PORTUGALIAE MATHEMATICA
Vol. 64 Fasc. 2 – 2007
Nova Série

CONGRUENCE AND $A^{-1}A^*$

Susana Furtado * and Charles R. Johnson

Abstract: A canonical form under congruence for nonsingular complex matrices may be deduced from the canonical pair form for Hermitian matrices in a straightforward way. This allows the study of $A^{-1}A^*$ and what it says about the congruence class of A. In turn, alternate congruential forms may be given along with a further generalization of Sylvester's law of inertia and an application to C-S equivalence.

1 – Introduction

The matrix $B \in M_n(\mathbb{C})$, or M_n for short, is *congruent* to $A \in M_n$ if there is a nonsingular $C \in M_n$ such that $B = C^*AC$. Congruence is an equivalence relation on M_n that arises in a variety of ways (including study of the algebraic Riccati equation and indefinite scalar products), and we have been interested in this subject in [JF, FJ1, FJ2, FJ3].

Our purpose here is to study some basic ideas about congruent matrices in the nonsingular case. A simple calculation shows that if B is congruent to a nonsingular A, then $B^{-1}B^*$ is similar to $A^{-1}A^*$. The converse is not generally true, but we describe precisely which congruence classes are identified with the similarity class of $A^{-1}A^*$. In the process, we identify a canonical form for congruence. As $A = H + iS$, the unique decomposition of a general $A \in M_n$ into an ordered pair of Hermitian matrices H and S, this form may be deduced directly from the canonical pair form for Hermitian matrices, e.g. [T1, T2], but our intent, for a variety of purposes, is to view A as a single matrix, rather than as an ordered pair of Hermitian matrices. This allows us, in particular, to continue the

Received: March 18, 2004.
Keywords: congruence; congruential canonical form; C-S equivalence; dubloid; generalized Cayley transform; unitoid.
* This work was done within the activities of *Centro de Estruturas Lineares e Combinatórias da Universidade de Lisboa*.

generalization of Sylvester's law of inertia begun in [JF] and to give other special congruential forms not so obviously related to Hermitian pair forms.

The *field of values* of A is

$$F(A) \equiv \Big\{ x^* A x : \; x \in \mathbb{C}^n, \; x^* x = 1 \Big\} ,$$

a compact convex subset of the complex plane [HJ2]. The set

$$F'(A) \equiv \Big\{ x^* A x : \; 0 \neq x \in \mathbb{C}^n \Big\} ,$$

is the smallest angular sector that contains $F(A)$ and is called the *angular field of values* of A. Congruent matrices share the same angular field of values, though the field of values can vary. Whether or not 0 lies in the interior of, on the boundary of, or outside $F(A)$ is a congruential invariant that is important to our considerations.

Given $A \in M_n$ nonsingular, we denote $A^{-1} A^*$ by $\Phi(A)$.

2 – A congruential canonical form for nonsingular matrices

For positive integers p, q and r, consider the blocks

$$D_p(\gamma) = \begin{bmatrix} 0 & \cdots & \cdots & 0 & \gamma - i \\ \vdots & & \cdot^{\cdot^{\cdot}} & \cdot^{\cdot^{\cdot}} & 1 \\ \vdots & \cdot^{\cdot^{\cdot}} & \cdot^{\cdot^{\cdot}} & \cdot^{\cdot^{\cdot}} & 0 \\ 0 & \cdot^{\cdot^{\cdot}} & \cdot^{\cdot^{\cdot}} & \cdot^{\cdot^{\cdot}} & \vdots \\ \gamma - i & 1 & 0 & \cdots & 0 \end{bmatrix} \in M_p , \tag{1}$$

$$G_q = \begin{bmatrix} 0 & \cdots & \cdots & 0 & -1 \\ \vdots & & \cdot^{\cdot^{\cdot}} & \cdot^{\cdot^{\cdot}} & i \\ \vdots & \cdot^{\cdot^{\cdot}} & \cdot^{\cdot^{\cdot}} & \cdot^{\cdot^{\cdot}} & 0 \\ 0 & \cdot^{\cdot^{\cdot}} & \cdot^{\cdot^{\cdot}} & \cdot^{\cdot^{\cdot}} & \vdots \\ -1 & i & 0 & \cdots & 0 \end{bmatrix} \in M_q , \tag{2}$$

$$E_r(\gamma) = \begin{bmatrix} 0 & D_r(\gamma) \\ D_r(\overline{\gamma}) & 0 \end{bmatrix} \in M_{2r} , \tag{3}$$

with $\gamma \in \mathbb{C}$ and $i = \sqrt{-1}$. In particular, $D_1(\gamma) = \big[\, \gamma - i \,\big]$ and $G_1 = \big[\, -1 \,\big]$.

The next result, also mentioned in [FJ3], is a simple consequence of the simultaneous canonical form for a pair of Hermitian matrices given in [T1, T2].

