THE MATRIX EQUATION $F(A)X - XA = 0$

Thesis by

Joseph A Parker Jr.

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

1976

(Submitted May 21, 1976)

ACKNOWLEDGMENTS

I want to thank my wife Kathy for her love and encouragement.

The topic of this thesis was suggested to me by my adviser, Dr. Olga Taussky Todd. I would like to thank her for her patience and the effort she spent on my work. I have enjoyed working with her.

Finally, I would like to thank the California Institute of Technology for their financial support through research and teaching assistantships.

ABSTRACT

In this work, all matrices are assumed to have complex entries. The cases of $F(A)X - XA = 0$ where $F(A)$ is a polynomial over $\mathbb{C}$ in A and $F(A) = (A^*)^{-1}$ are investigated. Canonical forms are derived for solutions X to these equations. Other results are given for matrices of the form $A^{-1}A^*$.

Let a set of solutions $\{X_i\}$ be called a <u>tower</u> if $X_{i+1} = F(X_i)$. It is shown that towers occur for all nonsingular solutions of $(A^*)^{-1}X - XA = 0$ if and only if A is normal. In contrast to this, there is no polynomial for which only normal matrices A imply the existence of towers for all solutions X of $P(A)X - XA = 0$. On the other hand, conditions are given for polynomials P, dependent upon the spectrum of A, for which only diagonalizable matrices A imply the existence of towers for all solutions X of $P(A)X - XA = 0$.

TABLE OF CONTENTS

INTRODUCTION

In this work, all matrices are assumed to have complex entries and all polynomials are assumed to have coefficients in the complex field. Many of the results hold for other fields, but no mention or use of this will be made here.

This work originated from a study of the matrix equation

$$(1.1)\qquad A^{*^{-1}}X - XA = 0$$

where A and X are n×n complex matrices. This equation arose in the work of DePrima and Johnson [3] where it is shown that a matrix A is a cosquare, ie. $A = B^{-1}B^{*}$ for some matrix B, if and only if there exists a nonsingular solution X to (1.1). In the present investigation, all solutions of this equation are considered.

In Chapter II, canonical forms are developed for solutions to (1.1). It is also shown that if the condition of DePrima and Johnson is weakened to one requiring only the existence of a normal, singular solution X to (1.1) then there exist matrices B and C where B is a cosquare, but C is not a cosquare such that

$$B^{*^{-1}}X - XB = 0$$

and

$$C^{*^{-1}}X - XC = 0.$$

Thus it is shown that this weakened condition does not imply the existance of a nonsingular solution to (1.1) and that even if a nonsingular solution of (1.1) exists, there may still be interesting singular solutions to the equation.

In the work of DePrima and Johnson [3] they study the matrices A and B where $A = B^{-1}B^*$. In Chapter III we concentrate on all the nonsingular solutions to (1.1), including those solutions X such that $A = X^{-1}X^*$. Included are spectral restrictions on solutions to (1.1) in the case that A is a cosquare and conditions on the set of all solutions to (1.1) which imply that A is unitary or normal. In [1] Choi has the result that a cosquare $A = B^{-1}B^*$ is normal if and only if B^2 is normal. In Chapter III we give another condition. In particular, the cosquare A is normal if and only if for each nonsingular solution X of (1.1) it follows that ${X^*}^{-1}$ is also a solution of (1.1). This result may be used to prove the result of Choi without the use of the Fuglede-Putnam theorem [9] which Choi uses. In light of the above theorem, the following problems arise.

Consider the matrix equation

1.2) $$P(A)X - XA = 0$$

where P is a polynomial.

<u>Problem 1.3</u>: For what polynomials P does the fact that the set of all solutions to (1.2) is closed under the operation $X \to P(X)$ imply that A is normal?

In Chapter IV we develop the tools to attack this problem by giving canonical forms for the solutions to (1.2). This chapter also gives spectral restrictions on such solutions. This work is an extension of the work of Drazin [5] who dealt with the equation

$$\epsilon AX - XA = 0$$

where ϵ is a complex scalar.

In Chapter V Problem 1.3 is answered negatively. Theorem 5.2 shows that there are no polynomials which satisfy the conditions of Problem 1.3. However, there are polynomials which satisfy the following revised problem.

Problem 1.4: For what polynomials P does the fact that the set of all solutions to (1.2) is closed under the operation $X \to P(X)$ imply that A is diagonalizable?

We continue in Chapter V to give sufficient conditions on P, depending on the spectrum of A, in order that P satisfy Problem 1.4. However, these conditions are shown to be not necessary. Other conditions are given which are necessary, but not sufficient.

CHAPTER I

NOTATIONS AND PRELIMINARIES

The following definitions, notations and theorems are assumed in this work.

Definition 1: Let A and B be square matrices, then the direct sum of A and B, denoted by $A \oplus B$ is the matrix

$$\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}$$

Likewise, if $A_1, \ldots, A_n$ are square matrices, then

$$\bigoplus_{i=1}^{n} A_i = \left(\bigoplus_{i=1}^{n-1} A_i \right) \oplus A_n.$$

Definition 2: The Kronecker product of $A = (a_{ij})$ and $B = (b_{ij})$, denoted by $A \otimes B$ is the matrix, $C = (c_{ij}) = (a_{ij}B)$.

Definition 3: The spectrum of a matrix A is the set of eigenvalues of A and will be written $\sigma(A)$.

Definition 4: The matrix A is called a cosquare if there exists a nonsingular matrix B such that $A = B^{-1}B^{*}$. (As used in [1]).

Definition 5: The matrix A is called a block monomial matrix if $A = (A_{ij})$ is a block matrix with at most one nonzero block in each row and column.

Definition 6: A matrix of the form

$$\begin{bmatrix} a_{11} & \cdots & a_{1r} \\ \vdots & & \vdots \\ a_{r1} & & \\ 0 & \ddots & \\ \vdots & & \\ 0 & \cdots\ 0 & a_{kr} \end{bmatrix}_{k\times r} \quad \text{or} \quad \begin{bmatrix} 0 & \cdots & 0 & a_{11} & \cdots & a_{1r} \\ \vdots & & & \ddots & \ddots & \vdots \\ 0 & & \cdots & & 0 & a_{rr} \end{bmatrix}_{r\times k} \quad \text{for } k > r$$

is called upper triangular

Definition 7: The matrix A is called nonderogatory if the Jordan normal form of A contains only one Jordan block for each distinct eigenvalue of A. Otherwise, A is called derogatory.

Definition 8: A set of solutions $\{X_i\}$ is a tower starting with X_1 for the equation $F(A)X - XA = 0$ if $X_{i+1} = F(X_i)$ for all i.

Notation 1: The matrix $A^{*^{-1}}$ will be denoted by A^{-*}.

Notation 2: The n×n identity matrix will be denoted by I_n.

Notation 3: The n×n matrix consisting of ones in the $(i,i+1)$ position for $i = 1,\ldots,n-1$ and zeroes elsewhere will be denoted by U_n.

Theorem 1: The matrices A and B are similar if and only if A and B have the same elementary divisors.

Theorem 2: A is a cosquare if and only if A^{-1} is similar to A^*. (see [3]).

Theorem 3: For given $n\times n$ matrices A and B, the map $X \to AXB$, determines a linear transformation on the set of $n\times n$ matrices, that is, a vector space of dimension n^2. If the matrix X is rewritten as an $n^2\times 1$ vector of its columns, then the matrix representative of this transformation can be expressed as $B^T \otimes A$. (see [10]).

Theorem 4: Given matrices A and B, $m\times m$ and $n\times n$ resp., then the equation

$$AX - XB = 0$$

has a nontrivial solution X if and only if A and B have an eigenvalue in common. (see [10]).

Theorem 5: The cosquare $A = B^{-1}B^*$ is normal if and only if B^2 is normal. (see [1]).

Theorem 6: Let A be a nonsingular complex $n\times n$ matrix. Let the eigenvalues of AA^* be $\lambda_1^2,\ldots,\lambda_n^2$ with $\lambda_1 > 0,\ldots,\lambda_n > 0$. Let γ be a nonzero number. Then the matrix

$$\begin{bmatrix} 0 & \gamma A \\ A^* & 0 \end{bmatrix}$$

is similar to a diagonal matrix and its eigenvalues are

$$\gamma^{\frac{1}{2}}\lambda_1,\ -\gamma^{\frac{1}{2}}\lambda_1,\ldots,\gamma^{\frac{1}{2}}\lambda_n,-\ \gamma^{\frac{1}{2}}\lambda_n.$$

(see [16]).

CHAPTER II

THE EQUATION $A^{-*}X - XA = 0$

In the work of DePrima and Johnson [3] properties are derived for the matrices A and B where $A = B^{-1}B^{*}$. If $A = B^{-1}B^{*}$ then it is clear that B is a nonsingular solution of the equation (1.1) $A^{-*}X - XA = 0$. In this chapter we investigate all solutions of this equation both in the case that A is a cosquare and in the case that A is not. We also deal with special subsets of solutions other than the set of solutions such that $A = X^{-1}X^{*}$.

Theorem 2.1: Let A be an nxn normal cosquare with distinct eigenvalues $\lambda_1,\bar{\lambda}_1^{-1},\ldots,\lambda_k,\bar{\lambda}_k^{-1},\lambda_{k+1},\ldots,\lambda_s$ where $|\lambda_i| \neq 1$ for $i = 1,\ldots,k$ and $|\lambda_i| = 1$ for $i = k+1,\ldots,s$.

Then, X is a solution of (1.1) if and only if X is unitarily similar by the similarity diagonalizing A to the form

$$\begin{bmatrix} 0 & X_1 \\ Y_1 & 0 \end{bmatrix} \oplus \cdots \oplus \begin{bmatrix} 0 & X_k \\ Y_k & 0 \end{bmatrix} \oplus X_{k+1} \oplus \cdots \oplus X_s$$

where $\dim X_i = \dim Y_i = m_i$, the multiplicity of λ_i in A, for $i = 1,\ldots,k$ and $\dim X_i = m_i$ for $i = k+1,\ldots,s$. Otherwise, X_i and Y_i are arbitrary.

Proof: By the work of DePrima and Johnson [3], A a cosquare implies that

$$A^{*} = SA^{-1}S^{-1}$$

for some matrix S. Thus if λ is an eigenvalue of A, then $\bar{\lambda}^{-1}$ is also an eigenvalue of A with the same multiplicity. Therefore, by a simultaneous unitary similarity of A and X, we may assume that

$$A = \begin{bmatrix} \lambda_1 I_{m_1} & 0 \\ 0 & \bar{\lambda}_1^{-1} I_{m_1} \end{bmatrix} \oplus \cdots \oplus \begin{bmatrix} \lambda_k I_{m_k} & 0 \\ 0 & \bar{\lambda}_k^{-1} I_{m_k} \end{bmatrix} \oplus \lambda_{k+1} I_{m_{k+1}} \oplus \cdots \lambda_s I_{m_s}$$

and

$$X = \left[X_{ij}\right]$$

for i, j = i,...,s where X is partitioned according to the partition of A induced by the direct sum representation given above.

