UNIVERSITE D'OTTAWA
Ecole des Etudes Supérieures

UNIVERSITY OF OTTAWA
School of Graduate Studies

Title of thesis: SYNTHESIS OF LINEAR SEQUENTIAL MACHINES USING CYCLIC SUBSPACES

Name of candidate: FARLEY, J.W. Jacques

Degree: M.A. Sc. (Electrical Engineering)

Date: January 25th, 1971

This thesis has been prepared under my supervision

M. Krieger
(Thesis Supervisor)

January 25th, 1971
(Date)

We, the undersigned, certify that we have approved this thesis, and that the candidate has defended it successfully.

1. C.L. Sheng .... Date: January 25th, 1971
   (Signature of examiners)
2. E. Bainbridge .... Date: January 25th, 1971
3. S.R. Das .... Date: January 25th, 1971
4. .... Date: ....
5. .... Date: ....
6. .... Date: ....
7. .... Date: ....

....
(Dean of Graduate Studies)

SYNTHESIS OF LINEAR SEQUENTIAL MACHINES USING CYCLIC SUBSPACES

by

J.W. Jacques Farley

Submitted to the Department of Electrical Engineering
in partial fulfillment of the requirements
for the degree of Master of Science

Department of Electrical Engineering
Faculty of Pure and Applied Science
The University of Ottawa
Ottawa, Canada

1970

ACKNOWLEDGEMENT

I wish to acknowledge the insight and inspiration which my God and Saviour has given me in the natural sciences, without which this thesis would not have been possible.

The author sincerely thanks Professor M. Krieger for the guidance and many suggestions during the numerous hours spent discussing and developing this thesis.

The author wishes to thank Miss Phyllis C. Harris for printing and proofreading the thesis material and for the great help and encouragement she has been in the development of this thesis.

The author also acknowledges the financial support of the National Research Council.

ABSTRACT

The material presented in this thesis can be divided into two major topics :

Controllability and observability of linear systems, where these principles are presented in terms of cyclic subspaces. This is a new approach that gives further insight in the physical aspects of controllability and observability and it is also used for the construction of controllable and observable linear sequential machines.

Shift register realization of linear sequential machines, where a general realization algorithm is presented. The algorithm was implemented for parallel shift register realization.

CONTENTS

CHAPTER I

INTRODUCTION

In recent years with the advent of microcircuits, a number of studies were initiated to consider modular synthesis of sequential circuits. Such realizations are advantageous from the point of view of cost, speed, weight, compactness and machine diagnosis. In the field of linear sequential machine the emphasis was on the shift register realization.

There exists a number of algorithms on shift register realizations such as Davis[5], Nichols[17], Johnson and O'Keefe[12]. These algorithms are based on the partitioning theory of sequential machines and yield shift register realizations whenever possible. The major difficulty in implementing these algorithms is that they do not consider minimality and require as a first step the computation of partition pairs which in itself is a major computational problem in large machines.

Another algorithm based on the input-output relation of a sequential machine was developed by Gill[10]. Here also a computational problem arises as all the input-output pairs must be generated and very large resulting matrices must be manipulated.

The purpose of this thesis was to develop an algorithm which can be implemented on a digital computer. The actual investigation was based on the fact that shift register realizations correspond to a canonical structure in the state space. Namely the companion or hyper companion matrix form. Thus the shift register realization of a linear sequential machine corresponds to the derivation of a suitable similarity transformation.

In line with the above, the thesis is divided in the following parts :

i - Concepts of linear vector spaces: In this part the existing algebraic concepts of invariant subspaces and canonical representation of linear transformations were introduced.

ii - Controllability and observability of linear systems in terms of cyclic subspaces: In this part we start with existing definitions of controllability and observability and develop a new approach in terms of cyclic subspaces. The advantages of this approach are twofold. On one hand we obtain a better insight in the physical aspects of controllability and observability. On the other hand the analysis aspect is reduced as only $n \times n$, $n \times r$ and $n \times p$ matrices, instead of $n \times n \cdot r$ and $n \times n \cdot p$ matrices, are to be considered, where "n" is the order of the system, "r" is the number of inputs and "p" is the number of outputs.

iii - Shift register synthesis of linear sequential systems: In this part some of the results of the previous parts and the numerical methods developed in the appendices are used to present an algorithm for shift register realization. The resulting algorithm is quite general as it applies to different types of shift register realization whenever such a realization is possible. Furthermore it is also minimal in terms of the number of memory elements (flip-flops or delays) used. The actual implementation of the algorithm was presented for parallel shift register realization where each shift register corresponds to an elementary divisor of the state transition matrix.

iv - Synthesis of linear time-invariant systems: In this part the construction of a controllable and/or observable system from an uncontrollable and/or unobservable system, by adding carefully chosen inputs and/or outputs, was investigated. This technique was developed for the identification, diagnosis and initialization of linear sequential machines.

Finally in Chapter VI of the thesis the conclusions and topics for further research are presented.

CHAPTER II

CONCEPTS OF LINEAR VECTOR SPACES

The purpose of this chapter is to introduce some of the basic concepts of linear vector spaces required in the study of linear systems. First some basic definitions are introduced then the characteristics of linear transformations are considered in more detail. Finally the concepts of cyclic subspaces are used to present the canonical decomposition theorems. It should be noted that in the following sections only the proofs, yielding more insight into the concepts under consideration, shall be considered.

II-1: BASIC DEFINITIONS.

Definition 2.1: A linear vector space $\mathcal{V}$ is a mathematical system consisting of an additive abelian group $(V, +)$ defined over a field consisting $(F, +, \cdot)$ where V is called the set of vectors and F is called the set of scalars such that:

a) if $\underline{v} \in V$ and $a \in F$ then $a \cdot \underline{v} = \underline{v} \cdot a \in V$

b) if $\underline{w}, v \in V$ and $a, b, e \in F$ then

-i- $a \cdot (\underline{v} + \underline{w}) = a \cdot \underline{v} + a \cdot \underline{w}$

-ii- $(a + b) \cdot \underline{v} = a \cdot \underline{v} + b \cdot \underline{v}$

-iii- $(a \cdot b) \cdot \underline{v} = a \cdot (b \cdot \underline{v})$

-iv- $e \cdot \underline{v} = \underline{v}$

Definition 2.2: A set of vectors $\{\underline{v}_1, \underline{v}_2, \ldots, \underline{v}_n\}$ is linearly dependent if and only if there exists a set of scalars $\{a_1, a_2, \ldots, a_n\}$ NOT ALL ZERO such that $\sum_{i=1}^{n} a_i \cdot \underline{v}_i = \underline{o} \in V$ otherwise the set of vectors is linearly independent.

Definition 2.3: A set $S = \{\underline{s}_1, \underline{s}_2, \ldots, \underline{s}_r\} \subseteq V$ is a basis of $\mathcal{V}$ if and only if:

a) S spans $\mathcal{V}$

b) ALL the vectors in S are linearly independent.

The number "r" of vectors in S is called the dimension of $\mathcal{V}$.

Definition 2.4: Let S be a basis for $\mathcal{V}$ of dimension n. For any $\underline{v} \in V$ there exists a unique set of scalars $\{a_1, a_2, \ldots, a_n\}$, called the coordinates of $\underline{v}$ such that

$$\underline{v} = a_1 \cdot \underline{s}_1 + a_2 \cdot \underline{s}_2 + \ldots + a_n \cdot \underline{s}_n$$

The vector $\underline{v}$ can be represented as an n-tuple $\underline{v} \equiv (a_1, a_2, \ldots, a_n)_S$ or $\underline{v} \equiv (a_1, a_2, \ldots, a_n)$ where the set S is understood to be the basis. In an n-dimensional vector space $\mathcal{V}$ any set of n linearly independent vectors forms a valid basis. The criterion for selecting the "best" basis depends largely on the problem under consideration.

II-2: COORDINATE TRANSFORMATION

Definition 2.5: A linear transform or vector space homomorphism is a mapping β between the vector spaces $\mathcal{V}$ and $\mathcal{W}$, denoted $\beta : \mathcal{V} \longrightarrow \mathcal{W}$, such that:

a) if $\underline{v}_1, \underline{v}_2 \in V$ then $\beta(\underline{v}_1 + \underline{v}_2) = \beta(\underline{v}_1) + \beta(\underline{v}_2) \in W$

b) if $\underline{v} \in V$, $a \in F$ then $\beta(a \cdot \underline{v}) = a \cdot \beta(\underline{v}) \in W$

Let $\mathcal{V}$ be an n-dimensional vector space with the basis $\{\underline{v}_1, \underline{v}_2, \ldots, \underline{v}_n\}$ and $\mathcal{W}$ be an m-dimensional vector space with the basis $\{\underline{w}_1, \underline{w}_2, \ldots, \underline{w}_m\}$. For every $\underline{x} = \sum_{i=1}^{n} c_i \cdot \underline{v}_i \in \mathcal{V}$ there exists a $\underline{y} = \beta(\underline{x}) = \sum_{j=1}^{m} b_j \cdot \underline{w}_j \in \mathcal{W}$. Since every $\underline{v}_i$ in the basis of $\mathcal{V}$ has a corresponding

$$\beta(\underline{v}_i) = \underline{z}_i = \sum_{j=1}^{m} a_{ji} \cdot \underline{w}_j \quad \text{then} \quad \beta(\underline{x}) = \beta\left(\sum_{i=1}^{n} c_i \cdot \underline{v}_i\right)$$

$$= \sum_{i=1}^{n} c_i \sum_{j=1}^{m} a_{ji} \cdot \underline{w}_j = \sum_{j=1}^{m} \sum_{i=1}^{n} c_i \cdot a_{ji} = \underline{w}_j$$

$$= \sum_{j=1}^{m} b_j \cdot \underline{w}_j = \underline{y}$$

Hence $b_j = \sum_{i=1}^{n} a_{ji} \cdot c_i$ for $j = 1, 2, \ldots, m$. In matrix notation this becomes:

$$\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix} = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ a_{21} & \ddots & \vdots \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} \quad \text{or} \quad \underline{B} = \underline{A}\ \underline{C}$$

where $\underline{B}$ and $\underline{C}$ are the vertical coordinate representation of the vectors $\underline{y}$ and $\underline{x}$ respectively. Thus the matrix $\underline{A}$ transforms the coordinates of $\underline{x} \in \mathcal{V}$ into the coordinates of $\underline{y} \in \mathcal{W}$. The matrix $\underline{A}$ is called the matrix representation of β with respect to the basis in $\mathcal{V}$ and $\mathcal{W}$.

Of particular interest in this report is the linear transformation $\beta : \mathcal{V} \rightarrow \mathcal{V}$ which is a coordinate transformation. In this case, the matrix representing the mapping is square.

Definition 2.6: The characteristic polynomial $p(x)$ of a linear transform $\underline{A}_{n\times n}$ is defined to be

$$p(x) \doteq \text{determinant}\ (x\,\underline{I} - \underline{A}) = \left| x\,\underline{I} - \underline{A} \right|$$

By the Caley-Hamilton theorem $p(\underline{A}) = \underline{O}$

Definition 2.7: The minimum polynomial of a linear transform $\underline{A}_{nxn}$ is a monic polynomial $m(x)$ such that $m(\underline{A}) = \underline{0}$ and if there exists a polynomial $q(x)$ such that $q(\underline{A}) = \underline{0}$ then $m(x)$ is a factor of $q(x)$ denoted $m(x) \mid q(x)$.

Definition 2.8: The minimum polynomial of a vector $\underline{v} \in V$ is a monic polynomial $n(x)$ such that $n(\underline{A})\,\underline{v} = \underline{o}$ and if there exists a polynomial $p(x)$ such that $p(\underline{A})\,\underline{v} = \underline{o}$ then $n(x) \mid p(x)$.

In view of the above definitions, $m(x) \mid p(x)$ and $m(x)$ is the "LOWEST COMMON MULTIPLE" of the minimum polynomials $n_1(x)$, $n_2(x)$, ..., $n_r(x)$ of the vectors in $\mathcal{V}$.

Definition 2.9: A square matrix is non-derogatory if the degree of $p(x)$ is equal to the degree of $m(x)$. Otherwise, the matrix is derogatory.

Lemma 2.1: Let $S = \{\underline{s}_1, \underline{s}_2, \ldots, \underline{s}_n\}$ be a basis of the vector space $\mathcal{V}$ and $\beta : \mathcal{V} \to \mathcal{V}$ whose matrix representation is $\underline{A}$. If a "new" basis $N = \{\underline{n}_1, \underline{n}_2, \ldots, \underline{n}_n\}$ is defined on $\mathcal{V}$ then the matrix representation $\underline{A}^*$ of β with respect to N is a similarity transformation given by $\underline{A}^* = \underline{M}\ \underline{A}\ \underline{M}^{-1}$

Proof:

Since $\underline{s}_i \in \mathcal{V}$ then it is a linear combination of the elements of N such that

$$\underline{s}_i = \sum_{j=1}^{n} m_{ij} \cdot \underline{n}_j .$$ This relationship may be expressed in matrix form as follows:

$$[\underline{s}_1, \underline{s}_2, \ldots, \underline{s}_n] = [\underline{n}_1, \underline{n}_2, \ldots, \underline{n}_n] \begin{bmatrix} m_{11} & \cdots & m_{1n} \\ \vdots & \ddots & \vdots \\ m_{n1} & \cdots & m_{nn} \end{bmatrix}$$

$$= [\underline{n}_1, \underline{n}_2, \ldots, \underline{n}_n] \quad \underline{M}$$

Let v and v^* be the coordinate representation of $\underline{v}$ with respect to the bases S and N respectively. Hence $v^* = \underline{M}\, v$ and $v = \underline{M}^{-1} v^*$. If $\underline{v}_2 = \beta(\underline{v}_1)$ then $\underline{v}_2 = \underline{A}\, \underline{v}_1$.

$$\therefore\ \underline{M}^{-1} v_2^* = \underline{A}\ \underline{M}^{-1} v_1^*$$

$$\therefore\ v_2^* = \underline{M}\,\underline{A}\,\underline{M}^{-1} v_1^* = \underline{A}^* v_1^*$$

Thus the matrix representation of β with respect to N is $\underline{A}^* = \underline{M}\,\underline{A}\,\underline{M}^{-1}$.

<u>Definition 2.10</u>: Two n-square matrices $\underline{A}_1$ and $\underline{A}_2$ are similar over F if and only if their characteristic matrices $(x\underline{I} - \underline{A}_i)$ for $i = 1, 2$ have the <u>same</u> invariant factors <u>and</u> the <u>same</u> elementary divisors.

<u>II-3: INVARIANT SUBSPACES.</u>

<u>Definition 2.11:</u> If $W \subseteq V$ then $\mathcal{W}$ is a subspace of $\mathcal{V}$ over F if W is a subgroup of V.

<u>Definition 2.12</u>: Let $\underline{A}$ be the matrix representation of $\beta : \mathcal{V} \longrightarrow \mathcal{V}$. The subspace $\mathcal{W}$ is $\underline{A}$ - invariant if $\underline{w} \in \mathcal{W}$ implies $\beta(\underline{w}) \in \mathcal{W}$; thus $\beta(\mathcal{W}) \subseteq \mathcal{W}$.

<u>Definition 2.13</u>: A vector space $\mathcal{V}$ is the "direct-sum" of its subspaces $\mathcal{W}_1, \mathcal{W}_2, \ldots, \mathcal{W}_r$, denoted

$\mathcal{V} = \mathcal{W}_1 \otimes \ldots\ldots \otimes \mathcal{W}_r$, if every vector $\underline{v} \in \mathcal{V}$ can be <u>written uniquely</u> in the form

$$\underline{v} = \underline{w}_1 + \underline{w}_2 + \ldots\ldots + \underline{w}_r \quad \text{where } \underline{w}_i \in \mathcal{W}_i .$$

<u>Lemma 2.2</u>: Let $\underline{A}$ be the matrix representation of $\beta : \mathcal{V} \longrightarrow \mathcal{V}$ and $f(x)$ be any polynomial over F, then the Kernel of $f(\underline{A})$ is invariant under $\underline{A}$.

Proof:

$$\ker f(\underline{A}) = \{\underline{v} \in \mathcal{V} \mid f(\underline{A})\,\underline{v} = \underline{o}\}$$

Let $\underline{v} \in \ker f(\underline{A})$ such that $f(\underline{A})\,\underline{v} = \underline{o}$

But $f(\underline{A})\,\beta(\underline{v}) = f(\underline{A})\,\underline{A}\,\underline{v} = \underline{A}\,f(\underline{A})\,\underline{v} = \underline{o}$

$\therefore\ \beta(\underline{v}) \in \ker f(\underline{A})$.

Hence the Kernel of $f(\underline{A})$ is $\underline{A}$ - invariant.

Lemma 2.3 : Let $\underline{A}$ be the matrix representation of $\beta: \mathcal{V} \to \mathcal{V}$ and $\mathcal{V} = \mathcal{W}_1 \otimes \mathcal{W}_2 \otimes \ldots\ldots\ldots \otimes \mathcal{W}_r$. If $\underline{A}_i$ is the matrix representation of the restriction $\bar{\beta}_i$ of β to $\mathcal{W}_i$ then $\underline{A}$ may be represented by the block diagonal matrix. The subspaces $\mathcal{W}_i$ are invariant.

$$\underline{A}^* = \begin{bmatrix} \underline{A}_1 & & & \\ & \underline{A}_2 & & \\ & & \ddots & \\ & & & \underline{A}_r \end{bmatrix}$$ where $\underline{A}^*$ is the direct sum

of the matrices $\underline{A}_1, \underline{A}_2, \ldots, \underline{A}_r$ denoted $\underline{A}^* = \underline{A}_1 \otimes \underline{A}_2 \otimes \ldots \otimes \underline{A}_r$.

Proof :

Let $\{\underline{w}_{i1}, \underline{w}_{i2}, \ldots, \underline{w}_{ir_i}\}$ be a basis of $\mathcal{W}_i$

for $i = 1, 2, \ldots, r$ such that

$\{\underline{w}_{11}, \underline{w}_{12}, \ldots, \underline{w}_{1r_1}; \underline{w}_{21}, \underline{w}_{22}, \ldots, \underline{w}_{2r_2}; \ldots; \underline{w}_{r1}, \underline{w}_{r2}, \ldots, \underline{w}_{rr_r}\}$

is a basis of $\mathcal{V}$.

therefore,

$$\mathcal{W}_1 \left\{ \begin{array}{lllll} \bar{\beta}_1(\underline{w}_{11}) & = & \beta(\underline{w}_{11}) & = & a^1_{11}\,\underline{w}_{11} + \dots + a^1_{r_1 1}\,\underline{w}_{1r_1} \\ \vdots & & \vdots & & \vdots \qquad \ddots \qquad \vdots \\ \bar{\beta}_1(\underline{w}_{1r_1}) & = & \beta(\underline{w}_{1r_1}) & = & a^1_{1r_1}\,\underline{w}_{11} + \dots + a^1_{r_1 r_1}\,\underline{w}_{1r_1} \end{array} \right.$$

$$\mathcal{W}_2 \left\{ \begin{array}{lllll} \bar{\beta}_2(\underline{w}_{21}) & = & \beta(\underline{w}_{21}) & = & a^2_{11}\underline{w}_{21} + \dots + a^2_{r_2 1}\underline{w}_{2r_2} \\ \vdots & & \vdots & & \vdots \qquad \ddots \qquad \vdots \\ \bar{\beta}_2(\underline{w}_{2r_2}) & = & \beta(\underline{w}_{2r_2}) & = & a^2_{1r_2}\underline{w}_{21} + \dots + a^2_{r_2 r_2}\underline{w}_{2r_2} \end{array} \right.$$

$$\vdots$$

$$\mathcal{W}_r \left\{ \begin{array}{lllll} \bar{\beta}_r(\underline{w}_{r1}) & = & \beta(\underline{w}_{r1}) & = & a^r_{11}\underline{w}_{r1} + \dots + a^r_{r_r 1}\underline{w}_{rr_r} \\ \vdots & & \vdots & & \vdots \qquad \ddots \qquad \vdots \\ \bar{\beta}_r(\underline{w}_{rr_r}) & = & \beta(\underline{w}_{rr_r}) & = & a^r_{1r_r}\underline{w}_{r1} + \dots + a^r_{r_r r_r}\underline{w}_{rr_r} \end{array} \right.$$

Since the matrix representation of β is the transpose of the matrix of coefficients in the above system then the matrix representation of β with respect to the above basis is $\underline{A}^*$ where

$$\underline{A}_i = \begin{bmatrix} a^i_{11} & a^i_{12} & \dots & a^i_{1r_i} \\ a^i_{21} & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ a^i_{r_i 1} & \dots & \dots & a^i_{r_i r_i} \end{bmatrix}$$

<u>Lemma 2.4:</u> Let $\underline{A}$ be the matrix representation of $\beta : \mathcal{V} \to \mathcal{V}$ and $f(x) = g(x)\, h(x)$ such that $f(\underline{A}) = \underline{0}$ and $g(x)$ and $h(x)$ are relatively prime polynomials over F ; then $\mathcal{V} = \mathcal{S} \otimes \mathcal{W}$ where $\mathcal{S}$ = kernel $g(\underline{A})$ and $\mathcal{W}$ = kernel $h(\underline{A})$.

<u>Proof:</u>

$\mathcal{S}$ and $\mathcal{W}$ are $\underline{A}$- invariant subspaces of $\mathcal{V}$ by lemma 2.2. Since $g(x)$ and $h(x)$ are relatively prime there exist polynomials $c(x)$ and $g(x)$ such that

$$c(x)\, g(x) + d(x)\, h(x) = 1$$

$$\text{i.e.} \quad c(\underline{A})\, g(\underline{A}) + d(\underline{A})\, h(\underline{A}) = \underline{I} \qquad \text{------------------------} \qquad (1).$$

Let $\underline{v} \in \mathcal{V}$ $\therefore$ $c(\underline{A})\, g(\underline{A})\, \underline{v} + d(\underline{A})\, h(\underline{A})\, \underline{v} = \underline{v}$

Since $c(\underline{A})\, g(\underline{A})\, h(\underline{A})\, \underline{v} = h(\underline{A})\, c(\underline{A})\, g(\underline{A})\, \underline{v} = c(\underline{A})\, f(\underline{A})\, \underline{v} = \underline{0}$

$$\therefore\ c(\underline{A})\, g(\underline{A})\, \underline{v} \in \text{kernel } h(\underline{A})$$

Similarly $d(\underline{A})\, h(\underline{A})\, \underline{v} \in$ kernel $g(\underline{A})$.

To show $\mathcal{V} = \mathcal{S} \otimes \mathcal{W}$, it remains to prove the representation $\underline{v} = \underline{s} + \underline{w}$ where $\underline{s} \in \mathcal{S}$, $\underline{w} \in \mathcal{W}$ is <u>uniquely</u> determined by $\underline{v}$.

Since $\underline{v} = \underline{s} + \underline{w}$

$$\therefore\ c(\underline{A})\, g(\underline{A})\, \underline{v} = c(\underline{A})\, g(\underline{A})\, \underline{s} + c(\underline{A})\, g(\underline{A})\, \underline{w}$$

$$= c(\underline{A})\, g(\underline{A})\, \underline{w}$$

but by 1 above

$$\underline{w} = c(\underline{A})\, g(\underline{A})\, \underline{w} + d(\underline{A})\, h(\underline{A})\, \underline{w} = c(\underline{A})\, g(\underline{A})\, \underline{w}$$

$\therefore$ $\underline{w} = c(\underline{A})\, g(\underline{A})\, \underline{v}$ So $\underline{w}$ is uniquely determined by $\underline{v}$.

Similarly for $\underline{s}$.

Hence $\mathcal{V} = \mathcal{S} \otimes \mathcal{W}$.

Lemma 2.5 : In lemma 2.4 if $f(x)$ is the minimum polynomial of $\underline{A}$, then the monic polynomials $g(x)$ and $h(x)$ are the minimum polynomials of the restriction $\underline{A}_1$ of $\underline{A}$ to $\mathcal{S}$ and the restriction $\underline{A}_2$ of $\underline{A}$ to $\mathcal{W}$ respectively.

Proof:

Let $m_1(x)$ and $m_2(x)$ be the minimum polynomial of $\underline{A}_1$ and $\underline{A}_2$ respectively. Since $\underline{A}_1$ is the matrix representation of $\beta_1: \mathcal{S} \to \mathcal{S}$ then $\forall\, \underline{v} \in \mathcal{S}$, $g(\underline{A}_1)\,\underline{v} = \underline{0}$. Hence $g(\underline{A}_1) = \underline{0}$. Similarly $h(\underline{A}_2) = \underline{0}$. Since $m_1(\underline{A}_1) = \underline{0}$ and $m_2(\underline{A}_2) = \underline{0}$ then $m_1(x) \mid g(x)$ and $m_2(x) \mid h(x)$. Since $f(x)$ is least common multiple of $m_1(x)$ and $m_2(x)$ which are relatively prime, then $f(x) = m_1(x)\, m_2(x)$. But $f(x) = g(x)\, h(x)$ $\therefore$ $m_1(x)=g(x)$ and $m_2(x) = h(x)$.

Theorem 2.1: "PRIMARY DECOMPOSITION THEOREM"

If $\underline{A}$, the matrix representation of $\beta : \mathcal{V} \to \mathcal{V}$, has minimum polynomial $m(x) = f_1(x)^{n_1} f_2(x)^{n_2} \ldots\ldots f_r(x)^{n_r}$ where $f_i(x)$ are distinct, monic, irreducible polynomials, then $\mathcal{V} = \mathcal{W}_1 \otimes \ldots\ldots \otimes \mathcal{W}_r$ where $\mathcal{W}_i$ is the kernel of $f_i(x)^{n_i}$. The minimum polynomial of the restriction $\underline{A}_i$ of $\underline{A}$ to $\mathcal{W}_i$ is $f_i(x)^{n_i}$.

Proof:

The proof is by induction on "r". For $r=1$ the theorem is satisfied. Assume the theorem holds for $r-1$. By lemma 2.4 $\mathcal{V}$ can be written as the direct sum of $\mathcal{W}_1$ and $\mathcal{V}_1$ where $\mathcal{W}_1 = \text{kernel } f_1(\underline{A})^{n_1}$ and $\mathcal{V}_1 = \text{kernel } f_2(\underline{A})^{n_2} \ldots\ldots f_r(\underline{A})^{n_r}$. By lemma 2.5 $f_1(x)^{n_1}$ and $f_2(x)^{n_2} \ldots\ldots f_r(x)^{n_r}$ are the minimum polynomials of the restriction $\underline{A}_1$ and $\underline{A}_2$ of $\underline{A}$ to $\mathcal{W}_1$ and $\mathcal{V}_1$ respectively.

By assumption, $\mathcal{V}_1 = \mathcal{W}_2 \otimes \ldots\ldots \otimes \mathcal{W}_r$ where $\mathcal{W}_i = \text{kernel } f_i(\underline{A})^{n_i}$ and $f_i(x)^{n_i}$ is the minimum polynomial of the restriction of $\underline{A}_2$ to $\mathcal{W}_i$ for $i = 2, \ldots\ldots, r$. Since $f_i(x)^{n_i} \mid f_2(x)^{n_2} \ldots\ldots f_r(x)^{n_r}$ then kernel $f_i(\underline{A})^{n_i} \subseteq \mathcal{V}_1$.

Thus Kernel $f_i(\underline{A})^{n_i}$ = kernel $f_i(\underline{A}_2)^{n_i} = \mathcal{W}_i$. Thus the restriction of $\underline{A}$ to $\mathcal{W}_i$ is the same as the restriction of $\underline{A}_2$ to $\mathcal{W}_i$. Hence $f_i(x)^{n_i}$ is the minimum polynomial of the restriction of $\underline{A}$ to $\mathcal{W}_i$.

$$\therefore \quad \mathcal{V} = \mathcal{W}_1 \otimes \mathcal{W}_2 \otimes \ldots . \otimes \mathcal{W}_r .$$

II-4: CANONICAL REPRESENTATION OF LINEAR TRANSFORMATION

Definition 2.14: Let $\underline{A}$ be the matrix representation of $\beta : \mathcal{V} \longrightarrow \mathcal{V}$. The set of vectors

$Z(\underline{v}, \underline{A}) = \{ f(\underline{A})\, \underline{v} \mid \underline{v} \neq \underline{0}, \underline{v} \in V, f(x) \in F[x] \}$ is an $\underline{A}$-invariant subspace of $\mathcal{V}$ called the $\underline{A}$-cyclic subspace of $\mathcal{V}$ generated by $\underline{v}$.

Considering the generator $\underline{v}$, let k be the lowest integer such that

$$a_o \cdot \underline{v} + a_1 \cdot \underline{Av} + \ldots . . + a_{k-1} \cdot \underline{A}^{k-1} \underline{v} + \underline{A}^k \underline{v} = \underline{0}$$

Hence $m_v(x) = x^k + a_{k-1} x^{k-1} + \ldots + a_1 x + a_o$ is a unique, monic polynomial of lowest degree such that $m_v(\underline{A})\, \underline{v} = \underline{0}$. The polynomial $m_v(x)$ is sometimes referred to as the minimum polynomial of $\underline{v}$ and is called the $\underline{A}$-annibilator of $\underline{v}$ and $Z(\underline{v}, \underline{A})$. Furthermore, in terms of lemma 2.3, let $\underline{A}_v$ be the restriction of $\underline{A}$ to $Z(\underline{v}, \underline{A})$.

Lemma 2.6: Let $Z(\underline{v}, \underline{A})$ be an $\underline{A}$-cyclic subspace, $\underline{A}_v$ the restriction of $\underline{A}$ to $Z(\underline{v}, \underline{A})$ and $m_v(x) = x^k + a_{k-1} x^{k-1} + \ldots + a_1 x + a_o$ the $\underline{A}$-annibilator of $\underline{v}$. Then

a) $\{ \underline{v}, \underline{Av}, \ldots, \underline{A}^{k-1} \underline{v} \}$ is a basis of $Z(\underline{v}, \underline{A})$

b) $m_v(x)$ is the minimum polynomial of $\underline{A}_v$.

c) The matrix $\underline{A}_v$ with respect to the above basis (a) is :

$$\underline{A}_v = \begin{bmatrix} 0 \ldots\ldots & 0 & -a_o \\ 1 \quad 0 & \ldots . & -a_1 \\ \quad 1 & \ddots & \vdots \\ \quad \underline{0} & \ddots \quad 1 & -a_{k-1} \end{bmatrix}$$

Proof:

By definition of $m_v(x)$, $B = \{\underline{v}, \underline{A}\,\underline{v}, \ldots, \underline{A}^{k-1}\underline{v}\}$ is a linearly independent set. Let us show $Z(v, \underline{A}) = LS(B)$, the linear span of B. $\underline{A}^k\underline{v} \in LS(B)$. The proof is by induction. Assume $n > k$ and $\underline{A}^{n-1}\underline{v} \in LS(B)$. However $\underline{A}^n\underline{v} = \underline{A}(\underline{A}^{n-1}\underline{v})$ is a linear combination of $\{\underline{A}\underline{v}, \underline{A}^2\underline{v}, \ldots, \underline{A}^k\underline{v}\}$. Since $\underline{A}^k\underline{v} \in LS(B)$ Then $\underline{A}^n\underline{v} \in LS(B)$ for any n. Thus for any $f(x) \in F(x)$, $f(\underline{A})\underline{v} \in LS(B)$.

$\therefore$ $Z(\underline{v}, \underline{A}) = LS(B)$ and $\{\underline{v}, \underline{A}\,\underline{v}, \ldots, \underline{A}^{k-1}\underline{v}\}$ is a basis of $Z(\underline{v}, \underline{A})$.

Let $m(x) = x^s + b_{s-1}x^{s-1} + \ldots + b_1 x + b_o$ be the minimum polynomial of $\underline{A}_v$. Since $\underline{v} \in Z(\underline{v}, \underline{A})$ then

$$\underline{0} = m(\underline{A}_v)\underline{v} = m(\underline{A})\underline{v} = b_o \cdot \underline{v} + b_1 \cdot \underline{A}\,\underline{v} + \ldots + \underline{A}^s\underline{v}$$

Hence $\underline{A}^s\underline{v}$ is a linear combination of $\{\underline{v}, \underline{A}\,\underline{v}, \ldots, \underline{A}^{s-1}\underline{v}\}$.

This implies $k \leq s$. By definition $m_v(\underline{A}) = \underline{0}$. Hence $m_v(\underline{A}_v) = \underline{0}$. Then $m(x) \mid m_v(x)$ and $s \leq k$. Therefore $s = k$ and $m_v(x) = m(x)$.

$$\begin{array}{lllllll}
\beta(\underline{v}) = \underline{A}_v(\underline{v}) & = \underline{A}\underline{v} = & 0 & 1 \cdot \underline{A}\,\underline{v} & & & \\
\beta(\underline{A}\underline{v}) = \underline{A}_v(\underline{A}\underline{v}) & = \underline{A}^2\underline{v} = & 0 & 0 & 1 \cdot \underline{A}^2\underline{v} & & \\
\vdots \quad \vdots \quad \vdots & & & & \vdots & & \\
\beta(\underline{A}^{k-1}\underline{v}) = \underline{A}_v(\underline{A}^{k-1}\underline{v}) & = \underline{A}^k\underline{v} = & -a_o \cdot \underline{v} & - a_1 \cdot \underline{A}\,\underline{v} & \cdots & - a_{k-1} \cdot \underline{A}^{k-1}\underline{v} &
\end{array}$$

This relationship may be expressed in matrix form as follows:

$$[\beta(\underline{v}), \beta(\underline{A}\,\underline{v}), \ldots \beta(\underline{A}^{k-1}\underline{v})]$$

$$= [\underline{v}, \underline{A}\,\underline{v}, \ldots, \underline{A}^{k-1}\underline{v}] \begin{bmatrix} 0 & 0 & & & -a_o \\ 1 & 0 & & & -a_1 \\ \vdots & \vdots & \ddots & & \vdots \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & & 1 & -a_{k-1} \end{bmatrix}$$

$$= [\underline{v}, \underline{A}\,\underline{v}, \ldots\ldots, \underline{A}^{k-1}\underline{v}]\underline{A}_v$$

$\underline{A}_v$ is called the companion matrix of $m_v(x)$.

Definition 2.15: Let W be a normal subgroup of a group V. If $v \in V$ then $\underline{v} + W = \{ \underline{v} + \underline{w} \mid w \in W \}$ is called a coset of W in V.

Lemma 2.7: If $\mathcal{W} \subset \mathcal{V}$ then the set V/W consisting of the cosets of W in V defined over F form a vector space, called the "Quotient Space" of V by W and denoted $\mathcal{V}/\mathcal{W}$.

Proof:

Let us define the operation of vector addition and scalar multiplication in terms of cosets as follows:

Addition: $(\underline{v} + W) \otimes (\underline{u} + W) = \underline{v} + \underline{u} + W \quad \forall\ \underline{v}, \underline{u} \in V$

Multiplication: $k \odot (\underline{v} + W) = k \cdot \underline{v} + W \quad \forall\ \underline{v} \in V, \ \forall\ k \in F$.

These operations are clearly well-defined. Obviously V/W is closed under vector addition and scalar multiplication. Therefore V/W is a vector space.

Lemma 2.8: If $\mathcal{W}$ is an $\underline{A}$-invariant subspace and $\underline{A}$ is the matrix representation of $\beta : \mathcal{V} \to \mathcal{V}$ then $\underline{A}$ induces a linear operator $\overline{\underline{A}}$ on V/W defined by $\overline{\underline{A}}(\underline{v} + W) = \underline{A}\,\underline{v} + W$. The minimum polynomial of $\overline{\underline{A}}$ divides the minimum polynomial of $\underline{A}$.

Proof:

$\overline{\underline{A}}$ is the matrix representation of $\overline{\beta} : \mathcal{V}/\mathcal{W} \to \mathcal{V}/\mathcal{W}$. As defined above, $\overline{\beta}$ can readily be shown to be a vector space homomorphism. For any $\underline{u} + W = \overline{u} \in V/W$

$$\overline{\underline{A}}^2 (\underline{u} + W) = \underline{A}^2 \underline{u} + W = \underline{A}(\underline{A}\,\underline{u}) + W = \overline{\underline{A}}(\underline{A}\,\underline{u} + W) = \overline{\underline{A}^2}(\underline{u} + W)$$

Similarly $\overline{\underline{A}}^n = \overline{\underline{A}^n}$ for any n. For any polynomial $f(x) = a_n x^n + \ldots + a_1 x + a_0$

$$\overline{f(\underline{A})}\,(\underline{u} + W) = f(\underline{A})\,\underline{u} + W = \sum_{i=0}^{n} a_i \cdot \underline{A}^i \underline{u} + W = \sum_{i=0}^{n} a_i (\underline{A}^i \underline{u} + W)$$

$$= \sum_{i=0}^{n} a_i \cdot \overline{\underline{A}^i}(\underline{u} + W) = \sum_{i=0}^{n} a_i\, \overline{\underline{A}}^i (\underline{u} + W) = f(\overline{\underline{A}})\,(\underline{u} + W)$$

If $\underline{A}$ is the root of $f(x)$ then $f(\underline{A}) = \overline{0} = W = f(\overline{\underline{A}})$. Thus $\overline{\underline{A}}$ is also a root of $f(x)$. Hence the minimum polynomial of $\overline{\underline{A}}$ divides the minimum polynomial of $\underline{A}$.

Lemma 2.9: If $\mathcal{W} \subseteq \mathcal{V}$ such that $\{\underline{w}_1, \underline{w}_2, \ldots, \underline{w}_r\}$ is a basis of $\mathcal{W}$ and the set of cosets $\{\overline{v}_1, \overline{v}_2, \ldots, \overline{v}_s\}$ is a basis of V/W then $B = \{\underline{v}_1, \underline{v}_2, \ldots, \underline{v}_s, \underline{w}_1, \underline{w}_2, \ldots, \underline{w}_r\}$ is a basis of $\mathcal{V}$ and $\text{Dim } \mathcal{V} = \text{Dim } \mathcal{W} + \text{Dim } \mathcal{V}/\mathcal{W}$.

Proof:

If $\underline{u} \in V$ then $\overline{u} = \underline{u} + W = a_1 \cdot \overline{v}_1 + a_2 \cdot \overline{v}_2 + \ldots + a_s \cdot \overline{v}_s$. Therefore $\underline{u} = a_1 \cdot \underline{v}_1 + a_2 \cdot \underline{v}_2 + \ldots + a_s \cdot \underline{v}_s + \underline{w}$ where $\underline{w} \in W$. Hence

$$\underline{u} = a_1 \cdot \underline{v}_1 + a_2 \cdot \underline{v}_2 + \ldots + a_s \cdot \underline{v}_s + b_1 \cdot \underline{w}_1 + b_2 \cdot \underline{w}_2 + \ldots + b_r \cdot \underline{w}_r$$

Thus B spans $\mathcal{V}$. Assume $c_1 \cdot \underline{v}_1 + \ldots + c_s \cdot \underline{v}_s + d_1 \cdot \underline{w}_1 + \ldots + d_r \cdot \underline{w}_r = \underline{0}$. Then $c_1 \cdot \underline{v}_1 + \ldots + c_s \cdot \underline{v}_s + d_1 \cdot \underline{w}_1 + \ldots + d_r \cdot \underline{w}_r + W = \underline{0} + W = \overline{0} = W$. Hence $c_1 \odot \overline{v}_1 \otimes c_2 \odot \overline{v}_2 \otimes \ldots \otimes c_s \odot \overline{v}_s = \overline{0}$. Since $\{\underline{v}_1, \ldots, \underline{v}_s\}$ are linearly independent then $c_i = o$ for $i = 1, 2, \ldots, s$. Substituting in the above expression $d_1 \cdot \underline{w}_1 + \ldots + d_r \cdot \underline{w}_r = \underline{0}$. Since $\{\underline{w}_1, \ldots, \underline{w}_r\}$ are linearly independent then $d_j = o$ for $j = 1, 2, \ldots, r$. Thus B is a linearly indepedent set.

