

# **PRODUCTS OF DIAGONALIZABLE MATRICES**

by

**MAROUN CLIVE KHOURY**

*submitted in part fulfilment of the requirements for*

*the degree of*

**MASTER OF SCIENCE**

*in the subject*

**MATHEMATICS**

*at the*

**UNIVERSITY OF SOUTH AFRICA**

*Supervisor:      Professor J. D. Botha*

*Date:              SEPTEMBER 2002*

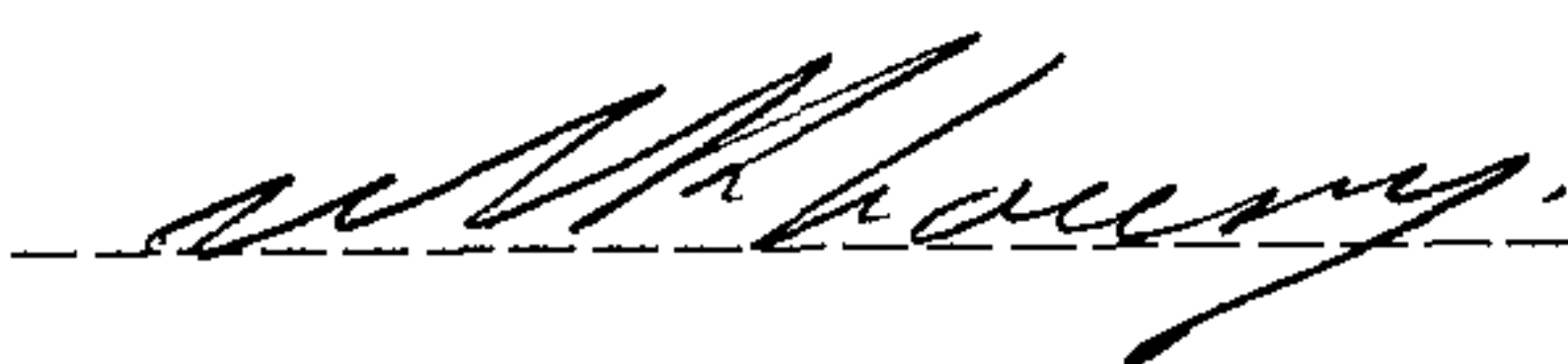
# INTRODUCTION

Matrix factorization is a very basic and indispensable tool in matrix theory, as is clear from well-known factorizations such as the LU, Polar and Elementary Matrix Factorizations. In the literature, much research has been conducted into the factorization of matrices, where the factors are restricted to certain classes which turned out to be diagonalizable; hence this dissertation which seeks:

1. to investigate the necessary and sufficient conditions for the factorization of a matrix into diagonalizable matrices, and
2. to consider the factorization of a matrix into certain subclasses of diagonalizable matrices,

paying special attention to the least number of factors required in all cases.

"I declare that: **Products of Diagonalizable Matrices** is my own work and that all sources that I have used or quoted have been indicated and acknowledged by means of complete references."

Signature 

# PRODUCTS OF DIAGONALIZABLE MATRICES

By M. C. KHOURY.

MASTER OF SCIENCE

MATHEMATICS

Supervisor: Professor J. D. Botha

## SUMMARY

Chapter 1 reviews better-known factorization theorems of a square matrix. For example, a square matrix over a field can be expressed as a product of two symmetric matrices; thus square matrices over real numbers can be factorized into two diagonalizable matrices. Factorizing matrices over complex numbers into Hermitian matrices is discussed. The chapter concludes with theorems that enable one to prescribe the eigenvalues of the factors of a square matrix, with some degree of freedom. Chapter 2 proves that a square matrix over arbitrary fields (with one exception) can be expressed as a product of two diagonalizable matrices. The next two chapters consider decomposition of singular matrices into Idempotent matrices, and of nonsingular matrices into Involutions. Chapter 5 studies factorization of a complex matrix into Positive-(semi)definite matrices, emphasizing the least number of such factors required.

### Key terms:

*Diagonalizable factorization of matrices; Hermitian factorization of matrices; Prescribing the eigenvalues of the factors of a square matrix; Idempotent factorization of matrices; Factorization of matrices into Involutions; Positive-definite and Positive-semidefinite factors.*

# ACKNOWLEDGEMENTS

I would like to thank my supervisor, Professor J. D. Botha of the University of South Africa, without whose help, continual encouragement, professional approach and knowledge of the subject this dissertation would not have been possible.

I would like to thank the National Research Foundation for the grant I received during the first three years of my studies towards this degree.

I would also like to thank my colleagues at St. Dunstan's College for their encouragement, support and understanding during this period. I would especially like to mention my headmaster Fr. Paterson for the above, as well as for the financial support he arranged for me, through the school, in my final year, and Mrs Langley for the many times she proof read my work for me.

Special thanks to Helen Khoury and my friends Thato Serumula and Jane O'Connor for their endless hours of proof reading.

Finally I would like to thank my family and friends for their encouragement and support.

# CONTENTS

<b>NOTATION</b>	<b>1</b>
<b>CHAPTER 1: SOME GENERAL MATRIX FACTORIZATION RESULTS.</b>	<b>3</b>
1.1 LU-FACTORIZATION.	3
1.2 THE POLAR DECOMPOSITION.	7
1.3 SINGULAR-VALUE DECOMPOSITION.	7
1.4 SYMMETRIC FACTORIZATION.	8
1.5 HERMITIAN FACTORIZATION.	12
1.6 PRESCRIBING THE EIGENVALUES.	23
1.7 COMMUTATORS.	39
<b>CHAPTER 2: DIAGONALIZABLE FACTORIZATION OVER AN ARBITRARY FIELD.</b>	<b>42</b>
2.1 DIAGONALIZABLE FACTORIZATION OVER A FIELD $F$ WHICH IS NOT $GF(3)$ AND $CHAR(F) \neq 2$ .	42
2.2 DIAGONALIZABLE FACTORIZATION OVER THE FIELD $GF(3)$ .	53
2.3 DIAGONALIZABLE FACTORIZATION OVER A FIELD OF CHARACTERISTIC 2 WITH AT LEAST FOUR ELEMENTS.	57
2.4 DIAGONALIZABLE FACTORIZATION OVER THE FIELD $GF(2)$ .	67
<b>CHAPTER 3: PRODUCTS OF IDEMPOTENT MATRICES.</b>	<b>69</b>
<b>CHAPTER 4: PRODUCTS OF INVOLUTIONS.</b>	<b>82</b>
4.1 PRODUCTS OF TWO INVOLUTIONS.	83
4.2 PRODUCTS OF FOUR INVOLUTIONS.	90
<b>CHAPTER 5: PRODUCTS OF POSITIVE-DEFINITE AND POSITIVE-SEMIDEFINITE MATRICES.</b>	<b>103</b>
5.1 PRODUCTS OF POSITIVE-DEFINITE MATRICES.	104
5.2 PRODUCTS OF POSITIVE-SEMIDEFINITE MATRICES.	114
5.3 CONCLUDING REMARKS.	119
<b>APPENDIX</b>	<b>122</b>
<b>Index of Theorems, Lemmas etc.</b>	<b>126</b>
<b>REFERENCES</b>	<b>127</b>



# NOTATION

$F^n$	The $n$ dimensional vector spaces over a field $F$ .
$\text{char}(F)$	The characteristic of a ring or field $F$ : the least positive integer $n$ such that $n \cdot a = 0$ for all $a \in F$ .
$M_{m \times n}(F)$	Vector space of all matrices of size $m \times n$ over a field $F$ . When referring to a specific field, for example the real numbers, $F$ will be replaced with the appropriate symbol, in this case $M_{m \times n}(R)$ .
$M_n(F)$	Vector space of all square matrices of size $n$ over a field $F$ .
$GL(n, F)$	Group of invertible matrices in $M_n(F)$ .
$SL(n, F)$	The subgroup of $GL(n, F)$ consisting of all matrices with determinant one.
$A^T$	The transpose of matrix $A$ . Vectors in $F^n$ will be considered as column vectors. A row vector in $F^n$ will be denoted by $x^T$ .
$A^*$	The conjugate transpose (adjoint) of matrix $A$ .
$\det(A)$	The determinant of the matrix $A$ .
$\text{nullity}(A)$	The dimension of the null space of $A$ .

$\sigma(A)$  The eigenvalues of  $A$  (spectrum). Always repeated according to algebraic multiplicity.

$A \sim B$  Matrix  $A$  is similar to matrix  $B$ .

$E_\lambda$  The eigenspace associated with the eigenvalue  $\lambda$  of a square matrix or linear operator.

$rsp[A]$  Row space of matrix  $A$ .

The  $n \times n$  Companion Matrix  $C_{a(x)}$  associated with the monic polynomial  $a(x) = x^n + a_{n-1}x^{n-1} \dots + a_2x^2 + a_1x + a_0$  will be represented as follows:

$$C_{a(x)} = \begin{bmatrix} 0 & 0 & 0 & \dots & \dots & \dots & 0 & -a_0 \\ 1 & 0 & 0 & \dots & \dots & \dots & 0 & -a_1 \\ 0 & 1 & 0 & \dots & \dots & \dots & 0 & -a_2 \\ \vdots & \vdots & \vdots & & & & \vdots & \vdots \\ \vdots & \vdots & \vdots & & & & \vdots & \vdots \\ \vdots & \vdots & \vdots & & & & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \dots & 1 & 0 & -a_{n-2} \\ 0 & 0 & 0 & \dots & \dots & 0 & 1 & -a_{n-1} \end{bmatrix}.$$

■ End of a proof.

□ End of example.



# CHAPTER 1

## SOME GENERAL MATRIX FACTORIZATION RESULTS.

The factorization of a matrix into factors having specific properties is a topic in which research is still in progress. It is applied in many fields such as the solving of systems of linear and differential equations, statistics and computer programming.

### 1.1 LU-FACTORIZATION.

The LU-factorization is a well-known and much used factorization of a square matrix  $A \in M_n(C)$ . The theorem states that if all the leading principal minors  $A \begin{pmatrix} 1 \dots i \\ 1 \dots i \end{pmatrix}$   $i = 1, \dots, n-1$  of a matrix  $A \in M_n(C)$  are nonzero, then there exist unique matrices  $L, U \in M_n(C)$  such that  $A = LU$ , where  $L$  is lower triangular with diagonal elements equal to one, and  $U$  is upper triangular. See [13], Theorem 1, page 62 for a proof.

When such a factorization exists, solving the linear system  $Ax = b$ , where  $x \in M_{n \times 1}(C)$  (if a solution exists), is relatively quick and simple. The method is suitable for computer programs since it saves storage space in memory. See [8], page 129.

The LU-Factorization also computes the determinant of a matrix efficiently, since, if  $A = LU$  then  $\det(A) = \det(L) \cdot \det(U)$ ,  $\det(L) = 1$  and  $\det(U)$  is equal to the product of the entries in the main diagonal.

This factorization can be extended to give the following factorization of a square matrix in the complex numbers.

### **Corollary 1.1**

(a) Under the hypotheses of the LU-Factorization, that is, all the leading principal minors  $A \begin{pmatrix} 1 \dots i \\ 1 \dots i \end{pmatrix}$   $i = 1, \dots, n-1$  of a matrix  $A \in M_n(C)$  are nonzero, there exist unique lower and upper triangular matrices  $L$  and  $U'$  respectively, with ones in the main diagonal of both matrices, and a unique diagonal matrix  $D$ , such that  $A = LDU'$ .

(b) If in addition to the hypotheses of the LU-factorization,  $A$  is Hermitian, then  $A = LDL^*$  where  $L$  is lower triangular with ones in the main diagonal, and  $D$  is a real diagonal matrix. Both  $L$  and  $D$  are unique. This is known as the Cholesky-factorization.

### Proof

(a) Partition  $A$  as  $\begin{bmatrix} A_{n-1} & a_1 \\ a_2^T & a_{nn} \end{bmatrix}$  where  $a_1, a_2 \in C^{n-1}$ , then by the LU-factorization,  $A = LU$ , that is

$$\begin{bmatrix} A_{n-1} & a_1 \\ a_2^T & a_{nn} \end{bmatrix} = \begin{bmatrix} L_{n-1} & 0 \\ d^T & 1 \end{bmatrix} \begin{bmatrix} U_{n-1} & c \\ 0 & u_{nn} \end{bmatrix},$$

where  $a_1, a_2, d, c \in M_{(n-1) \times 1}(C)$ .

Since  $A \begin{pmatrix} 1 \dots (n-1) \\ 1 \dots (n-1) \end{pmatrix} \neq 0$ ,  $A_{n-1} = L_{n-1} U_{n-1}$ , where both  $L_{n-1}$  and  $U_{n-1}$  are nonsingular and unique. So the main diagonal of  $L_{n-1}$  contains ones, and the main diagonal of  $U_{n-1}$  contains nonzero entries. Let  $U = [u_{ij}]_{i,j=1}^n$  ( $u_{ij} = 0$ , for  $i > j$ ). Then

$$D = \begin{bmatrix} u_{11} & & & 0 \\ & u_{22} & & \\ & & \ddots & \\ & & & \ddots & \\ 0 & & & & u_{nn} \end{bmatrix}$$

and

$$U' = \begin{bmatrix} 1 & u_{12}/u_{11} & u_{13}/u_{11} & \cdot & \cdot & \cdot & u_{1n}/u_{11} \\ 0 & 1 & u_{23}/u_{22} & \cdot & \cdot & \cdot & u_{2n}/u_{22} \\ \cdot & & & & & & \cdot \\ \cdot & & & \cdot & & & \cdot \\ \cdot & & & & \cdot & & \cdot \\ & & & & & \cdot & \\ 0 & & \cdot & \cdot & \cdot & 0 & 1 \end{bmatrix}.$$

The uniqueness follows from the uniqueness of the LU-factorization of  $A$ .

(b) By part (a)  $A = LDU$ , and since  $A$  is Hermitian,

$$A = A^* = U^* D^* L^*.$$

By the uniqueness of the factorization,  $L^* = U$  and  $D = D^*$ . Finally,  $D$  is real since it is diagonal and Hermitian. ■

The LU-factorization, and in particular the above theorem, gives the following factorization for a positive-definite matrix  $H$ :

$$H = GG^*,$$

where  $G$  is a unique lower triangular matrix with positive diagonal elements. This is known as the Cholesky factorization of positive-definite matrices.

## 1.2 THE POLAR DECOMPOSITION.

It is well known that for an arbitrary matrix,  $A \in M_{m \times n}(C)$ ,  $A^*A$  and  $AA^*$  are positive-semidefinite.

The following result is analogous to the polar form of a complex number:  $z = z_0 e^{i\theta}$  where  $z_0 \geq 0$  and  $0 \leq \theta < 2\pi$ .

Any matrix  $A \in M_n(C)$  can be expressed in the form

$$A = HU,$$

where  $H = (AA^*)^{1/2}$  and  $U$  is unitary. Thus  $H$  is unique and, if  $A$  is nonsingular,  $U$  is also unique.

The above decomposition is known as the polar decomposition. See [13], Theorem 1, page 190 for a proof.

## 1.3 SINGULAR-VALUE DECOMPOSITION.

The singular values of  $A$  are defined as the eigenvalues of  $(A^*A)^{1/2}$ . Thus the singular values of any matrix are nonnegative numbers.

The following decomposition of an arbitrary matrix in  $C$  (even rectangular) is known as the singular-value decomposition of the matrix. See [13], Theorem 2, page 192 for a proof.

Let  $A$  be an arbitrary matrix in  $M_{m \times n}(C)$  and let

$\{s_1, \dots, s_r\}$  be the nonzero singular values of  $A$ .

Then  $A$  can be represented in the form

$$A = UDV^*,$$

where  $U \in M_{m \times m}(C)$  and  $V \in M_{n \times n}(C)$  are unitary, and

$D$  has  $s_i$  in its  $(i, i)$  position and zeros elsewhere.

#### 1.4 SYMMETRIC FACTORIZATION.

One of the first matrix factorization problems to be investigated and solved, was that every square matrix over any field can be expressed as a product of two symmetric matrices. This was done by Frobenius in 1910. This result becomes significant when looking at matrices in the real numbers, since all symmetric matrices are then diagonalizable.

The proof below is a generalization of the one given by Halmos in [10]. Radjavi gave a similar proof in [14].

The following Lemma is necessary to prove the results of Frobenius.

**Lemma 1.2** (Halmos [10] and Radjavi [14])

Every companion matrix over an arbitrary field  $F$  is a product of two symmetric matrices.

**Proof**

Let  $C_{a(x)}$  be the companion matrix associated with the monic polynomial  $a(x)$ , and  $S_1$  and  $S_2^{-1}$  be the symmetric matrices

$$\begin{bmatrix} -a_0 & 0 & 0 & . & . & 0 & 0 \\ 0 & a_2 & a_3 & . & . & a_{n-1} & 1 \\ 0 & a_3 & & . & 1 & 0 & 0 \\ . & . & & . & 1 & 0 & 0 \\ . & . & . & . & & . & . \\ 0 & a_{n-1} & . & & & . & . \\ 0 & 1 & 0 & . & . & . & 0 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} a_1 & a_2 & a_3 & . & . & a_{n-1} & 1 \\ a_2 & a_3 & a_4 & . & . & a_{n-1} & 1 & 0 \\ a_3 & a_4 & a_5 & . & a_{n-1} & 1 & 0 & 0 \\ . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . \\ a_{n-1} & 1 & 0 & . & . & . & . & 0 \\ 1 & 0 & 0 & . & . & . & . & 0 \end{bmatrix}$$

respectively.



Then,

$$C_{a(x)} S_2^{-1} = \left[ \begin{array}{cccccc|c} 0 & 0 & 0 & . & . & . & 0 & -a_o \\ \hline & & & & & & & -a_1 \\ & & & & & & & -a_2 \\ & & & & & & & . \\ & & & & & & & . \\ & & & & & & & . \\ & & & & & & & a_{n-1} \\ & & & & & & & -a_{n-1} \end{array} \right] \left[ \begin{array}{cccccc|c|c} a_1 & a_2 & a_3 & . & . & . & a_{n-1} & 1 \\ a_2 & a_3 & a_4 & . & . & . & a_{n-1} & 1 & 0 \\ a_3 & a_4 & a_5 & . & . & . & 1 & 0 & 0 \\ . & . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . & . \\ a_{n-1} & 1 & 0 & . & . & . & . & . & 0 \\ \hline 1 & 0 & 0 & . & . & . & . & . & 0 \end{array} \right]$$

$$= \left[ \begin{array}{cccccc|c} -a_o & 0 & 0 & . & . & . & 0 & 0 \\ 0 & a_2 & a_3 & . & . & . & a_{n-1} & 1 \\ 0 & a_3 & . & . & . & . & 1 & 0 \\ . & . & . & . & . & . & 1 & 0 \\ . & . & . & . & . & . & . & . \\ 0 & a_{n-1} & . & . & . & . & . & . \\ 0 & 1 & 0 & . & . & . & . & 0 \end{array} \right] = S_1$$

$$\therefore C_{a(x)} = S_1 S_2 .$$

It remains to observe that the inverse of a symmetric matrix is also symmetric. ■

### **Theorem 1.3 (Frobenius)**

Every square matrix over an arbitrary field  $F$  is the product of two symmetric matrices over  $F$ .

### **Proof**

Let  $A \in M_n(F)$ . By the Rational Canonical Form Theorem,  $A$  is similar to the direct sum of companion matrices. Thus by Lemma A.2 it suffices to prove that a matrix  $A$  in the

form  $A = C_1 \oplus \dots \oplus C_k$ , where  $C_i$  is a companion matrix ( $1 \leq i \leq k$ ), can be expressed as a product of symmetric matrices. By Lemma 1.2,  $C_i = S_1^{(i)} S_2^{(i)}$ , where  $S_1^{(i)}$  and  $S_2^{(i)}$  are symmetric matrices ( $1 \leq i \leq k$ ).

$$\therefore A = S_1^{(1)} S_2^{(1)} \oplus \dots \oplus S_1^{(k)} S_2^{(k)} = (S_1^{(1)} \oplus \dots \oplus S_1^{(k)}) (S_2^{(1)} \oplus \dots \oplus S_2^{(k)})$$

Both factors on the right are symmetric. ■

The factorization of  $A \in M_n(R)$  into symmetric matrices, and hence into diagonalizable matrices, is not unique. This can be seen very easily from the fact that  $A = S_1 S_2 = (-S_1)(-S_2)$  for symmetric matrices  $S_1$  and  $S_2$ .

The example below is a direct application of the proof of Lemma 1.2 and Theorem 1.3.

#### **Example 1.4**

$$B = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 3 & 0 \\ 4 & 5 & 6 \end{bmatrix}$$

The characteristic polynomial of  $B$  is  $c_B(\lambda) = (\lambda - 1)(\lambda - 3)(\lambda - 6)$ .

$$E_1 = sp\left\{\begin{pmatrix} 5 \\ -5 \\ 1 \end{pmatrix}\right\}, \quad E_3 = sp\left\{\begin{pmatrix} 0 \\ -3 \\ 5 \end{pmatrix}\right\} \quad \text{and} \quad E_6 = sp\left\{\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}\right\}.$$

$$B = P^{-1} \begin{bmatrix} -1 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & -6 \end{bmatrix} (P^T)^{-1} P^T \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix} P \quad \text{where } P^{-1} = \begin{bmatrix} 5 & 0 & 0 \\ -5 & -3 & 0 \\ 1 & 5 & 1 \end{bmatrix}.$$

Multiplying together the first three factors, and the last three gives:

$$\begin{bmatrix} 1 & 0 & 0 \\ 2 & 3 & 0 \\ 4 & 5 & 6 \end{bmatrix} = \begin{bmatrix} -25 & 25 & -5 \\ 25 & -52 & 50 \\ -5 & 50 & -82 \end{bmatrix} \begin{bmatrix} -\frac{518}{225} & -\frac{23}{9} & -\frac{22}{15} \\ -\frac{23}{9} & -\frac{26}{9} & -\frac{5}{3} \\ -\frac{22}{15} & -\frac{5}{3} & -1 \end{bmatrix}.$$

See [10], page 2, for other symmetric factorizations of the matrix  $B$ . □

### 1.5 HERMITIAN FACTORIZATION.

Although the result of Theorem 1.3 is true over any field  $F$ , it is not the most useful factorization over the field of complex numbers. In this section we investigate the necessary and sufficient conditions under which a complex matrix can be expressed as a product of Hermitian matrices.

The first result, which is based on the work done by Halmos in [10], gives the necessary and sufficient condition for a matrix to be factorized into two Hermitian matrices.

The following well-known lemma is needed for the proof.

**Lemma 1.5**

Every nilpotent matrix is similar to its adjoint.

**Proof**

Let  $A$  be nilpotent; then zero is the only eigenvalue of  $A$ .

