Simultaneous Quasi-Diagonalization of Normal Matrices*

J. F. Watters
Department of Mathematics,
The University of Leicester,
LE1 7RH, England

Recommended by Marvin Marcus

ABSTRACT

We show that if S is a family of $n \times n$ normal matrices and $A_S$ is the algebra generated by S over the complex field, then the matrices in S are simultaneously unitarily similar to quasi-diagonal matrices if and only if

$$(AB - BA)^*Q = Q(AB - BA)^*$$

for all $A$ and $B \in A_S$, $Q \in S$. In fact, the domain of $B$ can be further restricted. For the purposes of this paper a quasi-diagonal matrix will be a matrix of block diagonal form $dg(D_1, D_2, \ldots, D_k)$ with each $D_i$ either $1 \times 1$ or $2 \times 2$, and zeros elsewhere.

1. INTRODUCTION

The literature contains a number of results which give necessary and sufficient conditions for a family of matrices to be reduced simultaneously, by some kind of similarity transformation, to some specific form, e.g., diagonal or triangular (see [2] and [5]). Among these is a theorem of Frobenius which states that a family of Hermitian matrices can be simultaneously diagonalized by a unitary matrix if and only if the members of the family commute pair-wise. In this paper we wish to establish an analogue of this result in which the form to be achieved is quasi-diagonal, i.e., with only $1 \times 1$ and/or $2 \times 2$ blocks on the diagonal, and zero elsewhere. If such a

*Some of the work for this paper was done while the author was in receipt of a grant from the University of Leicester Research Board.

quasi-diagonalization is possible, then the algebra $A_S$ generated by the given family $S$ is isomorphic to a subalgebra of a (finite) direct sum of complete $2 \times 2$ matrix algebras. Thus the algebra $A_S$ will satisfy all polynomial identities satisfied by the $2 \times 2$ matrix algebra. In particular $A_S$ will satisfy

\[(AB - BA)^2 C = C(AB - BA)^2\]

for all $A$, $B$, and $C$ in $A_S$.

The power of this identity lies in the fact that it produces central elements, viz. $[A, B]^2$, where $A$ and $B \in A_S$ and $[A, B] = AB - BA$. Using this we are able to show, in the main theorem, that if $S$ is a family of Hermitian matrices and $[[A, P]^2, Q] = 0$ for all $A \in A_S$, $P$ and $Q \in S$, then simultaneous quasi-diagonalization, by a unitary matrix, to the required form is possible. We can then extend the result to deal with normal matrices.

2. NOTATION AND TERMINOLOGY

Our principal interest is in $n \times n$ matrices over the complex field $C$ and, to a lesser extent, over the real field $R$. Where possible we consider $n \times n$ matrices over an arbitrary field $F$ and denote the algebra of all such matrices by $M_n(F)$. If $S \subseteq M_n(F)$, then the subalgebra generated by $S$ will be denoted by $A_S$. Since we are interested in finding conditions under which the matrices of $A$ can be quasi-diagonalized simultaneously, there is no loss of generality in assuming that the $n \times n$ identity matrix is in $S$ and so in $A_S$.

An $n \times n$ matrix is of type $(r_1, r_2, \ldots, r_k)$ if it is of block diagonal form

\[
\begin{bmatrix}
D_1 & \cdots & \cdots & \cdots \\
  & D_2 & \cdots & \cdots \\
  & \cdots & \ddots & \cdots \\
  & \cdots & \cdots & D_k
\end{bmatrix}
= \text{dg}(D_1; D_2, \ldots, D_k).
\]

where $D_i$ is an $r_i \times r_i$ matrix, the off-diagonal blocks are zeros, and $\sum_{i=1}^{k} r_i = n$. We shall say that a family $S$ (algebra $A_S$) is of similarity type $(r_1, r_2, \ldots, r_k)$ if the members of $S$ (of $A_S$) are simultaneously similar to matrices of type $(r_1, r_2, \ldots, r_k)$. A family $S$ or algebra $A_S$ of similarity type $(r_1, r_2, \ldots, r_k)$ is also of similarity type $(r_{\pi(1)}, r_{\pi(2)}, \ldots, r_{\pi(k)})$, where $\pi$ is any permutation of $\{1, 2, \ldots, k\}$. There is a permutation matrix which will transform a matrix of the first type
into one of the second type. In the case of matrices over \( \mathbb{C} \) (over \( \mathbb{R} \)), where the similarity transformation is effected by a unitary matrix (an orthogonal matrix), then we say that the family \( S \) or algebra \( A_S \) is of unitary type (orthogonal type) \((r_1, r_2, \ldots, r_k)\).

