Linear and Multilinear Algebra

ISSN: 0308-1087 (Print) 1563-5139 (Online) Journal homepage: http://www.tandfonline.com/loi/glma20

On simultaneous reduction of families of matrices to triangular or diagonal form by unitary congruences

Yoo Pyo Hong & Roger A. Horn

To cite this article: Yoo Pyo Hong & Roger A. Horn (1985) On simultaneous reduction of families of matrices to triangular or diagonal form by unitary congruences, Linear and Multilinear Algebra, 17:3-4, 271-288, DOI: 10.1080/03081088508817659

To link to this article: https://doi.org/10.1080/03081088508817659

Published online: 02 Apr 2008.

Submit your article to this journal

Article views: 43

View related articles

Citing articles: 10 View citing articles

Full Terms & Conditions of access and use can be found at
http://www.tandfonline.com/action/journalInformation?journalCode=glma20

Linear and Multilinear Algebra, 1985, Vol. 17, pp. 271–288
0308-1087/85/1704-0271 $20.00
© 1985 Gordon and Breach Science Publishers, Inc. and OPA Ltd.
Printed in the United States of America

On Simultaneous Reduction of Families of Matrices to Triangular or Diagonal Form by Unitary Congruences

YOO PYO HONG AND ROGER A. HORN
Department of Mathematical Sciences, The Johns Hopkins University, Baltimore, MD 21218

(Received May 14, 1984; in final form October 3, 1984)

Let $\{A_i\}$ and $\{B_i\}$ be two given families of n-by-n matrices. We give conditions under which there is a unitary U such that every matrix UA_iU^T is upper triangular. We give conditions, weaker than the classical conditions of commutativity of the whole family, under which there is a unitary U such that every matrix UB_jU^* is upper triangular. We also give conditions under which there is one single unitary U such that every UA_iU^T and every UB_jU^* is upper triangular. We give necessary and sufficient conditions for simultaneous unitary reduction to diagonal form in this way when all the A_i's are complex symmetric and all the B_j's are Hermitian.

1. INTRODUCTION

Motivated by the Grunsky inequalities from the theory of univalent analytic functions in the unit disc, we considered recently [3] the problem of simultaneous reduction of a pair of n-by-n complex matrices A (Hermitian) and B (symmetric) to diagonal form by nonsingular congruences that preserve the type of the matrix, i.e. $A \to SAS^*$ and $B \to SBS^T$. We also considered similar problems when A and B are both Hermitian or both symmetric.

In this paper, we consider unitary congruences and generalize our previous results in two ways: to families with an arbitrary number of matrices, and to the problem of simultaneous reduction to triangular form of families of matrices that need not be either Hermitian or symmetric.

2. NOTATION, DEFINITIONS AND BASIC CONCEPTS

We denote by M_n the set of n-by-n complex matrices. Matrices $A, B \in M_n$ are said to be *consimilar* if there is a nonsingular $R \in M_n$ such that $A = RB\bar{R}^{-1}$ [2]. Notice that a real consimilarity is just ordinary similarity ($A = SB\bar{S}^{-1} = SBS^{-1}$ if S is real); a unitary consimilarity is one type of congruence ($A = UB\bar{U}^{-1} = UBU^T$ if U is unitary); and a complex orthogonal consimilarity is another type of congruence, often called a conjunctivity ($A = QB\bar{Q}^{-1} = QBQ^*$ if Q is complex orthogonal, i.e. $Q^{-1} = Q^T$). The notion of consimilarity can be generalized by replacing the complex field with an arbitrary field F and replacing the operation of complex conjugation by an automorphism on F [6, p. 27].

A matrix $A \in M_n$ is said to be *contriangularizable* (respectively, *condiagonalizable*) if A is consimilar to an upper triangular (respectively, diagonal) matrix. A family of matrices $\mathscr{F} = \{A_i : i \in \mathscr{I}\} \subset M_n$ is said to be *simultaneously contriangularizable* (respectively, *simultaneously condiagonalizable*) if there is one nonsingular $R \in M_n$ such that $RA_i\bar{R}^{-1}$ is upper triangular (respectively, diagonal) for all $i \in \mathscr{I}$. If R can be chosen to be unitary, we say that $\mathscr{F}$ is *simultaneously unitarily contriangularizable* or *condiagonalizable*. The family $\mathscr{F}$ is said to be *concommuting* if $A_i\bar{A}_j = A_j\bar{A}_i$ for all $i, j \in \mathscr{I}$.

A subspace $S \subset \mathbb{C}^n$ is said to be *invariant* under $A \in M_n$ if $Ax \in S$ for every $x \in S$; it is *coninvariant* under A if $A\bar{x} \in S$ for every $x \in S$. If $\mathscr{F} = \{A_i : i \in \mathscr{I}\} \subset M_n$ is a given family of matrices, the subspace S is said to be $\mathscr{F}$*-invariant* (respectively, $\mathscr{F}$*-coninvariant*) if S is invariant (respectively, coninvariant) under A_i for all $i \in \mathscr{I}$.

A nonzero vector $x \in \mathbb{C}^n$ such that $A\bar{x} = \lambda x$ is said to be a *coneigenvector* of $A \in M_n$; the scalar λ is a *coneigenvalue* of A. If $A\bar{x} = \lambda x$, notice that $e^{-i\theta}A\bar{x} = A(\overline{e^{i\theta}x}) = (e^{-2i\theta}\lambda)(e^{i\theta}x)$ for all $\theta \in \mathbb{R}$, so a matrix with a nonzero coneigenvalue has infinitely many distinct coneigenvalues. Moreover, $A\bar{A}x = A(\overline{A\bar{x}}) = A(\overline{\lambda x}) = \bar{\lambda}A\bar{x} = \bar{\lambda}\lambda x = |\lambda|^2 x$, so $|\lambda|^2$ is necessarily an eigenvalue of the matrix $A\bar{A}$. If $A\bar{A}$ has no nonnegative eigenvalues, then A can have no coneigenvalues.

LEMMA 2.1 *If S is a subspace of $\mathbb{C}^n$ that is coninvariant under $A, B \in M_n$, then S is invariant under $A\bar{B}$.*

Proof If $x \in S$, then $y = B\bar{x} \in S$ and $A\bar{B}x = A(\overline{B\bar{x}}) \in S$. ∎

The converse of the assertion in the lemma is false, as may be seen from considering the example $B = A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, for which $A\bar{A} = -I$. The linear span of the vector $e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ is a subspace of $\mathbb{C}^2$ that is invariant under $A\bar{A}$ but is not coninvariant under A.

LEMMA 2.2 *Let $A \in M_n$ be given, and suppose $S \subset \mathbb{C}^n$ is a nonzero subspace that is coninvariant under A. If $A\bar{A}$ has an eigenvector in S whose associated eigenvalue is nonnegative, then S contains at least one coneigenvector of A.*

Proof By Lemma 2.1, S is invariant under $A\bar{A}$, so there is at least one eigenvector of $A\bar{A}$ in S. If $A\bar{A}x = \lambda x$ with $x \neq 0$, $x \in S$, and $\lambda \geqslant 0$, notice that $A\bar{A}(A\bar{x}) = A(\overline{A\bar{A}x}) = A(\lambda\bar{x}) = \lambda(A\bar{x})$. There are two possibilities: $A\bar{x}$ and x are dependent, or they are independent. In the first case, $A\bar{x} = \mu x$ for some $\mu \in \mathbb{C}$, in which case x is a coneigenvector of A. In the second case, $A\bar{x} + \mu x \neq 0$ for all $\mu \in \mathbb{C}$. But then for any μ such that $|\mu|^2 = \lambda$ we have $A(\overline{A\bar{x} + \mu x}) = A\bar{A}x + \bar{\mu}A\bar{x} = \lambda x + \bar{\mu}A\bar{x} = \mu\bar{\mu}x + \bar{\mu}A\bar{x} = \bar{\mu}(A\bar{x} + \mu x)$, so $A\bar{x} + \mu x$ is a coneigenvector of A. ∎

If $A\bar{A}$ has only nonnegative eigenvalues, the main hypothesis of the lemma is automatically satisfied. This is the case, for example, if A is symmetric (since $A\bar{A} = AA^*$ is positive semi-definite) or if A is Hermitian and positive definite (since $A\bar{A}$ is similar to $A^{1/2}\bar{A}A^{1/2}$, which is congruent to the positive definite matrix $\bar{A}$).

