

The intersection of some classical equivalence classes of matrices

by

Mark Alan Mills

A dissertation submitted to the graduate faculty
in partial fulfillment of the requirements for the degree of
DOCTOR OF PHILOSOPHY

Major: Mathematics

Major Professor: Bryan E. Cain

Iowa State University

Ames, Iowa

1999

Graduate College
Iowa State University

This is to certify that the Doctoral dissertation of
Mark Alan Mills
has met the dissertation requirements of Iowa State University

Major Professor

For the Major Program

For the Graduate College

TABLE OF CONTENTS

ACKNOWLEDGEMENTS	v
ABSTRACT	vi
CHAPTER 1 INTRODUCTION AND BACKGROUND	1
1.1 Equivalence Classes	1
1.2 Canonical Forms and Invariants	2
1.3 Unitary Similarity	5
1.4 The Problem	6
CHAPTER 2 THE CONJUNCTIVE-SIMILARITY EQUIVALENCE CLASS	7
2.1 $CS(A)$ and $\mathcal{U}(A)$	7
2.2 $CS(A)$: The $n \times n$ Case (Part 1)	8
2.3 $CS(A)$: The 1×1 Case	12
2.4 $CS(A)$: The 2×2 Case	12
2.5 $CS(A)$: The $n \times n$ Case (Part 2)	16
2.5.1 $Sim(A)$, $Conj(A)$, and $\mathcal{U}(A)$	16
2.5.2 Diagonal T and Triangular A	17
2.5.3 Direct Sum of 2×2 and Identity Matrices	18
2.5.4 AM Similar to A	18
2.5.5 A in Jordan Canonical Form	21
2.5.6 Conclusions	21
2.6 Invariants of $CS(A)$	21
2.7 The Converse Problem	23

CHAPTER 3 THE SIMILARITY-UNITARY EQUIVALENCE CLASS	27
3.1 UES(A) and $\mathcal{U}(A)$	27
3.2 UES(A): The $n \times n$ Case	28
3.3 UES(A): The 1×1 Case	34
3.4 UES(A): The 2×2 Case	35
3.5 U $\text{Equiv}(A)$ and $\mathcal{U}(A)$	35
CHAPTER 4 CONCLUSIONS AND FUTURE DIRECTIONS	37
4.1 CS(A)	37
4.2 UES(A)	37
APPENDIX	39
BIBLIOGRAPHY	41

ACKNOWLEDGEMENTS

First, I thank God for His guidance throughout this process and for the abilities and strength He has given me to be able to reach this far. When I consider God's creation of mathematics, Psalm 19:1–4 (New International Version) is certainly correct when it says: “The heavens declare the glory of God; the skies proclaim the work of his hands. Day after day they pour forth speech; night after night they display knowledge. There is no speech or language where their voice is not heard. Their voice goes out into all the earth, their words to the ends of the world.”

Next, I thank the many family members and friends who have supported and encouraged me during this time. Their confidence in me allowed me to reach further than I ever thought I could. I am especially indebted to my parents for their support (both emotional and monetary) that has allowed me to achieve a dream. Without them, none of this would have been possible.

Finally, I thank Dr. Bryan Cain for his mathematical guidance which led to this dissertation. His initial interest in this problem and his insights, encouragement, and friendship throughout this process have been invaluable to me.

ABSTRACT

Let A be an $n \times n$ complex matrix. Let $\text{Sim}(A)$ denote the similarity equivalence class of A , $\text{Conj}(A)$ denote the conjunctivity equivalence class of A , $\text{UEquiv}(A)$ denote the unitary-equivalence equivalence class of A , and $\mathcal{U}(A)$ denote the unitary similarity equivalence class of A . Each of these equivalence classes has been studied for some time and is generally well-understood. In particular, canonical forms have been given for each equivalence class. Since the intersection of any two equivalence classes of A is again an equivalence class of A , we consider two such intersections: $\text{CS}(A) \equiv \text{Sim}(A) \cap \text{Conj}(A)$ and $\text{UES}(A) \equiv \text{Sim}(A) \cap \text{UEquiv}(A)$. Though it is natural to first think that each of these is simply $\mathcal{U}(A)$, for each A , we show by examples that this is not the case. We then try to classify which matrices A have $\text{CS}(A) = \mathcal{U}(A)$. For matrices having $\text{CS}(A) \neq \mathcal{U}(A)$, we try to count the number of disjoint unitary similarity classes contained in $\text{CS}(A)$. Though the problem is not completely solved for $\text{CS}(A)$, we reduce the problem to non-singular, non-co-Hermitian matrices A . A similar analysis is performed for $\text{UES}(A)$, and a (less simple) reduction of the problem is also achieved.

CHAPTER 1

INTRODUCTION AND BACKGROUND

1.1 Equivalence Classes

Throughout mathematics, the notion of an equivalence class is a basis for understanding the structure of various algebraic objects (e.g., cosets in group theory). If S is a set, then recall that $R \subset S \times S$ is an equivalence relation on S provided R is:

reflexive : $(a, a) \in R$ for all $a \in S$;

symmetric : $(a, b) \in R \implies (b, a) \in R$; and

transitive : $(a, b) \in R$ and $(b, c) \in R \implies (a, c) \in R$.

Two elements $a, b \in S$ are said to be *equivalent*, if $(a, b) \in R$. We can then define the *equivalence class* of a particular element $a \in S$ as the set of all elements in S that are equivalent to a via the equivalence relation R . If there are two different equivalence relations R_1 and R_2 on S , then, for a fixed element $a \in S$, the intersection of the two equivalence classes of a is again an equivalence class of a with respect to the new equivalence relation $R_1 \cap R_2$ on S .

The set S is the union of the equivalence classes of a (possibly small) subset of the elements in S , and this gives S some structure. Certainly this is true in group theory, where a group is the union of its cosets, which are equivalence classes of a particular subset of group elements. This is also true for the set of all $n \times n$ complex matrices (denoted by M_n), where there are a number of well-known equivalence relations.

Recall that two matrices $A, B \in M_n$ are said to be *similar* if there exists an $n \times n$ invertible matrix S so that $B = S^{-1}AS$. (I will denote the set of $n \times n$ complex invertible matrices by GL_n .) Similarity is an equivalence relation in M_n , and we can consider the similarity

equivalence class of a matrix A , denoted by $\text{Sim}(A)$. If A and B are similar, then they have the same eigenvalues (counting algebraic multiplicities). (From here on, the term *multiplicity* when referring to eigenvalues will refer to algebraic multiplicity, not geometric multiplicity.) However, the converse is not true, as is shown by the matrices $\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ and $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ that both have the eigenvalue 0 with algebraic multiplicity 2, but are not similar.

Two matrices $A, B \in M_n$ are said to be *conjunctive* (or **congruent* or *Hermitian congruent*), if there exists $T \in GL_n$ so that $B = T^*AT$, where T^* denotes the conjugate transpose of the matrix T . Conjunctivity is an equivalence relation in M_n , and we can consider the conjunctive equivalence class of a matrix A , denoted by $\text{Conj}(A)$.

We will say that two matrices $A, B \in M_n$ are *unitarily equivalent*, if there exist unitary matrices U and V so that $B = U^*AV$. (I will denote the set of all $n \times n$ unitary matrices by U_n .) Again, unitary equivalence is an equivalence relation in M_n , and we can consider the unitary equivalence class of a matrix A , denoted by $\text{UEquiv}(A)$. Two $n \times n$ matrices are unitarily equivalent if and only if they have the same singular values (counting multiplicities).

Two matrices $A, B \in M_n$ will be said to be *unitarily similar*, if there exists $U \in U_n$ so that $B = U^*AU$. As expected, unitary similarity is an equivalence relation in M_n , and we can consider the unitary similarity equivalence class of a matrix A , denoted by $\mathcal{U}(A)$.

(It should be noted that some of the terminology within this dissertation is not necessarily consistent with other sources within the field of linear algebra. In particular, [HJ1] uses the term ‘unitarily equivalent’ to refer to what here has been defined as ‘unitary similarity’. Also, they would likely use the term ‘equivalent via unitaries’ to refer to what here has been defined as ‘unitarily equivalent’.)

1.2 Canonical Forms and Invariants

Often a “simple” representative of an equivalence class is sought. This representative is called a *canonical* representative, and its form is chosen so that there is a unique representative (up to the ordering of the components of the canonical form) for each equivalence class and so that different (i.e., disjoint) equivalence classes have different representatives. Each of the

above equivalence classes has such a canonical form.

Another thing that is often desired for an equivalence class is a set of *invariants*. These invariants are functions that are constant throughout all members of an equivalence class. Each of the above equivalence classes has a set of invariants.

For $\text{Sim}(A)$, the canonical form is the well-known Jordan canonical form. Consider the matrix pencil $A - \lambda I$, for variable λ . Let the $k_i \times k_i$ matrix

$$J_i(\lambda_i) = \begin{bmatrix} \lambda_i & 1 & & \mathbf{0} \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ \mathbf{0} & & & \lambda_i \end{bmatrix}$$

be associated with the elementary divisor $(\lambda - \lambda_i)^{k_i}$ of the matrix pencil. Then A is similar to a diagonal block matrix

$$\begin{bmatrix} J_1(\lambda_1) & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & J_p(\lambda_p) \end{bmatrix},$$

where $\lambda_1, \dots, \lambda_p$ are (not necessarily distinct) eigenvalues of the matrix A , and $k_1 + k_2 + \dots + k_p = n$. This form is unique up to the ordering of the λ_i . The elementary divisors of each similarity equivalence class are invariants.

For $\text{Conj}(A)$, the canonical form is less well-known and much less simple. For $A \in M_n$, let $\text{Re}(A) = \frac{A + A^*}{2}$ and $\text{Im}(A) = \frac{A - A^*}{2i}$ and consider the matrix pencil $\mu \text{Re}(A) + \lambda \text{Im}(A)$, for variables μ and λ . Then by Theorem 1 of [Th] we know that each matrix $A \in M_n$ is conjunctive to a direct sum of the following types of matrices:

(i)

$$\epsilon D_e(\gamma) = \epsilon \begin{bmatrix} \mathbf{0} & & & \gamma + i \\ & \cdot & & 1 \\ & & \cdot & \\ & & & \cdot \\ \gamma + i & 1 & & \mathbf{0} \end{bmatrix},$$

an $e \times e$ matrix with real γ associated with the (finite) elementary divisor $(\mu\gamma - \lambda)^e$ of the matrix pencil and $\epsilon = \pm 1$ the signature associated with γ ;

- (ii) $\begin{bmatrix} 0 & D_E(\Gamma) \\ D_E(\bar{\Gamma}) & 0 \end{bmatrix}$, a $2E \times 2E$ matrix with non-real Γ associated with the (finite) elementary divisors $(\mu\Gamma - \lambda)^E$ and $(\mu\bar{\Gamma} - \lambda)^E$ of the matrix pencil, and D_E defined by (i);

- (iii)

$$\epsilon D_e(\infty) = \epsilon \begin{bmatrix} \mathbf{0} & & & -1 \\ & & & \cdot & -i \\ & & & \cdot & \cdot \\ & & & \cdot & \cdot \\ & & & \cdot & \cdot \\ -1 & -i & & & \mathbf{0} \end{bmatrix},$$

an $e \times e$ matrix associated with the (infinite) elementary divisor μ^e of the matrix pencil and with signature $\epsilon = \pm 1$; and

- (iv)

$$D_{\mathcal{E}} = \left[\begin{array}{ccc|ccc} & & & i & & \mathbf{0} \\ & & & 1 & \ddots & \\ & & \mathbf{0}_{\mathcal{E}} & & \ddots & i \\ & & & \mathbf{0} & & 1 \\ \hline i & 1 & \mathbf{0} & & & \\ & \ddots & \ddots & & & \mathbf{0}_{\mathcal{E}-1} \\ \mathbf{0} & & i & 1 & & \end{array} \right],$$

a $(2\mathcal{E} - 1)$ -square matrix where $\mathbf{0}_{\mathcal{E}}$ denotes an \mathcal{E} -square zero matrix, and where \mathcal{E} is a minimal index of the matrix pencil.

