\xi I_n - A), b] = n$ .

$13^0 \Rightarrow 14^0$ : Suppose that the matrix  $[(\xi I_n - A), b]$  has the Smith canonical form  $[Z(\xi), 0]$ , where  $Z(\xi) = z_1(\xi) \oplus z_2(\xi) \oplus \dots \oplus z_n(\xi)$ . If  $z_n(\lambda) = 0$ , then  $\lambda$  is a zero of the  $n$ th determinantal divisor of  $[(\xi I_n - A), b]$  and hence  $\lambda$  is an eigenvalue of  $\xi I_n - A$ . Furthermore, the rank of  $[Z(\xi), 0]$  is the same as the rank of  $[\xi I_n - A, b]$  for all  $\xi$ . Since  $\text{rank} [(\xi I_n - A), b] = n$ , by hypothesis, it follows that  $z_n(\xi) = 1$  and hence  $Z(\xi) = I_n$ .

$14^0 \Rightarrow 15^0$ : We prove the chain of implications  $14^0 \Rightarrow 13^0 \Rightarrow 4^0 \Rightarrow 5^0 \Rightarrow 15^0$  of which  $4^0 \Rightarrow 5^0$  has already been proved. To show that  $14^0 \Rightarrow 13^0$ , assume that  $\xi I_n - A$  and  $b$  are coprime. Then the Smith canonical form of  $[\xi I_n - A, b]$  is  $[I_n, 0]$  which implies that the  $n$ th determinantal divisor of  $[\xi I_n - A, b]$  is unity. Therefore, the matrix  $[\xi I_n - A, b]$  has rank  $n$  for all  $\xi$ . The implication  $13^0 \Rightarrow 4^0$  follows from the fact that if the matrix  $[\xi I_n - A, b]$  has rank  $n$ , then its rows are linearly independent and hence for any vector  $v$ ,  $v^T [\xi I_n - A, b] = 0$  implies that  $v = 0$ . Therefore, there exists no nonzero vector which simultaneously satisfies  $v^T (\xi I_n - A) = 0$  and  $v^T b = 0$ . To show the last implication  $5^0 \Rightarrow 15^0$ , we observe that any expression of the form

$$P = \sum_{i=1}^n a_i A^{i-1}, \quad a_i \in GF(q) \quad (6.1.50)$$

satisfies the equation  $PA - AP = 0$ . Substituting (6.1.50) into  $Pb = d$  yields

$$Pb = \sum_{i=1}^n A^{i-1} b a_i = [b, Ab, A^2 b, \dots, A^{n-1} b] \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = d \quad (6.1.51)$$

From equation (6.1.51) it is clear that the solution (6.1.50) is unique if and only if  $[b, Ab, A^2 b, \dots, A^{n-1} b]$  is nonsingular. Hence  $5^0 \Leftrightarrow 15^0$ .

$15^0 \Rightarrow 16^0$ : We prove the chain of implications  $15^0 \Rightarrow 4^0 \Rightarrow 5^0 \Rightarrow 6^0 \Rightarrow 16^0$  of which only  $15^0 \Rightarrow 4^0$  and  $6^0 \Rightarrow 16^0$  have not yet been proved. To show that  $15^0 \Rightarrow 4^0$ , suppose that there exists a nonzero eigenvector  $z$  of the matrix  $A$  which is orthogonal to the vector  $b$ , that is,  $z^T A = \lambda z^T$ ,  $z^T b = 0$ , and  $z \neq 0$ . Hence  $z^T (A - \lambda I_n) = 0$  and, therefore,  $\det(A - \lambda I_n) = 0$ . But this conclusion implies that there exists a nonzero vector  $v$  such that  $Av = \lambda v$ . Now if we let  $P_0 \equiv vz^T \neq 0$ , then it is clear that  $P_0$  satisfies the relations  $P_0 A - AP_0 = 0$  and  $P_0 b = 0$ . Therefore, if  $P$  is a solution of equations (6.1.10), then so is  $P + P_0$  which contradicts the uniqueness condition of property  $15^0$ . Thus  $15^0 \Rightarrow 4^0$ . Finally, to prove that  $6^0 \Rightarrow 16^0$ , we recall that by the Cayley-Hamilton Theorem the matrices  $A$  and  $\tilde{A}$  satisfy their own characteristic equations which, by (6.1.12), are identical. Thus we can express  $A^n$  and  $\tilde{A}^n$  as

$$A^n = \sum_{i=1}^n a_i A^{i-1}, \quad \tilde{A}^n = \sum_{i=1}^n a_i \tilde{A}^{i-1}, \quad a_i \in GF(q) \quad (6.1.52)$$

Let us introduce the matrix

$$N \equiv \begin{pmatrix} 0 & 0 & \dots & 0 & a_1 \\ 1 & 0 & \dots & 0 & a_2 \\ 0 & 1 & \dots & 0 & a_3 \\ \vdots & \vdots & & \vdots & \vdots \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \dots & 1 & a_n \end{pmatrix} \quad (6.1.53)$$

which satisfies the equations

$$K_1 N = A K_1 \quad (6.1.54)$$

$$\tilde{K}_1 N = \tilde{A} \tilde{K}_1 \quad (6.1.55)$$

Since  $\tilde{K}_1$  is nonsingular by hypothesis and  $K_1$  is nonsingular by property  $6^0$ , in view of equations (6.1.54) and (6.1.55), we have  $\tilde{A} = P A P^{-1}$ , where  $P = \tilde{K}_1 K_1^{-1}$ . Furthermore,  $P K_1 = \tilde{K}_1$  and, in particular,  $P b = \tilde{b}$ .

$16^0 \Rightarrow 17^0$ : It can be easily checked directly that the pair  $(\tilde{A}, \tilde{b})$  defined by (6.1.14) satisfies all the conditions of property  $16^0$ , that is,  $\det[\tilde{b}, \tilde{A}\tilde{b}, \dots, \tilde{A}^{n-1}\tilde{b}] \neq 0$  and the characteristic polynomial of  $\tilde{A}$  is  $(\lambda)^n - \sum_{i=1}^n a_i(\lambda)^{i-1}$ . Therefore,  $16^0 \Rightarrow 17^0$ .

$17^0 \Rightarrow 18^0$ : We will prove that  $17^0 \Rightarrow 6^0 \Rightarrow 16^0 \Rightarrow 18^0$ . The implication  $17^0 \Rightarrow 6^0$  follows immediately since  $\text{rank}[\tilde{b}, \tilde{A}\tilde{b}, \dots, \tilde{A}^{n-1}\tilde{b}] = \text{rank } P[b, Ab, \dots, A^{n-1}b]P^{-1}$ . The implication  $6^0 \Rightarrow 16^0$  was proved as part of the proof of the implication  $15^0 \Rightarrow 16^0$ . Therefore, it remains to prove only the implication  $16^0 \Rightarrow 18^0$ . This can be accomplished by showing that the pair  $(\tilde{A}, \tilde{b})$  defined by (6.1.15) satisfies the conditions of property  $16^0$ . From (6.1.18) it is clear that the characteristic polynomials of  $\tilde{A}$  and  $A$  are identical. To check the second condition, we need to see if the matrix  $[\tilde{b}, \tilde{A}\tilde{b}, \dots, \tilde{A}^{n-1}\tilde{b}]$ , where  $\tilde{A}$  and  $\tilde{b}$  are given by (6.1.15) - (6.1.17), is nonsingular. First, we will show that the pair  $(\tilde{A}, \tilde{b})$  satisfies property  $4^0$ , that is, there exists no nonzero eigenvector of  $\tilde{A}$  which is orthogonal to  $\tilde{b}$ . From

(6.1.15) and (6.1.16) it follows that the matrix  $\tilde{A}$  has only  $\underline{v}$  eigenvectors of the form

$$v^i = \begin{pmatrix} w^{1i} \\ w^{2i} \\ \cdot \\ \cdot \\ w^{\ell i} \end{pmatrix}, \quad i \in \underline{v}$$

where  $w^{ji}$  are  $n_j$ -vectors given by the relations

$$w^{ji} = \begin{cases} \begin{pmatrix} 0 \\ 0 \\ \cdot \\ \cdot \\ 0 \end{pmatrix} & \text{if } i \neq j \\ \begin{pmatrix} 0 \\ 0 \\ \cdot \\ \cdot \\ a \end{pmatrix} & a \neq 0, \text{ for } i = j \end{cases}$$

Now using (6.1.15) and (6.1.17), we see that all the eigenvectors  $v^i$  of the matrix  $\tilde{A}$  satisfy the condition  $v^{iT} \tilde{b} \neq 0, i \in \underline{v}$ . Since  $4^0 \Rightarrow 6^0$ , it follows that the matrix  $[\tilde{b}, \tilde{A}\tilde{b}, \dots, \tilde{A}^{n-1}\tilde{b}]$  is nonsingular.

$18^0 \Rightarrow 19^0$ : We will consider the chain of implications  $18^0 \Rightarrow 6^0 \Rightarrow 17^0 \Rightarrow 19^0$  of which  $18^0 \Rightarrow 6^0$  was proved in the preceding discussion when we showed that the matrix  $[\tilde{b}, \tilde{A}\tilde{b}, \dots, \tilde{A}^{n-1}\tilde{b}]$  is



nonsingular, and  $6^0 \Rightarrow 17^0$  has already been proved. Therefore, we need to show that  $17^0 \Rightarrow 19^0$ . Assuming the pair  $(A, b)$  has been transformed to  $(\tilde{A}, \tilde{b})$  given by (6.1.14), consider an arbitrary  $n$ -vector  $\tilde{v}^T = (v_1, v_2, \dots, v_n)$ . Using (6.1.14), we obtain

$$\tilde{A} + \tilde{b}\tilde{v}^T = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & & 1 \\ a_1+v_1 & a_2+v_2 & a_3+v_3 & \dots & a_n+v_n \end{pmatrix} \quad (6.1.56)$$

The matrix (6.1.56) has the following characteristic polynomial:

$$f(\lambda) = (\lambda)^n - \sum_{i=1}^n (a_i + v_i)(\lambda)^{i-1} \quad (6.1.57)$$

Now if a polynomial  $f_1(\lambda)$  of the form (6.1.20) is specified, then we can choose

$$v_i = b_i - a_i, \quad i \in \underline{n}$$

so that

$$\det[\lambda I_n - (\tilde{A} + \tilde{b}\tilde{v}^T)] = f_1(\lambda) \quad (6.1.58)$$

Using the relations  $\tilde{A} \equiv PAP^{-1}$  and  $\tilde{b} \equiv Pb$ , and (6.1.58), we find that

$$\begin{aligned}
f(\lambda_1) &= \det(\lambda I_n - PAP^{-1} - Pb\tilde{v}^T) \\
&= \det P(\lambda I_n - A - b\tilde{v}^T P)P^{-1} \\
&= \det(\lambda I_n - A - b\tilde{v}^T P) \\
&= \det(\lambda I_n - A - bv^{0T})
\end{aligned}$$

where  $v^{0T} \equiv \tilde{v}^T P$ .

$19^0 \Rightarrow 20^0$ : We will show that  $19^0 \Rightarrow 4^0 \Rightarrow 17^0 \Rightarrow 20^0$ .

Since the implication  $4^0 \Rightarrow 17^0$  has already been proved, it remains to prove the implications  $19^0 \Rightarrow 4^0$  and  $17^0 \Rightarrow 20^0$ . To prove  $19^0 \Rightarrow 4^0$ , assume that property  $4^0$  is not satisfied. Then there exists a vector  $z$  such that  $z^T A = \lambda_0 z^T$ ,  $z^T b = 0$ , and  $z \neq 0$ . Therefore, for any vector  $v$  we have  $z^T(A + bv^T) = z^T A = \lambda_0 z^T$ . But this conclusion implies that  $\lambda_0$  is an eigenvalue of the matrix  $(A + bv^T)$  for any vector  $v$ . Therefore, if we choose a polynomial of the form (6.1.20) which does not vanish for  $\lambda = \lambda_0$ , then (6.1.21) is violated for every vector  $v$ . This obviously contradicts property  $19^0$ . Hence  $19^0 \Rightarrow 4^0$ . The implication  $17^0 \Rightarrow 20^0$  is easily proved by using the relations (6.1.14) and directly evaluating  $(\lambda I_n - \tilde{A})^{-1}b$  to obtain

$$(\lambda I_n - \tilde{A})^{-1}b = \frac{w(\lambda)}{\det(\lambda I_n - \tilde{A})}$$

where  $w(\lambda)$  is defined by (6.1.23). Now in view of the relations  $\tilde{A} \equiv PAP^{-1}$  and  $\tilde{b} \equiv Pb$ , it follows that

$$(\lambda I_n - A)^{-1} = \frac{P^{-1} w(\lambda)}{\det(\lambda I_n - A)} \quad (6.1.59)$$

$20^0 \Rightarrow 21^0$ : The polynomial  $f_2(\lambda)$  defined by (6.1.24) can be written in terms of  $w(\lambda)$  as

$$f_2(\lambda) = d^T w(\lambda) \quad (6.1.60)$$

where  $d^T \equiv (d_1, d_2, \dots, d_n)$ . Using (6.1.59) and (6.1.60), we can express (6.1.25) equivalently as

$$v^T P^{-1} w(\lambda) = d^T w(\lambda) \quad (6.1.61)$$

From (6.1.61) it follows that  $v^T P^{-1} = d^T$ , and consequently the required vector is given by  $v = P^T d$ .

$21^0 \Rightarrow 22^0$ : This implication is obvious since we can choose a polynomial  $f_2(\lambda)$  such that  $f_2(\lambda)/\det(\lambda I_n - A)$  is irreducible and then apply property  $21^0$  to determine a vector  $v$  which satisfies property  $22^0$ .

$22^0 \Rightarrow 23^0$ : Let the polynomials  $f(\lambda)$  and  $g(\lambda)$  in (6.1.27) have the forms

$$f(\lambda) = b_0 + b_1 \lambda + b_2 (\lambda)^2 + \dots + b_\mu (\lambda)^\mu \quad (6.1.62)$$

and

$$g(\lambda) = a_0 + a_1 \lambda + a_2 (\lambda)^2 + \dots + a_\nu (\lambda)^\nu \quad (6.1.63)$$

Suppose that  $f(\lambda)$  and  $g(\lambda)$  are not coprime. Then there exists a common root  $a$  and consequently  $f(\lambda)$  and  $g(\lambda)$  can be written as follows:

$$f(\lambda) = (\lambda - a)(-\bar{b}_0 - \bar{b}_1\lambda - \dots - \bar{b}_{\mu-1}(\lambda)^{\mu-1}) \quad (6.1.64)$$

$$g(\lambda) = (\lambda - a)(-\bar{a}_0 - \bar{a}_1\lambda - \dots - \bar{a}_{v-1}(\lambda)^{v-1}) \quad (6.1.65)$$

Eliminating the common factor  $(\lambda - a)$  in (6.1.64) and (6.1.65), we get

$$g(\lambda)(\bar{b}_0 + \bar{b}_1\lambda + \dots + \bar{b}_{\mu-1}(\lambda)^{\mu-1}) + f(\lambda)(\bar{a}_0 + \bar{a}_1\lambda + \dots + \bar{a}_{v-1}(\lambda)^{v-1}) = 0 \quad (6.1.66)$$

Substituting for  $f(\lambda)$  and  $g(\lambda)$  from (6.1.62) and (6.1.63) into (6.1.66), and equating the coefficients of each power of  $\lambda$  to zero, we obtain the following set of linear equations:

$$\begin{aligned} a_v \bar{b}_{\mu-1} + b_\mu \bar{a}_{v-1} &= 0 \\ a_{v-1} \bar{b}_{\mu-1} + a_v \bar{b}_{\mu-2} + b_\mu \bar{a}_{v-2} + b_{\mu-1} \bar{a}_{v-1} &= 0 \\ &\vdots \\ a_0 \bar{b}_0 + b_0 \bar{a}_0 &= 0 \end{aligned} \quad (6.1.67)$$

It is clear that the set of equations (6.1.67) will have a nonzero solution  $\bar{a}_0, \bar{a}_1, \dots, \bar{a}_{v-1}, \bar{b}_0, \bar{b}_1, \dots, \bar{b}_{\mu-1}$ , if and only if the determinant of the coefficient matrix which is precisely the matrix  $R$  of property 23<sup>0</sup>, is zero. Therefore,  $22^0 \Leftrightarrow 23^0$ .

$23^0 \Rightarrow 24^0$ : Since  $22^0 \Leftrightarrow 23^0$ , we will show that  $22^0 \Rightarrow 24^0$  by using the chain of implications  $22^0 \Rightarrow 2^0 \Rightarrow 21^0 \Rightarrow 24^0$  of which only the implications  $22^0 \Rightarrow 2^0$  and  $21^0 \Rightarrow 24^0$  are not yet proved. To prove the implication  $22^0 \Rightarrow 2^0$ , suppose that property  $2^0$  is not satisfied. Then the pair can be brought into the form (6.1.1). If we write the arbitrary vector  $v^T \in GF(q)^{1 \times n}$  in the form  $v^T \equiv (v^{1T}, v^{2T})$ , where  $v^{1T} \in GF(q)^{1 \times r}$ , and use the relations defined by (6.1.1), we obtain

$$v^T(\lambda I_n - \tilde{A})^{-1} \tilde{b} = [v^{1T}, v^{2T}] \begin{bmatrix} (\lambda I_r - \tilde{A}_{11})^{-1} & (\lambda I_r - \tilde{A}_{11})^{-1} \tilde{A}_{12} (\lambda I_{n-r} - \tilde{A}_{22})^{-1} \\ 0 & (\lambda I_{n-r} - \tilde{A}_{22})^{-1} \end{bmatrix} \begin{bmatrix} b^1 \\ 0 \end{bmatrix}$$

$$= v^{1T}(\lambda I_r - \tilde{A}_{11})^{-1} \tilde{b}^1 \quad (6.1.68)$$

$$\equiv \frac{h(\lambda)}{\det(\lambda I_r - \tilde{A}_{11})} \quad (6.1.69)$$

Since in (6.1.69) the degree of the polynomial  $\det(\lambda I_r - \tilde{A}_{11})$  is strictly less than  $n$ , it follows that any expression of the form  $v^T(\lambda I_n - \tilde{A})^{-1} \tilde{b}$  is reducible. But  $v^T(\lambda I_n - \tilde{A})^{-1} \tilde{b} = v^T(\lambda I_n - PAP^{-1})^{-1} Pb = v^T P(\lambda I_n - A)^{-1} b$ . Hence any expression of the form  $\hat{v}^T(\lambda I_n - A)^{-1} b$ , where  $\hat{v}^T \equiv v^T P$ , is also reducible. This conclusion obviously contradicts property  $22^0$ . To prove the implication  $21^0 \Rightarrow 24^0$ , we observe that the identity

$$\hat{v}^T(\lambda I_n - A)b = 0 \quad (6.1.70)$$

is satisfied for  $\hat{v} = 0$ . From the uniqueness of the vector  $\hat{v}$  which satisfies the conditions of property 21<sup>0</sup>, it follows that the vector  $\hat{v} = 0$  is the only vector which satisfies (6.1.70).

24<sup>0</sup>  $\Rightarrow$  25<sup>0</sup>: Suppose that property 25<sup>0</sup> is not satisfied.

Then there exists a vector  $v \neq 0$  such that

$$v^T(\lambda_i I_n - A)^{-1}b = 0, \quad i \in \underline{n}; \quad \det(\lambda_i I_n - A) \neq 0 \quad (6.1.71)$$

Since  $v^T(\lambda_i I_n - A)^{-1}b$  can be written as

$$v^T(\lambda_i I_n - A)^{-1}b \equiv \frac{\hat{f}(\lambda)}{\det(\lambda I_n - A)} \quad (6.1.72)$$

where  $\deg \hat{f}(\lambda) \leq n-1$ , in view of (6.1.71), it follows that

$$\hat{f}(\lambda_i) = 0, \quad i \in \underline{n}$$

Since  $\deg \hat{f}(\lambda) \leq n-1$  and  $\hat{f}$  vanishes at  $n$  distinct scalars  $\lambda_i$ , it must be identically zero. This conclusion implies that in (6.1.72),  $v^T(\lambda_i I_n - A)^{-1}b = 0$ . However, by hypothesis  $v^T \neq 0$ , and therefore it contradicts property 24<sup>0</sup>.

25<sup>0</sup>  $\Rightarrow$  1<sup>0</sup>: We will prove that 25<sup>0</sup>  $\Rightarrow$  2<sup>0</sup>  $\Rightarrow$  6<sup>0</sup>  $\Rightarrow$  1<sup>0</sup>.

Suppose that property 2<sup>0</sup> is not satisfied. Then the LSM  $(A, b)$  can be brought into the form (6.1.1) which clearly shows that the last  $n-r$  components of the vector  $(\lambda I_n - \tilde{A})^{-1}\tilde{b}$  are zero. Therefore, all vectors of this form belong to a subspace of a dimension smaller than  $n$  and consequently we cannot find  $n$  linearly independent vectors of this form. Furthermore, since  $(\lambda I_n - \tilde{A})^{-1}\tilde{b} = (\lambda I_n - PAP^{-1})^{-1}Pb = P(\lambda I_n - A)^{-1}b$ ,

it follows that there do not exist  $n$  linearly independent vectors of the form  $(\lambda I_n - A)^{-1}b$ , which contradicts property  $25^0$ , and proves the implication  $25^0 \Rightarrow 2^0$ . The implication  $6^0 \Rightarrow 1^0$  is obvious since  $\text{rank } K_1 = n$  implies that the range  $R(K_1)$  of the linear map  $K_1 : U^* \rightarrow X$ , which is the set of all reachable states of the LSM  $M_1 = (A, b)$ , is equal to  $X$ .  $\square$

## 6.2. Equivalence Classes of State Reachable Single-Input LSMs

In Section 5.1 we discussed the role of a particular transformation group, namely the group  $G_1$  given by (5.1.35), in identifying certain canonical forms for state reachable LSMs. In this section, we will briefly discuss the equivalence relation  $G_1$ -equivalence in relation to state reachable single-input LSMs.

Consider the following relation  $\rho_1$  on the set of all single-input LSMs  $(A, b)$  of the same order  $n$ :  $(A_1, b^1) \rho_1 (A_0, b^0)$ , that is, the LSM  $(A_1, b^1)$  is related to the LSM  $(A_0, b^0)$  under the relation  $\rho_1$ , if and only if there exists an isomorphism  $P_0 : X \rightarrow X$ ,  $P \in \text{GF}(n, q)$ , and a vector  $v^0 \in \text{GF}(q)^n$  such that

$$A_1 = P_0(A_0 + b^0 v^{0T})P_0^{-1}; \quad b^1 = P_0 b^0 \quad (6.2.1)$$

We want to show that  $\rho_1$  is an equivalence relation and hence allows the partition of the set of all reachable LSMs  $(A, b)$  into equivalence classes.

Theorem 6.2.1.  $\rho_1$  is an equivalence relation.

Proof. We need to show that the relation  $\rho_1$  is reflexive, symmetric, and transitive. Taking  $P_0 = I_n$  and  $v^0 = 0$ , shows that  $(A_0, b^0)_{\rho_1} (A_0, b^0)$  and hence  $\rho_1$  is reflexive. To show that  $\rho_1$  is symmetric, let  $A_0 = P_1(A_1 + b^1 v^{1T})P_1^{-1}$  and  $b^0 = P_1 b^1$ , where  $P_1 \equiv P_0^{-1}$  and  $v^1 \equiv -(P_0^{-1})^T v^0$ . Then  $(A_1, b^1)_{\rho_1} (A_0, b^0) \iff (A_0, b^0)_{\rho_1} (A_1, b^1)$ . Finally, suppose that  $(A_1, b^1)_{\rho_1} (A_0, b^0)$  and  $(A_2, b^2)_{\rho_1} (A_1, b^1)$ . We want to show that  $(A_2, b^2)_{\rho_1} (A_0, b^0)$ . Let  $(A_2, b^2)$  be related to  $(A_1, b^1)$  by the relations

$$A_2 = P_1(A_1 + b^1 v^{1T})P_1^{-1}; \quad b^2 = P_1 b^1 \quad (6.2.2)$$

Substituting into (6.2.2) for  $A_1$  and  $b^1$  from (6.2.1), we obtain

$$\begin{aligned} A_2 &= P_1 [P_0(A_0 + b^0 v^{0T})P_0^{-1} + P_0 b^0 v^{0T}]P_1^{-1} \\ &= (P_1 P_0) [A_0 + b^0 (v^0 + P_0^T v^1)^T] (P_1 P_0)^{-1} \end{aligned} \quad (6.2.3)$$

Similarly,

$$b^2 = (P_1 P_0) b^0 \quad (6.2.4)$$

From (6.2.3) and (6.2.4) we see that the LSM  $(A_2, b^2)$  is related to the LSM  $(A_0, b^0)$  by the relations

$$A_2 = P_2(A_0 + b^0 v^{2T})P_2^{-1}; \quad b^2 = P_2 b^0$$

where  $P_2 \equiv P_1 P_0$  and  $v^2 \equiv v^0 + P_0^T v^1$ .  $\square$

Corollary 6.2.1. The relation  $(A_1, b^1)_{\rho_1} (A_0, b^0)$ , where  $A_1 = P_0 A P_0^{-1}$  and  $b^1 = P_0 b^0$ , is an equivalence relation.



Corollary 6.2.2. The relation  $(A_1, b^1) \rho_1'' (A_0, b^0)$ , where  $A_1 = A_0 + b^0 v^{0T}$  and  $b^1 = b^0$ , is an equivalence relation.

The above result points out the existence of a strong link among state reachable single-input LSMs as shown in the following theorem.

Theorem 6.2.2. All state reachable single-input LSMs  $M = (A, b)$  of the same dimension  $n$  belong to an equivalence class of LSMs with respect to the relation defined by (6.2.1). That is, if  $M_0 = (A_0, b^0)$  is a reachable LSM, then all the LSMs  $M_1 = (A_1, b^1)$ , obtained by the relations

$$A_1 = P_0(A_0 + b^0 v^{0T})P_0^{-1}; \quad b^1 = P_0 b^0 \quad (6.2.5)$$

are reachable; conversely, if  $M_0 = (A_0, b^0)$  and  $M_1 = (A_1, b^1)$  are two reachable LSMs of the same dimension, then there exists an isomorphism  $P_0 : X \rightarrow X$ ,  $P_0 \in \text{GF}(n, q)$ , and a vector  $v^0 \in \text{GF}(q)^n$  such that the LSMs are related by (6.2.5).

Proof. From (6.2.5) it is easy to see that the following identity holds:

$$A_0 + b^0 v^{0T} = P_0^{-1} [A_1 + b^1 (v^T - v^{0T}) P_0^{-1}] P_0$$

which implies that

$$\det(\lambda I_n - A_0 - b^0 v^{0T}) = \det[\lambda I_n - A_1 - b^1 (v^T - v^{0T}) P_0^{-1}] \quad (6.2.6)$$

If the LSM  $M_0 = (A_0, b^0)$  is reachable, then by property 19<sup>o</sup> of Theorem 6.1.1,  $\det(\lambda I_n - A_0 - b^0 v^T)$  can be made equal to any monic polynomial of the form

$$f(\lambda) = (\lambda)^n - \sum_{i=1}^n a_i(\lambda)^{i-1}$$

by an appropriate choice of the vector  $v$ . In other words, by an appropriate choice of the vector  $v^{1T} = (v^T - v^{0T})P_0^{-1}$ , the right-hand side of (6.2.6) can be made equal to any monic polynomial of degree  $n$ . Applying again property 19<sup>o</sup> of Theorem 6.1.1, it follows that the LSM  $M_1 = (A_1, b^1)$  is also reachable. Conversely, if the LSMs  $M_0 = (A_0, b^0)$  and  $M_1 = (A_1, b^1)$  are reachable and of the same dimension, then by virtue of property 17<sup>o</sup> of Theorem 6.1.1, there exist isomorphisms  $\tilde{P}_0, \tilde{P}_1 : X \rightarrow X$  such that the isomorphic LSMs  $\tilde{M}_0 = (\tilde{A}_0, \tilde{b}^0) \equiv (\tilde{P}_0 A_0 \tilde{P}_0^{-1}, \tilde{P}_0 b^0)$  and  $\tilde{M}_1 = (\tilde{A}_1, \tilde{b}^1) \equiv (\tilde{P}_1 A_1 \tilde{P}_1^{-1}, \tilde{P}_1 b^1)$  have the forms

$$\tilde{A}_0 \equiv \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ a_1 & a_2 & a_3 & \dots & a_n \end{pmatrix}; \quad \tilde{b}^0 \equiv \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} \quad (6.2.7)$$

where  $a_i, i \in \underline{n}$ , are the coefficients of the characteristic polynomial

$$f_0(\lambda) = (\lambda)^n - \sum_{i=1}^n a_i(\lambda)^{i-1}$$

of the matrix  $A_0$ ; and

$$\tilde{A}_1 \equiv \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ b_1 & b_2 & b_3 & \dots & b_n \end{pmatrix}; \quad \tilde{b}^1 \equiv \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} \quad (6.2.8)$$

where  $b_i$ ,  $i \in \underline{n}$ , are the coefficients of the characteristic polynomial

$$f_1(\lambda) = (\lambda)^n - \sum_{i=1}^n b_i(\lambda)^{i-1}$$

of the matrix  $A_1$ . We observe that the matrices  $\tilde{A}_0$  and  $\tilde{A}_1$  differ only in the last row. Let  $(a_1, a_2, \dots, a_n)$  denote the last row of  $\tilde{A}_0$  and  $(b_1, b_2, \dots, b_n)$  the last row of  $\tilde{A}_1$ . Then the following equality holds:

$$\tilde{A}_1 = \tilde{A}_0 + \tilde{b}^0 \tilde{v}^T \quad (6.2.9)$$

where the components of the vector  $\tilde{v}^T \equiv (v_1, v_2, \dots, v_n)$  are given by the equalities

$$v_i = b_i - a_i, \quad i \in \underline{n}$$

Replacing  $\tilde{A}_1$ ,  $\tilde{A}_0$ , and  $\tilde{b}^0$  in (6.2.9) by their equivalent expressions in terms of  $A_1$ ,  $A_0$ , and  $b^0$ , we obtain the equality

$$\tilde{P}_1 A_1 \tilde{P}_1^{-1} = \tilde{P}_0 A_0 \tilde{P}_0^{-1} + \tilde{P}_0 b^0 \tilde{v}^T$$

from which it follows that

$$A_1 = \tilde{P}_1^{-1} \tilde{P}_0 (A_0 + b^0 v^T \tilde{P}_0) \tilde{P}_0^{-1} \tilde{P}_1 \quad (6.2.10)$$

From (6.2.7) and (6.2.8) we have  $\tilde{b}^1 = \tilde{b}^0$ . By expressing  $\tilde{b}^1$  and  $\tilde{b}^0$  in terms of  $b^1$  and  $b^0$ , we obtain

$$b^1 = \tilde{P}_1^{-1} \tilde{P}_0 b^0 \quad (6.2.11)$$

Equations (6.2.10) and (6.2.11) imply that the LSM  $M_1 = (A_1, b^1)$  is related to the LSM  $M_0 = (A_0, b^0)$  by the relations  $A_1 = \hat{P}(A_0 + b^0 \hat{v}^T) \hat{P}^{-1}$  and  $b^1 = \hat{P} b^0$ , where  $\hat{P} \equiv \tilde{P}_1^{-1} \tilde{P}_0$  and  $\hat{v}^T \equiv \tilde{v}^T \tilde{P}_0$ . Therefore,  $(A_1, b^1) \rho_1 (A_0, b^0)$ , and the proof is complete.  $\square$

Corollary 6.2.3. All state reachable single-input LSMs  $M = (A, b)$  of the same dimension belong to an equivalence class of LSMs with respect to the equivalence relation  $\rho_1'$  of Corollary 6.2.1.

Corollary 6.2.4. All state reachable single-input LSMs  $M = (A, b)$  of the same dimension belong to an equivalence class of LSMs with respect to the equivalence relation  $\rho_1''$  of Corollary 6.2.2.

The equivalence relation defined by equations (6.2.1) may be interpreted as a relation resulting from the sequential application of state feedback and nonsingular state transformation. In order to further explain this point of view, consider the state equation of the LSM  $M_0 = (A_0, b^0)$

$$x(k+1) = A_0 x(k) + b^0 u(k) \quad (6.2.12)$$

Introducing the state feedback law  $u(k) = v^{0T} x(k) + w(k)$ , where  $w(k)$  is a new external input and  $v^0 \in GF(q)^n$  is a vector, (6.2.12) becomes

$$x(k+1) = (A_0 + b^0 v^{0T})x(k) + b^0 w(k) \quad (6.2.13)$$

Now if we consider a state isomorphism  $P_0 : X \longrightarrow X$ ,  $x(k) \longmapsto P_0 x(k) \equiv \tilde{x}(k)$ , then the LSM (6.2.13) is transformed to the following isomorphic LSM:

$$\tilde{x}(k+1) = [P_0(A_0 + b^0 v^{0T})P_0^{-1}]\tilde{x}(k) + P_0 b^0 w(k)$$

In view of the invariance property of reachability under state feed-back homomorphism and state isomorphism, proved in Theorem 5.3.1 and Theorem 5.1.7, respectively, it follows that the LSM  $M_0 = (A_0, b^0)$  is reachable if and only if the LSM  $M_1 = (A_1, b^1) = (P_0(A_0 + b^0 v^{0T})P_0^{-1}, P_0 b^0)$  is reachable.

Theorem 6.2.2 can be used to state the criterion of reachability in alternative forms equivalent to the properties of Theorem 6.1.1. To accomplish this, it is sufficient to replace the LSM  $(A_0, b^0)$  by the LSM  $(A_1, b^1) = (P_0(A_0 + b^0 v^{0T})P_0^{-1}, P_0 b^0)$  in any of the properties of Theorem 6.1.1. For examples, property 2<sup>o</sup> of Theorem 6.1.1 can be equivalently restated as follows:

For any isomorphism  $P_0 : X \longrightarrow X$ ,  $P_0 \in GF(n, q)$ , and any vector  $v^0 \in GF(q)^n$ , the LSM  $(A, b)$  is reachable if and only if there does not exist any isomorphism  $P : X \longrightarrow X$ ,  $P \in GF(n, q)$ , such that the isomorphic LSM  $(\tilde{A}, \tilde{b}) \equiv (P[P_0(A + b v^{0T})P_0^{-1}], P(P_0 b))$  will have the form

$$\left( \begin{bmatrix} \tilde{A}_{11} & \tilde{A}_{12} \\ 0 & \tilde{A}_{22} \end{bmatrix}, \begin{bmatrix} \tilde{b}^1 \\ 0 \end{bmatrix} \right)$$

where  $\tilde{A}_{11} \in \text{GF}(q)^{r \times r}$ ,  $\tilde{A}_{12} \in \text{GF}(q)^{r \times (n-r)}$ ,  $\tilde{A}_{22} \in \text{GF}(q)^{(n-r) \times (n-r)}$ , and  $\tilde{b}^1 \in \text{GF}(q)^r$  with  $r < n$ .

### 6.3. Eighteen Equivalent Criteria for the State Reachability Property of Multi-Input LSMs

Theorem 6.3.1. For the multivariable LSM  $M = (A, B)$  the following statements are equivalent:

1<sup>o</sup>. The LSM  $M = (A, B)$  is state reachable.

2<sup>o</sup>. There does not exist an isomorphism  $P : X \rightarrow X$ ,  $P \in \text{GF}(n, q)$ , such that the characterizing matrices of the isomorphic machine  $\tilde{M} = (\tilde{A}, \tilde{B}) \equiv (PAP^{-1}, PB)$  will have the forms

$$\tilde{A} = \begin{bmatrix} \tilde{A}_{11} & \tilde{A}_{12} \\ 0 & \tilde{A}_{22} \end{bmatrix}; \quad \tilde{B} = \begin{bmatrix} \tilde{B}_1 \\ 0 \end{bmatrix} \quad (6.3.1)$$

where  $\tilde{A}_{11} \in \text{GF}(q)^{r \times r}$  ( $r < n$ , possibly  $r = 0$ ),  $\tilde{A}_{12} \in \text{GF}(q)^{r \times (n-r)}$ ,  $\tilde{A}_{22} \in \text{GF}(q)^{(n-r) \times (n-r)}$ , and  $\tilde{B}_1 \in \text{GF}(q)^{r \times m}$ .

Or

There does not exist any LSM  $\tilde{M} = (PAP^{-1}, PB)$ , isomorphic to  $M = (A, B)$ , whose state equations will have the following form:

$$\tilde{x}^I(k+1) = \tilde{A}_{11}\tilde{x}^I(k) + \tilde{A}_{12}\tilde{x}^{II}(k) + \tilde{B}_1 u(k) \quad (6.3.2a)$$

$$\tilde{x}^{II}(k+1) = \tilde{A}_{22}\tilde{x}^{II}(k) \quad (6.3.2b)$$

where  $\tilde{x}^I(k) \in GF(q)^r$  and  $\tilde{x}^{II}(k) \in GF(q)^{n-r}$ .

3°. There exists no  $A$ -invariant subspace of  $X$  of dimension smaller than  $n$ , containing  $R(B)$ .

4°. There exists no nonzero eigenvector  $v \in GF(q)^n$  of the matrix  $A^T$  which satisfies the relation  $v^T B = 0$ . That is, there exists no vector  $v \in GF(q)^n$  which will simultaneously satisfy the following relations:

$$v^T(\lambda I_n - A) = 0, \quad v^T B = 0, \quad v \neq 0 \quad (6.3.3)$$

5°. A subspace of  $X$  orthogonal to  $R(B)$  does not contain an  $A^T$ -invariant subspace.

6°. The following matrix  $K \in GF(q)^{n \times n}$  has rank  $n$ :

$$K \equiv [B, AB, A^2B, \dots, A^{n-1}B] \quad (6.3.4)$$

Or

The linear map

$$K \equiv [B, AB, A^2B, \dots, A^{n-1}B] : U^* \longrightarrow X$$

is an epimorphism.

7°. The rank of the matrix

$$KK^T = \sum_{j=0}^{n-1} A^{n-j-1} B B^T (A^T)^{n-j-1} \in GF(q)^{n \times n} \quad (6.3.5)$$

where  $K$  is defined by (6.3.4), is equal to  $n$ .

8°. The matrix equation

$$(A + BF)^n = 0 \quad (6.3.6)$$

has a unique solution with respect to  $F \in GF(q)^{m \times n}$ , and  $A$  is nonsingular.

9°. The set of matrix equations

$$FA^{-1}B = I_m \quad (6.3.7)$$

$$FA^{-i}B = 0, \quad i = 2, 3, \dots, n$$

has a unique solution with respect to  $F \in GF(q)^{m \times n}$ .

10°. The following matrix  $E \in GF(q)^{n^2 \times n(n+m-1)}$  has rank  $n^2$ :

$$E \equiv \begin{pmatrix} I_n & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & B \\ -A & I_n & \dots & 0 & 0 & 0 & \dots & 0 & B & 0 \\ 0 & -A & \dots & 0 & 0 & 0 & \dots & B & 0 & 0 \\ \cdot & \cdot & & \cdot & \cdot & \cdot & & \cdot & \cdot & \cdot \\ \cdot & \cdot & & \cdot & \cdot & \cdot & & \cdot & \cdot & \cdot \\ \cdot & \cdot & & \cdot & \cdot & \cdot & & \cdot & \cdot & \cdot \\ 0 & 0 & \dots & I_n & 0 & B & \dots & 0 & 0 & 0 \\ 0 & 0 & \dots & -A & B & 0 & \dots & 0 & 0 & 0 \end{pmatrix} \quad (6.3.8)$$

11°. Given any polynomial vector  $z(\xi) \in GF(q)[\xi]^{n \times 1}$  with elements of degree  $n-1$  or less, there exist a polynomial vector  $x(\xi) \in GF(q)[\xi]^{n \times 1}$  with elements of degree  $n-2$  or less, and a polynomial vector  $y(\xi) \in GF(q)[\xi]^{m \times 1}$  with elements of degree  $n-1$  or less, such that

$$(\xi I_n - A)x(\xi) + By(\xi) = z(\xi) \quad (6.3.9)$$



12<sup>0</sup>. There exist a polynomial matrix  $X(\xi) \in GF(q)[\xi]^{n \times n}$  with elements of degree  $n-2$  or less, and a polynomial matrix  $Y(\xi) \in GF(q)[\xi]^{m \times n}$  with elements of degree  $n-1$  or less, such that

$$(\xi I_n - A)X(\xi) + BY(\xi) = I_n \quad (6.3.10)$$

13<sup>0</sup>. The matrix  $[\xi I_n - A, B] \in GF(q)[\xi]^{n \times (n+m)}$  has rank  $n$  for all  $\xi$ .

14<sup>0</sup>. The polynomial matrices  $\xi I_n - A, B$  are coprime, that is,  $[\xi I_n - A, B]$  has the Smith canonical form  $[I_n, 0]$ .

15<sup>0</sup>. There exists an isomorphism  $P : X \rightarrow X$ ,  $P \in GF(n, q)$ , such that the characterizing matrices of the isomorphic LSM  $\tilde{M} = (\tilde{A}, \tilde{B}) \equiv (PAP^{-1}, PB)$  have the following forms:

$$\tilde{A} \equiv \begin{pmatrix} \tilde{A}_{11}^0 & \tilde{A}_{12} & \dots & \tilde{A}_{1\ell} \\ 0 & \tilde{A}_{22}^0 & \dots & \tilde{A}_{2\ell} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \tilde{A}_{\ell\ell}^0 \end{pmatrix}; \quad \tilde{B} \equiv \begin{pmatrix} \tilde{b}^{01} & 0 & 0 & \dots & 0 & \tilde{B}_1 \\ 0 & \tilde{b}^{02} & 0 & \dots & 0 & \tilde{B}_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \tilde{b}^{0\ell} & \tilde{B}_\ell \end{pmatrix} \quad (6.3.11)$$

$$\tilde{A}_{ii}^0 \equiv \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ a_{i1} & a_{i2} & a_{i3} & \dots & a_{in} \end{pmatrix}; \quad \tilde{b}^{0i} \equiv \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} \quad (6.3.12)$$

$i \in \underline{\ell}$ ,  $\ell \leq m$ ; in the matrices  $\tilde{A}_{it}$ ,  $i < t$ , only the entries in the first column may be nonzero, all other entries being zero;  $\tilde{A}_{ii}^0 \in GF(q)^{n_i \times n_i}$ ,  $n_1 + n_2 + \dots + n_\ell = n$ ,  $\tilde{A}_{it} \in GF(q)^{n_i \times n_t}$ ,  $\tilde{b}_{0i} \in GF(q)^{n_i}$  and  $\tilde{B}_i \in GF(q)^{n_i \times (m-t)}$ .

16<sup>0</sup>. Suppose that  $P : X \rightarrow X$  is an isomorphism such that the characterizing matrices of the isomorphic LSM  $\tilde{M} = (\tilde{A}, \tilde{B})$  have the following forms:

$$\tilde{A} \equiv \begin{pmatrix} \tilde{A}_1 & & & \\ & \tilde{A}_2 & & \\ & & \ddots & \\ & & & \tilde{A}_v \end{pmatrix}; \quad \tilde{B} \equiv \begin{pmatrix} \tilde{B}_1 \\ \tilde{B}_2 \\ \vdots \\ \tilde{B}_v \end{pmatrix} \quad (6.3.13)$$

where

$$\tilde{A}_i \equiv \begin{pmatrix} \lambda_i & 1 & & & \\ & \lambda_i & 1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ & & & & \lambda_i \end{pmatrix}, \quad i \in \underline{v} \quad (6.3.14)$$

that is,  $\tilde{A}$  is in the Jordan canonical form;  $\tilde{A}_i \in GF(q)^{n_i \times n_i}$ ,  $\tilde{B}_i \in GF(q)^{n_i \times m}$ , and  $n_1 + n_2 + \dots + n_v = n$ . Then all rows of the matrix  $\tilde{B}$  corresponding to the last rows of the Jordan blocks containing the same eigenvalue are linearly independent.

17<sup>0</sup>. For any given monic polynomial of degree  $n$  of the form

$$f(\lambda) = (\lambda)^n - \sum_{i=1}^n a_i(\lambda)^{i-1} \quad (6.3.15)$$

there exists a matrix  $F \in GF(q)^{m \times n}$  such that the characteristic polynomial of the matrix  $A + BF \in GF(q)^{n \times n}$  is equal to  $f(\lambda)$ , that is,

$$\det[\lambda I_n - (A + BF)] = f(\lambda) \quad (6.3.16)$$

18<sup>0</sup>. There does not exist a nonzero vector  $v \in GF(q)^n$  such that the expression

$$v^T(\lambda I_n - A)^{-1}B \quad (6.3.17)$$

is identically equal to zero. That is, the rows of the polynomial matrix  $(\lambda I_n - A)^{-1}B \in GF(q)[\lambda]^{n \times m}$  are linearly independent over  $GF(q)$ .

19<sup>0</sup>. For every set of distinct scalars  $\lambda_i$ ,  $i \in \underline{n}$ , different from the eigenvalues of the matrix  $A$ , the following matrix  $T \in GF(q)^{n \times mn}$  has rank  $n$ :

$$T \equiv [(\lambda_1 I_n - A)^{-1}B, (\lambda_2 I_n - A)^{-1}B, \dots, (\lambda_n I_n - A)^{-1}B] \quad (6.3.18)$$

Proof. We want to show that the following chain of sequential implications is closed:  $1^0 \Rightarrow 2^0 \Rightarrow 3^0 \Rightarrow 4^0 \Rightarrow 5^0 \Rightarrow 6^0 \Rightarrow 7^0 \Rightarrow 8^0 \Rightarrow 9^0 \Rightarrow 10^0 \Rightarrow 11^0 \Rightarrow 12^0 \Rightarrow 13^0 \Rightarrow 14^0 \Rightarrow 15^0 \Rightarrow 16^0 \Rightarrow 17^0 \Rightarrow 18^0 \Rightarrow 19^0 \Rightarrow 1^0$ .

The chain of implications  $1^0 \Rightarrow 2^0 \Rightarrow 3^0 \Rightarrow 4^0 \Rightarrow 5^0 \Rightarrow 6^0 \Rightarrow 7^0 \Rightarrow 8^0 \Rightarrow 9^0 \Rightarrow 10^0 \Rightarrow 11^0 \Rightarrow 12^0 \Rightarrow 13^0 \Rightarrow 14^0$  can be proved, with due consideration of the slight modifications resulting from the replacement of the input vector  $b$  by the input matrix  $B$ , in precisely the same manner as in the proof of the corresponding chain of implications in Theorem 6.1.1. To demonstrate this, we will provide the proof of the chain of implications  $1^0 \Rightarrow 2^0 \Rightarrow 3^0 \Rightarrow 4^0 \Rightarrow 5^0$  by imitating the proof of the corresponding chain of implications in Theorem 6.1.1, and incorporating the appropriate changes. Therefore, to avoid excessive repetition, the proof of the chain  $5^0 \Rightarrow 6^0 \Rightarrow 7^0 \Rightarrow 8^0 \Rightarrow 9^0 \Rightarrow 10^0 \Rightarrow 11^0 \Rightarrow 12^0 \Rightarrow 13^0 \Rightarrow 14^0$  will not be reproduced. Similar comments apply to the chain  $18^0 \Rightarrow 19^0 \Rightarrow 1^0$  which corresponds to the chain  $24^0 \Rightarrow 25^0 \Rightarrow 1^0$  in Theorem 6.1.1. However, properties  $15^0$  and  $16^0$  are considerably different in the multi-input case from the corresponding properties  $17^0$  and  $18^0$  in Theorem 6.1.1, and, therefore, we need to prove the chain of implications  $14^0 \Rightarrow 15^0 \Rightarrow 16^0 \Rightarrow 17^0 \Rightarrow 18^0$ .