3. SIMULTANEOUS UNITARY CONTRIANGULARIZATION OF A FAMILY OF MATRICES

Our first objective is to establish general sufficient conditions for a family of matrices to be simultaneously unitarily contriangularizable. We then specialize to obtain a simple necessary and sufficient condition in the case of a family of complex symmetric matrices. The simultaneous contriangularization theorem is an immediate consequence of a sufficient condition for a family of matrices to have a common coneigenvector.

LEMMA 3.1 *Let $\mathscr{F} = \{A_i: i \in \mathscr{I}\} \subset M_n$ be a given family of matrices, let $\mathscr{G} = \{A_i\bar{A}_j: i, j \in \mathscr{I}\}$, and assume that*

(a) *$\mathscr{G}$ is a commuting family,*

(b) *for all $i \in \mathscr{I}$, $A_i\bar{A}_i$ has only nonnegative eigenvalues, and*

c) *whenever $i, j \in \mathscr{I}$ and $x \neq 0$ is a vector such that $A_i\bar{A}_j x = \lambda x$ and $A_j\bar{A}_i x = \mu x$, then $|\lambda| = |\mu|$.*

Then every nonzero $\mathscr{F}$-coninvariant subspace S of $\mathbb{C}^n$ contains a common coneigenvector for $\mathscr{F}$, i.e. there is some nonzero $x \in S$ such that $A_i\bar{x} = \lambda_i x$ for all $i \in \mathscr{I}$.

Proof Let $S \subset \mathbb{C}^n$ be a nonzero subspace that is $\mathscr{F}$-coninvariant. Lemma 2.1 guarantees that S is $\mathscr{G}$-invariant, so there is a common (nonzero) eigenvector $x \in S$ for $\mathscr{G}$. Suppose $A_i\bar{A}_j x = \lambda_{ij} x$ for all $i, j \in \mathscr{I}$, so that $|\lambda_{ij}| = |\lambda_{ji}|$ for all $i, j \in \mathscr{I}$ by (c).

There are two cases to consider:

(1) $A_i\bar{A}_i x = 0$ for all $i \in \mathscr{I}$, or

(2) $A_0\bar{A}_0 x \neq 0$ for some $A_0 \in \mathscr{F}$.

Consider case (1) first. There are two possibilities:

(1′) $A_i\bar{x} = 0$ for all $i \in \mathscr{I}$, or

(1″) $A_0\bar{x} \neq 0$ for some $A_0 \in \mathscr{F}$.

In case (1′), x is a common coneigenvector for $\mathscr{F}$ and we are done. In case (1″), use the commutativity of $\mathscr{G}$ to compute

$$0 = A_0\bar{A}_i A_i\bar{A}_0 A_0\bar{x} = A_0\bar{A}_0 A_0\bar{A}_i A_i\bar{x} = A_0\bar{A}_i A_0\bar{A}_0 A_i\bar{x}$$
$$= A_0\bar{\lambda}_{i0}\bar{\lambda}_{0i}\bar{x} = \bar{\lambda}_{i0}\bar{\lambda}_{0i}(A_0\bar{x})$$

which implies that $\lambda_{i0} = \lambda_{0i} = 0$ since $|\lambda_{i0}| = |\lambda_{0i}|$. Thus, $A_i(\bar{A}_0 x) = 0$ for all $i \in \mathscr{I}$, and the nonzero vector $A_0\bar{x}$ is a common coneigenvector for $\mathscr{F}$.

The final case to consider is (2), so we assume that $A_0\bar{A}_0 x = \lambda x$ with $\lambda > 0$. Let $S' \equiv \{z \in S: A_0\bar{A}_0 z = \lambda z\} \neq \{0\}$, and let $\mathscr{G}_0 = \{A_0\bar{A}_i, A_i\bar{A}_0: i \in \mathscr{I}\}$. If $z \in S'$, $A_0\bar{A}_0(A_i\bar{A}_0)z = A_i\bar{A}_0 A_0\bar{A}_0 z = \lambda(A_i\bar{A}_0 z)$, so S' is invariant under $A_i\bar{A}_0$. A similar argument shows that S' is invariant under $A_0\bar{A}_i$, so S' is $\mathscr{G}_0$-invariant. Since $\mathscr{G}_0$ is commutative, there exists a common (nonzero) eigenvector $y \in S'$ for $\mathscr{G}_0$. Suppose $A_i\bar{A}_0 y = \mu_{i0} y$ and $A_0\bar{A}_i y = \mu_{0i} y$ for all $i \in \mathscr{I}$, so that $|\mu_{0i}| = |\mu_{i0}|$ by (c). Define

$$S_0 \equiv \bigcap_{i \in \mathscr{I}} \{\{z \in S': A_0\bar{A}_i z = \mu_{0i} z\} \cap \{z \in S': A_i\bar{A}_0 z = \mu_{i0} z\}\}$$

and notice that $S' \supset S_0 \supset \{y\} \neq \{0\}$.

We claim that S_0 is coninvariant under A_0. First observe that if $z \in S'$, we have $A_0\bar{A}_0(A_0\bar{z}) = A_0(\overline{A_0\bar{A}_0 z}) = A_0(\overline{\lambda z}) = \lambda A_0\bar{z}$, so $A_0\bar{z} \in S'$, i.e. S' is coninvariant under A_0. Next observe that if $z \in S'$, then

$$z = \frac{1}{\lambda} A_0\bar{A}_0 z = A_0\left(\frac{1}{\lambda}\bar{A}_0 z\right) = A_0\left(\overline{\frac{1}{\bar{\lambda}} A_0\bar{z}}\right),$$

i.e.

$$\text{for every } z \in S' \text{ there is some } w \in S' \text{ such that } z = A_0\bar{w}. \tag{3.2}$$

Using this fact, compute

$$A_0\bar{A}_0 A_i\bar{z} = A_0\bar{A}_0 A_i\bar{A}_0 w = (A_0\bar{A}_0)(A_i\bar{A}_0)w = (A_i\bar{A}_0)(A_0\bar{A}_0)w$$
$$= A_i\bar{A}_0 A_0(\overline{A_0\bar{w}}) = A_i\bar{A}_0 A_0\bar{z},$$

i.e.

$$(A_i\bar{A}_0 A_0)\bar{z} = (A_0\bar{A}_0 A_i)\bar{z} \qquad \text{for all } z \in S' \text{ and all } i \in \mathscr{I}. \tag{3.3}$$

Now let $z \in S'$ and use (3.2), (3.3), and (a) to compute

$$A_i\bar{A}_i A_0\bar{z} = A_i\bar{A}_i A_0\bar{A}_0 w = A_i(\overline{A_i\bar{A}_0 A_0\bar{w}}) = A_i(\overline{A_0\bar{A}_0 A_i\bar{w}})$$
$$= (A_i\bar{A}_0)(A_0\bar{A}_i)w = A_0\bar{A}_i A_i(\overline{A_0\bar{w}}) = A_0\bar{A}_i A_i\bar{z},$$

i.e.

$$(A_0\bar{A}_i A_i)\bar{z} = (A_i\bar{A}_i A_0)\bar{z} \qquad \text{for all } z \in S' \text{ and all } i \in \mathscr{I}. \tag{3.4}$$

Let $z \in S'$ again and use (3.2), (a), and (3.4) to compute

$$A_i\bar{A}_j A_j\bar{A}_i z = A_i\bar{A}_j A_j\bar{A}_i(A_0\bar{w}) = A_i\overline{(A_j\bar{A}_j)(A_i\bar{A}_0)}\bar{w} = A_i\bar{A}_i(A_0\bar{A}_j A_j)\bar{w}$$
$$= A_i\bar{A}_i(A_j\bar{A}_j A_0)\bar{w} = A_i\bar{A}_i A_j\bar{A}_j z,$$

i.e.