Since $A^{-*}X - XA = 0$,

$$1) \qquad \begin{bmatrix} \bar{\lambda}_i^{-1} I_{m_i} & 0 \\ 0 & \lambda_i I_{m_i} \end{bmatrix} X_{ij} - X_{ij} \begin{bmatrix} \lambda_j I_{m_j} & \\ & \bar{\lambda}_j^{-1} I_{m_j} \end{bmatrix} = 0$$

for i, j = 1,...,k.

$$2) \qquad \begin{bmatrix} \bar{\lambda}_i^{-1} I_{m_i} & 0 \\ 0 & \lambda_i I_{m_i} \end{bmatrix} X_{ij} - X_{ij} \begin{bmatrix} \lambda_j I_{m_j} \end{bmatrix} = 0$$

for i = 1,...,k and j = k+1,...,s.

$$3) \qquad \begin{bmatrix} \lambda_i I_{m_i} \end{bmatrix} X_{ij} - X_{ij} \begin{bmatrix} \lambda_j I_{m_j} & 0 \\ 0 & \bar{\lambda}_j^{-1} I_{m_j} \end{bmatrix} = 0$$

for i = k+1,...,s and j = 1,...,k.

$$4) \qquad \left[\lambda_i I_{m_i}\right] X_{ij} - X_{ij} \left[\lambda_j I_{m_j}\right] = 0$$

for i, j = k+1,...,s.

Furthermore, since the λ_i are distinct for $i \neq j$, the only solution of equations (1) - (4) is the zero matrix. If i = j, then equation (1) is satisfied if and only if

$$X_{ii} = \begin{bmatrix} 0 & X_i \\ Y_i & 0 \end{bmatrix}$$

where X_i and Y_i have dimension $m_i \times m_i$ and are otherwise arbitrary, and equation (4) is satisfied for any arbitrary X_{ii}. Therefore, X is transformed into the required form.

Conversely if A is diagonal and X is in the form derived above

$$A^{-*}X - XA = 0$$

by the derivation above. Therefore, if U is any unitary matrix

$$U(A^{-*}X - XA)U^* = (UAU^*)^{-*}(UXU^*) - (UXU^*)(UAU^*) = 0.$$

Thus, UXU^* is a solution of (1.1) for UAU^*.

<u>Theorem 2.2</u>: Let A be an n×n complex, nonsingular matrix with distinct eigenvalues $\lambda_1, \bar{\lambda}_1^{-1}, \ldots, \lambda_k, \bar{\lambda}_k^{-1}, \lambda_{k+1}, \ldots, \lambda_s, \lambda_{s+1}, \ldots, \lambda_t$ where $|\lambda_i| \neq 1$ for i = 1,...,k, and $|\lambda_i| = 1$ for i = k+1,...,s, and $\lambda_{s+1}, \ldots, \lambda_t$ not covered by the previous two cases.

Then if X is a solution of (1.1), X is congruent to

$$X_1 \oplus X_2 \oplus [0]$$

where

$$X_1 = \bigoplus_{i=1}^{k} \begin{bmatrix} 0 & Z_i \\ Y_i & 0 \end{bmatrix}$$

where Y_i and Z_i are of dimensions $p_i \times m_i$ and $m_i \times p_i$ respectively with m_i and p_i being the multiplicities of λ_i and $\bar{\lambda}_i^{-1}$ in $\sigma(A)$ respectively,

$$X_2 = \bigoplus_{i=k+1}^{s} Z_i$$

where the Z_i are of dimension $m_i \times m_i$, and $[0]$ is of the proper dimension.

Proof: Let S transform A^* into the direct sum of Jordan blocks

$$S^{-1}A^*S = \bigoplus_{i=1}^{t} (\gamma_i I_{n_i} + U_{n_i})$$

where the γ_i are the eigenvalues of A^*, not necessarily distinct.

Then, considering X as an $1 \times n^2$ vector $\bar{x}$ of its columns, the equation $X - A^*XA = 0$ can be rewritten via Theorem 3 as

$$(I_{n^2} - A^T \otimes A^*)\bar{x} = 0.$$

Following the treatment of Davis [2],

$$[(S^{-1} \otimes S^{-1})(I_{n^2} - A^T \otimes A)(S \otimes S)][(S^{-1} \otimes S^{-1})\bar{x}] = 0$$

where

$$(S^{-1} \otimes S^{-1})(I_{n^2} - A^T \otimes A^*)(S \otimes S) = I_{n^2} - [\bigoplus_{i=1}^{t} \gamma_i I_{n_i^2} + U_{n_i}]$$

and $(\bar{S}^{-1} \otimes S^{-1})\bar{x}$ corresponds to $(\bar{S}^{-1})X(\bar{S}^{-1})^*$. Thus, it is sufficient to consider

$$(I_{n^2} - [\bigoplus_{i=1}^{t} \gamma_i I_{n_i} + U_{n_i}] \otimes [\bigoplus_{i=1}^{t} \gamma_i I_{n_i} + U_{n_i}])\bar{x} = 0$$

equivalently

$$X = [\bigoplus_{i=1}^{t} \gamma_i I_{n_i} + U_{n_i}]X[\bigoplus_{i=1}^{t} \gamma_i I_{n_i} + U_{n_i}^T].$$

Let $X = (X_{ij})$ be a partition of X corresponding to the Jordan normal form of A. Then,

$$X_{ij} = B_i X_{ij} B_j^*$$

where

$$B_i = \gamma_i I_{n_i} + U_{n_i}.$$

We may assume that the Jordan blocks of A are ordered according to the distinct eigenvalues of A and that the blocks corresponding to pairs λ_i and $\bar{\lambda}_i^{-1}$ are adjacent. In otherwords, we may partition A into the form

$$(\bigoplus_{i=1}^{k} B_i) \oplus (\bigoplus_{i=k+1}^{k} B_i)$$

where

$$B_i = \begin{bmatrix} B_{i1} & 0 \\ 0 & B_{i2} \end{bmatrix}$$

for $i = 1,\ldots,k$ such that $\sigma(B_{i1}) = \lambda_i$ and $\sigma(B_{i2}) = \bar{\lambda}_i^{-1}$.

Partition $X = (X_{ij})$ according to the B_i. Then, $X_{ij} = 0$ for $i \neq j$ as in Theorem 2.1. Now partition each X_{ii} according to the natural partition of B_i,

$$X_{ii} = \begin{bmatrix} X_{ii}^{(1)} & X_{ii}^{(2)} \\ X_{ii}^{(3)} & X_{ii}^{(4)} \end{bmatrix}.$$

Then

$$B_i^* X_{ii} B_i = \begin{bmatrix} B_{i1}^* A_{ii}^{(1)} B_{i1} & B_{i1}^* A_{ii}^{(2)} B_{i2} \\ B_{i2}^* A_{ii}^{(3)} B_{i1} & B_{i2}^* A_{ii}^{(4)} B_{i2} \end{bmatrix}.$$

Thus $X - A^* X A = 0$ implies that $X_{ii}^{(1)}$ and $X_{ii}^{(4)} = 0$ for $i = 1,\ldots,k$. Therefore, X is congruent to the desired form.

Theorem 2.3: There exists a nonsingular, normal matrix X satisfying (1.1) if and only if A is unitarily similar to a $k \times k$ block matrix $A = (A_{ij})$ where k is the number of distinct eigenvalues $\lambda_1,\ldots,\lambda_k$ of X and the dimension of A_{ii} is the multiplicity of λ_i, and letting $A^{-*} = (B_{ij})$ for $i,j = 1,\ldots,k$ be a conforming partition of A^{-*},

$$B_{ij} = \gamma_{ij} A_{ij}$$

with $\gamma_{ij} = \lambda_i / \lambda_j$ for $i,j = 1,\ldots,k$.

Furthermore, the above similarity diagonalizes the nonsingular, normal solution X.

Proof: Assume that there exists a nonsingular, normal solution X of (1.1). Then, by a unitary transformation of A and X, the matrix X may be assumed to be in diagonal form with the eigenvalues of X ordered such that equal eigenvalues are adjacent on the diagonal.

Partition A and A^{-*} according to the blocks of equal eigenvalues in X. Let $A = (A_{ij})$ and $A^{-*} = (B_{ij})$. Then (1.1) implies

1) $$A^{-*}D - DA = 0$$

where A and D are the transformed A and X respectively. Thus

$$A_{ij} = \lambda_j/\lambda_i B_{ij}$$

where λ_i for $i = 1,\ldots,k$ is the ordering of the eigenvalues of X given by the particular unitary transformation.

Conversely, let

$$X = \bigoplus_{i=1}^{k} \lambda_i I_{m_i}$$

with λ_i distinct and nonzero for $i = 1,\ldots,k$. Let A be in the k×k block matrix form of the theorem. Then,

$$A^{-*}X - XA = 0.$$

If U is any unitary matrix, then

$$A_1^{-*}X_1 - X_1A_1 = 0$$

where $A_1 = UAU^*$ and $X_1 = UXU^*$. Furthermore, X_1 is normal.

<u>Theorem 2.4</u>: There exists a normal, singular matrix X satisfying (1.1) if and only if A is unitarily similar to

$$A = \begin{bmatrix} A_1 & 0 \\ A_3 & A_4 \end{bmatrix}$$

where A_1 and A_4 are square, nonsingular matrices and the dimension of A_4 is the nullity of X and A^{-*} is simultaneously similar to

$$A^{-*} = \begin{bmatrix} A_1^{-*} & B_2 \\ 0 & A_4^{-*} \end{bmatrix}$$

where A^{-*} is partitioned as A. Here A_1 and A_1^{-*} have the same structure as A and A^{-*} in Theorem 2.3 with respect to the nonzero eigenvalues of X and A_4 and A_4^{-*} correspond to the zero eigenvalues of X.

In addition,

$$A_1^{-*}A_3^{*} + B_2A_4^{*} = 0$$

and

$$A_1^{*}B_2 + A_3^{*}A_4^{-*} = 0.$$

Furthermore, the above similarity diagonalizes X.