$1^0 \Rightarrow 2^0$ : Suppose that property  $2^0$  is not satisfied. Then there exists an isomorphism  $P : X \rightarrow X$  such that the isomorphic LSM  $\tilde{M} = (\tilde{A}, \tilde{B}) \equiv (PAP^{-1}, PB)$  has the form (6.3.2). From equation (6.3.2b) it follows that  $\tilde{x}^{II}(k') = 0$ , for any clock period  $k'$ , implies the equality  $\tilde{x}^{II}(k'') = 0$  for any other clock period  $k''$ . Therefore,  $\tilde{M} = (\tilde{A}, \tilde{B})$  and consequently  $M = (A, B)$  cannot be reachable. Hence  $1^0 \Rightarrow 2^0$ .

$2^0 \not\Rightarrow 3^0$ : Suppose that property  $3^0$  is not satisfied. Then  $R(B)$  is contained in an  $A$ -invariant subspace  $S \subset X$  of dimension  $r < n$ . Let

$$\{s^1, s^2, \dots, s^r, s^{r+1}, \dots, s^n\} \quad (6.3.19)$$

be a basis for  $X$  such that  $\{s^1, s^2, \dots, s^r\}$  forms a basis for  $S$ . Then any vector  $x \in S$  can be uniquely expressed as

$$x = a_1 s^1 + a_2 s^2 + \dots + a_r s^r \quad (6.3.20)$$

for appropriate scalars  $a_i \in GF(q)$ . Since  $S \subset X$  is  $A$ -invariant, the vector  $Ax$  is in  $S$  and has the form

$$\begin{aligned} Ax &= b_1 As^1 + b_2 As^2 + \dots + b_r As^r \\ &= b_1 As^1 + b_2 As^2 + \dots + b_r As^r + 0s^{r+1} + \dots + 0s^n \end{aligned} \quad (6.3.21)$$

for appropriate  $b_i \in GF(q)$ . From (6.3.21) it follows that the matrix representations  $\tilde{A}$  and  $\tilde{B}$  of  $A$  and  $B$  with respect to the basis (6.3.19) have the forms given by (6.3.1), and hence property  $2^0$  is not satisfied. Therefore,  $2^0 \not\Rightarrow 3^0$ . Clearly  $P^{-1} = [s^1, s^2, \dots, s^n]$ .

$3^0 \Rightarrow 4^0$ : If property  $4^0$  is not satisfied, then there exists a nonzero vector  $v \in GF(q)^n$  such that

$$v^T A = \lambda v^T \quad (6.3.22)$$

$$v^T B = 0 \quad (6.3.23)$$

Suppose that  $W \subseteq X$  is orthogonal to  $v$  so that  $v^T w = 0$  for all  $w \in W$ . Now from equation (6.3.22) it follows that  $v^T A w = \lambda v^T w = 0$  and hence  $W$  is  $A$ -invariant. Since  $v \neq 0$ ,  $\dim W < n$ . Moreover, equation (6.3.23) shows that  $R(B) \subseteq W$ . Therefore, there exists an  $A$ -invariant subspace of dimension smaller than  $n$  containing  $R(B)$ . This conclusion contradicts property  $3^0$ , and hence  $3^0 \implies 4^0$ .

$4^0 \implies 5^0$ : Suppose that property  $5^0$  is not satisfied. Then there exists a subspace  $V \subseteq X$  such that  $v^T B = 0$  for all  $v \in V$ . Furthermore, there exists an  $A^T$ -invariant subspace  $W \subseteq V$ . Therefore, for any  $v \in W$ , we have  $A^T v \in W$ . Now we need to show that  $W$  contains an eigenvector of  $A^T$ . This can be accomplished in precisely the same way as in the proof of the implication  $4^0 \implies 5^0$  in Theorem 6.1.1.

$14^0 \implies 15^0$ : We prove this implication indirectly by considering the chain of implications  $14^0 \implies 4^0 \implies 5^0 \implies 6^0 \implies 15^0$  of which the implications  $14^0 \implies 4^0$  and  $6^0 \implies 15^0$  are not yet proved. The implication  $14^0 \implies 4^0$  follows easily from the fact that if the polynomial matrices  $\lambda I_n - A$  and  $B$  are coprime, then the matrix  $[\lambda I_n - A, B]$  has rank  $n$  for all  $\lambda$ , and hence no vector  $v \in GF(q)^n$  exists such that the relations (6.3.3) are simultaneously satisfied. To prove the implication  $6^0 \implies 15^0$ , let  $b^i$ ,  $i \in \underline{m}$ , denote the columns of the input matrix  $B$ . If property  $6^0$  is satisfied, then there exist integers  $0 = r_0 < r_1 < \dots < r_\ell = n$ ,  $\ell \leq m$ , and a set of linearly independent vectors  $v^i$ ,  $i \in \underline{\ell}$ , such that

$$\begin{aligned}
v^{r_i} &= b^i \\
v^{r_{i-1}} &= Ab^i - c_{in_i} b^i \\
v^{r_{i-2}} &= A^2 b^i - c_{in_i} Ab^i - c_{i,n_i-1} b^i \\
&\vdots
\end{aligned}
\tag{6.3.24}$$

$$v^{r_{i-(n_i-1)}} = A^{n_i} b^i - c_{in_i} A^{n_i-1} b^i - c_{i,n_i-1} A^{n_i-2} b^i - \dots - c_{i2} b^i;$$

$$n_i = r_i - r_{i-1}, \quad i \in \underline{\ell}$$

$$A^{n_i} b^i - c_{in_i} A^{n_i-1} b^i - c_{i,n_i-1} A^{n_i-2} b^i - \dots - c_{i2} b^i = \sum_{t=1}^{r_{i-1}} d_{it} v^t
\tag{6.2.25}$$

Assume that (6.3.24) and (6.3.25) are satisfied for any  $i < v-1$ ,  $v \in \underline{n+1}$ . We want to show that they also hold for  $i = v$ , and that the inequality  $r_{v-1} < n$  is also satisfied. Furthermore, we define the integer  $r_v > r_{v-1}$  and the vectors  $v^{r_{v-1}+1}$ ,  $v^{r_{v-1}+2}$ ,  $\dots$ ,  $v^{r_v}$ , so that (6.3.24) and (6.3.25) be satisfied for  $i = 1$ . Clearly we must have  $v-1 < m$ . From the original hypothesis, that is,  $\text{rank } [B, AB, A^2B, \dots, A^{n-1}B] = n$ , and the form of the relations (6.3.24) and (6.3.25), it follows that among the  $m - (v-1)$  columns of the  $B$  matrix there exists a column which contains at least one nonzero element, that is, this column cannot be expressed as a linear combination of the vectors  $v^t$ ,  $t = 1, 2, \dots, r_{v-1}$ . To see this, assume the contrary, that is,

that all columns of the matrix  $B$  are linear combinations of the previously selected vectors  $v^t$ ,  $t = 1, 2, \dots, r_{v-1}$ . Then from (6.3.24) and (6.3.25) it is clear that all columns of the matrices  $A^s B$ , where  $s$  is any positive integer, would also be expressed as linear combinations of the vectors  $v^t$ ,  $t = 1, 2, \dots, r_{v-1}$ . However, from the assumption  $r_{v-1} < n$ , it follows that  $\text{rank} [B, AB, A^2 B, \dots, A^{n-1} B] < n$  which obviously contradicts the original assumption. Therefore, by permuting, if necessary, the order of the last  $m - (v-1)$  columns of  $B$ , we can obtain a vector  $b^v$  which is not a linear combination of the vectors  $v^t$ ,  $t = 1, 2, \dots, r_{v-1}$ . In view of this observation, we can find an integer  $n_v$  such that the vectors

$$v^1, v^2, \dots, v^{r_{v-1}}, b^v, Ab^v, A^2 b^v, \dots, A^{n_v-1} b^v \quad (6.3.26)$$

are linearly independent and the vector  $A^{n_v} b^v$  is a linear combination of the vectors (6.3.26) of the form (6.3.25) for  $v = i$ . Now if we use the coefficients of this linear combination, and relations (6.3.24) to define the new set of vectors

$$v^{r_{v-1}+1}, v^{r_{v-1}+2}, \dots, v^{r_v} \quad (6.3.27)$$

where  $r_v = r_{v-1} + n_v$ , then it can be easily seen that these vectors are linearly independent. To see this, assume that the vectors (6.3.27) are linearly dependent. That is, a linear combination of these vectors is equal to zero and there exists at least one nonzero coefficient. Substituting for  $v^t$ ,  $t = r_{v-1}+1, r_{v-1}+2, \dots, r_v$ , in this linear combination their equivalent expressions from (6.3.24) for  $i = v$ , we will



obtain a linear combination of the vectors (6.3.26) which is equal to zero and has at least one nonzero coefficient, thus contradicting the independence of the vectors (6.3.26). Applying the preceding scheme iteratively, we can see that the number of the vectors  $v^t$  increases at each iteration by at least one and thus after a finite number of iterations we will obtain the greatest possible number of linearly independent vectors  $v^t$  after which the procedure can no longer be applied. Clearly in the final step we will have  $r_\ell = n$ . From (6.3.24) and (6.3.25) we obtain the following relations:

$$v^{r_i} = b^i \quad (6.3.28)$$

$$\begin{aligned} Av^{r_i} &= v^{r_i-1} + c_{in_i} v^{r_i} \\ Av^{r_i-1} &= v^{r_i-2} + c_{in_i} v^{r_i} \\ &\vdots \\ &\vdots \\ &\vdots \end{aligned} \quad (6.3.29)$$

$$Av^{r_i-(n_i-1)} = c_{i1} v^{r_i} + \sum_{t=1}^{r_i-1} d_{it} v^t, \quad i \in \underline{\ell}$$

From (6.3.28) and (6.3.29) it follows that the representations  $\tilde{A}$  and  $\tilde{B}$  of the matrices  $A$  and  $B$  with respect to the basis  $\{v^1, v^2, \dots, v^n\}$  have the form given by (6.3.11) and (6.3.12), and  $P^{-1} = [v^1, v^2, \dots, v^n]$ .

$15^0 \Rightarrow 16^0$ : We will prove the chain of implications  $15^0 \Rightarrow 4^0 \Rightarrow 16^0$ . To show that  $15^0 \Rightarrow 4^0$ , suppose that property  $4^0$  is not satisfied. Then there exists a vector  $v \in GF(q)^n$  such that

$$v^T \tilde{A} = \lambda v^T; \quad v^T \tilde{B} = 0; \quad v \neq 0 \quad (6.3.30)$$

If we rewrite  $v^T$  as  $(v^{1T}, v^{2T}, \dots, v^{\ell T})$ , where  $v^i$  has the same number of components as  $\tilde{b}^{0i}$ ,  $i \in \underline{\ell}$ , then in view of the special form of the vectors  $\tilde{b}^{0i}$ , defined by (6.3.11), it follows that the equality

$$v^T \tilde{B} = [v^{1T} \tilde{b}^{01}, v^{2T} \tilde{b}^{02}, \dots, v^{\ell T} \tilde{b}^{0\ell}, \sum_{i=1}^{\ell} v^{iT} \tilde{B}_i] = 0$$

will hold if and only if the last component of each of the subvectors  $v^i$  is equal to zero. Using this conclusion and observing the form of the matrices  $\tilde{A}_{ii}^0$  given by (6.3.11), it follows that the equality  $v^T \tilde{A} = \lambda v^T$  is possible only if  $v = 0$ , which contradicts the original hypothesis, and hence  $15^0 \Rightarrow 4^0$ . In order to prove the implication  $4^0 \Rightarrow 16^0$ , suppose that  $s$  Jordan blocks of  $\tilde{A}$ , given by (6.3.13), are associated with the same eigenvalue  $\lambda_0$ . Furthermore, let the last rows of the matrices  $\tilde{B}_1, \tilde{B}_2, \dots, \tilde{B}_s$  be denoted by  $\tilde{b}_\ell^{1T}, \tilde{b}_\ell^{2T}, \dots, \tilde{b}_\ell^{sT}$ . Now if property  $16^0$  is not satisfied, then  $\text{rank}[\tilde{b}_\ell^{1T}, \tilde{b}_\ell^{2T}, \dots, \tilde{b}_\ell^{sT}] < s$ , that is, there exist scalars  $c_i \in GF(q)$ ,  $i \in \underline{s}$ , not all zero, such that

$$\sum_{i=1}^s c_i \tilde{b}_\ell^{iT} = 0 \quad (6.3.31)$$

Recalling that the square matrices  $\tilde{A}_i$  have dimensions  $n_i$ ,  $i \in \underline{s}$ , it can be easily checked that the vector

$$v^T = [0, 0, \dots, c_1; 0, 0, \dots, c_2; \dots; 0, 0, \dots, c_s; 0, 0, \dots, 0]$$

$$\begin{matrix} n_1 & n_2 & n_s & s \\ & & & n - \sum_{i=1}^s n_i \end{matrix}$$

$$(6.3.32)$$

satisfies the equalities

$$v^T \tilde{A} = \lambda_0 v^T \quad (6.3.33)$$

and

$$v^T \tilde{B} = \sum_{i=1}^s c_i \tilde{b}_\ell^i \quad (6.3.34)$$

In view of (6.3.31) and (6.3.34), we obtain

$$v^T \tilde{B} = 0 \quad (6.3.35)$$

However, since  $v \neq 0$ , relations (6.3.33) and (6.3.35) show that property  $4^0$  is not satisfied. Therefore,  $4^0 \Rightarrow 16^0$ .

$16^0 \Rightarrow 17^0$ : We will prove the chain of implications  $16^0 \Rightarrow 4^0 \Rightarrow 15^0 \Rightarrow 17^0$  of which only the implications  $16^0 \Rightarrow 4^0$  and  $15^0 \Rightarrow 17^0$  are not yet proved. In order to prove that  $16^0 \Rightarrow 4^0$ , we will show that for any nonzero eigenvector  $v$  of the matrix  $\tilde{A}^T$  we have  $v^T \tilde{B} \neq 0$ . Assume, as in the proof of the implication  $15^0 \Rightarrow 16^0$ , that only the first  $s$  Jordan blocks of the matrix  $\tilde{A}$ , given by (6.3.13) are associated with the same eigenvalue  $\lambda_0$ . Then it follows that any eigenvector of the matrix  $\tilde{A}$  corresponding to the eigenvalue  $\lambda_0$  has the form (6.3.32), and hence  $v^T \tilde{B} = \sum_{i=1}^s c_i \tilde{b}_\ell^i$ , where  $\tilde{b}_\ell^i$  are the last rows

of the matrices  $\tilde{B}_i$ ,  $i \in \underline{s}$ . Now if property  $16^0$  is satisfied, then  $\text{rank} [\tilde{b}_\ell^{1T}, \tilde{b}_\ell^{2T}, \dots, \tilde{b}_\ell^{sT}] = s$  and consequently  $v^T \tilde{B} \neq 0$ . By appropriately permuting the order of the blocks, if necessary, it is seen that this conclusion is valid for any set of Jordan blocks associated with the same eigenvalue. Therefore, it follows that  $v^T \tilde{B} \neq 0$  for all eigenvectors  $v^T$  of the matrix  $\tilde{A}^T$ . In view of the relations  $\tilde{A} = PAP^{-1}$  and  $\tilde{B} = PB$ , it follows that  $v^T B \neq 0$  for any eigenvector of  $A^T$ . Hence  $16^0 \implies 4^0$ . It remains to be shown that  $15^0 \implies 17^0$ . Consider the matrix

$$\tilde{W} \equiv \begin{pmatrix} \tilde{w}^1 & 0 & 0 & \dots & 0 & 0 \\ 0 & \tilde{w}^2 & 0 & \dots & 0 & 0 \\ \cdot & \cdot & \cdot & & \cdot & \cdot \\ \cdot & \cdot & \cdot & & \cdot & \cdot \\ \cdot & \cdot & \cdot & & \cdot & \cdot \\ 0 & 0 & 0 & \dots & \tilde{w}^\ell & 0 \end{pmatrix} \quad (6.3.36)$$

where  $\tilde{w}^i \in GF(q)^{n_i}$ ,  $i \in \underline{\ell}$ ; and the last column of (6.3.36) whose elements are zero has the same dimensions as the last column of the matrix  $\tilde{B}$  given by (6.3.11). Since the matrix

$$\tilde{A} + \tilde{B}\tilde{W}^T = \begin{pmatrix} \tilde{A}_{11}^0 + \tilde{b}^{01} \tilde{w}^{1T} & \tilde{A}_{12} & \dots & \tilde{A}_{1\ell} \\ & \tilde{A}_{22}^0 + \tilde{b}^{02} \tilde{w}^{2T} & \dots & \tilde{A}_{2\ell} \\ & & \ddots & \vdots \\ & & & \tilde{A}_{\ell\ell}^0 + \tilde{b}^{0\ell} \tilde{w}^{\ell T} \end{pmatrix} \quad (6.3.37)$$

is block triangular, it can be easily shown by induction that its characteristic polynomial is equal to the product of the characteristic polynomials of the matrices  $\tilde{A}_{ii}^0 + \tilde{b}^0 \tilde{w}^i \tilde{w}^{iT}$ . By property 19<sup>0</sup> of Theorem 6.1.1, the vectors  $\tilde{w}^i$  can be chosen such that the characteristic polynomial of the matrix  $\tilde{A}_{ii}^0 + \tilde{b}^0 \tilde{w}^i \tilde{w}^{iT}$  be equal to any given monic polynomial of degree  $n_i$ ,  $i \in \underline{\ell}$ . Therefore, all the roots of the characteristic polynomial of the matrix (6.3.37) can be arbitrarily fixed. Using the relations  $\tilde{A} = PAP^{-1}$  and  $\tilde{B} = PB$ , we see that the characteristic polynomial of  $A + BF$  becomes equal to an arbitrary monic polynomial of degree  $n$  if we choose  $F = \tilde{W}^T P$  since  $\tilde{A} + \tilde{B} \tilde{W}^T = P(A + BF)P^{-1}$ . Hence  $15^0 \implies 17^0$ .

$17^0 \implies 18^0$ : We will prove the chain of implications  $17^0 \implies 4^0 \implies 15^0 \implies 18^0$  of which only the implications  $17^0 \implies 4^0$  and  $15^0 \implies 18^0$  are not yet proved. To prove the implication  $17^0 \implies 4^0$ , suppose property  $4^0$  is not satisfied. Then there exists a nonzero vector  $v \in GF(q)^n$  such that  $v^T(\lambda I_n - A) = 0$  and  $v^T B = 0$ , which imply that  $v^T(A + BF) = v^T A = \lambda v^T$  for any matrix  $F \in GF(q)^{m \times n}$ . Therefore, regardless of the properties of the matrix  $F$ , the scalar  $\lambda$  will be a root of the characteristic polynomial of the matrix  $A + BF$ , and hence property  $17^0$  cannot be satisfied. Thus  $17^0 \implies 4^0$ . To prove the implication  $15^0 \implies 18^0$ , suppose that property  $18^0$  is not satisfied so that there exists a vector  $v \in GF(q)^n$  such that

$$v^T(\lambda I_n - A)^{-1} B = 0 \quad (6.3.38)$$

However, using property 15<sup>o</sup>, we can bring the LSM (A, B) to the form  $(\tilde{A}, \tilde{B}) \equiv (PAP^{-1}, PB)$ , defined by (6.3.11) and (6.3.12). From (6.3.38) it follows that the vector

$$\tilde{v} = (P^{-1})^T v \neq 0 \quad (6.3.39)$$

satisfies the identity

$$\tilde{v}^T (\lambda I_n - \tilde{A})^{-1} \tilde{B} = 0 \quad (6.3.40)$$

Since the matrix  $\lambda I_n - \tilde{A}$  is block triangular, its inverse  $(\lambda I_n - \tilde{A})^{-1}$  is also block triangular having the matrices  $(\lambda I_{n_i} - \tilde{A}_{ii}^0)^{-1}$  as its diagonal elements. Writing the vector  $\tilde{v}^T$  as  $[\tilde{v}^{1T}, \tilde{v}^{2T}, \dots, \tilde{v}^{\ell T}]$ , where  $\tilde{v}^i$  has the same number of components as the vector  $b^{0i}$  given by (6.3.11), the matrix product in (6.3.40) becomes

$$\begin{bmatrix} \tilde{v}^{1T} & \tilde{v}^{2T} & \dots & \tilde{v}^{\ell T} \end{bmatrix} \begin{bmatrix} (\lambda I_{n_1} - \tilde{A}_{11}^0)^{-1} & R_{12} & R_{13} & \dots & R_{1\ell} \\ & (\lambda I_{n_2} - \tilde{A}_{22}^0)^{-1} & R_{23} & \dots & R_{2\ell} \\ & & \ddots & \ddots & \vdots \\ & & & (\lambda I_{n_\ell} - \tilde{A}_{\ell\ell}^0)^{-1} \end{bmatrix} \times$$

$$\begin{bmatrix} \tilde{b}^{01} & 0 & 0 & \dots & 0 & \tilde{B}_1 \\ & \tilde{b}^{02} & 0 & \dots & 0 & \tilde{B}_2 \\ & & \ddots & \ddots & \vdots & \vdots \\ & & & \tilde{b}^{0\ell} & \tilde{B}_\ell \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad (6.3.41)$$

where  $R_{ij}$  in  $(\lambda I_n - \tilde{A})^{-1}$  indicate the off-diagonal submatrices whose explicit forms are not needed for our purposes. Multiplying the matrices  $(\lambda I_n - \tilde{A})^{-1}$  and  $\tilde{B}$ , (6.3.41) reduces to

$$\begin{aligned}
 & \begin{bmatrix} \tilde{v}^{1T}, \tilde{v}^{2T}, \dots, \tilde{v}^{\ell T} \end{bmatrix} \begin{bmatrix} (\lambda I_{n_1} - \tilde{A}_{11}^0)^{-1} \tilde{b}^{01} & s_{12} & \dots & s_{1\ell} & s_{1,\ell+1} \\ & (\lambda I_{n_2} - \tilde{A}_{22}^0)^{-1} \tilde{b}^{02} & \dots & s_{2\ell} & s_{2,\ell+1} \\ & & \ddots & \vdots & \vdots \\ & & & (\lambda I_{n_\ell} - \tilde{A}_{\ell\ell}^0)^{-1} \tilde{b}^{0\ell} & s_{\ell,\ell+1} \end{bmatrix} \\
 & = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ \vdots \\ 0 \end{bmatrix} \quad (6.3.42)
 \end{aligned}$$

Again the explicit forms of  $s_{ij}$  are not needed for our purposes. From (6.3.42) we obtain the following set of identities:

$$\tilde{v}^{1T} (\lambda I_{n_1} - \tilde{A}_{11}^0)^{-1} \tilde{b}^{01} = 0 \quad (6.3.43)$$

$$\tilde{v}^{1T} s_{12} + \tilde{v}^{2T} (\lambda I_{n_2} - \tilde{A}_{22}^0)^{-1} \tilde{b}^{02} = 0$$

$$\tilde{v}^{1T} s_{13} + \tilde{v}^{2T} s_{23} + \tilde{v}^{3T} (\lambda I_{n_3} - \tilde{A}_{33}^0)^{-1} \tilde{b}^{03} = 0 \quad (6.3.44)$$

•  
•  
•

$$\tilde{v}^{1T} s_{1\ell} + \tilde{v}^{2T} s_{2\ell} + \dots + \tilde{v}^{(\ell-1)T} s_{\ell-1,\ell} + \tilde{v}^{\ell T} (\lambda I_{n_\ell} - \tilde{A}_{\ell\ell}^0)^{-1} \tilde{b}^{0\ell} = 0$$

Now examining the form of the submachine  $(\tilde{A}_{11}^0, \tilde{b}^{01})$  given by (6.3.12), we see that it is exactly the same as (6.1.14). This implies that the submachine  $(\tilde{A}_{11}^0, \tilde{b}^{01})$  satisfies property  $17^0$  and, therefore, also the equivalent property  $24^0$  of Theorem 6.1.1, which, in turn, implies that the subvector  $\tilde{v}^{1T}$  in (6.3.43) is equal to zero. This conclusion reduces the first equation of (6.3.44) to

$$\tilde{v}^{2T}(\lambda I_{n_2} - \tilde{A}_{22}^0)^{-1}\tilde{b}^{02} = 0 \quad (6.3.45)$$

Again since the submachine  $(\tilde{A}_{22}^0, \tilde{b}^{02})$  has the exact form (6.1.14), it follows that it satisfies property  $17^0$  and, therefore, the equivalent property  $24^0$  of Theorem 6.1.1, which implies that  $\tilde{v}^{2T} = 0$  in (6.3.45). Applying the above argument to the remaining equations of (6.3.44), we arrive at the conclusion that  $\tilde{v}^{1T} = \tilde{v}^{2T} = \dots = \tilde{v}^{\ell T} = 0$ , that is,  $\tilde{v} = 0$  which obviously contradicts (6.3.39). Hence  $15^0 \implies 18^0$ .

As pointed out earlier, the chain of implications  $18^0 \implies 19^0 \implies 1^0$  can be proved in precisely the same manner as in the proof of the corresponding chain of implications  $24^0 \implies 25^0 \implies 1^0$  in Theorem 6.1.1.  $\square$

#### 6.4. Equivalence Classes of State Reachable Multi-Input LSMs

As in the case of single-input LSMs, we will briefly discuss the equivalence relation  $G_1$ -equivalence in relation to the class of state reachable multi-input LSMs, where  $G_1$  is a transformation group defined by (5.1.35).



Consider the following relation  $\rho$  on the set of all LSMs  $(A, B)$  of the same dimension  $n$ :  $(A_1, B_1)\rho(A_0, B_0)$ , that is, the LSM  $(A_1, B_1)$  is related to the LSM  $(A_0, B_0)$  under the relation  $\rho$ , if and only if there exist isomorphisms  $P_0 : X \longrightarrow X$ ,  $G_0 : U \longrightarrow U$ ,  $P_0 \in \text{GF}(n, q)$ ,  $G_0 \in \text{GF}(m, q)$ , and a constant linear map  $F_0 : X \longrightarrow U$  such that

$$A_1 = P_0(A + BF_0)P_0^{-1}; \quad B_1 = P_0BG_0 \quad (6.4.1)$$

In the following theorem we will show that  $\rho$  is an equivalence relation and, therefore, allows the partition of the set of all state reachable LSMs  $(A, B)$  into equivalence classes.

Theorem 6.4.1.  $\rho$  is an equivalence relation.

Proof. We need to show that the relation  $\rho$  defined by (6.4.1) is reflexive, symmetric, and transitive. Letting  $P_0 = G_0 = I_n$  and  $F_0 = 0$ , it follows that  $(A_0, B_0)\rho(A_0, B_0)$  and hence  $\rho$  is reflexive. To show that  $\rho$  is symmetric, let  $A_0 = P_1(A_1 + B_1F_1)P_1^{-1}$  and  $B_0 = P_1B_1G_1$ , where  $P_1 \equiv P_0^{-1}$ ,  $F_1 \equiv G_0^{-1}F_0P_0^{-1}$ , and  $G_1 \equiv G_0^{-1}$ . Then it is straightforward to show that  $(A_1, B_1)\rho(A_0, B_0) \iff (A_0, B_0)\rho(A_1, B_1)$ . Finally, to show that  $\rho$  is transitive, that is,  $(A_1, B_1)\rho(A_0, B_0)$ ,  $(A_2, B_2)\rho(A_1, B_1) \implies (A_2, B_2)\rho(A_0, B_0)$ , suppose that  $(A_1, B_1)\rho(A_0, B_0)$  and  $(A_2, B_2)\rho(A_1, B_1)$ , that is,

$$A_1 = P_0(A_0 + B_0F_0)P_0^{-1}; \quad B_1 = P_0B_0G_0 \quad (6.4.1)$$

$$A_2 = P_1(A_1 + B_1F_1)P_1^{-1}; \quad B_2 = P_1B_1G_1 \quad (6.4.2)$$

Substituting for  $A_1$  and  $B_1$  from (6.4.1) into (6.4.2), we obtain

$$\begin{aligned}
 A_2 &= P_1 [P_0 (A_0 + B_0 F_0) P_0^{-1} + P_0 B_0 G_0 F_1] P_1^{-1}; \quad B_2 = P_1 P_0 B_0 G_0 G_1 \\
 &= P_1 P_0 (A_0 + B_0 F_0) P_0^{-1} P_1^{-1} + P_1 P_0 B_0 G_0 F_1 P_1^{-1} \\
 &= P_1 P_0 (A_0 + B_0 F_0) P_0^{-1} P_1^{-1} + P_1 P_0 B_0 G_0 F_1 P_0 P_0^{-1} P_1^{-1} \\
 &= (P_1 P_0) [(A_0 + B_0 F_0) + B_0 G_0 F_1 P_0] (P_1 P_0)^{-1} \\
 &= (P_1 P_0) [A_0 + B_0 (F_0 + G_0 F_1 P_0)] (P_1 P_0)^{-1}; \quad B_2 = (P_1 P_0) B_0 (G_0 G_1)
 \end{aligned}
 \tag{6.4.3}$$

From (6.4.3) it follows that the LSM  $(A_2, B_2)$  is related to the LSM  $(A_0, B_0)$  by the relations

$$A_2 = P_2 (A_0 + B_0 F_2) P_2^{-1}; \quad B_2 = P_2 B_0 G_2$$

where  $P_2 \equiv P_1 P_0$ ,  $G_2 \equiv G_0 G_1$ , and  $F_2 \equiv F_0 + G_0 F_1 P_0$ .  $\square$

Corollary 6.4.1. The relation  $(A_1, B_1) \rho'(A_0, B_0)$ , where  $A_1 = P_0 A P_0^{-1}$  and  $B_1 = P_0 B_0$ , is an equivalence relation.

Corollary 6.4.2. The relation  $(A_1, B_1) \rho''(A_0, B_0)$ , where  $A_1 = A_0 + B_0 F_0$  and  $B_1 = B_0$ , is an equivalence relation.

Corollary 6.4.3. The relation  $(A_1, B_1) \rho'''(A_0, B_0)$ , where  $A_1 = A_0$  and  $B_1 = B G_0$ , is an equivalence relation.

Theorem 6.4.2. All state reachable multi-input LSMs  $M = (A, B)$  of the same dimension  $n$  belong to an equivalence class of LSMs with respect to the relation defined by (6.4.1). That is, if  $M_0 = (A_0, B_0)$  is a reachable LSM, then all the LSMs  $M_1 = (A_1, B_1)$  obtained by the relations

$$A_1 = P_0(A_0 + B_0 F_0)P_0^{-1}; \quad B_1 = P_0 B_0 G_0 \quad (6.4.4)$$

are reachable; conversely, if  $M_0 = (A_0, B_0)$  and  $M_1 = (A_1, B_1)$  are any two reachable LSMs of the same dimension, then there exist isomorphisms  $P_0 : X \rightarrow X$ ,  $G_0 : U \rightarrow U$ ,  $P_0 \in GF(n, q)$ ,  $G_0 \in GF(m, q)$ , and a linear map  $F_0 : X \rightarrow U$  such that the LSMs are related by the relations (6.4.4).

Proof. Relations (6.4.4) can be viewed as relations resulting from the sequential application of state feedback homomorphism and state isomorphism since under the action of the feedback law  $u(k) = F_0 x(k) + G_0 w(k)$ , the LSM  $(A_0, B_0)$  is transformed to  $(A_0 + B_0 F_0, B_0 G_0)$ ; applying a state isomorphism  $P_0 : X \rightarrow X$ ,  $x(k) \mapsto P_0 x(k)$ , to the LSM  $(A_0 + B_0 F_0, B_0 G_0)$  transforms it to the isomorphic LSM  $(P_0(A_0 + B_0 F_0)P_0^{-1}, P_0 B_0 G_0)$ . Now the result follows immediately from Theorem 5.3.1 and Theorem 5.1.7.  $\square$

Corollary 6.4.4. All state reachable multivariable LSMs  $M = (A, B)$  of the same dimension  $n$  belong to an equivalence class of LSMs with respect to the relation  $\rho'$  of Corollary 6.4.1.

Corollary 6.4.5. All state reachable multivariable LSMs  $M = (A, B)$  of the same dimension  $n$  belong to an equivalence class of LSMs with respect to the relation  $\rho''$  of Corollary 6.4.2.

Corollary 6.4.6. All state reachable multivariable LSMs  $M = (A, B)$  of the same dimension  $n$  belong to an equivalence class of LSMs with respect to the relation  $\rho'''$  of Corollary 6.4.3.

From Theorem 6.4.2. it follows that we can state the criterion of state reachability for multivariable LSMs in alternative forms, equivalent to the properties of Theorem 6.3.1, by simply replacing the LSM  $(A_0, B_0)$  by the LSM  $(A_1, B_1) = (P_0(A + BF_0)P_0^{-1}, P_0BG_0)$  in any of the properties of Theorem 6.3.1. For example, property 2<sup>o</sup> of this theorem can be equivalently restated as follows:

For any isomorphisms  $P_0 : X \longrightarrow X$ ,  $G_0 : U \longrightarrow U$ ,  $P_0 \in \text{GF}(n, q)$ ,  $G_0 \in \text{GF}(m, q)$ , and any linear map  $F_0 : X \longrightarrow U$ , the LSM  $(A, B)$  is state reachable if and only if there does not exist any isomorphism  $P : X \longrightarrow X$ ,  $P \in \text{GF}(n, q)$ , such that the isomorphic LSM  $(\tilde{A}, \tilde{B}) \equiv (P[P_0(A + BF_0)P_0^{-1}]P^{-1}, P(P_0BG_0))$  will have the form

$$\left( \begin{bmatrix} \tilde{A}_{11} & \tilde{A}_{12} \\ 0 & \tilde{A}_{22} \end{bmatrix}, \begin{bmatrix} \tilde{B}_1 \\ 0 \end{bmatrix} \right)$$

where  $\tilde{A}_{11} \in \text{GF}(q)^{r \times r}$  ( $r < n$ , possibly  $r = 0$ ),  $\tilde{A}_{12} \in \text{GF}(q)^{r \times (n-r)}$ ,  $\tilde{A}_{22} \in \text{GF}(q)^{(n-r) \times (n-r)}$ , and  $\tilde{B}_1 \in \text{GF}(q)^{r \times m}$ .

### Summary and Conclusions

This chapter was devoted to a reexamination of state reachability of LSMs. Particular emphasis was directed towards reformulating this property in various other forms. Consequently, twenty-four equivalent criteria for the state reachability property of single-input LSMs

and eighteen equivalent criteria for the state reachability property of multi-input LSMs were stated, and in each case, equivalence of the stated criteria was proved. Furthermore, equivalence classes of state reachable single- and multi-input LSMs under certain transformation groups were identified (cf. [1], [2], [3], [15], [17], [18], [21], [37], [58], [61], [88], [91], [97], [99], [105], [107], [109], [110]).

## CHAPTER VII

THE JORDAN CANONICAL FORM AND SELECTIVE  
STATE REACHABILITY OF LSMS

In this chapter attention will be focused on the eigenstructures, that is, eigenvalues, (generalized) eigenvectors, and (generalized) eigenspaces associated with the LSM  $M = (A, B)$ , in the framework of the Jordan canonical form. Using the eigenproperties of  $M$ , some additional state reachability criteria which explicitly involve the Jordan canonical representation of  $M$  will be given. Furthermore, we will introduce and develop in detail the concept of selective state reachability for LSMs, making heavy use of the eigenstructures of  $M$ .

7.1. The Jordan Canonical Form for LSMs

This is one of the most well known canonical forms that can be rigorously discussed in terms of the theory of cyclic decomposition of  $X$ . However, since we will make heavy use of the Jordan canonical form in our development of the concept of selective state reachability which is primarily based on the eigenvalue-eigenvector structure of the characteristic matrix  $A$  of an LSM  $(A, B)$ , we will use the notion of generalized eigenvectors rather than a thorough and highly algebraic viewpoint, to briefly describe this canonical form. In order to motivate the need for generalized eigenvectors, initially we consider the case where  $A \in GF(q)^{n \times n}$  has  $n$  distinct eigenvalues, denoted by

$\lambda_1, \lambda_2, \dots, \lambda_n$ . Let  $v^1, v^2, \dots, v^n$  denote the corresponding eigenvectors, that is,  $Av^i = \lambda_i v^i, i \in \underline{n}$ . Since the set of eigenvectors associated with distinct eigenvalues is linearly independent, the set  $\{v^i : i \in \underline{n}\}$  forms a basis for  $X$  with respect to which  $A$  has the following simple representation:

$$\tilde{A} = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix} \quad (7.1.1)$$

Since

$$\begin{aligned} AV &\equiv A[v^1, v^2, \dots, v^n] = [Av^1, Av^2, \dots, Av^n] \\ &= [\lambda_1 v^1, \lambda_2 v^2, \dots, \lambda_n v^n] \equiv V\tilde{A} \end{aligned}$$

we obtain  $\tilde{A} = V^{-1}AV$ . Therefore, if  $A$  has  $n$  distinct eigenvalues, there exists an isomorphism  $V : X \longrightarrow X, V \in \text{GF}(n, q)$ , such that  $\tilde{A} \equiv V^{-1}AV$  has the diagonal form (7.1.1). However, if the eigenvalues of  $A$  are not distinct, then the set of eigenvectors associated with the subset of distinct eigenvalues will not be sufficient in number to constitute a basis for  $X$ . In order to extend this incomplete set of vectors to a basis with respect to which  $A$  will have a new representation, we have to resort to generalized eigenvectors.

Definition 7.1.1. A vector  $v \in GF(q)^n$  is said to be a *generalized eigenvector* or *supereigenvector* of index  $\ell$  (an integer  $\geq 1$ ), associated with the eigenvalue  $\lambda$  of  $A$ , if and only if  $(\lambda I_n - A)^r v = 0$  for all  $r \geq \ell$ , and  $(\lambda I_n - A)^{\ell-1} v \neq 0$ .

Clearly for  $\ell=1$ , the above definition reduces to the definition of an ordinary eigenvector.

Let  $v$  be a generalized eigenvector of  $A$  associated with the eigenvalue  $\lambda$ , and define

$$\begin{aligned} v^\ell &\equiv v \\ v^{\ell-1} &\equiv (\lambda I_n - A)v = (\lambda I_n - A)v^\ell \\ v^{\ell-2} &\equiv (\lambda I_n - A)^2 v = (\lambda I_n - A)v^{\ell-1} \\ &\vdots \\ v^1 &\equiv (\lambda I_n - A)^{\ell-1} v = (\lambda I_n - A)v^2 \end{aligned} \tag{7.1.2}$$

It can be easily checked that for each  $i \in \underline{\ell}$ ,  $v^i$  is a generalized eigenvector of index  $i$ .

Theorem 7.1.1. The generalized eigenvectors defined by (7.1.2) are linearly independent.

Proof. Suppose that  $v^1, v^2, \dots, v^\ell$  are linearly dependent. Then there exist  $a_i \in GF(q)$ ,  $i \in \underline{\ell}$ , not all zero, such that

$$\sum_{i=1}^{\ell} a_i v^i = 0 \tag{7.1.3}$$



Premultiplying (7.1.3) by  $(\lambda I_n - A)^{\ell-1}$ , and substituting for  $v^i$ ,  $i \in \underline{\ell}$ , from (7.1.2), we obtain

$$\sum_{i=1}^{\ell} a_i (\lambda I_n - A)^{\ell-1} (\lambda I_n - A)^{\ell-i} v = \sum_{i=1}^{\ell} a_i (\lambda I_n - A)^{2\ell-(i+1)} v = 0 \quad (7.1.4)$$

But  $(\lambda I_n - A)^{2\ell-(i+1)} v = 0$  for  $i \in \underline{\ell-1}$ . Therefore, (7.1.4) reduces to

$$a_{\ell} (\lambda I_n - A)^{\ell-1} v = 0$$

However, by the definition of a generalized eigenvector of index  $\ell$ ,  $a_{\ell} (\lambda I_n - A)^{\ell-1} v \neq 0$ , hence  $a_{\ell} = 0$ . If we multiply (7.1.3) by  $(\lambda I_n - A)^{\ell-j}$ ,  $j = 2, 3, \dots, \ell$ , and repeat the above sequence of steps, we will arrive at the conclusion that  $a_i = 0$  for  $i \in \underline{\ell}$ , which is obviously a contradiction.  $\square$

If we consider the generalized eigenvectors defined by (7.1.2) in the order  $v^1, v^2, \dots, v^{\ell}$ , then we can express the vectors  $Av^i$ ,  $i \in \underline{\ell}$ , as follows:

$$Av^1 = \lambda v^1 = [v^1, v^2, \dots, v^\ell] \begin{pmatrix} \lambda \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{pmatrix}$$

$$Av^2 = v^1 + \lambda v^2 = [v^1, v^2, \dots, v^\ell] \begin{pmatrix} 1 \\ \lambda \\ 0 \\ \cdot \\ \cdot \\ 0 \end{pmatrix}$$

$$Av^3 = v^2 + \lambda v^3 = [v^1, v^2, \dots, v^\ell] \begin{pmatrix} 0 \\ 1 \\ \lambda \\ 0 \\ \cdot \\ \cdot \\ 0 \end{pmatrix}$$

⋮

$$Av^\ell = v^{\ell-1} + \lambda v^\ell = [v^1, v^2, \dots, v^\ell] \begin{pmatrix} 0 \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \\ 1 \\ \lambda \end{pmatrix}$$

Therefore, the Jordan block  $\tilde{A}_1$  of  $A$ , associated with the eigenvalue  $\lambda$ , has the following canonical representation with respect to the new partial basis  $\{v^i, i \in \underline{\ell}\}$ :

$$\tilde{A}_1 = \begin{pmatrix} \lambda & 1 & & & & \\ & \lambda & 1 & & & \\ & & \lambda & 1 & & \\ & & & \ddots & \ddots & \\ & & & & \ddots & \\ & & & & & \lambda & 1 \\ & & & & & & \lambda \end{pmatrix} \quad (7.1.5)$$

It is clear that different arrangements of  $v^i, i \in \underline{\ell}$ , result into different forms for the block  $\tilde{A}_1$ .

**Theorem 7.1.2.** The generalized eigenvectors of  $A \in GF(q)^{n \times n}$ , associated with different eigenvalues, are linearly independent.

**Proof.** Suppose that  $\lambda_1$  and  $\lambda_2, \lambda_1 \neq \lambda_2$ , are two eigenvalues of  $A$ . Let  $v$  and  $w$  be generalized eigenvectors of indices  $\ell$  and  $m$  associated with  $\lambda_1$  and  $\lambda_2$ , respectively. If we define  $v^\ell \equiv v, v^i \equiv (\lambda_1 I_n - A)v^{i+1} = (\lambda_1 I_n - A)^{\ell-i}v, i \in \underline{\ell-1}$ ;  $w^m \equiv w, w^j \equiv (\lambda_2 I_n - A)w^{j+1} = (\lambda_2 I_n - A)^{m-j}w, j \in \underline{m-1}$ , then from Theorem 7.1.1 it follows that the sets  $\{v^i, i \in \underline{\ell}\}$  and  $\{w^j, j \in \underline{m}\}$  are separately linearly independent. We want to show that  $\{v^i, i \in \underline{\ell}\} \cup \{w^j, j \in \underline{m}\}$  is linearly independent. Suppose that  $\{v^i, i \in \underline{\ell}\}$  is linearly dependent on  $\{w^j, j \in \underline{m}\}$ . Then there exist  $a_{ij} \in GF(q), i \in \underline{\ell}, j \in \underline{m}$ , not all zero, such that

$$v^i = \sum_{j=1}^m a_{ij} w^j, i \in \underline{\ell} \quad (7.1.6)$$

Premultiplying (7.1.6) by  $(\lambda_1 I_n - A)^i$ , we obtain

$$\begin{aligned} (\lambda_1 I_n - A)^i v^i &= (\lambda_1 I_n - A)^i (\lambda_1 I_n - A)^{\ell-i} v = 0 \\ &= (\lambda_1 I_n - A)^i \sum_{j=1}^m a_{ij} w^j \end{aligned} \quad (7.1.7)$$

Premultiplying (7.1.6) by  $(\lambda_2 I_n - A)^{m-1}$ , we get

$$0 = \sum_{j=1}^m a_{ij} (\lambda_2 I_n - A)^{m-1} (\lambda_1 I_n - A)^i w^j$$

Since functions of the same matrix always commute, we have

$$\begin{aligned} 0 &= \sum_{j=1}^m a_{ij} (\lambda_1 I_n - A)^i (\lambda_2 I_n - A)^{m-1} w^j \\ &= a_{im} (\lambda_1 I_n - A)^i (\lambda_2 I_n - A)^{m-1} w^m \\ &= a_{im} (\lambda_1 I_n - A)^i w^1 \end{aligned} \quad (7.1.8)$$

because  $(\lambda_2 I_n - A)^{m-1} w^j = (\lambda_2 I_n - A)^{2m-(j+1)} = 0$ ,  $j \in \underline{m-1}$ . Equation (7.1.8) and  $(\lambda_2 I_n - A)w^1 = 0$  yield the following relation:

$$a_{im} (\lambda_2 - \lambda_1)^1 w^1 = 0$$

which, in view of the assumption that  $\lambda_1 \neq \lambda_2$ , implies that  $a_{im} = 0$ ,  $i \in \underline{\ell}$ . Repeating the above procedure, we will find that  $a_{ij} = 0$ ,  $i \in \underline{\ell}$ ,  $j \in \underline{m}$ , which clearly contradicts the hypothesis that not all  $a_{ij} = 0$ , and hence  $\{v^i, i \in \underline{\ell}\}$  is linearly independent of  $\{w^j, j \in \underline{m}\}$ . In a similar manner, if we assume that the set  $\{w^j, j \in \underline{m}\}$  is linearly

dependent on  $\{v^i, i \in \underline{\ell}\}$ . we will arrive at a contradiction. Therefore,  $\{v^1, v^2, \dots, v^\ell, w^1, w^2, \dots, w^m\}$  is a linearly independent set.  $\square$

Theorem 7.1.3. Let  $v$  and  $w$  be generalized eigenvectors of indices  $\ell$  and  $m$ , respectively, associated with the same eigenvalue  $\lambda$  of  $A \in GF(q)^{n \times n}$ . Define  $v^i \equiv (\lambda I_n - A)^{\ell-i} v$ ,  $i \in \underline{\ell}$ , and  $w^j \equiv (\lambda I_n - A)^{m-j} w$ ,  $j \in \underline{m}$ . If  $v^1$  and  $w^1$  are linearly independent, then the set  $\{v^i, i \in \underline{\ell}\} \cup \{w^j, j \in \underline{m}\}$  is linearly independent.

