Linear Algebra and its Applications 434 (2011) 44–67

Contents lists available at ScienceDirect

Linear Algebra and its Applications

journal homepage: www.elsevier.com/locate/laa

The solution of the equation $XA + AX^T = 0$ and its application to the theory of orbits ☆

Fernando De Terán [a], Froilán M. Dopico [b,*]

[a] *Departamento de Matemáticas, Universidad Carlos III de Madrid, Avda. Universidad 30, 28911 Leganés, Spain*

[b] *Instituto de Ciencias Matemáticas CSIC-UAM-UC3M-UCM and Departamento de Matemáticas, Universidad Carlos III de Madrid, Avda. Universidad 30, 28911 Leganés, Spain*

A R T I C L E I N F O

Article history:
Received 25 May 2010
Accepted 3 August 2010

Submitted by V. Sergeichuk

AMS classification:
15A24
15A21

Keywords:
Canonical forms for congruence
Congruence
Codimension
Matrix equations
Orbits
Palindromic pencils

A B S T R A C T

We describe how to find the general solution of the matrix equation $XA + AX^T = 0$, with $A \in \mathbb{C}^{n\times n}$, which allows us to determine the dimension of its solution space. This result has immediate applications in the theory of congruence orbits of matrices in $\mathbb{C}^{n\times n}$, because the set $\{XA + AX^T : X \in \mathbb{C}^{n\times n}\}$ is the tangent space at A to the congruence orbit of A. Hence, the codimension of this orbit is precisely the dimension of the solution space of $XA + AX^T = 0$. As a consequence, we also determine the generic canonical structure of matrices under the action of congruence. All these results can be directly extended to palindromic pencils $A + \lambda A^T$.

1. Introduction

We are interested in the solution of the matrix equation

$$XA + AX^T = 0, \tag{1}$$

☆ This work was partially supported by the Ministerio de Ciencia e Innovación of Spain through Grant MTM-2009-09281.

* Corresponding author.
E-mail addresses: fteran@math.uc3m.es (F. De Terán), dopico@math.uc3m.es (F.M. Dopico).

0024-3795/$ - see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.laa.2010.08.005

where $A \in \mathbb{C}^{n\times n}$ is a given matrix. This equation is apparently similar to the particular *Sylvester equation* $XA - AX = 0$, whose solutions are well known [11, Chapter VIII, Sections 1 and 2], [15, Section 4.4]. However, the transposition of the unknown X in (1) leads to a completely different problem. A tentative approach to reduce (1) to a Sylvester equation, for A nonsingular, may be the following: from (1) we have $X = -AX^TA^{-1}$ and, by transposition, we get $X^T = -A^{-T}XA^T$. Now, substitute this expression for X^T in (1) and obtain $XA - AA^{-T}XA^T = 0$, which is equivalent to the Sylvester equation

$$XB - BX = 0, \tag{2}$$

with $B = AA^{-T}$. Hence, if X is a solution of (1), then X is solution of (2). But the converse is not true in general. Consider, for instance, $A = I$: in this case, every $X \in \mathbb{C}^{n\times n}$ is a solution of (2) whereas only skew-symmetric matrices are solutions of (1).

The relationship between Eq. (1) and the Sylvester equation $XA - AX = 0$ goes further than looking similar. To show this relationship let us first recall the notions of *congruence* and *similarity*. Two matrices $A, B \in \mathbb{C}^{n\times n}$ are said to be *congruent*[1] if there exists a nonsingular matrix $P \in \mathbb{C}^{n\times n}$ such that $PAP^T = B$, and they are said to be *similar* if there exists a nonsingular $P \in \mathbb{C}^{n\times n}$ such that $PAP^{-1} = B$. Accordingly, the actions

$$\begin{array}{ccc} GL(n,\mathbb{C}) \times \mathbb{C}^{n\times n} & \longrightarrow & \mathbb{C}^{n\times n} \\ (P, A) & \longmapsto & PAP^T \end{array} \quad \text{and} \quad \begin{array}{ccc} GL(n,\mathbb{C}) \times \mathbb{C}^{n\times n} & \longrightarrow & \mathbb{C}^{n\times n} \\ (P, A) & \longmapsto & PAP^{-1} \end{array}$$

of the general linear group of nonsingular $n \times n$ matrices, $GL(n, \mathbb{C})$, on $\mathbb{C}^{n\times n}$ are called, respectively, the action of *congruence* and the action of *similarity*. Also, the orbits associated with these actions will be called, respectively, the *congruence orbits* and the *similarity orbits*. It is well known that these orbits are differentiable manifolds in the vector space $\mathbb{C}^{n\times n}$ [3] and that the set [5]

$$\left\{XA - AX \;:\; X \in \mathbb{C}^{n\times n}\right\}$$

is the tangent space to the similarity orbit of A at the point A. This means that the dimension of the set of solutions of the Sylvester equation $XA - AX = 0$ is the codimension of the similarity orbit of A. Something similar occurs with Eq. (1) and the action of congruence. The congruence orbit of $A \in \mathbb{C}^{n\times n}$ is

$$\mathcal{O}(A) = \left\{PAP^T \;:\; P \in \mathbb{C}^{n\times n} \text{ nonsingular}\right\}.$$

Since this orbit is a differentiable manifold, its *tangent space* is well defined and has the same dimension at each point of the orbit. At the point A, the tangent space of $\mathcal{O}(A)$ is [14]

$$\left\{XA + AX^T \;:\; X \in \mathbb{C}^{n\times n}\right\}.$$

As a consequence, the codimension of the orbit $\mathcal{O}(A)$ is precisely the dimension of the solution space of the matrix equation (1).

In this work we compute the dimension of the solution space of (1) and describe how to find this solution space through the following procedure: (a) we show how the solution space of (1) is transformed under congruence of A and prove that its dimension remains invariant; (b) we transform A into its canonical form for congruence [17] (see also [18,23]); and, (c) we solve Eq. (1) for the canonical form of A. In terms of orbits, the invariance under congruence of A of the dimension of the solution space of (1) is equivalent to the fact that the dimension of the tangent space of $\mathcal{O}(A)$ is the same at all points of the orbit.

We want to emphasize that congruence of matrices is nowadays a subject related to important applications, since it is the base of structure preserving numerical methods for solving the eigenvalue problem of *palindromic* pencils $A + \lambda A^T$. These eigenvalue problems arise in a number of applications and are receiving a considerable attention in the last years (see, for instance [13,19,21,22] and the references therein). In this context, the congruence orbit of $A \in \mathbb{C}^{n\times n}$ can be identified with the

[1] Some authors refer to this definition as *T-congruence* to avoid confusion with $*$-*congruence*, i.e., $PAP^* = B$. Since there is no risk of confusion in this paper, we have preferred to use only *"congruence"* for simplicity.

congruence orbit of the palindromic pencil $A + \lambda A^T$ and the codimensions of both orbits are the same. For matrix pencils, the congruence relation is defined in a similar way as for matrices: given $A, B, C, D \in \mathbb{C}^{n\times n}$, the matrix pencils $A + \lambda B$ and $C + \lambda D$ are *congruent* if there exists a nonsingular $P \in \mathbb{C}^{n\times n}$ such that $P(A + \lambda B)P^T = C + \lambda D$. Note that the congruence relation in matrix pencils is a particular case of the *strict equivalence* relation as defined in [11, Chapter XII].

The theory of orbits of matrices by *similarity* and matrix pencils by *strict equivalence* is a classical area of research with an intense activity in the last decades (see, for instance, [1,2,5–10,12,24] and the references therein). One of the most relevant applications of this theory has been the recent development of reliable numerical algorithms for computing the Jordan Canonical form of matrices and the Kronecker Canonical form of matrix pencils [9,20], where several questions related to orbits, as their dimensions, their genericity and their inclusion relationships, have played a paramount role. By contrast none of these problems has been yet considered for orbits of matrices by *congruence*, and this paper can be seen as a very first step where the codimension of matrix orbits by congruence is computed. The knowledge of these codimensions may have different applications but, for brevity, here we will restrict to determine the highest possible dimension of a congruence orbit and, using this and the concept of *bundle* for the action of congruence, we will show what is the generic canonical structure of a matrix by congruence. This will be extended to palindromic pencils.

We emphasize the lack of references on Eq. (1), which is in stark contrast with the abundant bibliography about Sylvester equation. We have not found any explicit reference to Eq. (1). The only reference somewhat related is [4], where the author solves $A^TX \pm X^TA = B$ in terms of a certain generalized inverse of A. In the last part of [4], the equation $AX - X^TC = B$ is introduced as a generalization of $A^TX - X^TA = B$ and the author comments: *"I don't know of a simple explicit solution to this equation at present"*.

Finally, observe that Eq. (1) looks like $XA + AX^* = 0$, where X^* denotes the conjugate transpose of X. However, this equation is not linear in $\mathbb{C}$ (though it is linear in $\mathbb{R}$) whereas Eq. (1) is. So, the solution of $XA + AX^* = 0$ presents important differences with respect (1) and will be addressed in a subsequent paper.

The paper is organized as follows. Section 2 presents some preliminaries and introduces the canonical form for congruence [17]. A summary of the main results is presented in Theorem 2 in Section 3, whose proof is developed in Sections 4 and 5, where the solution of Eq. (1) is found. In Section 6 we obtain the lowest possible codimension for a congruence orbit in $\mathbb{C}^{n\times n}$ and determine the generic canonical structure under congruence for matrices and palindromic matrix pencils. Conclusions and lines of future research are included in Section 7. Finally, a technical result needed in the paper is proved in Appendix A.

2. Canonical form for congruence and tangent space

Our approach to find the solution space of (1) and its dimension is based on Lemma 1.

Lemma 1. *Let $A, B \in \mathbb{C}^{n\times n}$ be two congruent matrices such that $B = PAP^T$. Let $Y \in \mathbb{C}^{n\times n}$ and $X := P^{-1}YP$. Then Y is a solution of $YB + BY^T = 0$ if and only if X is a solution of $XA + AX^T = 0$. Therefore the linear mapping $Y \mapsto P^{-1}YP$ is an isomorphism between the solution space of $YB + BY^T = 0$ and the solution space of $XA + AX^T = 0$, and, as a consequence, both spaces have the same dimension.*

Proof. Let $A, B, P, Y, X \in \mathbb{C}^{n\times n}$ be as in the statement. Then

$$YB + BY^T = YPAP^T + PAP^TY^T = PXAP^T + PAX^TP^T = P(XA + AX^T)P^T.$$

Hence, $YB + BY^T = 0$ if and only if $XA + AX^T = 0$. Clearly the mapping $Y \mapsto P^{-1}YP$ is linear, surjective and injective and the result follows. $\square$

Lemma 1 shows that the dimension of the solution space of Eq. (1) is invariant under congruence of A. It also shows how to obtain the solution space of (1) for A from the solution space of (1) for any matrix congruent to A. Then, to solve (1) we will reduce A to its canonical form under congruence, denoted by C_A, and we will solve (1) with C_A instead of A. We consider the canonical form for congruence as it appears in Ref. [17] by Horn and Sergeichuk, where the authors provide an explicit matrix method to determine C_A, at least theoretically. This canonical form was originally introduced by Sergeichuk in [23] for matrices over any field $\mathbb{F}$ with characteristic not 2 up to classification of Hermitian forms over finite extensions of $\mathbb{F}$ (see also the Refs. [16,18] by these authors).