Theorem 1. *Let $A \in M_n$ be a nonsingular matrix. Then, apart from the order of the direct summands, A is congruent to one and only one matrix of the form $Q_1 \oplus Q_2 \oplus Q_3$, with*

$$\begin{aligned} Q_1 &= \varepsilon_1 D_{p_1}(\beta_1) \oplus \cdots \oplus \varepsilon_k D_{p_k}(\beta_k) \ , \\ Q_2 &= \delta_1 G_{q_1} \oplus \cdots \oplus \delta_l G_{q_l} \ , \\ Q_3 &= E_{r_1}(\gamma_1) \oplus \cdots \oplus E_{r_m}(\gamma_m) \ , \end{aligned}$$

$\beta_1, \ldots, \beta_k \in \mathbb{R}$, $\gamma_1, \ldots, \gamma_m \in \mathbb{C} \backslash (\mathbb{R} \cup \{-i, i\})$, $i\,m(\gamma_j) > 0$, $j = 1, \ldots, m$, $\varepsilon_1, \ldots, \varepsilon_k$, $\delta_1, \ldots, \delta_l \in \{-1, 1\}$ *and* $p_1 + \cdots + p_k + q_1 + \cdots + q_l + 2r_1 + \cdots + 2r_m = n$.

Proof: The matrix A has a unique decomposition $A = H + iS$, with H and S Hermitian, namely $H = (A + A^*)/2$ and $S = (A - A^*)/2\,i$. The claim then follows from the unique simultaneous canonical form for the pair H, S given, for example, in [T1, lemma 2 and theorem 1]. ∎

We henceforth refer to $D_p(\beta)$ with $\beta \in \mathbb{R}$, G_q, and $E_r(\gamma)$ with $\gamma \in \mathbb{C} \backslash (\mathbb{R} \cup \{-i, i\})$ and $i\,m(\gamma) > 0$, those cases of (1), (2) and (3) that arise in the theorem, as "canonical blocks".

Note that in the statement of theorem 1 we assume that the imaginary part of γ_j, $j = 1, \ldots, m$, is positive in order to have uniqueness, because $E_r(\gamma)$ is congruent (via a permutation matrix) to $E_r(\overline{\gamma})$.

As mentioned in [FJ3, lemma 3], 0 is in the interior of the field of values of canonical blocks of the form (3), and likewise for canonical blocks of the forms (1) and (2) of sizes greater than 2. Also, 0 is on the boundary of the field of values of 2-by-2 canonical blocks of the forms (1) and (2).

In [JF] a matrix $A \in M_n$ is called *unitoid* if it is diagonalizable under congruence. In case it is nonsingular, this means that A is congruent to a diagonal unitary matrix. Here, we call A *dubloid* if A is non-unitoid but is congruent to a direct sum of blocks of size at most 2. The unitoid and dubloid matrices include all nonsingular $A \in M_n$ such that $0 \notin \operatorname{int} F(A)$. It is known [DJ], and also follows from theorem 1, that if $0 \notin F(A)$, then A is nonsingular unitoid. However, it may happen that $0 \in \partial F(A)$ and A is unitoid. Also 0 may be in the interior of $F(A)$ with A unitoid or dubloid.

3 – Congruential classes and $A^{-1}A^*$

The following proposition is easily verified.

Proposition 2. *Suppose that $A \in M_n$ is nonsingular. If $B \in M_n$ is congruent to A then $\Phi(B)$ is similar to $\Phi(A)$.* ∎

The purpose of this section is to study the converse of proposition 2. What can be said about the congruential relationship between A and B if $\Phi(A)$ is similar to $\Phi(B)$.

The proofs of the next three lemmas are straightforward calculations.

Lemma 3. *For $\beta \in \mathbb{R}$ the matrix $\Phi(D_p(\beta))$ is similar to the Jordan block of size p associated with the eigenvalue $\frac{\beta+i}{\beta-i}$.* ∎

Lemma 4. *The matrix $\Phi(G_q)$ is similar to the Jordan block of size q associated with the eigenvalue 1.* ∎

Lemma 5. *For $\gamma \in \mathbb{C}\backslash(\mathbb{R}\cup\{-i,i\})$ the matrix $\Phi(E_r(\gamma))$ is similar to the direct sum of the Jordan blocks of size r associated with the eigenvalues $\frac{\gamma+i}{\gamma-i}$ and $\frac{\overline{\gamma}+i}{\overline{\gamma}-i}$.* ∎

Suppose that $\beta \in \mathbb{R}$ and $\gamma \in \mathbb{C}\backslash(\mathbb{R}\cup\{-i,i\})$. It follows from lemmas 3 and 4 that the eigenvalues of the matrices $\Phi(D_p(\beta))$ and $\Phi(G_q)$ are unit modulus. Also, any unit modulus complex number, except 1, is an eigenvalue of $\Phi(D_p(\beta))$ for exactly one $\beta \in \mathbb{R}$. The matrix $\Phi(E_r(\gamma))$ has no unit modulus eigenvalues. Since $\det(\Phi(E_r(\gamma)))$ is unit modulus, clearly one of its 2 distinct eigenvalues has modulus greater than 1, while the other has modulus less than 1. Moreover, any nonzero non-unit modulus complex number is an eigenvalue of $\Phi(E_r(\gamma))$ for exactly one $\gamma \in \mathbb{C}\backslash(\mathbb{R}\cup\{-i,i\})$ with $i\,m(\gamma) > 0$. In particular, from these considerations and the lemmas above, it follows that if A is known to be a canonical block, then $\Phi(A)$ uniquely determines A.

Given $A \in M_n$ nonsingular we call A *circular* if all the eigenvalues of $\Phi(A)$ are unit modulus; we call A *off-circular* if $\Phi(A)$ has no unit modulus eigenvalues. If A is neither circular nor off-circular we call A *semicircular*. Note that the size of an off-circular matrix is always even.

The next result follows from theorem 1 and lemmas 3, 4, and 5.