If \( A \in M_n(F) \) and \( T \) is a nonsingular matrix in \( M_n(F) \), then we shall write \( A^T \) for the conjugate \( T^{-1}AT \). The algebra

\[
\{ A^T : A \in A_S \}
\]

is isomorphic to \( A_S \) and denoted by \( A_S^T \). If \( F = \mathbb{C} \) and \( T \) is a unitary matrix, we shall say that \( A_S^T \) is unitarily isomorphic to \( A_S \).

The transpose of a matrix \( A \in M_n(F) \) is denoted by \( A' \) and the complex conjugate transpose of \( A \in M_n(\mathbb{C}) \) by \( A^* \).

3. MAIN THEOREM

If a family \( S \) of matrices in \( M_n(F) \) is of similarity type \((r_1, \ldots, r_k)\), then \( A_S \), whose elements are of the form \( p(P_1, \ldots, P_m) \), where \( p(x_1, \ldots, x_m) \) is a polynomial over \( F \) in the noncommuting variables \( x_1, \ldots, x_m \) and \( P_1, \ldots, P_m \in S \), is also of similarity type \((r_1, \ldots, r_k)\). Therefore the algebra \( A_S \) can be embedded in the direct sum \( \sum_{i=1}^{k} M_{r_i}(F) \).

An arbitrary algebra \( A \) over the field \( F \) is said to satisfy a polynomial identity if, for some \( m \), there is a nonzero element \( p \) of \( F[x_1, \ldots, x_m] \), the free algebra over \( F \) in the noncommuting variables \( x_1, \ldots, x_m \), such that \( p(a_1, \ldots, a_m) = 0 \) for all \( a_1, \ldots, a_m \) in \( A \) (see [4], p. 53). When \( s < r \), \( M_s(F) \) is embeddable in \( M_r(F) \). Further, a direct sum of algebras satisfies the polynomial identities satisfied by all of its component algebras.

REMARK 1. If \( S \) is of similarity type \((r_1, \ldots, r_k)\), then \( A_S \) satisfies all the polynomial identities satisfied by \( M_s(F) \), where \( r = \max \{ r_i : i = 1, \ldots, k \} \).

If \( r = 1 \), then we are in the classical situation of having matrices in \( S \) simultaneously similar to diagonal matrices. Hence \( A_S \) satisfies the polynomial identity

\[
[x_1, x_2] = x_1x_2 - x_2x_1 = 0,
\]

since this identity is satisfied by \( F \). Many known theorems give this as a sufficient condition for the family \( S \) to be of similarity type \((1, \ldots, 1)\) provided the members of \( S \) are chosen suitably.

The main aim of this paper is to find such a condition in the case \( r = 2 \).
Among the polynomial identities satisfied by $M_{2}(F)$ is the identity
\[ \left[ [x_1, x_2]^2, x_3 \right] = 0. \]

We shall show in the main theorem that when $F = \mathbb{C}$, and the matrices in $S$ are Hermitian, and $A_S$ satisfies this identity, then $A_S$ is of unitary type $(2, 2, \ldots, 2, \delta)$, where $\delta = 1$ or $2$ depending on the parity of $n$.

Before proving the main theorem it is convenient to deal first with the special situation in which all the Hermitian matrices central in $A_S$ are scalar. It is then possible to present a straightforward proof, by induction, of the main theorem.

**Lemma 1.** Let $S$ be a family of Hermitian matrices in $M_n(\mathbb{C})$ such that $A_S$ is noncommutative,
\[ \left[ [A, P]^2, Q \right] = 0 \]
for all $A \in A_S$, $P$ and $Q \in S$, and every Hermitian matrix in the center of $A_S$ is a scalar matrix. Then $n$ is even.