In order to recall the canonical form for congruence, let us define the following $k \times k$ matrices as in [17]

$$\Gamma_k = \begin{bmatrix} 0 & & & & & (-1)^{k+1} \\ & & & & \cdot^{\cdot^{\cdot}} & (-1)^k \\ & & & -1 & \cdot^{\cdot^{\cdot}} & \\ & & 1 & 1 & & \\ & -1 & -1 & & & \\ 1 & 1 & & & & 0 \end{bmatrix} \quad (\Gamma_1 = [1]), \tag{3}$$

and the $k \times k$ Jordan block with eigenvalue λ

$$J_k(\lambda) = \begin{bmatrix} \lambda & 1 & & 0 \\ & \lambda & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda \end{bmatrix} \quad (J_1(\lambda) = [\lambda]). \tag{4}$$

Also, we define, for each $\mu \in \mathbb{C}$, the $2k \times 2k$ matrix

$$H_{2k}(\mu) = \begin{bmatrix} 0 & I_k \\ J_k(\mu) & 0 \end{bmatrix} \quad \left(H_2(\mu) = \begin{bmatrix} 0 & 1 \\ \mu & 0 \end{bmatrix} \right). \tag{5}$$

Theorem 1 (Canonical form for congruence [17, Theorem 1.1]). *Each square complex matrix is congruent to a direct sum, uniquely determined up to permutation of summands, of canonical matrices of the three types*:

Type 0	$J_k(0)$
Type I	Γ_k
Type II	$H_{2k}(\mu)$, $0 \neq \mu \neq (-1)^{k+1}$ μ *is determined up to replacement by* μ^{-1}

(6)

Lemma 2 establishes the connection between equation (1) and the tangent space to the orbit of A by congruence that was discussed in Section 1.

Lemma 2. *Let $A \in \mathbb{C}^{n\times n}$ be given and let $\mathcal{O}(A)$ be the orbit of A under the action of congruence. Then the tangent space of $\mathcal{O}(A)$ at A is*

$$\mathcal{T}_A = \left\{ XA + AX^T \ : \ X \in \mathbb{C}^{n\times n} \right\}. \tag{7}$$

Proof. We follow the same proof as the one in [5, p. 71] for the action of similarity. Consider the congruence of A by $I + \delta X$, where δ is a small scalar. This yields

$$(I + \delta X)A(I + \delta X)^T = A + \delta(XA + AX^T) + O(\delta^2),$$

and the result follows. □

Lemma 2 is proved also in [14] using a longer proof. We have included here a proof for completeness. As a consequence of Lemma 2, the dimension of the solution space of (1) is the codimension of $\mathcal{O}(A)$. This motivates Definition 1 that allows us to be more concise in the rest of the paper.

Definition 1. Let $A \in \mathbb{C}^{n\times n}$. The *codimension* of A is the codimension of its orbit by congruence $\mathcal{O}(A)$ (this codimension coincides with the dimension of the solution space of $XA + AX^T = 0$).

3. Main results

The codimension of any matrix $A \in \mathbb{C}^{n\times n}$ is equal to the codimension of its canonical form for congruence C_A, which is a certain direct sum of the canonical blocks of Type 0, I and II in Theorem 1. The codimension of C_A is a sum of terms coming from two sources: (a) the codimension of each individual canonical block in C_A; and, (b) the codimension due to *interactions* between pairs of canonical blocks in C_A. To understand this fact, let us consider equation (1) with A being a direct sum of two square diagonal blocks D_1 and D_2. If we partition the unknown matrix X accordingly to the partition of the matrix A, then (1) is equivalent to

$$\begin{bmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{bmatrix}\begin{bmatrix} D_1 & 0 \\ 0 & D_2 \end{bmatrix} + \begin{bmatrix} D_1 & 0 \\ 0 & D_2 \end{bmatrix}\begin{bmatrix} X_{11}^T & X_{21}^T \\ X_{12}^T & X_{22}^T \end{bmatrix} = 0, \tag{8}$$

and, equating by blocks, this is equivalent to the system of matrix equations

$$\begin{aligned} X_{11}D_1 + D_1X_{11}^T &= 0 \\ X_{22}D_2 + D_2X_{22}^T &= 0 \\ X_{21}D_1 + D_2X_{12}^T &= 0 \\ X_{12}D_2 + D_1X_{21}^T &= 0. \end{aligned} \tag{9}$$

Then X_{11} and X_{22} are solutions of Eq. (1) with A replaced, respectively, by D_1 and D_2, and X_{12}, X_{21} are solutions of the system of two matrix equations given by the last two equations of (9). Hence, the codimension of $\begin{bmatrix} D_1 & 0 \\ 0 & D_2 \end{bmatrix}$ is given by the sum of three terms: (i) the dimension of the solution space of $X_{11}D_1 + D_1X_{11}^T = 0$, i.e., the codimension of D_1; (ii) the dimension of the solution space of $X_{22}D_2 + D_2X_{22}^T = 0$, i.e., the codimension of D_2; and (iii) the dimension of the solution space of the system of two equations $X_{21}D_1 + D_2X_{12}^T = 0$ and $X_{12}D_2 + D_1X_{21}^T = 0$ for the unknowns X_{21} and X_{12}. This motivates the following definition.

Definition 2. Let $M \in \mathbb{C}^{m\times m}$ and $N \in \mathbb{C}^{n\times n}$. Then the *interaction* between M and N, denoted by $\operatorname{inter}(M, N)$, is the dimension of the solution space of the linear system

$$\begin{aligned} XM + NY^T &= 0 \\ YN + MX^T &= 0, \end{aligned} \tag{10}$$

for the unknowns $X \in \mathbb{C}^{n\times m}$ and $Y \in \mathbb{C}^{m\times n}$.

The 2×2 block diagonal case considered in (8) and (9) can be directly extended to any number of diagonal blocks $D = \operatorname{diag}(D_1, D_2, \ldots, D_p)$ in such a way that the codimension of D is given by Lemma 3.

Lemma 3. *The codimension of the block diagonal matrix* $D = \operatorname{diag}(\mathrm{D}_1, \mathrm{D}_2, \ldots, \mathrm{D}_\mathrm{p})$ *is the sum of the codimensions of the diagonal blocks* D_i *for all* $i = 1, \ldots, p$, *and the sum of the interactions between* D_i *and* D_j *for all* $i < j$.

The calculation of the codimensions of the individual canonical blocks in Theorem 1 is the subject of Section 4, and interactions between pairs of canonical blocks are considered in Section 5. In these sections we also show how to find the solutions of the equations related to codimensions and interactions of canonical blocks, which provides a theoretical way to solve (1) assuming that a nonsingular matrix P such that $C_A = PAP^T$ is known. In Theorem 2, we state how to compute the codimension of a matrix A as a consequence of the results in Sections 4 and 5. Here and hereafter, given a real number

q, $\lfloor q \rfloor$ (resp. $\lceil q \rceil$) is the largest (resp. smallest) integer that is less (resp. greater) than or equal to q. In addition, we will use the symbol $\oplus$ for the direct sum of matrices, i.e., $A \oplus B = \operatorname{diag}(A, B)$.

Theorem 2 (Breakdown of the Codimension Count)**.** *Let $A \in \mathbb{C}^{n\times n}$ be a matrix with canonical form for congruence*

$$\begin{aligned} C_A = & J_{p_1}(0) \oplus J_{p_2}(0) \oplus \cdots \oplus J_{p_a}(0) \\ & \oplus \Gamma_{q_1} \oplus \Gamma_{q_2} \oplus \cdots \oplus \Gamma_{q_b} \\ & \oplus H_{2r_1}(\mu_1) \oplus H_{2r_2}(\mu_2) \oplus \cdots \oplus \ 0H_{2r_c}(\mu_c), \end{aligned}$$

where $p_1 \geqslant p_2 \geqslant \cdots \geqslant p_a$. Then the codimension of the orbit of A for the action of congruence, i.e., the dimension of the solution space of (1), depends only on C_A. It can be computed as the sum

$$c_{\text{Total}} = c_0 + c_1 + c_2 + c_{00} + c_{11} + c_{22} + c_{01} + c_{02} + c_{12}$$

whose components are given by:

1. *The codimension of the Type 0 blocks*
$$c_0 = \sum_{i=1}^{a} \left\lceil \frac{p_i}{2} \right\rceil.$$
2. *The codimension of the Type I blocks*
$$c_1 = \sum_{i=1}^{b} \left\lfloor \frac{q_i}{2} \right\rfloor.$$
3. *The codimension of the Type II blocks*
$$c_2 = \sum_{i=1}^{c} r_i + 2\sum_{j} \left\lceil \frac{r_j}{2} \right\rceil,$$
where the second sum is taken over those blocks $H_{2r_j}((-1)^{r_j})$ in C_A.
4. *The codimension due to interactions between Type 0 blocks*
$$c_{00} = \sum_{\substack{i,j=1 \\ i<j}}^{a} \operatorname{inter}(J_{p_i}(0), J_{p_j}(0)),$$
where
$$\operatorname{inter}(J_{p_i}(0), J_{p_j}(0)) = \begin{cases} p_j, & \text{if } p_j \text{ is even}, \\ p_i, & \text{if } p_j \text{ is odd and } p_i \neq p_j, \\ p_i + 1, & \text{if } p_j \text{ is odd and } p_i = p_j. \end{cases}$$
5. *The codimension due to interactions between Type I blocks*
$$c_{11} = \sum \min\{q_i, q_j\},$$
where the sum runs over all pairs of blocks $(\Gamma_{q_i}, \Gamma_{q_j})$, $i < j$, in C_A such that q_i and q_j have the same parity (both odd or both even).
6. *The codimension due to interactions between Type II blocks*
$$c_{22} = 2\sum \min\{r_i, r_j\} + 4\sum \min\{r_s, r_t\},$$
where the first sum is taken over all pairs $(H_{2r_i}(\mu_i), H_{2r_j}(\mu_j))$, $i < j$, of blocks in C_A such that "$\mu_i \neq \mu_j$ and $\mu_i\mu_j = 1$" or $\mu_i = \mu_j \neq \pm 1$; and the second sum is taken over all pairs $(H_{2r_s}(\mu_s), H_{2r_t}(\mu_t))$, $s < t$, of blocks in C_A such that $\mu_s = \mu_t = \pm 1$.

7. *The codimension due to interactions between Type 0 and Type I blocks*

$$c_{01} = N_{\text{odd}} \cdot \sum_{i=1}^{b} q_i,$$

where N_{odd} is the number of Type 0 blocks with odd size in C_A.

8. *The codimension due to interactions between Type 0 and Type II blocks*

$$c_{02} = N_{\text{odd}} \cdot \sum_{i=1}^{c} 2r_i,$$

where N_{odd} is the number of Type 0 blocks with odd size in C_A.

9. *The codimension due to interactions between Type I and Type II blocks*

$$c_{12} = 2 \sum \min\{k, \ell\},$$

where the sum is taken over all pairs $(\Gamma_k, H_{2\ell}((-1)^{k+1}))$ of blocks in C_A.

The codimension of congruence orbits in Theorem 2 is much more complicated than the codimension of similarity orbits of matrices (compare with [1, p. 35] or [5, Theorem 2.1]). Theorem 2 is complicated due to the possible presence in C_A of blocks $J_k(0)$ (in particular, the ones with odd size) and to the possible presence of the special Type II blocks $H_{2k}((-1)^k)$. But for most matrices, these blocks are not in the canonical form for congruence[2] and then the codimension count is much simpler. This is stated in Corollary 1, whose proof is omitted since follows directly from Theorem 2. In Corollary 1, we need to separate Type I blocks of even and odd sizes, and group together the Type II blocks with the same μ value (recall that Type II blocks in Theorem 1 are determined up to replacement of μ by $1/\mu$).

