COMMUTATORS WHICH COMMUTE WITH ONE FACTOR

Helene Shapiro

Dedicated to the memory of Olga Taussky-Todd

Let A and B be $n \times n$ matrices, let $C = ABA^{-1}B^{-1}$ be the multiplicative commutator of A and B, and assume $AC = CA$. Olga Taussky (1961) examined the structure of A and B for the case where A and B are unitary. Marcus and Thompson (1966) generalized her results to the case where A and C are assumed normal. We look at this problem for general A, with particular attention to the cases where A is diagonalizable or nonderogatory.

Now let $[A, B] = AB - BA$ be the additive commutator of A and B and assume A commutes with $[A, B]$. The finite-dimensional version of a theorem of Putnam tells us that if A is normal, then A and B commute. We show that the same conclusion holds when A is diagonalizable. If A is nonderogatory, then A and B can be simultaneously triangularized.

1. Introduction.

A 1911 paper of Frobenius contains the following theorem.

Theorem A (Frobenius, [**9**]). *Let $C = ABA^{-1}B^{-1}$ be the commutator of the two unitary $n \times n$ matrices A and B. Assume that the characteristic roots of B are included in an arc less than a semicircle. If A commutes with C, then A commutes with B.*

Frobenius used this result to further simplify a proof Bieberbach gave of a theorem of Jordan about matrix groups [**8**, **9**, **2**, **22**]. We use the term "cramped" for a unitary matrix whose chacteristic roots are included in an arc less than a semicircle [**1**, **26**].

Olga Taussky [**25**], in her 1961 paper, "Commutators of unitary matrices which commute with one factor", examined the following question: "What is the structure of B if A does not commute with B?" Zassenhaus made further remarks in [**32**] and Wielandt reformulated Taussky's theorem in his review of her paper [**31**].

Marcus and Thompson [**16**] then generalized both theorem A and Taussky's result to normal matrices. By using the field of values, they found an appropriate generalization for the condition that B be cramped and also found a simpler proof than the computation used by Frobenius.

Definition.   If M is an $n \times n$ complex matrix, then the *field of values, or numerical range of M is the set* $F(M) = \{x^*Mx | x \in C^n \text{ and } x^*x = 1\}$.

The set $F(M)$ is a set of complex numbers which may be viewed as a set of points in R^2 by using the usual correspondence between a complex number $a + ib$ and the point (a, b). The field of values has been much studied [**13**, Chapter 1]. It is a closed, convex set in R^2. If λ is an eigenvalue of M, with eigenvector x, where $x^*x = 1$, then $x^*Mx = \lambda$, so $F(M)$ contains the eigenvalues of M. Hence, it contains the polygon formed by taking the convex hull of the eigenvalues. For any unitary matrix U, we have $F(U^*MU) = F(M)$, so $F(M)$ is invariant under unitary similarity. When M is normal, we can diagonalize M with a unitary similarity and then show that the numerical range is just the convex hull of the eigenvalues. This leads to the connection between the field of values and cramped matrices. Unitary matrices are normal, so the field of values of a unitary matrix will be the convex hull of its eigenvalues. These eigenvalues lie on the unit circle, so the matrix is cramped if and only if 0 is not in the field of values.

Theorem A$'$ (Marcus and Thompson [**16**]).   *Let $C = ABA^{-1}B^{-1}$ be the commutator of the $n \times n$ matrices A and B. Assume that A and C are normal and that* 0 *is not in $F(B)$. Then $AC = CA$ implies that $AB = BA$.*

The proof of Marcus and Thompson uses the fact that commuting normal matrices can be simultaneously diagonalized with a unitary similarity, together with the fact that $F(B)$ contains the diagonal entries of B. (If $x = e_i$ is the i'th unit coordinate vector, then $x^*Mx = m_{ii}$.)

This result has also been studied for bounded operators on a Hilbert space. Putnam [**19**] proved the Frobenius theorem for unitary operators on a Hilbert space, and DePrima and Richard [**3**] established the following improvement of the Marcus and Thompson theorem.

Theorem A$''$.   *Let A and B be bounded, invertible linear operators on a complex Hilbert space and let $C = ABA^{-1}B^{-1}$. If A is normal, then either $AB = BA$ or* 0 *is in the closure of $F(B)$.*

Marcus and Thompson [**16**] also generalized results in Taussky's 1961 paper with the following theorem.

Theorem B.   *Let $C = ABA^{-1}B^{-1}$ be the commutator of the $n \times n$ matrices A and B. Assume that A and C are normal and that $AC = CA$. After a*

simultaneous unitary similarity, we may assume that $A = \alpha_1 I_{n_1} \oplus \alpha_2 I_{n_2} \oplus \cdots \oplus \alpha_t I_{n_t}$ where $\alpha_1, \alpha_2, \cdots, \alpha_t$ are distinct and that C is diagonal. Then $B = PW$, where P is a permutation matrix and $W = W_1 \oplus W_2 \oplus \cdots \oplus W_t$ where W_i is $n_i \times n_i$, and $C = APA^{-1}P^{-1}$. Conversely, if A and B are in the form described, then $AC = CA$, and $C = APA^{-1}P^{-1}$ is diagonal.

