1951

Extension of normal theory to general matrices

Robert J. Lambert

Iowa State College

Follow this and additional works at: https://lib.dr.iastate.edu/rtd

Part of the Mathematics Commons

Recommended Citation

Lambert, Robert J., "Extension of normal theory to general matrices " (1951). Retrospective Theses and Dissertations. 12856.

https://lib.dr.iastate.edu/rtd/12856

This Dissertation is brought to you for free and open access by the Iowa State University Capstones, Theses and Dissertations at Iowa State University Digital Repository. It has been accepted for inclusion in Retrospective Theses and Dissertations by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.
EXTENSION OF NORMAL THEORY TO GENERAL MATRICES

by

Robert J. Lambert

A Dissertation Submitted to the Graduate Faculty in Partial Fulfillment of The Requirements for the Degree of DOCTOR OF PHILOSOPHY

Major Subject: Mathematics

Approved:

Signature was redacted for privacy.

In Charge of Major Work

Signature was redacted for privacy.

Head of Major Department

Signature was redacted for privacy.

Dean of Graduate College

Iowa State College

1951
INFORMATION TO USERS

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleed-through, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

UMI Microform DP11918
Copyright 2005 by ProQuest Information and Learning Company. All rights reserved. This microform edition is protected against unauthorized copying under Title 17, United States Code.

ProQuest Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346
# TABLE OF CONTENTS

I. INTRODUCTION ........................................ 1

II. PROPERTIES OF NORMAL MATRICES AND REDUCTION
   OF A MATRIX TO NORMAL FORM .......................... 3

III. THE NORMAL FORM AND p(A) .......................... 6

IV. THE ALGEBRA OF REAL POLYNOMIALS IN A
    REAL MATRIX ........................................ 14
   A. Automorphisms of R(A) .............................. 14
   B. Elements of R(A) Which are Similar to A .......... 15
   C. Star (Prime) Commutativity with a
      Positive Definite Hermitian Matrix ............... 18
   D. The Algebra H(A) ................................. 20
   E. A Basis for R(A) and H(A) ........................ 21
   F. An Idempotent Basis for H(A) ........................ 25
   G. Isomorphisms of R(A) and R(A') .................. 27

V. CONSTRUCTION OF p(A) IF AN IDEMPOTENT BASIS
   FOR H(A) IS KNOWN ................................... 35

VI. CONSTRUCTION OF AN IDEMPOTENT BASIS FOR H(A) .... 37
   A. General Problem of Constructing a Basis .......... 37
   B. A Solution for Rank One Idempotents ............. 38
   C. A General Method for Obtaining
      Idempotents .................................. 39
   D. Properties of the Si's ............................ 44
   E. Numerical Example ................................. 45

VII. BIBLIOGRAPHY ......................................... 48

VIII. ACKNOWLEDGMENT ...................................... 51
I. INTRODUCTION

The problem discussed in this thesis arose from an examination of the methods used to compute the complex characteristic roots of a matrix. It occurred to the author that the relative simplicity of the computation of the characteristic roots in the case of a normal matrix indicates that some suitable extension of the normal matrix theory may apply to general matrices.

By a general matrix is meant an arbitrary real matrix $A$ with distinct non-zero characteristic roots. The theory developed here breaks down for the complex case and the repeated root case and is incomplete for the singular case.

A special reduction of a general matrix to a normal matrix leads to a polynomial $p(A)$ which has properties analogous to those of the conjugate transpose of a normal matrix. The reduction of the general matrix can be made without specific knowledge of the characteristic roots, but the principal aim of this investigation is to study the properties of $p(A)$ and the application of those properties to the characteristic root problem.

The philosophy behind the application of this theory to the characteristic root problem is to carry out the
analysis in the real number field. Clearly, this cannot be done in linear fashion. However, the methods given here bring the non-linear aspects of this particular theory within the scope of practical solution.

It is helpful to know the nature of the roots of $A$, that is, the number of complex roots. This is easily accomplished by the following well known criterion:

The number of distinct real characteristic roots of $A$ is equal to the signature of the matrix

\[
\begin{pmatrix}
   s_1 & s_2 & \ldots & s_{n-1} \\
   s_2 & s_3 & \ldots & s_n \\
   \vdots & \vdots & \ddots & \vdots \\
   s_{n-1} & s_n & s_{n+1} & \ldots & s_{2n-2}
\end{pmatrix}
\]

where $s_1$ is the trace of $A$.

The results from Theorem 2.6 on are apparently original although some of the results of Chapter IV are implicit in the general structure theory of algebras.

In regard to notation, $A'$ is the transpose of $A$, $A^* = A'$ is the conjugate transpose of $A$, and $\vec{x}$ is a vector $(x_1, x_2, x_3, \ldots, x_n)$.

---

II. PROPERTIES OF NORMAL MATRICES AND REDUCTION OF A MATRIX TO NORMAL FORM

The definition and some special properties of normal matrices are included in this chapter. In addition it is proved that a matrix with linear elementary divisors can be reduced to a normal matrix by a positive definite hermitian similarity transformation. Specifically, the purpose of this chapter is twofold: to prove Theorem 2.6, and ultimately to make clear the analogy between $A^*$ when $A$ is normal and the polynomial $p(A)$ of Chapter III when $A$ is general.

Definition 2.1. A matrix $A$ is normal if and only if $AA^* = A^*A$.

Theorem 2.1. A necessary and sufficient condition that $U^*AU = D$, where $U$ is unitary and $D$ is diagonal, is that $A$ be normal.¹

Theorem 2.2. If $A$ is normal, and $U^*AU = D$ is diagonal with $U$ unitary, then

$U(A^*A^*U^* U = \mathcal{Q}(D)$, $U(A^*A^*)^* = \mathcal{J}(D)$,

and
\[ U(AA^*)U^* = U(A^*A)U^* = (\text{mod } D)^2 \]
where \( R(D), I(D), \) and \( (\text{mod } D)^2 \) are respectively the real part of \( D \), the imaginary part of \( D \), and \( DD^* \).\(^1\)

**Theorem 2.3.** If \( A \) is normal, then \( A^k \) is a polynomial in \( A \) determined by the characteristic roots of \( A \), and, conversely, if \( A^k \) is a polynomial in \( A \), then \( A \) is normal.\(^2\)

If \( A \) has distinct roots, the polynomial is unique.

**Theorem 2.4.** If \( A \) is normal and non-singular, then the characteristic roots of \( A^{**-1} \) are \( e^{2i\theta_1}, e^{2i\theta_2}, \ldots, e^{2i\theta_n} \), where \( \theta_i \) is the argument of the \( i \)-th characteristic root of \( A \).

**Proof:** Let \( UAU^* = D \), where \( D \) is diagonal. Then \( UA^{-1}U^* = D^{-1} \). Let \( D = RE \) in which \( R = \text{diag}(|\lambda_1|, |\lambda_2|, \ldots, |\lambda_n|) \) and \( E = \text{diag}(e^{i\theta_1}, e^{i\theta_2}, \ldots, e^{i\theta_n}) \). Then one has
\[ UAA^{**-1}U^* = UAU^*UA^{-1}U^* = DD^{-1} = RER^{-1}E^{**-1} = EE^{**-1}. \]
The matrix \( EE^{**-1} = \text{diag}(e^{2i\theta_1}, e^{2i\theta_2}, \ldots, e^{2i\theta_n}) \), and the proof is thus completed.

**Theorem 2.5.** If \( A \) is non-singular, it can be uniquely expressed as the product \( UH \) of a unitary matrix by a positive definite hermitian matrix. If \( A \) is also normal, \( U \) and \( H \) commute.\(^3\)

---


\(^3\)MacDuffee, op. cit., p. 77.
Corollary 2.5. If $A$ is real non-singular, it can be expressed uniquely as the product $OS$ of an orthogonal matrix by a positive definite symmetric matrix.

Theorem 2.6. If $A$ is any matrix with linear elementary divisors, there exists a positive definite hermitian matrix $H$ such that

$$HAH^{-1} = N,$$

where $N$ is normal, and conversely. $N$ will be called a normal form of $A$.

Proof: $A$ has linear elementary divisors. Therefore there exists a non-singular $Q$ such that

$$QAQ^{-1} = D$$

with $D$ diagonal. By Theorem 2.5, the matrix $Q$ can be written

$$Q = UH$$

where $U$ is unitary and $H$ is positive definite hermitian. Thus

$$UHAH^{-1}U^* = D$$

$$HAH^{-1} = U^*DU = N.$$  

$N$ is normal because $D$ is normal and a unitary transformation of a normal matrix remains normal.
III. THE NORMAL FORM AND p(A)

In Theorem 2.6 there is defined a normal matrix N which is similar to the matrix A provided the latter has linear elementary divisors. The similarity transformation is effected by a positive definite hermitian matrix. In this chapter the polynomial p(A) is derived and its elementary properties are established.