Corollary 1. *Let $A \in \mathbb{C}^{n\times n}$ be nonsingular with canonical form for congruence*

$$C_A = \bigoplus_{i=1}^{t} \mathcal{H}(\mu_i) \oplus \bigoplus_{k=1}^{g_{t+1}} \Gamma_{2r_{t+1,k}} \oplus \bigoplus_{k=1}^{g_{t+2}} \Gamma_{2r_{t+2,k+1}}, \text{ with } |\mu_i| \neq 1, \mu_i \neq \mu_j, \mu_i \neq 1/\mu_j \text{ if } i \neq j,$$

where

$$\mathcal{H}(\mu_i) = H_{2r_{i,1}}(\mu_i) \oplus H_{2r_{i,2}}(\mu_i) \oplus \cdots \oplus H_{2r_{i,g_i}}(\mu_i), \qquad \text{for } i = 1, \ldots, t,$$

and $r_{i,1} \geqslant r_{i,2} \geqslant \cdots \geqslant r_{i,g_i}$ for $i = 1, 2, \ldots, t+2$. Then the codimension of the orbit of A for the action of congruence is

$$c_{Total} = \frac{g_{t+2}(g_{t+2}-1)}{2} + \sum_{i=1}^{t+2} \left(r_{i,1} + 3\, r_{i,2} + 5\, r_{i,3} + \cdots + (2g_i - 1)\, r_{i,g_i}\right). \tag{11}$$

Eq. (11) resembles the codimension count for similarity orbits given in [1, p. 35] and [5, Theorem 2.1]. Moreover, recall that if *A is nonsingular*, then the blocks in the canonical form for congruence of A are in one-to-one correspondence with the blocks in the Jordan canonical form of the cosquare $A^{-T}A$ [17, p. 1016]. More precisely, each Type I block Γ_k of A is in one-to-one correspondence with a block $J_k((-1)^{k+1})$ in the Jordan canonical form of $A^{-T}A$, and each Type II block $H_{2k}(\mu)$ of A is in one-to-one correspondence with a pair of blocks $J_k(\mu) \oplus J_k(\mu^{-1})$ in the Jordan canonical form of $A^{-T}A$. Therefore, it is tempting to think that for a nonsingular $A \in \mathbb{C}^{n\times n}$, the codimension of the congruence orbit of A could be obtained from the well-known codimension of the similarity orbit of $A^{-T}A$. A general rule for doing this does not seem possible even in simple cases. For instance, the codimension of a single 2×2 Type II block $H_2(3)$ is 1, whereas the codimension of the similarity orbit of the associated Jordan

[2] Note that only singular matrices have Type 0 blocks in the canonical form for congruence.

blocks $J_1(3) \oplus J_1(1/3)$ is 2, i.e., double than for congruence. However, the congruence codimension of $H_{10}(-1)$ is 11, while the codimension of the associated Jordan blocks $J_5(-1) \oplus J_5(-1)$ is 20. A more striking example is $A = I_n = \Gamma_1 \oplus \Gamma_1 \oplus \cdots \oplus \Gamma_1$: the codimension of the congruence orbit is $n(n-1)/2$, while the codimension of the similarity orbit is n^2 (dimension 0). Observe that the orbits by congruence and by similarity of I_n are, respectively, $\mathcal{O}_{cong}(A) = \{PP^T : P \text{ invertible}\}$ and $\mathcal{O}_{sim}(A) = \{I_n\}$. This example clearly shows that there is no a one-to-one correspondence between the elements in the congruence orbit of A and those in the similarity orbit of $A^{-T}A$.

4. Codimension of canonical blocks

In this section we compute the codimension of the Type 0, I and II blocks in the canonical form for congruence given in Theorem 1 and show how to find the solution of the corresponding equations (1). This section and the next one are of a technical nature and include many details that can be skipped in a first reading. The main results obtained in Sections 4 and 5 are stated in a series of lemmas that have already been summarized in Theorem 2.

4.1. Type 0 blocks

Lemma 4. *The codimension of an individual Type 0 block is*

$$\operatorname{codim}(J_k(0)) = \left\lceil \frac{k}{2} \right\rceil.$$

Proof. We want to calculate the dimension of the solution space of the matrix equation

$$XJ_k(0) + J_k(0)X^T = 0. \tag{12}$$

If we set $X = [x_{ij}]_{i,j=1}^k$, then (12) is equivalent to

$$\begin{bmatrix} 0 & x_{11} & \dots & x_{1,k-1} \\ 0 & x_{21} & \dots & x_{2,k-1} \\ \vdots & \vdots & & \vdots \\ 0 & x_{k1} & \dots & x_{k,k-1} \end{bmatrix} + \begin{bmatrix} x_{12} & x_{22} & \dots & x_{k2} \\ \vdots & \vdots & & \vdots \\ x_{1k} & x_{2k} & \dots & x_{kk} \\ 0 & 0 & \dots & 0 \end{bmatrix} = 0_{k\times k},$$

and this is in turn equivalent to the following system of equations

(a) $x_{12} = x_{13} = \cdots = x_{1k} = 0$,
(b) $x_{k1} = x_{k2} = \cdots = x_{k,k-1} = 0$,
(c) $x_{i,j-1} + x_{j,i+1} = 0$, for $i = 1, \dots, k-1$ and $j = 2, \dots, k$.

Now, if we write (c) above for the index $j+1$ instead of j and, on the other hand, for indices $i = j+1$ and $j = i+2$, we achieve

$$x_{ij} = -x_{j+1,i+1} = x_{i+2,j+2}, \quad \text{for } i = 1, \dots, k-2, \ j = 1, \dots, k-2. \tag{13}$$

This implies, in particular, that the matrix X is completely determined by its first two rows and columns. Next we will prove that, in fact, X is completely determined only by $x_{22}, x_{23}, \dots, x_{2k}$. To this purpose, we prove that the remaining entries of the first two rows and columns of X are known or determined by $x_{22}, x_{23}, \dots, x_{2k}$. By (c) above (with $j = 2$), $x_{i1} = -x_{2,i+1}$, for $i = 1, \dots, k-1$. Also, (a) and (c) (with $i = 1$) together imply $x_{i2} = 0$, for $i = 3, \dots, k$, and, by (a) and (b) we have $x_{k1} = 0$ and $x_{1j} = 0$ for $j = 2, \dots, k$. Hence, we conclude that X is completely determined by the entries $x_{22}, x_{23}, \dots, x_{2k}$.

Next, we find which entries among $x_{22}, x_{23}, \dots, x_{2k}$ are necessarily zero. Starting from the last row of X and the set of equations (b) above, we apply (13) upwards recursively to get, for the off-diagonal entries of the first column of X

$$0 = x_{k-2,1} = x_{k-4,1} = x_{k-6,1} = \cdots$$

and this in turn implies

$$0 = x_{2,k-1} = x_{2,k-3} = x_{2,k-5} = \cdots$$

Hence, we conclude that if k is even, then $x_{2,k-1} = x_{2,k-3} = \cdots = x_{23} = 0$, and that, if k is odd, then $x_{2,k-1} = x_{2,k-3} = \cdots = x_{24} = 0$. Notice that x_{2k} remains in both cases as a free parameter.

Therefore X must be of the form

$$X = \begin{bmatrix} x_1 & 0 & 0 & 0 & 0 & 0 & \dots & 0 \\ 0 & -x_1 & 0 & x_2 & 0 & x_3 & \dots & x_{\frac{k}{2}} \\ -x_2 & 0 & x_1 & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & -x_1 & 0 & x_2 & \dots & x_{\frac{k}{2}-1} \\ -x_3 & 0 & -x_2 & 0 & x_1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ -x_{\frac{k}{2}} & 0 & -x_{\frac{k}{2}-1} & 0 & -x_{\frac{k}{2}-2} & 0 & \dots & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \dots & -x_1 \end{bmatrix} \quad (k \text{ even}),$$

$$X = \begin{bmatrix} x_1 & 0 & 0 & 0 & 0 & \dots & 0 \\ -x_2 & -x_1 & x_2 & 0 & x_3 & \dots & x_{\frac{k+1}{2}} \\ 0 & 0 & x_1 & 0 & 0 & \dots & 0 \\ -x_3 & 0 & -x_2 & -x_1 & x_2 & \dots & x_{\frac{k-1}{2}} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ -x_{\frac{k+1}{2}} & 0 & -x_{\frac{k-1}{2}} & 0 & -x_{\frac{k-3}{2}} & \dots & x_2 \\ 0 & 0 & 0 & 0 & 0 & \dots & x_1 \end{bmatrix} \quad (k \text{ odd}),$$

for some parameters $x_1, x_2, \dots, x_{\lceil \frac{k}{2} \rceil}$. On the other hand, every matrix X as the one above is a solution of (12) because it satisfies conditions (a), (b), and (c). Then the general solution of (12) depends on $\lceil \frac{k}{2} \rceil$ free parameters. $\square$

4.2. Type I blocks

Lemma 5. *The codimension of an individual Type I block is*

$$\operatorname{codim}(\Gamma_k) = \left\lfloor \frac{k}{2} \right\rfloor.$$

Proof. We want to calculate the number of linearly independent solutions of

$$X\Gamma_k = -\Gamma_k X^T. \tag{14}$$

We will consider separately the cases k even and k odd. The argument is the same in both cases with minor variations. For brevity, we present with detail the even case, while for k odd we only show the final result.

► ***k* even**. Set $X = [x_{ij}]_{i,j=1}^k$. Equating entries in (14) we get for k even

$$\begin{bmatrix} x_{1k} & x_{1k} - x_{1,k-1} & -x_{1,k-1} + x_{1,k-2} & \dots & x_{12} - x_{11} \\ x_{2k} & x_{2k} - x_{2,k-1} & -x_{2,k-1} + x_{2,k-2} & \dots & x_{22} - x_{21} \\ \vdots & \vdots & \vdots & & \vdots \\ x_{kk} & x_{kk} - x_{k,k-1} & -x_{k,k-1} + x_{k,k-2} & \dots & x_{k2} - x_{k1} \end{bmatrix}$$

$$= \begin{bmatrix} x_{1k} & x_{2k} & \dots & x_{kk} \\ -x_{1k} - x_{1,k-1} & -x_{2k} - x_{2,k-1} & \dots & -x_{kk} - x_{k,k-1} \\ x_{1,k-1} + x_{1,k-2} & x_{2,k-1} + x_{2,k-2} & \dots & x_{k,k-1} + x_{k,k-2} \\ \vdots & \vdots & & \vdots \\ -x_{12} - x_{11} & -x_{22} - x_{21} & \dots & -x_{k2} - x_{k1} \end{bmatrix} \quad (k \text{ even}).$$

For simplicity, we adopt the following convention in the next equations: an entry x_{pq} such that p does not satisfy $1 \leqslant p \leqslant k$ or q does not satisfy $1 \leqslant q \leqslant k$ is defined as zero. Note that, for $1 \leqslant i, j \leqslant k$, the $(k-j+1, i)$ entries of $X\Gamma_k$ and $-\Gamma_k X^T$ are

$$(X\Gamma_k)(k-j+1,i) = (-1)^i(x_{k-j+1,k-i+2} - x_{k-j+1,k-i+1})$$
$$(-\Gamma_k X^T)(k-j+1,i) = (-1)^j(x_{ij} + x_{i,j+1}),$$

and the $(i, k-j+1)$ entries are

$$(X\Gamma_k)(i,k-j+1) = (-1)^{k-j+1}(-x_{ij} + x_{i,j+1})$$
$$(-\Gamma_k X^T)(i,k-j+1) = (-1)^{k-i+1}(x_{k-j+1,k-i+2} + x_{k-j+1,k-i+1}).$$

Then, equating the corresponding entries from (14) we get

$$\begin{aligned} (-1)^i(x_{k-j+1,k-i+2} - x_{k-j+1,k-i+1}) &= (-1)^j(x_{ij} + x_{i,j+1}) \\ (-1)^{k-i+1}(x_{k-j+1,k-i+2} + x_{k-j+1,k-i+1}) &= (-1)^{k-j+1}(-x_{ij} + x_{i,j+1}). \end{aligned} \tag{15}$$

Now, if we add up and subtract the previous equations we obtain that (15) is equivalent to

$$\begin{aligned} (-1)^{i+1}x_{k-j+1,k-i+1} &= (-1)^j x_{ij} \\ (-1)^i x_{k-j+1,k-i+2} &= (-1)^j x_{i,j+1}. \end{aligned} \tag{16}$$

If we write the second equation in (16) for $i+1$ instead of i we reach

$$(-1)^{i+1}x_{k-j+1,k-i+1} = (-1)^j x_{ij}$$
$$(-1)^{i+1}x_{k-j+1,k-i+1} = (-1)^j x_{i+1,j+1},$$

and equating both expressions for $(-1)^{i+1}x_{k-j+1,k-i+1}$ we obtain $x_{i+1,j+1} = x_{ij}$, for $i, j = 1, \dots, k-1$, that is, X is a *Toeplitz matrix*.

