JOURNAL OF RESEARCH of the National Bureau of Standards — B. Mathematics and Mathematical Physics
Vol. 71B, Nos. 2 and 3, April–September 1967

# Two Classical Theorems on Commuting Matrices

**Morris Newman**

**Institute for Basic Standards, National Bureau of Standards, Washington, D.C. 20234**

**(April 3, 1967)**

Simple proofs are given of the following classical theorems: (1) An arbitrary set of commuting matrices may be simultaneously brought to triangular form by a unitary similarity. (2) An arbitrary set of commuting normal matrices may be simultaneously brought to diagonal form by a unitary similarity.

Key Words: Commuting matrices, group representations, normal matrices, Schur's lemma, simultaneous triangularization and diagonalization.

## 1. Introduction

The purpose of this note, which is expository, is to present proofs of two fundamental theorems on sets of commuting matrices. The theorems are classical, but existing proofs tend to be unnecessarily complicated and furthermore are difficult to find in the literature. For these reasons the following simple proofs (along the lines set down by Frobenius and Schur in their original memoirs on group representations) are of interest. In fact the basic tools are the concept of irreducibility and a simplified form of Schur's lemma, both from the theory of group representations. In addition some special information concerning normal matrices will be required, which we summarize briefly below. (Complete proofs may be found in MacDuffee's book.)[1] For simplicity, all the matrices considered below are over the complex field.

A matrix $A$ is *normal* if it commutes with $A^*$, the conjugate transpose of $A$. A matrix $U$ is *unitary if* $U^*=U^{-1}$. Unitary matrices are themselves normal matrices, the unitary matrices form a group, and $U^*AU$ is normal if and only if $A$ is normal. It is known that $A$ is normal if and only if $A=U^*DU$, where $U$ is unitary and $D$ diagonal.

We also need

LEMMA 1. *Let* A *be an arbitrary* n × n *complex matrix. Then there is a unitary matrix* U *and an upper triangular matrix* T *such that* A = UT.

PROOF. Let $A_1$ be the first column of $A$ and suppose that $A_1 \neq 0$. Put $V_1=A_1/N(A_1)$ ($N(A_1)$ is the euclidean norm of $A_1$) and complete $V_1$ by the Gram-Schmidt process to an orthonomal basis $V_1, V_2, \ldots, V_n$. Then the matrix $V=(V_1, V_2, \ldots, V_n)$ is unitary, and the first column of $V^*A$ is

$$\begin{bmatrix} V_1^*A_1 \\ V_2^*A_1 \\ \cdots \\ V_n^*A_1 \end{bmatrix} = \begin{bmatrix} N(A_1) \\ 0 \\ \cdots \\ 0 \end{bmatrix}$$

Define

$$W=\begin{cases} I & A_1=0 \\ V & A_1\neq 0. \end{cases}$$

Then $W$ is unitary, and the first column of $W^*A$ has all elements below the diagonal element 0. The process may now be repeated with the matrix $W^*A$ using unitary transformations of the form $(1)+V$, etc. We ultimately find that there is a unitary matrix $U$ such that $U^*A=T$ is upper triangular. Hence $A=UT$, and the proof of the lemma is concluded.

We also require the fact that if $A$ is any matrix such that $AA^*$ has zero trace[2] then $A$ must be the zero matrix. For if $A=(a_{ij})$ then

$$tr(AA^*)=\sum_{i,j} |a_{ij}|^2=N(A)^2,$$

and so $tr(AA^*)=0$ implies $A=0$.

## 2. The Theorems and Their Proofs

A set (finite or infinite) of $n\times n$ matrices $\mathfrak{A}=\{A\}$ is said to be *reducible* if fixed positive integers $p$, $q$,

[1] C. C. MacDuffee, The Theory of Matrices, New York, 1946.

[2] The trace of a matrix $A$ is the sum of its diagonal elements, and is a similarity invariant; i.e., remains unchanged under transformations of the form $S^{-1}AS$, where $S$ is any nonsingular matrix.



263-048 O - 67 - 2

and a fixed nonsingular matrix $S$ exist such that for each $A \epsilon \mathfrak{A}$,

$$(1) \qquad S^{-1}AS = \begin{pmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{pmatrix},$$

where $A_{11}$ is a $p \times p$ matrix, $A_{12}$ a $p \times q$ matrix and $A_{22}$ a $q \times q$ matrix. Otherwise $\mathfrak{A}$ is said to be *irreducible*. If the form (1) can be achieved with $A_{12} = 0$ as well for all $A \epsilon \mathfrak{A}$, then $\mathfrak{A}$ is said to be *fully reducible*.