Theorem 6. *Let $A \in M_n$ be nonsingular. Then A is congruent to a direct sum of a circular matrix and an off-circular matrix.* ∎

If a nonsingular $A \in M_n$ is congruent to $B_1 \oplus B_2$, with B_1 circular and B_2 off-circular, we call B_1 a *circular part* of A and B_2 an *off-circular part* of A.

Note that 0 is in the interior of the field of values of the off-circular part of A. Thus, if there is an eigenvalue of $\Phi(A)$ not on the unit circle then $0 \in \operatorname{int} F(A)$ because the off-circular part of A is non-empty.

Given $A, B \in M_n$, we say that A is *quasi-congruent* to B if the congruential canonical forms given in theorem 1 of A and B differ only in the ± 1 factors $\varepsilon_1, \ldots, \varepsilon_k, \delta_1, \ldots, \delta_l$.

Theorem 7. *Let $A, B \in M_n$. Then B is quasi-congruent to A if and only if $\Phi(B)$ is similar to $\Phi(A)$.*

Proof: Let A' and B' be matrices of the form described in theorem 1 that are congruent to A and B, respectively. Note that, by proposition 2, $\Phi(A)$ is similar to $\Phi(A')$ and $\Phi(B)$ is similar to $\Phi(B')$. If B is quasi-congruent to A, then, apart from the order of the direct summands, $\Phi(A')$ and $\Phi(B')$ are the same matrix, so that $\Phi(B)$ is similar to $\Phi(A)$. Now suppose that $\Phi(B)$ is similar to $\Phi(A)$. Then $\Phi(B')$ is similar to $\Phi(A')$ and, from the discussion after lemma 5, A' and B' can differ only in the ± 1 factors $\varepsilon_1, \ldots, \varepsilon_k, \delta_1, \ldots, \delta_l$ and, thus, these matrices are quasi-congruent. ∎

We have the following consequence of theorem 7.

Corollary 8. *Let $A \in M_n$ be an off-circular matrix and let $B \in M_n$. Then B is congruent to A if and only if $\Phi(B)$ is similar to $\Phi(A)$.* ∎

Note that for $C = I_r \oplus (-I_r)$, we have $C^*(-E_r(\gamma))\, C = E_r(\gamma)$, so that, as expected, the matrices $E_r(\gamma)$ and $-E_r(\gamma)$ are congruent.

An interesting fact is that if A is off-circular, then there is a neighborhood of A in which all matrices are off-circular. In fact, if $\Phi(A)$ has no eigenvalues on the unit circle the same is true for matrices sufficiently close to A because of the continuous dependence of the eigenvalues on the entries of a matrix.

The remarks of this section allow us to characterize all matrices of the special form $\Phi(A)$, sometimes called the *generalized Cayley transform* of A. Thus, the current work may be viewed, in part, as a continuation of work begun in [DJ].

Since a nonsingular A is congruent to a matrix of the form described in theorem 1, by proposition 2, $\Phi(A)$ is similar to a direct sum of blocks of the forms $\Phi(D_p(\beta))$, $\Phi(G_q)$, $\Phi(E_r(\gamma))$, with $\beta \in \mathbb{R}$ and $\gamma \in \mathbb{C}\backslash(\mathbb{R}\cup\{-i,i\})$. It follows that if $B=\Phi(A)$, then the eigenvalues of B not on the unit circle occur in pairs of like Jordan structure, each pair fully determined by one of its members. If λ_1 is an eigenvalue of $\Phi(A)$ not on the unit circle then, by lemma 5, $\lambda_1=\frac{\gamma+i}{\gamma-i}$ for some (unique) $\gamma \in \mathbb{C}\backslash(\mathbb{R}\cup\{-i,i\})$. By lemma 5, the mate of λ_1 is $\lambda_2=\frac{\overline{\gamma}+i}{\overline{\gamma}-i}$. It is easily checked that if $\lambda_1=k\,e^{i\theta}$, with $k>0$ and $k\neq 1$, then $\lambda_2=\frac{1}{k}\,e^{i\theta}$. On the other hand, unit modulus eigenvalues of B occur independently and with arbitrary Jordan structure. Also, it is clear from lemmas 3, 4 and 5 that any nonsingular B satisfying these two statements is similar to $\Phi(A)$, for some $A \in M_n$. If $B = C^{-1}\Phi(A)\,C$, with $C \in M_n$ nonsingular, then $B=\Phi(C^*AC)$. Taken together, these observations fully describe what matrices can be $\Phi(A)$. A consequence is that if n is odd, $\Phi(A)$ must have at least one unit modulus eigenvalue.

Theorem 9. *Let $B \in M_n$ be a nonsingular matrix. There is an $A \in M_n$ such that $B=\Phi(A)$ if and only if B is similar to a matrix of the form $B_1 \oplus B_2$, in which $B_1 \in M_{n_1}$ and $B_2 \in M_{n-n_1}$, $0 \le n_1 \le n$, are such that*

a) *all the eigenvalues of B_1 lie on the unit circle, and*

b) *the distinct eigenvalues of B_2 are $\lambda_1=s_1\,e^{i\theta_1}$, $\lambda_1'=\frac{1}{s_1}\,e^{i\theta_1}$, ..., $\lambda_m=s_m\,e^{i\theta_m}$, $\lambda_m'=\frac{1}{s_m}\,e^{i\theta_m}$, for some $\theta_1,\dots,\theta_m \in \mathbb{R}$, $s_1,\dots,s_m>1$, and the Jordan structure of B_2 associated with λ_j and λ_j' is the same, $j=1,\dots,m$.* ∎

We note that theorem 9 above may be seen to be equivalent to theorem 1 of [DJ]. However, the approach here and the related ideas are very different from those in [DJ].