<u>Proof</u>: The proof of Theorem 2.4 follows that of Theorem 2.3 until equation (1) $A^{-*}D - DA = 0$. In this case,

$$D = \left(\bigoplus_{i=1}^{k} \lambda_i I_{m_i} \right) \oplus \begin{bmatrix} 0 \end{bmatrix} = \begin{bmatrix} D_1 & 0 \\ 0 & 0 \end{bmatrix}.$$

Let

$$A = \begin{bmatrix} A_1 & A_2 \\ A_3 & A_4 \end{bmatrix}$$

and

$$A^{-*} = \begin{bmatrix} B_1 & B_2 \\ B_3 & B_4 \end{bmatrix}$$

be partitioned as D. Thus $A^{-*}D - DA = 0$ implies that

1) $$B_1 D_1 - D_1 A_1 = 0$$

2) $$0 - D_1 A_2 = 0$$

3) $$B_3 D_1 - 0 = 0$$

4) $$0 - 0 = 0.$$

Equations (2) and (3) give the correct zero blocks in A and A^{-*}, and the dimension of A_4 is the multiplicity of 0 as an eigenvalue of X.

Furthermore,

$$A^* A^{-*} = A^{-*} A^* = I.$$

Thus

$$A^{-*} = \begin{bmatrix} A_1^{-*} & B_2 \\ 0 & A_4^{-*} \end{bmatrix}$$

where

$$A_1^{-*} A_3^* + B_2 A_4^* = 0$$

and

$$A_1^* B_2 + A_3^* A_4^{-*} = 0$$

Then (1) and the fact that $B_1 = A_1^{-*}$ implies

$$A_1^{-*}D_1 - D_1A = 0.$$

Thus A_1 and A_1^{-*} have the same structure as A and A^{-*} in Theorem 2.3 for the nonzero eigenvalues of X.

Conversely, let

$$X = \left(\bigoplus_{i=1}^{k} \lambda_i I_{m_i} \right) \oplus \left[0 \right]$$

with all λ_i distinct and nonzero for i = 1,...,k. Let A have the block matrix form of the theorem. Then,

$$A^{-*}X - XA = 0.$$

If U is any unitary matrix, then

$$A_1^{-*}X_1 - X_1A_1 = 0$$

where $A_1 = UAU^*$ and $X_1 = UXU^*$. Furthermore, X is normal and singular.

<u>Corollary 2.5</u>: Let X be a normal, singular matrix satisfying equation (1.1) for some matrix A. Then there exist matrices B and C such that

$$B^{-*}X - XB = 0$$

and

$$C^{-*}X - XC = 0$$

with B a cosquare, but C not a cosquare.

<u>Proof</u>: In Theorem 2.4, let A_3 and B_2 be zero matrices. Choose

$$B = U \begin{bmatrix} A_1 & 0 \\ 0 & A_4 \end{bmatrix} U^*$$

with A_1 and A_4 cosquares. Then

$$B^{-*}X - XB = 0$$

and B is a cosquare.

Choose

$$C = U \begin{bmatrix} A_1 & 0 \\ 0 & A_4 \end{bmatrix} U^*$$

with A_1 a cosquare and A_4 nonsingular, but not a cosquare. Then

$$C^{-*}X - XC = 0$$

but C is not a cosquare.

CHAPTER III

COSQUARES

The aim of this chapter is to apply the results of Chapter II to the special case of solutions X such that $A = X^{-1}X^*$. Notice that this is a nontrivial restriction since for a cosquare A, there are always solutions Y of (1.1) such that $A \neq Y^{-1}Y^*$. This is easily seen since in (1.1) the set of solutions is closed under scalar multiplication, but the set of all matrices X such that $A = X^{-1}X^*$ is not closed under scalar multiplication.

Theorem 3.1 is a special case of Theorem 2.1. This result is then used to derive spectral restrictions for the components of a cosquare.

<u>Theorem 3.1</u>: Let $N = A^{-1}A^*$ then N is normal if and only if A is unitarily similar to

$$\bigoplus_{i=1}^{k} \begin{bmatrix} 0 & \gamma_i B_i \\ B_i^* & 0 \end{bmatrix} \oplus \bigoplus_{i=k+1}^{s} \varphi_i I_{m_i}$$

where B_i is a nonsingular matrix for all i, and γ_i all distinct and γ_i and φ_i nonzero for all i.

<u>Proof</u>: By the work of DePrima and Johnson [3], $N = A^{-1}A^*$ implies $N^* = SN^{-1}S^{-1}$ for some matrix S. Thus if γ is an eigenvalue of N, then $\bar{\gamma}^{-1}$ is an eigenvalue of N with the same multiplicity. Therefore,

since N is normal, N may be assumed, by a unitary similarity, to be of the form

$$\begin{bmatrix} \gamma_1 I_{m_1} & 0 \\ 0 & \bar{\gamma}_1^{-1} I_{m_1} \end{bmatrix} \oplus \cdots \oplus \begin{bmatrix} \gamma_k I_{m_k} & 0 \\ 0 & \bar{\gamma}_k^{-1} I_{m_k} \end{bmatrix} \oplus \xi_{k+1} I_{m_{k+1}} \oplus \cdots \oplus \xi_s I_{m_s}$$

with γ_i and ξ_i all distinct and $|\gamma_i| \neq 1$ and $|\xi_i| = 1$.

Notice that $A^{-1}N^{-*} = NA^{-1}$, thus

$$A^{-1} = \bigoplus_{i=1}^{k} \begin{bmatrix} 0 & A_i \\ B_i & 0 \end{bmatrix} \oplus A_{k+1} \oplus \cdots \oplus A_s$$

by Theorem 2.1. Furthermore, $N^{-1}A^{-1} = A^{-*}$. Therefore, $B_i^* = \gamma_i^{-1} A_i$ for $i = 1,\ldots,k$ and $\xi_i^{-1} A_i = A_i^*$ for $i = k+1,\ldots,s$. Thus A^{-1} is unitarily similar to

$$\bigoplus_{i=1}^{k} \begin{bmatrix} 0 & A_i \\ \gamma_i^{-1} A_i^* & 0 \end{bmatrix} \oplus \bigoplus_{i=k+1}^{s} \varphi_i^{-1} I_{m_i}$$

where $\varphi_j = \bar{\xi}_j / \xi_j$ and, A is unitarily similar to

$$\bigoplus_{i=1}^{k} \begin{bmatrix} 0 & \gamma_i A_i^{-*} \\ A_i^{-1} & 0 \end{bmatrix} \oplus \bigoplus_{i=k+1}^{s} \varphi_i I_{m_i} .$$

Conversely, let UAU^* be of the form required by the theorem.

Then

$$UA^{-1}A^{*}U^{*} = (UAU^{*})^{-1}(UAU^{*})^{*} =$$

$$= \left(\bigoplus_{i=1}^{k} \begin{bmatrix} 0 & B_i \\ \gamma^{-1}B_i^{*} & 0 \end{bmatrix} \oplus \bigoplus_{i=k+1}^{s} \varphi_i^{-1} I_{m_i} \right) \times$$

$$\left(\bigoplus_{i=1}^{k} \begin{bmatrix} 0 & B_i^{-*} \\ \bar{\gamma}_i B_i^{-1} & 0 \end{bmatrix} \oplus \bigoplus_{i=k+1}^{s} \bar{\varphi}_i I_{m_i} \right) =$$

$$= \bigoplus_{i=1}^{k} \begin{bmatrix} \bar{\gamma}_i I_{m_i} & 0 \\ 0 & \gamma_i^{-1} I_{m_i} \end{bmatrix} \oplus \bigoplus_{i=k+1}^{s} \frac{\bar{\varphi}_{k+1}}{\varphi_{k+1}} .$$

Therefore, $A^{-1}A^{*} = N$ is normal.

Corollary 3.2: Let $N = A^{-1}A^{*}$ be normal. Then

1) $$\sigma(N) = \{\gamma_1, \bar{\gamma}_1^{-1}, \ldots, \gamma_k, \bar{\gamma}_k^{-1}, \xi_{k+1}, \ldots, \xi_s\}$$

where $|\xi_i| = 1$ for $i = k+1, \ldots, s$ and

2) $$\sigma(A) = \{\lambda_{1j}\gamma_1^{\frac{1}{2}}, -\lambda_{1j}\gamma_1^{\frac{1}{2}}, \ldots, \lambda_{kj}\gamma_k^{\frac{1}{2}}, -\lambda_{kj}\gamma_k^{\frac{1}{2}}, \varphi_{k+1}, \ldots, \varphi_s\}$$

where $j = 1, \ldots, m_i$ and $\lambda_{ij} \in R^{+}$ for $i = 1, \ldots, k$ and $j = 1, \ldots, m_j$.

Proof: Part (1) follows from the work of DePrima and Johnson [3]. Part (2) follows from the theorem and Lemma 3.5 of Thompson [16] (see Theorem 6 of chapter 1 above).

The final two theorems of this chapter give necessary and sufficient conditions for a cosquare to be normal or unitary in terms of the set of solutions of (1.1).

<u>Theorem 3.3</u>: Let A be a cosquare. Let $\Omega = \{X: X^{-1}AX = A^{-*}\}$. Then the following are equivalent.

1) A is unitary.

2) There exists a pair of matrices S and $S^{-1}S^{*} \in \Omega$.

3) For each $S \in \Omega$, the matrix $S^{-1}S^{*} \in \Omega$.

<u>Proof</u>: Clearly, if either of (1) or (3) holds, then (2) holds.

Conversely, let $A^{-*} = SAS^{-1}$ then

$$A = S^{-1}A^{-*}S \text{ and } A = S^{-*}A^{-*}S^{*}.$$

Thus

$$A^{-*} = S^{*}AS^{-*} \text{ and } A^{-*} = SAS^{-1}.$$

Therefore

$$S^{*}AS^{-*} = SAS^{-1}$$

or

$$S^{-1}S^{*}A = AS^{-1}S^{*}.$$

Therefore, if $A^{-*} = S^{-1}S\,A(S^{-1}S^{*})^{-1}$, then $A^{-*} = A$ and A is unitary.

In otherwords, if there exists a single pair S and $S^{-1}S^{*} \in \Omega$ then A is unitary. Therefore, (2) $\Rightarrow$ (1).

Let $T \in \Omega$ be another matrix. Then the above argument shows that

$$T^{-1}T^{*}A = AT^{-1}T^{*}$$

or

$$T^{-1}T^{*}A(T^{-1}T^{*})^{-1} = A.$$

However A is unitary. Thus

$$T^{-1}T^{*}A(T^{-1}T^{*})^{-1} = A^{-*}$$

and $T^{-1}T^{*} \in \Omega$. Therefore, (3) holds.

Theorem 3.4: Let $A = B^{-1}B^{*}$. Let $\Omega = \{X: X^{-1}AX = A^{-*}\}$. Then the following are equivalent.

1) A is normal.

2) B^2 is normal.

3) For each $X \in \Omega$, the matrix $X^{-*} \in \Omega$.

Proof: (1) $\Rightarrow$ (3). Taking the $*^{-1}$ of a solution preserves the structure required by Theorem 2.1. Therefore, for each $X \in \Omega$, the matrix $X^{-*} \in \Omega$.