Since any matrix is the product of a unitary by an upper triangular matrix (lemma 1) and a similarity transformation by an upper triangular matrix retains the block form of (1), the matrix $S$ may be chosen unitary. From this remark it follows that $\mathfrak{A}$ is reducible if and only if it it unitarily reducible, so that these are equivalent concepts.

The basic lemma is the following:

*Lemma 2 (Schur's lemma, specialized). Let* $\mathfrak{A} = \{A\}$ *be an irreducible set of* $n \times n$ *matrices, and let* $M$ *be a fixed matrix such that for each* $A \epsilon \mathfrak{A}$, *there is a matrix* $\tilde{A}$ *satisfying*

$$AM = M\tilde{A}.$$

*Then either* $M = 0$ *or* $M$ *is nonsingular. Furthermore if* $\tilde{A} = A$ *(so that* $M$ *commutes with each element of* $\mathfrak{A}$*) then* $M$ *is scalar.*[3]

PROOF. Suppose that the rank of $M$ is $r$, and write

$$M = P \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} Q$$

where $P$, $Q$ are nonsingular and $I_r$ is the $r \times r$ identity matrix. Then for each $A \epsilon \mathfrak{A}$,

$$(2) \qquad (P^{-1}AP) \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} (Q\tilde{A}Q^{-1}).$$

Put

$$P^{-1}AP = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix},$$

$$Q\tilde{A}Q^{-1} = \begin{pmatrix} \tilde{A}_{11} & \tilde{A}_{12} \\ \tilde{A}_{21} & \tilde{A}_{22} \end{pmatrix}$$

where $A_{11}$, $\tilde{A}_{11}$ are $r \times r$ matrices, $A_{12}$, $\tilde{A}_{12}$ $r \times (n-r)$ matrices, $A_{21}$, $\tilde{A}_{21}$ $(n-r) \times r$ matrices and $A_{22}$, $\tilde{A}_{22}$ $(n-r) \times (n-r)$ matrices. Then (2) implies that

$$\begin{pmatrix} A_{11} & 0 \\ A_{21} & 0 \end{pmatrix} = \begin{pmatrix} \tilde{A}_{11} & \tilde{A}_{12} \\ 0 & 0 \end{pmatrix}.$$

[3] That is, a multiple of the identity matrix.

Thus $A_{21} = 0$, an impossibility since $\mathfrak{A}$ is irreducible. Hence $r$ must be 0 or $n$, and so either $M = 0$ or $M$ is nonsingular. This proves the first part of the lemma.

Now suppose that $M$ commutes with each element of $\mathfrak{A}$, and choose $\lambda$ as any eigenvalue of $M$. Then $M - \lambda I$ is singular and also commutes with each element of $\mathfrak{A}$. Hence $M - \lambda I = 0$, $M = \lambda I$, and the second part of the lemma is proved.

We note that the lemma remains true if $\mathfrak{A}$ is assumed unitarily irreducible. This remark will find application later.

Now suppose that $\mathfrak{A} = \{A\}$ is any set of $n \times n$ matrices. It is clear that after a suitable similarity has been performed, the matrices $A$ may be taken so that with respect to some fixed partitioning,

$$A = (A_{ij})$$

where $A_{ij} = 0$ for $i > j$, and for each $i$ the set $\mathfrak{A}_i = \{A_{ii}\}$ is irreducible. If we assume in addition that $\mathfrak{A}$ is a set of commuting matrices, then it follows that for each $i$ $\mathfrak{A}_i$ is also a set of commuting matrices, and hence by lemma 2 that $\mathfrak{A}_i$ consists entirely of scalar matrices. Hence we have proved the first of the two theorems:

THEOREM 1. *Let* $\mathfrak{A} = \{A\}$ *be any set of commuting matrices. Then there is a fixed nonsingular matrix* $S$ *(which may be chosen unitary) such that* $S^{-1}AS$ *is upper triangular for each* $A \epsilon \mathfrak{A}$.

Now let $\mathfrak{A} = \{A\}$ be any set of $n \times n$ normal matrices. We first prove

LEMMA 3. *If* $\mathfrak{A}$ *is unitarily reducible then it is unitarily fully reducible.*

PROOF. Suppose that $\mathfrak{A}$ is unitarily reducible and let $U$ be a unitary matrix such that with respect to some fixed partitioning,

$$U^{-1}AU = \begin{pmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{pmatrix}$$

for each $A \epsilon \mathfrak{A}$.