Lemma 2.10: If $\mathcal{W} = Z(\underline{v}, \underline{A})$ and $f(x)^n$ is the $\underline{A}$-annihilator of $\underline{v} \in \mathcal{V}$ where $f(x)$ is a monic, irreducible polynomial of degree d then $f(\underline{A})^s(\mathcal{W})$ is an $\underline{A}$-cyclic subspace generated by $f(\underline{A})^s \underline{v}$ having dimension $d(n-s)$ if $n > s$ and zero if $n \leq s$.

Proof:

Let $G = f(\underline{A})^s(\mathcal{W}) = \{f(\underline{A})^s \underline{w} \mid \underline{w} \in \mathcal{W}\}$. If $f(\underline{A})^s \underline{w}_1$, $f(\underline{A})^s \underline{w}_2 \in G$ and $a_1, a_2 \in F$ then $a_1 \cdot f(\underline{A})^s \underline{w}_1 + a_2 \cdot f(\underline{A})^s \underline{w}_2 =$
$f(\underline{A})^s (a_1 \cdot \underline{w}_1 + a_2 \cdot \underline{w}_2) = f(\underline{A})^s \underline{w}_3 \in G$

$\therefore$ G is a subspace.

If $f(\underline{A})^s \underline{w}_1 \in G$ such that $\underline{w}_1 = g(\underline{A})\underline{v}$ where $g(x) \in F[x]$

then
$$\begin{aligned}\underline{A}(f(\underline{A})^s \underline{w}_1) &= \underline{A} f(\underline{A})^s g(\underline{A})\underline{v}\\ &= f(\underline{A})^s h(\underline{A})\underline{v}\\ &= f(\underline{A})^s \underline{w}_2 \in G\end{aligned}$$

$\therefore$ G is an $\underline{A}$ - invariant subspace.

Let $H = Z(f(\underline{A})^s \underline{v}, \underline{A})$

Let $\underline{g} = f(\underline{A})^s \underline{w}_1 \in G$

$$\underline{g} = f(\underline{A})^s \underline{w}_1 = f(\underline{A})^s g(\underline{A})\underline{v} = g(\underline{A})(f(\underline{A})^s \underline{v}) \in H$$

$\therefore$ $G \subseteq H$

Similarly let $\underline{h} = g(\underline{A})(f(\underline{A})^s \underline{v}) \in H$

$$\underline{h} = g(\underline{A}) f(\underline{A})^s \underline{v} = f(\underline{A})^s (g(\underline{A})\underline{v}) = f(\underline{A})^s \underline{w}_1 \in G$$

$\therefore$ $H \subseteq G$

Hence $H = G$ and G is an $\underline{A}$-cycle subspace generated by $f(\underline{A})^s \underline{v}$. Since $f(x)^n$ is the A-annihilator of $\underline{v}$ then $f(\underline{A})^n \underline{v} = f(\underline{A})^{n-s}(f(\underline{A})^s \underline{v}) = \underline{0}$. Thus the $\underline{A}$-annihilator of $f(\underline{A})^s \underline{v}$ is $f(x)^{n-s}$ and the dimension of G for $n > s$ is $d(n-s)$. If $n \leq s$ the dimension of G is zero since the subspace is non-existent.

<u>Lemma 2.11 :</u> If $\underline{A}$ is the matrix representation of $\beta : \mathcal{V} \longrightarrow \mathcal{V}$ whose minimum polynomial is $f(x)^n$ where $f(x)$ is a monic, irreducible polynomial over F then $\mathcal{V}$ is the direct sum

$\mathcal{V} = Z(\underline{v}_1, \underline{A}) \otimes Z(\underline{v}_2, \underline{A}) \otimes \dots \otimes Z(\underline{v}_r, \underline{A})$ of $\underline{A}$ -cyclic subspaces $Z(\underline{v}_i, \underline{A})$ with corresponding $\underline{A}$ -annihilators $f(x)^{n_1}, f(x)^{n_2}, \dots, f(x)^{n_r}$ where $n = n_1 \geq n_2 \geq \dots \geq n_r$. Hence $\underline{A} = \underline{A}_1 \otimes \underline{A}_2 \otimes \dots \otimes \underline{A}_r$ where $\underline{A}_i$ is the companion matrix of $f(x)^{n_i}$ for $i = 1, 2, \dots, r$.

<u>Proof:</u>

The proof is by induction on the dimension of $\mathcal{V}$. If $\operatorname{Dim} \mathcal{V} = 1$

then $\mathcal{V}$ is $\underline{A}$-cyclic and the lemma holds. Assume the lemma holds for vector spaces of dimension less than that of $\mathcal{V}$.

Since $f(x)^n$ is the minimum polynomial of $\underline{A}$ then there exists a vector $\underline{v}_1 \in V$ having $f(x)^n$ as $\underline{A}$-annihilator. Let $Z_1 = Z(\underline{v}_1, \underline{A})$, $\overline{V} = V/Z_1$ and $\overline{\underline{A}}$ be the matrix representation of the linear operator induced on $\overline{V}$ by $\underline{A}$. By lemma 2.8 the minimum polynomial of $\overline{\underline{A}}$ divides $f(x)^n$. By induction $\overline{V}$ is the direct sum of $\overline{\underline{A}}$ - cyclic subspaces $\overline{\mathcal{V}} = Z(\overline{v}_2, \overline{\underline{A}}) \oplus Z(\overline{v}_3, \overline{\underline{A}}) \oplus \ldots \oplus Z(\overline{v}_r, \overline{\underline{A}})$ where the corresponding $\overline{\underline{A}}$-annihilators are $f(x)^{n_2}, f(x)^{n_3}, \ldots, f(x)^{n_r}$ where $n \geq n_2 \geq n_3 \geq \ldots \geq n_r$.

Let us show the existence of a vector $\underline{v}_2 \in \overline{v}_2$ whose $\underline{A}$-annihilator is $f(x)^{n_2}$. Let $\underline{w}$ be any vector in $\overline{v}_2$, then $f(\underline{A})^{n_2} \underline{w} \in Z_1$. Hence there exists a polynomial $g(x)$ such that $f(\underline{A})^{n_2} \underline{w} = g(\underline{A}) \underline{v}_1$. Since $f(x)^n$ is the minimum polynomial of $\underline{A}$ then

$$\underline{0} = f(\underline{A})^n \underline{w} = f(\underline{A})^{n-n_2} g(\underline{A})\underline{v}_1 \text{ implies } f(x)^n \mid g(x) f(x)^{n-n_2} \text{ and}$$

$g(x) = h(x) f(x)^{n_2}$ for some $h(x) \in F[x]$.

$$\text{Let } \underline{v}_2 = \underline{w} - h(\underline{A}) \underline{v}_1$$

Since $\underline{w} - \underline{v}_2 = h(\underline{A}) \underline{v}_1 \in Z_1$ then $\underline{v}_2 \in \overline{v}_2$.

$$\therefore \quad f(\underline{A})^{n_2} \underline{v}_2 = f(\underline{A})^{n_2} (\underline{w} - h(\underline{A})\underline{v}_1) = f(\underline{A})^{n_2} \underline{w} - g(\underline{A}) \underline{v}_1 = \underline{0}$$

Thus the $\underline{A}$-annihilator of $\underline{v}_2$ is $f(x)^{n_2}$.

Similarly there exists vectors $\underline{v}_3, \ldots, \underline{v}_r \in V$ such that $\underline{v}_i \in \overline{v}_i$ and the $\underline{A}$-annihilator of $\underline{v}_i$ is $f(x)^{n_i}$, the $\overline{\underline{A}}$-annihilator of $\overline{v}_i$.

Let $Z_i = Z(\underline{v}_i, \underline{A})$ for $i = 2, \ldots, r$. If the degree of $f(x)$ is d then $\{\underline{v}_i, \underline{A}\underline{v}_i, \ldots, \underline{A}^{dn_i - 1} \underline{v}_i\}$ and $\{\overline{v}_i, \overline{\underline{A}}\,\overline{v}_i, \ldots, \overline{\underline{A}}^{dn_i - 1} \overline{v}_i\}$ are bases for $Z(\underline{v}_i, \underline{A})$ and $Z(\overline{v}_i, \overline{\underline{A}})$ respectively for $i = 2, \ldots, r$. Thus

$\{\bar{v}_2, \ldots, \bar{A}^{dn_2-1}\bar{v}_2; \ldots; \bar{v}_r, \ldots, \bar{A}^{dn_r-1}\bar{v}_r\}$ is a basis for $\bar{\mathcal{V}}$. By lemmas 2.8 and 2.9 it follows that $\{\underline{v}_1, \ldots, \underline{A}^{dn_1-1}\underline{v}_1; \underline{v}_2, \ldots, \underline{A}^{dn_2-1}\underline{v}_2; \ldots$ $\ldots; \underline{v}_r, \ldots, \underline{A}^{dn_r-1}\underline{v}_r\}$ is a basis for $\mathcal{V}$.

$\therefore \mathcal{V} = Z(\underline{v}_1, \underline{A}) \otimes Z(\underline{v}_2, \underline{A}) \otimes \cdots \otimes Z(\underline{v}_r; \underline{A})$.

Let us show the exponents $n_1, n_2, \ldots, n_r$ are uniquely determined by $\underline{A}$ and $\mathcal{V}$.

Dim $\mathcal{V} = d(n_1 + n_2 + \ldots + n_r)$ and Dim $Z_i = dn_i$ for $i = 1, 2, \ldots, r$. Any vector $\underline{v} \in \mathcal{V}$ can be uniquely represented in the form $\underline{v} = \underline{w}_1 + \underline{w}_2 + \ldots + \underline{w}_r$. Hence for $s \geq o$

$$f(\underline{A})^s \underline{v} = f(\underline{A})^s \underline{w}_1 + f(\underline{A})^s \underline{w}_2 + \ldots + f(\underline{A})^s \underline{w}_r$$

$\therefore \quad f(\underline{A})^s (\mathcal{V}) = f(\underline{A})^s (Z_1) \otimes f(\underline{A})^s (Z_2) \otimes \ldots \otimes f(\underline{A})^s (Z_r)$ where

$f(\underline{A})^s \underline{w}_i \in f(\underline{A})^s (Z_i)$. Let q be an integer, dependent on s, so that $n_1 > s, n_2 > s, \ldots, n_q > s, n_{q+1} \leq s, \ldots$.

By lemma 2.10 Dim $f(\underline{A})^s (\mathcal{V}) = d\,[\,(n_1 - s) + (n_2 - s) + \ldots + (n_q - s)\,]$

For a given value of s Dim $f(\underline{A})^s (\mathcal{V})$ depends only on $\underline{A}$ and $\mathcal{V}$. If set $s = n-j$ then the above expression determines the number of n_i equal to $(n-j+1)$. Repeat the process until $s = o$. Thus n_i are uniquely determined by $\underline{A}$ and $\mathcal{V}$.

Theorem 2.2 : "SECONDARY DECOMPOSITION THEOREM"

If $\underline{A}$ is the matrix representation of $\beta : \mathcal{V} \to \mathcal{V}$ whose minimum polynomial is $m(x) = f_1(x)^{m_1} f_2(x)^{m_2} \ldots f_s(x)^{m_s}$ where $f_i(x)$ are distinct, monic, irreducible polynomials, then $\underline{A}$ has a unique block diagonal matrix representation

$$\begin{bmatrix} \underline{A}_{11} & & & & & \\ & \ddots & & & \underline{O} & \\ & & \underline{A}_{1r_1} & & & \\ & & & \ddots & & \\ & \underline{O} & & & \underline{A}_{s1} & \\ & & & & & \ddots \\ & & & & & & \underline{A}_{sr_s} \end{bmatrix}$$

where $\underline{A}_{ij}$ are the "companion matrices" of the polynomials $f_i(x)^{n_{ij}}$ where

$$m_1 = n_{11} = n_{12} \geq \ldots \geq n_{1r_1}; \ldots ; \; m_s = n_{s_1} \geq n_{s_2} \geq \ldots \geq n_{sr_s}.$$

<u>Proof:</u>

The theorem follows directly from lemma 2.11 and theorem 2.1. The above matrix representation of $\underline{A}$ is known as the "Rational Canonical Form". The polynomials $f_i(x)^{n_{ij}}$ are called the "elementary divisors" of $\underline{A}$.

<u>Lemma 2.12:</u> Let $\underline{A}$ be the matrix representation of $\beta : \mathcal{V} \longrightarrow \mathcal{V}$ whose minimum polynomial is $m(x) = f(x)^e = (x^r + c_{r-1}x^{r-1} + \ldots + c_1 x + c_o)^e$ where $f(x)$ is a monic, irreducible polynomial over F. The matrix representation $\underline{A}^*$ of β with respect to the basis $\mathcal{B}$ is:

$$\underline{A}^* = \begin{bmatrix} \underline{C}^T & & & \\ \underline{N} & \underline{C}^T & & \underline{0} \\ & & \ddots & \\ & & & \underline{C}^T \\ & \underline{0} & & \underline{N} & \underline{C}^T \end{bmatrix} \quad \text{where} \quad N = \begin{bmatrix} 0 & \cdots & \cdots & 1 \\ & \ddots & & \vdots \\ \underline{0} & & \ddots & \vdots \\ & & & 0 \end{bmatrix}$$

and $\underline{C}$ is the companion matrix of $f(x)$.

Proof:

Let $\underline{v} \in \mathcal{V}$ be a non-zero vector whose minimum polynomial is $m(x)$.

Let $B = \{\underline{b}_1, \underline{b}_2, \ldots, \underline{b}_r; \underline{b}_{r+1}, \ldots, \underline{b}_{2r}; \ldots; \underline{b}_{(e-1)r+1}, \ldots, \underline{b}_{er}\}$

where

$$\begin{array}{llll}
\underline{b}_1 = \underline{v}, & \underline{b}_2 = \underline{A}\,\underline{v}, & \ldots\ldots & \underline{b}_r = \underline{A}^{r-1}\,\underline{v} \\
\underline{b}_{r+1} = f(\underline{A})\underline{v}, & \underline{b}_{r+2} = \underline{A}f(\underline{A})\,\underline{v}, & \ldots\ldots & \underline{b}_{2r} = \underline{A}^{r-1}f(\underline{A})\,\underline{v} \\
\vdots & \vdots & & \vdots \\
\underline{b}_{(e-1)r+1} = f(\underline{A})^{e-1}\underline{v}, & \underline{b}_{(e-1)r+2} = \underline{A}f(\underline{A})^{e-1}\underline{v} & \ldots\ldots & \underline{b}_{er} = \underline{A}^{r-1}f(\underline{A})^{e-1}\underline{v}
\end{array}$$

Assume the elements of B are linearly dependent. Then there exists a set of scalars $\{a_1, a_2, \ldots, a_{er}\}$ NOT ALL ZERO such that:

$$a_1 \cdot \underline{b}_1 + a_2 \cdot \underline{b}_2 + \ldots\ldots + a_{er} \cdot \underline{b}_{er} = \underline{0}$$

$$a_1 \cdot \underline{v} + a_2 \cdot \underline{A}\,\underline{v} + \ldots + a_r \cdot \underline{A}^{r-1}\underline{v} + a_{r+1} \cdot f(\underline{A})\,\underline{v} + \ldots\ldots a_{er} \cdot \underline{b}_{er} = \underline{0}$$

$$a_1 \cdot \underline{v} + a_2 \cdot \underline{A}\underline{v} + \ldots + a_r \cdot A^{r-1}\underline{v} + a_{r+1} \cdot [\underline{A}^r + c_{r-1}\underline{A}^{r-1} + \ldots + c_1\underline{A} + c_0\underline{I}]\underline{v} + \ldots + a_{er} \cdot \underline{b}_{er} = \underline{0}$$

$$d_1 \cdot \underline{v} + d_2 \cdot \underline{A}\underline{v} + \ldots + d_r \cdot \underline{A}^{r-1}\underline{v} + a_{r+1} \cdot \underline{A}^r\underline{v} + \ldots\ldots + a_{er} \cdot \underline{b}_{er} = \underline{0}$$

where $d_i = a_i + a_{r+1}\, c_{i-1}$ for $i = 1, 2, \ldots, r$. Expanding the remaining terms in a similar manner we obtain $e_1 \cdot \underline{v} + e_2\,\underline{A}\,\underline{v} + \ldots + e_{er} \cdot \underline{A}^{er-1}\underline{v} = \underline{0}$ where e_i for $i = 1, 2, \ldots, er$ is a linear combination of a_j for $j = 1, \ldots, er$.

$\therefore\ (e_1\underline{I} + e_2\underline{A} + \ldots + e_{er}\,\underline{A}^{er-1})\,\underline{v} = \underline{0}$. Therefore there exists a polynomial $p(x) = c_1 + c_2 x + \ldots + c_{er} x^{er-1}$ of degree less than $e \cdot r$ such that $p(\underline{A})\,\underline{v} = \underline{0}$. This is a contradiction. Hence the elements of B are linearly independent and may be used as a basis for $\mathcal{V}$.

Let us show that the matrix representation of β with respect to the basis B is $\underline{A}^*$.

$$\beta\ (\underline{b}_1) = \underline{A}\,\underline{b}_1 = \underline{b}_2$$

$$\beta\ (\underline{b}_2) = \underline{A}\underline{b}_2 = \underline{b}_3$$

$$\vdots \qquad\qquad \vdots$$

$$\beta\ (\underline{b}_r) = \underline{A}\underline{b}_r = \underline{A}^r\underline{v} = [\, f(\underline{A}) - c_o\underline{I} - c_1 \cdot \underline{A} \ldots\ldots - c_{r-1} \cdot \underline{A}^{r-1}\,]\,\underline{v}$$

$$= f\,(\underline{A})\,\underline{v} - c_o \cdot \underline{v} - c_1 \cdot \underline{A}\underline{v} \ldots\ldots - c_{r-1} \cdot \underline{A}^{r-1}\,\underline{v}$$

$$= \underline{b}_{r+1} - c_o \cdot \underline{b}_1 - c_1 \cdot \underline{b}_2 \ldots\ldots - c_{r-1} \cdot \underline{b}_r$$

$$\beta\ (\underline{b}_{r+1}) = \underline{A}\,\underline{b}_{r+1} = \underline{b}_{r+2}$$

$$\beta\ (\underline{b}_{r+2}) = \underline{A}\underline{b}_{r+2} = \underline{b}_{r+3}$$

$$\vdots \qquad\qquad \vdots$$

$$\beta\ (\underline{b}_{2r}) = \underline{A}\underline{b}_{2r} = \underline{A}^r f(\underline{A})\underline{v} = [f(\underline{A}) - c_o\underline{I} - c_1\,\underline{A} \ldots - c_{r-1}\underline{A}^{r-1}]\,f(\underline{A})\,\underline{v}$$

$$= f(\underline{A})^2\underline{v} - c_o \cdot f(\underline{A})\underline{v}. - c_1 \cdot \underline{A}f(\underline{A})\underline{v} \ldots - c_{r-1} \cdot \underline{A}^{r-1} f(\underline{A})\underline{v}$$

$$= \underline{b}_{2r+1} - c_o \cdot \underline{b}_{r+1} \ldots\ldots \qquad - c_{r-1} \cdot \underline{b}_{2r}$$

$$\vdots \qquad\qquad \vdots$$

$$\beta(\underline{b}_{(e-1)r+1}) = \underline{A}\underline{b}_{(e-1)r+1} = \underline{b}_{(e-1)r+2}$$

$$\beta(\underline{b}_{(e-1)r+2}) = \underline{A}\underline{b}_{(e-1)r+2} = \underline{b}_{(e-1)r+3}$$

$$\vdots \qquad\qquad \vdots$$

$$\beta(\underline{b}_{er}) = \underline{A}\underline{b}_{er} = \underline{A}^r f(\underline{A})^{e-1}\,\underline{v} = [f(\underline{A}) - c_o\underline{I} - c_1\underline{A} \ldots - c_{r-1}\underline{A}^{r-1}]f(\underline{A})^{e-1}\underline{v}$$

$$= f(\underline{A})^e\underline{v} - c_o \cdot f(\underline{A})^{e-1}\underline{v} \ldots - c_{r-1} \cdot \underline{A}^{r-1} f(\underline{A})^{e-1}\underline{v}$$

$$= -c_o \cdot \underline{b}_{(e-1)r+1} \ldots - c_{r-1} \cdot \underline{b}_{er}$$

The above equations can be expressed in matrix form as follows:

$$
\begin{bmatrix}
\beta(\underline{b}_1) \\
\beta(\underline{b}_2) \\
\vdots \\
\beta(\underline{b}_r) \\
\beta(\underline{b}_{r+1}) \\
\vdots \\
\beta(\underline{b}_{(e-1)r+1}) \\
\vdots \\
\beta(\underline{b}_{er})
\end{bmatrix}
=
\left[\begin{array}{ccccc:ccc:ccccc}
0 & 1 & \cdots & & 0 & & & & & & & & \\
\vdots & \ddots & & & \vdots & & & & & & & & \\
\vdots & \underline{0} & \ddots & & \vdots & & & & & & \underline{O} & & \\
\vdots & & & \ddots & 1 & & & & & & & & \\
-a_o & \cdots & \cdots & \cdots & -a_{r-1} & 1 & & & & & & & \\
\hdashline
 & & & & & \ddots & & & & & & & \\
 & & & & & & \ddots & & & & & & \\
 & & & & & & & 1 & & & & & \\
\hdashline
 & & & & & & & & 0 & 1 & \cdots & & 0 \\
 & & \underline{O} & & & & & & \vdots & \ddots & & & \vdots \\
 & & & & & & & & \vdots & \underline{0} & \ddots & & 1 \\
 & & & & & & & & -a_o & \cdots & \cdots & \cdots & a_{r-1}
\end{array}\right]
\begin{bmatrix}
\underline{b}_1 \\
\underline{b}_2 \\
\vdots \\
\underline{b}_r \\
\vdots \\
\underline{b}_{er}
\end{bmatrix}
$$

$$
=
\begin{bmatrix}
\underline{C} & \underline{N}^T & & & \\
 & \underline{C} & \underline{N}^T & & \underline{0} \\
 & & \ddots & \ddots & \\
 & \underline{0} & & \ddots & \underline{N}^T \\
 & & & \underline{C} & \underline{N}^T
\end{bmatrix}
\begin{bmatrix}
\underline{b}_1 \\
\underline{b}_2 \\
\vdots \\
\underline{b}_{er}
\end{bmatrix}
$$

The above matrix of coefficients is known as the " JACOBSON CANONICAL FORM" or "HYPERCOMPANION MATRIX" of all matrices similar to $\underline{A}$.

In view of our vertical coordinate representation for vectors, the matrix representation $\underline{A}^*$ of β with respect to the basis B is the transpose of the above matrix.

$$\text{Hence } \underline{A}^* = \begin{bmatrix} \underline{C}^T & & & & \\ \underline{N} & \underline{C}^T & & & \\ & \ddots & \ddots & & \\ & & \ddots & \underline{C}^T & \\ & & & \underline{N} & \underline{C}^T \end{bmatrix}$$

Since the number of references used in this chapter are very large and overlapping, they were not presented with the theorems and definitions. The references are [2, 3, 6, 14, 16] as per the Bibliography.

CHAPTER III

CONTROLLABILITY AND OBSERVABILITY OF LINEAR SYSTEMS

The purpose of this chapter is to discuss the concepts of controllability and observability of linear systems. First we define systems in their most general form and then we consider the three classes of systems, continuous, sampled-data and discrete. The concepts of controllability and observability, from the point of view of cyclic subspaces, are developed in detail for linear time-invariant systems.

III-1 SYSTEM DEFINITION

Definition 3.1: A system is a mathematical structure $S = < \mathcal{U}, \mathcal{Y}, \mathcal{X}, \delta, \lambda >$ consisting of $\mathcal{U}$: An input space which leads in a natural way to an input segment space $\mathcal{U}^*$ where an input segment is a mapping $\mu : (t_o, t) \rightarrow \mathcal{U}^*$; T: A set of values of time at which the behaviour of the system is defined:

$\mathcal{Y}$: an output space

$\mathcal{X}$: a state space

δ : an output function, $\delta : \mathcal{X} \times \mathcal{U} \times T \longrightarrow \mathcal{Y}$

λ : a state transition function, $\lambda : \mathcal{X} \times \mathcal{U}^* \times T \longrightarrow \mathcal{X}$

Hence for any t, $t_o \in T$, any initial state $\underline{x}(t_o) = \underline{x}_o$ and any admissible input segment $\underline{\mu}(t_o, t)$:

$$\underline{x}(t) = \lambda(\underline{x}_o, \underline{\mu}(t_o, t), t)$$

$$\underline{y}(t) = \delta(\underline{x}(t), \underline{\mu}(t), t)$$

These are the "state equations" of the system S.

Eliminating $\underline{x}(t)$ from the above expression for $\underline{y}(t)$ we obtain:

$$\underline{y}(t) = F(\underline{x}_o, \underline{\mu}(t_o, t), t)$$

This is the "input-output" equation of the system S.

<u>Definition 3.2:</u> A system is time-invariant if

$\underline{\mu}(t_o, t) = \underline{\mu}(t_o + \tau, t + \tau)$ implies

$F(\underline{x}_o, \underline{\mu}(t_o, t), t) = F(\underline{x}_o, \underline{\mu}(t_o + \tau, t + \tau), t + \tau)$ for all inital states $\underline{x}_o \in \mathcal{X}$, all $\underline{\mu} \in \mathcal{U}^*$, all $t_o, t \in T$ and $\tau \geq 0$ such that $t_o + \tau$, $t + \tau \in T$.

<u>Definition 3.3:</u> A state $\underline{z}$ of a system S is a zero state if $\underline{y}(t) = F(\underline{z}, \underline{0}_t, t) = \underline{0}_t$ for all $t \in T$ where $\underline{0}_t$ is the zero input segment of duration t.

<u>Definition 3.4</u>: A system S is linear if it satisfies the following properties:

<u>1 - Decomposition Property:</u>

$$F(\underline{x}_o, \underline{\mu}(t_o, t), t) = F(\underline{x}_o, \underline{0}_t, t) + F(\underline{z}, \underline{\mu}(t_o, t), t)$$

For all $\underline{x}_o \in \mathcal{X}$, $\underline{\mu} \in \mathcal{U}^*$.

<u>2 - Zero State Linearity:</u>

$$F(\underline{z}, k(\underline{\mu}_1(t_o, t) - \underline{\mu}_2(t_o, t)), t) = k\, F(\underline{z}, \underline{\mu}_1(t_o, t), t) - k\, F(\underline{z}_1, \underline{\mu}_2(t_o, t), t)$$

For all $\underline{\mu}_1, \underline{\mu}_2 \in \mathcal{U}^*$ and all k for which the products are defined.

<u>3 - Zero Input Linearity:</u>

$$F(k(\underline{x}_1 - \underline{x}_2), \underline{0}_t, t) = k\, F(\underline{x}_1, \underline{0}_t, t) - k\, F(\underline{x}_2, \underline{0}_t, t)$$

For all $\underline{x}_1, \underline{x}_2 \in \mathcal{X}$ and all k for which the products are defined.

At this point we shall restrict our discussion to the different types of linear time-invariant systems:

<u>i - Continuous Systems:</u>

A continuous linear, time-invariant system is generally described by a set of linear, first order differential equations with constant coefficients called the state equations of the system of the form

$$\dot{\underline{x}}(t) = \frac{d}{dt}\underline{x}(t) = \underline{A}\,\underline{x}(t) + \underline{B}\,\underline{\mu}(t) \qquad \underline{x}(t_o) = \underline{x}_o$$

$$\underline{y}(t) = \underline{C}\,\underline{x}(t) + \underline{D}\,\underline{\mu}(t)$$

where $\underline{x}(t)$ is an n-dimensional state vector

$\underline{y}(t)$ is a p-dimensional output vector

$\underline{\mu}(t)$ is a r-dimensional input vector

and $\underline{A}$, $\underline{B}$, $\underline{C}$ and $\underline{D}$ are n x n, n x r, p x n and p x r constant matrices respectively. The set $T = [\,o, \infty)$ is an interval of the real line and since the input, output and state spaces are Euclidean spaces then S is known as a continuous-state system. For such systems both time and the input segments are continuous.

ii - Sampled-Data Systems.

A linear, time-invariant sampled-data system is generally described by a set of first order difference equations with constant coefficients, called the state equations of the system of the form:

$$\underline{x}\,[\,(k+1)T\,] = \underline{A}\,\underline{x}\,(kT) + \underline{B}\,\underline{\mu}\,(kT) \qquad \underline{x}(t_o) = \underline{x}(o) = \underline{x}_o$$

$$\underline{y}(kT) = \underline{C}\underline{x}\,(kT) + \underline{D}\underline{\mu}\,(kT)$$

where T is the sampling rate in units of time per interval and $k = 0, 1, 2, \ldots, \infty$. Unless otherwise stated the symbols above have the same meaning as in the previous system. In view of the "sample and hold" operation, the set $T = \{\,0, T, 2T, \ldots.\}$ and at certain points within the system the signals are discrete or discontinuous although the input and output segments of the system are continuous. The input, output and state spaces are Euclidean spaces. For such systems, the time is discrete while the input and output segments are continuous.

iii - Discrete Systems.

A linear, time invariant discrete system known in automata theory as a "linear sequential machine" is generally described by a set of first order difference equations with constant coefficients, called the state equations of the system of the form:

$$\underline{x}'(t) = \underline{x}(t+1) = \underline{A}\,\underline{x}(t) + \underline{B}\,\underline{\mu}(t) \qquad \underline{x}(t_o) = \underline{x}_o$$

$$\underline{y}(t) = \underline{C}\,\underline{x}(t) + \underline{D}\,\underline{\mu}(t)$$

Unless otherwise stated, the symbols above have the same meaning as in the previous system. The set $T = \{0, 1, 2, \ldots, \infty\}$. Since the input segments are discrete, the input segment space $\mathcal{U}^*$ is defined, using the star operator, as follows:

$\mathcal{U}^* = \{\lambda + \mathcal{U} + \mathcal{U}^2 + \ldots\ldots\}$ where λ is the null sequence of zero length (no input). Since the input, output and state spaces are finite discrete sets then S is known as a discrete state system.

In view of the similarity in the state equations of the above systems, the concepts of controllability and observability to be developed are applicable to all of them.

III-2 CONTROLLABILITY:

Definition 3.5: A state $\underline{x}_2 \in \mathcal{X}$ of S is reachable from a state $\underline{x}_1 \in \mathcal{X}$ of S if there exists an input segment $\underline{\mu} \in \mathcal{U}^*$ such that $\lambda(\underline{x}_1, \underline{\mu}) = \underline{x}_2$.

Definition 3.6: A system is strongly connected if every state is reachable from every other state.

Although it is beyond the scope of this thesis, it may readily be shown that a system is controllable if and only if it is strongly connected.

Theorem 3.1 [8]: A linear, time-invariant system S is controllable if and only if the matrix

$$\underline{K} = [\underline{B}, \underline{A}\,\underline{B}, \ldots, \underline{A}^{n-1}\,\underline{B}]$$

has rank n.

From a synthesis point of view, the following sufficient conditions hold.

<u>Theorem 3.2 :</u> Let S be a linear, time-invariant system with $\underline{B} = [\underline{b}_1, \underline{b}_2, \ldots, \underline{b}_s]$ where $s \geq r$ and $\underline{A}$ the matrix representation of $\beta : \mathcal{X} \to \mathcal{X}$ whose elementary divisors are $\ell_1(x), \ell_2(x), \ldots, \ell_r(x)$. S is controllable if there exist r elements $\underline{b}_j \in \mathcal{X}$ for $1 \leq j \leq s$ such that

$$\mathcal{X} = Z(\underline{b}_{j1}, \underline{A}) \otimes Z(\underline{b}_{j2}, \underline{A}) \otimes \ldots \otimes Z(\underline{b}_{jr}, A)$$

<u>Proof:</u>

Assume there exist r elements $\underline{b}_{j1}, \underline{b}_{j2}, \ldots, \underline{b}_{jr}$ appearing as columns of $\underline{B}$ such that

$$\mathcal{X} = Z(\underline{b}_{j1}, \underline{A}) \otimes Z(\underline{b}_{j2}, \underline{A}) \otimes \ldots \otimes Z(\underline{b}_{jr}, \underline{A})$$

By theorem 2.2 $\underline{A} = \underline{A}_1 \otimes \underline{A}_2 \otimes \ldots \otimes \underline{A}_r$ where $\underline{A}_i$, the restriction of $\underline{A}$ to $Z(\underline{b}_{ji}, \underline{A})$, is the companion matrix of $\ell_i(x) = x^{n_i} + a_{n_i-1} x^{n_i-1} + \ldots + a_1 x + a_o$ for $i = 1, 2, \ldots, r$. The matrix $\underline{K} = [\underline{B}, \underline{A}\underline{B}, \ldots, \underline{A}^{n-1}\underline{B}]$ may be rewritten as follows:

$$\underline{K}' = [\underline{b}_{j1}, \underline{A}\,\underline{b}_{j1}, \ldots, \underline{A}^{n-1}\underline{b}_{j1}; \underline{b}_{j2}, \underline{A}\,\underline{b}_{j2}, \ldots, \underline{A}^{n-1}\underline{b}_{j2}; \ldots$$
$$\ldots; \underline{b}_{js}, \underline{A}\underline{b}_{js}, \ldots, \underline{A}^{n-1}\underline{b}_{js} \quad]$$

Permuting the columns of $\underline{K}$ to obtain the above form does not affect the rank of $\underline{K}$. The set of vectors $\{\underline{b}_{ji}, \underline{A}\underline{b}_{ji}, \ldots, \underline{A}^{n_i-1}\underline{b}_{ji}\}$ is a basis for $Z(\underline{b}_{ji}, \underline{A})$ such that the set of vectors

$$\{\underline{b}_{j1}, \underline{A}\underline{b}_{j1}, \ldots, \underline{A}^{n_1-1}\underline{b}_{j1}; \underline{b}_{j2}, \underline{A}\underline{b}_{j2}, \ldots, \underline{A}^{n_2-1}\underline{b}_{j2}; \ldots; \underline{b}_{jr}, \underline{A}\underline{b}_{jr}, \ldots, \underline{A}^{n_r-1}\underline{b}_{jr}\}$$

is a basis for $\mathcal{X}$. The remaining columns of $\underline{K}'$ are linear combinations of the set of basis vectors for $\mathcal{X}$. Hence by a sequence of column operations, $\underline{K}'$ may be reduced to the form:

$$\underline{K}'' = [\underline{b}_{j1}, \underline{A}\underline{b}_{j1}, \ldots, \underline{A}^{n_1-1}\underline{b}_{j1}; \underline{b}_{j2}, \underline{A}\underline{b}_{j2}, \ldots, \underline{A}^{n_2-1}\underline{b}_{j2}; \ldots; \underline{b}_{jr}, \ldots, \underline{A}^{n_r-1}\underline{b}_{jr}; \ \underline{0}\ ]$$

From the form of $\underline{K}''$, $\underline{K}$ obviously has rank n and the system is controllable.

In terms of cyclic subspaces, the above theorem may be restated as follows: The system S is controllable if the r cyclic subspaces are represented respectively by r independent vectors acting as columns of $\underline{B}$.

Let us consider a few examples in which the matrix $\underline{A}$ is derogatory.

Example 3.1: Consider the following multi-variable control system defined over R, the field of real numbers.

$$\dot{\underline{x}}(t) = \underline{A}\underline{x}(t) + \underline{B}\underline{\mu}(t) = \begin{bmatrix} -1 & -1 & 1 & 0 & -1 \\ 0 & -1 & 0 & -1 & 0 \\ 0 & -1 & -1 & -1 & 0 \\ 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 & -1 \end{bmatrix} \underline{x}(t) + \begin{bmatrix} 0 & 0 & 3 \\ 0 & 0 & 3 \\ -2 & 1 & 3 \\ 0 & 0 & 0 \\ 0 & 1 & 3 \end{bmatrix} \begin{bmatrix} \mu_1(t) \\ \mu_2(t) \\ \mu_3(t) \end{bmatrix}$$

After calculation we find $m(x) = (x+1)^2 (x^2 + 2x+2)$ and $p(x) = m(x)(x+1)$. Thus $\underline{A}$ has three elementary divisors $x+1$, $(x+1)^2$ and $x^2 + 2x + 2$ such that $\mathcal{X} = Z(\underline{b}_1, \underline{A}) \otimes Z(\underline{b}_2, \underline{A}) \otimes Z(\underline{b}_3, \underline{A})$ where the minimum polynomial of $\underline{b}_1$, $\underline{b}_2$ and $\underline{b}_3$ is $(x+1)^2$, $x + 1$ and $x^2 + 2x + 2$ respectively. By theorem 3.2 the system is controllable. Upon verification,

$$\underline{K} = [\ \underline{B},\ \underline{A}\underline{B},\ \underline{A}^2\underline{B},\ \underline{A}^3\underline{B},\ \underline{A}^4\underline{B}\ ]$$

$$= \begin{array}{c} \left[\begin{array}{ccc|ccc|ccc|ccc|ccc} 0 & 0 & 3 & -2 & 0 & -6 & 0 & 0 & 6 & -2 & 0 & 0 & 0 & 0 & -12 \\ 0 & 0 & 3 & 0 & 0 & -3 & 0 & 0 & 0 & 0 & 0 & 6 & 0 & 0 & -12 \\ -2 & 1 & 3 & 2 & -1 & -6 & -2 & 1 & 6 & 2 & -1 & 0 & -2 & 1 & -12 \\ 0 & 0 & 0 & 0 & 0 & 3 & 0 & 0 & -6 & 0 & 0 & 6 & 0 & 0 & 0 \\ 0 & 1 & 3 & 0 & -1 & -3 & 0 & 1 & 0 & 0 & -1 & 6 & 0 & 1 & -12 \end{array}\right] \\ \quad \underline{B} \qquad\qquad \underline{A}\underline{B} \qquad\qquad \underline{A}^2\underline{B} \qquad\qquad \underline{A}^3\underline{B} \qquad\qquad \underline{A}^4\underline{B} \end{array}$$

Upon using a series of elementary column operations $\underline{K}$ reduces to:

$$\underline{K}' = \left[\begin{array}{cc|c|cc|c} 0 & -2 & 0 & 3 & -6 & \\ 0 & 0 & 0 & 3 & -3 & \\ -2 & 2 & 1 & 3 & -6 & \underline{0} \\ 0 & 0 & 0 & 0 & 3 & \\ 0 & 0 & 1 & 3 & -3 & \end{array}\right] = [\underline{b}_1 \;\; \underline{A}\underline{b}_1 \,\vdots\, \underline{b}_2 \,\vdots\, \underline{b}_3 \;\; \underline{A}\,\underline{b}_3 \,\vdots\, \underline{0}]$$

Since $\underline{K}$ has rank 5, the system is controllable by theorem 3.1.