$$(A_i\bar{A}_j)(A_j\bar{A}_i)z = (A_i\bar{A}_i)(A_j\bar{A}_j)z \qquad \text{for all } z \in S' \text{ and all } i, j \in \mathscr{I}. \tag{3.5}$$

Now let $z \in S_0$ be given. Notice that $A_i\bar{A}_i A_0\bar{A}_0 z = \mu_{ii}\mu_{00}z$ and $A_i\bar{A}_0 A_0\bar{A}_i z = \mu_{i0}\mu_{0i}z$ by the definition of S_0, and $\mu_{ii}\mu_{00}z = A_i\bar{A}_i A_0\bar{A}_0 z = A_i\bar{A}_0 A_0\bar{A}_i z = \mu_{i0}\mu_{0i}z$ by (3.5). But μ_{ii} and μ_{00} are nonnegative by assumption (b), so $\mu_{i0}\mu_{0i} = \mu_{ii}\mu_{00}$ is real and nonnegative for all $i \in \mathscr{I}$. Since $|\mu_{i0}| = |\mu_{0i}|$, it follows that $\mu_{i0} = \bar{\mu}_{0i}$ for all $i \in \mathscr{I}$.

We can now show that S_0 is coninvariant under A_0, for if $z \in S_0$ is given, we have

$$A_0\bar{A}_i(A_0\bar{z}) = A_0(\overline{\bar{A}_i A_0\bar{z}}) = A_0\bar{\mu}_{i0}\bar{z} = \mu_{0i}A_0\bar{z},$$

and

$$A_i\bar{A}_0(A_0\bar{z}) = (A_i\bar{A}_0A_0)\bar{z} = (A_0\bar{A}_0A_i)\bar{z} = A_0\bar{\mu}_{0i}\bar{z} = \mu_{i0}A_0\bar{z}$$

where we have used the identity (3.3) in the second calculation. Together, these two identities show that S_0 is coninvariant under A_0.

Since $A_0\bar{A}_0S_0 = \lambda S_0$ with $\lambda > 0$, S_0 is certainly invariant under $A_0\bar{A}_0$. By Lemma 2.2, A_0 must have a (nonzero) coneigenvector $x_0 \in S_0$, $A_0\bar{x}_0 = \mu x_0$, and $A_0\bar{A}_0x_0 = |\mu|^2x_0 = \lambda x_0$, so $\mu \neq 0$. Then

$$A_i\bar{x}_0 = A_i\frac{1}{\bar{\mu}}\bar{A}_0x_0 = \frac{1}{\bar{\mu}}A_i\bar{A}_0x_0 = (\mu_{i0}/\bar{\mu})x \qquad \text{for all} \quad i \in \mathscr{I},$$

and this vector x_0 is a common coneigenvector for $\mathscr{F}$. ■

THEOREM 3.6 *Let $\mathscr{F} = \{A_i: i \in \mathscr{I}\} \subset M_n$ be a given family, let $\mathscr{G} = \{A_i\bar{A}_j: i,j \in \mathscr{I}\}$, and assume that $\mathscr{G}$ is a commuting family. There exists There exists a unitary $U \in M_n$ such that UA_iU^T is upper triangular for all $i \in \mathscr{I}$ if and only if*

(a) *For all $i \in \mathscr{I}$, $A_i\bar{A}_i$ has only nonnegative eigenvalues, and*
(b) *For all $i,j \in \mathscr{I}$, $A_i\bar{A}_j + A_j\bar{A}_i$ has only real eigenvalues and $A_i\bar{A}_j - A_j\bar{A}_i$ has only imaginary eigenvalues.*

Proof If there exists a unitary $U \in M_n$ such that $UA_iU^T = \Delta_i$ is upper triangular for all $i \in \mathscr{I}$, then $A_i\bar{A}_i = U^*\Delta_i\bar{\Delta}_iU$, and $\Delta_i\bar{\Delta}_i$ has nonnegative main diagonal entries, which are the eigenvalues of $A_i\bar{A}_i$. Moreover, $A_i\bar{A}_j \pm A_j\bar{A}_i = U^*(\Delta_i\bar{\Delta}_j \pm \Delta_j\bar{\Delta}_i)U$ has only real (+) or imaginary (−) eigenvalues because the respective main diagonal entries of $\Delta_i\bar{\Delta}_j$ are the conjugates of those of $\Delta_j\bar{\Delta}_i$. The conditions (a) and (b) are therefore necessary.

To show that they are also sufficient, claim that it suffices to show that conditions (a) and (b) imply that there is a common coneigenvector for all $A_i \in \mathscr{F}$. If $x \neq 0$ is such that $A_i\bar{x} = \lambda_ix$ for all $i \in \mathscr{I}$, let $U \in M_n$ be a unitary matrix with first column $x/\|x\|_2$. Then

$$U^*A_i\bar{U} = \left(\begin{array}{c|c} \lambda_i & * \\ \hline 0 & A_i' \end{array}\right), \qquad A_i' \in M_{n-1}, \qquad \text{for all} \quad i \in \mathscr{I}. \tag{3.7}$$

We can now proceed to reduce the family $\mathscr{F}' = \{A_i': i \in \mathscr{I}\} \subset M_{n-1}$ in the same way if $\mathscr{F}'$ inherits the two properties (a) and (b) and if $\mathscr{G}' = \{A_i'\bar{A}_j': i,j \in \mathscr{I}\}$ is a commuting family. But

$$A_i\bar{A}_j = U\begin{pmatrix} \lambda_i\bar{\lambda}_j & * \\ 0 & A_i'\bar{A}_j' \end{pmatrix}U^*,$$

so the eigenvalues of $A_i\bar{A}_i$ are just $|\lambda_i|^2$ together with the eigenvalues of $A_i'\bar{A}_i'$ which must, therefore, be nonnegative because of (a). Similarly,

$$A_i\bar{A}_j \pm A_j\bar{A}_i = U\begin{pmatrix} \lambda_i\bar{\lambda}_j \pm \lambda_j\bar{\lambda}_i & * \\ 0 & A_i'\bar{A}_j' \pm A_j'\bar{A}_i' \end{pmatrix}U^*,$$

so the eigenvalues of $A_i\bar{A}_j \pm A_j\bar{A}_i$ are just $\lambda_i\bar{\lambda}_j \pm \lambda_j\bar{\lambda}_i$ together with those of $A_i'\bar{A}_j' \pm A_j'\bar{A}_i'$. The eigenvalues must, therefore, be real $(+)$ or imaginary $(-)$ because of (b) and the fact that $\lambda_i\bar{\lambda}_j + \lambda_j\bar{\lambda}_i = 2\,\mathrm{Re}(\lambda_i\bar{\lambda}_j)$ is real and $\lambda_i\bar{\lambda}_j - \lambda_j\bar{\lambda}_i = 2i\,\mathrm{Im}(\lambda_i\bar{\lambda}_j)$ is imaginary. The commutativity of $\mathscr{G}'$ is easily verified.

Finally, observe that if x is a common eigenvector of both $A_i\bar{A}_j$ and $A_j\bar{A}_i$, and if $A_i\bar{A}_jx = \lambda x$ and $A_j\bar{A}_ix = \mu x$, then $(A_i\bar{A}_j \pm A_j\bar{A}_i)x = (\lambda \pm \mu)x$. Assumption (b) says that $\lambda + \mu$ is real and $\lambda - \mu$ is imaginary, so $\lambda = \bar{\mu}$. In particular, $|\lambda| = |\mu|$, and hence assumption (c) of Lemma 3.1 is met. Since the other assumptions of the lemma are also met, we can apply Lemma 3.1 to the $\mathscr{F}$-coninvariant subspace $S = \mathbb{C}^n$ to ensure the existence of the desired common coneigenvector of $\mathscr{F}$. ■

If the family $\mathscr{F}$ consists of a single matrix, the assumption about commutativity of $\mathscr{G}$ is vacuous, and hypothesis (b) is subsumed under (a). Thus, the condition that $A\bar{A}$ has all nonnegative eigenvalues is necessary and sufficient for A to be unitarily contriangularizable. A more direct proof of this result is in [2]. If A is symmetric, then $A\bar{A} = AA^*$ is positive semi-definite and hence every symmetric matrix is unitarily contriangularizable. But UAU^T is symmetric whenever A is symmetric, and a symmetric triangular matrix is diagonal, so we conclude that every symmetric matrix is unitarily condiagonalizable (and conversely). This is a classical result of Takagi [9], Siegel [8], and Schur [7].