$\therefore A \sim J = \text{diag}[J_1, \dots, J_k]$  where each  $J_i$  is a Jordan block of the form

$$\begin{bmatrix} 0 & 1 & 0 & \cdot & \cdot & 0 \\ 0 & 0 & 1 & & & \cdot \\ \cdot & \cdot & \cdot & \cdot & & \cdot \\ \cdot & & \cdot & \cdot & \cdot & 0 \\ \cdot & & & \cdot & \cdot & 1 \\ 0 & \cdot & \cdot & \cdot & 0 & 0 \end{bmatrix}.$$

Therefore  $A = SJS^{-1}$  for some invertible matrix  $S$ , which implies that  $A^* = S^{*-1}J^T S^*$ .

It suffices to show that  $J_i \sim J_i^T$  since then,  $A \sim J \sim J^T \sim A^*$ .

Let  $P_n$  be the  $n \times n$  involution

$$\begin{bmatrix} 0 & & & & 1 \\ & \cdot & & & \\ & & \cdot & & \\ & & & \cdot & \\ 1 & & & & 0 \end{bmatrix};$$

then,

$$P_n^{-1} = P_n \text{ and } P_n J P_n = J^T.$$

This completes the proof. ■

**Theorem 1.6 (Halmos [10])**

Let  $A \in M_n(C)$ , then  $A$  is the product of two Hermitian matrices if, and only if,  $A$  is similar to its adjoint.

**Proof**

Suppose  $A = EF$ , where  $E$  and  $F$  are Hermitian. If  $E$  is nonsingular, then  $A = E(FE)E^{-1} = EA^*E^{-1}$  and the result follows. If  $F$  is nonsingular, then  $A = F^{-1}(FE)F = F^{-1}A^*F$  and again the result follows.

Suppose both  $F$  and  $G$  are singular, then by Lemmas A.3 and A.6 it suffices to prove the result for  $A$  in the form

$$\begin{bmatrix} N & 0 \\ 0 & K \end{bmatrix},$$

where  $N$  is nilpotent and  $K$  is nonsingular.

Since  $A^* = FE$ ,

$$A^m E = E A^{*m} \quad \dots \quad (1.7).$$

Let  $E = \begin{bmatrix} L & M \\ M^* & S \end{bmatrix}$  and  $F = \begin{bmatrix} P & Q \\ Q^* & R \end{bmatrix}$ , where all the blocks are

conformable with  $N \oplus K$  and  $L, S, P$  and  $R$  are Hermitian.

From equation 1.7:

$$\begin{bmatrix} N^m & 0 \\ 0 & K^m \end{bmatrix} \begin{bmatrix} L & M \\ M^* & S \end{bmatrix} = \begin{bmatrix} L & M \\ M^* & S \end{bmatrix} \begin{bmatrix} N^{*m} & 0 \\ 0 & K^{*m} \end{bmatrix},$$

which implies that  $K^m M^* = M^* N^{*m} = M^* N^{m*}$ . This holds for all  $m \in \mathbb{Z}^+$ , therefore  $K^m M^* = 0$  if  $m$  is sufficiently large.

So  $M^* = 0$  since  $K$  is nonsingular, and therefore  $E = \begin{bmatrix} L & 0 \\ 0 & S \end{bmatrix}$ .

Now

$$\begin{bmatrix} N & 0 \\ 0 & K \end{bmatrix} = \begin{bmatrix} L & 0 \\ 0 & S \end{bmatrix} \begin{bmatrix} P & Q \\ Q^* & R \end{bmatrix},$$

which implies that  $K = SR$  and so both  $S$  and  $R$  are nonsingular. From equation 1.7 with  $m = 1$  we get  $KS = SK^*$ . This implies that  $K^* = S^{-1}KS$ . By Lemma 1.5  $N^* = H^{-1}NH$  for some nonsingular matrix  $H$ .

Let  $T = \begin{bmatrix} H & 0 \\ 0 & S \end{bmatrix}$ , then

$$A^* = \begin{bmatrix} N^* & 0 \\ 0 & K^* \end{bmatrix} = \begin{bmatrix} H^{-1}NH & 0 \\ 0 & S^{-1}KS \end{bmatrix} = T^{-1}AT.$$

Conversely suppose that  $A \sim A^*$ , then there exists a nonsingular matrix  $T$  such that

$$AT = TA^* \quad \dots \quad (1.8)$$

and, by taking the adjoint of both sides,

$$AT^* = T^*A^* \quad \dots \quad (1.9).$$

Let  $u \in C$ , then

$$A(uT + \bar{u}T^*) = (uT + \bar{u}T^*)A^* \quad \dots \quad (1.8) \times u + (1.9) \times \bar{u}.$$

Therefore  $AE_u = E_u A^*$ , where  $E_u = (uT + \bar{u}T^*)$  and  $E_u$  is Hermitian, hence  $AE_u$  is also Hermitian. If  $u \neq 0$ , then

$$E_u = \left( TT^{*-1} + \frac{\bar{u}}{u} I \right) uT^*$$

and  $u$  can be chosen so that  $TT^{*-1} + \frac{\bar{u}}{u} I$  is nonsingular (this is always assured since the spectrum of  $TT^{*-1}$  is finite). Therefore  $E_u$  is nonsingular and so  $A = (AE_u)E_u^{-1}$ . This completes the proof. ■

If  $A$  is a real matrix, then in terms of Theorem 1.6,  $A = EF$ , where  $E$  and  $F$  are symmetric matrices if, and only if,  $A \sim A^T$ . There is nothing gained by this statement, since a matrix is always similar to its transpose and, as was shown in section 1.4, all matrices are products of two symmetric matrices. Therefore this theorem and most of the following results are of interest mainly for matrices over the complex numbers.

In the following theorem, originally proved by Radjavi [14], a square matrix can be expressed as the product of



four Hermitian matrices if, and only if, the determinant of the matrix is real. This result is a more general result than that of Theorem 1.6, since if  $A \sim A^*$ , then

$$\det A = \det A^* = \overline{\det A},$$

therefore the determinant of  $A$  is real.

**Theorem 1.10 (Radjavi [14])**

Let  $A$  be an  $n \times n$  matrix over any conjugate-closed subfield  $F$  of the complex numbers. Then  $A$  is the product of four Hermitian matrices over  $F$  if, and only if,  $\det(A) \in R$ .

**Proof**

By Lemma A.3 we can assume  $A$  is in rational canonical form; that is,  $A = \text{diag}[A_1, \dots, A_m]$  where  $A_j$  has the form

$$A_j = \begin{bmatrix} 0 & & & & & t_{j1} \\ 1 & 0 & & & & t_{j2} \\ & 1 & 0 & & & t_{j3} \\ & & \ddots & \ddots & & \vdots \\ & & & \ddots & \ddots & \vdots \\ & & & & 1 & 0 & t_{jn(j)-1} \\ & & & & & 1 & t_{jn(j)} \end{bmatrix}, \quad j = 1, \dots, m.$$

Then  $n = n(1) + \dots + n(m)$  and  $\det(A) = \pm t_{11} t_{21} \dots t_{m1}$ .

If  $\det(A) = 0$  then assume that  $t_{m1} = 0$ .

Let  $A = BK$ , where  $K$  is the permutation matrix

$$\begin{bmatrix} 0 & 0 & 0 & . & . & . & 0 & 1 \\ 1 & 0 & 0 & . & . & . & 0 & 0 \\ 0 & 1 & 0 & . & . & . & 0 & 0 \\ . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . \\ 0 & 0 & 0 & . & . & . & 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ I_{n-1} & 0 \end{bmatrix}.$$

Now by Theorem 1.3,  $K$  is the product of two real symmetric matrices, which are Hermitian. It remains to show that  $B$  is the product of two Hermitian matrices over  $F$ .

$$B = AK^{-1} = A \begin{bmatrix} 0 & I_{n-1} \\ 1 & 0 \end{bmatrix}.$$

So  $B$  is the permutation of the columns of matrix  $A$  with the last column first and the others shifted one to the right. That is

$$B = \begin{bmatrix} B_1 & C_1 & 0 & . & . & . & 0 \\ 0 & B_2 & C_2 & . & . & . & 0 \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ 0 & . & . & . & B_{m-1} & C_{m-1} \\ C_m & . & . & . & 0 & B_m \end{bmatrix},$$

where  $B_j = \text{diag}[0, 1, 1, \dots, 1] \in M_{n(j)}(F)$  and

$$C_j = \begin{bmatrix} t_{j1} & 0 & . & . & . & 0 \\ t_{j2} & 0 & . & . & . & 0 \\ . & & & & & \\ . & & & & & \\ . & & & & & \\ t_{jn(j)} & 0 & . & . & . & 0 \end{bmatrix} \in M_{n(j) \times n(j+1)}(F) ,$$

$j = 1, \dots, m$ . In the last formula, and in what follows, if  $j=m$  then  $j+1=1$ .

A diagonal matrix  $D$  is constructed as follows:

For each pair of integers  $(j, k)$ ,  $1 \leq j \leq m$  and  $1 \leq k \leq n(j)$  except  $(m, 1)$ , set

$$r_{jk} = \begin{cases} t_{jk} & \text{if } t_{jk} \neq 0 \\ 1 & \text{if } t_{jk} = 0 \end{cases}$$

and set

$$r_{m1} = \frac{1}{r_{11} r_{21} \dots r_{m-1, 1}} .$$

Let  $D = \text{diag}[D_1, \dots, D_m]$ , where

$$D_j = \text{diag} \left[ \frac{r_{j1}}{\prod_{i=1}^j r_{i1}}, \dots, \frac{r_{jn(j)}}{\prod_{i=1}^j r_{i1}} \right] .$$

Then

$$D^{-1}BD = \begin{bmatrix} D_1^{-1}B_1D_1 & D_1^{-1}C_1D_2 & 0 & \cdot & \cdot & \cdot & 0 \\ 0 & D_2^{-1}B_2D_2 & D_2^{-1}C_2D_3 & \cdot & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \cdot & \cdot & \cdot & D_{m-1}^{-1}B_{m-1}D_{m-1} & D_{m-1}^{-1}C_{m-1}D_m \\ D_m^{-1}C_mD_1 & \cdot & \cdot & \cdot & 0 & D_m^{-1}B_mD_m \end{bmatrix}.$$

So a typical block matrix of  $D^{-1}BD$  is of the form  $D_j^{-1}B_jD_j$  or  $D_j^{-1}C_jD_{j+1}$ .

Now  $D_j^{-1}B_jD_j = B_j$ , since diagonal matrices commute.

The only entries of  $D_i^{-1}C_iD_{i+1}$  that may possibly be nonzero are in the first column. The entry in the  $k^{\text{th}}$  row of the first column of  $D_j^{-1}C_jD_{j+1}$  is

$$t_{jk} \times \frac{\prod_{i=1}^j r_{i1}}{r_{jk}} \times \frac{r_{j+1,1}}{\prod_{i=1}^{j+1} r_{i1}}, \quad k = 1, \dots, n(j).$$

This expression is equal to 1 or 0 for all possible values of  $j$  and  $k$ , except when  $j = m$  and  $k = 1$ . In this case

$$t_{m1} \times \frac{\prod_{i=1}^m r_{i1}}{r_{m1}} \times \frac{r_{11}}{\prod_{i=1}^1 r_{i1}} = t_{m1} \prod_{i=1}^{m-1} r_{i1}.$$

If  $\det(A) = 0$ , then  $t_{m1} = 0$ , and so  $t_{m1} \prod_{i=1}^{m-1} r_{i1} = 0$ . If

$\det(A) \neq 0$ , then  $t_{m1} \prod_{i=1}^{m-1} r_{i1} = t_{11} t_{21} \dots t_{m1} = \pm \det(A) \in R$ .

Thus  $D^{-1}BD = S$ , where  $S \in M_n(R)$ , and so by Theorem 1.3,  $S = S_1 S_2$ , where  $S_1$  and  $S_2$  are real symmetric matrices. Therefore  $B = (DS_1 D^*)(D^{*-1} S_2 D^{-1})$  and both brackets are Hermitian.

The converse follows from the fact that the determinant of a Hermitian matrix is always real. ■

The example below is a direct application of the proof of Theorem 1.10 for a matrix in rational canonical form consisting of one companion matrix.

### **Example 1.11**

Let

$$A = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & i \\ 0 & 1 & -i \end{bmatrix},$$

then

$$A = \begin{bmatrix} 1 & 0 & 0 \\ i & 1 & 0 \\ -i & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} = BK.$$

$$D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & i & 0 \\ 0 & 0 & -i \end{bmatrix} \quad \text{and} \quad D^{-1}BD = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}.$$

The Rational Canonical Form of  $D^{-1} B D$  is  $D^{-1} B D = Q C Q^{-1}$  where

$$Q = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \quad \text{and} \quad C = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Applying Theorem 1.3 to the matrix  $C$

$$C = S_1 S_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & -1 & 0 \\ -1 & -2 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Thus  $B = [(DQ) S_1 (DQ)^*] \cdot [(DQ)^{-1} S_2 (DQ)^{-1}]$ .

$K = P E P^{-1}$ , where

$$P = \begin{bmatrix} 1 & -2 & 1 \\ 1 & 1 & 1 \\ -2 & 1 & 1 \end{bmatrix} \quad \text{and} \quad E = \begin{bmatrix} 0 & -1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Applying Theorem 1.3 to the matrix  $E$

$$E = S_3 S_4 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & -1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Thus  $K = [P S_3 P^T] \cdot [(P^T)^{-1} S_4 P^{-1}]$ .

Finally

$$A = BK$$

$$= [(DQ) S_1 (DQ)^*] \cdot [(DQ)^{-1} S_2 (DQ)^{-1}] \cdot [P S_3 P^T] \cdot [P^{-1} S_4 P^{-1}]$$

$$= \begin{bmatrix} 0 & i & -i \\ -i & -1 & 1 \\ i & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} 0 & i & 0 \\ -i & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} -2 & 4 & 1 \\ 4 & 1 & -2 \\ 1 & -2 & 4 \end{bmatrix} \cdot \begin{bmatrix} \frac{2}{9} & \frac{1}{9} & 0 \\ \frac{1}{9} & 0 & \frac{2}{9} \\ 0 & \frac{2}{9} & \frac{1}{9} \end{bmatrix}. \quad \square$$

The corollary below gives an almost Hermitian factorization for a complex matrix with non-real determinant.

**Corollary 1.12**

If  $A \in M_n(F)$ , where  $F$  is a conjugate-closed subfield of  $C$  and  $\det(A) = \alpha \notin R$ , then  $A = \frac{1}{z_0} H_1 H_2 H_3 H_4$ , where  $H_i$  are Hermitian,  $i = 1, 2, 3, 4$  and  $z_0$  is a root of the equation  $z^n - \bar{\alpha} = 0$ .

**Proof**

$\det(z_0 A) = z_0^n \alpha = \bar{\alpha} \alpha \in R$ ; therefore by Theorem 1.10,

$z_0 A = H_1 H_2 H_3 H_4$ , where  $H_i$  is Hermitian,  $i = 1, 2, 3, 4$ .

The result follows. ■

## 1.6 PRESCRIBING THE EIGENVALUES.

In [15], Sourour proved that a nonscalar, nonsingular matrix can be expressed as a product of two matrices with prescribed eigenvalues.

The following lemma by Fillmore [6] is required to prove Sourour's theorem.



**Lemma 1.13** (Fillmore [6])

Let  $A \in M_n(F)$  be a nonscalar matrix, then  $A$  is similar to a matrix whose  $(1,1)$  position is  $\alpha$ , for arbitrary fixed  $\alpha \in F$ .

**Proof**

Firstly note that every vector  $x \in F^n$  is an eigenvector of  $A$  if, and only if,  $A$  is a scalar matrix. Let  $e_1, \dots, e_n$  be any basis of  $F^n$  and suppose  $Ae_i = \lambda_i e_i$  for  $(1 \leq i \leq n)$ . Let  $x = \sum e_i$  and  $Ax = \lambda x$ . Then  $A \sum e_i = \lambda \sum e_i$ , which implies that  $\sum \lambda_i e_i = \sum \lambda e_i$ ; therefore  $\sum (\lambda_i - \lambda) e_i = 0$  and so  $\lambda = \lambda_i$ . The matrix is scalar since the basis was arbitrary. The converse is obvious.

Thus, since  $A$  is nonscalar, there exists a vector  $e_1$  which is not an eigenvector of  $A$  and therefore not of  $A - \alpha I$  for arbitrary  $\alpha \in F$ . Letting  $e_2 = (A - \alpha I)e_1$ , it is easily seen that  $e_1$  and  $e_2$  are linearly independent. Now choose an ordered basis of  $F^n$  whose first two elements are  $e_1$  and  $e_2$ . Then the matrix  $A_1$ , with respect to this basis, has a first column  $(\alpha, 1, 0, \dots, 0)$ . We conclude that  $A$  is similar to

$$A_1 = \begin{bmatrix} \alpha & y^T \\ x & R \end{bmatrix},$$

where  $x$  is the column vector with one in its first position and zeros elsewhere,  $y^T$  is a row vector, and  $R \in M_{n-1}(F)$ . ■

**Theorem 1.14 (Sourour [15])**

Let  $A \in M_n(F)$  be nonscalar, invertible and let  $\beta_j$  and  $\gamma_j$  ( $1 \leq j \leq n$ ) be elements of  $F$ , such that  $\prod_{j=1}^n \beta_j \gamma_j = \det A$ . Then there exist  $B, C \in M_n(F)$  with eigenvalues  $\beta_1, \dots, \beta_n$  and  $\gamma_1, \dots, \gamma_n$  respectively, such that  $A = BC$ . Furthermore,  $B$  and  $C$  can be chosen so that  $B$  is lower triangularizable and  $C$  is simultaneously upper triangularizable.

**Proof**

We use induction on  $n$ , the size of  $A$ .

For  $n=1$ , the result is vacuously true since there are no nonscalar matrices of size  $1 \times 1$ .

For  $n=2$ , let  $\beta_1 \beta_2 \gamma_1 \gamma_2 = \det A$ , where  $\beta_1, \beta_2, \gamma_1, \gamma_2 \in F$ . By

Lemma 1.13,  $A$  is similar to  $A_1 = \begin{bmatrix} \beta_1 \gamma_1 & y \\ x & r \end{bmatrix}$ , where  $x, y, r \in F$ .

Now  $\det A_1 = \det A$ , and so  $r = xy\beta_1^{-1}\gamma_1^{-1} + \beta_2\gamma_2$ .

$$\therefore A_1 = \begin{bmatrix} \beta_1 & 0 \\ \gamma_1^{-1}x & \beta_2 \end{bmatrix} \begin{bmatrix} \gamma_1 & \beta_1^{-1}y \\ 0 & \gamma_2 \end{bmatrix} = B_1 C_1,$$

and  $A = P^{-1} B_1 C_1 P = BC$ , where  $B = P^{-1} B_1 P$  and  $C = P^{-1} C_1 P$ .

This completes the proof of the theorem for  $n = 2$ .

Now let  $n \geq 3$  and assume the conclusion for all square matrices of size less than  $n$ . By Lemma 1.13,  $A$  is similar to

$$A_1 = \begin{bmatrix} \beta_1 \gamma_1 & y^T \\ x & R \end{bmatrix} \dots \quad (1.15),$$

where  $x, y \in F^{n-1}$  and  $R \in M_{n-1}(F)$ .

It can be shown that  $A \sim A_1 \sim A_2$ , where  $A_2 = \begin{bmatrix} \beta_1 \gamma_1 & z^T \\ x & S \end{bmatrix}$  and

$S - \beta_1^{-1} \gamma_1^{-1} x z^T$  is not a scalar matrix. (This is done at the end of the proof.)

Since  $A_2 = \begin{bmatrix} \beta_1 \gamma_1 & 0 \\ x & S - \beta_1^{-1} \gamma_1^{-1} x z^T \end{bmatrix} \begin{bmatrix} 1 & \beta_1^{-1} \gamma_1^{-1} z^T \\ 0 & I_{n-1} \end{bmatrix}$  and

$$\det(A_2) = \det(A),$$

$$\det(A_2) = \beta_1 \gamma_1 \det(S - \beta_1^{-1} \gamma_1^{-1} x z^T) = \det(A) = \prod_{j=1}^n \beta_j \gamma_j.$$

$$\therefore \det(S - \beta_1^{-1} \gamma_1^{-1} x z^T) = \beta_1^{-1} \gamma_1^{-1} \prod_{j=1}^n \beta_j \gamma_j = \prod_{j=2}^n \beta_j \gamma_j.$$

By the induction hypothesis there exist  $B_o, C_o \in M_{n-1}(F)$ ,

such that  $\sigma(B_o) = \{\beta_2, \dots, \beta_n\}$ ,  $\sigma(C_o) = \{\gamma_2, \dots, \gamma_n\}$  and

$S - \beta_1^{-1} \gamma_1^{-1} x z^T = B_o C_o$ . It follows that

$$A_2 = \begin{bmatrix} \beta_1 & 0 \\ \gamma_1^{-1} x & B_o \end{bmatrix} \begin{bmatrix} \gamma_1 & \beta_1^{-1} z^T \\ 0 & C_o \end{bmatrix}.$$

Again by the induction hypothesis there exists a nonsingular matrix  $Q_o \in M_{n-1}(F)$ , such that  $Q_o^{-1} B_o Q_o$  and  $Q_o^{-1} C_o Q_o$  are respectively lower triangular and upper

triangular. Let  $Q = \begin{bmatrix} 1 & 0 \\ 0 & Q_o \end{bmatrix}$ ; then

$$Q^{-1} A_2 Q = \begin{bmatrix} \beta_1 & 0 \\ \xi & Q_o^{-1} B_o Q_o \end{bmatrix} \cdot \begin{bmatrix} \gamma_1 & \eta^T \\ 0 & Q_o^{-1} C_o Q_o \end{bmatrix},$$

where  $\xi, \eta \in F^{n-1}$ . Since  $Q^{-1} A_2 Q \sim A_2 \sim A$ , this completes the induction.

To complete the proof of this theorem we need to show the existence of the matrix  $A_2$ .

Suppose that for  $A_1$  in equation 1.15,  $R - \beta_1^{-1} \gamma_1^{-1} x y^T = \alpha I_{n-1}$  for some  $\alpha \in F$ . Since  $A$  is of full rank,  $[x \ R]$  must be of full row rank while the columns of  $[x \ R]$  are linearly dependent. There exists  $w \in F^{n-1}$ , such that  $w^T x = 0$  and  $w^T R \neq 0$ , since according to the proof of Lemma 1.13 it may be assumed that  $x \neq 0$ . Define linear transformations  $T_x: F^{n-1} \rightarrow F$  and  $T_R: F^{n-1} \rightarrow F^n$  as follows:  $T_x(w) = w^T x$  and  $T_R(w) = w^T [x \ R]$ ; then  $nullity(T_x) = n-2$  and  $nullity(T_R) = 0$ . Choose any nonzero vector  $w \in \ker(T_x)$ .