(3) $\Rightarrow$ (2). Since B is a solution of (1.1), B^{-*} is also a solution of (1.1). Therefore,

$$(B^{*}B^{-1})B^{-*} - B^{-*}(B^{-1}B^{*}) = 0$$

or

$$(B^{*})^{2}B^{-1} - B^{-1}(B^{*})^{2} = 0.$$

Thus

$$B(B^{*})^{2} - (B^{*})^{2}B = 0.$$

Therefore, B^2 is normal.

(2) $\Rightarrow$ (1). From Choi [1].

Choi [1] shows that A is normal if and only if B^2 is normal. In his proof he applies the Fuglede-Putnam theorem [9] to show that

(1) $\Rightarrow$ (2). Theorem 3.4 gives an alternate proof of this result in the finite dimensional case without the use of the Fuglede-Putnam theorem [9].

CHAPTER IV

THE EQUATION $P(A)X - XA = 0$

The aim of this chapter is to develop decomposition theorems like those of Chapter II for the matrix equation

$$P(A)X - XA = 0 \tag{1.2}$$

where P is a polynomial. These results will be used in Chapter V to prove an analogue of Theorem 3.4. We also develop spectral restrictions for solutions of (1.2). In all of these theorems, we will assume that there exists a nonsingular solution of (1.2). As a result we will use the following lemma.

Lemma 4.1: Let A be a matrix, $P(X)$ a polynomial such that equation (1.2) has a nonsingular solution. Then $P(x)$ is a one - to - one mapping of $\sigma(A) = \{\lambda_i: i=1,\ldots,s \text{ and } \lambda_i \text{ distinct}\}$ onto itself. In particular, if $P(\lambda_i) = \lambda_j$ then the multiplicity of λ_i equals the multiplicity of λ_j, and the Jordan normal forms of A and $P(A)$ are the same.

Proof: If $P(A)X - XA = 0$ has a nonsingular solution, then $P(A)$ and A are similar. Therefore, $\{P(\lambda_i): i=1,\ldots,s\} = \{\lambda_i: i=1,\ldots,s\}$ including multiplicities. Furthermore, the Jordan normal form is preserved.

Theorem 4.2: Let A be a diagonalizable matrix, $P(X)$ a polynomial such that equation (1.2) has a nonsingular solution and $\sigma(A) = \{\lambda_i: i=1,\ldots,s\}$ with λ_i distinct and the multiplicity of λ_i being m_i. Let α be a permutation of $1,\ldots,s$ defined by $P(\lambda_i) = \lambda_{\alpha(i)}$, with cycle decomposition $\alpha_1\ldots\alpha_k$. Assume that the eigenvalues of A are numbered according to this cycle decomposition. Let S be a similarity which diagonalizes A and preserves the ordering of the eigenvalues, $\lambda_1,\ldots,\lambda_s$.

Then X is a solution of (1.2) if and only if SXS^{-1} is some block monomial matrix corresponding to α with diagonal blocks of dimension m_i.

<u>Proof</u>: The assumption that A is diagonalizable implies that, by a similarity S, the equation $P(A)X - XA = 0$ may be reduced to

$$P(D)Y - YD = 0$$

where $Y = SXS^{-1}$ and D is a diagonal matrix of the eigenvalues of A with eigenvalues arranged according to the ordering $\lambda_1,\ldots,\lambda_s$. Partition Y into a block matrix (Y_{ij}) and D into (D_i) according to the blocks of equal eigenvalues in D. Thus $P(D)Y - YD = 0$ implies

$$P(\lambda_i)Y_{ij} - Y_{ij}\lambda_j = 0$$

for $i,j = 1,\ldots,s$. Therefore, $Y_{ij} = 0$ for $P(\lambda_i) \neq \lambda_j$, ie. $j \neq \alpha(i)$ and $i,j = 1,\ldots,s$. This is the desired form for Y and results in the desired form for X.

Notice that the particular form of S is not used. Therefore, any similarity of this type results in the same decomposition.

Conversely, if $Y_{ij} = 0$ for $j \neq \alpha(i)$ and $i,j = 1,\ldots,s$, then

$$P(\lambda_i)Y_{ij} - Y_{ij}\lambda_j = 0$$

for $i,j = 1,\ldots,s$. Therefore, $P(D)Y - YD = 0$ and hence $P(A)X - XA = 0$

where $X = S^{-1}DS$ for some matrix S.

Theorem 4.3: Let A and P(A) be as in Theorem 4.2 except that A is not assumed to be diagonalizable. Let S be a similarity transforming A into Jordan normal form and arranging the Jordan blocks according to the ordering $\lambda_1,\ldots,\lambda_s$.

Then X is a solution of equation (1.2) if and only if $SXS^{-1} = (Y_i)$ is a block monomial matrix corresponding to α with diagonal blocks of dimension m_i, and

$$P(A_i)Y_i - Y_iA_j = 0$$

for $\alpha(i) \neq j$. Here A_i is the submatrix of the Jordan normal form of A consisting of the direct sum of the Jordan blocks corresponding to the eigenvalue λ_i.

Proof: Let $P(A)X - XA = 0$, then

$$SP(A)XS^{-1} - SXAS^{-1} = 0$$

or

$$(SP(A)S^{-1})(SXS^{-1}) - (SXS^{-1})(SAS^{-1}) = 0.$$

Therefore, A may be assumed to be in Jordan normal form with eigenvalue blocks arranged according to the ordering $\lambda_1,\ldots,\lambda_s$. P(A) will be in a corresponding block diagonal form. Partitioning $A = (A_i)$ and $Y = (Y_{ij})$ according to the blocks of equal eigenvalues in A implies that

$$P(A_i)Y_{ij} - Y_{ij}A_j = 0$$

for $i,j = 1,\ldots,s$. Then, $Y_{ij} = 0$ for $\alpha(i) \neq j$, since in this case $P(A_i)$ and A_j have no eigenvalues in common.

Conversely, if Y is a block monomial matrix corresponding to the permutation α and

$$P(A_i)Y_i - Y_iA_j = 0$$

for $\alpha(i) = j$ where $i,j = 1,\ldots,s$ then $P(A)Y - YA = 0$ and thus $P(A)X - XA = 0$ where $X = S^{-1}YS$.

In [15], Taussky-Todd investigates the connection between equation 1.2 and Galois theory. The following example illustrates this connection.

Example: Let ξ be a seventh root of unity, then ξ satisfies:

$$x^6 + x^5 + x^4 + x^3 + x^2 + x + 1 = 0.$$

Furthermore, $\alpha_1 = \xi + \bar{\xi}$ will satisfy:

$$f(x) = x^3 + x^2 - 2x - 1 = 0$$

and there exist rational polynomials $p_1(x)$ and $p_2(x)$ such that $f(x)$ has roots α_1, $p_1(\alpha_1)$, and $p_2(\alpha_1)$ where

$$p_1(x) = x^2 - 2$$

and

$$p_2(x) = -x^2 - x + 3.$$

$p_1(x)$ and $p_2(x)$ act as cyclic permutations on the roots of $f(x)$. Thus if A is any 3×3 matrix with characteristic polynomial $f(x)$, then by the above theorem $p_i(A)X - XA = 0$ has solutions X where

$$S_i X S_i^{-1} = \begin{bmatrix} 0 & x_1 & 0 \\ 0 & 0 & x_2 \\ x_3 & 0 & 0 \end{bmatrix}$$

and

$$S_i A S_i^{-1} = \begin{bmatrix} \alpha_1 & 0 & 0 \\ 0 & p_i(\alpha_1) & 0 \\ 0 & 0 & p_i(p_i(\alpha_1)) \end{bmatrix}$$

with x_1, x_2 and x_3 arbitrary, $i=1,2$.

We will now derive spectral restrictions for the solutions of equation (1.2). For this we need the following lemma.

Lemma 4.4: Let $A = (A_{ij})$ be a block monomial matrix corresponding to a cyclic permutation. Let the diagonal blocks of A be square of dimension n_i for $i=1,\ldots,s$. Then

$$\det(\lambda I - A) = \lambda^{n_2+\cdots+n_s-(s-1)n_1} \det\left(\lambda^s I_{n_1} - \left(\prod_{i=1}^{s-1} A_{i,i+1}\right) A_{s1}\right).$$

Proof: By a similarity transformation, we may replace A by

$$A' = SAS^{-1} = \begin{bmatrix} 0 & A_{12} & 0 & \cdots & 0 \\ \vdots & 0 & & \ddots & \vdots \\ & & & \ddots & 0 \\ & & \ddots & & \\ 0 & & & & A_{s-1,s} \\ A_{s1} & 0 & & \cdots & 0 \end{bmatrix}$$

Then

$$\lambda I - A' = \begin{bmatrix} \lambda I_{n_1} & -A_{12} & 0 & & 0 \\ 0 & & & \ddots & \vdots \\ \vdots & & & \ddots & 0 \\ 0 & & \ddots & \ddots & -A_{s-1,s} \\ -A_{s1} & 0 & \cdots & 0 & \lambda I_{n_s} \end{bmatrix}.$$

Multiplying on the right by

$$\begin{bmatrix} I_{n_1} & 0 & & \cdots & 0 \\ 0 & & & & \vdots \\ \vdots & & & \ddots & \\ 0 & & \ddots & & 0 \\ \lambda^{-1}A_{s1} & 0 & \cdots & 0 & I_{n_s} \end{bmatrix}$$

gives

$$|\lambda I - A'| = \begin{vmatrix} \lambda I_{n_1} & -A_{12} & 0 & \cdots & 0 \\ 0 & & & \ddots & \vdots \\ \vdots & & & \ddots & \\ 0 & & \ddots & & 0 \\ -A_{s1}^{(1)} & & & & -A_{s-1,s} \\ 0 & & \cdots & 0 & \lambda I_{n_s} \end{vmatrix}$$

where $-A_{s1}^{(1)} = -\lambda^{-1} A_{s-1,s}A_{s1}$. Notice that the $(s-1)\times(s-1)$ principal submatrix of the right hand side is an $(s-1)\times(s-1)$ block matrix of the same type as $\lambda I - A$. Therefore,

$$|\lambda I - A| = |\lambda I_{n_s}| \begin{vmatrix} \lambda I_{n_1} & -A_{12} & 0 & \cdots & 0 \\ 0 & & & \ddots & \vdots \\ \vdots & & \ddots & & 0 \\ 0 & & \ddots & & -A_{s-2,s-1} \\ -A_{s1}^{(1)} & 0 & \cdots & 0 & \lambda I_{n_{s-1}} \end{vmatrix} .$$