4 – Alternate congruential forms

Lemma 10. *For every $\beta \in \mathbb{R}$ there is one and only one $\alpha \in (0,\pi)$ such that $D_q(\beta)$ is quasi-congruent to $e^{i\alpha}G_q$.*

Proof: By theorem 7, it is enough to show that there is one and only one $\alpha \in (0,\pi)$ such that $\Phi(D_q(\beta))$ and $e^{-2i\alpha}\,\Phi(G_q)$ are similar or, equivalently (because $\Phi(D_q(\beta))$ and $\Phi(G_q)$ are each similar to a Jordan block), have the same eigenvalues. According to lemma 4, the eigenvalues of $\Phi(G_q)$ are 1. By lemma 3, the eigenvalues of $\Phi(D_q(\beta))$ are $\frac{\beta+i}{\beta-i}$. Thus, $\Phi(D_q(\beta))$ and $e^{-2i\alpha}\,\Phi(G_q)$ have the same eigenvalues if and only if $\alpha = -\frac{1}{2}\arg\bigl(\frac{\beta+i}{\beta-i}\bigr) + p\pi,\ p \in \mathbb{Z}$. Because $\frac{\beta+i}{\beta-i}$ is different from 1, the existence of a unique α in $(0,\pi)$ follows. ∎

By theorem 1 G_q and $-G_q$ are not congruent. Thus, from lemma 10 there is one and only one $\alpha \in (0,2\pi)\backslash\{\pi\}$ such that $D_q(\beta)$ is congruent to $e^{i\alpha}G_q$.

With lemma 10 we then can rephrase theorem 1 as follows.

Theorem 11. *Let $A \in M_n$ be a nonsingular matrix. Then, apart from the order of the direct summands, A is congruent to one and only one matrix of the form $Q_1' \oplus Q_2'$, with*

$$Q_1' = e^{i\alpha_1}G_{q_1} \oplus \cdots \oplus e^{i\alpha_l}G_{q_l}\ ,$$
$$Q_2' = E_{r_1}(\gamma_1) \oplus \cdots \oplus E_{r_m}(\gamma_m)\ ,$$

$\alpha_1,\ldots,\alpha_l \in [0,2\pi),\ \gamma_1,\ldots,\gamma_m \in \mathbb{C}\backslash(\mathbb{R}\cup\{-i,i\}),\ i\,m(\gamma_j) > 0,\ j=1,\ldots,m,$ *and* $q_1+\cdots+q_l+2r_1+\cdots+2r_m = n$. ∎

Lemma 12. *There is an upper triangular $J_q \in M_q$ with all eigenvalues* 1 *such that G_q is congruent to $e^{i\theta_q}J_q$, for some $\theta_q \in [0,2\pi)$.*

Proof: If $q=1$ the result is trivial. Suppose that $q \geq 2$. Let $\theta_2 = \frac{\pi}{2}$ and for $q \geq 3$ let $\theta_q = \frac{1}{q}\bigl(\arg(\det(G_q)) + 2\,p_q\pi\bigr)$, for some arbitrary $p_q \in \mathbb{Z}$. Note that for any $q \geq 2$, $e^{i\theta_q} \in F'(G_q)$ and, modulo 2π, $q\,\theta_q = \arg(\det(G_q))$. According to [FJ1], if $q \geq 3$, or [J, FJ3], if $q=2$, G_q is congruent to a matrix B with all eigenvalues equal to $e^{i\theta_q}$. By Schur's unitary triangularization theorem there is a unitary matrix $U \in M_q$ such that $U^*BU = e^{i\theta_q}J_q$, in which J_q is upper triangular with all eigenvalues 1. ∎

For example, since $\Phi(G_2)$ is similar to $\Phi(e^{i\frac{\pi}{2}}J_2)$ for

$$J_2 = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix},$$

the matrices G_2 and $e^{i\frac{\pi}{2}}J_2$ are quasi-congruent. Moreover, because of the location of the field of values of both matrices, it follows that G_2 and $e^{i\frac{\pi}{2}}J_2$ are congruent. As we will see later, the choice of J_2 is unique up to a diagonal unitary similarity; the choice of θ_2 is unique modulo 2π.

From the proof of lemma 12, if $q > 2$ then the value of θ_q is not intrinsically determined by G_q. However, θ_q is uniquely determined by G_q once J_q has been chosen. For example, G_3 is congruent to both

$$J_3 = \begin{bmatrix} 1 & 3 & 3 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{bmatrix}$$

and $e^{i2\pi/3}J_3'$, with

$$J_3' = \begin{bmatrix} 1 & 3 & 3 \\ 0 & 1 & \frac{9}{2} + \frac{2\sqrt{15}}{3} - \frac{\sqrt{3}}{6}\,i \\ 0 & 0 & 1 \end{bmatrix}.$$

In fact, $\Phi(G_3)$ is similar to both $\Phi(J_3)$ and $\Phi(e^{i2\pi/3}J_3')$, so that G_3 is congruent to either J_3 or $-J_3$. Also G_3 is congruent to either $e^{i2\pi/3}J_3'$ or $-e^{i2\pi/3}J_3'$. The claim about the displayed matrices follows because congruence preserves the argument of the determinant. (We note that for $q > 3$ no particular explicit form for J_q is as simple to describe.) However, for a fixed J_q, θ_q is unique (modulo 2π). In fact, suppose that G_q is congruent to both $e^{i\theta_q}J_q$ and $e^{i\theta_q'}J_q$. Then $e^{-2i\theta_q}\Phi(J_q)$ is similar to $e^{-2i\theta_q'}\Phi(J_q)$, which implies $\theta_q = \theta_q' + p\pi$, for some $p \in \mathbb{Z}$. But, since G_q and $-G_q$ are not congruent, then, modulo 2π, θ_q must be θ_q'.