Let  $P = \begin{bmatrix} 1 & w^T \\ 0 & I_{n-1} \end{bmatrix}$ ; then

$$A_2 = P^{-1} A_1 P = \begin{bmatrix} \beta_1 \gamma_1 & z^T \\ x & S \end{bmatrix},$$

where

$$z^T = y^T + \beta_1 \gamma_1 w^T - w^T R$$

and

$$S = R + x w^T.$$

And so

$$\begin{aligned} S - \beta_1^{-1} \gamma_1^{-1} x z^T &= R - \beta_1^{-1} \gamma_1^{-1} x y^T + \beta_1^{-1} \gamma_1^{-1} x w^T R \\ &= \alpha I_{n-1} + \beta_1^{-1} \gamma_1^{-1} x w^T R. \end{aligned}$$

Since  $x \neq 0$ , and  $w^T R \neq 0$ , the matrix  $\beta_1^{-1} \gamma_1^{-1} x w^T R$  is of rank one, therefore  $S - \beta_1^{-1} \gamma_1^{-1} x z^T$  is not scalar. ■

### **Corollary 1.16**

Let  $A$  be a nonsingular matrix of size  $n$  over a field  $F$ . Then  $A$  can always be expressed as a product of two diagonalizable matrices, provided that the field has at least  $n + 2$  elements.

### **Proof**

Let  $\beta_1, \dots, \beta_{n-1}$  be any distinct elements of  $F$  not equal to

zero, one or  $\det(A)$ , and let  $\beta_n = \det(A)$ . Finally, let

$$\gamma_1 = \frac{1}{\beta_1}, \dots, \gamma_{n-1} = \frac{1}{\beta_{n-1}} \text{ and } \gamma_n = 1. \text{ Then } \prod_{i=1}^n \beta_i \gamma_i = \det(A)$$

and so by Theorem 1.14, there exist matrices  $B$  and  $C$  with distinct eigenvalues, such that  $A = BC$ . Since  $B$  and  $C$  have distinct eigenvalues they are diagonalizable. ■

From the corollary above it is clear that a nonsingular matrix over a field with infinitely many elements can be factorized into a product of two diagonalizable matrices.

### Example 1.17

$$A = \begin{bmatrix} 2 & 1 & 0 \\ 2 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix} \in M_3(R).$$

Let  $\beta_1 = 1$ ,  $\beta_2 = 2$ ,  $\beta_3 = 3$ ,  $\gamma_1 = 1$ ,  $\gamma_2 = \frac{1}{2}$  and  $\gamma_3 = \frac{1}{3}$ ;

then  $\prod_{i=1}^3 \beta_i \gamma_i = 1 = \det(A)$ .

$$A - \beta_1 \gamma_1 I_3 = \begin{bmatrix} 1 & 1 & 0 \\ 2 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$$

and it follows from the proof of Lemma 1.13, that  $e_1 = [1 \ 0 \ 0]^T$  is a suitable (*non-eigenvector*) choice.

$$\text{Thus } e_2 = \begin{bmatrix} 1 & 1 & 0 \\ 2 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \text{ and take } e_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

Then

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 1 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 2 & 1 & 0 \\ 2 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & \frac{3}{2} & -\frac{1}{2} \\ 1 & \frac{5}{2} & \frac{1}{2} \\ 0 & -\frac{1}{2} & \frac{1}{2} \end{bmatrix} = \begin{bmatrix} \beta_1 & \gamma_1 & y^T \\ x & R \end{bmatrix}$$

$$\text{and } R - \beta_1^{-1} \gamma_1^{-1} x y^T = \begin{bmatrix} \frac{5}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{bmatrix} - \begin{bmatrix} 1 \\ 0 \end{bmatrix} \begin{bmatrix} \frac{3}{2} & -\frac{1}{2} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ -\frac{1}{2} & \frac{1}{2} \end{bmatrix}.$$

The last matrix is nonscalar, so

$$A_2 = A_1 = \left[ \begin{array}{c|cc} 1 & 0 & 0 \\ \hline 1 & B_o & \\ \hline 0 & & \end{array} \right] \left[ \begin{array}{c|cc} 1 & \frac{3}{2} & -\frac{1}{2} \\ \hline 0 & C_o & \\ \hline 0 & & \end{array} \right]$$

where

$$B_o C_o = \begin{bmatrix} 1 & 1 \\ -\frac{1}{2} & \frac{1}{2} \end{bmatrix}.$$

At this point the process is repeated on the  $2 \times 2$  matrix  $B_o C_o$ . Note that  $\beta_2 \gamma_2$  is already in the (1,1) position, therefore

$$B_o C_o = \begin{bmatrix} 2 & 0 \\ -1 & 3 \end{bmatrix} \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ 0 & \frac{1}{3} \end{bmatrix},$$

and

$$A_1 = A_2 = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 2 & 0 \\ 0 & -1 & 3 \end{bmatrix} \begin{bmatrix} 1 & \frac{3}{2} & -\frac{1}{2} \\ 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & \frac{1}{3} \end{bmatrix} = BC.$$



Finally,  $A = P A_1 P^{-1} = P B C P^{-1} = (P B P^{-1})(P C P^{-1})$

where  $P = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 1 & 1 \end{bmatrix}$  and so

$$A = \begin{bmatrix} 2 & 0 & 0 \\ 2 & 1 & 0 \\ 1 & -\frac{3}{2} & 3 \end{bmatrix} \begin{bmatrix} 1 & \frac{1}{2} & 0 \\ 0 & 0 & 1 \\ 0 & -\frac{1}{6} & \frac{5}{6} \end{bmatrix}. \quad \square$$

Theorem 1.14 has many applications, for example.

- (i) It can be used to determine a diagonalizable factorization of a matrix over a field  $F$ , as was illustrated in the example above. The only restriction being on the number of elements in the field. This restriction was dealt with by Botha in [3] and [4] and will be discussed in the next chapter.
- (ii) It gives a simple characterization of the commutators of the group  $GL(n, F)$ , which will be discussed in the next section.
- (iii) It can be used to show that if the determinant of a square matrix of size  $n$  over a field  $F$  is  $\pm 1$ , then it can be expressed as the product of four

involutions. However, there is a restriction on the proof given by Sourour, namely the field must contain at least  $n + 2$  elements. The general case was dealt with by Gustafson, Halmos and Radjavi in [9] and will be discussed in a later chapter.

In [17], Theorem 1, Sourour and Tang proved that a singular matrix can be expressed as a product of two matrices with prescribed eigenvalues. In the case of a nonsingular matrix the only restriction is that the product of the eigenvalues of the factors must be equal to the determinant of the original matrix. In the case of a singular matrix there is an additional restriction on the choice of prescribed eigenvalues, namely the number of zeros must be equal to, or exceed, the nullity of the original matrix.

A few preliminaries need to be established before the theorem by Sourour and Tang can be proved.

**Lemma 1.18** (Sourour and Tang [17])

(a) If  $T = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \in M_n(F)$  where  $A$  is a nonsingular square matrix, then  $\text{nullity}(T) = \text{nullity}(D - CA^{-1}B)$ .

(b) Let  $D \in M_n(F)$  be nonsingular and  $Y \in M_{k \times n}(F)$  then,

$$\begin{bmatrix} 0 & 0 \\ 0 & D \end{bmatrix} \sim \begin{bmatrix} 0 & Y \\ 0 & D \end{bmatrix}.$$

**Proof**

(a)  $T = \begin{bmatrix} A & 0 \\ C & I \end{bmatrix} \begin{bmatrix} I & A^{-1}B \\ 0 & D - CA^{-1}B \end{bmatrix}$ . The first factor is nonsingular

and so

$$\text{rank}(T) = \text{rank} \begin{bmatrix} I & A^{-1}B \\ 0 & D - CA^{-1}B \end{bmatrix} = n - \text{nullity}(D - CA^{-1}B).$$

Therefore  $\text{nullity}(D - CA^{-1}B) = n - \text{rank}(T) = \text{nullity}(T)$ .

$$(b) \quad \begin{bmatrix} I & YD^{-1} \\ 0 & I \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & D \end{bmatrix} \begin{bmatrix} I & -YD^{-1} \\ 0 & I \end{bmatrix} = \begin{bmatrix} 0 & Y \\ 0 & D \end{bmatrix}. \quad \blacksquare$$

The following theorem is used in the proof of the theorem by Sourour and Tang. It was first proved in [20] by Wu for a matrix over the complex numbers. In [12] Laffey and in [16] Sourour proved the result independently for a matrix over an arbitrary field.

**Theorem 1.19** (Wu [20], Sourour [16], Laffey [12])

Let  $A$  be a singular matrix over an arbitrary field. Then  $A$  is a product of two nilpotent matrices if, and only if,  $A$  is not a nonzero  $2 \times 2$  nilpotent matrix.

**Theorem 1.20** (Sourour and Tang [17])

Let  $A \in M_n(F)$  be a singular matrix over a field  $F$ , and  $\beta_j$  and  $\gamma_j$  ( $1 \leq j \leq n$ ) be elements of  $F$ . If  $A$  is not a nonzero  $2 \times 2$  nilpotent matrix, then  $A$  can be expressed as a product of two matrices  $B$  and  $C$  with  $\sigma(B) = \{\beta_1, \dots, \beta_n\}$  and  $\sigma(C) = \{\gamma_1, \dots, \gamma_n\}$  if, and only if, the number of zeros  $m$  in the set  $\{\beta_1, \gamma_1, \dots, \beta_n, \gamma_n\}$  is not less than the nullity of  $A$ . If  $A$  is a nonzero  $2 \times 2$  nilpotent matrix, then  $A$  can be factored as above if, and only if,  $1 \leq m \leq 3$ .

**Note** The following notation will be used in the proof:  $\sigma(B) = \{\beta_1, \dots, \beta_n\}$  will be denoted by  $\beta$  and  $\sigma(C) = \{\gamma_1, \dots, \gamma_n\}$  will be denoted by  $\gamma$ ;  $\beta = 0$  if  $\beta_1 = \beta_2 = \dots = \beta_n = 0$ . The reduced list  $\{\beta_2, \dots, \beta_n\}$  is denoted by  $\beta'$ .

**Proof**

Suppose  $A = BC$  with  $\sigma(B) = \{\beta_1, \dots, \beta_n\}$  and  $\sigma(C) = \{\gamma_1, \dots, \gamma_n\}$ . Let  $m_1$  and  $m_2$  be the number of zeros in  $\sigma(B)$  and  $\sigma(C)$  respectively. Then

$$\text{nullity}(A) = \text{nullity}(BC) \leq \text{nullity}(B) + \text{nullity}(C) \leq m_1 + m_2 = m.$$

If  $A$  is a  $2 \times 2$  nonzero nilpotent matrix, then from the above  $m \geq 1$  and  $m \neq 4$ . Since if  $m = 4$ , then  $B$  and  $C$  would

both be nilpotent, and by Theorem 1.19 this is not possible. So  $1 \leq m \leq 3$ .

Conversely, suppose that  $m \geq \text{nullity}(A)$ .

The proof is by induction on  $n$ . If  $n = 1$ , then  $A$  is the zero matrix and the result is trivial. Assume the result is true for all singular square matrices of size less than  $n$ . Let  $A$  be an  $n \times n$  singular matrix and  $\beta_1, \gamma_1, \dots, \beta_n, \gamma_n$  be elements of  $F$ ,  $m$  of which are zero.

Case 1  $m < n$ .

In this case,  $\beta \neq 0$  and  $\gamma \neq 0$ . Assume without loss of generality, that  $\beta_1 \neq 0$  and  $\gamma_1 \neq 0$ . By Lemma 1.13

$$A \sim A_1 = \begin{bmatrix} \beta_1 \gamma_1 & y^T \\ x & D \end{bmatrix},$$

where  $x, y \in F^{n-1}$  and  $D \in M_{n-1}(F)$ . It suffices to establish the factorization for  $A_1$ . By Lemma 1.18(a),  $\text{nullity}(D - \beta_1^{-1} \gamma_1^{-1} x y^T) = \text{nullity} A_1 \leq m$ . Since the number of zeros in the reduced sets  $\beta'$  and  $\gamma'$  is still  $m$ , the induction hypothesis may be applied. Thus there exist matrices  $B_o, C_o \in M_{n-1}(F)$ , such that  $\sigma(B_o) = \beta'$  and  $\sigma(C_o) = \gamma'$ .

Now  $D - \beta_1^{-1} \gamma_1^{-1} x y^T = B_o C_o$  and  $A_1 = \begin{bmatrix} \beta_1 & 0 \\ \gamma_1^{-1} x & B_o \end{bmatrix} \begin{bmatrix} \gamma_1 & \beta_1^{-1} y^T \\ 0 & C_o \end{bmatrix}$ , which

completes the proof for this case.

Case 2  $n \leq m \leq 2n - 1$ .

Assume  $\beta \neq 0$  and take  $\beta_1 \neq 0$ . Since  $m \geq n$  at least one  $\gamma_i = 0$ ; take  $\gamma_1 = 0$ . An application of the Rational Canonical Form Theorem shows that  $A \sim A_1$ , where

$$A_1 = \begin{bmatrix} 0 & y^r \\ 0 & D \end{bmatrix},$$

$y \in F^{n-1}$  and  $D \in M_{n-1}(F)$ . The number of zeros in  $\beta' \cup \gamma'$  is  $m - 1$  where  $m - 1 \geq n - 1 \geq \text{nullity}(D)$ .

If  $D$  is a nonsingular matrix other than a  $2 \times 2$  matrix of the form  $\begin{bmatrix} 0 & a \\ b & 0 \end{bmatrix}$ , then let  $R \in M_{n-1}(F)$ , where  $R$  agrees with

$D$  in one row and all the other rows are zero. Then  $D - R$  is singular. However if  $D = \begin{bmatrix} 0 & a \\ b & 0 \end{bmatrix}$  then take  $R = \begin{bmatrix} -b & a \\ 0 & 0 \end{bmatrix}$ .

Now  $\text{nullity}(D - R) \leq n - 1 \leq m - 1$ , so by the induction hypothesis, there exist matrices  $B_o, C_o \in M_{n-1}(F)$ , such that  $\sigma(B_o) = \beta'$ ,  $\sigma(C_o) = \gamma'$  and

$$B_o C_o = D - R \dots (1.21).$$

$R$  is an  $n - 1 \times n - 1$  matrix with only one nonzero row. Let this be the  $i^{\text{th}}$  row, and define  $z$  to be the column vector with one in the  $i^{\text{th}}$  position and zero elsewhere; then



$R = z w^T$  where  $w^T$  is the  $i^{\text{th}}$  row of  $R$ . Equation (1.21) now gives  $D = z w^T + B_o C_o$ . Let  $B$  and  $C$  be the  $n \times n$  matrices

$$\begin{bmatrix} \beta_1 & 0 \\ z & B_o \end{bmatrix} \text{ and } \begin{bmatrix} 0 & w^T \\ 0 & C_o \end{bmatrix}$$

respectively.

Then  $\sigma(B) = \beta$ ,  $\sigma(C) = \gamma$  and  $BC = \begin{bmatrix} 0 & \beta_1 w^T \\ 0 & D \end{bmatrix}$ . By Lemma

1.18(b), both  $A_1$  and  $BC$  are similar to  $\begin{bmatrix} 0 & 0 \\ 0 & D \end{bmatrix}$ .

If  $D$  is singular and  $A$  is not a  $3 \times 3$  nilpotent matrix with  $\beta_2 = \beta_3 = \gamma_2 = \gamma_3 = 0$ , then we may apply the induction hypothesis to  $D$ . That is, there exist  $B_o, C_o \in M_{n-1}(F)$ , such that  $\sigma(B_o) = \beta'$ ,  $\sigma(C_o) = \gamma'$  and  $D = B_o C_o$ . The matrix  $A_1$  can now be factored as follows

$$A_1 = \begin{bmatrix} \beta_1 & 0 \\ 0 & B_o \end{bmatrix} \begin{bmatrix} 0 & \beta_1^{-1} y^T \\ 0 & C_o \end{bmatrix}.$$

If  $A$  is a  $3 \times 3$  nilpotent matrix of rank one, then

$$A \sim \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \text{ and } \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} \beta_1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & \beta_1^{-1} & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

If  $A$  is a  $3 \times 3$  nilpotent matrix of rank 2, and

$\beta_2 = \beta_3 = \gamma_2 = \gamma_3 = 0$ , then



$$A \sim \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \\ 1 & \beta_1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}.$$

The last thing to note in this case is that if  $\beta = 0$  then  $\gamma \neq 0$  since  $m \leq 2n - 1$ . Then we may proceed, as above, to factorize  $A^T = RS$ , where  $\sigma(R) = \gamma$  and  $\sigma(R) = \beta$ . The result then follows, since  $A = S^T R^T$ .

Case 3      $m = 2n$ .

In this case  $A$  is being factorized into two nilpotent matrices. This is the result of Theorem 1.19.

This completes the proof. ■

A direct application of Theorem 1.20 does not necessarily give a diagonalizable decomposition of a singular matrix, since the eigenvalue zero may have to occur more than once in the spectrum of either  $B$  or  $C$  (or both) depending on the nullity of  $A$ . In which case the minimal polynomials of  $B$  or  $C$  may have  $x$  as a factor with multiplicity larger than one, and will therefore not be diagonalizable. If however  $\text{nullity}(A)$  is 1 or 2, then Theorem 1.20 will give a diagonalizable decomposition by choosing each factor with at most one eigenvalue equal to zero, and the other

eigenvalues distinct, provided that the field is sufficiently large.

Let  $A \in M_n(F)$ . If  $\text{nullity}(A) = 1$  or  $2$ , and  $F$  has at least  $n$  elements, then take  $\beta = \gamma$  where  $\beta_1 = 0$  and  $\beta_2, \dots, \beta_n$  are distinct nonzero elements of  $F$ . Then by Theorem 1.20  $A = BC$  where both matrices are diagonalizable.

Although Theorem 1.20 has limitations when looking for a diagonalizable factorization of a matrix, it is very useful when considering the factorization of a matrix into special types of diagonalizable matrices over the real or complex fields. An example of this is the factorization of a matrix into positive-definite, or positive-semidefinite, matrices. This will be discussed in a later chapter.

## 1.7 COMMUTATORS.

If  $G$  is any group, then the commutator of  $b, c \in G$  is defined as the element  $bcb^{-1}c^{-1}$  of  $G$ . The commutator subgroup of  $G$  is the subgroup generated by all the commutators of  $G$ .

The following are some well known properties of the commutator subgroup  $H$  of the group  $G$ .

- (i)  $H$  is a normal subgroup of  $G$ .
- (ii) The factor group  $G/H$  is abelian.
- (iii)  $G$  is abelian if, and only if,  $H = \{1\}$ .

The result of Theorem 1.14 can be used to give relatively simple proofs regarding the set of commutators of the group  $GL(n, F) = \{A \in M_n(F) : \det(A) \neq 0\}$ . The result was first proved by Shoda for an algebraically closed field  $F$ . In [19], Thompson showed the same results for all fields  $F$  with one exception, namely when  $F$  has only two elements and  $n = 2$ . Theorem 1.22 shows that these results can be obtained using the result of Sourour's main theorem in [15], provided the field contains sufficiently many elements.

**Theorem 1.22 (Shoda – Thompson)**

If  $F$  has at least  $n + 1$  elements, then the set of commutators of  $GL(n, F)$  is  $SL(n, F) = \{A \in M_n(F) : \det(A) = 1\}$ .

**Proof**

Let  $A \in SL(n, F)$ . If  $A$  is not a scalar matrix, then (since  $\det(A) = 1$ ), by Theorem 1.14,  $A = BE$  where  $B$  has distinct

eigenvalues  $\beta_1, \dots, \beta_n$ , and the eigenvalues of  $E$  are  $\beta_1^{-1}, \dots, \beta_n^{-1}$ . Now  $B = P D_1 P^{-1}$  and  $E = Q D_2 Q^{-1}$  where  $D_1 = \text{diag}[\beta_1, \dots, \beta_n]$  and  $D_2 = \text{diag}[\beta_1^{-1}, \dots, \beta_n^{-1}]$ . Therefore  $D_2 = D_1^{-1} = P^{-1} B^{-1} P$ . So  $E = C B^{-1} C^{-1}$  where  $C = Q P^{-1}$ , and the result follows. If  $A = \alpha I$  then  $\det(A) = \alpha^n = 1$ ; so let  $B = \text{diag}[\alpha, \alpha^2, \dots, \alpha^n]$  and  $E = [\alpha^n, \alpha^{n-1}, \dots, \alpha]$ , then  $A = BE$  and  $E = C B^{-1} C^{-1}$  where  $C = \begin{bmatrix} 0 & 1 \\ I_{n-1} & 0 \end{bmatrix}$ , and once again the result follows.

On the other hand, if  $A$  is a commutator of  $GL(n, F)$ , then there exist matrices  $B$  and  $C$  in  $M_n(F)$ , such that  $A = BCB^{-1}C^{-1}$ , from which it follows that  $\det(A) = 1$ . ■

# CHAPTER 2

## DIAGONALIZABLE FACTORIZATION OVER AN ARBITRARY FIELD.

In this chapter, the factorization of a square matrix into diagonalizable matrices over an arbitrary field is considered, paying special attention to the least number of such factors required.

In [3], Botha showed that for a field  $F$ , such that  $\text{char}(F) \neq 2$  and  $F \neq GF(3)$ , then any matrix  $A \in M_n(F)$  is the product of two diagonalizable matrices, and if  $F = GF(3)$  then every matrix over  $F$  is the product of three diagonalizable matrices. In [4], the same author proved a similar result for a matrix over a field of characteristic two, provided that the field has at least four elements.

### 2.1 DIAGONALIZABLE FACTORIZATION OVER A FIELD $F$ WHICH IS NOT $GF(3)$ AND $\text{CHAR}(F) \neq 2$ .

The following Lemma is needed to prove Botha's results.



Then  $A_n$  is diagonalizable over  $F$  if, and only if,  $\text{char}(F) \neq 2$ , and in this case  $\sigma(A_n) = \{\pm a_1, \dots, \pm a_{n/2}\}$  for  $n$  even and  $\sigma(A_n) = \{\pm a_1, \dots, \pm a_{(n-1)/2}, b\}$  for  $n$  odd.

(c) Let  $B$  be an  $n \times n$  diagonalizable matrix over a field  $F$ , and suppose  $a \in F$  is not an eigenvalue of  $B$ . Then

$$A = \begin{bmatrix} B & x \\ 0 & a \end{bmatrix} \text{ is diagonalizable, where } x \in F^n.$$

### **Proof**

(a) For  $A \in M_n(F)$ , let  $E_\lambda(A)$  denote the eigenspace of  $A$  associated with  $\lambda \in F$  [ $E_\lambda(A) = \{0\}$  if  $\lambda \notin \sigma(A)$ ], and  $m_\lambda(A)$  denote the algebraic multiplicity of  $A$  with respect to  $\lambda$ .