Again, this form is unique up to the ordering of the direct summands. The finite and infinite elementary divisors, signatures, and minimal indices for each conjunctive equivalence class are invariants. (See the Appendix for definitions and discussion of elementary divisors and minimal indices.)

The canonical form for $\text{UEquiv}(A)$ is the well-known singular-value decomposition. Every matrix $A \in M_n$ is unitarily equivalent to a unique diagonal matrix of the form

$$\begin{bmatrix} \sigma_1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \sigma_n \end{bmatrix},$$

where $\sigma_1 \geq \dots \geq \sigma_n \geq 0$ are the singular values of A and each σ_i is a square root of an eigenvalue of the matrix A^*A . The singular values (counting multiplicities) of each unitary equivalence class are invariants.

For the purposes of this dissertation, we need not consider canonical forms for $\mathcal{U}(A)$. In the next section, we will present a set of unitary invariants that will be used throughout this dissertation. However, a development of canonical forms and other invariants can be found in [Sh].

1.3 Unitary Similarity

As noted previously, the intersection of two different equivalence classes of a fixed element is again an equivalence class of this element. This dissertation will consider two such intersections of matrix equivalence classes:

$$\text{Sim}(A) \cap \text{Conj}(A) \quad \text{and} \quad \text{Sim}(A) \cap \text{UEquiv}(A). \quad (1.1)$$

The first reaction might be that one or both of these intersections is simply the unitary similarity class $\mathcal{U}(A)$. We will see that this is true for some matrices A , but certainly not all.

However, this raises the issue of how to determine when two matrices are unitarily similar. Theoretically, we could place each one in its canonical form for unitary similarity, but this is generally not easy to do. We seek some other criteria to determine unitary similarity more simply. Such criteria can be found in Theorems 6.1 and 6.3 from the survey paper [Sh], or in Theorems 2.2.6 and 2.2.8 in [HJ1]. (Note that the trace of the matrix A is denoted by $\text{tr}(A)$.) For the purposes of the following two theorems, a *word* in the non-commuting variables x and y is a finite formal product of non-negative powers of x and y , and a word's *degree* is the sum of all its powers of x and y .

Theorem 1.3.1 (Specht, [Sp]). Let $A, B \in M_n$. Then A and B are unitarily similar if and only if $\text{tr}(\omega(A, A^*)) = \text{tr}(\omega(B, B^*))$, for every word $\omega(x, y)$ in non-commuting variables x and y .

While the previous theorem says that we must consider an infinite number of words, the following theorem shows that it is enough to just consider a finite number of words to determine unitary similarity.

Theorem 1.3.2 (Percy, [Pe]). If $A, B \in M_n$, and if $\text{tr}(\omega(A, A^*)) = \text{tr}(\omega(B, B^*))$ for every word $\omega(x, y)$ of degree less than or equal to $2n^2$ in non-commuting variables x and y , then A and B are unitarily similar.

For the 2×2 case, it is even simpler.

Theorem 1.3.3 ([Sh], Theorem 2.4). Let $A \in M_2$ with eigenvalues λ_1 and λ_2 , which may or may not be distinct. Let

$$r = \sqrt{\text{tr}(A^*A) - |\lambda_1|^2 - |\lambda_2|^2}.$$

Then A is unitarily similar to a matrix of the form $\begin{bmatrix} \lambda_1 & r \\ 0 & \lambda_2 \end{bmatrix}$. Furthermore, if A is unitarily similar to any triangular matrix $T = (t_{ij})$ with λ_1 and λ_2 on the main diagonal, then $|t_{12}| = r$.

1.4 The Problem

With this introduction and background in place, we are now ready to examine the two intersections of equivalence classes (1.1). The original problem was to find a canonical form for $\text{Sim}(A) \cap \text{Conj}(A)$. However, before this could be done, we needed to understand better what would make a matrix both similar and conjunctive to a matrix A . The results of this analysis make up Chapter 2. In Chapter 3, a similar analysis of $\text{Sim}(A) \cap \text{UEquiv}(A)$ is presented. Open problems and directions for future research appear in Chapter 4.

CHAPTER 2

THE CONJUNCTIVE-SIMILARITY EQUIVALENCE CLASS

For $A \in M_n$, consider the equivalence class intersection $\text{Sim}(A) \cap \text{Conj}(A)$, which we will denote by $\text{CS}(A)$. This intersection is all matrices $B \in M_n$ such that $B = T^*AT = S^{-1}AS$, for some $T, S \in GL_n$. Of course, this intersection is non-empty because $A \in \text{CS}(A)$, and if $A = \alpha I$, then $\text{CS}(A) = A$. Also, because $U^* = U^{-1}$ for all $U \in U_n$, we know that $\mathcal{U}(A) \subset \text{CS}(A)$.

2.1 $\text{CS}(A)$ and $\mathcal{U}(A)$

As was mentioned before, a natural first thought is that perhaps $\text{CS}(A) = \mathcal{U}(A)$ for all A . However, the following example shows that this is not true.

Example 2.1.1. For $n \geq 3$, consider the $n \times n$ permutation matrix (which is both normal and unitary)

$$P = \begin{bmatrix} 0 & 0 & \dots & 0 & 1 \\ 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \ddots & 0 & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & 0 & \dots & 1 & 0 \end{bmatrix}.$$

Let $D \in GL_n$ be the diagonal matrix

$$\left[\begin{array}{cc|c} a & 0 & \mathbf{0} \\ 0 & \frac{1}{a} & \\ \hline \mathbf{0} & & I_{n-2} \end{array} \right],$$

where $a > 0$ and I_k denotes the $k \times k$ identity matrix. Note that then $D^* = D$ and the determinant $\det(D) = 1$. Also note that P has characteristic polynomial $p(\lambda) = \lambda^n - 1$, so that

its n eigenvalues are the n th roots of unity $\{e^{\frac{2\pi j}{n}i} \mid j = 0, 1, \dots, n-1\}$, and hence are distinct. The matrix $B = D^*PD$ (or just DPD) has the same characteristic polynomial as P . So these two matrices have the same eigenvalues, and since the eigenvalues are distinct, we know that P and B are similar. So $B \in \text{CS}(P)$. We see that $\text{tr}(P^*P) = n$ and $\text{tr}(B^*B) = n - 2 + a^2 + \frac{1}{a^2}$, and for $a > 1$ these two traces are not equal. Using the criteria from Theorem 1.3.1, we see that P and B are not unitarily similar when $a > 1$. So $\text{CS}(P)$ contains more than just $\mathcal{U}(P)$. In fact, this example shows that $\text{CS}(P)$ contains an uncountable number of disjoint unitary similarity classes, since for $a > 1$, we can get a continuum of values for $\text{tr}(B^*B)$.

In light of this example, it is natural to ask what matrices A have $\text{CS}(A) = \mathcal{U}(A)$. Certainly, this is not true for all normal (or unitary or permutation) matrices, as the example shows. But is it true of a smaller class of matrices?

2.2 $\text{CS}(A)$: The $n \times n$ Case (Part 1)

In the general $n \times n$ case, we seek to characterize those matrices A for which $\text{CS}(A) = \mathcal{U}(A)$. For those matrices not satisfying this equation, we then want to know how many disjoint unitary similarity classes are contained in $\text{CS}(A)$. If Example 2.1.1 is any guide, when $\text{CS}(A) \neq \mathcal{U}(A)$ we might expect to get uncountably many disjoint unitary similarity classes.

Recall that a matrix $A \in M_n$ is *Hermitian*, if $A^* = A$. (\mathcal{H}_n will denote the set of $n \times n$ Hermitian matrices.) We will say that a matrix B is *co-Hermitian*, if $B = \alpha A$ for some $\alpha \in \mathbb{C}$ and $A \in \mathcal{H}_n$. The skew-Hermitian matrices, where $A^* = -A$, are co-Hermitian.

We need a lemma to help us on our way.

Lemma 2.2.1. Let $A, B \in \mathcal{H}_n$. Then $B = U^*AU$ for some $U \in U_n$ if and only if A and B have the same eigenvalues (counting multiplicities).

Proof. First suppose that $B = U^*AU$ for some $U \in U_n$. Since $U^* = U^{-1}$, we have that A and B are similar. Therefore, they have the same eigenvalues (counting multiplicities).

Now suppose A and B have the same eigenvalues (counting multiplicities). Then there exist $V, W \in U_n$ with $V^*AV = W^*BW = \Lambda$, where Λ is a diagonal matrix. Therefore,

$B = (VW^*)^*A(VW^*)$, and A and B are unitarily similar. \square

With this lemma in place, we are now ready to look at the Hermitian matrices.

Theorem 2.2.2. If $A \in \mathcal{H}_n$, then $\text{CS}(A) = \mathcal{U}(A)$.

Proof. Since $\mathcal{U}(A) \subset \text{CS}(A)$, we only need to show the other containment. Let $B \in \text{CS}(A)$ and let $T, S \in GL_n$ be such that $B = T^*AT = S^{-1}AS$. Since $B = T^*AT$ and A is Hermitian, B is Hermitian. Since $B = S^{-1}AS$, A and B are similar, and so have the same eigenvalues (counting multiplicities). By Lemma 2.2.1, A and B are unitarily similar, and $\text{CS}(A) \subset \mathcal{U}(A)$. Therefore, $\text{CS}(A) = \mathcal{U}(A)$. \square

Corollary 2.2.3. If A is co-Hermitian, then $\text{CS}(A) = \mathcal{U}(A)$.

Proof. Since A is co-Hermitian, $A = \alpha B$ for some $\alpha \in \mathbb{C}$ and $B \in \mathcal{H}_n$. If $A = 0$, then this corollary is trivially true. So suppose $B \neq 0$ and $\alpha \neq 0$. Then $B = \frac{1}{\alpha}A$ is Hermitian, and so $\text{CS}(B) = \mathcal{U}(B)$ by Theorem 2.2.2. However, because scalar multiples α do not affect unitary similarity, similarity, or conjunctivity among two matrices, we have $\text{CS}(A) = \mathcal{U}(A)$. \square

The next natural step is to try to prove a similar result for the normal matrices. However, Example 2.1.1 has already shown that such a result is not possible. So what we would like to do is to reduce the problem to a very explicit case.

Proposition 2.2.4. Let $A \in M_n$, $n \geq 2$. Suppose A is conjunctive and similar to a block matrix $\begin{bmatrix} B & C \\ 0 & 0 \end{bmatrix}$, where $B \in M_r$, for $0 < r < n$, and $C \neq 0$. Then $\text{CS}(A)$ contains uncountably many disjoint unitary similarity classes.

Proof. Without loss of generality, we can assume that $A = \begin{bmatrix} B & C \\ 0 & 0 \end{bmatrix}$. Let $\alpha \in \mathbb{C}$ with $\alpha \neq 0$, and consider the block matrix $M = \begin{bmatrix} I_r & 0 \\ 0 & \alpha I_{n-r} \end{bmatrix}$. Then $M^*AM = M^{-1}AM = \begin{bmatrix} B & \alpha C \\ 0 & 0 \end{bmatrix} \in \text{CS}(A)$. Therefore, $\text{tr}((M^*AM)^*(M^*AM)) = \text{tr}(B^*B) + |\alpha|^2 \text{tr}(C^*C)$, and for $\alpha \neq 0$ we get a continuum of values since $\text{tr}(C^*C) \neq 0$. Therefore, by Theorem 1.3.1, $\text{CS}(A)$ contains uncountably many disjoint unitary similarity classes. \square

Proposition 2.2.5. Let $A, B \in GL_n$, $m \in \mathbb{N}$. Then $B \in \text{CS}(A)$ if and only if the block matrix $\begin{bmatrix} B & 0 \\ 0 & 0_m \end{bmatrix} \in \text{CS}\left(\begin{bmatrix} A & 0 \\ 0 & 0_m \end{bmatrix}\right)$.