Proof. From Theorem 7.1.1 it follows that the sets  $\{v^i, i \in \underline{\ell}\}$  and  $\{w^j, j \in \underline{m}\}$  are separately linearly independent. Therefore, we need only show that these sets are linearly independent of each other. This can be easily accomplished by an argument similar to that used in the proof of Theorem 7.1.2, and hence will not be repeated here.  $\square$

The preceding results can be incorporated in an algorithm for generating the Jordan canonical form of any matrix  $A \in GF(q)^{n \times n}$ .

For the purpose of future reference, we will summarize the above results in the most general and detailed form in terms of LSMs.

Theorem 7.1.4. For any arbitrary LSM  $M = (A, B)$  there exists an isomorphism  $V : X \rightarrow X$ ,  $V \in GF(n, q)$ , such that the isomorphic LSM  $\tilde{M} = (\tilde{A}, \tilde{B}) \equiv (V^{-1}AV, V^{-1}B)$  has the following form:

$$\begin{pmatrix} \tilde{x}^1(k+1) \\ \tilde{x}^2(k+1) \\ \vdots \\ \tilde{x}^v(k+1) \end{pmatrix} = \begin{pmatrix} \tilde{A}_1(\lambda_1) & & & \\ & \tilde{A}_2(\lambda_2) & & \\ & & \ddots & \\ & & & \tilde{A}_v(\lambda_v) \end{pmatrix} \begin{pmatrix} \tilde{x}^1(k) \\ \tilde{x}^2(k) \\ \vdots \\ \tilde{x}^v(k) \end{pmatrix} + \begin{pmatrix} \tilde{B}^1 \\ \tilde{B}^2 \\ \vdots \\ \tilde{B}^v \end{pmatrix} \begin{pmatrix} u_1(k) \\ u_2(k) \\ \vdots \\ u_m(k) \end{pmatrix} \quad (7.1.9)$$

where

$$\tilde{A}_i(\lambda_i) \equiv \begin{bmatrix} \tilde{A}_{i1}(\lambda_i) & & & \\ & \tilde{A}_{i2}(\lambda_i) & & \\ & & \ddots & \\ & & & \tilde{A}_{i\mu(i)}(\lambda_i) \end{bmatrix}, \quad \tilde{B}^i \equiv \begin{bmatrix} \tilde{B}_1^i \\ \tilde{B}_2^i \\ \vdots \\ \tilde{B}_v^i \end{bmatrix} \quad (7.1.10)$$

$$\tilde{A}_i(\lambda_i) \in GF(q)^{n_i \times n_i}, \quad \tilde{B}^i \in GF(q)^{n_i \times m}; \quad i \in \underline{v}$$

$$\tilde{A}_{ij}(\lambda_i) \equiv \begin{bmatrix} \lambda_i & 1 & & & \\ & \lambda_i & 1 & & \\ & & \ddots & \ddots & \\ & & & \lambda_i & 1 \\ & & & & \lambda_i \end{bmatrix}; \quad \tilde{B}_j^i \equiv \begin{bmatrix} b_{1j}^{(i)} \\ b_{2j}^{(i)} \\ \vdots \\ b_{j_{ij},j}^{(i)} \end{bmatrix} \quad (7.1.11)$$

$$\tilde{A}_{ij}(\lambda_i) \in GF(q)^{n_{ij} \times n_{ij}}, \quad \tilde{B}_j^i \in GF(q)^{n_{ij} \times m}, \quad j \in \underline{\mu(i)}, \quad i \in \underline{v}$$

$$\tilde{B}_j^i \equiv \begin{pmatrix} \tilde{b}_{1j1}^{(i)} & \tilde{b}_{1j2}^{(i)} & \dots & \tilde{b}_{1jm}^{(i)} \\ \tilde{b}_{2j1}^{(i)} & \tilde{b}_{2j2}^{(i)} & \dots & \tilde{b}_{2jm}^{(i)} \\ \vdots & \vdots & & \vdots \\ \tilde{b}_{n_{ij},j1}^{(i)} & \tilde{b}_{n_{ij},j2}^{(i)} & \dots & \tilde{b}_{n_{ij},jm}^{(i)} \end{pmatrix} \quad (7.1.12)$$

$$n = \sum_{i=1}^v n_i = \sum_{i=1}^v \sum_{j=1}^{\mu(i)} n_{ij}$$

Sometimes we will use the direct sum notation and write  $\tilde{A}$  as

$$\begin{aligned} \tilde{A} \equiv & \tilde{A}(n_{11}, \lambda_1) \oplus \tilde{A}(n_{12}, \lambda_1) \oplus \dots \oplus \tilde{A}(n_{1\mu(1)}, \lambda_1) \\ & \oplus \tilde{A}(n_{21}, \lambda_2) \oplus \tilde{A}(n_{22}, \lambda_2) \oplus \dots \oplus \tilde{A}(n_{2\mu(2)}, \lambda_2) \\ & \oplus \dots \oplus \tilde{A}(n_{v1}, \lambda_v) \oplus \tilde{A}(n_{v2}, \lambda_v) \\ & \oplus \dots \oplus \tilde{A}(n_{v\mu(v)}, \lambda_v) \end{aligned}$$

Therefore, in terms of the above notation, the original LSM is isomorphic to the following  $n$  coupled linear submachines:

$$\begin{aligned} \tilde{x}_{\ell s}^{(i)}(k+1) &= \lambda_i \tilde{x}_{\ell s}^{(i)}(k) + \tilde{x}_{\ell+1,s}^{(i)}(k) + \sum_{t=1}^m \tilde{b}_{\ell st}^{(i)} u_t(k) \\ \tilde{x}_{n_{is},s}^{(i)}(k+1) &= \lambda_i \tilde{x}_{n_{is},s}^{(i)}(k) + \sum_{t=1}^m \tilde{b}_{n_{is},st}^{(i)} u_t(k) \end{aligned} \quad (7.1.13)$$

$$\ell \in \underline{n_{i\ell}}, s \in \underline{\mu(i)}, i \in \underline{v}$$

In order to see how the state equations (7.1.13) are identified with respect to the block and sub-block structure of the Jordan canonical form (7.1.9), we will write out the  $\sum_{t=1}^{\mu(1)} n_{1t}$  state equations corresponding to the Jordan block  $\tilde{A}(n_{1t}, \lambda_t)$  as follows:

$$\tilde{A}(n_{11}, \lambda_1) : \tilde{x}_{\ell 1}^{(1)}(k+1) = \lambda_1 \tilde{x}_{\ell 1}^{(1)}(k) + \tilde{x}_{\ell+1,1}^{(1)}(k) + \sum_{t=1}^m \tilde{b}_{\ell 1 t}^{(1)} u_t(k)$$

$$\ell \in \underline{n_{11}-1}$$

$$\tilde{x}_{n_{11},1}^{(1)}(k+1) = \lambda_1 \tilde{x}_{n_{11},1}^{(1)}(k) + \sum_{t=1}^m \tilde{b}_{n_{11},1 t}^{(1)} u_t(k)$$

$$\tilde{A}(n_{12}, \lambda_1) : \tilde{x}_{\ell 2}^{(1)}(k+1) = \lambda_1 \tilde{x}_{\ell 2}^{(1)}(k) + \tilde{x}_{\ell+1,2}^{(1)}(k) + \sum_{t=1}^m \tilde{b}_{\ell 2 t}^{(1)} u_t(k)$$

$$\ell \in \underline{n_{12}-1}$$

$$\tilde{x}_{n_{12},2}^{(1)}(k+1) = \lambda_1 \tilde{x}_{n_{12},2}^{(1)}(k) + \sum_{t=1}^m \tilde{b}_{n_{12},2 t}^{(1)} u_t(k)$$

⋮

$$\tilde{A}(n_{1\mu(1)}, \lambda_1) : \tilde{x}_{\ell, \mu(1)}^{(1)}(k+1) = \lambda_1 \tilde{x}_{\ell, \mu(1)}^{(1)}(k) + \tilde{x}_{\ell+1, \mu(1)}^{(1)}(k) + \sum_{t=1}^m \tilde{b}_{\ell \mu(1) t}^{(1)} u_t(k)$$

$$\ell \in \underline{n_{1\mu(1)}-1}$$

$$\tilde{x}_{n_{1\mu(1)}, \mu(1)}^{(1)}(k+1) = \lambda_1 \tilde{x}_{n_{1\mu(1)}, \mu(1)}^{(1)}(k) + \sum_{t=1}^m \tilde{b}_{n_{1\mu(1)}, \mu(1) t}^{(1)} u_t(k)$$

(7.1.14)



Similar scalar state equations can be explicitly written for the remaining Jordan blocks  $\tilde{A}(n_{ij}, \lambda_i)$ ,  $i = 2, 3, \dots, v$ ;  $j = 2, 3, \dots, \mu(i)$ .

Now we would like to present some further state reachability criteria that explicitly involve the Jordan canonical form of LSMs.

Theorem 7.1.5. The LSM  $M = (A, B)$  is state reachable if and only if all rows of the matrix  $\tilde{B} \equiv V^{-1}B$  corresponding to the last rows of the Jordan blocks associated with the same eigenvalue are linearly independent. That is,

$$\text{rank} \begin{pmatrix} \tilde{b}_{n_{i1},11}^{(i)} & \tilde{b}_{n_{i1},12}^{(i)} & \dots & \tilde{b}_{n_{i1},1m}^{(i)} \\ \tilde{b}_{n_{i2},21}^{(i)} & \tilde{b}_{n_{i2},22}^{(i)} & \dots & \tilde{b}_{n_{i2},2m}^{(i)} \\ \vdots & \vdots & & \vdots \\ \tilde{b}_{n_{i\mu(i),\mu(i)1}}^{(i)} & \tilde{b}_{n_{i\mu(i),\mu(i)2}}^{(i)} & \dots & \tilde{b}_{n_{i\mu(i),\mu(i)m}}^{(i)} \end{pmatrix} = \mu(i), \quad i \in \underline{v} \quad (7.1.15)$$

Proof. This theorem was proved as part of Theorem 6.3.1. Here we will present a completely different proof (cf. [27]). In Theorem 6.3.1 it was shown that the LSM  $(A, B)$  is state reachable if and only if  $\text{rank}[\lambda I_n - A, B] = n$  for all  $\lambda \in \text{GF}(q)$ . If  $V$  is the matrix of the isomorphism yielding the Jordan canonical form, that is, the matrix of generalized eigenvectors of  $A$ , then  $\tilde{A} = V^{-1}AV$  or  $V\tilde{A} = AV$ . Let

$$R(\lambda) \equiv [\lambda I_n - A, B] = V[(\lambda I_n - \tilde{A})V^{-1}, V^{-1}B]$$

and

$$\Delta(n_{ij}, \lambda_i, \lambda) \equiv [\lambda I_n - \tilde{A}(n_{ij}, \lambda_i)]$$

$$= \begin{pmatrix} \lambda - \lambda_i & -1 & & & \\ & & & & \\ & & \lambda - \lambda_i & -1 & \\ & & & \cdot & \\ & & & \cdot & \\ & & & & \cdot \\ & & & & \lambda - \lambda_i & -1 \\ & & & & & \\ & & & & & \lambda - \lambda_i \end{pmatrix}$$

Therefore,

$$\lambda I_n - \tilde{A} = \Delta(n_{11}, \lambda_1, \lambda) \oplus \Delta(n_{12}, \lambda_1, \lambda) \oplus \dots$$

$$\oplus \Delta(n_{1\mu(1)}, \lambda_1, \lambda) \oplus \Delta(n_{21}, \lambda_2, \lambda) \oplus \Delta(n_{22}, \lambda_2, \lambda)$$

$$\oplus \dots \oplus \Delta(n_{2\mu(2)}, \lambda_2, \lambda) \oplus \dots \oplus \Delta(n_{v1}, \lambda_v, \lambda)$$

$$\oplus \Delta(n_{v2}, \lambda_v, \lambda) \oplus \dots \oplus \Delta(n_{v\mu(v)}, \lambda_v, \lambda)$$

$$\text{rank } \Delta(n_{ij}, \lambda_i, \lambda) = n_{ij}, \lambda \neq \lambda_i, i \in \underline{v}, j \in \underline{\mu(i)}$$

$$\text{rank } \Delta(n_{ij}, \lambda_i, \lambda_i) = n_{ij} - 1 = \ell_{ii}, i \in \underline{v}, j \in \underline{\mu(i)}$$

$$\text{rank } \Delta(n_{ij}, \lambda_i, \lambda_s) = n_{ij} = \ell_{is}, i \neq s; i, s \in \underline{v}, j \in \underline{\mu(i)}$$

$\Rightarrow$

$$\begin{aligned}
 \text{rank}[\lambda I_n - \tilde{A}] &= \sum_{i=1}^v \sum_{j=1}^{\mu(i)} \text{rank } \Delta(n_{ij}, \lambda_i, \lambda) \\
 &= \sum_{i=1}^v \sum_{j=1}^{\mu(i)} n_{ij}, \lambda \neq \lambda_i, i \in \underline{v}. \\
 \text{rank}[\lambda_s I_n - \tilde{A}] &= \sum_{i=1}^v \sum_{j=1}^{\mu(i)} \text{rank } \Delta(n_{ij}, \lambda_i, \lambda_s) \\
 &= \sum_{\substack{i=1 \\ i \neq s}}^v \sum_{j=1}^{\mu(i)} n_{ij} + \sum_{j=1}^{\mu(s)} (n_{sj} - 1) \\
 &= n - \mu(s), s \in \underline{v}
 \end{aligned}$$

The rank deficiency  $\mu(s)$  of the matrix  $[\lambda I_n - \tilde{A}]$ ,  $\lambda = \lambda_s$ ,  $s \in \underline{v}$ , occurs because the last row of each of the  $\mu(s)$  matrices  $\Delta(n_{sj}, \lambda_s, \lambda_s)$  is clearly null. Since the matrix  $R(\lambda)$  is of dimensions  $n \times (n + m)$ , the only possible way in which  $\text{rank } R(\lambda) = n$  for all  $\lambda \in \text{GF}(q)$ , is that all the rows of the matrix  $\tilde{B}$  corresponding to the last rows of the Jordan blocks associated with the same eigenvalue are linearly independent.  $\square$

Corollary 7.1.1. The single-input LSM  $M_1 = (A, b)$  is state reachable if and only if no two Jordan blocks of  $\tilde{A}$  are associated with the same eigenvalue.

Corollary 7.1.2. If the characteristic matrix  $A$  of the LSM  $M = (A, B)$  has  $n$  distinct eigenvalues, then  $M$  is state reachable if and only if all the rows of  $\tilde{B} \equiv V^{-1}B$  are nonzero.

Corollary 7.1.3. The minimum number of inputs required for state reachability of the LSM  $M = (A, B)$  is equal to the greatest number of Jordan blocks associated with the same eigenvalue.

We will use the result of Theorem 7.1.5 along with some additional properties of the Jordan canonical form (7.1.9) - (7.1.12), to derive a simple unreachability criterion for the LSM  $(A, B)$ .

The index  $v_i$  of the eigenvalue  $\lambda_i$  is defined as follows:

$$v_i \equiv \min\{\ell : N[(A - \lambda_i I_n)^\ell] = N[(A - \lambda_i I_n)^{\ell+1}]\}, i \in \underline{v}$$

It is known [118] that the multiplicity of  $\lambda_i$  in the minimal polynomial of  $A$  is equal to  $v_i$ , for all  $i$ , and

$$n_{ij} \leq v_i, i \in \underline{v}, j \in \underline{\mu(i)} \quad (7.1.16)$$

If we let  $[[d]]$  denote the greatest integer  $\leq d$ , then by virtue of (7.1.16) we have the inequality

$$\mu(i) \geq \left[ \left[ \frac{n_i}{v_i} \right] \right], i \in \underline{v} \quad (7.1.17)$$

That is,

$$\begin{aligned} \mu(i) &\geq \frac{n_i}{v_i} && \text{if } \frac{n_i}{v_i} \text{ is an integer} \\ \mu(i) &\geq \left[ \left[ \frac{n_i}{v_i} \right] \right] + 1 && \text{if } \frac{n_i}{v_i} \text{ is a fraction} \end{aligned}$$

It is easy to check that these two inequalities imply the following equality:

$$\mu(i) = \left\{ \left( \frac{n_i + v_i - 1}{v_i} \right) \right\}, \quad i \in \underline{v}$$

Since  $\text{rank } \tilde{B} = \text{rank } P^{-1}B = \text{rank } B$ , if there exist integers  $i \in \underline{v}$  such that

$$\left\{ \left( \frac{n_i + v_i - 1}{v_i} \right) \right\} > \text{rank } B$$

then the rows of the matrix (7.1.15) will be linearly dependent and hence by Theorem 7.1.5, the LSM  $(A, B)$  cannot be reachable. Therefore, we have proved the following sufficient condition for the unreachability property of the LSM  $(A, B)$ .

Theorem 7.1.6. (cf. [65]) If there exist integers  $i \in \underline{v}$  such that

$$\left\{ \left( \frac{n_i + v_i - 1}{v_i} \right) \right\} > \text{rank } B$$

then the LSM  $(A, B)$  is unreachable.

Corollary 7.1.4. If there exist integers  $i \in \underline{v}$  such that

$$\left\{ \left( \frac{n_i + v_i - 1}{v_i} \right) \right\} > m$$

then the LSM  $(A, B)$  is unreachable.

It is interesting to note that for checking the unreachability of an LSM by the above criterion, it is sufficient to know only the numbers  $\{n_i, v_i, i \in \underline{v}\}$  which can be obtained from the Smith canonical form of the characteristic matrix  $A$ . Obviously, knowledge of the isomorphism  $P$  and the eigenvalues  $\lambda_i$  is not needed for the application of the criterion.

Theorem 7.1.7. The LSM  $M = (A, B)$  is state reachable if and only if there does not exist an isomorphism  $P : X \rightarrow X, P \in GF(n, q)$ , such that the isomorphic LSM  $\tilde{M} = (\tilde{A}, \tilde{B}) \equiv (P^{-1}AP, P^{-1}B)$  will have  $\tilde{x}_i(k+1) = \lambda \tilde{x}_i(k), \lambda \in GF(q)$ , as one of its state equations.

Proof. The necessity part of the theorem is obvious; to prove sufficiency, assume that  $M$  is not state reachable. Then by Theorem 4.2.1 there exists an isomorphism  $P : X \rightarrow X, P \in GF(n, q)$ , such that the isomorphic LSM  $\tilde{M} = (\tilde{A}, \tilde{B}) \equiv (P^{-1}AP, P^{-1}B)$  has the form

$$\left( \begin{bmatrix} \tilde{A}_{11} & \tilde{A}_{12} \\ 0 & \tilde{A}_{22} \end{bmatrix}, \begin{bmatrix} \tilde{B}_1 \\ 0 \end{bmatrix} \right)$$

Now consider a further isomorphism  $\hat{P} : X \rightarrow X, \hat{P} \in GF(n, q)$ , of the form  $\hat{P} = \hat{P}_{11} \oplus \hat{P}_{22}$ , and choose  $\hat{P}_{22}$  such that the submatrix  $\hat{P}_{22}^{-1} \tilde{A}_{22} \hat{P}_{22}$  will have the Jordan canonical form, denoted by  $J[\tilde{A}_{22}]$ . That is, the new isomorphic LSM  $\tilde{\tilde{M}} = (\tilde{\tilde{A}}, \tilde{\tilde{B}}) \equiv (\hat{P}^{-1} \tilde{A} \hat{P}, \hat{P}^{-1} \tilde{B})$  is given by

$$\begin{bmatrix} \tilde{\tilde{x}}^I(k+1) \\ \tilde{\tilde{x}}^{II}(k+1) \end{bmatrix} = \begin{bmatrix} \hat{P}_{11}^{-1} \tilde{A}_{11} \hat{P}_{11} & \hat{P}_{11}^{-1} \tilde{A}_{12} \hat{P}_{22} \\ 0 & J[\tilde{A}_{22}] \end{bmatrix} \begin{bmatrix} \tilde{\tilde{x}}^I(k) \\ \tilde{\tilde{x}}^{II}(k) \end{bmatrix} + \begin{bmatrix} \hat{P}_{11}^{-1} \tilde{B}_1 \\ 0 \end{bmatrix} u(k)$$

From this representation and the structure of Jordan blocks, it is clear that there is at least one component of the state equations which is of the form  $\tilde{x}_\ell(k+1) = \lambda \tilde{x}_\ell(k)$ . Therefore,  $\tilde{M}$  cannot be state reachable. But this implies that  $\tilde{M}$  and consequently  $M$  cannot be state reachable.  $\square$

The reachability property of an LSM  $M = (A, B)$  can often be effectively maintained by means of a scalar input sequence, that is, by means of  $w(0)w(1) \dots w(\ell-1) \in GF(q)^*$ ,  $w(k) \in GF(q)$ , instead of a vector sequence  $u(0)u(1) \dots u(\ell-1) \in \mathcal{U}^*$ ,  $u(k) \in \mathcal{U}$ , where  $u(k) = vw(k)$  for a constant vector  $v \in GF(q)^m$ . Defining  $b \equiv Bv$ ,  $M$  reduces to the single-input LSM  $M_1 = (A, b)$ . This possibility of reducing the input sequence space  $\mathcal{U}^*$  to  $GF(q)^*$  can be characterized in terms of the structure of the Jordan canonical LSM  $\tilde{M} = (\tilde{A}, \tilde{B})$  as shown in the following theorem.

Theorem 7.1.8. If the LSM  $M = (A, B)$  is state reachable, then there exists a constant vector  $v \in GF(q)^m$  such that the single-input LSM  $M_1 = (A, Bv)$  is state reachable if and only if any two blocks in the Jordan canonical form of  $A$  are associated with unequal eigenvalues of  $A$ .

Proof. Let  $V : X \rightarrow X$ ,  $V \in GF(n, q)$ , be the isomorphism that yields the isomorphic LSM  $\tilde{M} = (\tilde{A}, \tilde{B}) \equiv (V^{-1}AV, V^{-1}B)$  having the Jordan canonical form

$$\tilde{A} \equiv \tilde{A}_1 \oplus \tilde{A}_2 \oplus \dots \oplus \tilde{A}_v, \quad \tilde{B} \equiv \begin{pmatrix} \tilde{B}^1 \\ \tilde{B}^2 \\ \vdots \\ \tilde{B}^v \end{pmatrix}$$

where

$$\tilde{A}_i \equiv \begin{pmatrix} \lambda_i & 1 & & & \\ & \lambda_i & 1 & & \\ & & \ddots & \ddots & \\ & & & \lambda_i & 1 \\ & & & & \lambda_i \end{pmatrix} \in GF(q)^{n_i \times n_i}, i \in \underline{v}$$

$$\tilde{B}^i \equiv \begin{pmatrix} \tilde{b}_1^i \\ \tilde{b}_2^i \\ \vdots \\ \tilde{b}_{n_i}^i \end{pmatrix} \in GF(q)^{n_i \times 1}, i \in \underline{v}$$

$$\sum_{i=1}^v n_i = n$$

Since  $M$  is assumed to be state reachable, by Theorem 7.1.5 we have

$$b_{n_i}^i \neq 0, i \in \underline{v}$$

Suppose that two blocks, say  $\tilde{A}_1$  and  $\tilde{A}_2$ , are associated with the same eigenvalue  $\lambda_1 = \lambda_2$ , and let  $\tilde{b} \in GF(q)^n$  and  $w(k) \in GF(q)$ . Then it is easy to see that the single-input LSM  $\tilde{M}_1 = (\tilde{A}, \tilde{b})$  is not state reachable, that is,  $\text{rank } \tilde{K}_1 \equiv [\hat{b}, \hat{A}\tilde{b}, \dots, \hat{A}^{n-1}\tilde{b}] < n$  since the  $n_1$ th and



$(n_1 + n_2)$ th rows of  $\tilde{K}_1$  are just  $[\hat{b}_{1n_1}, \lambda_1 \hat{b}_{1n_1}, \dots, (\lambda_1)^{n-1} \hat{b}_{1n_1}]$  and  $[\hat{b}_{2n_2}, \lambda_2 \hat{b}_{2n_2}, \dots, (\lambda_2)^{n-1} \hat{b}_{2n_2}]$ , which are linearly dependent if  $\lambda_1 = \lambda_2$ . Therefore, in this case, the state reachability property of  $M$  cannot be maintained by a scalar input. To prove the converse, suppose that the blocks  $\tilde{A}_i$  are associated with the distinct eigenvalues  $\lambda_i, i \in \underline{v}$ . Choose a vector  $v \in GF(q)^m$  such that

$$\tilde{B}V \equiv \bar{b} = \begin{bmatrix} \bar{b}_1 \\ \bar{b}_2 \\ \vdots \\ \bar{b}_v \end{bmatrix}, \quad \bar{b}_i \equiv \begin{bmatrix} \bar{b}_{i1} \\ \bar{b}_{i2} \\ \vdots \\ \bar{b}_{in_i} \end{bmatrix}, \quad i \in \underline{v}$$

has  $\bar{b}_{in_i} \neq 0, i \in \underline{v}$ . This can be done, for instance, by choosing the entries of  $v$  to be algebraically independent of all the entries of  $\tilde{B}$ .

We will show that  $\text{rank} [\bar{b}, \tilde{A}\bar{b}, \dots, \tilde{A}^{n-1}\bar{b}] = n$ . It is clear that

$$R([\bar{b}, (\tilde{A} - \lambda_v I_n)\bar{b}, (\tilde{A} - \lambda_v I_n)^2\bar{b}, \dots, (\tilde{A} - \lambda_v I_n)^{n_v-1}\bar{b}]) = R([\bar{b}, \tilde{A}\bar{b}, \tilde{A}^2\bar{b}, \dots, \tilde{A}^{n_v-1}\bar{b}])$$

Let

$$z \equiv (\tilde{A} - \lambda_v I_n)^{n_v}\bar{b} = \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_v \end{bmatrix}, \quad z_i \equiv \begin{bmatrix} z_{i1} \\ z_{i2} \\ \vdots \\ z_{in_i} \end{bmatrix}, \quad i \in \underline{v}$$

so that

$$z = \tilde{A}^{\tilde{n}_{\tilde{v}}-1} \tilde{b} + \sum_{i=0}^{\tilde{n}_{\tilde{v}}-1} a_i \tilde{A}^i \tilde{b}, \quad a_i \in \text{GF}(q)$$

Then

$$\begin{aligned} & R([z, (\tilde{A}-\lambda_{\tilde{v}-1} I_{\tilde{n}})z, (\tilde{A}-\lambda_{\tilde{v}-1} I_{\tilde{n}})^2 z, \dots, (\tilde{A}-\lambda_{\tilde{v}-1} I_{\tilde{n}})^{\tilde{n}_{\tilde{v}-1}-1} z]) \\ &= R([z, \tilde{A}z, \tilde{A}^2 z, \dots, \tilde{A}^{\tilde{n}_{\tilde{v}}-1} z]) \end{aligned}$$

Repeating the above procedure, we can show that

$$\begin{aligned} & R([\bar{b}, (\tilde{A}-\lambda_{\tilde{v}} I_{\tilde{n}})\bar{b}, (\tilde{A}-\lambda_{\tilde{v}} I_{\tilde{n}})^2 \bar{b}, \dots, (\tilde{A}-\lambda_{\tilde{v}} I_{\tilde{n}})^{\tilde{n}_{\tilde{v}}-1} \bar{b}, \\ & (\tilde{A}-\lambda_{\tilde{v}} I_{\tilde{n}})^{\tilde{n}_{\tilde{v}}-1} \bar{b}, (\tilde{A}-\lambda_{\tilde{v}-1} I_{\tilde{n}})(\tilde{A}-\lambda_{\tilde{v}} I_{\tilde{n}})^{\tilde{n}_{\tilde{v}}-1} \bar{b}, \dots, \\ & (\tilde{A}-\lambda_{\tilde{v}-1} I_{\tilde{n}})^{\tilde{n}_{\tilde{v}-1}-1} (\tilde{A}-\lambda_{\tilde{v}} I_{\tilde{n}})^{\tilde{n}_{\tilde{v}}-1} \bar{b}, \dots, \\ & (\tilde{A}-\lambda_2 I_{\tilde{n}})^{\tilde{n}_2} \dots (\tilde{A}-\lambda_{\tilde{v}} I_{\tilde{n}})^{\tilde{n}_{\tilde{v}}-1} \bar{b}, (\tilde{A}-\lambda_1 I_{\tilde{n}})(\tilde{A}-\lambda_2 I_{\tilde{n}})^{\tilde{n}_2} \dots (\tilde{A}-\lambda_{\tilde{v}} I_{\tilde{n}})^{\tilde{n}_{\tilde{v}}-1} \bar{b}, \\ & \dots, (\tilde{A}-\lambda_1 I_{\tilde{n}})^{\tilde{n}_1-1} (\tilde{A}-\lambda_2 I_{\tilde{n}})^{\tilde{n}_2} \dots (\tilde{A}-\lambda_{\tilde{v}} I_{\tilde{n}})^{\tilde{n}_{\tilde{v}}-1} \bar{b}]) \\ &= R([\bar{b}, \tilde{A}\bar{b}, \tilde{A}^2 \bar{b}, \dots, \tilde{A}^{\tilde{n}-1} \bar{b}]) \end{aligned}$$

The last  $\tilde{n}_{\tilde{v}}$  rows of the vectors

$$(\tilde{A}-\lambda_{\tilde{v}} I_{\tilde{n}})^{\tilde{n}_{\tilde{v}}-1} \bar{b}, \dots, (\tilde{A}-\lambda_{\tilde{v}} I_{\tilde{n}})^2 \bar{b}, (\tilde{A}-\lambda_{\tilde{v}} I_{\tilde{n}}) \bar{b}, \bar{b}$$

form the triangular matrix

$$\begin{pmatrix} b_{vn_v} & * & * & . & . & . & * & * \\ & b_{vn_v} & * & . & . & . & * & * \\ & & b_{vn_v} & & & & * & * \\ & & & . & & & . & . \\ & & & & . & & . & . \\ & & & & & . & . & . \\ & & & & & & b_{vn_v} & * \\ & & & & & & & b_{vn_v} \end{pmatrix}$$

where  $b_{vn_v} \neq 0$  and  $*$ 's denote the possibly nonzero entries. Note that the vector  $z$  has zeros as the last  $n_v$  entries, and that

$$z_{v-1, n_{v-1}} = (\lambda_{v-1} - \lambda_v)^{n_{v-1}} b_{v-1, n_{v-1}} \neq 0$$

Now by direct computation it can be shown that the matrix

$$T \equiv [(\tilde{A} - \lambda_1 I_n)^{n_1-1} \tilde{A}^{n_1-1} \tilde{b}, \dots, (\tilde{A} - \lambda_v I_n)^{n_v-1} \tilde{b}, \dots, (\tilde{A} - \lambda_v I_n)^{n_v-1} \tilde{b}, \tilde{b}]$$

has a triangular form with diagonal elements

$$(\lambda_1 - \lambda_2)^{n_2} \dots (\lambda_1 - \lambda_v)^{n_v} b_{1n_1} \neq 0, \dots, b_{vn_v} \neq 0$$

Since  $R(T) = R([\tilde{b}, \tilde{A}\tilde{b}, \dots, \tilde{A}^{n-1}\tilde{b}])$  and  $\det T \neq 0$ , it follows that  $\det[\tilde{b}, \tilde{A}\tilde{b}, \dots, \tilde{A}^{n-1}\tilde{b}] \neq 0$  which implies that  $\det[\tilde{B}v, \tilde{A}\tilde{B}v, \dots, \tilde{A}^{n-1}\tilde{B}v] = \det[Bv, ABv, \dots, A^{n-1}Bv] \neq 0$ . Therefore, the single-input LSM  $(A, Bv) = (A, b)$  is state reachable.  $\square$

## 7.2. Selective State Reachability of LSMs

So far, we have considered the influencibility of the entire state vector  $x(k)$  as a point in the state space  $X$  of the LSM  $M = (A, B)$  by input vector sequences  $u(0)u(1) \dots u(\ell-1) \in U^*$ . However, if we take a "microscopic" viewpoint by considering the reachability of state components  $x_i(k)$ ,  $i \in \underline{n}$ , by input components  $u_j(k)$ ,  $j \in \underline{m}$ , then many additional types of reachability become available. Some of these are formalized in the following definitions.

Definition 7.2.1. The  $i$  th component  $x_i^1(k)$  of the state  $x^1(k) \in X$  is said to be *selectively  $\ell$ -reachable* by the  $j$  th component  $u_j(k)$  of the input  $u(k) \in U$  if there exists a scalar sequence  $u_j(0)u_j(1) \dots u_j(\ell-1) \in GF(q)^*$  which transfers  $x_i^1(k)$  to  $x_i^2(k)$  for any  $x^2 \in X$ , where  $u_s(k) = 0$  for all  $s \neq j$ . If every component  $x_i^1(k)$ ,  $i \in \underline{n}$ , of  $x^1(k) \in X$  is selectively  $\ell$ -reachable, then the LSM is said to be *selectively  $\ell$ -state reachable*.

Definition 7.2.2. If every component  $x_i(k)$ ,  $i \in \underline{n}$ , of the state  $x(k) \in X$  is selectively  $\ell$ -reachable at every clock period  $k \in K' \subseteq K$ , where  $K'$  is the admissible clock period set, then the LSM is said to be *selectively completely  $\ell$ -state reachable*.

Definition 7.2.3. If the  $i$  th component  $x_i(k)$  of the state  $x(k) \in X$  is selectively  $\ell$ -reachable by all input component sequences  $u_j(0)u_j(1) \dots u_j(\ell-1)$ ,  $j \in \underline{m}$ , separately, then it is said to be *strongly  $\ell$ -reachable*. If every component  $x_i(k)$ ,  $i \in \underline{n}$ , of the state  $x(k) \in X$  is strongly  $\ell$ -reachable, then the LSM is said to be *strongly  $\ell$ -state reachable*.

Definition 7.2.4. If an LSM is strongly  $\ell$ -state reachable at every clock period  $k \in K' \subseteq K$ , where  $K'$  is the admissible clock period set, then it is said to be *strongly completely  $\ell$ -state reachable*.

Definition 7.2.5. If an LSM is  $\ell$ -state reachable by each of the input sequences  $u_j(0)u_j(1) \dots u_j(\ell-1)$ ,  $j \in \underline{m}$ , separately, then it is said to be *state normal*.

From Definition 7.2.5, it follows that an LSM  $M = (A, B)$  is state normal if and only if  $\text{rank } K_j = \text{rank}[b^j, Ab^j, \dots, A^{n-1}b^j] = n$ ,  $j \in \underline{m}$ , where  $b^j$  denotes the  $j$ th column of the input matrix  $B$ , that is, if and only if the  $m$  single-input LSMs  $x(k+1) = Ax(k) + b^j u_j(k)$ ,  $j \in \underline{m}$ , are state reachable. Therefore, in view of Theorem 6.1.1, there exists a large number of state reachability criteria for characterizing the normality property of LSMs.

We assume that the above definitions of reachability based on the state components and input component sequences clearly indicate the possibility for defining additional types of reachability and hence their explicit formalizations will not be pursued any further. In order to avoid excessive detail, we will adopt a similar point of view with respect to the theoretical development of the concepts introduced in the above definitions in the sense that we will concentrate only on the property of selective state reachability and assume that our results, if needed, can be easily stated for other types of state reachability.

In Theorem 6.3.1 we provided a large number of criteria for ascertaining the state reachability property of the LSM  $M = (A, B)$ .

However, it is clear that none of these criteria provides any information about the selective reachability of a particular component of the state vector. In order to investigate the possibility for developing necessary and sufficient conditions for selective state reachability, consider the LSM  $M$  in the form

$$x(k+1) = Ax(k) + \sum_{j=1}^m b^j u_j(k) \quad (7.2.1)$$

where  $b^j$ ,  $j \in \underline{m}$ , denote the columns of  $B$  and  $u_j(k)$ ,  $j \in \underline{m}$ , are the components of the input vector  $u(k) \in U$ . Assuming that  $u_j(k) = 0$  for all  $j \neq s$ , the zero-state solution of the state equation (7.2.1) can be written as

$$\begin{aligned} x(k) &= \sum_{i=0}^{k-1} A^{k-i-1} b^s u_s(i) \\ &= [b^s, Ab^s, \dots, A^{k-1}b^s] \begin{bmatrix} u_s(k-1) \\ u_s(k-2) \\ \vdots \\ u_s(0) \end{bmatrix} \\ &= \begin{bmatrix} b_1^s & \sum_{i=1}^n a_{1i}^{(1)} b_i^s & \dots & \sum_{i=1}^n a_{1i}^{(k-1)} b_i^s \\ b_2^s & \sum_{i=1}^n a_{2i}^{(1)} b_i^s & \dots & \sum_{i=1}^n a_{2i}^{(k-1)} b_i^s \\ \vdots & \vdots & & \vdots \\ b_n^s & \sum_{i=1}^n a_{ni}^{(1)} b_i^s & \dots & \sum_{i=1}^n a_{ni}^{(k-1)} b_i^s \end{bmatrix} \begin{bmatrix} u_s(k-1) \\ u_s(k-2) \\ \vdots \\ u_s(0) \end{bmatrix} \end{aligned} \quad (7.2.2)$$

where  $a_{uv}^{(\ell)}$  are the entries in the matrix  $A^\ell$ . From (7.2.2) it is clear that the  $i$ th component  $x_i(k)$  of the state  $x(k)$  can be influenced by the  $s$ th component sequence  $u_s(0)u_s(1) \dots u_s(k-1)$  if and only if there is at least one nonzero element in the  $i$ th row of the matrix  $K_s \equiv [b^2, Ab^s, \dots, A^{k-1}b^s]$ . This observation clearly indicates the key role that will be played by the matrices  $K_j \equiv [b^j, Ab^j, \dots, A^{n-1}b^j]$ ,  $j \in \underline{m}$ , in the investigation of the concept of selective state reachability. For this reason,  $K_j$  will be called the *selective state reachability matrix*. In the ensuing discussion, we will explore further properties of this matrix in conjunction with the Jordan canonical form of the characteristic matrix  $A$  to develop selective state reachability criteria for the LSM  $M = (A, B)$ . We will break down our discussion into a number of cases.

Case 1.  $A$  has  $n$  distinct eigenvalues.

If we denote the  $n$  distinct eigenvalues of  $A$  by  $\lambda_1, \lambda_2, \dots, \lambda_n$ , and let the matrix of the isomorphism  $V_1 : X \rightarrow X$  consist of the corresponding  $n$  linearly independent eigenvectors  $v^1, v^2, \dots, v^n$ , that is,  $V_1 \equiv [v^1, v^2, \dots, v^n]$ , then the isomorphic LSM  $\tilde{M} = (\tilde{A}, \tilde{B}) \equiv (V_1^{-1}AV_1, V_1^{-1}B)$  will have the form

$$\begin{pmatrix} \tilde{x}_1(k+1) \\ \tilde{x}_2(k+1) \\ \vdots \\ \tilde{x}_n(k+1) \end{pmatrix} = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix} \begin{pmatrix} \tilde{x}_1(k) \\ \tilde{x}_2(k) \\ \vdots \\ \tilde{x}_n(k) \end{pmatrix} + \begin{pmatrix} \tilde{b}_{11} & \tilde{b}_{12} & \dots & \tilde{b}_{1m} \\ \tilde{b}_{21} & \tilde{b}_{22} & \dots & \tilde{b}_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ \tilde{b}_{n1} & \tilde{b}_{n2} & \dots & \tilde{b}_{nm} \end{pmatrix} \begin{pmatrix} u_1(k) \\ u_2(k) \\ \vdots \\ u_m(k) \end{pmatrix} \quad (7.2.3)$$

which is equivalent to the following  $n$  uncoupled submachines:

$$\tilde{x}_i(k+1) = \lambda_i \tilde{x}_i(k) + \sum_{j=1}^m \tilde{b}_{ij} u_j(k), \quad i \in \underline{n} \quad (7.2.4)$$

From (7.2.4) it is clear that the  $j$ th component of the input vector can affect the  $i$ th component  $\tilde{x}_i(k)$  of the state vector if and only if

$$\tilde{b}_{ij} \neq 0, \quad i \in \underline{n}, \quad j \in \underline{m} \quad (7.2.5)$$

We summarize the above observation in the following theorem.

Theorem 7.2.1. If the characteristic matrix  $A$  of the LSM  $M = (A, B)$  is cyclic, then the  $i$ th component  $x_i(k)$  of the state  $x(k)$  of  $M$  is selectively reachable by the  $j$ th component  $u_j(k)$  of the input  $u(k)$  if and only if  $\tilde{b}_{ij} \neq 0$ , where  $\tilde{b}_{ij}$  is defined by (7.2.3).

Since  $\tilde{B} = V_1^{-1}B$ , we can write

$$B = [b^1, b^2, \dots, b^m] = \tilde{V}_1 B = [v^1, v^2, \dots, v^n] \begin{pmatrix} \tilde{b}_{11} & \tilde{b}_{12} & \dots & \tilde{b}_{1m} \\ \tilde{b}_{21} & \tilde{b}_{22} & \dots & \tilde{b}_{2m} \\ \vdots & \vdots & & \vdots \\ \tilde{b}_{n1} & \tilde{b}_{n2} & \dots & \tilde{b}_{nm} \end{pmatrix}$$

so that

$$b^j = \sum_{i=1}^n \tilde{b}_{ij} v^i, \quad j \in \underline{m} \quad (7.2.6)$$



Now

$$Ab^j = \sum_{i=1}^n \tilde{b}_{ij} Av^i, \quad j \in \underline{m} \quad (7.2.7)$$

In view of the eigenvalue-eigenvector relationships  $Av^i = \lambda_i v^i$ ,  $i \in \underline{n}$ , (7.2.7) becomes

$$Ab^j = \sum_{i=1}^n \tilde{b}_{ij} \lambda_i v^i, \quad j \in \underline{m}$$

Similarly,

$$\begin{aligned} A^2 b^j &= A(Ab^j) = \sum_{i=1}^n \tilde{b}_{ij} \lambda_i Av^i = \sum_{i=1}^n \tilde{b}_{ij} (\lambda_i)^2 v^i, \quad j \in \underline{m} \\ &\vdots \\ A^{n-1} b^j &= \sum_{i=1}^n \tilde{b}_{ij} (\lambda_i)^{n-1} v^i, \quad j \in \underline{m} \end{aligned}$$

Therefore, the selective state reachability matrix  $K_j$  can be expressed as

$$\begin{aligned} K_j &\equiv [b^j, Ab^j, \dots, A^{n-1} b^j] \\ &= \left[ \sum_{i=1}^n \tilde{b}_{ij} v^i, \sum_{i=1}^n \tilde{b}_{ij} (\lambda_i)^2 v^i, \dots, \sum_{i=1}^n \tilde{b}_{ij} (\lambda_i)^{n-1} v^i \right], \quad j \in \underline{m} \end{aligned} \quad (7.2.8)$$

Since

$$\begin{aligned}
 b^j &= \sum_{i=1}^n \tilde{b}_{ij} v^i = [v^1, v^2, \dots, v^n] \begin{pmatrix} \tilde{b}_{1j} \\ \tilde{b}_{2j} \\ \vdots \\ \tilde{b}_{nj} \end{pmatrix} \\
 &= [v^1, v^2, \dots, v^n] \begin{pmatrix} \tilde{b}_{1j} & & & \\ & \tilde{b}_{2j} & & \\ & & \ddots & \\ & & & \tilde{b}_{nj} \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}
 \end{aligned}
 \tag{7.2.9}$$

we have

$$\begin{aligned}
 {}^A \ell b^j &= \sum_{i=1}^n \tilde{b}_{ij} (\lambda_i)^\ell v^i = [v^1, v^2, \dots, v^n] \begin{pmatrix} \tilde{b}_{1j} (\lambda_1)^\ell \\ \tilde{b}_{2j} (\lambda_2)^\ell \\ \vdots \\ \tilde{b}_{nj} (\lambda_n)^\ell \end{pmatrix} \\
 &= [v^1, v^2, \dots, v^n] \begin{pmatrix} \tilde{b}_{1j} & & & \\ & \tilde{b}_{2j} & & \\ & & \ddots & \\ & & & \tilde{b}_{nj} \end{pmatrix} \begin{pmatrix} (\lambda_1)^\ell \\ (\lambda_2)^\ell \\ \vdots \\ (\lambda_n)^\ell \end{pmatrix}, \quad \begin{matrix} \ell \in \underline{n-1}, \\ j \in \underline{m} \end{matrix}
 \end{aligned}
 \tag{7.2.10}$$

In view of (7.2.9) and (7.2.10),  $K_j$  can be written as a product of three matrices as follows:

$$\begin{aligned}
 K_j &\equiv [b^j, Ab^j, \dots, A^{n-1}b^j] \\
 &= \begin{bmatrix} v_{11} & v_{12} & \dots & v_{1n} \\ v_{21} & v_{22} & \dots & v_{2n} \\ \vdots & \vdots & & \vdots \\ v_{n1} & v_{n2} & \dots & v_{nn} \end{bmatrix} \begin{bmatrix} \tilde{b}_{1j} \\ \tilde{b}_{2j} \\ \vdots \\ \tilde{b}_{nj} \end{bmatrix} \begin{bmatrix} 1 & \lambda_1 & (\lambda_1)^2 & \dots & (\lambda_1)^{n-1} \\ 1 & \lambda_2 & (\lambda_2)^2 & \dots & (\lambda_2)^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & \lambda_n & (\lambda_n)^2 & \dots & (\lambda_n)^{n-1} \end{bmatrix} \\
 &\equiv V_1 D_{1j} W_1, \quad j \in \underline{m} \tag{7.1.11}
 \end{aligned}$$

The matrix  $W_1 \in GF(q)^{n \times n}$  is known as the Vandermonde matrix. It can be shown that the determinant of  $W$  is equal to  $\prod_{1 \leq i < j \leq n} (\lambda_i - \lambda_j)$ , and hence  $W$  is singular if  $\lambda_i = \lambda_j$  for  $i \neq j$ . However, in our case  $\lambda_i$ ,  $i \in \underline{n}$ , are distinct and, therefore,  $W_1$  is nonsingular. Since  $V_1$  is the matrix of linearly independent eigenvectors of  $A$ , from (7.2.11) it follows that

$$\text{rank } K_j = \text{rank } D_{1j}, \quad j \in \underline{m} \tag{7.2.12}$$

The rank of the diagonal matrix  $D_{1j}$  is clearly equal to the number of nonzero elements  $\tilde{b}_{ij}$ ,  $i \in \underline{n}$ ,  $j \in \underline{m}$ . Since by Theorem 7.2.1 the  $i$ th state component  $x_i(k)$  of the state vector  $x(k)$  is selectively reachable by the  $j$ th component  $u_j(k)$  of the input vector  $u(k)$  if and only if  $\tilde{b}_{ij} \neq 0$ , then (7.2.12) implies that the rank of  $K_j$  is equal to the

number  $n_j$  of the state vector components that can be selectively affected by  $u_j(k)$ . We formalize this result in the following theorem.

Theorem 7.2.2. If  $A$  has  $n$  distinct eigenvalues, then the rank of the selective state reachability matrix  $K_j \equiv [b^j, Ab^j, \dots, A^{n-1}b^j]$  is equal to the number of state vector components that are selectively reachable by the  $j$ th component  $u_j(k)$  of the input vector  $u(k)$ .