Upon application of the condition of normality to the matrix N, there results

\[ NN^* = N^*N \]

\[ HAH^{-1}H^{-1}A^*H = H^{-1}A^*HHAH^{-1}. \]

If the latter equation is multiplied left and right by H and \( H^{-1} \) respectively and if \( H^2 = J \), the equation becomes

\[ JAJ^{-1}A^* = A^*JAJ^{-1}. \]

This equation shows that \( JAJ^{-1} \) commutes with \( A^* \). Upon further restricting the matrix \( A^* \) to be non-derogatory, one can conclude that \( JAJ^{-1} \) is a scalar polynomial in \( A^* \).\(^1\)

Denoting this polynomial in \( A^* \) by \( p(A^*) \) one obtains

\[ JAJ^{-1} = p(A^*). \quad (2) \]

---

If A is of order n, \( p(A^*) \) can be assumed to be of degree \((n-1)\) since any higher degree terms can be reduced by the minimum function of A.

**Theorem 3.1.** \( A = p(\mathcal{P}(A)) \).  

*Proof:* From Equation (2),

\[ JAJ^{-1} = p(A^*) \]

and

\[ J^{-1}A^*J = \mathcal{P}(A). \]

Solving the latter equation for \( A^* \), one obtains

\[ A^* = J\mathcal{P}(A)J^{-1} = \mathcal{P}(JAJ^{-1}). \]

Substituting for \( A^* \) in \( JAJ^{-1} = p(A^*) \) gives

\[ JAJ^{-1} = p(J\mathcal{P}(JAJ^{-1})) = Jp(\mathcal{P}(A))J^{-1}. \]

Multiplying left and right by \( J^{-1} \) and \( J \) respectively gives

the required result (3).

In the following text, the matrices \( A, J, p(A), H, N, \ldots \), etc., are as defined above.

**Theorem 3.2.** If \( B \) is any matrix similar to \( A \), then \( p(B^*) \) is similar to \( B \) by a positive definite hermitian transformation where \( JAJ^{-1} = p(A^*) \).

*Proof:* Write \( QAQ^{-1} = B \), then \( A = Q^{-1}BQ \). Substitute in the equation \( JAJ^{-1} = p(A^*) \). Thus

\[ JQ^{-1}BQJ^{-1} = p(Q*B^*Q^{-1}) = Q*p(B^*)Q^{-1}. \]

Multiply left and right by \( Q^{-1} \) and \( Q^* \) respectively and

obtain

\[ Q^{-1}JQ^{-1}BQJ^{-1}Q^* = p(B^*). \]

The matrix \( Q^{-1}JQ^{-1} = J_0 \) is positive definite since \( J \)
is positive definite, and the theorem is proved.
From Theorem 3.2, it is evident that \( p(x) \) is invariantly associated with the matrix \( A \) under similarity. The transforming matrix \( J \), of course, changes but remains positive definite.

**Theorem 3.3.** If \( A \) has distinct characteristic roots, invariantly associated with the matrix \( A \) is a unique polynomial \( p(x) \) of degree \( n-1 \) or less such that

\[
J A J^{-1} = p(A^*) .
\]

**Proof:** Let \( H A H^{-1} = N \) so that \( A = H^{-1} N H \). Then

\[
J H^{-1} N H J^{-1} = p(H N^* H^{-1})
\]

\[
H^{-1} J H^{-1} N H J^{-1} H = p(N^*) .
\]

Since \( J = H^2 \), this last equation becomes

\[
N = p(N^*) ,
\]

where \( N \) is normal. By Theorem 2.3, a normal matrix \( N \) is uniquely expressible as a rational function of \( N^* \).

Of particular interest in this discussion is the case where the matrix \( A \) is general, as defined in Chapter I.

**Theorem 3.4.** If \( A \) is a real matrix with distinct characteristic roots, then the coefficients of \( p(A^*) \) are real.

**Proof:** Let \( Q A Q^{-1} = \Lambda \) where

\[
\Lambda = \begin{pmatrix}
\lambda_1 & 0 \\
0 & \lambda_2 \\
& \ddots \\
& & 0 \\
& & & \lambda_n
\end{pmatrix}
\]

is the diagonal form of \( A \). Because \( \Lambda \) is normal, one can write

\[
\Lambda = p(\Lambda^*) = p(\Lambda) .
\]
If $\lambda_i = \alpha + i\beta$ is a complex characteristic root and $\lambda_j$ is its conjugate, one obtains the scalar equations

$$\lambda_i = p(\lambda_j)$$
$$\lambda_j = p(\lambda_i).$$

The conjugate of the second equation is

$$\lambda_i = \overline{p}(\lambda_j).$$

Therefore, $\overline{p}(\lambda_j) = p(\lambda_j)$ for all $j = 1, 2, 3, \ldots, n$.

Thus,

$$\overline{p}(\Lambda) = p(\Lambda).$$

Now transform back to the matrix $A$ and obtain

$$\overline{p}(QAQ^{-1}) = p(QAQ^{-1})$$

or

$$\overline{p}(A) = p(A). \quad (4)$$

The matrix $A$ is real so that the coefficients of $p(A)$ are real.

It is now evident that the matrix $p(A)$ plays a role analogous to the conjugate transpose of a normal matrix.

**Theorem 3.5.** The characteristic roots of $\frac{A + p(A)}{2}$, $A - p(A)$, $Ap(A)$, and $A[p(A)]^{-1}$ are respectively the real part, the imaginary part, the square of the moduli, and the exponential of twice the argument of the roots of $A$.

**Proof:** $HAH^{-1} = N$ is normal so that

$$\frac{N + N^*}{2} = \frac{HAH^{-1} + H^{-1}A^*H}{2}.$$

If the matrix on the right hand side is multiplied left and right by $H^{-1}$ and $H$ respectively, the characteristic roots are unaltered. One obtains

$$\frac{A + J^{-1}A^*J}{2} = \frac{A + p(A)}{2},$$
but the roots of $\frac{N+N^*}{2}$ are the real parts of the roots of $A$ by Theorem 2.2. The proofs of the other parts of the theorem are similar.

**Theorem 3.6.** If $A$ is real, then $J_0A J_0^{-1} = p(A^*)$, where $J_0$ can be chosen as $\mathcal{R}(J)$ and is positive definite.

**Proof:** Let $J = J_0+iJ_1$, then

$$(J_0+iJ_1)A = p(A^*)(J_0+iJ_1).$$

By equating real parts of this equation, there results

$$J_0A = p(A^*)J_0.$$

To prove $J_0$ is positive definite, consider $\bar{x}J_0x'$, for every real vector $x \neq 0$. $\bar{x}J_0x'$ is real and greater than zero because $J$ is positive definite. Therefore

$$\bar{x}(J_0+iJ_1)x' = \bar{x}J_0x' > 0.$$

The following illustrative example for a $2 \times 2$ matrix makes use of much of the material derived in this chapter.

Let the matrix $A$ be given, and let $\lambda^2-C_1\lambda-C_2 = 0$ be its characteristic equation.

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}.$$  

The polynomial $p(A)$ is of first degree. Let

$$p(A) = \alpha A + \beta I,$$

where $\alpha$ and $\beta$ are real scalars to be determined. From Equation (3), $p[p(A)] = A$. Thus

$$p[p(A)] \equiv \alpha(\alpha A + \beta I) + \beta I = A$$

or

$$(\alpha^2-1)A + (\alpha \beta + \beta)I = 0.$$
A is assumed non-derogatory and hence satisfies a second degree minimum equation so that the coefficients of the above equation must vanish. Therefore

\[ \alpha^2 = 1 \quad \text{and} \quad \beta(\alpha + 1) = 0. \]

Two possible solutions arise from this equation, namely

\[
\begin{align*}
\alpha &= 1 \\
\beta &= 0 \\
&\quad \quad \quad \text{and} \\
\alpha &= -1 \\
\beta &= c_1.
\end{align*}
\]

The values \( \beta = c_1 \) is necessary so that the trace of \( p(A) \) is equal to the trace of \( A \). The cases corresponding to these two solutions are discussed separately.

Case I. If \( \alpha = 1, \beta = 0 \), then \( p(A) = A' \). Therefore

\[ JA = A'J. \]

Now let

\[ J = \begin{pmatrix} 1 & j_2 \\ j_2 & j_1 \end{pmatrix} \]

be the positive definite hermitian matrix. From the equation

\[
\begin{pmatrix} 1 & j_2 \\ j_2 & j_1 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{pmatrix} \begin{pmatrix} 1 & j_2 \\ j_2 & j_1 \end{pmatrix}
\]

one can obtain sufficient linear equations in the unknowns \( j_1 \) and \( j_2 \) to find \( J \). This matrix equation is satisfied if one picks \( j_1 > 0 \) and

\[ j_2 = \frac{a_{21}j_1 - a_{12}}{a_{22} - a_{11}}. \]

One needs only to insure that \( J \) be positive definite with the choice made for \( j_1 \). By direct computation one can verify that \( j_1 \) must be chosen so that
\[
\frac{(a_{11}-a_{22})^2+2a_{12}a_{21}}{2a_{21}} - \frac{1}{4}a_{12}a_{21}\left(\frac{a_{11}-a_{22}}{2a_{21}}\right)^2 < j_1 < \\
\frac{a_{11}^2-2a_{12}a_{21}+a_{22}^2}{2a_{21}} + \frac{1}{4}a_{12}a_{21}\left(\frac{a_{11}-a_{22}}{2a_{21}}\right)^2.
\]

This selection for \( j_1 \) is possible only if \((a_{11}-a_{22})^2 + 4a_{12}a_{21} > 0\) which is precisely the condition that the roots of \( A \) be real.