On the other hand, if we replace j by k in (15) we have

$$(-1)^{k+i}(x_{1,k-i+2} - x_{1,k-i+1}) = x_{ik}$$
$$(-1)^{k+i}(-x_{1,k-i+2} - x_{1,k-i+1}) = x_{ik},$$

and subtracting we conclude that $x_{1,k-i+2} = 0$, for $i = 2, \dots, k$. Since X is Toeplitz this implies that X is *lower triangular*.

Now, if we set $j = 1$ in (15), we get

$$(-1)^i(x_{k,k-i+2} - x_{k,k-i+1}) = -x_{i1} - x_{i2}$$
$$(-1)^i(-x_{k,k-i+2} - x_{k,k-i+1}) = -x_{i1} + x_{i2},$$

and we sum up these equations to reach

$$(-1)^i x_{k,k-i+1} = x_{i1}, \quad \text{for } i = 1, \dots, k. \tag{17}$$

For odd i in (17) we have $x_{i1} = -x_{k,k-i+1}$. Since, on the other hand, X is a Toeplitz matrix, we have $x_{i1} = x_{k,k-i+1}$. Hence $x_{i1} = 0$ for odd i. We have, so far, that if X is a solution of (14) for k even, then it has the following structure:

$$X = \begin{bmatrix} 0 & & & & & & 0 \\ x_1 & 0 & & & & & \\ 0 & x_1 & 0 & & & & \\ x_2 & 0 & x_1 & 0 & & & \\ 0 & x_2 & 0 & x_1 & 0 & & \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \\ x_{\frac{k}{2}} & \dots & 0 & x_2 & 0 & x_1 & 0 \end{bmatrix} \tag{18}$$

for some parameters $x_1, x_2, \dots, x_{\frac{k}{2}}$. Finally, we have to prove that X in (18) is the general solution of (14), i.e., that any matrix X as in (18), with $x_1, x_2, \dots, x_{\frac{k}{2}}$ free parameters, is a solution of (14). For this purpose, simply check that X in (18) satisfies (14).

▶k **odd**. The arguments are the same as in the even case and allow us to prove that the general solution of (14) is

$$X = \begin{bmatrix} 0 & & & & & & & 0 \\ x_1 & 0 & & & & & & \\ 0 & x_1 & 0 & & & & & \\ x_2 & 0 & x_1 & 0 & & & & \\ 0 & x_2 & 0 & x_1 & 0 & & & \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & & \\ x_{\frac{k-1}{2}} & \dots & 0 & x_2 & 0 & x_1 & 0 & \\ 0 & x_{\frac{k-1}{2}} & \dots & 0 & x_2 & 0 & x_1 & 0, \end{bmatrix}$$

where $x_1, x_2, \dots, x_{\frac{k-1}{2}}$ are free parameters. $\square$

4.3. Type II blocks

Lemma 6. *The codimension of an individual Type II block is*

$$\operatorname{codim}(H_{2k}(\mu)) = \begin{cases} k + 2\left\lceil \frac{k}{2} \right\rceil, & \text{if } \mu = (-1)^k \\ k, & \text{otherwise.} \end{cases}$$

Proof. We want to find out the number of linearly independent solutions of the equation

$$XH_{2k}(\mu) + H_{2k}(\mu)X^T = 0. \tag{19}$$

We begin by partitioning the unknown matrix X conformally with the partition of $H_{2k}(\mu)$, that is, $X = \begin{bmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{bmatrix}$, where $X_{ij} \in \mathbb{C}^{k\times k}$, for $i, j = 1, 2$. Then (19) is equivalent to

$$X_{12}J_k(\mu) = -X_{12}^T, \tag{20}$$

$$X_{21} = -J_k(\mu)X_{21}^T, \tag{21}$$

$$X_{11} = -X_{22}^T, \text{ and } \; X_{22}J_k(\mu) = -J_k(\mu)X_{11}^T. \tag{22}$$

So (19) decouples into the three independent linear systems, (20), (21) and (22), that we will solve separately.

(i) We start with (22). Note that (22) is equivalent to $X_{11} = -X_{22}^T$ and $X_{22}J_k(\mu) = J_k(\mu)X_{22}$. By the first equation, X_{11} is determined by X_{22}, so we just have to solve the Sylvester equation $X_{22}J_k(\mu) = J_k(\mu)X_{22}$, whose general solution can be found in [11, Chapter VIII, Section 1]. It is an arbitrary $k \times k$ upper triangular Toeplitz matrix, that is

$$X_{22} = \begin{bmatrix} x_1 & x_2 & \dots & x_k \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & x_2 \\ 0 & \dots & 0 & x_1 \end{bmatrix}.$$

(ii) Consider now (20): $X_{12}J_k(\mu) = -X_{12}^T$. We will separate the proof in two cases: $\mu \neq (-1)^k$ and $\mu = (-1)^k$. We have not found a simple approach to deal with the case $\mu = (-1)^k$. It will be addressed in Appendix A, where an algorithm to find the general solution of $X_{12}J_k((-1)^k) = -X_{12}^T$ is presented and the dimension of its solution space is determined.

- ▷ $\mu \neq (-1)^k$. By Theorem 1, $\mu \neq (-1)^{k+1}$, then $\mu \neq \pm 1$, and $\mu \neq 0$. The matrix $J_k(\mu)$ is invertible, so $X_{12} = -X_{12}^T J_k(\mu)^{-1}$. Then $X_{12}^T = -J_k(\mu)^{-T}X_{12}$ and substituting in (20) we conclude that X_{12} satisfies the Sylvester equation
$$X_{12}J_k(\mu) = J_k(\mu)^{-T}X_{12}. \tag{23}$$
The Jordan canonical form of $J_k(\mu)^{-T}$ is $J_k(1/\mu)$. Since $\mu \neq \mu^{-1}$ ($\mu \neq \pm 1$), the unique solution of (23) is $X_{12} = 0$ [11, Chapter VIII, Section 1].
- ▷ $\mu = (-1)^k$. We have $X_{12}J_k((-1)^k) = -X_{12}^T$. We will see in Appendix A that the general solution X_{12} depends on $\lceil \frac{k}{2} \rceil$ free variables.

(iii) Finally, we consider (21): $X_{21} = -J_k(\mu)X_{21}^T$. This equation reduces to the equation $X_{12}J_k(\mu) = -X_{12}^T$ analyzed in (ii) above. To see this, let

$$R := \begin{bmatrix} 0 & & 1 \\ & \iddots & \\ 1 & & 0 \end{bmatrix}$$

be the $k \times k$ reverse identity. Now, $X_{21} = -J_k(\mu)X_{21}^T$ is equivalent to $X_{21}^T = -X_{21}J_k(\mu)^T$ and this is in turn equivalent to $(RX_{21}R)^T = -(RX_{21}R)(RJ_k(\mu)^TR)$. Note, finally, that $J_k(\mu) = RJ_k(\mu)^TR$.

As a conclusion of items (i), (ii) and (iii) above, the general solution of (19) is

$$X = \left[\begin{array}{cccc|cccc} -x_1 & 0 & \dots & 0 & & & & \\ -x_2 & -x_1 & \ddots & \vdots & & X_{12} & & \\ \vdots & \ddots & \ddots & 0 & & & & \\ -x_k & \dots & -x_2 & -x_1 & & & & \\ \hline & & & & x_1 & x_2 & \dots & x_k \\ & X_{21} & & & 0 & \ddots & \ddots & \vdots \\ & & & & \vdots & \ddots & x_1 & x_2 \\ & & & & 0 & \dots & 0 & x_1 \end{array}\right],$$

where $x_1, x_2, \dots, x_k$ are free parameters and, if $\mu \neq (-1)^k$, $X_{12} = X_{21} = 0$. If $\mu = (-1)^k$, then X_{12} depends on $\lceil \frac{k}{2} \rceil$ free parameters and X_{21} depends on a different set of $\lceil \frac{k}{2} \rceil$ free parameters. This completes the proof. □

5. Interactions between canonical blocks

The *interaction* between two square matrices was introduced in Definition 2. In this section we compute the interactions between pairs of blocks of Type 0, I and II in Theorem 1 and show how to find the solutions of the corresponding Eq. (10). We use MATLAB notation for submatrices, i.e., $A(i:j, k:l)$ denotes the submatrix of A consisting of rows i through j and columns k through l, $A(i:j, :)$ denotes the submatrix of A consisting of rows i through j, and $A(:, k : l)$ denotes the submatrix of A consisting of columns k through l.

5.1. Type 0 blocks

Lemma 7. *The interaction between two Type 0 blocks $J_k(0)$ and $J_\ell(0)$ with $k \geqslant \ell$, is*

$$\operatorname{inter}(J_k(0), J_\ell(0)) = \begin{cases} \ell, & \text{if } \ell \text{ is even}, \\ k, & \text{if } \ell \text{ is odd and } k \neq \ell, \\ k+1, & \text{if } \ell \text{ is odd and } k = \ell. \end{cases}$$

Proof. According to Definition 2, the interaction between $J_k(0)$ and $J_\ell(0)$ is the number of linearly independent solutions (X, Y) of the system of equations

$$\begin{aligned} XJ_k(0) &= -J_\ell(0)Y^T \\ YJ_\ell(0) &= -J_k(0)X^T. \end{aligned} \tag{24}$$

Without loss of generality, we will assume that $k \geqslant \ell$, as in the statement. Set $X = [x_{ij}]_{\ell \times k}$ and $Y = [y_{ij}]_{k \times \ell}$. The solution of (24) for $\ell = 1$ is immediate: (a) if $\ell = 1$ and $k > \ell$, then $X = 0$ and Y arbitrary, so the general solution of (24) depends on k free variables; and, (b) if $\ell = k = 1$, then X and Y are arbitrary scalars, so the number of free variables in this case is 2. In the rest of the proof we consider that $\ell > 1$.

If we transpose the second equation in (24), then (24) is equivalent to the system consisting of the following two matrix equations

$$\begin{bmatrix} 0 & x_{11} & x_{12} & \dots & x_{1,k-1} \\ 0 & x_{21} & x_{22} & \dots & x_{2,k-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & x_{\ell 1} & x_{\ell 2} & \dots & x_{\ell,k-1} \end{bmatrix} = - \begin{bmatrix} y_{12} & y_{22} & \dots & y_{k2} \\ y_{13} & y_{23} & \dots & y_{k3} \\ \vdots & \vdots & & \vdots \\ y_{1\ell} & y_{2\ell} & \dots & y_{k\ell} \\ 0 & 0 & \dots & 0 \end{bmatrix} \tag{25}$$

and

$$\begin{bmatrix} x_{12} & x_{13} & \dots & x_{1k} & 0 \\ x_{22} & x_{23} & \dots & x_{2k} & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ x_{\ell 2} & x_{\ell 3} & \dots & x_{\ell k} & 0 \end{bmatrix} = - \begin{bmatrix} 0 & 0 & \dots & 0 \\ y_{11} & y_{21} & \dots & y_{k1} \\ y_{12} & y_{22} & \dots & y_{k2} \\ \vdots & \vdots & & \vdots \\ y_{1,\ell-1} & y_{2,\ell-1} & \dots & y_{k,\ell-1} \end{bmatrix}. \tag{26}$$

Note that (25) and (26) imply, in particular, that Y is completely determined by X. So, we will focus in determining X. Equate the entries in (25) and (26) that do not correspond to identically zero rows and columns. These give

$$\begin{aligned} x_{i,j-1} &= -y_{j,i+1}, \quad \text{for } 1 \leqslant i \leqslant \ell-1, \quad 2 \leqslant j \leqslant k, \\ x_{i,j+1} &= -y_{j,i-1}, \quad \text{for } 2 \leqslant i \leqslant \ell, \quad 1 \leqslant j \leqslant k-1. \end{aligned} \tag{27}$$