In this paper we shall examine the case where A is not normal and obtain generalizations of Theorem B for the cases where A is diagonalizable and A is nonderogatory.

There is an analagous problem for the additive commutator $[A, B] = AB - BA$. If A commutes with $[A, B]$, what can we say about the pair A and B? We examine this in Section 4 in connection with a theorem of Putnam [**20**].

2. Preliminaries.

We work over the complex numbers $\mathbb{C}$, however, most of our results hold over an algebraically closed field of characteristic 0. If M is an $n \times n$ matrix, then $\operatorname{spec}(M)$ denotes the set of eigenvalues (spectrum) of M. Throughout Sections 2 and 3, A and B denote nonsingular, $n \times n$ matrices and $C = ABA^{-1}B^{-1}$ is the multiplicative commutator of A and B. We set $\operatorname{spec}(A) = \{\alpha_1, \alpha_2, \dots, \alpha_t\}$, where $\alpha_1, \alpha_2, \dots, \alpha_t$ are distinct and let n_i denote the algebraic multiplicity of the eigenvalue α_i. We can then apply a similarity to put A into a block diagonal form, $A_1 \oplus A_2 \oplus \cdots \oplus A_t$, where A_i is $n_i \times n_i$ and $\operatorname{spec}(A_i) = \{\alpha_i\}$. This block diagonal form for A comes from decomposing $\mathbb{C}^n$ into a direct sum of the generalized eigenspaces of A. Thus, if $V_i = \{x \in \mathbb{C}^n | (A - \alpha_i)^k x = 0$ for some positive integer $k\}$ is the generalized eigenspace of A, corresponding to α_i, then V_i is an invariant subspace of A, with $\dim(V_i) = n_i$, and $\mathbb{C}^n = V_1 \oplus V_2 \oplus \cdots \oplus V_t$. The i'th diagonal block A_i in the block diagonal form of A represents the action of A on the subspace V_i.

If we apply a similarity, S, to both A and B, then $S^{-1}CS$ is the multiplicative commutator of $S^{-1}AS$ and $S^{-1}BS$; note also that A and C commute if and only if $S^{-1}AS$ and $S^{-1}CS$ commute. So, without loss of generality, we may assume that $A = A_1 \oplus A_2 \oplus \cdots \oplus A_t$. Any matrix M which commutes with A must then have the form $M_1 \oplus M_2 \oplus \cdots \oplus M_t$, where M_i is $n_i \times n_i$. This can be proven by using the well known result that the matrix equation $RX - XT = 0$ has nontrivial solutions, X, if and only if R and T have a common eigenvalue [**10**, **13**, **23**]. Partition M into blocks conformal with the diagonal blocks of A. The equation $AM = MA$ then gives $A_i M_{ij} = M_{ij} A_j$. Since A_i and A_j do not have a common eigenvalue unless $i = j$, we must have $M_{ij} = 0$ whenever $i \neq j$. Alternatively, one can easily show that if M

commutes with A, then the generalized eigenspaces V_i, of A, are invariant subspaces of M, and hence M must be in the block diagonal form.

We use term "triangular" to mean upper triangular; note that the diagonal entries of a triangular matrix are its eigenvalues. Matrices which commute can be put in triangular form with a simultaneous similarity [7], furthermore, it is possible to do this with a unitary similarity. The proof of this depends on the fact that if $RT = TR$ and ρ is an eigenvalue of R, then the eigenspace for ρ is invariant under T. Hence, one can choose a vector, x, in this eigenspace which is also an eigenvector of T. Let x be the first vector in a basis for C^n; then the similarity transformation corresponding to this new basis will put R and T in the forms

$$
R \longrightarrow \begin{pmatrix} \rho & * * \cdots * \\ 0 & \\ 0 & \widetilde{R} \\ 0 & \end{pmatrix} \quad T \longrightarrow \begin{pmatrix} \tau & * * \cdots * \\ 0 & \\ 0 & \widetilde{T} \\ 0 & \end{pmatrix}
$$

Since $\widetilde{R}\widetilde{T} = \widetilde{T}\widetilde{R}$, the argument can be repeated until R and T have been triangularized. By choosing x to be a vector of length 1, and using it as the first vector in an orthonormal basis, one can show that R and S can be simultaneously triangularized with a unitary similarity. Finally, note that one can specify the order of the diagonal entries for the matrix R.

Applying this to the situation $AC = CA$, we see that after a simultaneous similarity, we may assume $A = A_1 \oplus A_2 \oplus \cdots \oplus A_t$, where $\operatorname{spec}(A_i) = \{\alpha_i\}$, and $C = C_1 \oplus C_2 \oplus \cdots \oplus C_t$, where C_i is $n_i \times n_i$ and A_i and C_i are triangular.