Case II. If \( a = -1, \beta = c_1, \) then \( p(A') = -A' + c_1 I. \)

Let \( JA = (-A' + c_1 I)J. \) \( (c_1 = a_{11} + a_{22}) \)

Proceeding as in Case I gives

\[
\begin{pmatrix}
1 & j_2 \\
j_2 & j_1
\end{pmatrix}
\begin{pmatrix}
a_{11} & a_{12} \\
a_{21} & a_{22}
\end{pmatrix}
= \begin{pmatrix}
a_{22} & -a_{21} \\
-a_{12} & a_{11}
\end{pmatrix}
\begin{pmatrix}
1 & j_2 \\
j_2 & j_1
\end{pmatrix}
\]

as the matrix equation to be solved for \( J. \) To satisfy this matrix equation, one must have

\( j_1 = -\frac{a_{12}}{a_{21}} \) and \( j_2 = \frac{a_{22} - a_{11}}{2a_{21}} \) where \( a_{21} \neq 0. \)

Therefore the matrix \( J \) is given by

\[
J = \begin{pmatrix}
1 & \frac{a_{22}-a_{11}}{2a_{21}} \\
\frac{a_{22}-a_{11}}{2a_{21}} & -\frac{a_{12}}{a_{21}}
\end{pmatrix}
\]

For \( J \) to be positive definite, the condition

\[-4a_{12}a_{21} > (a_{22}-a_{11})^2\]

must be satisfied. Note that this condition implies that \( a_{21} \neq 0 \) as was required above. This is the precise condition that \( A \) have complex roots.
The results of this example suggest, in the general case, that $p(A) = A$ when $A$ has real roots. If $A$ has any complex roots, then $p(A)$ is of degree $n-1$ or less. Both of these facts can be easily proved by use of the interpolation formula.
IV. THE ALGEBRA OF REAL POLYNOMIALS IN A REAL MATRIX

The class of all real scalar polynomials in a given real matrix constitutes a matric algebra over the field of reals. This algebra will be denoted by \( R(A) \). The degree of the polynomials in \( R(A) \) can be assumed to be less than or equal to \( u-1 \), where \( u \) is the degree of the minimum function of \( A \). It will be assumed throughout this chapter that \( A \) has distinct characteristic roots.

The study of the polynomial \( p(A) \) of the previous chapter leads to the study of a special aspect of \( R(A) \), namely, that \( R(A) \) contains a sub-algebra \( H(A) \) which consists of all matrices in \( R(A) \) with real characteristic roots, so if the matrix \( A \) has real characteristic roots then \( H(A) \subseteq R(A) \). However, if \( A \) has any pairs of complex roots then \( H(A) \) is a proper sub-algebra of \( R(A) \). A detailed analysis is made of possible bases for \( H(A) \). There will also be derived some relations between the algebras \( R(A) \) and \( R(A') \).

A. Automorphisms of \( R(A) \)

Associated with the algebra \( R(A) \) is the group of automorphisms of \( R(A) \). Let \( \sigma[R(A)] \) denote the image of
R(A) under an automorphism \( \sigma \). All these automorphisms of \( R(A) \) can be found by studying the images of \( A \) itself. Assuming that \( A \) has minimum equation \( \phi(A) = 0 \), one observes that

\[
A \leftrightarrow \sigma(A)
\]

\[
\phi(A) \leftrightarrow \sigma[\phi(A)] = \phi[\sigma(A)]
\]

Therefore the mapping is a similarity transformation \( \sigma(A) = SAS^{-1} \) since it preserves the characteristic roots of \( A \). Conversely, any real polynomial in \( A \) which is similar to \( A \) defines an automorphism of \( R(A) \). These facts are summarized in the following theorem:

**Theorem 4.1.** There is a one-to-one correspondence between the automorphisms of \( R(A) \) and the elements of \( R(A) \) which are similar to \( A \).

**B. Elements of \( R(A) \) Which Are Similar to \( A \)**

For consideration of polynomials in \( A \) which are similar to \( A \), there is no loss of generality in replacing \( A \) by \( A = \Lambda T \Lambda^{-1} \), where \( \Lambda \) is diagonal, as the algebra \( R(\Lambda) \) is isomorphic to the algebra \( R(A) \).

Let \( q(\Lambda) \) be an element in \( R(\Lambda) \) which is similar to \( \Lambda \). Then

\[
\Lambda \leftrightarrow q(\Lambda) = EAE^{-1}
\]

is an automorphism of \( R(\Lambda) \). All of the images of \( \Lambda \) under the automorphisms of \( R(\Lambda) \) are matrices \( \Lambda_p \), where \( \Lambda_p \) is derived from \( \Lambda \) by permutations of the diagonal elements and \( \Lambda_p \) is in \( R(\Lambda) \). The following theorem tells explicitly which
permutations these are.

Theorem 4.2. Every automorphism of \( R(\Lambda) \) is determined by a permutation of the real diagonal elements of \( \Lambda \) among themselves, a permutation of complex pairs among themselves, and an interchange of conjugate elements, performed in any possible manner.

Proof: Let \( \Lambda = \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n) \). Any particular polynomial in \( \Lambda \) which is similar to \( \Lambda \) can be found by solving the equations

\[
\sum_{i=0}^{n-1} a_i \lambda_i^j = \lambda_k^j \quad (j = 1, 2, \ldots, n)
\]

for the coefficients \( a_i \) by Cramer's rule. Among the \( n! \) solutions of these systems, only the real solutions are elements of \( R(\Lambda) \). Let \( \lambda_1 \) be a real root. Then no real linear combination of the powers of \( \lambda_1 \) could be equal to a complex root. Since none of the real roots can generate a complex root, the complex roots must generate themselves.

More abstractly, \( R(\Lambda) \) is a direct sum of complex and real fields. Then any automorphism is a correspondence among the complex fields together with a correspondence among the real fields and automorphisms of the complex fields within themselves.

As an illustration, suppose \( \Lambda \) to be fifth order, non-derogatory, with complex roots \( \lambda, \bar{\lambda} \) and real roots \( u, v, w \).
The matrices $q_0, q_1, q_2,$ and $q_3$ are each in $R(\Lambda)$. The polynomial $q_4(\Lambda)$ is not in $R(\Lambda)$ for, if it were, then
\[ a_0 + a_1u + a_2u^2 + a_3u^3 + a_4u^4 = \lambda \]
would possess a real solution for the $a_i$'s. The polynomial $q_1(\Lambda)$ is the $p(\Lambda)$ of the previous chapter. One can see that one similarity which takes $\Lambda$ into $q_1(\Lambda)$ is
\[
\begin{bmatrix}
0 & 1 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 1
\end{bmatrix} \begin{bmatrix}
\lambda & 0 & 0 & 0 & 0 \\
0 & \lambda & 0 & 0 & 0 \\
0 & 0 & u & 0 & 0 \\
0 & 0 & 0 & v & 0 \\
0 & 0 & 0 & 0 & w
\end{bmatrix} \begin{bmatrix}
0 & 1 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 1
\end{bmatrix} = \begin{bmatrix}
\lambda & 0 & 0 & 0 & 0 \\
0 & \lambda & 0 & 0 & 0 \\
0 & 0 & u & 0 & 0 \\
0 & 0 & 0 & v & 0 \\
0 & 0 & 0 & 0 & w
\end{bmatrix}.
\]
Written more briefly, this equation is
\[ E_1 \Lambda E_1^{-1} = q_1(\Lambda) . \]

**Theorem 4.3.** Each $q_1(A)$ in $R(A)$ is given explicitly by
\[ q_1(A) = Q_1 A Q_1^{-1} , \]
where
\[ Q_1 = T^{-1} E_1 T, \ T \Lambda T^{-1} = \Lambda, \text{ and } E_1^{-1} q_1(\Lambda) E_1 = \Lambda. \]
The proof follows by an easy substitution and will be omitted. The matrix Q₁ in general is not real. A real form will be displayed later.

We note that some of these automorphisms are involutory, that is, two applications of these automorphisms returns each element to itself. In particular, \( A \leftrightarrow p(A) \) is clearly such an involutory automorphism. Thus, \( R(A) \) is a simple illustration of an involutory algebra and, because it is commutative, every involutory mapping onto itself is an automorphism.