The previous observation, together with theorem 11 and lemma 12, implies theorem 13, in which we have chosen a particular J_q for each positive integer q.

Theorem 13. *Let $A \in M_n$ be a circular matrix. Then, apart from the order of the direct summands, A is congruent to one and only one matrix of the form*

$$e^{i\theta_1}J_{q_1} \oplus \cdots \oplus e^{i\theta_j}J_{q_j}\ ,$$

with $\theta_1, \ldots, \theta_j \in [0, 2\pi)$ and $q_1 + \cdots + q_j = n$. ∎

Note that theorem 13 applies only to circular matrices. A block of the form $E_r(\gamma)$ is also congruent to $e^{i\theta}B$, with B upper triangular with eigenvalues 1, but in this case B cannot be taken to be a single particular matrix independent of γ; B can only be given parametrically. Thus γ determines both θ and B. However, when $r = 1$ the parametric description of B is particularly simple which is a motivation to study the dubloid case.

5 – An "inertia" law for dubloid matrices

A nonsingular unitoid matrix $A \in M_n$ is congruent to a diagonal unitary matrix U. The arguments of the principal entries of U are called the *canonical angles* for A and according to [JF], and also theorem 1, are a congruential invariant. The nonsingular A is unitoid if and only if $\Phi(A)$ is similar to a unitary matrix. The lines through the origin on which the canonical angles lie are determined by the spectrum of $\Phi(A)$, though the canonical angles are not. If $A \in M_n$ is such that $0 \notin F(A)$ then A is unitoid (but not conversely). In this event the canonical angles of A are determined by the spectrum of $\Phi(A)$ and the location of $F(A)$.

In [FJ3] it is shown that if $A \in M_n$ is a nonsingular non-unitoid matrix such that $0 \in \partial F(A)$ then, up to a permutation, A is congruent to a unique direct sum of a diagonal unitary matrix U and copies of the block

$$e^{i\theta}\begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix},$$

in which $\theta \in \mathbb{R}$ is fixed and such that the canonical angles for U lie in $\left[\theta - \frac{\pi}{2}, \theta + \frac{\pi}{2}\right]$.

In this section we generalize the above mentioned results in [JF] and [FJ3] to general nonsingular dubloid matrices. Note that a nonsingular A is dubloid if and only if in the Jordan form of $\Phi(A)$, the Jordan blocks associated with unit modulus eigenvalues are of size at most 2, the Jordan blocks associated with non-unit modulus eigenvalues are of size 1 and either there are non-unit modulus eigenvalues or there is at least one 2-by-2 Jordan block. The unitoid and dubloid matrices include all nonsingular $A \in M_n$ such that $\Phi(A)$ is diagonalizable by similarity.

Lemma 14. *For $\theta \in \mathbb{R}$ and $k > 0$ let*

$$A(\theta, k) = e^{i\theta}\begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix}. \tag{4}$$

a) *G_2 is congruent to exactly one block of the form (4), namely $A(\frac{\pi}{2}, 2)$;*

b) *For $\gamma \in \mathbb{C}\backslash(\mathbb{R} \cup \{-i, i\})$, $E_1(\gamma)$ is congruent to exactly two blocks of the form (4), namely $A(\theta_1, k_1)$ and $A(\theta_1 + \pi, k_1)$, with $\theta_1 = \frac{\pi - \arg(\lambda_1)}{2}$, $k_1 = \left(\frac{|\lambda_1|^2 + 1}{|\lambda_1|} + 2\right)^{1/2}$ and $\lambda_1 = \frac{\gamma + i}{\gamma - i}$.*

Proof: It is easily checked that $\Phi(G_2)$ is similar to $\Phi(A(\frac{\pi}{2},2))$ and $\Phi(E_1(\gamma))$ is similar to $\Phi(A(\theta_1,k_1))$. According to corollary 8, $E_1(\gamma)$ is congruent to $A(\theta_1,k_1)$ and $A(\theta_1+\pi,k_1)$. According to theorem 7, G_2 is quasi-congruent to $A(\frac{\pi}{2},2)$, and, thus, G_2 is congruent to either $A(\frac{\pi}{2},2)$ or $A(-\frac{\pi}{2},2)$. The "uniqueness" follows because $\Phi(A(\theta,k))$ is similar to $\Phi(A(\theta',k'))$ if and only if $k=k'$ and $\theta=\theta'+p\pi$, $p\in\mathbb{Z}$. Since the angular fields of values of $A(-\frac{\pi}{2},2)$ and G_2 are distinct it follows that G_2 is not congruent to $A(-\frac{\pi}{2},2)$ and, then, G_2 is congruent to $A(\frac{\pi}{2},2)$. ∎

We now give an alternate canonical form for dubloid matrices.