Proof. First suppose that $B \in \text{CS}(A)$. There exist $T, S \in GL_n$ so that $B = T^*AT = S^{-1}AS$.

So

$$\begin{bmatrix} B & 0 \\ 0 & 0_m \end{bmatrix} = \begin{bmatrix} T & 0 \\ 0 & I_m \end{bmatrix}^* \begin{bmatrix} A & 0 \\ 0 & 0_m \end{bmatrix} \begin{bmatrix} T & 0 \\ 0 & I_m \end{bmatrix} = \begin{bmatrix} S & 0 \\ 0 & I_m \end{bmatrix}^{-1} \begin{bmatrix} A & 0 \\ 0 & 0_m \end{bmatrix} \begin{bmatrix} S & 0 \\ 0 & I_m \end{bmatrix}.$$

Conversely, let $T = \begin{bmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{bmatrix}$ and $S = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix}$ be non-singular block matrices, with $T_{11}, S_{11} \in M_n$, so that

$$\begin{bmatrix} B & 0 \\ 0 & 0_m \end{bmatrix} = T^* \begin{bmatrix} A & 0 \\ 0 & 0_m \end{bmatrix} T = \begin{bmatrix} T_{11}^*AT_{11} & T_{11}^*AT_{12} \\ T_{12}^*AT_{11} & T_{12}^*AT_{12} \end{bmatrix}$$

and

$$S \begin{bmatrix} B & 0 \\ 0 & 0_m \end{bmatrix} = \begin{bmatrix} S_{11}B & 0 \\ S_{21}B & 0_m \end{bmatrix} = \begin{bmatrix} AS_{11} & AS_{12} \\ 0 & 0_m \end{bmatrix} = \begin{bmatrix} A & 0 \\ 0 & 0_m \end{bmatrix} S.$$

By examining the (1,1)-entry of conjunctivity, we get $T_{11}^*AT_{11} = B$. However, since $A, B \in GL_n$, we know $T_{11} \in GL_n$, and so A and B are conjunctive. By examining the (1,2) and (2,1)-entries of similarity, we get $AS_{12} = 0$ and $S_{21}B = 0$. Again, since $A, B \in GL_n$, we know $S_{12} = 0$ and $S_{21} = 0$, and S is block diagonal with $S^{-1} = \begin{bmatrix} S_{11}^{-1} & 0 \\ 0 & S_{22}^{-1} \end{bmatrix}$. So $S_{11}^{-1}AS_{11} = B$, and A and B are similar. Therefore, $B \in \text{CS}(A)$. \square

Proposition 2.2.6. Let $A, B \in M_n$ and $C \in M_p$. Then $B \in \mathcal{U}(A)$ if and only if the block matrix $\begin{bmatrix} B & 0 \\ 0 & C \end{bmatrix} \in \mathcal{U}\left(\begin{bmatrix} A & 0 \\ 0 & C \end{bmatrix}\right)$.

Proof. By Theorem 1.3.1, $B \in \mathcal{U}(A)$ if and only if

$$\text{tr}(\omega(A, A^*)) = \text{tr}(\omega(B, B^*)), \quad (2.1)$$

for every word $\omega(x, y)$ in non-commuting variables x and y . Since

$$\omega\left(\begin{bmatrix} X & 0 \\ 0 & Y \end{bmatrix}, \begin{bmatrix} X & 0 \\ 0 & Y \end{bmatrix}^*\right) = \begin{bmatrix} \omega(X, X^*) & 0 \\ 0 & \omega(Y, Y^*) \end{bmatrix},$$

for any word $\omega(x, y)$, (2.1) is true if and only if

$$\operatorname{tr}(\omega\left(\begin{bmatrix} A & 0 \\ 0 & C \end{bmatrix}, \begin{bmatrix} A & 0 \\ 0 & C \end{bmatrix}^*\right)) = \operatorname{tr}(\omega\left(\begin{bmatrix} B & 0 \\ 0 & C \end{bmatrix}, \begin{bmatrix} B & 0 \\ 0 & C \end{bmatrix}^*\right)).$$

And by Theorem 1.3.1, this is true if and only if $\begin{bmatrix} B & 0 \\ 0 & C \end{bmatrix} \in \mathcal{U}\left(\begin{bmatrix} A & 0 \\ 0 & C \end{bmatrix}\right)$. \square

The following theorem shows that if the matrix is block diagonal, with one of the diagonal blocks being a zero matrix and the other being non-singular, then we need only focus our attention on the non-zero block.

Theorem 2.2.7. Let $A \in GL_n$, $m \in \mathbb{N}$. Then $\operatorname{CS}(A) = \mathcal{U}(A)$ if and only if

$$\operatorname{CS}\left(\begin{bmatrix} A & 0 \\ 0 & 0_m \end{bmatrix}\right) = \mathcal{U}\left(\begin{bmatrix} A & 0 \\ 0 & 0_m \end{bmatrix}\right).$$

Proof. Suppose $\operatorname{CS}(A) = \mathcal{U}(A)$. Without loss of generality, consider

$\begin{bmatrix} B & C \\ 0 & D \end{bmatrix} \in \operatorname{CS}\left(\begin{bmatrix} A & 0 \\ 0 & 0_m \end{bmatrix}\right)$, where $B \in GL_n$. (If the matrix is not in this block triangular form, then it can be placed in this form by unitary similarity. Since $A \in GL_n$, we can choose this unitary similarity to give $B \in GL_n$.) So there exist invertible matrices $S = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix}$

and $T = \begin{bmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{bmatrix}$, with $S_{11}, T_{11} \in M_n$, so that

$$\begin{bmatrix} B & C \\ 0 & D \end{bmatrix} = T^* \begin{bmatrix} A & 0 \\ 0 & 0 \end{bmatrix} T = \begin{bmatrix} T_{11}^* A T_{11} & T_{11}^* A T_{12} \\ T_{12}^* A T_{11} & T_{12}^* A T_{12} \end{bmatrix}$$

and

$$S \begin{bmatrix} B & C \\ 0 & D \end{bmatrix} = \begin{bmatrix} S_{11} B & S_{11} C + S_{12} D \\ S_{21} B & S_{21} C + S_{22} D \end{bmatrix} = \begin{bmatrix} A S_{11} & A S_{12} \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} A & 0 \\ 0 & 0 \end{bmatrix} S.$$

From the (1,1)-entry in conjunctivity, we have $B = T_{11}^* A T_{11}$, so that $T_{11} \in GL_n$. Looking at the (2,1)-entry of conjunctivity, we must have $T_{12} = 0$, and this gives $C = 0$ and $D = 0$.

Looking now at the (1,2) and (2,1)-entries of similarity, we see that $S_{12} = 0$ and $S_{21} = 0$, so

that we must have $S_{11} \in GL_n$. So $B = S_{11}^{-1} A S_{11} = T_{11}^* A T_{11}$, and hence $B \in \operatorname{CS}(A) = \mathcal{U}(A)$.

So by Proposition 2.2.6, $\begin{bmatrix} B & 0 \\ 0 & 0_m \end{bmatrix} \in \mathcal{U}\left(\begin{bmatrix} A & 0 \\ 0 & 0_m \end{bmatrix}\right)$ and $\operatorname{CS}\left(\begin{bmatrix} A & 0 \\ 0 & 0_m \end{bmatrix}\right) = \mathcal{U}\left(\begin{bmatrix} A & 0 \\ 0 & 0_m \end{bmatrix}\right)$.

Now suppose that $\text{CS}\left(\begin{bmatrix} A & 0 \\ 0 & 0_m \end{bmatrix}\right) = \mathcal{U}\left(\begin{bmatrix} A & 0 \\ 0 & 0_m \end{bmatrix}\right)$, and let $B \in \text{CS}(A)$. Then by Proposition 2.2.5 $\begin{bmatrix} B & 0 \\ 0 & 0_m \end{bmatrix} \in \text{CS}\left(\begin{bmatrix} A & 0 \\ 0 & 0_m \end{bmatrix}\right)$. So by Proposition 2.2.6, $B \in \mathcal{U}(A)$. Therefore, $\text{CS}(A) = \mathcal{U}(A)$. \square

At first glance, it would seem that the 0_m block in Theorem 2.2.7 can be replaced by any $m \times m$ matrix (e.g., I_m) with the result preserved. However, we will see later in Section 2.5.3 that this is not the case.

The result of our analysis so far is to reduce the problem to examining non-singular, non-co-Hermitian matrices. Any singular matrix has at least one zero eigenvalue, and so can be triangularized via unitary similarity to have one of the zero eigenvalues in the (n, n) -entry. If this places the matrix in the form of Proposition 2.2.4, then we know that $\text{CS}(A) \neq \mathcal{U}(A)$ and we are done. Otherwise, the matrix is in the form of Proposition 2.2.5, and Proposition 2.2.6 and Theorem 2.2.7 say we need only focus on the non-singular block matrix in the upper-left corner. Since we already know the result for co-Hermitian matrices, this leaves the non-co-Hermitian matrices to be understood.

Before we try to understand the general $n \times n$ case, let's look at a couple of specific sizes.

2.3 $\text{CS}(A)$: The 1×1 Case

Because of the commutativity of 1×1 matrices, this case is simple.

Proposition 2.3.1. If $A \in M_1$, then $\text{CS}(A) = \mathcal{U}(A) = A$.

Proof. Because of the commutativity of 1×1 matrices, A is the only matrix in $\text{CS}(A)$ and $\mathcal{U}(A)$. \square

The only possible canonical form for $\text{CS}(A)$ is A .

2.4 $\text{CS}(A)$: The 2×2 Case

By Corollary 2.2.3, we already know that $\text{CS}(A) = \mathcal{U}(A)$ for co-Hermitian matrices A . So for the 2×2 case, we want to understand the non-co-Hermitian matrices.

Before we go any further, it may be good to get another characterization of co-Hermitian matrices.

Lemma 2.4.1. If a matrix $A \in M_n$ is co-Hermitian, then all its eigenvalues are collinear in \mathbb{C} on a line passing through the origin. In particular, any singular, normal 2×2 matrix is co-Hermitian.

Proof. A proof of this lemma can be found as part of Proposition 2.7.1. If one eigenvalue of $A \in M_2$ is 0, then the two eigenvalues of A are collinear, hence any singular, normal 2×2 matrix is co-Hermitian. \square

With this lemma in place, we can now work to understand the 2×2 case.

Proposition 2.4.2. Let $A \in M_2$ be non-co-Hermitian and normal. Then $\text{CS}(A) = \mathcal{U}(A)$.

Proof. Without loss of generality, we may assume $A = \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}$, where $ab \neq 0$, and a and b are linearly independent over \mathbb{R} . Let $T, S \in GL_2$, with $T = \begin{bmatrix} x & y \\ z & w \end{bmatrix}$, be matrices such that $T^*AT = S^{-1}AS$. Then

$$T^*AT = \begin{bmatrix} x\bar{x}a + z\bar{z}b & \bar{x}ya + \bar{z}wb \\ x\bar{y}a + z\bar{w}b & y\bar{y}a + w\bar{w}b \end{bmatrix}.$$

Since T^*AT is similar to A , we must have that $\text{tr}(T^*AT) = \text{tr}(A)$ and $\det(T^*AT) = \det(A)$.

From the trace equation, we get

$$(x\bar{x} + y\bar{y})a + (z\bar{z} + w\bar{w})b = a + b.$$

But since a and b are linearly independent over \mathbb{R} , we must have $x\bar{x} + y\bar{y} = z\bar{z} + w\bar{w} = 1$, so that the rows of T are normal.