C. Star (prime) Commutativity With a Positive Definite Hermitian Matrix

If the matrix \( A \) is real, Equation (2) and the results of Theorems 3.1 and 3.6 can be written as follows:

\[
JAJ^{-1} = p(A') \tag{5}
\]
\[
A = p[p(A)] \tag{6}
\]

with the matrix \( J \) real positive definite symmetric.

By taking the transpose of Equation (5) and solving for \( A' \), there results

\[
Jp(A)J^{-1} = A'. \tag{7}
\]

The addition of Equation (5) and (7) gives

\[
J[A+p(A)]J^{-1} = A'+p(A') \tag{8}
\]

while the multiplication of these two gives

\[
JAp(A)J^{-1} = A'p(A') \tag{9}
\]
Equations (8) and (9) suggest the following useful definition:

**Definition 4.1.** If matrices $X$ and $Y$ with complex coefficients are so related that $YX = X^*Y$ (or $YX = X'Y$), then $X$ is said to star commute (prime commute in the real case) with $Y$. It is to be noted that $X$ star commuting with $Y$ does not imply that $Y$ star commutes with $X$.

On the basis of the definition of prime commutativity, it is easily seen from Equations (8) and (9) that the matrices $A+p(A)$ and $A_p(A)$ each prime commute with $J$.

**Theorem 4.4.** Any real matrix $B$ with linear elementary divisors which prime commutes with a positive definite hermitian matrix $C$ has real characteristic roots.

**Proof:** It is given that

$$BC = CB^*.$$ Let $TBT^{-1} = \Lambda$ (diagonal). Then

$$TBT^{-1}CT^* = TCT^*B^*T^*$$

$$\Lambda C_0 = C_0\Lambda^*$$

where $C_0 = TCT^*$ remains positive definite. The last equation may be written

$$\sum_k \lambda_{ik}c_{kj} = \sum_k c_{ik}\bar{\lambda}_{jk},$$

which, because $\lambda_{ik} = 0$ if $i \neq k$, becomes

$$\lambda_{ii}c_{ij} = c_{ij}\bar{\lambda}_{jj}$$

or

$$(\lambda_{ii} - \bar{\lambda}_{jj})c_{ij} = 0 \text{ for all } i \text{ and } j.$$
If \( i = j \), and because \( C_{ii} > 0 \) in a positive definite hermitian matrix, the last equations prove that
\[
\lambda_{ii} = \bar{\lambda}_{ii}.
\]
If the \( \lambda_{ii} \) are distinct, then it follows from
\[
(\lambda_{ii} - \lambda_{jj})C_{ij} = 0
\]
that \( C_{ij} = 0 \) when \( i \neq j \), hence the matrix \( C_0 \) is positive definite, real and diagonal.

As a direct result of Theorem 4.4, the matrices \( A + p(A) \) and \( A + p(A) \) must have real characteristic roots. This fact was shown by other considerations in Chapter III.

D. The Algebra \( H(A) \)

Real linear combinations of the matrices \( A + p(A) \) and \( A + p(A) \), which are of course elements of \( R(A) \), and their powers will result in elements of \( R(A) \) with real characteristic roots. This fact leads to the study of the class of all elements in \( R(A) \) which have real characteristic roots.

**Theorem 4.5.** The set of all elements of \( R(A) \) which prime commute with \( J \) form a sub-algebra \( H(A) \).

Proof: Let \( h_1(A) \) and \( h_2(A) \) be any two elements of \( R(A) \) which prime commute with \( J \). Then the equations
\[
Jh_1(A) = h_1(A')J
\]
\[
Jh_2(A) = h_2(A')J
\]
implies that
\[
J[\alpha h_1(A) + \beta h_2(A)] = [\alpha h_1(A') + \beta h_2(A')]J
\]
\[
J[h_1(A)h_2(A)] = [h_1(A')h_2(A')]J,
\]
where \( \alpha \) and \( \beta \) are real scalars. The set evidently contains the identity element.
The sub-algebra $H(A)$ of the above theorem contains only elements which have real characteristic roots by Theorem 4.4. It will be shown in the next section that every element of $R(A)$ which has real characteristic roots will prime commute with $J$ and, hence, is in $H(A)$.

E. A Basis For $R(A)$ and $H(A)$

The following theorem is useful for further study and application of the algebras $R(A)$ and $H(A)$ and in particular for computing $p(A)$. The special involution $A \leftrightarrow p(A)$ as defined earlier will be used often in the following sections, so for brevity let $p(A) \equiv P$.

**Theorem 4.6.** If the matrix $A$ has characteristic roots with real parts of the complex roots distinct from real roots and each other, and imaginary parts distinct, then $I, (A+P), (A+P)^2, \ldots, (A+P)^{n-k-1}, (A-P), (A-P)^3, \ldots, (A-P)^{2k-1}$ is a basis for $R(A)$, where $k$ is the number of pairs of complex roots.

**Proof:** Consider $A$ and $P$ in diagonal form $\Lambda$ and $\bar{\Lambda}$. 
Then
\[ T(A+P)T^{-\frac{1}{2}}\Lambda\Lambda^{-1} = \begin{pmatrix}
2\alpha_1 & 2\alpha_1 & 2\alpha_2 & \cdots & 0 \\
2\alpha_1 & 2\alpha_2 & \cdots & & \\
\vdots & \vdots & \ddots & \ddots & \vdots \\
0 & \cdots & \cdots & 2\alpha_k & 2\alpha_k \\
& & & 2\alpha_k & 2\alpha_k \lambda_1 \\
& & & \cdots & 2\alpha_k \lambda_2 \\
& & & \cdots & \cdots & \cdots & \cdots & 2\lambda_{n-2k}
\end{pmatrix} \]

where \( a_j \pm i\beta_j \) are complex roots and the \( \lambda_i \)'s are real roots.

This matrix is derogatory with index \( n-k \). Therefore, I, (\( \Lambda+\Lambda \)), (\( \Lambda+\Lambda \))^2, ..., (\( \Lambda+\Lambda \))^{n-k-1} are independent.

Now consider the matrix
\[ T(A-P)T^{-\frac{1}{2}}\Lambda\Lambda^{-1} = \begin{pmatrix}
2i\beta_1 & 2i\beta_1 & 2i\beta_2 & \cdots & 0 \\
2i\beta_1 & 2i\beta_2 & \cdots & & \\
\vdots & \vdots & \ddots & \ddots & \vdots \\
0 & \cdots & \cdots & 2i\beta_k & 2i\beta_k \\
& & & 2i\beta_k & 2i\beta_k \lambda_1 \\
& & & \cdots & 2i\beta_k \lambda_2 \\
& & & \cdots & \cdots & \cdots & \cdots & 2\lambda_{n-2k}
\end{pmatrix} \]

This matrix is derogatory with index \( 2k \). Therefore, I, \( \Lambda-\Lambda \), (\( \Lambda-\Lambda \))^2, ..., (\( \Lambda-\Lambda \))^{2k-1} are independent. The matrices (\( \Lambda-\Lambda \))^{2l-1} with \( l \leq k \) each have pure imaginary characteristic roots so that each of these are independent of the matrices (\( \Lambda+\Lambda \))^j, \( j \leq n-k-1 \), which have real roots.
The total number of independent matrices in these two sets is \( l+(n-k-1)+k = n \). This number is sufficient to generate \( R(A) \), and hence \( R(A) \).

**Corollary 1.** The matrices \( I, (A+P), (A+P)^2, \ldots, (A+P)^{n-k-1} \) are a basis for \( H(A) \).

**Proof:** Let \( h_1(A) \) be an element of \( H(A) \). Then

\[
h_1(A) = a_0 I + a_1 (A+P) + \cdots + a_{n-k-1} (A+P)^{n-k-1} + b_1 (A-P) + b_3 (A-P)^3 + \cdots + b_{2k-1} (A-P)^{2k-1}.
\]

\( h_1(A) \) has real characteristic roots so that \( b_1 = b_3 = \cdots = b_{2k-1} = 0 \).

**Corollary 2.** Let \( K(A) \) be the vector sub-space of \( R(A) \) such that each element \( k_1(A) \) of \( K(A) \) has pure imaginary characteristic roots, then \( (A-P), (A-P)^3, \ldots, (A-P)^{2k-1} \) is a basis for \( K(A) \).

The proof is similar to that of Corollary 1.

By Theorem 4.6, every element of \( R(A) \) which has real characteristic roots can be expressed as a linear combination of the basis elements listed in Corollary 1. This proves the following theorem:

**Theorem 4.7.** For a matrix \( A \) satisfying the conditions of Theorem 4.6, every element of \( R(A) \) which has real characteristic roots is an element of \( H(A) \).