**Proof.** Since $A_S$ is noncommutative there are Hermitian matrices $H \in A_S$ and $P \in S$ with $[H, P] \neq 0$. If we so wished we could take both $H$ and $P$ from $S$. However, it is more convenient, for later use, not to make this restriction. Now $[H, P]$ is unitarily similar to a diagonal matrix $\operatorname{dg}(\lambda_1, \ldots, \lambda_n)$ with $\sum_{i=1}^{n} \lambda_i = 0$. But the Hermitian matrix $[H, P]^2$ is in the center of $A_S$ and so a scalar matrix. Thus $\lambda_1^2 = \lambda_2^2 = \cdots = \lambda_n^2 \neq 0$. Hence $n$ is even and $[H, P]$ is unitarily similar to
\[ \begin{bmatrix} \lambda I & 0 \\ 0 & -\lambda I \end{bmatrix}, \]
with $I$ the $n/2 \times n/2$ identity matrix.

The next step is to show that each Hermitian matrix in $A_S$ has minimum polynomial of degree $\leq 2$. This we achieve in Lemma 2. The algebraic manipulation involved in the proof of the lemma is identical with that given by Hall in [3], Lemma 1, p. 262. We omit this manipulation here. The set $S$ and algebra $A_S$ are as in Lemma 1.

**Lemma 2.** If $H$ is any Hermitian matrix in $A_S$, then
\[ H^2 + \alpha H + \beta I = 0 \]
for some $\alpha, \beta \in \mathbb{C}$. 

Proof. If $H$ is central, then, by hypothesis, $H$ is scalar and the conclusion clearly holds. If $H$ is not central, then $\exists P \in S$ with $[H, P] \neq 0$. The matrix $[H, P]^2$ will also be nonzero and, by hypothesis, a scalar matrix. By considering the central matrices $[H, P]^2$, $[H^2, P]^2$, and $[H + H^2, P]^2$, it can be shown, as in Hall's proof, that

$$H^2 + \alpha H + \beta I = 0$$

for some $\alpha, \beta \in \mathbb{C}$.

From this lemma it follows that a noncentral Hermitian matrix will have at most two, and hence exactly two, distinct characteristic roots. In fact, these two roots have equal multiplicity. This result is not required in the present situation; however, its analogue is needed, and will be proved, when we consider real symmetric matrices in the next section. We now go some way towards characterizing the algebra $A_S$ considered in the preceding two lemmas.

**Theorem 1.** Let $S$ be a family of Hermitian matrices in $M_n(\mathbb{C})$ such that $A_S$ is noncommutative, but

$$[[A, P], Q] = 0$$

for all $A \in A_S$, $P$ and $Q \in S$. If every Hermitian matrix in the center of $A_S$ is a scalar matrix, then $n$ is even and $A_S$ is either of unitary type $(n/2, n/2)$ or unitarily isomorphic to the algebra generated by matrices

$$J = \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix}, \quad K = \begin{bmatrix} 0 & L \\ L^* & 0 \end{bmatrix},$$

where $0 \neq L \in M_{n/2}(\mathbb{C})$, $K^2$ is a nonzero scalar matrix, and $I$ is the $n/2 \times n/2$ identity matrix.

**Proof.** In the statement of the theorem and throughout its proof all blocks in partititoned matrices will be $n/2 \times n/2$. Also $I$ will be used to denote both the $n \times n$ and $n/2 \times n/2$ identity matrices. It will be clear from the context which is intended.

From Lemma 1 we have that $n$ is even. From the proof of Lemma 1 we have that there is a unitarily isomorphic copy of $A_S$ containing $\text{dg}(\lambda I, -\lambda I)$, $\lambda \neq 0$.