From the determinant equation, we get $\det(T^*T) = 1$ since A is non-singular. Now

$$T^*T = \begin{bmatrix} x\bar{x} + y\bar{y} & x\bar{z} + y\bar{w} \\ \bar{x}z + \bar{y}w & z\bar{z} + w\bar{w} \end{bmatrix} = \begin{bmatrix} 1 & x\bar{z} + y\bar{w} \\ \bar{x}z + \bar{y}w & 1 \end{bmatrix},$$

so

$$\det(T^*T) = 1 - |x\bar{z} + y\bar{w}|^2 = 1.$$

So $x\bar{z} + y\bar{w} = 0$, and the rows of T are orthonormal. Hence, by Theorem 2.1.4 in [HJ1], T is unitary and $\text{CS}(A) \subset \mathcal{U}(A)$. Therefore, $\text{CS}(A) = \mathcal{U}(A)$. \square

Proposition 2.4.3. Let $A \in M_2$ be non-singular and non-normal. Then $\text{CS}(A) = \mathcal{U}(A)$.

Proof. Since $\mathcal{U}(A) \subset \text{CS}(A)$ and A is non-normal, by Theorem 1.3.3 we may assume that A is a triangular matrix $\begin{bmatrix} a & r \\ 0 & b \end{bmatrix}$, with $ab \neq 0$ and $r > 0$. Let $T = \begin{bmatrix} x & y \\ z & w \end{bmatrix} \in GL_2$ and note that

$$T^*AT = \begin{bmatrix} ax\bar{x} + r\bar{x}z + bz\bar{z} & a\bar{x}y + r\bar{x}w + b\bar{z}w \\ ax\bar{y} + r\bar{y}z + bz\bar{w} & ay\bar{y} + r\bar{y}w + bw\bar{w} \end{bmatrix}. \quad (2.2)$$

Because T^*AT is similar to A , without loss of generality, we may assume that T has been chosen to make $T^*AT = \begin{bmatrix} a & s \\ 0 & b \end{bmatrix}$, for some $s \in \mathbb{C}$. (If T does not make T^*AT upper-triangular, then we can find $U \in U_n$ so that U^*T^*ATU is upper-triangular and relabel the matrix TU as T .)

Also, since $A, T \in GL_2$, we may normalize to get $\det(T) = 1$.

Looking at the (1, 1), (2, 1), and (2, 2)-entries in equation (2.2), along with the determinant condition on T , we get four equations that must be true for such a T :

$$ax\bar{x} + r\bar{x}z + bz\bar{z} = a \quad (2.3)$$

$$ax\bar{y} + r\bar{y}z + bz\bar{w} = 0 \quad (2.4)$$

$$ay\bar{y} + r\bar{y}w + bw\bar{w} = b \quad (2.5)$$

$$xw - yz = 1 \quad (2.6)$$

We first consider the case where $y \neq 0$. What we will show is that, for any such T , we must have $|s| = r$. Then, by Theorem 1.3.3, we will have $\text{CS}(A) = \mathcal{U}(A)$.

From (2.6), we get

$$z = \frac{xw - 1}{y}. \quad (2.7)$$

Substituting this into (2.4) and using (2.5), we get

$$x = \frac{r\bar{y} + b\bar{w}}{b}. \quad (2.8)$$

(Recall that $b \neq 0$, since $A \in GL_n$.) Also, if we multiply (2.5) by \bar{b} and subtract its complex conjugate from itself, we get the equation

$$r(\bar{b}y\bar{w} - by\bar{w}) + y\bar{y}(a\bar{b} - \bar{a}b) = 0. \quad (2.9)$$

We now want to look at the $(1, 2)$ -entry of T^*AT . (We will put in $[\dots]$ those expressions that will be replaced using (2.5) and in $\{\dots\}$ an expression that will be replaced using (2.9).)

$$\begin{aligned}
s &= a\bar{x}y + r\bar{x}w + b\bar{z}w \\
&= ay \left(\frac{ry + \bar{b}w}{\bar{b}} \right) + rw \left(\frac{ry + \bar{b}w}{\bar{b}} \right) + bw \left(\frac{\bar{x}w - 1}{\bar{y}} \right) \\
&= \frac{ary^2\bar{y} + a\bar{b}y\bar{y}w + r^2y\bar{y}w + \bar{b}r\bar{y}w^2 + bw[ry\bar{w} + \bar{b}w\bar{w} - \bar{b}]}{\bar{b}\bar{y}} \quad (\text{by (2.7)}) \\
&= \frac{ry[ay\bar{y} + r\bar{y}w] + a\bar{b}y\bar{y}w + \bar{b}r\bar{y}w^2 - a\bar{b}y\bar{y}w}{\bar{b}\bar{y}} \\
&= \frac{rby + w\{r(\bar{b}\bar{y}w - by\bar{w}) + y\bar{y}(a\bar{b} - \bar{a}b)\}}{\bar{b}\bar{y}} \\
&= \frac{by}{\bar{b}\bar{y}} r.
\end{aligned}$$

So we have $|s| = r$, and $\text{CS}(A) \subset \mathcal{U}(A)$. Therefore, $\text{CS}(A) = \mathcal{U}(A)$.

Now let $y = 0$. We will show in this case that such a T is always unitary, and so $\text{CS}(A) = \mathcal{U}(A)$.

Equations (2.3)–(2.6) simplify to:

$$ax\bar{x} + r\bar{x}z + bz\bar{z} = a \quad (2.10)$$

$$bz\bar{w} = 0 \quad (2.11)$$

$$bw\bar{w} = b \quad (2.12)$$

$$xw = 1 \quad (2.13)$$

From (2.12) and $b \neq 0$, we get $|w| = 1$. So (2.11) gives $z = 0$. Then equation (2.13) gives $|x| = 1$. Therefore, T is unitary, and $\text{CS}(A) = \mathcal{U}(A)$. \square

Proposition 2.4.4. Let $A \in M_2$ be singular and non-normal. Then $\text{CS}(A)$ contains uncountably many unitary similarity classes.

Proof. If A is singular and non-normal, then it can be unitarily triangularized to $\begin{bmatrix} a & r \\ 0 & 0 \end{bmatrix}$, for $r \neq 0$ and possibly $a = 0$. Then, by Theorem 2.2.4, $\text{CS}(A)$ contains uncountably many disjoint unitary similarity classes. \square

With these propositions in place, we are now ready to completely characterize the 2×2 case. Combining Corollary 2.2.3 and Propositions 2.4.2, 2.4.3, and 2.4.4, we get the following theorem.

Theorem 2.4.5. Let $A \in M_2$. If A is non-singular or normal (or both), then $\text{CS}(A) = \mathcal{U}(A)$. Otherwise, $\text{CS}(A)$ contains uncountably many disjoint unitary similarity classes.

This will not completely generalize to $n \times n$. Example 2.1.1 gives a 3×3 (or larger) normal matrix for which $\text{CS}(A) \neq \mathcal{U}(A)$.

Canonical forms for $\text{CS}(A)$ can be seen from the previous theorem. If A is non-singular or normal, then A is similar and conjunctive to a matrix of the form $\begin{bmatrix} \lambda_1 & r \\ 0 & \lambda_2 \end{bmatrix}$, where $r \geq 0$ and $\lambda_1, \lambda_2 \in \mathbb{C}$ are the eigenvalues of A . If A is singular and non-normal, then A is similar and conjunctive to a matrix of the form $\begin{bmatrix} \lambda & 1 \\ 0 & 0 \end{bmatrix}$, where λ (possibly 0) and 0 are the eigenvalues of A .

2.5 $\text{CS}(A)$: The $n \times n$ Case (Part 2)

After having settled the 1×1 and 2×2 cases, we would like to settle the $n \times n$ non-singular, non-co-Hermitian case. We have made some progress in this direction. What is presented below are the various approaches that we took and what we learned from each.

2.5.1 $\text{Sim}(A)$, $\text{Conj}(A)$, and $\mathcal{U}(A)$

Since $\text{CS}(A) = \text{Sim}(A) \cap \text{Conj}(A)$, we might look at $\text{Sim}(A)$ and $\text{Conj}(A)$ individually to see how many disjoint unitary similarity classes each contains. Perhaps this will shed some light on $\text{CS}(A)$.

Proposition 2.5.1. If $A \in M_n$ and $A \neq \alpha I$, then $\text{Sim}(A)$ contains uncountably many disjoint unitary similarity classes.

Proof. Let $A = (a_{ij})$. Since $A \neq \alpha I$, we know that $a_{kl} \neq 0$, for some $k \neq l$. Without loss of generality, we may assume $a_{12} \neq 0$. Let $S = \begin{bmatrix} x & \mathbf{0} \\ \mathbf{0} & I_{n-1} \end{bmatrix} \in GL_n$, for $x > 0$, and let

$B = S^{-1}AS = (b_{ij})$. Then $b_{1j} = \frac{1}{x}a_{1j}$, for $j = 2, \dots, n$, and $b_{i1} = xa_{i1}$, for $i = 2, \dots, n$. So

$$\operatorname{tr}(BB^*) = \operatorname{tr}(AA^*) + \left(\frac{1}{x^2} - 1\right) \sum_{j=2}^n |a_{1j}|^2 + (x^2 - 1) \sum_{i=2}^n |a_{i1}|^2.$$

Since $a_{12} \neq 0$, $\operatorname{tr}(BB^*)$ depends on x , and we can get a continuum of values for $\operatorname{tr}(BB^*)$. Therefore, by Theorem 1.3.1, $\operatorname{Sim}(A)$ contains uncountably many disjoint unitary similarity classes. \square

Proposition 2.5.2. If $A \in M_n$ and $A \neq 0$, then $\operatorname{Conj}(A)$ contains uncountably many disjoint unitary similarity classes.

Proof. Let $T = \alpha I$, for $\alpha > 0$, and let $B = T^*AT \in \operatorname{Conj}(A)$. So if $A = (a_{ij})$ and $B = (b_{ij})$, then $b_{ij} = \alpha^2 a_{ij}$. Therefore, $\operatorname{tr}(B) = \alpha^2 \operatorname{tr}(A)$, and, for $\alpha > 1$, we get a continuum of values for $\operatorname{tr}(B)$. Therefore, by Theorem 1.3.1, $\operatorname{Conj}(A)$ contains uncountably many disjoint unitary similarity classes. \square

We now know that for any Hermitian matrix $A \neq \alpha I$, $\operatorname{CS}(A) = \mathcal{U}(A)$. However, from the previous two propositions, we know that $\operatorname{Sim}(A) \neq \mathcal{U}(A)$ and $\operatorname{Conj}(A) \neq \mathcal{U}(A)$. Therefore, it seems that our individual knowledge of $\operatorname{Sim}(A)$ and $\operatorname{Conj}(A)$ does not shed any light on our problem.

2.5.2 Diagonal T and Triangular A

In the equation $T^*AT = S^{-1}AS$, we now suppose that T is a diagonal matrix and A is an upper-triangular matrix. Because $A, T \in GL_n$, we may assume that $\det(T) = 1$, so the product of the diagonal elements of T is 1.

In order for T^*AT to be similar to A , at the very least their eigenvalues must be the same (counting multiplicities). This means that T^*AT can only differ from A on the diagonal by a permutation of the diagonal elements. Let a_{jj} and t_{jj} denote the diagonal elements of A and T , respectively. The diagonal elements of T^*AT are $\overline{t_{jj}} a_{jj} t_{jj} = r_j a_{jj}$, for some $r_j \in \mathbb{R}$. Therefore, the only diagonal entries of A that can be permuted are those that are real multiples of one another. Assuming all the diagonal entries of A are real multiples of one another, this would give at most $n!$ possible diagonal matrices T .

One might think that the number of possible diagonal matrices T is equal to the number of unitary similarity classes contained in $\text{CS}(A)$. However the permutation matrix in Example 2.1.1, when put into a diagonal matrix, has diagonal entries $e^{\frac{2\pi j}{n}i}$, for $j = 0, 1, \dots, n-1$, and none of these entries is a real multiple of any other. So only the diagonal $T = I$ works to maintain similarity. However, $\text{CS}(A)$ contains uncountably many unitary similarity classes.