Iterating this procedure s - 2 times yields

$$|\lambda I - A| = \lambda^{n_s+n_{s-1}+\cdots+n_3} \begin{vmatrix} \lambda I_{n_1} & -A_{12} \\ -A_{s1}^{(s-2)} & \lambda I_{n_2} \end{vmatrix}$$

where $-A_{s1}^{(s-2)} = -\lambda^{-(s-2)} \prod_{i=1}^{s-2} (A_{i+1,i+2}) A_{s1}$. Multiplying

$$\begin{bmatrix} \lambda I_{n_1} & -A_{12} \\ -A_{s1}^{(s-2)} & \lambda I_{n_2} \end{bmatrix}$$

on the right by

$$\begin{bmatrix} I_{n_1} & 0 \\ \lambda^{-1} A_{s1}^{(s-2)} & I_{n_2} \end{bmatrix}$$

gives

$$\begin{vmatrix} \lambda I_{n_1} & -A_{12} \\ -A_{s1}^{(s-2)} & \lambda I_{n_2} \end{vmatrix} = \begin{vmatrix} \lambda I_{n_1} - A_{12}\lambda^{-1} A_{s1}^{(s-2)} & -A_{12} \\ 0 & \lambda I_{n_2} \end{vmatrix} .$$

Therefore,

$$\begin{aligned} |\lambda I - A| &= \lambda^{n_s + n_{s-1} + \dots + n_3} \lambda^{n_2} |\lambda I_{n_1} - \lambda^{-1} A_{12} A_{s1}^{(s-2)}| = \\ &= \lambda^{n_s + n_{s-1} + \dots + n_2} |\lambda I_{n_1} - \lambda^{-s+1} \prod_{i=1}^{s-1} A_{i,i+1} A_{s1}| = \\ &= \lambda^{n_2 + \dots + n_s - (s-1) n_1} |\lambda^s I_{n_1} - \prod_{i=1}^{s-1} A_{i,i+1} A_{s1}| . \end{aligned}$$

<u>Lemma 4.5</u>: Let A and B be $\ell \times \ell$ Jordan blocks for the eigenvalues λ_1 and λ_2 respectively. Let $P(x)$ be a polynomial such that $P(\lambda_1) = \lambda_2$.

Then $X = (x_{ij})$ is a solution of $P(A)X - XB = 0$ if and only if X satisfies the following.

1) X is upper triangular.

2) $X_{ii} = P'(\lambda_1)X_{i+1,i+1}$ for $i = 1,\ldots,\ell-1$.

3) $$X_{i,i+j} = P'(\lambda_1)X_{i+1,i+j+1} + \sum_{k=2}^{j+1} \frac{P^{(k)}(\lambda_1)}{k!} X_{i+k,i+j+1}$$

for $i = 1,\ldots,\ell-1$ and $j = 1,\ldots,\ell-i-1$.

<u>Proof</u>: Let J_j be the matrix consisting of all zeroes except for ones on the jth upper diagonal. Then

$$P(A) = P(\lambda_1)I + P'(\lambda_1)J + \frac{P^{(2)}(\lambda_1)}{2!} J_2 + \ldots + \frac{P^{(\ell-1)}(\lambda_1)}{(\ell-1)!} J_{\ell-1} .$$

Thus $P(A)X - XB = 0$ implies

$$\left(P(\lambda_1)I + \ldots + \frac{P^{(\ell-1)}(\lambda_1)}{(\ell+1)!} J_{\ell-1}\right) X - X\left(\lambda_2 I + J_1\right) = 0$$

or

$$\left(P'(\lambda_1)J_1 + \ldots + \frac{P^{(\ell-1)}(\lambda_1)}{(\ell-1)!} J_{\ell-1}\right) X - XJ_1 = 0.$$

Comparing the entries of the matrices on the left and right hand sides row by row, starting with the last row and working up, shows that X satisfies conditions (1), (2) and (3).

Conversely, these conditions may be seen to be sufficient by multiplying out the left hand side.

Theorem 4.6: Let A and α be as in Theorem 4.2. In addition let $n^{(i)}$ be the multiplicity of any eigenvalue corresponding to the cycle α_i. Let X be any particular solution of $P(A)X - XA = 0$. Then there exist complex scalars b_{ij} for $i = 1,\dots,k$ and $j = 1,\dots,n^{(i)}$ such that

$$\sigma(X) = \bigcup_i \left(\bigcup_j \{\lambda : \lambda^{|\alpha_i|} = b_{ij}\} \right)$$

Conversely, for any set of $b_{ij} \in \mathbb{C}$, for $i = 1,\dots,k$ and $j = 1,\dots,n^{(i)}$ there exists a solution Y_1 having this set as its spectrum. In particular, for any prescribed set of $\sum_{i=1}^{k} n^{(i)}$ eigenvalues, there exists a solution Y_2 with $\sigma(Y_2)$ containing these eigenvalues.

Theorem 4.7: Let A be as in Theorem 4.6 except that A is nonderogatory, but not assumed to be diagonalizable. Let X be any particular solution of $P(A)X - XA = 0$. Then there exist complex scalars b_{ij} for $i = 1,\dots,k$ and $j = 1,\dots,n^{(i)}$ such that

$$\sigma(X) = \bigcup_i \left(\bigcup_j \{\lambda : \lambda^{|\alpha_i|} = b_{ij}\}\right) .$$

Conversely, for any set of $b_{ij} \in \mathbb{C}$ with $i = 1,\dots,k$ and $j = 1,\dots,n^{(i)}$ and

$$b_{ij} = \prod_{\ell=1}^{|\alpha_i|} c_{i\ell} P'(\lambda_{i\ell})^{n^{(i)}-j}$$

where $c_{i\ell} \in \mathbb{C}$, and $\lambda_{i\ell}$ is the ordering of the eigenvalues of A based on the cycle decomposition $\alpha_1 \cdots \alpha_k$, there exists a solution Y_1 having this set as its spectrum. In particular, for any prescribed set of k eigenvalues, there exists a nondiagonalizable solution Y_2, with $\sigma(Y_2)$ containing these eigenvalues.

<u>Proof of Theorem 4.6 and Theorem 4.7</u>: As in Theorem 4.3, reduce X to the direct sum of block monomial matrices, each corresponding to a single cycle α_i. A further permutation similarity will put each $X^{(i)}$ into the form

$$\begin{bmatrix} 0 & X_1^{(i)} & 0 & \cdots & 0 \\ \vdots & & & \ddots & \vdots \\ & & & \ddots & \\ & & \ddots & & 0 \\ 0 & & & & X^{(i)}_{|\alpha_i|-1} \\ X_{|\alpha_i|} & 0 & & \cdots & 0 \end{bmatrix}$$

where each $X_j^{(i)}$ is square, and s×s by Lemma 4.5. Therefore, Lemma 4.4 states that the characteristic polynomial of $X^{(i)}$ is

$$\lambda^{n_i-|\alpha_i|m_i} \left| \lambda^{|\alpha_i|} I_s - \prod_{j=1}^{|\alpha_i|} X_j^{(i)} \right| .$$

If A is diagonalizable, then the $X_j^{(i)}$ may be chosen arbitrarily and Theorem 4.6 follows.

If A is nonderogatory, then each $X_j^{(i)}$ is triangular with diagonal

$$(cP'(\lambda)^{n^{(i)}-1},\ cP'(\lambda)^{n^{(i)}-2},\ \ldots\ ,\ c)$$

by Theorem 4.3 and Lemma 4.5. Therefore, $\prod_{j=1}^{|\alpha_i|} X_j^{(i)}$ is triangular and has diagonal

$$\left(\prod_{j=1}^{|\alpha_i|} c_{ij}P'(\lambda_{ij})^{n^{(i)}-1},\ \ldots\ ,\ \prod_{j=1}^{|\alpha_i|} c_{ij}\right)$$

where $c_{ij} \in \mathbb{C}$. Thus Theorem 4.7 follows.

<u>Theorem 4.8</u>: Let A be as in Theorem 4.7 except that A is derogatory. In addition, let m_i be the number of elementary divisors for any eigenvalue corresponding to the cycle α_i and $m = \sum_{i=1}^{k} m_i$. Then for any prescribed set of m complex numbers, there exists a matrix Y with $P(A)Y - YA = 0$ and $\sigma(Y)$ containing this set.

<u>Proof</u>: Reduce X as in Theorem 4.3. Then, by Lemma 4.4, the characteristic polynomial of $X^{(i)}$ is

$$\lambda^{n_i-|\alpha_i|m_i} \left| \lambda^{|\alpha_i|} I_s - \prod_{j=1}^{|\alpha_i|} X_j^{(i)} \right|$$

where $X_j^{(i)}$ for $j = 1,\ldots,|\alpha_i|$ are the blocks of the block monomial matrix $X^{(i)}$ which are not required to be zero. By Lemma 4.1, these are all square and of the same dimension. Partition each $X_j^{(i)}$ according to the Jordan structure of the corresponding eigenvalue in A. Then by Theorem 4.3 and Lemma 4.5, each block in $X_j^{(i)}$ is an upper triangular matrix whose final column may be chosen arbitrarily and still give a solution of $P(A)X - XA = 0$.

In particular, by choosing the final columns of the lower blocks to be $(0, \ldots, 0)^T$, the resulting matrix will be upper triangular with each diagonal block contributing one arbitrary element to the diagonal. Therefore, m_i of the eigenvalues of a solution may be chosen arbitrarily for each cycle α_i. Thus a matrix Y may be constructed with any set of m eigenvalues in $\sigma(Y)$ and $P(A)Y - YA = 0$.

CHAPTER V

TOWERS IN $P(A)X - XA = 0$

In this chapter, we wish to derive a result for the equation

$$P(A)X - XA = 0 \tag{1.2}$$

which is analogous to Theorem 3.4. The natural analogue is to consider

Problem 1.3: For what polynomials P does the fact that there exists a tower for each solution of (1.2) imply that A is normal?

To attack this problem, we derive conditions on P which are necessary and sufficient for the existence of towers.

Theorem 5.1: Let A be a diagonalizable matrix, $P(X)$ a polynomial such that equation (1.2) has a nonsingular solution, and $\sigma(A) = \{\lambda_i : i = 1,\dots,s\}$ with λ_i distinct. Then there exists a tower for each solution of (1.2) if and only if the following two conditions hold.

1) $P(\lambda_i) = \lambda_{\alpha(i)}$ for some permutation α with order say t.

2) $P(X) = \sum_{i=0}^{\ell} a_i X^{it+1}$ for all solutions X of (1.2).

Proof: Since $P(A)X - XA = 0$ for some nonsingular X, condition (1) is necessary. By Theorem 4.2, $P(X)$ is a solution for all solutions X if and only if $P(X)$ and X are simultaneously similar, by the similarity given in Theorem 4.2, to the same block monomial form. In other words, there exists a matrix S such that

$$\left[SP(X)S^{-1}\right]_{ij} = 0$$

for all $i,j = 1,\ldots,s$ with $i \neq \alpha(j)$ and for all solutions X of (1.2).