We begin with a simple matrix proof for the finite dimensional case of Theorem A''. This proof resembles the original proof of Marcus and Thompson gave for Theorem A' and is also similar to the argument in Remark 4 of [3, p. 254]. We use the fact that if $\widetilde{M}$ is a principal submatrix of M, then $F(\widetilde{M}) \subseteq F(M)$.

Theorem 1. *Assume A is normal and 0 is not in $F(B)$. Then $AC = CA$ if and only if $AB = BA$.*

Proof. If $AB = BA$, then $C = I$, so obviously $AC = CA$. Now assume $AC = CA$. Since A is normal, there is a unitary matrix U such that $U^*AU = \alpha_1 I_{n_1} \oplus \alpha_2 I_{n_2} \oplus \cdots \oplus \alpha_t I_{n_t}$. Applying U to A, B and C, we may assume A is already in this diagonal form. Then $CA = AC$ implies that $C = C_1 \oplus C_2 \oplus \cdots \oplus C_t$, where C_i is $n_i \times n_i$.

From $C = ABA^{-1}B^{-1}$ we get $CBA = AB$. Partition B conformally with the blocks of A and C and denote the ij block by B_{ij}. Then the i'th diagonal block of CBA is $\alpha_i C_i B_{ii}$, while the i'th diagonal block of AB is

$\alpha_i B_{ii}$. Hence, $C_i B_{ii} = B_{ii}$. Now, $F(B_{ii}) \subseteq F(B)$, so 0 is not in $F(B_{ii})$. So B_{ii} is nonsingular, and the equation $C_i B_{ii} = B_{ii}$ tells us C_i is an identity matrix. Therefore $C = I$ and $AB = BA$. $\qquad\square$

We now drop the assumption that A is normal and see what the hypothesis $AC = CA$ tells us about the structure of A and B. Most of our results deal with the cases where A is diagonalizable or nonderogatory, and make no assumptions about B. Notice that while $F(B) = F(U^{-1}BU)$ for a unitary similarity, U, in general $F(B) \neq F(S^{-1}BS)$ for a general similarity S. Hence the condition $0 \notin F(B)$ will not be preserved if we apply a general similarity.

Our first result concerns the eigenvalues of C.

Theorem 2. *If $AC = CA$ then each eigenvalue of C is a quotient of eigenvalues of A. Thus, if λ is an eigenvalue of C, then $\lambda = \alpha_i/\alpha_j$ for some i and j.*

Proof. Since $AC = CA$, we can assume A and C are upper triangular; the diagonal entries will be the eigenvalues. From $C = ABA^{-1}B^{-1}$, we get $A^{-1}C = BA^{-1}B^{-1}$, so $A^{-1}C$ is similar to A^{-1}. But $A^{-1}C$ is upper triangular, and for some i, has a diagonal entry λ/α_i. So λ/α_i must be an eigenvalue of A^{-1}. Therefore, $\lambda/\alpha_i = 1/\alpha_j$ for some j, and $\lambda = \alpha_i/\alpha_j$. $\qquad\square$

Observe that $BAB^{-1} = C^{-1}A$, and so A commutes with C if and only if A commutes with BAB^{-1}, which is similar to A. Note that $B(V_i)$ is the generalized eigenspace of BAB^{-1}, for the eigenvalue α_i. Hence, when $AC = CA$, the space $B(V_i)$ will be invariant under A. Our next result is stated in matrix language, but actually tells us that $\mathrm{spec}(C) = \{1\}$ if and only if $B(V_i) = V_i$ for each i.

Theorem 3. *Assume $AC = CA$. Then there is a similarity which will simultaneously put A in the form $A_1 \oplus A_2 \oplus \cdots \oplus A_t$, where A_i is $n_i \times n_i$, and $\mathrm{spec}(A_i) = \{\alpha_i\}$, and B in the form $B_1 \oplus B_2 \oplus \cdots \oplus B_t$, where B_i is $n_i \times n_i$, if and only if $\mathrm{spec}(C) = \{1\}$.*

Proof. Since $AC = CA$, we may assume $A = A_1 \oplus A_2 \oplus \cdots \oplus A_t$ where A_i is $n_i \times n_i$ and $\mathrm{spec}(A_i) = \{\alpha_i\}$ and $C = C_1 \oplus C_2 \oplus \cdots C_t$ where C_i is $n_i \times n_i$. Furthermore, since $A_i C_i = C_i A_i$, we may assume the blocks A_i and C_i are upper triangular.

Suppose $\mathrm{spec}(C) = \{1\}$. Partition B conformally and compute the ij block on both sides of the equation $CBA = AB$. This gives $C_i B_{ij} A_j = A_i B_{ij}$, or $B_{ij} A_j = C_i^{-1} A_i B_{ij}$. Now A_i and C_i are upper triangular, and since $\mathrm{spec}(C) = \{1\}$, each diagonal entry of C_i (and thus of C_i^{-1}) is a one. Therefore, every diagonal entry of the triangular matrix $C_i^{-1} A_i$ is α_i and so

$\mathrm{spec}(C_i^{-1}A_i) = \{\alpha_i\}$. Hence, whenever $i \neq j$, the equation $XA_j = C_i^{-1}A_iX$ has no nontrivial solutions. Therefore, $B_{ij} = 0$ if $i \neq j$, and B has the desired block diagonal form.