2.5.3 Direct Sum of 2×2 and Identity Matrices

Since we already understand the 2×2 case, we might consider matrices that are a direct sum of a 2×2 matrix and an identity matrix to see what we can get. However, we again look to the permutation matrix in Example 2.1.1 to show that this does not work. Consider the 3×3 permutation matrix

$$\begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}.$$

Since it is normal, it can be unitarily diagonalized to

$$A = \begin{bmatrix} e^{\frac{2\pi}{3}i} & 0 & 0 \\ 0 & e^{\frac{4\pi}{3}i} & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

which is a direct sum of the 2×2 matrix $B = \begin{bmatrix} e^{\frac{2\pi}{3}i} & 0 \\ 0 & e^{\frac{4\pi}{3}i} \end{bmatrix}$ and the 1×1 identity matrix. However, since B is non-co-Hermitian and normal, we know by Proposition 2.4.2 that $\text{CS}(B) = \mathcal{U}(B)$. But this is not true of A .

2.5.4 AM Similar to A

Suppose $T^*AT = S^{-1}AS$, for $A, S, T \in GL_n$. Then with a little algebra we can change this equation into

$$A(TT^*) = (ST^*)^{-1}A(ST^*). \quad (2.14)$$

Since $T \in GL_n$, the matrix TT^* is positive definite (i.e., TT^* is Hermitian with all eigenvalues positive). So we want to understand what positive definite matrices M would make AM similar to A . Of course, we can quickly find a necessary and sufficient condition on such matrices M not assumed to be positive definite.

Proposition 2.5.3. Let $A, M \in GL_n$. AM is similar to A if and only if $M = A^{-1}S^{-1}AS$, for some $S \in GL_n$.

Proof. If $AM = S^{-1}AS$, then $M = A^{-1}S^{-1}AS$. Conversely, if $M = A^{-1}S^{-1}AS$, then $AM = S^{-1}AS$. \square

So we see that a necessary and sufficient condition for such a matrix M is that it be a multiplicative commutator involving the matrix A . The natural next question to ask is, which multiplicative commutators are positive definite, or for which matrices $S \in GL_n$ is $A^{-1}S^{-1}AS$ positive definite?

We have some specific examples of such matrices S , though no general solution. An example of matrices S for which the multiplicative commutator $A^{-1}S^{-1}AS$ is positive definite is when $S = A^m B$, for integer m , and $B \in GL_n$ with $AB = BA$. Of course, this choice of S causes $A^{-1}S^{-1}AS$ to just collapse to the identity matrix I .

Because M is positive definite, it can be unitarily diagonalized as $M = U^*DU$, for some $U \in U_n$ and diagonal matrix D with positive real diagonal. So we can change equation (2.14) into

$$A'D = (UST^*U^*)^{-1}A'(UST^*U^*), \quad (2.15)$$

where $A' = UAU^*$, and we can now search for positive real diagonal matrices D for which $A'D$ is similar to A' . One thought is that perhaps the number of such diagonal matrices D equals the number of unitary similarity classes in $\text{CS}(A)$. However, the following examples show there is generally no such equality.

Example 2.5.4. Let $A = \begin{bmatrix} i & -i \\ -1 & 1 \end{bmatrix}$ which is non-normal and singular. So by Theorem 2.4.5, $\text{CS}(A)$ contains uncountably many unitary similarity classes. Let $D = \begin{bmatrix} x & 0 \\ 0 & y \end{bmatrix}$, for

$x, y > 0$. In this case, we can actually solve for the values of x and y that will make AD similar to A . $AD = \begin{bmatrix} ix & -iy \\ -x & y \end{bmatrix}$ and $\det(AD) = xy \det(A) = 0$. Therefore, for our choice of A , all we need to maintain similarity is for $\text{tr}(AD) = \text{tr}(A)$, which gives us $ix + y = i + 1$. Looking at the real and imaginary parts of this equation, we get immediately that $x = y = 1$. Therefore, the identity matrix I is the only diagonal matrix for which AD is similar to A , and the number of possible diagonal matrices is not equal to the number of unitary similarity classes contained in $\text{CS}(A)$.

Example 2.5.5. Let $A = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 0 & 1 \\ 2 & 1 & 0 \end{bmatrix}$. By Corollary 2.2.3, we know that $\text{CS}(A) = \mathcal{U}(A)$.

Let

$$D = \begin{bmatrix} x & 0 & 0 \\ 0 & y & 0 \\ 0 & 0 & z \end{bmatrix},$$

for $x, y, z > 0$, and consider the matrix AD . If we assume that AD is similar to A , then the two characteristic polynomials will be the same. Recall that the coefficient of λ^k , for $k < n$, in the characteristic polynomial is the sum of the $(n - k) \times (n - k)$ principal minors. Equating coefficients from the characteristic polynomials of AD and A gives:

$$0 = 0$$

$$-xy - 4xz - yz = -6$$

$$xyz = 1$$

Solving for z in the last equation, substituting into the second equation, and solving for y gives

$$y = \frac{(6x - 1) \pm \sqrt{-16x^3 + 36x^2 - 12x + 1}}{2x^2}.$$

Since we wish to have $x, y, z > 0$, we want both the radicand and $6x - 1$ to be positive. The zeros of the radicand are $x = \frac{1}{4}$ and $x = \frac{2 \pm \sqrt{3}}{2}$. By graphing the radicand, we see that it is positive for $0 < x < \frac{2 - \sqrt{3}}{2}$ and $\frac{1}{4} < x < \frac{2 + \sqrt{3}}{2}$. Also, we have $6x - 1 > 0$ for $x > \frac{1}{6}$. Therefore, we get positive x, y, z for $\frac{1}{4} < x < \frac{2 + \sqrt{3}}{2}$. However, what is more important is

that we get uncountably many possible diagonal matrices D for which AD is similar to A . Therefore, the number of unitary similarity classes contained in $\text{CS}(A)$ is not equal to the number of possible diagonal matrices.

Example 2.5.6. Let's consider the permutation matrix example (2.1.1) and ask what positive diagonal matrices D will make AD similar to A . If we let the diagonal entries of D be d_i , for $1 \leq i \leq n$, then the characteristic polynomial for AD can be seen to be $\lambda^n - (d_1 d_2 \cdots d_n)$. So AD is similar to A if and only if $\det(D) = 1$. Here is one case where the number of such diagonal matrices does match the number of unitary similarity classes contained in $\text{CS}(A)$, but the previous two examples show that this is not generally true.

2.5.5 A in Jordan Canonical Form

Suppose that $R \in GL_n$ is the matrix for which RAR^{-1} is the Jordan canonical form for the matrix A . From $T^*AT = S^{-1}AS$, we get

$$(TS^*R^*)^*A(TS^*R^*) = (RAR^{-1})(RSS^*R^*)$$

and the matrix RSS^*R^* is positive definite. Therefore, we want to know what positive definite matrices M will make JM conjunctive to A , where J is the Jordan canonical form of A . However, this was not pursued further, since it did not seem to shed any real light on our problem.

2.5.6 Conclusions

The evidence suggests that if A is non-singular and non-co-Hermitian, then $\text{CS}(A)$ will contain uncountably many unitary similarity classes. However, as of this writing, we have been unable to prove this.

2.6 Invariants of $\text{CS}(A)$

In Chapter 1, we discussed the invariants of both $\text{Sim}(A)$ and $\text{Conj}(A)$. Certainly, the union of the invariants of $\text{Sim}(A)$ and $\text{Conj}(A)$ serve as invariants for $\text{CS}(A)$. However, many

of these are difficult to calculate. We will present a more usable set of invariants when $n = 2$. But first we present a lemma that will help with these invariants.

Lemma 2.6.1. Let $A, B \in M_2$. The following conditions are equivalent:

- (a) A is normal.
- (b) Every matrix $B \in \text{CS}(A)$ is normal.
- (c) There exists some normal matrix $B \in \text{CS}(A)$.

Proof.

- (a) \Rightarrow (b): If A is normal, then, by Theorem 2.4.5, $\text{CS}(A) = \mathcal{U}(A)$. Therefore, any $B \in \text{CS}(A)$ is normal, since $B = U^*AU$, for some unitary U .
- (b) \Rightarrow (c): Obvious.
- (c) \Rightarrow (a): Suppose there exists a normal matrix $B \in \text{CS}(A)$. Recall that $\text{CS}(B) = \text{CS}(A)$, and since B is normal, $\text{CS}(B) = \mathcal{U}(B)$. So $A = U^*BU$, for some unitary U , and A is normal.

□

Proposition 2.6.2. Let $A, B \in M_2$.

- (1) If both A and B are non-singular or normal (or both), then $B \in \text{CS}(A)$ if and only if:
 - (a) $\text{tr}(A) = \text{tr}(B)$;
 - (b) $\text{tr}(A^2) = \text{tr}(B^2)$; and
 - (c) $\text{tr}(AA^*) = \text{tr}(BB^*)$.
- (2) If both A and B are singular and non-normal, then $B \in \text{CS}(A)$ if and only if $\text{tr}(A) = \text{tr}(B)$.

Proof. First note that Lemma 2.6.1 says that when $n = 2$, for non-normal B and normal A (or vice versa), we cannot have $B \in \text{CS}(A)$. So it is appropriate to divide this Proposition into

the case where both A and B are non-singular or normal (or both) and the case where both A and B are singular and normal.

- (1) Since both A and B are non-singular or normal (or both), we know by Theorem 2.4.5 that $\text{CS}(A) = \mathcal{U}(A)$. Therefore, we need only apply the 2×2 unitary invariants given in [Mu], which are the trace conditions (a), (b), and (c).
- (2) Two singular and non-normal 2×2 matrices are similar if and only if they have the same eigenvalues. But they each already have an eigenvalue of 0, so that the other eigenvalue is the trace. By Theorem 1.3.3, we may assume that the two matrices are in the form $A = \begin{bmatrix} \lambda & r \\ 0 & 0 \end{bmatrix}$ and $B = \begin{bmatrix} \lambda & q \\ 0 & 0 \end{bmatrix}$, with $r, q > 0$. Then the matrix $T = \begin{bmatrix} 1 & 0 \\ 0 & \frac{q}{r} \end{bmatrix}$ gives $B = T^*AT$. Therefore, two 2×2 singular, non-normal matrices that are similar are always conjunctive. Hence, we need to only check that $\text{tr}(A) = \text{tr}(B)$.

□

It should be noted that, in the $n \times n$ co-Hermitian case, the matrix pencil considered in the canonical form for conjunctivity is of a much simpler form. Let $A \in \mathcal{H}_n$ and $\alpha \in \mathbb{C}$. Then the matrix pencil $\mu\text{Re}(\alpha A) + \lambda\text{Im}(\alpha A)$ equals $(\mu\text{Re}(\alpha) + \lambda\text{Im}(\alpha))A$. Therefore, the elementary divisors are simply n copies of $\mu\text{Re}(\alpha) + \lambda\text{Im}(\alpha)$. Perhaps this may relate to the fact that $\text{CS}(\alpha A) = \mathcal{U}(\alpha A)$.

2.7 The Converse Problem

We know that if A is co-Hermitian, then $\text{CS}(A) = \mathcal{U}(A)$. Here we look for a converse. In the process of trying to prove the converse, some properties of co-Hermitian matrices are needed.

Proposition 2.7.1. Let $A \in M_n$. The following conditions are equivalent:

- (a) A is co-Hermitian.
- (b) $V(A) = \{x^*Ax \mid x \in \mathbb{C}^n\} \subset \alpha\mathbb{R}$, for some $\alpha \in \mathbb{C}$.

- (c) A is normal and all the eigenvalues of A lie on a line in \mathbb{C} that passes through the origin.
- (d) S^*AS is co-Hermitian for all $S \in M_n$.

Proof.

(a) \Rightarrow (b): Since A is co-Hermitian, there exist $\alpha \in \mathbb{C}$ and $H \in \mathcal{H}_n$ with $A = \alpha H$. So, for any $x \in \mathbb{C}^n$, $x^*Ax = \alpha x^*Hx \in \alpha\mathbb{R}$.