Let $P(X) = \sum_{k=0}^{\ell} c_k X^k$ for $\ell \in I^+$. Each X^k must reduce to the same block monomial form as X and P(X). Otherwise, $(SX^kS^{-1})_{ij} \neq 0$ for some i, j, k and X with $i \neq \alpha(j)$. However, since SXS^{-1} is a block monomial matrix, each block $(SX^kS^{-1})_{ij}$ is the product of blocks of SXS^{-1}. In particular, if $(SX^kS^{-1})_{ij} \neq 0$, the blocks are some of the nonzero blocks $X_1, \ldots, X_s$ of SXS^{-1}. However, by Theorem 4.2, these blocks may be chosen arbitrarily. Therefore, if $(SX^kS^{-1})_{ij} \neq 0$ for this i and j, then $P(X) = 0$. This is a contradiction. Since X^k must reduce to the same block monomial form as X, we have $k = it+1$ where t is the order of the permutation α, and $i \in I^+$. Therefore, condition (2) is necessary.

Conversely, conditions (1) and (2) result in a matrix P(X) which reduces to the same block monomial form as X. Therefore, by Theorem 4.2, P(X) is also a solution.

Corollary 5.2: Let P be a polynomial and A be a normal matrix with $\sigma(A) = \{\lambda_i : i = 1, \ldots, s\}$ with λ_i distinct such that there exists a tower for each solution of (1.2). Then there exists a diagonalizable, but nonnormal matrix B such that there exists a tower for each solution of $P(B)X - XB = 0$.

Proof: Since towers exist for solutions of $P(A)X - XA = 0$, P satisfies conditions (1) and (2) of the theorem for some permutation α and set of eigenvalues $\{\lambda_i : i = 1,\ldots,s\}$. Let B be a nonnormal, diagonalizable matrix with the same spectrum as A. Then P satisfies conditions (1) and (2) of the theorem for $\sigma(B)$ and α. Therefore, there exists a tower for each solution of $P(B)X - XB = 0$.

<u>Corollary 5.3</u>: Let A and P(X) be as in the theorem, except that A is not assumed to be diagonalizable. If there exists a tower for each solution of (1.2), then conditions (1) and (2) of the theorem hold.

<u>Proof</u>: In the proof of the theorem, to show the necessity of conditions (1) and (2), we needed that there exists a nonsingular solution, X reduces to a block monomial form, and there exists a tower for each solution X. By Theorem 4.3, X also reduces to a block monomial form in the case that A is not assumed to be diagonalizable. Therefore, the proof of the theorem holds in this case.

Corollary 5.2 shows that there are no polynomials which satisfy Problem 1.3. However, Theorem 5.1 does suggest another problem.

<u>Problem 1.4</u>: For what polynomials P does the fact that there exists a tower for each solution of (1.2) imply that A is diagonalizable?

To attack this problem, we need results like Theorem 5.1 for nondiagonalizable matrices.

<u>Theorem 5.4</u>: Let P(X) and A be as in Theorem 5.1 except that A is nonderogatory, but not assumed to be diagonalizable. If there exists

a tower for every solution of (1.2), then in addition to conditions (1) and (2) of Theorem 5.1, one of the following two conditions holds.

1) $P(x) = ax$.

2) For each λ_i, one of the following holds.

a) $P^{(k)}(\lambda_i) = 0$ for $k=1,\ldots,m_i-1$.

b) $P'(\lambda_i) = 1$ and $P^{(k)}(\lambda_i) = 0$ for $k=2,\ldots,m_i-1$ and λ_i fixed by α.

c) $\prod_{j=1}^{|\alpha'|} P'(\lambda_j)^{\frac{|\alpha|}{|\alpha'|}} = 1$.

where m_i is the multiplicity of λ_i in A and $\{ \lambda_j \}$ is the set of all eigenvalues in the same cycle α' of α as λ_i.

Furthermore, there exists a tower for each solution of (1.2) if, in addition to the conditions of Theorem 5.1, one of the following two conditions holds.

1) $P(x) = ax$.

2) For each λ_i, one of the following conditions holds.

a) $P^{(k)}(\lambda_i) = 0$ for $k=1,\ldots,m_i-1$.

b) $P'(\lambda_i) = 1$ and $P^{(k)}(\lambda_i) = 0$ for $k=2,\ldots,m_i-1$ and λ_i fixed by α.

c) $\prod_{j=1}^{|\alpha'|} P'(\lambda_j)^{\frac{|\alpha|}{|\alpha'|}} = 1$ and $P^{(k)}(\lambda_i) = 0$ for $k = 2,\ldots,m(\lambda_i)$.

Proof By Corollary 5.3, conditions (1) and (2) of Theorem 5.1 are necessary. Let $X = \bigoplus_{i=1}^{s} X_i$ be in Jordan normal form and look at each X_i.

Let $\lambda = \lambda_i$, $Y = (y_{jk}) = X_i$ where $j,k = 1,\ldots,\ell$, and $\ell = m_i$. Then,

1') Y is upper triangular

2') $y_{jj} = P'(\lambda)y_{j+1,j+1}$ for $j = 1,\ldots,\ell-1$

3') $$y_{j,j+k} = P'(\lambda)y_{j+1,j+k+1} + \sum_{m=2}^{k+1} \frac{P^{(m)}(\lambda)}{m!} y_{j+m,j+k+1}$$

for $j = 1,\ldots,\ell-1$ and $k = 1,\ldots,\ell-j-1$

by Theorem 4.3 and lemma 4.5, where X_i is the Y_i of the theorem. In particular, the diagonal of Y equals

$$y_{\ell\ell}P'(\lambda)^{\ell-1},\ y_{\ell\ell}P'(\lambda)^{\ell-2},\ldots,y_{\ell\ell}.$$

We begin by considering the case of λ fixed by α. If λ is fixed by α, then Y will lie on the diagonal of X. Therefore, the corresponding block of $SP(X)S^{-1}$ is $P(Y)$. Thus, the diagonal of $P(Y)$ is

$$(P(y_{\ell\ell}P'(\lambda)^{\ell-1}),\ P(y_{\ell\ell}P'(\lambda)^{\ell-2}),\ldots,P(y_{\ell\ell}).$$

The matrix $P(Y)$ must also satisfy (2'). In particular

$$P(y_{\ell\ell})P'(\lambda)^{\ell-1} = P(y_{\ell\ell}P'(\lambda)^{\ell-1}).$$

Thus, since by condition (2) of Theorem 5.1 P has no constant term, $P'(\lambda) =$, 0 or $y_{\ell\ell} = 0$ or $\ell = 1$ or $P(X) = aX$ for some a such that a is an ℓth root of 1. However, $y_{\ell\ell}$ is arbitrary. Therefore, $P'(\lambda) = 1,0$ or λ is a simple eigenvalue or $P(x) = ax$.

Consider the off diagonal elements of Y. Let $Y = D + C$ where $D = (d_i)$ is the diagonal of Y. There are two cases depending on $P'(\lambda)$.

Case 1: Let $P'(\lambda) = 1$. Thus, (2') implies $D = d_\ell I$. Then, (3') applied to Y implies

4) $$c_{i,i+1} = c_{i+1,i+2} + \frac{P^{(2)}(\lambda)}{2} d_\ell$$

Therefore,

$$(D + C)^k = \bigoplus_{i=0}^{k} \binom{k}{i} d^i C^{k-i}$$

and the first upper diagonal of $\binom{k}{k-1} d_\ell^{k-1} C$ is the first upper diagonal of $(D + C)^k$ since C is upper triangular and has zero diagonal. Therefore,

$$\left((D + C)^k\right)_{i,i+1} = \binom{k}{k-1} d_\ell^{k-1} c_{i,i+1} .$$

Thus, by (3') it is necessary that

$$\binom{k}{k-1} d_\ell^{k-1} c_{i,i+1} = \binom{k}{k-1} d_\ell^{k-1} c_{i+1,i+2} + \frac{P^{(2)}(\lambda)}{2} d_\ell'$$

where $d_\ell' = (D + C)^k_{\ell\ell}$.

Combining with (4) above,

$$\binom{k}{k-1} d_\ell^{k-1} c_{i+1,i+2} + \frac{P^{(2)}(\lambda)}{2} d_\ell = \binom{k}{k-1} d_\ell^{k-1} c_{i+1,i+2} + \frac{P^{(2)}(\lambda)}{2} d_\ell'$$

or

$$\binom{k}{k-1} d_\ell^{k-1} \frac{P^{(2)}(\lambda)}{2} d_\ell = \frac{P^{(2)}(\lambda)}{2} d_\ell' .$$

Thus, $P^{(2)}(\lambda) = 0$ or $d_\ell = \binom{k}{k-1} d_\ell^k$. However, from consideration of the diagonals, $d_\ell' = d_\ell^k$. Therefore, since d_ℓ is arbitrary, $P^{(2)}(\lambda) = 0$.

Proceed to the next upper diagonal where (3') now implies

$$c_{i,i+2} = c_{i+1,i+3} + \frac{P^{(3)}(\lambda)}{3!} d_\ell$$

and a similar argument gives $P^{(3)}(\lambda) = 0$. Likewise for each upper diagonal, $P^{(i)}(\lambda) = 0$ for $i = 2,\ldots,\ell-1$.

<u>Case 2</u>: Let $P'(\lambda) = 0$, then (2') implies the diagonal of Y is $(0,\ldots,0,d_\ell)$. Thus, $(Y^k)_{\ell-2,\ell-1} = 0$ for $k>1$. From (3')

$$(Y^k)_{\ell-2,\ell-1} = \frac{P^{(2)}(\lambda)}{2} d_\ell^k .$$

Therefore, $P^{(2)}(\lambda) = 0$ for $k>1$. Substituting this back into Y and checking the next upper diagonal gives $P^{(3)}(\lambda) = 0$. Likewise, $P^{(i)}(\lambda) = 0$ for $i = 1,\ldots,\ell-1$. This concludes the case of λ fixed by α.

Now consider the case of λ not fixed by α. Then, from condition (2) of Theorem 5.1, $P(X) = g(X^{|\alpha|})X$ where g is a polynomial in $X^{|\alpha|}$. Thus, using the fact that Y is a block monomial matrix,

$$P(Y_1) = g\left(\left(\prod_{i=1}^{|\alpha'|} Y_i\right)^{\frac{|\alpha|}{|\alpha'|}}\right) Y_1$$

where Y_i for $i = 1,\ldots,|\alpha'|$ are the blocks of SXS^{-1} corresponding to all the eigenvalues λ_i in the same cycle α' of α as $Y = Y_1$ and $\lambda = \lambda_1$.