Conversely, if A and B have the given block diagonal forms, then $C = C_1 \oplus C_2 \oplus \cdots \oplus C_t$ where $C_i = A_iB_iA_i^{-1}B_i^{-1}$ and $A_iC_i = C_iA_i$. Theorem 2 then tells us $\mathrm{spec}(C_i) = \{1\}$ for each i, and hence $\mathrm{spec}(C) = \{1\}$. □

The following examples are instructive.

Example 1.   Let $A = \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$ and $B = \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$. Then $C = -I$ so A and B both commute with C. However, $AB \neq BA$. Observe that B interchanges the two eigenspaces of A. Also, A and B are unitary, but $0 \in F(B)$.

Example 2.   Let $A = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$ and $B = \left(\begin{smallmatrix} b_1 & x \\ 0 & b_2 \end{smallmatrix}\right)$. Then C will have the form $\left(\begin{smallmatrix} 1 & y \\ 0 & 1 \end{smallmatrix}\right)$, for some y, and $AC = CA$. Whenever $b_1 \neq b_2$, the matrices A and B will not commute.

3. Generalizations of Theorem B.

The method used in [16] to prove Theorem B serves to establish the following generalization.

Theorem 4.   *Assume A is diagonalizable. Then $AC = CA$ if and only if there is a simultaneous similarity which puts A in the form $\alpha_1 I_{n_1} \oplus \alpha_2 I_{n_2} \oplus \cdots \oplus \alpha_t I_{n_t}$ and B in the form $P(W_1 \oplus W_2 \oplus \cdots \oplus W_t)$, where P is an $n \times n$ permutation matrix and W_i is $n_i \times n_i$. We also have $C = APA^{-1}P^{-1}$.*

Proof. Suppose $AC = CA$. Then we may assume $A = \alpha_i I_{n_1} \oplus \alpha_2 I_{n_2} \oplus \cdots \oplus \alpha_t I_{n_t}$ and $C = C_1 \oplus C_2 \oplus \cdots \oplus C_t$. Partition B conformally. From $C = ABA^{-1}B^{-1}$ we get $A^{-1}C = BA^{-1}B^{-1}$. So the matrix $A^{-1}C = (\alpha_1)^{-1}C_1 \oplus (\alpha_2)^{-1}C_2 \oplus \cdots \oplus (\alpha_t)^{-1}C_t$ is similar to the diagonal matrix A^{-1}. Therefore, each C_i can be diagonalized. Let S_i be a similarity that diagonalizes C_i, then $S = S_1 \oplus S_2 \oplus \cdots \oplus S_t$ will diagonalize C and $S^{-1}AS = A$. So, without loss of generality, we may assume $A = \alpha_1 I_{n_1} \oplus \alpha_2 I_{n_2} \oplus \cdots \oplus \alpha_t I_{n_t}$ and C is diagonal. Invert both sides of $A^{-1}C = BA^{-1}B^{-1}$ to get $C^{-1}A = BAB^{-1}$. The diagonal matrix $C^{-1}A$ is then similar to the diagonal matrix A. Hence, there is a permutation matrix P such that $C^{-1}A = PAP^{-1}$. So, we have $PAP^{-1} = C^{-1}A = BAB^{-1}$, which gives $(P^{-1}B)A = A(P^{-1}B)$. Since $P^{-1}B$ commutes with A, we must have $P^{-1}B = W_1 \oplus W_2 \oplus \cdots \oplus W_t$, where W_i is $n_i \times n_i$.

Conversely, suppose $A = \alpha_1 I_{n_1} \oplus \alpha_2 I_{n_2} \cdots \oplus \alpha_t I_{n_t}$ and $B = PW$ where $W = W_1 \oplus W_2 \oplus \cdots W_t$. Then $C = ABA^{-1}B^{-1} = APWA^{-1}W^{-1}P^{-1} =$

$APA^{-1}P^{-1}$. Now $PA^{-1}P^{-1}$ is diagonal, because the effect of the permutation similarity P is to rearrange the diagonal entries of A^{-1}, so C is diagonal and thus commutes with A. $\square$

Remark. If A has n distinct eigenvalues ($t = n$), then W will be diagonal, so B will be the product of a permutation matrix and a diagonal matrix (sometimes called a monomial matrix).

In Theorem 4, the equation $C = APA^{-1}P^{-1}$ shows us how the eigenvalues of C are quotients of eigenvalues of A, as we saw in Theorem 2. When $\mathrm{spec}(C) = \{1\}$, we get a stronger conclusion.

Theorem 5. *Assume A is diagonalizable and $AC = CA$. Then $AB = BA$ if and only if $\mathrm{spec}(C) = \{1\}$.*

Proof. If $AB = BA$, then $C = I$, so $\mathrm{spec}(C) = \{1\}$.