Theorem 15. *Let $A\in M_n$ be a nonsingular matrix. If A is dubloid, then, apart from the order of the direct summands, A is congruent to one and only one matrix of the form $Q_1\oplus Q_2\oplus Q_3$, with*

$$Q_1 = \operatorname{diag}\left(e^{i\theta_1},\ldots,e^{i\theta_{j_1}}\right),$$

$$Q_2 = e^{i\beta_1}\begin{bmatrix}1 & 2\\ 0 & 1\end{bmatrix}\oplus\cdots\oplus e^{i\beta_{j_2}}\begin{bmatrix}1 & 2\\ 0 & 1\end{bmatrix}$$

and

$$Q_3 = e^{i\gamma_1}\begin{bmatrix}1 & k_1\\ 0 & 1\end{bmatrix}\oplus\cdots\oplus e^{i\gamma_{j_3}}\begin{bmatrix}1 & k_{j_3}\\ 0 & 1\end{bmatrix},$$

$\theta_1,\ldots,\theta_{j_1},\beta_1,\ldots,\beta_{j_2}\in[0,2\pi)$, $\gamma_1,\ldots,\gamma_{j_3}\in[0,\pi)$, $k_1,\ldots,k_{j_3}>2$, $j_1<n$ *and* $j_1+2\,j_2+2\,j_3=n$.

Proof: Let $A\in M_n$ be a nonsingular dubloid matrix. By definition, A is non-unitoid and is congruent to a direct sum of blocks of size at most 2. But, by theorem 11, this implies that A is congruent to a unique (apart from the order of the blocks) direct sum of a unitary diagonal matrix and 2-by-2 blocks of the forms $e^{i\theta}G_2$ and $E_1(\gamma)$, $0\le\theta<2\pi$, $\gamma\in\mathbb{C}\backslash(\mathbb{R}\cup\{-i,i\})$, $im(\gamma)>0$, with at least a 2-by-2 block. Now the claimed existence is a simple consequence of lemma 14. The uniqueness follows because two distinct matrices with the form described in the statement of the theorem have distinct canonical forms according to theorem 1 (or theorem 11) and then are not congruent. ∎

It follows from theorem 15 that the congruence class of a nonsingular dubloid matrix is uniquely determined by the canonical angles $\theta_1,\ldots,\theta_{j_1}$, the angles $\beta_1,\ldots,\beta_{j_2}$ and $\gamma_1,\ldots,\gamma_{j_3}$, that generalize the notion of canonical angles, and the magnitudes $k_1,\ldots,k_{j_3}$.

6 – An application to C-S equivalence

In [FJ2] we say that $A \in M_n$ is C-S equivalent to $B \in M_n$ if A is both congruent and similar to B, a notion introduced by M. Mills [M1]. For $A \in M_n$ normal, the number of unitary similarity classes in the C-S equivalence class of A is studied and completely determined as a function of the location of $F(A)$. Also, it is shown that for $A \in M_2$ nonsingular there is just one unitary similarity class in the C-S equivalence class of A. Generalizing this result, we may use ideas developed here to show that for $A \in M_n$ nonsingular and such that its minimum polynomial has degree 2, there is just one unitary similarity class in the C-S equivalence class of A.

Lemma 16. *Let $a, b \in \mathbb{C}\backslash\{0\}$. Let*

$$A = \bigoplus_{i=1}^{k} \begin{bmatrix} a & c_i \\ 0 & b \end{bmatrix} \qquad \text{and} \qquad B = \bigoplus_{i=1}^{k} \begin{bmatrix} a & d_i \\ 0 & b \end{bmatrix},$$

with $c_1 \geq \cdots \geq c_k \geq 0$ and $d_1 \geq \cdots \geq d_k \geq 0$. Then, A and B are congruent if and only if $c_i = d_i$, $i = 1, \ldots, k$.

Proof: Suppose that A and B are congruent. For $x \in \mathbb{C}$, let

$$D_x = \begin{bmatrix} a & x \\ 0 & b \end{bmatrix}. \tag{5}$$

First, note that if $|x_1| \neq |x_2|$ then $\Phi(D_{x_1})$ is not similar to $\Phi(D_{x_2})$ and, therefore, D_{x_1} is not congruent to D_{x_2}. Also, if γ_1 is known to be a canonical angle of a block of the form (5), then γ_1 determines both γ_2 and the $x \geq 0$ for which γ_1 and γ_2 are the canonical angles of D_x; γ_2 is such that $\gamma_1 + \gamma_2 = \arg(ab)$ and $x \geq 0$ (unique) is such that $\Phi(D_x)$ has eigenvalues $e^{-2\gamma_1}$, $e^{-2\gamma_2}$. Denote by D'_x the matrix of the form described in theorem 1 congruent to D_x. Clearly, either D'_x is diagonal or it is quasi-congruent to a 2-by-2 canonical block. Then A is congruent to $A' = \bigoplus_{i=1}^{k} D'_{c_i}$ and B is congruent to $B' = \bigoplus_{i=1}^{k} D'_{d_i}$. By theorem 1, up to a permutation of the blocks, A' and B' are the same matrix. Because of the assumed form of the c's and d's, it follows that $c_i = d_i$, $i = 1, \ldots, k$. Of course, the converse is immediate. ∎

Theorem 17. *Let $A \in M_n$ be a nonsingular matrix and suppose that the minimum polynomial of A has degree 2. Let $B \in M_n$ be congruent and similar to A. Then B is unitarily similar to A.*