(b) \Rightarrow (c): Without loss of generality, we may assume $|\alpha| = 1$. Let $\sigma(M)$ denote the set of eigenvalues of a matrix $M \in M_n$ and note that $\sigma(M) \subset V(M)$. Hence $\sigma(A) \subset \alpha\mathbb{R}$. Now $V(\bar{\alpha}A) \subset \mathbb{R}$ and hence $V(\alpha A^*) \subset \mathbb{R}$, so that $\bar{\alpha}A = H \in \mathcal{H}_n$. Therefore, $A = \alpha H$ is normal.

(c) \Rightarrow (a): We have

$$A = U^* \begin{bmatrix} \alpha r_1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \alpha r_n \end{bmatrix} U = \alpha U^* \begin{bmatrix} r_1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & r_n \end{bmatrix} U,$$

for some $U \in U_n$, $\alpha \in \mathbb{C}$, and $r_1, \dots, r_n \in \mathbb{R}$. Therefore, A is co-Hermitian.

(a) \Leftrightarrow (d): Obvious. □

There is another property of co-Hermitian matrices that we will use to prove the converse to Corollary 2.2.3 in a special case.

Lemma 2.7.2. Let $A = U^*\Lambda U$, for some $U \in U_n$ and diagonal $\Lambda \neq 0$. Let $\mathcal{M} \subseteq M_n$ be a set of matrices such that, for every $j = 2, \dots, n$, there is a matrix $M \in \mathcal{M}$ with $(UMU^*)_{1j} \neq 0$. Then A is co-Hermitian if and only if $A^*MA = AMA^*$, for every $M \in \mathcal{M}$.

Proof. Suppose A is co-Hermitian. So there exists $\alpha \in \mathbb{C}$ and $H \in \mathcal{H}_n$ with $A = \alpha H$. Then $A^*MA = \bar{\alpha}\alpha HMH = AMA^*$, for every $M \in \mathcal{M}$.

Now suppose the converse and fix k , with $k \neq 1$, and let $M \in \mathcal{M}$ be such that $(UMU^*)_{1k} \neq$

0. Let

$$\Lambda = \begin{bmatrix} \lambda_1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \lambda_n \end{bmatrix},$$

with $\lambda_1 \neq 0$. Since

$$U^*\Lambda^*UMU^*\Lambda U = A^*MA = AMA^* = U^*\Lambda UMU^*\Lambda^*U,$$

we have that

$$\Lambda^*(UMU^*)\Lambda = \Lambda(UMU^*)\Lambda^*.$$

Let $UMU^* = B = (b_{ij})$. So $\Lambda^*B\Lambda = (\overline{\lambda_i}b_{ij}\lambda_j)$ and $\Lambda B\Lambda^* = (\lambda_i b_{ij} \overline{\lambda_j})$. Looking at the $(1, k)$ -entry of each matrix, we see that we must have $\overline{\lambda_1}b_{1k}\lambda_k = \lambda_1 b_{1k} \overline{\lambda_k}$. If $\lambda_k = 0$, then clearly λ_k and λ_1 are collinear in \mathbb{C} . If $\lambda_k \neq 0$, then because $b_{1k} \neq 0$ (by our choice of M), we must have $\overline{\lambda_1}\lambda_k = \lambda_1\overline{\lambda_k} = \overline{\overline{\lambda_1}\lambda_k}$. So $\overline{\lambda_1}\lambda_k \in \mathbb{R}$. But this can happen if and only if λ_k and λ_1 are collinear in \mathbb{C} , or $\lambda_k = r\lambda_1$ for some $r \in \mathbb{R}$. Since k is arbitrary, we see that all the eigenvalues of A lie on the line through the origin in \mathbb{C} that contains λ_1 . So the eigenvalues of A all lie on a line through the origin in \mathbb{C} , and therefore, by Proposition 2.7.1, A is co-Hermitian. \square

This lemma allows us to get a small result as the converse to Corollary 2.2.3.

Proposition 2.7.3. Let $A = U^*\Lambda U$, for some $U \in U_n$ and diagonal $\Lambda \neq 0$. Suppose that $\text{CS}(A) = \mathcal{U}(A)$, and that for any $j = 2, \dots, n$, there is a matrix $T \in GL_n$ so that $T^*AT \in \text{CS}(A)$ and $(UTT^*U^*)_{1j} \neq 0$. Then A is co-Hermitian.

Proof. Fix k , with $k \neq 1$, and choose $T \in GL_n$ so that $T^*AT \in \text{CS}(A)$ and $(UTT^*U^*)_{1k} \neq 0$. Since A is normal and $\text{CS}(A) = \mathcal{U}(A)$, we know that $T^*AT = V^*AV$, for some $V \in U_n$, and so T^*AT is also normal. Therefore, $(T^*AT)(T^*A^*T) = (T^*A^*T)(T^*AT)$, and because $T \in GL_n$ we have

$$ATT^*A^* = A^*TT^*A.$$

If we let $TT^* = M$, then we have that $AMA^* = A^*MA$, and $(UMU^*)_{1k} \neq 0$. However, since k is arbitrary, this can be done for any such k . Therefore, by Lemma 2.7.2, A is co-Hermitian. \square

CHAPTER 3

THE SIMILARITY-UNITARY EQUIVALENCE EQUIVALENCE CLASS

For $A \in M_n$, consider the equivalence class intersection $\text{Sim}(A) \cap \text{UEquiv}(A)$, which we will denote by $\text{UES}(A)$. This intersection is all matrices $B \in M_n$ such that $B = S^{-1}AS = U^*AV$, for some $S \in GL_n$ and $U, V \in U_n$. Again, we know that this intersection is non-empty because $A \in \text{UES}(A)$, and if $A = \alpha I$, then A is the only element of $\text{UES}(A)$. Of course, we again have that $\mathcal{U}(A) \subset \text{UES}(A)$.

3.1 $\text{UES}(A)$ and $\mathcal{U}(A)$

As was discussed in Chapter 1, a natural first thought is that perhaps $\text{UES}(A) = \mathcal{U}(A)$, for all A . However, as in the $\text{CS}(A)$ case, we have an example to show that this is not true.

Example 3.1.1. Let $A = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 2 \end{bmatrix}$, $W_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & e^{i\theta} & 0 \\ 0 & 0 & e^{i\theta} \end{bmatrix}$, and $W_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & e^{i\theta} \end{bmatrix}$.

Then

$$B_\theta = W_1^* A W_2 = \begin{bmatrix} 1 & 2 & e^{i\theta} \\ 0 & 0 & 1 \\ 0 & 0 & 2 \end{bmatrix},$$

and B_θ is both similar and unitarily equivalent to A . However, $\text{tr}(B_\theta^2 B_\theta^*) = 18 + 2e^{-i\theta}$, so we can get a continuum of values according to the choice of θ . Therefore, by Theorem 1.3.1, $\text{UES}(A)$ contains uncountably many disjoint unitary similarity classes.

Considering this example, it is natural to ask what matrices A have $\text{UES}(A) = \mathcal{U}(A)$.

3.2 UES(A): The $n \times n$ Case

In the general $n \times n$ case, we seek to characterize those matrices A for which $\text{UES}(A) = \mathcal{U}(A)$. For those matrices not satisfying this equation, we then want to know how many disjoint unitary similarity classes are contained in $\text{UES}(A)$. If the previous example is any guide, when $\text{UES}(A) \neq \mathcal{U}(A)$ we might expect to get uncountably many disjoint unitary similarity classes.

We need a lemma to help us get started with this case. Recall that a matrix A is *positive definite* (*positive semi-definite*), if $A \in \mathcal{H}_n$ and all the eigenvalues of A are positive (non-negative). Note that the set of positive semi-definite matrices contains the set of positive definite matrices.

Lemma 3.2.1. Let $A \in M_n$. Then A is positive semi-definite if and only if the eigenvalues and singular values of A coincide (counting multiplicities).

Proof. Suppose A is positive semi-definite. Then there exists $U \in U_n$ with

$$U^*AU = \begin{bmatrix} \lambda_1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \lambda_n \end{bmatrix},$$

where $\lambda_1 \geq \dots \geq \lambda_n \geq 0$ are the eigenvalues of A . The singular values are the non-negative square roots of the eigenvalues of the matrix

$$AA^* = U^* \begin{bmatrix} \lambda_1^2 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \lambda_n^2 \end{bmatrix} U,$$

which are $\lambda_1^2, \dots, \lambda_n^2$. So the singular values are also $\lambda_1, \dots, \lambda_n$.

Conversely, suppose the eigenvalues and singular values coincide (counting multiplicities), and let them be represented by $\lambda_1 \geq \dots \geq \lambda_n \geq 0$. A can be triangularized via unitary similarity and, without loss of generality, let

$$A = \begin{bmatrix} \lambda_1 & & a_{ij} \\ & \ddots & \\ \mathbf{0} & & \lambda_n \end{bmatrix}.$$

If $a_{ij} = 0$ for $1 \leq i < j \leq n$, then A is Hermitian and so is positive semi-definite. Since the singular values are the square roots of the eigenvalues of AA^* , we know that AA^* has the characteristic polynomial

$$p(\lambda) = (\lambda - \lambda_1^2) \dots (\lambda - \lambda_n^2).$$

Now the coefficient of λ^{n-1} in the characteristic polynomial is known to be

$$-\operatorname{tr}(AA^*) = -\sum_{i=1}^n \lambda_i^2 - \sum_{1 \leq i < j \leq n} |a_{ij}|^2.$$

However, we see from $p(\lambda)$ that the coefficient of λ^{n-1} is $-\sum_{i=1}^n \lambda_i^2$. Therefore, $\sum_{1 \leq i < j \leq n} |a_{ij}|^2 = 0$, and $a_{ij} = 0$, for $1 \leq i < j \leq n$. Therefore, A is Hermitian and hence positive semi-definite. \square

Proposition 3.2.2. Let $A \in M_n$ be positive semi-definite. Then $\operatorname{UES}(A) = \mathcal{U}(A)$.

Proof. Clearly, $\mathcal{U}(A) \subset \operatorname{UES}(A)$. Let $B \in \operatorname{UES}(A)$. Then by Lemma 3.2.1, B is also positive semi-definite, and so it is Hermitian. By Lemma 2.2.1, since A and B share eigenvalues (counting multiplicities), $B \in \mathcal{U}(A)$. Therefore, $\operatorname{UES}(A) = \mathcal{U}(A)$. \square

Corollary 3.2.3. Let $A \in M_n$ be positive semi-definite and $\alpha \in \mathbb{C}$. Then $\operatorname{UES}(\alpha A) = \mathcal{U}(\alpha A)$.

Proof. If $\alpha = 0$, then this Corollary is true. So assume $\alpha \neq 0$ and let $B \in \operatorname{UES}(\alpha A)$. Then $\frac{1}{\alpha}B \in \operatorname{UES}(A)$ and, by Proposition 3.2.2, $\frac{1}{\alpha}B \in \mathcal{U}(A)$. So $B \in \mathcal{U}(\alpha A)$ and $\operatorname{UES}(\alpha A) \subset \mathcal{U}(\alpha A)$. Therefore, $\operatorname{UES}(\alpha A) = \mathcal{U}(\alpha A)$. \square

So we see that, for any matrix A that is a scalar multiple of a positive semi-definite matrix, $\operatorname{UES}(A)$ is simply the unitary similarity class of A . We would like to extend this result to a different class of matrices, and with the help of a lemma, we can do so.

Lemma 3.2.4. Let $A \in M_n$, $S \in GL_n$, and let S have the polar decomposition $S = PW$, for some positive definite matrix P and unitary matrix W . If A and P commute, then $S^{-1}AS = W^*AW$.

Proof. $S^{-1}AS = W^*P^{-1}APW = W^*P^{-1}PAW = W^*AW$. \square

With this lemma in place, we can now understand the case when A is unitary.

Proposition 3.2.5. If $A \in U_n$, then $\text{UES}(A) = \mathcal{U}(A)$.