Example 3.2: Consider the following linear sequential machine defined over $GF(3) = \{0, 1, 2\}$.

$$\underline{x}'(t) = \underline{A}\,\underline{x}(t) + \underline{B}\underline{\mu}(t) = \begin{bmatrix} 2 & 0 & 0 & 1 \\ 0 & 0 & 0 & 2 \\ 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 2 \end{bmatrix} \underline{x}(t) + \begin{bmatrix} 2 & 0 & 2 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \mu_1(t) \\ \mu_2(t) \\ \mu_3(t) \end{bmatrix}$$

After calculation we find $m(x) = (x+2)^2 (x+1)$ and $p(x) = m(x)\,(x+2)$. Thus $\underline{A}$ has three elementary divisors $(x+2)^2$, $x+2$ and $x+1$ such that $\mathcal{X} = Z(\underline{b}_1, \underline{A}) \otimes Z(\underline{b}_2, \underline{A}) \otimes Z(\underline{b}_3, \underline{A})$ where the minimum polynomial of $\underline{b}_1$, $\underline{b}_2$ and $\underline{b}_3$ is $(x+2)^2$, $x + 2$ and $x+1$ respectively. By theorem 3.2 the system is controllable. Upon verification,

$$\underline{K} = [\underline{B},\ \underline{A}\,\underline{B},\ \underline{A}^2\,\underline{B},\ \underline{A}^3\,\underline{B}]$$

$$= \left[\begin{array}{ccc|ccc|ccc|ccc} 2 & 0 & 2 & 1 & 0 & 1 & 0 & 0 & 2 & 2 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 2 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 0 & 2 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 2 & 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

$$\underline{B} \qquad\qquad \underline{A}\,\underline{B} \qquad\qquad \underline{A}^2\,\underline{B} \qquad\qquad \underline{A}^3\,\underline{B}$$

Upon using a series of elementary column operations $\underline{K}$ reduces to:

$$\underline{K}' = \left[\begin{array}{cc|c|c|c} 2 & 0 & 2 & 1 & \\ 0 & 0 & 1 & 0 & \underline{0} \\ 0 & 1 & 0 & 1 & \\ 0 & 0 & 0 & 1 & \end{array}\right] = [\underline{b}_1 \;\; \underline{A}\,\underline{b}_1 \mid \underline{b}_2 \mid \underline{b}_3 \mid \underline{0}\,]$$

Since K has rank 4, the system is controllable by theorem 3.1.

Let us now consider a few examples in which the matrix $\underline{A}$ is non-derogatory.

Example 3.3: Consider the following linear sequential machine defined over GF(2).

$$\underline{x}'(t) = \underline{A}\,\underline{x}\,(t) + \underline{B}\,\underline{\mu}(t) = \begin{bmatrix} 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{bmatrix} \underline{x}(t) + \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 1 \\ 0 & 1 \end{bmatrix} \underline{\mu}\,(t)$$

$m(x) = p(x) = (x+1)^2\,(x^2+x+1)$ and $\mathcal{X} = Z(\underline{b}_1, \underline{A}) \oplus Z(\underline{b}_2, \underline{A})$ where $\underline{b}_1$ and $\underline{b}_2$ have minimum polynomials $(x+1)^2$ and $x^2 + x + 1$ respectively. By theorem 3.2 the system is controllable. Upon verfication,

$$\underline{K} = [\underline{B}\,,\ \underline{A}\,\underline{B},\ \underline{A}^2\,\underline{B},\ \underline{A}^3\,\underline{B}\,]$$

$$= \begin{array}{c} \left[\begin{array}{cc|cc|cc|cc} 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \end{array}\right] \\ \begin{array}{cccc} \underline{B} & \quad\underline{A}\,\underline{B} & \quad\underline{A}^2\underline{B} & \quad\underline{A}^3\underline{B} \end{array} \end{array}$$

Using a sequence of column operations $\underline{K}$ reduces to:

$$\underline{K}' = \left[\begin{array}{cc|cc|c} 0 & 1 & 0 & 0 & \\ 0 & 0 & 0 & 1 & \\ 1 & 0 & 1 & 1 & \underline{0} \\ 0 & 0 & 1 & 0 & \end{array}\right] = [\underline{b}_1\ \underline{A}\underline{b}_1 \mid \underline{b}_2\ \underline{A}\underline{b}_2 \mid \underline{0}\,]$$

Since $\underline{K}$ has rank 4, the system is controllable..

Example 3.4: Consider the following multi-variable control system defined over R.

$$\dot{\underline{x}}(t) = \underline{A}\underline{x}(t) + \underline{B}\underline{\mu}(t) = \begin{bmatrix} -3 & 1 & 0 & 0 \\ 0 & -3 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -2 & -2 \end{bmatrix} \underline{x}(t) + \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} \underline{\mu}(t)$$

$m(x) = (x+3)^2(x^2+2x+2) = p(x)$ and $\mathcal{X} = Z(\underline{b}_1, \underline{A}) \oplus Z(\underline{b}_2, \underline{A})$ where $\underline{b}_1$ and $\underline{b}_2$ have minimum polynomials $(x+3)^2$ and $x^2 + 2x + 2$ respectively. By theorem 3.2 the system is controllable. Upon verification,

$$\underline{K} = [\underline{B}, \underline{A}\underline{B}, \underline{A}^2\underline{B}, \underline{A}^3\underline{B}]$$

$$= \left[\begin{array}{cc|cc|cc|cc} 1 & 0 & -2 & 0 & 3 & 0 & 0 & 0 \\ 1 & 0 & -3 & 0 & 9 & 0 & -27 & 0 \\ 0 & 1 & 0 & 1 & 0 & -4 & 0 & 6 \\ 0 & 1 & 0 & -4 & 0 & 6 & 0 & -4 \end{array}\right]$$

Using a sequence of column operations, $\underline{K}$ reduces to :

$$\underline{K}' = \left[\begin{array}{cc|cc|c} 1 & -2 & 0 & 0 & \\ 1 & -3 & 0 & 0 & \underline{0} \\ 0 & 0 & 1 & 1 & \\ 0 & 0 & 1 & -4 & \end{array}\right] = [\underline{b}_1 \ \underline{A}\underline{b}_1 \mid \underline{b}_2 \ \underline{A}\underline{b}_2 \mid \underline{0}]$$

Since $\underline{K}$ has rank 4, the system is controllable.

Theorem 3.3: Let S be a linear time-invariant system with $\underline{B} = [\underline{b}_1 \underline{b}_2 \ldots \underline{b}_s]$ where $s \geq 1$ and $\underline{A}$ a non-derogatory matrix representation of $\beta : \mathcal{X} \to \mathcal{X}$ whose minimum polynomial is $m(x) = m_1(x)^{e_1} m_2(x)^{e_2} \ldots m_r(x)^{e_r}$. The system S is controllable if there exists <u>at least one</u> element $\underline{b}_j \in \mathcal{X}$ for $1 \leq j \leq s$ such that $\mathcal{X} = Z(\underline{b}_j, \underline{A})$.

Proof: By definition of a minimum polynomial there exists at least one element $\underline{b}_j \in \mathcal{X}$ whose minimum polynomial is $m(x)$. Hence by definition of an $\underline{A}$-cyclic subspace generated by $\underline{b}_j$, the set $\{\underline{b}_j, \underline{A}\,\underline{b}_j, \ldots, \underline{A}^{n-1}\underline{b}_j\}$ is a basis for $\mathcal{X}$ such that $\mathcal{X} = Z(\underline{b}_j, \underline{A})$. As in Theorem 3.2 if $\underline{b}_j$ is a column of $\underline{B}$ then the system S is controllable.

In other words, the system S is controllable if $\underline{B}$ contains at least one column whose minimum polynomial is $m(x)$.

Consider the following examples:

Example 3.5: Consider the following linear sequential machine with binary logic.

$$\underline{x}'(t) = \underline{A}\underline{x}(t) + \underline{B}\mu(t) = \begin{bmatrix} 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{bmatrix} \underline{x}(t) + \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \mu(t)$$

$p(x) = m(x) = (x+1)^2\,(x^2+x+1)$ and $\mathcal{X} = Z(\underline{b}, \underline{A})$ where the minimum polynomial of $\underline{b}$ is $p(x)$.

$$\text{The matrix } \underline{K} = [\underline{b},\ \underline{A}\underline{b},\ \underline{A}^2\underline{b},\ \underline{A}^3\underline{b}] = \begin{bmatrix} 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \end{bmatrix}$$

obviously has rank 4. Hence the system is controllable.

Example 3.6: Consider the following linear sequential machine with binary logic.

$$\underline{x}'(t) = \underline{A}\underline{x}(t) + \underline{B}\,\underline{\mu}(t) = \begin{bmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{bmatrix} \underline{x}(t) + \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \underline{\mu}(t)$$

$p(x) = m(x) = (x+1)^4$ and $\mathcal{X} = Z(\underline{b}, \underline{A})$ where the minimum polynomial of $\underline{b}$ is $p(x)$.

The matrix $\underline{K} = [\underline{b}, \underline{A}\,\underline{b}, \underline{A}^2\,\underline{b}, \underline{A}^3\,\underline{b}] = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

obviously has rank 4. Hence the system is controllable.

Example 3.7: Consider the following multi-variable control system defined over R.

$$\dot{\underline{x}}(t) = \underline{A}\,\underline{x}(t) + \underline{B}\,\underline{\mu}(t) = \begin{bmatrix} -3 & 1 & 0 & 0 \\ 0 & -3 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -2 & -2 \end{bmatrix} \underline{x}(t) + \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \underline{\mu}(t)$$

$p(x) = m(x) = (x+3)^2 (x+3)^2 (x^2 + 2x + 2)$ and $\mathcal{X} = Z(\underline{b}, \underline{A})$, where the minimum polynomial of $\underline{b}$ is $p(x)$.

The matrix

$$\underline{K} = [\underline{b}\,; \underline{A}\underline{b}, \underline{A}^2\underline{b}, \underline{A}^3\underline{b} \mid \underline{0}] = \left[\begin{array}{cccc|c} 1 & -2 & 3 & 0 & \\ 1 & -3 & 9 & -27 & \\ 1 & 1 & -4 & 6 & \underline{0} \\ 1 & -4 & 6 & -4 & \end{array}\right]$$ obviously has rank 4.

Hence the system is controllable.

Theorem 3.4: If the minimum polynomial of the non-derogatory transition matrix $\underline{A}$ is a monic, irreducible polynomial over F then <u>any</u> single non trivial input ($\underline{B} \neq \underline{0}$) ensures that the system is controllable.

Proof:

Let $m(x) = x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0$ be the minimum polynomial of $\underline{A}$. Since $m(x)$ is irreducible, then for all $\underline{b} \in \mathcal{X}$; $\{\underline{b}, \underline{A}\,b, \dots, \underline{A}^{n-1}\,\underline{b}\}$

is a linearly independent set such that $\mathcal{X} = Z(\underline{b}, \underline{A})$. By theorem 2.3 it follows that the system is controllable, for all $\underline{b} \in \mathcal{X}$ such that $\underline{B} = [\underline{b}]$.

If F is a finite field of q elements then there exist $(q^n - 1)$ such elements. If F is an infinite field then there are infinitely many such elements.

<u>Corollary 3.4</u>: A <u>necessary</u> condition for a single input system to be controllable is that the state transition matrix $\underline{A}$ be non-derogatory.

Consider the following examples:

<u>Example 3.8</u>: Consider the following linear sequential machine with binary logic.

$$\underline{x}'(t) = \underline{A}\underline{x}(t) + \underline{B}\,\underline{\mu}(t) = \begin{bmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 \end{bmatrix} \underline{x}(t) + \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix} \underline{\mu}(t)$$

$p(x) = m(x) = x^4 + x^3 + x^2 + x + 1$ and $\mathcal{X} = Z(\underline{b}, \underline{A})$ where the minimum polynomial of $\underline{b}$ is $p(x)$. The matrix

$$\underline{K} = [\,\underline{b}, \underline{A}\underline{b}, \underline{A}^2\underline{b}, \underline{A}^3\underline{b}\,] = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix}$$ obviously has rank 4.

Hence the system is controllable.

<u>Example 3.9:</u> Consider the following multivariable control system defined over R.

$$\dot{\underline{x}}(t) = \underline{A}\,\underline{x}(t) + \underline{B}\,\underline{\mu}(t) = \begin{bmatrix} 0 & 1 \\ -7 & -4 \end{bmatrix} \underline{x}(t) + \begin{bmatrix} 1 \\ 1 \end{bmatrix} \underline{\mu}(t)$$

$p(x) = m(x) = x^2 + 4x + 7$ and $\mathcal{X} = Z(\underline{b}, \underline{A})$ where the minimum polynomial of $\underline{b}$ is $p(x)$.

The matrix $\underline{K} = [\,\underline{b}, \underline{A}\,\underline{b}\,] = \begin{bmatrix} 1 & 1 \\ 1 & -11 \end{bmatrix}$ obviously has rank 2.

Hence the system is controllable.

The above example is rather simple since the state space $\mathcal{X}$ has only dimension 2. However, a polynomial of degree n over R has n roots in the field C of complex numbers and complex roots occur as complex conjugate pairs. From this it follows that any irreducible polynomical over R has degree 1 or 2. This explains why such a simple example was chosen.

Theorem 3.5: Let S be a linear, time-invariant system where $\underline{A}$ is the matrix representation of $\beta : \mathcal{X} \to \mathcal{X}$ whose minimum polynomial is $m(x) = m_1^e(x)\ m_2(x) \ldots\ m_r(x)$ such that $m_i(x)$ for $1 \leq i \leq r$ are monic and irreducible polynomials over F. The elementary divisors, divisible by $m_1(x)$, are $m_1(x)^{e_1}$, $m_1(x)^{e_2}, \ldots, m_1(x)^{e_\ell}$ such that $e = e_1 \geq e_2 \geq \ldots \geq e_\ell$ where "ℓ" is the number of elementary divisors divisible by $m_1(x)$. A necessary condition for the system to be controllable is the existence of at least "ℓ" inputs.

Proof:

The proof is by construction. Since $\underline{A}$ has $(\ell + r - 1)$ elementary divisors then $\mathcal{X}$ may be expressed as the direct-sum of $(\ell + r - 1)$ $\underline{A}$-cyclic subspaces as follows:

$$\mathcal{X} = Z(\underline{b}_{11}, \underline{A}) \otimes Z(\underline{b}_{12}, \underline{A}) \otimes \ldots \otimes Z(\underline{b}_{1\ell}, \underline{A}) \otimes Z(\underline{b}_2, \underline{A}) \otimes \ldots \otimes Z(\underline{b}_r, \underline{A})$$

where the minimum polynomial of $Z(\underline{b}_{1i}, \underline{A})$ is $m_1(x)^{e_i}$ for $1 \leq i \leq \ell$ and $Z(\underline{b}_j, \underline{A})$ is $m_j(x)$ for $2 \leq j \leq r$. Every generator of the above $\underline{A}$-cyclic subspaces is an element of the null space of the elementary divisor which is the minimum polynomial of that $\underline{A}$-cyclic subspace. Since the null spaces of the elementary divisors are not necessarily disjoint then every generator must be selected such that the set of generators forms a linearly independent set.

In order to express $\mathcal{X}$ as a direct-sum of fewer subspaces, the $\underline{A}$-cyclic subspaces must be combined in some legitimate manner. Let us combine $Z(\underline{b}_i, \underline{A})$ and $Z(\underline{b}_j, \underline{A})$ whose minimum polynomial is

$m_i(x)^{e_i}$ and $m_j(x)^{e_j}$ respectively for $i \neq j$, $1 \leq i,j \leq r$ and $e_i, e_j \geq 1$. Let $\underline{b}_* = \underline{b}_i + \underline{b}_j$. Since $m_i(\underline{A})^{e_i} m_j(\underline{A})^{e_j} \underline{b}_* = m_i(\underline{A})^{e_i} m_j(\underline{A})^{e_j} (\underline{b}_i + \underline{b}_j) = \underline{0}$ then the minimum polynomial of $\underline{b}_*$ and $Z(\underline{b}_*, \underline{A})$ is $m_i(x)^{e_i} m_j(x)^{e_j}$ where $m_i(x)^{e_i}$ and $m_j(x)^{e_j}$ are relatively prime. Therefore by lemma 2.4

$Z(\underline{b}_*, \underline{A}) = Z(\underline{b}_i, \underline{A}) \otimes Z(\underline{b}_j, \underline{A})$. Let us combine $Z(\underline{b}_{1i}, \underline{A})$ and $Z(\underline{b}_{1j}, \underline{A})$ whose minimum polynomial is $m_1(x)^{e_i}$ and $m_1(x)^{e_j}$ respectively for $i \neq j$, $1 \leq i,j \leq \ell$ and $e_i \geq e_j$. Let $\underline{b}_* = \underline{b}_{1i} + \underline{b}_{1j}$. Since $m_1(x)^{e_j} \mid m_1(x)^{e_i}$ then the minimum polynomial of $\underline{b}_*$ and $Z(\underline{b}_*, \underline{A})$ is $m_1(x)^{e_i}$ and $\dim Z(\underline{b}_*, \underline{A}) = \dim Z(\underline{b}_{1i}, \underline{A}) < \{\dim Z(\underline{b}_{1i}, \underline{A}) + \dim Z(\underline{b}_{1j}, \underline{A})\}$.
Therefore $Z(\underline{b}_*, \underline{A}) \neq Z(\underline{b}_{1i}, \underline{A}) \otimes Z(\underline{b}_{1j}, \underline{A})$. Clearly any $\underline{A}$-cyclic subspaces whose minimum polynomials are relatively prime may be combined. Since the minimum polynomial of $Z(\underline{b}_{1i}, \underline{A})$ for $1 \leq i \leq \ell$ are <u>not</u> relatively prime then these subspaces <u>cannot</u> be combined. Thus the minimal direct-sum representation of $\mathcal{X}$ contains "ℓ" $\underline{A}$-cyclic subspaces.

There are many different ways of combining the $\underline{A}$-cyclic subspaces in order to obtain a minimal direct-sum representation for $\mathcal{X}$. Without any loss of generality let us combine the $\underline{A}$-cyclic subspaces as follows:

$$\text{Let } Z(\underline{b}_{11} + \underline{b}_2 + \ldots + \underline{b}_r, \underline{A}) = Z(\underline{b}_{11}, \underline{A}) \otimes Z(\underline{b}_2, \underline{A}) \otimes \ldots \otimes Z(\underline{b}_r, \underline{A}),$$

$$\underline{b}_{j1} = \underline{b}_{11} + \underline{b}_2 + \ldots + \underline{b}_r \quad \text{and} \quad \underline{b}_{ji} = \underline{b}_{1i} \quad \text{for } 2 \leq i \leq \ell .$$

Hence $\mathcal{X}$ may be expressed as the direct-sum of "ℓ" $\underline{A}$-cyclic subspaces as follows:

$$\mathcal{X} = Z(\underline{b}_{j1}, \underline{A}) \otimes Z(\underline{b}_{j2}, \underline{A}) \otimes \ldots \otimes Z(\underline{b}_{j\ell}, \underline{A})$$

As in Theorem 3.2 if $\underline{B} = [\underline{b}_{j1}, \underline{b}_{j2}, \ldots \underline{b}_{j\ell}]$ then the system is controllable with "ℓ" inputs. Since this direct-sum representation for $\mathcal{X}$ is minimal then the existence of at <u>least</u> "ℓ" inputs is a necessary requirement for the system S to be controllable.

Consider the following examples :

Example 3.10: Consider the multi-variable control system defined in Example 3.1. $m(x) = (x+1)^2(x^2+2x+2)$ and $p(x) = m(x)(x+1)$. Thus $\mathcal{X} = Z(\underline{b}_1, \underline{A}) \otimes Z(\underline{b}_2, \underline{A}) \otimes Z(\underline{b}_3, \underline{A})$ where the minimum polynomial of $\underline{b}_1$, $\underline{b}_2$, and $\underline{b}_3$ is $(x+1)^2$, $x+1$ and x^2+2x+2 respectively. As in theorem 3.5 $\mathcal{X} = Z(\underline{b}_1 + \underline{b}_3, \underline{A}) \otimes Z(\underline{b}_2, \underline{A})$ where $Z(\underline{b}_1 + \underline{b}_3, \underline{A}) = Z(\underline{b}_1, \underline{A}) \otimes Z(\underline{b}_3, \underline{A}) = Z(\underline{b}_{13}, \underline{A})$

$$\text{Let } \underline{B} = [\,\underline{b}_1 + \underline{b}_3,\ \underline{b}_2\,] = \begin{bmatrix} 3 & 0 \\ 3 & 0 \\ 1 & 1 \\ 0 & 0 \\ 3 & 1 \end{bmatrix}$$

Upon verification,

$$\underline{K} = [\,\underline{B},\ \underline{A}\,\underline{B},\ \underline{A}^2\,\underline{B},\ \underline{A}^3\,\underline{B},\ \underline{A}^4\,\underline{B}\,]$$

$$= \begin{array}{c} \left[\begin{array}{cc|cc|cc|cc|cc} 3 & 0 & -8 & 0 & 10 & 0 & -6 & 0 & -4 & 0 \\ 3 & 0 & -3 & 0 & 0 & 0 & 6 & 0 & -12 & 0 \\ 1 & 1 & -4 & -1 & 4 & 1 & 2 & -1 & -14 & 1 \\ 0 & 0 & 3 & 0 & -6 & 0 & 6 & 0 & 0 & 0 \\ 3 & 1 & -3 & -1 & 0 & 1 & 6 & -1 & -12 & 1 \end{array}\right] \\ \underline{B} \qquad\quad \underline{A}\,\underline{B} \qquad\quad \underline{A}^2\underline{B} \qquad\quad \underline{A}^3\underline{B} \qquad\quad \underline{A}^4\underline{B} \end{array}$$

Using a sequence of elementary column operations, $\underline{K}$ reduces to:

$$\underline{K}' = \left[\begin{array}{cccc|c|c} 3 & -8 & 10 & -6 & 0 & \\ 3 & -3 & 0 & 6 & 0 & \\ 1 & -4 & 4 & 2 & 1 & \underline{0} \\ 0 & 3 & -6 & 6 & 0 & \\ 3 & -3 & 0 & 6 & 1 & \end{array}\right] = [\underline{b}_{13},\ \underline{A}\underline{b}_{13},\ \underline{A}^2\underline{b}_{13},\ \underline{A}^3\underline{b}_{13};\underline{b}_2 \mid \underline{0}]$$

Since $\underline{K}$ has rank 5, the system is controllable.

Example 3.11: Consider the linear sequential machine defined in Example 3.2. $m(x) = (x+2)^2 \cdot (x+1)$ and $p(x) = m(x)(x+2)$. Thus $\mathcal{X} = Z(\underline{b}_1, \underline{A}) \otimes Z(\underline{b}_2, \underline{A}) \otimes Z(\underline{b}_3, \underline{A})$ where the minimum polynomial of $\underline{b}_1$, $\underline{b}_2$ and $\underline{b}_3$ is $(x+2)^2$, $(x+2)$ and $(x+1)$ respectively. As in theorem 3.5, $\mathcal{X} = Z(\underline{b}_1 + \underline{b}_3, \underline{A}) \otimes Z(\underline{b}_2, \underline{A})$ where

$$Z(\underline{b}_1 + \underline{b}_3, \underline{A}) = Z(\underline{b}_1, \underline{A}) \otimes Z(\underline{b}_3, \underline{A}) = Z(\underline{b}_{13}, \underline{A}).$$

Let $\underline{B} = [\underline{b}_1 + \underline{b}_3, \underline{b}_2] = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$

Upon verification,

$$\underline{K} = [\underline{B}, \underline{A}\,\underline{B}, \underline{A}^2\underline{B}, \underline{A}^3\underline{B}]$$

$$= \begin{array}{c} \left[\begin{array}{cc|cc|cc|cc} 1 & 0 & 2 & 0 & 2 & 0 & 3 & 0 \\ 1 & 0 & 0 & 0 & 2 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 2 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 2 & 0 & 0 & 0 \end{array}\right] \\ \quad \underline{B} \qquad \underline{A}\underline{B} \qquad \underline{A}^2\underline{B} \qquad \underline{A}^3\underline{B} \end{array}$$

Using a sequence of column operations, $\underline{K}$ reduces to :

$$\underline{K}' = \left[\begin{array}{ccc|c|c} 1 & 2 & 2 & 0 & \\ 1 & 0 & 2 & 0 & \\ 0 & 1 & 2 & 1 & \underline{0} \\ 0 & 1 & 2 & 0 & \end{array}\right] = [\underline{b}_{13}, \underline{A}\,\underline{b}_{13}, \underline{A}^2\underline{b}_{13} \mid \underline{b}_2 \mid \underline{0}]$$

Since $\underline{K}$ has rank 4, the system is controllable.

Theorem 3.6: Let S be a linear, time-invariant system where $\underline{A}$ is the matrix representation of $\beta : \mathcal{X} \to \mathcal{X}$ whose minimum polynomial is $m(x) = m_1(x)^{e_1} m_2(x)^{e_2} \ldots m_r(x)^{e_r}$ such that $m_i(x)$ for $1 \le i \le r$ are monic and irreducible polynomials over F. The elementary divisors, divisible by $m_i(x)$ for $1 \le i \le r$, are $m_i(x)^{e_{i1}}, m_i(x)^{e_{i2}}, \ldots, m_i(x)^{e_{i\ell_i}}$ where $e_i = e_{i1} \ge e_{i2} \ge \ldots \ge e_{i\ell_i}$ and "ℓ_i" is the number of elementary divisors divisible by $m_i(x)$ such that $\ell_1 \ge \ell_2 \ge \ldots \ge \ell_r$.

A necessary condition for the system to be controllable is the existence of at least "ℓ_1" inputs.

Proof: This theorem is a generalization of theorem 3.5. The proof is by construction. Since $\underline{A}$ has $\sum_{i=1}^{r} \ell_i$ elementary divisors then $\mathcal{X}$ may be expressed as the direct-sum of as many $\underline{A}$-cyclic subspaces as follows:

$$\mathcal{X} = Z(\underline{b}_{11},\underline{A}) \otimes \ldots \otimes Z(\underline{b}_{1\ell_1},\underline{A}) \otimes Z(\underline{b}_{21},\underline{A}) \otimes \ldots \otimes Z(\underline{b}_{r1},\underline{A}) \otimes \ldots \otimes Z(\underline{b}_{r\ell_r},\underline{A})$$

where the minimum polynomial of $Z(\underline{b}_{ij},\underline{A})$ is $m_i(x)^{e_{ij}}$ for $1 \le i \le r$ and $1 \le j \le \ell_i$. Every generator of the above $\underline{A}$-cyclic subspaces is an element of the null space of the elementary divisor which is the minimum polynomial of that $\underline{A}$-cyclic subspace. Since the null spaces of the elementary divisors are not necessarily disjoint then every generator must be selected such that the resulting set of generators forms a linearly independent set.

In order to express $\mathcal{X}$ as the direct sum of fewer subspaces, the $\underline{A}$-cyclic subspaces must be combined in some legitimate manner.

In view of Theorem 3.5, any $\underline{A}$-cyclic subspaces whose minimum polynomials are relatively prime may be combined. For any given "i" $1 \le i \le r$ the minimum polynomials of $Z(\underline{b}_{ij}, \underline{A})$ for $1 \le j \le \ell_i$ are not relatively prime. Therefore the $\underline{A}$-cyclic subspaces in $Z(\underline{b}_{i1}, \underline{A})$, $Z(\underline{b}_{i2},\underline{A})$, ..., $Z(\underline{b}_{i\ell_i}, \underline{A})$ cannot be combined. Hence the minimal direct-sum representation of $\mathcal{X}$ contains "ℓ_1" $\underline{A}$-cyclic subspaces where ℓ_1 is the largest number of $\underline{A}$-cyclic subspaces which cannot be combined.

There are many different ways of combining the $\underline{A}$-cyclic subspaces in order to obtain a minimal direct-sum representation for $\mathcal{X}$. Without any loss of generality, let us combine the $\underline{A}$-cyclic subspaces as follows:

Let $Z(\underline{b}_{11} + \underline{b}_{21} + \ldots + \underline{b}_{r1}, \underline{A}) = Z(\underline{b}_{11}, \underline{A}) \otimes Z(\underline{b}_{21}, A) \otimes \ldots \otimes Z(\underline{b}_{r1}, \underline{A})$

$$Z(\underline{b}_{12} + \underline{b}_{22} + \ldots + \underline{b}_{r2}, \underline{A}) = Z(\underline{b}_{12}, \underline{A}) \otimes Z(\underline{b}_{22}, \underline{A}) \otimes \ldots \otimes Z(\underline{b}_{r2}, \underline{A})$$

$$\vdots$$

$$Z(\underline{b}_{1\ell_1} + \quad \ldots \quad , \underline{A}) = Z(\underline{b}_{1\ell_1}, \underline{A}) \otimes \ldots$$

When the magnitude of $\ell_1, \ell_2, \ldots, \ell_r$ is known, the above direct-sums of the $\underline{A}$-cyclic subspaces are precisely defined.

Let $\underline{b}_{j1} = \underline{b}_{11} + \underline{b}_{21} + \ldots + \underline{b}_{r1}$

$$\underline{b}_{j2} = \underline{b}_{12} + \underline{b}_{22} + \ldots + \underline{b}_{r2}$$

$$\vdots$$

$$\underline{b}_{j\ell_1} = \underline{b}_{1\ell_1} + \ldots$$

Hence $\mathcal{X}$ may be expressed as the direct-sum of "ℓ_1" $\underline{A}$-cyclic subspaces asf follows:

$$\mathcal{X} = Z(\underline{b}_{j1}, \underline{A}) \otimes Z(\underline{b}_{j2}, A) \otimes \ldots \otimes Z(\underline{b}_{j\ell_1}, \underline{A})$$

As in Theorem 3.2 if $\underline{B} = [\,\underline{b}_{j1}, \underline{b}_{j2}, \ldots, \underline{b}_{j\ell_1}\,]$ then the system is controllable with "ℓ_1" inputs. Since this direct-sum representation for $\mathcal{X}$ is minimal then the existence of at <u>least</u> "ℓ_1" inputs is a necessary requirement for the system S to be controllable.

<u>Example 3.12</u>: Consider the following multi-variable control system defined over R.

$$\dot{\underline{x}}(t) = \underline{A}\,\underline{x}(t) + \underline{B}\,\underline{\mu}(t) = \begin{bmatrix} 1 & 2 & 1 & 0 & -2 \\ 0 & 2 & 0 & -2 & 0 \\ 0 & 2 & 1 & -1 & -1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & -2 & 0 \end{bmatrix} \underline{x}(t) + \begin{bmatrix} 0 & -1 & 3 & 2 \\ 0 & 1 & 0 & 2 \\ 2 & -1 & 3 & 2 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 3 & 2 \end{bmatrix} \underline{\mu}(t)$$

$m(x) = (x-1)^2\, x(x-2)$ and $p(x) = m(x)\,x$.

$\mathcal{X} = Z(\underline{b}_1, \underline{A}) \otimes Z(\underline{b}_2, \underline{A}) \otimes Z(\underline{b}_3, \underline{A}) \otimes Z(\underline{b}_4, \underline{A})$ where the minimum polynomial of $\underline{b}_1$, $\underline{b}_2$, $\underline{b}_3$ and $\underline{b}_4$ is $(x-1)^2$, x, x and $(x-2)$ respectively.

By theorem 3.2 the system is controllable. As in theorem 3.6,

$$\mathcal{X} = Z(\underline{b}_1 + \underline{b}_2 + \underline{b}_4, \underline{A}) \otimes Z(\underline{b}_3, \underline{A}) = Z(\underline{b}_{124}, \underline{A}) \otimes Z(\underline{b}_3, \underline{A}).$$

Let $B = [\underline{b}_{124}, \underline{b}_3] = \begin{bmatrix} 1 & 3 \\ 3 & 0 \\ 3 & 3 \\ 1 & 0 \\ 2 & 3 \end{bmatrix}$

Upon verification,

$$K = [\underline{B}, \underline{A}\,\underline{B}, \underline{A}^2\underline{B}, \underline{A}^3\underline{B}, \underline{A}^4\,\underline{B}]$$

$$= \begin{array}{c} \left[\begin{array}{cc|cc|cc|cc|cc} 1 & 3 & 6 & 0 & 12 & 0 & 22 & 0 & 40 & 0 \\ 3 & 0 & 4 & 0 & 8 & 0 & 16 & 0 & 32 & 0 \\ 3 & 3 & 6 & 0 & 10 & 0 & 18 & 0 & 34 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 2 & 3 & 4 & 0 & 8 & 0 & 16 & 0 & 32 & 0 \end{array}\right] \\ \begin{array}{ccccc} \underline{B} & \quad \underline{AB} & \quad \underline{A}^2\underline{B} & \quad \underline{A}^3\underline{B} & \quad \underline{A}^4\underline{B} \end{array} \end{array}$$

Using a sequence of elementary column operations, $\underline{K}$ reduces to:

$$\underline{K}' = \left[\begin{array}{cccc|c|c} 1 & 6 & 12 & 22 & 3 & \\ 3 & 4 & 8 & 16 & 0 & \\ 3 & 6 & 10 & 18 & 3 & \underline{0} \\ 1 & 0 & 0 & 0 & 0 & \\ 2 & 4 & 8 & 16 & 3 & \end{array}\right] = [\underline{b}_{124}, \underline{A}\underline{b}_{124}, \underline{A}^2\underline{b}_{124}, \underline{A}^3\underline{b}_{124} \mid \underline{b}_3 \mid \underline{0}]$$

Since $\underline{K}$ has rank 5, the system is controllable.

<u>Example 3. 13:</u> Consider the following linear sequential machine defined over GF(3).

$$\underline{x}'(t) = \underline{A}\,\underline{x}(t) + \underline{B}\,\underline{\mu}(t) = \begin{bmatrix} 1 & 0 & 0 & 2 & 0 \\ 0 & 2 & 1 & 0 & 0 \\ 0 & 0 & 2 & 1 & 0 \\ 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \underline{x}(t) + \begin{bmatrix} 2 & 2 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \underline{\mu}\,(t)$$

$$m(x) = (x+1)^3\,(x+2) \quad \text{and} \quad p(x) = m(x)\,(x+2).$$

$\mathcal{X} = Z(\underline{b}_1, \underline{A}) \otimes Z(\underline{b}_2, \underline{A}) \otimes Z(\underline{b}_3, \underline{A})$ where the minimum polynomial of $\underline{b}_1, \underline{b}_2$ and $\underline{b}_3$ is $(x+1)^3$, $(x+2)$ and $(x+2)$ respectively. By theorem 3. 2 the system is controllable. As in theorem 3. 6

$$\mathcal{X} = Z(\underline{b}_1 + \underline{b}_2,\ \underline{A}) \otimes Z(\underline{b}_3, \underline{A}) = Z(\underline{b}_{12},\ \underline{A}) \otimes Z(\underline{b}_3, \underline{A})$$

Let $\underline{B} = [\underline{b}_{12},\ \underline{b}_3\,] = \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$. Upon verification,

$$\underline{K} = [\underline{B},\ \underline{A}\,\underline{B},\ \underline{A}^2\underline{B},\ \underline{A}^3\underline{B},\ \underline{A}^4\underline{B}\ ]$$

$$= \begin{array}{c} \left[\begin{array}{cc|cc|cc|cc|cc} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 2 & 0 \\ 1 & 0 & 2 & 0 & 1 & 0 & 2 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{array}\right] \\ \begin{array}{ccccc} \underline{B} \quad & \quad \underline{A}\underline{B} \quad & \quad \underline{A}^2\underline{B} \quad & \quad \underline{A}^3\underline{B} \quad & \quad \underline{A}^4\underline{B} \end{array} \end{array}$$

Using a sequence of elementary column operations, $\underline{K}$ reduces to :

$$\underline{K}' = \left[\begin{array}{cccc|c|c} 1 & 0 & 1 & 0 & 0 & \\ 0 & 0 & 1 & 0 & 0 & \\ 0 & 1 & 1 & 0 & 0 & \underline{0} \\ 1 & 2 & 1 & 2 & 0 & \\ 0 & 0 & 0 & 0 & 1 & \end{array}\right] = [\underline{b}_{12},\ \underline{A}\underline{b}_{12},\ \underline{A}^2\underline{b}_{12},\ \underline{A}^3\underline{b}_{12} \mid\ \underline{b}_3 \mid \underline{0}\,]$$

Since $\underline{K}$ has rank 5, the system is controllable.

If this system had less than 2 inputs, it would be uncontrollable! See the corollary to theorem 3.6.

III-3 OBSERVABILITY :

Definition 3.7 : A linear, time-invariant system S is observable if for some $t > o$ and all initial states $\underline{x}_o \in \mathcal{X}$, a knowledge of $\underline{A}$ and $\underline{C}$ and the zero input response $\underline{y}(t) = \underline{C}\,\underline{x}(t)$ is sufficient to determine the initial state $\underline{x}_o$.

Theorem 3.7 [8]. A linear, time-invariant system S is observable if and only if the matrix,

$$\underline{L} = [\underline{C}^*, \ \underline{A}^* \underline{C}^*, \dots, \ \underline{A}^{*(n-1)} \underline{C}^*]$$

where * denotes the complex conjugate of the transpose of the matrix, has rank n.

Note the similarity between theorems 3.1 and 3.7. In considering observability, the transpose of the rows of $\underline{C}$ will be treated as elements of $\mathcal{X}$. For controllability, the columns of $\underline{B}$ were dealt with whereas for observability we deal with the columns of $\underline{C}^*$. When determining what elements in $\mathcal{X}$, if used as columns of $\underline{C}^*$, will yield an observable system; the transpose of $\underline{A}$ must be used as indicated clearly in theorem 3.7.

Since we are dealing essentially with finite and infinite fields in which the elements are real numbers the complex conjugate of the transpose of the matrices $\underline{A}$ and $\underline{C}$ is $\underline{A}^T$ and $\underline{C}^T$ respectively.

With these differences well in mind, remembering we are interested primarily in synthesis, the following sufficient conditions for observability hold. The proofs are omitted in each case since it is analogous to that of the corresponding theorem for the controllability of such systems.

Theorem 3.8: Let S be a linear, time-invariant system with $\underline{C}^T = [\underline{c}_1^T, c_2^T, \dots, \underline{c}_s^T]$ where $s \geq r$ and $\underline{A}$

the matrix representation of $\beta : \mathcal{X} \longrightarrow \mathcal{X}$ whose elementary divisors are $\ell_1(x), \ell_2(x), \ldots, \ell_r(x)$. The system is observable if there exists r elements $\underline{c}_j^T \in \mathcal{X}$ for $1 \le j \le s$ such that:

$$\mathcal{X} = Z(\underline{c}_{j1}^T, \underline{A}^T) \oplus Z(\underline{c}_{j2}^T, \underline{A}^T) \oplus \ldots \oplus Z(\underline{c}_{jr}^T, \underline{A}^T)$$

In terms of cyclic subspaces, the above theorem may be restated as follows: the system S is observable if the r cyclic subspaces are represented respectively by r <u>independent</u> vectors acting as rows of $\underline{C}$.