Proof: Using Schur's triangularization theorem, it is straightforward to observe that A and B are unitarily similar to matrices of the forms

$$A' = \begin{bmatrix} aI_{n_1} & X \\ 0 & bI_{n_2} \end{bmatrix}$$

and

$$B' = \begin{bmatrix} aI_{n_1} & Y \\ 0 & bI_{n_2} \end{bmatrix},$$

respectively, with $a, b \in \mathbb{C}\backslash\{0\}$, $X, Y \in M_{n_1,n_2}$ and $n_1 + n_2 = n$. Let $U, V \in M_{n_1}$ and $U', V' \in M_{n_2}$ be unitary matrices such that

$$UXU'^* = \begin{bmatrix} 0 & 0 \\ 0 & \mathrm{diag}(c_1, \dots, c_k) \end{bmatrix} \quad \text{and} \quad VYV'^* = \begin{bmatrix} 0 & 0 \\ 0 & \mathrm{diag}(d_1, \dots, d_{k'}) \end{bmatrix}$$

with $c_1 \geq \cdots \geq c_k > 0$ and $d_1 \geq \cdots \geq d_{k'} > 0$, the singular values of X and Y, respectively. Without loss of generality, suppose that $k \geq k'$ and let $d_{k'+1} = \cdots = d_k = 0$. The matrix A' is unitarily similar to $A'' = (U \oplus U')\, A'(U^* \oplus U'^*)$ and B' is unitarily similar to $B'' = (V \oplus V')\, B'(V^* \oplus V'^*)$. The matrices A'' and B'' are unitarily similar (via a permutation matrix) to $A''' = aI_{n_1-k} \oplus bI_{n_2-k} \oplus A_0$ and $B''' = aI_{n_1-k} \oplus bI_{n_2-k} \oplus B_0$, respectively, with

$$A_0 = \bigoplus_{i=1}^{k} \begin{bmatrix} a & c_i \\ 0 & b \end{bmatrix} \quad \text{and} \quad B_0 = \bigoplus_{i=1}^{k} \begin{bmatrix} a & d_i \\ 0 & b \end{bmatrix}.$$

Since A''' and B''' are congruent, it follows from the uniqueness of the canonical form given in theorem 1 that A_0 and B_0 are congruent. According to lemma 16, $c_i = d_i$, for $i \in \{1, \dots, k\}$. Therefore, $A''' = B'''$ and B is unitarily similar to A. ∎

It follows from [FJ2] that if A is a nonsingular normal matrix with at most 2 distinct eigenvalues then the C-S equivalence class of A contains just one unitary similarity class (in this event $0 \notin \mathrm{int}\, F(A)$). A normal matrix with 2 distinct eigenvalues necessarily has a quadratic minimum polynomial. Theorem 17 shows that the normality assumption is not necessary, but having quadratic minimum polynomial is sufficient for the existence of just one unitary similarity class in the C-S equivalence class of the nonsingular A. Theorem 17 also can be seen as a generalization of the 2-by-2 case discussed in [M1, M2] and [FJ2]. It is also interesting to notice that for any integer $n > 1$, there is a non-normal $A \in M_n$ such that the C-S equivalence class of A contains just one unitary similarity class.

7 – An example

The matrix

$$A_n = \begin{bmatrix} 1 & 2 & \cdots & 2 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 2 \\ 0 & \cdots & 0 & 1 \end{bmatrix} \in M_n ,$$

$n \geq 2$, arises in a variety of ways. In particular, it plays an important role in the work in [KRS, CSS] as in our continuing investigation of C-S equivalence.

It is easily seen that $A_n + A_n^*$ is a positive semi-definite rank 1 matrix. Therefore, $0 \notin \operatorname{int} F(A_n)$, otherwise $A_n + A_n^*$ would be indefinite. Thus, A_n is either unitoid or dubloid. Since $F(Q) \subset F(A_n)$, with

$$Q = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix},$$

and $0 \in \partial F(Q)$, then $0 \in \partial F(A_n)$. Moreover, the field of values of A_n lies in the closed right half plane.

The purpose of this section is to show that our methods can be used to determine the congruential canonical form of A_n, which is, otherwise, not so simple.

A calculation shows that the Jordan structure of $\Phi(A_n)$ depends on n. If n is odd, $\Phi(A_n)$ has eigenvalues 1 and -1. The eigenvalue 1 is simple and the eigenvalue -1 has multiplicity $n-1$. Because $\operatorname{rank}\big(\Phi(A_n) + I_n\big) = 1$, all the Jordan blocks associated with the eigenvalue -1 have size 1. If n is even, -1 is an eigenvalue of $\Phi(A_n)$ with multiplicity n. Because $\operatorname{rank}\big(\Phi(A_n) + I_n\big) = 1$, there is one Jordan block of size 2 and $n-2$ Jordan blocks of size 1 associated with the eigenvalue -1. As a consequence of theorem 7, for n odd, A_n is unitoid and is congruent to

$$B_n = \varepsilon \begin{bmatrix} 1 \end{bmatrix} \oplus (-i) I_{s_1} \oplus i I_{n-s_1-1} ,$$

for some $\varepsilon \in \{-1, 1\}$ and $0 \leq s_1 \leq n-1$; for n even, A_n is dubloid and is congruent to

$$B_n = \varepsilon \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} \oplus (-i) I_{s_2} \oplus i I_{n-s_2-2} ,$$

for some $\varepsilon \in \{-1, 1\}$ and $0 \leq s_2 \leq n-2$.