Proof. Let $S \in GL_n$ and $U, V \in U_n$ be such that $B = S^{-1}AS = U^*AV$. Then B is also unitary. So $B^{-1} = B^*$ and we get that $S^{-1}A^{-1}S = S^*A^*S^{-*}$. But $A^* = A^{-1}$ so that $A^*(SS^*) = (SS^*)A^*$ or $(SS^*)A = A(SS^*)$, and A commutes with SS^* . However, in the polar decomposition of $S = PW$, we have that $P = (SS^*)^{\frac{1}{2}}$. Also, by Theorem 7.2.6 in [HJ1], we know that P is a polynomial in SS^* . Therefore, A and P commute. So by Lemma 3.2.4, $B = W^*AW$ and $\text{UES}(A) \subset \mathcal{U}(A)$. Therefore, $\text{UES}(A) = \mathcal{U}(A)$. \square

Corollary 3.2.6. If $A \in U_n$ and $\alpha \in \mathbb{C}$, then $\text{UES}(\alpha A) = \mathcal{U}(\alpha A)$.

Proof. If $\alpha = 0$, then this Corollary is true. So assume $\alpha \neq 0$ and let $B \in \text{UES}(\alpha A)$. Then $\frac{1}{\alpha}B \in \text{UES}(A)$, and by Proposition 3.2.5, $\frac{1}{\alpha}B = W^*AW$, for some $W \in U_n$. So $B = W^*(\alpha A)W$, and $\text{UES}(\alpha A) \subset \mathcal{U}(\alpha A)$. Therefore, $\text{UES}(\alpha A) = \mathcal{U}(\alpha A)$. \square

So we now know that if a matrix A is a multiple of a unitary matrix, then $\text{UES}(A)$ is the unitary similarity class of A . At this point, we understand what $\text{UES}(A)$ is for a subset of the normal matrices. With the help of a lemma, we can actually understand $\text{UES}(A)$ for all the normal matrices A .

Lemma 3.2.7. Let $A \in M_n$ be normal. If $B \in \text{UES}(A)$, then B is normal.

Proof. Since A is normal, we may assume that

$$A = \begin{bmatrix} \lambda_1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \lambda_n \end{bmatrix}.$$

Let $B \in \text{UES}(A)$ and note that, since A and B are similar, we may assume that

$$B = \begin{bmatrix} \lambda_1 & & b_{ij} \\ & \ddots & \\ \mathbf{0} & & \lambda_n \end{bmatrix}.$$

However, since $B \in \text{UES}(A)$, A and B must also share singular values (counting multiplicities).

So AA^* and BB^* must have the same characteristic polynomial

$$p(\lambda) = (\lambda - |\lambda_1|^2) \dots (\lambda - |\lambda_n|^2).$$

The coefficient of λ^{n-1} in $p(\lambda)$ can be seen to be $-\sum_{i=1}^n |\lambda_i|^2$. But we also know that this

coefficient is $-\text{tr}(BB^*) = -\sum_{i=1}^n |\lambda_i|^2 - \sum_{1 \leq i < j \leq n} |b_{ij}|^2$. Therefore, $b_{ij} = 0$, for $1 \leq i < j \leq n$, and

B is also normal. \square

Theorem 3.2.8. If $A \in M_n$ is normal, then $\text{UES}(A) = \mathcal{U}(A)$.

Proof. By Lemma 3.2.7, we know that every matrix in $\text{UES}(A)$ is normal. So if $B \in \text{UES}(A)$, then

$$U^*AU = V^*BV = \begin{bmatrix} \lambda_1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \lambda_n \end{bmatrix},$$

for some $U, V \in U_n$. So $B = (UV^*)^*A(UV^*)$, and $\text{UES}(A) \subset \mathcal{U}(A)$. Therefore, $\text{UES}(A) = \mathcal{U}(A)$. \square

Note that this Theorem has as corollaries Corollaries 3.2.3 and 3.2.6. Also, this characterization includes all co-Hermitian matrices. So, unlike $\text{CS}(A)$, we are able to understand $\text{UES}(A)$ for all normal matrices.

In an attempt to reduce the remainder of the problem to a special type of matrix, we present two propositions.

Proposition 3.2.9. If $A \in M_n$ and $B \in \text{UES}(A)$ with $m \in \mathbb{N}$, then $\begin{bmatrix} B & 0 \\ 0 & 0_m \end{bmatrix} \in \text{UES}\left(\begin{bmatrix} A & 0 \\ 0 & 0_m \end{bmatrix}\right)$.

Proof. Let $S \in GL_n$ and $U, V \in U_n$ be such that $B = S^{-1}AS = U^*AV$. Then

$$\begin{bmatrix} B & 0 \\ 0 & 0_m \end{bmatrix} = \begin{bmatrix} S & 0 \\ 0 & I_m \end{bmatrix}^{-1} \begin{bmatrix} A & 0 \\ 0 & 0_m \end{bmatrix} \begin{bmatrix} S & 0 \\ 0 & I_m \end{bmatrix} = \begin{bmatrix} U & 0 \\ 0 & I_m \end{bmatrix}^* \begin{bmatrix} A & 0 \\ 0 & 0_m \end{bmatrix} \begin{bmatrix} V & 0 \\ 0 & I_m \end{bmatrix}.$$

\square

Proposition 3.2.10. Let $A \in GL_n$, $m \in \mathbb{N}$. If $\begin{bmatrix} A & 0 \\ 0 & 0_m \end{bmatrix}$ is both similar and unitarily equivalent to the block matrix $\begin{bmatrix} B & 0 \\ 0 & 0_m \end{bmatrix}$ with $B \in M_n$, then $B \in \text{UES}(A)$, and hence is non-singular.

Proof. That B is non-singular is a consequence of the similarity of the two block matrices. Let $S = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix}$, $U = \begin{bmatrix} U_{11} & U_{12} \\ U_{21} & U_{22} \end{bmatrix}$, and $V = \begin{bmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{bmatrix}$, with S non-singular, U, V unitary, and $S_{11}, U_{11}, V_{11} \in M_n$, so that

$$\begin{bmatrix} S_{11}B & 0 \\ S_{21}B & 0_m \end{bmatrix} = S \begin{bmatrix} B & 0 \\ 0 & 0_m \end{bmatrix} = \begin{bmatrix} A & 0 \\ 0 & 0_m \end{bmatrix} S = \begin{bmatrix} AS_{11} & AS_{12} \\ 0 & 0_m \end{bmatrix}$$

and

$$\begin{bmatrix} B & 0 \\ 0 & 0_m \end{bmatrix} = U^* \begin{bmatrix} A & 0 \\ 0 & 0_m \end{bmatrix} V = \begin{bmatrix} U_{11}^*AV_{11} & U_{11}^*AV_{12} \\ U_{12}^*AV_{11} & U_{12}^*AV_{12} \end{bmatrix}.$$

Looking first at the (1,1)-entry of unitary equivalence, we see that $U_{11}AV_{11} = B$, so that $U_{11}, V_{11} \in GL_n$. The (2,1)-entry of unitary equivalence gives $U_{12}^*AV_{11} = 0$ and the (1,2)-entry gives $U_{11}^*AV_{12} = 0$. So we know $U_{12} = 0$ and $V_{12} = 0$, and U and V are block lower-triangular. But this makes both U_{11} and V_{11} unitary, and so A and B are unitarily equivalent.

Looking at the (1,2) and (2,1)-entries of similarity, we see that $AS_{12} = 0$ and $S_{21}B = 0$. So $S_{12} = 0$ and $S_{21} = 0$, and S is block diagonal with $S^{-1} = \begin{bmatrix} S_{11}^{-1} & 0 \\ 0 & S_{22}^{-1} \end{bmatrix}$. Hence, A and B are similar.

Therefore, $B \in \text{UES}(A)$. □

Combining these two propositions, we get the following theorem.

Theorem 3.2.11. Let $A \in GL_n$, $m \in \mathbb{N}$. Then $B \in \text{UES}(A)$ if and only if

$$\begin{bmatrix} B & 0 \\ 0 & 0_m \end{bmatrix} \in \text{UES}\left(\begin{bmatrix} A & 0 \\ 0 & 0_m \end{bmatrix}\right).$$

One more step in the reduction process is gained by the following lemma, corollary, and propositions. We will let $\sigma(A)$ denote the spectrum or set of eigenvalues of the matrix A .

Lemma 3.2.12. Consider the block matrix $M = \begin{bmatrix} A & B \\ 0 & C \end{bmatrix} \in M_n$, with $A \in M_k$ and $C \in$

M_{n-k} , for some $1 \leq k < n$. If $\sigma(A) \cap \sigma(C) = \emptyset$, then M is similar to the matrix $\begin{bmatrix} A & 0 \\ 0 & C \end{bmatrix}$.

Proof. Let $S = \begin{bmatrix} I_k & X \\ 0 & I_{n-k} \end{bmatrix}$. Then

$$S^{-1}MS = \begin{bmatrix} A & AX + B - XC \\ 0 & C \end{bmatrix}.$$

We would like to find X so that

$$AX + B - XC = 0$$

or

$$AX - XC = -B. \quad (3.1)$$

Since $\sigma(A) \cap \sigma(C) = \emptyset$, by Theorem 4.4.6 of [HJ2] we can find a matrix X that satisfies (3.1).

So for this choice of X , we get that $S^{-1}MS = \begin{bmatrix} A & 0 \\ 0 & C \end{bmatrix}$. \square

Corollary 3.2.13. Two block upper-triangular matrices $\begin{bmatrix} A & B \\ 0 & C \end{bmatrix}$ and $\begin{bmatrix} X & Y \\ 0 & Z \end{bmatrix}$, with $\sigma(A) \cap \sigma(C) = \sigma(X) \cap \sigma(Z) = \emptyset$, are similar if and only if A and X are similar and C and Z are similar.

Proof. By Lemma 3.2.12, we can place each block matrix into block diagonal form via similarity.

Then they are similar if and only if these block diagonal matrices are similar. \square

Proposition 3.2.14. Let $A \in M_n$ be unitarily similar to

$$\left[\begin{array}{c|c|c} E & v_0 & V \\ \hline 0 & 0_1 & b \\ \hline 0 & 0 & L \end{array} \right],$$

where $0 < k \leq n-1$, $E \in GL_k$, $L \in M_{n-k-1}$, $v_0 \in \mathbb{C}^{k \times 1}$, $b \in \mathbb{C}^{1 \times (n-k-1)}$, $\sigma(E) \cap \sigma(L) = \emptyset$, and $bV^*v_0 \neq 0$. Then $\text{UES}(A)$ contains uncountably many unitary similarity classes.

Proof. Let $W_1 = \begin{bmatrix} I_k & 0 \\ 0 & e^{i\alpha}I_{p+1} \end{bmatrix}$ and $W_2 = \begin{bmatrix} I_{k+1} & 0 \\ 0 & e^{i\alpha}I_p \end{bmatrix}$. So

$$B = W_1^*AW_2 = \left[\begin{array}{c|c|c} E & v_0 & e^{i\alpha}V \\ \hline 0 & 0_1 & b \\ \hline 0 & 0 & L \end{array} \right]$$

is similar to A by Corollary 3.2.13, since $\sigma(E) \cap \sigma\left(\begin{bmatrix} 0_1 & b \\ 0 & L \end{bmatrix}\right) = \emptyset$. But

$\text{tr}(B^2B^*) = z_0 + e^{-i\alpha}bV^*v_0$, for some constant $z_0 \in \mathbb{C}$, and since $bV^*v_0 \neq 0$, this trace is dependent on α . Therefore, we can get a continuum of values for $\text{tr}(B^2B^*)$ and, by Theorem 1.3.1, $\text{UES}(A)$ contains uncountably many disjoint unitary similarity classes. \square

While the condition $bV^*v_0 \neq 0$ in Proposition 3.2.14 does ensure that A is non-normal, it is not chosen for that purpose. Instead, as seen in the proof, this condition forces $\text{tr}(B^2B^*)$ to be dependent on α , hence allowing for a continuum of values and uncountably many disjoint unitary similarity classes in $\text{UES}(A)$.