We can actually identify the state components that are reachable as a result of the rank condition of the matrix  $K_j$ . To see this, we can use (7.2.8) to write

$$A^\ell b^j = \sum_{i=1}^{n_j} \tilde{b}_{ij}(\lambda_i)^\ell v^i, \quad \ell \in \underline{n_j-1}, \quad j \in \underline{m} \quad (7.2.13)$$

Since  $\text{rank } K_j = n_j$ ,  $j \in \underline{m}$ , we can express  $A^\ell b^j$ ,  $\ell \geq n_j$ , as a linear combination of the linearly independent columns of  $K_j$ . In particular,

$$A^{n_j} b^j = \sum_{s=0}^{n_j-1} a_{sj} A^s b^j, \quad a_{sj} \in \text{GF}(q) \quad (7.2.14)$$

In view of (7.2.13) and (7.2.14), we have the following equality:

$$\sum_{i=1}^{n_j} \tilde{b}_{ij}(\lambda_i)^{n_j} v^i = \sum_{s=0}^{n_j-1} a_{sj} \left( \sum_{i=1}^{n_j} \tilde{b}_{ij}(\lambda_i)^{s+n_j} v^i \right)$$

or

$$\sum_{i=1}^{n_j} (\tilde{b}_{ij}(\lambda_i)^{n_j}) v^i = \sum_{i=1}^{n_j} \left( \sum_{s=0}^{n_j-1} a_{sj} \tilde{b}_{ij}(\lambda_i)^s \right) v^i \quad (7.2.15)$$

Equating the coefficients of  $v^i$ ,  $i \in \underline{n}$ , in (7.2.15), we obtain

$$\begin{aligned} \tilde{b}_{ij}(\lambda_i)^{n_j} &= \sum_{s=0}^{n_j-1} a_{sj} \tilde{b}_{ij}(\lambda_i)^s \\ (\lambda_i)^{n_j} &= \sum_{s=0}^{n_j-1} a_{sj} (\lambda_i)^s, \quad i \in \underline{n_j}, \quad j \in \underline{m} \end{aligned} \quad (7.2.16)$$

since  $\tilde{b}_{ij} \neq 0$ . Equation (7.2.16) is an  $n_j$ th order monic polynomial whose roots are the  $n_j$  eigenvalues associated with the selectively reachable state components. Therefore, in order to identify these components, it is necessary to determine  $a_{sj}$  in equation (7.2.14), and then solve equation (7.2.16). We will illustrate this procedure by an example. Consider the following LSM over GF(3):

$$\begin{bmatrix} x_1(k+1) \\ x_2(k+1) \\ x_3(k+1) \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1(k) \\ x_2(k) \\ x_3(k) \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} u_1(k) \\ u_2(k) \end{bmatrix}$$

$$K_1 \equiv [b^1, Ab^1, A^2b^1] = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix}$$

$$K_2 \equiv [b^2, Ab^2, A^2b^2] = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix}$$

Since  $\text{rank } K_1 = 2$ , that is  $n_1 = 2$ , and  $\text{rank } K_2 = 1$ , that is  $n_2 = 1$ , there are two state components which are selectively reachable by

$u_1(k)$ , and one state component that is selectively reachable by  $u_2(k)$ . In order to identify these state components, we need to solve the following sets of equations:

$$A^2 b^1 = a_{01} b^1 + a_{11} A b^1$$

$$A b^2 = a_{02} b^2$$

which in view of the given data, reduce to

$$\begin{pmatrix} 0 \\ 2 \\ 0 \end{pmatrix} = a_{01} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + a_{11} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

$$\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = a_{02} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

Solving these equations, we obtain  $a_{01} = 0$ ,  $a_{11} = 2$ , and  $a_{02} = 1$ . Substituting in (7.2.16), we get  $(\lambda)^2 = 2\lambda$  for  $j = 1$ , and  $\lambda = 1$  for  $j = 2$ . Therefore,  $\lambda_1 = 0$ ,  $\lambda_2 = 0$  for  $j = 1$ , and  $\lambda_3 = 1$  for  $j = 2$ . Thus the state components  $x_1(k)$  and  $x_2(k)$  are selectively reachable only by  $u_1(k)$ , and  $x_3(k)$  is selectively reachable only by  $u_2(k)$ .

Case 2. A has  $n$  repeated eigenvalues.

If we let the matrix of the isomorphism  $V_2 : X \rightarrow X$  consist of the  $n$  linearly independent generalized eigenvectors  $v^1, v^2, \dots, v^n$  of the form (7.1.2), then the isomorphic LSM  $\tilde{M} = (\tilde{A}, \tilde{B}) \equiv (V_2^{-1} A V_2, V_2^{-1} B)$  will have the following form:

$$\begin{bmatrix} \tilde{x}_1(k+1) \\ \tilde{x}_2(k+1) \\ \vdots \\ \tilde{x}_{n-1}(k+1) \\ \tilde{x}_n(k+1) \end{bmatrix} = \begin{bmatrix} \lambda_1 & 1 & & & \\ & \lambda_1 & 1 & & \\ & & \ddots & \ddots & \\ & & & \lambda_1 & 1 \\ & & & & \lambda_1 \end{bmatrix} \begin{bmatrix} \tilde{x}_1(k) \\ \tilde{x}_2(k) \\ \vdots \\ \tilde{x}_{n-1}(k) \\ \tilde{x}_n(k) \end{bmatrix} + \begin{bmatrix} \tilde{b}_{11} & \tilde{b}_{12} & \dots & \tilde{b}_{1m} \\ \tilde{b}_{21} & \tilde{b}_{22} & \dots & \tilde{b}_{2m} \\ \vdots & \vdots & & \vdots \\ \tilde{b}_{n1} & \tilde{b}_{n2} & \dots & \tilde{b}_{nm} \end{bmatrix} \begin{bmatrix} u_1(k) \\ u_2(k) \\ \vdots \\ u_m(k) \end{bmatrix} \quad (7.2.17)$$

which is equivalent to the following  $n$  coupled linear submachines:

$$\tilde{x}_i(k+1) = \lambda_1 \tilde{x}_i(k) + \tilde{x}_{i+1}(k) + \sum_{j=1}^m \tilde{b}_{ij} u_j(k), \quad i \in \underline{n-1} \quad (7.2.18a)$$

$$\tilde{x}_n(k+1) = \lambda_1 \tilde{x}_n(k) + \sum_{j=1}^m \tilde{b}_{nj} u_j(k) \quad (7.2.18b)$$

From (7.2.18) it is evident that the  $j$ th input component  $u_j(k)$  can selectively affect the  $\ell$ th state component  $\tilde{x}_\ell(k)$  if and only if

$$\tilde{b}_{\ell j} \neq 0, \quad \ell \in \underline{n}$$

and

$$\tilde{b}_{sj} = 0, \quad s = \ell+1, \ell+2, \dots, n \quad (7.2.19)$$

It is also clear that the overall LSM  $M = (A, B)$  under consideration is state reachable by the  $j$ th input component  $u_j(k)$  if and only if  $\tilde{b}_{nj} \neq 0$  since by Theorem 7.1.6 if  $\tilde{b}_{nj} = 0$ , then  $M$  cannot be state reachable. Furthermore,  $M$  is not state reachable by  $u_j(k)$  if and only if  $\tilde{b}_{sj} = 0, s \in \underline{n}$ . We summarize the above observations in the following theorem.

Theorem 7.2.3. If the characteristic matrix of the LSM  $M = (A, B)$  has  $n$  repeated eigenvalues, then  $x_i(k)$ ,  $i \in \underline{n}$ , is selectively reachable by  $u_j(k)$ ,  $j \in \underline{m}$ , if and only if (7.2.19) holds, where  $\tilde{b}_{ij}$  is defined by (7.2.17). Furthermore,  $M$  is state reachable by  $u_j(k)$ ,  $j \in \underline{m}$ , if and only if  $\tilde{b}_{nj} \neq 0$ , and  $M$  is state unreachable by  $u_j(k)$  if and only if  $\tilde{b}_{ij} = 0$ ,  $i \in \underline{n}$ .

In order to further investigate the selective reachability properties of this special class of LSMs, we need to take a closer look at the selective state reachability matrix  $K_j$ . From the relation  $AV_2 = V_2\tilde{A}$  and the special form of  $\tilde{A}$  given by (7.2.17), we obtain the following eigenvalue-eigenvector relations:

$$Av^1 = \lambda_1 v^1 \quad (7.2.20a)$$

$$Av^i = \lambda_1 v^{i-1}, \quad i = 2, 3, \dots, n \quad (7.2.20b)$$

These relations will be used to derive general expressions for the columns  $A^\ell b^j$ ,  $\ell \in \underline{n-1}$ , of  $K_j$  in terms of  $\lambda_1$  and  $v^i$ ,  $i \in \underline{n}$ , which will consequently result into a decomposition of  $K_j$  into a product of matrices.

Since  $B = V_2\tilde{B}$ , we can express the columns  $b^j$ ,  $j \in \underline{m}$ , of  $B$  as

$$b^j = \sum_{i=1}^n \tilde{b}_{ij} v^i, \quad j \in \underline{m} \quad (7.2.21)$$

Thus

$$Ab^j = \sum_{i=1}^n \tilde{b}_{ij} Av^i, \quad j \in \underline{m}$$



By virtue of (7.2.20),  $Ab^j$  can be expressed as follows:

$$\begin{aligned}
 Ab^j &= \tilde{b}_{1j} \lambda_1 v^1 + \tilde{b}_{2j} (\lambda_1 v^2 + v^1) + \tilde{b}_{3j} (\lambda_1 v^3 + v^2) + \dots + \tilde{b}_{nj} (\lambda_1 v^n + v^{n-1}) \\
 &= \sum_{i=1}^n \tilde{b}_{ij} \lambda_1 v^i + \sum_{i=2}^n \tilde{b}_{ij} v^{i-1} \\
 &= [v^1, v^2, \dots, v^n] \begin{pmatrix} \tilde{b}_{1j} & \tilde{b}_{2j} \\ \tilde{b}_{2j} & \tilde{b}_{3j} \\ \vdots & \vdots \\ \tilde{b}_{n-1,j} & \tilde{b}_{nj} \\ \tilde{b}_{nj} & 0 \end{pmatrix} \begin{pmatrix} \lambda_1 \\ 1 \end{pmatrix} \quad (7.2.22)
 \end{aligned}$$

Similarly,

$$A^2 b^j = A(Ab^j) = \sum_{i=1}^n \tilde{b}_{ij} \lambda_1 A v^i + \sum_{i=2}^n \tilde{b}_{ij} A v^{i-1}$$

Again using (7.2.20), we obtain

$$\begin{aligned}
A^2 b^j &= \sum_{i=1}^n \tilde{b}_{ij} (\lambda_1)^2 v^i + 2 \sum_{i=2}^n \tilde{b}_{ij} \lambda_1 v^{i-1} + \sum_{i=3}^n \tilde{b}_{ij} \lambda_1 v^{i-2} \\
&= [v^1, v^2, \dots, v^n] \begin{pmatrix} \tilde{b}_{1j} & \tilde{b}_{2j} & \tilde{b}_{3j} \\ \tilde{b}_{2j} & \tilde{b}_{3j} & \tilde{b}_{4j} \\ \vdots & \vdots & \vdots \\ \tilde{b}_{n-2,j} & \tilde{b}_{n-1,j} & \tilde{b}_{nj} \\ \tilde{b}_{n-1,j} & \tilde{b}_{nj} & 0 \\ \tilde{b}_{nj} & 0 & 0 \end{pmatrix} \begin{pmatrix} (\lambda_1)^2 \\ 2\lambda_1 \\ 1 \end{pmatrix} \quad (7.2.23)
\end{aligned}$$

Combining (7.2.21), (7.2.22), and (7.2.23), the first three columns of  $K_j$  can be written in the following form:

$$[b^j, Ab^j, A^2 b^j] = [v^1, v^2, \dots, v^n] \begin{pmatrix} \tilde{b}_{1j} & \tilde{b}_{2j} & \tilde{b}_{3j} \\ \tilde{b}_{2j} & \tilde{b}_{3j} & \tilde{b}_{4j} \\ \vdots & \vdots & \vdots \\ \tilde{b}_{n-2,j} & \tilde{b}_{n-1,j} & \tilde{b}_{nj} \\ \tilde{b}_{n-1,j} & \tilde{b}_{nj} & 0 \\ \tilde{b}_{nj} & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & \lambda_1 & (\lambda_1)^2 \\ 0 & 1 & 2\lambda_1 \\ 0 & 0 & 1 \end{pmatrix}$$

The last expression essentially exhibits the pattern that will emerge if we continue in the preceding manner to express  $A^\ell b^j$ ,  $\ell = 3, 4, \dots, n$ , in terms of  $\lambda_1$  and  $v^i$ ,  $i \in \underline{n}$ . Therefore, it is clear that  $K_j$  can be written as a product of three matrices as follows:

$$K_j \equiv [b^j, Ab^j, \dots, A^{n-1}b^j] = V_2 D_{2j} W_2 \quad (7.2.24)$$

where

$$D_{2j} \equiv \begin{pmatrix} \tilde{b}_{1j} & \tilde{b}_{2j} & \tilde{b}_{3j} & \dots & \tilde{b}_{nj} \\ \tilde{b}_{2j} & \tilde{b}_{3j} & \tilde{b}_{4j} & \dots & 0 \\ & & \vdots & & \vdots \\ & & \vdots & & \vdots \\ & \cdot & \tilde{b}_{nj} & \dots & 0 \\ & \cdot & & & \cdot \\ \cdot & \tilde{b}_{nj} & 0 & \dots & 0 \\ \cdot & & & & \\ \tilde{b}_{nj} & 0 & 0 & \dots & 0 \end{pmatrix}, \quad j \in \underline{m} \quad (7.2.25)$$

$$W_2 \equiv \begin{pmatrix} 1 & \begin{pmatrix} 1 \\ 0 \end{pmatrix} \lambda_1 & \begin{pmatrix} 2 \\ 0 \end{pmatrix} (\lambda_1)^2 & \begin{pmatrix} 3 \\ 0 \end{pmatrix} (\lambda_1)^3 & \dots & \begin{pmatrix} n-1 \\ 0 \end{pmatrix} (\lambda_1)^{n-1} \\ & 1 & \begin{pmatrix} 2 \\ 1 \end{pmatrix} \lambda_1 & \begin{pmatrix} 3 \\ 1 \end{pmatrix} (\lambda_1)^2 & \dots & \begin{pmatrix} n-1 \\ 1 \end{pmatrix} (\lambda_1)^{n-2} \\ & & 1 & \begin{pmatrix} 3 \\ 2 \end{pmatrix} \lambda_1 & \dots & \begin{pmatrix} n-1 \\ 2 \end{pmatrix} (\lambda_1)^{n-3} \\ & & & 1 & \dots & \begin{pmatrix} n-1 \\ 3 \end{pmatrix} (\lambda_1)^{n-4} \\ & & & \cdot & \cdot & \vdots \\ & & & & \cdot & \vdots \\ & & & & & 1 \end{pmatrix} \quad (7.2.26)$$

where

$$\binom{s}{t} \equiv \frac{s!}{t!(s-t)!}, \quad s! \equiv s(s-1)(s-2) \dots 1, \quad 0! \equiv 1$$

Since  $V_2$  and  $W_2$  are nonsingular, from (7.2.24) it follows that

$$\text{rank } K_j = \text{rank } D_{2j} \equiv m_j, \quad j \in \underline{m}$$

where

$$\tilde{b}_{m_j, j} \neq 0, \quad m_j \in \underline{n} \quad (7.2.27)$$

$$\tilde{b}_{ij} = 0, \quad i = m_j + 1, m_j + 2, \dots, n; \quad j \in \underline{m}$$

From Theorem 7.2.3 it follows that  $m_j$  is equal to the number of state components that are selectively reachable by  $u_j(k)$ ,  $j \in \underline{m}$ .

Theorem 7.2.4. If the characteristic matrix of the LSM  $M = (A, B)$  has  $n$  repeated eigenvalues, then the rank of the selective state reachability matrix  $K_j \equiv [b^j, Ab^j, \dots, A^{n-1}b^j]$  is equal to the number of state components  $x_i(k)$  that are selectively reachable by the  $j$ th component of the input vector  $u(k)$ .

Equations (7.2.27) together with (7.2.6) yield the following expression for  $b^j$ :

$$b^j = \sum_{i=1}^{m_j} \tilde{b}_{ij} v^i, \quad j \in \underline{m} \quad (7.2.28)$$

Using (7.2.28) and (7.2.24), it is seen that columns of  $K_j$  have the form

$$A_{b^j}^{\ell} = \sum_{i=1}^{m_j} \sum_{s=0}^{m_j-i} \binom{\ell}{s} \tilde{b}_{i+s,j} (\lambda_1)^{\ell-s} v^i, \quad j \in \underline{m} \quad (7.2.29)$$

Since  $\text{rank } K_j = m_j$ , we can express the vectors  $A_{b^j}^{\ell}$  as linear combinations of the  $m_j$  linearly independent columns of  $K_j$  as follows:

$$A_{b^j}^{\ell} = \sum_{\ell=0}^{m_j-1} a_{\ell j} A_{b^j}^{\ell}, \quad j \in \underline{m} \quad (7.2.30)$$

or

$$\sum_{\ell=0}^{m_j} a_{\ell j} A_{b^j}^{\ell} = 0, \quad a_{m_j,j} = -1 \quad (7.2.31)$$

Substituting (7.2.29) into (7.2.31) yields

$$\sum_{\ell=0}^{m_j} a_{\ell j} \left( \sum_{i=1}^{m_j} \sum_{s=0}^{m_j-i} \binom{\ell}{s} \tilde{b}_{i+s,j} (\lambda_1)^{\ell-s} v^i \right) = 0$$

$$\sum_{i=1}^{m_j} \left( \sum_{\ell=0}^{m_j} \sum_{s=0}^{m_j-i} a_{\ell j} \binom{\ell}{s} \tilde{b}_{i+s,j} (\lambda_1)^{\ell-s} \right) v^i = 0$$

with  $a_{m_j,j} = -1$ . Now independence of  $v^i$ ,  $i \in \underline{m_j}$ , implies that

$$\tilde{b}_{m_j,j} (\lambda_1)^{m_j} = \tilde{b}_{m_j,j} \sum_{\ell=0}^{m_j-1} a_{\ell j} (\lambda_1)^{\ell}$$

or

$$(\lambda_1)^{m_j} = \sum_{\ell=0}^{m_j-1} a_{\ell j} (\lambda_1)^\ell \quad (7.2.32)$$

since  $\tilde{b}_{m_j, j} \neq 0$ ,  $m_j \in \underline{n}$ . There are also some other relations that result from the independence of  $v^i$ , but will not be needed for our analysis. Therefore, in order to identify the  $m_j$  selectively reachable state components, we need to determine the scalars  $a_{\ell j}$  from (7.2.30) and then find the roots of the polynomial (7.2.32). Obviously, in this special case all these roots will be equal to  $\lambda_1$ .

Case 3. Confluent eigenvalues of  $A$  are associated with distinct Jordan blocks of  $\tilde{A}$ .

If we let the matrix of the isomorphism  $V_3 : X \rightarrow X$  consist of sets of generalized eigenvectors of the form (7.1.2), one set for each eigenvalue  $\lambda_i$ ,  $i \in \underline{v}$ , then the isomorphic LSM  $\tilde{M} = (\tilde{A}, \tilde{B})$   $(V_3^{-1}AV_3, V_3^{-1}B)$  will have the form

$$\left( \begin{bmatrix} \tilde{A}_1(\lambda_1) & & & \\ & \tilde{A}_2(\lambda_2) & & \\ & & \ddots & \\ & & & \tilde{A}_v(\lambda_v) \end{bmatrix}, \begin{bmatrix} \tilde{B}^1 \\ \tilde{B}^2 \\ \vdots \\ \tilde{B}^v \end{bmatrix} \right) \quad (7.2.33)$$

where

$$\tilde{A}_i(\lambda_i) \equiv \begin{pmatrix} \lambda_i & 1 & & & \\ & \lambda_i & 1 & & \\ & & \ddots & \ddots & \\ & & & \lambda_i & 1 \\ & & & & \lambda_i \end{pmatrix} \in GF(q)^{n_i \times n_i}, i \in \underline{v} \quad (7.2.34)$$

$\lambda_i \neq \lambda_j$  for  $i \neq j$ ;  $i, j \in \underline{v}$ .

$$\tilde{B}^i \equiv \begin{pmatrix} \tilde{b}_{11}^{(i)} & \tilde{b}_{12}^{(i)} & \dots & \tilde{b}_{1m}^{(i)} \\ \tilde{b}_{21}^{(i)} & \tilde{b}_{22}^{(i)} & \dots & \tilde{b}_{2m}^{(i)} \\ \vdots & \vdots & & \vdots \\ \tilde{b}_{n_i,1}^{(i)} & \tilde{b}_{n_i,2}^{(i)} & \dots & \tilde{b}_{n_i,m}^{(i)} \end{pmatrix} \in GF(q)^{n_i \times m}, i \in \underline{v} \quad (7.2.35)$$

Clearly  $\tilde{M}$  is equivalent to the following  $v$  uncoupled submachines:

$$\tilde{x}^i(k+1) = \tilde{A}_i(\lambda_i) \tilde{x}^i(k) + \tilde{B}_u^i(k), i \in \underline{v} \quad (7.2.36)$$

The  $i$ th submachine is composed of the following set of coupled submachines:

$$\tilde{x}_{\ell}^i(k+1) = \lambda_i \tilde{x}_{\ell}^i(k) + \tilde{x}_{\ell+1}^i(k) + \sum_{j=1}^m \tilde{b}_{sj}^{(i)} u_j(k) \quad (7.2.37)$$

$$\ell, s \in \underline{n_i - 1}, i \in \underline{v}$$

$$\tilde{x}_{n_i}^i(k+1) = \lambda_i \tilde{x}_{n_i}^i(k) + \sum_{j=1}^m \tilde{b}_{n_i, j}^{(i)} u_j(k), i \in \underline{v} \quad (7.2.38)$$

From (7.2.37) and (7.2.38) it is clear that the  $j$ th input component  $u_j(k)$  can influence the  $\ell$ th state component  $x_{\ell}^i(k)$  if and only if

$$\begin{aligned} \tilde{b}_{\ell j}^{(i)} &\neq 0, \ell \in \underline{n_i} \\ \tilde{b}_{sj}^{(i)} &= 0, s = \ell+1, \ell+2, \dots, n_i; i \in \underline{v}, j \in \underline{m} \end{aligned} \quad (7.2.39)$$

We formalize this observation in the following theorem.

**Theorem 7.2.5.** If in the Jordan canonical form of the LSM  $M = (A, B)$  confluent eigenvalues are associated with distinct blocks, then the  $\ell$ th state component  $x_{\ell}^i(k)$ ,  $\ell \in \underline{n_i}$ ,  $i \in \underline{v}$ , is selectively reachable by the  $j$ th input component  $u_j(k)$  if and only if (7.2.39) holds, where  $\tilde{b}_{\ell j}^{(i)}$  is defined by (7.2.35).

Let the generalized eigenvector matrix  $V_3$  be partitioned as follows:

$$V_3 \equiv [V_{31} \quad V_{32} \quad \dots \quad V_{3v}] \quad (7.2.40)$$

where each  $V_{3j} \in GF(q)^{n \times n_j}$  has the form

$$V_{3j} \equiv [v_1^{(j)}, v_2^{(j)}, \dots, v_{n_i}^{(i)}]; i, j \in \underline{v} \quad (7.2.41)$$



In view of the relation  $B = V_3 \tilde{B}$ , the first column  $b^j$  of the selective state reachability matrix  $K_j$  can be expressed as

$$b^j = \sum_{i=1}^v \sum_{t=1}^{n_i} \tilde{b}_{tj}^{(i)} v_t^{(i)}, \quad j \in \underline{m} \quad (7.2.42)$$

By virtue of the relation  $AV_3 = V_3 \tilde{A}$ , for the present situation the following eigenvalue-eigenvector equations hold:

$$\left. \begin{aligned} Av_1^{(r)} &= \lambda_r v_1^{(r)} \\ Av_i^{(r)} &= \lambda_r v_i^{(r)} + v_{i-1}^{(r)} \end{aligned} \right\} \begin{array}{l} r \in \underline{v} \\ i = 2, 3, \dots, v \end{array} \quad (7.2.43)$$

where  $v_i^{(r)}$  is the generalized eigenvector associated with the  $r$ th eigenvalue  $\lambda_r$  and the  $i$ th subblock  $\tilde{A}_i(\lambda_r)$ .

Using (7.2.42), we obtain the following expression for the second column of  $K_j$ :

$$\begin{aligned} Ab^j &= \sum_{i=1}^v \sum_{\ell=1}^{n_i} \tilde{b}_{\ell j}^{(i)} Av_{\ell}^{(i)} \\ &= \sum_{i=1}^v \tilde{b}_{1j}^{(i)} Av_1^{(i)} + \sum_{i=1}^v \tilde{b}_{2j}^{(i)} Av_2^{(i)} \\ &\quad + \sum_{i=1}^v \tilde{b}_{3j}^{(i)} Av_3^{(i)} + \dots + \sum_{i=1}^v \tilde{b}_{n_i, j}^{(i)} Av_{n_i}^{(i)} \end{aligned} \quad (7.2.44)$$

Using (7.2.43), (7.2.44) becomes

$$\begin{aligned}
Ab^j &= \sum_{i=1}^v \tilde{b}_{1j}^{(i)} \lambda_i v_1^{(i)} + \left\langle \sum_{i=1}^v \tilde{b}_{2j}^{(i)} \lambda_i v_2^{(i)} + \sum_{i=1}^v \tilde{b}_{2j}^{(i)} v_1^{(i)} \right\rangle \\
&\quad + \left\langle \sum_{i=1}^v \tilde{b}_{3j}^{(i)} \lambda_i v_3^{(i)} + \sum_{i=1}^v \tilde{b}_{3j}^{(i)} v_2^{(i)} \right\rangle \\
&\quad + \dots + \left\langle \sum_{i=1}^v \tilde{b}_{n_i, j}^{(i)} \lambda_i v_{n_i}^{(i)} + \sum_{i=1}^v \tilde{b}_{n_i, j}^{(i)} v_{n_i-1}^{(i)} \right\rangle \\
Ab^j &= \left\langle \sum_{i=1}^v \tilde{b}_{1j}^{(i)} \lambda_i v_1^{(i)} + \sum_{i=1}^v \tilde{b}_{2j}^{(i)} \lambda_i v_2^{(i)} + \sum_{i=1}^v \tilde{b}_{3j}^{(i)} \lambda_i v_3^{(i)} \right. \\
&\quad \left. + \dots + \sum_{i=1}^v \tilde{b}_{n_i, j}^{(i)} \lambda_i v_{n_i}^{(i)} \right\rangle \\
&\quad + \left\langle \sum_{i=1}^v \tilde{b}_{2j}^{(i)} v_1^{(i)} + \sum_{i=1}^v \tilde{b}_{3j}^{(i)} v_2^{(i)} + \dots + \sum_{i=1}^v \tilde{b}_{n_i, j}^{(i)} v_{n_i-1}^{(i)} \right\rangle \\
Ab^j &= \sum_{i=1}^v \sum_{t=0}^{n_i} \tilde{b}_{tj}^{(i)} \lambda_i v_t^{(i)} + \sum_{i=1}^v \sum_{t=2}^{n_i} \tilde{b}_{tj}^{(i)} v_{t-1}^{(i)} \tag{7.2.45}
\end{aligned}$$

In a similar manner, we can show that

$$\begin{aligned}
A^2 b^j &= \sum_{i=1}^v \sum_{t=1}^{n_i} \tilde{b}_{tj}^{(i)} (\lambda_i)^2 v_t^{(i)} + 2 \sum_{i=1}^v \sum_{t=2}^{n_i} \tilde{b}_{tj}^{(i)} \lambda_i v_{t-1}^{(i)} \\
&\quad + \sum_{i=1}^v \sum_{t=3}^{n_i} \tilde{b}_{tj}^{(i)} v_{t-2}^{(i)} \tag{7.2.46}
\end{aligned}$$

Now we are ready to indicate the form of the decomposition for the first three columns of  $K_j$  as follows:

$$[b^j, Ab^j, A^2 b^j] = \left\{ \begin{aligned} & \sum_{i=1}^v \sum_{t=1}^{n_i} b_{tj}^{(i)} v_t^{(i)}, \quad \sum_{i=1}^v \sum_{t=1}^{n_i} b_{tj}^{(i)} \lambda_i v_t^{(i)} + \sum_{i=1}^v \sum_{t=2}^{n_i} b_{tj}^{(i)} v_{t-1}^{(i)}, \\ & \sum_{i=1}^v \sum_{t=1}^{n_i} b_{tj}^{(i)} (\lambda_i)^2 v_t^{(i)} + 2 \sum_{i=1}^v \sum_{t=2}^{n_i} b_{tj}^{(i)} \lambda_i v_{t-1}^{(i)} \\ & + \sum_{i=1}^v \sum_{t=3}^{n_i} b_{tj}^{(i)} v_{t-2}^{(i)} \end{aligned} \right\}$$

$$[b^j, Ab^j, A^2 b^j] =$$

$$\left( v_1^{(1)}, v_2^{(1)}, \dots, v_{n_1}^{(1)}; v_1^{(2)}, v_2^{(2)}, \dots, v_{n_2}^{(2)}; \dots; v_1^{(v)}, v_2^{(v)}, \dots, v_{n_v}^{(v)} \right) \times$$

$$\begin{pmatrix} \begin{array}{ccc|c} \tilde{b}_{1j}^{(1)} & \tilde{b}_{2j}^{(1)} & \tilde{b}_{3j}^{(1)} & \\ \tilde{b}_{2j}^{(1)} & \tilde{b}_{3j}^{(1)} & \tilde{b}_{4j}^{(1)} & \\ & & \vdots & \\ & & \vdots & \\ & & \tilde{b}_{n_1,j}^{(1)} & \\ & \cdot & & \\ & \tilde{b}_{n_1,j}^{(1)} & 0 & \\ \cdot & & & \\ \cdot & & & \\ \tilde{b}_{n_1,j}^{(1)} & 0 & 0 & \end{array} & \begin{array}{ccc|c} 1 & \lambda_1 & (\lambda_1)^2 & \\ 0 & 1 & 2\lambda_1 & \\ & & \vdots & \\ & & \vdots & \\ & & \tilde{b}_{n_1,j}^{(1)} & \\ & \cdot & & \\ & \cdot & & \\ & \cdot & & \\ & 0 & 0 & 0 & \\ \hline & \cdot & & \\ & \cdot & & \\ & \cdot & & \\ & \cdot & & \\ \hline & \tilde{b}_{1j}^{(v)} & \tilde{b}_{2j}^{(v)} & \tilde{b}_{3j}^{(v)} & 0 & \lambda_v & (\lambda_v)^2 \\ & \tilde{b}_{2j}^{(v)} & \tilde{b}_{3j}^{(v)} & \tilde{b}_{4j}^{(v)} & 0 & 1 & 2\lambda_v \\ & & & \vdots & & & \\ & & & \vdots & & & \\ & & & \tilde{b}_{n_v,j}^{(v)} & & & \\ & & \cdot & & & & \\ & & \tilde{b}_{n_v,j}^{(v)} & 0 & & \cdot & \cdot & \cdot \\ & \cdot & & & & \cdot & \cdot & \cdot \\ & \cdot & & & & \cdot & \cdot & \cdot \\ & \tilde{b}_{n_v,j}^{(v)} & 0 & 0 & & 0 & 0 & 0 \end{array} \end{pmatrix}$$

(7.2.47)

Proceeding as above, we can show that, in general, the selective state reachability matrix  $K_j$  can be written as a product of three matrices as follows:

$$K_j = V_3 D_{3j} W_3, \quad j \in \underline{m} \quad (7.2.48)$$

where

$$D_{3j} \equiv D_j^{(1)} \oplus D_j^{(2)} \oplus \dots \oplus D_j^{(v)}, \quad j \in \underline{m} \quad (7.2.49)$$

$$D_j^{(i)} \equiv \begin{pmatrix} \tilde{b}_{1j}^{(i)} & \tilde{b}_{2j}^{(i)} & \dots & \tilde{b}_{n_i,j}^{(i)} \\ \tilde{b}_{2j}^{(i)} & \tilde{b}_{3j}^{(i)} & \dots & 0 \\ & \vdots & & \vdots \\ & \tilde{b}_{n_i,j}^{(i)} & \dots & 0 \\ \tilde{b}_{n_i,j}^{(i)} & 0 & \dots & 0 \end{pmatrix}, \quad i \in \underline{v}, \quad j \in \underline{m} \quad (7.2.50)$$

$$W_3 \equiv \begin{pmatrix} W_1 \\ W_2 \\ \vdots \\ W_v \end{pmatrix} \quad (7.2.51)$$

$$W_i \equiv \begin{pmatrix} 1 & \begin{pmatrix} 1 \\ 0 \end{pmatrix} \lambda_i & \begin{pmatrix} 2 \\ 0 \end{pmatrix} (\lambda_i)^2 & \begin{pmatrix} 3 \\ 0 \end{pmatrix} (\lambda_i)^3 & \dots & \begin{pmatrix} n-1 \\ 0 \end{pmatrix} (\lambda_i)^{n-1} \\ 0 & 1 & \begin{pmatrix} 2 \\ 1 \end{pmatrix} \lambda_i & \begin{pmatrix} 3 \\ 1 \end{pmatrix} (\lambda_i)^2 & \dots & \begin{pmatrix} n-1 \\ 1 \end{pmatrix} (\lambda_i)^{n-2} \\ 0 & 0 & 1 & \begin{pmatrix} 3 \\ 2 \end{pmatrix} \lambda_i & \dots & \begin{pmatrix} n-1 \\ 2 \end{pmatrix} (\lambda_i)^{n-3} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \dots & \begin{pmatrix} n-1 \\ n_i-1 \end{pmatrix} (\lambda_i)^{n-n_i} \end{pmatrix} \in GF(q)^{n_i \times n}, i \in \underline{v}$$

(7.2.52)

From the form of  $D_{3j}$  it is clear that this matrix is nonsingular.

Since  $V_3$  is also nonsingular, from (7.2.48) it follows that

$$\text{rank } K_j = \text{rank } D_{3j}, j \in \underline{m}$$

If we define  $\text{rank } D_j^{(i)} \equiv m_{ij}$ , then

$$\text{rank } K_j = \sum_{i=1}^v m_{ij} \quad (7.2.53)$$

where

$$\tilde{b}_{m_{ij},j}^{(i)} \neq 0, m_{ij} \in \underline{n_i} \quad (7.2.54)$$

$$\tilde{b}_{tj}^{(i)} = 0, t = m_{ij}+1, m_{ij}+2, \dots, n_i; i \in \underline{v}, j \in \underline{m}$$

From the Theorem 7.2.5 it follows that  $m_{ij}$  is equal to the number of state components that are selectively reachable by  $u_j(k)$ ,  $j \in \underline{m}$ .

Therefore, we can state the following theorem.

Theorem 7.2.6. If in the Jordan canonical form of  $A$  confluent eigenvalues are associated with distinct Jordan blocks, then the rank of the selective state reachability matrix  $K_j \equiv [b^j, Ab^j, \dots, A^{n-1}b^j]$  is equal to the number of state components that are selectively reachable by the  $j$ th input component  $u_j(k)$ ,  $j \in \underline{m}$ .

Next, we want to show how the selectively reachable state components can be identified. In view of (7.2.54), (7.2.42) can be expressed as

$$b^j = \sum_{i=1}^v \sum_{t=1}^{m_{ij}} \tilde{b}_{tj}^{(i)} v_t^{(i)}, \quad j \in \underline{m} \quad (7.2.55)$$

From the general form of the matrix  $K_j$  given by (7.2.48), it follows that an arbitrary column  $A^s b^j$  can be written as

$$A^s b^j = \sum_{\ell=1}^v \sum_{i=1}^{m_{ij}} \sum_{t=1}^{m_{ij}-1} \begin{pmatrix} s \\ t \end{pmatrix} \tilde{b}_{i+t,j}^{(\ell)} (\lambda_i)^{s-t} v_i^{(\ell)} \quad (7.2.56)$$

On the other hand, we observe from (7.2.53) that any column  $A^h b^j$ ,  $h \geq m_j$ , can be expressed as a linear combination of the  $m_j$  linearly independent columns  $A^s b^j$ ,  $s \in \underline{m_j-1}$ , of the matrix  $K_j$ , where  $m_j \equiv \text{rank } K_j = \sum_{i=1}^v m_{ij}$ . In particular, we have

$$A^{m_j} b^j = \sum_{s=0}^{m_j-1} a_{sj} A^s b^j, \quad a_{sj} \in GF(q)$$

or equivalently,

$$\sum_{s=0}^{m_j} a_{sj} A^s b^j = 0, \quad a_{m_j, j} = -1 \quad (7.2.57)$$

Substituting (7.2.56) into (7.2.57) and interchanging the order of summations, we obtain

$$\sum_{i=1}^{m_{ij}} \left( \sum_{\ell=1}^v \sum_{s=0}^{m_j} \sum_{t=1}^{m_{ij}-1} a_{sj} \binom{s}{t} \tilde{b}_{i+t, j}^{(\ell)} (\lambda_i)^{s-t} \right) v_i^{(\ell)} = 0$$

Independence of the generalized eigenvectors  $v_i^{(\ell)}$  implies that

$$\tilde{b}_{m_{ij}, j}^{(i)} (\lambda_i)^{m_j} = \tilde{b}_{m_{ij}, j}^{(i)} \sum_{s=0}^{m_j-1} a_{sj} (\lambda_i)^s$$

or

$$(\lambda_i)^{m_j} = \sum_{s=0}^{m_j-1} a_{sj} (\lambda_i)^s, \quad i \in \underline{v}, \quad j \in \underline{m} \quad (7.2.58)$$

since  $\tilde{b}_{m_{ij}, j}^{(i)} \neq 0$ .

Therefore, as in previous cases, equation (7.2.57) can be used to determine the constants  $a_{sj}$ ,  $s \in \underline{m_j-1}$ ,  $j \in \underline{m}$ , and the polynomial (7.2.58) may be solved to identify the selectively reachable state components.

Case 4. Confluent eigenvalues of  $A$  are associated with a number of nondistinct Jordan blocks of  $\tilde{A}$ .

Finally, we will consider the most general case in which the isomorphic LSM  $\tilde{M} = (\tilde{A}, \tilde{B})$  has the form given by (7.1.9) - (7.1.14). The steps required for the analysis of this case are, in principle, identical to those taken in the analysis of the previous cases.



However, the notation is somewhat more involved in the present situation. For the sake of completeness, we will, for the last time, repeat our routine analysis.

From the set of scalar state-input equations (7.1.13) it is easily seen that the state components  $x_{\ell s}^{(i)}$ ,  $\ell \in \underline{n}_{is}$ ,  $s \in \mu(i)$ ,  $i \in \underline{v}$ , can be selectively affected by the  $j$ th input component  $u_j(k)$ ,  $j \in \underline{m}$ , if and only if

$$\tilde{b}_{\ell s j}^{(i)} \neq 0, \ell \in \underline{n}_{is}$$

$$\tilde{b}_{tsj}^{(i)} = 0, t = \ell+1, \ell+2, \dots, n_{is}; s \in \mu(i), i \in \underline{v}, j \in \underline{m}$$

We formalize this observation in the following theorem.

Theorem 7.2.7. If in the Jordan canonical form of  $A$  confluent eigenvalues are associated with a number of nondistinct Jordan blocks, then the  $\ell$ th state component in the  $s$ th subblock of the  $i$ th block,  $x_{\ell s}^{(i)}(k)$ ,  $\ell \in \underline{n}_{is}$ ,  $s \in \mu(i)$ ,  $i \in \underline{v}$ , is selectively reachable by the  $j$ th input component  $u_j(k)$ ,  $j \in \underline{m}$ , if and only if (7.2.59) holds, where  $\tilde{b}_{\ell s j}^{(i)}$  is defined by (7.1.12).

In order to develop a procedure for determining the number of selectively reachable state components and a method for identifying these components, we will imitate the corresponding sequence of steps used in the previous three cases.

Let the matrix  $V_4$  of generalized eigenvectors be partitioned as follows:

$$V_4 \equiv \left( v_1^{(1)}, v_2^{(1)}, \dots, v_{\mu(1)}^{(1)}; v_1^{(2)}, v_2^{(2)}, \dots, v_{\mu(2)}^{(2)}; \right. \\ \left. \dots; v_1^{(\nu)}, v_2^{(\nu)}, \dots, v_{\mu(\nu)}^{(\nu)} \right) \quad (7.2.60)$$

where

$$v_s^{(i)} \equiv \left( v_{s1}^{(i)}, v_{s2}^{(i)}, \dots, v_{s n_{is}}^{(i)} \right), \quad s \in \underline{\mu(i)}, \quad i \in \underline{\nu} \quad (7.2.61)$$

In terms of the above notation and in view of the relation  $AV_4 = V_4 \tilde{A}$ , we have the following set of eigenvalue-eigenvector equations:

$$Av_{s1}^{(i)} = \lambda_i v_{s1}^{(i)} \\ Av_{s\ell}^{(i)} = \lambda_i v_{s\ell}^{(i)} + v_{s,\ell-1}^{(i)} \quad (7.2.62)$$

$$\ell \in \underline{n_{is}}, \quad s \in \underline{\mu(i)}, \quad i \in \underline{\nu}$$

Using the relation  $B = V_4 \tilde{B}$ , we can express the columns  $A^h b^j$ ,  $h \in \underline{n-1}$ ,  $j \in \underline{m}$ , of the selective reachability matrix  $K_j$  as follows:

$$b^j = \sum_{i=1}^{\nu} \sum_{s=1}^{\mu(i)} \sum_{\ell=1}^{n_{is}} \tilde{b}_{\ell sj}^{(i)} v_{s\ell}^{(i)} \quad (7.2.63)$$

$$Ab^j = \sum_{i=1}^{\nu} \sum_{s=1}^{\mu(i)} \sum_{\ell=1}^{n_{is}} \tilde{b}_{\ell sj}^{(i)} (Av_{s\ell}^{(i)})$$

Substituting for  $Av_{s\ell}^{(i)}$  from (7.2.62) and simplifying, we obtain

$$\begin{aligned}
Ab^j = & \sum_{i=1}^v \sum_{s=1}^{\mu(i)} \sum_{\ell=1}^{n_{is}} b_{\ell sj}^{(i)} \lambda_i v_{s\ell}^{(i)} \\
& + \sum_{i=1}^v \sum_{s=1}^{\mu(i)} \sum_{\ell=2}^{n_{is}} b_{\ell sj}^{(i)} v_{s,\ell-1}^{(i)}, \quad j \in \underline{m}
\end{aligned} \tag{7.2.64}$$

Similarly, it can be shown that

$$\begin{aligned}
A^2 b^j = & \sum_{i=1}^v \sum_{s=1}^{\mu(i)} \sum_{\ell=1}^{n_{is}} b_{\ell sj}^{(i)} (\lambda_i)^2 v_{s\ell}^{(i)} \\
& + 2 \sum_{i=1}^v \sum_{s=1}^{\mu(i)} \sum_{\ell=2}^{n_{is}} b_{\ell sj}^{(i)} \lambda_i v_{s,\ell-1}^{(i)} \\
& + \sum_{i=1}^v \sum_{s=1}^{\mu(i)} \sum_{\ell=3}^{n_{is}} b_{\ell sj}^{(i)} v_{s,\ell-2}^{(i)}, \quad j \in \underline{m}
\end{aligned} \tag{7.2.65}$$

Now using (7.2.63) - (7.2.65), the first three columns  $b^j$ ,  $Ab^j$ , and  $A^2 b^j$  of  $K_j$  can be written in matrix notation as follows:

$$[b^j, Ab^j, A^2 b^j] = v_4 \hat{D}_j \hat{W} \tag{7.2.66}$$

where

$$\hat{D}_j \equiv \sum_{i=1}^v \sum_{s=1}^{\mu(i)} \oplus \hat{D}_{sj}^{(i)} \tag{7.2.67}$$

$$\hat{D}_{sj}^{(i)} \equiv \begin{pmatrix} \tilde{b}_{1sj}^{(i)} & \tilde{b}_{2sj}^{(i)} & \tilde{b}_{3sj}^{(i)} \\ \tilde{b}_{2sj}^{(i)} & \tilde{b}_{3sj}^{(i)} & \tilde{b}_{4sj}^{(i)} \\ & & \vdots \\ & & \vdots \\ & \cdot & \tilde{b}_{n_{is},sj}^{(i)} \\ & \cdot & \\ & \cdot & \\ \cdot & \tilde{b}_{n_{is},sj}^{(i)} & 0 \\ \cdot & & \\ \cdot & & \\ \tilde{b}_{n_{is},sj}^{(i)} & 0 & 0 \end{pmatrix}, \quad s \in \underline{\mu(i)}, \quad i \in \underline{\nu}, \quad j \in \underline{m}$$

(7.2.68)

$$\hat{W} \equiv \begin{pmatrix} \hat{W}_1^{(1)} \\ \hat{W}_2^{(1)} \\ \cdot \\ \cdot \\ \cdot \\ \hat{W}_{\mu(1)}^{(1)} \\ \cdot \\ \cdot \\ \cdot \\ \hat{W}_{\mu(\nu)}^{(\nu)} \\ \hat{W}_{\mu(\nu)}^{(\nu)} \end{pmatrix}$$

(7.2.69)

$$\hat{W}_s^{(i)} \equiv \begin{bmatrix} 1 & \lambda_i & (\lambda_i)^2 \\ 0 & 1 & 2\lambda_i \\ 0 & 0 & 0 \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ 0 & 0 & 0 \end{bmatrix}, s \in \mu(i), i \in \underline{v} \quad (7.2.70)$$

If we continue to generate the remaining columns  $A^{t_b j}$ ,  $t = 3, 4, \dots, n-1$ ,  $j \in \underline{m}$ , in the above manner, it will be seen that for the matrix  $K_j$  the pattern exhibited by (7.2.66) - (7.2.70) can be generalized and thus  $K_j$  can be expressed as a product of three matrices as follows:

$$K_j \equiv [b^j, Ab^j, \dots, A^{n-1} b^j] = v_4 D_{4sj}^{(i)} w_4 \quad (7.2.71)$$

where

$$\begin{aligned} D_{4sj}^{(i)} \equiv & D_{1j}^{(1)} \oplus D_{2j}^{(1)} \oplus \dots \oplus D_{\mu(1)j}^{(1)} \\ & \oplus D_{1j}^{(2)} \oplus D_{2j}^{(2)} \oplus \dots \oplus D_{\mu(2)j}^{(2)} \\ & \oplus \dots \oplus D_{1j}^{(v)} \oplus D_{2j}^{(v)} \oplus \dots \oplus D_{\mu(v)j}^{(v)}, j \in \underline{m} \end{aligned} \quad (7.2.72)$$

$$D_{sj}^{(i)} \equiv \begin{pmatrix} \tilde{b}_{1sj}^{(i)} & \tilde{b}_{2sj}^{(i)} & \dots & \tilde{b}_{n_{is},sj}^{(i)} \\ \tilde{b}_{2sj}^{(i)} & \tilde{b}_{3sj}^{(i)} & \dots & 0 \\ & \cdot & & \cdot \\ & \cdot & & \cdot \\ \cdot & \tilde{b}_{n_{is},sj}^{(i)} & \dots & 0 \\ \cdot & & & \\ \tilde{b}_{n_{is},sj}^{(i)} & 0 & \dots & 0 \end{pmatrix}, \quad i \in \underline{v}, \quad s \in \mu(i)$$

(7.2.73)

$$W_4 \equiv \begin{pmatrix} w_1^{(1)} \\ w_2^{(1)} \\ \cdot \\ \cdot \\ \cdot \\ w_{\mu(1)}^{(1)} \\ \cdot \\ \cdot \\ \cdot \\ w_{\mu(v)}^{(v)} \\ w_{\mu(v)}^{(v)} \end{pmatrix}$$

(7.2.74)

$$W_s^{(i)} = \begin{pmatrix} 1 & \begin{pmatrix} 1 \\ 0 \end{pmatrix} \lambda_i & \begin{pmatrix} 2 \\ 0 \end{pmatrix} (\lambda_i)^2 & \begin{pmatrix} 3 \\ 0 \end{pmatrix} (\lambda_i)^3 & \dots & \begin{pmatrix} n-1 \\ 0 \end{pmatrix} (\lambda_i)^{n-1} \\ 0 & 1 & \begin{pmatrix} 2 \\ 1 \end{pmatrix} \lambda_i & \begin{pmatrix} 3 \\ 1 \end{pmatrix} (\lambda_i)^2 & \dots & \begin{pmatrix} n-1 \\ 1 \end{pmatrix} (\lambda_i)^{n-2} \\ 0 & 0 & 1 & \begin{pmatrix} 3 \\ 2 \end{pmatrix} \lambda_i & \dots & \begin{pmatrix} n-1 \\ 2 \end{pmatrix} (\lambda_i)^{n-3} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \dots & \begin{pmatrix} n-1 \\ n_{is}-1 \end{pmatrix} (\lambda_i)^{n-n_{is}} \end{pmatrix}$$

$$s \in \underline{\mu(i)}, i \in \underline{\nu} \quad (7.2.75)$$

From the form of the matrix  $W_4$  it is evident that it is nonsingular. Since the matrix  $V_4$  of generalized eigenvectors is also nonsingular, from (7.2.71) it follows that

$$\begin{aligned} \text{rank } K_j &= \text{rank } D_{4sj}^{(i)} = \sum_{i=1}^{\nu} \sum_{s=1}^{\mu(i)} \text{rank } D_{sj}^{(i)} \\ &= \sum_{i=1}^{\nu} \sum_{s=1}^{\mu(i)} m_{isj}, j \in \underline{m} \end{aligned} \quad (7.2.76)$$

where

$$\tilde{b}_{m_{isj},sj}^{(i)} \neq 0, m_{isj} \in \underline{n_{is}}$$

$$\tilde{b}_{tsj}^{(i)} = 0, t = m_{isj}+1, m_{isj}+2, \dots, n_{is}; s \in \underline{\mu(i)}, i \in \underline{\nu}$$

The above conclusion together with Theorem 7.2.7 lead to the following result.