Let $H = (h_{jk}) = \left(\prod_{i=1}^{|\alpha'|} Y_i\right)^{\frac{|\alpha|}{|\alpha'|}}$ and $G = (g_{jk}) = g(H)$ for $j,k = 1,\ldots,\ell$. Since X is a solution of (1.2), each Y_i is upper triangular by (1').

$$P(Y_1)_{ii} = y_{ii}g_{ii}$$

for $i = 1,\ldots,\ell$. From (2') it is necessary that

$$y_{ii}g_{ii} = P'(\lambda_1)^{\ell-i}y_{\ell\ell}g_{ii}.$$

Thus, again using (2'),

$$P'(\lambda_1)^{\ell-i}y_{\ell\ell}g_{ii} = P'(\lambda_1)^{\ell-i}y_{\ell\ell}g_{\ell\ell}.$$

Since $y_{\ell\ell}$ and $g_{\ell\ell}$ are arbitrary, either

5) $P'(\lambda_1) = 0$

or

6) $g_{ii} = g_{\ell\ell}$

for $i = 1,\ldots,\ell$.

Since each Y_i is upper triangular,

$$g_{ii} = g\Big(\big(\prod_{j=1}^{|\alpha'|} P'(\lambda_j) \prod_{j=1}^{|\alpha'|} (Y_j)_{\ell\ell}\big)^{\frac{|\alpha|}{|\alpha'|}}\Big)$$

for $i=1,\ldots,$. Therefore, (6) implies

$$7) \qquad \Big[\prod_{j=1}^{|\alpha'|} P'(\lambda_j)\Big]^{\frac{|\alpha|}{|\alpha'|}} = 1.$$

Therefore, the first set of conditions of the theorem are necessary for the existence of a tower for each solution of (1.2)

If P satisfies conditions (1) and (2) of Theorem 5.1, P(X) will reduce to the correct block monomial form. Therefore, let P(X) satisfy the second set of conditions of the theorem and consider P(X) one block at a time to show that conditions (1'), (2'), and (3') hold for P(X).

If $P(x) = ax$, then P(X) is clearly a solution for all solutions X.

If $P^{(k)}(\lambda_i) = 0$ for $k=1,\ldots,m_i-1$, then by (1'), (2'), and (3') Y_i has the form

$$1) \qquad \begin{bmatrix} & & y_1 \\ & 0 & \vdots \\ & & y_\ell \end{bmatrix}$$

where y_i are arbitrary for $i=1,\ldots\ell$. Since Y is a block monomial matrix, $P(Y)_i$ will consist of sums and products of blocks of this type. Thus, $P(Y)_i$ is again of this type.

If $P'(\lambda_i) = 1$ and $P^{(k)}(\lambda_i) = 0$ for $k=2,\ldots,m_i-1$ and λ_i is fixed by α, then it is clear from the proof of the necessity of these conditions that (1'), (2'), and (3') are satisfied. Thus $P(Y)_i$ is of the correct form.

If $\left[\prod_{j=1}^{|\alpha'|} P'(\lambda_j) \right]^{\frac{|\alpha|}{|\alpha'|}} = 1$ and $P^{(k)}(\lambda_i) = 0$ for $k=2,\ldots,m_i-1$ then

$P(Y)_i$ will consist of sums and products of the form

$$
2) \quad \begin{bmatrix}
P'^{\ell-1}(\lambda)y_{\ell\ell} & P'^{\ell-2}(\lambda)y_{\ell-1,\ell-1} & \cdots & & y_{1\ell} \\
 & \ddots & \ddots & & \vdots \\
 & & & P'(\lambda)y_{\ell-1,\ell-1} & \vdots \\
 & & & P'(\lambda)y_{\ell\ell} & y_{\ell-1,\ell} \\
 & & & & y_{\ell\ell}
\end{bmatrix}
$$

By condition (2) of Theorem 5.1, $P(X) = Xg(X^{|\alpha|})$ and the products in g will consist of powers of the product of the blocks of Y in the same cycle of α as λ_i. Thus, it can be shown that if

$$\left[\prod_{j=1}^{|\alpha'|} P'(\lambda_j) \right]^{\frac{|\alpha|}{|\alpha'|}} = 1$$

$P(Y)_i$ will again be of type 2.

Therefore, the second set of conditions of the theorem are sufficient to insure the existence of a tower for each solution of equation 1.2.

<u>Theorem 5.5</u>: Let P(X) and A be as in Theorem 5.1, except that A is derogatory, not assumed to be diagonalizable. If there exists a tower for each solution of (1.2), then P satisfies conditions (1) and (2) of Theorem 5.1 and one of the following two conditions holds.

1) $P(x) = ax$.

2) For each eigenvalue λ_i, one of the following holds.

a) $P^{(k)}(\lambda_i) = 0$ for $k=1,\ldots,m(\lambda_i)$.

b) $P'(\lambda_i) = 1$ and $P^{(k)}(\lambda_i) = 0$ for $k=2,\ldots,m(\lambda_i)$ and λ_i is fixed by α and all the Jordan blocks of A corresponding to λ_i have the same dimension.

c) $\left[\prod_{j=1}^{|\alpha'|} P'(\lambda_j) \right]^{\frac{|\alpha|}{|\alpha'|}} = 1.$

where $m(\lambda_i) + 1$ is the maximum dimension of the Jordan blocks of A corresponding to λ_i and $\{\lambda_j\}$ is the set of all eigenvalues in the same cycle α' of α as λ_i.

Furthermore, there exists a tower for each solution of (1.2) if, in addition to conditions (1) and (2) of Theorem 5.1, one of the following two conditions holds.

1) $P(x) = ax$.

2) For each eigenvalue λ_i, one of the following conditions holds.

   a) $P^{(k)}(\lambda_i) = 0$ for $k=1,\ldots,m(\lambda_i)$.

   b) $P'(\lambda_i) = 1$ and $P^{(k)}(\lambda_i) = 0$ for $k=2,\ldots,m(\lambda_i)$ and λ_i is fixed by α and all the Jordan blocks of A corresponding to λ_i have the same dimension.

   c) $\left[\prod_{j=1}^{|\alpha'|} P'(\lambda_j) \right]^{\frac{|\alpha|}{|\alpha'|}} = 1$ and $P^{(k)}(\lambda_i) = 0$ for $k=2,\ldots,m(\lambda_i)$ and all the Jordan blocks of A corresponding to λ_i have the same dimension.

Proof: By Corollary 5.3, conditions (1) and (2) of Theorem 5.1 are necessary. By Theorem 4.3, it is necessary that $P(A_i)Y_i - Y_iA_j = 0$ for $\alpha(i) = j$, where A_i is the submatrix of the Jordan normal form of A consisting of the direct sum of the Jordan blocks corresponding to λ_i. Let Y be any of the Y_i, and let the corresponding A_i and A_j be

$\bigoplus_{k=1}^{s'} A'_k$ and $\bigoplus_{\ell=1}^{s''} A''_\ell$ respectively, where each A'_k and A''_ℓ is a Jordan block.

Then by Lemma 4.1, the resulting partitions of A_i and A_j above are the same. Let $Y = (Y_{k\ell})$ be the corresponding partition of Y. Then

$$P(A'_k)Y_{k\ell} - Y_{k\ell}A''_\ell = 0$$

for $k = 1,\ldots,s'$ and $\ell = 1,\ldots,s'$. Applying Lemma 4.5 to this equation implies that for each $Z = Y_{k\ell} = (z_{ij})$ for $i = 1,\ldots,m$ and $j = 1,\ldots,n$

1") Z is upper triangular

2") $z_{m-i,n-i} = P'(\lambda)z_{m-i+1,n-i+1}$ for $i = 1,\ldots,\min(m,n)-1$

3") $z_{m-i,n-i+t} = P'(\lambda)z_{m-i+1,n-i+t+1} + \sum_{j=2}^{t+1} \frac{P^{(j)}(\lambda)}{j!} z_{m-1+j,n-i+t+1}$

for $i = 1,\ldots,\min(m,n)-1$ and $t = 1,\ldots,\min(m,n)-i-1$. Thus, the final column of each Z is arbitrary.

If there exists a tower for all solutions of (1.2), then it is necessary that P(X) is a solution for all choices of the final columns of the $Y_{k\ell}$. In particular, assume that A'_1 and A''_1 are the largest of the blocks A'_k and A''_ℓ respectively, and choose the final columns of the

$Y_{k\ell}$ to be zero for all $(k,\ell) \neq (1,1)$. Because of the relations (2") and (3"), this choice results in $Y_{k\ell} = 0$ for all $(k,\ell) \neq (1,1)$. Thus

$$Y = \left[\begin{array}{c|c} Y_{11} & 0 \\ \hline 0 & 0 \end{array}\right] \quad \text{and } P(Y) = \left[\begin{array}{c|c} P(Y_{11}) & 0 \\ \hline 0 & 0 \end{array}\right]$$

Therefore, the problem of finding solutions $P(Y)$ is contained in the problem of finding solutions $P(Y_{11})$. However, this is the case of Theorem 5.4, since A_1' and A_1'' are nonderogatory. Therefore, it is necessary that $P^{(k)}(\lambda) = 0$ for $k = 1,\ldots,m(\lambda)$

or $\left[\prod_{j=1}^{|\alpha'|} P'(\lambda_j)\right]^{\frac{|\alpha|}{|\alpha'|}} = 1$ or λ is fixed by α and $P'(\lambda) = 1$ and

$P^{(k)}(\lambda) = 0$ for $k = 2,\ldots,m(\lambda)$ or $P(x) = ax$ where $m(\lambda) + 1$ is the maximum dimension of the Jordan blocks of A corresponding to λ and $\{\lambda_j\}$ is the set of all eigenvalues in the same cycle α' as $\lambda = \lambda_1$.

Consider the case of $P'(\lambda) = 1$ and λ fixed by α and not all of the Jordan blocks of A corresponding to λ having the same dimension. Again, we may choose the final columns of $Y_{k\ell}$ to be $(0,\ldots,0)^T$ for all $(k,\ell) \neq (1,1)$, $(2,2)$, $(1,2)$ or $(2,1)$ where the dimension of Y_{11} is assumed to be greater than the dimension of Y_{22}. Then

$$
\left[\begin{array}{c|c|c|c}
\begin{matrix} y_{11} & & * \\ & \ddots & \\ & & y_{11} \end{matrix} & * & * & 0 \\
\hline
0 & \begin{matrix} y_{11} & & * \\ & \ddots & \\ 0 & & y_{11} \end{matrix} & \begin{matrix} y_{12} & & * \\ & \ddots & \\ 0 & & y_{12} \end{matrix} & 0 \\
\hline
0 & \begin{matrix} y_{21} & & * \\ & \ddots & \\ 0 & & y_{21} \end{matrix} & \begin{matrix} y_{22} & & * \\ & \ddots & \\ 0 & & y_{22} \end{matrix} & 0 \\
\hline
0 & 0 & 0 & 0
\end{array}\right.
$$

where this is a new partition of Y obtained by partitioning Y_{11} into four blocks with the second diagonal block having the same dimension as Y_{22}. Let the dimension of Y_{11} and Y_{22} be n_1' and n_2' for this new partition. Then

$$(Y^k)_{ii} = y_{11}^k$$

for $i = 1,\ldots,n_1'$ and

$$(Y^k)_{ii} = y_{11}^k + f(y_{12},y_{21})$$

for $i = n_1'+1,\ldots,n_1'+n_2'$, where f is a polynomial in y_{12} and y_{21} with positive coefficients.