For convenience, let us identify $A_S$ with this unitarily isomorphic copy. Thus $A_S$ contains $\mathbb{C}^2$, the algebra of all matrices $\text{dg}(\lambda_1 I, \lambda_2 I), \lambda_1, \lambda_2 \in \mathbb{C}$, since every such matrix can be realized as a linear combination of the $n \times n$...
identity matrix and \( \text{dg}(\lambda I, -\lambda I), \lambda \neq 0 \). In particular the Hermitian matrix \( J \in A_S \). Now, from Lemma 2, if \( H \) is any Hermitian matrix in \( A_S \)

\[
H^2 + \alpha H + \beta I = 0
\]

\[
(J + H)^2 + \alpha_1 (J + H) + \beta_1 I = 0.
\]

Of course, \( J^2 = J \) and therefore

\[
JH + HJ + (\alpha_1 - \alpha) H + (\alpha_1 - 1) J + (\beta_1 - \beta) I = 0.
\]

If we partition \( H = \begin{bmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{bmatrix} \) and examine the blocks in Eq. (1) we find that:

\[
(2 + \alpha_1 - \alpha) H_{11} + (\alpha_1 + 1 + \beta_1 - \beta) I = 0;
\]

\[
(\alpha_1 - \alpha) H_{22} + (\beta_1 - \beta) I = 0;
\]

and

\[
(1 + \alpha_1 - \alpha) H_{12} = 0;
\]

\[
(1 + \alpha_1 - \alpha) H_{21} = 0.
\]

Hence, either

\[
1 + \alpha_1 - \alpha = 0,
\]

or \( H_{12} = H_{21} = 0 \). Also, since we cannot have both \( 2 + \alpha_1 - \alpha = 0 \) and \( \alpha_1 - \alpha = 0 \), either \( H_{11} \) or \( H_{22} \) will be scalar. In the case when \( 1 + \alpha_1 - \alpha = 0 \), both \( H_{11} \) and \( H_{22} \) will be scalar.

Summarizing, we have that if the Hermitian matrix \( H \in A_S \setminus C^2 \), then, modulo \( C^2 \), \( H \) is congruent to a Hermitian matrix of one, and only one, of the following three forms:

(a) \( \begin{bmatrix} M_{11} & 0 \\ 0 & 0 \end{bmatrix} \), \( M_{11} \) not scalar;

(b) \( \begin{bmatrix} 0 & 0 \\ 0 & M_{22} \end{bmatrix} \), \( M_{22} \) not scalar;

(c) \( \begin{bmatrix} 0 & M_{12} \\ M_{12} & 0 \end{bmatrix} \), \( M_{12} \neq 0 \).
Furthermore, since the sum of Hermitian matrices is Hermitian, any two Hermitian matrices in $A_S \setminus C^2$ must take the same form modulo $C^2$.

If the Hermitian members of $A_S \setminus C^2$ are all of form (a) or all of form (b) modulo $C^2$, then $A_S$ is of type $(n/2, n/2)$. It remains to consider the case when all the Hermitian members of $A_S \setminus C^2$ are of form (c). Since $A_S$ is noncommutative, there will be one such matrix and suppose it is congruent, modulo $C^2$, to

$$K = \begin{bmatrix} 0 & L \\ L^* & 0 \end{bmatrix}, \quad I \neq 0.$$  

Lemma 2 implies that $K$ is a (nonzero) scalar matrix. We now show that if $M$ is a second such matrix in $A_S$, then $M$ is of the form $dg(\gamma I, \overline{\gamma} I)K$ for some $\gamma \in C$.

Suppose

$$M = \begin{bmatrix} 0 & N \\ N^* & 0 \end{bmatrix} \in A_S,$$

then the Hermitian matrices $[K, M] \sqrt{-1}$ and $KM + MK$ are also in $A_S$. They clearly cannot be of form (c), modulo $C^2$, so they belong to $C^2$ and the matrices $LN^* \pm NL^*, L^*N \pm N^*L$ are scalars. Thus $0 \neq LL^*$ and $L^*N$ are scalar matrices, so $N$ is a multiple of $L$, say, $N = \gamma L$. Then $M = dg(\gamma I, \overline{\gamma} I)K$.

From these remarks we deduce, as follows, that $A_S$ is generated, as a $C$-algebra, by $K$ and $J = dg(I, 0)$. Every matrix $dg(\mu_1 I, \mu_2 I)$ is a linear combination of $J$ and $K^2$. Since any Hermitian matrix in $A_S \setminus C^2$ is of form (c), modulo $C^2$, it will be a linear combination of $J$, $K$, $K^2$ and $JK$. The algebra $A_S$ is generated by Hermitian matrices and so can be generated by $J$ and $K$. Hence, in this case, the original algebra is unitarily isomorphic to the algebra generated by $J$ and $K$. This completes the proof of Theorem 1.