Theorem 7.2.8. If in the Jordan canonical form  $\tilde{A}$  of  $A$ , confluent eigenvalues are associated with a number of nondistinct Jordan blocks, then the rank of the selective state reachability matrix  $K_j$  of the LSM  $(A, B)$  is equal to the number of state components that are selectively reachable by the  $j$ th input component  $u_j(k)$ ,  $j \in \underline{m}$ .

The procedure for the identification of selectively reachable state components characterized in Theorem 7.2.8 can be derived in precisely the same manner as in the previous cases.

### Summary and Conclusions

Throughout this chapter, the Jordan canonical form of the characteristic matrix of the LSM  $M = (A, B)$  played a central role in the investigation of various aspects of the state reachability property of  $M$ . Exploiting the generalized eigenproperties of  $M$ , additional state reachability criteria in terms of the Jordan canonical representation of  $M$  were formulated. Furthermore, the possibility of controlling a multi-input LSM by a scalar control sequence was characterized in terms of its Jordan canonical form (cf. [21], [27], [37], [58], [61], [65], [99], [107], [118]).

Also in this chapter, the concept of selective state reachability for LSMs was introduced. Consideration of this idea for LSMs was motivated in part by the somewhat analogous notion of modal analysis in conventional linear control systems [92], [99]. Since due to lack of order in  $GF(q)$  the eigenvalues of an LSM, in contrast to those of a conventional linear system, cannot be related to any physical quantities such as power, state evolution modes, etc., the concept of modal



control was reinterpreted as selective state reachability for LSMs. The generalized eigenproperties of LSMs in the framework of their Jordan canonical forms were utilized in the investigation of selective state reachability properties of LSMs (cf. [92], [99], [121]).

## CHAPTER VIII

## PROJECTIVE-GEOMETRIC STRUCTURES AND LSMS

In early 1970, Wonham and Morse [111], [112], and independently Basile and Marro [5], [6], [7] introduced a new approach for the study of linear continuous-time systems, which makes heavy use of the abstract geometrical structures of finite-dimensional vector spaces over the fields of real and complex numbers. Generally speaking, the essence of the geometric approach is to first characterize solvability of the problem of interest as a verifiable property of some subspace of the state or output space of the linear system under consideration and then translate the subspace solution into matrix operations [114]. Wonham and Morse explored their geometric theory extensively and with considerable success initially in the areas of decoupling and pole assignment which have been problems of longstanding interest in the area of linear systems. Their formulation of the decoupling problems relied heavily on the concept of a generalized controllability subspace, which is a subspace that satisfies certain restrictive conditions. Solvability of decoupling problems then became equivalent to finding suitable sets of controllability subspaces. They obtained powerful and general criteria for decoupling by state variable feedback that subsumed most of the earlier results. Later these authors and others applied the Wonham-Morse-Basile-Marro geometric theory to many other aspects of linear systems such as disturbance localization, decoupling

by output feedback, stabilization, tracking and regulation, feedback invariants and canonical forms, dynamic observer design, decentralized control, and so forth.

Although the computational efficiency of the geometric method is debatable at this stage of its development, its elegance and generality are certainly very appealing.

Some aspects of the Wonham-Morse-Basile-Marro geometric system theory are essentially coordinate-free and hence certain portions of it remain valid on arbitrary fields and, in particular, over the finite field  $\text{GF}(q)$ . Therefore, a geometric theory can be developed for the investigation of certain structural properties of LSMs. In fact, such a theory is motivated by many important interrelationships existing between LSMs and coding theory. A linear code by definition is a subspace of  $\text{GF}(q)^n$ . In the area of coding and decoding theory highly geometric concepts have already been employed for the development of certain classes of codes which are primarily based on the properties of finite affine and projective geometries and their associated combinatorial structures. On the other hand, close relationships have been discovered between the burst correction properties of convolutional codes and the controllability and observability properties of LSMs [77]. In view of these relationships and the central role played by LSMs in coding, decoding, and other significant computational tasks on finite fields, there seems to be ample justification and incentive for exploring the possibility of developing a geometric theory of the Wonham-Morse-Basile-Marro type for LSMs. Therefore, we intend to initiate

the research in this direction by providing the rudiments of a projective-geometric theory for LSMs in the present chapter. The guiding source for the material in this chapter is the research monograph by Wonham [114]. However, it should be borne in mind that we are concerned with a different mathematical framework.

In this chapter familiarity with the basic concepts, terminology, and notation concerning finite affine and projective geometries, provided in the Appendix, is needed.

### 8.1. Geometric Definition of State Reachability of LSMs

Our purpose in this section is to present a geometric definition of reachability of an LSM  $M = (A, B)$  and then discuss some related results that are of a geometric nature.

We recall from Section 4.1 that the reachable flat

$$\begin{aligned}
 R(K) &= R([B, AB, A^2B, \dots, A^{n-1}B]) \\
 &= R(B) + AR(B) + A^2R(B) + \dots + A^{n-1}R(B) \\
 &\equiv \{A \mid R(B)\}
 \end{aligned} \tag{8.1.1}$$

of an LSM  $M = (A, B)$  is the smallest  $A$ -invariant flat in  $P(X)$  that contains  $R(B)$ . Now let  $P : P(X) \longrightarrow A(x + R(K))$ ,  $\langle x \rangle \longmapsto P \langle x \rangle \equiv \langle \bar{x} \rangle$ , be the canonical projection. Since  $R(K) \subseteq N(P)$ , we obtain the following autonomous representation for the originally nonautonomous LSM  $M = (A, B)$ :

$$\bar{x}(k+1) = \bar{A} \bar{x}(k) \tag{8.1.2}$$

where  $\bar{A}$  is the map induced in  $A(x + R(K))$  by  $A$ . From (8.1.2) it is clear that the control sequence  $u(k) \in U^*$  has no influence on the coset of  $x(k) \bmod R(K)$ . This shows that a necessary and sufficient condition for the reachability of all states of an LSM from the zero state is that  $A(x + R(K)) = 0$ , that is,  $P(X) = P(R(K))$ . We will use this characterization as the definition of state reachability.

Theorem 8.1.1. The LSM  $M = (A, B)$  is state reachable if and only if for any given irreducible polynomial  $g \in GF(q)[\lambda]$  of degree  $n$ , there exists a state feedback homomorphism  $F : P(X) \rightarrow P(U)$  such that the characteristic polynomial of  $A + BF$  is precisely  $g$ .

In order to be able to prove this theorem, we will need some auxiliary results.

Lemma 8.1.1. If  $S \in P_A(X)$ , then  $A$  induces an affinity  $\bar{A} : A(x + S) \rightarrow A(x + S)$  defined by  $\bar{A}(x + S) \equiv Ax + S$ . Moreover, if  $A$  is the zero of any polynomial, then so is  $\bar{A}$ . Thus the minimal polynomial of  $\bar{A}$  divides the minimal polynomial of  $A$ .

Proof. First we need to show that  $\bar{A}$  is well-defined, that is, if  $x' + S = x'' + S$ , then  $\bar{A}(x' + S) = \bar{A}(x'' + S)$ . If  $x' + S = x'' + S$ , then  $x' - x'' \in S$  and, since  $S \in P_A(X)$ ,  $A(x' - x'') \in S$ . Hence  $\bar{A}(x' + S) = Ax' + S = Ax'' + S$ . We next show that  $\bar{A}$  is linear. Let  $x', x'' \in X$  and  $a \in GF(q)$ . Then we have

$$\begin{aligned} A((x' + S) + (x'' + S)) &= \bar{A}(x' + x'' + S) = \bar{A}(x' + x'') + S \\ &= \bar{A}x' + \bar{A}x'' + S \\ &= \bar{A}x' + S + \bar{A}x'' + S \\ &= \bar{A}(x' + S) + \bar{A}(x'' + S) \end{aligned}$$

and

$$\begin{aligned}\overline{A}(a(x' + S)) &= \overline{A}(ax' + S) = a\overline{A}x' + S \\ &= a(\overline{A}x' + S) = a\overline{A}(x' + S)\end{aligned}$$

Now for any coset  $x' + S \in A(x + S)$ ,

$$\begin{aligned}\overline{A^2}(x' + S) &= A^2x' + S = A(Ax') + S \\ &= \overline{A}(Ax' + S) = \overline{A^2}(x' + S)\end{aligned}$$

Hence  $\overline{A^2} = \overline{A}^2$ . Similarly, we can show that  $\overline{A^n} = \overline{A}^n$  for any integer  $n$ .

Thus for any polynomial  $f(\lambda) = \sum_{i=0}^n a_i(\lambda)^i$ ,

$$\begin{aligned}\overline{f(A)}(x' + S) &= f(A)x' + S = \sum_{i=0}^n a_i A^i x' + S \\ &= \sum_{i=0}^n a_i (A^i x' + S) = \sum_{i=0}^n a_i \overline{A^i}(x' + S) \\ &= \sum_{i=0}^n a_i \overline{A}^i(x' + S) = \left( \sum_{i=0}^n a_i \overline{A}^i \right) (x' + S) \\ &= f(\overline{A})(x' + S)\end{aligned}$$

and so  $\overline{f(A)} = f(\overline{A})$ . Accordingly, if  $A$  is a root of  $f(\lambda)$  then  $\overline{f(A)} = \overline{0} = S = f(\overline{A})$ , that is,  $\overline{A}$  is also a root of  $f(\lambda)$ .  $\square$

Proof of Theorem 8.1.1. The necessity part of the theorem was proved in Theorem 5.3.4 (in fact, this result was proved as property 17<sup>o</sup> of Theorem 6.3.1). To prove sufficiency, first of all notice that we can construct an irreducible polynomial of degree  $n$  since irreducible

polynomials of all degrees  $\geq 2$  exist in  $GF(q)[\lambda]$ . In view of the invariance property of state reachability under the action of a state feedback homomorphism  $F : P(X) \longrightarrow P(U)$  (Theorem 5.3.1), it is clear that  $R(K) \in P_{A+BF}(X)$ . Therefore, by Lemma 8.1.1 the map  $A + BF$  induces an affinity  $\overline{A + BF} : A(x + R(K)) \longrightarrow A(x + R(K))$  defined by  $\overline{A + BF}(x' + R(K)) \equiv (A + BF)x' + R(K)$  such that  $(\overline{A + BF})^n = \overline{(A + BF)^n}$ . Let  $f(\lambda) = \sum_{i=0}^n a_i(\lambda)^i$  be the irreducible characteristic polynomial of  $A + BF$ . Then by the Cayley-Hamilton Theorem,  $f(A + BF) = 0$ . Hence by Lemma 8.1.1 we also have  $f(\overline{A + BF}) = 0$ , that is, the zero map in  $A(x + R(K))$ . Let  $\hat{f}_m(\lambda)$  be the minimal polynomial of  $\overline{A + BF}$ . Then  $\hat{f}_m$  divides  $f$  since  $f(\overline{A + BF}) = 0$ . Since by hypothesis  $f$  is irreducible, either  $\hat{f}_m = 1$  or  $\hat{f}_m = \pm f$ . Since  $B \neq 0$ ,  $\deg \hat{f}_m < n = \deg f$ , so  $\hat{f}_m = 1$ . But  $\hat{f}_m(\overline{A + BF}) = 0$  which means that the identity map on  $A(x + R(K))$  is equal to the zero map and hence  $A(x + R(K)) = P(R(K))$ . Therefore,  $P(R(K)) = P(X)$  and  $M$  is reachable.  $\square$

The next result shows that machine reachability in the projective geometry  $P(X)$  implies reachability in a certain affine geometry.

Theorem 8.1.2. Suppose that  $P(R(K)) = P(X)$ , and let  $S \in P_A(X)$ .

Then

$$\{\overline{A} \mid \overline{R}(B)\} = A(x + S)$$

where  $\overline{A}$  is the map induced by  $A$  in  $A(x + S)$ , and  $\overline{R}(B) \equiv A(z + S)$ ,  $z \in R(B) + S$ .

Proof. Let the homomorphism  $P : P(X) \longrightarrow A(x + S)$  be the canonical projection. Thus  $\overline{R}(B) = PR(B)$  and  $\overline{AP} = PA$ . Then

$$\begin{aligned}
A(x + S) &= P\{A \mid R(B)\} = P\left(\sum_{i=0}^{n-1} A^i R(B)\right) \\
&= \sum_{i=0}^{n-1} \overline{A}^i P R(B) \\
&= \sum_{i=0}^{n-1} \overline{A}^i \overline{R}(B) \\
&\equiv \{\overline{A} \mid \overline{R}(B)\} \square
\end{aligned}$$

The homomorphism relations used in the above theorem are exhibited in the following commutative diagram:

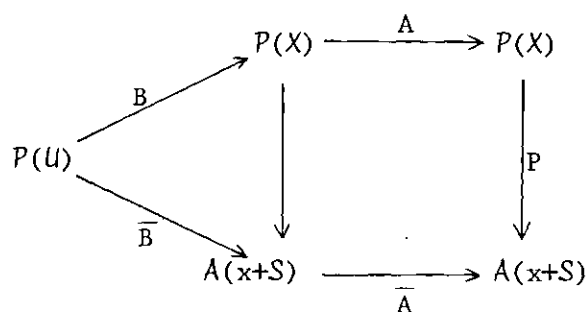


Fig. 8.1.1. Diagram of Homomorphisms for Theorem 8.1.2.

Theorem 8.1.3. Consider the LSM  $(A, B)$  and let  $P(\{A \mid R(B)\}) = P(X)$ . Furthermore, assume that  $S \in P(X)$  decomposes the endomorphism  $A : P(X) \rightarrow P(X)$ , that is,  $S \in P_A(X)$  and there exists a flat  $T \in P_A(X)$  such that  $P(S) \oplus P(T) = P(X)$ . Then

$$P(S) = P(\{A \mid QR(B)\})$$

where  $Q : P(X) \rightarrow P(S)$  is the projection on  $P(S)$  along  $P(T)$ .



Proof. Since  $QA = AQ$ , we have

$$\begin{aligned}
 P(S) &= QP(X) = QP\left(\sum_{i=0}^{n-1} A^i R(B)\right) \\
 &= P\left(\sum_{i=0}^{n-1} QA^i R(B)\right) \\
 &= P\left[\sum_{i=0}^{n-1} A^i (QR(B))\right] \\
 &\equiv P(\{A \mid QR(B)\}) \square
 \end{aligned}$$

## 8.2. (A, B)-Invariant Flats

In this section we will introduce two important dually isomorphic sets of flats of  $P(X)$  and  $P^0(X)$ , and discuss their properties. Later in this chapter these sets will be employed to solve an output invariance problem and some decoupling problems in LSMs.

Consider the following set

$$I(A, B; X) \equiv \{S \in P(X) : AS \subseteq S + R(B)\} \quad (8.2.1)$$

The elements of  $I(A, B; X)$ , or simply  $I(X)$ , will be called *(A, B)-invariant flats*. It is clear that  $I(X)$  contains  $P_A(X)$  as a subset. Furthermore, we notice that if  $R(B) = 0$  or  $R(B) \subseteq S$ , then  $I(X) = P_A(X)$ ; and if  $R(B) = X$  or  $S + R(B) = X$ , then  $I(X) = P(X)$ .

The real importance of  $I(X)$  lies in the fact that its elements can be made  $(A + BF)$ -invariant for a suitable choice of a state feedback homomorphism as shown in the following theorem.

Theorem 8.2.1. Let  $S \in \mathcal{P}(X)$ . Then there exists a state feedback homomorphism  $F : \mathcal{P}(X) \rightarrow \mathcal{P}(U)$  such that  $S \in \mathcal{P}_{A+BF}(X)$  if and only if  $S \in I(X)$ .

Proof. Suppose that there exists a state feedback homomorphism  $F : \mathcal{P}(X) \rightarrow \mathcal{P}(U)$  such that  $S \in \mathcal{P}_{A+BF}(X)$ , and let  $s \in S$ . Then  $(A + BF)s = s'$  for some  $s' \in S$ , or  $As = s' - BF s \in S + R(B)$ . To prove the converse, suppose that  $S \in I(X)$  having the basis  $\{s^1, s^2, \dots, s^\ell\}$ . Then there exist  $z^i \in S$  and  $u^i \in U$ ,  $i \in \underline{\ell}$ , such that  $As^i = z^i - Bu^i$ ,  $i \in \underline{\ell}$ . Now if we define  $F_0 : \mathcal{P}(S) \rightarrow \mathcal{P}(U)$  by  $F_0 s^i \equiv u^i$ ,  $i \in \underline{\ell}$ , and let  $F$  be any extension of  $F_0$  to  $\mathcal{P}(X)$ , then it is clear that  $(A + BF)S \subseteq S$ .  $\square$

It is easy to see that  $I(X)$  is closed under the operation of flat addition since  $S_1, S_2 \in I(X) \implies A(S_1 + S_2) = AS_1 + AS_2 \subseteq S_1 + S_2 + R(B)$  and hence  $S_1 + S_2 \in I(X)$ . However, it can be readily verified that  $S_1, S_2 \in I(X) \not\implies A(S_1 \cap S_2) \subseteq (S_1 \cap S_2) + R(B)$ , that is,  $I(X)$  is not closed under the operation of flat intersection. Therefore, in spite of the fact that  $I(X)$  contains  $\mathcal{P}_A(X)$  as a subset, the totality of elements of  $I(X)$  does not form a projective geometry. Later it will be shown that under a certain restriction  $I(X)$  becomes closed under intersection and hence a finite projective geometry.

Let

$$I^O(A, B; X) \equiv \{T \in \mathcal{P}^O(X) : A(T \cap Z(B)) \subseteq T\} \quad (8.2.2)$$

where  $Z(B)$  is a flat dependent on  $B$ . Clearly  $I^0(A, B; X)$ , or simply  $I^0(X)$ , contains  $P_A^0(X)$  as a subset. Moreover, we observe that if  $Z(B) = X$  or  $Z(B) \subseteq T$ , then  $I^0(X) = P_A^0(X)$ . It is also easy to see that  $I^0(X)$  is closed under the operation of flat intersection since

$$\begin{aligned} T_1, T_2 \in I^0(X) &\implies A[(T_1 \cap T_2) \cap Z(B)] \\ &= [A(T_1 \cap Z(B))] \cap [A(T_2 \cap Z(B))] \\ &\subseteq T_1 \cap T_2 \end{aligned}$$

However,  $T_1, T_2 \in I^0(X) \not\Rightarrow A[(T_1 + T_2) \cap Z(B)] \subseteq T_1 + T_2$ , and hence  $I^0(X)$  is not a projective geometry.

In order to discuss the relationship between the sets  $I(X)$  and  $I^0(X)$ , we need the following lemma whose simple proof is omitted.

Lemma 8.2.1. For any  $S, T \in P(X)$ ,  $AS \subseteq T$  if and only if  $A^T T^\perp \subseteq S^\perp$ .

Theorem 8.2.2. Let  $S \in P(X)$ . If  $AS \subseteq S + R(B)$ , then  $A^T(S^\perp \cap R(B)^\perp) \subseteq S^\perp$ .

Proof. By Lemma 8.2.1,  $AS \subseteq S + R(B)$  implies that  $A^T(S + R(B))^\perp \subseteq S^\perp$ . By the duality of the projective geometries  $P(X)$  and  $P^0(X)$ ,  $(S + R(B))^\perp = S^\perp \cap R(B)^\perp$ . Therefore, the theorem is proved.  $\square$

Theorem 8.2.3. Let  $T \in P(X)$ . If  $A(T \cap Z(B)) \subseteq T$ , then  $A^T T^\perp \subseteq T^\perp + Z(B)^\perp$ .

Proof. By Lemma 8.2.1,  $A(T \cap Z(B)) \subseteq T$  implies that  $A^T T^\perp \subseteq (T \cap Z(B))^\perp$ . By the duality of the projective geometries  $P(X)$  and  $P^0(X)$ ,  $(T \cap Z(B))^\perp = T^\perp + Z(B)^\perp$ , and hence the theorem is proved.  $\square$

Theorem 8.2.4. The sets  $I(X)$  and  $I^0(X)$  are dually isomorphic.

Proof. It follows from the preceding two theorems.  $\square$

In many applications it becomes necessary to determine the supremal or infimal element of a given class of flats of  $P(X)$ . Since the set of all flats of  $P(X)$  are partially ordered by the inclusion relation, a supremal element  $S_{\text{sup}}$  of a given class of flats, say  $I(X)$ , is obviously that element of  $I(X)$  which contains every other member of  $I(X)$ . Thus  $S_{\text{sup}} \in I(X)$  and if  $S \in I(X)$ , then  $S \subseteq S_{\text{sup}}$ . Clearly  $S_{\text{sup}}$  is unique.

The existence of  $S_{\text{sup}} \equiv \sup I(X)$  is rather obvious since  $I(X)$  is closed under addition and, therefore, due to the fact that  $I(X)$  contains a finite number of elements, a strictly ascending chain of elements of  $I(X)$  of the form  $S_1 \subset S_1 + S_2 \subset S_1 + S_2 + S_3 \subset \dots$ ,  $S_i \in I(X)$ , will terminate after a certain number, say  $\ell$ , of terms. Clearly then  $S_{\text{sup}} = S_1 + S_2 + \dots + S_\ell$ . By duality it is clear that the set  $I^0(X)$  always has a unique infimal element  $T_{\text{inf}}$ , namely  $S_{\text{sup}}^\perp = S_1^\perp \cap S_2^\perp \cap \dots \cap S_\ell^\perp$ .

Let  $I(A, B; V)$ , or simply  $I(V)$ , denote the subclass of  $(A, B)$ -invariant flats contained in  $V$ , that is,

$$I(V) \equiv \{S : S \in I(X) \text{ and } S \subseteq V\} \quad (8.2.3)$$

The following iterative scheme generates  $\sup I(V)$ .

Theorem 8.2.5. Let  $V \in P(X)$  and define the sequence  $\{S^j\}$  according to

$$S^0 \equiv V$$

$$S^j \equiv V \cap A^{-1*}(R(B) + S^{j-1}), \quad j \in \underline{n} \quad (8.2.4)$$

where  $A^{-1*}S \equiv \{s : As \in S\}$ . Then  $S^j \subseteq S^{j-1}$  and for some  $\ell \leq n$ ,  $S^\ell = \sup I(V)$ .

Proof. First we show by induction that  $\{S^j\}$  is a nonincreasing sequence. It is clear that  $S^1 \subseteq S^0$ . Suppose that  $S^j \subseteq S^{j-1}$ . Then

$$S^{j+1} = V \cap A^{-1*}(R(B) + S^j) \subseteq V \cap A^{-1*}(R(B) + S^{j-1}) = S^j$$

Therefore, for some  $\ell \leq \dim V$ ,  $S^j = S^\ell$ ,  $j \geq \ell$ . Now from the definition (8.2.3) of  $I(V)$ , it is clear that an arbitrary  $S \in I(V)$  if and only if

$$S \subseteq V, \quad S \subseteq A^{-1*}(R(B) + S) \quad (8.2.5)$$

From (8.2.5),  $S \subseteq S^0$ , and if  $S \subseteq S^{j-1}$ ,

$$S \subseteq V \cap A^{-1*}(R(B) + S) \subseteq V \cap A^{-1*}(R(B) + S^{j-1}) = S^j$$

Therefore,  $S \subseteq S^\ell \in I(V)$  and consequently  $S^\ell = \sup I(V)$ .  $\square$

In a similar manner, we can devise an algorithm to generate the infimal element of the set  $I^0(A, B; W)$ , or simply  $I^0(W)$ , where

$$I^0(W) \equiv \{T : T \in I^0(X) \text{ and } T \supseteq W\} \quad (8.2.6)$$

This is accomplished in the following theorem.

Theorem 8.2.6. Let  $W \in P^0(X)$  and define the sequence  $\{T^j\}$  according to

$$T^0 \equiv W$$

$$T^j \equiv W + A(Z(B) \cap T^{j-1}), \quad j \in \underline{n} \quad (8.2.7)$$

Then  $T^{j-1} \subseteq T^j$  and for some  $r \leq n$ ,  $T^r = \inf I^O(W)$ .

Proof. To show that the sequence  $\{T^j\}$  given by (8.2.7) is nondecreasing, we observe that  $T^0 \subseteq T^1$ . Suppose that  $T^{j-1} \subseteq T^j$ . Then

$$T^{j+1} = W + A(Z(B) \cap T^j) \supseteq W + A(Z(B) \cap T^{j-1}) = T^j$$

Therefore, there exists a positive integer  $r$  such that  $T^j = T^r$  for  $j \geq r$ . Let  $T$  be an arbitrary element of  $I^O(W)$ . Thus

$$W \subseteq T, \quad A(Z(B) \cap T) \subseteq T \quad (8.2.8)$$

From (8.2.8) it is clear that  $T^0 \subseteq T$ , and if  $T \supseteq T^{j-1}$ , then

$$T \supseteq W + A(Z(B) \cap T) \supseteq W + A(Z(B) \cap T^{j-1}) = T^j$$

Since  $T$  is an arbitrary element, it follows that  $T \supseteq T^r$ , and hence  $T^r = \inf I^O(W)$ .  $\square$

It is also easy to see that the algorithm of Theorem 8.2.6 can be obtained from the algorithm of Theorem 8.2.5 by dualizing the sequence (8.2.4). To illustrate this alternative procedure we need some simple preliminary results.

Lemma 8.2.2. Let  $E : P(X) \longrightarrow P(U)$  be a linear map. Then

$$R(E)^\perp = N(E^\top)$$

Proof. Let  $z \in R(E)^\perp$ . This implies that  $z^\top y = 0$  for all  $y \in R(E)$ . Since  $y$  can be expressed as  $y = Eu$  for some  $u \in U$ , we have  $z^\top Eu = 0$  which is the same as  $(Eu)^\top z = 0$  or  $u^\top E^\top z = 0$ . Hence  $E^\top z = 0$  and consequently  $z \in N(E^\top)$ . Thus  $R(E)^\perp \subseteq N(E^\top)$ . The reverse inclusion can be easily established by essentially reversing the preceding sequence of steps.  $\square$

Lemma 8.2.3. For all  $S \in P(X)$ ,  $(A^{-1*}S)^\perp = (A^\top)^{-1*}S^\perp$ , where  $A^{-1*}S \equiv \{s : As \in S\}$ .

Proof. Let  $S$  be a basis matrix of  $S$ . Then in view of Lemma 8.2.2, we have

$$\begin{aligned} (AR(S))^\perp &= (R(AS))^\perp = N(S^\top A^\top) \\ &= (A^\top)^{-1*} N(S^\top) \\ &= (A^\top)^{-1*} R(S)^\perp \quad \square \end{aligned}$$

Now using Lemma 8.2.3 and dualizing the sequence (8.2.4), we obtain

$$\begin{aligned} S^{0\perp} &= V^\perp \\ S^{j\perp} &= V^\perp + A^\top (R(B)^\perp \cap S^{(j-1)\perp}), \quad j \in \underline{n} \quad (8.2.9) \\ S^{j\perp} &\supseteq S^{(j-1)\perp}, \quad j \in \underline{n} \end{aligned}$$

Replacing  $V^\perp$  by  $W$ ,  $S^{0\perp}$  by  $T^0$ ,  $A^\top$  by  $A$ ,  $R(B)^\perp$  by  $Z(B)$ , and  $S^{j\perp}$  by  $T^j$ ,  $j \in \underline{n}$ , yields (8.2.7). Similarly, dualizing (8.2.7) results into (8.2.4). Consequently the following relationships hold:

$$\begin{aligned}
& [\sup\{S : AS \subseteq S + R(B) \text{ and } S \subseteq V\}]^\perp \\
& = \inf\{S^\perp : A^T(S^\perp \cap R(B)^\perp) \subseteq S^\perp \text{ and } S^\perp \supseteq V^\perp\} \quad (8.2.10)
\end{aligned}$$

and

$$\begin{aligned}
& [\inf\{T : A(T \cap Z(B)) \subseteq T \text{ and } T \supseteq W\}]^\perp \\
& \sup\{T^\perp : A^T T^\perp \subseteq T^\perp + Z(B)^\perp \text{ and } T^\perp \subseteq W^\perp\} \quad (8.2.11)
\end{aligned}$$

In general, the algorithm of Theorem 8.2.5 does not lead to a closed form expression for  $S_{\sup} = \sup I(V)$ . However, if  $V$  is a hyperplane of  $P(X)$ , then it can be shown that  $S_{\sup}$  has a particularly simple closed form.

Let  $V$  be a hyperplane and  $Z$  a point of  $P(X)$  such that  $Z = \langle z \rangle$ ,  $0 \neq z \in \text{GF}(q)^n$ , and  $N(z^T) = V$ . Furthermore, let

$$\begin{aligned}
d & \equiv \min \{i : z^T A^i B \neq 0\} \text{ if } i \text{ is a positive integer} \\
& = n-1 \text{ if } z^T A^\ell B = 0 \text{ for all positive integers } \ell \quad (8.2.12)
\end{aligned}$$

Theorem 8.2.7. If  $V = N(z^T)$ ,  $z \neq 0$ , then  $S_{\sup} \equiv \sup I(V)$  is given by

$$S_{\sup} = (Z + A^T Z + \dots + (A^T)^d Z)^\perp$$

Proof. Applying (8.2.4) to this special case, we get

$$S^0 = N(z^T)$$

$$S^j = N(z^T) \cap A^{-1*}(S^{j-1} + R(B)), \quad j \in \underline{n}$$



Taking orthogonal complements and using Lemma 8.2.2 and Lemma 8.2.3, we have

$$S^{0\perp} = (N(Z^T))^{\perp} = Z \quad (8.2.13)$$

$$S^{j\perp} = Z + A^T(S^{(j-1)\perp} \cap N(B^T)), \quad j \in \underline{n}$$

By definition of  $d$ ,

$$\sum_{j=1}^i (A^T)^{j-1} Z \subseteq N(B^T), \quad i \in \underline{d} \quad (8.2.14)$$

and

$$(A^T)^d Z \cap N(B^T) = 0 \quad (8.2.15)$$

From (8.2.13) and (8.2.14) it follows that

$$\begin{aligned} S^{1\perp} &= Z \\ S^{2\perp} &= Z + A^T Z \\ &\vdots \\ S^{d\perp} &= Z + A^T Z + \dots + (A^T)^d Z \end{aligned} \quad (8.2.16)$$

From (8.2.13) and (8.2.16),

$$S^{(d+1)\perp} = Z + A^T(N(B^T) \cap \sum_{j=0}^d (A^T)^j Z) \quad (8.2.17)$$

In view of (8.2.14) and (8.2.15), (8.2.17) reduces to

$$\begin{aligned}
 S^{(d+1)\perp} &= Z + A^\top \left( \sum_{j=0}^{d-1} (A^\top)^j Z \right) \\
 &= Z + \sum_{j=0}^d (A^\top)^j Z \\
 &= S^{d\perp}
 \end{aligned}$$

Continuing in this manner, we will see that  $S^{j\perp} = S^{d\perp}$  for  $j \geq d$ .

Since  $\dim V = n-1$ , we must have  $d \leq n-1$ . Therefore,  $S_{\sup} = S^{d\perp}$ .  $\square$

Corollary 8.2.1. If  $z \in \text{GF}(q)^n$ ,  $z \neq 0$ ,  $\langle z \rangle \equiv Z$ , then

$T_{\inf} = \inf I^0(Z)$  is given by

$$T_{\inf} = \sum_{j=0}^d (A^\top)^j Z$$

Now we will show that the above result can be used to derive a sufficient condition for the closure of the set  $I(A, B; X)$  under the operation of intersection.

Theorem 8.2.8. Let  $V \in P(X)$  be of dimension  $n-\ell$  and let  $\{z^1, z^2, \dots, z^\ell\}$  be a basis for  $V^\perp$ . Further, let  $V_i \equiv N(z^{i\top})$ , with  $S_i^{**} \equiv \sup I(A, B; V_i)$ ,  $i \in \underline{\ell}$ . For  $V_i$  let  $d_i$  be the feedback invariants defined by (7.2.12). If

$$\text{rank} \begin{pmatrix} z^{1\top} & A^{d_1} & B \\ z^{2\top} & A^{d_2} & B \\ \vdots & \vdots & \vdots \\ z^{\ell\top} & A^{d_\ell} & B \end{pmatrix} = \ell \quad (8.2.18)$$

then

$$S^{**} \equiv \bigcap_{i=1}^{\ell} S_i^{**} = \sup I(A, B; V)$$

Proof. From (8.2.4),  $S_i^{**}$  can be computed as follows:

$$S_i^0 = V_i$$

$$S_i^j = V_i \cap A^{-1*}(S_i^{j-1} + R(B)), \quad j \in \underline{n} \quad (8.2.19)$$

$$S_i^{**} = S_i^n, \quad i \in \underline{\ell}$$

Therefore,  $S^{**} \subseteq S_i^{**}$ ,  $i \in \underline{\ell}$ , which implies that  $S^{**} \subseteq \bigcap_{i=1}^{\ell} S_i^{**}$ . To prove the reverse inclusion, it suffices, due to maximality of  $S^{**}$ , to show that  $\bigcap_{i=1}^{\ell} S_i^{**} \in I(A, B; V)$ . To accomplish this, let  $T \equiv \bigcap_{i=1}^{\ell} S_i^{**}$  so that

$$T^{\perp} = \sum_{i=1}^{\ell} S_i^{**\perp} \quad (8.2.20)$$

Clearly  $T \subseteq V$ , so it remains to prove that  $T \subseteq A^{-1*}(T + R(B))$  or equivalently,  $A^T(T^{\perp} \cap R(B)^{\perp}) \subseteq T^{\perp}$ . From Theorem 8.2.7 it follows that

$$S_i^{**} = \left( \sum_{j=0}^{di} (A^T)^j Z_i \right)^{\perp}, \quad i \in \underline{\ell}$$

where  $Z_i \equiv \langle z_i^1 \rangle$ ,  $i \in \underline{\ell}$ , so that

$$S_i^{**\perp} = \sum_{j=0}^{di} (A^T)^j Z_i, \quad i \in \underline{\ell} \quad (8.2.21)$$

Substituting (8.2.21) into (8.2.20), we obtain

$$\begin{aligned} T^\perp &= \sum_{i=1}^{\ell} \sum_{j=0}^{d_i} (A^T)^j z_i \\ &= \langle \bigcup_{i=1}^{\ell} \{z_i^1, A^T z_i^1, \dots, (A^T)^{d_i} z_i^1\} \rangle \end{aligned} \quad (8.2.22)$$

Therefore, if  $x \in T^\perp$ , then  $x$  can be expressed as

$$x = \sum_{i=1}^{\ell} \sum_{j=0}^{d_i} a_{ij} (A^T)^j z_i^1, \quad a_{ij} \in \text{GF}(q) \quad (8.2.23)$$

Suppose that  $x$  is also in  $\mathcal{R}(B)^\perp = N(B^T)$ . Then  $B^T x = 0$ . Since by the definition of reachability indices  $d_i$ ,

$$Z_i \subseteq N(B^T)$$

$$Z_i + A^T Z_i \subseteq N(B^T)$$

.

$$Z_i + A^T Z_i + \dots + (A^T)^{d_i-1} Z_i \subseteq N(B^T), \quad i \in \underline{\ell}$$

it follows that

$$\sum_{i=1}^{\ell} a_{id_i} z_i^{1T} A^{d_i} B = 0 \quad (8.2.24)$$

Since by hypothesis (8.2.18), the row  $m$ -vectors  $z_i^{1T} A^{d_i} B$ ,  $i \in \underline{\ell}$ , are linearly independent, (8.2.24) implies that

$$a_{id_i} = 0, \quad i \in \underline{\ell} \quad (8.2.25)$$

In view of (8.2.25), (8.2.23) reduces to

$$x = \sum_{i=1}^{\ell} \sum_{j=0}^{d_i-1} a_{ij} (A^T)^j z^i$$

which, when premultiplied by  $A^T$ , yields  $A^T x = \sum_{i=1}^{\ell} \sum_{j=0}^{d_i-1} a_{ij} (A^T)^{j+1} z^i$ , that is,  $A^T x \in T^\perp$ . Since it was assumed that  $x \in R(B)^\perp$ , we conclude that  $A^T(T^\perp \cap R(B)^\perp) \subseteq T^\perp$ .  $\square$

Corollary 8.2.2. Let  $W \in P(X)$  be of dimension  $\ell$  with a basis  $\{z^1, z^2, \dots, z^\ell\}$ ,  $w_i = \langle z^i \rangle \equiv Z_i$ , and  $T_{**i} \equiv \inf I^0(A, B; w_i)$ ,  $i \in \underline{\ell}$ . Further, let  $d_i$  be the feedback invariants associated with  $w_i$  and defined by (8.2.12). If the condition (8.2.18) holds, then

$$T_{**} \equiv \sum_{i=1}^{\ell} T_{**i} = \inf I^0(A, B; W)$$

Therefore, the set of elements  $S_i^{**}$  of the set  $I(A, B; V)$  satisfying the conditions of Theorem 8.2.8 is closed under intersection. Since the set  $I(A, B; V)$  is always closed under addition,  $\{S_i^{**}\}$  forms a finite projective geometry  $\hat{P}$ . Similarly,  $\{T_{**i}\} \subseteq I^0(A, B; W)$ , specified in Corollary 8.2.2, is closed under the operations of addition and intersection, and hence forms a finite projective geometry  $\hat{P}^0$ . Clearly the geometries  $\hat{P}$  and  $\hat{P}^0$  are dually isomorphic.

### 8.3. (A, B)-Invariant Flats and Output Invariance

(A, B)-invariant flats can be utilized to derive necessary and sufficient conditions for output invariance of LSMs with respect to undesirable disturbance inputs. Consider the perturbed LSM

$$x(k+1) = Ax(k) + bu(k) + Ev(k) \quad (8.3.1)$$

$$y(k) = Cx(k)$$

The term  $Ev(k)$  represents an external disturbance which is assumed not to be directly measurable by the controller. The output invariance problem is to find, if possible, a feedback map  $F : P(X) \longrightarrow P(U)$  such that  $v(k)$  has no influence on the controlled output  $y(k)$ . Before attempting a solution to this problem, we need to make the notion of output invariance more precise.

An output component  $y_i(k)$  is said to be invariant with respect to the  $j$ th component  $v_j(k)$  of disturbance input if the zero state output component  $y_{oi}(k)$  with respect to  $v_j(k)$ , that is, with  $u(k) = 0$ , is identically zero for all  $k$ . If all components  $y_i(k)$ ,  $i \in \underline{r}$ , of the output vector  $y(k)$  are invariant with respect to all the components  $v_j(k)$ ,  $j \in \underline{s}$ , of the disturbance input vector  $v(k)$ , then the LSM  $(A, B, C)$  is said to be output invariant with respect to  $v(k)$ . That is,

$$y_o(k) = C \sum_{j=0}^{k-1} A^{k-j-1} Ev(k) = 0 \quad (8.3.2)$$

for all clock periods  $k$ . Expression (8.3.2) can be equivalently written as

$$C[E, AE, A^2E, \dots, A^{k-1}E] \begin{bmatrix} v(k-1) \\ v(k-2) \\ \vdots \\ v(0) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad (8.3.3)$$

From (8.3.3) it follows that the LSM  $(A, B, C)$  is output invariant with respect to  $v(k)$  if and only if

$$\{A \mid R(E)\} \subseteq N(C)$$

In order to isolate the effect of the disturbance input, we need to introduce the feedback law  $u(k) = Fx(k)$  so that (8.3.1) becomes

$$\begin{aligned} x(k+1) &= (A + BF)x(k) + Ev(k) \\ y(k) &= Cx(k) \end{aligned} \quad (8.3.4)$$

Therefore, the output invariance problem can be formulated as a feedback synthesis problem as follows: Given  $A : P(X) \rightarrow P(X)$ ,  $B : P(U) \rightarrow P(X)$ , and  $R(E), N(C) \in P(X)$ , find a state feedback homomorphism  $F : P(X) \rightarrow P(U)$  such that

$$\{A + BF \mid R(E)\} \subseteq N(C)$$

Consequently this problem is solvable if and only if the largest  $(A + BF)$ -invariant flat that contains  $R(E)$  is part of  $N(C)$ . A glance at the properties of  $(A, B)$ -invariant flats suggests the following solvability condition.

Theorem 8.3.1. The output invariance problem is solvable if and only if

$$S_{\text{sup}} \equiv \sup I(A, B; N(C)) \supseteq R(E)$$

Proof. Suppose that  $S_{\text{sup}} \supseteq R(E)$ . Then by Theorem 8.2.1 we can choose  $F$  such that  $S_{\text{sup}} \in \mathcal{P}_{A+BF}(X)$ . Hence

$$\{A + BF \mid R(E)\} \subseteq \{A + BF \mid S_{\text{sup}}\} = S_{\text{sup}} \subseteq N(C)$$

Conversely, if  $F$  solves the output invariance problem, then  $\{A + BF \mid R(E)\} \in I(A, B; N(C))$ , and therefore  $S_{\text{sup}} \supseteq \{A + BF \mid R(E)\} \supseteq R(E)$ .  $\square$

This result together with Theorem 8.2.5 constitutes a constructive solution to the output invariance problem.

#### 8.4. Reachability Flats

Let  $A : P(X) \rightarrow P(X)$  and  $B : P(U) \rightarrow P(X)$ . A flat  $R \in P(X)$  is said to be a *reachability flat* of the LSM  $(A, B)$  if there exist homomorphisms  $F : P(X) \rightarrow P(U)$  and  $G : P(U) \rightarrow P(U)$  such that

$$R = \{A + BF \mid R(BG)\} \quad (8.4.1)$$

Thus  $R$  is precisely the reachable flat of the LSM  $(A + BF, BG)$ .

The set of reachability flats of a fixed LSM  $(A, B)$  is, in general, a proper subset of  $I(A, B; X)$ , and thus is also related to the set  $I^0(A, B; X)$ . The interrelationships among these sets will be discussed later in this section. The real importance of reachability flats lies in the fact that the restriction of  $A + BF$  to an  $(A + BF)$ -invariant reachability flat can be assigned an arbitrary polynomial



by a suitable choice of  $F$  (see Theorem 5.3.4 and Theorem 8.1.1).

In this chapter we will first discuss the basic properties of reachability flats and then demonstrate their application to some decoupling problems.

First of all, we will replace (8.4.1) by an equivalent expression without the explicit appearance of  $G$ .

Theorem 8.4.1. If  $Z \subseteq R(B)$  and  $\{A \mid Z\} \equiv R$ , then  $\{A \mid R(B) \cap R\} = R$ . Conversely, if  $\{A \mid R(B) \cap R\} = R$ , there exists a  $G : P(U) \rightarrow P(U)$  such that  $\{A \mid R(BG)\} = R$ .

Proof. If  $\{A \mid Z\} = R$ , then  $Z \subseteq R$ , that is,  $Z \subseteq R(B) \cap R$ ,  $R = \{A \mid Z\} \subseteq \{A \mid R(B) \cap R\}$ . To prove the reverse inclusion, notice that  $AR \subseteq R$ , and hence  $\{A \mid R(B) \cap R\} \subseteq R$ , and thus  $\{A \mid R(B) \cap R\} = R$ . For the converse let  $\{b^1, b^2, \dots, b^r\}$  be a basis for  $R(B) \cap R$ . Then  $b^i = Bu^i$ , where  $u^i \in U$ ,  $i \in \underline{r}$ , are linearly independent. Let  $\{u^1, u^2, \dots, u^m\}$  be a basis for  $U$ , and define

$$Gu^i \equiv u^i, \quad i \in \underline{r}$$

$$Gu^i \equiv 0, \quad i = r+1, r+2, \dots, m$$

Then  $R(BG) = R(B) \cap R$ .  $\square$

As an immediate consequence of this result, we have the following characterization of a reachability flat.

Theorem 8.4.2. A flat  $R \in P(X)$  is a reachability flat of an LSM  $(A, B)$  if and only if

$$R = \{A + BF \mid R(B) \cap R\} \quad (8.4.2)$$

Let

$$R(A, B; X) \equiv \{R \in P(X) : R = \{A + BF \mid R(B) \cap R\}\} \quad (8.4.3)$$

and

$$F(R) \equiv \{F : P(X) \longrightarrow P(U) : R \in P_{A+BF}(X)\} \quad (8.4.4)$$

Theorem 8.4.3. If  $R \in R(A, B; X)$ , then  $R = \{A + BF \mid R(B) \cap R\}$  for every map  $F \in F(R)$ .