The conditions of Theorem 4.6 and Theorem 4.7 on the matrix \( A \) can be removed if one uses in addition to powers of \( A+P \) also the powers of the matrix \( AP \) for a basis of \( H(A) \) and odd powers of \( A^2-P^2 \) in the basis of \( K(A) \). Hereafter, the matrix \( A \) is assumed to fulfill the conditions of Theorem 4.6.
An easy consequence of the basis Theorem 4.6 is the following characterization of the automorphism $A \leftrightarrow P$:

**Theorem 4.8.** The automorphism $A \leftrightarrow P$ is the only one other than the identity which leaves the elements of $H(A)$ elementwise invariant.

Proof: From Equation (6)

$$A = p[p(A)].$$

Under the automorphism $A \leftrightarrow P$,

$$(A+P)^k \leftrightarrow (P+A)^k,$$

so that the automorphism $A \leftrightarrow P$ leaves the basis elements invariant. On the other hand, if $A \leftrightarrow q(A) \neq P$, then

$$(A+P) \leftrightarrow q(A)+p[q(A)] \neq (P+A).$$

For certain simple algebras the concept of "even" and "odd" elements is found useful.\(^1\) In the algebra $R(A)$, which is generally not simple, an analogous concept presents itself naturally. We will call the elements of $H(A)$ "even" elements and the elements of $K(A)$ "odd" elements. The following theorem is easily proved.

**Theorem 4.9.** Every element $r(A)$ of $R(A)$ can be written uniquely as $r(A) = E+0 = E+QE_1$, where $E, E_1$ are even, $0, Q$ are odd and $Q = A-P$ is therefore a fixed matrix.

Furthermore,

$$E = \frac{r(A)+r[p(A)]}{2},$$

$$0 = \frac{r(A)-r[p(A)]}{2}.$$

F. An Idempotent Basis For \( H(A) \)

It is evident that any idempotent elements of \( R(A) \) are elements of \( H(A) \). The existence of idempotent elements in \( H(A) \) is proved in the next theorem.

**Theorem 4.10.** The algebra \( H(A) \) contains exactly \( n-k \) unique primitive\(^1\) orthogonal idempotent elements.

**Proof:** By Theorem 4.6, the matrices \( I, (A+P), (A+P)^2, \ldots, (A+P)^{n-k-1} \) are a basis for \( H(A) \). The matrix \( A+P \) satisfies a minimum equation of degree \( n-k \). Using the notation of Theorem 4.6, this minimum equation may be written

\[
(x-2\alpha_1)(x-2\alpha_2) \cdots (x-2\alpha_k)(x-2\lambda_1) \cdots (x-2\lambda_{n-k}) = 0.
\]

Let

\[
f_i(x) = \frac{(x-2\alpha_1)(x-2\alpha_2) \cdots (x-2\alpha_k)(x-2\lambda_1) \cdots (x-2\lambda_{n-k})}{(x-2\alpha_i)} \quad (i=1, 2, \ldots, k) \tag{12}
\]

\[
g_i(x) = \frac{(x-2\alpha_1)(x-2\alpha_2) \cdots (x-2\alpha_k)(x-2\lambda_1) \cdots (x-2\lambda_{n-k})}{(x-2\lambda_i)} \quad (i=1, 2, \ldots, n-2k) \tag{13}
\]

then the matrices

\[
F_i = \frac{f_i(A+P)}{\frac{1}{2} \text{tr}[f_i(A+P)]} \quad (i=1, 2, \ldots, k) \tag{14}
\]

are rank two idempotent elements, and the matrices

\[
G_i = \frac{g_i(A+P)}{\text{tr}[g_i(A+P)]} \quad (i=1, 2, \ldots, n-2k) \tag{15}
\]

are rank one idempotent elements. This latter fact is easily verified by inspection of the diagonal form of \( A+P \), namely, Equation (10). The orthogonality properties of these

\(^1\)not the sum of two other idempotents of smaller rank.
idempotent elements is evident because the product of any two distinct matrices of the $F_i$'s and $G_i$'s contains the minimum equation of $A+P$ as a factor.

The $n-k$ matrices $F_i$'s and $G_i$'s, are independent of each other so that they may be used as a basis for $R(A)$.

To show uniqueness of the primitive idempotents, suppose $Q$ is another primitive idempotent which is a polynomial in $A$. Then

\[ Q = \sum_{i=1}^{k} a_i F_i + \sum_{i=1}^{n-2k} b_i G_i. \]

Because $Q$ is assumed idempotent, $Q^2 = Q$ so that

\[ Q = \sum_{i=1}^{k} a_i^2 F_i + \sum_{i=1}^{n-2k} b_i^2 G_i. \]

By multiplying these two equations successively by $F_1, F_2, \ldots, F_k, G_1, G_2, \ldots, G_{n-2k}$ one obtains the equations

- $QF_1 = a_1 F_1 = a_1^2 F_1$
- $QF_2 = a_2 F_2 = a_2^2 F_2$
- $\vdots$
- $QF_k = a_k F_k = a_k^2 F_k$
- $QG_1 = b_1 G_1 = b_1^2 G_1$
- $\vdots$
- $QG_{n-2k} = b_{n-2k} G_{n-2k} = b_{n-2k}^2 G_{n-2k}$.

From these equations, one can deduce that the $a_i$'s and $b_i$'s are either zero or one. Since $Q$ is assumed primitive, only one of the $a_i$'s and $b_i$'s is non-zero. Therefore $Q$ is equal to one of the other idempotents.
G. Isomorphisms of \( R(A) \) and \( R(A') \)

The basic equation \( JAJ^{-1} = p(A') \) defines an isomorphism with \( R(A') \) which is generally distinct from \( WAW^{-1} = A' \) where \( W \) is real. This suggests that \( R(A) \) should be studied together with \( R(A') \).

\[
\begin{align*}
A & \rightarrow JAJ^{-1} = p(A') \rightarrow \cdots \rightarrow p(A') \\
R(A) & \rightarrow \cdots \rightarrow R(A') \\
p(A) & \rightarrow \cdots \rightarrow \cdots \rightarrow p(A') \\
H(A) & \cap H' \\
A+p(A) & \rightarrow \cdots \rightarrow A'+p(A') \\
\end{align*}
\]

Figure 4.1. The algebras \( R(A) \) and \( R(A') \).

If there is a non-trivial intersection of \( R(A) \) and \( R(A') \), it will be of value in constructing a basis for \( H(A) \). The theory of this section is deficient in that it does not exhibit a method for finding a matrix \( B \) similar to \( A \) such that \( R(B) \cap R(B') \) is non-trivial.
To facilitate calculations, the matrix $A$ will be transformed under similarity to the rational canonical form. Let $VAV^{-1} = M$, where

$$M = \begin{pmatrix}
0 & 1 & 0 & \ldots & 0 & 0 & 0 \\
0 & 0 & 1 & \ldots & 0 & 0 & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & \ldots & 0 & 1 & 0 \\
0 & 0 & 0 & \ldots & 0 & 0 & 1 \\
c_n & c_{n-1} & c_{n-2} & \ldots & c_3 & c_2 & c_1
\end{pmatrix}. \quad (16)
$$

The characteristic equation of this matrix is

$$\lambda^n - c_1 \lambda^{n-1} - c_2 \lambda^{n-2} - \ldots - c_{n-1} \lambda - c_n = 0. \quad (16')$$

The powers of $M$ are readily computed. If $R(M)$ and $R(M')$ have any elements in common other than the trivial ones (scalar multiples of the identity) there will be solutions to the $n^2$ linear equations given by $r_1(M) = r_2(M')$.

Written explicitly, this matrix equation is

$$x_1 M^{n-1} + x_2 M^{n-2} \ldots x_{n-1} M + x_n I = y_1 M'^{n-1} + y_2 M'^{n-2} \ldots y_{n-1} M' + y_n I. \quad (17)$$

Solutions other than

$$x_n = y_n, \quad x_1 = x_2 = \ldots = x_{n-1} = y_1 = y_2 = \ldots = y_{n-1} = 0$$

and

$$x_1 = x_2 = \ldots = x_n = y_1 = y_2 = \ldots = y_n = 0$$

are desired. In general no other solutions are possible as will be illustrated with a $4 \times 4$ matrix. Let

$$M = \begin{pmatrix}
0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1 \\
c_4 & c_3 & c_2 & c_1
\end{pmatrix}. \quad (17)$$
Then the higher powers of $M$ are

$$M^2 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ c_4 & c_3 & c_2 & c_1 \\ c_1 c_4 + c_1 c_3 + c_4 c_1 c_2 + c_3 c_1^2 + c_2 \\ \end{pmatrix}$$

$$M^3 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ c_4 & c_3 & c_2 & c_1 \\ c_1 c_4 & c_1 c_3 + c_4 & c_1 c_2 + c_3 & c_1^2 + c_2 \\ c_1 c_4 + c_2 c_4 & c_1 c_3 + c_2 c_3 & c_1 c_2 + c_2 & c_1^3 + 2 c_1 c_2 \\ \end{pmatrix}$$