Conversely, suppose $\mathrm{spec}(C) = \{1\}$. Applying a similarity, we may assume A and B are in the block diagonal form given by Theorem 3. However, since A is diagonalizable, each diagonal block A_i is actually the scalar matrix $\alpha_i I_{n_i}$. (Equivalently, each generalized eigenspace V_i is actually an eigenspace.) So, we get $AB = BA$. $\square$

If the generalized eigenspaces are orthogonal to each other, then we can get A into the block diagonal form with a unitary similarity; in this case the hypothesis $0 \notin F(B)$ does tell us something.

Theorem 6. *Suppose $AC = CA$ and there is a unitary matrix U such that $U^*AU = A_1 \oplus A_2 \oplus \cdots \oplus A_t$ where A_i is $n_i \times n_i$ and $\mathrm{spec}(A_i) = \{\alpha_i\}$. Then if $0 \notin F(B)$, we have $\mathrm{spec}(C) = 1$.*

Proof. As in the proof of Theorem 3, but using a unitary similarity, we may assume $A = A_1 \oplus A_2 \oplus \cdots \oplus A_t$ with $\mathrm{spec}(A_i) = \{\alpha_i\}$ and $C = C_1 \oplus C_2 \oplus \cdots \oplus C_t$. Partition B conformally and use $CBA = AB$ to get $C_iB_{ii}A_i = A_iB_{ii}$. Since $0 \notin F(B)$, we know B_{ii} is nonsingular, so $C_i = A_iB_{ii}A_i^{-1}B_{ii}^{-1}$. But $C_iA_i = A_iC_i$ and $\mathrm{spec}(A_i) = \{\alpha_i\}$ so Theorem 2 tells us $\mathrm{spec}(C) = \{1\}$. $\square$

Remark. Theorem 1 now follows from Theorems 5 and 6.

Recall that a matrix is said to be *nonderogatory* if its minimal polynomial equals its characteristic polynomial. This is equivalent to saying that the Jordan canonical form for the matrix has exactly one Jordan block for each eigenvalue; it is also equivalent to saying that each eigenvalue has a one dimensional eigenspace. When A is nonderogatory, we obtain results in Theorems 7 and 8 which are analogous to Theorems 4 and 5.

Lemma. *Suppose* $A = \begin{pmatrix} A_1 & X \\ 0 & A_2 \end{pmatrix}$ *where the diagonal blocks* A_1 *and* A_2 *are square of sizes* k *and* $n - k$, *respectively, and* X *is* $k \times (n - k)$, *while* 0 *represents an* $(n - k) \times k$ *block of zeroes. If* A *is nonderogatory, then so are* A_1 *and* A_2.

Proof. Let $m_i(x)$ be the minimal polynomial of A_i for $i = 1, 2$. Then $\deg(m_1) \leq k$ and $\deg(m_2) \leq n - k$. A computation shows $m_1(A)m_2(A) = 0$, and hence A satisfies the polynomial $m_1(x)m_2(x)$. Since A is nonderogatory, its minimal polynomial has degree n, so we must have $\deg(m_1) = k$ and $\deg(m_2) = n - k$. Hence, each block A_i is nonderogatory. □

Definition. A square matrix which has exactly one nonzero entry in each row and column is called a *monomial* matrix. Equivalently, a monomial matrix is a matrix which is the product of a permutation matrix and a nonsingular diagonal matrix.

Note that the inverse of a monomial matrix is also monomial.

Theorem 7. *Suppose* $AC = CA$ *and* A *is nonderogatory. There is a simultaneous similarity which puts* A *into triangular form and* B *into the form* PT, *where* P *is a permutation matrix and* T *is upper triangular.*

Proof. Let $\text{spec}(A) = \{\alpha_1, \alpha_2, \ldots, \alpha_t\}$ and assume $n_1 \geq n_2 \geq \cdots \geq n_t$. Let x_i be an eigenvector for α_i. Since A is nonderogatory, x_i will span the eigenspace for α_i. Hence, since $AC = CA$, the vector x_i is also an eigenvector of C. Let $Cx_i = \gamma_i x_i$. Let V be the subspace spanned by $\{x_1, x_2, \ldots, x_t\}$. From $C = ABA^{-1}B^{-1}$ we get $AB^{-1}A^{-1}C = B^{-1}$, so $A(B^{-1}x_i) = (\alpha_i/\gamma_i)(B^{-1}x_i)$. So $B^{-1}x_i$ is an eigenvector of A and thus must be a multiple of x_j for some j. The space V is then invariant under B^{-1}, and the matrix for the action of B^{-1} on V, with respect to the basis $\{x_1, x_2, \ldots, x_t\}$, will be monomial. But then V is also invariant under B, and the matrix for the action of B on V is also monomial. Now choose a basis for $\mathbb{C}^n$ in which the first t elements are $x_1, x_2, \ldots, x_t$ and let S be the similarity transformation corresponding to this new basis. Then

$$S^{-1}AS = \begin{pmatrix} D_1 & A_{12} \\ 0 & A_2 \end{pmatrix} \quad \text{and} \quad S^{-1}BS = \begin{pmatrix} M_1 & B_{12} \\ 0 & B_2 \end{pmatrix}$$