Proof. By Theorem 8.4.2, there exists a map  $F_0 : P(X) \longrightarrow P(U)$  such that  $R = \{A + BF_0 \mid R(B) \cap R\}$ . Clearly  $F_0 \in F(R)$ . Let  $F_1 \in F(R)$  and define  $R_1 \equiv \{A + BF_1 \mid R(B) \cap R\}$ . Then  $R_1 \subseteq R$ . To show the reverse inclusion, suppose

$$(A + BF_0)^{i-1}(R(B) \cap R) \subseteq R_1, \quad i \in \underline{\ell} \quad (8.4.5)$$

for some  $\ell \in \underline{n}$ . With (8.4.5) as induction hypothesis, we have

$$\sum_{i=1}^{\ell+1} (A + BF_0)^{i-1}(R(B) \cap R) = R(B) \cap R + (A + BF_0) \sum_{i=1}^{\ell} (A + BF_0)^{i-1}(R(B) \cap R)$$

$$\subseteq R(B) \cap R + (A + BF_0)R_1$$

$$= R(B) \cap R + [A + BF_1 + B(F_0 - F_1)]R_1$$

$$\subseteq R(B) \cap R + (A + BF_1)R_1 + B(F_0 - F_1)R_1$$

Let  $x \in R_1$ . Then  $B(F_0 - F_1)x \in R(B)$  and, since  $R_1 \subseteq R$ , (8.4.6)

$$B(F_0 - F_1)x = (A + BF_0)x - (A + BF_1)x \in R$$

Therefore,

$$R(B) \cap R + (A + BF_1)R_1 + B(F_0 - F_1)R_1 \subseteq R(B) \cap R + (A + BF_1)R_1 \subseteq R_1 \quad (8.4.7)$$

By (8.4.6) and (8.4.7),

$$(A + BF_0)^\ell (R(B) \cap R) \subseteq R_1$$

Therefore, (8.4.5) is true for any  $\ell \in \underline{n}$ . Thus  $R \subseteq R_1$ .  $\square$

The above result can be used to check whether a given flat  $R \in \mathcal{P}(X)$  is a reachability flat. This can be done by first examining if  $R \in I(A, B; X)$ . If  $R \notin I(A, B; X)$ , then obviously  $R$  is not a reachability flat since it cannot be made  $(A + BF)$ -invariant. On the other hand, if  $R \in I(A, B; X)$ , then it must pass the additional test of satisfying the relation  $\{A + BF \mid R(B) \cap R\} = R$  to qualify for a reachability flat.

Reachability flats of the LSM  $(A, B)$  can also be characterized in terms of polynomial matrices. In Theorem 6.3.1 a reachability criterion was formulated in terms of the singular pencil of matrices  $[A - \lambda I_n, B]$ . In [108], regarding  $[A - \lambda I_n, B]$  as a mapping, a characterization of the controllability subspace of a linear system in terms of the elements of the null space of  $[A - \lambda I_n, B]$  has been given. This result remains valid over  $GF(q)$  and is given in the following theorem.

Theorem 8.4.4. [108] A flat  $R \in \mathcal{P}(X)$  of dimension  $r \geq 1$  is a reachability flat of the LSM  $(A, B)$  if and only if there exist polynomial vectors  $x(\lambda) \in GF(q)[\lambda]^n$  and  $u(\lambda) \in GF(q)[\lambda]^m$  such that

- (i)  $\deg u(\lambda) = \ell-1$ , for some  $\ell \geq r$ ;
- (ii)  $(A - \lambda I_n)x(\lambda) = Bu(\lambda)$ ;
- (iii) If  $x(\lambda) = \sum_{i=1}^{\ell} (\lambda)^{i-1} x^{i-1}$ , then  $R = \langle x^0, x^1, \dots, x^{\ell-1} \rangle$ .

### 8.5. Eigenvalue Assignability

The set of reachability flats can also be characterized by the eigenvalue assignability property of reachable LSMs  $(A, B)$ .

Theorem 8.5.1. Let  $R \in \mathcal{R}(A, B; X)$  with  $\dim R = r \geq 1$ . Let  $0 \neq b \in R(B) \cap R$ . Then for every set  $\Lambda$  of  $r$  elements of the field  $\text{GF}(q)$ , there exists a map  $F : P(X) \rightarrow P(U)$  such that  $\Lambda = \{A + BF \mid \langle b \rangle\}$  and  $E[(A + BF) \mid R] = \Lambda$ , where  $E[(A + BF) \mid R]$  is the set of eigenvalues of  $(A + BF) \mid R$ , the restriction of  $A + BF$  to  $R$ .

Proof. Suppose

$$R \equiv \{A + BF_0 \mid R(B) \cap R\} \quad (8.5.1)$$

and choose  $G : P(U) \rightarrow P(U)$  such that

$$R(BG) = R(B) \cap R \quad (8.5.2)$$

If we define  $A_0 : P(R) \rightarrow P(R)$  and  $B_0 : P(U) \rightarrow P(R)$  according to  $A_0 \equiv (A + BF_0) \mid R$ ,  $B_0 \equiv BG$ ; then by (8.5.1) and (8.5.2), we have  $\{A_0 \mid R(B_0)\} = R$ . Then application of Theorem 8.1.1 to the pair  $(A_0, B_0)$  yields the existence of  $F_1 : P(R) \rightarrow P(U)$ , such that  $R = \{A_0 + B_0 F_1 \mid \langle b \rangle\}$  and  $E(A_0 + B_0 F_1) = \Lambda$ . Let  $F_2 : P(X) \rightarrow P(U)$  be any extension of  $F_1$  from  $P(R)$  to  $P(X)$ . Then  $F \equiv F_0 + GF_2$  is a map with the required properties.  $\square$

Theorem 8.5.2. Let  $R \in \mathcal{P}(X)$  with  $\dim R = r \geq 1$ . Suppose that for every set  $\Lambda$  of  $r$  elements of the field  $\text{GF}(q)$  there exists a map  $F : \mathcal{P}(X) \rightarrow \mathcal{P}(U)$  such that

$$R \in \mathcal{P}_{A+BF}(X), E[(A + BF) \mid R] = \Lambda \quad (8.5.3)$$

Then  $R \in \mathcal{R}(A, B; X)$ .

Proof. Fix  $F_0 \in F(R)$  and write  $A_0 \equiv (A + BF) \mid R$ . We have  $F \in F(R)$  if and only if  $B(F - F_0)R \subseteq R(B) \cap R$ . Let  $B_0 : \mathcal{P}(U) \rightarrow \mathcal{P}(R)$  be an arbitrary map with  $R(B_0) = R(B) \cap R$ . Then if  $F \in F(R)$ , there exists  $F_1 : \mathcal{P}(R) \rightarrow \mathcal{P}(U)$  such that  $B_0 F_1 = B(F - F_0) \mid R$ . Thus (8.5.3) implies that for every  $\Lambda$  there exists an  $F_1$  such that  $E(A_0 + BF_1) = \Lambda$ . By Theorem 8.1.1, the LSM  $(A, B)$  is reachable. Hence  $R = \{A_0 \mid R(B_0)\} = \{A + BF_0 \mid R(B) \cap R\}$  and thus  $R \in \mathcal{R}(A, B; X)$ .  $\square$

### 8.6. Reachability Flat Algorithm (RFA)

In this section, we will introduce an algorithm that computes the reachability flat  $R$  of a given LSM  $(A, B)$  without explicitly constructing  $F \in F(R)$ . This algorithm will be used to identify further properties of reachability flats.

Let  $R \in \mathcal{P}(X)$  and define

$$\underline{S} \equiv \{S \in \mathcal{P}(X) : S = R \cap (AS + R(B))\} \quad (8.6.1)$$

Later it will be shown that the least  $(A, B)$ -invariant element of this set is the reachability flat of  $(A, B)$ . First we will compute this least element and discuss some other preliminary results.

Theorem 8.6.1. The set  $\underline{S}$  defined by (8.6.1) contains an infimal element  $S_{\inf} \equiv \inf \underline{S}$ .

Proof. Define a sequence  $\{S^j\}$ ,  $S^j \in P(X)$  according to

$$S^0 \equiv \{0\}; S^j \equiv R \cap (AS^{j-1} + R(B)), j \in \underline{n} \quad (8.6.2)$$

We will first show by induction that  $\{S^j\}$  is nondecreasing. Clearly  $S^1 \supseteq S^0$ . Suppose that  $S^j \supseteq S^{j-1}$ . Then

$$S^{j+1} = R \cap (AS^j + R) \supseteq R \cap (AS^{j-1} + R(B)) = S^j$$

Thus there exists an  $\ell \in \underline{n}$  such that  $S^j = S^\ell$  for all  $j \geq \ell$ . To see that  $S^\ell$  is the infimal element of  $\underline{S}$ , let  $S \in \underline{S}$  be an arbitrary flat. Clearly  $S \supseteq S^0$ , and if  $S \supseteq S^j$ , we have

$$S = R \cap (AS + R(B)) \supseteq R \cap (AS^j + R(B)) = S^{j+1}$$

Hence  $S \supseteq S^j$  for all  $j$ , and thus  $S \supseteq S^\ell = S_{\inf}$ .  $\square$

Therefore, the RFA computes  $\inf \underline{S}$  in at most  $n$  steps so that

$$S_{\inf} = \lim_j S^j = S^n \quad (8.6.3)$$

Lemma 8.6.1. Let  $R \in I(A, B; X)$ . If  $F \in F(R)$  and  $\hat{R} \subseteq R$ , then

$$R(B) \cap R + (A + BF)\hat{R} = R \cap (A\hat{R} + R(B))$$

Proof. Clearly  $(A + BF)\hat{R} \subseteq R$  and  $A\hat{R} + R(B) = (A + BF)\hat{R} + R(B)$ .

By the modular distributive law for  $P(X)$ ,

$$R \cap (A\hat{R} + R(B)) = R \cap [(A + BF)\hat{R} + R(B)] = (A + BF)\hat{R} + R(B) \cap R. \quad \square$$

Lemma 8.6.2. Let  $R \in I(A, B; X)$ , let  $F \in F(R)$  and define  $S^j$  by the RFA. Then

$$S^j = \sum_{i=1}^j (A + BF)^{i-1} (R(B) \cap R), \quad j \in \underline{n} \quad (8.6.4)$$

Proof. Clearly (8.6.4) is true for  $j = 1$ . Suppose it holds for  $j = \ell$ . Then

$$\begin{aligned} \sum_{i=1}^{\ell+1} (A + BF)^{i-1} (R(B) \cap R) &= R(B) \cap R + (A + BF)S^\ell \\ &= R \cap (AS^\ell + R(B)) \quad (\text{by Lemma 8.6.1}) \\ &= S^{\ell+1} \end{aligned}$$

Thus the result is proved by induction.  $\square$

Theorem 8.6.2. Let  $R \in P(X)$  and define  $\underline{S}$  by (8.6.1). Then  $R \in R(A, B; X)$  if and only if

$$R \in I(A, B; X) \quad (8.6.5)$$

and

$$R = \inf \underline{S} \equiv S_{\inf} \quad (8.6.6)$$

Proof. Suppose that (8.6.5) and (8.6.6) are true. Then  $F(R) \neq \emptyset$ . Taking  $F \in F(R)$ , we have from (8.6.6), (8.6.3), and (8.6.4)

$$R = S_{\inf} = S^n = \{A + BF \mid R(B) \cap R\}$$

and thus  $R \in R(A, B; X)$ . Conversely, if  $R \in R(A, B; X)$ , then  $F(R) \neq \emptyset$ , so that (8.6.5) is true; and if  $F \in F(R)$ ,

$$R = \{A + BF \mid R(B) \cap R\} = S^n = S_{\inf}$$

by (8.6.4) and (8.6.3).  $\square$

### 8.7. Supremal Reachability Flats

Supremal reachability flats will play an important part in the applications of geometric method. In this section, we will discuss their existence and present some algorithms for their computation.

Theorem 8.7.1. The set of flats  $R(A, B; X)$  is closed under the operation of additon.

Proof. Let  $R_i \in R(A, B; X)$ . Then  $A(R_1 + R_2) \subseteq R_1 + R_2 + R(B)$  and by the RFA

$$R_i = S_i^n, \quad i \in \underline{2}$$

where

$$S_i^0 \equiv \{0\}; \quad S_i^j \equiv R_i \cap (AS_i^{j-1} + R(B)), \quad i \in \underline{2}, \quad j \in \underline{n}$$

Define  $S^j$  according to

$$S^0 \equiv \{0\}; \quad S^j \equiv (R_1 + R_2) \cap (AS^{j-1} + R(B)), \quad j \in \underline{n}$$

Thus  $S^0 = \{0\} = S_i^0$ ,  $i \in \underline{2}$ , and if  $S^j \supseteq S_i^j$ , then

$$S^{j+1} \supseteq R_i \cap (AS_i^j + R(B)) = S_i^{j+1}, \quad i \in \underline{2}$$



and so  $S^{j+1} \supseteq S_1^j + S_2^j$ . Therefore,

$$R_1 + R_2 = S_1^n + S_2^n \subseteq S^n \subseteq R_1 + R_2$$

Hence  $R_1 + R_2 = S^n$ , and the result follows by Theorem 8.6.2.  $\square$

$R(A, B; X)$  is not closed under intersection and thus it is not a projective geometry. Let

$$R(A, B; V) \equiv \{R : R \in R(A, B; X) \text{ and } R \subseteq V \in P(X)\}$$

Theorem 8.7.2. Every flat  $V \in P(X)$  contains a unique reachability flat, denoted by  $\sup R(A, B; V)$ .

Proof. Since  $R(A, B; V) \neq \emptyset$  and is closed under addition, the result follows from finiteness of  $R(A, B; V)$ .  $\square$

Next, two methods will be presented for the computation of  $\sup R(A, B; V)$ .

Theorem 8.7.3. Let  $V_{\sup} \equiv \sup I(A, B; V)$  and  $\sup R(A, B; V) \equiv R_{\sup}$ . If  $F \in F(A, B; V_{\sup})$ , then

$$R_{\sup} = \{A + BF \mid R(B) \cap V_{\sup}\} \quad (8.7.1)$$

This theorem will be proved with the aid of the following two lemmas.

Lemma 8.7.1. Let  $V \in I(A, B; X)$ ,  $R(B_0) \subseteq R(B) \cap V$ ,  $F_0 \in F(V)$ , and define  $R \equiv \{A + BF_0 \mid R(B_0)\}$ . If  $F \in F(V)$  and  $B(F - F_0)V \subseteq R(B_0)$ , then  $R = \{A + BF \mid R(B_0)\}$ .

Proof. Let  $R_1 \equiv \{A + BF \mid R(B_0)\}$  and

$$V^i \equiv \sum_{j=1}^i (A + BF_0)^{j-1} R(B_0), \quad i \in \underline{n}$$

Then  $V^1 = R(B_0) \subseteq R_1$ . Suppose  $V^i \subseteq R_1$ . Then

$$V^{i+1} = R(B_0) + (A + BF_0)V^i \subseteq R(B_0) + (A + BF)V^i + B(F - F_0)V^i$$

Since  $F \in F(R_1)$ ,  $(A + BF)V^i \subseteq R_1$  and because  $F \in F(V)$  and  $R(B_0) \subseteq V$ , we have that  $R_1 \subseteq V$ , and hence

$$B(F - F_0)V^i \subseteq B(F - F_0)R_1 \subseteq R(B_0) \subseteq R_1$$

Therefore,  $V^{i+1} \subseteq R_1$ , so that  $V^i \subseteq R_1$ ,  $i \in \underline{n}$ , and

$$R = V^n \subseteq R_1$$

By interchanging the roles of  $F$  and  $F_0$ , we can infer that  $R_1 \subseteq R$ , and the result follows.  $\square$

Lemma 8.7.2. Let  $R \subseteq V \in I(A, B; X)$  and suppose that the inner product bilinear form is nondegenerate on  $R$ . If  $F_0 \in F(R)$ , then there exists an  $F \in F(V) \cap F(R)$  such that

$$F \mid R = F_0 \mid R$$

Proof. Let  $R \oplus S = V$  for some  $S \in P(X)$  (see Lemma 4.2.1), and let  $\{s^1, s^2, \dots, s^\ell\}$  be a basis for  $S$ . Then  $As^i = v^i + Bu^i$ ,  $i \in \underline{\ell}$ , for some  $v^i \in V$  and  $u^i \in U$ . Let  $F : P(X) \rightarrow P(U)$  be any map such that  $Fx = F_0x$  ( $x \in R$ ) and  $Fs^i = -u^i$ ,  $i \in \underline{\ell}$ . Then  $F$  has the required properties.  $\square$

Proof of Theorem 8.7.2. Let  $F \in F(V_{\text{sup}})$  and

$$R \equiv \{A + BF \mid R(B) \cap V_{\text{sup}}\}$$

Since  $R(B) \cap V_{\text{sup}} = R(BG)$  for some  $G : P(U) \rightarrow P(U)$ , and since

$$(A + BF)^{j-1} (R(B) \cap V_{\text{sup}}) \subseteq V_{\text{sup}} \subseteq V, \quad j \in \underline{n}$$

it is clear that  $R \in R(A, B; V)$ . Let  $R_0 \in R(A, B; V)$  be arbitrary.

Then

$$R_0 = \{A + BF_0 \mid R(B) \cap R_0\}$$

for some  $F_0 : P(X) \rightarrow P(U)$ . Since  $R_0 \in P_{A+BF}(X)$ , clearly

$$R_0 \subseteq \sup I(A, B; V) = V_{\text{sup}}$$

choose, by Lemma 8.7.2,  $F_1 \in F(R_0) \cap F(V_{\text{sup}})$  such that  $F_1 \mid R_0 = F_0 \mid R_0$ . If  $x \in V_{\text{sup}}$ , then

$$B(F - F_1)x = (A + BF)x - (A + BF_1)x \in V_{\text{sup}}$$

so that

$$R_0 = \{A + BF_1 \mid R(B) \cap R_0\} \subseteq \{A + BF_1 \mid R(B) \cap V_{\text{sup}}\}$$

$$= \{A + BF \mid R(B) \cap V_{\text{sup}}\} \quad (\text{by Lemma 8.7.1})$$

$$= R$$

Therefore,  $R \in R(A, B; V)$  is supremal and so  $R = R_{\text{sup}}$ .

A generalization of RFA provides a second method for the computation of  $\sup R(A, B; V_{\text{sup}})$  which does not require prior computation of  $F \in F(V_{\text{sup}})$

Theorem 8.7.4. Define the sequence  $\{S^j\}$  according to

$$S^0 \equiv \{0\}; S^j \equiv V_{\text{sup}} \cap (AS^{j-1} + R(B)), j \in \underline{n} \quad (8.7.2)$$

Then  $S^j = R_{\text{sup}}$  for  $j \geq \dim V_{\text{sup}}$ .

Proof. It can be easily shown by induction that the sequence  $\{S^j\}$  is monotonically nondecreasing and thus  $S^j = S^\ell$  for  $j \geq \ell = \dim V_{\text{sup}}$ . Since  $S^j \subseteq V_{\text{sup}} \in I(A, B; X)$ , we have

$$\begin{aligned} AS^j &\subseteq (V_{\text{sup}} + R(B)) \cap (AS^j + R(B)) \\ &= V_{\text{sup}} \cap (AS^j + R(B) + R(B)) \\ &= S^{j+1} + R(B) \end{aligned}$$

so that  $AS^\ell \subseteq S^\ell + R(B)$ . Since  $S^j \subseteq S^\ell \subseteq V_{\text{sup}}$ ,  $j \in \underline{n}$ , (8.7.2) implies that

$$S^j = S^\ell \cap (AS^{j-1} + R(B)), j \in \underline{n}$$

By Theorem 8.6.2,  $S^\ell \in R(A, B; V)$ , and thus  $S^\ell \subseteq R_{\text{sup}}$ . On the other hand,  $R_{\text{sup}} = R^n$ , where

$$R^0 \equiv \{0\}; R^j \equiv R_{\text{sup}} \cap (AR^{j-1} + R(B)), j \in \underline{n}$$

Since  $R_{\text{sup}} \subseteq V_{\text{sup}}$ , it follows by induction on  $j$  that  $R^j \subseteq S^j$ ,  $j \in \underline{n}$ , and therefore  $R_{\text{sup}} = S^\ell$ .  $\square$

### 8.8. Noninteraction in LSMs

Noninteraction or decoupling has been a longstanding problem of theoretical and practical interest in the area of dynamical systems. Although this concept has never been applied to LSMs in any practical context, it is conceivable that in the future the incorporation of noninteracting controls will be of major importance in certain aspects of the design process of large scale LSMs and other models in the area of automata theory. Here we will present a brief discussion of a geometric formulation of the decoupling problem for LSMs and some related results for the purpose of indicating an application of supremal reachability flats discussed in the previous sections.

Roughly speaking, a multi-input multi-output LSM is decoupled if each output can be independently controlled by a corresponding input. To make this notion more precise, let the output vector  $y(k)$  and the matrix  $C$  of the LSM  $M = (A, B, C)$  be partitioned as

$$y(k) = \begin{bmatrix} y^1(k) \\ y^2(k) \\ \vdots \\ y^\ell(k) \end{bmatrix}, \quad C = \begin{bmatrix} C_1 \\ C_2 \\ \vdots \\ C_\ell \end{bmatrix} \quad (8.8.1)$$

where  $y^i(k) \in GF(q)^{s_i}$  and  $C_i \in GF(q)^{s_i \times n}$ ,  $i \in \underline{\ell}$ ,  $s_1 + s_2 + \dots + s_\ell = r \equiv \dim V$ . Thus the output relation of  $M$  can be expressed as

$$y^i(k) = C_i x(k), \quad i \in \underline{\ell}$$

Consider the feedback law

$$u(k) = Fx(k) + \sum_{i=1}^{\ell} G_i v^i(k)$$

Then a solution to the decoupling problem consists of finding matrices  $F$  and  $G_i$ ,  $i \in \underline{\ell}$ , such that input  $v^i(k)$  can control output  $y^i(k)$  without affecting any other output  $y^j(k)$ ,  $j \neq i$ .

In order to give a geometric formulation of this problem, let  $R^i$  denote the reachability flat generated by  $v^i(k)$ , that is,

$$R^i \equiv \{A + BF \mid R(BG_i)\} \equiv \{A + BF \mid R(B) \cap R(G_i)\}, \quad i \in \underline{\ell}$$

Since the output  $y^i(k)$  is to be controlled completely by the input  $v^i(k)$ , we must have

$$C_i R^i = R(G_i), \quad i \in \underline{\ell}$$

For  $v^i(k)$  to leave the outputs  $y^j(k)$ ,  $j \neq i$ , unaffected it must be required that

$$C_j R^i = \{0\}, \quad j \neq i, \quad i \in \underline{\ell}$$

Using the above observations, the decoupling problem can be stated in geometric terms as follows: Given  $A$ ,  $B$ , and  $N(C_i)$ ,  $i \in \underline{\ell}$ , determine

a feedback homomorphism  $F : P(X) \longrightarrow P(U)$  and reachability flats

$R^i \in P(X)$ ,  $i \in \underline{\ell}$ , such that

$$R^i = \{A + BF \mid R(B) \cap R(G_i)\}, \quad i \in \underline{\ell}$$

$$P(R^i + N(C_i)) = P(X), \quad i \in \underline{\ell}$$

$$R^i \subseteq \bigcap_{\substack{j=1 \\ j \neq i}}^{\ell} N(C_j), \quad i \in \underline{\ell}$$

A set of reachability flats  $R^i \in P(X)$ ,  $i \in \underline{\ell}$  satisfying these conditions is called a solution to the decoupling problem.

We will state necessary and sufficient conditions for the solvability of the decoupling problem for LSMs in certain special cases. The proofs of these assertions are essentially the same as those given in [111] for conventional continuous-time linear systems over the field of real numbers and thus will not be reproduced here.

In the remainder of this section, let  $R_{\sup}^i$  denote the supremal reachability flat such that

$$R_{\sup}^i \subseteq \bigcap_{\substack{j=1 \\ j \neq i}}^{\ell} N(C_j), \quad i \in \underline{\ell}$$

Case 1. Rank  $C = n = r$

This assumption means that there is a one-to-one correspondence between state variables and output variables. Furthermore, in view of (8.8.1), rank  $C = n$  implies that

$$\bigcap_{i=1}^{\ell} N(C_i) = \{0\} \quad (8.8.2)$$

Theorem 8.8.1. Suppose that (8.8.2) holds. Then a solution to the decoupling problem exists if and only if

$$P(R_{\text{sup}}^i + N(C_i)) = P(X), \quad i \in \underline{\ell}$$

$$\text{Case 2. Rank } G = \text{Rank}[G_1, G_2, \dots, G_{\ell}] = m \quad (8.8.3)$$

This assumption is equivalent to

$$R(B) = \sum_{i=1}^{\ell} R(B) \cap R^i \quad (8.8.4)$$

where  $R^i$ ,  $i \in \underline{\ell}$ , is any fixed solution of the decoupling problem for which (8.8.3) holds. The equivalence of (8.8.3) and (8.8.4) follows from the fact that  $\text{rank } G = m$  implies that  $\dim R(G) = m$ , or  $R(G) = U$  and thus

$$R(B) = BU = BR(G) = B \sum_{i=1}^{\ell} R(G_i) \subseteq R(B) \cap R^i \subseteq R(B)$$

Theorem 8.8.2. Suppose that (8.8.4) holds. Then a solution to the decoupling problem exists if and only if

$$R(B) = \sum_{i=1}^{\ell} R(B) \cap R_{\text{sup}}^i$$



Case 3. Rank  $B = \ell$

This assumption means that there is a one-to-one correspondence between the inputs and the outputs. Rank  $B = \ell$  implies that

$$\dim R(B) = \ell \quad (8.8.5)$$

Theorem 8.8.3. If (8.8.5) holds, then the decoupling problem has a solution if and only if

$$R(B) = \sum_{i=1}^{\ell} R(B) \cap R_{\text{sup}}^i$$

We will conclude this chapter with a final observation: It can be shown that every projective geometry is a modular lattice, and every affine geometry is a semimodular lattice. Thus, the contents of this chapter can also be couched in the language of lattice theory (cf. [119]). It appears that an effective exploitation of the interconnections among linear sequential machines, coding theory, finite projective and affine geometries, and lattice theory warrants much further research.

Summary and Conclusions

In this chapter, using the language of finite affine and projective geometries, certain state reachability aspects of a geometric theory which was recently introduced by Wonham and Morse [111], [112], [113], and independently by Basile and Marro [5], [6], [7] for continuous-time linear systems, were adapted and specialized for LSMs (cf. [5], [6], [7], [8], [9], [10], [80], [81], [83], [84], [85], [108], [111], [112], [113], [114]).

## CHAPTER IX

## OUTPUT REACHABILITY AND OUTPUT CONTROLLABILITY OF LSMs

Similar to the concepts of state reachability and state controllability, it makes sense to define the notions of output reachability and output controllability for LSMs which, generally speaking, refer to the transferability of initial outputs to final outputs by input sequences of finite lengths. In view of the fact that states and outputs of the LSM  $M = (A, B, C)$  are connected by the linear map  $C : X \rightarrow Y$ , it is natural to expect that most of the results developed for the properties of state reachability and state controllability may be easily checked to see if they can be appropriately modified to yield similar results for output reachability and output controllability. Therefore, based on this understanding, in this chapter we will only very briefly discuss the concept of output reachability of LSMs.

9.1. Output Reachability of LSMs

This section is essentially an imitation of Section 4.1 with the aim of pointing out the type of modifications that have to be made for converting certain state reachability criteria to corresponding output reachability criteria.

Throughout this chapter,  $r \equiv \dim Y$ .

Definition 9.1.1. An output  $y^1 \neq 0_Y$  of the LSM  $M = (A, B, C)$  is said to be *reachable* from the output  $y^0 \in Y$  if there exists an

input sequence  $w \in U^*$  which transfers  $y^0$  to  $y^1$ ; if  $lg(w) = \ell$ , then  $y^1$  is said to be  $\ell$ -reachable from  $y^0$ . The LSM  $M$  is said to be  $\ell$ -output reachable if every output of  $M$  is  $\ell$ -reachable for at least one particular  $\ell$ . The smallest integer  $\ell$  for which  $M$  is  $\ell$ -output reachable is called the *output reachability index* of  $M$ .

Now if we define

$$y(y^0) \equiv \{\text{outputs reachable from } y^0\}$$

$$e_j(y^0) \equiv \{\text{outputs reachable from } y^0 \text{ in exactly } j \text{ clock periods}\}$$

$$y_j(y^0) \equiv \{\text{outputs reachable from } y^0 \text{ in at most } j \text{ clock periods}\}$$

then Lemma 4.1.1 - Lemma 4.1.4 also hold for the above output sets and can be restated as follows:

Lemma 9.1.1.  $y(y^0) = \bigcup_{j=0}^{\infty} e_j(y^0) = \bigcup_{j=0}^{\infty} y_j(y^0)$

Lemma 9.1.2.  $y_0(y^0) \subseteq y_1(y^0) \subseteq \dots \subseteq y(y^0)$

Lemma 9.1.3. If there exists an integer  $t$  such that  $y_{t'}^{(y^0)} = y_t^{(y^0)}$  for all  $t' \geq t$ , then

$$y(y^0) = \bigcup_{j=0}^{\infty} y_j(y^0) = y_t^{(y^0)} = \bigcup_{j=0}^t e_j(y^0)$$

Lemma 9.1.4. Let  $M = (A, B, C)$  be an LSM with an output  $y^0 \in Y$  for which there exists an integer  $j$  such that

$$y_j^{(y^0)} = y_{j+1}^{(y^0)}$$

Then

$$y_j^{(y^0)} = y_j^{(y^0)}$$

With the aid of the above results and notation, the following theorem can be proved in precisely the same manner as Theorem 4.1.3.

Theorem 9.1.1. Every output of the  $n$ -dimensional LSM  $M = (A, B, C)$ , reachable from the zero output  $0_y$ , can be reached in at most  $r$  clock periods, that is,  $y_r^{(0_y)} = y^{(0_y)}$ .

Theorem 9.1.2. For the LSM  $M = (A, B, C)$  the set of all reachable outputs from the zero output  $0_y$  in at most  $\ell$  clock periods is the range of the linear map

$$K_o \equiv [CA^{\ell-1}B, CA^{\ell-2}B, \dots, CAB, CB] : U^* \longrightarrow Y$$

That is,

$$y_\ell^{(0_y)} = R(K_o)$$

Proof. Since a zero input leaves the zero state  $0_x$  unchanged and  $0_y = C0_x$ , if an output  $\tilde{y}$  can be reached from the zero output by applying an input sequence  $u(0)u(1) \dots u(j-1)$  of length  $j < \ell$ , then  $\tilde{y}$  can also be reached from  $0_y$  by first applying the input sequence  $0^{\ell-j}$ , that is, a string of  $\ell-j$  successive zero inputs, and then applying  $u(0)u(1) \dots u(j-1)$ . Thus, for all  $u(0)u(1) \dots u(j-1) \in U^*$  and all  $j < \ell$ ,  $y_\ell^{(\tilde{y})} = y_\ell^{e(\tilde{y})}$ . Therefore,  $\tilde{y} \in Y^{(0_y)}$  if and only if

there exists an input sequence  $u(0)u(1) \dots u(\ell-1)$  of length exactly  $\ell$  such that

$$\tilde{y} = CA^{\ell}0_X + \sum_{j=0}^{\ell-1} CA^{\ell-j-1}Bu(j)$$

or

$$\tilde{y} = [CA^{\ell-1}B, CA^{\ell-2}B, \dots, CAB, CB] \begin{bmatrix} u(0) \\ u(1) \\ \vdots \\ u(\ell-1) \end{bmatrix}$$

Thus  $\tilde{y} \in R(K_0)$ .  $\square$

With the aid of Theorem 9.1.1, we can derive a simple criterion for checking the output reachability of an LSM.

Theorem 9.1.3. The LSM  $M = (A, B, C)$  is output reachable if and only if

$$\text{rank}[CA^{n-1}B, CA^{n-2}B, \dots, CAB, CB] = r \equiv \dim Y$$

Proof. The LSM  $M = (A, B, C)$  is output reachable if and only if  $Y = Y_{(0_Y)}$  or, since  $Y_{(0_Y)} = Y_r^{(0_Y)}$  by Theorem 9.1.1, if and only if,

$$\dim Y = \dim Y_r^{(0_Y)}$$

$$r = \dim R([CA^{n-1}B, CA^{n-2}B, \dots, CAB, CB])$$

$$r = \text{rank}[CA^{n-1}B, CA^{n-2}B, \dots, CAB, CB] \quad \square$$

A trivial consequence of this theorem is that every single-output LSM is output reachable.

It is clear that the output reachability criterion of Theorem 9.1.3 is similar to the state reachability criterion of Theorem 4.1.4. In fact, these criteria become identical if  $r = n$  and  $C = I_n$ . Therefore, output reachability implies state reachability for the LSM  $M = (A, B, C)$  only if  $r = n$  and  $C$  is nonsingular. On the other hand, it is easy to see that state reachability implies output reachability if and only if  $\text{rank } C = r$ . Thus, in general, state reachability is neither necessary nor sufficient for output reachability.

Corollary 9.1.1. The LSM  $M = (A, B, C)$  is output reachable if and only if the matrix

$$K_o K_o^T = \sum_{j=0}^{n-1} CA^{n-j-1} BB^T (A^T)^{n-j-1} C^T \in GF(q)^{r \times r}$$

is nonsingular.

Corollary 9.1.2. The LSM  $M = (A, B, C)$  is  $\ell$ -output reachable  $\ell \leq n$ , if and only if

$$\text{rank}[CA^{\ell-1}B, CA^{\ell-2}B, \dots, CAB, CB] = r$$

Corollary 9.1.3. Let the minimal polynomial of  $A$  be of degree  $s \leq n$ . Then the LSM  $M = (A, B, C)$  is  $\ell$ -output reachable for some  $\ell \geq s$ , if and only if it is  $s$ -output reachable.

In applying Theorem 9.1.3 it is required to compute the entire output reachability matrix  $K_o$ . However, in many cases the output reachability property can be checked by considering a matrix of relatively smaller size. This simplification is based on the following result.

Theorem 9.1.4. If  $j$  is the least integer such that  $\text{rank}[CB, CAB, \dots, CA^j B] = \text{rank}[CB, CAB, \dots, CA^{j+1} B]$ , then  $\text{rank}[CB, CAB, \dots, CA^\ell B] = \text{rank}[CB, CAB, \dots, CA^j B]$  for all  $\ell < j$ , and  $j \leq \min\{n-s, \bar{n}-1\}$ , where  $s = \text{rank } CB$  and  $\bar{n}$  is the degree of the minimal polynomial of  $A$ .

Corollary 9.1.4. (Simplified Output Reachability Criterion)  
If  $\text{rank } CB = s$ , then the LSM  $M = (A, B, C)$  is output reachable if and only if  $\text{rank}[CB, CAB, \dots, CA^{n-s} B] = r$ .

Corollary 9.1.5. If  $\text{rank } CB = s$ , then the LSM  $M = (A, B, C)$  is output reachable if and only if  $\begin{matrix} \vee & \vee \\ K_O & K_O^T \end{matrix} \in \text{GF}(q)^{r \times r}$  is nonsingular, where  $\begin{matrix} \vee \\ K_O \end{matrix} \equiv [CB, CAB, \dots, CA^{n-s} B]$ .

In a similar manner various other concepts and results pertaining to the properties of state reachability and state controllability presented in the previous chapters may be appropriately modified and adapted for output reachability and output controllability of LSMs.

### Summary and Conclusions

In this chapter it was briefly demonstrated that although state reachability and output reachability are essentially distinct concepts in the sense that, in general, one does not imply the other, nevertheless the simple mathematical relationship between the state and output of a Mealy LSM may be utilized to see, in a straightforward manner, if any given state reachability criterion can be restated in terms of the output reachability property of LSMs. In view of this fact and in the interest of avoiding unnecessary repetition, an extensive development and documentation of the concepts of output reachability and output controllability was not pursued.

## CHAPTER X

### OBSERVABILITY AND STATE OBSERVER DESIGN FOR LSMS

This chapter will be devoted to a brief discussion of the concept of state observability of LSMS and some of its applications. A complete duality relationship will be established between the properties of state reachability and state observability which can be used to restate, in a straightforward manner, all the results pertaining to reachability in terms of observability. The important role of the property of observability will be illustrated by presenting some design procedures for Luenberger type state observers of full and reduced order for single-output and general multivariable LSMS.

#### 10.1. State Observability of LSMS

Closely linked with the notion of state reachability of LSMS is the dual notion of *observability*, or *diagnosability*. Loosely speaking, observability refers to the possibility of reconstructing the state from output measurements. Thus, the dual relationship between reachability and observability is intuitively clear: an LSM is reachable if every state can be reached by a suitable choice of input sequences; it is observable if every state can be computed by suitable processing of outputs.

In discussing state reachability and other state related concepts, it was always implicitly assumed that the entire state vector



is available. Suppose, however, that the state vector or some sub-vector of it is not directly accessible for measurement. Then the natural question of interest is how to "observe" the dynamical behavior of the entire state vector if the only available measurements are the output components  $y_i(k)$ ,  $i \in \underline{r}$ . In other words, suppose that a "black box" having the string  $u(0)u(1) \dots u(\ell-1)$  as its input and  $y(1)y(2) \dots y(\ell)$  as its output is in state  $x^0$ . How can we verify that the "box" is in fact in state  $x^0$  without "opening" it? This state determination problem can be solved by applying inputs to the given LSM and checking that the resulting outputs are indeed appropriate to the LSM started in state  $x^0$ , that is, by feeding sequences  $w \in U^*$  and checking the responses  $\rho_{x^0}(w)$ , where

$$\rho_{x^0} : U^* \longrightarrow Y, w \longmapsto \rho_{x^0}(w) \equiv \eta(\phi(x^0), w)$$

If two states  $x^0$  and  $x^1$  of the LSM have identical response functions, then there is no way of telling "from the outside" whether the LSM is in state  $x^0$  or in state  $x^1$ . This suggests the following definition of observability.

Definition 10.1.1. The LSM  $M = (A, B, C)$  is *observable* if for every pair of distinct states  $x^0$  and  $x^1$  there exists at least one input sequence to which they respond differently, that is, there exists  $w \in U^*$  such that  $\rho_{x^0}(w) \neq \rho_{x^1}(w)$ .

This result implies state determinability in the following stronger sense.

Definition 10.1.2. The LSM  $M = (A, B, C)$  is said to be  $\ell$ -*observable* if and only if every initial state  $x(0)$  of  $M$  can be determined from the input-output record  $\{u(0)u(1) \dots u(\ell-1)\}$ ,  $\{y(1)y(2) \dots y(\ell)\}$ . The smallest integer  $\ell_0$  for which  $M$  is  $\ell_0$ -observable is called the *observability index* of  $M$ .

Therefore, the property of observability refers to deducing the present state of an LSM from *future* output observations. However, there is a complementary state determination problem called *reconstructibility*, *determinability*, or *identifiability* which refers to deducing the present state from the *past* output record. This property is essential in data filtering problems since usually past output values are available in such situations. It turns out that for time-invariant LSMs state observability and state reconstructibility imply each other.

In order to derive explicit criteria for checking observability and also for the purpose of establishing the relationship between the properties of state observability and state reachability, we will exploit the connection between the notions of observability and indistinguishability established in Definition 10.1.1.

In Section 3.4, it was noted that two states  $x^1$  and  $x^2$  of the LSM  $M = (A, B, C)$  are  $\ell$ -indistinguishable if and only if  $x^1 - x^2 \in N(L)$ , where

$$L \equiv \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{\ell-1} \end{bmatrix}$$

But by Lemma 8.2.2,  $N(L) = R(L^T)^\perp$  so that we have the following equivalent characterization of  $\ell$ -indistinguishability.

Lemma 10.1.1. Two states  $x^1$  and  $x^2$  of the LSM  $M = (A, B, C)$  are  $\ell$ -indistinguishable if and only if  $x^1 - x^2 \in R(L^T)^\perp$ , that is, if and only if the vector  $x^1 - x^2$  is orthogonal to

$$R(L^T) = R([C^T, A^T C^T, (A^T)^2 C^T, \dots, (A^T)^{\ell-1} C^T]) \quad (10.1.1)$$

Comparing (10.1.1) with the state reachability matrix  $K \equiv [B, AB, A^2 B, \dots, A^{\ell-1} B]$ , we immediately see that (10.1.1) can be interpreted as the set of all states of the LSM

$$z(k+1) = A^T z(k) + C^T v(k) \quad (10.1.2)$$

reachable, in at most  $\ell$  steps, from the zero state of (10.1.2). Now letting  $S_\ell^0$  denote the set of states of (10.1.2) that are reachable from the zero state in at most  $\ell$  steps, and  $S^0$  be the set of all states reachable from the zero state, then clearly  $S_{\ell-1}^0 \subseteq S_\ell^0 \subseteq S_n^0 \subseteq S^0$ , and Lemma 10.1.1 can be equivalently restated as follows.

Lemma 10.1.2. Two states  $x^1$  and  $x^2$  of the LSM  $M = (A, B, C)$  are  $\ell$ -indistinguishable if and only if  $x^1 - x^2 \in S_\ell^{0\perp}$ .

The above discussion leads to the following important result.

Theorem 10.1.1. The LSM  $M = (A, B, C)$  is observable if and only if the LSM  $M^0 : z(k+1) = A^T z(k) + C^T v(k)$  is reachable.

Proof. Let  $x^1$  and  $x^2$  be any two arbitrary states of  $M$ . Then by definition  $M$  is observable if and only if indistinguishability of  $x^1$  and  $x^2$  implies that  $x^1 = x^2$ . Thus in view of Lemma 10.1.2,  $M$  is observable if and only if  $(x^1 - x^2) \in S_\ell^{0\perp}$  for all  $\ell \geq 0$  implies that  $x^1 = x^2$ . That is,  $M$  is observable if and only if  $(x^1 - x^2) \in (S_n^0 = S^0)^\perp$  implies that  $x^1 = x^2$ . But this implies that  $M$  is observable if and only if  $S^{0\perp} = \{0_X\}$ , that is, the only vector orthogonal to all the reachable states of  $M^0$  is  $0_X$ . Thus  $S^0 = \{0_X\}^\perp = X^0$ . Therefore,  $M$  is observable if and only if all states of  $M^0$  are reachable.  $\square$

Corollary 10.1.1. The LSM  $M = (A, B, C)$  of dimension  $n$  is observable if and only if

$$\text{rank}([C^\top, A^\top C^\top, \dots, (A^\top)^{n-1} C^\top]) = n \quad (10.1.3)$$

In terms of the condition (10.1.3), the observability index  $\ell_0$  of Definition 10.1.2 can be characterized as

$$\ell_0 = \min\{j : \text{rank}([C^\top, A^\top C^\top, \dots, (A^\top)^{j-1} C^\top]) = n\} \quad (10.1.4)$$

If we define the dual of the LSM  $M = (A, B, C)$  to be the LSM  $M^0 = (A^\top, C^\top, B^\top)$ , then we have proved the celebrated Kalman Duality Theorem of conventional linear systems for linear machines.

Theorem 10.1.2. (Duality Theorem) The LSM  $M = (A, B, C)$  is observable (reachable) if and only if its dual  $M^0 = (A^\top, C^\top, B^\top)$  is reachable (observable).

Before we discuss the implications of this duality relationship for LSMs, we want to reexamine the link between the concepts of indistinguishability and observability in geometric terms.

From Definition 10.1.1, it follows that a state  $\bar{x}$  of the LSM  $M = (A, B, C)$  is unobservable if and only if  $\bar{x} \sim 0$ . But

$$\bar{x} \sim 0 \iff \bar{x} \in N(L)$$

$$\iff \begin{pmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{pmatrix} \bar{x} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

$$\iff CA^{j-1}\bar{x} = 0, \quad j \in \underline{n}$$

$$\iff \bar{x} \in N(CA^{j-1}), \quad j \in \underline{n}$$

$$\iff \bar{x} \in \bigcap_{j=1}^n N(CA^{j-1})$$

Therefore,

$$Z \equiv \bigcap_{j=1}^n N(CA^{j-1}) \quad (10.1.5)$$

is the unobservable subspace of the LSM  $M = (A, B, C)$ . Clearly  $AZ \subseteq Z$ . Thus the orthogonal complement  $Z^\perp$  of (10.1.5) is the set of observable states of  $M$ . Consequently  $M$  is observable if and only if  $Z = \{0_X\}$  or equivalently, if and only if

$$Z^\perp = \{0_X\}^\perp = X^o \quad (10.1.6)$$

Using the above results, we will present a different proof of the Duality Theorem.

Theorem 10.1.3. (Duality Theorem) Let  $C : X \longrightarrow Y$  and  $A : X \longrightarrow X$  be linear maps with duals  $C^T : Y^O \longrightarrow X^O$  and  $A^T : X^O \longrightarrow X^O$ . Then the LSM  $M = (A, B, C)$  is observable (reachable) if and only if the dual LSM  $M^O = (A^T, C^T, B^T)$  is reachable (observable).

Proof. In view of (10.1.5) and Lemma 8.2.2, we have

$$\begin{aligned} N_o^\perp &\equiv \left[ \begin{array}{c} n \\ \cap \\ j=1 \end{array} N(CA^{j-1}) \right]^\perp = \sum_{j=1}^n \left[ N(CA^{j-1}) \right]^\perp \\ &= \sum_{j=1}^n (A^T)^{j-1} R(C^T) = R([C^T, C^T A^T, \dots, (A^T)^{n-1} C^T]) \\ &\equiv \{A^T \mid R(C^T)\} \end{aligned} \quad (10.1.7)$$

Therefore,  $N_o = \{0_X\}$ , that is,  $M$  is observable if and only if  $\{A^T \mid R(C^T)\} = X^O$ , that is, if and only if  $M^O$  is reachable.  $\square$

The observability index  $\ell_o$  given by (10.1.4) can be alternatively characterized in terms of (10.1.5) and (10.1.7) as follows:

$$\ell_o = \min\{j : \bigcap_{i=1}^j N(CA^{i-1}) = \{0\}\} \quad (10.1.8)$$

$$\ell_o = \min\{j : \sum_{i=1}^j (A^T)^{i-1} R(C^T) = X^O\} \quad (10.1.9)$$

### 10.2. Consequences of the Duality Theorem

The Duality Theorem is certainly a fortunate result in that it obviates the necessity of developing a separate detailed observability theory for LSMs. In view of this duality relationship, all the results in Chapters IV, V, VI, VII, and VIII pertaining to state reachability can be dualized, in a straightforward manner, to yield the corresponding results in terms of observability. For example, all the 24 state reachability criteria of Theorem 6.1.1 for the single-input LSM  $(A, b, c^T)$  can be dualized by simply replacing  $A$  by  $A^T$  and  $b$  by  $(c^T)^T = c$  to yield 24 observability criteria for the single-output LSM  $(A, b, c^T)$ . Likewise, it is sufficient to replace  $A$  by  $A^T$  and  $B$  by  $C^T$  in any of the state reachability properties of Theorem 6.3.1 in order to obtain various equivalent statements for the observability of the LSM  $(A, B, C)$ . Thus such a simple rule makes it unnecessary to formulate explicitly all the various equivalent forms of the property of observability.