<u>Theorem 3.9:</u> Let S be a linear, time invariant system with $\underline{C}^T = [\underline{c}_1^T, \underline{c}_2^T, \ldots, \underline{c}_s^T]$ where $s \ge 1$ and the matrix $\underline{A}$ a non-derogatory matrix representation of $\beta : \mathcal{X} \longrightarrow \mathcal{X}$ whose minimum polynomial is $m(x) = m_1(x)^{e_1} m_2(x)^{e_2} \ldots m_r(x)^{e_r}$. The system S is observable if there exists <u>at least one</u> element $\underline{c}_j^T \in \mathcal{X}$ for $1 \le j \le s$ such that: $\mathcal{X} = Z(\underline{c}_j^T, \underline{A}^T)$.

In other words, the system S is observable if $\underline{C}$ contains <u>at least one</u> column whose minimum polynomial is $m(x)$.

<u>Theorem 3.10</u>: If the minimum polynomial of the non-derogatory transition matrix $\underline{A}$ is a monic irreducible polynomial over F then <u>any</u> single non-trivial output ($\underline{C} \neq \underline{0}$) ensures the system is observable.

<u>Corollary 3.10:</u> A necessary condition for a single output system to be observable is that the state transition matrix $\underline{A}$ be non-derogatory.

<u>Theorem 3.11:</u> Let S be a linear, time invariant system where $\underline{A}$ is the matrix representation of $\beta : \mathcal{X} \longrightarrow \mathcal{X}$ whose minimum polynomial is $m(x) = m_1(x)^e m_2(x) \ldots m_r(x)$ such that $m_i(x)$ for $1 \le i \le r$ are monic and irreducible polynomials over F. The elementary divisors, divisible by $m_1(x)$, are $m_1(x)^{e_1}, m_1(x)^{e_2}, \ldots, m_1(x)^{e_\ell}$

such that $e = e_1 \geq e_2 \geq \dots e_\ell$ where "ℓ" is the number of elementary divisors divisible by $m_1(x)$. A necessary condition for the system to be observable is the existence of at least "ℓ" outputs.

<u>Theorem 3.12:</u> Let S be a linear, time-invariant system where $\underline{A}$ is the matrix representation of $\beta : \mathcal{X} \rightarrow \mathcal{X}$ whose minimum polynomial is $m(x) = m_1(x)^{e_1} m_2(x)^{e_2} \dots m_r(x)^{e_r}$ such that $m_i(x)$ for $1 \leq i \leq r$ are monic and irreducible polynomials over F. The elementary divisors, divisible by $m_i(x)$ for $1 \leq i \leq r$, are $m_i(x)^{e_{i1}}, m_i(x)^{e_{i2}}, \dots, m_i(x)^{e_{i\ell_i}}$ where $e_i = e_{i1} \geq e_{i2} \geq \dots \geq e_{i\ell_i}$, and "$\ell_i$" is the number of elementary divisors divisible by $m_i(x)$ such that $\ell_1 \geq \ell_2 \geq \dots \geq \ell_r$. A necessary condition for the system to be observable is the existence of at <u>least</u> "ℓ_1" outputs.

No examples are included since these would be essentially the same as for controllability with the slight difference that rows of $\underline{C}$ are manipulated rather than columns of $\underline{B}$ and $\underline{A}^T$-cyclic subspaces are formed rather than $\underline{A}$-cyclic subspaces.

III-4 CONCLUSION

This chapter introduced and developed a method for the study of controllability and observability of general real linear time-invariant systems. In contrast with existing methods where these concepts are developed by considering the ranks of large combined matrices, see theorems 3.1 and 3.7, this method is based on the characteristics of cyclic spaces.

The method presented in this chapter has the advantage that it provides a physical insight into these concepts, since the elements of the system matrices are considered separately. This eliminates the need to combine the system matrices.

The approach taken was to subdivide the state space into a number of cyclic subspaces. This involves a thorough investigation of the state transition matrix $\underline{A}$. Once this is obtained the concept of controllability reduces to the verification of whether or not columns of the state input matrix $\underline{B}$ are generators of these subspaces. Similarly, observability reduces to the verification of whether or not the rows of the state output matrix $\underline{C}$ are generators of these subspaces.

Theorems 3.2 and 3.8 yield the general results for controllability and observability, respectively. More detailed results were obtained for non-derogatory systems. For controllability these results were presented in theorems: 3.8, 3.4 and for observability in theorems 3.9, 3.10. Finally the minimality requirements were presented in theorems 3.5, 3.6, 3.11, and 3.12. It should be noted that these theorems are independent of the form of the systems state matrix.

Illustrative examples to clarify these ideas where presented for both continuous and discrete systems.

CHAPTER IV

SHIFT REGISTER SYNTHESIS OF LINEAR SEQUENTIAL CIRCUITS

This chapter introduces an algorithm for shift register realization of linear sequential machines whenever such a realization is possible. The algorithm makes use of the fact that a shift register realization corresponds to a particular canonical form of the state equations. The algorithm derives the family of similarity transformation matrices $\{\underline{M}_i\}$ that transform the given state transition matrix $\underline{A}$ into the corresponding companion matrix form $\underline{A}_c$. Finally it chooses those $\underline{M}_i$ that satisfy the additional requirements of shift register realizability.

IV-1: SYSTEM ISOMORPHISM

Definition 4.1: A system $S = < \mathcal{U}, \mathcal{Y}, \mathcal{X}, \delta, \lambda >$ defined over GF(p), where p is a prime number, is minimal if $n \geq \log_p q$ where n is the smallest such integer and q is the number of non-equivalent states.

Definition 4.2: Two minimal systems S_1 and S_2 are isomorphic if and only if their state spaces are equidimensional and their input-output behaviours are identical.

System isomorphism implies a definite relationship between the state spaces of the individual systems. To obtain the mathematical characteristics of this relationship, consider two isomorphic sequential systems S_1 and S_2; where $S_i = < \mathcal{U}, \mathcal{Y}, \mathcal{X}_i, \delta_i, \lambda_i >$, the dimension of $\mathcal{X}_i$ is n for $i = 1, 2$ and $\underline{M}$ is an nxn matrix representing a mapping $\gamma : \mathcal{X}_1 \longrightarrow \mathcal{X}_2$ such that $\underline{x}_2(t) = \underline{M}\,\underline{x}_1(t)$. Let the behaviour of S_i be described by the following state equations:

$$\underline{x}_i(t+1) = \underline{A}_i\,\underline{x}_i(t) + \underline{B}_i\,\underline{\mu}(t)$$

$$\underline{y}_i(t) = \underline{C}_i\,\underline{x}_i(t) + \underline{D}_i\,\underline{\mu}(t)$$

Since S_1 and S_2 are assumed isomorphic, their input-output behaviour is to be identical.

Hence, $\underline{x}_2(t+1) = \underline{M}\,\underline{x}_1(t+1) = \underline{M}\,[\,\underline{A}_1\underline{x}_1(t) + \underline{B}_1\,\underline{\mu}(t)\,]$

$$= \underline{M}\,\underline{A}_1\underline{M}^{-1}\,\underline{x}_2(t) + \underline{M}\underline{B}_1\,\underline{\mu}(t)$$

$$= \underline{A}_2\,\underline{x}_2(t) + \underline{B}_2\,\underline{\mu}(t)$$

$$\underline{y}_2(t) = \underline{y}_1(t) = \underline{C}_1\,\underline{M}_1^{-1}\,\underline{x}_2(t) + \underline{D}_1\,\underline{\mu}(t)$$

$$= \underline{C}_2\,\underline{x}_2(t) + \underline{D}_2\,\underline{\mu}(t)$$

From these matrix equations we conclude that S_1 and S_2 are isomorphic if and only if there exists a non-singular matrix $\underline{M}$ representing the mapping $\gamma : \mathcal{X}_1 \longrightarrow \mathcal{X}_2$ such that:

$$\underline{A}_2 = \underline{M}\;\underline{A}_1\;\underline{M}^{-1} \qquad\qquad \underline{B}_2 = \underline{M}\,\underline{B}_1$$

$$\underline{C}_2 = \underline{C}_1\,\underline{M}^{-1} \qquad\qquad \underline{D}_2 = \underline{D}_1$$

In other words if two systems are isomporphic there exists a similarity relationship between their corresponding state transition matrices.

Thus given any linear sequential system S, all the similarity transformation matrices of its state transition matrix yield isomorphic systems. Since the matrix entries simply represent interconnections between different circuit variables; a modification of these matrices, as defined above, represents a change in the way these elements are interconnected to yield the <u>same</u> input-output behaviour. Clearly machine isomorphism affects in <u>no way</u> the input-output behaviour of a system.

IV-2: CHARACTERISTICS OF SHIFT REGISTER REALIZATIONS.

It is our objective to modify the matrix entries in such a way that the resulting state equations may be synthesized using shift registers as opposed to single delay elements. As in chapter II, $\underline{M}$ is called the transition matrix from the set of basis vectors in $\mathcal{X}_1$ to the set of basis vectors in $\mathcal{X}_2$.

Clearly by defining a new basis on the state space $\mathcal{X}$ we obtain a machine S' which is isomorphic to the machine S.

Consider the state equation of a linear sequential machine S:

$$\underline{x}(t+1) = \underline{A}\,\underline{x}(t) + \underline{B}\,\underline{\mu}(t)$$
$$\underline{y}(t) = \underline{C}\,\underline{x}(t) + \underline{D}\,\underline{\mu}(t)$$

$\underline{A}$ is the matrix representation of the linear transformation $\beta : \mathcal{X} \longrightarrow \mathcal{X}$. If the elementary divisors of $\underline{A}$ are $\ell_1(x), \ell_2(x), \dots, \ell_r(x)$ then there exists a matrix $\underline{A}_c$, similar to $\underline{A}$, such that:

$$\underline{A}_c = \underline{M}\ \underline{A}\ \underline{M}^{-1} = \begin{bmatrix} \underline{C}_1 & & & \\ & \underline{C}_2 & & \underline{0} \\ & \underline{0} & \ddots & \\ & & & \underline{C}_r \end{bmatrix}$$

where $\underline{C}_i$ is the companion matrix of $\ell_i(x)$ for $i = 1, 2, \dots, r$. Generally there exist more than one $\underline{M}$ such that $\underline{A}_c = \underline{M}\ \underline{A}\ \underline{M}^{-1}$. As was seen in chapter two, for every such $\underline{M}$, there exists a set of basis vectors $\{\underline{n}_1, \dots, \underline{n}_n\}$ such that

$[\underline{e}_i] = \underline{M}[\underline{n}_i]$ where $\{\underline{e}_1 \dots \underline{e}_n\}$ is the elementary basis of $\mathcal{X}$ where $e_i \equiv (0 \dots 1_i \dots 0)$. For any such $\underline{M}$, the system S', described by the following state equations, is isomorphic to S.

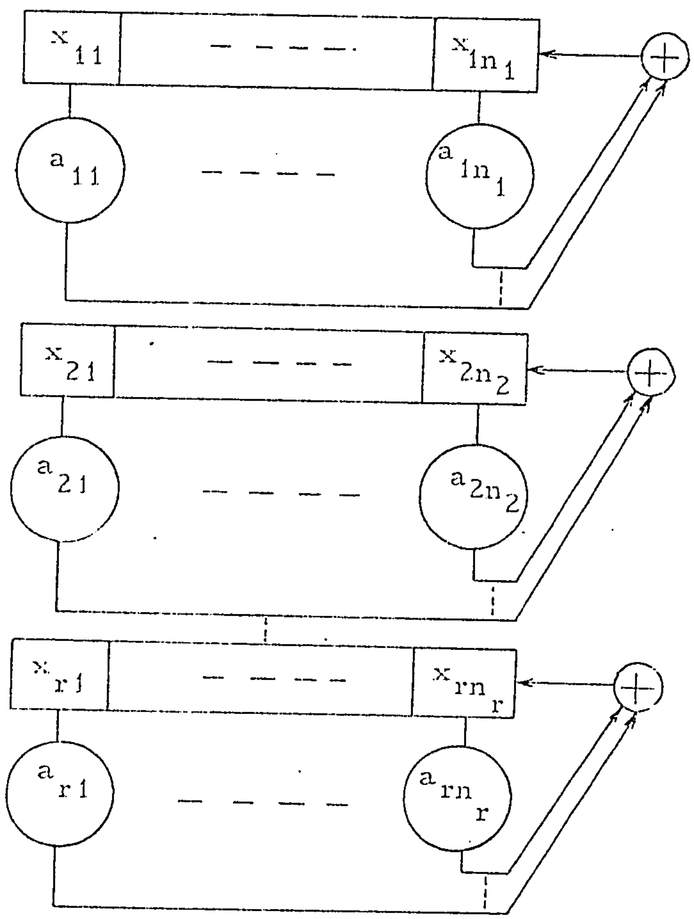
$$\begin{aligned} \underline{x}'(t+1) &= \underline{M}\ \underline{A}\ \underline{M}^{-1}\ \underline{x}'(t) + \underline{M}\ \underline{B}\ \underline{\mu}(t) \\ &= \underline{A}_c\ \underline{x}'(t) + \underline{B}'\ \underline{\mu}(t) \\ \underline{y}'(t) &= \underline{C}\ \underline{M}^{-1}\ \underline{x}'(t) + \underline{D}\ \underline{\mu}(t) \\ &= \underline{C}'\ \underline{x}'(t) + \underline{D}'\ \underline{\mu}(t) \end{aligned}$$

Let us show now a "canonical shift-register realization" is obtained by selectively choosing the set of basis vectors yielding $\underline{M}$.

For simplicity, consider only the state behaviour or the autonomous behaviour of machine S'.

That is $\underline{B} = \underline{0}$ and $\underline{x}'(t+1) = \underline{A}_c \; \underline{x}'(t)$.

The state behaviour of machine S' can now be synthesized using modulo adders and shift registers as shown in Figure 4.1.



where

$$x_{1i}(t+1) = x_{1,i+1}(t)$$

$$x_{1n_1}(t+1) = a_{11}x_{11} + \ldots + a_{1n_1}x_{1n_1}$$

For $i = 1, 2, \ldots, n_1 - 1$

$$x_{2i}(t+1) = x_{2,\, i+1}(t)$$

$$x_{2n_2}(t+1) = a_{21}\, x_{21} + \ldots + a_{2n_2}\, x_{2n_2}$$

For $i = 1, 2, \ldots, n_2 - 1$

$$x_{r_i}(t+1) = x_{r,i+1}(t)$$

$$x_{rn_r}(t+1) = a_{r1}\, x_{r1} + \ldots + a_{rn_r} x_{rn_r}$$

For $i = 1, 2, \ldots, n_r - 1$

Figure 4.1: Autonomous Realization of S'.

Obviously such a state behaviour can be realized using "r" shift-registers varying in length from n_1 to n_r. Since such a realization depends largely on the rational canonical form of $\underline{A}$, it is called the "canonical shift-register realization" of S'. Hence the following lemma can be stated without proof.

Lemma 4.1: The autonomous behaviour (zero-input) of a linear sequential machine is "canonical shift-register realizable".

Consider the situation where $\underline{B} = [\underline{b}_1 \ldots \underline{b}_s] \neq \underline{0}$, $s \geq 1$. In machine S', $\underline{B}' = \underline{M}\ \underline{B} = \underline{M}[\underline{b}_1 \ldots \underline{b}_s]$

$$= [\underline{M}\ \underline{b}_1, \underline{M}\underline{b}_2, \ldots, \underline{M}\underline{b}_s] \qquad \text{where}$$

$\underline{M}\ \underline{b}_i$ is the vertical coordinate representation of $\underline{b}_i$ with respect to the basis $\{\underline{n}_1, \ldots, \underline{n}_n\}$. From Figure 4.1 it is clear that S' is canonical shift register realizable (C.S.R.R) if only the state variables $\underline{x}_{in_i}$ for $i = 1, 2, \ldots, r$ are accessed. In other words S' is canonical shift register realizable if there exists a mapping $\gamma : \mathcal{X} \longrightarrow \mathcal{X}'$ which maps the vectors $\underline{b}_i$ on the vectors $\underline{b}'_i$ for $i = 1, 2, \ldots, s$ whose coordinates $n_1, n_1+n_2, \ldots, n = \sum_{i=1}^{r} n_i$ are the only nonzero coordinates.

Example 4.1: Let S be a binary sequential system described by the following state equations:

$$\underline{x}(t+1) = \underline{A}\ \underline{x}(t) + \underline{B}\ \underline{\mu}(t)$$

$$= \begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \underline{x}(t) + \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \underline{\mu}(t)$$

$$\underline{y}(t) = \underline{C}\ \underline{x}(t) + \underline{D}\ \underline{\mu}(t)$$

This system is readily synthesized using modulo-2 adders and delay elements.

A partial realization is shown in Figure 4.2.

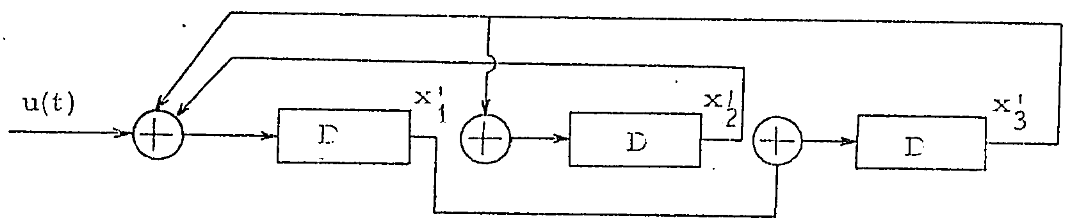


Figure 4.2: Delay Line Realization of Example 4.1.

Consider the similarity transformation matrix

$$\underline{M} = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

The corresponding isomorphic system S' is described by the following state equations:

$$\underline{x}'(t+1) = \underline{M}\ \underline{A}\ \underline{M}^{-1}\ \underline{x}'(t) + \underline{M}\ \underline{B}\ \underline{\mu}(t)$$

$$= \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} \underline{x}'(t) + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \underline{\mu}(t)$$

$$\underline{y}'(t) = \underline{C}\ \underline{M}^{-1}\ \underline{x}'(t) + \underline{D}'\ \underline{\mu}(t)$$

Therefore the system S' may be synthesized using modulo-2 adders and one shift register. A partial realization is shown in Figure 4.3.

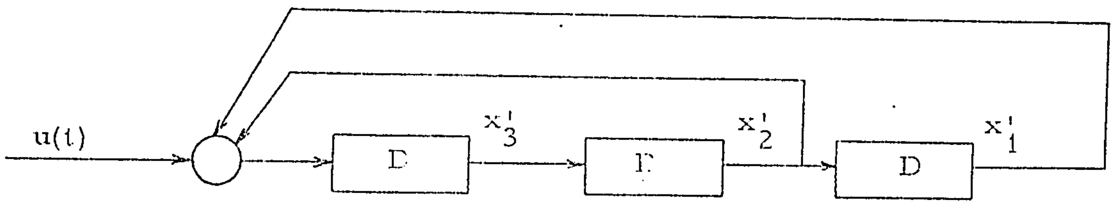


Figure 4.3: Shift Register Realization of Example 4.1.

Note the simplicity of this realization as compared to the realization of S. It is noteworthy that the shift register realizability of a sequential system is completely independent of the output equation $\underline{y}(t) = \underline{C}\,\underline{x}(t) + \underline{D}\,\underline{\mu}(t)$. This is due to the fact that the output variables are functions of the input and state variables but not vice versa.

The question now arises as to the realizability of such a realization. It was seen earlier that all zero input linear sequential machines are C.S.R.R. . Let us now consider the feasibility of such a realization when the system has a number of independent inputs.

Theorem 4.1: A necessary condition for a sequential system S to be C.S.R.R. is that $2^r - s \geq 1$, $s \leq p$ where $r \geq 1$ is the number of shift registers, s is the number of distinct columns of $\underline{B}$ and p is the number of independent inputs.

Proof: Let $\underline{B} = [\underline{b}_1, \underline{b}_2, \ldots, \underline{b}_p]$ $s \geq 1$, be the state input matrix of the sequential machine S'. Consider the reduced state-input matrix $\underline{B}_r = [\underline{b}_1 \ldots \underline{b}_s]$ containing only the distinct columns of $\underline{B}$. Since S can be synthesized using "r" shift registers, then only r state variables may be accessed. Recall that S is C.S.R.R. if there exists a mapping $\gamma : \mathfrak{X} \rightarrow \mathfrak{X}'$ which maps the vectors $\underline{b}_i$ on the vectors $\underline{M}\ \underline{b}_i = \underline{b}'_i$

for $i = 1, 2, \ldots, s$ where only "r" of their coordinates may be nonzero. $\underline{M}$ is the matrix representation of γ and $\mathcal{X}'$ is the state space for which a new basis has been defined. Hence γ can map any vector $\underline{b}_i$ on at most $(2^r - 1)$ elements of $\mathcal{X}'$ satisfying the coordinate restraint.

This is shown schematically in Figure 4.4.

$$\underline{M}\,\underline{B}_r = \underline{B}'_r = \begin{array}{c} \underline{b}'_1 \quad\; \underline{b}'_2 \cdots\cdots \quad \underline{b}'_s \\ \begin{bmatrix} 0 & 0 & \cdots\cdots & 0 \\ \vdots & \vdots & & \vdots \\ * & * & & * \\ \vdots & \vdots & & \vdots \\ 0 & 0 & & 0 \\ \vdots & \vdots & & \vdots \\ * & * & \cdots\cdots & * \end{bmatrix} \end{array} \left.\begin{array}{l} \\ \\ \\ \\ \\ \\ \end{array}\right\} \begin{array}{l} \text{"r" accessible} \\ \text{state variables} \end{array}$$

where the symbol $*$ represents a zero or a one.

Figure 4.4: Schematic Diagram of $\underline{B}'_r$.

Thus γ may map ---

$$\begin{array}{llll} \underline{b}_1 & \text{on at most} & (2^r - 1) & \text{vectors in } \mathcal{X}' \\ \underline{b}_2 & \text{on at most} & (2^r - 1) - 1 = 2^r - 2 & \text{vectors in } \mathcal{X}' \\ \vdots & & \qquad\qquad\quad \vdots & \\ \underline{b}_s & \text{on at most} & (2^r - 1) - (s - 1) = 2^r - s & \text{vectors in } \mathcal{X}' . \end{array}$$

Thus the vector $\underline{b}_s$ is mapped on at most $(2^r - s)$ vectors in $\mathcal{X}'$. For S to be C.S.R.R. there must exist at least one vector in $\mathcal{X}'$ on which $\underline{b}_s$ is mapped.

$$\therefore \quad 2^r - s \geq 1$$

Note that by mapping the columns of $\underline{B}$ in the above order, there is no loss of generality.

IV-3: ALGORITHM FOR SHIFT REGISTER REALIZATION.

Using the results developed in the previous sections, the shift register realization algorithm is as follows:

Step 1: Determine the elementary divisors $\ell_1(x), \ell_2(x), \ldots, \ell_r(x)$ of the state transition matrix $\underline{A}$. Let the degree of $\ell_i(x)$ be n_i for $i = 1, 2, \ldots, r$.

Step 2: Obtain the matrices $\ell_i(\underline{A})$.

Step 3: Obtain the vectors in the null space of $\ell_i(\underline{A})$ for $i = 1, 2, \ldots, r$. Let S_i be the set of basis vectors for the null space $\mathcal{N}_i$ of $\ell_i(\underline{A})$. S_i is the set of columns of the matrix $\underline{N}_i$ (Appendix B).

Step 4: Let $LS(S_i)$ be the linear span of S_i. Form the matrix $\underline{W}$ as follows:

$$\underline{W} = [\, LS(S_1) \mid LS(S_2) \mid \ \ldots \ \mid LS(S_r) \,]$$

Recall that $\underline{A}_c^T$ is the matrix representation of $\beta : \mathcal{X} \to \mathcal{X}$ with respect to the basis

$$\{ \underline{b}_1, \underline{A}\underline{b}_1, \ldots, \underline{A}^{n_1-1}\underline{b}_1; \underline{b}_2, \ldots ; \underline{b}_r, \underline{A}\underline{b}_r, \ldots, \underline{A}^{n_r-1}\underline{b}_r \}$$

$$= \{ \underline{v}_1, \underline{v}_2, \ldots, \underline{v}_n \}$$

such that $\mathcal{X} = Z(\underline{b}_1, \underline{A}) \otimes Z(\underline{b}_2, \underline{A}) \otimes \ldots \otimes Z(\underline{b}_r, \underline{A})$

where $\underline{b}_i \in \ell_i(\underline{A})$, $\ell_i(x)$ is the minimum polynomial of $Z(\underline{b}_i, \underline{A})$ and $[\, \underline{e}_1, \ldots, \underline{e}_n \,] = [\, \underline{v}_1, \ldots, \underline{v}_n \,] \, \underline{M}_B$.

Step 5: Choose one element $\underline{b}_i$ from each null space $\mathcal{N}_i$. Obtain the vectors $\underline{b}_i, \underline{A}\,\underline{b}_i, \ldots, \underline{A}^{n_i-1}\underline{b}_i$. Form the matrix $\underline{M}_B^{-1}$ as follows:

$$\underline{M}_B^{-1} = [\, \underline{b}_1, \underline{A}\,\underline{b}_1, \ldots, \underline{A}^{n_1-1}\underline{b}_1; \ \underline{b}_2, \ldots; \ \underline{b}_r, \underline{A}\underline{b}_r, \ldots, \underline{A}^{n_r-1}\underline{b}_r \,]$$

If $\underline{M}_B^{-1}$ is nonsingular then its' columns form a basis for $\mathcal{X}$. The matrix representation of $\beta : \mathcal{X} \rightarrow \mathcal{X}$, with respect to this basis, is $\underline{A}_c^T = \underline{M}_B \, \underline{A} \, \underline{M}_B^{-1}$.

Step 6: Using the algorithm given in Appendix B, obtain a matrix $\underline{M}_T$ such that $\underline{A}_c = \underline{M}_T \; \underline{A}_c^T \; \underline{M}_T^{-1}$.

Step 7: Obtain the matrix $\underline{M} = \underline{M}_T \; \underline{M}_B$ where $\underline{A}_c = \underline{M} \; \underline{A} \; \underline{M}^{-1}$.

Step 8: Obtain the matrix $\underline{M} \; \underline{B}$ and verify that $n_1, n_1 + n_2, \ldots, n$ are the only nonzero rows. If this test is satisfied go to Step 11; otherwise go to Step 9.

Step 9: Repeat Steps 6, 7 and 8 for all the matrices $\underline{M}_T$ generated by the algorithm in Appendix B. When the set of matrices $\underline{M}_T$ is exhausted, go to Step 10.

Step 10: Repeat Steps 5, 6, 7 and 8 for all the sets $\{\underline{b}_1, \underline{A}\underline{b}_1, \ldots, \underline{A}^{n_1 - 1}\underline{b}_1$ $\underline{b}_2, \ldots; \; \underline{b}_r, \underline{A}\underline{b}_r, \ldots, \underline{A}^{n_r - 1}\underline{b}_r\}$ where $\underline{b}_i \in \eta_i$ for $i = 1, 2, \ldots, r$. Having exhausted all the possible combinations, go to Step 12.

Step 11: The state equations of the isomorphic linear sequential machine S' which is C.S.R.R. are :

$$\underline{x}'(t+1) = \underline{A}_c \; \underline{x}'(t) + \underline{M} \; \underline{B} \; \underline{\mu}(t)$$

$$\underline{y}'(t) = \underline{C} \; \underline{M}^{-1} \underline{x}'(t) + \underline{D} \; \underline{\mu}(t)$$

Step 12: End of algorithm.

To illustrate this algorithm, consider the following example.

Example 4.2 : Consider the linear sequential machine S defined over GF = {0, 1} and described by the following state equations:

$$\underline{x}(t+1) = \underline{A}\,\underline{x}(t) + \underline{B}\,\underline{\mu}(t)$$

$$= \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ \cdot \\ x_6(t) \end{bmatrix} + \begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} \mu_1(t) \\ \mu_2(t) \\ \vdots \\ \vdots \\ \mu_4(t) \end{bmatrix}$$

$$\underline{y}(t) = \underline{C}\,\underline{x}(t) + \underline{D}\,\underline{\mu}(t)$$

$$\begin{bmatrix} y_1(t) \\ y_2(t) \\ \vdots \\ y_4(t) \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 \end{bmatrix} \underline{x}(t) + \begin{bmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 \end{bmatrix} \underline{\mu}(t)$$

This system is readily synthesized using modulo - 2 adders and delay elements. A partial realization is shown in Figure 4.8 .

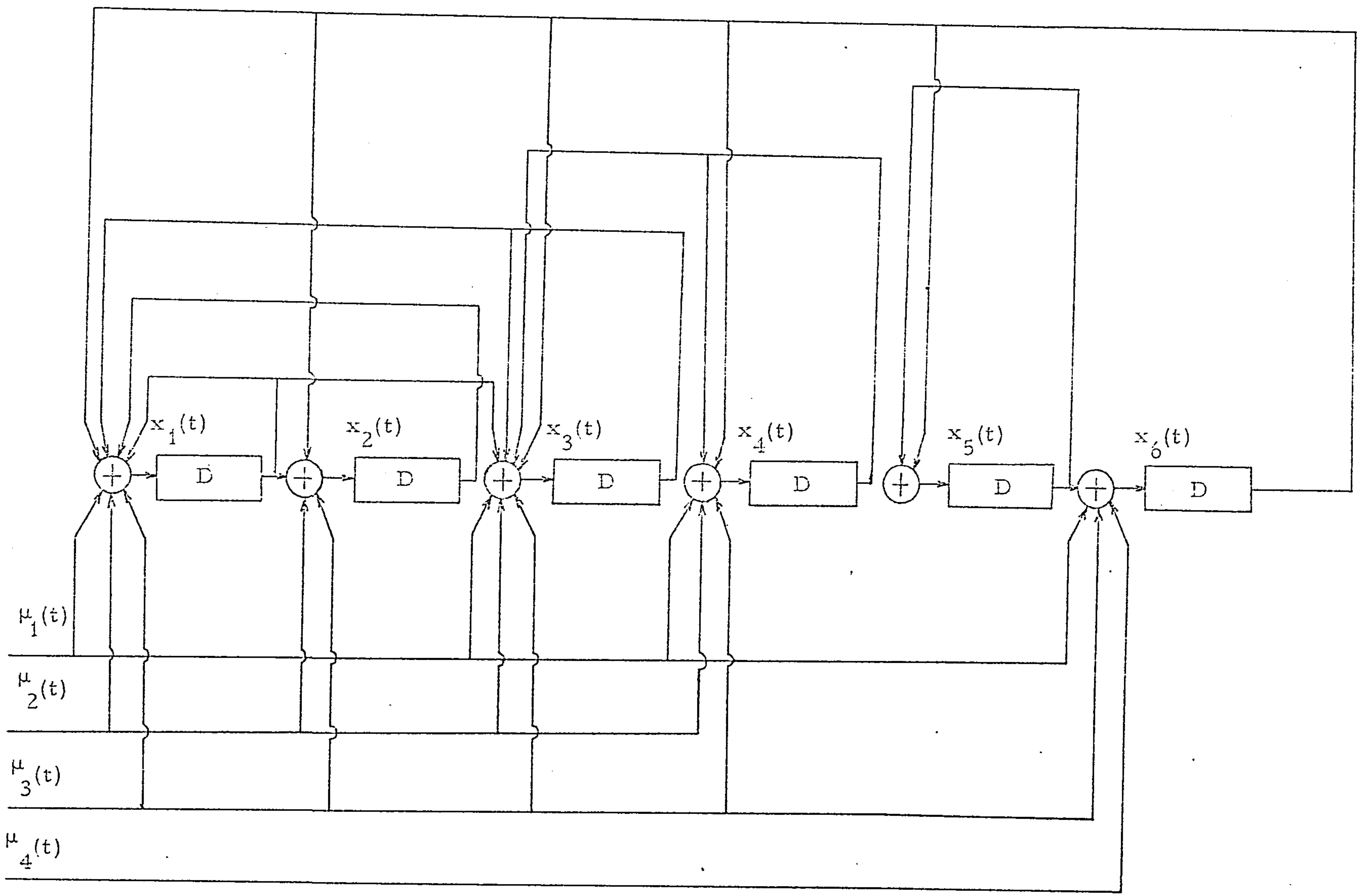


Figure 4.8 : Delay Line Realization of Example 4.2.

The elementary divisors of the state transition matrix $\underline{A}$ of S are $\ell_1(x) = 1 + x + x^3$, $\ell_2(x) = 1 + x$ and $\ell_3(x) = 1 + x + x^2$. Hence using the above algorithm it is possible to obtain a linear sequential machine S' isomorphic to S and realizable using 3 shift registers of length 1, 2 and 3 delay elements. The following is one of the many linear sequential machines S' yielded by the algorithm.

S' is a linear sequential machine defined over $GF = \{0, 1\}$ isomorphic to S and described by the following state equations :

$$\underline{x}'(t+1) = \underline{M}\,\underline{A}\,\underline{M}^{-1}\,\underline{x}'(t) + \underline{M}\,\underline{B}\,\underline{\mu}(t)$$

$$= \left[\begin{array}{ccc|c|cc} 0 & 1 & 0 & & & \\ 0 & 0 & 1 & & \underline{0} & \\ 1 & 1 & 0 & & & \\ \hline & & & 1 & & \\ \hline & \underline{0} & & & 0 & 1 \\ & & & & 1 & 1 \end{array}\right] \underline{x}'(t) + \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 \end{bmatrix} \underline{\mu}(t)$$

$$\underline{y}'(t) = \underline{C}\,\underline{M}^{-1}\,\underline{x}(t) + \underline{D}'\,\underline{\mu}(t)$$

$$= \begin{bmatrix} 1 & 1 & 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 & 1 & 0 \end{bmatrix} \underline{x}(t) + \begin{bmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 \end{bmatrix} \underline{\mu}(t)$$

A partial realization of this system, using modulo-2 adders and delay elements, is shown in Figure 4.9.

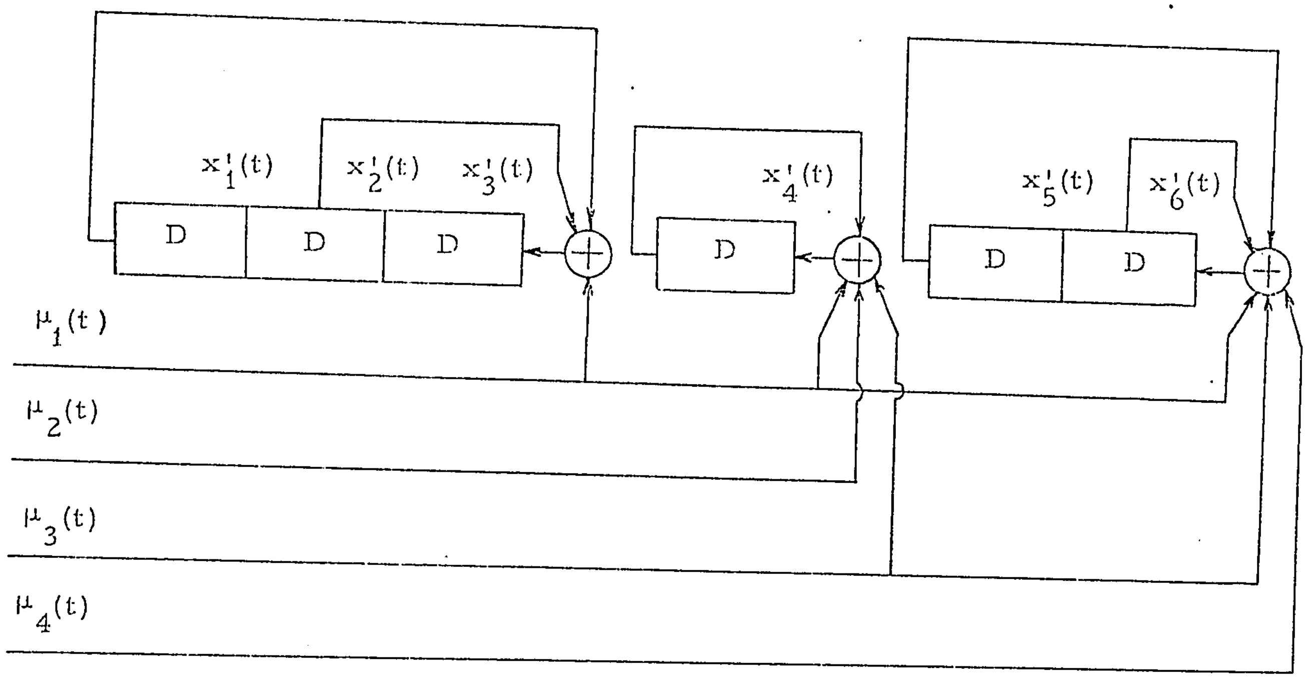


Figure 4.9 : Delay Line Realization of Machine S' .

Note the simplicity of this shift register realization as compared to the realization given in Figure 4.8 .

IV-4 : SINGLE SHIFT REGISTER REALIZABILITY

In this section additional results are presented for single shift register realizability when the reduced state input matrix $\underline{B}_r$ has a single column. This case is of special interest as it corresponds to single input systems.

Theorem 4.2 : A linear sequential machine S whose state transition matrix $\underline{A}$ is nonderogatory and has an irreducible minimum polynomial is single shift register realizable if and only if the reduced state input matrix $\underline{B}_r$ has a single column.