The next result is our main theorem.

**Theorem 2.** Let $S$ be a family of Hermitian matrices in $M_n(C)$. The algebra $A_S$ is of unitary type $(2, 2, \ldots, 2, \delta)$, where $\delta = 1$ or $2$, depending on the parity of $n$, if and only if

$$[[A, P] Q]_2 = 0$$

for all $A \in A_S$, $P$ and $Q \in S$. 

Proof. Suppose $A_S$ is of unitary type $(2, \ldots, 2, \delta)$. Then, by Remark 1, $A_S$ satisfies all the polynomial identities satisfied by $M^2(C)$. In particular

$$[[A, P]^2, Q] = 0$$

for all $A \in A_S$, $P$ and $Q \in S$.

The converse is proved by induction on $n$. The result is trivial when $n = 2$. Suppose that the result holds for all $k < n$ and all Hermitian families $S \subseteq M_k(C)$. If $A_S$ is a commutative algebra, then the result follows from the well-known theorem of Frobenius. From now on in this proof we take $A_S$ to be noncommutative.

If $A_S$ contains an Hermitian matrix $N$, which is central in $A_S$, but not a scalar matrix, then there is a unitary matrix $U$ such that $N^U = \text{deg}(\lambda_1 I_{n_1}, \ldots, \lambda_t I_{n_t})$ with $t > 1$ and distinct $\lambda_i$'s. The matrix $N^U$ is central in $A_S^U$, so every matrix in $A_S^U$ of type $(n_1, n_2, \ldots, n_t)$. Now $t > 1$ so $n_i < n$ for $i = 1, 2, \ldots, t$.

Applying the induction hypothesis to these blocks and, if necessary, rearranging the blocks using permutation matrices, we find that $A_S$ is of unitary type $(2, 2, \ldots, 2, \delta)$.

Finally, if every central Hermitian matrix is scalar we can use Theorem 1. Either $A_S$ is of unitary type $(n/2, n/2)$ so the result follows from the induction hypothesis, or $A_S$ is unitarily isomorphic to the algebra generated by $J$ and $K$, as in Theorem 1.

In the latter case, let us identify $A_S$ with the algebra generated by $J$ and $K$. Consider $\text{Sp}\{e_1, Ke_1\}$ where $e_1 = (1, 0, \ldots, 0)'$. This space is $A_S$-invariant (i.e., $A$-invariant for all $A \in A_S$) since $Je_1 = e_1, JKe_1 = 0$, and $K^2 e_1$ is a scalar multiple of $e_1$. Moreover, $e_1$ and $Ke_1$ are linearly independent, since $K \neq 0$, and so $\text{Sp}\{e_1, Ke_1\}$ has dimension 2. By taking orthonormal bases for $\text{Sp}\{e_1, Ke_1\}$ and its orthogonal complement in the space $C^n$, we can find a unitary matrix $U$ so that each matrix in $A_S$ is of unitary type $(2, n - 2)$. The induction hypothesis can then be used to complete the proof.

It is now a fairly simple matter to extend this result to deal with normal matrices. A characteristic property of a normal matrix $A$ is that $A^\ast$ can be expressed as a polynomial in $A$ over $C$.

Theorem 3. Let $S$ be a family of normal matrices in $M_n(C)$. The algebra $A_S$ is of unitary type $(2, 2, \ldots, 2, \delta)$, where $\delta = 1$ or 2 depending on the parity of $n$, if and only if

$$[[A, P + P^\ast]^2, Q] = 0$$

for all $A \in A_S$, $P$ and $Q \in S$. 

Proof. The necessity of the commutator condition follows from Remark 1.