$A$  is diagonalizable

$$\Leftrightarrow \dim E_\lambda(A) = m_\lambda(A) \quad \forall \lambda \in F.$$

(See [7], Theorem 5.14 page 257)

$$\Leftrightarrow \sum_{i=1}^k \dim E_\lambda(A_i) = \sum_{i=1}^k m_\lambda(A_i) \quad \forall \lambda \in F.$$

$$\Leftrightarrow \dim E_\lambda(A_i) = m_\lambda(A_i) \quad \forall \lambda \in F, 1 \leq i \leq k.$$

(Since  $\dim E_\lambda(A_i) \leq m_\lambda(A_i) \quad \forall \lambda \in F$  [13], Theorem 1 page 160)

$$\Leftrightarrow A_i \text{ is diagonalizable, } 1 \leq i \leq k.$$



(b) First prove by induction that if  $n \geq 1$  then:

$$A_n \sim \bigoplus_{i=1}^{n/2} \begin{bmatrix} 0 & a_i^2 \\ 1 & 0 \end{bmatrix} \text{ if } n \text{ is even and}$$

$$A_n \sim [b] \oplus \bigoplus_{i=1}^{(n-1)/2} \begin{bmatrix} 0 & a_i^2 \\ 1 & 0 \end{bmatrix} \text{ if } n \text{ is odd.}$$

For  $n = 1$  or  $n = 2$  the result is trivial.

Suppose that  $n > 2$ , and the result is true for all square matrices of size  $n - 2$  over  $F$ .

$$\text{Let } Y = \begin{bmatrix} 0 & I_{n-2} & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}; \text{ then } Y^{-1} = \begin{bmatrix} 0 & 1 & 0 \\ I_{n-2} & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \text{ and}$$

$$Y A_n Y^{-1} = Y \begin{bmatrix} 0 & 0 & a_1^2 \\ 0 & A_{n-2} & 0 \\ 1 & 0 & 0 \end{bmatrix} Y^{-1} = \begin{bmatrix} A_{n-2} & 0 & 0 \\ 0 & 0 & a_1^2 \\ 0 & 1 & 0 \end{bmatrix},$$

and the result follows.

Hence by part (a) of this lemma it suffices to show

that  $A_i = \begin{bmatrix} 0 & a_i^2 \\ 1 & 0 \end{bmatrix}$  is diagonalizable.

The characteristic polynomial of  $A_i$  is

$$x^2 - a_i^2 = (x - a_i)(x + a_i),$$

and so  $A_i$  has distinct eigenvalues  $\pm a_i$ , provided that

$a_i \neq -a_i$ . This will be the case if  $\text{char}(F) \neq 2$ .

This completes the proof that  $A_n$  is diagonalizable over any field other than one of characteristic two, with the stated eigenvalues.

Now suppose that  $\text{char}(F) = 2$  and  $A_i$  is diagonalizable with eigenvalues  $a_i = -a_i$ . Then  $A_i \sim a_i I_2$  and so there exists a nonsingular matrix  $P_i$ , such that  $A_i = P_i(a_i I_2)P_i^{-1} = a_i I_2$ . This is a contradiction since  $A_i$  is not scalar.

(c) Since  $a$  is not an eigenvalue of  $B$ ,  $B - aI$  is nonsingular. Therefore there exists  $y \in F^n$ , such that  $(B - aI)y = x$ .

Let  $X = \begin{bmatrix} I_n & -y \\ 0 & 1 \end{bmatrix}$ , then

$$X^{-1} = \begin{bmatrix} I_n & y \\ 0 & 1 \end{bmatrix}$$

and

$$X^{-1}AX = \begin{bmatrix} B & 0 \\ 0 & a \end{bmatrix}.$$

By part (a) of this lemma this matrix is diagonalizable. ■

**Theorem 2.2** (Botha [3])

Let  $F$  be any field, such that  $\text{char}(F) \neq 2$  and  $F \neq GF(3)$ .  
Then every square matrix over  $F$  is a product of two diagonalizable matrices.

**Proof**

By Lemma A.1 it suffices to prove the theorem for a matrix  $A$  in rational canonical form,  $A = C_1 \oplus \dots \oplus C_k$ , where each  $C_i$  is a companion matrix. By Lemma 2.1(a) we need only prove the result for one companion matrix  $C$ .

Let

$$C = \begin{bmatrix} 0 & . & . & . & 0 & a_0 \\ 1 & . & . & . & . & . \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ . & . & . & 0 & a_{n-2} & . \\ 0 & . & . & 1 & a_{n-1} & . \end{bmatrix},$$

define  $P_n$  to be the  $n \times n$  matrix

$$P_n = \begin{bmatrix} 0 & & & 1 \\ & . & & . \\ & & . & . \\ & & & . \\ 1 & & & 0 \end{bmatrix},$$

and note that  $P_n^2 = I_n$ .

If  $a_0 \neq \pm 1$ , then  $C$  can be factorize as follows

$$C = \begin{bmatrix} 0 & 1 \\ P_{n-1} & 0 \end{bmatrix} \begin{bmatrix} P_{n-1} & \underline{a} \\ 0 & a_o \end{bmatrix} = P_n \begin{bmatrix} P_{n-1} & \underline{a} \\ 0 & a_o \end{bmatrix} \quad \dots \quad (2.3)$$

where  $\underline{a} = [a_{n-1}, \dots, a_1]^T$ .

By Lemma 2.1(b)  $P_n$  is diagonalizable with eigenvalues  $\pm 1$  and by Lemma 2.1(c) the second factor of equation 2.3 is also diagonalizable.

If  $a_o = \pm 1$  then consider the following two cases:

Case 1  $|F| > 5$ .

It is possible in this case to choose  $a \in F$  ( $a \neq 0$ ), such that  $a^2 \neq \pm a_o$ . Then  $C$  can be factorized as follows

$$C = \begin{bmatrix} 0 & a^2 \\ P_{n-1} & 0 \end{bmatrix} \begin{bmatrix} P_{n-1} & \underline{a} \\ 0 & a_o/a^2 \end{bmatrix}$$

The first factor is diagonalizable by Lemma 2.1(b) and the second is diagonalizable by Lemma 2.1(c), since  $a_o/a^2 \neq \pm 1$ .

Case 2  $F = GF(5)$ .

If  $n$  is even, then  $C$  can be factorized as follows

$$C = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & P_{\frac{n}{2}-1} & 0 \\ 0 & -\frac{1}{a_o} & 0 & 0 \\ -P_{\frac{n}{2}-1} & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & -P_{\frac{n}{2}-1} & -\underline{a}_1 \\ 0 & -a_o & 0 & -a_o a_{n/2} \\ P_{\frac{n}{2}-1} & 0 & 0 & \underline{a}_2 \\ 0 & 0 & 0 & a_o \end{bmatrix} \quad \dots \quad (2.4)$$

where  $\underline{a}_1 = [a_{n-1}, \dots, a_{\frac{n}{2}+1}]^T$  and  $\underline{a}_2 = [a_{\frac{n}{2}-1}, \dots, a_1]^T$ .

The first factor on the right of equation 2.4 is similar to

$$\begin{bmatrix} 0 & 0 & -\frac{1}{a_o} \\ 0 & -P_{\frac{n}{2}-1} & 0 \\ P_{\frac{n}{2}} & 0 & 0 \end{bmatrix}.$$

This is because one can multiply both sides by

$$X = X^{-1} = \begin{bmatrix} 0 & I_{\frac{n}{2}} \\ I_{\frac{n}{2}} & 0 \end{bmatrix}.$$

Now  $\gamma_{a_o} = \pm 1$ :

Hence if  $\gamma_{a_o} = -1$ , then this matrix is equal to

$$\begin{bmatrix} 0 & 0 & 1 \\ 0 & 2^2 P_{\frac{n}{2}-1} & 0 \\ P_{\frac{n}{2}} & 0 & 0 \end{bmatrix}.$$

If  $\gamma_{a_o} = 1$ , then it is equal to

$$\begin{bmatrix} 0 & 2^2 P_{\frac{n}{2}} \\ P_{\frac{n}{2}} & 0 \end{bmatrix},$$

since  $2^2 = -1$ .

In either case, by Lemma 2.1(b) the first factor on the right of equation (2.4) is diagonalizable.

Partition the second factor on the right of equation (2.4) as follows

$$\left[ \begin{array}{ccc|c} 0 & 0 & -P_{\frac{n}{2}-1} & -\underline{a}_1 \\ 0 & -a_o & 0 & -a_o a_{\frac{n}{2}} \\ P_{\frac{n}{2}-1} & 0 & 0 & \underline{a}_2 \\ \hline 0 & 0 & 0 & a_o \end{array} \right] = \left[ \begin{array}{ccc|c} 0 & 0 & 2^2 P_{\frac{n}{2}-1} & -\underline{a}_1 \\ 0 & -a_o & 0 & -a_o a_{\frac{n}{2}} \\ P_{\frac{n}{2}-1} & 0 & 0 & \underline{a}_2 \\ \hline 0 & 0 & 0 & a_o \end{array} \right].$$

By Lemma 2.1(b) the (1,1) block is diagonalizable with eigenvalues  $\pm 2$  and  $-a_o$ . Thus  $a_o$  is not an eigenvalue of this block, and so by Lemma 2.1(c) the whole matrix is diagonalizable.

If  $n$  is odd, then  $C$  can be factorized as follows

$$\left[ \begin{array}{cc} 0 & P_{\frac{n+1}{2}} \\ -P_{\frac{n-1}{2}} & 0 \end{array} \right] \left[ \begin{array}{ccc} 0 & -P_{\frac{n-1}{2}} & -\underline{a}_3 \\ P_{\frac{n-1}{2}} & 0 & \underline{a}_4 \\ 0 & 0 & a_o \end{array} \right],$$

where  $\underline{a}_3 = [a_{n-1}, \dots, a_{\frac{n+1}{2}}]^T$  and  $\underline{a}_4 = [a_{\frac{n-1}{2}}, \dots, a_1]^T$ .

It follows as before that both factors are diagonalizable. ■

In the example below, a  $4 \times 4$  companion matrix over  $GF(5)$  is factorized into two diagonal matrices. It is a direct application of Theorem 2.2.

**Example 2.5**

$$\text{Let } C = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 3 \end{bmatrix} \in M_4(GF(5)).$$

Then,

$$C = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & -1 & -3 \\ 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 2 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 4 & 0 & 0 \\ 4 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 4 & 2 \\ 0 & 4 & 0 & 0 \\ 1 & 0 & 0 & 2 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$= AB.$$

$A$  has two eigenvalues, namely 3 and 2. Their respective eigenspaces are

$$E_3 = \text{Span} \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \\ 3 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 3 \\ 0 \end{bmatrix} \right\} \quad \text{and} \quad E_2 = \text{Span} \left\{ \begin{bmatrix} 3 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 3 \\ 1 \\ 0 \end{bmatrix} \right\}.$$

Thus  $A$  is diagonalizable, and  $P^{-1}AP = \text{diag}[3, 3, 2, 2]$ ,  
where



$$P = \begin{bmatrix} 1 & 0 & 3 & 0 \\ 0 & 1 & 0 & 3 \\ 0 & 3 & 0 & 1 \\ 3 & 0 & 1 & 0 \end{bmatrix} \quad \text{and} \quad P^{-1} = \begin{bmatrix} 3 & 0 & 0 & 1 \\ 0 & 3 & 1 & 0 \\ 1 & 0 & 0 & 3 \\ 0 & 1 & 3 & 0 \end{bmatrix}.$$

$B$  has four distinct eigenvalues, namely 1, 2, 3 and 4.

Thus  $B$  is diagonalizable, and  $Q^{-1}BQ = \text{diag}[1, 2, 3, 4]$ , where

$$Q = \begin{bmatrix} 0 & 2 & 3 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 \\ 3 & 0 & 0 & 0 \end{bmatrix} \quad \text{and} \quad Q^{-1} = \begin{bmatrix} 0 & 0 & 0 & 2 \\ 4 & 0 & 3 & 4 \\ 1 & 0 & 3 & 4 \\ 0 & 1 & 0 & 0 \end{bmatrix}.$$

□

The following corollary is true for all fields except those of characteristic 2.

**Corollary 2.6 (Botha [3])**

Any nilpotent matrix  $A$  over a field  $F$  with  $\text{char}(F) \neq 2$  is a product of two diagonalizable matrices.

**Proof**

By the Rational Canonical Form Theorem

$$A \sim A_1 \oplus \dots \oplus A_k,$$

$$\text{where } A_i = \begin{bmatrix} 0 & 0 \\ I_{n_i-1} & 0 \end{bmatrix}, \quad i = 1, \dots, k.$$

The result follows from the first part of the proof of Theorem 2.2, since  $a_0 = 0$ , that is  $a_0 \neq \pm 1$  (see equation 2.3 in the proof of Theorem 2.2). ■

## 2.2 DIAGONALIZABLE FACTORIZATION OVER THE FIELD $GF(3)$ .

This case was also investigated and solved in [3] by Botha. It turns out that the only time a nonsingular matrix over  $GF(3)$  can be factorized into two diagonalizable matrices is if, and only if, the matrix is similar to its inverse. This follows from [11] by Hoffman and Paige, where it was proved that a nonsingular matrix  $A$  over an arbitrary field is the product of two involutions if, and only if,  $A \sim A^{-1}$ . The result of [11] by Hoffman and Paige will be treated in a later chapter of this dissertation.

It remains to note that an involution over any field, other than a field of characteristic two, is always diagonalizable. Since if  $A \in M_n(F)$  is an involution, then

$$a(x) = x^2 - 1 = (x - 1)(x + 1)$$

is an annihilating polynomial of  $A$ .

If  $A = \pm I$ , then the result is clear.

If  $A \neq \pm I$ , then  $a(x)$  is the minimal polynomial of  $A$ .

If  $\text{char}(F) \neq 2$ , then the multiplicity of each linear factor is one. If  $\text{char}(F) = 2$ , then  $1 = -1$ , therefore  $a(x) = x^2 - 1 = (x - 1)^2$ , and so the minimal polynomial contains a linear factor of multiplicity larger than one, and hence is not diagonalizable.

**Theorem 2.7 (Botha [3])**

Let  $A \in M_n(\text{GF}(3))$ :

- (a) Then  $A$  can be written as a product of three diagonalizable matrices;
- (b) If  $A$  is nonsingular, then  $A$  is a product of two diagonalizable matrices if, and only if,  $A \sim A^{-1}$ .

**Proof**

- (a) Firstly suppose that  $A$  is nonsingular; then  $A$  is a direct sum of companion matrices of the form

$$C = \begin{bmatrix} 0 & . & . & . & 0 & a_o \\ 1 & . & . & . & . & . \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ . & . & . & 0 & a_{n-2} & . \\ 0 & . & . & 1 & a_{n-1} & . \end{bmatrix}, \quad a_o \neq 0.$$

$C$  can be factorized as follows

$$C = P_n \begin{bmatrix} -P_{n-1} & 0 \\ 0 & a_o \end{bmatrix} \begin{bmatrix} -I_{n-1} & -\underline{a} \\ 0 & 1 \end{bmatrix},$$

where  $P_n$  was defined in the proof of Theorem 2.2 and  $\underline{a} = [a_1, \dots, a_{n-1}]$ .

The first factor is diagonalizable by Lemma 2.1(b), the second by Lemma 2.1(a) and the third by Lemma 2.1(c).

If  $A$  is singular, then by Lemma A.6,  $A = N \oplus K$  where  $N$  is nilpotent and  $K$  is nonsingular. By Corollary 2.6,  $N$  is a product of two diagonalizable matrices  $N_1$  and  $N_2$  (say) and by the first part of this proof,  $K$  is a product of three diagonalizable matrices  $K_1$ ,  $K_2$  and  $K_3$ . Thus

$$A = (N_1 N_2) \oplus (K_1 K_2 K_3) = (N_1 \oplus K_1)(N_2 \oplus K_2)(I \oplus K_3),$$

where  $I$  is the identity matrix the same size as  $N$ . The last three factors above are all diagonalizable by Lemma 2.1(a).

In general, the number three is minimal.

For example if

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}, \quad \text{then} \quad A^{-1} = \begin{bmatrix} -1 & 1 \\ 1 & 0 \end{bmatrix},$$

and these matrices are not similar, since their characteristic polynomials differ. Hence by part (b) of this theorem,  $A$  cannot be expressed as a product of two diagonalizable matrices.

- (b) Suppose  $A = EF$ , where  $E$  and  $F$  are both diagonalizable. Then there exist invertible matrices  $P$  and  $Q$  in  $M_n(GF(3))$ , such that  $PEP^{-1} = D_1$  and  $QFQ^{-1} = D_2$ , where  $D_i$  ( $i = 1, 2$ ) are diagonal matrices. The only nonsingular, diagonal matrices of  $M_n(GF(3))$  have  $+1$  and  $-1$  in the diagonal, which makes them involutions. Thus  $E$  and  $F$  are involutions. By the note above,  $A \sim A^{-1}$ . The converse also follows from the preceding note. ■

### 2.3 DIAGONALIZABLE FACTORIZATION OVER A FIELD OF CHARACTERISTIC 2 WITH AT LEAST FOUR ELEMENTS.

In [4], Botha showed that any square matrix over a field of characteristic two, with at least four elements, is the product of two diagonalizable matrices.

#### Definition 2.8

A square matrix  $L = [l_{ij}]_{i,j=1}^n$  is called  $S_2$ -lower triangular if  $l_{ij} = 0$  for  $i - 1 \leq j$ . That is

$$L = \begin{bmatrix} 0 & & & & & 0 \\ 0 & 0 & & & & \\ l_{31} & 0 & 0 & & & \\ \cdot & & \cdot & \cdot & & \cdot \\ \cdot & & & \cdot & \cdot & \cdot \\ \cdot & & & & \cdot & \cdot \\ l_{n1} & \cdot & \cdot & \cdot & l_{nn-2} & 0 & 0 \end{bmatrix}.$$

#### Lemma 2.9 (Botha [4])

Let  $A = C + L$  be a square matrix over an arbitrary field  $F$ , with  $C$  a companion matrix and  $L$   $S_2$ -lower triangular. Then  $A$  is similar to a companion matrix. Moreover, the similarity  $X^{-1}AX$  can be achieved by a lower triangular matrix  $X$  with 1's on the main diagonal and last row equal to  $[0 \dots 0 \ 1]$ .

**Proof**

For  $n = 2$  the result is trivial.

$$\text{For } n = 3, \text{ then } A = \begin{bmatrix} 0 & 0 & c_0 \\ 1 & 0 & c_1 \\ a & 1 & c_2 \end{bmatrix} \text{ and } X = \begin{bmatrix} 1 & 0 & 0 \\ -a & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

completes the proof in this case.

Let  $A \in M_n(F)$ ,  $n \geq 4$  and assume that the result holds for matrices of the size  $n - 1$ .

Let  $A'$  denote the matrix obtained from  $A$  by deleting the first row and column. Then there exists a nonsingular matrix  $X' \in M_{n-1}(F)$  of the specified form, such that

$$(X')^{-1} A' X' = C',$$

where  $C'$  is a companion matrix.

$$\text{Let } X = \left[ \begin{array}{c|cccc} 1 & 0 & & & 0 \\ \hline x_1 & & & & \\ \cdot & & & & \\ \cdot & & X' & & \\ \cdot & & & & \\ x_{n-2} & & & & \\ 0 & & & & \end{array} \right]$$

$$\text{and } \begin{pmatrix} y_1 \\ \cdot \\ \cdot \\ \cdot \\ y_{n-2} \\ y_{n-1} \end{pmatrix} = -(X')^{-1} \begin{pmatrix} x_1 \\ \cdot \\ \cdot \\ \cdot \\ x_{n-2} \\ 0 \end{pmatrix}.$$



$$\begin{aligned}
X^{-1}AX &= \left[ \begin{array}{c|cccc} 1 & 0 & . & . & 0 \\ \hline y_1 & & & & \\ . & & & & \\ . & & & & \\ . & & & & \\ y_{n-2} & & & & \\ y_{n-1} & & & & \end{array} \right] \left[ \begin{array}{c|cccc} 0 & 0 & . & . & 0 & c_0 \\ \hline 1 & & & & & \\ a_{31} & & & & & \\ . & & & & & \\ . & & & & & \\ . & & & & & \\ a_{n1} & & & & & \end{array} \right] \left[ \begin{array}{c|cccc} 1 & 0 & . & . & 0 \\ \hline x_1 & & & & \\ . & & & & \\ . & & & & \\ . & & & & \\ x_{n-2} & & & & \\ 0 & & & & \end{array} \right] \\
&= \left[ \begin{array}{c|cccc} 0 & 0 & . & . & 0 & c_0 \\ \hline \alpha_1 & & & & & \\ . & & & & & \\ . & & & & & \\ . & & & & & \\ . & & & & & \\ \alpha_{n-1} & & & & & \end{array} \right] B + C',
\end{aligned}$$

where

$$B = \begin{bmatrix} y_1 \\ . \\ . \\ . \\ y_{n-1} \end{bmatrix} \begin{bmatrix} 0 & . & . & . & 0 & c_0 \end{bmatrix}$$

and

$$\begin{bmatrix} \alpha_1 \\ . \\ . \\ . \\ \alpha_{n-1} \end{bmatrix} = (X')^{-1} \begin{bmatrix} 1 \\ a_{31} \\ . \\ . \\ . \\ a_{n1} \end{bmatrix} + A' \begin{bmatrix} x_1 \\ . \\ . \\ . \\ x_{n-2} \\ 0 \end{bmatrix}.$$

$B + C'$  is clearly a companion matrix.

It remains to show that  $\begin{bmatrix} x_1 \\ . \\ . \\ . \\ x_{n-2} \\ 0 \end{bmatrix}$  can be chosen so that,  $\begin{bmatrix} \alpha_1 \\ . \\ . \\ . \\ \alpha_{n-1} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ . \\ . \\ 0 \end{bmatrix}.$

Let  $\begin{bmatrix} x_1 \\ \cdot \\ \cdot \\ \cdot \\ x_{n-2} \end{bmatrix} = D^{-1} \begin{bmatrix} x_{31} - a_{31} \\ \cdot \\ \cdot \\ \cdot \\ x_{n-1,1} - a_{n-1,1} \\ -a_{n1} \end{bmatrix}$ , where  $\begin{bmatrix} 1 \\ x_{31} \\ \cdot \\ \cdot \\ x_{n-1,1} \\ 0 \end{bmatrix}$  is the first column

of  $X'$  and  $D = \begin{bmatrix} 1 & 0 & \cdot & \cdot & \cdot & 0 \\ a_{42} & 1 & 0 & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & 0 \\ a_{n2} & \cdot & \cdot & \cdot & a_{nn-2} & 1 \end{bmatrix}$ .

**Lemma 2.10 (Botha [4])**

The following holds for an arbitrary field  $F$ .