Proof: This class of linear sequential machines is characterized by the following state input equation:

$$x(t+1) = \underline{A}\underline{x}(t) + \underline{B}\underline{\mu}(t)$$

where $p(x) = m(x) = x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_o$ is irreducible over F and $\underline{B} = [\underline{b}_1 \underline{b}_2 \ldots \underline{b}_p]$ where $\underline{b}_1 = \underline{b}_2 = \ldots = \underline{b}_p$ such that $\underline{B}_r = \underline{b}_1$. Note that the companion matrix form of $\underline{A}$ is $\underline{A}_c$ such that

$$\underline{A}_c = \begin{bmatrix} 0 & 1 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & 1 \\ -a_o & -a_1 & \cdots & -a_{n-1} \end{bmatrix}$$

A - Let us prove that the machine S is single shift register realizable if $\underline{B}_r$ has a single column. The proof is completed in two parts. First it will be shown that different isomorphic mappings map a given vector into a different set of coordinates. Secondly we proceed to show that the machine S is single shift register realizable.

i - Let $\underline{M}_i$ be the matrix representation of $\gamma_i : \mathcal{X} \rightarrow \mathcal{X}_i$ for $i = 1, 2$, where $\mathcal{X}_i$ is the state space $\mathcal{X}$ for which a new basis has been defined and $\underline{A}_c = \underline{M}_1 \underline{A} \underline{M}_1^{-1} = \underline{M}_2 \underline{A} \underline{M}_2^{-1}$ where $\underline{M}_1 \neq \underline{M}_2$. Let $\underline{x} \in \mathcal{X}$ such that $\underline{x}_i = \underline{M}_i \underline{x} \in \mathcal{X}_i$ for $i = 1, 2$. Assuming $\underline{x}_1 = \underline{x}_2$

then $\underline{x}_i = \underline{M}_1 \underline{x} = \underline{M}_2 \underline{x} = \underline{x}_2$. Since γ_1 and γ_2 are isomorphic mappings then $\underline{x} = \underline{M}_1^{-1} \underline{M}_2 \underline{x}$ such that $\underline{M}_1 = \underline{M}_2$. This is a contradiction since $\underline{M}_1 \neq \underline{M}_2$. Hence two different isomorphic mappings map a given vector into a different set of coordinates.

ii - By lemma 2.6, for all $\underline{x} \neq \underline{o}$, $\underline{x} \in \mathcal{X}$ $S_x = \{\underline{x}, \underline{A}\underline{x}, \ldots, \underline{A}^{n-1}\underline{x}\}$ is a basis of $\mathcal{X}$ such that the matrix representation of $\beta : \mathcal{X} \rightarrow \mathcal{X}$ with respect to S_x is $\underline{A}_c^T$. Since there are $(2^n - 1)$ nonzero elements in $\mathcal{X}$ then there exist $(2^n - 1)$ distinct matrices $\underline{M}_x$ representing the isomorphic mapping $\gamma_x : \mathcal{X} \rightarrow \mathcal{X}_x$ such that $\underline{A}_c^T = \underline{M}_x \underline{A} \underline{M}_x^{-1}$ where $\mathcal{X}_x$ is the state space $\mathcal{X}$ with respect to S_x. Since $\underline{A}_c^T$ and $\underline{A}_c$ are similar, as shown in Appendix A, then there exists at least one matrix $\underline{M}^*$ such that $\underline{A}_c = \underline{M}^* \underline{A}_c^T \underline{M}^{*-1}$. Thus there exists $(2^n - 1)$ matrices $\bar{\underline{M}}_x = \underline{M}^* \underline{M}_x$ representing the isomorphic mapping $\bar{\gamma}_x : \mathcal{X} \rightarrow \mathcal{X}_x$ such that $\underline{A}_c = \bar{\underline{M}}_x \underline{A} \bar{\underline{M}}_x^{-1}$.

Since the number of mappings $\bar{\gamma}_x$ is equal to the number of nonzero vectors in $\mathcal{X}$ and since a vector $\underline{x} \in \mathcal{X}$ is mapped on a different set of coordinates by each mapping then for all $\underline{x}_1, \underline{x}_2 \in \mathcal{X}$ there exists a mapping $\bar{\gamma}_x$ such that $\bar{\gamma}_x(\underline{x}_1) = \underline{x}_2$. Setting $\underline{x}_1 = \underline{b}_1$ and $\underline{x}_2 = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}$ then the above class of linear sequential machines is single shift register realizable. A partial shift register realization is shown in figure 4.8 .

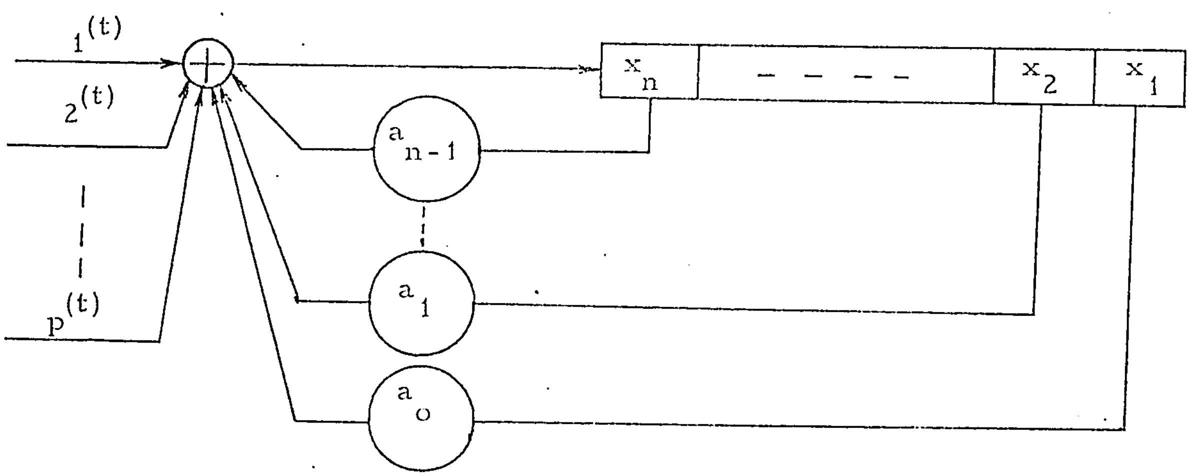


Figure 4.10: Shift register realization of machine S.

B - Let us prove that $\underline{B}_r$ has a single column if the machine S is single shift register realizable. The proof is by contradiction. Assume $\underline{B}_r = \{ \underline{b}_1 \underline{b}_2 \}$ where $\underline{b}_1 \neq \underline{b}_2$. Since m(x) is irreducible and the machine S is single shift register realizable then only the n^{th} state variable may be accessed. Thus $\underline{b}_1$ and $\underline{b}_2$ must be mapped by a similarity transformation matrix into the vector $\begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$. As seen in part A-i of this proof this is possible only if $\underline{b}_1 = \underline{b}_2$. This is a contradiction. Hence $\underline{B}_r = \underline{b}_1$ and $\underline{B}_r$ has a single column.

An extension of this theorem can be obtained by using theorem A.5 of appendix A.

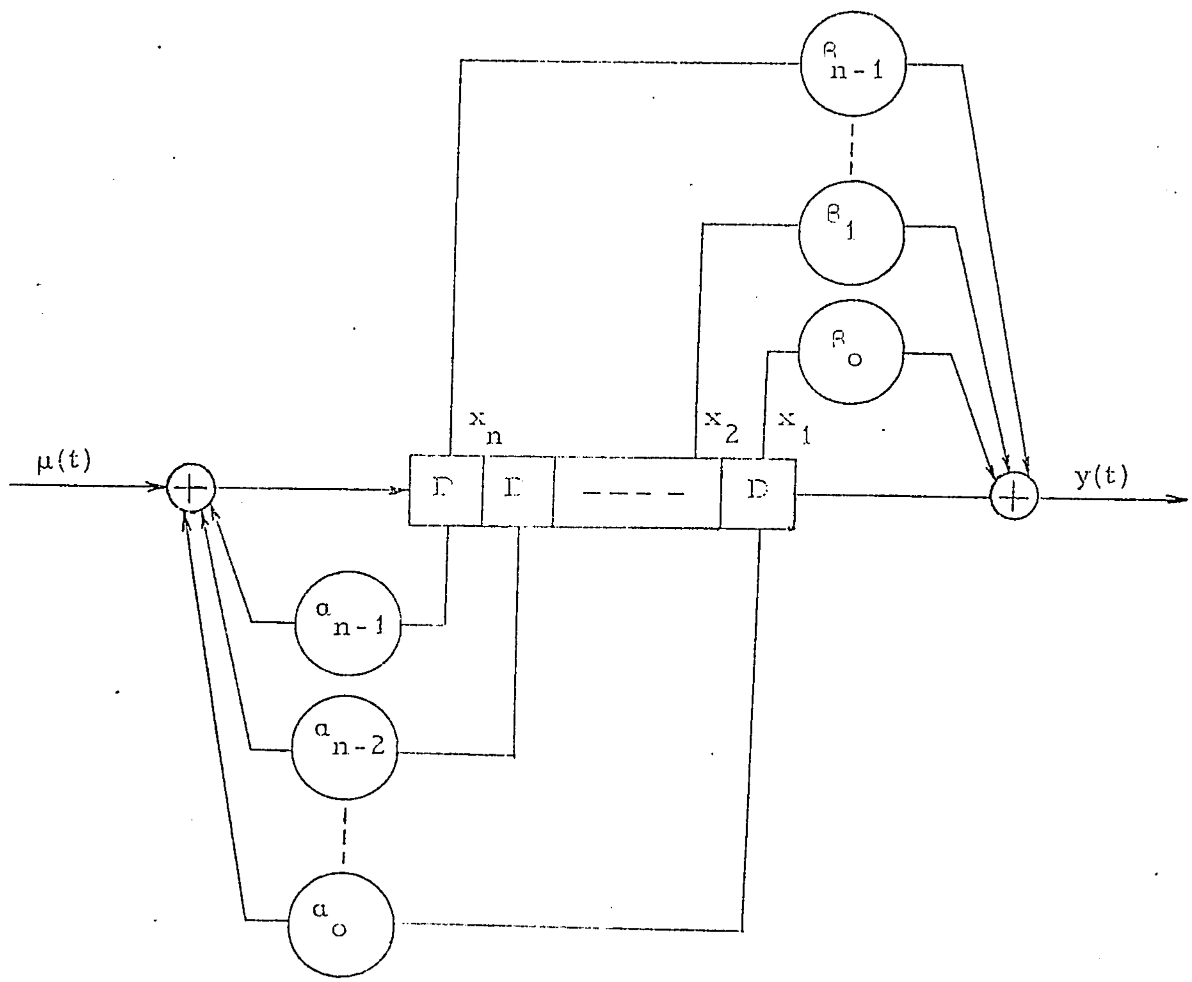
Hypothesis 4.1 : A linear sequential machine S whose state transition matrix $\underline{A}$ is nonderogatory is single shift register realizable if and only if the reduced state input matrix $\underline{B}_r$ has a single column.

In view of theorem A.5 of appendix A, given a system S, there exists a similarity transformation matrix $\underline{M}$ which yields an isomorphic system S' whose state transition matrix $\underline{A}^* = \underline{M}\,\underline{A}\,\underline{M}^{-1}$ is the companion matrix of m(x). There exists at least one such $\underline{M}$ matrix. If $\underline{B}_r$ has a single column then S is single shift register realizable if there exists a matrix $\underline{M}$ which maps this column vector into the vector $\begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$. On the other hand if S is single shift register realizable then necessarily $\underline{B}_r$ has a single column as in part B of the proof of theorem 4.2.

A further extension of these results to include all single input single output systems yields the following hypothesis.

Hypothesis 4.2 : All single input, single output linear sequential machines S are single shift register realizable.

Consider the following general shift register realization of a single input, single output system.



This system is described by the following state equations :

$$\underline{x}(t+1) = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} 0 & 1 & & & \\ 0 & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & & 1 \\ \alpha_o & \alpha_1 & \cdots & \cdots & \alpha_{n-1} \end{bmatrix} \underline{x}(t) + \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} \underline{\mu}(t)$$

$$y(t) = [\, \beta_0 \, \beta_1 \, \ldots\ldots \, \beta_{n-1} \,] \, \underline{x}(t)$$

where α_i, β_i are elements in the field over which S is defined.

Hence the transfer function representation of S using the delay operator D is:

$$\frac{Y(D)}{U(D)} = \underline{C}\,[D^{-1}\underline{I} + \underline{A}]^{-1}\,\underline{B}$$

The matrix $[D^{-1}\underline{I} + \underline{A}]_{nxn} = \dfrac{[n_{ij}(D^{-1})]}{|D^{-1}\underline{I} + \underline{A}|}$ for $i, j = 1, 2, \ldots, n$

where $|D^{-1}\underline{I} + \underline{A}| = (D^{-1})^n + a_{n-1}(D^{-1})^{n-1} + \ldots + a_1(D^{-1}) + a_o$ and $n_{ij}(D^{-1})$ are polynomials of degree less than n.

Furthermore,

$$[D^{-1}\underline{I} + \underline{A}] = \begin{bmatrix} D^{-1} & 1 & & & \\ 0 & D^{-1} & 1 & \ddots & \underline{0} \\ \vdots\ \underline{0} & \vdots & & \ddots & \\ \vdots & \vdots & & & 1 \\ a_o & a_1 & \ldots & & D^{-1}+a_{n-1} \end{bmatrix}$$

such that

$$[D^{-1}\underline{I} + \underline{A}]^{-1} = \frac{1}{|D^{-1}\underline{I} + \underline{A}|}\begin{bmatrix} & 1 \\ & (D^{-1}) \\ n_{ij}(D^{-1}) & (D^{-1})^2 \\ & \vdots \\ & (D^{-1})^{n-1} \end{bmatrix}$$

Since $\underline{B}$ is of the form shown then

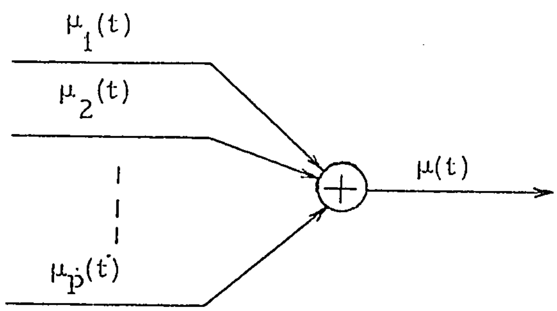
$$[D^{-1}\underline{I} + \underline{A}]^{-1}\,\underline{B} = \frac{1}{|D^{-1}\underline{I} + \underline{A}|}\begin{bmatrix} 1 \\ (D^{-1}) \\ \vdots \\ (D^{-1})^{n-1} \end{bmatrix}.$$

Since $\underline{C}$ is of the form shown then :

$$\frac{Y(D)}{U(D)} = \frac{\beta_{n-1}(D^{-1})^{n-1} + \ldots\ldots + \beta_1(D^{-1}) + \beta_o}{(D^{-1})^n + \alpha_{n-1}(D^{-1})^{n-1} + \ldots\ldots + \alpha_1(D^{-1}) + \alpha_o}$$

This is the general expression for the transfer function representation of any single input, single output sequential machine. Hence all single input, single output sequential machines are single shift register realizable. This proof is by construction as a mathematical proof could not be found.

This indicates that theorem 4.2 may be generalized to include all the linear sequential machines whose $\underline{B}_r$ contains a single column. This is the equivalent of a single input machine as shown in figure 4.9.



where $\underline{B} = [\underline{b}_1, \underline{b}_2 \ldots\ldots \underline{b}_p]$

and $\underline{B}_r = [\underline{b}_1]$ since

$\underline{b}_1 = \underline{b}_2 = \ldots\ldots = \underline{b}_p$.

Figure 4.11: Single input representative of S.

Lemma 4.2: A necessary condition for a sequential machine S to be single shift register realizable is that $\underline{B}_r$ contain a single column.

Proof: The proof follows directly from part B of the proof of theorem 4.2.

It should be noted that the autonomous behaviour of a linear sequential machine need not be restricted to a zero input. Unless otherwise stated, the antonomous behaviour of a linear sequential machine refers to the response of the machine to a constant input combination. With this generalization in mind, consider the following theorem.

Theorem 4.3: The generalized autonomous behaviour of a linear sequential machine is single shift register realizable if its' state transition matrix is non-derogatory and has an irreducible minimum polynomial over F.

Proof: Let the state equation of a linear sequential machine S be:

$$\underline{x}(t+1) = \underline{A}x(t) + [\,\underline{b}_1\ \underline{b}_2\ \ldots\ \underline{b}_p\,] \begin{bmatrix} \mu_1(t) \\ \mu_2(t) \\ \vdots \\ \mu_p(t) \end{bmatrix}$$

Construct the following linear sequential machine S' :

$$\underline{x}(t+1) = \underline{A}\,\underline{x}(t) + [b]\mu(t) \quad \text{where}$$

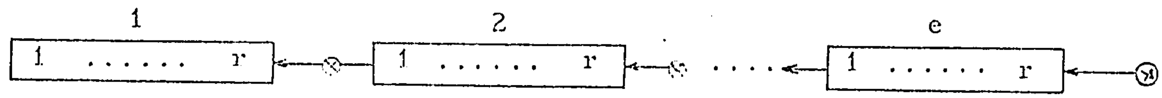
$b = b_1 \mu_1 + b_2 \mu_2 + \ldots\ldots + b_p\ \mu_p$. By theorem 4.2 S' is single shift register realizable. The input-output behaviour of S and S' are identical for any constant input combination $\{\mu_1 \mu_2 \ldots \mu_p\}$ and $\mu(t) = 1$ for all t. Since S' is single shift register realizable then S is single shift register realizable for any constant input combination.

It is expected that a generalization of theorem 4.3 holds if hypothesis 2 holds.

IV-5 : CONCLUDING REMARKS.

The purpose of this chapter was to study the characteristics of shift register realizable machines and to develop an algorithm for this realization.

The algorithm, excluding step 1, was implemented on the IBM system 360 model 65 at the Ottawa University Computing Center. The programming language used was Fortran IV-G level 18. The details of theprogram and the flow charts are presented in Appendix C. This algorithm yielded a parallel shift register realization. Using the results in appendix A, if $p(x) = m(x) = (f(x))^{c}$ then it is possible to obtain a serial shift register realization, using "c" registers of length " r " where r is the degree of $f(x)$, as follows:



Using the results in Appendix A it is also possible to realize a sequential machine using a single shift register.

The algorithms for such serial and single shift register realizations were not implemented due to lack of time.

In section IV-4 we introduced additional results for single shift register realization where it was shown that this is possible if $\underline{B}_r$ has a single column.

CHAPTER V

SYNTHESIS OF LINEAR, TIME-INVARIANT SYSTEMS

The theorems developed in chapter III define the conditions under which a linear, time-invariant system is controllable and/or observable. Given an $\underline{A}$-matrix and using these theorems, it is quite easy to construct a controllable and/or observable system. In fact this is how the examples in chapter III were constructed.

The purpose of this chapter is to introduce algorithms for constructing controllable and observable systems from given uncontrollable and/ or unobservable systems by adding additional inputs and/ or outputs. The cases of the nonderogatory and derogatory systems are treated separately.

V-I: ALGORITHM FOR NON-DEROGATORY SYSTEMS

Let $m(x) = p(x)$ be the minimum polynomial of $\underline{A}$, the matrix representation of the state transition $\beta : \mathcal{X} \to \mathcal{X}$. By definition of a minimum polynomial, there exists at least one vector whose minimum polynomial is $m(x)$. By theorem 3.3 the system is controllable if matrix $\underline{B}$ contains as one of its columns at least one element $\underline{b} \in \mathcal{X}$ whose minimum polynomial is $m(x)$. Similarly, by theorem 3.9 the system is observable if matrix $\underline{C}^T$ contains as one of its columns at least one element $\underline{c}^T \in \mathcal{X}$ whose minimum polynomial is $m(x)$.

Thus the problem reduces to finding a vector $\underline{v}$ that satisfies the required conditions. However the set of elements $\underline{v} \in \mathcal{V}$, whose minimum polynomial is $q(x)$ such that $q(\underline{A})\,\underline{v} = \underline{0}$, is the kernel of $q(\underline{A})$ or the null space of $q(\underline{A})$.

Thus the addition of an element from the null space of $m(\underline{A})$, whose minimum polynomial is $m(x)$, as a column of $\underline{B}$ yields a controllable system. Likewise the addition of an element from the null space of $m(\underline{A}^T)$, whose minimum polynomial is $m(x)$, as a column of $\underline{C}^T$ yields an observable system.

Therefore the algorithm for constructing a controllable (observable) system is as follows:

1. Find the null space of each elementary divisor of $\underline{A}$ ($\underline{A}^T$).
2. Choose from each null space a vector $\underline{v}_i$, such that the resulting set of vectors $\{ \underline{v}_1, \underline{v}_2, \ldots, \underline{v}_r \}$ are linearly independent.
3. Let the required modification $\underline{b}_o$ ($\underline{c}_o^T$) be $\sum_{i=1}^{r} \underline{v}_i$.

The required null space of step 1 can be computed using the algorithm given in Appendix B.

In this way it is <u>always</u> possible to transform a system whose state transition matrix is non-derogatory, into a controllable and/or observable system by increasing respectively the number of inputs and/ or outputs by <u>one</u> only.

To see the required modification let the state equations of the uncontrollable and unobservable system S be:

$$\underline{x}'' = \underline{A}_{n\times n}\, \underline{x} + [\, \underline{b}_1\, \underline{b}_2 \ldots \underline{b}_r \,] \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_r \end{bmatrix} .$$

$$\underline{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_p \end{bmatrix} = [\underline{c}_1^T ,\, \underline{c}_2^T \ldots \underline{c}_p^T]^T \underline{x} + [\underline{d}_1\, \underline{d}_2 \ldots \underline{d}_r] \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_r \end{bmatrix}$$

where $\underline{x}''$ denotes either $\dot{\underline{x}}$ or $\underline{x}'$, for continuous or discrete systems respectively.

Let $\underline{b}_o$ and $\underline{c}_o^T$ be the results of the above algorithm. The addition of $\underline{b}_o$ and $\underline{c}_o^T$ to $\underline{B}$ and $\underline{C}^T$ respectively yields the following state equations for the corresponding controllable and observable system S'.

$$\underline{x}'' = \underline{A}_{nxn} \ \underline{x} + [\underline{b}_1 \ \underline{b}_2 \cdots \underline{b}_r \mid \underline{b}_o] \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_r \\ \mu_{r+1} \end{bmatrix}$$

$$\underline{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_p \\ y_{p+1} \end{bmatrix} = [\underline{c}_1^T \ \underline{c}_2^T \cdots \underline{c}_p^T \mid \underline{c}_o^T]^T \underline{x} + [\underline{d}_1 \ \underline{d}_2 \cdots \underline{d}_r \mid \underline{o}] \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_r \\ \mu_{r+1} \end{bmatrix}$$

Consider the following examples.

Example 5.1: Let the state equations of a linear sequential machine, defined over GF = { 0, 1}, be:

$$\underline{x}'(t) = \underline{A}\,\underline{x}(t) + \underline{B}\,\underline{\mu}(t) = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix} \underline{x}(t) + \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \mu_1(t)$$

$$\underline{y}(t) = \underline{C}\,\underline{x}(t) + \underline{D}\,\underline{\mu}(t)$$

where $\underline{A}$ is a non-derogatory matrix whose minimum polynomial is $m(x) = (x+1)(x^2+x+1)$. This system is uncontrollable and its state diagram is represented in Figure 5.1.

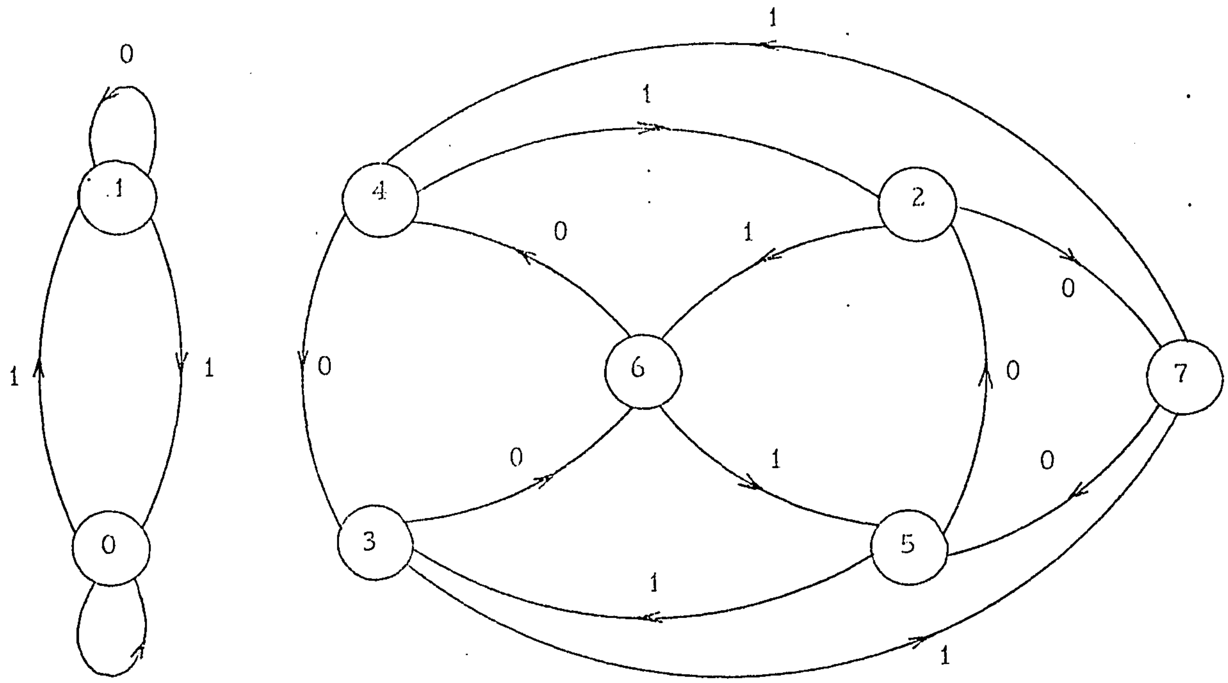


Figure 5. 1 : State Diagram.

In Figure 5. 1 the decimal code representation of a state is obtained from the binary number corresponding to (x_3, x_2, x_1). Note that the state diagram is disconnected since the set of states $\{0, 1\}$ forms one sub-state diagram, while $\{2, 3, 4, 5, 6, 7\}$ forms another.

Let us go through the steps of the algorithm:

Step 1: The elementary divisors of $\underline{A}$ are $m_1(x) = 1 + x$ and $m_2(x) = 1 + x + x^2$. The vectors in the null space of $m_1(\underline{A})$ are $\left\{\begin{matrix} 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{matrix}\right\}$

The vectors in the null space of $m_2(\underline{A})$ are $\left\{\begin{matrix} 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{matrix}\right\}$

Step 2: Choosing one element from every null space the resulting sets of linear independent vectors are obtained.

$$S_1 = \left\{\begin{matrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{matrix}\right\} ; \quad S_2 = \left\{\begin{matrix} 1 & 1 \\ 0 & 0 \\ 0 & 1 \end{matrix}\right\} \quad \text{and} \quad S_3 = \left\{\begin{matrix} 1 & 1 \\ 0 & 1 \\ 0 & 1 \end{matrix}\right\}$$

Step 3: Thus the following $\underline{b}$'s are valid modifications.

$\underline{b}_{01} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$; $\underline{b}_{02} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ and $\underline{b}_{03} = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$. Let $\underline{B} = \left[\begin{array}{c|c} 1 & \\ 0 & \underline{b}_{01} \\ 0 & \end{array}\right] = \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$

and Figure 5.2 represents the state diagram of the resulting system.

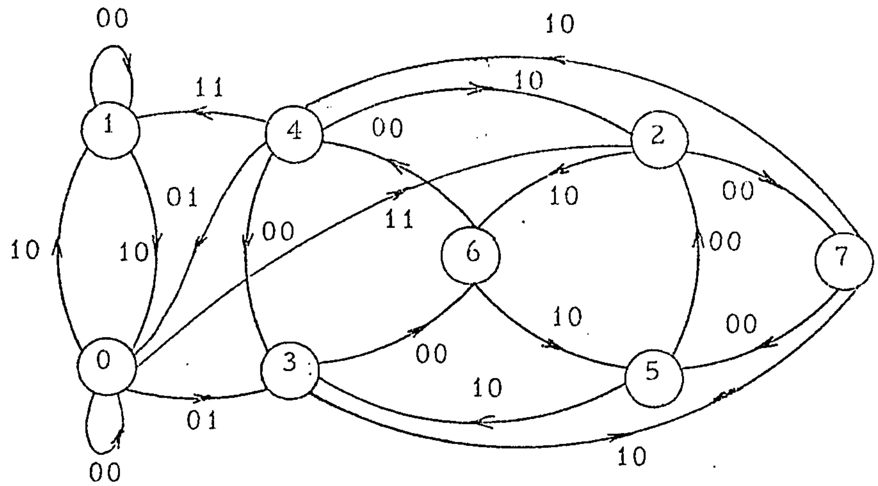


Figure 5.2 : State Diagram.

Let $\underline{B} = \left[\begin{array}{c|c} \begin{matrix}1\\0\\0\end{matrix} & \underline{b}_{o2} \end{array}\right] = \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix}$ and Figure 5.3 represents

the state diagram of the resulting system.

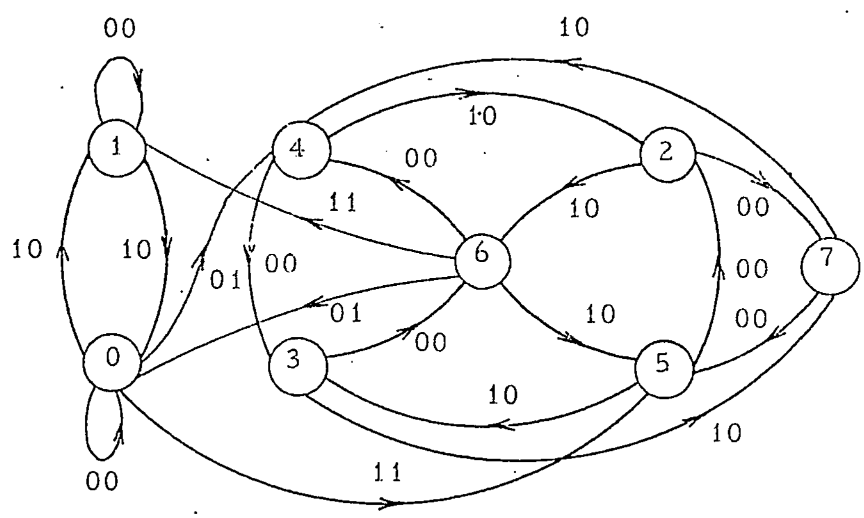


Figure 5.3 : State Diagram.

Let $\underline{B} = \left[\begin{array}{c|c} 1 & \\ 0 & \underline{b}_{03} \\ 0 & \end{array}\right] = \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 0 & 1 \end{bmatrix}$ and Figure 5.4 represents the state diagram of the resulting system.

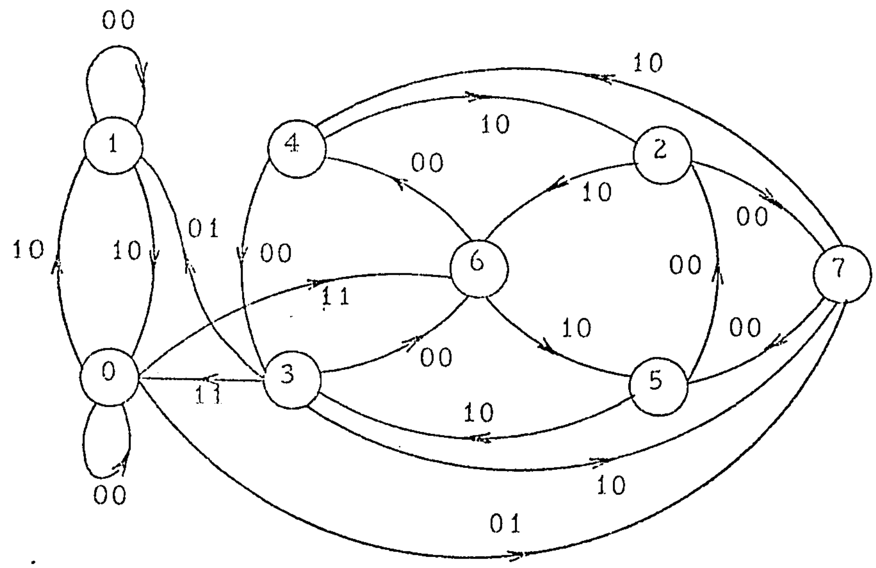


Figure 5.4 : State Diagram.

Note in Figures 5.2, 5.3 and 5.4 that the addition of $\underline{b}_{oi}$ for $i = 1, 2, 3,$ to our system introduced additional state transition lines which linked the sub-state diagrams together. Clearly these machines are strongly connected, hence controllable.

Example 5.2 Let the state equation of a linear sequential machine, defined over $GF = \{0, 1\}$ be:

$$x'(t) = \underline{A}\,\underline{x}(t) + \underline{B}\,\underline{\mu}(t) = \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix} \underline{x}(t) + \begin{bmatrix} 1 \\ 0 \end{bmatrix} \mu_1(t)$$

$\underline{y}(t) = y_1(t) = [1 \;\; 0]\, x(t)$ where $\underline{D} = \underline{0}$ and $\underline{A}$ is a non-derogatory matrix whose minimum polynomial is $m(x) = x(1+x)$. This system is unobservable and its state diagram is represented in Figure 5.5.

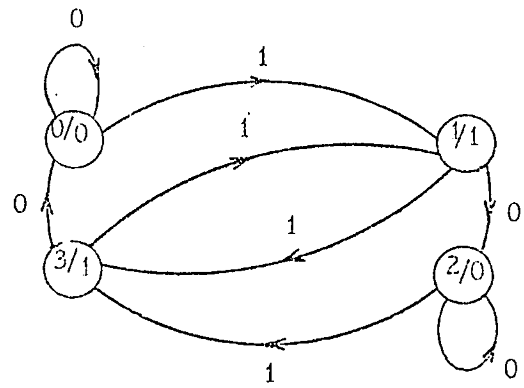


Figure 5. 5: State Diagram.

Since $\underline{D} = \underline{0}$ the output is associated with the state only and is independent of the input. Note that the system is either in states 0 or 2 if $y_1(t) = 0$ and either in states 1 or 3 if $y_1(t) = 1$. Hence it is impossible to determine the initial state of the system by examining the output sequence.

Let us go through the steps of the algorithm:

Step 1: The elementary divisors of $\underline{A}^T$ are $m_1(x) = x$ and $m_2(x) = 1 + x$. The vectors in the null space of $m_1(\underline{A}^T)$ are $\left\{\begin{matrix}0 & 1\\ 0 & 0\end{matrix}\right\}$. The vectors in the null space of $m_2(\underline{A}^T)$ are $\left\{\begin{matrix}0 & 1\\ 0 & 1\end{matrix}\right\}$.

Step 2: Choosing one element from every null space the resulting set of linear indépendent vectors are obtained.

$$S = \begin{pmatrix}0\\1\end{pmatrix}$$

Step 3: Thus the following $\underline{c}_o^T$ is a valid modification $\underline{c}_o^T = \begin{pmatrix}0\\1\end{pmatrix}$.

Let $\underline{C}^T = \left[\begin{array}{c|c}1 & \underline{c}_o^T\\ 0 & \end{array}\right] = \begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}$ and Figure 5. 6 represents the diagram of the resulting system.

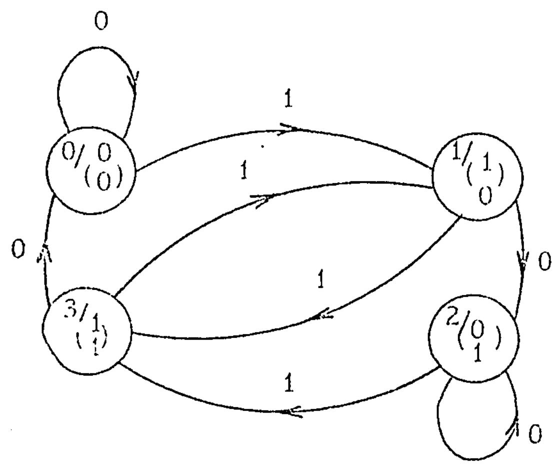


Figure 5.6: State Diagram.

Note that a unique set of outputs is associated with each state. The matrix $[\underline{C}^T, \underline{A}^T \underline{C}^T] = \left[\begin{array}{cc|cc} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \end{array}\right]$ obviously has rank 2. Hence the system is observable.

V-2: ALGORITHM FOR DEROGATORY SYSTEMS

Let $m(x) \neq p(x)$ be the minimum polynomial of $\underline{A}$, the matrix representation of the state transition $\beta: \mathcal{X} \longrightarrow \mathcal{X}$. Let $m(x) = m_1^{e_1}(x)\, m_2^{e_2}(x) \ldots \ldots m_r^{e_r}(x)$ such that the elementary divisors of $\underline{A}\,(\underline{A}^T)$ divisible by the monic, irreducible polynomial $m_i(x)$ are $m_i^{e_{i1}}(x)$, $m_i^{e_{i2}}(x), \ldots, m_i^{e_{i\ell_i}}(x)$ where $e_i = e_{i1} \geq e_{i2} \geq e_{i\ell_i}$ for $i = 1, 2, \ldots, r$. With no loss of generality, let $\ell_1 \geq \ell_2 \geq \ldots \geq \ell_r$. By theorem 3.6 the system is controllable if the matrix $\underline{B}$ contains "ℓ_1", linearly independent columns representing the "ℓ_1", $\underline{A}$-cyclic subspaces defined in the proof of the theorem. Similarly by theorem 3.12, the system is observable if the matrix $\underline{C}^T$ contains "ℓ_1", linearly independent columns representing the "ℓ_1" $\underline{A}^T$-cyclic subspaces defined in the proof of the theorem.

Thus the addition of "ℓ_1", linearly independent vectors, satisfying the required conditions, as columns of $\underline{B}$ yields a controllable system. Likewise, the addition of "ℓ_1", linearly independent vectors, satisfying the required conditions, as columns of $\underline{C}^T$ yields an observable system.

Therefore the algorithm for constructing a controllable (observable) system is as follows:

1. Find the null space of each elementary divisor of $\underline{A}$ ($\underline{A}^T$).

2. Choose from each null space a vector $\underline{v}_i$, such that the resulting set of vectors $\{\underline{v}_{11}, \underline{v}_{12}, \ldots, \underline{v}_{1\ell_1};$ $\underline{v}_{21}, \underline{v}_{22}, \ldots, \underline{v}_{2\ell_2}; \ldots; \underline{v}_{r1}, \underline{v}_{r2}, \ldots, \underline{v}_{r\ell_r}\}$ are linearly independent.

3. Add the vectors $\underline{v}_{i1}$ together to form the following set of vectors:
$\{\underline{v}_{11} + \underline{v}_{21} + \ldots + \underline{v}_{r1};\ \underline{v}_{12} + \underline{v}_{22} + \ldots + \underline{v}_{r2}; \ldots ; \underline{v}^*\}$ where the exact structure of $\underline{v}^*$ depends on the nature of the elementary divisors of the matrix $\underline{A}$.

4. This set of "ℓ_1" linearly independent vectors represents a valid modification.

The required null space of step 1 can be computed using the algorithm given in appendix B.

In this way it is always possible to transform a system, whose state transition matrix is derogatory, into a controllable and/or observable system by increasing respectively the number of inputs and/or outputs by "ℓ_1".

To see the required modification let the state equations of the uncontrollable and unobservable system S be:

$$\underline{x}'' = \underline{A}_{n\times n}\,\underline{x} + [\underline{b}_1 \underline{b}_2 \ldots \underline{b}_r] \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_r \end{bmatrix}$$

$$\underline{y} = [\underline{c}_1^T \; \underline{c}_2^T \ldots \underline{c}_p^T]^T \; \underline{x} + [\underline{d}_1 \; \underline{d}_2 \ldots \underline{d}_r] \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_r \end{bmatrix}$$

where $\underline{x}''$ denotes either $\dot{\underline{x}}$ or $\underline{x}'$, for continuous or discrete, systems respectively.