where $D_1 = \text{diag}\,(\alpha_1, \alpha_2, \cdots, \alpha_t)$, the matrix M_1 is a monomial matrix of size t, and A_2 and B_2 are square blocks of size $n - t$. Since A is nonderogatory, so is A_2. We have $S^{-1}CS = \begin{pmatrix} C_1 & C_{12} \\ 0 & C_2 \end{pmatrix}$ where $C_2 = A_2 B_2 A_2^{-1} B_2^{-1}$. Now $AC = CA$ implies $A_2 C_2 = C_2 A_2$ so the argument may be repeated on A_2 and B_2. After n_1 applications of the process, we will have a simultaneous

similarity which puts A into triangular form and B into a block triangular form with diagonal blocks $M_1, M_2, \ldots M_n$ which are monomial matrices. In fact, if $r_1, r_2, \ldots, r_{n_1}$ is the conjugate partition of $n_1, n_2, \ldots, n_t$, then M_i has size r_i and the corresponding diagonal block of A is the diagonal matrix $\text{diag}(\alpha_1, \alpha_2, \ldots \alpha_{r_i})$. (Note $r_1 = t$.) Let $M_i = P_i E_i$ where P_i is a permutation matrix of size r_i and E_i is a diagonal matrix size of r_i. Let $P = P_1 \oplus P_2 \oplus \cdots \oplus P_t$. The effect of multiplying a matrix on the left by P is to rearrange the first r_1 rows by the permutation P_1, the next r_2 rows by the permutation P_2 and so on. Hence, B can be factored as PT, where T is triangular with the diagonal matrix E_i in the i'th diagonal block. □

Theorem 8. *Suppose $AC = CA$ and A is nonderogatory. Then A and B can be simultaneously triangularized if and only if $\text{spec}(C) = \{1\}$.*

Proof. If A and B are both upper triangular, then C is also upper triangular with each diagonal entry equal to 1. So $\text{spec}(C) = \{1\}$.

Conversely, suppose $\text{spec}(C) = \{1\}$. Then, in the proof of Theorem 7, we have $\gamma_i = 1$ for each i, and thus $A(B^{-1}x_i) = (\alpha_i)(B^{-1}x_i)$. So $B^{-1}x_i$ is a multiple of x_i and the matrix M_1 will actually be a diagonal matrix. Using an induction argument shows that each block M_i is a diagonal matrix, and hence the permutation matrix P of Theorem 7 is just the identity. □

A natural question is whether the hypothesis that A is nonderogatory is necessary in Theorems 7 and 8. Here is an example of two 4×4 matrices A and B such that $AC = CA$, and $\text{spec}(C) = \{1\}$, but A and B do not have a common eigenvector. Hence, they cannot be simultaneously triangularized.

Example 3. Let

$$A = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 2 & 0 & 0 \\ 0 & 0 & 1 & 2 \end{pmatrix}.$$

Let $R = BAB^{-1}$; we have observed that A commutes with C if and only if A commutes with R. By direct calculation, we get

$$R = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad \text{and} \quad AR = RA.$$

Any eigenvector of A has the form $(x, y, 0, 0)$, but no vector of this form is an eigenvector of B. Hence, A and B have no common eigenvector. However, $AC = CA$ and from Theorem 2 we know $\text{spec}(C) = \{1\}$.

The referee has pointed out that Theorem 8 can be proved independently of Theorem 7 and can be generalized. Note that A is nonderogatory if and only if the centralizer of A consists of polynomials in A. Hence, Theorem 8 is a special case of the following result (statement and proof by the referee).

Theorem 8′. *Let A and B be invertible. If $C = ABA^{-1}B^{-1}$ is a polynomial in A, then A and B are simultaneously triangularizable if and only if C is unipotent.*

Proof. The forward implication is trivial (and holds without the assumption on C). For the converse, let H be the unit group of the ring $F[A]$, the ring of polyonimals in A over the field F. Let H_1 be the set of unipotent elements in H; since H can be triangularized, the set H_1 is a subgroup of H. Since $B^{-1}AB = C^{-1}A$, we see $B^{-1}AB$ is a polynomial in A and hence B normalizes the group H and the subgroup H_1. Let G be the subgroup of $GL_n(F)$ generated by H_1, A and B. Note that H_1 is normal in G and, since $C \in H_1$, the group G/H_1 is abelian. Hence, the commutator subgroup G' of G is unipotent. This is a necessary and sufficient condition for a group to be triangularizable [**11**, **15**, **30**]. □

4. The additive commutator.

In this section, A and B denote $n \times n$ matrices (not necessarily nonsingular) and $[A, B] = AB - BA$ is their additive commutator. We can use the methods of the previous section to generalize the finite dimensional case of the following theorem of Putnam [**17**]. Putnam's theorem was for a pair of bounded operators on a Hilbert space, but we shall only deal with the finite dimensional cae.