For  $n$  even, let

$$A = \begin{bmatrix} a_1 & & & & & & & b_1 \\ & \cdot & & & & & & \\ & & \cdot & & & & & \\ & & & \cdot & & & & \\ & & & & \cdot & & & \\ & & & & & a_{n/2} & b_{n/2} & \\ & & & & & c_{n/2} & d_{n/2} & \\ & & & & & & \cdot & \\ & & & & & & & \cdot \\ & & & & & & & \\ & & & & & & & \\ & & & & & & & \\ c_1 & & & & & & & d_1 \end{bmatrix}$$

then  $A \sim \bigoplus_{i=1}^{n/2} \begin{bmatrix} a_i & b_i \\ c_i & d_i \end{bmatrix}$ .

For  $n$  odd, let

$$A = \begin{bmatrix} a_1 & & & & & & & b_1 \\ & \ddots & & & & & & \\ & & \ddots & & & & & \\ & & & \ddots & & & & \\ & & & & a_{(n-1)/2} & 0 & b_{(n-1)/2} \\ & & & & 0 & e & 0 \\ & & & & c_{(n-1)/2} & 0 & d_{(n-1)/2} \\ & & & & & \ddots & & \\ & & & & & & \ddots & \\ c_1 & & & & & & & d_1 \end{bmatrix}$$

$$\text{then } A \sim [e] \oplus \left( \bigoplus_{i=1}^{(n-1)/2} \begin{bmatrix} a_i & b_i \\ c_i & d_i \end{bmatrix} \right).$$

### **Proof**

The result is trivial for  $n = 1$  or  $n = 2$ . Let  $n > 2$  and suppose the result is true for all matrices of size  $n - 2$ .

That is, if  $A \in M_{n-2}(F)$  and  $n$  is even, then

$$A \sim \bigoplus_{i=1}^{(n-2)/2} \begin{bmatrix} a_i & b_i \\ c_i & d_i \end{bmatrix},$$

and if  $n$  is odd, then

$$A \sim [e] \oplus \left( \bigoplus_{i=1}^{(n-3)/2} \begin{bmatrix} a_i & b_i \\ c_i & d_i \end{bmatrix} \right).$$

$$\text{Let } X = \begin{bmatrix} 0 & 1 & 0 \\ I_{n-2} & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \text{ then } X^{-1} = \begin{bmatrix} 0 & I_{n-2} & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \text{ and}$$

$$X^{-1} A X = B \oplus \begin{bmatrix} a_1 & b_1 \\ c_1 & d_1 \end{bmatrix},$$

where  $B \in M_{n-2}(F)$  is obtained from  $A$  by deleting the first and last rows and columns of  $A$ . The result now follows from the induction hypothesis. ■

In Lemma 2.11 and Theorem 2.14 which follow,  $F$  is a field of characteristic two containing at least four elements. That is,  $2 \cdot b = 0$  for all  $b \in F$ , and  $F$  contains the subfield  $GF(4) = \{0, 1, a, a^{-1}\}$ , where  $a + a^{-1} = 1$  and  $a^2 = a^{-1}$ . It is worth noting that the following  $2 \times 2$  matrices over  $GF(4)$  are diagonalizable

$$\begin{bmatrix} 0 & 1 \\ a^{-1} & a \end{bmatrix}, \begin{bmatrix} a^{-1} & a \\ 1 & 0 \end{bmatrix} \text{ and } \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}.$$

This fact will be used in the proof of Lemma 2.11. It is easily verified by determining the eigenvalues for each matrix. In each case they are distinct.

**Lemma 2.11 (Botha [4])**

Every companion matrix  $C$  over a characteristic two field  $F$  with at least four elements is a product of two diagonalizable matrices.



Let  $Y$  denote the first matrix on the left. By Lemma 2.10

$$Y \sim \begin{bmatrix} 0 & 1 \\ v^{-1} & v \end{bmatrix} \oplus \begin{bmatrix} 0 & 1 \\ u^{-1} & u \end{bmatrix} \oplus \dots \oplus \begin{bmatrix} 0 & 1 \\ u^{-1} & u \end{bmatrix}.$$

By the note preceding this lemma, each of these  $2 \times 2$  matrices are diagonalizable, and hence by Lemma 2.1(a)  $Y$  is diagonalizable.

Let  $Z$  denote the second matrix on the left. By Lemma 2.10 the (1,1) block of  $Z$  is similar to

$$[v] \oplus \begin{bmatrix} u^{-1} & u \\ 1 & 0 \end{bmatrix} \oplus \dots \oplus \begin{bmatrix} u^{-1} & u \\ 1 & 0 \end{bmatrix}.$$

Again by the note preceding this lemma, each  $2 \times 2$  matrix is diagonalizable, and hence by Lemma 2.1(a) the (1,1) block of  $Z$  is diagonalizable. By the way  $u$  and  $v$  were chosen  $c_o$  is not an eigenvalue of this matrix. Since if  $c_o = a^{-1}$  (or  $a$ ) the eigenvalues are  $a$  and 1 (respectively  $a^{-1}$  and 1); if  $c_o = 1$  the eigenvalues are  $a$  and  $a^{-1}$ ; if  $c_o \notin GF(4)^\times$  then this matrix has eigenvalues 1,  $a$  and  $a^{-1}$ . Therefore by Lemma 2.1(c),  $Z$  is diagonalizable.

Let  $V$  denote the matrix on the right. By Lemma 2.9

$X^{-1} V X = C'$ , where  $C'$  is a companion matrix and  $X$  is a lower triangular matrix with 1's on the main diagonal and the last row of the form  $[0 \dots 0 \ 1]$ .

$$\therefore X^{-1} Y Z X = C'.$$

It remains to show that the values  $x_1, \dots, x_{n-1}$  in the last column of  $Z$  can be chosen so that  $C' = C$ .

Because of the nature of  $X$ , the product  $ZX$  does not affect the last column of  $Z$ . Therefore the last column of  $C'$  is given by

$$X^{-1} Y \begin{bmatrix} x_1 \\ \cdot \\ \cdot \\ \cdot \\ x_{n-1} \\ c_o \end{bmatrix} \quad \dots \quad (2.12).$$

Since  $X^{-1}$  and  $Y$  are both nonsingular, there exists a unique vector  $x = [x_1, \dots, x_{n-1}, x_n]^T$ , such that

$$Y^{-1} X \begin{bmatrix} c_o \\ \cdot \\ \cdot \\ \cdot \\ c_{n-2} \\ c_{n-1} \end{bmatrix} = \begin{bmatrix} x_1 \\ \cdot \\ \cdot \\ \cdot \\ x_{n-1} \\ x_n \end{bmatrix} \quad \dots \quad (2.13).$$

A direct calculation of  $Y^{-1} X$  shows that the last row is  $[1 \ 0 \ \dots \ 0]$ , thus  $x_n = c_o$ . Putting equation 2.13 into equation 2.12, the last column of  $C'$  is  $[c_o, \dots, c_{n-1}]^T$ , and hence  $C = C'$ .





## 2.4 DIAGONALIZABLE FACTORIZATION OVER THE FIELD $GF(2)$ .

In [2], Ballantine proved that any matrix  $A \in M_n(F)$ , where  $F$  is an arbitrary field, can be expressed as a product of  $k$  idempotent matrices if, and only if,  $\text{rank}(I - A) \leq k \cdot \text{nullity}(A)$ . The proof of Ballantine's result will be discussed in a later chapter. However it does give some insight into the diagonalizable factorization of a matrix over the field  $GF(2)$ .

Before making use of this result we note the following:

- (i) An idempotent matrix is diagonalizable over any field. This can be seen since if  $A$  is an idempotent matrix over any field  $F$ , and  $A$  is not the zero or identity matrix, then  $m(\lambda) = \lambda(\lambda - 1)$  is the minimal polynomial of  $A$ . Thus the minimal polynomial of  $A$  can be expressed as linear factors of multiplicity one over  $F$ .
- (ii) Clearly from (i), the identity matrix is the only nonsingular idempotent matrix.

(iii) If  $A \in M_n(GF(2))$ , then the concept of idempotence and diagonalizability are equivalent. Suppose that  $A$  is diagonalizable. Then there exists a nonsingular matrix  $P \in M_n(GF(2))$ , such that  $A = P D P^{-1}$ , where  $D$  is a diagonal matrix. But the diagonal of  $D$  consists of zeros and ones and so is idempotent. The result now follows from Lemma A.4.

The following results now follow from Ballantine's theorem and the notes above.

**Theorem 2.15 (Botha [3])**

A matrix  $A$  over  $F = GF(2)$  is a product of  $k \geq 1$  diagonalizable matrices if, and only if,

$$\text{rank}(I - A) \leq k \cdot \text{nullity}(A).$$

**Corollary 2.16 (Botha [3])**

The only nonsingular matrix over  $F = GF(2)$  expressible as a product of diagonalizable matrices, is the identity matrix.

# CHAPTER 3

## PRODUCTS OF IDEMPOTENT MATRICES.

### Definition 3.1

A matrix  $S \in M_n(F)$ , where  $F$  is an arbitrary field, is idempotent if  $S^2 = S$ .

The following well-known lemma gives a necessary and sufficient condition for a matrix to be idempotent. It will be used to prove the main result in this chapter.

### Lemma 3.2

Let  $S \in M_n(F)$ . Then  $S$  is idempotent if, and only if,  $\ker(S) = \text{Im}(I - S)$ .

### Proof

Firstly for any matrix  $S$  over  $F$ ,  $\ker(S) \subseteq \text{Im}(I - S)$ . Since if  $x \in \ker(S)$  then  $(I - S)x = Ix - Sx = x$ . That is  $x \in \text{Im}(I - S)$ . Now suppose that  $S^2 = S$  and let  $x \in \text{Im}(I - S)$ . Then there exists  $y \in F^n$ , such that  $(I - S)y = x$ , and so  $S(I - S)y = Sx$ , which implies that  $Sy - S^2y = Sx$ . Therefore  $Sx = 0$ .

Conversely let  $x \in F^n$ , then there exists  $y \in \text{Im}(I - S)$ , such that  $(I-S)x=y$ , which implies that  $S(I-S)x = Sy = 0$ . Therefore  $Sx - S^2x = 0$ , and the result follows. ■

In [5], Erdos proved that every singular matrix  $A$  over any field can be written as a product of idempotent matrices. In [2], Ballantine established a more detailed result by prescribing the number of idempotent matrices required in the factorization. Here follows a proof of Ballantine's results.

In what follows a row-partitioned matrix will be denoted as

$$\begin{bmatrix} A \\ B \end{bmatrix} = \text{col}[A \ B], \quad \begin{bmatrix} A \\ B \\ C \end{bmatrix} = \text{col}[A \ B \ C], \quad \text{etc} \dots$$

so that in the expression  $\text{col}[A \ B \ C]$  it is understood that  $A$ ,  $B$  and  $C$  all have the same number of columns, but not necessarily the same number of rows.

**Lemma 3.3**      (Ballantine [2])

Let  $S$  be an  $n \times n$  matrix,  $c = \text{nullity}(S)$  and  $d = \text{nullity}(I - S)$ . Then  $c + d \leq n$  and  $S$  is similar to a matrix  $T$ , such that the last  $c$  columns of  $T$  and the first  $d$  rows of  $I-T$  are zero.

**Proof**

The only vector common to both  $\ker(S)$  and  $\ker(I-S)$  is the zero vector, therefore  $c + d \leq n$ .

Now  $S$  can be reduced by elementary column operations to a matrix with the last  $c$  columns being zero. Therefore there exists an invertible matrix  $P$  (consisting of products of elementary matrices), such that  $SP$  has zeros in the last  $c$  columns. Then  $P^{-1}SP$  is also of this form, that is,  $P^{-1}SP = [S_o \ 0]$ , where  $S_o \in M_{n \times (n-c)}(F)$  and  $0$  is the  $n \times c$  zero matrix.

Partition  $P^{-1}SP$  as follows

$$P^{-1}SP = \begin{bmatrix} S_1 & 0 \\ S_2 & 0 \end{bmatrix} \begin{matrix} (n-c) \\ (c) \end{matrix},$$

where the size of each partition is indicated in parentheses on the right and the diagonal blocks are square.

Now

$$P^{-1}(I-S)P = \begin{bmatrix} I_{n-c} - S_1 & 0 \\ -S_2 & I_c \end{bmatrix} \begin{matrix} (n-c) \\ (c) \end{matrix}$$

and since  $d \leq n-c$ , there exists a nonsingular matrix  $P_1 \in M_{n-c}(F)$ , such that the first  $d$  rows of  $P_1(I_{n-c} - S_1)$  are zero, therefore

$$P_1(I_{n-c} - S_1)P_1^{-1} = \begin{bmatrix} 0 \\ S_3 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ S_4 & S_5 \end{bmatrix} \begin{matrix} (d) \\ (n-c-d) \end{matrix} \dots \quad (3.4).$$

Let  $Q = P(P_1^{-1} \oplus I_c)$  then:

$$T = Q^{-1} S Q = \begin{bmatrix} P_1 S_1 P_1^{-1} & 0 \\ S_2 P_1^{-1} & 0 \end{bmatrix} \quad \dots \quad (3.5).$$

From equation (3.4):

$$P_1 S_1 P_1^{-1} = \begin{bmatrix} I_d & 0 \\ -S_4 & I_{n-d-c} - S_5 \end{bmatrix} = \begin{bmatrix} I_d & 0 \\ A_o & A \end{bmatrix} \quad \begin{matrix} (d) \\ (n-c-d) \end{matrix}$$

Substitute this into equation (3.5), with  $S_2 P_1^{-1}$  partitioned as  $[C_o \ C]$

$$T = \begin{bmatrix} I_d & 0 & 0 \\ A_o & A & 0 \\ C_o & C & 0 \end{bmatrix} \quad \begin{matrix} (d) \\ (a) \\ (c) \end{matrix} \quad \text{where } a = n - d - c. \quad \blacksquare$$

**Lemma 3.6 (Ballantine [2])**

Let  $A \in M_{a \times (a+b)}(F)$ ,  $B \in M_{b \times (a+b)}(F)$  and  $C \in M_{c \times (a+b)}(F)$  with  $a \leq c$  and either  $A=0$  or  $\text{rank}(B) = b$ . Then there is a matrix  $D \in M_{c \times a}(F)$ , such that the matrices

$$\begin{bmatrix} B \\ C + D A \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} A \\ B \\ C \end{bmatrix}$$

have the same row space.

**Proof**

If  $A = 0$  then take  $D = 0$ , so assume  $\text{rank}(B) = b$ .



Let  $x_1, \dots, x_b$  be the rows of  $B$ . Then since  $B$  has full row rank, these vectors are linearly independent. Let  $P \in M_c(F)$  and  $Q \in M_a(F)$  be permutation matrices so that if  $y_1, \dots, y_c$  are the rows of  $PC$  then  $x_1, \dots, x_b, y_1, \dots, y_l$

$(l \leq c)$  is a basis for  $\text{rsp} \begin{bmatrix} B \\ C \end{bmatrix}$  and if  $z_1, \dots, z_a$  are the rows of

$QA$  then  $x_1, \dots, x_b, y_1, \dots, y_l, z_1, \dots, z_m$  ( $m \leq a$ ) is a

basis for  $\text{rsp} \begin{bmatrix} A \\ B \\ C \end{bmatrix}$ . Since the dimension of the row space of a

matrix is equal to the dimension of its column space,

$b + l + m \leq a + b$ , which implies that  $l + m \leq a \leq c$ .

Let  $E \in M_{c \times a}(F)$  be the matrix  $E = \begin{bmatrix} 0 & 0 \\ I_m & 0 \\ 0 & 0 \end{bmatrix} \begin{matrix} (l) \\ (m) \\ (c - l - m) \end{matrix}$ .

Then  $PC + EQA = \text{col}[y_1, \dots, y_l, y_{l+1} + z_1, \dots, y_{l+m} + z_m, y_{l+m+1}, \dots, y_c]$ ,

therefore  $C + P^{-1}EQA$  is a permutation of the matrix on the

right. Now  $x_1, \dots, x_b, y_1, \dots, y_l, y_{l+1} + z_1, \dots, y_{l+m} + z_m$  are

linearly independent and therefore constitute a basis for

$\text{rsp} \begin{bmatrix} A \\ B \\ C \end{bmatrix}$ , but at the same time constitutes a basis for

$\text{rsp} \begin{bmatrix} B \\ C + DA \end{bmatrix}$ , where  $D = P^{-1}EQ$ . ■

**Note** If  $b = 0$  in the above lemma, then the matrix  $B$  is missing; but the result is still true, namely

$$\text{rsp}[C + DA] = \text{rsp} \begin{bmatrix} A \\ C \end{bmatrix}.$$

**Theorem 3.7** (Ballantine [2])

Let  $F$  be an arbitrary field and  $S \in M_n(F)$ . Then  $S = P_1 P_2 \dots P_k$  where  $P_i$  ( $1 \leq i \leq k$ ) is idempotent if, and only if,  $\text{rank}(I - S) \leq k \cdot \text{nullity}(S)$ .

**Note**

1. The theorem is trivial (but of little interest) if  $S$  is nonsingular. Since then  $\text{nullity}(S) = 0$ , and the last inequality implies that  $S = I$  which can trivially be expressed as the product of any  $k \geq 1$  copies of  $I$  (the only idempotent matrix which is also nonsingular).
2. It was shown in chapter 2 that an idempotent matrix is diagonalizable over any field  $F$ . Thus this theorem gives a factorization of a matrix (in particular a singular matrix) into diagonalizable matrices of a specific type, and the minimum number of such factors required. It is for this reason that in [3] Botha was able to use this result to factorize a singular matrix over the field  $F$  of order 2.

### Proof of Theorem 3.7

Suppose that  $P_1, P_2, \dots, P_k$  are idempotent matrices and that  $S = P_1 P_2 \dots P_k$ . Then:

$$\begin{aligned}
 & \text{rank}(I - S) \\
 &= \text{rank}((I - P_1) + (P_1 - P_1 P_2) + (P_1 P_2 - P_1 P_2 P_3) + \dots + (P_1 \dots P_{k-1} - P_1 \dots P_k)) \\
 &\leq \text{rank}(I - P_1) + \text{rank } P_1(I - P_2) + \text{rank } P_1 P_2(I - P_3) + \dots + \text{rank } P_1 \dots P_{k-1}(I - P_k) \\
 &\leq \sum_{j=1}^k \text{rank}(I - P_j), \quad \text{since } \text{rank}(AB) \leq \min\{\text{rank}(A), \text{rank}(B)\} \text{ for} \\
 &\quad \text{any } A, B \in M_n(F) \\
 &= \sum_{j=1}^k \text{nullity}(P_j), \quad \text{by Lemma 3.2} \\
 &\leq k \cdot \text{nullity}(P_1 \dots P_k) \\
 &= k \cdot \text{nullity}(S).
 \end{aligned}$$

Conversely, suppose that  $\text{rank}(I - S) \leq k \cdot \text{nullity}(S)$

For the case  $k = 1$ :

That is,  $\text{rank}(I - S) \leq \text{nullity}(S)$ , but  $\text{nullity}(S) \leq \text{rank}(I - S)$ . Thus  $\text{nullity}(S) = \text{rank}(I - S)$  and, by Lemma 3.2,  $S$  is idempotent.

For the case  $k=2$ :

By Lemmas A.4 and 3.3, it suffices to prove the theorem

$$\text{for the matrix } S = \begin{bmatrix} I_d & 0 & 0 \\ A_o & A & 0 \\ C_o & C & 0 \end{bmatrix} \quad \begin{matrix} (d) \\ (a) \\ (c) \end{matrix}, \quad \text{where } c = \text{nullity}(S),$$

$$d = \text{nullity}(I - S) \quad \text{and} \quad a = n - d - c.$$

By the hypothesis,  $\text{rank}(I - S) \leq 2 \cdot \text{nullity}(S)$ .

That is  $n - d \leq 2c$  or  $n - d - c \leq c$ , therefore  $a \leq c$ .

So we may apply Lemma 3.6 (with  $B$  missing) to  $\begin{bmatrix} A \\ C \end{bmatrix}$ .

Replace  $S$  with the similarity copy,

$$\begin{aligned} T &= \begin{bmatrix} I_d & 0 & 0 \\ 0 & I_a & 0 \\ 0 & D & I_c \end{bmatrix} \cdot \begin{bmatrix} I_d & 0 & 0 \\ A_o & A & 0 \\ C_o & C & 0 \end{bmatrix} \cdot \begin{bmatrix} I_d & 0 & 0 \\ 0 & I_a & 0 \\ 0 & -D & I_c \end{bmatrix} \\ &= \begin{bmatrix} I_d & 0 & 0 \\ A_o & A & 0 \\ C_o + DA_o & C + DA & 0 \end{bmatrix}, \end{aligned}$$

where  $D$  is the matrix given by Lemma 3.6, and

$\text{rsp}[C + DA] = \text{rsp}\begin{bmatrix} A \\ C \end{bmatrix}$ . In particular,  $\text{rsp}[A] \subseteq \text{rsp}[C + DA]$ ;

hence there exists an  $a \times c$  matrix  $H$ , such that

$$A = H(C + DA).$$

Then

$$T = \begin{bmatrix} I_d & 0 & 0 \\ A_o - HC_o - HDA_o & 0 & H \\ 0 & 0 & I_c \end{bmatrix} \cdot \begin{bmatrix} I_d & 0 & 0 \\ 0 & I_a & 0 \\ C_o + DA_o & C + DA & 0 \end{bmatrix}$$

and both factors are idempotent.