To prove the converse, let us suppose that $S$ is a family of normal matrices satisfying the given condition. Put $T = \{ P + P^*, (P - P^*)/\sqrt{(-1)} \}$. Each $P^*$ is a polynomial in $P$ so $A_T \subseteq A_S$. But clearly $S \subseteq A_T$ so $A_T = A_S$. All the matrices in $T$ are Hermitian and it remains to check that $A_T$ satisfies the identity of Theorem 2.

Since $Q^*$ is a polynomial in $Q$, we have both

$$[[A, P + P^*]^2, Q] = 0$$

and

$$[[A, P + P^*]^2, Q^*] = 0.$$ 

It follows that the condition of Theorem 2 holds for the family $T$ and the algebra $A_T$. Thus $A_T = A_S$ is of unitary type $(2, 2, \ldots, 2, \delta)$.

4. REAL SYMMETRIC MATRICES

If $S$ is a family of real symmetric matrices, then it is well-known that the matrices in $S$ can be simultaneously diagonalized by an orthogonal matrix if and only if the elements of $S$ commute. It is natural to ask if there is an analogue of this theorem in the case where we require the matrices to be diagonalized to type $(2, 2, \ldots, 2, \delta)$. The following example shows that this is not possible in general.

Example 1. Let

$$J = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \text{ and } S = \left\{ \begin{bmatrix} I & 0 \\ 0 & J \end{bmatrix}, \begin{bmatrix} 0 & J \\ J' & 0 \end{bmatrix}, \begin{bmatrix} 0 & J' \\ J & 0 \end{bmatrix} \right\}$$

where all the blocks are $2 \times 2$.

The map

$$\alpha + \beta \sqrt{-1} \rightarrow \begin{bmatrix} \alpha & \beta \\ -\beta & \alpha \end{bmatrix}$$

is not a homomorphism.
is a monomorphism from $C$ into $M_2(R)$ (as rings), which induces a monomorphism from $M_2(C)$ into $M_2(M_2(R)) = M_4(R)$. Under this latter monomorphism, the elements of $S$ are the images of

\[
\begin{bmatrix}
1 & 0 \\
0 & 0
\end{bmatrix}, \quad \begin{bmatrix}
0 & 1+i \\
1-i & 0
\end{bmatrix}, \quad \text{and} \quad \begin{bmatrix}
0 & 1-i \\
1+i & 0
\end{bmatrix}.
\]

Thus $A_S$ is isomorphic to a subring of $M_2(C)$. Hence $A_S$ satisfies all the polynomial identities satisfied by $M_2(C)$, in particular

\[
[[A, P]^2, Q] = 0
\]

for all $A \in A_S$, $P$ and $Q \in S$.

On the other hand, if $O \neq x = (x_1, x_2, x_3, x_4)' \in \mathbb{R}^4$, then $x$ and the three images of $x$ under the action of the members of $S$ span a space of dimension $\geq 3$. It is not difficult to see that in fact this space must be the whole of $\mathbb{R}^4$. Thus, there are no 2-dimensional $A_S$-invariant subspaces of $\mathbb{R}^4$ and $A_S$ is not of similarity type $(2,2)$.

Note that over the complex field the unitary matrix

\[
U = \begin{bmatrix}
1 & i & 0 & 0 \\
0 & 0 & 1 & i \\
i & 1 & 0 & 0 \\
0 & 0 & i & 1
\end{bmatrix}
\]

is such that $UPU^*$ is of type $(2,2)$ for any $P \in S$.

This example involves an algebra with 3 generators. If the family has only 2 members then there is a diagonalization theorem for symmetric matrices. As a preliminary, we look at a result in the situation analogous to that in Theorem 1.

**Theorem 4.** Let $S$ be a family of symmetric matrices in $M_n(R)$, and suppose that $A_S$ is noncommutative but such that

\[
[[A, P]^2, Q] = 0
\]

for all $A \in A_S$, $P$ and $Q \in S$. If every symmetric matrix in the center of $A_S$ is
scalar, then \( n \) is even and every noncentral symmetric matrix in \( A_s \) has exactly two distinct characteristic roots and these have equal multiplicity.

Proof. The arguments used in the proofs of Lemmas 1 and 2 apply just as well in this context. Thus \( n \) is even and every noncentral symmetric matrix in \( A_s \) has exactly two distinct characteristic roots. It remains to establish their equal multiplicity.