If we write the second set of equations in (27) for the index $i+2$ instead of i and equate the corresponding expressions for $-y_{j,i+1}$ for both sets of equations, then we get $x_{i,j-1} = x_{i+2,j+1}$, for $1 \leqslant i \leqslant \ell - 2$ and $2 \leqslant j \leqslant k-1$, that can be written as

$$x_{i,j} = x_{i+2,j+2}, \quad \text{for } 1 \leqslant i \leqslant \ell-2, \quad 1 \leqslant j \leqslant k-2. \tag{28}$$

Also, from the last row in (25) and the first row in (26), we have

$$x_{\ell 1} = x_{\ell 2} = \cdots = x_{\ell,k-1} = 0, \tag{29}$$

$$x_{12} = x_{13} = \cdots = x_{1k} = 0. \tag{30}$$

From the first column in (25) and the last column in (26), we get $y_{12} = y_{13} = \cdots = y_{1\ell} = 0$ and $y_{k1} = y_{k2} = \cdots = y_{k,\ell-1} = 0$ which implies, by (27),

$$x_{32} = x_{42} = \cdots = x_{\ell 2} = 0, \tag{31}$$

$$x_{1,k-1} = x_{2,k-1} = \cdots = x_{\ell-2,k-1} = 0. \tag{32}$$

Now we will construct the most general $X \in \mathbb{C}^{\ell \times k}$ that satisfies (28), (29), (30), (31) and (32). The key fact is that the first and second rows and columns of X completely determine X through (28). Then, we just have to find out which entries in $X(1:2,:)$ and $X(:,1:2)$ are free variables. By (30) and (31), we have $X(1, 2:k) = 0$ and $X(3:\ell, 2) = 0$, therefore the free variables of $X(1:2,:)$ and $X(:,1:2)$ have to be found among the entries

$$\{x_{11}, x_{21}, x_{31}, \ldots, x_{\ell 1}\} \cup \{x_{22}, x_{23}, x_{24}, \ldots, x_{2k}\}. \tag{33}$$

But some of these entries are not free because they are zero by (29), (32) and (28). To determine which entries in (33) can be free, we distinguish between ℓ even and odd.

▶ ℓ *even.* Consider (29) and apply (28) "backwards" to get

$$X(\ell, 1:k-1) = 0,\ X(\ell-2, 1:k-3) = 0,\ X(\ell-4, 1:k-5) = 0,$$
$$\ldots, X(2, 1:k-\ell+1) = 0,$$

which implies that the following entries in (33) are zero

$$x_{\ell 1} = x_{\ell-2,1} = x_{\ell-4,1} = \cdots = x_{41} = 0 \quad \text{and} \quad x_{21} = x_{22} = x_{23} = \cdots = x_{2,k-\ell+1} = 0. \tag{34}$$

Next, if $\ell > 2$, consider (32) and apply (28) "backwards" to get

$$X(1:\ell-2, k-1) = 0,\ X(1:\ell-4, k-3) = 0,\ X(1:\ell-6, k-5) = 0,$$
$$\ldots, X(1:2, k-\ell+3) = 0,$$

which implies that the following entries in (33) are zero

$$x_{2,k-1} = x_{2,k-3} = x_{2,k-5} = \cdots = x_{2,k-\ell+3} = 0. \tag{35}$$

As a consequence of (34) and (35), the entries in (33) that can be free are

$$x_{11}, x_{31}, x_{51}, \ldots, x_{\ell-1,1} \quad \text{and} \quad x_{2k}, x_{2,k-2}, x_{2,k-4}, \ldots, x_{2,k-\ell+2}. \tag{36}$$

Note that in (36) there are precisely ℓ entries. We will see below that they are indeed free parameters, which will prove Lemma 7 for ℓ even.

▶ ℓ *odd.* Consider (29) and apply (28) "backwards" to get

$$X(\ell, 1:k-1) = 0,\ X(\ell-2, 1:k-3) = 0, \ldots, X(3, 1:k-\ell+2) = 0,\ X(1, 1:k-\ell) = 0,$$

where $X(1, 1:k-\ell) = 0$ only appears if $k > \ell$. This implies that the following entries in (33) are zero

$$x_{\ell 1} = x_{\ell-2,1} = x_{\ell-4,1} = \cdots = x_{31} = x_{11} = 0, \tag{37}$$

where $x_{11} = 0$ only appears if $k > \ell$. Next consider (32) and apply (28) "backwards" to get,

$$X(1:\ell-2, k-1) = 0,\ X(1:\ell-4, k-3) = 0,$$
$$\ldots, X(1:3, k-\ell+4) = 0, X(1:1, k-\ell+2) = 0$$

which implies that the following entries in (33) are zero

$$x_{2,k-1} = x_{2,k-3} = x_{2,k-5} = \cdots = x_{2,k-\ell+4} = 0. \tag{38}$$

As a consequence of (37) and (38), if $k > \ell$, then the entries in (33) that can be free are

$$\begin{array}{l} x_{21}, x_{41}, x_{61}, \ldots, x_{\ell-1,1}, \quad \text{and} \\ x_{22}, x_{23}, x_{24}, \ldots, x_{2,k-\ell+3}, \quad \text{and} \\ x_{2,k-\ell+5},\ x_{2,k-\ell+7}, \ldots, x_{2,k-2},\ x_{2,k}. \end{array} \tag{39}$$

If $k = \ell$, then x_{11} has to be added to the set (39) of variables. Note that in (39) there are k entries. We will see below that they are indeed free, which will prove Lemma 7 for ℓ odd.

Before finishing the proof, let us summarize what we have proved so far. We have proved that any matrix X satisfying (24) is determined by $X(1:2,:)$ and $X(:,1:2)$ according to (28), and that those entries of $X(1:2,:)$ and $X(:,1:2)$ that are different from the ones in (36) for ℓ even or different from the ones in (39) for ℓ odd (if $k = \ell$, add x_{11} to (39)) are zero. But given an arbitrary matrix X with these properties, it remains to prove that it is always a solution of (24), i.e., we have to prove that the entries in (36) and (39) are really free parameters. For this purpose, define Y as follows

$$Y^T = \left[\begin{array}{c|c} -X(2:\ell, 2:k) & \begin{matrix} 0 \\ \vdots \\ 0 \end{matrix} \\ \hline 0 \quad -X(\ell-1, 1:k-1) & \end{array}\right],$$

where the 0's are scalars, and check that (X, Y) satisfies (24). This is immediate when one realizes that X fulfills (29) and (32). $\square$

5.2. Type I blocks

Lemma 8. *The interaction between two Type I blocks is*

$$\operatorname{inter}(\Gamma_k, \Gamma_\ell) = \begin{cases} 0, & \text{if } k, \ell \text{ have different parity} \\ \min\{k, \ell\}, & \text{if } k, \ell \text{ have the same parity}. \end{cases}$$

Proof. According to Definition 2, the interaction between Γ_k and Γ_ℓ is the number of linearly independent solutions of the system of equations

$$\begin{array}{l} X\Gamma_k = -\Gamma_\ell Y^T \\ Y\Gamma_\ell = -\Gamma_k X^T. \end{array} \tag{40}$$

Since Γ_ℓ is invertible, we find $Y = -\Gamma_k X^T \Gamma_\ell^{-1}$ and, taking transposes, $Y^T = -\Gamma_\ell^{-T} X \Gamma_k^T$. Replacing this expression for Y^T in the first equation of (40) we get the system of equations

$$X\Gamma_k\Gamma_k^{-T} = \Gamma_\ell\Gamma_\ell^{-T} X \tag{41}$$

$$Y = -\Gamma_k X^T \Gamma_\ell^{-1}, \tag{42}$$

which is equivalent to (40). To solve (41) and (42), we just have to solve (41) for X and then to obtain Y from (42). Note that (41) is a Sylvester equation. To solve it, we recall that $\Gamma_s\Gamma_s^{-T}$ is similar to $J_s((-1)^{s+1})$ [17, p. 1016], and consider the following two cases.

- k, ℓ have different parity. $\Gamma_k\Gamma_k^{-T}$ and $\Gamma_\ell\Gamma_\ell^{-T}$ have no common eigenvalues, so the solution of (41) is $X = 0$ [[11, Chapter VIII, Section 1]], and this implies that $Y = 0$ by (42).
- k, ℓ have the same parity. Without loss of generality, we assume that $k \geqslant \ell$. The eigenvalues of $\Gamma_k\Gamma_k^{-T}$ and $\Gamma_\ell\Gamma_\ell^{-T}$ coincide (they are both 1 or both -1). We reduce $\Gamma_k\Gamma_k^{-T} = PJ_k((-1)^{k+1})P^{-1}$ and $\Gamma_\ell\Gamma_\ell^{-T} = QJ_\ell((-1)^{\ell+1})Q^{-1}$ to their Jordan canonical forms and we write (41) in the following equivalent form

$$(Q^{-1}XP)J_k((-1)^{k+1}) = J_\ell((-1)^{\ell+1})(Q^{-1}XP).$$

Then the general solution of (41) is [11, Chapter VIII, Section 1]

$$X = Q \left[\begin{array}{ccc|cccc} 0 & \dots & 0 & x_1 & x_2 & \dots & x_\ell \\ 0 & \dots & 0 & 0 & x_1 & \ddots & \vdots \\ \vdots & & \vdots & \vdots & & \ddots & x_2 \\ 0 & \dots & 0 & 0 & \dots & 0 & x_1 \end{array}\right] P^{-1},$$

where $x_1, x_2, \dots, x_\ell$ are free parameters. $\square$

5.3. Type II blocks

Lemma 9. *The interaction between two Type II blocks is*

$$\operatorname{inter}(H_{2k}(\mu), H_{2\ell}(\widetilde{\mu})) = \begin{cases} 4\min\{k,\ell\}, & \text{if } \mu = \widetilde{\mu} = \pm 1 \\ 2\min\{k,\ell\}, & \text{if } \mu = \widetilde{\mu} \neq \pm 1 \\ 2\min\{k,\ell\}, & \text{if } \mu \neq \widetilde{\mu},\ \mu\widetilde{\mu} = 1 \\ 0, & \text{if } \mu \neq \widetilde{\mu},\ \mu\widetilde{\mu} \neq 1 \end{cases}.$$

Proof. According to Definition 2, we have to determine the number of linearly independent solutions of the system of equations

$$\begin{aligned} XH_{2k}(\mu) &= -H_{2\ell}(\widetilde{\mu})Y^T \\ YH_{2\ell}(\widetilde{\mu}) &= -H_{2k}(\mu)X^T. \end{aligned} \tag{43}$$

$H_{2\ell}(\widetilde{\mu})$ is invertible, so the second equation in (43) is equivalent to $Y = -H_{2k}(\mu)X^T H_{2\ell}(\widetilde{\mu})^{-1}$. Obtain Y^T from here and substitute in the first equation of (43) to get

$$\begin{aligned} XH_{2k}(\mu)H_{2k}(\mu)^{-T} &= H_{2\ell}(\widetilde{\mu})H_{2\ell}(\widetilde{\mu})^{-T}X \\ Y &= -H_{2k}(\mu)X^T H_{2\ell}(\widetilde{\mu})^{-1}, \end{aligned} \tag{44}$$

which is equivalent to (43). The first equation in (44) is a Sylvester equation. We just have to solve this equation and then obtain Y from the second one. For each $\lambda \in \mathbb{C}$

$$H_{2s}(\lambda)H_{2s}(\lambda)^{-T} = \left[\begin{array}{c|c} J_s(\lambda)^{-T} & 0 \\ \hline 0 & J_s(\lambda), \end{array}\right]$$

and $J_s(\lambda)^{-T}$ is similar to $J_s(1/\lambda)$, so the Jordan canonical form of $H_{2s}(\lambda)H_{2s}(\lambda)^{-T}$ is $J_s(\lambda) \oplus J_s(1/\lambda)$. Therefore the solution of the first equation in (44) depends on the equality relationships between the numbers $\mu, 1/\mu, \widetilde{\mu}, 1/\widetilde{\mu}$. The explicit solution can be found in [11, Chapter VIII, Section 1], and it depends on the number of free variables indicated in the statement. $\square$

5.4. Blocks of different type

We calculate in this section the interactions between blocks of different type and show how to solve the corresponding equations (10). These interactions are stated in Lemma 10.