Since normality is preserved under unitary similarities, the matrices $U^{-1}AU$ are normal. Hence

$$\begin{pmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{pmatrix} \begin{pmatrix} A_{11}^* & 0 \\ A_{12}^* & A_{22}^* \end{pmatrix} = \begin{pmatrix} A_{11}^* & 0 \\ A_{12}^* & A_{22}^* \end{pmatrix} \begin{pmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{pmatrix},$$

and it follows that

$$A_{12}A_{12}^* = A_{11}^*A_{11} - A_{11}A_{11}^*.$$

Thus

$$tr(A_{12}A_{12}^*) = 0,$$

since $A_{11}^*A_{11}$ and $A_{11}A_{11}^*$ have the same trace. By the remarks made in the introduction, it follows that $A_{12}=0$. This completes the proof of the lemma.

From lemma 3 we obtain

LEMMA 4. *There is a unitary matrix* U *such that for each* $A\epsilon\mathfrak{A}$,

$$\text{(3)} \qquad U^{-1}AU = A_{11} \dot{+} A_{22} \dot{+} \ldots \dot{+} A_{rr}$$

*where the set* $\mathfrak{A}_i = \{A_{ii}\}$ *is a unitarily irreducible set of normal matrices,* $1 \leq i \leq r$.

Combining these lemmas, we obtain the second of the two theorems:

THEOREM 2. *Let* $\mathfrak{A} = \{A\}$ *be a set of commuting normal matrices. Then there is a fixed unitary matrix* U *such that* $U^{-1}AU$ *is diagonal for each* $A\epsilon\mathfrak{A}$.

PROOF. Choose $U$ so that the form (3) is achieved. Then the sets $\mathfrak{A}_i = \{A_{ii}\}$, $1 \leq i \leq r$ are unitarily irreducible sets of commuting normal matrices. Lemma 2 now implies that $A_{ii}$ is scalar for each $A_{ii}\epsilon\mathfrak{A}_i$, $1 \leq i \leq r$ from which the theorem follows.

## 3. Consequences of the Theorems and a Problem

These theorems have many important consequences, of which we mention two:

(4) Let $A_j$, $1 \leq j \leq p$ be commuting $n \times n$ matrices and let $f = f(x_1, x_2, \ldots, x_p)$ be an arbitrary polynomial in $x_1, x_2, \ldots, x_p$. Then there is a fixed ordering of the eigenvalues of $A_j$, say $\lambda_j(1), \lambda_j(2), \ldots, \lambda_j(n)$, $1 \leq j \leq p$ (which does not depend on $f$) such that the eigenvalues of $f(A_1, A_2, \ldots, A_p)$ are precisely $f(\lambda_1(i), \lambda_2(i), \ldots, \lambda_p(i))$, $1 \leq i \leq n$.

(5) The irreducible representations of an abelian group are all of degree 1.

There are also important applications in quantum mechanics, in the theory of the Hecke operators, and of course in group representations.

The following problem has some interest: Give conditions for the simultaneous diagonalizability of a given set $\mathfrak{A}$ of commuting $n \times n$ matrices. One such criterion is furnished by Theorem 2. Another sufficient condition is that $\mathfrak{A}$ contain a diagonalizable nonderogatory matrix (one whose characteristic and minimal polynomials coincide); for example, one with distinct eigenvalues. An inductive solution is as follows: If $\mathfrak{A}$ consists entirely of scalar matrices, we are through. If not, $\mathfrak{A}$ must contain a diagonalizable nonscalar matrix $B$, and after a suitable similarity has been performed we may assume that

$$\text{(6)} \qquad B = \lambda_1 I \dot{+} \lambda_2 I \dot{+} \ldots \dot{+} \lambda_r I$$

where $r > 1$ and $\lambda_i = \lambda_j$ if and only if $i = j$. Next (6) and the fact that the elements of $\mathfrak{A}$ commute imply that if $A$ is any element of $\mathfrak{A}$ then

$$A = A_{11} \dot{+} A_{22} \dot{+} \ldots \dot{+} A_{rr}$$

where the partitioning is that imposed by the form $B$. The problem is now reduced to the study of the $r$ commuting sets $\mathfrak{A}_i = \{A_{ii}\}$, $1 \leq i \leq r$, each of smaller dimension than $n$, to which the procedure described above may be applied again, etc. The difficulty of course lies in recognizing when a given set of matrices contains a nonscalar diagonalizable element.

(Paper 71B2 & 3–201)