Let $\underline{b}_{01}\; \underline{b}_{02}, \ldots, b_{0\ell_1}$ be "ℓ_1" linearly independent vectors representing the "ℓ_1" $\underline{A}$-cyclic subspaces defined in theorem 3.6. Similarly let $\underline{c}_{01}^T, \underline{c}_{02}^T, \ldots, \underline{c}_{0\ell_1}^T$ be "ℓ_1" $\underline{A}^T$-cyclic subspaces defined in theorem 3.12. The addition of these elements to $\underline{B}$ and $\underline{C}^T$ respectively yields the following state equations for the corresponding controllable and observable system S' :

$$\underline{x}'' = \underline{A}_{nxn}\, \underline{x} + [\underline{b}_1\; \underline{b}_2 \ldots \underline{b}_r\; \underline{b}_{01}\; \underline{b}_{02} \ldots \underline{b}_{0\ell_1}] \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_r \\ \mu_{r+1} \\ \vdots \\ \mu_{r+\ell_1} \end{bmatrix}$$

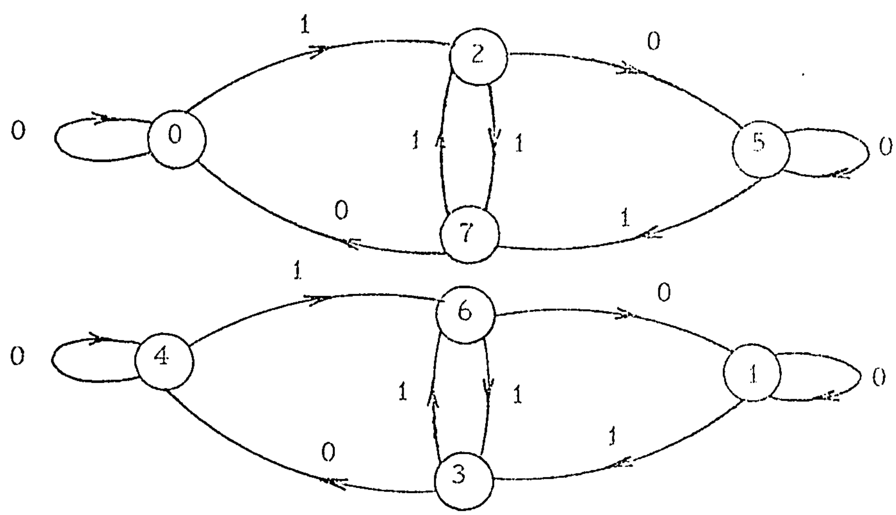
$$\underline{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_p \\ y_{p+1} \\ \vdots \\ y_{p+\ell_1} \end{bmatrix} = [\underline{c}_1^T \ldots \underline{c}_p^T\; \underline{c}_{01}^T \ldots \underline{c}_{0\ell_1}^T]^T \underline{x} + [\underline{d}_1\; \underline{d}_2 \ldots \underline{d}_r\; \underline{0}] \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_r \\ \mu_{r+1} \\ \vdots \\ \mu_{r+\ell_1} \end{bmatrix}$$

Consider the following example.

Example 5.3: Let the state equations of a linear sequential machine, defined over $GF = \{0, 1\}$, be:

$$\underline{x}'(t) = \underline{A}\,\underline{x}(t) + \underline{B}\,\underline{\mu}(t) = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix} \underline{x}(t) + \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \mu_1(t)$$

$\underline{y}(t) = \underline{C}\,\underline{x}(t) + \underline{D}\,\underline{\mu}(t)$ where $\underline{A}$ is a derogatory matrix whose minimum polynomial is $m(x) = x(1+x)$ and $p(x) = (1+x)\, m(x)$. This system is uncontrollable and its state diagram is represented in Figure 5. 7.



<u>Figure 5. 7:</u> State Diagram

Note that the state diagram is disconnected, since the set of states $\{0, 2, 5, 7\}$ and $\{1, 3, 4, 6\}$ form sub-state diagrams.

Let us go through the steps of the algorithm:

<u>Step 1:</u> The elementary divisors of $\underline{A}$ are $m_1(x) = x$, $m_2(x) = 1+x$ and $m_3(x) = 1+x$. The vectors in the null space of $m_1(\underline{A})$ are $\left\{\begin{matrix} 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{matrix}\right\}$.

The vectors in the null space of $m_2(\underline{A})$ and $m_3(\underline{A})$ are $\left\{\begin{matrix} 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{matrix}\right\}$.

<u>Step 2:</u> Choosing one element from every null space, the following sets of linearly independent vectors are obtained.

$$V_1 = \left\{\begin{matrix} 1 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 1 & 1 \end{matrix}\right\}; \quad V_2 = \left\{\begin{matrix} 1 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 1 & 1 \end{matrix}\right\}; \quad V_3 = \left\{\begin{matrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 1 & 0 & 1 \end{matrix}\right\};$$

$$V_4 = \left\{\begin{matrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \end{matrix}\right\}; \quad V_5 = \left\{\begin{matrix} 1 & 1 & 1 \\ 0 & 0 & 1 \\ 1 & 0 & 1 \end{matrix}\right\}; \quad V_6 = \left\{\begin{matrix} 1 & 0 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \end{matrix}\right\}$$

Step 3:

Adding the vectors within these sets as in step 3 of the algorithm yields the following sets of linearly independent vectors.

$$S_1 = \left\{\begin{matrix} 0 & 0 \\ 1 & 0 \\ 1 & 1 \end{matrix}\right\}; \quad S_2 = \left\{\begin{matrix} 0 & 1 \\ 1 & 0 \\ 1 & 1 \end{matrix}\right\}; \quad S_3 = \left\{\begin{matrix} 1 & 1 \\ 1 & 0 \\ 0 & 0 \end{matrix}\right\}$$

$$S_4 = \left\{\begin{matrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{matrix}\right\}; \quad S_5 = \left\{\begin{matrix} 0 & 1 \\ 1 & 0 \\ 0 & 0 \end{matrix}\right\}; \quad S_6 = \left\{\begin{matrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{matrix}\right\}$$

Let $\underline{B} = \left[\begin{array}{c|c} 0 & \\ 1 & \underline{b}_{01}, \underline{b}_{02} \\ 0 & \end{array}\right] = \left[\begin{array}{c|cc} 0 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{array}\right]$ and Figure 5.7

represents the state diagram of the resulting system.

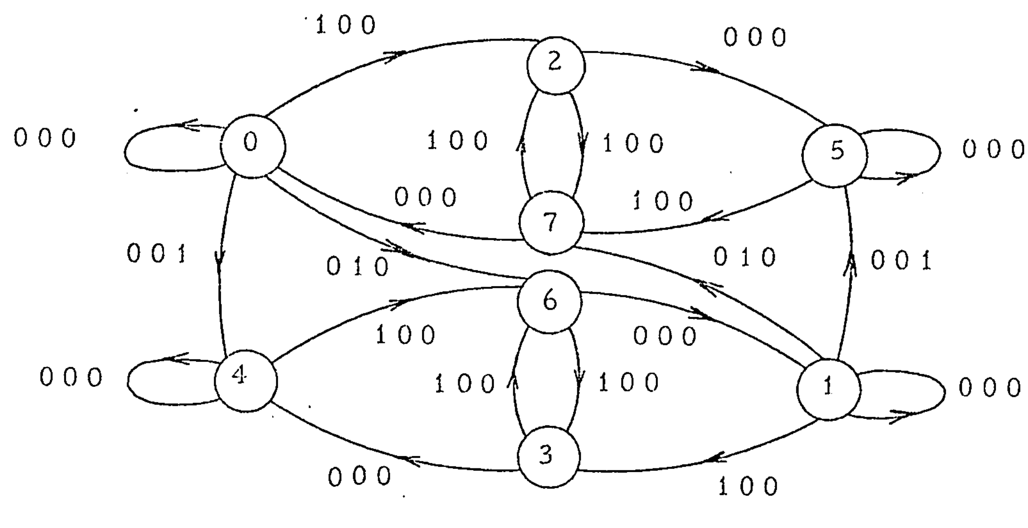


Figure 5.8 : State Diagram

Let $\underline{B} = \left[\begin{array}{c|c} 0 & \\ 1 & \underline{b}_{01}\ \underline{b}_{02} \\ 0 & \end{array}\right] = \left[\begin{array}{c|cc} 0 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{array}\right]$ and Figure 5.8

represents the state diagram of the resulting system.

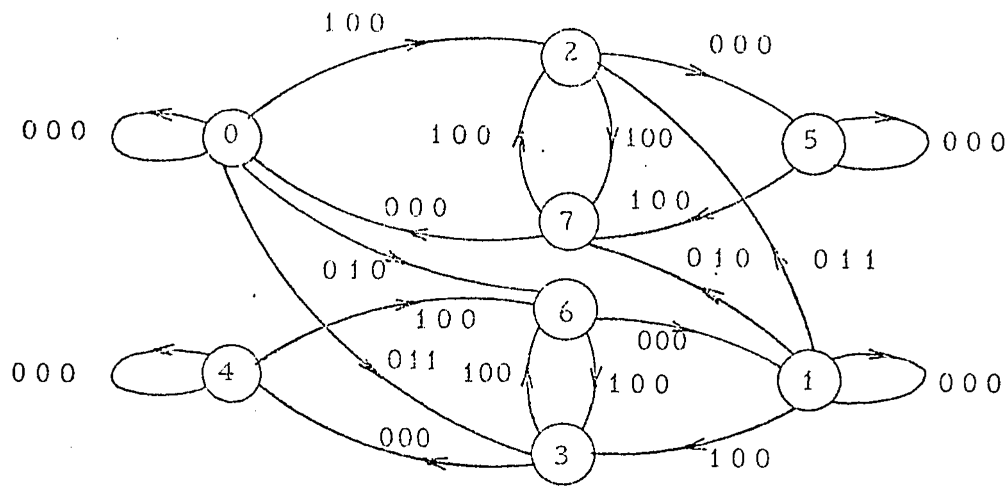


Figure 5.9: State Diagram

Note in figures 5.7 and 5.8 that the addition of two inputs to our system introduced additional state transition lines which linked the sub-state diagrams together. Clearly these machines are strongly connected, hence controllable.

Similarly it is possible to construct an observable system given an unobservable system whose state transition matrix is derogatory.

V-3: CONCLUSIONS

As shown in the previous sections a controllable and/or observable real, linear, time invariant system can always be synthesized from a given real linear, time invariant uncontrollable and/or unobservable system given its state transition matrix. It is noteworthy that the resulting controllable and/or observable system is NOT isomorphic to the original system. To illustrate this fact let us define Δ as a linear operator such that

$\Delta \ \underline{\bar{X}}(t) = \underline{x}(t)$ where $\underline{\bar{X}}(t)$ represents $\dot{\underline{x}}(t)$ or $\underline{x}'(t)$, for continuous or discrete systems respectively. Using this linear operator and signal flow graph techniques, the following well-known transfer function representation for real, linear time invariant systems is obtained:

$$\underline{Y}(\Delta) = [\underline{U}[\Delta^{-1}\underline{I} - \underline{A}]^{-1}\underline{B} + \underline{D}]\ \underline{U}(\Delta) = \underline{H}(\Delta)\,\underline{U}(\Delta)$$

The addition of inputs and outputs to the system modify the state input, state output and input-output transition matrices respectively as follows:

$$\underline{B}^{*} = [\, \underline{B} \mid \underline{B}_o \,] ; \quad \underline{C}^{*} = \left[\begin{array}{c} \underline{C} \\ \hline \underline{C}_o \end{array}\right] \quad \text{and} \quad \underline{D}^{*} = \left[\begin{array}{c|c} \underline{D} & \underline{0} \\ \hline \multicolumn{2}{c}{\underline{0}} \end{array}\right]$$

The transfer function of the resultant controllable and observable system is:

$$\underline{Y}^{*}(\Delta) \quad \underline{Y} \quad = \quad [\, \underline{C}^{*} [\Delta^{-1} \underline{I} - \underline{A}]^{-1} \underline{B}^{*} + \underline{D}^{*} \,] \, \underline{U}^{*}(\Delta) = \underline{H}^{*}(\Delta) \underline{U}^{*}(\Delta)$$

$$\underline{Y}^{*}(\Delta) = \begin{bmatrix} \underline{Y}(\Delta) \\ \underline{Y}_o(\Delta) \end{bmatrix} = \left[\left[\begin{array}{c} \underline{C} \\ \hline \underline{C}_o \end{array}\right] [\Delta^{-1} \underline{I} - \underline{A}]^{-1} [\underline{B} \mid \underline{B}_o] + \left[\begin{array}{c|c} \underline{D} & \underline{0} \\ \hline \multicolumn{2}{c}{\underline{0}} \end{array}\right] \right] \begin{bmatrix} \underline{U}(\Delta) \\ \underline{U}_o(\Delta) \end{bmatrix}$$

$$= \left[\begin{array}{c|c} \underline{C} [\Delta^{-1} \underline{I} - \underline{A}]^{-1} \underline{B} + \underline{D} & \underline{C} [\Delta^{-1} \underline{I} - \underline{A}]^{-1} \underline{B}_o \\ \hline \underline{C}_o [\Delta^{-1} \underline{I} - \underline{A}]^{-1} \underline{B} & \underline{C}_o [\Delta^{-1} \underline{I} - \underline{A}]^{-1} \underline{B}_o \end{array}\right] \begin{bmatrix} \underline{U}(\Delta) \\ \\ \underline{U}_o(\Delta) \end{bmatrix}$$

For continuous systems $\Delta^{-1} = S$, the Laplace operator, whereas for discrete systems $\Delta = D$, the Delay operator. Clearly the two systems have different input-output behaviours, hence are <u>not</u> isomorphic, since $\underline{H}^{*}(\Delta) \neq \underline{H}(\Delta)$. However, if $\underline{u}_o(t) = \underline{0}$ and the additional outputs $\underline{y}_o(t)$ are disregarded then

$$\underline{Y}(\Delta) \quad = \quad [\, \underline{C} [\Delta^{-1} \underline{I} - \underline{A}]^{-1} \underline{B} + \underline{D} \,] \, \underline{U}(\Delta) = \underline{H}(\Delta) \underline{\mu}(\Delta) .$$

Under this condition the two systems are isomorphic.

Thus in practical application such as identification, initial machine installation and machine diagnosis the additional inputs may take on nonzero values. During this phase, the additional outputs convey information about the initial state of the system. During normal system operation, the additional inputs are set to zero and the additional outputs are disregarded.

CHAPTER VI

CONCLUSIONS AND SUGGESTIONS FOR FURTHER RESEARCH

The contributions of this thesis are in the areas of controllability and observability of linear systems, modular synthesis of linear sequential machines and numerical methods in similarity transformations.

In terms of controllability and observability a new approach, based on cyclic subspaces, was developed. This approach is based on the study of the properties of the state transition matrix $\underline{A}$. The method involves the partitioning of the state space $\mathcal{X}$ into a number of $\underline{A}$-cyclic or $\underline{A}^T$-cyclic subspaces. Existing methods determine controllability and observability by considering the rank of large combined matrices. The approach presented in this thesis investigates controllability and/or observability by determining whether the columns of $\underline{B}$ and/or $\underline{C}^T$ are generators of the $\underline{A}$-cyclic and/or $\underline{A}^T$-cyclic subspaces respectively. This approach provides physical insight into these concepts since it considers controllability and observability from the point of view of being able to access different cyclic subspaces.

No attempt was made to extend these results to nonlinear and/or time-varying systems. This is left as a topic for further research.

In terms of modular synthesis, the shift register synthesis of sequential machines was investigated. The algorithm is based on finding a suitable similarity transformation matrix that will lead to a companion matrix form. The result was a parallel shift register realization. However using the results of Appendix A it is quite easy to modify the algorithm to obtain a single or a serially comected shift register realization. Only the companion matrix form of the state space matrix was dealt with, since the author was mainly interested in shift register realizations. A consideration of other canonical forms may lead to other interesting forms of modular synthesis of linear sequential machines. There is reason to believe that all controllable and observable linear sequential machines are shift register realizable. Also in view of the results for single shift register realization, it seems that all single input, single output linear sequential machines are single shift register realizable. Hence the following question arises: Given a linear sequential machine which is not realizable using shift registers; is it possible, by splitting the states and/or adding more inputs or outputs, to obtain a shift register realization ? These thoughts are left as topics for further research.

In terms of numerical methods in similarity transformations, Appendix A contains several algorithms that construct a similarity transformation matrix that enables the conversion from one canonical form to another. Nevertheless the exact number of such similarity transformation matrices remains nebulous. This is left as a topic for further research.

In this thesis it was assumed that the linear sequential machine was given in state space representation. Thus a topic for further research is the obtention of a generalized procedure yielding the state space representation of a linear sequential machine from the transition table. Another interesting topic is to relate the characteristics of the state space representation to the decomposition of linear sequential machines. A recent paper in this area is by Gallaire and Harrison[9].

APPENDIX A

SOME NUMERICAL RESULTS IN SIMILARITY TRANSFORMATIONS

This appendix includes a set of numerical algorithms for forming the transformation matrix $\underline{M}$ in a similarity transformation for the following cases:

1. Transform the companion matrix to the transpose of the companion matrix or vice versa.
2. Transform the hypercompanion matrix to the transpose of the hypercompanion matrix or vice versa.
3. Transform the companion matrix to the transpose of the hypercompanion matrix or vice versa.

Lemma A.1: An n-square matrix $\underline{A}$ is similar to its transpose $\underline{A}^T$.

Proof:

Let $x\underline{I} - \underline{A}$ be the characteristic matrix of $\underline{A}$. By a series of row and column operations $(x\underline{I} - \underline{A})$ can be reduced to a diagonal matrix $\underline{D} = \underline{\text{Diag}}\ (f_1(x)\ f_2(x) \dots f_n(x))$ known as the Smith Normal Form of $\underline{A}$ where $f_i(x)$ for $i = 1, 2, \dots, n$ are the invariant factors (similarity invariants) of $\underline{A}$ such that $f_i(x) / f_{i+1}(x)$ for $i = 1, 2, \dots, n-1$. The row and column operations are achieved by pre and post multiplying $(x\underline{I} - \underline{A})$ by carefully chosen non-singular matrices $\underline{P}$ and $\underline{Q}$ respectively.

Hence $\underline{P}(x)\,(x\underline{I} - \underline{A})\,\underline{Q}(x) = \underline{\text{Diag}}\ (f_1(x)\ f_2(x) \dots f_n(x)) = \underline{D}$

Taking the transpose of the above expression:

$$\begin{aligned}\underline{D}^T = \underline{\text{Diag}}\ (f_1(x)\, f_2(x) \dots f_n(x))^T &= [\,\underline{P}(x)(x\underline{I} - \underline{A})\,\underline{Q}(x)\,]^T \\ &= [\,(x\underline{I} - \underline{A})\,\underline{Q}(x)\,]^T\,\underline{P}(x)^T \\ &= \underline{Q}(x)^T\,(x\underline{I} - \underline{A}^T)\,\underline{P}(x)^T\end{aligned}$$

$\underline{Q}(x)^T$ and $\underline{P}(x)^T$ are non-singular matrices since $\left|\underline{Q}(x)\right| = \left|\underline{Q}(x)^T\right| \neq 0$ and $\left|\underline{P}(x)\right| = \left|\underline{P}(x)^T\right| \neq 0$. Since the characteristic matrix of $\underline{A}^T$ is reducible to the same Smith Normal Form then $\underline{A}$ and $\underline{A}^T$ are similar.

<u>Theorem A.1:</u> Let $\underline{A}_c$ be the companion matrix of $p(x) = m(x) = x^n - a_{n-1}x^{n-1} - \ldots, - a_1x - a_o$ a monic, polynomial over F. Since $\underline{A}_c$ and $\underline{A}_c^T$ are similar then there exists a non-singular matrix $\underline{M} = [m_{ij}]_{nxn}$ such that $\underline{A}_c = \underline{M}\,\underline{A}_c^T\,\underline{M}^{-1}$. Let us describe an algorithm to construct $\underline{M}$ given $\underline{A}_c^T$.

Since $\underline{A}_c = \underline{M}\,\underline{A}_c^T\,\underline{M}^{-1}$ then $\underline{M}\,\underline{A}_c^T = \underline{A}_c\,\underline{M}$

Hence

$$\begin{bmatrix} m_{11} & \cdots\cdots & m_{1n} \\ m_{21} & \ddots & \vdots \\ \vdots & & \vdots \\ m_{n1} & \cdots\cdots & m_{nn} \end{bmatrix} \begin{bmatrix} 0 & \cdots\cdots & & a_o \\ 1 & 0 & & a_1 \\ 0 & 1 & & \vdots \\ & \underline{0} & \ddots & \vdots \\ & & 1 & a_{n-1} \end{bmatrix} = \begin{bmatrix} 0 & 1 & & & \\ 0 & 0 & 1 & & \underline{0} \\ \vdots & & & \ddots & \\ \vdots & & & & 1 \\ a_o & a_1 & \cdots\cdots & & a_{n-1} \end{bmatrix} \begin{bmatrix} m_{11} & \cdots\cdots & m_{1n} \\ m_{21} & & \vdots \\ \vdots & & \vdots \\ m_{n1} & \cdots\cdots & m_{nn} \end{bmatrix}$$

$$\begin{bmatrix} m_{12} & m_{13} & \cdots m_{1n} & a_om_{11}+\ldots+a_{n-1}m_{1n} \\ m_{22} & m_{23} & \cdots m_{2n} & a_om_{21}+\ldots+a_{n-1}m_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ & \ddots & & \\ m_{n2} & m_{n3} & \cdots m_{nn} & a_om_{n1}+\ldots+a_{n-1}m_{nn} \end{bmatrix} = \begin{bmatrix} m_{21} & m_{22} & \cdots\cdots & m_{2n} \\ m_{31} & m_{32} & \cdots\cdots & m_{3n} \\ \vdots & \vdots & \ddots & \vdots \\ m_{n1} & m_{n2} & \cdots\cdots & m_{nn} \\ (a_om_{11}+\ldots+a_{n-1}m_{n1}) & \cdots & & (a_om_{1n}+\ldots+a_{n-1}m_{nn}) \end{bmatrix}$$

Comparing corresponding matrix entries we obtain the following set of independent equations :

$$m_{i+1,n} = \sum_{j=1}^{n} - a_{j-1}\, m_{ij} \qquad \text{for } i = 1,2,\ldots, n-1$$

$$m_{i_1j_1} = m_{i_2j_2} \quad \text{if } i_1 + j_1 = i_2 + j_2 \quad \text{for } i,j = 1,2,\ldots, n.$$

Hence the sum of the matrix entry subscripts defines an equivalence relation S on the set of matrix entries.

For more insight into the entries of $\underline{M}$, as developed in Theorem A. 1, see Figure A. 1.

$$\begin{bmatrix} m_{11} & m_{12} & \cdots & m_{1n} \\ m_{21} & & & m_{2n} \\ \vdots & & & \vdots \\ & & & m_{n-1,n} \\ m_{n1} & \cdots & m_{n,n-1} & m_{nn} \end{bmatrix} \begin{array}{l} \\ \rightarrow m_{2n} = f_2(m_{11}, m_{12}, \ldots, m_{1n}) \\ \quad\vdots \\ \\ \rightarrow m_{nn} = f_n(m_{n-1,1}, m_{n-1,2}, \ldots, m_{n-1,n}) \end{array}$$

<u>Figure A. 1:</u> Schematic Diagram of the Entries of $\underline{M}$.

All the matrix entries on the same diagonal are equal. Upon determining the first row entries of $\underline{M}$ the value of the remaining entries may be calculated as in the above theorem. Thus if $\underline{M}$ is defined over a finite field F of q elements then there exist <u>at most</u> $(q^n - 1)$ $\underline{M}$ matrices such that $\underline{A}_c = \underline{M}\,\underline{A}_c^T\,\underline{M}^{-1}$.

Consider the following examples.

<u>Example A. 1:</u> Let $\underline{A}_c$ be the companion matrix of a monic polynomial $m(x) = x^4 + 2x^3 - x^2 + 3x$ defined over R.

Let $(m_{11}\ m_{12}\ m_{13}\ m_{14}) = (0\ 0\ 0\ 1)$. By the previous theorem

$$M = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & -2 \\ 0 & 1 & -2 & 5 \\ 1 & -2 & 5 & -15 \end{bmatrix} \quad \text{and} \quad |\underline{M}| = 1 \quad \therefore \ \underline{M}^{-1} \text{ exists. After}$$

calculation

$$\underline{M}^{-1} = \begin{bmatrix} 3 & -1 & 2 & 1 \\ -1 & 2 & 1 & 0 \\ 2 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}$$

$$\underline{M}\,\underline{A}_c^T\,\underline{M}^{-1} = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & -2 \\ 0 & 1 & -2 & 5 \\ 1 & -2 & 5 & -15 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & -3 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & -2 \end{bmatrix} \begin{bmatrix} 3 & -1 & 2 & 1 \\ -1 & 2 & 1 & 0 \\ 2 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & -3 & 1 & -2 \end{bmatrix} = \underline{A}_c$$

Example A.2: Let $\underline{A}_c$ be the companion matrix of a monic polynomial $m(x) = x^3 + x^2 + 2x + 2$ defined over $GF(3) = \{0, 1, 2\}$. Let $(m_{11}\ m_{12}\ m_{13}) = (1\ 0\ 0)$. By previous theorem

$$\underline{M} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 2 \end{bmatrix}$$ and $\underline{M} = 2$. $\therefore$ $\underline{M}^{-1}$ exists. After calculation,

$$\underline{M}^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix}.$$ Upon verification,

$$\underline{M}\,\underline{A}_c^T\,\underline{M}^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 2 \end{bmatrix} = \underline{A}_c$$

Theorem A.2: Let $\underline{A}_H$ be the hypercompanion matrix of $p(x) = m(x) = f(x)^e = (x^r + a_{r-1}x^{r-1} + \dots + a_1 x + a_o)^e$ where $f(x)$ is a monic irreducible polynomial over F. Since $\underline{A}_H^T$ and $\underline{A}_H$ are similar then there exists a non-singular matrix $\underline{M} = [m_{ij}]_{nxn}$ such that

$\underline{A}_H = \underline{M}\,\underline{A}_H^T\,\underline{M}^{-1}$. Let us describe an algorithm to construct $\underline{M}$ given $\underline{A}_H^T$.

Since $\underline{A}_H = \underline{M}\,\underline{A}_H^T\,\underline{M}^{-1}$ then $\underline{M}\underline{A}_H^T = \underline{A}_H\underline{M}$. Recall that

$$\underline{A}_H = \begin{bmatrix} \underline{C} & \underline{N}^T & & & \\ & \underline{C} & \underline{N}^T & & \underline{0} \\ & & \ddots & \ddots & \\ & \underline{0} & & \underline{C} & \underline{N}^T \\ & & & & \underline{C} \end{bmatrix} \qquad \text{where } \underline{N}^T = \begin{bmatrix} 0 & & & \\ \vdots & \ddots & & \underline{0} \\ \vdots & & \ddots & \\ 1 & \dots & & 0 \end{bmatrix}$$

and $\underline{C}$ is the companion matrix of $f(x)$.

Hence,

$$\begin{bmatrix} m_{11} & \dots & m_{1n} \\ \vdots & \ddots & \vdots \\ m_{n1} & \dots & m_{nn} \end{bmatrix} \begin{bmatrix} \underline{C} & \underline{N}^T & & \\ & \underline{C} & \ddots & \\ & & \ddots & \\ & & \underline{C} & \underline{N}^T \\ & & & \underline{C} \end{bmatrix}^T = \begin{bmatrix} \underline{C} & \underline{N}^T & & \\ & \underline{C} & \underline{N}^T & \\ & & \ddots & \ddots \\ & & \underline{C} & \underline{N}^T \\ & & & \underline{C} \end{bmatrix} \begin{bmatrix} m_{11} & \dots & m_{1n} \\ \vdots & \ddots & \vdots \\ m_{n1} & \dots & m_{nn} \end{bmatrix}$$

Upon multiplying the matrices and carefully comparing corresponding matrix entries; we obtain the following set of independent equations.

If

$$\underline{M}_{nxn} = \begin{bmatrix} \underline{M}_{11} & \dots & \underline{M}_{1e} \\ \underline{M}_{21} & \ddots & \vdots \\ \vdots & & \vdots \\ \underline{M}_{e1} & \dots & \underline{M}_{ee} \end{bmatrix} \qquad \text{where the}$$

$\underline{M}_{k\ell} = [\,m_{ij}\,]_{rxr}$ for $i,j = 1,2,\dots\dots, r$; $k, \ell = 1,2,\dots\dots, e$

Then i - $\underline{M}_{k1,\ell 1} = \underline{M}_{k2,\ell 2}$ if $k_1+\ell_1 = k_2+\ell_2$ for $k,\ell = 1,2,\dots\dots,e$

ii - $\underline{M}_{k\ell} = \underline{0}$ if $k+\ell \geq e+2$

iii - $m_{i+1,r}(k,\ell) = \sum_{j=1}^{r} - a_{j-1}\, m_{ij}(k,\ell) + m_{i,1}(k,\ell+1)$

For $i = 1,2,\dots\dots, r-1$; $k = 1,2,\dots\dots, e$; $\ell = 1,2,\dots\dots, e-1$

iv - $m_{i+1,r}(k,e) = \sum_{j=1}^{r} -a_{j-1}\, m_{ij}\,(k,e)$.

for $i = 1, 2, \ldots, r$; $k = 1, 2, \ldots, e$.

v. $m_{i_1, j_1}(k, \ell) = m_{i_2 j_2}(k, \ell)$ if $i_1 + j_1 = i_2 + j_2$ for $i, j = 1, 2, \ldots, r$.

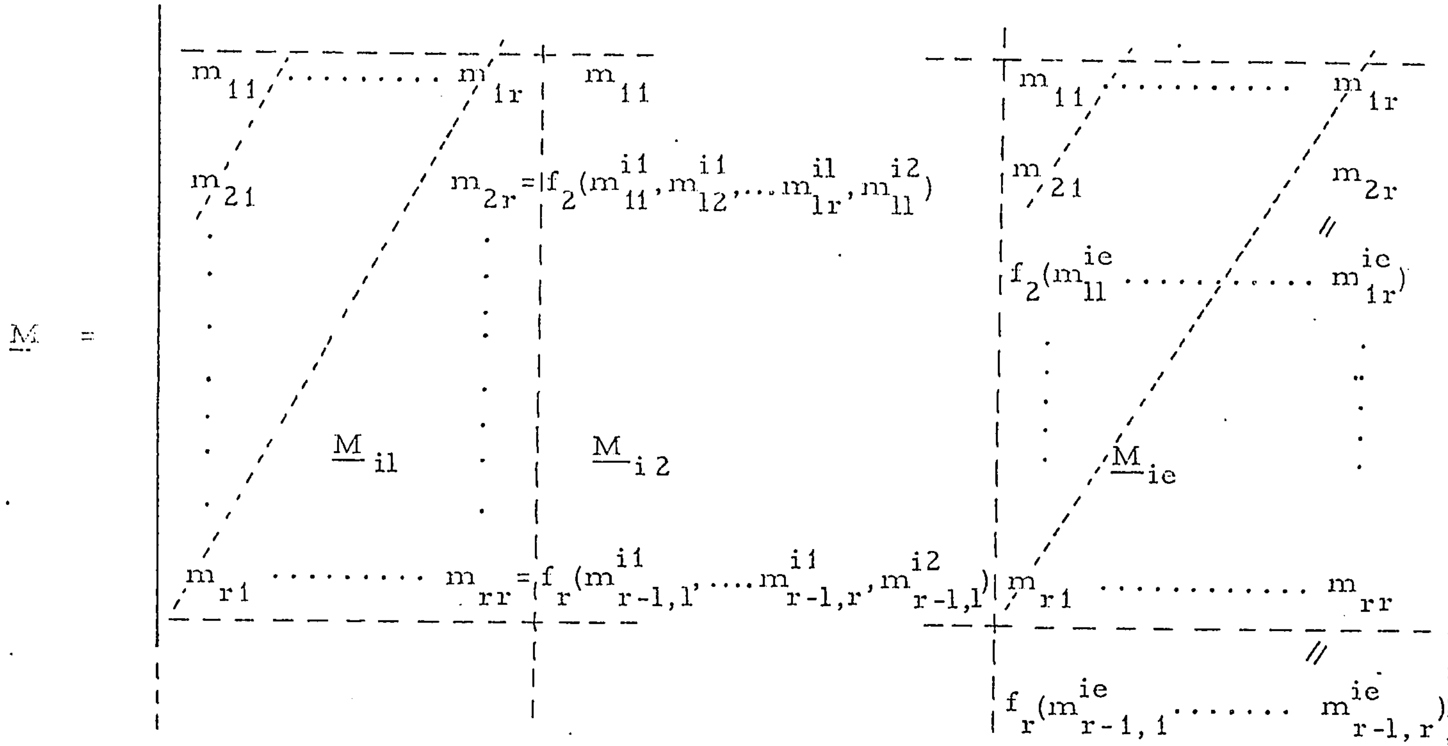
Note that the sum of the submatrix subscripts defines an equivalence relation S_1 on the set of submatrices $\underline{M}_{k\ell}$. Upon considering a submatrix, $\underline{M}_{k\ell} = [m_{ij}]_{r \times r}$, the sum of the submatrix entry subscripts defines an equivalence relation S_2 on the set of submatrix entries.

Upon assigning the first row entries of $\underline{M}$ the value of the remaining entries may be calculated using the above equations. If $\underline{M}$ is defined over a finite field F of q elements then there exist at <u>most</u> $(q^n - 1)$ $\underline{M}$ matrices such that $\underline{A}_H = \underline{M}\,\underline{A}_H^T\,\underline{M}^{-1}$.

For more insight into the entries of $\underline{M}$, as developed in Theorem A.2, see Figure A.2.

$$\underline{M} = \begin{bmatrix} \underline{M}_{11} & \underline{M}_{12} & & \underline{M}_{1e} \\ \underline{M}_{21} & & & \underline{0} \\ & & \underline{0} & \vdots \\ \underline{M}_{e1} & \underline{0} & \cdots & \underline{0} \end{bmatrix}$$

All the submatrices on the same diagonal are equal. Consider an enlarged section of $\underline{M}$.



Within a submatrix $\underline{M}_{k\ell}$ all the elements on the same diagonal are equal.

Figure A.2 : Schematic Diagram of the entries of M.

Consider the following examples:

Example A.3: Let $\underline{A}_H$ be the hypercompanion matrix of $m(x) = (x^2 + 2x + 2)^2$ where $(x^2 + 2x + 2)$ is a monic, irreducible polynomial over the field of real numbers R. Arbitrarily set $(m_{11}\ m_{12}\ m_{13}\ m_{14}) = (1\ 1\ 1\ 1)$.
By theorem A.2, $\underline{M}_{11} = \underline{M}_{21}$, $\underline{M}_{22} = \underline{0}$.

For $\underline{M}_{11}$: $m_{12} = m_{21}$, $\underline{m}_{22} = a_o\ m_{11} - a_1\ m_{12} + m_{13} = -3$

For $\underline{M}_{12}$ $m_{23} = m_{14}$; $m_{24} = -a_o\ m_{13} - a_1\ m_{14} = -4$

$$\therefore \underline{M} = \begin{bmatrix} \underline{M}_{11} & \underline{M}_{12} \\ \underline{M}_{21} & \underline{M}_{22} \end{bmatrix} = \left[\begin{array}{cc|cc} 1 & 1 & 1 & 1 \\ 1 & -3 & 1 & -4 \\ \hline 1 & 1 & \multicolumn{2}{c}{\underline{0}} \\ 1 & -4 & & \end{array}\right] \quad \text{and} \quad |\underline{M}| = 25.$$

If the $|\underline{M}| = 0$ then choose a different first row vector and repeat the calculations as in theorem A. 2. In this example $|\underline{M}| \neq 0$ and $\underline{M}^{-1}$ exists.

$$\underline{M}^{-1} = \left[\begin{array}{cc|cc} 0 & 0 & 4/5 & 1/5 \\ 0 & 0 & 1/5 & -1/5 \\ \hline 4/5 & 1/5 & -21/25 & -4/25 \\ 1/5 & -1/5 & -4/25 & 4/25 \end{array}\right]$$

Upon verification,

$$\underline{M}\,\underline{A}_H^T\,\underline{M}^{-1} = \left[\begin{array}{cc|cc} 1 & 1 & 1 & 1 \\ 1 & -3 & 1 & -4 \\ \hline 1 & 1 & 0 & 0 \\ 1 & -4 & 0 & 0 \end{array}\right] \left[\begin{array}{cc|cc} 0 & -2 & 0 & 0 \\ 1 & -2 & 0 & 0 \\ \hline 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & -2 \end{array}\right] \left[\begin{array}{cc|cc} 0 & 0 & 4/5 & 1/5 \\ 0 & 0 & 1/5 & -1/5 \\ \hline 4/5 & 1/5 & -21/25 & -4/25 \\ 1/5 & -1/5 & -4/25 & 4/25 \end{array}\right] = \left[\begin{array}{cc|cc} 0 & 1 & 0 & 0 \\ -2 & -2 & 1 & 0 \\ \hline 0 & 0 & 0 & 1 \\ 0 & 0 & -2 & -2 \end{array}\right] = \underline{A}_H$$

<u>Example A. 4</u>: Let $\underline{A}_H$ be the hypercompanion matrix of $m(x) = (x+2)^3$ where $(x+2)$ is a monic, irreducible polynomial over R. Arbitrarily set $(m_{11}\ m_{12}\ m_{13}) = (0\ 0\ 1)$.

By Theorem A. 2

$$\underline{M} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \qquad \text{and} \quad |\underline{M}| = 1 \ \text{ such that } \ \underline{M} = \underline{M}^{-1}.$$

Upon verification,

$$\underline{M}\,\underline{A}_H^T\,\underline{M}^{-1} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} -2 & 0 & 0 \\ 1 & -2 & 0 \\ 0 & 1 & -2 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} = \begin{bmatrix} -2 & 1 & 0 \\ 0 & -2 & 1 \\ 0 & 0 & -2 \end{bmatrix} = \underline{A}_H$$

<u>Example A. 5:</u> Let $\underline{A}_H$ be the hypercompanion matrix of $m(x) = (x+1)^3$ where $(x+1)$ is a monic, irreducible polynomial over $GF(2) = \{0, 1\}$.

Let $(m_{11}\ m_{12}\ m_{13}) = (0\ 1\ 0)$. By Theorem A. 2

$$\underline{M} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad \text{and} \quad |\underline{M}| = 0. \quad \text{Thus } \underline{M}^{-1} \text{ doesn't exist.}$$ Here is

an example where the selected first row <u>fails</u> to yield a suitable $\underline{M}$.

Let $(m_{11} \; m_{12} \; m_{13}) = (1 \; 0 \; 1)$. In this case,

$$\underline{M} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \quad \text{and} \quad |\underline{M}| = 1. \quad \text{Thus } \underline{M}^{-1} \text{ exists.}$$

After calculation $$\underline{M}^{-1} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$$

Upon verification,

$$\underline{M}\underline{A}_H^T \underline{M}^{-1} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} = \underline{A}_H$$

<u>Example A. 6:</u> Let $\underline{A}_H$ be the hypercompanion matrix of $m(x) = (x^2 + x + 1)^2$ where $(x^2 + x + 1)$ is a monic, irreducible polynomial over $GF(2) = \{0, 1\}$.