Lemma 10. *The interactions between blocks of different types in Theorem 1 are*:

(i) *For Type 0 and Type I blocks,*

$$\operatorname{inter}(J_k(0), \Gamma_\ell) = \begin{cases} 0, & \text{if } k \text{ is even} \\ \ell, & \text{if } k \text{ is odd} \end{cases}.$$

(ii) *For Type 0 and Type II blocks,*

$$\operatorname{inter}(J_k(0), H_{2\ell}(\mu)) = \begin{cases} 0, & \text{if } k \text{ is even} \\ 2\ell, & \text{if } k \text{ is odd.} \end{cases}$$

(iii) *For Type I and Type II blocks,*

$$\operatorname{inter}(\Gamma_k, H_{2\ell}(\mu)) = \begin{cases} 2\min\{k, \ell\}, & \text{if } \mu = (-1)^{k+1} \\ 0, & \text{if } \mu \neq (-1)^{k+1}. \end{cases}$$

Proof. Since Γ_ℓ and $H_{2\ell}(\mu)$ are invertible, (i) and (ii) are consequence of the following result: let $F \in \mathbb{C}^{p\times p}$ be any invertible matrix, then

$$\operatorname{inter}(J_k(0), F) = \begin{cases} 0, & \text{if } k \text{ is even} \\ p, & \text{if } k \text{ is odd.} \end{cases}$$

To show this, let $M = J_k(0)$ and $N = F$ in (10). We want to solve the system of equations

$$\begin{aligned} XJ_k(0) &= -FY^T \\ YF &= -J_k(0)X^T. \end{aligned} \tag{45}$$

From the second equation, we get

$$Y^T = -F^{-T}XJ_k(0)^T, \tag{46}$$

and introducing this expression in the first equation gives

$$XJ_k(0) = FF^{-T}XJ_k(0)^T. \tag{47}$$

It can be checked that (45) is equivalent to the system consisting of equations (46) and (47), so we have to solve (47) for X and then get Y from (46). To solve (47) for $X \in \mathbb{C}^{p\times k}$ we write this equation by columns:

$$0 = FF^{-T}X(:,2)\,; \quad X(:,j-1) = FF^{-T}X(:,j+1), \quad j = 2,\ldots,k-1\,; \quad X(:,k-1) = 0.$$

Now, we use that FF^{-T} is nonsingular and distinguish between k even and k odd.

1. k **even.** We have

   $$0 = X(:,2) = X(:,4) = \cdots = X(:,k) \;\text{ and }\; 0 = X(:,k-1) = X(:,k-3) = \cdots = X(:,1),$$

   hence $X = 0$ and $Y = 0$ by (46).
2. k **odd.** We have

   $$0 = X(:,2) = X(:,4) = \cdots = X(:,k-1),$$

   and $X(:,1), X(:,3), \ldots, X(:,k-2)$ are determined by $X(:,k)$. The entries of $X(:,k)$ can be chosen arbitrarily and are the p free variables in the general solution of (47).

(iii) Let $M = \Gamma_k$ and $N = H_{2\ell}(\mu)$ in (10). We want to solve the system of equations

$$\begin{aligned} X\Gamma_k &= -H_{2\ell}(\mu)Y^T \\ YH_{2\ell}(\mu) &= -\Gamma_k X^T. \end{aligned} \tag{48}$$

Since Γ_k and $H_{2\ell}(\mu)$ are nonsingular, we can proceed as in Sections 5.2 and 5.3, to prove that (48) is equivalent to the system

$$\begin{aligned} Y &= -\Gamma_k X^T H_{2\ell}(\mu)^{-1} \\ X\Gamma_k\Gamma_k^{-T} &= H_{2\ell}(\mu)H_{2\ell}(\mu)^{-T}X. \end{aligned} \tag{49}$$

Therefore, we have to solve the second equation of (49) for X and to get Y from the first equation. The second equation in (49) is a Sylvester equation, whose solution is known. To find this solution, recall from Section 5.2 that the Jordan canonical form of $\Gamma_k\Gamma_k^{-T}$ is $J_k((-1)^{k+1})$ and from Section 5.3 that the Jordan canonical form of $H_{2\ell}(\mu)H_{2\ell}(\mu)^{-T}$ is $J_\ell(1/\mu)\oplus J_\ell(\mu)$ and apply [11, Theorem 1, Chapter VIII, Section 1]. Then we get: (a) if $\mu=(-1)^{k+1}$ the dimension of the solution space of the second equation in (49) is $2\min\{k,\ell\}$; (b) if $\mu\neq(-1)^{k+1}$ the unique solution of the second equation in (49) is $X=0$. $\square$

6. Minimal codimension of orbits and generic structure

This section is devoted to find the minimal possible codimension of an orbit for the action of congruence and to determine the *generic* canonical structure of matrices under congruence. We understand by "generic canonical structure" the canonical structure for congruence of a certain set of matrices, to be defined below, that has codimension zero, i.e., it has the same dimension as the whole space $\mathbb{C}^{n\times n}$ and, therefore, contains almost all matrices.

Theorem 3. *The minimal codimension for a congruence orbit in $\mathbb{C}^{n\times n}$ is $\lfloor n/2\rfloor$.*

Proof. For any $A\in\mathbb{C}^{n\times n}$, the codimension c_{Total} of its congruence orbit $\mathcal{O}(A)$ is given by Theorem 2. In the rest of the proof we follow the notation used in Theorem 2 and assume that C_A in Theorem 2 is the canonical form for congruence of A. The proof has three steps: (1) note that $c_{\text{Total}}\geqslant c_0+c_1+c_2+c_{11}$; (2) we will prove that $c_0+c_1+c_2+c_{11}\geqslant\lfloor n/2\rfloor$; and, finally, (3) we will find a matrix $B\in\mathbb{C}^{n\times n}$ such that $c_{\text{Total}}(B)=\lfloor n/2\rfloor$. We will use the inequalities

$$\lceil x\rceil+\lceil y\rceil\geqslant\lceil x+y\rceil\quad\text{and}\quad\lceil x\rceil+\lfloor y\rfloor\geqslant\lfloor x+y\rfloor, \tag{50}$$

where x and y are any real numbers. Assume, as in Theorem 2, that the canonical form of A consists of a blocks of Type 0 with sizes $p_1,\ldots,p_a$, b blocks of Type I with sizes $q_1,\ldots,q_b$, and c blocks of Type II with sizes $2r_1,\ldots,2r_c$. Define $k_0:=\sum_{i=1}^a p_i$, $k_1:=\sum_{i=1}^b q_i$ and $k_2:=\sum_{i=1}^c r_i$, and note that $k_0+k_1+2k_2=n$. Observe that

$$c_0=\sum_{i=1}^a\left\lceil\frac{p_i}{2}\right\rceil\geqslant\left\lceil\frac{k_0}{2}\right\rceil\quad\text{and}\quad c_2\geqslant\sum_{i=1}^c r_i=k_2, \tag{51}$$

where we have used the first inequality in (50) to get the lower bound for c_0.

Next, we will get a joint lower bound for c_1+c_{11}. For this purpose, assume that there are b_o Type I blocks in C_A with odd size and b_e with even size. So $b=b_o+b_e$. Then

$$c_1=\sum_{i=1}^b\frac{q_i}{2}-\frac{b_o}{2}\quad\text{and}\quad c_{11}\geqslant\frac{b_o(b_o-1)}{2}+\frac{b_e(b_e-1)}{2}\geqslant\frac{b_o^2-b_o}{2},$$

and

$$c_1+c_{11}\geqslant\frac{k_1}{2}+\frac{b_o(b_o-2)}{2}\geqslant\left\lfloor\frac{k_1}{2}\right\rfloor, \tag{52}$$

because if $b_o\geqslant0$ and $b_0\neq1$, then $b_o(b_o-2)\geqslant0$, and if $b_o=1$, then $b_o(b_o-2)/2=-1/2$ and k_1 is odd, so $k_1/2-1/2=\lfloor k_1/2\rfloor$. Now, we combine (51) with (52) and use (50) to get

$$c_{\text{Total}}\geqslant c_0+c_1+c_2+c_{11}\geqslant\left\lceil\frac{k_0}{2}\right\rceil+\left\lfloor\frac{k_1}{2}\right\rfloor+k_2\geqslant\left\lfloor\frac{k_0+k_1+2k_2}{2}\right\rfloor=\left\lfloor\frac{n}{2}\right\rfloor.$$

Finally, note that the matrix $B=\Gamma_n$ satisfies $c_{\text{Total}}(B)=\lfloor n/2\rfloor$. $\square$

Note that the minimal codimension given by Theorem 3 can be reached by orbits corresponding to different canonical forms. For instance, it can be reached with only one block in C_A: one block Γ_n, as in the proof of Theorem 3, one block $J_n(0)$ if n is even, or one block $H_{2n/2}(\mu)$ if n is even and $\mu\neq\pm1$. It can also be reached with exactly two blocks in C_A: if n is odd, by two Type I blocks $\Gamma_{k_1}\oplus\Gamma_{k_2}$ with $k_1+k_2=n$ and k_1,k_2 having different parity, or, if n is even, by two Type I blocks $\Gamma_{n-1}\oplus\Gamma_1$.

Observe also that Theorem 3 states that there are no orbits for congruence of codimension zero (except in the trivial case $n = 1$). Therefore, to determine the generic canonical structure for congruence, we need to consider sets of matrices larger than orbits. To find adequate sets, we look for inspiration in the action of similarity: recall [1,5] that the minimal possible codimension of an orbit by similarity in $\mathbb{C}^{n\times n}$ is n, that is always greater that zero, and so there are no generic orbits by similarity. However, it is well known that matrices in $\mathbb{C}^{n\times n}$ have, generically, n distinct eigenvalues corresponding to n Jordan blocks with size 1×1, which gives the generic Jordan canonical form. This can be made rigorous by considering the notion of *bundle* by the action of similarity introduced by Arnold in [1]. To define an appropriate notion of *bundle for the action of congruence*, we need to specify the Type II blocks in the canonical form for congruence C_A of $A \in \mathbb{C}^{n\times n}$ with more detail, so we write

$$C_A = \bigoplus_{i=1}^{a} J_{p_i}(0) \ \oplus \ \bigoplus_{i=1}^{b} \Gamma_{q_i} \ \oplus \ \bigoplus_{i=1}^{t} \mathcal{H}(\mu_i), \qquad \text{with } \mu_i \neq \mu_j \text{ and } \mu_i \neq 1/\mu_j \text{ if } i \neq j, \tag{53}$$

where

$$\mathcal{H}(\mu_i) = H_{2r_{i,1}}(\mu_i) \oplus H_{2r_{i,2}}(\mu_i) \oplus \cdots \oplus H_{2r_{i,g_i}}(\mu_i), \qquad \text{for } i = 1, \ldots, t, \tag{54}$$

(recall that the Type II blocks in C_A are determined up to replacement of μ by $1/\mu$). Then the *bundle* $\mathcal{B}(A)$ of A for the action of congruence is defined by the following union of congruence orbits

$$\mathcal{B}(A) = \bigcup_{\substack{\mu_i' \in \mathbb{C},\ i=1,\ldots,t \\ \mu_i' \neq \mu_j',\ \mu_i'\mu_j' \neq 1, i \neq j}} \mathcal{O}\left(\bigoplus_{i=1}^{a} J_{p_i}(0) \ \oplus \ \bigoplus_{i=1}^{b} \Gamma_{q_i} \ \oplus \ \bigoplus_{i=1}^{t} \mathcal{H}(\mu_i') \right). \tag{55}$$

Note that all orbits in the union in (55) have the same sizes of the canonical blocks (same p_i, q_i and r_{ij}), so the bundles are built up from orbits that have the same canonical form for congruence except that the pairwise distinct μ-values of the Type II blocks are different. In terms of algebraic geometry, it is said that a bundle is a *fibre space* whose *fibres* are the orbits appearing in (55) [1]. Then, we may talk about the (co)dimension of a bundle by relating it with the (co)dimension of anyone of its fibre orbits. More precisely, following [1],

$$\operatorname{codim}(\mathcal{B}(A)) = \operatorname{codim}(\mathcal{O}(A)) - t. \tag{56}$$

With this definition of codimension of bundles in mind we can state the following result, which gives us the generic canonical structure of matrices for the action of congruence.