However, it should be pointed out that there is one minor exception: observability is not, in general, invariant under the action of state feedback. To see this, suppose that the LSM  $M = (A, B, C)$  is observable, that is,  $\text{rank}[C^T, A^T C^T, \dots, (A^T)^{n-1} C^T] = n$ . Then if the feedback homomorphism  $F$  can be chosen such that  $A = -BF$ , then the observability matrix of the feedback compensated LSM  $\bar{M} = (A + BF, B, C) = (0, B, C)$  becomes  $[C^T, 0, \dots, 0]$ . Hence if  $\text{rank } C < n$ , then  $\bar{M}$  is not observable.

### 10.3. State Minimization and Observability of LSMs

From Section 2.4, we recall that a minimal LSM is one in which no two states are indistinguishable. This, in turn, implies that a minimal LSM  $M = (A, B, C)$  is observable, that is,

$$\text{rank } L \equiv \text{rank} \begin{pmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{pmatrix} = n$$

If, on the other hand, the LSM  $M = (A, B, C)$  is not observable, that is,  $\text{rank } L < n$ , then the minimal form  $\overset{\vee}{M} = (\overset{\vee}{A}, \overset{\vee}{B}, \overset{\vee}{C})$  of  $M$  which will be of smaller dimensionality than  $M$  can be determined by some simple procedures [22], [41]. The process by which the characterizing matrices of  $\overset{\vee}{M}$  are computed is referred to as the *minimization* of  $M$ .

A slightly more general notion related to minimality is that of *irreducibility* of an LSM. In Theorem 4.2.1 it was established that if an LSM  $M$  is not state reachable, then it can be *reduced* to a reachable LSM  $\bar{M}$  which is of smaller dimension than and zero-state equivalent to  $M$ . A special state transformation was introduced in Theorem 4.2.2 to transform  $M$  to the unreachable isomorphic form (4.2.1) from which  $\bar{M}$  can be extracted. Obviously, similar reducibility results can be asserted in terms of the observability property by virtue of the Duality Theorem.

The dual of Theorem 4.2.1 can be stated as follows:



Theorem 10.3.1. If the LSM  $M = (A, B, C)$  is not observable, that is,  $\text{rank } L = s < n$ , then there exists an isomorphism  $P : X \rightarrow X$ ,  $P \in \text{GF}(n, q)$ , such that the isomorphic LSM  $\tilde{M} = (\tilde{A}, \tilde{B}, \tilde{C}) \equiv (PAP^{-1}, PB, CP^{-1})$  has the form

$$\begin{bmatrix} \tilde{x}^1(k+1) \\ \tilde{x}^2(k+1) \end{bmatrix} = \begin{bmatrix} \tilde{A}_{11} & 0 \\ \tilde{A}_{21} & \tilde{A}_{22} \end{bmatrix} \begin{bmatrix} \tilde{x}^1(k) \\ \tilde{x}^2(k) \end{bmatrix} + \begin{bmatrix} \tilde{B}_1 \\ \tilde{B}_2 \end{bmatrix} u(k)$$

$$y(k) = [\tilde{C}_1, 0] \begin{bmatrix} \tilde{x}^1(k) \\ \tilde{x}^2(k) \end{bmatrix}$$

where  $\tilde{x}^1(k) \in \text{GF}(q)^s$ ,  $\tilde{A}_{11} \in \text{GF}(q)^{s \times s}$ ,  $\tilde{B}_1 \in \text{GF}(q)^{s \times m}$ , and  $\tilde{C}_1 \in \text{GF}(q)^{r \times s}$ . Furthermore, the  $s$ -dimensional LSM  $\tilde{M} = (\tilde{A}_{11}, \tilde{B}_1, \tilde{C}_1)$  is observable and hence is the minimal form of and zero-state equivalent to  $M$ .

As an important consequence of this theorem, we conclude that the dual of Theorem 4.2.2 provides a new state minimization algorithm for LSMs.

The above results lead to the following irreducibility criterion for LSMs.

Theorem 10.3.2. The LSM  $M = (A, B, C)$  is irreducible if and only if it is both reachable and observable.

Proof. If  $M$  is either unreachable or unobservable then by Theorem 4.2.1 and Theorem 10.3.1 it is reducible. Therefore, assume that  $M$  is both reachable and observable, and that there exists an LSM

$\overset{V}{M} = (\overset{V}{A}, \overset{V}{B}, \overset{V}{C})$  of dimension  $n_1 < n$  that is zero-state equivalent to  $M$ , that is,

$$\sum_{j=0}^{n-1} CA^{n-j-1}Bu(j) = \sum_{j=0}^{n-1} \overset{VV}{CA}^{n-j-1}\overset{V}{Bu}(j)$$

which implies that

$$CA^{\ell-1}B = \overset{VV}{CA}^{\ell-1}\overset{V}{B}, \ell \in \underline{n} \quad (10.3.1)$$

Consider the product of the observability and reachability matrices of  $M$

$$\begin{aligned} LK &\equiv \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{bmatrix} [B, AB, \dots, A^{n-1}B] \\ &= \begin{bmatrix} CB & CAB & \dots & CA^{n-1}B \\ CAB & CA^2B & \dots & CA^nB \\ \vdots & \vdots & & \vdots \\ CA^{n-1}B & CA^nB & \dots & CA^{2(n-1)}B \end{bmatrix} \end{aligned} \quad (10.3.2)$$

In view of (10.3.1), (10.3.2) becomes

$$\begin{aligned}
LK &= \begin{bmatrix} \overset{VV}{CB} & \overset{VVV}{CAB} & \dots & \overset{VV}{CA^{n-1}B} \\ \overset{VVV}{CAB} & \overset{VV}{CA^2B} & \dots & \overset{VV}{CA^nB} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \overset{VV}{CA^{n-1}B} & \overset{VV}{CA^nB} & \dots & \overset{VV}{CA^{2(n-1)}B} \end{bmatrix} \\
&= \begin{bmatrix} \overset{V}{C} \\ \overset{VV}{CA} \\ \cdot \\ \cdot \\ \cdot \\ \overset{VV}{CA^{n-1}} \end{bmatrix} \begin{bmatrix} \overset{V}{B}, \overset{VV}{AB}, \dots, \overset{V}{A^{n-1}B} \end{bmatrix} \\
&\equiv \overset{VV}{LK} \tag{10.3.3}
\end{aligned}$$

By hypothesis,  $\text{rank } L = \text{rank } K = n$  and hence  $\text{rank } LK = n$ . Since  $\overset{VV}{LK} \in \text{GF}(q)^{n \times n_1}$ , the maximum rank of  $\overset{VV}{LK}$  can be  $n_1$ . But in view of (10.3.3),  $\text{rank } LK = \overset{VV}{\text{rank } LK}$ , that is,  $n = n_1$  which is obviously a contradiction since  $n_1 < n$ . Therefore, if an LSM is both reachable and observable, then it is irreducible.  $\square$

#### 10.4. State Observer Design for LSMs

The utilization of the important concept of state feedback in the design of various compensation schemes for LSMs obviously hinges upon the availability of all the state variables. However, in practice some or all of the state variables may not be available because they may not be accessible for direct measurement or the number of measuring devices may be limited. Therefore, if the state of the LSM is to be

used in some synthesis or design process, then a reasonable substitute for the state vector must be found. This is usually accomplished by designing an auxiliary data processor, called a *state observer*, a *state estimator*, or a *state reconstructor*. A state observer is essentially an LSM which uses the input and output sequences of the original LSM and after a finite number of clock periods reconstructs the state vector of the given LSM without error regardless of the error in the initial estimate of the state vector. The property of observability plays a central role in this state reconstruction process. In order to demonstrate this process for LSMs, we will first discuss a state observer design procedure for single-output LSMs and then show that for any observable multivariable LSM a state observer can be constructed which would consist of an assemblage of state observers for single-output LSMs.

Consider the single-output LSM

$$M_1 : x(k+1) = Ax(k) + bu(k) \quad (10.4.1a)$$

$$y(k) = c^T x(k) \quad (10.4.1b)$$

and assume that  $A$ ,  $b$ , and  $c$  are completely known, but the state vector  $x(k)$  is not accessible for direct measurement. Then the problem of state observer design is to determine another LSM

$$\hat{M}_1 : \hat{x}(k+1) = \hat{A}\hat{x}(k) + \hat{b}u(k) + \hat{h}y(k) \quad (10.4.2)$$

which accepts  $u(k)$  and  $y(k)$  of  $M_1$  as its inputs and after a finite number of clock periods produces an estimate  $\hat{x}(k)$  of  $x(k)$  without error regardless of the error in the initial estimate  $\hat{x}(0)$  of  $x(0)$ . Clearly if we can determine  $\hat{A}$ ,  $\hat{b}$ , and  $h$  such that the error vector

$$\bar{x}(k) = x(k) - \hat{x}(k) \quad (10.4.3)$$

approaches the zero vector in a finite number of steps, then  $\hat{x}(k)$  is the desired estimate of  $x(k)$ , and thus the design problem is solved.

From (10.4.1) - (10.4.3) it follows that the error dynamics can be described as

$$\bar{x}(k+1) = \hat{A}\bar{x}(k) + (A - \hat{A} - hc^T)\bar{x}(k) + (b - \hat{b})u(k) \quad (10.4.4)$$

Since we want  $\bar{x}(k)$  to be an equilibrium state of (10.4.4), we must choose

$$\hat{A} = A - hc^T, \quad \hat{b} = b \quad (10.4.5)$$

which reduces (10.4.4) to the autonomous LSM

$$\bar{x}(k+1) = (A - hc^T)\bar{x}(k) \quad (10.4.6)$$

In view of (10.4.5), the observer (10.4.2) becomes

$$\hat{x}(k+1) = \hat{A}\hat{x}(k) + \hat{b}u(k) + h(y(k) - c^T\hat{x}(k)) \quad (10.4.7)$$

A combined realization diagram of (10.4.1) and (10.4.7) is shown in Fig. 10.4.1.

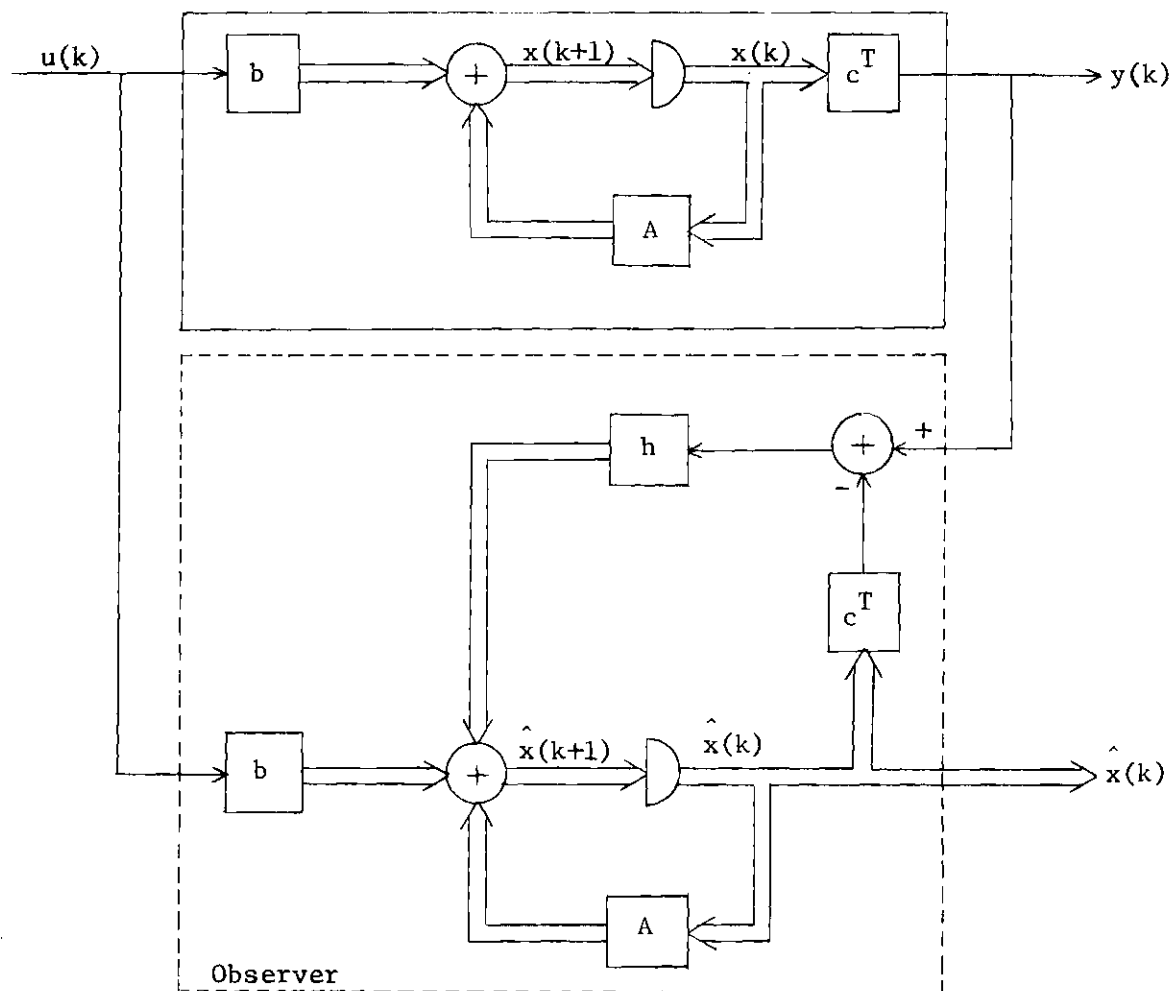


Fig. 10.4.1. State Observer Realization Diagram.

In order to choose  $h$  to force  $\hat{x}(k)$  to zero in a finite number of steps, we consider the following two possibilities for the known initial estimate  $\hat{x}(0)$  of  $x(0)$ ;

- (i) if  $\hat{x}(0) = x(0)$ , then solving (10.4.1) and (10.4.7) recursively, we obtain

$$y(0) - c^T \hat{x}(0) = 0, \quad \bar{x}(1) = A\hat{x}(0) + bu(0) = x(1)$$

$$y(1) - c^T \hat{x}(1) = 0, \quad \bar{x}(2) = A\hat{x}(1) + bu(1) = x(2)$$

⋮

$$y(\ell) - c^T \hat{x}(\ell) = 0, \quad \bar{x}(\ell+1) = A\hat{x}(\ell) + bu(\ell) = x(\ell+1)$$

Therefore, in the present case the observer (10.4.7) generates the exact value of the state regardless of the choice of  $h$ .

(ii) if  $\hat{x}(0) \neq x(0)$ , then solving (10.4.6) recursively, we obtain after the  $\ell$  th iteration

$$\bar{x}(\ell) = (A - hc^T)^{\ell} \bar{x}(0) \quad (10.4.8)$$

The error vector  $\bar{x}$  in (9.4.8) will approach the zero vector if and only if there exists a vector  $h$  such that

$$(A - hc^T)^{\ell} = 0 \quad (10.4.9)$$

But this is precisely the dual statement of the reachability property 8<sup>o</sup> of Theorem 6.1.1. That is, there exists a vector  $h$  such that (10.4.9) holds if and only if the LSM (10.4.1) is observable. Furthermore, the smallest integer  $\ell$  for which (10.4.9) holds is the observability index of  $(A, b, c^T)$ . Therefore, for an observable LSM  $M_1 = (A, b, c^T)$  it is always possible to construct the desirable observer. In fact, this state observer will be time-optimal in the sense that it will produce an errorless estimate of the state vector in the smallest number of clock periods, equal to the observability index of the given LSM  $M_1 = (A, b, c^T)$ .

Suppose that the LSM  $M_1 = (A, b, c^T)$  is  $n$ -observable. In order to determine the required observer gain  $h$ , we apply the matrix identity (5.3.4) to  $(A - hc^T)^n = 0$  and obtain

$$(A - hc^T)^n = A^n - [(A - hc^T)^{n-1}h, (A - hc^T)^{n-2}h, \dots, (A - hc^T)h, h] \begin{pmatrix} c^T \\ c^T A \\ \vdots \\ c^T A^{n-1} \end{pmatrix} = 0$$

Thus

$$[(A - hc^T)^{n-1}h, (A - hc^T)^{n-2}h, \dots, (A - hc^T)h, h] = A^n \begin{pmatrix} c^T \\ c^T A \\ \vdots \\ c^T A^{n-1} \end{pmatrix}^{-1} \quad (10.4.10)$$

Therefore, choosing  $h$  to be the  $n$ th column of the matrix on the right side of (10.4.10), gives the desired observer gain.

We will illustrate the above procedure by an example. Consider the following LSM  $M_1 = (A, b, c^T)$  over  $GF(2)$ :



$$\begin{pmatrix} x_1(k+1) \\ x_2(k+1) \\ x_3(k+1) \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} x_1(k) \\ x_2(k) \\ x_3(k) \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} u(k)$$

(10.4.11)

$$y(k) = [1 \quad 0 \quad 1] \begin{pmatrix} x_1(k) \\ x_2(k) \\ x_3(k) \end{pmatrix}$$

Since

$$\text{rank} \begin{pmatrix} c^T \\ c^T A \\ c^T A^2 \end{pmatrix} = \text{rank} \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \end{pmatrix} = 3$$

the given LSM is 3-observable. In order to determine  $h$ , we need to compute the third column of the matrix

$$A^3 \begin{pmatrix} c^T \\ c^T A \\ c^T A^2 \end{pmatrix}^{-1}$$

By direct computation, we find that  $A^3 = I_3$  and hence

$$A^3 \begin{pmatrix} c^T \\ c^T A \\ c^T A^2 \end{pmatrix}^{-1} = \begin{pmatrix} c^T \\ c^T A \\ c^T A^2 \end{pmatrix}^{-1} = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix}$$

Therefore,

$$h = \begin{pmatrix} h_1 \\ h_2 \\ h_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

Thus the time-optimal observer is given by

$$\begin{pmatrix} \hat{x}_1(k+1) \\ \hat{x}_2(k+1) \\ \hat{x}_3(k+1) \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} \hat{x}_1(k) \\ \hat{x}_2(k) \\ \hat{x}_3(k) \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} u(k) \\ + \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} (y(k) - \hat{x}_1(k) - \hat{x}_3(k))$$

A combined realization circuit for the given LSM and its associated state observer is shown in Fig. 10.4.2.

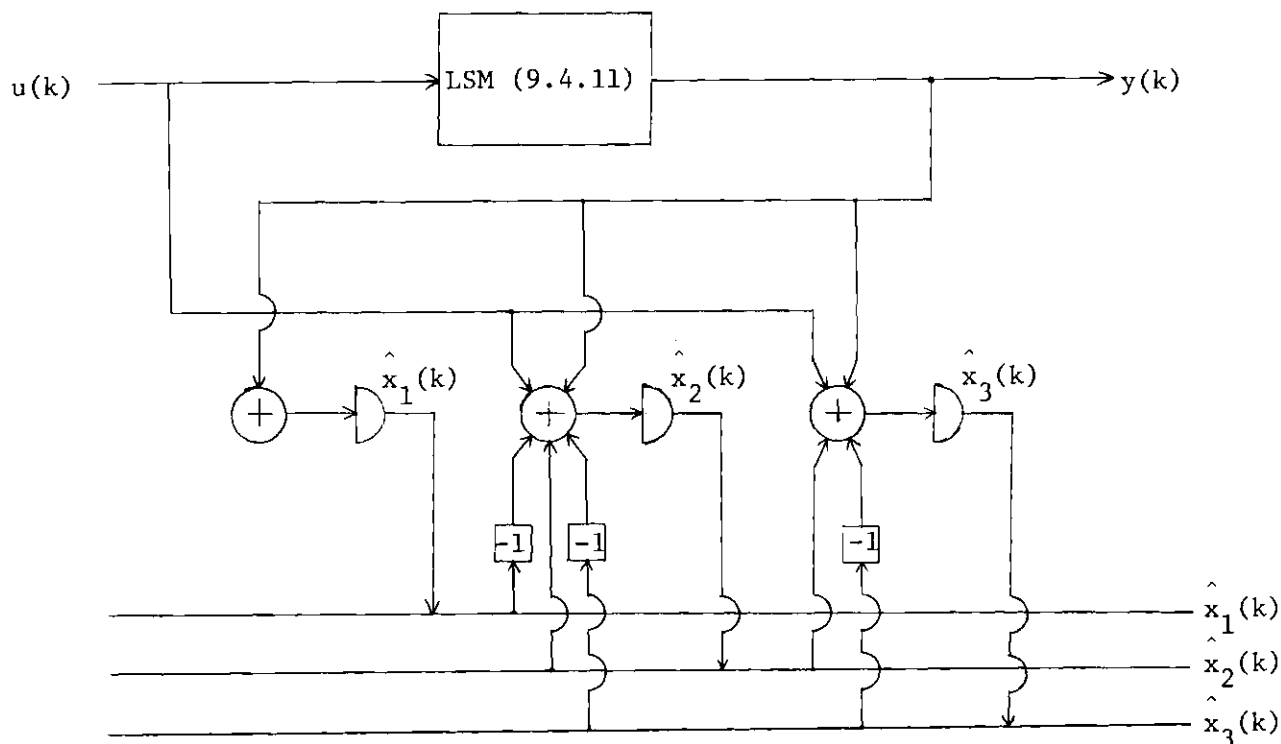


Fig. 10.4.2. Realization Circuit for the LSM of Example 10.4.1.

At this point, we would like to pause momentarily for consideration of some questions that arise naturally concerning the effect of feedback that employs the estimated state vector  $\hat{x}(k)$  instead of the original state vector  $x(k)$ . Let the LSM  $M_1(A, b, c^T)$  be reachable and observable. Then two questions of interest are: (i) Will the characteristic polynomial of  $M_1$  be preserved under a state feedback of the form

$$u(k) = f^T \hat{x}(k) + v(k) \quad (10.4.12)$$

where the vector  $f$  has been chosen with respect to  $x(k)$ ? (ii) What is the overall effect of introducing the observer in the LSM  $M_1$ ?

In order to resolve these questions, we need to determine the overall description of the given LSM and its associated observer. Substituting (10.4.1b) and (10.4.12) into (10.4.1a) and (10.4.7), we obtain the desired combined representation as follows:

$$\begin{bmatrix} x(k+1) \\ \hat{x}(k+1) \end{bmatrix} = \begin{bmatrix} A+bf^T & 0 \\ hc^T & A-hc^T+bf \end{bmatrix} \begin{bmatrix} x(k) \\ \hat{x}(k) \end{bmatrix} + \begin{bmatrix} b \\ b \end{bmatrix} v(k)$$

Now using the special state isomorphism  $x \mapsto Px \equiv \begin{bmatrix} x \\ \hat{x} \end{bmatrix}$ , where

$$P = P^{-1} \begin{bmatrix} I_n & 0 \\ 0 & I_n \end{bmatrix}$$

(10.4.13) is transformed to the following isomorphic representation:

$$\begin{bmatrix} \tilde{x}(k+1) \\ \tilde{\hat{x}}(k+1) \end{bmatrix} = \begin{bmatrix} A+bf^T & -bf^T \\ 0 & A-hc^T+bf \end{bmatrix} \begin{bmatrix} \tilde{x}(k) \\ \tilde{\hat{x}}(k) \end{bmatrix} + \begin{bmatrix} b \\ 0 \end{bmatrix} v(k) \quad (10.4.14)$$

From (10.4.14) it follows that the characteristic polynomial of the combined LSM (10.4.13) is equal to the product of the characteristic polynomials of the given LSM  $M_1$  and the associated observer  $\hat{M}_1$ .

Therefore, there is no difference in state feedback between using  $\hat{x}(k)$  or  $x(k)$ , and the characteristic polynomial of  $M_1$  remains invariant.

In view of the above *separation property*, the design of a state feedback and that of a state observer can be carried out independently.

#### (n-1)-Dimensional State Observer for Single-Output LSMs

The state observer  $\hat{M}_1$  given by (10.4.7) is clearly of dimension  $n$ , that is,  $\hat{M}_1$  estimates all the  $n$  components  $x_i(k)$  of  $M_1$ . However, with the help of the following result, it is always possible to replace the  $n$ -dimensional observer (10.4.7) by an  $(n-1)$ -dimensional one and still obtain an estimate of the entire state vector. This reduction in dimensionality results into considerable computational and storage savings for observable multivariable LSMs.

Theorem 10.4.1. Suppose that the LSMs  $M_1 = (A, b, c^T)$  and  $\tilde{M}_1 = (\tilde{A}, \tilde{b}, \tilde{c}^T) \equiv (PAP^{-1}, Pb, c^T P^{-1})$  are isomorphic under the isomorphism  $P : X \rightarrow X$ ,  $P \in GF(n, q)$ . Then their observer gains are related as  $\tilde{h} = Ph$ .

Proof. The observer (10.4.7) for the LSM  $\tilde{M}_1$  becomes

$$\tilde{x}(k+1) = \tilde{A}\tilde{x}(k) + \tilde{b}u(k) + \tilde{h}(y(k) - \tilde{c}^T\tilde{x}(k)) \quad (10.4.15)$$

On the other hand, the isomorphism  $\hat{x}(k) \mapsto P\hat{x}(k) \equiv \tilde{\hat{x}}(k)$  transforms (10.4.7) to

$$\tilde{\hat{x}}(k+1) = PAP^{-1}\tilde{\hat{x}}(k) + pbu(k) + Ph(y(k) - c^T P^{-1}\tilde{\hat{x}}(k)) \quad (10.4.16)$$

Comparing (10.4.15) and (10.4.16) yields the desired result.  $\square$

Therefore, by virtue of this result we can use any canonical form of the given LSM to design a state observer. This freedom of choice of LSM representation can obviously lead to considerable computational efficiency. For instance, if we choose to work with the canonical form  $(\tilde{A}, \tilde{b}, \tilde{c}^T)$ , where

$$\tilde{A} \equiv \begin{pmatrix} 0 & 1 & 0 & 0 & \dots & 0 \\ 0 & 0 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & 0 & \dots & 1 \\ a_0 & a_1 & a_2 & a_3 & \dots & a_{n-1} \end{pmatrix}; \quad \tilde{c}^T \equiv [1 \ 0 \ \dots \ 0]$$

then the observability matrix appearing in (10.4.11) will be an identity matrix which obviously simplifies the computation of  $\tilde{h}$ .

However, if we employ the above procedure, then the resulting state observer will produce an estimate  $\hat{\tilde{x}}$  of the transformed state  $\tilde{x}$  and not of the state  $x$  of the given LSM. Since  $x$  and  $\tilde{x}$  are related by  $x = P^{-1}\tilde{x}$ , if the estimate  $\hat{\tilde{x}}$  is passed through a device with gain  $P^{-1}$ , then the output  $\hat{x} = P^{-1}\hat{\tilde{x}}$  of this device will give the desired estimate of  $x$ .

In the ensuing discussion of designing an  $(n-1)$ -dimensional state observer we will make use of a particular canonical form which can be obtained by dualizing (5.2.3) as follows:

Theorem 10.4.2. If the LSM  $M_1 = (A, b, c^T)$  is observable, then there exists an isomorphism  $P : X \rightarrow X$ ,  $P \in GF(n, q)$ , such that the isomorphic LSM  $\tilde{M}_1 = (\tilde{A}, \tilde{b}, \tilde{c}^T) \equiv (PAP^{-1}, Pb, c^TP^{-1})$  has the canonical form

$$\begin{pmatrix} \tilde{x}_1(k+1) \\ \tilde{x}_2(k+1) \\ \vdots \\ \tilde{x}_n(k+1) \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & \dots & 0 & a_0 \\ 1 & 0 & 0 & \dots & 0 & a_1 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & a_{n-1} \end{pmatrix} \begin{pmatrix} \tilde{x}_1(k) \\ \tilde{x}_2(k) \\ \vdots \\ \tilde{x}_n(k) \end{pmatrix} + \begin{pmatrix} \tilde{b}_1 \\ \tilde{b}_2 \\ \vdots \\ \tilde{b}_n \end{pmatrix} u(k) \quad (10.4.17)$$

$$y(k) = [0, 0, \dots, 0, 1] \begin{pmatrix} \tilde{x}_1(k) \\ \tilde{x}_2(k) \\ \vdots \\ \tilde{x}_n(k) \end{pmatrix}$$

where  $f_c(\lambda) = (\lambda)^n - a_{n-1}(\lambda)^{n-1} - \dots - a_1\lambda - a_0$  is the characteristic polynomial of  $M_1$ . The isomorphism  $P$  is given by

$$P \equiv \begin{pmatrix} -a_1 & -a_2 & \dots & -a_{n-1} & 1 \\ -a_2 & -a_3 & \dots & 1 & \\ \vdots & \vdots & & \vdots & \\ -a_{n-1} & 1 & & & \\ 1 & & & & \end{pmatrix} \begin{pmatrix} c^T \\ c^T A \\ \vdots \\ c^T A^{n-2} \\ c^T A^{n-1} \end{pmatrix}$$

From (10.4.17) we see that  $y(k) = \tilde{x}_n(k)$ , and hence the last state component is known and measurable. Thus there is no need to estimate  $\tilde{x}_n(k)$ . Consequently we need only estimate the first  $(n-1)$  components of  $\tilde{x}(k)$ .

We now claim that the following  $(n-1)$ -dimensional LSM is a state observer for (10.4.17):

$$\begin{pmatrix} \hat{x}_1(k+1) \\ \hat{x}_2(k+1) \\ \hat{x}_3(k+1) \\ \vdots \\ \hat{x}_{n-1}(k+1) \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & \dots & 0 \\ 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix} \begin{pmatrix} \hat{x}_1(k) \\ \hat{x}_2(k) \\ \hat{x}_3(k) \\ \vdots \\ \hat{x}_{n-1}(k) \end{pmatrix} + \begin{pmatrix} \tilde{b}_1 \\ \tilde{b}_2 \\ \tilde{b}_3 \\ \vdots \\ \tilde{b}_{n-1} \end{pmatrix} u(k) + \begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ \vdots \\ a_{n-2} \end{pmatrix} y(k) \quad (10.4.18)$$

To see this, let

$$z(k) \equiv \begin{pmatrix} \tilde{x}_1(k) \\ \tilde{x}_2(k) \\ \vdots \\ \tilde{x}_{n-1}(k) \end{pmatrix}; \quad \bar{A} \equiv \begin{pmatrix} 0 & 0 & 0 & \dots & 0 \\ 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix}; \quad \bar{b} \equiv \begin{pmatrix} \tilde{b}_1 \\ \tilde{b}_2 \\ \vdots \\ \tilde{b}_{n-1} \end{pmatrix} \quad (10.4.19)$$

$$\hat{z}(k) \equiv \begin{pmatrix} \hat{x}_1(k) \\ \hat{x}_2(k) \\ \vdots \\ \hat{x}_{n-1}(k) \end{pmatrix}; \quad \bar{a} \equiv \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_{n-2} \end{pmatrix}$$



In terms of the above notation and  $\tilde{x}_n(k)$ , the inaccessible states of the LSM (10.4.17) can be expressed as follows:

$$\tilde{z}(k+1) = \bar{A}\tilde{z}(k) + \bar{a}\tilde{x}_n(k) + \bar{b}u(k) \quad (10.4.20)$$

Since from (10.4.17)  $y(k) = \tilde{x}_n(k)$ , (10.4.20) becomes

$$\tilde{z}(k+1) = \bar{A}\tilde{z}(k) + \bar{a}y(k) + \bar{b}u(k) \quad (10.4.21)$$

Similarly, rewriting the observer description (10.4.18) in terms of the notation (10.4.19), we obtain

$$\hat{z}(k+1) = \bar{A}\hat{z}(k) + \bar{b}u(k) + \bar{a}y(k) \quad (10.4.22)$$

In view of (10.4.21) and (10.4.22), the error can be expressed as

$$\tilde{z}(k+1) - \hat{z}(k+1) = \bar{A}[\tilde{z}(k) - \hat{z}(k)] \quad (10.4.23)$$

Solving (10.4.23) recursively, we obtain

$$\tilde{z}(k) - \hat{z}(k) = \bar{A}^\ell [\tilde{z}(0) - \hat{z}(0)], \ell = 0, 1, \dots$$

But  $\bar{A}^\ell = 0$  for  $\ell \geq n-1$ , that is, the matrix  $\bar{A}$  is  $(n-1)$ -nilpotent since its characteristic polynomial is  $f_c(\lambda) = (\lambda)^{n-1}$  and by the Cayley-Hamilton Theorem  $f_c(\bar{A}) = \bar{A}^{n-1} = 0$ . Therefore, the estimate  $\hat{z}$  becomes equal to the true value  $\tilde{z}$  of the state after the  $n$ th iteration regardless of the quality of the initial estimate  $\hat{z}(0)$ . Thus (10.4.18) is an  $(n-1)$ -dimensional state observer for the LSM (10.4.17).

Recall that the LSM  $\tilde{M}_1 = (\tilde{A}, \tilde{b}, \tilde{c}^T)$  given by (10.4.17) was obtained from the original LSM  $M_1 = (A, b, c^T)$  by the state transformation

$x \mapsto Px \equiv \tilde{x}$ , that is,  $\tilde{M}_1 = (\tilde{A}, \tilde{b}, \tilde{c}^T) \equiv (PAP_o, Pb, c^TP^{-1})$ . Therefore, the state observer (10.4.18) generates  $\hat{\tilde{x}}_i(k)$ ,  $i \in \underline{n-1}$ . In order to obtain  $\hat{x}_i(k)$ ,  $i \in \underline{n-1}$ , we simply multiply  $\hat{\tilde{x}}_i(k)$  by  $P^{-1}$ , that is,  $\hat{x}_i(k) = P^{-1}\hat{\tilde{x}}_i(k)$ ,  $i \in \underline{n-1}$ .

To illustrate the above procedure, we will construct a 2-dimensional state observer for the LSM (10.4.11). The first step is to transform the LSM (10.4.11) to (10.4.17) using the isomorphism  $P$  specified in Theorem 10.4.2. Thus we compute

$$P = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}; \quad P^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}$$

The LSM (10.4.11) is transformed to

$$\begin{bmatrix} \tilde{x}_1(k+1) \\ \tilde{x}_2(k+1) \\ \tilde{x}_3(k+1) \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \tilde{x}_1(k) \\ \tilde{x}_2(k) \\ \tilde{x}_3(k) \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} u(k)$$

(10.4.24)

$$y(k) = [0 \quad 0 \quad 1] \begin{bmatrix} \tilde{x}_1(k) \\ \tilde{x}_2(k) \\ \tilde{x}_3(k) \end{bmatrix}$$

Therefore, the 2-dimensional state observer is

$$\begin{bmatrix} \hat{x}_1(k+1) \\ \hat{x}_2(k+1) \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \hat{x}_1(k) \\ \hat{x}_2(k) \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \end{bmatrix} u(k) + \begin{bmatrix} 1 \\ 0 \end{bmatrix} y(k)$$

A realization circuit for the generation of  $\hat{x}(k)$  is shown in Fig. 10.4.3.

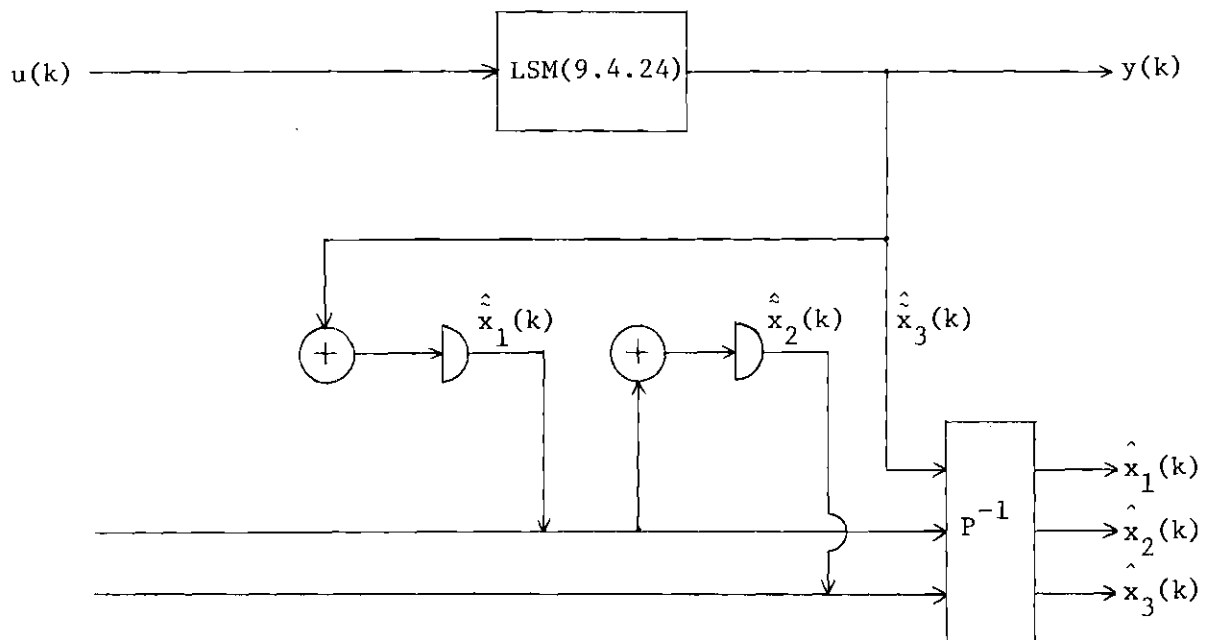


Fig. 10.4.3. Realization Circuit for the 2-Dimensional State Observer for the LSM of Example 10.4.11.

(n-r)-Dimensional State Observer for Multivariable LSMs

The results derived for the single-output LSM can be easily extended to the more general multivariable case. To this end, we will choose a particular quasi-canonical form for the LSM  $M = (A, B, C)$ . Assuming that  $M$  is observable and  $C$  has full rank, the reachable quasi-canonical representation (5.1.27) can be dualized as follows: Let the matrix  $C \in GF(q)^{r \times n}$  be partitioned as

$$C \equiv \begin{pmatrix} c^1 \\ c^2 \\ \cdot \\ \cdot \\ \cdot \\ c^r \end{pmatrix}, \quad c^i \in GF(q)^{1 \times n}, \quad i \in \underline{r} \quad (10.4.25)$$

Because of the observability assumption, the observability matrix

$$\begin{pmatrix} C \\ CA_2 \\ CA^2 \\ \cdot \\ \cdot \\ CA^{n-1} \end{pmatrix} = \begin{pmatrix} c^1 \\ c^2 \\ \cdot \\ \cdot \\ \cdot \\ c^r \\ c^1 A \\ c^2 A \\ \cdot \\ \cdot \\ \cdot \\ c^r A \\ \cdot \\ \cdot \\ \cdot \\ c^r A^{n-1} \end{pmatrix}$$

has  $n$  linearly independent rows. We first choose these independent rows in the order  $c^1, c^2, \dots, c^r, c^1A, c^2A, \dots, c^rA, c^1A^2, \dots$ , retaining the independent ones and discarding those that are linearly dependent on their predecessors, and then rearrange them to form the following matrix:

$${}^v_P \equiv \begin{pmatrix} c^1 \\ c^1A \\ \cdot \\ \cdot \\ \cdot \\ c^1A^{m_1-1} \\ c^2A \\ c^2A^2 \\ \cdot \\ \cdot \\ \cdot \\ c^2A^{m_2-1} \\ \cdot \\ \cdot \\ \cdot \\ c^rA^{m_r-1} \end{pmatrix}$$

where  $m_i$  are the observability indices of  $M = (A, B, C)$ , and thus satisfy the relation  $m_1 + m_2 + \dots + m_r = n$ .

Let  $w^{ij}$ ,  $i \in \underline{r}$ ,  $j \in \underline{m_i}$ , denote the columns of  ${}^v_{P^{-1}}$  and write  ${}^v_{P^{-1}}$  in terms of its columns as follows:

$${}^v_{P^{-1}} \equiv [w^{11}, w^{12}, \dots, w^{1m_1}; w^{21}, w^{22}, \dots, w^{2m_2}; \dots; w^{rm_r}]$$

Let  $\mu_\ell = m_1 + m_2 + \dots + m_\ell$ ,  $\ell \in \underline{r}$ . Then using the  $\mu_\ell$ -th columns of  $P^{-1}$ , we form the following matrix:

$$P \equiv \begin{bmatrix} w^{1m_1}, Aw^{1m_1}, \dots, A^{m_1-1} w^{1m_1}; w^{2m_2}, Aw^{2m_2}, \dots, A^{m_2-1} w^{2m_2}; \\ \dots; w^{rm_r}, Aw^{rm_r}, \dots, A^{m_r-1} w^{rm_r} \end{bmatrix}$$

Since  $PP^{-1} = I_n$  and consequently

$$\begin{aligned} c^s A_w^{t,ij} &= 1 \text{ if } i = s \text{ and } j = t+1 \\ &= 0 \text{ otherwise} \end{aligned}$$

it is easy to show that the columns of  $P$  are linearly independent and thus form a basis for  $X$ . By direct computation it is easy to see that the isomorphic LSM  $\tilde{M} = (\tilde{A}, \tilde{B}, \tilde{C}) \equiv (PAP^{-1}, PB, CP^{-1})$  has the following quasi-canonical form:

$$\begin{aligned} \begin{bmatrix} \tilde{x}^1(k+1) \\ \tilde{x}^2(k+1) \\ \vdots \\ \tilde{x}^r(k+1) \end{bmatrix} &= \begin{bmatrix} \tilde{A}_{11} & \tilde{A}_{12} & \dots & \tilde{A}_{1r} \\ \tilde{A}_{21} & \tilde{A}_{22} & \dots & \tilde{A}_{2r} \\ \vdots & \vdots & & \vdots \\ \tilde{A}_{r1} & \tilde{A}_{r2} & \dots & \tilde{A}_{rr} \end{bmatrix} \begin{bmatrix} \tilde{x}^1(k) \\ \tilde{x}^2(k) \\ \vdots \\ \tilde{x}^r(k) \end{bmatrix} + \begin{bmatrix} \tilde{B}_1 \\ \tilde{B}_2 \\ \vdots \\ \tilde{B}_r \end{bmatrix} u(k) \quad (10.4.26) \\ \begin{bmatrix} y_1(k) \\ y_2(k) \\ \vdots \\ y_r(k) \end{bmatrix} &= \begin{bmatrix} 0 & 0 & \dots & 0 & 1 & 0 & 0 & \dots & 0 & 0 & \vdots & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 1 & \vdots & 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots & \ddots & \vdots & \vdots & & \vdots & \vdots \\ \vdots & \vdots & & \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots & & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 & \vdots & 0 & 0 & \dots & 0 & 1 \end{bmatrix} \begin{bmatrix} \tilde{x}^1(k) \\ \tilde{x}^2(k) \\ \vdots \\ \tilde{x}^r(k) \end{bmatrix} \quad (10.4.27) \end{aligned}$$

where

$$\tilde{x}^i(k) \equiv \begin{pmatrix} \tilde{x}_{i1}(k) \\ \tilde{x}_{i2}(k) \\ \vdots \\ \tilde{x}_{im_i}(k) \end{pmatrix}, \quad i \in \underline{r} \quad (10.4.28)$$

$$\tilde{A}_{ii} \equiv \begin{pmatrix} 0 & 0 & \dots & 0 & * \\ 1 & 0 & \dots & 0 & * \\ 0 & 1 & \dots & 0 & * \\ \vdots & \vdots & & \vdots & \vdots \\ \vdots & \vdots & & \vdots & \vdots \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \dots & 1 & * \end{pmatrix} \in GF(q)^{m_i \times m_i}, \quad i \in \underline{r} \quad (10.4.29)$$

$$\tilde{A}_{ij} \equiv \begin{pmatrix} 0 & 0 & \dots & 0 & * \\ 0 & 0 & \dots & 0 & * \\ \vdots & \vdots & & \vdots & \vdots \\ \vdots & \vdots & & \vdots & \vdots \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \dots & 0 & * \end{pmatrix} \in GF(q)^{m_i \times m_j}, \quad i \in \underline{r}, \quad j \in \underline{m_i} \quad (10.4.30)$$

and \*'s denote possibly nonzero entries.

Now consider any  $m_i$ -dimensional,  $i \in \underline{r}$ , submachine of the above quasi-canonical LSM with state vector of the form (10.4.28). Then the  $i$ -th submachine can be expressed as follows:

$$\tilde{x}^i(k+1) = \tilde{A}_{ii} \tilde{x}^i(k) + \sum_{\substack{j=1 \\ j \neq i}}^r \tilde{A}_{ij} \tilde{x}^j(k) + \tilde{B}_i u(k), \quad i \in \underline{r} \quad (10.4.31)$$

However, from (10.4.30) it is clear that the first  $(m_i-1)$  columns of the matrices  $\tilde{A}_{ij}$ ,  $i \in \underline{r}$ ,  $j \in \underline{m_i}$ , are identically zero and thus

$$\tilde{A}_{ij} \tilde{x}^i(k) = \tilde{a}_{m_j}^i \tilde{x}_{jm_j}^i(k), \quad i \in \underline{r}, \quad j \in \underline{m_i} \quad (10.4.32)$$

where  $\tilde{a}_{m_j}^i$  is the  $m_j$ th column of  $\tilde{A}_{ij}$ . Substituting (10.4.32) into the second term on the right side of (10.4.31) and observing that in view of (10.4.27),  $y_i(k) = \tilde{x}_{im_i}^i(k)$ ,  $i \in \underline{r}$ , we have

$$\begin{pmatrix} \tilde{x}_{i1}^i(k+1) \\ \tilde{x}_{i2}^i(k+1) \\ \tilde{x}_{i3}^i(k+1) \\ \vdots \\ \tilde{x}_{im_i}^i(k+1) \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & \dots & 0 & * \\ 1 & 0 & 0 & \dots & 0 & * \\ 0 & 1 & 0 & \dots & 0 & * \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & * \end{pmatrix} \begin{pmatrix} \tilde{x}_{i1}^i(k) \\ \tilde{x}_{i2}^i(k) \\ \tilde{x}_{i3}^i(k) \\ \vdots \\ \tilde{x}_{im_i}^i(k) \end{pmatrix} + \sum_{\substack{j=1 \\ j \neq i}}^r \tilde{a}_{m_j}^i y_j(k) + \tilde{B}_i u(k), \quad i \in \underline{r} \quad (10.4.33)$$

Therefore, the general multivariable LSM (10.4.26) - (10.4.30) has been reduced to  $r$  individual multi-input single-output submachines of the form (10.4.33), each driven by the directly measurable signals  $u(k)$  and  $y(k)$ , and each characterized by a known state component  $\tilde{x}_{im_i}^i(k) = y_i(k)$ ,  $i \in \underline{r}$ . Since the input does not play any role in the design of state observers, and the submachines (10.4.33) are single-output and  $m_i$ -observable,  $i \in \underline{r}$ , we can employ the  $(n-1)$ -dimensional state observer



design procedure, developed in the preceding subsection for single-output LSMs,  $r$  times for (10.4.33). Since each submachine of (10.4.33) requires an  $(m_i-1)$ -dimensional observer, an overall observer of total dimension  $\sum_{i=1}^r (m_i-1) = n-r$  can be employed to estimate the entire state vector  $x(k)$  of the given multivariable LSM.