The left hand side of Equation (17) becomes

$$r_1(M) = x_1 M^2 + x_2 M^3 + x_3 M + x_4 I$$

or explicitly, $r_1(M) =$

$$\begin{pmatrix} x_4 & x_3 & x_2 & x_1 \\ x_1 c_4 & x_1 c_3 + x_4 & x_1 c_2 + x_3 & x_1 c_1 + x_2 \\ x_1 c_1 c_4 & x_1 (c_1 c_3 + c_4) & x_1 (c_1 c_2 + c_3) & x_1 (c_1^2 + c_2) \\ + x_2 c_4 & + x_2 c_3 & + x_2 c_2 + x_4 & + x_2 c_1 + x_3 \\ x_1 (c_1 c_4 + c_2 c_4) & x_1 (c_1 c_3 + c_2 c_3 + c_1 c_4) & x_1 (c_1 c_2 + c_3 + c_3 x_1 c_1 c_2 + c_3) & x_1 (c_1^3 + 2 c_1 c_2 + c_3) \\ + x_2 c_1 c_4 & + x_2 (c_1 c_3 + c_4) & + x_2 (c_1 c_2 + c_3) & + x_2 (c_1^2 + c_2) \\ + x_3 c_4 & + x_3 c_3 & + x_3 c_2 + c_4 & + x_3 c_1 + x_4 \end{pmatrix}$$

the right hand side of Equation (17) is
By equating $r_1(M) = r_2(M')$, there results sixteen homogeneous linear equations. In attempting to solve this system, one forces rather stringent conditions to be met by the non-arbitrary invariants $c_i$ of the matrix $M$. A few cases where solutions will exist are listed below.

In general, however, the only solutions will be multiples of the identity.

The question arises in attempting to solve Equation (17) whether any non-symmetric solutions exist when non-trivial symmetric ones do not exist. The following theorem answers the question whenever $A$ has at least one real root. In the case of all complex roots, the theorem is not true.
Theorem 4.11. If the matrix $A$ has a real root and $H(A) \cap H(A') = \{ \alpha I \}$, then the only elements in $R(A) \cap R(A')$ are $\{ \alpha I \}$.

Proof: Suppose $r_1(A) = r_2(A')$ are elements not in $H(A)$. Then $r_1(A') = r_2(A)$, so that $r_1(A') \in R(A)$. Hence

$$r_1(A) \cdot r_1(A') \in H(A) \quad (20)$$
$$r_1(A) + r_1(A') \in H(A) \quad (21)$$

Since $r_1(A)$ is not the zero matrix, the matrix (20) is a positive multiple of the identity, say

$$r_1(A) \cdot r_1(A') = \alpha I. \quad (22)$$

Similarly matrix (21) becomes

$$r_1(A) + r_1(A') = \beta I \quad (23)$$

where $\beta$ may be zero. Equation (22) implies

$$\frac{1}{\sqrt{\alpha}} r_1(A) \cdot \frac{1}{\sqrt{\alpha}} r_1(A') = I \quad (24)$$

and Equation (23) implies

$$\frac{1}{\sqrt{\alpha}} r_1(A) \cdot \frac{1}{\sqrt{\alpha}} r_1(A') = \frac{\beta}{\sqrt{\alpha}} I. \quad (25)$$

Equation (24) shows that $\frac{1}{\sqrt{\alpha}} r_1(A)$ is an orthogonal matrix.

Multiplying Equation (25) by $\frac{1}{\sqrt{\alpha}} r_1(A)$ and transposing gives

$$\frac{1}{\alpha} [r_1(A)]^2 - \frac{\beta}{\alpha}[r_1(A)] + I = 0. \quad (26)$$

Equation (26) implies that $r_1(A)$ has only two distinct roots, actually a complex pair because $r_1(A)$ was assumed not to be in $H(A)$. 

The matrix $r_1(A)$ must have multiple pairs of complex roots so in particular it is of even degree. Thus one can conclude that if $A$ is odd degree, no such element $r_1(A)$ exists. If $A$ is of even degree and has one real root, it therefore has two real roots. Hence, $r_1(A)$ would have at least two real roots and thus cannot have only multiple complex pairs.

The example, $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$, demonstrates that Theorem 4.11 does not hold if a matrix has only complex roots.

It can be proved that $H(A) \cap H(A')$ is the algebra generated by the symmetric idempotents. It can also be proved that if $A$ has all complex roots and $r_1(A) = r_2(A')$, then $r_1(A) = r_2[p(A)]$.

An illustrative example will now be given which will employ much of the material of this chapter.

Let $A$ be similar to $M$ where

$$M = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ c_3 & c_2 & c_1 \end{pmatrix}.$$ 

The characteristic equation of $M$ is

$$\phi(\lambda) = \lambda^3 - c_1\lambda^2 - c_2\lambda - c_3 = 0. \quad (27)$$

The algebra $H(M)$ for this case can have only two independent basis elements if $M$ is assumed to have a pair of complex roots. Then by Theorem 4.6, the matrix $M_p(M)$ is a linear combination of the basis elements $I$ and $M+p(M)$. Thus

$$M_p(M) = a[M+p(M)] + aI \quad (28)$$
where \(a\) and \(a\) are to be determined. Equation (28), when solved for \(p(M)\) is

\[
p(M) = (M-aI)^{-1}(aM+aI)\]

\[
p(M) = aI+(a^2+a)(M-aI)^{-1},
\]

where \(a\) is not a characteristic root of \(M\). Let \(a^2+a = \beta\) and the above equation becomes

\[
p(M) = aI+\beta(M-aI)^{-1}. \quad (29)
\]

From the characteristic equation of \(M\), one can show that

\[
(M-aI)^{-1} = \frac{1}{\phi(a)}[(M-aI)^2+(3a-c_1)(M-aI)+(3a^2-2ac_1-c_2)I]. \quad (30)
\]

Substituting in Equation (29) gives

\[
p(M) = aI-\frac{\beta}{\phi(a)}[(M-aI)^2+(3a-c_1)(M-aI)+(3a^2-2ac_1-c_2)I] \quad (31)
\]

as the structure of the matrix \(p(M)\).

By equating the traces\(^1\) of both sides of Equation (31) one obtains

\[
\beta = (3a-c_1)\frac{\phi(a)}{\phi'(a)}. \quad (32)
\]

To get another relationship between \(a\) and \(\beta\), one can use the square of Equation (31) and equate traces again. After considerable computation and reduction, the following equation in \(a\) results:

\[
(3a-c_1)^3 \phi(a) = \left[\phi'(a)\right]^3. \quad (33)
\]

Equation (33) when multiplied out and simplified, gives a polynomial in \(a\) with coefficients in terms of the invariants of the matrix \(M\). This polynomial is

\[1 \text{ tr } p(M) = \text{ tr } M.\]
\[(2c_1^3 + 9c_1c_2 + 27c_3) \alpha^3 - (c_1^4 + 3c_1^2c_2 + 27c_1c_3 - 9c_2) \alpha^2 \\
- (c_1^3c_2 - 9c_1^2c_3 + 6c_1c_2^2) \alpha - (c_2^3 + c_1^3c_3) = 0. \quad (34)\]

One needs only to solve for the real root of Equation (34) and hence to find \( \beta \). As a consequence of Equations (32) and (33),
\[ \beta = \frac{3}{\sqrt[3]{\phi(\alpha)^2}}. \quad (35)\]

After finding \( \alpha \) and \( \beta \) as outlined above, the polynomial \( p(M) \) can be constructed from Equation (31). After some simplification this equation becomes
\[ p(M) = -\frac{1}{3\sqrt[3]{\phi(\alpha)}} \left[ M^2 + (\alpha - c_1)M + (\alpha^2 - \alpha c_1 - c_2)I \right] + \alpha I. \quad (36)\]

This direct method of computing \( p(M) \) can be applied to a 4x4 matrix with two pairs of complex roots as the algebra \( H(M) \) would still only have two independent basis elements. The direct method as outlined above becomes quite tedious for higher order cases where the number of basis elements in \( H(M) \) is larger than two. Another method of solution by finding rank one and rank two idempotents will be given later.
V. CONSTRUCTION OF \( p(A) \) IF AN IDEMPOTENT BASIS FOR \( H(A) \) IS KNOWN

In this chapter it will be assumed that an idempotent basis for the algebra \( H(A) \) is known. Let \( h(A) \) be an element of \( H(A) \), then

\[
H(A) = \sum_{i=1}^{k} a_i F_i + \sum_{i=1}^{n-2k} b_i G_i,
\]

where the \( F_i \)'s are rank two idempotents and the \( G_i \)'s are rank one idempotents. In particular,

\[
A + p(A) = \sum_{i=1}^{k} a_i F_i + \sum_{i=1}^{n-2k} b_i G_i,
\]

where the \( a_i \)'s and \( b_i \)'s must be determined. The following theorem will be proved, using the terminology above:

**Theorem 5.1.**

\[
\text{tr}[p(A)F_i] = \text{tr}[AF_i]
\]

\[
\text{tr}[p(A)G_i] = \text{tr}[AG_i]
\]

where \( \text{tr} X \) means trace of \( X \).