Theorem 4 (Generic canonical form for congruence)**.**

1. *Let n be even and $A \in \mathbb{C}^{n\times n}$ be a matrix whose canonical form for congruence is*

$$G_A = H_2(\mu_1) \oplus H_2(\mu_2) \oplus \cdots \oplus H_2(\mu_{n/2}), \tag{57}$$

with $\mu_i \neq \pm 1, i = 1, \ldots, n/2, \mu_i \neq \mu_j$ and $\mu_i \neq 1/\mu_j$ if $i \neq j$. Then $\operatorname{codim}(\mathcal{B}(A)) = 0$. Therefore, we can say that the generic canonical form for congruence of a matrix in $\mathbb{C}^{n\times n}$ is the one in (57) with unspecified values $\mu_1, \mu_2, \ldots, \mu_{n/2}$.

2. *Let n be odd and $A \in \mathbb{C}^{n\times n}$ be a matrix whose canonical form for congruence is*

$$G_A = H_2(\mu_1) \oplus H_2(\mu_2) \oplus \cdots \oplus H_2(\mu_{(n-1)/2}) \oplus \Gamma_1, \tag{58}$$

with $\mu_i \neq \pm 1$, $i = 1, \ldots, (n-1)/2$, $\mu_i \neq \mu_j$ and $\mu_i \neq 1/\mu_j$ if $i \neq j$. Then $\operatorname{codim}(\mathcal{B}(A)) = 0$. Therefore, we can say that the generic canonical form for congruence of a matrix in $\mathbb{C}^{n\times n}$ is the one in (58) with unspecified values $\mu_1, \mu_2, \ldots, \mu_{(n-1)/2}$.

Proof. Use Theorem 2 to prove that $\operatorname{codim}(\mathcal{O}(A)) = \lfloor n/2 \rfloor$ both for n even and odd, and then apply (56) with $t = \lfloor n/2 \rfloor$. $\square$

We have already mentioned at the end of Section 3 that, if A *is* nonsingular, then the blocks in the canonical form for congruence of A are in one-to-one correspondence with the blocks in the Jordan canonical form of the cosquare $A^{-T}A$ [17, p.1016]. So the generic Jordan canonical form of cosquares follows from Theorem 4. Observe that if n is odd, then necessarily there exists a block Γ_k with k odd in the canonical form for congruence *of* any nonsingular $A \in \mathbb{C}^{n\times n}$, i.e., $\lambda = 1$ is an eigenvalue of $A^{-T}A$ for any A. This is no surprising, because if n is odd, then $A - A^T$ is singular (it is skew-symmetric) and this implies that $A^{-T}A - I$ is singular. This makes natural the presence of the block Γ_1 in (58).

6.1. Generic Kronecker form of palindromic matrix pencils

In this section we extend the previous results on generic canonical forms for congruence of matrices to complex palindromic matrix pencils. First, note that a canonical form for congruence of palindromic pencils follows immediately from Theorem 1 by taking into account that $A \in \mathbb{C}^{n\times n}$ is congruent to $B \in \mathbb{C}^{n\times n}$ if and only if $A + \lambda A^T$ is congruent to $B + \lambda B^T$.

Theorem 5. *Each palindromic matrix pencil $A + \lambda A^T$, with $A \in \mathbb{C}^{n\times n}$, is congruent to a direct sum, uniquely determined up to permutation of summands, of canonical palindromic pencils of the following three types*

Type 0	$J_k(0) + \lambda J_k(0)^T$	
Type I	$\Gamma_k + \lambda \Gamma_k^T$	(59)
Type II	$H_{2k}(\mu) + \lambda H_{2k}(\mu)^T,\ \ 0 \neq \mu \neq (-1)^{k+1}$ μ *is determined up to replacement by* μ^{-1}	

We may establish a bijection $A \mapsto A + \lambda A^T$ between the set $\mathbb{C}^{n\times n}$ of matrices and the set of palindromic pencils $\{A + \lambda A^T : A \in \mathbb{C}^{n\times n}\}$, which induces a bijection between the congruence orbit of a given matrix A and the orbit of $A + \lambda A^T$ under congruence, i.e., $\{P(A + \lambda A^T)P^T : \quad P \text{ nonsingular}\}$. Hence Theorem 4 implies that the generic canonical form for congruence of palindromic pencils is

$$G_A + \lambda\, G_A^T, \tag{60}$$

where G_A is given by (57), if n is even, or by (58), if n is odd, and the μ_i are unspecified numbers that satisfy $\mu_i \neq \pm 1, \mu_i \neq \mu_j$ and $\mu_i \neq 1/\mu_j$ if $i \neq j$. From (60), we can get the generic Kronecker canonical form for strict equivalence [11, Chapter XII] of palindromic pencils as follows: (60) is strictly equivalent to $G_A^{-T}\, G_A + \lambda\, I_n$ and recall that the Jordan canonical form of $H_{2k}(\mu)^{-T}\, H_{2k}(\mu)$ is $J_k(\mu) \oplus J_k(\mu^{-1})$. This leads to Theorem 6.

Theorem 6. *The generic Kronecker canonical form of palindromic pencils in $\mathbb{C}^{n\times n}$ is*

1. *If n is even*:

$$\begin{aligned}&(\lambda + \mu_1) \oplus (\lambda + 1/\mu_1) \oplus (\lambda + \mu_2) \oplus (\lambda + 1/\mu_2) \oplus \cdots \\ &\quad \oplus (\lambda + \mu_{n/2}) \oplus (\lambda + 1/\mu_{n/2}),\end{aligned}$$

where $\mu_1, \ldots, \mu_{n/2}$ are unspecified complex numbers such that $0 \neq \mu_i \neq \pm 1$, $i = 1, \ldots, n/2$, $\mu_i \neq \mu_j$ and $\mu_i \neq 1/\mu_j$ if $i \neq j$.

2. *If n is odd*:

$$\begin{aligned}&(\lambda + \mu_1) \oplus (\lambda + 1/\mu_1) \oplus (\lambda + \mu_2) \oplus (\lambda + 1/\mu_2) \oplus \cdots \\ &\quad \oplus (\lambda + \mu_{(n-1)/2}) \oplus (\lambda + 1/\mu_{(n-1)/2}) \oplus (\lambda + 1),\end{aligned}$$

where $\mu_1, \ldots, \mu_{(n-1)/2}$ are unspecified complex numbers such that $0 \neq \mu_i \neq \pm 1, i = 1, \ldots, (n-1)/2$, $\mu_i \neq \mu_j$ and $\mu_i \neq 1/\mu_j$ if $i \neq j$.

Notice that the presence of the block $\lambda + 1$ associated with the eigenvalue -1 for n odd is not surprising. Any palindromic matrix pencil with odd size has always the eigenvalue -1, because $A - A^T$ is singular, as pointed out before.

7. Conclusions and future work

In this paper we have obtained the dimension of the solution space of the matrix equation $XA + AX^T = 0$, with $A \in \mathbb{C}^{n\times n}$, in terms of C_A, the canonical form for congruence of A, and we have shown how to find the general solution of this equation assuming that a nonsingular matrix P such that $C_A = PAP^T$ is known. This has allowed us to use C_A for computing the codimension of the orbit of A under the action of congruence. As a consequence, we have determined the generic canonical structure for congruence of matrices in $\mathbb{C}^{n\times n}$. These results can be directly extended to palindromic pencils $A + \lambda A^T$. This is the first step in describing the structure of the set of congruence orbits of matrices and palindromic matrix pencils. The following step would be to determine the inclusion relationships existing between the closures of these orbits. A description in the spirit of the one provided by Edelman, Elmroth and Kågström in [8,9] for the similarity orbits of matrices and the equivalence orbits of matrix pencils is in the aim of the authors, and remains as an open question and a field of future research. We also plan to extend the results in this paper to the matrix equation $XA + AX^* = 0$, an equation that is not linear in $\mathbb{C}$ and whose solution presents differences with the solution of $XA + AX^T = 0$.

Appendix A. The solution of $XJ_k((-1)^k) = -X^T$

This appendix is devoted to prove that the general solution of $XJ_k((-1)^k) = -X^T$ depends on $\lceil k/2 \rceil$ free parameters, a result that was used in the proof of Lemma 6. This result relies in a simple algorithm to determine the general solution of $XJ_k((-1)^k) = -X^T$ dealing with anti-diagonals, i.e., the sets of entries $\mathcal{L}_s = \{x_{ij} : i + j = s\}$ for $s = 2, 3, \ldots, 2k$. The strategy will be to prove that $x_{11} = 0$ and then to compute the entries in $\mathcal{L}_s$ from those in $\mathcal{L}_{s-1}$, which will require to consider some particular entries of X as free variables. We will present in detail the case k even, while, for brevity, we only state the main results for k odd.

A.1. Solution for k even

We present first necessary and sufficient conditions in terms of entries for a matrix X being solution of $XJ_k(1) = -X^T$.

Lemma 11. *Let $k > 0$ be an even number. A matrix $X = [x_{ij}]_{i,j=1}^k \in \mathbb{C}^{k\times k}$ is a solution of $XJ_k(1) = -X^T$ if and only if X satisfies the following four conditions*

$$x_{ij} = 0 \quad \text{if } 2 \leqslant i + j \leqslant k, \tag{61}$$

$$x_{k1} + x_{1k} = 0, \tag{62}$$

$$x_{ij} + x_{ji} = -x_{i,j-1} \quad \text{if } k + 1 \leqslant i + j \leqslant 2k \text{ and } 2 \leqslant j \leqslant i \leqslant k, \tag{63}$$

$$x_{i,j-1} = x_{j,i-1} \quad \text{if } k + 1 \leqslant i + j - 1 \leqslant 2k - 1 \text{ and } 2 \leqslant j < i \leqslant k. \tag{64}$$

Note, in particular, that every solution of $XJ_k(1) = -X^T$ is lower anti-triangular by (61).