If \( H \) is a noncentral symmetric matrix in \( A_s \), the \( \exists P \in S \) with \([H, P] \neq 0\). Hence \( \exists \) an orthogonal matrix \( U \) such that

\[
[H, P]^U = \text{dg}(\lambda I, -\lambda I), \lambda \neq 0.
\]

From Lemma 2

\[
[H^2, P] + \alpha [H, P] = 0
\]

for some \( \alpha \in \mathbb{C} \).

Let

\[
H^U = \begin{bmatrix}
H_{11} & H_{12} \\
H_{21} & H_{22}
\end{bmatrix}.
\]

Then

\[
\]

\[
= \begin{bmatrix}
2\lambda H_{11} & 0 \\
0 & -2\lambda H_{22}
\end{bmatrix}
\]

and, from (1),

\[
[H^2, P]^U = -\alpha [H, P]^U
\]

\[
= \text{dg}(-\alpha \lambda I, \alpha \lambda I).
\]

Therefore \( H_{11} = H_{22} = (\alpha / 2)I \). Now

\[
(H^U)^2 + \alpha H^U + \beta I = 0,
\]

so that the matrices \( H_{12} H_{21} \) and \( H_{21} H_{12} \) are scalar. But \( H_{21} = H_{12} \) and \( H_{12} \neq 0 \), since \( H \) is not a scalar matrix, so \( H_{12} H_{21} \neq 0 \). Thus \( H_{12} \) is a nonsingular matrix and the rank of \( H + \gamma I \), for any \( \gamma \in \mathbb{C} \), is \( > n / 2 \). Therefore the nullity of \( H + \gamma I \) is \( < n / 2 \) and so any characteristic root of \( H \) has geometric multiplicity \( < n / 2 \). However, the algebraic multiplicity of a characteristic root of the
symmetric matrix $H$ coincides with its geometric multiplicity, so $H$ has two characteristic roots of equal multiplicity, $n/2$.

**Theorem 5.** Let $X$ and $Y$ be symmetric matrices in $M_n(\mathbb{R})$ and $S = \{X, Y\}$. The algebra $A_S$ is of orthogonal type $(2, \ldots, 2, \delta)$, where $\delta = 1$ or 2 depending on the parity of $n$, if and only if

$$\left[(A, P)^\delta, Q\right] = 0$$

for all $A \in A_S$, $P$ and $Q \in S$.

**Proof.** If $A_S$ is of the given orthogonal type, then, by Remark 1, the conditions above are satisfied.

The converse is proved by induction on $n$. The result is trivial when $n = 2$.

Suppose that the result holds for all $k < n$ and all pairs of symmetric matrices $\{X, Y\} \subseteq M_k(\mathbb{R})$. If $A_S$ contains a symmetric matrix $N$, which is central in $A_S$, but not a scalar matrix, then there is an orthogonal matrix $Z$ such that $N^\delta = \text{diag}(I, I, \ldots, I)$ with $\delta > 1$ and distinct $\lambda_i$'s. The matrix $N^\delta$ is central in $A_S^\delta$ so every matrix is of type $(n_1, n_2, \ldots, n_t)$. Each $n_i < n$ so applying the induction hypothesis to these blocks, and, if necessary, rearranging the blocks using permutation matrices, we find that $A_S$ is of orthogonal type $(2, \ldots, 2, \delta)$.

If $A_S$ is noncommutative and every central symmetric matrix is scalar we can use Theorem 4. The matrices $X$ and $Y$ cannot be central, otherwise $A_S$ would be commutative, so there is an orthogonal matrix $U$, such that

$$X^U = \begin{bmatrix} \lambda_1 I & 0 \\ 0 & \lambda_2 I \end{bmatrix}$$

with $\lambda_1 \neq \lambda_2$, and

$$Y^U = \begin{bmatrix} Y_{11} & Y_{12} \\ Y_{21} & Y_{22} \end{bmatrix}$$

We can diagonalize $Y_{11}$ and $Y_{22}$ by orthogonal matrices without affecting $X^U$, so let $U$ be chosen so that $Y_{11} = \mu_1 I$ and $Y_{22} = \mu_2 I$. Then