For the case  $k > 2$

Assume that  $R$  is a product of  $k - 1$  idempotent matrices, if

$$\text{rank}(I - R) \leq (k - 1) \cdot \text{nullity}(R).$$

Now suppose that  $S \in M_n(F)$  and  $\text{rank}(I - S) \leq k \cdot \text{nullity}(S)$ . Let  $c = \text{nullity}(S)$  and  $d = \text{nullity}(I - S)$ , then  $n - d \leq kc$ , which implies that  $n - d - (k-1)c \leq c$ . If  $a = n - d - (k-1)c \leq 0$ , then  $\text{rank}(I - S) \leq (k-1) \cdot \text{nullity}(S)$ . By the hypothesis,  $S$  is already a product of  $k-1$  idempotent matrices. So we assume that  $a = n - d - (k-1)c > 0$ . By Lemma A.4 and 3.3 it can be assumed that  $S$  has the form

$$S = \begin{bmatrix} I_d & 0 & 0 & 0 \\ A_o & A_1 & A_2 & 0 \\ B_o & B_1 & B_2 & 0 \\ C_o & C_1 & C_2 & 0 \end{bmatrix} \begin{matrix} (d) \\ (a) \\ (b) \\ (c) \end{matrix}$$

where the diagonal blocks are square, with the dimensions indicated in parenthesis. Now  $a + b + c + d = n$ ; therefore  $b = (k-2)c$ . Let  $A = [A_1 \ A_2]$ ,  $B = [B_1 \ B_2]$  and  $C = [C_1 \ C_2]$ , then we may assume that either  $A = 0$  or  $B$  has full rank. (Since if this is not the case then there exists a nonsingular matrix  $P \in M_{a+b}(F)$ , such that  $P \begin{bmatrix} A \\ B \end{bmatrix}$  has zeros in its top rows, and the remaining rows are linearly independent, and then  $PS P^{-1}$  has the required form).

$$a = n - d - (k-1)c = \text{rank}(I - S) - (k-1)c \leq kc - (k-1)c = c,$$

so we can apply Lemma 3.6, that is

$$rsp \begin{bmatrix} A \\ B \\ C \end{bmatrix} = rsp \begin{bmatrix} B \\ C + D A \end{bmatrix}$$

for some matrix  $D \in M_{c \times a}(F)$ . Replace  $S$  with the similarity copy  $T$ :

$$T = \begin{bmatrix} I_d & 0 & 0 & 0 \\ 0 & I_a & 0 & 0 \\ 0 & 0 & I_b & 0 \\ 0 & D & 0 & I_c \end{bmatrix} \cdot \begin{bmatrix} I_d & 0 & 0 & 0 \\ A_o & A_1 & A_2 & 0 \\ B_o & B_1 & B_2 & 0 \\ C_o & C_1 & C_2 & 0 \end{bmatrix} \cdot \begin{bmatrix} I_d & 0 & 0 & 0 \\ 0 & I_a & 0 & 0 \\ 0 & 0 & I_b & 0 \\ 0 & -D & 0 & I_c \end{bmatrix}$$

$$= \begin{bmatrix} I_d & 0 & 0 & 0 \\ A_o & A_1 & A_2 & 0 \\ B_o & B_1 & B_2 & 0 \\ D A_o + C_o & D A_1 + C_1 & D A_2 + C_2 & 0 \end{bmatrix}.$$

In particular  $rsp[A] \subseteq rsp \begin{bmatrix} B \\ C + D A \end{bmatrix}$ . Thus in what follows we

may assume that  $S$  is such that  $rsp[A] \subseteq rsp \begin{bmatrix} B \\ C \end{bmatrix}$ , hence there

exists  $H \in M_{a \times (b+c)}(F)$ , such that  $A = H \begin{bmatrix} B \\ C \end{bmatrix}$ . Partition  $H$  as

$H = [H_1 \ H_2]$  where  $H_1 \in M_{a \times b}(F)$  and  $H_2 \in M_{a \times c}(F)$ .

Then

$$[A_1 \ A_2] = [H_1 \ H_2] \cdot \begin{bmatrix} B_1 & B_2 \\ C_1 & C_2 \end{bmatrix} = [H_1 B_1 + H_2 C_1 \quad H_1 B_2 + H_2 C_2],$$

therefore  $A_1 = H_1 B_1 + H_2 C_1$  and  $A_2 = H_1 B_2 + H_2 C_2$ .

$$S = PR = \begin{bmatrix} I_d & 0 & 0 & 0 \\ A_o - H_1 B_o - H_2 C_o & 0 & H_1 & H_2 \\ 0 & 0 & I_c & 0 \\ 0 & 0 & 0 & I_b \end{bmatrix} \cdot \begin{bmatrix} I_d & 0 & 0 & 0 \\ 0 & I_a & 0 & 0 \\ B_o & B_1 & B_2 & 0 \\ C_o & C_1 & C_2 & 0 \end{bmatrix}$$

$P$  is idempotent and  $\text{rank}(I - R) \leq b + c = (k - 1)c = (k - 1) \cdot \text{nullity}(R)$ . Therefore by the induction hypothesis,  $R$  is a product of  $k-1$  idempotent matrices. This completes the proof. ■

The proof of Ballantine's theorem gives a way to determining these factors.

### Example 3.8

$$S = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ \hline 1 & 2 & 3 & 0 & 0 & 0 & 0 & 0 \\ 5 & 6 & 7 & 8 & 0 & 0 & 0 & 0 \\ 9 & 10 & 11 & 12 & 0 & 0 & 0 & 0 \\ 13 & 14 & 15 & 16 & 0 & 0 & 0 & 0 \\ 17 & 18 & 19 & 20 & 0 & 0 & 0 & 0 \end{bmatrix},$$

then  $\text{rank}(I - S) = 5 \leq 2 \cdot \text{nullity}(S) = 8$ .

Therefore there are two idempotent factors. These factors are not unique, since any  $H = [a \ b \ c \ d]$  where  $8a + 12b + 16c + 20d = 0$  will suffice according to the above proof. If  $H = [0 \ 0 \ 0 \ 0]$  then  $S = PR$  where



$$P = \left[ \begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ \hline 1 & 2 & 3 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{array} \right] \quad \text{and} \quad R = \left[ \begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 5 & 6 & 7 & 8 & 0 & 0 & 0 & 0 \\ 9 & 10 & 11 & 12 & 0 & 0 & 0 & 0 \\ 13 & 14 & 15 & 16 & 0 & 0 & 0 & 0 \\ 17 & 18 & 19 & 20 & 0 & 0 & 0 & 0 \end{array} \right]$$

If  $H = [-10 \ 2 \ 1 \ 2]$  then  $S = QR$  where

$$Q = \left[ \begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ \hline -14 & -8 & -2 & 0 & -10 & 2 & 1 & 2 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{array} \right].$$

□

When a singular matrix is expressed as a product of idempotent factors  $P_i$ , and each  $P_i$  is expressed as  $P_i = Q_i D_i Q_i^{-1}$  where  $D_i$  is a diagonal matrix, the only nonzero value in the diagonal is one.

Since it follows easily that  $\text{rank}(I - S) \leq n \leq n.\text{nullity}(S)$  for any singular matrix of order  $n$ , we have by Ballantine's Theorem that:

**Corollary 3.9**

Every singular matrix in  $M_n(F)$ , where  $F$  is an arbitrary field, is a product of at most  $n$  idempotent matrices. ■

**Note** There are singular  $n \times n$  matrices over  $F$  that are not a product of less than  $n$  idempotent matrices. For example the  $n \times n$  nilpotent Jordan block:

$$J_o = \begin{bmatrix} 0 & 1 & & 0 \\ & 0 & 1 & \\ & & \ddots & \ddots \\ & & & 0 & 1 \\ 0 & & & & 0 \end{bmatrix},$$

then  $nullity(S) = 1$  and so  $rank(I - J_o) = n = n \cdot nullity(S)$ .

## CHAPTER 4

### PRODUCTS OF INVOLUTIONS.

#### Definition 4.1

A square matrix  $A \in M_n(F)$ , where  $F$  is an arbitrary field, is called an involution, if  $A^2 = I$ .

For any positive integer  $n$ ,  $I_n$ ,  $-I_n$  and  $P_n$ , where  $P_n$  was defined in the proof of Theorem 2.2, are examples of involutions. Other than  $I_2$  and  $-I_2$ , all other  $2 \times 2$  involutions over  $F$  have the form:

$$\begin{bmatrix} a & b \\ c & -a \end{bmatrix} \text{ where } a^2 + bc = 1.$$

#### Lemma 4.2

Let  $F$  be a field of characteristic different from two. If  $A$  is an involution over  $F$ , then  $A$  is diagonalizable. Furthermore  $A \sim I_k \oplus -I_{n-k}$  for  $0 \leq k \leq n$  if, and only if,  $A$  is an involution over  $F$ .

#### Proof

The first part was shown in Section 2.2. If  $A \sim I_k \oplus -I_{n-k}$  then there exists an invertible matrix  $P \in M_n(F)$ , such that  $A = P [I_k \oplus -I_{n-k}] P^{-1}$ , and the result follows.

Conversely, if  $A = I_n$ , then  $k = n$ , if  $A = -I_n$ , then  $k = 0$ , otherwise the minimal polynomial of  $A$  is  $\lambda^2 - 1 = (\lambda - 1)(\lambda + 1)$  and thus  $A$  is diagonalizable with 1 and  $-1$  as the only eigenvalues. ■

The factorization of a matrix  $A$  into involutions is a specific type of diagonalizable factorization if  $\text{char}(F) \neq 2$ .

#### 4.1 PRODUCTS OF TWO INVOLUTIONS.

In [11], Hoffman and Paige showed that a necessary and sufficient condition for a nonsingular matrix to be a product of two involutions, is that the matrix be similar to its inverse. A proof of this result is given in this section. The following fact is used in the proof: the property of being the product of involutions is invariant under similarity. The proof is similar to that of Lemma A.1.

##### Definition 4.3

Let  $g(x) = x^m + a_{m-1}x^{m-1} + \dots + a_1x + a_0$  be a monic polynomial of degree  $m$  over a field  $F$ , such that  $a_0 \neq 0$ . Define  $\tilde{g}(x)$  to be the monic polynomial

$$\tilde{g}(x) = \frac{x^m}{a_0} \cdot g\left(\frac{1}{x}\right) = x^m + \frac{a_1}{a_0}x^{m-1} + \dots + \frac{a_{m-1}}{a_0}x + \frac{1}{a_0}.$$

**Lemma 4.4**

- (a)  $g(w) = 0$  if, and only if,  $\tilde{g}(\frac{1}{w}) = 0$ .
- (b) If  $g(x)$  is irreducible over  $F$  then so is  $\tilde{g}(x)$ .
- (c) If  $g(x) = h(x)k(x)$  then  $\tilde{g}(x) = \tilde{h}(x)\tilde{k}(x)$ .

**Proof:**

- (a) Since  $a_o \neq 0$ , if  $g(w) = 0$  then  $w \neq 0$ . Therefore

$$g(w) = 0 \text{ if, and only if, } \tilde{g}(\frac{1}{w}) = \frac{1}{a_o w^m} \cdot g(w) = 0.$$

- (b) Suppose that  $g(x)$  is irreducible and

$$\tilde{g}(x) = \frac{1}{a_o} (b_o x^p + \dots + b_{p-1} x + b_p) (c_o x^q + \dots + c_{q-1} x + c_q),$$

where  $p + q = m$  and  $p, q \geq 1$  (that is,  $\tilde{g}(x)$  is reducible).

$$\begin{aligned} g(x) &= a_o x^m \tilde{g}\left(\frac{1}{x}\right) \\ &= a_o x^m \frac{1}{a_o} \left( \frac{b_o}{x^p} + \dots + \frac{b_{p-1}}{x} + b_p \right) \left( \frac{c_o}{x^q} + \dots + \frac{c_{q-1}}{x} + c_q \right) \\ &= (b_o + \dots + b_{p-1} x^{p-1} + b_p x^p) (c_o + \dots + c_{q-1} x^{q-1} + c_q x^q) \end{aligned}$$

This is a contradiction since  $g(x)$  was assumed to be irreducible.

- (c) Let  $h(x) = \sum_{i=0}^p b_i x^i$  and  $k(x) = \sum_{i=0}^q c_i x^i$ , where  $p + q = m$ ,

then  $a_o = b_o c_o$  and:

$$\begin{aligned}
\tilde{h}(x)\tilde{k}(x) &= \frac{x^p}{b_o}h\left(\frac{1}{x}\right) \cdot \frac{x^q}{c_o}k\left(\frac{1}{x}\right) \\
&= \frac{x^{p+q}}{b_o c_o}h\left(\frac{1}{x}\right)k\left(\frac{1}{x}\right) \\
&= \frac{x^m}{a_o}g\left(\frac{1}{x}\right) \\
&= \tilde{g}(x)
\end{aligned}$$

■

**Definition 4.5**

The monic polynomial  $g(x)$  is said to be symmetric over a field  $F$  if  $g(x) = \tilde{g}(x)$ .

**Note:**

1. If  $g(x)$  is symmetric and  $a_o = 1$  then:

$$x^m + a_{m-1}x^{m-1} + \dots + a_1x + 1 = x^m + a_1x^{m-1} + \dots + a_{m-1}x + 1, \quad \text{and}$$

so  $a_{m-i} = a_i$ , where  $1 \leq i \leq m-1$ .

2. If  $g(x)$  is symmetric and  $a_o = -1$  then:

$$x^m + a_{m-1}x^{m-1} + \dots + a_1x - 1 = x^m - a_1x^{m-1} - \dots - a_{m-1}x - 1, \quad \text{and}$$

so  $a_{m-i} = -a_i$ , where  $1 \leq i \leq m-1$ . If, in addition,  $m$  is

even, then  $a_{m/2} = -a_{m/2}$ , that is,  $a_{m/2} = 0$ .

**Theorem 4.6 (Hoffman and Paige [11])**

Let  $A \in M_n(F)$  be invertible ( $n > 1$ ), then  $A$  is the product of two involutions if, and only if,  $A$  is similar to  $A^{-1}$ .

**Proof**

First suppose that  $A \sim A^{-1}$ . Since a product of involutions is invariant under similarity, we may assume that  $A = C_{h_1(x)} \oplus \dots \oplus C_{h_t(x)}$ , where the polynomials  $h_i(x)$  ( $1 \leq i \leq t$ ) are the (non-trivial) invariant factors of  $A$ .

Let  $h_i(x) = x^m + a_{m-1}x^{m-1} + \dots + a_1x + a_0$ , and define  $P_m \in M_m(F)$  to be the involution

$$P_m = \begin{bmatrix} 0 & . & . & 0 & 1 \\ . & & & 1 & 0 \\ . & & . & & \\ 0 & . & & & \\ 1 & 0 & . & . & 0 \end{bmatrix}.$$

Then

$$\begin{aligned} P_m^{-1} C_{\tilde{h}_i(x)} P_m &= \begin{bmatrix} 0 & P_{m-1} \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & -1/a_0 \\ I_{m-1} & b \end{bmatrix} \begin{bmatrix} 0 & P_{m-1} \\ 1 & 0 \end{bmatrix} \\ &= (C_{\tilde{h}_i(x)})^{-1}, \end{aligned}$$

where  $b = \left[ -\frac{a_{m-1}}{a_0}, \dots, -\frac{a_1}{a_0} \right]^T$ , that is,  $C_{\tilde{h}_i(x)} \sim (C_{h_i(x)})^{-1}$ .

Now,



$$A = \begin{bmatrix} C_{h_1(x)} & & \\ & \ddots & \\ & & C_{h_t(x)} \end{bmatrix} \sim A^{-1} = \begin{bmatrix} C_{h_1(x)}^{-1} & & \\ & \ddots & \\ & & C_{h_t(x)}^{-1} \end{bmatrix}$$

$$\sim \begin{bmatrix} C_{\tilde{h}_1(x)} & & \\ & \ddots & \\ & & C_{\tilde{h}_t(x)} \end{bmatrix},$$

and so we can conclude that  $\tilde{h}_j(x) = h_j(x)$  ( $1 \leq j \leq t$ ), since the invariant factors of  $A^{-1}$  are  $\tilde{h}_1(x), \dots, \tilde{h}_t(x)$  and must be the same as those for  $A$ .

Consider one companion matrix  $C_{m(x)}$  where  $m(x) = \tilde{m}(x)$ . Then when  $m(x)$  is expressed as the product of powers of prime polynomials, these primes must either be symmetric, or occur in pairs,  $q_i(x)$  and  $\tilde{q}_i(x)$ . That is

$$m(x) = \prod_{i=1}^r [p_i(x)]^{n_i} \cdot \prod_{i=1}^s [q_i(x)\tilde{q}_i(x)]^{m_i},$$

where the  $p_i(x)$ ,  $q_i(x)$  and  $\tilde{q}_i(x)$  are all distinct and irreducible and  $p_i(x)$  ( $1 \leq i \leq n_r$ ) are symmetric. This follows from Lemma 4.4(c).

By the Primary Decomposition Theorem

$$C_{m(x)} \sim \bigoplus_{i=1}^r C_{p_i(x)^{n_i}} \oplus \bigoplus_{i=1}^s \begin{bmatrix} C_{q_i(x)^{m_i}} & 0 \\ 0 & C_{\tilde{q}_i(x)^{m_i}} \end{bmatrix}.$$

By Lemma 4.4(c),  $\tilde{q}_i(x)^{m_i} = q_i(x)^{m_i}$ , therefore

$$C_{\tilde{q}_i(x)^{m_i}} = C_{q_i(x)^{m_i}} \sim (C_{q_i(x)^{m_i}})^{-1}.$$

Thus

$$C_{m(x)} \sim \bigoplus_{i=1}^r C_{p_i(x)^{n_i}} \oplus \bigoplus_{i=1}^s \begin{bmatrix} C_{q_i(x)^{m_i}} & 0 \\ 0 & (C_{q_i(x)^{m_i}})^{-1} \end{bmatrix}.$$

It remains to show that  $C_{p_i(x)^{n_i}}$  and  $\begin{bmatrix} C_{q_i(x)^{m_i}} & 0 \\ 0 & (C_{q_i(x)^{m_i}})^{-1} \end{bmatrix}$  are both

products of two involutions.

Now

$$\begin{bmatrix} C_{q_i(x)^{m_i}} & 0 \\ 0 & (C_{q_i(x)^{m_i}})^{-1} \end{bmatrix} = P \left( P \begin{bmatrix} C_{q_i(x)^{m_i}} & 0 \\ 0 & (C_{q_i(x)^{m_i}})^{-1} \end{bmatrix} \right),$$

where

$$P = \begin{bmatrix} 0 & I_{n_i} \\ I_{n_i} & 0 \end{bmatrix} \quad \text{with} \quad n_i = \deg q_i(x)^{m_i}$$

and it follows readily that both  $P$  and

$$P \begin{bmatrix} C_{q_i(x)^{m_i}} & 0 \\ 0 & \left(C_{q_i(x)^{m_i}}\right)^{-1} \end{bmatrix} = \begin{bmatrix} 0 & \left(C_{q_i(x)^{m_i}}\right)^{-1} \\ C_{q_i(x)^{m_i}} & 0 \end{bmatrix}$$

are involutions.

We have seen earlier in the proof, that  $P_m C_{\tilde{h}(x)} P_m = (C_{h(x)})^{-1}$ .

If  $h(x) = (p_i(x))^{n_i}$ , then  $h(x)$  is symmetric and so

$P_m C_{h(x)} P_m = (C_{h(x)})^{-1}$ . Hence

$$I_m = (P_m C_{h(x)} P_m) C_{h(x)} = (P_m C_{h(x)})^2,$$

that is,  $P_m C_{h(x)}$  is an involution. It follows as above that

$C_{h(x)} = P_m (P_m C_{h(x)})$  is a product of two involutions.

Therefore there exist involutions  $E_i$  and  $F_i$  such that

$$A = \bigoplus_i E_i F_i = \left( \bigoplus_i E_i \right) \left( \bigoplus_i F_i \right).$$

To conclude this part, note that  $\left( \bigoplus_i E_i \right)$  and  $\left( \bigoplus_i F_i \right)$  are also involutions.

Conversely, if  $A = ST$ , where  $S$  and  $T$  are involutions, then

$TAT^{-1} = TS = (ST)^{-1} = A^{-1}$ . That is,  $A \sim A^{-1}$ . ■

## 4.2 PRODUCTS OF FOUR INVOLUTIONS.

In [9] Gustafson, Halmos and Radjavi showed that a matrix  $T$  over an arbitrary field  $F$  is a product of not more than four involutions if, and only if,  $\det(T) = \pm 1$ . The result of Section 4.1, Theorem 4.6 is a special case of this, since if  $A \sim A^{-1}$ , then  $\det(A) = \pm 1$ .

The following lemmas will simplify the proof of the Gustafson, Halmos and Radjavi Theorem.

### Lemma 4.7

Let  $n = n_1 + \dots + n_k + m$ , with  $n_i \geq 2$  and  $m \geq 0$ ,  $P_2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ ,

and let  $B \in M_n(F)$  be the permutation matrix

$$B = \text{diag}[I_{n_1-1} \ P_2 \ I_{n_2-2} \ P_2 \ \dots \ P_2 \ I_{n_k-2} \ P_2 \ P_2 \ \dots \ P_2 \ 1]$$

if  $m$  is even, and without the last diagonal entry, if  $m$  is odd. Then  $B$  is an involution.

### Proof

The result follows from a direct computation of  $B$ . ■

Note      If  $m = 0$  in Lemma 4.7, then  $B$  becomes:

$$B = \text{diag}[I_{n_1-1} \ P_2 \ I_{n_2-2} \ P_2 \ \dots \ P_2 \ I_{n_k-1}].$$

**Definition 4.8**

Let  $S \in M_n(F)$ , then  $S$  is said to be a *weighted permutation matrix* if each row and each column of  $S$  contains exactly one nonzero entry.

Permutation matrices  $S = [s_{ij}]_{i=1}^n_{j=1}^n$  (weighted or not) are in a natural correspondence with the indices of the nonzero entries;  $j$  is mapped onto  $i$  where the nonzero entry is in column  $j$  and row  $i$ . That is,  $s_{ij} \neq 0$ .

For example, consider the matrix  $S_1 \in M_5(F)$ :

$$S_1 = \begin{bmatrix} 0 & 0 & 0 & 0 & s_{15} \\ 0 & 0 & 0 & s_{24} & 0 \\ s_{31} & 0 & 0 & 0 & 0 \\ 0 & 0 & s_{43} & 0 & 0 \\ 0 & s_{52} & 0 & 0 & 0 \end{bmatrix} \quad \text{then} \quad \begin{array}{l} 1 \rightarrow 3 \\ 2 \rightarrow 5 \\ 3 \rightarrow 4 \\ 4 \rightarrow 2 \\ 5 \rightarrow 1 \end{array} .$$

We will refer to a weighted permutation matrix in which the associated permutation is a full cycle, as a *cyclic matrix*. (Note: A full cycle on  $1, 2, \dots, n$  is a cycle of length  $n$ ). The example  $S_1$  above is a cyclic matrix. The associated permutation of  $S_1$  is the full cycle  $(1, 3, 4, 2, 5)$ .

**Note** Without loss of generality, it can always be assumed that the cycle starts with 1.

**Lemma 4.9**

Let  $C \in M_n(F)$  and  $D \in M_m(F)$  be cyclic matrices of order  $m, n \geq 1$  respectively, and let  $A = C \oplus D$ , then  $BA$  is cyclic, where  $B$  is the involution  $\text{diag}[I_{n-1}, P_2, I_{m-1}]$ , where  $P_2$  was defined in Lemma 4.7.

(If  $m$  or/and  $n$  is/are equal to one, then the respective identity matrix/matrices in  $B$  is/are missing.)

**Proof**

Let  $(1, i_2, \dots, i_n)$  be the associated permutation of  $C$ , then there is an integer  $s, 1 \leq s \leq n$ , such that  $i_s = n$ .

For  $m = 1$  it is easily seen that  $BA$  is a cyclic matrix with associated permutation:

$$(1, i_2, \dots, i_{s-1}, n+1, i_s, i_{s+1}, \dots, i_n).$$

For  $m > 1$  let  $(j_1, j_2 = 1, \dots, j_m)$  be the associated permutation of  $D$ , then  $A$  consists of two cycles namely:

$$(1, i_2, \dots, i_n)$$

and

$$(j_1 + n, j_2 + n = 1 + n, j_3 + n, \dots, j_m + n),$$

and  $BA$  is a cyclic matrix with the associated permutation

$$(1, \dots, i_{s-1}, 1+n, j_3+n, \dots, j_m+n, j_1+n, i_s=n, i_{s+1}, \dots, i_n). \blacksquare$$

The next lemma gives a more general result than that of Lemma 4.9, with one restriction. The dimension of each cyclic matrix must not be less than two. The proof is by induction.