We are also interested in simultaneous unitary condiagonalization of a family of symmetric matrices, and Theorem 3.6 provides a necessary and sufficient condition.

COROLLARY 3.8 *Let $\mathscr{F} = \{A_i : i \in \mathscr{I}\} \subset M_n$ be a given family of symmetric matrices, and let $\mathscr{G} = \{A_i\bar{A}_j : i, j \in \mathscr{I}\}$. There exists a unitary $U \in M_n$ such that UA_iU^T is diagonal for all $i \in \mathscr{I}$ if and only if $\mathscr{G}$ is a commuting family. In this event, every product $A_i\bar{A}_j$ is normal.*

Proof If there is a unitary $U \in M_n$ such that $A_i = U\Lambda_i U^T$ for all $i \in \mathscr{I}$ and if each Λ_i is diagonal, then each $A_i\bar{A}_j = U(\Lambda_i\bar{\Lambda}_j)U^*$, and $\mathscr{G}$ is a commutative family of normal matrices since it is simultaneously unitarily diagonalizable.

For the converse, observe that $(A_i\bar{A}_j)^* = A_j\bar{A}_i$. Thus, hypothesis (b) of Theorem 3.6 is satisfied because for any $C \in M_n$, $C + C^*$ is Hermitian and has only real eigenvalues while $C - C^*$ is skew-Hermitian and has only imaginary eigenvalues. Hypothesis (a) is satisfied because $A_i\bar{A}_i = A_iA_i^*$ is Hermitian and positive semi-definite. Thus, the theorem guarantees that there is a unitary $U \in M_n$ such that $UA_iU^T = \Delta_i$ is upper triangular for every $i \in \mathscr{I}$. But then $\Delta_i^T = UA_i^TU^T = UA_iU^T = \Delta_i$, so each Δ_i is actually diagonal. ■

A commuting family $\mathscr{F} = \{A_i: i \in \mathscr{I}\} \subset M_n(\mathbb{R})$ of real matrices with real eigenvalues is simultaneously upper triangularizable by a single real orthogonal similarity, i.e. a (real) unitary consimilarity. The following theorem is a simple generalization of this fact to simultaneous unitary contriangularization of a family of complex matrices. Its hypotheses are in part stronger than those of Theorem 3.6; the assumption that $\mathscr{F}$ is concommutative ($A_i\bar{A}_j = A_j\bar{A}_i$ for all $i, j \in \mathscr{I}$) implies that $\mathscr{G}$ is commutative. But it makes no assumption on the eigenvalues of $A_i\bar{A}_j \pm A_j\bar{A}_i$, so its hypotheses are in part weaker than those of Theorem 3.6.

THEOREM 3.9 *Let $\mathscr{F} = \{A_i: i \in \mathscr{I}\} \subset M_n$ and assume that $A_i\bar{A}_j = A_j\bar{A}_i$ for all $i, j \in \mathscr{I}$. There exists a unitary $U \in M_n$ such that UA_iU^T is upper triangular for all $i \in \mathscr{I}$ if and only if $A_i\bar{A}_i$ has only non-negative eigenvalues for all $i \in \mathscr{I}$. In this event, there is a unitary diagonal matrix $D \in M_n$ such that $V = DU$ is a unitary matrix for which VA_iV^T is upper triangular and has real main diagonal entries for all $i \in \mathscr{I}$.*

Proof The condition on the spectrum of $A_i\bar{A}_i$ is necessary as before. To show that it is sufficient, notice that the assumption that $\mathscr{F}$ is concommutative implies that $\mathscr{G} = \{A_i\bar{A}_j: i, j \in \mathscr{I}\}$ is commutative, so assumption (a) of Lemma 3.1 is satisfied. We are explicitly assuming (b) of the Lemma, and (c) is trivially satisfied since $A_i\bar{A}_j = A_j\bar{A}_i$. Thus, all the assumptions of Lemma 3.1 are satisfied and, by a normalization if necessary, we are guaranteed the existence of a unit vector u such that $A_i\bar{u} = \lambda_i u$ for all $i \in \mathscr{I}$. Let $U \in M_n$ be a unitary matrix whose first column is u. Then $U^*A_i\bar{U}$ has the form (3.7) and one shows, just as in the proof of

(3.6), that the family $\mathscr{F}' = \{A_i': i \in \mathscr{I}\} \subset M_{n-1}$ is a concommuting family such that every $A_i'\bar{A}_i'$ has only nonnegative eigenvalues. The desired reduction to upper triangular form now follows after at most $n-2$ repetitions of this reduction, and hence there is one unitary $U \in M_n$ such that $UA_iU^T = \Delta_i$ is upper triangular for all $i \in \mathscr{I}$.

The hypothesis of concommutativity of $\mathscr{F}$ implies that the family of Δ_i's is also concommutative, since $\Delta_i\bar{\Delta}_j = UA_iU^T\bar{U}\bar{A}_jU^* = UA_i\bar{A}_jU^* = UA_j\bar{A}_iU^* = UA_jU^T\bar{U}\bar{A}_iU^* = \Delta_j\bar{\Delta}_i$. If a set of complex numbers $\{\delta_i: i \in \mathscr{I}\}$ has the property that $\delta_i\bar{\delta}_j = \delta_j\bar{\delta}_i$ for all $i, j \in \mathscr{I}$, then either all $\delta_i = 0$ or for some $k \in \mathscr{I}$, $\delta_k \neq 0$. In the latter case, $\delta_i\bar{\delta}_k/|\delta_k|$ is real for all $i \in \mathscr{I}$. In either case, therefore, we can be sure there is a complex number d of absolute value one such that $d^2\delta_i$ is real for all $i \in \mathscr{I}$. Applying this observation to the respective main diagonal entries of the matrices Δ_i, we see that there is a unitary diagonal matrix $D = \operatorname{diag}(d_1, \ldots, d_n)$ such that $D\Delta_iD$ has real main diagonal entries for all $i \in \mathscr{I}$, and hence $(DU)A_i(DU)^T = D\Delta_iD$ has real main diagonal entries for all $i \in \mathscr{I}$. The unitary matrix $V = DU$ has the asserted properties. ■

The hypothesis of concommutativity of $\mathscr{F}$ in (3.9) is, like the hypothesis of commutativity of $\mathscr{G}$ in (3.6), sufficient but not necessary, e.g. $\mathscr{F} = \{\left(\begin{smallmatrix}0&0\\0&1\end{smallmatrix}\right), \left(\begin{smallmatrix}0&1\\0&0\end{smallmatrix}\right)\}$.

The previous theorem implies a sharpening of the conclusion of Corollary 3.8 which could have been proved directly.