Theorem (Putnam, [**20**])**.** *If A is normal and A commutes with $[A, B]$, then A commutes with B.*

In Theorem 10, we will show that "A is normal" can be weakened to "A is diagonalizable". First, a result for general A.

Theorem 9. *If $\mathrm{spec}(A) = \{\alpha_1, \alpha_2, \dots, \alpha_t\}$ and A commutes with $[A, B]$, then there is a nonsingular matrix S such that $S^{-1}AS = A_1 \oplus A_2 \oplus \cdots \oplus A_t$, where A_i is $n_i \times n_i$ and $\mathrm{spec}(A_i) = \{\alpha_i\}$, and $S^{-1}BS = B_1 \oplus B_2 \oplus \cdots \oplus B_t$ where B_i is $n_i \times n_i$. Equivalently, each of the generalized eigenspaces of A is invariant under B.*

Proof. Applying a similarity to A, B, and $[A, B]$, we may assume that A is in the desired form. Then, since A commutes with $[A, B]$, we must have

$[A, B] = R_1 \oplus R_2 \oplus \cdots \oplus R_t$ where R_i is $n_i \times n_i$. Partition B conformally. Since $AB - BA$ is block diagonal, when $i \neq j$ we must have $A_i B_{ij} = B_{ij} A_j$. Since A_i and A_j have no common eigenvalue, $B_{ij} = 0$ when $i \neq j$ and hence B is in the desired block diagonal form. $\qquad \square$

Theorem 10. *If A is diagonalizable and A commutes with $[A, B]$, then A commutes with B.*

Proof. Applying a similarity to A, B, and $[A, B]$, we may assume A and B have the forms given in Theorem 9. However, since A is diagonalizable, we can actually assume each diagonal block A_i is a scalar matrix, and so we get $AB = BA$. $\qquad \square$

Our next result deals with the case where A is nonderogatory. The proof uses the Jacobson Lemma [**14**].

Jacobson Lemma. *If a and b are elements of a finite dimensional algebra over a field of characteristic zero, and a commutes with $[a, b]$, then $[a, b]$ is nilpotent.*

Theorem 11. *If A is nonderogatory and A commutes with $[A, B]$, then A and B can be simultaneously triangularized.*

Proof. Let α be an eigenvalue of A. Since A is nonderogatory, the eigenspace for α is one dimensional. Let x span this eigenspace. Since A commutes with $[A, B]$, this eigenspace is invariant under $[A, B]$ and hence x is also an eigenvector of $[A, B]$. But the Jacobson Lemma tells us that $[A, B]$ is nilpotent. Hence every eigenvalue of $[A, B]$ is zero and $[A, B]x = 0$. So $A(Bx) = BAx = \alpha Bx$ and Bx must be a multiple of x. Thus, x is an eigenvector of B. Using x as the first vector in a basis for C^n, the corresponding similarity will put A and B into the forms

$$A \longrightarrow \begin{pmatrix} \alpha & * \; * \cdots * \\ \hline 0 & \\ 0 & \widetilde{A} \\ 0 & \end{pmatrix} \qquad B \longrightarrow \begin{pmatrix} \beta & * \; * \cdots * \\ \hline 0 & \\ 0 & \widetilde{B} \\ 0 & \end{pmatrix}.$$

Now $\widetilde{A}$ will be nonderogatory and will commute with $[\widetilde{A}, \widetilde{B}]$, so we may repeat the process until A and B have been triangularized. $\qquad \square$

The following example shows that the assumption that A is nonderogatory is needed in Theorem 11.

Example 4.   Let $A = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ and let B be any 3×3 matrix such that $b_{31} = 0$. A calculation shows that $[A, B]$ is strictly upper triangular and commutes

with A. However, if b_{21} and b_{32} are both nonzero, then A and B have no common eigenvector, and hence cannot be simultaneously triangularized. Note that A is not nonderogatory; its minimal polynomial is x^2.

The referee has pointed out that Theorem 11, like Theorem 8, can be generalized as follows:

Theorem 11′. *Let F be an algebraically closed field of characteristic 0 and let $A, B \in M_n(F)$. Suppose $[A, B] = f(A)$ for some polynomial F. Then A and B can be simultaneously triangularized.*

Proof. It follows by induction on i that $[A^i, B] = f_i(A)$ for some polynomial f_i. $(A^i B = A(BA^{i-1} + f_{i-1}(A)) = BA^i + f(A)A^{i-1} + Af_{i-1}(A).)$ Thus, the Lie algebra generated by A and B is contained in the Lie algebra $L :=$ $F[A] + F \cdot B$ and $F[A]$ is an abelian ideal of L. Hence, the derived algebra of L is abelian and L is solvable. By Lie's theorem, L is triangularizable.   □

If both A and B commute with $[A, B]$, then A and B are said to be quasi-commutative matrices. Results of McCoy in [**17**] and [**18**] show that quasi-commutative matrices can be simultaneously triangularized. Drazin generalized this result in [**4**]. See also [**5**] and [**24**]. The analogous problem for the multiplicative commutator has been analyzed by R.C. Thompson [**27**, **28**, **29**] and by I. Sinha, [**24**].