**Lemma 4.10**

Let  $R = C_1 \oplus \dots \oplus C_k$  with  $C_i$  a cyclic matrix of size  $n_i$ . If  $n_i \geq 2$  for  $i = 1, \dots, k$ , then  $BR$  is cyclic, where  $B$  is the involution defined in Lemma 4.7 (with  $m = 0$ ).

**Proof**

For  $k = 1$  the result is trivial since  $B = I$ , and for  $k = 2$  the result follows from Lemma 4.9. Suppose  $k > 2$  and the result is true for a matrix that is the direct sum of  $k - 1$  cyclic matrices.

Now

$$BR = B_1(B_2 R)$$

where

$$B_1 = \text{diag}[I_t, P_2, I_{n_k-1}], \quad t = \left( \sum_{s=1}^{k-1} n_s \right) - 1$$

and

$$B_2 = \text{diag}[I_{n_1-1}, P_2, \dots, I_{n_{k-1}-1}, I_{n_k}].$$

$B_2 R = C \oplus C_k$ , and by the hypothesis,  $C$  is cyclic. By

Lemma 4.9,  $B_1(C \oplus C_k) = BR$  is cyclic. ■



The corollary below shows that the restriction on the size of the last cyclic matrix in Lemma 4.10 can be relaxed.

**Corollary 4.11**

Let  $R = C_1 \oplus \dots \oplus C_k$  with  $C_i$  a cyclic matrix of size  $n_i$ . If  $n_i \geq 2$  for  $i = 1, \dots, k-1$  and  $n_k = 1$ , then  $BR$  is cyclic, where  $B$  is the involution defined in Lemma 4.7.

**Proof**

Let  $B_2 = \text{diag}[I_{n_1-1}, P_2, I_{n_2-2}, P_2, \dots, I_{n_{k-1}-1}, 1]$ ,

then

$$B_2 R = C \oplus C_k$$

and by Lemma 4.10  $C$  is cyclic. Let

$$B_1 = \text{diag}[I_t, P_2] \text{ where } t = \sum_{s=1}^{k-1} n_s - 1,$$

then by Lemma 4.9,  $B_1(C \oplus C_k) = B_1 B_2 R = BR$  is cyclic. ■

**Lemma 4.12**

Let  $S \in M_n(F)$  be a cyclic matrix and  $\det(S) = \pm 1$ . Then  $S$  is a product of two involutions.

**Proof**

Let  $(1, i_2, i_3, \dots, i_n)$  be the permutation associated with  $S$  and let  $\{e_i\}_{i=1}^n$  denote the basis that is obtained after

ordering the standard basis of  $F^n$  according to this permutation.

Then:

$$\left. \begin{aligned} S(e_k) &= s_{i_{k+1} i_k} \cdot e_{k+1} \quad \text{for } 1 \leq k \leq n-1 \\ S(e_n) &= s_{1 i_n} \cdot e_1 \end{aligned} \right\} \dots \quad (4.13)$$

For simplicity denote  $s_{i_{k+1} i_k}$  by  $s_k$ , then (4.13) becomes:

$$S(e_k) = s_k \cdot e_{k+1} \quad \text{for } 1 \leq k \leq n-1 \quad \text{and} \quad S(e_n) = s_n \cdot e_1.$$

Let  $\{f_1, \dots, f_n\}$  be the ordered basis of  $F^n$  defined by:

$$f_1 = e_1, \quad f_2 = s_1 e_2, \quad f_3 = s_1 s_2 e_3, \quad \dots, \quad f_n = s_1 \dots s_{n-1} e_n.$$

Then:

$$S(f_1) = S(e_1) = s_1 e_2 = f_2, \quad S(f_2) = s_1 S(e_2) = s_1 s_2 e_3 = f_3, \quad \dots,$$

$$S(f_n) = s_1 \dots s_n e_1 = \varepsilon f_1, \quad \text{where } \varepsilon = s_1 \dots s_n = \pm 1.$$

The last equation follows since  $\det(S) = \pm 1$ .

So  $S = P^{-1} S_1 P$ , where

$$S_1 = \begin{bmatrix} 0 & 0 & 0 & \dots & \dots & \dots & \varepsilon \\ 1 & 0 & 0 & \dots & \dots & \dots & 0 \\ 0 & 1 & 0 & & & & 0 \\ \vdots & & \vdots & & & & \vdots \\ \vdots & & & \ddots & & & \vdots \\ \vdots & & & & \ddots & & \vdots \\ 0 & 0 & 0 & & & 1 & 0 \end{bmatrix}$$

and  $P$  is the transition matrix from the standard ordered basis to  $\{f_i\}$ .

$$S_1 = \begin{bmatrix} 0 & & & & 1 \\ & & & 1 & \\ & & \ddots & & \\ & & & \ddots & \\ & 1 & & & \\ 1 & & & & 0 \end{bmatrix} \begin{bmatrix} 0 & & & & 1 & 0 \\ & & & \ddots & & \\ & & & \ddots & & \\ & & 1 & & & \\ & 1 & & & & \\ 0 & & & & & \varepsilon \end{bmatrix} = C_1 D_1.$$

Both  $C_1$  and  $D_1$  on the right are involutions. The result now follows from the note at the beginning of Section 4.1. ■

**Theorem 4.14 (Gustafson, Halmos and Radjavi [9])**

Let  $T \in M_n(F)$ , then  $T$  is the product of not more than four involutions if, and only if,  $\det(T) = \pm 1$ .

**Proof**

First suppose that  $\det(T) = \pm 1$ .

It is sufficient to prove the theorem for the matrix  $T$  in the form  $T = P_1 \oplus \dots \oplus P_k \oplus Q$ , where  $P_i$  ( $1 \leq i \leq k$ ) is a companion matrix of size  $n_i \geq 2$  and  $Q$  is diagonal :

$$P_i = \begin{bmatrix} 0 & & & & 0 & -p_{i0} \\ & & & & & -p_{i1} \\ & & & & & \vdots \\ & & I_{n_i-1} & & & \vdots \\ & & & & & \vdots \\ & & & & & -p_{in_i-1} \end{bmatrix} \quad \text{and} \quad Q = \begin{bmatrix} q_1 & & & \\ & \ddots & & \\ & & \ddots & \\ & & & q_m \end{bmatrix}.$$

Let  $A = P'_1 \oplus \dots \oplus P'_k \oplus Q'$  where,

$$P'_i = \begin{bmatrix} -1 & 0 & . & . & . & 0 \\ -\frac{p_{i1}}{p_{i0}} & & & & & \\ & & & & & \\ . & & & & & \\ & & I_{n_i-1} & & & \\ . & & & & & \\ & & & & & \\ -\frac{p_{in_i-1}}{p_{i0}} & & & & & \end{bmatrix} \quad \text{and} \quad Q' = \begin{bmatrix} 0 & 1 & & & & \\ 1 & 0 & & & & \\ & & 0 & 1 & & \\ & & 1 & 0 & & \\ & & & & \ddots & \\ & & & & & 0 & 1 \\ & & & & & 1 & 0 \\ & & & & & & 1 \end{bmatrix}$$

(the last row and column of  $Q'$  are missing if  $m$  is even).

A direct calculation shows that  $A$  is an involution and

$$T = AR,$$

where  $R = P'_1 \oplus \dots \oplus P'_k \oplus Q'$ ,

$$P'_i = \begin{bmatrix} 0 & . & . & . & 0 & p_{i0} \\ & & & & 0 & \\ & & & & . & \\ & & I_{n_i-1} & & . & \\ & & & & . & \\ & & & & 0 & \end{bmatrix} \quad \text{and} \quad Q'' = \begin{bmatrix} 0 & q_2 & & & & \\ q_1 & 0 & & & & \\ & & 0 & q_4 & & \\ & & q_3 & 0 & & \\ & & & & \ddots & \\ & & & & & 0 & q_{m-1} \\ & & & & & q_{m-2} & 0 \\ & & & & & & q_m \end{bmatrix}$$

Once again, if  $m$  is even, then the last row and column of  $Q''$  are missing, and the subscripts of the last diagonal block are increased by one.

$$\det(R) = \pm p_{10} \dots p_{k0} q_1 \dots q_m = \pm \det(A) = \pm 1$$

Let  $B$  be the permutation matrix defined in Lemma 4.7 and  $R = BS$ . Then  $S$  can be obtained from  $R$  by the same

permutations. Also  $\det(B) = \pm 1$ , therefore  $\det(S) = \pm 1$  since  $\det(R) = \pm 1$ .  $S = BR$  since  $B$  is an involution, by Lemma 4.7.  $R$  is a direct sum of cyclic matrices of size larger than two with the exception of the last diagonal block which may be of size one. Therefore, by Lemma 4.10 or Corollary 4.11,  $S$  is a cyclic matrix, and so by Lemma 4.12  $S = CD$ , where both  $C$  and  $D$  are involutions. This completes the necessary part of the proof.

The converse is immediate since the determinant of an involution is  $\pm 1$ . ■

The following example is a direct application of the proofs of Theorem 4.14 and the preceding lemmas.

**Example 4.15**

Let

$$T = \begin{bmatrix} 0 & 1 & & & \\ 1 & 2 & & & \\ & & 0 & -0,1 & \\ & & 1 & 0 & \\ & & & & 5 & 0 \\ & & & & 0 & 2 \end{bmatrix}$$

then

$$A = \begin{bmatrix} -1 & 0 & & & & \\ -2 & 1 & & & & \\ & & -1 & 0 & & \\ & & 0 & 1 & & \\ & & & & 0 & 1 \\ & & & & 1 & 0 \end{bmatrix} \quad \text{and} \quad R = \begin{bmatrix} 0 & -1 & & & & \\ 1 & 0 & & & & \\ & & 0 & 0,1 & & \\ & & 1 & 0 & & \\ & & & & 0 & 2 \\ & & & & 5 & 0 \end{bmatrix}.$$

$R = BS$  where

$$B = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad \text{and} \quad S = \begin{bmatrix} 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0,1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 5 & 0 \end{bmatrix}.$$

The cycle associated with  $S$  is  $(1,3,5,6,4,2)$  and so the basis  $\{e_1, \dots, e_6\}$  as defined in Lemma 4.12 is:

$$e_1 = c_1, \quad e_2 = c_3, \quad e_3 = c_5, \quad e_4 = c_6, \quad e_5 = c_4, \quad e_6 = c_2$$

where  $c_i$  is the  $i^{\text{th}}$  column of  $I_6$ .

Then:

$$Se_1 = e_2, \quad Se_2 = e_3, \quad Se_3 = 5e_4, \quad Se_4 = 2e_5, \quad Se_5 = 0,1e_6, \quad Se_6 = -e_1.$$

Let

$$f_1 = e_1, \quad f_2 = e_2, \quad f_3 = e_3, \quad f_4 = 5e_4, \quad f_5 = 10e_5, \quad f_6 = e_6.$$

The transformation matrix from the basis  $\{f_1, \dots, f_6\}$  to the standard basis for  $R^6$  is

$$P = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0,2 \\ 0 & 0 & 0 & 0,1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

$S = P^{-1} S_1 P = (P^{-1} C_1 P)(P^{-1} D_1 P)$ , where  $S_1$ ,  $C_1$  and  $D_1$  were defined in Lemma 4.12

Therefore:

$$C = P^{-1} \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} P = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0,1 & 0 & 0 \\ 0 & 0 & 10 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0,2 \\ 0 & 0 & 0 & 0 & 5 & 0 \end{bmatrix}$$

$$D = P^{-1} \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 \end{bmatrix} P = \begin{bmatrix} 0 & 0 & 0 & 0,1 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0,2 \\ 10 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 5 & 0 & 0 & 0 \end{bmatrix};$$

finally  $T = ABCD$  and  $A$ ,  $B$ ,  $C$  and  $D$  are involutions.  $\square$

In [15, Theorem 5] Sourour gave a simpler proof for Theorem 4.14. However this proof does not work for an arbitrary field, only for a field with at least  $n+2$  elements in the case of  $n \times n$  matrices.



The following corollary gives another factorization of a nonsingular matrix over the real field into diagonalizable matrices, some of which are involutions.

**Corollary 4.16**

Every nonsingular matrix  $T \in M_n(R)$  is the product of three involutions and a diagonalizable matrix. Moreover the diagonalizable matrix has at most two eigenvalues  $\sqrt[n]{|\det(T)|}$  and  $-\sqrt[n]{|\det(T)|}$ .

**Proof**

Let  $\det(T) = \alpha$ . The case where  $\alpha = \pm 1$  was dealt with in Theorem 4.14. If  $\alpha \neq \pm 1$ , then consider the matrix  $\frac{1}{\sqrt[n]{|\alpha|}} T$ .

The determinant of this matrix is  $\pm 1$  and so, by Theorem 4.14, there exist four involutory matrices  $A, B, C$  and  $D$  such that  $\frac{1}{\sqrt[n]{|\alpha|}} T = ABCD$ , therefore

$$T = ABC(\sqrt[n]{|\alpha|} D).$$

Since  $D$  is an involution, by Lemma 4.2 it is diagonalizable and has at most two eigenvalues, namely 1 and  $-1$ . Let  $n_1$  be the algebraic multiplicity of 1 then,

$$\sqrt[n]{|\alpha|} \, D = P^{-1} \begin{bmatrix} \sqrt[n]{|\alpha|} \, I_{n_1} & 0 \\ 0 & -\sqrt[n]{|\alpha|} \, I_{n-n_1} \end{bmatrix} P,$$

for some invertible matrix  $P \in M_n(R)$ . ■

## **CHAPTER 5**

# **PRODUCTS OF POSITIVE-DEFINITE AND POSITIVE-SEMIDEFINITE MATRICES.**

In Chapter 1 it was proved that every square matrix over the field of real numbers can be expressed as a product of two symmetric matrices. It is a well known fact that the spectrum of a symmetric matrix over the field of real numbers is real and the matrix is diagonalizable. Similarly, for a square matrix over the field of complex numbers, it was proved that it is the product of at most four Hermitian matrices, provided that the determinant of the matrix is real. Like a symmetric matrix over the field of real numbers, the spectrum of a Hermitian matrix is real and the matrix is diagonalizable. Thus in both cases it is reasonable to ask under what conditions will these factors have positive (non-negative) eigenvalues. A Hermitian matrix (symmetric matrix over the real numbers) is said to be positive-definite (positive-semidefinite) if the spectrum consists of positive (non-negative) real numbers.

In Section 5.1 we consider the case of factorizing a square matrix over the field of real or complex numbers into positive-definite matrices. In Section 5.2 we consider the similar case for positive-semidefinite factors. In both cases a necessary condition for a square matrix  $A$  to be factorized into positive-(semi)definite matrices, is that  $\det(A)$  must be positive (non-negative).

In what follows,  $F$  will denote the field of real or complex numbers unless stated otherwise. For simplicity we will continue to use the terms Hermitian matrix and unitary matrix, keeping in mind however that for  $F = R$  these concepts are equivalent to symmetric matrix and orthogonal matrix respectively.

## 5.1 PRODUCTS OF POSITIVE-DEFINITE MATRICES.

Most of the results in this section are due to Ballantine [1]; however some of the proofs can be derived more easily using the results of Sourour's main theorem in [15].

Lemma 5.1(a) below is true for all matrices (even rectangular) over the field  $F$ .

**Lemma 5.1**

- (a) For any matrix  $A$  over the field  $F$ ,  $AA^*$  and  $A^*A$  are positive-semidefinite.
- (b) Let  $A \in M_n(F)$ , then  $A^*A$  and  $AA^*$  are positive-definite if, and only if,  $A$  is nonsingular.

**Proof**

- (a) Let  $A \in M_{m \times n}(F)$ , then  $A^*A \in M_n(F)$ . Let  $(\cdot, \cdot)_1$  be the standard inner product over  $F^n$  and  $(\cdot, \cdot)_2$  the standard inner product over  $F^m$  then:

$$(A^*Ax, x)_1 = (Ax, Ax)_2 \geq 0 \quad \forall x \in F^n.$$

The result follows from [13] Theorem 3, page 179.

The proof is similar for  $AA^*$ .

- (b) The proof is immediate from the result of part (a). ■

**Note:**

- (i) If  $H$  is a Hermitian matrix it has the following similarity relation

$$H = UDU^* \quad \dots \quad (5.2)$$

where  $U$  is a unitary matrix (i.e.  $UU^* = I$ ) and  $D$  is a diagonal matrix of real numbers. By definition,  $H$  is positive-definite (positive-semidefinite) if, and only

if, the diagonal entries of  $D$  in equation (5.2) are positive (non-negative).

- (ii) If  $H$  is positive-definite (positive-semidefinite) then so is  $VHV^*$  for any nonsingular matrix  $V$ .

### **Product of two positive-definite matrices.**

Lemma 5.3 below forms part of Theorem 2 of Ballantine's results in [1]. Although his proof is restricted to the field of real numbers, its generalization to the field  $F$  is simple. The proof given below is by Taussky [18]. The actual result of Lemma 5.3 does not give the best classification of matrices over  $F$  that can be factorized into two positive-definite matrices; its usefulness is in the proof of the next theorem.

#### **Lemma 5.3 (Taussky [18])**

Let  $A \in M_n(F)$  (where  $F$  is the field of real or complex numbers) and  $\det(A) > 0$ . Then  $A$  is a product of two positive-definite matrices if, and only if,  $A$  is similar to a positive-definite matrix.

**Proof**

Suppose  $A = P_1 P_2$  where  $P_1$  and  $P_2$  are both positive-definite matrices. Then  $P_i = U_i D_i U_i^*$ ,  $i = 1, 2$ , where  $U_i$  is unitary and  $D_i$  is a diagonal matrix with positive real numbers in the main diagonal. Now  $A = P_1^{\frac{1}{2}} (P_1^{\frac{1}{2}} P_2 P_1^{\frac{1}{2}}) P_1^{-\frac{1}{2}}$  where  $P_1^{\frac{1}{2}} = U_1 D_o U_1^*$ ,  $D_o^2 = D_1$  and:

$$\begin{aligned} P_1^{\frac{1}{2}} P_2 P_1^{\frac{1}{2}} &= (U_1 D_o U_1^*) P_2 (U_1 D_o U_1^*) \\ &= (U_1 D_o U_1^*) P_2 (U_1 D_o U_1^*)^* . \end{aligned}$$

Therefore by note (ii) above,  $P_1^{\frac{1}{2}} P_2 P_1^{\frac{1}{2}}$  is positive-definite and so  $A$  is similar to a positive-definite matrix.

Conversely, suppose that  $A$  is similar to  $H$  where  $H$  is a positive-definite matrix. Then there exists a unitary matrix  $U$ , such that  $H = U^* D U$ , where  $D$  is a diagonal with positive entries. Therefore:

$$\begin{aligned} A &= R^{-1} U^* D U R \\ &= (R^{-1} R^{*-1}) R^* U^* D U R \\ &= (R^{-1} (R^{-1})^*) ((U R)^* D (U R)) \end{aligned}$$

By Lemma 5.1(b) the first factor on the right is positive-definite. By the note following Lemma 5.1, the second is positive-definite, since  $A$  is nonsingular. ■



Theorem 5.4 below was proved by Ballantine in [1]. In that paper he restricted the results to a square matrix over the field of real numbers. The proof below shows a similar result holds for a matrix over  $F$ .

**Theorem 5.4 (Ballantine [1])**

$A$  is a product of two positive-definite matrices if, and only if,  $A$  is unitarily similar to a diagonalizable lower triangular matrix of positive diagonal.

**Proof**

Suppose  $A$  is a product of two positive-definite matrices, then by Lemma 5.3  $A = RPR^{-1}$ , where  $P$  is positive-definite. Hence the spectrum of  $A$  is full and consists of positive real numbers. Now by [7] Theorem 6.21 page 363,  $A^* = UT^*U^*$ , where  $U$  is unitary and  $T^*$  is upper triangular, therefore  $A = UTU^*$  (note that this result also holds over  $F = R$  since the characteristic polynomial of  $A$  splits). The diagonal elements of  $T$  are its eigenvalues which are in turn the eigenvalues of  $A$ , which are positive real numbers. Thus  $T$  is of the required form.

Conversely, suppose that  $A = UTU^*$ , where  $U$  is unitary and  $T$  is a diagonalizable lower triangular matrix with

positive entries in the main diagonal. Since  $T$  is diagonalizable,  $T \sim D$  where  $D$  is a diagonal matrix. Since the spectrum of  $T$  is the diagonal entries of  $T$ ,  $D$  is a positive diagonal matrix and hence is positive-definite. But  $A \sim T$  therefore  $A \sim D$  and the result now follows from Lemma 5.3. ■

### Product of three positive-definite matrices.

The lemma below is based on a proof by Ballantine in [1]. As in the previous section, Ballantine showed the results are true for a square matrix over the real numbers. In what follows, similar results are proved for a square matrix over  $F$ .

The following facts are used in the proof below:

- (i) The property of being the product of two positive-definite matrices is invariant under similarity. This is easily seen since if  $A = P_1 P_2$  where  $P_i$  is positive-definite  $i = 1, 2$  then,

$$R A R^{-1} = (R P_1 R^*) ((R^{-1})^* P_2 R^{-1})$$

for any nonsingular matrix  $R \in M_n(F)$ .

- (ii)  $P \in M_n(F)$  is positive-definite if, and only if,  $P = BB^*$  for a nonsingular matrix  $B \in M_n(F)$ . See, for example, [13], Colollary 1, page 185.

**Lemma 5.5** (Ballantine. [1])

Let  $A$  be a nonsingular matrix in  $M_n(F)$ , then the following statements are equivalent:

- (a)  $A$  is a product of three positive-definite matrices.
- (b)  $A$  is congruent to a matrix that is a product of three positive-definite matrices.
- (c)  $A$  is congruent to a matrix that is a product of two positive-definite matrices.

**Proof**

All matrices in this proof are nonsingular and  $P_i$  is positive-definite ( $i = 1, 2, \dots$ ).

(a)  $\Rightarrow$  (b) Suppose that  $A = P_1 P_2 P_3$  and let  $U$  be an unitary matrix. Then  $A = U(U^* P_1 U)(U^* P_2 U)(U^* P_3 U)U^*$ .

(b)  $\Rightarrow$  (c) Suppose  $A = C P_1 P_2 P_3 C^*$  and  $P_1 = BB^*$ . Then

$$A = C(BB^*)P_2 P_3 ((B^*)^{-1} B^*)C^* = (CB)(B^* P_2 P_3 (B^*)^{-1})(CB)^*.$$

(c)  $\Rightarrow$  (a)  $A = C P_1 P_2 C^* = C P_1^{1/2} P_1^{1/2} P_2 C^* =$

$$(C P_1^{1/2} C^*)((C^*)^{-1} P_1^{1/2} C^{-1})(C P_2 C^*)$$

■

**Theorem 5.6 (Ballantine [1])**

Let  $A \in M_n(F)$ , then  $A$  is a product of three positive-definite matrices if, and only if,  $A$  is congruent to a lower triangular matrix with positive diagonal.