Proof. $XJ_k(1) = -X^T$ is equivalent to $X(I + J_k(0)) = -X^T$, which is equivalent to $X + X^T = -XJ_k(0) = -[0\, X(:, 1 : k-1)]$, where we use MATLAB notation for submatrices. From this last equation, it follows that X is solution of $XJ_k(1) = -X^T$ if and only if X satisfies the following two conditions

$$x_{i1} + x_{1i} = 0 \quad \text{if } 1 \leqslant i \leqslant k, \tag{65}$$

$$x_{ij} + x_{ji} = -x_{i,j-1} \quad \text{if } 1 \leqslant i \leqslant k \text{ and } 2 \leqslant j \leqslant k. \tag{66}$$

Combining (65) with (66) evaluated at $i = 1$, we get that (65) and (66) are equivalent to the following three conditions

$$x_{i1} + x_{1i} = 0 \quad \text{if } 1 \leqslant i \leqslant k, \tag{67}$$
$$x_{1j} = 0 \quad \text{if } 1 \leqslant j \leqslant k-1, \tag{68}$$
$$x_{ij} + x_{ji} = -x_{i,j-1} \quad \text{if } 2 \leqslant i \leqslant k \text{ and } 2 \leqslant j \leqslant k. \tag{69}$$

Note that (69) implies that $x_{i,j-1} = x_{j,i-1}$ for $2 \leqslant i \leqslant k$ and $2 \leqslant j \leqslant k$. Then, some other elementary arguments allow us to prove that (67)–(69) are equivalent to

$$x_{i1} = x_{1i} = 0 \quad \text{if } 1 \leqslant i \leqslant k-1, \tag{70}$$
$$x_{k1} + x_{1k} = 0, \tag{71}$$
$$x_{ij} + x_{ji} = -x_{i,j-1} \quad \text{if } 2 \leqslant j \leqslant i \leqslant k, \tag{72}$$
$$x_{i,j-1} = x_{j,i-1} \quad \text{if } 2 \leqslant j < i \leqslant k. \tag{73}$$

We have proved so far that X is solution of $XJ_k(1) = -X^T$ if and only if X satisfies (70)–(73). On the other hand, it is immediate to see that conditions (61)–(64) in the statement of Lemma 11 imply (70)–(73), because (62)–(64) are precisely (71)–(73) for the entries in the lower anti-triangular part, and (61) implies that all entries in the strictly upper anti-triangular part are zero, so they satisfy the remaining equations (70)–(73). To complete the proof, we have to show that (70)–(73) imply (61)–(64), which reduces to get only (61) from (70)–(73). For this purpose, note that $x_{11} = 0$ from (70), and proceed by induction on anti-diagonals $\mathcal{L}_s = \{x_{ij} : i + j = s\}$. We assume that $\mathcal{L}_s = \{0\}$ for some $2 \leqslant s < k$ and we will prove that $\mathcal{L}_{s+1} = \{0\}$. From (70), we get $x_{1s} = x_{s1} = 0$, and from (73) to (72)

$$x_{2,s-1} = 0 \quad \text{and} \quad x_{s-1,2} = 0.$$

Repeatedly applying (73)–(72), we get

$$x_{3,s-2} = 0 \quad \text{and} \quad x_{s-2,3} = 0\,, \quad x_{4,s-3} = 0 \quad \text{and} \quad x_{s-3,4} = 0, \ldots,$$

i.e., $\mathcal{L}_{s+1} = \{0\}$. □

Observe that (61) amounts to $(k^2 - k)/2$ equations on the entries of X, (62) and (63) amount to $(k^2/4) + (k/2)$ equations, and (64) amounts to $(k^2/4) - (k/2)$ equations. This makes a total number of $k^2 - (k/2)$ equations in (61)–(64). Therefore, the general solution of $XJ_k(1) = -X^T$ depends on *at least* $(k/2)$ *free parameters* (it might depend on more that $(k/2)$ free parameters if equations (61)–(64) were linearly dependent). We will show in Lemma 12 that the general solution of $XJ_k(1) = -X^T$ depends precisely on $(k/2)$ free parameters, because if equations (61)–(64) are arranged in an appropriate order, then it is evident that certain $(k/2)$ entries of X determine uniquely the remaining ones. This appropriate order consists in ordering equations (62)–(64) by anti-diagonals in such a way that every anti-diagonal $\mathcal{L}_s$ is obtained from $\mathcal{L}_{s-1}$.

Lemma 12. *Let $k > 0$ be an even number, then the general solution X of $XJ_k(1) = -X^T$ depends on $k/2$ free variables. In particular, the entries*

$$x_{\frac{k+2}{2},\frac{k}{2}},\ x_{\frac{k+4}{2},\frac{k+2}{2}},\ x_{\frac{k+6}{2},\frac{k+4}{2}},\ldots,x_{k,k-1},$$

can be taken as free variables and then the remaining entries of X are uniquely determined by the following algorithm:

set $x_{ij} = 0$ if $2 \leqslant i + j \leqslant k$
for $s = k + 1 : 2k$
  if s is odd
    $h = \frac{s+1}{2}$
    $x_{h,h-1}$ is a free variable

$x_{h-1,h} = -x_{h,h-1} - x_{h,h-2}$
else
$h = \frac{s}{2}$
$x_{h,h} = -(x_{h,h-1})/2$
endif
for $i = h+1 : k$
$x_{i,s-i} = x_{s-(i-1),i-1}$
$x_{s-i,i} = -x_{i,s-i} - x_{i,s-i-1}$
endfor
endfor

For simplicity, in this algorithm we define $x_{k0} \equiv 0$ and it is understood that the inner loop "for $i = h+1 : k$" is not performed if $h+1 > k$.

Proof. Note that the algorithm arranges all the equations in (61)–(64) in an order that allows to compute each entry from entries that are already known. We only remark that $x_{h,h} = -(x_{h,h-1})/2$ is (63) with $i = j = h$, that $x_{h-1,h} = -x_{h,h-1} - x_{h,h-2}$ and $x_{s-i,i} = -x_{i,s-i} - x_{i,s-i-1}$ are (63) with appropriate indices, and that $x_{i,s-i} = x_{s-(i-1),i-1}$ is (64). Since we have already established that the general solution of $XJ_k(1) = -X^T$ depends on at least $(k/2)$ free parameters, and all the equations in (61)–(64) are satisfied in the algorithm in a unique way for any selection of arbitrary values of the $(k/2)$ entries $x_{h,h-1}$, for $h = (k+2)/2, (k+4)/2, (k+6)/2, \ldots, k$, then the number of free variables is precisely $k/2$. $\square$

A.2. Solution for k odd

We state without proofs counterparts of Lemmas 11 and 12. The proofs are similar to those of Lemmas 11 and 12 with the corresponding variations.

Lemma 13. *Let $k > 0$ be an odd number. A matrix $X = [x_{ij}]_{i,j=1}^{k} \in \mathbb{C}^{k\times k}$ is a solution of $XJ_k(-1) = -X^T$ if and only if X satisfies the following five conditions*

$$\begin{aligned}
&x_{ij} = 0 && \text{if } 2 \leqslant i+j \leqslant k,\\
&x_{k1} - x_{1k} = 0, && \\
&x_{i,i-1} = 0 && \text{if } i = \tfrac{k+3}{2}, \tfrac{k+3}{2}+1, \ldots, k,\\
&x_{ij} - x_{ji} = x_{i,j-1} && \text{if } k+1 \leqslant i+j \leqslant 2k-1 \text{ and } 2 \leqslant j < i \leqslant k,\\
&x_{i,j-1} = -x_{j,i-1} && \text{if } k+1 \leqslant i+j-1 \leqslant 2k-1 \text{ and } 2 \leqslant j < i \leqslant k.
\end{aligned}$$

Note, in particular, that every solution of $XJ_k(-1) = -X^T$ is lower anti-triangular.

Lemma 14. *Let $k > 0$ be an odd number, then the general solution X of $XJ_k(-1) = -X^T$ depends on $(k+1)/2$ free variables. In particular, the entries*

$$x_{\frac{k+1}{2},\frac{k+1}{2}},\ x_{\frac{k+3}{2},\frac{k+3}{2}},\ x_{\frac{k+5}{2},\frac{k+5}{2}},\ldots,x_{k,k},$$

can be taken as free variables and then the remaining entries of X are uniquely determined by the following algorithm:

set $x_{ij} = 0$ if $2 \leqslant i+j \leqslant k$
for $s = k+1 : 2k$
if s is odd
$h = \frac{s+1}{2}$
$x_{h,h-1} = 0$
$x_{h-1,h} = x_{h,h-1} - x_{h,h-2}$
else

$h = \frac{s}{2}$
$x_{h,h}$ is a free variable
endif
for $i = h + 1 : k$
$x_{i,s-i} = -x_{s-(i-1),i-1}$
$x_{s-i,i} = x_{i,s-i} - x_{i,s-i-1}$
endfor
endfor

For simplicity, in this algorithm we define $x_{k0} \equiv 0$ *and it is understood that the inner loop "for* $i = h + 1 : k$*" is not performed if* $h + 1 > k$.

Acknowledgements

The authors thank Peter Benner and Ulrike Baur for pointing out Ref. [4].

References

[1] V.I. Arnold, On matrices depending on parameters, Russian Math. Surveys 26 (1971) 29–43.
[2] M.A. Beitia, J.M. Gracia, Sylvester matrix equations for matrix pencils, Linear Algebra Appl. 232 (1996) 155–197.
[3] W.M. Boothby, An Introduction to Differentiable Manifolds and Riemannian Geometry, Academic Press, New York, 1975.
[4] H.W. Braden, The equations $A^TX \pm X^TA = B$, SIAM J. Matrix Anal. Appl. 20 (1998) 295–302.
[5] J.W. Demmel, A. Edelman, The dimension of matrices (matrix pencils) with given Jordan (Kronecker) canonical forms, Linear Algebra Appl. 230 (1995) 61–87.
[6] I. De Hoyos, Points of continuity of the Kronecker canonical form, SIAM J. Matrix Anal. Appl. 11 (1990) 278–300.
[7] F. De Terán, F.M. Dopico, A note on generic Kronecker orbits of matrix pencils with fixed rank, SIAM J. Matrix Anal. Appl. 30 (2008) 491–496.
[8] A. Edelman, E. Elmroth, B. Kågström, A geometric approach to perturbation theory of matrices and matrix pencils. Part I: Versal deformations, SIAM J. Matrix Anal. Appl. 18 (1997) 653–692.
[9] A. Edelman, E. Elmroth, B. Kågström, A geometric approach to perturbation theory of matrices and matrix pencils. Part II: A stratification-enhanced staircase algorithm, SIAM J. Matrix Anal. Appl. 20 (1999) 667–699.
[10] E. Elmroth, S. Johansson, B. Kågström, Stratification of controllability and observability pairs – theory and use in applications, SIAM J. Matrix Anal. Appl. 31 (2009) 203–226.
[11] F.R. Gantmacher, The Theory of Matrices, Chelsea, New York, 1959.
[12] M.I. García-Planas, M.D. Magret, Stratification of linear systems. Bifurcation diagrams for families of linear systems, Linear Algebra Appl. 297 (1999) 23–56.
[13] M. Hofer, N. Finger, J. Schöberl, S. Zaglmayr, U. Langer, R. Lerch, Finite element simulation of wave propagation in periodic piezoelectric SAW structures, IEEE Trans. UFFC 53 (2006) 1192–1201.
[14] Y.P. Hong, R.A. Horn, C.K. Li, Linear operators preserving t-congruence on matrices, Linear Algebra Appl. 175 (1992) 191–211.
[15] R.A. Horn, C.R. Johnson, Topics in Matrix Analysis, Cambridge University Press, Cambridge, 1994 (Corrected reprint of the 1991 original).
[16] R.A. Horn, V.V. Sergeichuk, Congruences of a square matrix and its transpose, Linear Algebra Appl. 389 (2004) 347–353.
[17] R.A. Horn, V.V. Sergeichuk, Canonical forms for complex matrix congruence and $*$congruence, Linear Algebra Appl. 416 (2006) 1010–1032.
[18] R.A. Horn, V.V. Sergeichuk, Canonical matrices of bilinear and sesquilinear forms, Linear Algebra Appl. 428 (2008) 193–223.
[19] T.M. Huang, W-W. Lin, J. Qiang, Structure preserving algorithms for palindromic quadratic eigenvalue problems arising from vibration of fast trains, SIAM J. Matrix Anal. Appl. 30 (2009) 1566–1592.
[20] P. Johansson, StratiGraph Software Design and Algorithms, UMINF 06.16, Department of Computing Science, Umeå University, Sweden, 2006.
[21] D.S. Mackey, N. Mackey, C. Mehl, V. Mehrmann, Structured polynomial eigenvalue problems: good vibrations from good linearizations, SIAM J. Matrix Anal. Appl. 28 (2006) 1029–1051.
[22] D.S. Mackey, N. Mackey, C. Mehl, V. Mehrmann, Numerical methods for palindromic eigenvalue problems: computing the anti-triangular Schur form, Numer. Linear Algebra Appl. 16 (2009) 63–86.
[23] V.V. Sergeichuk, Classification problems for systems of forms and linear mappings, Math. USSR Izv. 31 (3) (1988) 481–501.
[24] W.C. Waterhouse, The codimension of singular matrix pairs, Linear Algebra Appl. 57 (1984) 227–245.