The result of Theorems 3.2.8 and 3.2.11 and Proposition 3.2.14 is to reduce the problem to:

- (1) non-singular, non-normal matrices A ;
- (2) singular, non-normal matrices A that are unitarily similar to a matrix of the form

$$\left[\begin{array}{c|c|c} E & v_0 & V \\ \hline 0 & 0_1 & b \\ \hline 0 & 0 & L \end{array} \right],$$

where $0 < k \leq n-1$, $E \in GL_k$, $L \in M_{n-k-1}$, $\sigma(E) \cap \sigma(L) = \emptyset$, and $bV^*v_0 = 0$; and

- (3) singular, non-normal matrices A that are unitarily similar to a matrix of the form $\left[\begin{array}{c|c} 0_1 & b \\ \hline 0 & L \end{array} \right]$, where $L \in M_{n-1}$ has only one distinct eigenvalue.

This is as far as we have gotten with the general $n \times n$ case. However, we can completely classify the 1×1 and 2×2 cases.

3.3 $\text{UES}(A)$: The 1×1 Case

Because these matrices are normal, $\text{UES}(A) = \mathcal{U}(A) = A$.

The only possible canonical form for $\text{UES}(A)$ is A .

3.4 UES(A): The 2×2 Case

Thanks to Theorem 1.3.3, the 2×2 case is also simple.

Proposition 3.4.1. If $A \in M_2$, then $\text{UES}(A) = \mathcal{U}(A)$.

Proof. Since $\mathcal{U}(A) \subset \text{UES}(A)$, we may assume $A = \begin{bmatrix} \lambda_1 & x \\ 0 & \lambda_2 \end{bmatrix}$, for $\lambda_1, \lambda_2 \in \mathbb{C}$ and $x \geq 0$. Let $B \in \text{UES}(A)$. Since A and B are similar, they have the same eigenvalues. So we may assume that $B = \begin{bmatrix} \lambda_1 & y \\ 0 & \lambda_2 \end{bmatrix}$, for $y \geq 0$. Since A and B are unitarily equivalent, they have the same singular values. So the matrices AA^* and BB^* have the same eigenvalues, and we must have $\text{tr}(AA^*) = \text{tr}(BB^*)$. But $AA^* = \begin{bmatrix} \lambda_1\bar{\lambda}_1 + x^2 & \bar{\lambda}_2x \\ \lambda_2x & \lambda_2\bar{\lambda}_2 \end{bmatrix}$ and $BB^* = \begin{bmatrix} \lambda_1\bar{\lambda}_1 + y^2 & \bar{\lambda}_2y \\ \lambda_2y & \lambda_2\bar{\lambda}_2 \end{bmatrix}$, so that $\text{tr}(AA^*) = \lambda_1\bar{\lambda}_1 + \lambda_2\bar{\lambda}_2 + x^2$ and $\text{tr}(BB^*) = \lambda_1\bar{\lambda}_1 + \lambda_2\bar{\lambda}_2 + y^2$. For these traces to be equal, we must have $x = y$. So by Theorem 1.3.3, A and B are unitarily similar, and $\text{UES}(A) \subset \mathcal{U}(A)$. Therefore, $\text{UES}(A) = \mathcal{U}(A)$. \square

A canonical form for $\text{UES}(A)$ can then be seen to be $\begin{bmatrix} \lambda_1 & r \\ 0 & \lambda_2 \end{bmatrix}$, for $r \geq 0$.

3.5 UEquiv(A) and $\mathcal{U}(A)$

As we tried in the general $n \times n$ case for $\text{CS}(A)$, we might also try to see how many disjoint unitary similarity classes are contained in $\text{Sim}(A)$ and $\text{UEquiv}(A)$ individually. Perhaps then this will tell us something about how many disjoint unitary similarity classes are contained in $\text{UES}(A)$. We know from Proposition 2.5.1 that $\text{Sim}(A)$ contains uncountably many such classes, if $A \neq \alpha I$. We seek to understand this for $\text{UEquiv}(A)$.

Proposition 3.5.1. If $A \in M_n$ and $A \neq 0$, then $\text{UEquiv}(A)$ contains uncountably many disjoint unitary similarity classes.

Proof. Without loss of generality, we may assume that

$$A = \begin{bmatrix} \sigma_1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \sigma_n \end{bmatrix},$$

with $\sigma_1 \geq \cdots \geq \sigma_n \geq 0$ and $\sigma_1 > 0$. Let $V = \begin{bmatrix} e^{i\alpha} & \mathbf{0} \\ \mathbf{0} & I_{n-1} \end{bmatrix}$ and consider the matrix $AV \in \text{UEquiv}(A)$. Then $\text{tr}(AV) = e^{i\alpha}\sigma_1 + \sigma_2 + \cdots + \sigma_n$, and this depends on α . Therefore, we can get a continuum of values for $\text{tr}(AV)$, and by Theorem 1.3.1, $\text{UEquiv}(A)$ contains uncountably many disjoint unitary similarity classes. \square

Unfortunately, knowing about $\text{Sim}(A)$ and $\text{UEquiv}(A)$ individually does not seem to shed any light on $\text{UES}(A)$. We know that for any non-zero normal matrix $A \in M_n$, with $n \geq 3$ and $A \neq \alpha I$, $\text{UES}(A) = \mathcal{U}(A)$. However, $\text{Sim}(A) \neq \mathcal{U}(A)$ and $\text{UEquiv}(A) \neq \mathcal{U}(A)$.

CHAPTER 4

CONCLUSIONS AND FUTURE DIRECTIONS

We include this chapter only to summarize some open problems remaining for the two equivalence classes examined in this dissertation.

4.1 $CS(A)$

The main open problem for this equivalence class is to determine how many unitary similarity classes are contained in $CS(A)$ for non-singular, non-co-Hermitian $A \in M_n$, with $n \geq 3$. The next task will be to find a canonical form for $CS(A)$ that generalizes those presented for the 1×1 and 2×2 cases.

Another tangential open problem is to classify all positive definite matrices M such that AM is similar to A .

4.2 $UES(A)$

The main open problems for this equivalence class are to determine how many unitary similarity classes are contained in $UES(A)$ for:

- (1) non-singular, non-normal matrices A ;
- (2) singular, non-normal matrices A that are unitarily similar to a matrix of the form

$$\left[\begin{array}{c|c|c} E & v_0 & V \\ \hline 0 & 0_1 & b \\ \hline 0 & 0 & L \end{array} \right],$$

where $0 < k \leq n - 1$, $E \in GL_k$, $L \in M_{n-k-1}$, $\sigma(E) \cap \sigma(L) = \emptyset$, and $bV^*v_0 = 0$; and

- (3) singular, non-normal matrices A that are unitarily similar to a matrix of the form $\begin{bmatrix} 0_1 & b \\ 0 & L \end{bmatrix}$, where $L \in M_{n-1}$ has only one distinct eigenvalue.

Then the next task will be to find a canonical form for $\text{UES}(A)$ that generalizes those presented for the 1×1 and 2×2 cases.

APPENDIX

Because the definitions of the elementary divisors and the minimal indices of a matrix pencil are not as easily stated as most other ideas in this dissertation, we will use this appendix for this purpose. For this discussion, we will consider the matrix pencil $C(\mu, \lambda) = \mu A + \lambda B$, for $A, B \in M_n$ and variables μ and λ . (For details beyond what are discussed here, see [Ga] and [Tu].)

Elementary Divisors

For $k = 0, 1, \dots, n$, let $D_k(\mu, \lambda)$ denote the greatest common divisor of all $k \times k$ minors of the matrix pencil $C(\mu, \lambda)$, where we define $D_0(\mu, \lambda) \equiv 1$. We then obtain the *invariant polynomials* of the matrix pencil by the formula $i_k(\mu, \lambda) = \frac{D_{n-k+1}(\mu, \lambda)}{D_{n-k}(\mu, \lambda)}$, for $k = 1, 2, \dots, n$. Note that the invariant polynomials are homogeneous in the variables μ and λ (i.e., the sum of the powers of μ and λ in each monomial is constant throughout a given polynomial). We can then split the invariant polynomials into powers of homogeneous polynomials irreducible over \mathbb{C} , and these are the *elementary divisors* of the matrix pencil.

If $\det(B) \neq 0$, then all the elementary divisors are of the form $(\mu + \alpha\lambda)^e$, for some $\alpha \in \mathbb{C}$ and $e > 0$. These are called the *finite* elementary divisors of the matrix pencil. However, if $\det(B) = 0$, then there also exist elementary divisors of the form μ^e , for $e > 0$, and these are called the *infinite* elementary divisors of the matrix pencil.

Minimal Indices

Suppose that the matrix pencil has $\det(C(\mu, \lambda)) \equiv 0$. Then we can find vectors $x(\mu, \lambda)$ for which $C(\mu, \lambda)x(\mu, \lambda) = 0$, and each entry of $x(\mu, \lambda)$ is a homogeneous polynomial in μ and λ

of degree m (i.e., the sum of the powers of μ and λ in each monomial is m). We say that such vectors $x_1(\mu, \lambda), \dots, x_k(\mu, \lambda)$ are linearly dependent if there exist homogeneous polynomials $p_1(\mu, \lambda), \dots, p_k(\mu, \lambda)$, not all identically zero, so that $p_1(\mu, \lambda)x_1(\mu, \lambda) + \dots + p_k(\mu, \lambda)x_k(\mu, \lambda) \equiv 0$.

Among all the solutions of $C(\mu, \lambda)x(\mu, \lambda) = 0$, we choose a non-zero solution $x_1(\mu, \lambda)$ of minimal degree m_1 . From all the solutions of $C(\mu, \lambda)x(\mu, \lambda) = 0$ that are linearly independent from $x_1(\mu, \lambda)$, we choose a solution $x_2(\mu, \lambda)$ of minimal degree m_2 . (Note that $m_1 \leq m_2$.) We then continue this process by choosing a solution linearly independent from $x_1(\mu, \lambda)$ and $x_2(\mu, \lambda)$ and of minimal degree m_3 , and so on. Doing this, we obtain a maximal linearly independent set of solutions

$$x_1(\mu, \lambda), x_2(\mu, \lambda), \dots, x_p(\mu, \lambda)$$

having degrees

$$m_1 \leq m_2 \leq \dots \leq m_p.$$

While the choice of the solutions $x_k(\mu, \lambda)$ is not unique, their number, p , and their degrees are unique, and these degrees are called the *minimal indices* for the matrix pencil.

BIBLIOGRAPHY

- [Ga] Gantmacher, F. R., *The Theory of Matrices*, Chelsea Publishing Company, New York, 1959.
- [HJ1] Horn, R. A. and Johnson, C. R., *Matrix Analysis*, Cambridge University Press, Cambridge, England, 1985.
- [HJ2] Horn, R. A. and Johnson, C. R., *Topics in Matrix Analysis*, Cambridge University Press, Cambridge, England, 1991.
- [Mu] Murnaghan, F. D., On the unitary invariants of a square matrix, *Proc. Nat. Acad. Sci. U.S.A.* 18:185–189 (1932).
- [Pe] Percy, C., A complete set of unitary invariants for operators generating finite W^* -algebras of type I, *Pacific J. Math.* 12:1405–1416 (1962).
- [Sh] Shapiro, H., A survey of canonical forms and invariants for unitary similarity, *Linear Algebra Appl.* 147:101–168 (1991).
- [Sp] Specht, W., Zur Theorie der Matrizen, II, *Jahresber. Deutsch. Math.-Verein.* 50:19–23 (1940).
- [Th] Thompson, R. C., The characteristic polynomial of a principal subpencil of a Hermitian matrix pencil, *Linear Algebra Appl.* 14:135–177 (1976).
- [Tu] Turnbull, H. W., On the equivalence of pencils of Hermitian matrices, *Proc. Lond. Math. Soc.* 39:232–248 (1935).