**Proof**

If  $A$  is a product of three positive-definite matrices then, by Lemma 5.5 (c),  $A = RP_1P_2R^*$  where  $P_1$  and  $P_2$  are positive-definite and  $R$  is nonsingular. By Theorem 5.4,  $P_1P_2 = UTU^*$ , where  $U$  is unitary and  $T$  is lower triangular with positive diagonal. The result now follows since unitarily congruent and unitarily similar are the same.

Conversely, if  $A = RTR^*$ , where  $R$  is nonsingular and  $T$  is a lower triangular matrix with positive diagonal, then  $DT$  is similar to a diagonal matrix with positive diagonal, for suitably chosen diagonal matrix  $D$ . By Lemma 5.3  $DT = P_1P_2$ , where  $P_1$  and  $P_2$  are positive-definite. The result follows from Lemma 5.5 (b) with  $A = D^{-1}P_1P_2$ , where  $D^{-1}$  is a diagonal matrix with positive diagonal. ■

In [1] Ballantine showed that the following two statements are also equivalent to that of a matrix  $A \in M_n(R)$  being a product of three positive-definite matrices:

- (i)  $A$  is congruent to a matrix all of whose leading principal minors are positive.
- (ii)  $A$  is positive-definite itself, or its symmetric part  $\frac{1}{2}(S + S^T)$  is not nonpositive-definite.

**Product of four positive-definite matrices.**

The conditions under which a matrix over the field of real numbers is a product of four or five positive-definite matrices was also proved by Ballantine in [1]. He proved similar results for a matrix over the complex numbers. The proofs below are a result of Sourour's main theorem [15]. The matrix is not restricted to one over the real numbers.

**Theorem 5.7** (Sourour [15])

Let  $A \in M_n(F)$ , then  $A$  is a product of four positive-definite matrices if, and only if,  $\det(A) > 0$  and  $A$  is not a scalar matrix  $\alpha I$ , unless  $\alpha > 0$ .

**Proof**

Assume  $\det(A) > 0$ . The case  $A = \alpha I$ ,  $\alpha > 0$ , is trivial. So suppose that  $A$  is not scalar, then by Theorem 1.14 there exist distinct, positive real numbers  $\beta_1, \gamma_1, \dots, \beta_n, \gamma_n$ ,

such that  $\det(A) = \prod_{i=1}^n \beta_i \gamma_i$  and  $A = BC$  where  $\sigma(B) = \{\beta_1, \dots, \beta_n\}$

and  $\sigma(C) = \{\gamma_1, \dots, \gamma_n\}$ . Now  $B = R^{-1}DR$  where  $D$  is a diagonal matrix consisting of the positive numbers  $\beta_1, \dots, \beta_n$ .

Therefore  $B = (R^{-1}R^{*-1})(R^*DR)$  and both factors on the right are positive-definite. Similarly  $C$  is a product of two positive-definite matrices. This completes the proof of the necessary condition.

Conversely, suppose that  $A = P_1 P_2 P_3 P_4$ , where  $P_1, P_2, P_3$  and  $P_4$  are positive-definite, then clearly  $\det(A) > 0$ . Furthermore, if  $A = \alpha I$  then  $P_1 P_2 = \alpha P_4^{-1} P_3^{-1}$ ,  $P_1 P_2$  and  $P_4^{-1} P_3^{-1}$  are similar to positive-definite matrices and hence have positive spectrum. Therefore  $\alpha > 0$ . ■

### Product of five positive-definite matrices.

#### Theorem 5.8

$A$  is a product of five positive-definite matrices if, and only if,  $\det(A) > 0$ .

#### Proof

If  $A$  is nonscalar the result follows from Theorem 5.7, thus we need only consider the case where  $A = \alpha I$ . Let  $P$  be a



nonscalar positive-definite matrix, then  $A = \alpha P^{-1} P$  and  $\det(\alpha P^{-1}) = \alpha^n \det(P^{-1}) = \det(A) \det(P^{-1}) > 0$  since  $\det(A) > 0$ . Therefore by Theorem 5.7,  $\alpha P^{-1}$  can be expressed as the product of four positive-definite matrices and the result follows. ■

## 5.2 PRODUCTS OF POSITIVE-SEMIDEFINITE MATRICES.

In [21] Wu treated the case of factorizing a matrix over the field of complex numbers into positive-semidefinite matrices. Clearly this case is only of interest if the matrix is singular, otherwise the results of Section 5.1 may be applied. Wu's results for a square matrix over the field of complex numbers into three positive-semidefinite matrices, is inconclusive and is omitted here. As in Section 5.1, some of the results proved by Wu are more easily obtained using the main theorem of Sourour and Tang in [17]. It was shown in this paper that a singular matrix can be factorized into two matrices with prescribed eigenvalues. The only restriction is that the number of zero eigenvalues in the two factors must not be less than the nullity of the original matrix.



**Product of two positive-semidefinite matrices.**

In what follows we use the terms positive-semidefinite and non-negative interchangeably.

**Lemma 5.9**

The only nilpotent matrix in  $M_n(F)$  that is the product of two non-negative matrices is the zero matrix.

**Proof**

Let  $T = AB$  where  $A, B \geq 0$  and  $T$  is nilpotent, therefore:

$$\begin{aligned}\sigma(B^{\frac{1}{2}}(AB^{\frac{1}{2}})) &= \sigma((AB^{\frac{1}{2}})B^{\frac{1}{2}}) && \text{by [13] Exercise 10 page 165} \\ &= \sigma(T) \\ &= \{0\}\end{aligned}$$

But  $B^{\frac{1}{2}}AB^{\frac{1}{2}}$  is Hermitian, so it is diagonalizable with spectrum  $\{0\}$ ; this implies that  $B^{\frac{1}{2}}AB^{\frac{1}{2}} = 0$ .

Since

$$B^{\frac{1}{2}}AB^{\frac{1}{2}} = (A^{\frac{1}{2}}B^{\frac{1}{2}})^*(A^{\frac{1}{2}}B^{\frac{1}{2}}) = 0,$$

we have that

$$A^{\frac{1}{2}}B^{\frac{1}{2}} = 0;$$

hence

$$T = A^{\frac{1}{2}}(A^{\frac{1}{2}}B^{\frac{1}{2}})B^{\frac{1}{2}} = 0. \quad \blacksquare$$

**Theorem 5.10**

Let  $T \in M_n(F)$ . Then the following statements are equivalent:

- (a)  $T$  is the product of two non-negative matrices.
- (b)  $T$  is the product of a positive-definite matrix and a non-negative matrix.
- (c)  $T$  is similar to a non-negative matrix.

**Proof**

(b)  $\Rightarrow$  (c). Let  $T = AB$ , where  $A > 0$  and  $B \geq 0$ . Then since  $A$  is nonsingular so is  $A^{1/2}$ , therefore  $T \sim A^{-1/2} T A^{1/2} = A^{1/2} B A^{1/2}$  which is non-negative.

(c)  $\Rightarrow$  (b). Suppose  $T = X^{-1} C X$ , where  $X$  is nonsingular and  $C$  is non-negative, then  $T = (X^{-1} (X^{-1})^*) (X^* C X)$ . The first factor on the right is positive-definite and the second is non-negative.

(b)  $\Rightarrow$  (a). This is trivial.

(a)  $\Rightarrow$  (c). Let  $T = AB$ , where  $A, B \geq 0$ . Since the property of being the product of two non-negative matrices is invariant under similarity, we may (by Lemma A.6) assume

$$T = \begin{bmatrix} N & 0 \\ 0 & K \end{bmatrix}$$

on the space  $F^n = H_1 \oplus H_2$ , where  $N$  is nilpotent, and  $K$  is nonsingular.

Let

$$A = \begin{bmatrix} A_1 & X \\ X^* & A_2 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} B_1 & Y \\ Y^* & B_2 \end{bmatrix},$$

where all the blocks are conformable with  $N \oplus K$  and  $A_1$ ,  $A_2$ ,  $B_1$  and  $B_2$  are non-negative. Since  $A$  and  $B$  are Hermitian in particular, it follows exactly as in the proof of Theorem 1.6 that  $X = 0$  and  $A_2$ ,  $B_2$  are nonsingular.

Thus

$$T = \begin{bmatrix} N & 0 \\ 0 & K \end{bmatrix} = \begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix} \begin{bmatrix} B_1 & Y \\ Y^* & B_2 \end{bmatrix},$$

so that  $A_2 Y^* = 0$ ; hence  $Y = 0$  also.

Therefore

$$T = \begin{bmatrix} N & 0 \\ 0 & K \end{bmatrix} = \begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix} \begin{bmatrix} B_1 & 0 \\ 0 & B_2 \end{bmatrix} \dots \quad 5.11$$

which implies that  $N = A_1 B_1$  and  $K = A_2 B_2$ .

Since  $N$  is nilpotent and  $A_1$  and  $B_1$  are non-negative, by Lemma 5.9,  $N = 0$ .

Since  $K$  is nonsingular  $A_2$  and  $B_2$  are both positive, thus  $K$  is similar to a non-negative matrix ((b)  $\Rightarrow$  (c)). That is,  $K = X^{-1} C X$  where  $C \geq 0$ . Equation 5.11 now becomes

$$T = \begin{bmatrix} 0 & 0 \\ 0 & K \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & X^{-1} C X \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & X^{-1} \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & C \end{bmatrix} \begin{bmatrix} I & 0 \\ 0 & X \end{bmatrix}. \quad \blacksquare$$

### Product of four positive-semidefinite matrices.

#### Theorem 5.12 (Sourour and Tang [17])

Let  $A \in M_n(F)$  be a singular matrix, then  $A$  is a product of four positive-semidefinite matrices, three of which may be taken to be positive-definite.

#### Proof

First suppose that  $\text{nullity}(A) = 1$ . Then by Theorem 1.20  $A = BC$ , where the eigenvalues of  $B$  are distinct and positive, and those of  $C$  are distinct, one of which is zero and the rest are positive. Therefore  $B = T^{-1}PT$ , where  $P = \text{diag}[p_1, \dots, p_n]$  ( $p_i > 0, i = 1, \dots, n$ ) and  $C = R^{-1}SR$ , where  $S = \text{diag}[s_1, \dots, s_n]$  ( $s_1 = 0$  and  $s_i > 0, i = 2, \dots, n$ ).

The equation

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

gives  $A$  as a product of three positive-definite factors, and one positive-semidefinite factor.

If  $\text{nullity}(A) = k, k \geq 1$ , then using standard canonical forms,

$$A \sim A_1 \oplus \dots \oplus A_k,$$

where  $\text{nullity}(A_i) = 1$  for  $1 \leq i \leq k$ . The first part of the proof may now be applied to each  $A_i$ , from which the result will follow. That is,

$$A_i = P_{i1} P_{i2} P_{i3} P_{i4}, \text{ where } P_{i1}, P_{i2}, P_{i3} > 0 \text{ and } P_{i4} \geq 0,$$

and

$$\begin{aligned} A &\sim (P_{11} P_{12} P_{13} P_{14}) \oplus \dots \oplus (P_{k1} P_{k2} P_{k3} P_{k4}) \\ &= (P_{11} \dots P_{k1}) \oplus (P_{12} \dots P_{k2}) \oplus (P_{13} \dots P_{k3}) \oplus (P_{14} \dots P_{k4}). \quad \blacksquare \end{aligned}$$

### 5.3 CONCLUDING REMARKS.

So far we have shown that for a square matrix  $A$  over the field of real or complex numbers:

If  $\det(A) \in R$ :

Then  $A$  is a product of four Hermitian matrices.

If  $\det(A) > 0$ :

If  $A$  is not a scalar matrix  $\alpha I$  (unless  $\alpha > 0$ ) then  $A$  can be expressed as the product of positive-definite matrices, four being the maximum number of factors

required. Otherwise five positive-definite factors are required.

If  $\det(A) = 0$ :

Then  $A$  can be expressed as a product of positive-semidefinite matrices, with the maximum number of such factors required being four.

For  $\det(A) < 0$  and  $\det(A) \notin R$  we offer the following results:

### **Corollary 5.13**

Let  $A = M_n(F)$  and  $\det(A) < 0$  then,  $A = P \left( \prod_{i=1}^t P_i \right) = \left( \prod_{i=1}^t P_i \right) P$

where  $P_i$  are positive-definite,  $t$  is at most 5 and  $P$  is an involution that exchanges two rows or columns of  $I_n$ .

### **Proof**

Let  $A' = PA$ , where  $P$  is the involution described above. Then  $\det(A') > 0$  and so, by Theorems 5.7 or 5.8, it follows that

$$A' = \prod_{i=1}^t P_i,$$

where  $P_i$  are positive-definite. The result follows since  $P$  is an involution. ■

**Corollary 5.14**

If  $A \in M_n(C)$  and  $\det(A) = \alpha \notin R$ , then  $A = \frac{1}{z_o} \prod_{i=1}^t P_i$ , where  $P_i$  are positive-definite matrices,  $t$  is at most five and  $z_o$  is a root of the equation  $z^n - \bar{\alpha} = 0$ .

**Proof**

$\det(z_o A) = z_o^n \alpha = \bar{\alpha} \alpha > 0$ . The result follows from Theorem 5.7 or 5.8. ■



## APPENDIX

The following well known lemmas show that certain properties of products of matrices are invariant under similarity. They are used frequently in this dissertation and enable us to restrict our study to simpler canonical forms of a matrix in order to obtain the required factorizations.

### Lemma A.1

The property of being the product of  $k$  diagonalizable matrices over a field  $F$  is invariant under similarity over  $F$ .

### Proof

Suppose that  $A = D_1 D_2 \dots D_k$ , where  $D_i$  is a diagonalizable matrix ( $1 \leq i \leq k$ ) and  $A \sim B$ . Then  $B = P^{-1} A P$  for some invertible matrix  $P$ .

$$B = P^{-1} D_1 (P P^{-1}) D_2 (P P^{-1}) \dots (P P^{-1}) D_k P = (P^{-1} D_1 P) (P^{-1} D_2 P) \dots (P^{-1} D_k P)$$

and diagonalizability of the factors is preserved by similarity. ■

**Lemma A.2**

The property of being the product of symmetric matrices over a field  $F$  is invariant under similarity over  $F$ .

**Proof**

Suppose that  $A = S_1 S_2 \dots S_k$ , where  $S_i$  is a symmetric matrix ( $1 \leq i \leq k$ ) and  $A \sim B$ . Then  $B = P^{-1} A P$  for some invertible matrix  $P$ .

If  $k$  is even:

$$\begin{aligned} B &= P^{-1} S_1 ((P^T)^{-1} P^T) S_2 (P P^{-1}) \dots ((P^T)^{-1} P^T) S_k P \\ &= (P^{-1} S_1 (P^T)^{-1}) (P^T S_2 P) \dots (P^T S_k P) \end{aligned}$$

Each factor on the right is symmetric.

If  $k$  is odd:

$$\begin{aligned} B &= P^{-1} S_1 ((P^T)^{-1} P^T) S_2 (P P^{-1}) \dots (P P^{-1}) S_k ((P^T)^{-1} P^T) P \\ &= (P^{-1} S_1 (P^T)^{-1}) (P^T S_2 P) \dots (P^{-1} S_k (P^T)^{-1}) (P^T P) \end{aligned}$$

Again each factor on the right is symmetric. ■

**Note** The number of symmetric factors of a matrix similar to  $A$  in Lemma A.2 will be the same as the number of symmetric matrices that make up  $A$ , if this number is even. It will be one more if  $A$  is a product of an odd number of symmetric matrices.

**Lemma A.3**

The property of being the product of Hermitian matrices over a field  $F$  is invariant under similarity over  $F$ .

**Proof**

The proof is similar to that of Lemma A.2. Insert the appropriate products of  $P^*$ ,  $P^{*-1}$ ,  $P$  and  $P^{-1}$  between the factors of  $A = H_1 H_2 \dots H_k$ , where  $H_i$  is Hermitian ( $1 \leq i \leq k$ ). ■

**Lemma A.4**

The property of being the product of idempotent matrices over a field  $F$  is invariant under similarity over  $F$ .

**Proof**

The proof is similar to that of Lemma A.1 ■

**Lemma A.5**

The property of being the product of nilpotent matrices over a field  $F$  is invariant under similarity over  $F$ .

**Proof**

Let  $A = N_1 \dots N_k$ , where  $N_i$  ( $1 \leq i \leq k$ ) are nilpotent matrices. Let the degree of nilpotency of each matrix  $N_i$  be

$m_i$ . If  $A_1 \sim A$  then  $A_1 = P^{-1}AP = (P^{-1}N_1P) \cdot (P^{-1}N_2P) \dots (P^{-1}N_kP)$   
 and  $(P^{-1}N_iP)^{m_i} = N_i^{m_i} = 0$ . ■

We also refer to the following well known result whose proof follows from the Rational Canonical Form Theorem.

**Lemma A.6**

Every matrix  $A \in M_n(F)$  is similar to a matrix of the form

$\begin{bmatrix} N & 0 \\ 0 & K \end{bmatrix}$ , where  $N$  is nilpotent and  $K$  is nonsingular.

**Proof**

$A \sim \text{diag}[C_1, \dots, C_k]$ , where each  $C_i$  is the companion matrix associated with an elementary divisor of  $A$ . Let  $C_1, \dots, C_p$  denote the companion matrices associated with polynomials which are powers of  $x$ . Then  $N = \text{diag}[C_1, \dots, C_p]$  and  $K = \text{diag}[C_{p+1}, \dots, C_k]$  and the result follows. ■

Index of Theorems, Lemmas etc.*Chapter 1*Section 1.1

Corollary 1.1 4

Section 1.4

Lemma 1.2 9

Theorem 1.3 10

Example 1.4 11

Section 1.5

Lemma 1.5 13

Theorem 1.6 14

Equation 1.7 14

Equation 1.8 15

Equation 1.9 15

Theorem 1.10 17

Example 1.11 21

Corollary 1.12 23

Section 1.6

Lemma 1.13 24

Theorem 1.14 25

Equation 1.15 26

Corollary 1.16 28

Example 1.17 29

Lemma 1.18 32

Theorem 1.19 33

Theorem 1.20 34

Equation 1.21 36

Section 1.7

Theorem 1.22 40

*Chapter 2*Section 2.1

Lemma 2.1 43

Theorem 2.2 47

Equation 2.3 48

Equation 2.4 48

Example 2.5 51

Corollary 2.6 52

Section 2.2

Theorem 2.7 54

Section 2.3

Definition 2.8 57

Lemma 2.9 57

Lemma 2.10 60

Lemma 2.11 62

Equation 2.12 65

Equation 2.13 65

Theorem 2.14 66

Section 2.4

Theorem 2.15 68

Corollary 2.16 68

*Chapter 3*

Definition 3.1 69

Lemma 3.2 69

Lemma 3.3 70

Equation 3.4 71

Equation 3.5 72

Lemma 3.6 72

Theorem 3.7 74

Example 3.8 79

Corollary 3.9 81

*Chapter 4*

Definition 4.1 82

Lemma 4.2 82

Section 4.1

Definition 4.3 83

Lemma 4.4 84

Definition 4.5 85

Theorem 4.6 86

Section 4.2

Lemma 4.7 90

Definition 4.8 91

Lemma 4.9 92

Lemma 4.10 93

Corollary 4.11 94

Lemma 4.12 94

Equation 4.13 95

Theorem 4.14 96

Example 4.15 98

Corollary 4.16 101

*Chapter 5*Section 5.1

Lemma 5.1 105

Equation 5.2 105

Lemma 5.3 106

Theorem 5.4 108

Lemma 5.5 111

Theorem 5.6 112

Theorem 5.7 112

Theorem 5.8 113

Section 5.2

Lemma 5.9 115

Theorem 5.10 116

Equation 5.11 117

Theorem 5.12 118

Section 5.3

Corollary 5.13 120

Corollary 5.14 121

*Appendix*

Lemma A1 122

Lemma A2 123

Lemma A3 124

Lemma A4 124

Lemma A5 124

Lemma A6 125

## **REFERENCES:**

1. C. S. Ballantine, Products of Positive Definite Matrices III, *Journal of Algebra* 10: 174-182(1968).
2. C. S. Ballantine, Products of idempotent matrices, *Linear Algebra and Its Applications* 19: 81-86(1978).
3. J. D. Botha, Products of diagonalizable matrices, *Linear Algebra and Its Applications* 273: 65-82(1998).
4. J. D. Botha, Products of diagonalizable matrices over a field of characteristic two, *Linear Algebra and Its Applications* 286: 37-44(1999).
5. J. A. Erdos, On products of idempotent matrices, *Glasgow Mathematical Journal* 8: 118-122(1967).
6. P. A. Fillmore, On similarity and the diagonal matrix, *American Mathematical Monthly* 76: 167-169(1969).
7. S. H. Friedberg, A. J. Insel and L. E. Spence, *Linear Algebra*. Third Edition.
8. C. F. Gerald and P. O. Wheatly, *Applied Numerical Analysis*. Fifth Edition.
9. W. H. Gustafson, P. R. Halmos, H Radjavi, Products of involutions, *Linear Algebra and Its Applications* 13: 157-162(1976).
10. P. R. Halmos, Bad products of good matrices, *Linear and Multilinear Algebra* 29: 1-20(1991).
11. F. Hoffman and E. C. Paige, Products of two involutions in the general linear group, *Indiana University Mathematical Journal* 20: 1017-1020(1971).
12. T. J. Laffey, Products of matrices, *Generators and Relations in Groups and Geometries*, *NATO ASI Series*, Kluwer, 1991.
13. P. Lancaster and M. Tismenetsky, *The Theory Of Matrices*. Second Edition.
14. H. Radjavi, Products of Hermitian matrices and symmetries, *Proceedings of the American Mathematical Society* 21: 369-372(1969).
15. A. R. Sourour, A factorization theorem for matrices, *Linear and Multilinear Algebra* 19: 141-147(1986).



16. A. R. Sourour, Nilpotent factorization of matrices,  
*Linear and Multilinear Algebra* 31: 303-308(1992).
17. A. R. Sourour and K Tang, Factorization of singular matrices,  
*Proceedings of the American Mathematical Society* 116: 629-634(1992).
18. O. Taussky, Positive-definite matrices and their role in the study of the  
characteristic roots of general matrices,  
*Advances in Mathematics* 2: 175-186(1968).
19. R. C. Thompson, Commutators in the special and general linear groups,  
*Transactions of the American Mathematical Society* 101: 16-33 (1961).
20. P. Y. Wu, Products of nilpotent matrices,  
*Linear Algebra and Its Applications* 96: 227-232(1987).
21. P. Y. Wu, Products of Positive Semidefinite Matrices,  
*Linear Algebra and Its Applications* 111: 53-61(1988).