Corollary 3.10 *Let $\mathscr{F} = \{A_i: i \in \mathscr{I}\}$ be a given family of symmetric matrices. There exists a unitary $U \in M_n$ such that UA_iU^T is a real diagonal matrix for all $i \in \mathscr{I}$ if and only if $A_i\bar{A}_j = A_j\bar{A}_i$ for all $i, j \in \mathscr{I}$, i.e. each $A_i\bar{A}_j$ is Hermitian.*

Proof If there is a unitary $U \in M_n$ such that $A_i = U\Lambda_iU^T$ for all $i \in \mathscr{I}$, where each $\Lambda_i \in M_n$ is real and diagonal, then $A_i\bar{A}_j = U\Lambda_iU^T\bar{U}\Lambda_jU^* = U\Lambda_i\Lambda_jU^* = U\Lambda_j\Lambda_iU^* = U\Lambda_jU^T\bar{U}\Lambda_iU^* = A_j\bar{A}_i$. The converse follows directly from the theorem, since each $A_i\bar{A}_i = A_iA_i^*$ is positive semidefinite. ■

The equimodular eigenvalue property (c) in Lemma 3.1 arose in this section in the context of assumption (b) of Theorem 3.6, but there are many other natural assumptions that imply it. We list several that are easily verified. The conditions (dx) played a role in this section; the conditions (ex) play a role in the next section.

PROPOSITION 3.11 *Let $A, B \in M_n$. Each of the following conditions is sufficient for a family $\mathscr{G} = \{C_i : i \in \mathscr{I}\} \subset M_n$ to have the property that if $x \neq 0$ and $C_i x = \lambda_i x$ for all $i \in \mathscr{I}$ then $|\lambda_i| = |\lambda_j|$ for all $i, j \in \mathscr{I}$.*

(a1) $\mathscr{G} = \{A_1, A_2\}$, *where* $A_1 = A_2$
(a2) $\mathscr{G} = \{A_1 A_2, A_2 A_1\}$, *where A_1 and A_2 commute*
(a3) $\mathscr{G} = \{A_1 \bar{A}_2, A_2 \bar{A}_1\}$, *where A_1 and A_2 concommute*
(b1) $\mathscr{G} = \{A_1, A_1^*\}$, *where A_1 is normal*
(b2) $\mathscr{G} = \{A_1 A_2, A_2 A_1\}$, *where A_1 and A_2 are Hermitian and $A_1 A_2$ commutes with $A_2 A_1$*
(b3) $\mathscr{G} = \{A_1 \bar{A}_2, A_2 \bar{A}_1\}$, *where A_1 and A_2 are symmetric and $A_1 \bar{A}_2$ commutes with $A_2 \bar{A}_1$*
(c) $\mathscr{G} = \{A_1, A_2\}$, *where $|x^* A_1 x| = |x^* A_2 x|$ for all $x \in \mathbb{C}^n$*
(d1) $\mathscr{G} = \{A_1, A_2\}$, *where $A_1 + A_2$ has only real eigenvalues and $A_1 - A_2$ has only imaginary eigenvalues*
(d2) $\mathscr{G} = \{A_1, A_2\}$, *where $A_1 + A_2$ is Hermitian and $A_1 - A_2$ is skew-Hermitian*
(e1) $\mathscr{G} = \{A_1, A_2\}$, *where $A_1 - A_2$ is nilpotent*
(e2) $\mathscr{G} = \{A_1 A_2, A_2 A_1\}$, *where A_1 and A_2 are simultaneously triangularizable*
(e3) $\mathscr{G} = \{A_1 \bar{A}_2, A_2 \bar{A}_1\}$, *where A_1 and A_2 are simultaneously contriangularizable*

Each of the conditions in the proposition not only implies the stated eigenvalue property, but also has the property that it is inherited under any unitary partial triangularization. That is, if $\mathscr{G}$ is any one of the cited families with the corresponding conditions, and if U is a unitary matrix such that UA_iU^* has the form

$$UA_iU^* = \left(\begin{array}{c|c} \lambda_1^{(i)} & * \\ \hline 0 & \\ \vdots & A_i' \\ 0 & \end{array}\right), \qquad A_i' \in M_{n-1},\ i = 1, 2, \ldots$$

then each A_i' inherits the corresponding property, and hence the family $\mathscr{G}'$ formed from A_1' and A_2' (and their successors from further unitary reductions) has the equimodular eigenvalue property. This is exactly the situation one needs for a successful simultaneous unitary triangularization and it is the reason why some sort of unitarily inheritable property is assumed in each of our simultaneous unitary triangularization theorems.

4. SIMULTANEOUS TRIANGULARIZATION OF A FAMILY BY UNITARY SIMILARITY

If $\mathscr{F} = \{A_i: i \in \mathscr{I}\} \subset M_n$ is a given family, a well-known sufficient condition for $\mathscr{F}$ to be simultaneously unitarily triangularizable is that $\mathscr{F}$ be commutative. Weaker conditions are sufficient for this same conclusion, however.

The key observation is that if $\Delta_1, \Delta_2, \ldots, \Delta_k \in M_n$ are any k upper triangular matrices, then for any permutation π of the integers $1, 2, \ldots, k$, the difference of products $\Delta_1\Delta_2 \cdots \Delta_k - \Delta_{\pi(1)}\Delta_{\pi(2)} \cdots \Delta_{\pi(k)}$ always has a zero main diagonal and hence is nilpotent. Thus, if a family $\{A_i\}$ is simultaneously triangularizable, every difference of products $A_{i_1}A_{i_2} \cdots A_{i_k} - A_{i_{\pi(1)}}A_{i_{\pi(2)}} \cdots A_{i_{\pi(k)}}$ must be nilpotent. In the following lemma, we use this necessary condition to find a sequence of sufficient conditions for a given family to have a common eigenvector, and then use the common eigenvector to construct the desired common unitary similarity. The case $k = 2$ of the Lemma parallels the statement and proof of Lemma 3.1; the case $k = 1$ is just the classical sufficient condition that the family $\mathscr{F}$ is commutative.

LEMMA 4.1 *Let $\mathscr{F} = \{A_i: i \in \mathscr{I}\} \subset M_n$ be a given family, let k be a given positive integer, assume that the family of all possible k-fold products $\mathscr{G}_k = \{A_{i_1}A_{i_2} \cdots A_{i_k}: i_1, i_2, \ldots, i_k \in \mathscr{I}\}$ is commutative, and assume that the difference of products $A_{i_1}A_{i_2} \cdots A_{i_k} - A_{i_{\pi(1)}}A_{i_{\pi(2)}} \cdots A_{i_{\pi(k)}}$ is nilpotent for every $i_1, i_2, \ldots, i_k \in \mathscr{I}$ and every permutation π of the integers $1, 2, \ldots, k$. Then every hon-zero $\mathscr{F}$-invariant subspace of S of $\mathbb{C}^n$ contains a common eigenvector for $\mathscr{F}$, i.e. there is some nonzero $x \in S$ such that $A_ix = \lambda_ix$ for all $i \in \mathscr{I}$.*

Proof Let $S \subset \mathbb{C}^n$ be a given nonzero $\mathscr{F}$-invariant subspace, which must therefore also be $\mathscr{G}_k$-invariant. Since $\mathscr{G}_k$ is commutative, there is a common (nonzero) eigenvector $x \in S$ for $\mathscr{G}_k$. If $k = 1$, we are done, so assume $k \geqslant 2$. Suppose $A_{i_1}A_{i_2} \cdots A_{i_k}x = \lambda_{i_1 i_2 \cdots i_k}x$ for all $i_1, i_2, \ldots, i_k \in \mathscr{I}$. By the nilpotence assumption, every eigenvalue of the difference $A_{i_1}A_{i_2} \cdots A_{i_k} - A_{i_{\pi(1)}}A_{i_{\pi(2)}} \cdots A_{i_{\pi(k)}}$ is zero, and since $\lambda_{i_1 i_2 \cdots i_k} - \lambda_{i_{\pi(1)} i_{\pi(2)} \cdots i_{\pi(k)}}$ is the eigenvalue of this difference associated with the common eigenvector x, we have the identity

$$\lambda_{i_1 i_2 \cdots i_k} = \lambda_{i_{\pi(1)} i_{\pi(2)} \cdots i_{\pi(k)}}$$
$$\text{for all } i_1, i_2, \ldots, i_k \in \mathscr{I} \text{ and every permutation } \pi. \quad (4.2)$$

There are two cases to consider:

(a) $A_i^k x = 0$ for all $i \in \mathscr{I}$, or
(b) There is some $A_0 \in \mathscr{F}$ such that $A_0^k x \neq 0$.