Because of the location of the field of values of A_n, it follows that in any case $\varepsilon = 1$. We just need to determine the numbers s_1 and s_2 for which A_n and B_n are congruent. First consider the case in which n is odd. We show that $s_1 = \frac{n-1}{2}$. If we measure the canonical angles in $[-\frac{\pi}{2}, \frac{\pi}{2}]$, the canonical angles for B_n (and A_n) are $\theta_1 = \cdots = \theta_{s_1} = -\frac{\pi}{2}$, $\theta_{s_1+1} = 0$ and $\theta_{s_1+2} = \cdots = \theta_n = \frac{\pi}{2}$. Since the arguments of the eigenvalues of A_n are all 0, it follows from [FJ1] that

$$\begin{aligned} &\theta_1 + \cdots + \theta_i \le 0 , \quad i \in \{1, \ldots, n-1\} , \\ &\theta_1 + \cdots + \theta_n = 0 . \end{aligned}$$

In particular, these requirements imply $s_1 = \frac{n-1}{2}$.

Now consider the case in which n is even. We show that $s_2 = \frac{n-2}{2}$. Using the terminology defined in [FJ3], the canonical angles for A_n are $\theta_1 = \cdots = \theta_{s_2} = -\frac{\pi}{2}$, $\theta_{s_2+3} = \cdots = \theta_n = \frac{\pi}{2}$ while the limit canonical angles are $\theta_{s_2+1} = -\frac{\pi}{2}$ and $\theta_{s_2+2} = \frac{\pi}{2}$. By [FJ3]

$$\begin{aligned} &\theta_1 + \cdots + \theta_i \le 0 , \quad i \in \{1, \ldots, s_2\} , \\ &\theta_1 + \cdots + \theta_i < 0 , \quad i \in \{s_2+1, s_2+2\} , \\ &\theta_1 + \cdots + \theta_i \le 0 , \quad i \in \{s_2+3, \ldots, n-1\} , \\ &\theta_1 + \cdots + \theta_n = 0 . \end{aligned}$$

In particular, these requirements imply $s_2 = \frac{n-2}{2}$.

Thus, for n odd, A_n is congruent to

$$B_n = \begin{bmatrix} 1 \end{bmatrix} \oplus (-i)\, I_{\frac{n-1}{2}} \oplus i\, I_{\frac{n-1}{2}} ,$$

and for n even, A_n is congruent to

$$B_n = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} \oplus (-i)\, I_{\frac{n-2}{2}} \oplus i\, I_{\frac{n-2}{2}} .$$

It should be noted that this example shows that any relation between the congruential canonical form of a matrix and those of its principal submatrices of size one smaller is weak. In particular, a non-unitoid matrix may be bordered to give a unitoid one.

REFERENCES

[CSS] Caston, L.; Savova, M.; Spitkovsky, I. and Zobin, N. – On eigenvalues and boundary curvature of the numerical range, *Linear Algebra and Its Applications,* 322 (2001), 129–140.

[**DJ**] DePrima, C.R. and Johnson, C.R. – The range of $A^{-1}A^*$ in $GL(n,C)$, *Linear Algebra and Its Applications,* 9 (1974), 209–222.

[**FJ1**] Furtado, S. and Johnson, C.R. – Spectral variation under congruence, *Linear and Multilinear Algebra,* 49 (2001), 243–256.

[**FJ2**] Furtado, S. and Johnson, C.R. – On the number of unitary similarity classes in a C-S equivalence class: the normal case, *Linear Algebra and Its Applications,* 348 (2002), 193–202.

[**FJ3**] Furtado, S. and Johnson, C.R. – Spectral variation under congruence for a nonsingular matrix with 0 on the boundary of its field of values, *Linear Algebra and Its Applications,* 359 (2003), 67–78.

[**HJ1**] Horn, R. and Johnson, C.R. – *Matrix Analysis*, Cambridge University Press, NY, 1985.

[**HJ2**] Horn, R. and Johnson, C.R. – *Topics in Matrix Analysis*, Cambridge University Press, NY, 1991.

[**J**] Johnson, C.R. – The field of values and spectra of positive definite matrices, *J. Res. NBS B,* 78B (1974), 197–198.

[**JF**] Johnson, C.R. and Furtado, S. – A generalization of Sylvester's law of inertia, *Linear Algebra and Its Applications,* 338 (2001), 287–290.

[**KRS**] Keeler, D.; Rodman, L. and Spitkovsky, I. – The numerical range of 3×3 matrices, *Linear Algebra and Its Applications,* 252 (1997), 115–139.

[**M1**] Mills, M. – *The intersection of some classical equivalence classes of matrices*, Ph.D. Thesis, Iowa State University, 1999.

[**M2**] Mills, M. – The intersection of the similarity and conjunctivity equivalence classes, *Linear Algebra and Its Applications,* 362 (2003), 129–136.

[**T1**] Thompson, R.C. – The characteristic polynomial of a principal subpencil of a hermitian matrix pencil, *Linear Algebra and Its Applications,* 14 (1976), 135–177.

[**T2**] Thompson, R.C. – Pencils of complex and real symmetric and skew matrices, *Linear Algebra and Its Applications,* 147 (1991), 323–371.

Susana Furtado,
Faculdade de Economia, Universidade do Porto,
Rua Dr. Roberto Frias, 4200 Porto – PORTUGAL

and

Charles R. Johnson,
Department of Mathematics, College of William and Mary,
Williamsburg, VA 23187-8795 – USA