Since it is necessary that $(Y^k)_{ii} = (Y^k)_{i+1,i+1}$ for $i = 1,\ldots,n_1'+n_2'$, $f(y_{12},y_{21}) = 0$ for all y_{12},y_{21}. However,

$f(y_{12},y_{21}) \neq 0$ for all y_{12},y_{21} if $k > 1$. Therefore, only $P(x) = ax$ preserves the diagonal. Therefore, conditions(1) and (2) are necessary.

Conversely, as in Theorem 5.4, conditions (1) and (2) of Theorem 5.1 are necessary and it is only necessary to check that the transformation $X \to P(X)$ preserves the internal structure of each Y_k. If $P(x) = ax$ then clearly towers exist for all solutions X. If $P^{(k)}(\lambda) = 0$ for $k = 1,\ldots,m(\lambda)$, then all blocks Y_k are of the form

$$\begin{bmatrix} & y_1 \\ 0 & \vdots \\ & y_{m'} \end{bmatrix}$$

where $m' = m(\lambda)+1$. Sums and products of blocks of this type are again of this type. Therefore towers exist for all solutions X. If $P'(\lambda) = 1$ and $P^{(k)}(\lambda) = 0$ for $k=2,\ldots,m(\lambda)$ and λ is fixed by α, then all blocks Y_k are of the form

$$\begin{bmatrix} y_{m'} & \cdots & y_2 & y_1 \\ & & & y_2 \\ & \ddots & & \vdots \\ 0 & & & y_{m'} \end{bmatrix}$$

where $m' = m(\lambda)+1$. Sums and products of blocks of this type are again of this type. Therefore towers exist for all solutions X. If

$[\prod_{j=1}^{|\alpha'|} P'(\lambda)]^{\frac{|\alpha|}{|\alpha'|}} = 1$ and $P^{(k)}(\lambda) = 0$ for $k=2,\ldots,m(\lambda)$ then all blocks

Y_k are of the form

$$\begin{bmatrix} P'(\lambda)^{m'-1}y_{m'} & P'(\lambda)^{m'-2}y_{m'-1} & \cdots & & y_1 \\ & & \ddots & & \vdots \\ & & & P'(\lambda)y_{m'-1} & \\ & & & P'(\lambda)y_{m'} & y_{m'-1} \\ & & & & y_{m'} \end{bmatrix}$$

Since condition (2) of Theorem 5.1 is satisfied, the sums and products of this type occuring in P(X) will be of this type. Therefore towers exist for all solutions X.

For a polynomial P and a matrix A to satisfy Problem 1.4, it is necessary that P and A satisfy the conditions of Theorem 5.1 but fail to satisfy the necessary conditions of Theorem 5.4 and Theorem 5.5.

Examples: Consider the following polynomials

$$P_1(x) = -x^3$$

$$P_2(x) = \frac{1}{2}(x^3 - 3x)$$

$$P_3(x) = \frac{1}{8}(-3x^5 + 10x^3 - 15x)$$

Each of these polynomials acts as a permutation of 1 and -1. Let A be any matrix with 1 and -1 as its only eigenvalues where the Jordan structures for 1 and -1 are the same. Then each of the polynomials above satisfies the conditions of Theorem 5.1.

$$P_1'(1) = -3$$

$$P_1'(-1) = -3$$

Therefore, P_1 does not satisfy the necessary conditions of Theorem 5.4 and Theorem 5.5.

$$P_2'(1) = 0$$

$$P_2'(-1) = 0$$

$$P_2''(1) = 3$$

$$P_2''(-1) = -3$$

Therefore P_2 does not satisfy the necessary conditions of Theorem 5.4 and Theorem 5.5 if there is a Jordan block in A with dimension greater than 2. In particular, P_2 does not satisfy Problem 1.4 if the dimension of A is less than or equal to 4.

$P_3'(1) = 0$

$P_3'(-1) = 0$

$P_3''(1) = 0$

$P_3''(-1) = 0$

$P_3^{(3)}(1) = -15$

$P_3^{(3)}(-1) = -15$

Therefore, P_3 satisfies Problem 1.4 if there exists a Jordan block in A of dimension greater than 3.

In general, if P is a polynomial of degree n then $P^{(n)}(\lambda) \neq 0$. Thus, if $P'(\lambda) = 0$ or 1 for some eigenvalue λ of A and the dimension of A is less than or equal to 2n, then P will not satisfy Problem 1.4 for A.

<u>Theorem 5.6</u>: Let $\Omega_t = \{\sum_{i=0}^{\ell} a_i x^{it+1}$ with $t, \ell \in I^+$ and a_i complex$\}$. Let $\theta_{\alpha,\lambda_1,\ldots,\lambda_s} = \{P \in \Omega_{|\alpha|} : P(\lambda_i) = \lambda_{\alpha(i)}$ for $i = 1,\ldots,s\}$. Let $\Psi_{\alpha,\lambda_1,\ldots,\lambda_s} = \{P : xP \in \Omega_{|\alpha|}$ and $P(\lambda_i) = 0$ for $i = 1,\ldots,s\}$. Let $\Phi_{\alpha,\lambda_1,\ldots,\lambda_s} = \{P \in \Psi_{\alpha,\lambda_1,\ldots,\lambda_s}$: P has minimal degree, $q_0|\alpha|\}$. Then $\theta_{\alpha,\lambda_1,\ldots,\lambda_s} = \{P_0 + xP_1Q_0 + xQ_1\}$ where $P_0 \in \theta_{\alpha,\lambda_1,\ldots,\lambda_s}$ and $Q_0, Q_1 \in \Phi_{\alpha,\lambda_1,\ldots,\lambda_s}$ and $xP_1 \in \Omega_{|\alpha|}$.

<u>Proof</u>: We drop the subscripts $\alpha,\lambda_1,\ldots,\lambda_s$ for clarity. First, it is clear that $\theta = \{P_0 + xQ : P_0 \in \theta$ and $Q \in \Psi\}$. Therefore, it is sufficient to show that $\Psi = \{P_1Q_0 + Q_1 : xP_1 \in \Omega$ and $Q_0, Q_1 \in \Phi\}$.

Let $Q \in \Psi$ with degree $q|\alpha|$ and $Q_0 \in \Phi$. Let $R_1 = Q - c_1 x^{(q-q_0)|\alpha|} Q_0$ where c_1 is the leading coefficient of Q. Then, R_1 is in Ψ and the degree of R_1 is less than the degree of Q. Let $R_2 = R_1 - c_2 x^{(q-q_0-1)|\alpha|} Q_0$ where c_2 is the leading coefficient of R_1, then $R_\ell = R_{\ell-1} \sum_{i=1}^{\ell} a_i x^{b_i|\alpha|} Q_0$ is of degree $q_0|\alpha|$, the minimal degree. Let

$Q_1 = R_\ell$. Then $Q_1 \in \Phi$ and $Q = \sum_{i=1}^{\ell} a_i x^{b_i|\alpha|} Q_0 + Q_1 = P_1 Q_0 + Q_1$ with

$xP_1 \in \Omega$. Thus $\Psi = \{P_1 Q_0 + Q_1 : xP_1 \in \Omega \text{ and } Q_0, Q_1 \in \Phi\}$.

Conversely, if $xP_1 \in \Omega$ and $Q_0 \in \Phi$ then $P_1 Q_0 \in \Psi$. Furthermore, if $Q_1 \in \Phi$ and $Q_2 \in \Psi$ then $Q_1 + Q_2 \in \Psi$. Therefore, $P_1 Q_0 + Q_1 \in \Psi$ for all $xP_1 \in \Omega$ and $Q_0 Q_1 \in \Phi$. Thus $\Psi = \{P_1 Q_0 + Q_1 : xP_1 \in \Omega \text{ and } Q_0 Q_1 \in \Phi\}$.

REFERENCES

1. Choi, Man-Duen. "Adjunction and Inversion of Invertible Hilbert-Space Operators." Indiana Univ. Math. Jour. 23 (1973) : 413-419.

2. Davis, C.W. "Elementary Divisors of the Stein Transformation $X \to X - CXC^*$." SIAM J. Appl. Math. 27 (1974) : 492-494.

3. DePrima, C.R., and Johnson, C.R. "The Range of $A^{-1}A^*$ in GL(n,C)." Linear Algebra and Appl. 9 (1974) : 209-222.

4. Djokovic, D.V. "On Some Representations of Matrices." To appear in Linear and Multilinear Algebra.

5. Drazin, M.P. "A Reduction of the Matrix Equation $AB = \epsilon BA$." Proc. Cambridge Philos. Soc. 47 (1951) : 7-10.

6. Fan, Ky. "Generalized Cayley Transforms and Strictly Dissipative Matrices." Linear Algebra and Appl. 5 (1972) : 155-172.

7. Fan, Ky. "On Real Matrices with Positive Definite Symmetric Component." Linear and Multilinear Algebra. 1 (1973) 1-4.

8. Fan, Ky. "On Strictly Dissipative Matrices." Linear Algebra and Appl. 9 (1974) : 223-241.

9. Fuglede, B. "A Commutativity Theorem for Normal Operators." Proc. Nat. Acad. Sci. 36 (1950) : 35-40.

10. Givens, W. Elementary Divisors and Some Properties of the Lyapunov Mapping $X \to AX + XA^*$. Argonne National Laboratory Report 6456 (1961).

11. Roth, W.E. "On Direct Product Matrices." Bull. Amer. Math. Soc. 40 (1934) : 461-468.

12. Singh, U.N., and Mangla, K. "Operators with Inverses Similar to their Adjoints." PAMS 38 (1973) : 258-260.

13. Taussky, O. "A Remark Concerning the Similarity of a Finite Matrix A and A^*." Math. Z. 117 (1970) : 189-190.

14. Taussky, O. "Hilbert's Theorem 90 in Matrix Rings." Linear and Multilinear Algebra. 1 (1973) : 5-8.

15. Taussky, O. "A Matrix Generalization of Galois Theory." In preparation.

16. Thompson, R.C. "Some Matrix Factorization Theorems." Pacific Journal of Math. 33 (1970) : 763-810.

17. Thompson, R.C. "Dissipative Matrices and Related Results." Linear Algebra and Appl. 11 (1975) : 155-169.