### Summary and Conclusions

In this chapter the concept of state observability of LSMs which is dual to that of state reachability, and some other related topics were investigated. More specifically, after demonstrating the duality relationship between state observability and state reachability, the duality theorem of Kalman [61] was proved for LSMs, and its consequences were discussed. As an illustration of the significance of state observability property in estimating the inaccessible state components of an LSM, some design procedures for full- and reduced order state observers of Luenberger type for both single- and multi-input LSMs were presented (cf. [69], [70], [71], [106]).

## CHAPTER XI

## CONCLUSIONS AND RECOMMENDATIONS

The primary objective of the research reported in this dissertation was to investigate the possibility of developing a structure theory for the important class of finite-state linear sequential machines in the framework of modern multivariable control theory. Examining the parallelisms and interconnections between the disciplines of sequential machines and control systems, selecting the state reachability property as the pivotal component of this structure theory, and adopting a finite-geometric point of view, we have introduced and developed in detail a number of concepts for linear sequential machines that may be summarized as follows:

The important concepts of state reachability and state controllability were extensively investigated. Twenty-four state reachability criteria for single-input LSMs and eighteen state reachability criteria for multi-input LSMs were formulated. In each case the equivalence of these criteria was proved. Furthermore, some equivalence classes of state reachable LSMs with respect to certain transformation groups were identified. The implications of the state reachability property relative to some structural invariants, canonical forms, and state variable feedback were discussed. In particular, the application of the reachability indices in the invariant description of LSMs via Brunovsky's canonical form was demonstrated.

Exploiting the eigenstructures of LSMs in the framework of generalized eigenvectors and Jordan canonical forms, additional state reachability criteria which explicitly involve the Jordan canonical form were presented. The possibility of maintaining the state reachability property of a reachable multivariable LSM by scalar control sequences was characterized in terms of the Jordan canonical representation of an LSM. The concept of selective state reachability which makes heavy use of the generalized eigenproperties was introduced and developed in detail for LSMs.

Further aspects of the concept of state reachability, in conjunction with state feedback, were studied in the context of the finite projective geometry and certain classes of flats related to the structural properties of LSMs were characterized, their applications to some areas of LSMs were demonstrated, and algorithms for their computation were discussed.

Finally, the concept of state observability, which is dual to that of state reachability, was investigated. A complete duality relationship between the state reachability and state observability properties was established. In view of this duality, all the results pertaining to state reachability can be restated in terms of observability in a simple and direct manner. As an illustration of the central role of the property of observability in the state reconstruction problem, some design procedures for full- and reduced-order state observers for both single- and multi-input LSMs were presented.

### Recommendations for Further Research

For the purpose of developing an integrated multivariable machine control theory, our results can be supplemented by investigating other aspects of time-invariant LSMs such as realization, non-interaction, inversion, identification, decentralization, optimal regulation, and so forth. Similar investigations need to be carried out for deterministic time-varying LSMs, stochastic time-invariant LSMs, and various types of stochastic time-varying LSMs. Most of these areas of LSM are virtually untouched.

The use of formal polynomials and polynomial matrices in the study of various aspects of LSMs, as pointed out in Section 3.6, seems to be a promising area of research. This approach which does not involve the use of any transform techniques can lead to the development of a comprehensive theory for time-invariant LSMs paralleling an existing theory for linear systems promoted by Rosenbrock [97], Wolovich [110], and others, which is based on the Laplace transform method.

In our study of LSMs no combinatorial analysis was utilized. It seems that many combinatorial structures of the finite projective geometry can be used to characterize certain aspects of LSMs. This area, especially in conjunction with geometric coding theory, warrants much further research.

As was demonstrated throughout this dissertation, there exist many similarities between linear machines and conventional linear systems. Unfortunately, at present very few such similarities exist between nonlinear control systems and general automata, and thus there

does not seem to exist any appreciable interchange of ideas between these disciplines. We believe that the investigation of a special class of finite-state quasi-linear systems which we will term *bilinear machines*, will contribute to a better rapprochement between these two areas of dynamical systems. Recently, conventional bilinear systems due to their modeling capability, vast areas of applicability and mathematical tractability have attracted a lot of attention.

A deterministic time-invariant finite-state bilinear sequential machine is described by the following vector difference equations over  $GF(q)$ :

$$x(k+1) = Ax(k) + \sum_{i=1}^m N_i u_i(k)x(k) + Bu(k)$$

$$y(k) = Cx(k)$$

where  $A \in GF(q)^{n \times n}$ ,  $N_i \in GF(q)^{n \times n}$ ,  $i \in \underline{m}$ ,  $B \in GF(q)^{n \times m}$ , and  $C \in GF(q)^{r \times n}$ .

This model can be transformed to an equivalent homogeneous-in-the-state model by defining the following new state variables and characterizing matrices:

$$z(k) \equiv \begin{bmatrix} x(k) \\ 1 \end{bmatrix}; \quad F \equiv \begin{bmatrix} A & 0 \\ 0 & 0 \end{bmatrix}$$

$$G_i \equiv \begin{bmatrix} N_i & b^i \\ 0 & 0 \end{bmatrix}; \quad H \equiv [C \quad 0]$$

where  $b^i \in GF(q)^n$ ,  $i \in \underline{m}$ , are the columns of  $B$ . Thus we have

$$z(k+1) = \left( F + \sum_{i=1}^m G_i u_i(k) \right) z(k)$$

$$y(k) = Hz(k)$$

Due to its "quasi-linear" nature, the class of bilinear machines will constitute a transitional link between linear sequential machines which, as we have seen in this research, lend themselves to thorough analysis, and general nonlinear sequential machines for which no general theory exists. Moreover, from a modeling point of view, bilinear machines seem to be a reasonable compromise between the conflicting demands of accuracy and simplicity. This class of machines will make it possible to overcome, on the one hand, the representability and precision limitations imposed by the linear models and, on the other hand, the theoretical and computational complications associated with more highly nonlinear sequential machine models. Furthermore, due to their intrinsic variable structure and adaptivity characteristics, bilinear sequential machines are more controllable compared to linear machines. This property of bilinear machines is due to the existence of the multiplicative control action on the evolution of the state dynamics of the machine. That is, the input sequence  $u(k) \in U^*$  controls the state dynamics not only additively by means of the term  $Bu(k)$ , but also in a multiplicative way by means of the term  $\sum_{i=1}^m N_i u_i(k)x(k)$ . Thus, intuitively, for a bilinear machine the control should be more effective than for a linear machine with only an additive

control action. Clearly the most desirable feature of bilinear machines is the "quasi-linear" form of their nonlinearity which seems to admit the extension of certain aspects of linear machine theory to bilinear machines.

## APPENDIX

Finite Projective and Affine Spaces and Geometries

For the purpose of easy reference, we have collected, in this appendix, a few basic definitions and facts concerning finite projective and affine geometries. Comprehensive coverage of these subjects is available in [30].

Projective Spaces

Definition 1. A triple  $(P, L, *)$ , where  $P$  and  $L$  are disjoint sets and  $*$  is a relation on  $P \times L$ , that is,  $P \cap L = \emptyset$  and  $* \subseteq P \times L$ , is called an *incidence structure*. The elements of  $P$  are called "*points*," those of  $L$  "*lines*;" and  $*$  is called the "*incidence*" relation, where the terms "point," "line," and "incidence" are undefined. If the ordered pair  $(r, s)$  is an element of  $*$ , we will write  $r * s$ , which is to be read " $r$  is on  $s$ ," or " $r$  and  $s$  are incident," or " $s$  passes through  $r$ ." The incidence structure  $(P, L, *)$  is called *finite* if the sets  $P$  and  $L$  are finite.

If certain conditions, in the form of axioms, are imposed on an incidence structure  $(P, L, *)$ , we will obtain a mathematical structure, called a *projective space* which is defined next.

Definition 2. An incidence structure  $(P, L, *)$  is called a *projective space* if the following conditions are satisfied:



- P1. If  $r, s \in P$  such that  $r \neq s$ , then there exists exactly one line  $\ell \in L$  such that  $r * \ell$  and  $s * \ell$ . This is usually denoted by  $r + s = \ell$ .
- P2. For  $\ell \in L$ , there exist distinct points  $r, s, t \in P$  such that  $r + s + t = \ell$ .
- P3. If  $r, s, t, u, v \in P$  are distinct points such that  $r, s$ , and  $t$  are noncollinear,  $u \in r + s$ , and  $v \in r + t$ , then  $(s + t) \cap (u + v) \in P$ . That is, any line passing through two sides of a triangle at points other than vertices intersect the third side.

Next we will present some illustrative examples of projective spaces.

Example 1. Let  $(F, +, \cdot)$  be a field,  $S$  any set containing at least three elements, and  $P \equiv \{f : S \rightarrow F : f(x) \neq 0 \ \forall x \in S\}$ . For  $f, g \in P$  and  $a \in F$ , define  $f + g$  and  $af$  by  $(f + g)(x) = f(x) + g(x)$  and  $(af)(x) = af(x) \ \forall x \in S$ , respectively, and let  $f = ag$ , denoted by  $f \approx g$ ,  $a \neq 0$ , imply that  $f, g \in P$ . Further, for  $f, g \in P$ , define  $f * g \equiv \{af + bg : a, b \in F \text{ and not both } a = 0, b = 0\}$ . Then for  $f, g, f', g' \in P$ ,  $f * g = f' * g' \iff f \approx f' \text{ and } g \approx g'$ . Finally, let  $L \equiv \{f * g : f, g \in P \text{ and } f \not\approx g\}$ . Then for  $f, g \in P$  such that  $f \approx g$  and  $h \in L$ ,  $f \in h \iff g \in h$ . Then the triple  $(P, L, \epsilon)$  is a projective space.

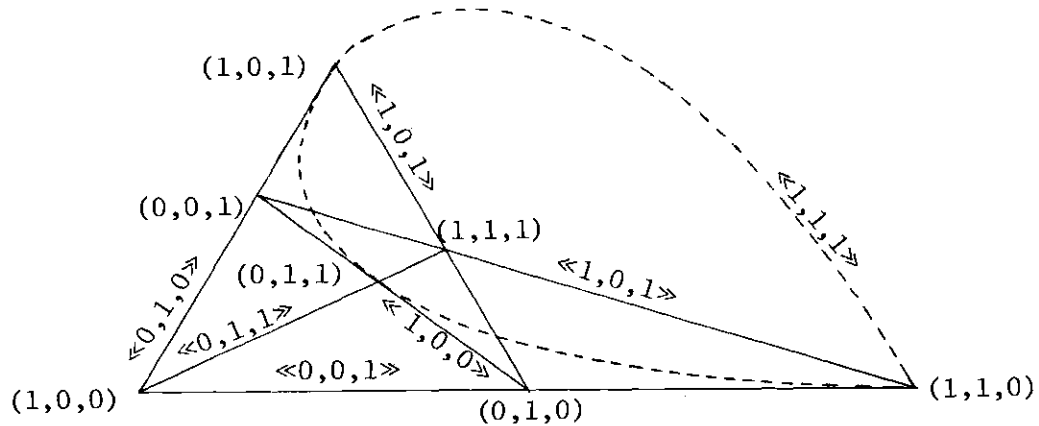
Example 2. Let  $P \equiv \{(x_1, x_2, \dots, x_n) \in F^n : (x_1, x_2, \dots, x_n) \neq (0, 0, \dots, 0) \text{ and } n \geq 3\}$ ,  $x_i = ay_i \implies (x_1, x_2, \dots, x_n) \approx (y_1, y_2, \dots, y_n) \forall (x_1, x_2, \dots, x_n), (y_1, y_2, \dots, y_n) \in F^n$ ,  $\forall a \in F, a \neq 0, i \in \underline{n}$ . Let  $L \equiv \{(ax_1 + by_1, ax_2 + by_2, \dots, ax_n + by_n) : a, b \in F, (a, b) \neq (0, 0), (x_1, x_2, \dots, x_n), (y_1, y_2, \dots, y_n) \in F^n, (x_1, x_2, \dots, x_n) \approx (y_1, y_2, \dots, y_n)\}$ . Then the incidence structure  $(P, L, \varepsilon)$  is a projective space. This space is included in that of Example 1 if we let  $S = \{1, 2, \dots, n\}$ .

Example 3. Let

$$P \equiv \{(1,0,0), (0,1,0), (0,0,1), (1,1,1), (1,1,0), (0,1,1), (1,0,1)\}$$

$$L \equiv \{\langle\langle 1,0,0 \rangle\rangle, \langle\langle 0,1,0 \rangle\rangle, \langle\langle 0,0,1 \rangle\rangle, \langle\langle 1,1,1 \rangle\rangle, \langle\langle 1,1,0 \rangle\rangle, \langle\langle 0,1,1 \rangle\rangle, \langle\langle 1,0,1 \rangle\rangle\}$$

and let the incidence relation  $*$  be as displayed in the following figure:



Then the incidence structure  $(P, L, *)$  is a projective space. This space for which  $F = GF(2)$  is clearly a special case of the projective space of Example 2.

### Finite Projective Geometry

Let  $P(X_{v+1}(q))$  denote the set of all subspaces of a  $(v+1)$ -dimensional vector space  $X_{v+1}(q)$  over  $GF(q)$ , and define  $P \equiv \{S \in P(X_{v+1}(q)) : \dim S = 1\}$  and  $L \equiv \{T \in P(X_{v+1}(q)) : \dim T = 2\}$ . Now if  $\subseteq$  denotes the set inclusion relation, then it can be shown that the incidence structure  $(P, L, \subseteq)$  forms a projective space.

The collection of all subspaces of  $X_{v+1}(q)$  together with the natural containment relation is called the  $v$ -dimensional finite projective geometry which will be denoted by  $P(X_{v+1}(q))$ , or simply by  $P_v(X)$ . If  $S$  is a subspace of  $X_{v+1}(q)$ , then the projective dimension of  $S$  is defined to be  $\dim S - 1$ , and will be denoted by  $p \dim S$ . The dimension of the geometry  $P_v(X)$ , written  $p \dim P_v(X)$ , will be  $p \dim X_{v+1}(q)$ . The elements of projective dimensions 0, 1, and 2 are called *projective points*, *projective lines*, and *projective planes*, respectively. An element of  $P_v(X)$  of projective dimension  $p \dim X_{v+1}(q) - 1$  is called a *hyperplane* in  $P_v(X)$ .

If we were to think of the points of  $P_v(X)$  as its most fundamental objects, then their intersection (or indeed, the intersection of any two of them) would be the "natural" empty set; so the zero subspace of  $X_{v+1}(q)$ , whose algebraic dimension is zero and whose projective dimension is  $-1$ , is the empty space in  $P_v(X)$ .

If  $S$  is a subspace of  $X_{v+1}(q)$ , then  $P(S)$ , the projective geometry generated by the collection of subspaces of  $S$ , is contained in a natural way in  $P_v(X)$ , and all the objects of  $P(S)$  are also objects of  $P_v(X)$ , with the same projective dimension, and  $P(S)$  is called a

*subgeometry* of  $P_V(X)$ . In particular, if  $S$  is a projective point, then  $P(S)$  consists only of the two elements  $\{0\}$  and  $S$ . Moreover, the point  $S$  is determined by any nonzero vector  $s \in S$ , that is,  $\langle s \rangle = S$ , and such a vector is called a *homogeneous* vector for  $S$ .

Definition 3. Let  $P(X)$  and  $P(X')$  be projective geometries over  $GF(q)$ . A map  $\pi : P(X) \rightarrow P(X')$  is called a *projectivity* of  $P(X)$  onto  $P(X')$  if

- (a)  $S \in P(X) \implies \pi(S) \in P(X')$
- (b)  $S \subseteq T \iff \pi(S) \subseteq \pi(T) \quad \forall S, T \in P(X)$
- (c) For every element  $S' \in P(X')$  there exists a unique element  $S \in P(X)$  such that  $\pi(S) = S'$ .

In particular, if  $X = X'$  then  $\pi$  is called an *autoprojectivity*, or a *collineation* of  $P(X)$ .

From the above definition it is clear that the identity map on a projective geometry, composition of projectivities, and inverses of projectivities are likewise well determined projectivities.

Theorem 1. Let  $X_i \in P(X)$ ,  $i \in \underline{\ell}$ , and let  $\pi$  be a projectivity of  $P(X)$  onto  $P(X')$ . Then

$$\pi \left( \begin{array}{c} \ell \\ \Sigma \\ i=1 \end{array} X_i \right) = \begin{array}{c} \ell \\ \Sigma \\ i=1 \end{array} \pi(X_i)$$

and

$$\pi \left( \begin{array}{c} \ell \\ \cap \\ i=1 \end{array} X_i \right) = \begin{array}{c} \ell \\ \cap \\ i=1 \end{array} \pi(X_i)$$

That is, projectivities preserve sum and intersection.

Theorem 2. A projectivity maps points on points.

Theorem 3. Projectivities preserve dimensions.

If  $P_v(X)$  is a projective geometry of dimension  $v$  over the field  $GF(q)$ , then a *coordinatization* of  $P_v(X)$ , or a *projective coordinate system* for  $P_v(X)$ , is any projectivity  $\pi : P_v(X) \rightarrow P_v(GF(q)^{v+1})$ . Let  $X_0, X_1, \dots, X_{v+1}$  be points of  $P_v(X)$ , no  $v+1$  of which lie in a hyperplane. Then there is a unique projectivity  $\pi : P_v(X) \rightarrow P_v(GF(q)^{v+1})$  such that  $\pi(X_i) = \langle e^i \rangle$ ,  $i \in \underline{v+1}$ , and  $\pi(X_0) = \langle \sum_{i=1}^{v+1} e^i \rangle$ , where  $\{e^i\}$ ,  $i \in \underline{v+1}$ , is the standard basis of  $GF(q)^{v+1}$ . The set  $\{X_0, X_1, \dots, X_{v+1}\}$  is called a *frame of reference* for  $P_v(X)$  with *unit point*  $X_0$  and *simplex*  $\{X_1, X_2, \dots, X_{v+1}\}$ . It is important to realize that for the unique determination of a projective coordinate system for  $P_v(X)$  we must augment a simplex of reference with the unit point. This requirement is due to an ambiguity in the choice of  $X_i = \langle x^i \rangle$ ,  $i \in \underline{v+1}$ , except in the case where the ground field is  $GF(2)$ , because each  $x^i$  may be multiplied by any arbitrary nonzero scalar in  $GF(q)$ . In a sense, the unit point "stiffens up" the simplex of reference.

### Dual Projective Geometries

A *duality*  $\delta$  of a projective geometry  $P(X)$  over the field  $GF(q)$  onto the projective geometry  $P(X')$  over the field  $GF(q)$  is a monotone decreasing monomorphism, that is, it is a monomorphism such that for  $R, S \in P(X)$ ,  $R \subseteq S \iff \delta(R) \supseteq \delta(S)$ . As immediate consequences of this definition, it can be easily seen that the inverse of a duality is a duality, well-defined composition of dualities is a projectivity, dualities interchange sums and intersections, and dualities interchange

points and hyperplanes. Dualities of  $P(X)$  onto itself are called *auto-dualities*, also *correlations*, of  $P(X)$ , and a projective geometry possessing an auto-duality is called *self-dual*. Two projective geometries are said to be duals of each other if there exists a duality between them.

Instead of discussing dualities in a general setting, it will suffice for our purposes to restrict our attention to a particular type of duality, namely the annihilator mapping connecting  $P(X)$  and the dual geometry  $P(X^0)$ , where  $X_n^0(q)$  is the canonical dual vector space of  $X_n(q)$ . It is easy to see that the annihilator mapping  $\sigma : S \longmapsto S^0$ , where  $S \subseteq X_n(q)$  and  $S^0 = \{f \in X_n^0(q) : f(s) = 0 \ \forall s \in S\}$  is a monomorphism of  $P(X)$  onto  $P(X^0)$ , and has the following properties:

$$\dim S^0 = \dim X_n(q) - \dim S, \quad S^{00} = S, \quad S \subseteq T \iff S^0 \supseteq T^0 \quad \forall S, T \in P(X),$$

$$\left( \sum_{i=1}^{\ell} S_i \right)^0 = \bigcap_{i=1}^{\ell} S_i^0, \quad \left( \bigcap_{i=1}^{\ell} S_i \right)^0 = \sum_{i=1}^{\ell} S_i^0, \quad \forall S_i \in P(X), \quad i \in \underline{\ell}, \quad \{0\}^0 = X_n(q), \quad \text{and} \quad (X_n(q))^0 = \{0\}.$$

The relationships between the geometries  $P(X)$  and  $P(X^0)$  can often be most easily discussed in terms of the dual coordinates. Let  $(X_0, X_1, \dots, X_{v+1})$  be a frame of reference for  $P(X)$  and  $(x^1, x^2, \dots, x^{v+1})$  be any basis of  $X_{v+1}(q)$  determining it. Then the corresponding dual basis  $(\hat{x}^1, \hat{x}^2, \dots, \hat{x}^{v+1})$  determines, through the annihilator mapping, a unique dual frame of reference  $(\hat{X}_0, \hat{X}_1, \dots, \hat{X}_{v+1})$  for  $P(X^0)$ .

Further relationships between  $P(X)$  and  $P(X^0)$  are provided by the duality principle of projective geometry. A proposition II in a  $v$ -dimensional projective geometry over the field  $GF(q)$  is a statement involving only the elements of the geometry and the underlying incidence

relations. It is usually phrased in terms of intersections, joins, and dimensions. The dual proposition  $\Pi^0$  is defined to be the statement obtained from  $\Pi$  by changing  $\subseteq$  to  $\supseteq$  throughout and hence replacing intersection, join, and dimension  $\ell$  by join, intersection, and dimension  $v-1-\ell$ , respectively. The principle of duality essentially states that if  $\Pi$  is a proposition which is true in all  $v$ -dimensional projective geometries over a given field  $GF(q)$ , then  $\Pi^0$  is also true in all  $v$ -dimensional projective geometries over  $GF(q)$ . Clearly this principle "doubles" the theorems at our disposal without our having to do any extra work.

#### Finite Affine Geometry

The collection of all cosets of  $X_n(q)$ , denoted by  $A(X)$ , forms a mathematical structure called a *finite affine geometry* whose points are the cosets of rank 0, that is, essentially the elements of  $X_n(q)$ , and whose lines are the cosets of rank 1. This geometry, like the projective geometry  $P(X)$ , has its own unique set of incidence axioms and incidence propositions which can be developed directly. However, it turns out that there exist strong connections between  $P(X)$  and  $A(X)$  which can be exploited to study the constructive interplay between these geometries, and in most cases, either one can be used to study the other; for instance, the propositions of incidence in  $A(X)$  can be easily deduced from those in  $P(X)$ . More specifically,  $A(X)$  can be embedded in  $P(X)$ .

Theorem 3.12.1. (The Embedding Theorem). If  $H$  is any hyperplane in  $P(X)$  and  $x \in X(q)$ ,  $x \notin H$ , then the mapping  $\hat{\phi} : A(x + H) \rightarrow P(X)$  induced by  $\phi : C \mapsto \langle C \rangle$ ,  $C \in A(x + H)$ , has the following properties:

- (1)  $\hat{\phi}$  is a monomorphism;
- (2)  $\hat{\phi}(A(x + H)) = P \equiv \{S \subseteq X(q) : S \not\subseteq H\} \subseteq P(X)$ ;
- (3)  $C \subseteq C' \iff \langle C \rangle \subseteq \langle C' \rangle \forall C, C' \in A(x + H)$ ;
- (4)  $C_i \in A(x + H)$ ,  $i \in \underline{\ell}$ ;  $\bigcap_{i=1}^{\ell} C_i \neq \emptyset \implies \langle \bigcap_{i=1}^{\ell} C_i \rangle = \bigcap_{i=1}^{\ell} \langle C_i \rangle$ ;
- (5)  $\dim C = p \dim \langle C \rangle \forall C \in A(x + H)$ ;
- (6)  $C \parallel C' \iff \langle C \rangle \cap H \subseteq \langle C' \rangle \cap H$  or  $\langle C \rangle \cap H \supseteq \langle C' \rangle \cap H$ ;  $C, C' \in A(x + H)$ .

Parts (1) and (2) of this theorem provide the essential link between  $A(X)$  and  $P(X)$  by using  $\hat{\phi}$  to carry over to  $\hat{P} \subseteq P(X)$  the basic geometrical notions in  $A(x + H)$ , such as inclusion, intersection, dimension, and parallelism. This scheme is frequently employed in geometry to generate more general geometrical structures.

Parts (3) - (6) of the above theorem ensure the compatibility of the inherent structure of  $\hat{A}$  as a subset of  $P(X)$  with those carried over from  $A(x + H)$  by  $\hat{\phi}$ .

In summary, the Embedding Theorem provides a procedure by which an affine geometry is obtained from a projective geometry. This is accomplished simply by deleting a hyperplane from  $P(X)$ . The deleted hyperplane is called the *hyperplane at infinity*. That all affine



geometries obtained by this process have the same structure follows from the fact that any two hyperplanes of  $P(X)$  are isomorphic to each other.

Other close relationships between projective and affine geometries exist with respect to their isomorphism structures. In particular, it can be shown that if  $H$  and  $H'$  are hyperplanes in  $P(X)$  over the field  $F$  and  $P(X')$  over  $F'$  determining the affine geometries  $A$  and  $A'$ , respectively, then any projective isomorphism  $\pi : P(X) \rightarrow P(X')$  for which  $\pi(H) = H'$ , restricts to an affine isomorphism  $\alpha : A \rightarrow A'$ . Conversely, any affine isomorphism  $\alpha : A \rightarrow A'$  is the restriction of just one such projective isomorphism  $\pi$ . If  $F = F'$ , then  $\pi$  is a projectivity if and only if  $\alpha$  is an affinity.

## BIBLIOGRAPHY

1. Arbib, M. A., "A Common Framework for Automata Theory and Control Theory," *SIAM J. Control*, Vol. 3, pp. 206-222, 1965.
2. Arbib, M. A., "Automata Theory and Control Theory: A Rapprochement," *Automatica*, Vol. 3, pp. 161-189, 1966.
3. Arbib, M. A. and H. P. Zeiger, "On the Relevance of Abstract Algebra to Control Theory," *Automatica*, Vol. 5, pp. 589-606, 1969.
4. Bartee, T. C. and D. I. Schneider, "Computation with Finite Fields," *Information and Control*, Vol. 6, pp. 79-98, 1963.
5. Basile, G. and G. Marro, "Controlled and Conditioned Invariant Subspaces in Linear System Theory," *J. Opt. Theory and Appl.*, Vol. 3, pp. 306-315, 1969.
6. Basile, G. and G. Marro, "On the Observability of Linear, Time-Invariant Systems with Unknown Inputs," *J. Opt. Theory and Appl.*, Vol. 3, pp. 410-415, 1969.
7. Basile, G. and G. Marro, "A State Space Approach to Non-interacting Controls," *Ricerche di Automatica*, Vol. 1, pp. 68-77, 1970.
8. Basile, G. and G. Marro, "On the Perfect Output Controllability of Linear Dynamic Systems," *Ricerche di Automatica*, Vol. 2, pp. 1-10, 1971.
9. Basile, G. and G. Marro, "A New Characterization of Some Structural Properties of Linear Systems: Unknown-Input Observability, Invertibility, and Functional Controllability," *Int. J. Control*, Vol. 17, pp. 931-943, 1973.
10. Bhattacharyya, S. P., "On Calculating Maximal (A, B)-Invariant Subspaces," *IEEE Trans. Aut. Control*, Vol. 20, pp. 264-265, 1975.
11. Blake, I. F. and R. C. Mullen, *The Mathematical Theory of Coding*, Academic Press, New York, 1975.
12. Bobrow, L. S. and M. A. Arbib, *Discrete Mathematics: Applied Algebra for Computer and Information Science*, Saunders, Philadelphia, 1974.

13. Bollman, D. A., "Some Periodicity Properties of Transformations on Vector Spaces Over Residue Class Rings," *SIAM J. Appl. Math.*, Vol. 13, pp. 902-912, 1965.
14. Booth, T. L., *Sequential Machines and Automata Theory*, Wiley, New York, 1967.
15. Brunovsky, P., "A Classification of Linear Controllable Systems," *Kybernetika*, Vol. 3, pp. 173-187, 1970.
16. Brzozowski, J. A. and W. A. Davis, "On the Linearity of Autonomous Sequential Machines," *IEEE Trans. Elec. Computers*, Vol. EC-13, pp. 673-678, 1964.
17. Bucy, R. S., "Canonical Forms for Multivariable Systems," *IEEE Trans. Aut. Control*, Vol. AC-13, pp. 567-569, 1968.
18. Cadzow, J. A., "Nilpotency Property of the Discrete Regulator," *IEEE Trans. Aut. Control*, Vol. AC-13, pp. 734-735, 1968.
19. Candy, J. V., M. E. Warren, and T. E. Bullock, "An Algorithm for the Determination of System Invariants and Canonical Forms," *Proc. Seventh Annual Southeastern Symp. on System Theory*, pp. 218-224, 1975.
20. Chen, C. T., C. A. Desoer, and A. Niederlinski, "Simplified Conditions for Controllability and Observability of Linear, Time-Invariant Systems," *IEEE Trans. Aut. Control*, Vol. AC-11, pp. 613-614, 1966.
21. Chen, C. T., *Introduction to Linear System Theory*, Holt, Rinehart and Winston, New York, 1970.
22. Chen, C. T., "Minimization of Linear Sequential Machines," *IEEE Trans. Computers*, Vol. C-23, pp. 93-95, 1974.
23. Cohn, M., "Controllability in Linear Sequential Networks," *IRE Trans. Circuit Theory*, Vol. CT-9, pp. 74-78, 1962.
24. Cohn, M., "A Theorem on Linear Automata," *IEEE Trans. Elec. Computers*, Vol. EC-13, pp. 52-53, 1964.
25. Cohn, M., "Properties of Linear Machines," *J. Assoc. Comp. Mach.*, Vol. 11, pp. 296-301, 1964.
26. Cohn, M. and S. Even, "Identification and Minimization of Linear Machines," *IEEE Trans. Elec. Computers*, Vol. EC-14, pp. 367-376, 1965.

27. Crossley, T. R. and B. Porter, "Simple Proof of the Simon-Mitter Controllability Theorem," *Electronics Letters*, Vol. 9, pp. 51-52, 1973.
28. Crowell, R. H., "Graphs of Linear Transformations Over Finite Fields," *SIAM J. Appl. Math.*, Vol. 10, pp. 103-112, 1962.
29. Davis, W. A. and J. A. Brzozowski, "On the Linearity of Sequential Machines," *IEEE Trans. Elec. Computers*, Vol. EC-15, pp. 21-29, 1966.
30. Dembowski, P., *Finite Geometries*, Springer-Verlag, New York, 1968.
31. Deuel, D. R., "Time-Varying Linear Sequential Machines. I," *J. Computer and System Sci.*, Vol. 3, pp. 93-118, 1969.
32. Elspas, B., "The Theory of Autonomous Sequential Networks," *IRE Trans. Circuit Theory*, Vol. CT-6, pp. 45-60, 1959.
33. Friedland, B., "Linear Modular Sequential Circuits," *IRE Trans. Circuit Theory*, Vol. CT-6, pp. 61-68, 1959.
34. Friedland, B. and T. E. Stern, "Linear Modular Sequential Circuits and Their Application to Multiple Level Coding," *IRE Natl. Conv. Record*, Vol. 7, pt. 2, pp. 40-48, 1959.
35. Fujimoto, S., "Various Properties of the Group of Nonsingular Linear Sequential Machines," *Systems, Computers, Controls*, Vol. 5, pp. 68-76, 1974.
36. Gallaire, H. and M. A. Harrison, "Decomposition of Linear Sequential Machines," *Math. Systems Theory*, Vol. 3, pp. 246-287, 1969.
37. Gilbert, E. G., "Controllability and Observability in Multivariable Control Systems," *SIAM J. Control*, Vol. 1, pp. 128-151, 1963.
38. Gill, A., *Introduction to the Theory of Finite-State Machines*, McGraw-Hill, New York, 1962.
39. Gill, A., "Analysis of Linear Sequential Circuits by Confluence Sets," *IEEE Trans. Elec. Computers*, Vol. EC-13, pp. 226-231, 1964.
40. Gill, A., "Analysis and Synthesis of Stable Linear Sequential Circuits," *J. Assoc. Comp. Mach.*, Vol. 12, pp. 141-149, 1965.
41. Gill, A., "The Minimization of Linear Sequential Circuits," *IEEE Trans Circuit Theory*, Vol. CT-12, pp. 292-294, 1965.

42. Gill, A., "The Reduced Form of a Linear Automaton," in E. R. Caianiello (ed.), *Automata Theory*, pp. 164-175, Academic Press, New York, 1966.
43. Gill, A., "State Graphs of Autonomous Linear Automata," in E. R. Caianiello (ed.), *Automata Theory*, pp. 176-180, Academic Press, New York, 1966.
44. Gill, A., "On Series-to-Parallel Transformation of Linear Sequential Circuits," *IEEE Trans. Elec. Computers*, Vol. EC-15, pp. 107-108, 1966.
45. Gill, A., "Graphs of Affine Transformations with Applications to Sequential Circuits," *IEEE Conf. Record, 7th Ann. Symp. Switching and Automata Theory*, pp. 127-135, 1966.
46. Gill, A., *Linear Sequential Circuits: Analysis, Synthesis, and Applications*, McGraw-Hill, New York, 1967.
47. Gill, A., "Linear Modular Systems," in L. A. Zadeh and E. Polak (eds.), *System Theory*, pp. 179-231, McGraw-Hill, New York, 1969.
48. Gökner, I. C., "A Characterization of Controllability Suitable for Signal-Flow Graph Applications," *Int. J. Systems Sci.*, Vol. 7, pp. 121-129, 1976.
49. Harrison, M. A., *Lectures on Linear Sequential Machines*, Academic Press, New York, 1969.
50. Hartmanis, J., "Linear Multivalued Sequential Coding Networks," *IRE Trans. Circuit Theory*, Vol. CT-6, pp. 69-74, 1959.
51. Hartmanis, J., "Two Tests for the Linearity of Sequential Machines," *IEEE Trans. Elec. Computers*, Vol. EC-14, pp. 781-786, 1965.
52. Hirvonen, J., H. Blomberg, and R. Ylinen, "An Algebraic Approach to Canonical Forms and Invariants for Linear Time-Invariant Differential and Difference Systems," *Int. J. Systems Sci.*, Vol. 6, pp. 1119-1134, 1975.
53. Hohn, F. E., "Tryon's Delay Operator and the Design of Synchronous Digital Circuits," in H. Aiken and W. F. Main (eds.), *Switching Theory in Space Technology*, Stanford University Press, Stanford, 1963.
54. Hotz, G., "On the Mathematical Theory of Linear Sequential Networks," in H. Aiken and W. F. Main (eds.), *Switching Theory in Space Technology*, pp. 11-19, Stanford University Press, Stanford, 1963.

55. Hsiao, M. Y. and K. Y. Sih, "Series-to-Parallel Transformation of Linear-Feedback Shift-Register Circuits," *IEEE Trans. Elec. Computers*, Vol. EC-13, pp. 738-740, 1964.
56. Huffman, D. A., "A Linear Circuit Viewpoint of Error-Correcting Codes," *IRE Trans. Information Theory*, Vol. IT-2, pp. 20-28, 1956.
57. Huffman, D. A., "The Synthesis of Linear Coding Networks," in C. Cherry (ed.), *Information Theory*, pp. 77-95, Academic Press, New York, 1956.
58. Kalman, R. E., Y. C. Ho, and K. S. Narendra, "Controllability of Linear Dynamical Systems," in *Contrib. to Diff. Equations*, Vol. 1, pp. 189-213, 1962.
59. Kalman, R. E., "Canonical Structure of Linear Dynamical Systems," *Proc. Natl. Acad. Sci.*, Vol. 48, pp. 596-600, 1962.
60. Kalman, R. E., "Mathematical Description of Linear Dynamical Systems," *SIAM J. Control*, Vol. 1, pp. 152-192, 1963.
61. Kalman, R. E., P. L. Falb, and M. A. Arbib, *Topics in Mathematical System Theory*, McGraw-Hill, New York, 1969.
62. Kalman, R. E. "Kronecker Invariants and Feedback," *Proc. Conf. on Ord. Diff. Eqns.*, NRL Math. Res. Center, 1961; in L. Weiss (ed.), *Ordinary Differential Equations*, pp. 459-471, Academic Press, New York, 1972.
63. Kamal, A. K., H. Singh, S. Puri, and N. K. Nanda, "On the Realization of Linear Sequential Machines From the Given Delay Transfer-Function Matrix," *Int. J. Systems Sci.*, Vol. 6, pp. 787-791, 1975.
64. Kamal, A. K., H. Singh, S. Puri, and N. K. Nanda, "On the Evaluation of Transition Matrices in Finite Fields," *Int. J. Systems Sci.*, Vol. 6, pp. 561-564, 1975.
65. Klamka, J., "Uncontrollability and Unobservability of Multivariable Systems," *IEEE Trans. Aut. Control*, Vol. AC-17, pp. 725-726, 1972.
66. Kreindler, E. and P. E. Sarachik, "On the Concept of Controllability and Observability in Linear Systems," *IEEE Trans. Aut. Control*, Vol. AC-9, pp. 129-136, 1964.
67. Lavallee, P., "Nonstable Cycle and Level Sets for Linear Sequential Machines," *IEEE Trans. Elec. Computers*, Vol. EC-14, pp. 957-959, 1965.

68. Lavallee, P., "Some New Group-Theoretic Properties of Singular Linear Sequential Machines," *IEEE Trans. Elec. Computers*, Vol. EC-14, pp. 959-961, 1965.
69. Luenberger, D. G., "Observing the State of a Linear System," *IEEE Trans Military Electronics*, Vol. MIL-8, pp. 74-80, 1964.
70. Luenberger, D. G., "Observers for Multivariable Systems," *IEEE Trans. Aut. Control*, Vol. AC-11, pp. 190-197, 1966.
71. Luenberger, D. G., "Canonical Forms for Linear Multivariable Systems," *IEEE Trans. Aut. Control*, Vol. AC-12, pp. 290-293, 1967.
72. MacLane, S. and G. Birkhoff, *Algebra*, MacMillan, New York, 1967.
73. Magidin, M. and A. Gill, "Singular Shift Registers Over Residue Class Rings," *Math. Syst. Theory*, Vol. 9, pp. 345-358, 1976.
74. Mandelbaum, D., "A Comparison of Linear Sequential Circuits and Arithmetic Sequences," *IEEE Trans. Elec. Computers*, Vol. EC-16, pp. 151-157, 1967.
75. Massey, J. L. and M. K. Sain, "Codes, Automata, and Continuous Systems: Explicit Interconnections," *IEEE Trans. Aut. Control*, Vol. AC-12, pp. 644-650, 1967.
76. Massey, J. L. and M. K. Sain, "Inverses of Linear Sequential Circuits," *IEEE Trans. Computers*, Vol. C-17, pp. 330-337, 1968.
77. Massey, J. L., "Application of Automata Theory in Coding," in J. T. Tou (ed.), *Applied Automata Theory*, pp. 125-146, Academic Press, New York, 1968.
78. Mealy, G. H., "A Method for Synthesizing Sequential Circuits," *Bell Syst. Tech. J.*, Vol. 34, pp. 1045-1079, 1955.
79. Mitter, S. K. and R. Foulkes, "Controllability and Pole Assignment for Discrete-Time Linear Systems Defined Over Arbitrary Fields," *SIAM J. Control*, Vol. 9, pp. 1-7, 1971.
80. Molinary, B. P., "Extended Controllability and Observability for Linear Systems," *IEEE Trans. Aut. Control*, Vol. AC-21, pp. 136-137, 1976.
81. Molinary, B. P., "A Strong Controllability and Observability in Linear Multivariable Control," *IEEE Trans. Aut. Control*, Vol. AC-21, pp. 761-764, 1976.

82. Moore, E. F., "Gedanken-Experiments on Sequential Machines," in C. E. Shannon and J. McCarthy (eds.), *Automata Studies*, pp. 129-153, Princeton Univ. Press, Princeton, 1956.
83. Morse, A. S., "Output Controllability and System Synthesis," *SIAM J. Control*, Vol. 9, pp. 143-148, 1971.
84. Morse, A. S. and W. M. Wonham, "Status of Noninteracting Control," *IEEE Trans. Aut. Control*, Vol. AC-16, pp. 568-581, 1971.
85. Morse, A. S., "Structural Invariants of Linear Multivariable Systems," *SIAM J. Control*, Vol. 11, pp. 446-465, 1973.
86. Nerode, A., "Linear Automaton Transformations," *Proc. Amer. Math. Soc.*, Vol. 9, pp. 541-544, 1958.
87. Olson, R. R., "Note on Feedforward Inverses for Linear Sequential Circuits," *IEEE Trans Computers*, Vol. C-19, pp. 1216-1221, 1970.
88. Padulo, L. and M. A. Arbib, *System Theory: A Unified State Space Approach to Continuous and Discrete Systems*, Saunder, Philadelphia, 1974.
89. Peterson, W. W. and E. J. Weldon, *Error Correcting Codes*, Second Edition, MIT Press, Cambridge, 1972.
90. Popov, V. M., "Invariant Description of Linear, Time-Invariant systems," *SIAM J. Control*, Vol. 10, pp. 252-264, 1972.
91. Popov, V. M., *Hyperstability of Control Systems*, Springer-Verlag, New York, 1973.
92. Porter, B. and R. Crossley, *Modal Control: Theory and Applications*, Taylor & Francis, London, 1972.
93. Preparata, F. P., "On the Realizability of Special Classes of Autonomous Sequential Circuits," *IEEE Trans. Elec. Computers*, Vol. EC-14, pp. 791-797, 1965.
94. Pugsley, J. H., "Sequential Functions and Linear Sequential Machines," *IEEE Trans. Elec. Computers*, Vol. EC-14, pp. 376-382, 1965.
95. Richalet, J., "Operational Calculus for Finite Rings," *IEEE Trans. Circuit Theory*, Vol. CT-12, pp. 558-570, 1965.
96. Rissanen, J., "Basis of Invariants and Canonical Forms for Linear Dynamic Systems," *Automatica*, Vol. 10, pp. 175-182, 1974.



97. Rosenbrock, H. H., *State Space and Multivariable Theory*, Wiley, New York, 1970.
98. Scherba, M. B. and R. B. Roesser, "Computation of the Transition Matrix of a Linear Sequential Circuit," *IEEE Trans. Computers*, Vol. C-22, pp. 427-428, 1973.
99. Simon, J. D. and S. K. Mitter, "A Theory of Modal Control," *Information and Control*, Vol. 13, pp. 316-353, 1968.
100. Srinivasan, C. V., "State Diagram of Linear Sequential Machines," *J. Franklin Inst.*, Vol. 273, pp. 383-418, 1962.
101. Stern, T. E. and B. Friedland, "The Linear Modular Sequential Circuit Generalized," *IRE Trans. Circuit Theory*, Vol. CT-8, pp. 79-80, 1961.
102. Tang, D. T., "Transfer Function Synthesis of Linear Shift Register Circuits," *Proc. Third Ann. Allerton Conf. Circuits and System Theory*, pp. 63-72, 1965.
103. Toda, I., "The Tree Set of a Linear Machine," *IEEE Trans. Elec. Computers*, Vol. EC-14, pp. 954-957, 1965.
104. Tsypkin, Ya. Z. and R. G. Faradzev, "Laplace-Galois Transform in the Theory of Sequential Machines," *DAN SSSR*, Vol. 166, 1966.
105. Tzafestas, S. G., "Concerning Controllability and Observability of Linear Sequential Machines," *Int. J. Systems Sci.*, Vol. 3, pp. 197-208, 1972.
106. Tzafestas, S. G., "State-Observer Design for Linear Sequential Machines," *Int. J. Systems Sci.*, Vol. 4, pp. 13-25, 1973.
107. Tzafestas, S. G., "Multivariable Control Theory in Linear Sequential Machines," *Int. J. Systems Sci.*, Vol. 4, pp. 363-396, 1973.
108. Warren, M. E. and A. E. Eckberg, Jr., "On the Dimensions of Controllability Subspaces: A Characterization Via Polynomial Matrices and Kronecker Invariants," *SIAM J. Control*, Vol. 13, pp. 434-445, 1975.
109. Willems, J. C. and S. J. Mitter, "Controllability, Observability, Pole Allocation, and State Reconstruction," *IEEE Trans. Aut. Control*, Vol. AC-16, pp. 582-595, 1971.
110. Wolovich, W. A., *Linear Multivariable Systems*, Springer-Verlag, New York, 1974.

111. Wonham, W. M. and A. S. Morse, "Decoupling and Pole Assignment in Linear Multivariable Systems: A Geometric Approach," *SIAM J. Control*, Vol. 8, pp. 1-18, 1970.
112. Wonham, W. M., "Dynamic Observers: Geometric Theory," *IEEE Trans. Aut. Control*, Vol. AC-15, pp. 258-259, 1970.
113. Wonham, W. M. and A. S. Morse, "Feedback Invariants of Linear Multivariable Systems," *Automatica*, Vol. 8, pp. 93-100, 1972.
114. Wonham, W. M., *Linear Multivariable Control: A Geometric Approach*, Springer-Verlag, New York, 1974.
115. Yau, S. S. and K. C. Wang, "Linearity of Sequential Machines," *IEEE Trans. Elec. Computers*, Vol. EC-15, pp. 337-354, 1966.
116. Yuan, F. M., "Minimal Memory Inverses of Linear Sequential Circuits," *IEEE Trans. Computers*, Vol. C-23, pp. 1155-1163, 1974.
117. Yuan, F. M., "Minimal Dimension Inverses of Linear Sequential Circuits," *IEEE Trans. Aut. Control*, Vol. AC-20, pp. 42-52, 1975.
118. Zadeh, L. A. and C. A. Desser, *Linear System Theory: The State Space Approach*, McGraw-Hill, New York, 1963.
119. Zalmai, G. J., *Lattice-Theoretic Characterization of Some Structural Properties of Linear Multivariable Dynamical Control Systems*, Master's Thesis, Georgia Institute of Technology, 1974.
120. Zierler, N., "Linear Recurring Sequences," *SIAM J. Appl. Math.*, Vol. 7, pp. 31-48, 1959.
121. Zunde, P., *Controllability, Invariance, and Interaction in Linear Dynamical Systems*, Ph.D. Dissertation, Georgia Institute of Technology, 1968.

## VITA

G. J. Zalmai was born in Kabul, Afghanistan, on March 28, 1946. He entered the Georgia Institute of Technology in September of 1966, and completed all the requirements for the degree Bachelor of Industrial Engineering (with honor) in December of 1969. From January of 1970 to January of 1971, he served as an instructor in the School of Engineering at Kabul University, Kabul, Afghanistan. In the spring of 1971, he resumed his graduate studies at Georgia Tech in the School of Industrial and Systems Engineering. Simultaneously pursuing two programs of study in the areas of Operations Research and Control Systems, he received the M.S.O.R. degree in March and the M.S. (E.E.) degree in August of 1974.

In the course of his graduate studies, he held the positions of Research Assistant, Teaching Assistant, and Instructor in the School of Industrial and Systems Engineering.

Since September of 1976, G. J. Zalmai has been pursuing the Ph.D. program in the Department of Pure and Applied Mathematics at Washington State University, Pullman, Washington.