$$[X, Y] = \begin{bmatrix} 0 & (\lambda_1 - \lambda_2) Y_{12} \\ (\lambda_2 - \lambda_1) Y_{21} & 0 \end{bmatrix}.$$
Consider $\text{Sp}(e_1, Ye_1)$ where $e_1 = (1, 0, \ldots, 0)'$. The matrix $Y$ has minimal polynomial of degree 2 so $Y^2e_1 \in \text{Sp}(e_1, Ye_1)$. Also $Xe_1 = \lambda_1 e_1$ and

$$XYe_1 = YXe_1 + [X, Y]e_1$$

$$= \lambda_1 Ye_1 + (\lambda_2 - \lambda_1)(Ye_1 - \mu_1 e_1),$$

so that $\text{Sp}(e_1, Ye_1)$ is $A_S$-invariant. By taking orthonormal bases for this space and its orthogonal complement in $\mathbb{R}^n$, we can find a (real) orthogonal matrix $V$ so that each matrix in $A_S^V$ is of type $(d, n - d)$, where $d = \dim \text{Sp}(e_1, Ye_1) \geq 1$. The induction hypothesis can then be used to complete the proof.

Finally, if $A_S$ is commutative then the result follows from Frobenius' Theorem.

5. FURTHER OBSERVATIONS

Remark 2. In the main theorem we required

$$[[A, P]^2, Q] = 0$$

for all $A \in A_S$, $P$ and $Q \in S$. A closer examination of the proof shows that we could get by if we restrict $A$ to be Hermitian. Unfortunately, it is not sufficient to have

$$[[P, Q]^2, R] = 0$$

for all $P, Q$ and $R \in S$. The following example shows this.

If $X = \text{dg}(1, 2, 3, 4)$ and

$$Y = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 3 & -1 \\ 1 & 3 & 1 & 1 \\ 1 & -1 & 1 & 1 \end{bmatrix},$$

then $A_S$ is not of similarity type $(2, 2)$. The only 2-dimensional spaces fixed by $X$ are $\text{Sp}(e_i, e_j)$, where $e_i$ and $e_j$ are different natural basis vectors of $\mathbb{C}^4$. However, $Ye_1 \not\in \text{Sp}(e_i, e_j)$, so this space is not fixed by $Y$. Hence $A_S$ is not of
similarity type \((2,2)\). But

\[ [X, Y]^2 = -14I_4, \]

so that \( [[P, Q]^2, R] = 0 \) for all \( P, Q \) and \( R \in S \).

**Remark 3.** The polynomial identity

\[ [[[x, y]^2, z]] = 0 \]

is not an identity of minimal degree satisfied by \( M_2(F) \). Another identity satisfied by this matrix ring is the so-called standard identity of degree 4, that is,

\[ [x_1, x_2, x_3, x_4] = \sum_{\pi \in S_4} \sum_{\sigma \in S_4} \pm x_{\pi(1)}x_{\pi(2)}x_{\sigma(3)}x_{\sigma(4)} = 0, \]

where the sign is + if \( \pi \) is an even permutation and – if \( \pi \) is an odd permutation. It is known (see [1]) that if \( x_1, \ldots, x_k \) is a given set of indeterminates, then minimal polynomials satisfied by \( M_2(F) \) and depending only on the \( x \)'s may be constructed if and only if \( k > 4 \). Any minimal polynomial in \( x_1, \ldots, x_k \) will then be a linear combination of the \( \binom{k}{4} \) standard polynomials \([x_{i_1}, \ldots, x_{i_4}]\), where \((i_1, \ldots, i_4)\) is an arbitrary combination of 4 letters out of \( k \). The only exception to this is when \( F \) is the field of 2 elements. This is not relevant to the present discussion.

It is now natural to ask if our main theorem still holds when the identity used there is replaced by the identity

\[ [A, B, C, D] = 0 \quad \text{for all} \quad A, B, C, D \in A_S. \]

The author is at present unable to answer this question.

I am grateful to the referee, whose careful reading of the first version of this paper has led to the removal of a number of obscurities and errors from this work.

**References**


*Received August 1973*