Let $(m_{11} \; m_{12} \; m_{13} \; m_{14}) = (0 \; 1 \; 0 \; 1)$. By Theorem A. 2

$$\underline{M} = \left[\begin{array}{cc|cc} 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 \\ \hline 0 & 1 & \multicolumn{2}{c}{\multirow{2}{*}{$\underline{0}$}} \\ 1 & 1 & & \end{array}\right] \quad \text{and} \quad |\underline{M}| = 1.$$

After calculation

$$\underline{M}^{-1} = \left[\begin{array}{cc|cc} \multicolumn{2}{c|}{\multirow{2}{*}{$\underline{0}$}} & 1 & 1 \\ & & 1 & 0 \\ \hline 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 \end{array}\right]$$

Upon verification,

$$\underline{M}\,\underline{A}_H^T\,\underline{M}^{-1} = \left[\begin{array}{cc|cc} 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 \\ \hline 0 & 1 & \multicolumn{2}{c}{\underline{0}} \\ 1 & 1 & & \end{array}\right] \left[\begin{array}{cc|cc} 0 & 1 & \multicolumn{2}{c}{\underline{0}} \\ 1 & 1 & & \\ \hline 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{array}\right] \left[\begin{array}{cc|cc} \multicolumn{2}{c|}{\underline{0}} & 1 & 1 \\ & & 1 & 0 \\ \hline 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 \end{array}\right] = \left[\begin{array}{cc|cc} 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ \hline \multicolumn{2}{c|}{\underline{0}} & 0 & 1 \\ & & 1 & 1 \end{array}\right] = \underline{A}_H$$

Let $\underline{A}_H$ be the matrix representation of $\beta : \mathcal{V} \to \mathcal{V}$ such that $p(x) = m(x) = (x-\lambda)^n$. In control $\underline{A}_H$ is generally denoted $\underline{J}$ and called " The Jordan Canonical Form" of all matrices similar to $\underline{A}_H$. By definition, λ is an eigenvalue of the linear operator β if and only if $\underline{A}_H \cdot \underline{v} = \lambda \cdot \underline{v}$ for some $\underline{v} \neq o, \underline{v} \in \mathcal{V}, \lambda \in F$. $\underline{A}_H$ is often called an n^{th} order Jordan Block of λ .

<u>Theorem A. 3:</u> If $\underline{J}$ is the Jordan Canonical Form of $p(x) = m(x) = (x-\lambda)^n$, where $x - \lambda$ is a monic, irreducible polynomial over R, then there exists a matrix

$$\underline{M} = \begin{bmatrix} \underline{0} & & 1 \\ & \diagup & \\ 1 & & \underline{0} \end{bmatrix}_{nxn} = \underline{M}^{-1} \quad \text{such that} \quad \underline{M}\,\underline{J}\,\underline{M}^{-1} = \underline{J}^T.$$

<u>Proof:</u>

Let $(m_{11}\; m_{12} \ldots . m_{1n}) = (0\; 0 \ldots .\; 0\; 1)$.

By theorem A. 2,

$$\underline{M} = \begin{bmatrix} \underline{0} & & 1 \\ & \diagup & \\ 1 & & \underline{0} \end{bmatrix}.$$

Since $\underline{M}\;\underline{M} = \underline{I}$, then $\underline{M} = \underline{M}^{-1}$. Since $\underline{M}$ is non-singular then it satisfies the conditions of theorem A. 2 such that $\underline{M}\;\underline{J}\;\underline{M}^{-1} = \underline{J}^T$. The determinant of $\underline{M}$ depends on the dimension of $\underline{M}$ such that if n is odd $n = 2k + 1$ or even $n = 2k$ where $k = 0, 1, 2, \ldots$ then

$$|\underline{M}| = 1 \qquad \text{for} \quad k = 0, 2, 4, \ldots .$$

$$|\underline{M}| = -1 \qquad \text{for} \quad k = 1, 3, 5, \ldots .$$

In chapter II we showed that the definition of a new basis over $\mathcal{V}$ resulted in a different matrix representation for $\beta: \mathcal{V} \to \mathcal{V}$.

Let $\underline{J}$ be the matrix representation of $\beta: \mathcal{V} \to \mathcal{V}$ with respect to the basis $\{\underline{e}_1 \underline{e}_2 \cdots\cdots \underline{e}_n\}$ and $\underline{J}^T$ be the matrix representation of β with respect to the basis $\{\underline{v}_1 \underline{v}_2 \cdots\cdot \underline{v}_n\}$. By Lemma 2.1 $\underline{J}^T = \underline{M}\ \underline{J}\ \underline{M}^{-1}$ where $[\underline{e}_1 \underline{e}_2 \cdots\cdot \underline{e}_n] = [\underline{v}_1 \underline{v}_2 \cdots\cdot \underline{v}_n]\ \underline{M}$. Let $\underline{M}$ be the matrix defined in theorem A.3. Thus $\underline{e}_1 = \underline{v}_n$; $\underline{e}_2 = \underline{v}_{n-1}; \ldots ; \underline{e}_n = \underline{v}_1$. Hence the set of basis vectors is the <u>same</u> except that the basis vectors are now taken in the <u>reverse</u> order, namely

$$\{\underline{e}_n, \ \underline{e}_{n-1}, \ldots, e_2, \ \underline{e}_1\}.$$

<u>Theorem A.4:</u> Let $\underline{A}_H$ be the hypercompanion matrix of $p(x) = m(x) = (x^r + a_{r-1}x^{n-1} + \ldots + a_1 x + a_o)^e = (f(x))^e$ where $f(x)$ is a monic, irreducible polynomial over F and $\underline{A}_c$ be the companion matrix of $m(x)$. Since $\underline{A}_c$ and $\underline{A}_H^T$ are similar then there exists a non-singular matrix $\underline{M} = [m_{ij}]_{nxn}$ such that $\underline{A}_c = \underline{M}\ \underline{A}_H^T\ \underline{M}^{-1}$. Let us describe an algorithm to construct $\underline{M}$.

Since $\underline{A}_c = \underline{M}\ \underline{A}_H^T \underline{M}^{-1}$ then $\underline{M}\ \underline{A}_H^T = \underline{A}_c \underline{M}$. Let $f(x)^e = b_o + b_1 x + \ldots + b_{n-1} x^{n-1} + x^n$. Thus

$$\begin{bmatrix} m_{11} & \cdots & m_{1n} \\ m_{21} & \ddots & \vdots \\ \vdots & & \vdots \\ m_{n1} & & m_{nn} \end{bmatrix} \begin{bmatrix} \underline{C}^T & & & \\ \underline{N} & \underline{C}^T & & \underline{0} \\ & \ddots & \ddots & \\ \underline{0} & & \ddots \ \ \underline{C}^T & \\ & & \underline{N}\,\underline{C}^T \end{bmatrix} = \begin{bmatrix} 0 & 1 & & \underline{0} \\ & \ddots & \ddots & \\ \underline{0} & & 0 & 1 \\ -b_o & -b_1 & \cdots & -b_{n-1} \end{bmatrix} \begin{bmatrix} m_{11} & \cdots & m_{1n} \\ m_{21} & \ddots & \vdots \\ \vdots & & \vdots \\ m_{n1} & \cdots & m_{nn} \end{bmatrix}$$

where $\underline{C}$ is the companion matrix of $f(x)$ and $N = \begin{bmatrix} 0 & \cdots & 1 \\ & \ddots & \vdots \\ \underline{0} & & 0 \end{bmatrix}_{rxr}$

Upon multiplying the matrices and carefully comparing corresponding matrix entries we obtain the following set of independent equations.

If

$$\underline{M} = \left[\underline{M}_1 \;\vdots\; \underline{M}_2 \;\vdots\; \cdots \;\vdots\; \underline{M}_e \right]$$

where $\underline{M}_k = [m_{ij}]$ for $k = 1, 2, \ldots, e$

$i = 1, 2, \ldots, e \cdot r$

$j = 1, 2, \ldots, r$

then

i. $m_{i_1, j_1}(k) = m_{i_2, j_2}(k)$ if $i_1 + j_1 = i_2 + j_2$

for $i = 1, 2, \ldots, e \cdot r$; $j = 1, 2, \ldots, r$.

ii. $m_{i+1, r}(k) = \sum_{j=1}^{r} - a_{j-1} m_{i,j}(k) + m_{i,1}(k+1)$

for $i = 1, 2, \ldots, r-1$; $k = 1, 2, \ldots, e-1$

iii. $m_{i+1, r}(e) = \sum_{j=1}^{r} - a_{j-1} m_{i,j}(e)$

for $i = 1, 2, \ldots, r-1$.

iv. $m_{e.r, i+1}(k) = \sum_{j=1}^{e.r} - b_{j-1} m_{j,i}(k)$

for $i = 1, 2, \ldots, r-1$; $k = 1, 2, \ldots, e$

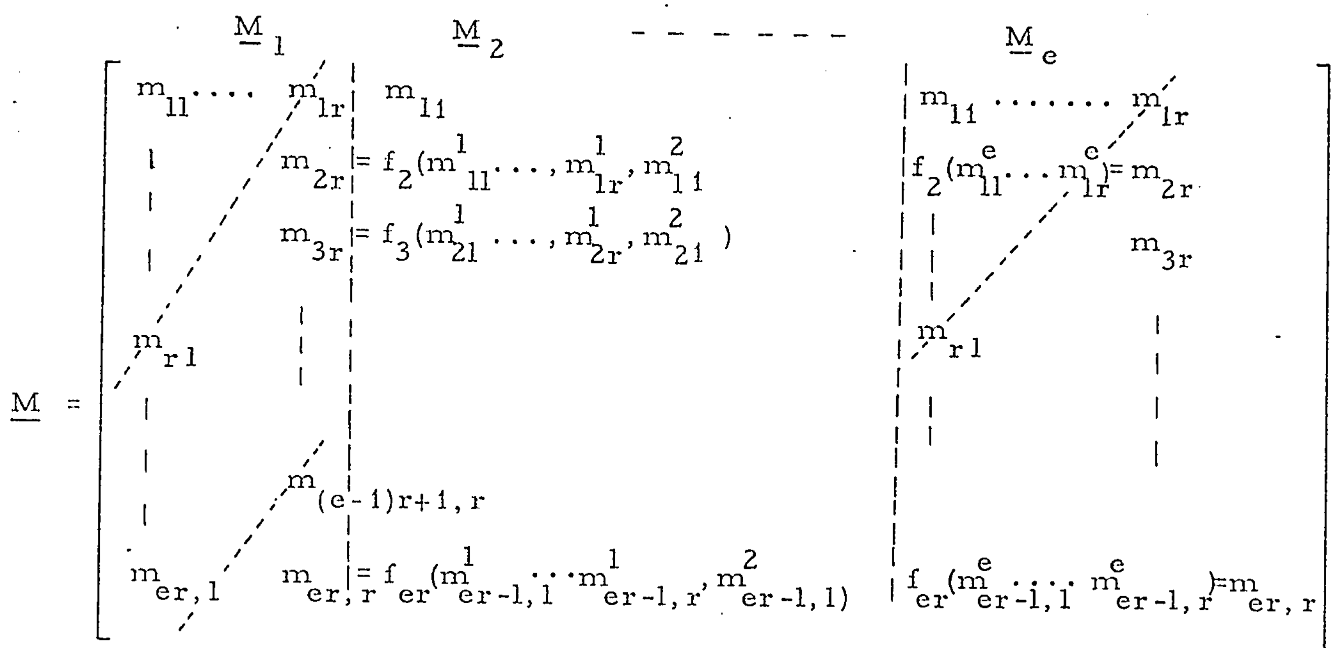
v. $\sum_{j=1}^{e.r} - b_{j-1} m_{j,r}(k) = \sum_{i=1}^{r} - a_{i-1} m_{e.r,i}(k) + m_{e.r,1}(k+1)$

for $k = 1, 2, \ldots, e-1$

vi. $\sum_{j=1}^{e.r} - b_{j-1} m_{j,r}(e) = \sum_{i=1}^{r} - a_{i-1} m_{e \cdot r, i}(e)$

Note that the sum of the submatrix entry subscripts defines an equivalence relation S on the set of submatrix entries. Upon assigning the first row entries of $\underline{M}$ the value of the remaining entries may be calculated using the above equations. If $\underline{M}$ is defined over a finite field F of q elements then there exists at most $(q^n - 1)$ $\underline{M}$ matrices such that $\underline{A}_c = \underline{M}\,\underline{A}_H^T\,\underline{M}^{-1}$.

For more insight into the entries of $\underline{M}$, as developed in Theorem A.4, see Figure A.3.



All the matrix entries on the same diagonal are equal.

Figure A.3: Schematic Diagram of the Entries of $\underline{M}$.

Consider the following examples.

Example A.7: Let $\underline{A}_H^T$ be the transpose of the hypercompanion matrix of $m(x) = (x-2)^3$ where $(x-2)$ is a monic, irreducible polynomial over R and let $\underline{A}_c$ be the companion matrix of $m(x) = x^3 - 6x^2 + 12x - 8$.

Let $(m_{11}\ m_{12}\ m_{13}) = (1\ 0\ 1)$. Hence by Theorem A.4

$$\underline{M} = \left[\begin{array}{c|c|c} 1 & 0 & 1 \\ 2 & 1 & 2 \\ 5 & 4 & 4 \end{array}\right] \quad \text{and} \quad |\underline{M}| = -1 \quad \therefore \underline{M}^{-1} \text{ exists,} \quad \text{After calculation}$$

$$\underline{M}^{-1} = \begin{bmatrix} 4 & -4 & 1 \\ -2 & 1 & 0 \\ -3 & 4 & -1 \end{bmatrix} .$$ Upon verification,

$$\underline{M}\underline{A}_H^T\underline{M}^{-1} = \begin{bmatrix} 1 & 0 & 1 \\ 2 & 1 & 2 \\ 5 & 4 & 4 \end{bmatrix}\begin{bmatrix} 2 & 0 & 0 \\ 1 & 2 & 0 \\ 0 & 1 & 2 \end{bmatrix}\begin{bmatrix} 4 & -4 & 1 \\ -2 & 1 & 0 \\ -3 & 4 & -1 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 8 & -12 & 6 \end{bmatrix} = \underline{A}_c$$

Example A.8: Let $\underline{A}_H^T$ be the transpose of the hypercompanion matrix of $m(x) = (x+1)^3$ where $(x+1)$ is a monic, irreducible polynomial over GF (2) and let $\underline{A}_c$ be the companion matrix of $m(x) = x^3 + x^2 + x + 1$.

Let $(m_{11} \; m_{12} \; m_{13}) = (1 \; 1 \; 1)$. Hence by theorem A.4

$$\underline{M} = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}$$ and $|\underline{M}| = 1$ $\therefore$ $\underline{M}^{-1}$ exists. After calculation, $$\underline{M}^{-1} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix} .$$ Upon verification,

$$\underline{M}\underline{A}_H^T\underline{M}^{-1} = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix} = \underline{A}_c$$

Lemma A.2: Let $\underline{A} = \underline{A}_1 \otimes \ldots\ldots \otimes \underline{A}_r$, $\underline{A}^* = \underline{A}_1^* \otimes \underline{A}_2^* \otimes \ldots \otimes \underline{A}_r^*$ and $\underline{A}_i$ be in canonical form such that $\underline{M}_i \; \underline{A}_i \; \underline{M}_i^{-1} = \underline{A}_i^*$, where $\underline{M}_i$ is constructed as in theorems A.1, A.2 and A.4 depending on the characteristic polynomial of $\underline{A}_i$ and the desired canonical form of $\underline{A}_i^*$. Since $\underline{A}$ and $\underline{A}^*$ are similar, there exists a nonsingular matrix $\underline{M}$ such that $\underline{M} \; \underline{A} \; \underline{M}^{-1} = \underline{A}^*$. Let us describe an algorithm to construct $\underline{M}$.

Let $\underline{M} = \underline{M}_1 \otimes \underline{M}_2 \otimes \ldots\ldots \otimes \underline{M}_r$. Therefore $\underline{M}^{-1} = \underline{M}_1^{-1} \otimes \underline{M}_2^{-1} \otimes \ldots\ldots \otimes \underline{M}_r^{-1}$ since $\underline{M} \; \underline{M}^{-1} = \underline{1}$.

Thus

$$\underline{M}\,\underline{A}\,\underline{M}^{-1} = \begin{bmatrix} \underline{M}_1 & & & \underline{0} \\ & \underline{M}_2 & & \\ & & \ddots & \\ \underline{0} & & & \underline{M}_r \end{bmatrix} \cdot \begin{bmatrix} \underline{A}_1 & & & \underline{0} \\ & \underline{A}_2 & & \\ & & \ddots & \\ \underline{0} & & & \underline{A}_r \end{bmatrix} \cdot \begin{bmatrix} \underline{M}_1^{-1} & & & \underline{0} \\ & \underline{M}_2^{-1} & & \\ & & \ddots & \\ \underline{0} & & & \underline{M}_r^{-1} \end{bmatrix}$$

$$= \begin{bmatrix} \underline{M}_1 \underline{A}_1 \underline{M}_1^{-1} & & & \underline{0} \\ & \underline{M}_2 \underline{A}_2 \underline{M}_2^{-1} & & \\ & & \ddots & \\ \underline{0} & & & \underline{M}_r \underline{A}_r \underline{M}_r^{-1} \end{bmatrix} = \begin{bmatrix} \underline{A}_1^* & & \\ & \underline{A}_2^* & \\ & & \ddots \\ & & & \underline{A}_r^* \end{bmatrix} = \underline{A}^*$$

Hence the required $\underline{M} = \underline{M}_1 \otimes \underline{M}_2 \otimes \ldots \otimes \underline{M}_r$.

<u>Theorem A.5:</u> Let $\underline{A}_c$ be the companion matrix of $p(x) = m(x) = m_1(x)^{e_1}\, m_2(x)^{e_2} \ldots m_r(x)^{e_r}$ where $m_i(x)$ is a monic, irreducible polynomial over F and $\underline{A}_{c_i}$ be the companion matrix of $m_i(x)$ for $i = 1, 2, \ldots, r$. Since $\underline{A}_c$ and $\underline{A}^* = \underline{A}_{c_1} \otimes \underline{A}_{c_2} \otimes \ldots \otimes \underline{A}_{c_r}$ are similar then there exists a nonsingular matrix $\underline{M} = [\, m_{ij} \,]_{n \times n}$ where $n = e \cdot r$ such that $\underline{A}_c = \underline{M}\; \underline{A}^* \underline{M}^{-1}$. Let us describe an algorithm to construct $\underline{M}$.

Since $\underline{A}_c = \underline{M}\; \underline{A}^*\; \underline{M}^{-1}$ then $\underline{M}\; \underline{A}^* = \underline{A}_c \underline{M}$. Let $m_i(x)^{e_i} = f_i(x) = x^{r_i} - \alpha_{i, r_i - 1}\, x^{r_i - 1} - \ldots - \alpha_{i1}\, x - \alpha_{io}$ for $i = 1, 2, \ldots, r$ and $m(x) = x^n - \beta_{n-1}\, x^{n-1} - \ldots - \beta_1 x - \beta_o$.

Thus

$$\begin{bmatrix} m_{11} & \ldots & m_{1n} \\ m_{21} & \ddots & \vdots \\ \vdots & & \vdots \\ m_{n1} & \ldots & m_{nn} \end{bmatrix} \begin{bmatrix} \underline{A}^T_{c_1} & & & \\ & \underline{A}^T_{c_2} & & \underline{0} \\ \underline{0} & & \ddots & \\ & & & \underline{A}^T_{c_r} \end{bmatrix} = \begin{bmatrix} 0 & 1 & & & \\ 0 & 0 & 1 & & \underline{0} \\ \vdots & \vdots & & \ddots & \\ \vdots & \vdots & & & 1 \\ \beta_0 & \beta_1 & \ldots & & \beta_{n-1} \end{bmatrix} \begin{bmatrix} m_{11} & \ldots & m_{1n} \\ m_{21} & \ddots & \vdots \\ \vdots & & \vdots \\ m_{n1} & \ldots & m_{nn} \end{bmatrix}$$

where

$$\underline{A}_{c_i} = \begin{bmatrix} 0 & 1 & & & \\ 0 & 0 & 1 & & \underline{0} \\ & \underline{0} & & \ddots & \\ & & & & 1 \\ \alpha_{io} & \alpha_{i1} & \cdots & \cdots & \alpha_{i,r_i-1} \end{bmatrix}$$

Upon multiplying the matrices and carefully comparing corresponding matrix entries we obtain the following set of independent equations.

If

$$\underline{M} = [\, \underline{M}_1 \quad \underline{M}_2 \quad \cdots \quad \underline{M}_r \,] \qquad \text{where}$$

$$\underline{M}_k = [\, m_{ij} \,] \qquad \text{for} \qquad \begin{aligned} k &= 1, 2, \ldots, r \\ i &= 1, 2, \ldots, n \\ j &= 1, 2, \ldots, r_k \end{aligned}$$

then

i - $m_{i1,j1}(k) = m_{i2,j2}(k)$ if $i_1 + j_1 = i_2 + j_2$

for $i = 1, 2, \ldots, n$; $j = 1, 2, \ldots, r_k$

ii - $m_{i+1,r_k}(k) = \sum_{j=1}^{r_k} \alpha_{k,j-1} \; m_{ij}(k)$

for $i = 1, 2, \ldots, n-1$; $k = 1, 2, \ldots, r$

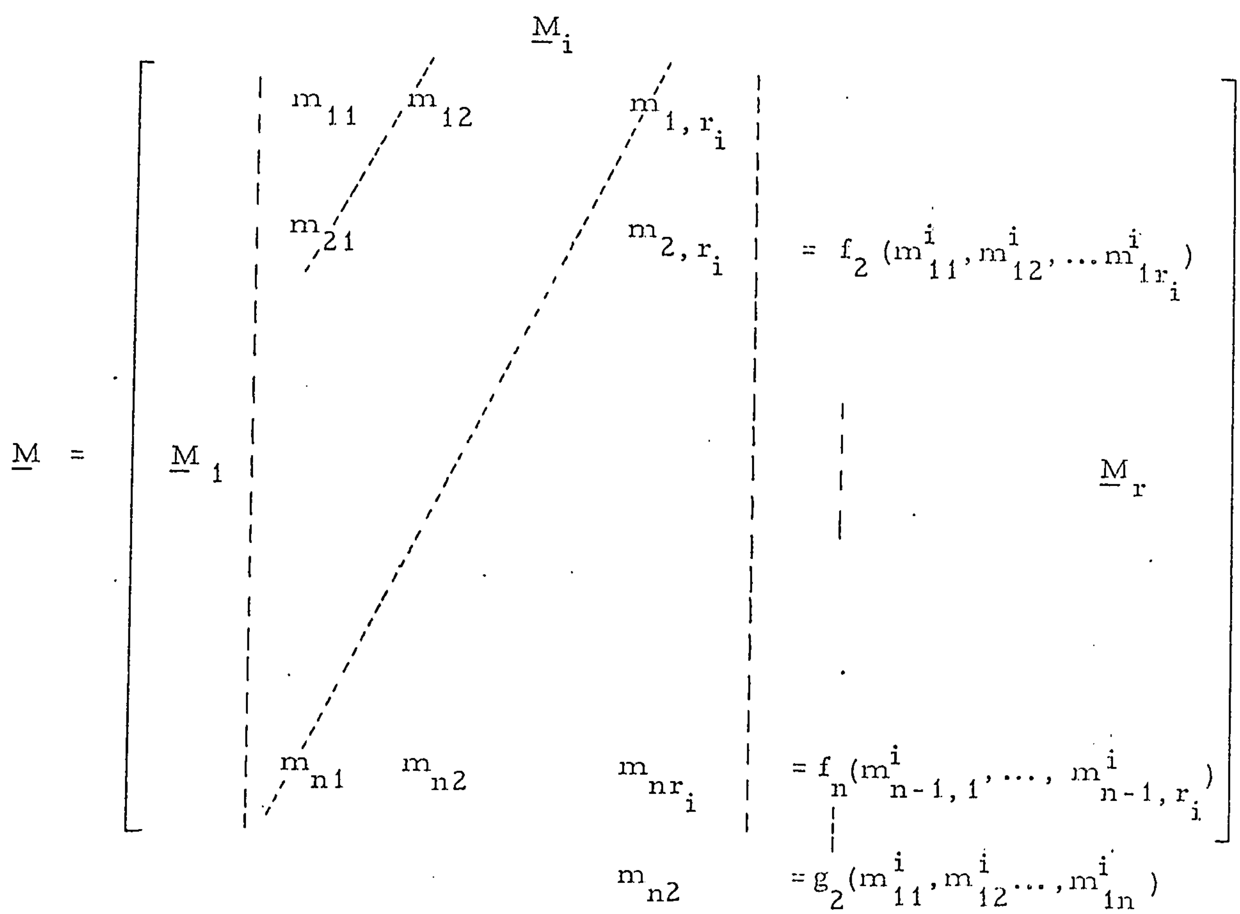
iii - $m_{n,j+1}(k) = \sum_{i=1}^{n} \beta_{i-1,j} \; m_{ij}$

for $k = 1, 2, \ldots, r$; $j = 1, 2, \ldots, r_k - 1$

Note that the sum of the submatrix entry subscripts defines an equivalence relation S on the set of submatrix entries. Upon assigning the first row entries of $\underline{M}_i$ for $i = 1, 2, \ldots, r$ the value of the remaining entries may be calculated using the above equations. Hence if $\underline{M}$ is defined over a finite field F of q elements and N is equal to the number of $\underline{M}$ matrices satisfying the above equations then the following inequality holds:

$$\prod_{i=1}^{r} (q^{r_i} - 1) \leq N \leq \prod_{i=1}^{r} (q^{n} - i)$$

For more insight into the entries of $\underline{M}$, as developed in theorem A.5, see figure A.4.



All the matrix entries on the same diagonal are equal.

Figure A.4: Schematic Diagram of the entries of $\underline{M}$.

Consider the following examples.

Example A.9: Let $\underline{A}_c$ be the companion matrix of $m(x) = (x+1)(x^2+x+1) = x^3+1$ and let $\underline{A}_{c_1}$ and $\underline{A}_{c_2}$ be the companion matrix of $x+1$ and x^2+x+1 over $GF = \{0, 1\}$ respectively.

Let $(m_{11} \; m_{12} \; m_{13}) = (1 \; 1 \; 1)$. Hence by theorem A.5

$$\underline{M} = \left[\begin{array}{c|cc} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{array}\right] \quad \text{and} \quad |\underline{M}| = 1 \quad \therefore \; \underline{M}^{-1} \text{ exists. After}$$

calculation

$$\underline{M}^{-1} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} \qquad \text{Upon verification}$$

$$\underline{M}\,\underline{A}^*\,\underline{M}^{-1} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \left[\begin{array}{c|cc} 1 & \multicolumn{2}{c}{\underline{O}} \\ \hline \underline{O} & 0 & 1 \\ & 1 & 1 \end{array}\right] \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} = \underline{A}_c$$

Example A.10: Let $\underline{A}_c$ be the companion matrix of $m(x) = x(x^2+1) = x^3+x$ and let $\underline{A}_{c_1}$ and $\underline{A}_{c_2}$ be the companion matrix of x and x^2+x+1 over $GF = \{0, 1\}$ respectively.

Let $(m_{11} \; m_{12} \; m_{13}) = (1 \; 0 \; 1)$. Hence by theorem A. 5

$$\underline{M} = \left[\begin{array}{c|cc} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array}\right] \quad \text{and} \quad |\underline{M}| = 1. \quad \therefore \; \underline{M}^{-1} \text{ exists. After}$$

calculation

$$\underline{M}^{-1} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad \text{Upon verification}$$

$$\underline{M}\ \underline{A}^{*}\ \underline{M}^{-1} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} = \underline{A}_c$$

The numerical algorithms developed in this appendix enable us to construct the transformation matrix in the following cases:

1. The matrix to be transformed is in canonical form.

2. The matrix to be transformed is the direct sum of a finite number of matrices in canonical form.

APPENDIX B

ALGORITHM FOR GENERATING THE NULL SPACE OF A MATRIX

This appendix includes an algorithm for finding all the nonzero elements in the null space of a matrix. The use of this algorithm is extended to the solution of a system of m equations in n unknowns where $m \leq n$.

B-1: Algorithm for generating the null space of a matrix

Definition B.1: An echelon matrix is a matrix in which the number of zeroes preceding the first nonzero entry of a row increases row by row until only zero rows remain. The first nonzero entry in any row is a distinguished element of the echelon matrix.

Definition B.2: A row reduced echelon matrix is an echelon matrix such that the distinguished elements are:

1. the only nonzero entries in their respective columns
2. each equal to 1.

Let $\underline{M}_{m \times n}$ be a matrix with coefficients defined over the field F. Let the null space of $\underline{M}$ have dimension q. Therefore the algorithm for finding all the nonzero elements in the null space of $\underline{M}$ is as follows:

1. Reduce $\underline{M}$ to its' row reduced echelon form $\underline{M}_r$, by a series of elementary row operations such as:

   i. interchange two rows
   ii. add one row to another
   iii. multiply one row by a constant.

2. Interchange, if necessary, the columns of $\underline{M}_r$ until it is in the form:

$$\underline{M}_r = \left[\begin{array}{c|c} \underline{I}_p & \underline{Q}_{p \times q} \\ \hline & \underline{0} \end{array}\right]_{m \times n} \qquad \text{where } p = n-q$$

3. Construct the matrix $\underline{N}$ as follows:

$$\underline{N} = \begin{bmatrix} -\underline{Q}_{p \times q} \\ \hline \underline{I}_q \end{bmatrix}_{n \times q}$$

4. Permute the rows of $\underline{N}$ in like manner to the columns of $\underline{M}_r$ in Step 2. If the j^{th} column of $\underline{M}_R$ was placed in the i^{th} column of $\underline{M}_R$ then the i^{th} row of $\underline{N}$ is placed in the j^{th} row of $\underline{N}$.

5. The columns of $\underline{N}$ form a basis of the null space of $\underline{M}$. Therefore the set of <u>all</u> the nonzero elements in the null space of $\underline{M}$ is equal to the set of all the linear combinations of the basis vectors.

<u>Example B.1:</u>

Let $$\underline{M} = \begin{bmatrix} 1 & 1 & 0 & 3 \\ -4 & -4 & 1 & -8 \\ 3 & 3 & -2 & 1 \\ 2 & 2 & 0 & 6 \end{bmatrix}$$

<u>Step 1:</u>

$$\underline{M}_r = \begin{bmatrix} 1 & 1 & 0 & 3 \\ 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

<u>Step 2:</u>

$$\underline{M}_r = \left[\begin{array}{cc|cc} 1 & 0 & 1 & 3 \\ 0 & 1 & 0 & 4 \\ \hline \multicolumn{4}{c}{\underline{0}} \end{array}\right]$$

<u>Step 3:</u>

$$\underline{N} = \left[\begin{array}{cc} -1 & -3 \\ 0 & -4 \\ \hline 1 & 0 \\ 0 & 1 \end{array}\right]$$

<u>Step 4:</u>

$$\underline{N} = \begin{bmatrix} -1 & -3 \\ 1 & 0 \\ 0 & -4 \\ 0 & 1 \end{bmatrix}$$

Therefore the set of all nonzero elements in the null space of $\underline{M}$ is equal to the set of all linear combinations of $\left\{ \begin{bmatrix} -1 & -3 \\ 1 & 0 \\ 0 & -4 \\ 0 & 1 \end{bmatrix} \right\}$.

B-2: Method for solving a system of linear equations

Consider the following system of m linear equations in n unknowns $x_1, x_2, \ldots, x_n$:

$$
\begin{aligned}
a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= b_1 \\
a_{21}x_1 + \ldots\ldots \qquad + a_{2n}x_n &= b_2 \\
\vdots \qquad \ddots \qquad \vdots \quad &\quad \vdots \\
a_{m1}x_1 + a_{m2}x_2 + \ldots\ldots + a_{mn}x_n &= b_m
\end{aligned}
$$

where $a_{ij}, b_i \in F$ for $1 \le i \le m$, $1 \le j \le n$.

The above equations may be represented in matrix form as follows:

$$
\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & \ddots & & \\ \vdots & & \ddots & \\ a_{m1} & \cdots & & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix} \quad \text{or} \quad \underline{A}\,\underline{x} = \underline{B}
$$

where $\underline{A}$ is called the coefficient matrix.

If $\underline{B} = \underline{0}$ then the system of linear equations is a homogeneous system of linear equations such that $\underline{A}\,\underline{x} = \underline{0}$. As in the previous section, the set of solutions to this system of equations is equal to the set of elements in the null space of $\underline{A}$.

If $\underline{B} \neq 0$ then the system of linear equations is a heterogeneous system of linear equations. Let $\underline{A}^* = \left[\begin{array}{c|c} A & B \\ \hline 0 \ldots & 0 \end{array}\right]$ and $x_{n+1} = -1$. $\underline{A}^*$ is called the augmented matrix of $\underline{A}$ while x_{n+1} is a dummy variable which enables us to represent the heterogeneous. system of equations as follows:

$$\underline{A}^* \underline{x}^* = \underline{0} \quad \text{where} \quad \underline{x}^* = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n+1} \end{bmatrix}.$$

As in the previous section the set of solutions to this system of equations is equal to the set of elements in the null space of $\underline{A}^*$ for which the $n+1^{th}$ column entry is equal to -1.

<u>Example B.2:</u> Consider the following system of 3 linear equations in 3 unknowns:

$$\begin{aligned} 3x_1 + x_2 + 6x_3 &= -3 \\ 2x_1 + x_2 + 4x_3 &= -2 \\ 2x_1 - x_2 + 3x_3 &= 1 \end{aligned}$$

In matrix form:

$$\begin{bmatrix} 3 & 1 & 6 \\ 2 & 1 & 4 \\ 2 & -1 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -3 \\ -2 \\ 1 \end{bmatrix} \quad \text{or} \quad \underline{A}\,\underline{x} = \underline{B}$$

$$\therefore \quad \underline{A}^* = \left[\begin{array}{c|c} \underline{A} & \underline{B} \\ \hline 0 \;.\;.\;.\;.\; & 0 \end{array}\right] = \begin{bmatrix} 3 & 1 & 6 & -3 \\ 2 & 1 & 4 & -2 \\ 2 & -1 & 3 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$ The row reduced echelon form of $\underline{A}^*$ is

$$\begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$ Hence the matrix

$$\underline{N} = \begin{bmatrix} -1 \\ 0 \\ 1 \\ 1 \end{bmatrix}.$$ The set of elements in the null space of $\underline{A}^*$ is given

by $\begin{bmatrix} -\beta \\ 0 \\ \beta \\ \beta \end{bmatrix}$ where $\beta \in F$. The solution to our system of equations is that element for which the $n+1^{th}$ column entry is -1, that is $\beta = -1$.

Hence the solution is $\begin{bmatrix} 1 \\ 0 \\ -1 \\ -1 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}$

Therefore $x_1 = 1$, $x_2 = 0$ and $x_3 = -1$.

Hence the algorithm in the previous sections permits us to solve any system of linear equations by finding the elements in the null space of the coefficient matrix or the augmented matrix.

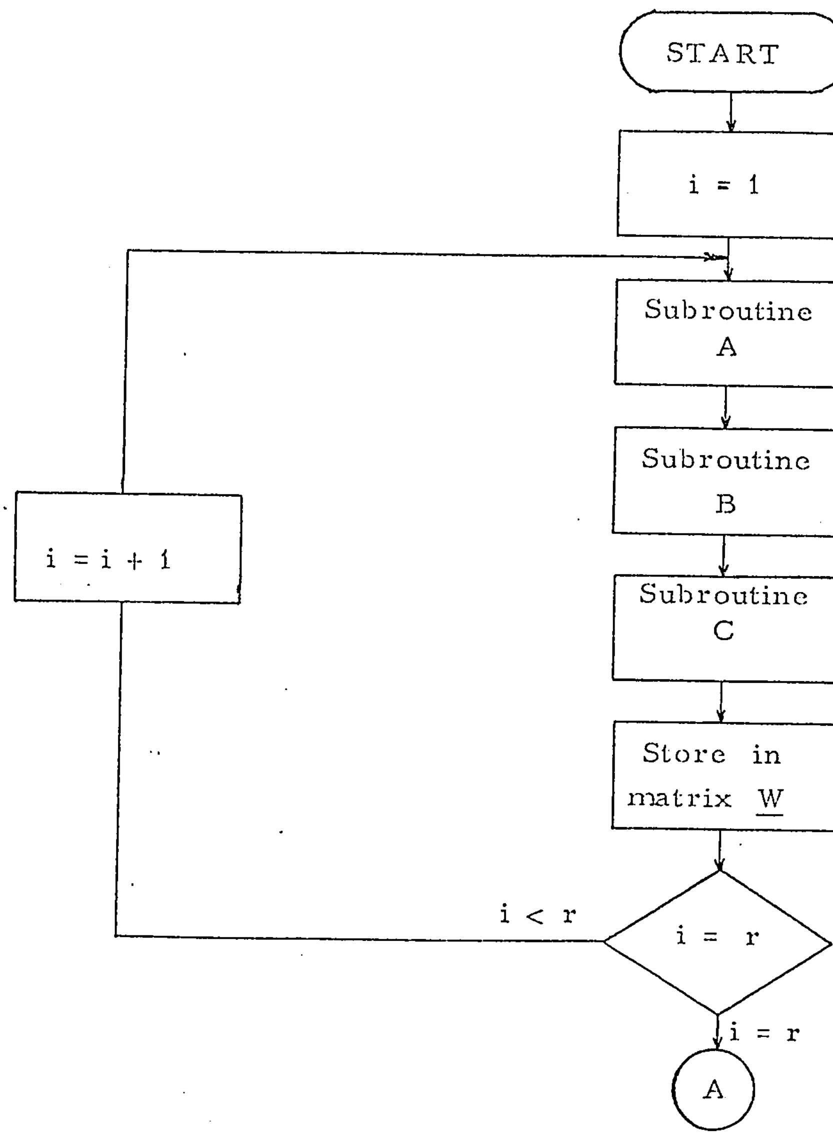
APPENDIX C

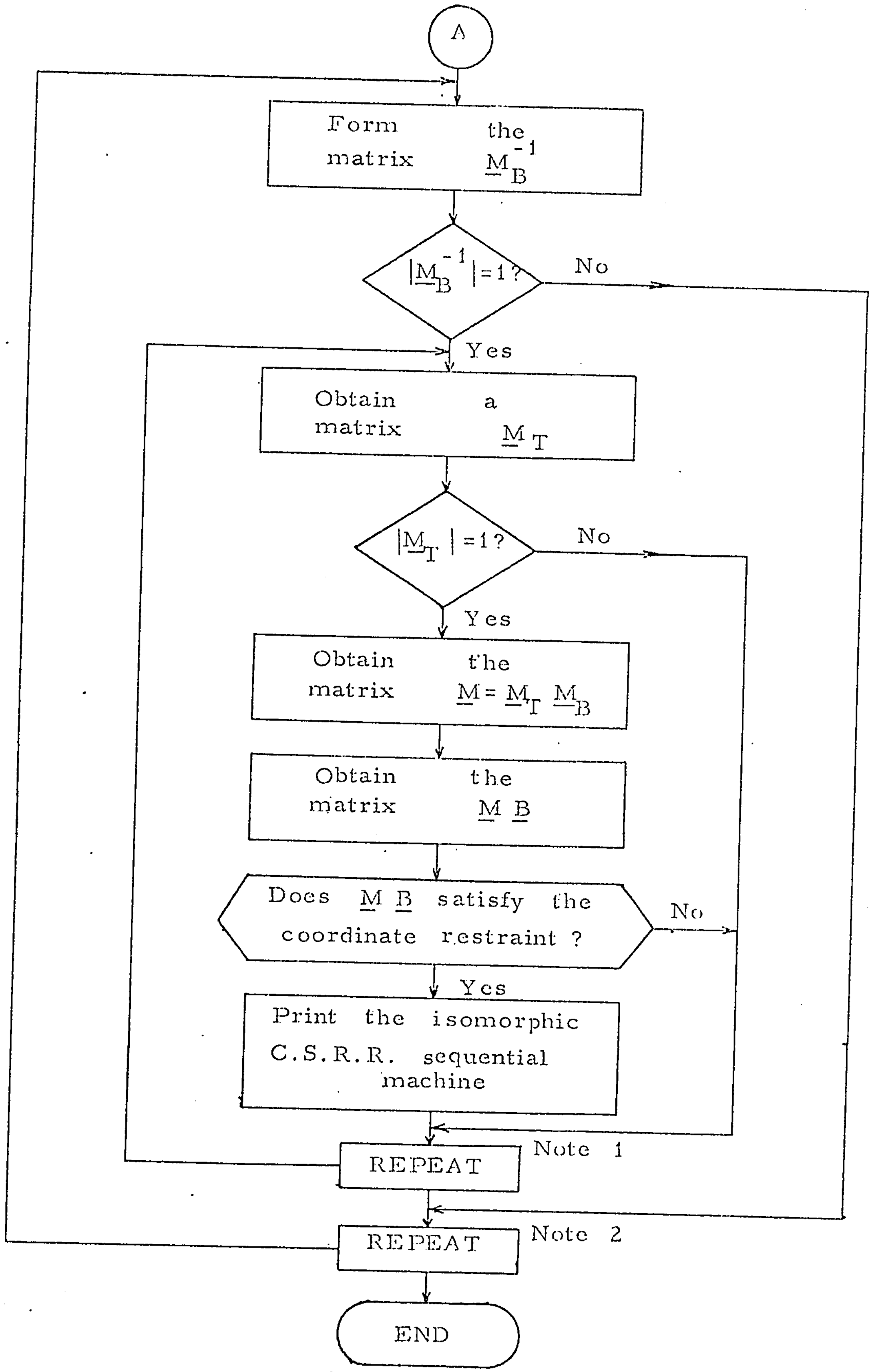
FLOW CHARTS AND PROGRAM LISTING

This appendix includes a main program flow chart, several subprogram flow charts and a listing of the program written to implement the shift register realization algorithm described in Section IV-3.