**Proof:**

\[
\text{tr}[p(A)F_i] = \text{tr}[J^{-1}A'JF_i] = \text{tr}[A'JF_iJ^{-1}]
\]

\[
= \text{tr}[A'F_i] = \text{tr}[AF_i],
\]

because \( F_i \) is a polynomial in \( A + p(A) \) and \( J[A + p(A)]J^{-1} = A' + p(A') \).

The proof for the idempotents \( G_i \) is similar.

Now, from Equation (37) above, one obtains, upon multiplying by \( F_j \),

\[
AF_j + p(A)F_j = a_j F_j.
\]
Equating the traces of both sides of Equation (39) and using Theorem 5.1. gives

$$\text{tr}[AF_j] + \text{tr}[p(A)F_j] = 2a_j,$$

or

$$a_j = \text{tr}[AF_j]. \quad (40)$$

Now multiply Equation (37) by $G_j$ and equate traces. This gives

$$b_j = 2 \text{tr}[AG_j]. \quad (41)$$

The $a_i$'s and $b_i$'s in Equation (37) are completely determined from Equation (40) and (41). By transposing the matrix $A$, one obtains the matrix $p(A)$.

The results of this chapter demonstrate the ease of obtaining $p(A)$ once the basis for $H(A)$ is known. However, obtaining this basis is not an easy matter if there are many complex roots. The next chapter will discuss the construction of a basis for $H(A)$. 
VI. CONSTRUCTION OF AN IDEMPOTENT BASIS FOR $H(A)$

In this chapter, methods for explicitly computing an idempotent basis for $H(A)$ are derived. A method for finding rank one idempotents is given which is simpler though equivalent to depressing the characteristic equation by its real roots. However, a special system of simultaneous quadratics is constructed which can be used to obtain all idempotents, in particular those of rank two. This system is solvable by Newton's method of iteration. In connection with these simultaneous quadratics, several symmetric matrices are obtained which transform the previously defined matrix $M$ into $M'$, and hence give an unexpected theory of the solutions of the matrix equation $AX =XA'$.

A. General Problem of Constructing a Basis

If one is able to find a polynomial $h(A)$ of $H(A)$ which has the maximum number of distinct roots, this matrix $h(A)$ will be derogatory (if $A$ has any complex roots). By obtaining the greatest common divisor of the characteristic equation of $h(A)$ and its derivative, one can find all of the rank two idempotents.
Of course, finding any non-trivial element of \( H(A) \) would be helpful. For this purpose, one can use the more general statement of the theorem quoted on page two of the introductory chapter as a criterion which insures real roots. The method, however, is very difficult to apply.

B. A Solution for Rank One Idempotents

The matrix \( M \) is used here to facilitate computation. A rank one matrix will be a polynomial in \( M \) of degree \( n-1 \). Let this polynomial be

\[
G = x_1 M^{n-1} + x_2 M^{n-2} + \ldots + x_{n-1} M + x_n I.
\]

Because the matrix \( G \) is rank one, every second order minor must vanish. The top two rows of the matrix \( G \) are

\[
G = \begin{pmatrix}
x_n & x_{n-1} & x_{n-2} & \ldots & x_3 & x_2 & x_1 \\
x_1 c_n & x_1 c_{n-1} + x_n & x_1 c_{n-2} + x_{n-1} & \ldots & x_1 c_3 + x_4 & x_1 c_2 x_3 & x_1 c_1 + x_2 \\
: & : & : & \ldots & : & : & : \\
: & : & : & \ldots & : & : & : \\
: & : & : & \ldots & : & : & : \\
\end{pmatrix},
\]

(42)

By fixing \( x_1 = 1 \), the matrix will still have rank one, but will not in general be idempotent. Starting with the upper right \( 2 \times 2 \) minor in \( G \) one obtains the equations

\[
\begin{align*}
(c_2 + x_3) &= x_2 (x_2 + c_1) \\
x_2 (c_3 + x_4) &= x_3 (x_3 + c_2) \\
x_3 (c_4 + x_5) &= x_4 (x_4 + c_3) \\
&\vdots \\
x_{n-2} (c_{n-1} + x_{n}) &= x_{n-1} (x_{n-1} + c_{n-2}) \\
x_{n-1} c_n &= x_n (x_n + c_{n-1}).
\end{align*}
\]
Substituting the left hand side of the first equation in the right hand side of the second, etc., one obtains
\[ \begin{align*}
C_2 + X_3 &= X_2 (x_2 + c_1) \\
C_3 + X_4 &= X_3 (x_2 + c_1) \\
C_4 + X_5 &= X_4 (x_2 + c_1) \\
& \quad \vdots \\
C_{n-1} + X_n &= X_{n-1} (x_2 + c_1) \\
C_n &= X_n (x_2 + c_1)
\end{align*} \] (43)

Next, successively eliminate \( x_3, x_4, \ldots, x_n \) and obtain a polynomial in \( x_2 \), namely
\[ (x_2 + c_1)^n - c_1 (x_2 + c_1)^{n-1} - c_2 (x_2 + c_1)^{n-2} \ldots - c_{n-1} (x_2 + c_1) - c_n = 0 \] (44)
This is the characteristic polynomial in the variable \( x_2 + c_1 \) from which a real \( x_2 \) can be found. Equations (43) will then successively yield \( x_3, x_4, \ldots, x_n \). These values, together with \( x_1 = 1 \), will explicitly determine \( G \). Then the matrix
\[ G_0 = \frac{G}{\text{tr}G} \] (45)
will be idempotent.

C. A General Method for Obtaining Idempotents

The method which will be derived in this section will first be demonstrated by an application to a general 4x4 matrix. Let \( A \) be similar to \( M \) where
\[ M = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ c_4 & c_3 & c_2 & c_1 \end{pmatrix} \]
The powers of $M$ are given on page 29. Also let

$$F = x_4^2 + x_3^2 + x_2^2 + x_1^2.$$ 

The explicit form of $F$ is

$$F = \begin{bmatrix} x_4^2 + x_1^2 & x_4^2 + x_2^2 & x_4^2 + x_3^2 & x_4^2 + x_4^2 \\ x_4^2 + x_1^2 & x_4^2 + x_2^2 & x_4^2 + x_3^2 & x_4^2 + x_4^2 \\ x_4^2 + x_1^2 & x_4^2 + x_2^2 & x_4^2 + x_3^2 & x_4^2 + x_4^2 \\ x_4^2 + x_1^2 & x_4^2 + x_2^2 & x_4^2 + x_3^2 & x_4^2 + x_4^2 \end{bmatrix}$$

By setting $F^2 = F$ and equating the entries of the first row of this matrix equation, one obtains the four equations

$$(x_1, x_3, x_3, x_4) \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & c_4 \\ 0 & 0 & c_4 & c_4 \\ 0 & c_4 & c_4 & c_4^2 + c_4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = x_1$$

$$(x_1, x_2, x_3, x_4) \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & c_3 \\ 0 & 0 & c_3 & c_3 + c_4 \\ 0 & c_3 & c_3 + c_4 & c_3 + c_4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = x_2$$

$$(x_1, x_2, x_3, x_4) \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & c_2 \\ 1 & 0 & c_2 & c_2 + c_3 \\ 0 & c_2 & c_2 + c_3 & c_2 + c_3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = x_3$$
\[(x_1, x_2, x_3, x_4) \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & c_1 \\ 0 & 1 & c_1 & c_1^2 + c_2 \\ 1 & c_1 & c_1^2 + c_2 & c_1^3 + 2c_1c_2 + c_3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = x_4 \]

For brevity these equations will be designated by
\[\bar{x}' S_1 \bar{x}' = x_i^1, \quad (i = 1, 2, 3, 4). \quad (48')\]

It is to be noted that the matrix \(S_1\) is made up of
the first column of \(I\), the first column of \(M\), the first
column of \(M^2\), and the first column of \(M^3\). The matrix
\(S_2\) is made up of the corresponding second columns, \(S_3\) of
the third columns and \(S_4\) the fourth columns.

The linear term on the right hand side of these four
equations can be removed by the substitution \(x_1 = y_1 + \frac{1}{2}\),
\(x_2 = y_2\), \(x_3 = y_3\), \(x_4 = y_4\). Denote the new vector by
\[\bar{y} = (y_1, y_2, y_3, y_4).\]

The equations (48) become
\[f_1 = \bar{y}' S_1 \bar{y}' = \frac{1}{4} \]
\[f_2 = \bar{y}' S_2 \bar{y}' = 0 \]
\[f_3 = \bar{y}' S_3 \bar{y}' = 0 \]
\[f_4 = \bar{y}' S_4 \bar{y}' = 0. \quad (49)\]

The four quadratic equations (49) can be solved
simultaneously for all of the sets of values of the \(y_i\)'s.
After finding each set of \(y_i\)'s, the corresponding \(x_i\)'s are
found and hence the corresponding idempotent matrix \(F\).
If it is desired to obtain only rank two idempotents from the Equations (49), one need only to require the trace of $F$ to be equal to two. That is

$$\text{tr} F = 4x_1 + c_1x_2 + (c_1^2 + 2c_2)x_3 + (c_1^3 + 3c_1c_2 + 3c_3)x_4 = 2.$$  

Upon changing to the variables $y_i$, one obtains the homogeneous linear equation

$$4y_1 + c_1y_2 + (c_1^2 + 2c_2)y_3 + (c_1^3 + 3c_1c_2 + 3c_3)y_4 = 0. \quad (50)$$

Equation (50) can be used in conjunction with Equations (49) to select only those solutions of (49) which will yield a rank two idempotent.