We shall show that there exists a common eigenvector for $\mathscr{F}$ in either case.

First consider case (a), in which there are two possibilities:

(a1) $A_i x = 0$ for all $i \in \mathscr{I}$, or
(a2) There is some $A_0 \in \mathscr{F}$ such that $A_0 x \neq 0$.

In case (a1) the vector x itself is a common eigenvector for $\mathscr{F}$ and we are done.

In case (a2), let $m_0 \equiv \max\{j: A_i^j x \neq 0 \text{ for all } i \in \mathscr{I}\}$, so that $1 \leqslant m_0 < k$ by assumption, $A_i^{m_0+1} x = 0$ for all $i \in \mathscr{I}$, and $y_0 \equiv A_0^{m_0} x \neq 0$ for some $A_0 \in \mathscr{F}$. Then either $A_i y_0 = 0$ for all $i \in \mathscr{I}$ (in which case we are done), or $A_i y_0 = A_i A_0^{m_0} x \neq 0$ for some $i \in \mathscr{I}$. In the latter case, observe that

$$A_i^{2k} A_0^{m_0} x = A_i^{m_0}(A_i^k)(A_i^{k-m_0} A_0^{m_0})x = A_i^{m_0}(A_i^{k-m_0} A_0^{m_0}) A_i^k x = 0$$

because $A_i^k x = 0$. Thus $m_1 \equiv \max\{j: A_i^j y_0 \neq 0 \text{ for all } i \in \mathscr{I}\}$ is finite and surely $2k \geqslant m_1 \geqslant 1$, $A_i^{m_1+1} y_0 = 0$ for all $i \in \mathscr{I}$, and $y_1 \equiv A_1^{m_1} y_0 = A_1^{m_1} A_0^{m_0} x \neq 0$ for some $A_1 \in \mathscr{F}$. We claim that this process will, after at most $k-1$ steps, produce a nonzero vector $y_i \equiv A_i^{m_i} y_{i-1}$ for which $A_j y_i = 0$ for all $j \in \mathscr{I}$, in which case we are done. If not, then for some integer p with $1 \leqslant p \leqslant k-1$ we shall have $m_0 + m_1 + \cdots + m_{p-1} < k$, $m_0 + m_1 + \cdots m_{p-1} + m_p \geqslant k$, and $A_p^{m_p} A_{p-1}^{m_{p-1}} \cdots A_1^{m_1} A_0^{m_0} x \neq 0$ for some $A_0, A_1, \ldots, A_p \in \mathscr{F}$. Let $q \equiv k - (m_0 + m_1 + \cdots + m_{p-1})$ so that $1 \leqslant q \leqslant m_p$. Notice that $A_p^q A_{p-1}^{m_{p-1}} \cdots A_1^{m_1} A_0^{m_0} \in \mathscr{G}_k$ and that $A_p^q A_{p-1}^{m_{p-1}} \cdots A_1^{m_1} A_0^{m_0} x \equiv \lambda x \neq 0$ because x is a common eigenvector for $\mathscr{G}_k$ and

$$0 \neq A_p^{m_p} A_{p-1}^{m_{p-1}} \cdots A_1^{m_1} A_0^{m_0} x = A_p^{m_p - q}(A_p^q A_{p-1}^{m_{p-1}} \cdots A_1^{m_1} A_0^{m_0} x).$$

Because of (4.2) we can permute the terms in the product to obtain

$$A_p^q A_{p-1}^{m_{p-1}} \cdots A_1^{m_1} A_0^{m_0} x = A_p^{q-1} A_{p-1}^{m_{p-1}} \cdots A_1^{m_1} A_0^{m_0} A_p x. \quad (4.3)$$

Using this fact and the commutativity of $\mathscr{G}_k$, we compute

$$
\begin{aligned}
\lambda^2 A_0 x &= A_0(\lambda^2 x) = A_0(A_p^q A_{p-1}^{m_{p-1}} \cdots A_1^{m_1} A_0^{m_0})(A_p^q A_{p-1}^{m_{p-1}} \cdots A_1^{m_1} A_0^{m_0})x \\
&= A_0(A_p^{q-1} A_{p-1}^{m_{p-1}} \cdots A_1^{m_1} A_0^{m_0} A_p)(A_p^q A_{p-1}^{m_{p-1}} \cdots A_1^{m_1} A_0^{m_0})x \\
&= (A_0 A_p^{q-1} A_{p-1}^{m_{p-1}} \cdots A_1^{m_1} A_0^{m_0})(A_p A_p^q A_{p-1}^{m_{p-1}} \cdots A_1^{m_1} A_0^{m_0-1})A_0 x \\
&= (A_p^{q+1} A_{p-1}^{m_{p-1}} \cdots A_1^{m_1} A_0^{m_0-1})(A_0 A_p^{q-1} A_{p-1}^{m_{p-1}} \cdots A_1^{m_1} A_0^{m_0})A_0 x \\
&= (A_p^{q+1} A_{p-1}^{m_{p-1}} \cdots A_1^{m_1} A_0^{m_0} A_p^{q-1} A_{p-1}^{m_{p-1}} \cdots A_1^{m_1})A_0^{m_0+1} x.
\end{aligned}
$$

But this leads to a contradiction, because $\lambda^2 A_0 x \neq 0$ and $A_0^{m_0+1}x = 0$ by construction. Thus, both possibilities have the desired outcome of a common eigenvector in case (a).

The final case to consider is (b), so we assume that $A_0^k x = \lambda x$, $x \neq 0$, $\lambda \neq 0$. Let S_λ denote the λ-eigenspace of A_0^k, i.e. $S_\lambda = \{z \in S: A_0^k z = \lambda z\} \supset \{x\} \neq \{0\}$. Let $\mathscr{G}_{k,0} \equiv \{A_0^{k-1}A_i,\ A_0^{k-2}A_iA_0,\ \ldots,\ A_0A_iA_0^{k-2},\ A_iA_0^{k-1}: i \in \mathscr{I}\} \subset \mathscr{G}_k$. The vector x is a common eigenvector for $\mathscr{G}_{k,0}$, and because of (4.2) we know that $A_0^{k-p}A_iA_0^{p-1}x = \mu_i x$ for $p = 1, 2, \ldots, k$ and all $i \in \mathscr{I}$, i.e. the eigenvalue does not depend on p because of the assumed permutation property. Now define

$$
S_0 \equiv \bigcap_{i \in \mathscr{I}} \bigcap_{p=1}^{k} \{z \in S_\lambda: A_0^{k-p}A_iA_0^{p-1}z = \mu_i z\},
$$

and notice that $S \supset S_\lambda \supset S_0 \supset \{x\} \neq \{0\}$. We claim that S_0 is invariant under A_0^{k-1}. To show this, let $z \in S_0$ be given. If $p = 1$, then

$$
A_0^{k-p}A_iA_0^{p-1}(A_0^{k-1}z) = A_0^{k-1}(A_iA_0^{k-1}z) = A_0^{k-1}(\mu_i z) = \mu_i(A_0^{k-1}z).
$$

If $k \geqslant p \geqslant 2$, then we can use the commutativity of $\mathscr{G}_k$ to compute

$$
\begin{aligned}
A_0^{k-p}A_iA_0^{p-1}(A_0^{k-1}z) &= \frac{1}{\lambda}A_0^{k-p}A_iA_0^{p-1}A_0^{k-1}(A_0^k z) \\
&= \frac{1}{\lambda}A_0^{k-p}A_iA_0^{p-1}A_0^kA_0^{k-1}z = \frac{1}{\lambda}(A_0^{k-p}A_iA_0^{p-1})(A_0^k)A_0^{k-1}z \\
&= \frac{1}{\lambda}(A_0^k)(A_0^{k-p}A_iA_0^{p-1})A_0^{k-1}z = \frac{1}{\lambda}(A_0^{2k-p}A_iA_0^{p-2})(A_0^k z) \\
&= A_0^{2k-p}A_iA_0^{p-2}z = A_0^{k-1}(A_0^{k-p+1}A_iA_0^{p-2}z) = A_0^{k-1}(\mu_i z) \\
&= \mu_i(A_0^{k-1}z).
\end{aligned}
$$