C-1: Flow Charts:

Main Program Flow Chart:



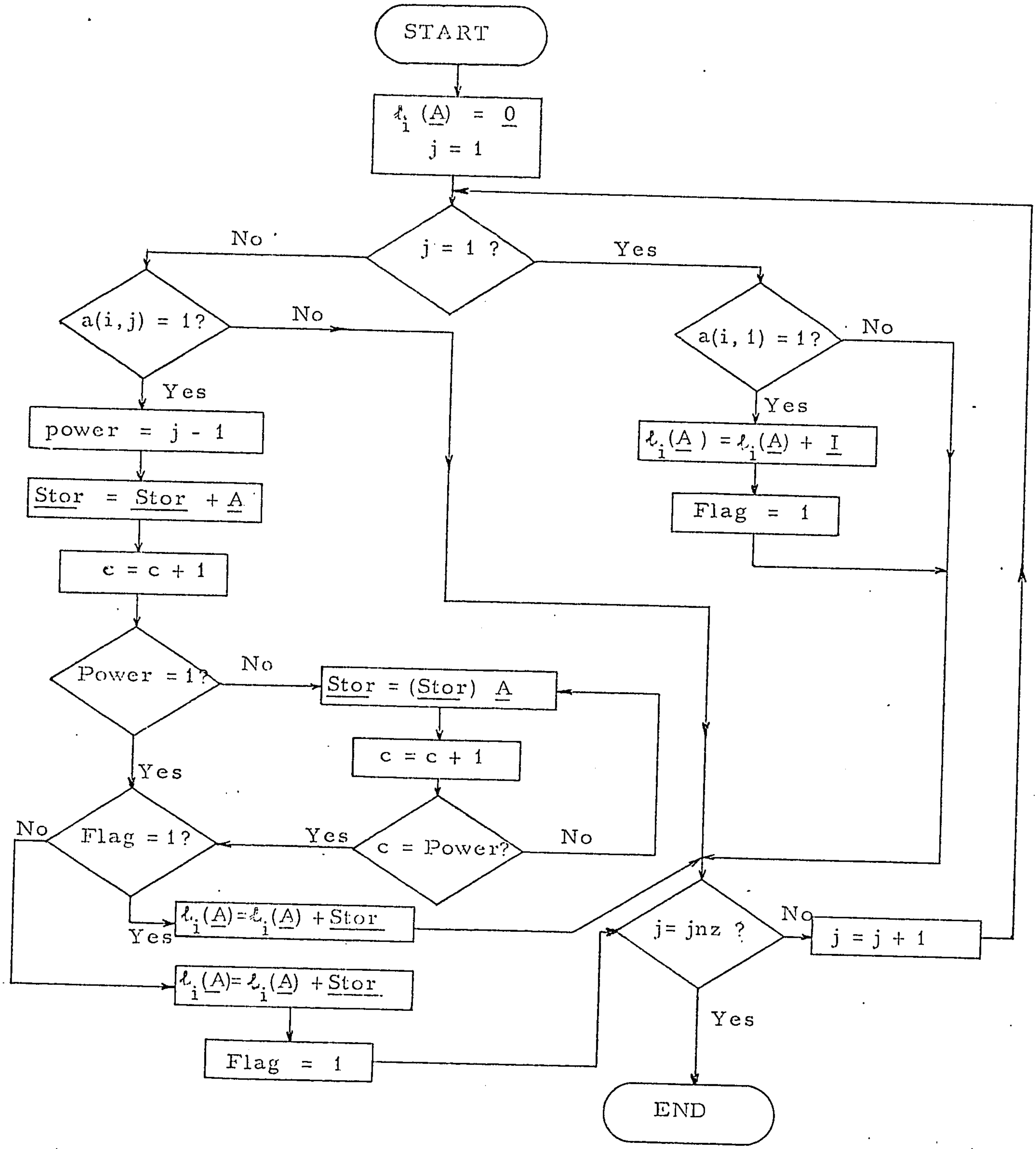


Note 1: Repeat as in Step 9 of the algorithm given in Chapter IV.

Note 2: Repeat as in Step 10 of the algorithm given in Chapter IV.

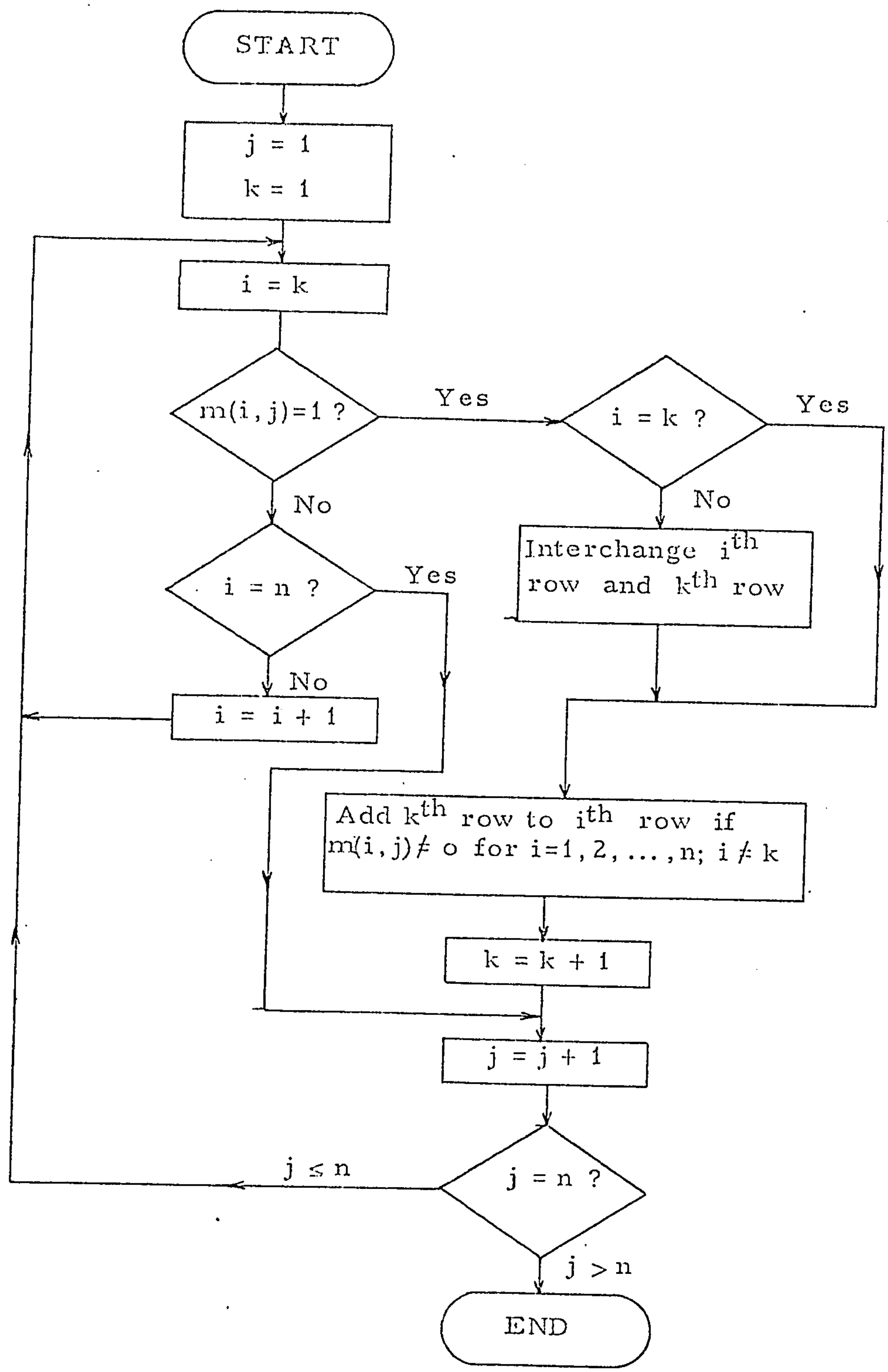
Subroutine A Flow Chart:

This subroutine computes the matrix $\ell_i(\underline{A})$ for any polynomial $\ell_i(x)$ and matrix $\underline{A}$ defined over GF = {0, 1}.



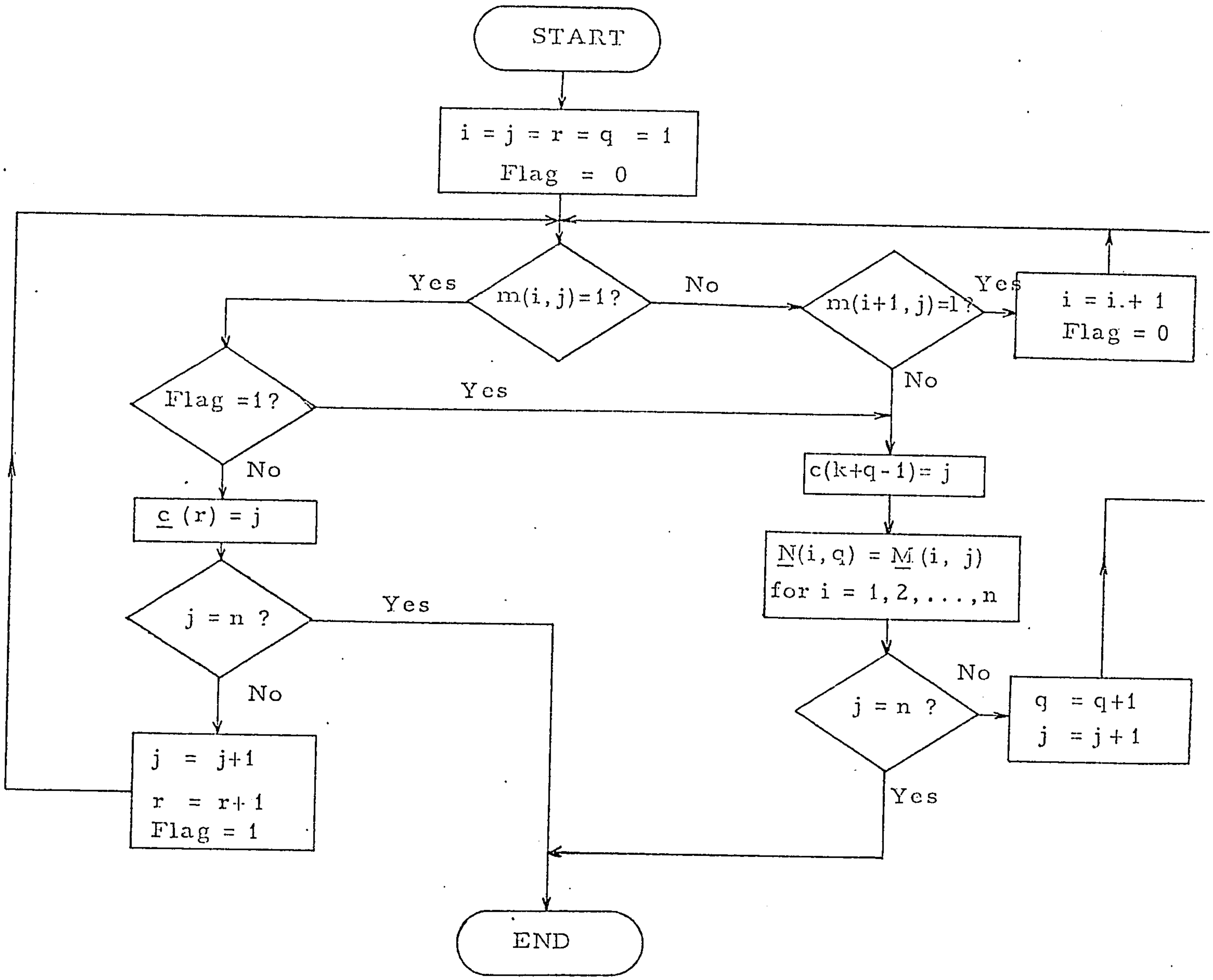
Subroutine B Flow Chart:

This subroutine converts a given matrix M defined over GF = {0, 1} to its row reduced echelon canonical form.



Subroutine C Flow Chart:

This subroutine computes the null space matrix $\underline{N}_{nxq}$ (Appendix B) of any matrix $\underline{M}_{nxn}$ defined over GF = {0, 1}.



C-2: Program Listing :

MAIN PROGRAM

```
      IMPLICIT INTEGER(A-Z)
      DIMENSION A(6,6),M(6,6),STOR(6,6),SCRAP(6,6),ELDIV(3,4),
     1          MDEG(3),C(6),NUSP(6,6),W(6,11),DIMINV(3),V(6),
     1          AV(6,1),BASIS(6,6),MB(6,6),IB(3),NN(3),
     1          MS(6,6),STOR1(1,6),MIS(6,6)
      DIMENSION AC(6,6),CSTAR(4,6),BSTAR(6,3),D(4,3),B(6,1),CI(4,6)
      N = 6
      Q = 1
      P = 4
      R = 3
      JNZL = 4
      DO 100 I = 1,N
      READ(1,105) (A(I,J),J=1,N)
  105 FORMAT(6I2)
  100 CONTINUE
      DO 110 I = 1,R
      READ(1,115) (ELDIV(I,J),J=1,JNZL)
  115 FORMAT(4I2)
  110 CONTINUE
      DO 400 I = 1,N
      READ(1,401) (B(I,J),J=1,Q)
  401 FORMAT(1I2)
  400 CONTINUE
      DO 402 I = 1,P
      READ(1,403) (CI(I,J),J=1,N)
  403 FORMAT(6I2)
  402 CONTINUE
      DO 404 I = 1,P
      READ(1,405) (D(I,J),J=1,Q)
  405 FORMAT(1I2)
  404 CONTINUE
      DO 1 I = 1,R
      IF(ELDIV(I,JNZL)) 2,2,3
    3 DEG = JNZL - 1
      GO TO 6
    2 Y = JNZL - 1
      DO 4 J = 1,Y
      JK = JNZL - J
      IF(ELDIV(I,JK)) 4,4,5
    4 CONTINUE
    5 DEG = JK - 1
    6 MDEG(I) = DEG
      CALL MELDIV(N,I,R,DEG,JNZL,A,M,ELDIV,STOR,SCRAP)
      CALL RRECF(N,RANKM,NUDI,M)
      Y = RANKM
      CALL NUSPM(N,Y,NUDI,M,C,STOR,NUSP)
      NELM = 2**NUDI - 1
      DIMINV(I) = NELM
      COUNT = 1
```

```
      DETERMINE POSITION OF VECTORS IN W MATRIX

      IF (I - 1 ) 7,7,8
    8 IM1 = I - 1
      Y=0
      DO 11 J = 1,IM1
   11 Y = Y + DIMINV(J)
      GO TO 21

      PUT ALL THE LINEAR COMBINATIONS OF THE COLUMNS OF NUSP IN W.

    7 Y = 0
   21 DO 15 J = 1,N
      DO 15 K = 1,NELM
   15 W(J,(Y+K)) = 0
      DO 13 L = 1,NELM
      ND = L
      CALL BCONV(NUDI,STOR1,ND)
      DO 14 J = 1,N
      DO 14 K = 1,NUDI
      MURES = BMULT(STOR1(1,K),NUSP(J,K))
   14 W(J,(Y+COUNT)) = BSUM(W(J,(Y+COUNT)),MURES)
   13 COUNT = COUNT + 1
    1 CONTINUE
      DIMW = Y + COUNT - 1
      DO 205 I = 1,R
      WRITE(3,210) DIMINV(I),MDEG(I)
  210 FORMAT(5X,'DIMINV = ',1I2,5X,'MDEG = ',1I2)
  205 CONTINUE
      DO 200 I = 1,N
      WRITE(3,201) (W(I,J),J=1,11)
  201 FORMAT(10X,'W = ',11I2)
  200 CONTINUE
      E1 = DIMINV(1)
      B2 = E1 + 1
      E2 = E1 + DIMINV(2)
      B3 = E2 + 1
      E3 = E2 + DIMINV(3)
      DO 9 I1 = 1,E1
      DO 9 I2 = B2,E2
      DO 9 I3 = B3,E3
      IB(1) = I1
      IB(2) = I2
      IB(3) = I3
      CHECK = 0
      COUNT = 1
      DO 10 L = 1,R
      DO 24 KK = 1,N
      BASIS(KK,COUNT) = W(KK,IB(L))
   24 V(KK) = W(KK,IB(L))
      COUNT = COUNT + 1
   19 IF(CHECK - MDEG(L) + 1 ) 12,20,20
   12 CALL MMULT(N,N,1,AV,A,V)
```

```
      DO 18 K = 1,N
      BASIS(K,COUNT) = AV(K,1)
   18 V(K) = AV(K,1)
      COUNT = COUNT + 1
      CHECK = CHECK + 1
      GO TO 19
   20 CHECK=0
   10 CONTINUE
      DO 125 I = 1,N
      WRITE(3,120) (BASIS(I,J),J=1,N)
  120 FORMAT(10X,'BASIS = ',6I2)
  125 CONTINUE
      CALL MINV(BASIS,MB,N,DETERM)
      IF(DETERM) 9,9,25
   25 WRITE(3,260) (IB(I),I=1,R)
  260 FORMAT(10X,3I5)

      OBTAIN THE MATRIX THAT ENABLES THE TRANSFORMATION FROM
      AC-TRANSPOSE TO AC.

   22 ND1 = 2**MDEG(1)-1
      ND2 = 2**MDEG(2)-1
      ND3 = 2**MDEG(3)-1
      ME1 = MDEG(1)
      ME2 = MDEG(2) + ME1
      ME3 = MDEG(3) + ME2
      MB2 = ME1 + 1
      MB3 = ME2 + 1
      C(1) = 0
      C(2) = ME1
      C(3) = ME2
      DO 60 N1 = 1,ND1
      DO 60 N2 = 1,ND2
      DO 60 N3 = 1,ND3
      DO 53 I = 1,N
      DO 53 J = 1,N
   53 MS(I,J) = 0
      NN(1)=N1
      NN(2)=N2
      NN(3) = N3
      DO 61 L = 1,R
      DEGL = MDEG(L)
      NL = NN(L)
      CALL BCONV(DEGL,STOR1,NL)
      DO 62 K = 1,DEGL
   62 MS((1+C(L)),(K+C(L))) = STOR1(1,K)
      COMPUTE THE SUB-MATRIX M ASSOCIATED WITH THE ELEMENTARY
      DIVISOR L.
      IF(MDEG(L) - 1) 61,61,63
   63 JN = MDEG(L)
      I = 1
      DO 81 J = 2,JN
      S = I+J
      IK = I + C(L)
      JK = J + C(L)
```

```
      DO 81 K = 3,S
      MS((IK+1),(JK-1)) = MS(IK,JK)
      IK=IK+1
      JK = JK-1
   81 CONTINUE
      Z=JN-1
      DO 82 I = 1,Z
      MS((I+1+C(L)),(JN+C(L))) = 0
      DO 83 K = 1,JN
      EL = ELDIV(L,K)
      MI = MS((I+C(L)),(K+C(L)))
      ST1 = BMULT(EL,MI)
      MI1 = MS((I+1+C(L)),(JN+C(L)))
   83 MS((I+1+C(L)),(JN+C(L))) = BSUM(ST1,MI1)
      IF(I+1-JN) 84,84,82
   84 S = JN + 1 - I
      IK = I + C(L)
      JK = JN + C(L)
      DO 86 K = 2,S
      MS((IK+1),(JK-1)) = MS(IK,JK)
      IK=IK+1
   86 JK = JK-1
   82 CONTINUE
   61 CONTINUE
      DO 26 I = 1,N
      DO 26 J = 1,N
   26 SCRAP(I,J) = MS(I,J)
      WRITE(3,300) N1,N2,N3
  300 FORMAT(10X,3I2)
      CALL MINV(MS,MIS,N,DETERM)
      IF(DETERM) 60,60,87
   87 CALL MMULT(N,N,N,M,SCRAP,MB)
      DO 29 I = 1,N
      WRITE(3,30) (M(I,J),J=1,N)
   30 FORMAT(10X,'M = ',6I2)
   29 CONTINUE
 9999 CONTINUE
      DO 89 I = 1,N
      DO 89 J = 1,Q
   89 BSTAR(I,J) = B(I,J)
      CALL MMULT(N,N,Q,BSTAR,M,B)
      CALL MMULT(N,N,N,STOR,M,A)
      CALL MINV(M,SCRAP,N,DETERM)
      CALL MMULT(N,N,N,AC,STOR,SCRAP)
      CALL MMULT(P,N,N,CSTAR,CI,SCRAP)
      CALL PRINT(N,P,Q,AC,BSTAR,CSTAR,D)
   60 CONTINUE
    9 CONTINUE
      DO 500 I = 1,R
      WRITE(3,505) (ELDIV(I,J),J=1,JNZL)
  505 FORMAT(10X,4I2)
  500 CONTINUE
      STOP
      END
```

SUBROUTINE A

```
      SUBROUTINE MELDIV(N,I,R,DEG,JNZL,A,M,ELDIV,STOR,SCRAP)

      THIS PROGRAM FINDS THE MATPIX M = L(A) WHERE L(X) IS AN
      ELEMENTARY DIVISOR

      IMPLICIT INTEGER(A-Z)
      DIMENSION A(N,N), M(N,N), STOR(N,N), SCRAP(N,N), ELDIV(R,JNZL)
 9999 CONTINUE
      FLAG = 0
      JNZ = DEG + 1
      DO 30 K = 1,N
      DO 30 L = 1,N
      STOR(K,L) = 0
      SCRAP(K,L) = 0
   30 M(K,L) = 0
      DO 1 J = 1,JNZ
      COUNT = 0
      IF (J-1) 2,2,3

      CONSIDER THE CASE WHEN J = 1

    2 IF(ELDIV(I,1)) 1,1,4
    4 DO 5 K = 1,N
    5 M(K,K) = 1
      FLAG = 1
      GO TO 1

      CONSIDER THE CASE WHEN J = 2,----,DEG + 1

    3 IF(ELDIV(I,J)) 1,1,6
    6 POWER = J - 1
      DO 7 K = 1,N
      DO 7 L = 1,N
    7 STOR(K,L) = A(K,L)
      COUNT = COUNT + 1
      IF ( POWER - 1 ) 8,8,9
    8 IF (FLAG) 10,10,11

      FIND STOR = ( A RAISED TO A POWER )

    9 DO 15 K = 1,N
      DO 15 L = 1,N
      SCRAP(K,L) = 0
      DO 15 Q = 1,N
      IF (STOR(K,Q)) 16,16,17
   17 IF (A(Q,L)) 16,16,18
```

```
 16 PROD = 0
    GO TO 19
 18 PROD = 1
 19 IF(SCRAP(K,L) - PROD ) 20,21,20
 20 SCRAP(K,L) = 1
    GO TO 15
 21 SCRAP(K,L) = 0
 15 CONTINUE
    DO 22 K = 1,N
    DO 22 J = 1,N
 22 STOR(K,J) = SCRAP(K,J)
    COUNT = COUNT + 1
    IF ( COUNT - POWER ) 9,8,8

    PLACE CONTENTS OF STOR IN M

 10 DO 25 K = 1,N
    DO 25 L = 1,N
 25 M(K,L) = STOR(K,L)
    FLAG = 1
    GO TO 1

    ADD CONTENTS OF STOR TO M

 11 DO 12 K = 1,N
    DO 12 L = 1,N
    IF ( M(K,L) - STOR(K,L) ) 13,14,13
 13 M(K,L) = 1
    GO TO 12
 14 M(K,L) = 0
 12 CONTINUE
  1 CONTINUE
    DO 200 II = 1,N
    WRITE(3,205) (M(II,J),J=1,N)
205 FORMAT(10X,'MELDIV-- M = ',6I2)
200 CONTINUE
    RETURN
    END



    INTEGER FUNCTION BMULT(J,K)
    IMPLICIT INTEGER(A-Z)
    IF (J) 2,2,1
  1 IF (K) 2,2,3
  2 BMULT = 0
    GO TO 4
  3 BMULT = 1
  4 RETURN
    END
```

SUBROUTINE B

```
      SUBROUTINE RRECF(N,RANKM,NUDI,M)
      PROGRAM TO REDUCE A MATRIX TO THE ROW REDUCED ECHELON CANONICAL FORM
      IMPLICIT INTEGER(A-Z)
      DIMENSION M(N,N)
 9999 CONTINUE
      K = 1
      DO 31 J = 1,N
      FIND A NON ZERO ELEMENT IN THE J-TH COLUMN
      I = K
    1 IF(M(I,J)) 2,2,5
    2 IF(I-N) 4,31,31
    4 I = I+1
      GO TO 1
    5 IF(I-K) 13,13,6
      INTERCHANGE THE I-TH AND THE K-TH ROW
    6 DO 7 L = J,N
      A= M(I,L)
      M(I,L) = M(K,L)
    7 M(K,L) = A
      ADD THE K-TH ROW TO ALL THE NECESSARY ROWS
   13 DO 8 L = 1,N
      IF(L-K) 9,8,9
    9 IF(M(L,J)) 8,8,10
   10 DO 8 JJ = J,N
      IF(M(L,JJ)-M(K,JJ)) 11,12,11
   11 M(L,JJ) = 1
      GO TO 8
   12 M(L,JJ) = 0
    8 CONTINUE
   30 K = K+1
   31 CONTINUE
      RANKM = K-1
      NUDI = N-RANKM
      DO 200 I = 1,N
      WRITE(3,205) (M(I,J),J=1,N)
  205 FORMAT(10X,'RRECF-- M = ',6I2)
  200 CONTINUE
      WRITE(3,210) NUDI
  210 FORMAT(40X,'NUDI = ',1I10)
      RETURN
      END
```

SUBROUTINE C

```
      SUBROUTINE NUSPM(N,RANKM,NUDI,M,C,STOR,NUSP)

      PRUPOSE:  CONVERT M MATRIX INTO THE NULL SPACE MATRIX NUSP
      M IS MATRIX IN ROW REDUCED ECHELON CANONICAL FORM
      NUSP IS NULL SPACE MATRIX
      C IS A COLUMN VECTOR WHICH STORES THE ORDER OF THE
      PERMUTATIONS OF THE COLUMNS OF M

      IMPLICIT INTEGER(A-Z)
      DIMENSION M(N,N), C(N), STOR(N,NUDI), NUSP(N,NUDI)
 9999 CONTINUE
      FLAG = 0
      R =1
      Q =1
      K = RANKM + 1

      STORE IDENTITY MATRIX IN STOR MATRIX BEFORE THE ROW PERMUTATIONS

      DO 1 I = K,N
      DO 1 J = 1,NUDI
    1 STOR(I,J) = 0
      I = K
      J = 1
    2 STOR(I,J) = 1
      IF(I-N) 3,4,4
    3 I = I+1
      J = J+1
      GO TO 2

      FORM THE STOR MATRIX:  SAME AS NUSP BEFORE ROW PREMUTATIONS

    4 I = 1
      J = 1
    5 IF(M(I,J)) 6,6,12
    6 IF(I-RANKM) 7,9,9
    7 IF(M((I+1),J)) 9,9,8
    8 I = I+1
      FLAG = 0
      GO TO 5
    9 C(J) = RANKM + Q

      PLACE CONTENTS OF J-TH COLUMN OF M IN Q-TH COLUMN OF STOR

      DO 10 L = 1,RANKM
   10 STOR(L,Q) = M(L,J)
      IF(J-N) 11,15,15
```

```
   11 J = J+1
      Q = Q+1
      GO TO 5
   12 IF (FLAG) 13,13,9
   13 C(J) = R
      IF(J-N) 14,15,15
   14 J = J+1
      R = R+1
      FLAG = 1
      GO TO 5

      PERMUTE ROWS OF STOR IN SAME MANNER AS COLUMNS OF
      M AS DEFINED BY C MATRIX

   15 DO 16 I = 1,N
      DO 16 J = 1,NUDI
   16 NUSP(I,J) = STOR(C(I),J)
      DO 200 I = 1,N
      WRITE(3,205) (NUSP(I,J),J=1,NUDI)
  205 FORMAT(10X,'NUSPM-- M = ',6I2)
  200 CONTINUE
      RETURN
      END

      SUBROUTINE BCONV(DEG,STOR1,ND)
      IMPLICIT INTEGER(A-Z)
      DIMENSION STOR1(1,DEG)
 9999 CONTINUE
      DO 1 K = 1,DEG
    1 STOR1(1,K) = 0
    2 IF(STOR1(1,DEG)) 3,3,5
    3 STOR1(1,DEG) = 1
    4 ND = ND - 1
      IF(ND) 9,9,2
    5 T = DEG
      STOR1(1,T) = 0
    6 IF(STOR1(1,(T-1))) 7,7,8
    7 STOR1(1,(T-1)) = 1
      GO TO 4
    8 STOR1(1,(T-1)) = 0
      T = T-1
      GO TO 6
    9 RETURN
      END

      INTEGER FUNCTION BSUM (J,K)
      IMPLICIT INTEGER(A-Z)
      IF (J-K) 1,2,1
    1 BSUM = 1
      GO TO 3
    2 BSUM = 0
    3 RETURN
      END
```

```
      SUBROUTINE MINV(A,AINV,N,DETERM)
      IMPLICIT INTEGER(A-Z)
      DIMENSION A(N,N),AINV(N,N), W(4,5)
      DO 1 I = 1,N
      DO 1 J = 1,N
    1 AINV(I,J) = 0
      DO 2 I = 1,N
    2 AINV(I,I) = 1
      I = 1
   17 IF(A(I,I)) 3,3,4
    4 DO 5 J = 1,N
      IF(J-I) 18,5,18
   18 IF(A(J,I)) 5,5,6
    6 DO 7 K = 1,N
      IF(A(J,K)-A(I,K)) 8,9,8
    9 A(J,K) = 0
      GO TO 12
    8 A(J,K) = 1
   12 IF(AINV(J,K)-AINV(I,K)) 10,11,10
   11 AINV(J,K) = 0
      GO TO 7
   10 AINV(J,K) = 1
    7 CONTINUE
    5 CONTINUE
      IF(I-N) 13,14,13
   13 I = I+1
      GO TO 17
    3 K =I+1
      DO 19 L = K,N
      IF(A(L,I)) 19,19,15
   19 CONTINUE
      GO TO 50
   15 DO 16 J =1,N
      V = A(K,J)
      T = AINV(K,J)
      A(K,J) = A(I,J)
      AINV(K,J) = AINV(I,J)
      A(I,J) = V
   16 AINV(I,J) = T
      GO TO 4
   50 DETERM = 0
      GO TO 21
   14 DETERM = 1
   21 RETURN
      END
```

```
      SUBROUTINE PRINT(N,P,R,AC,BSTAR,CSTAR,D)
      IMPLICIT INTEGER(A-Z)
      DIMENSION AC(N,N),BSTAR(N,R),CSTAR(P,N),D(P,R)
      EVEN = N/2
      IF(2*EVEN .EQ. N/2)  GO TO 1000
      DO 100 I = 1,N
      IF(I .EQ. (N/2 + 1))  GO TO 90
      WRITE(3,5) (AC(I,II),II=1,N),(BSTAR(I,II),II=1,R)
      GO TO 100
   90 WRITE(3,10) (AC(I,II),II=1,N),(BSTAR(I,II),II=1,R)
  100 CONTINUE
      GO TO 2000
 1000 DO 200 I =1,N
      IF(I .EQ. N/2)  GO TO 190
      WRITE(3,5) (AC(I,II),II=1,N),(BSTAR(I,II),II=1,R)
      GO TO 200
  190 WRITE(3,10) (AC(I,II),II=1,N),(BSTAR(I,II),II=1,R)
  200 CONTINUE
 2000 WRITE(3,39)
   39 FORMAT(1H0/////)
      EVEN = P/2
      IF(2*EVEN .EQ. P)  GO TO 3000
      DO 2100 I = 1,P
      IF(I .EQ. (P/2 + 1))  GO TO 2090
      WRITE(3,5) (CSTAR(I,II),II=1,N),(D(I,II),II=1,R)
      GO TO 2100
 2090 WRITE(3,30) (CSTAR(I,II),II=1,N),(D(I,II),II=1,R)
 2100 CONTINUE
 3000 DO 2200 I =1,P
      IF(I .EQ. P/2)  GO TO 2190
      WRITE(3,5) (CSTAR(I,II),II=1,N),(D(I,II),II=1,R)
      GO TO 2200
 2190 WRITE(3,30) (CSTAR(I,II),II=1,N),(D(I,II),II=1,R)
 2200 CONTINUE
    5 FORMAT(' ',18X,'|',6I2,' |',8X,'|',1I2,' |')
   10 FORMAT(' ',9X,'X(T+1) = |',6I2,' | X(T) + |',1I2,' | U(T)')
   30 FORMAT(' ',11X,'Y(T) = |',6I2,' | X(T) + |',1I2,' | U(T)')
      RETURN
      END
      SUBROUTINE MMULT(P,Q,R,A,B,C)
      IMPLICIT INTEGER(A-Z)
      DIMENSION A(P,R), B(P,Q), C(Q,R)
      DO 7 I = 1,P
      DO 7 J = 1,R
      A(I,J) = 0
      DO 7 K = 1,Q
      IF (B(I,K)) 2,2,1
    1 IF (C(K,J)) 2,2,3
    2 PROD = 0
      GO TO 4
    3 PROD = 1
    4 IF (A(I,J) - PROD ) 5,6,5
    5 A(I,J) = 1
      GO TO 7
    6 A(I,J) = 0
    7 CONTINUE
      RETURN
      END
```

BIBLIOGRAPHY

1. Arbib M.A., Falb P.L, Kalman R.E, "Topics in Mathematical System Theory", McGraw Hill Book Company, New York, N.Y., 1969.

2. Birkhoff G., MacLane S., "A Survey of Modern Algebra", Macmillan Co., New York, N.Y., 1965.

3. Booth T.L, "Sequential Machines and Automata Theory", John Wiley and Sons, New York, N.Y., 1967.

4. Cohn M., "Controllability in Linear Sequential Networks", IRE Trans. on Circuit Theory, Vol. CT-9, March 1962, pp. 74-78.

5. Davis W.A., "On the Linearity of Sequential Machines", Master's Thesis, Electrical Engineering Department, University of Ottawa, September 1963.

"Single Shift Register Realizations for Sequential Machines", IEEE Trans. on Electronic Computers, Vol. EC-17, May 1968, pp. 422 - 431.

"On Shift Register Realizations for Sequential Machines", Technical Report No. 65-10, Electrical Engineering Department, University of Ottawa, July 1965.

6. Distefano III, Stubberud, Williams, "Theory and Problems of Feedback and Control Systems", Schaum Publishing Co., New York, N.Y., 1967.

7. Friedland B., "Linear Modular Sequential Circuits", IRE Trans. on Circuit Theory, March 1959, pp. 61-68.

8. Fu K.S, Swain P.H, "On Relationships Between the Theories of Discrete-State and Continuous-State Linear Time-Invariant Systems", Proceedings of 4th Annual Allerton Conference on Circuit and Systems Theory, 1966, pp. 794 - 804.

9. Gallaire H., Harrison M.A, "Decomposition of Linear Sequential Machines", Math. Systems Theory 3, 1969, pp. 246-287.

10. Gill A., "Synthesis of Linear Sequential Circuits From Input-Output Relations", SIAM Journal of Applied Math., Vol. 16, 1968, pp. 216-227.

"Linear Sequential Circuits", McGraw Hill Book Company, New York, N.Y., 1966.

11. Golomb S.W, "Shift Register Sequences", Holden-Day Inc., Cambridge, Mass., 1967.

12. Johnson D.L, O'Keefe K.H, "The Application of Shift Registers to Secondary State Assignment: Part I", IEEE Trans. on Electronic Computers, Vol. C-17, October 1968, pp. 954-965.

"The Application of Shift Registers to Secondary State Assignment: Part II", IEEE Trans. on Electronic Computers, Vol. C-17, October 1968, pp. 966-977.

13. Kautz W.H, "Linear Sequential Switching Circuits", Holden-Day Inc., Cambridge, Mass., 1965.

14. Lipschutz S., "Theory and Problems of Linear Algebra", McGraw Hill Book Company, New York, N.Y., 1968.

15. Martin R.L, "Studies in Feedback-Shift-Register Synthesis of Sequential Machines", M.I.T. Press, Cambridge, Mass., 1969.

16. Morrey C.B, Protter M.H, "Modern Mathematical Analysis", Addison-Wesley Publishing Co., Mass., 1964.

17. Nichols A.J, "Minimal Shift-Register Realizations of Sequential Machines", IEEE Trans. on Electronic Computers, Vol. EC-14, October 1965, pp. 688-700.

18. Peterson W.W, "Error Correcting Codes", M.I.T. Press, Cambridge, Mass., 1961.

VITA

Name :	Joseph William Jacques Farley
Born :	Ottawa, Ontario, Canada, December 21, 1946.
Education :	
Primary:	St. Joseph's Separate School, Cyrville, Ontario.
Secondary:	Eastview High School, Eastview, Ontario.
University:	Ottawa University, Ottawa, Ontario (1964-1968).
Degree:	B.A.Sc. (1968).