Newton's method of iteration is suggested to solve the system (49). It will be shown that this iteration always converges in the general case if the initial values of the variables are sufficiently close to the solution.

It is of interest to note that the matrix $M'$ commutes with each of the matrices $S_i$ of Equations (48), that is

$$MS_i = S_iM'.$$

It will be shown that $M'$ commutes with each $S_i$ in the general case.

It will now be proved that the results of this example are true in general.

Let

$$F = \sum_{i=1}^{n} x_iM_i^{i-1} \quad (51)$$

and

$$F_0 = F - \frac{1}{2}I = \sum_{i=1}^{n} y_iM_i^{i-1} \quad (52)$$

Then the function $F^2 - F$ becomes

$$Y = F_0^2 - \frac{1}{4}I. \quad (53)$$
Consider the functions
\[ f_1 = \bar{\bar{y}} S_1 \bar{y}' - \frac{1}{4} \]
\[ f_1 = \bar{\bar{y}} S_1 \bar{y}' \quad (i = 2, 3, \ldots, n) \]
where the \( S_i \) and \( \bar{\bar{y}} \) are defined as before.

The differential of \( Y \) in Equation (53) is
\[ dY = 2F_0 dF_0 = 2dF_0 \cdot F_0 \] (55)
while the differential of \((f_1, f_2, \ldots, f_n)\) from (54) is
\[ d(f_1, f_2, f_3, \ldots, f_n) = 2(dy)(S_1 \bar{y}', S_2 \bar{y}', S_3 \bar{y}', \ldots, S_n \bar{y}'). \] (56)
The matrix \((S_1 \bar{y}', S_2 \bar{y}', \ldots, S_n \bar{y}')\) = \( F_0 \) by (52) and the definition of the \( S_i \). Furthermore, at \( \bar{\bar{y}} = (\frac{1}{2}, 0, 0, 0) \), \( Y = 0 \) and \( f_1 = 0 \), \( (i=1, 2, \ldots, n) \).

The first row of Equation (55) is the first row of
\[ 2\begin{pmatrix} dy_1 & dy_2 & \ldots & dy_n \\ * & * & \ldots & * \end{pmatrix} \cdot F_0 \]
and hence is the same as (56). Therefore the functions (54) are the same as the functions in the top row of (53).

The basic condition for applying the Newton method of iteration\(^1\) to the equations
\[ f_1 = \bar{\bar{y}} S_1 \bar{y}' - \frac{1}{4} = 0 \]
\[ f_1 = \bar{\bar{y}} S_1 \bar{y}' = 0 \quad (i=2, 3, \ldots, n) \]
requires the non-vanishing of the functional determinant, which in this case is \( 2|F_0| \), at the solutions. At the solutions, the functional determinant is equal to \( 2|F-\frac{1}{2} I| \) where

\(^1\)Ostrowski, A. gives an excellent exposition and improvement of the Newton method in Konvergenzdiskussion und Fehlerabschatzung für die Newton'sche Methode bei Gleichungssystemen, Commentarii Mathematici Helvetic. Vol. 9, p. 79-83, 1936-1937.
F is the required idempotent, and since F has characteristic roots which are 1's and 0's, the determinant is non-zero.

D. Properties of the $S_i$'s.

**Theorem 6.1.** $MS_i = S_iM'$. \hspace{1cm} (57)

**Proof:** The equation $MS_i = S_iM'$ is equivalent to the Kronecker product equation

$$(MxM^{-1}) \bar{s}_i' = \bar{s}_i'$$ \hspace{1cm} (58)

where

$$\bar{s}_i' = \begin{pmatrix} \bar{e}_1' \\ \bar{M} \bar{e}_1' \\ \vdots \\ \bar{e}_{-1-i}' \\ \bar{M}^{-1} \bar{e}_1' \end{pmatrix}$$

and

$$MxM^{-1} = \begin{pmatrix} 0 & M^{-1} & 0 & \ldots & 0 \\ 0 & 0 & M^{-1} & \ldots & 0 \\ 0 & 0 & 0 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \\ c_nM^{-1} & c_{n-1}M^{-1} & c_{n-2}M^{-1} & \ldots & c_1M^{-1} \end{pmatrix}$$

Direct multiplication verifies the theorem.

The number of independent solutions of (58) is exactly $n$ because $MxM^{-1}$ has exactly $n$ characteristic roots equal to one. Therefore the matrices $S_1, S_2, \ldots, S_n$, which are obviously linearly independent, are a basis for all solutions of (57). In fact the following theorem holds:
Theorem 6.2. Let $A = QMQ^{-1}$, $Q$ real, and let $T_i = QS_iQ'$. Then the real solutions of the matrix equation

$$AX = XA'$$

are the linear set $T_1, T_2, \ldots, T_n$.

Proof: $MS_i = S_iM'$ so that

$$Q^{-1}AQQ^{-1}T_iQ^{-1} = Q^{-1}T_iQ^{-1}Q'A'Q^{-1}.$$ 

Hence

$$AT_i = T_iA'.$$

E. Numerical Example

Consider the matrix

$$M = \begin{pmatrix}
0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1 \\
3 & 5 & 4 & 1
\end{pmatrix},$$

which has one pair of complex roots. The second and third powers of $M$ are

$$M^2 = \begin{pmatrix}
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1 \\
3 & 5 & 4 & 1 \\
3 & 8 & 9 & 5
\end{pmatrix},$$

$$M^3 = \begin{pmatrix}
0 & 0 & 0 & 1 \\
3 & 5 & 4 & 1 \\
3 & 8 & 9 & 5 \\
15 & 28 & 28 & 14
\end{pmatrix}.$$ 

From Equation (44), one obtains the polynomial

$$(x_2+1)^4 - (x_2+1)^3 - 4(x_2+1)^2 - 5(x_2+1) - 3 = 0,$$
which has the real roots $x_2 = 2$ and $x_2 = -2$. Corresponding to these two solutions are

$$
\begin{align*}
X_1 &= 1 & X_1 &= 1 \\
X_2 &= 2 & X_2 &= -2 \\
X_3 &= 2 & X_3 &= -2 \\
X_4 &= 1 & X_4 &= -3 
\end{align*}
$$

Therefore

$$G_1 = M^3 + 2M^2 + 2M + I$$

and

$$G_2 = M^3 - 2M^2 - 2M - 3I$$

are rank one matrices. Thus the matrices

$$G_{01} = \frac{G_1}{\text{tr}G_1} = \frac{1}{52} \begin{pmatrix} 1 & 2 & 2 & 1 \\ 3 & 6 & 6 & 3 \\ 9 & 18 & 18 & 9 \\ 27 & 54 & 54 & 27 \end{pmatrix}$$

and

$$G_{02} = \frac{G_2}{\text{tr}G_2} = -\frac{1}{4} \begin{pmatrix} -3 & -2 & -2 & 1 \\ 3 & 2 & 2 & -1 \\ -3 & -2 & -2 & 1 \\ 3 & 2 & 2 & -1 \end{pmatrix}$$

are rank one idempotents. The rank two idempotent $F$ can be found from

$$F = I - G_{01} - G_{02}$$

$$= \frac{1}{52} \begin{pmatrix} 12 & -28 & -28 & 12 \\ 36 & 72 & 20 & -16 \\ -48 & -44 & 8 & 4 \\ 12 & -28 & -28 & 12 \end{pmatrix}.$$
The matrix $M + p(M) = aF + b_1G_0 + b_2G_2$. By Equations (40) and (41), one obtains

\[ a = -1 \]
\[ b_1 = 6 \]
\[ b_2 = -2. \]

Therefore

\[
p(M) = \frac{1}{13} \left[ 5M^3 - 3M^2 - 16M - 211 \right]
\]
\[
= \frac{1}{13} \begin{pmatrix}
-21 & -16 & -3 & 5 \\
15 & 4 & 4 & 2 \\
6 & 25 & 12 & 6 \\
18 & 36 & 49 & 18
\end{pmatrix}.
\]
VII. BIBLIOGRAPHY


VIII. ACKNOWLEDGMENT

The author wishes to express his sincere appreciation to Bernard Vinograd for his valuable assistance and guidance during the preparation of this thesis.