These two calculations show that S_0 is invariant under A_0^{k-1}. Thus, there exists an eigenvector for A_0^{k-1} in S_0, i.e. there is a nonzero vector $w \in S_0$ such that $A_0^{k-1}w = \mu w$, and $\mu \neq 0$ because $\lambda w = A_0^k w = A_0 A_0^{k-1} w = A_0 \mu w = \mu A_0 w$ and $\lambda \neq 0$. But then $\mu_i w = A_i A_0^{k-1} w = A_i \mu w = \mu A_i w$ and $A_i w = (\mu_i/\mu) w$ for all $i \in \mathscr{I}$, so w is a common eigenvector for $\mathscr{F}$. ■

THEOREM 4.4 *Let $\mathscr{F} = \{A_i : i \in \mathscr{I}\} \subset M_n$ be a given family, let k be a given positive integer, and assume that the family $\mathscr{G}_k = \{A_{i_1} A_{i_2} \cdots A_{i_k} : i_1, i_2, \ldots, i_k \in \mathscr{I}\}$ of all k-fold products from $\mathscr{F}$ is commutative. There exists a unitary $U \in M_n$ such that UA_iU^* is upper triangular for all $i \in \mathscr{I}$ if and only if the difference of products $A_{i_1} A_{i_2} \cdots A_{i_k} - A_{i_{\pi(1)}} A_{i_{\pi(2)}} \cdots A_{i_{\pi(k)}}$ is nilpotent for every $i_1, i_2, \ldots, i_k \in \mathscr{I}$ and every permutation π of the integers $1, 2, \ldots, k$.*

Proof The necessity of the nilpotence condition has already been observed. To show its sufficiency, invoke Lemma 4.1 with $S = \mathbb{C}^n$ to produce a common eigenvector x for $\mathscr{F}$, which we may assume is a unit vector and satisfies $A_i x = \lambda_i x$ for all $i \in \mathscr{I}$. Let $U \in M_n$ be a unitary matrix with first column x, so that

$$U^* A_i U = \left(\begin{array}{c|c} \lambda_i & * \\ \hline 0 & A_i' \end{array}\right), \qquad A_i' \in M_{n-1},\ i \in \mathscr{I}.$$

We could now proceed to reduce the family $\mathscr{F}' \equiv \{A_i' : i \in \mathscr{I}\}$ in the same way if $\mathscr{F}'$ inherited from $\mathscr{F}$ the properties that $\mathscr{G}' = \{A'_{i_1} A'_{i_2} \cdots A'_{i_k} - A'_{i_{\pi(1)}} A'_{i_{\pi(2)}} \cdots A'_{i_{\pi(k)}}\}$ is commutative and $A'_{i_1} A'_{i_2} \cdots A'_{i_k} - A'_{i_{\pi(1)}} A'_{i_{\pi(2)}} \cdots A'_{i_{\pi(k)}}$ is nilpotent for all $i_1, i_2, \ldots, i_k \in \mathscr{I}$ and all permutations π. It is a straightforward computation to verify both of these properties, as in the proof of Theorem 3.6. ■

The kinship of this result with Theorem 3.6 is particularly apparent in the case $k = 2$, which we state separately for emphasis.

COROLLARY 4.5 *Let $\mathscr{F} = \{A_i : i \in \mathscr{I}\} \subset M_n$ be a given family and assume that the family $\mathscr{G}_2 \equiv \{A_i A_j : i, j \in \mathscr{I}\}$ is commutative. There exists a unitary $U \in M_n$ such that UA_iU^* is upper triangular for all $i \in \mathscr{I}$ if and only if every member of the family $\mathscr{H} \equiv \{A_i A_j - A_j A_i : i, j \in \mathscr{I}\}$ is nilpotent.*

The Corollary, and hence the Theorem, is stronger than the classical theorem on simultaneous unitary triangularization, which assumes that $\mathscr{F}$ is commutative. Consider the following example of a noncommutative $\mathscr{F}$:

$$\mathscr{F} = \left\{A_1 = \begin{pmatrix} 1 & -1 \\ 0 & -1 \end{pmatrix}, A_2 = \begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix}\right\}$$

$$A_1^2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I = A_2^2, \qquad A_1A_2 = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} \neq A_2A_1 = \begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix}$$

$$(A_1A_2)(A_2A_1) = I = (A_2A_1)(A_1A_2).$$

Thus, $\mathscr{G}_2 = \{A_1^2, A_1A_2, A_2A_1, A_2^2\}$ is commutative.

There may be a temptation to hope that commutativity of the family $\mathscr{G}_k$ for some, perhaps large, value of k might be necessary as well as sufficient for simultaneous unitary triangularization of $\mathscr{F}$ in Theorem 4.4, but this is not true. Consider the family

$$\mathscr{F} = \left\{A_1 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, A_2 = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}\right\}$$

for which

$$A_1^k = \begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad A_2^k = \begin{pmatrix} 1 & 0 \\ 0 & 2^k \end{pmatrix}.$$

Since

$$A_1^kA_2^k = \begin{pmatrix} 1 & k2^k \\ 0 & 2^k \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 1 & k \\ 0 & 2^k \end{pmatrix} = A_2^kA_1^k,$$

the family $\mathscr{G}_k$ is not commutative for any $k \geqslant 1$.

For completeness, we give an example [suggested by the referee] which shows that commutativity of the family $\mathscr{G}_k$ for $k \geqslant 2$ is, by itself, insufficient to imply simultaneous unitary triangularizability of the family $\mathscr{F}$, i.e. commutativity of $\mathscr{G}_k$ does not imply that the elements of $\mathscr{H}$ are nilpotent. For any $k \geqslant 2$, let $A = J_k(0)$ be the k-by-k Jordan block with zero eigenvalue and let $B = E_{k,1}$, where $E_{i,j}$ denotes the k-by-k matrix with a 1 in the i,j position and zero everywhere else. Then $\mathscr{G}_k = \{0, E_{1,1}, E_{2,2}, \ldots, E_{k,k}\}$ is commutative, but $A^{k-1}B - BA^{k-1} = E_{1,1} - E_{k,k}$ is not nilpotent.

The conditions we have given in Theorem 4.4 are merely *sufficient* for a given family of matrices to be simultaneously unitarily triangularizable. Necessary and sufficient conditions have been known since the work of McCoy [5] in 1936; see [1, 4] for a more recent perspective and further references.

5. SIMULTANEOUS TRIANGULARIZATION OF TWO FAMILIES BY UNITARY CONGRUENCES

We wish to generalize Theorems 3.6 and 4.4 to cover the case of two families of matrices $\{A_i\}$ and $\{B_j\}$ that are to be reduced simultaneously to upper triangular form by unitary consimilarity and unitary similarity, respectively. That is, we want to have a single unitary matrix U such that every UA_iU^T and UB_jU^* is upper triangular. Our first step is to obtain a version of Lemmata 3.1 and 4.1 that covers this combined situation.

LEMMA 5.1 *Let $\mathscr{F} = \{A_i: i \in \mathscr{I}\} \subset M_n$ and $\mathscr{H} = \{B_j: j \in \mathscr{J}\} \subset M_n$ be given families. Assume that every nonzero $\mathscr{F}$-coninvariant subspace of $\mathbb{C}^n$ contains a common coneigenvector for $\mathscr{F}$, and that every nonzero $\mathscr{H}$-invariant subspace of $\mathbb{C}^n$ contains a common eigenvector for $\mathscr{H}$. Assume also that*

(a) *B_j has only real eigenvalues for all $j \in \mathscr{J}$, and*
(b) *$A_i\bar{B}_j = B_jA_i$ for all $i \in \mathscr{I}$ and all $j \in \mathscr{J}$.*

Then every nonzero $\mathscr{H}$-invariant subspace S of $\mathbb{C}^n$ contains a nonzero vector x such that $A_i\bar{x} = \lambda_i x$ for all $i \in \mathscr{I}$ and $B_jx = \mu_jx$ for all $j \in \mathscr{J}$, i.e. there exists a nonzero vector in S that is a common coneigenvector for $\mathscr{F}$ as well as a common eigenvector for $\mathscr{H}$.