References

[1]   S.K. Berberian, *A note on operators unitarily equivalent to their adjoints*, J. London Math. Soc., **37** (1962), 403-404.

[2]   C.W. Curtis and I. Reiner, *Representation Theory of Finite Groups and Associative Algebras*, Interscience Publishers, John Wiley and Sons, (1962), 258-262.

[3]   C.R. DePrima and B.K. Richard, *An extension of the Frobenius group commutator theorem*, Indiana University Mathematics Journal, **23** (1973), 251-255.

[4]   M.P. Drazin, *Some generalizations of matrix commutativity*, Proc. London Math. Soc., **1**(**3**) (1951), 222-231.

[5]   M.P. Drazin, J.W. Dungey and K.W. Gruenberg, *Some theorems on commutative matrices*, J. London Math. Soc., **26** (1951), 221-228.

[6] G.A. Elliott, *An application of the holomorphic functional calculus to sharpen a classical matrix commutator theorem*, Indiana University Mathematics Journal, **20** (1970), 407-409.

[7] G. Frobenius, *Über vertauschbare Matrizen*, Sitzungsberichte der Königlich Preussischen Akademie der Wissenschaften zu Berlin, (1896), 601-614.

[8] ———, *Über der von L. Bieberbach gefundenen Beweis eines Satzes von C. Jordan*, Sitzungsber. Akad. Wiss. Berlin, (1911), 241-248.

[9] ———, *Über unitäre Matrizen*, Sitzungsberichte der Königlich Preussischen Akademie der Wissenschaften zu Berlin, (1911), 373-378.

[10] F.R. Gantmacher, *The Theory of Matrices*, Vol. I, Chelsea, New York, 1959, 220.

[11] R.M. Guralnick, *Triangularization of sets of matrices*, Linear and Multilinear Algebra, **9** (1980), 133-140.

[12] R. Guralnick and Olga Taussky, *A remark concerning unipotent matrix groups*, Linear and Multilinear Algebra, **7** (1979), 87-89.

[13] R. Horn and C. Johnson, *Topics in Matrix Analysis*, Cambridge University Press, Cambridge, (1991), 270.

[14] Nathan Jacobson, *Rational methods in the theory of Lie algebras*, Annals of Mathematics, **36** (1935), 875-881.

[15] E.R. Kolchin, *On certain concepts in the theory of algebraic matrix groups*, Ann. of Math., **49** (1948), 774-789.

[16] M. Marcus and R.C. Thompson, *On a classical commutator result*, Journal of Mathematics and Mechanics, **16** (1966), 583-588.

[17] N.H. McCoy, *On quasi-commutative matrices*, Trans. Amer. Math. Soc., **36** (1934), 327-340.

[18] ———, *On the characteristic roots of matrix polynomials*, Bull. Amer. Math. Soc., **42** (1936), 592-600.

[19] C.R. Putnam, *Group commutators of bounded operators in Hilbert space*, Michigan Mathematical Journal, **7** (1960), 229-232.

[20] ———, *On normal operators in Hilbert space*, American Journal of Mathematics, **73** (1971), 357-362.

[21] I. Sinha, *Commutators in nipotent linear groups of class two*, J. London Math. Soc., **41** (1966), 353-363.

[22] A. Speiser, *Theorie der Gruppen von Endlicher Ordnung*, Julius Springer, Berlin, (1927), 215-219.

[23] J.J. Sylvester, *Sur l'equation en matrices px = xq*, C.R. Acad. Sci. Paris, **99** (1884), 67-71, 115-116.

[24] Olga Taussky, *Commutativity in finite matrices*, American Mathematical Monthly, **45** (1957), 229-235.

[25] ———, *Commutators of unitary matrices which commute with one factor*, Journal of Mathematics and Mechanics, **10** (1961), 175-178.

[26] ———, *How I became a torchbearer for matrix theory*, American Mathematical Monthly, **95** (1988), 801-812.

[27] R.C. Thompson, *Multiplicative matrix commutators commuting with both factors*, Journal of Mathematical Analysis and Applications, **18** (1967), 315-335.

[28] ———, *A note on matrix commutators*, Amer. Math. Monthly, **74** (1967), 276-278.

[29] ———, *Some matrix factorization theorems*, Pacific J. of Math., **33** (1970), 763-820.

[30] D.B. Wales and H.J. Zassenhaus, *On L-Groups*, Mathematische Annalen, **198** (1972), 1-12.

[31] H. Wielandt, *Review of "Commutators of unitary matrices that commute with one factor"*, by O. Taussky, Zentralblatt für Mathematik und ihre Grenzgebiete, **96** (1962), 9.

[32] H. Zassenhaus, *A remark on a paper of O. Taussky*, Journal of Mathematics and Mechanics, **10** (1961), 179-180.

SWARTHMORE COLLEGE
SWARTHMORE, PA 19081
E-mail address: hshapir1@cc.swarthmore.edu