Proof Let S be a nonzero $\mathscr{H}$-invariant subspace, and let w be a common eigenvector for $\mathscr{H}$ in S, i.e. $0 \neq w \in S$ is such that $B_jw = \mu_jw$ for all $j \in \mathscr{J}$. Each μ_j is real by assumption (a). Let S' denote the (necessarily nonzero) subspace of S consisting of all the common eigenvectors of $\mathscr{H}$, with the same eigenvalues as w, i.e.

$$S \supset S' \equiv \bigcap_{j \in \mathscr{J}} \{x: B_jx = \mu_jx\} \supset \{w\} \neq \{0\}.$$

It is evident that S' is $\mathscr{H}$-invariant, but we claim that it is $\mathscr{F}$-coninvariant as well. If $x \in S'$, $B_jA_i\bar{x} = A_i\bar{B}_j\bar{x} = A_i(\overline{B_jx}) = A_i(\overline{\mu_jx}) = \mu_j(A_i\bar{x})$ because of the assumptions that $A_i\bar{B}_j = B_jA_i$ and each μ_j is real. This says that $A_i\bar{x} \in S'$ for all $i \in \mathscr{I}$, i.e. S' is $\mathscr{F}$-coninvariant. The hypotheses guarantee that there is a common coneigenvector for $\mathscr{F}$ in S', and this is the desired vector. ■

Our main result says that, subject to three conditions, the two families $\mathscr{F}$ and $\mathscr{H}$ are jointly simultaneously unitarily contriangularizable and triangularizable, respectively, if and only if they are separately simultaneously contriangularizable and simultaneously triangularizable, respectively.

THEOREM 5.2 *Let $\mathscr{F} = \{A_i: i \in \mathscr{I}\} \subset M_n$ and $\mathscr{H} = \{B_j: j \in \mathscr{J}\} \subset M_n$ be given families, let k be a given positive integer, let $\mathscr{G} = \{A_i\bar{A}_j: i, j \in \mathscr{I}\}$, and let $\mathscr{L}_k = \{B_{i_1}B_{i_2} \cdots B_{i_k}: i_1, i_2, \ldots, i_k \in \mathscr{I}\}$. Assume that*

(1) *Each of $\mathscr{G}$ and $\mathscr{L}_k$ is a commuting family,*
(2) *B_j has only real eigenvalues for all $j \in \mathscr{J}$, and*
(3) *$A_i\bar{B}_j = B_jA_i$ for all $i \in \mathscr{I}$ and all $j \in \mathscr{J}$.*

Then there exists a unitary $U \in M_n$ such that UA_iU^T is upper triangular for all $i \in \mathscr{I}$ and UB_jU^* is upper triangular for all $j \in \mathscr{J}$ if and only if

(a) *$A_i\bar{A}_i$ has only nonnegative eigenvalues for all $i \in \mathscr{I}$,*
(b) *$A_i\bar{A}_j + A_j\bar{A}_i$ has only real eigenvalues and $A_i\bar{A}_j - A_j\bar{A}_i$ has only imaginary eigenvalues for all $i, j \in \mathscr{I}$, and*
(c) *$B_{j_1}B_{j_2} \cdots B_{j_k} - B_{j_{\pi(1)}}B_{j_{\pi(2)}} \cdots B_{j_{(k)}}$ is nilpotent for all $j_1, j_2, \ldots, j_k \in \mathscr{J}$ and every permutation π of the integers $1, 2, \ldots, k$.*

Proof The necessity of conditions (a), (b), and (c) is easily checked; it also follows from Theorems 3.6 and 4.4. To establish the sufficiency of these three conditions it suffices to show, as in the proofs of Theorems 3.6 and 4.4, that there is some nonzero $x \in \mathbb{C}^n$ that is both a common coneigenvector for $\mathscr{F}$ and a common eigenvector for $\mathscr{H}$, and Lemmata 5.1, 3.1, and 4.1 guarantee the existence of such an x. ■

An important special case is the one in which the family $\mathscr{F}$ consists of complex symmetric matrices and $\mathscr{H}$ consists of Hermitian matrices. In this case, assumption (2) and conditions (a) and (b) are automatically satisfied.

COROLLARY 5.3 *Let $\mathscr{F} = \{A_i: i \in \mathscr{I}\} \subset M_n$ be a given family of symmetric matrices, let $\mathscr{H} = \{B_j: j \in \mathscr{J}\} \subset M_n$ be a given family of Hermitian matrices, and let $\mathscr{G} = \{A_i\bar{A}_j: i, j \in \mathscr{I}\}$. There exists a unitary $U \in M_n$ such that UA_iU^T is diagonal for all $i \in \mathscr{I}$ and UB_jU^* is diagonal for all $j \in \mathscr{J}$ if and only if each of $\mathscr{G}$ and $\mathscr{H}$ is a commuting family and B_jA_i is symmetric for all $i \in \mathscr{I}$ and all $j \in \mathscr{J}$.*

Proof The necessity of the stated conditions is easily verified. To show that they are also sufficient, we show that they imply the six assumptions and conditions in Theorem 5.2. Commutativity of $\mathscr{H}$ is a stronger condition than commutativity of $\mathscr{L}_k$ in Theorem 5.2, so assumption (1) is satisfied and (c) is trivially satisfied since all these differences vanish. Because $(B_jA_i)^T = A_i^TB_j^T = A_i\bar{B}_j$, the assumption that B_jA_i is symmetric is equivalent to (3) when B_j is Hermitian and A_i is symmetric. We have already observed that the remaining assumption (2) and conditions (a) and (b) are automatically satisfied. Thus, there exists a unitary $U \in M_n$ such that every UA_iU^T and UB_jU^* is upper triangular. But an upper triangular symmetric or Hermitian matrix must be diagonal, so the assertion is proved. ■

If we specialize the Corollary to the case in which the families $\mathscr{F}$ and $\mathscr{H}$ each contain only one element, we obtain a result from [3]: Let $A, B \in M_n$ be given, with A symmetric and B Hermitian. There exists a unitary $U \in M_n$ such that both UAU^T and UBU^* are diagonal if and only if BA is symmetric, i.e. $BA = A\bar{B}$.

References

[1] F. Gaines and R. C. Thompson, Sets of nearly triangular matrices, *Duke Math. J.* **35** (1968), 441–453.

[2] Y. P. Hong and R. A. Horn, On the reduction of a matrix to triangular or diagonal form by consimilarity, *SIAM J. Algebraic & Discrete Methods* (to appear).

[3] Y. P. Hong, R. A. Horn and C. R. Johnson, On the reduction of pairs of Hermitian or symmetric matrices to diagonal form by congruence, *J. Linear Algebra Appl.* (to appear).

[4] T. J. Laffey, Simultaneous triangularization of matrices—low rank cases and nonderogatory case, *Linear and Multilinear Algebra* **6** (1978), 269–305.

[5] N. H. McCoy, On the characteristic roots of matric polynomials, *Bull. Amer. Math. Soc.* **42** (1936), 592–600.

[6] N. Jacobson, *The Theory of Rings*, American Mathematical Society, New York, 1943.

[7] I. Schur, Ein Satz über quadratische Formen mit komplexen Koeffizienten, *Amer. J. Math.* **67** (1945), 472–480.

[8] C. L. Siegel, Symplectic geometry, *Amer. J. Math.* **65** (1943), 1–86.

[9] T. Takagi, On an algebraic problem related to an analytic theorem of Carathéodory and Fejer and on an allied theorem of Landau, *Japanese J. Math.* **1** (1925), 83–93.