Isotropy groups of the action of orthogonal similarity on symmetric matrices

Tadej Starčič

Published online: 04 Mar 2022.

Submit your article to this journal ↗

Article views: 715

View related articles ↗

View Crossmark data ↗

Taylor & Francis
Taylor & Francis Group

Isotropy groups of the action of orthogonal similarity on symmetric matrices

Tadej Starčič

Faculty of Education, University of Ljubljana, Ljubljana, Slovenia; Institute of Mathematics, Physics and Mechanics, Ljubljana, Slovenia

ABSTRACT
We find an algorithmic procedure that enables the computation and description of the structure of the isotropy subgroups of the group of complex orthogonal matrices with respect to the action of similarity on complex symmetric matrices. A key step in our proof is to solve a certain rectangular block upper triangular Toeplitz matrix equation.

1. Introduction

All matrices considered in this paper are complex unless otherwise stated. We use the notation $\mathbb{C}^{m \times n}$ for the set of matrices of size $m \times n$. By $S_n(\mathbb{C})$ we denote the vector space of all $n-$by$-n$ symmetric matrices; A is symmetric if and only if $A = A^T$. Let further $O_n(\mathbb{C})$ be the subgroup of orthogonal matrices in the group of nonsingular $n-$by$-n$ matrices $GL_n(\mathbb{C})$. A matrix Q is orthogonal if and only if $Q = (Q^T)^{-1}$. The action of *orthogonal similarity* on $S_n(\mathbb{C})$ is defined as follows:

$$\Phi \colon O_n(\mathbb{C}) \times S_n(\mathbb{C}) \to S_n(\mathbb{C}), \quad (Q, A) \mapsto Q^T A Q. \tag{1}$$

The *isotropy group* at $A \in S_n(\mathbb{C})$ with respect to action (1) is denoted by

$$\Sigma_A := \{Q \in O_n(\mathbb{C}) \mid Q^T A Q = A\}, \tag{2}$$

and matrices that are orthogonally similar to A form the *orbit* of A:

$$\mathrm{Orb}(A) := \{Q^T A Q \mid Q \in O_n(\mathbb{C})\}. \tag{3}$$

Isotropy groups corresponding to the same orbit are conjugated (isomorphic).

CONTACT Tadej Starčič ✉ tadej.starcic@pef.uni-lj.si 🖅 Faculty of Education, University of Ljubljana, Kardeljeva Ploščad 16, 1000Ljubljana, Slovenia Institute of Mathematics, Physics and Mechanics, Jadranska 19, 1000Ljubljana, Slovenia

Action (1) describes symmetries of $S_n(\mathbb{C})$. Hua's fundamental results [1–3] on the geometry of symmetric matrices assure that the study of symmetric matrices under T-congruence (which includes (1)) is quite general.

An important information concerning a group action is provided by its orbits and the corresponding isotropy groups (see textbooks [4,5]), and to find these for action (1) is the main purpose of this paper. In the (generic) case of matrices with all distinct eigenvalues, the isotropy groups are clearly trivial (Proposition 3.1), while the situation for (especially) nondiagonalizable matrices is more involved. We find an inductive procedure that enables the computation and description of the structure of the isotropy subgroups (Theorem 3.2). A key ingredient in the proof of the theorem is Lemma 4.1. It provides solutions of a certain rectangular upper triangular Toeplitz matrix equation ; hence, it might be also of independent interest in matrix analysis; e.g. in the paper by the author [6] (see Remark 5.1) a similar equation appeared.

To some extent (in lower dimensions), the isotropy groups are expected to be applied to tackle the problem of simultaneous reduction under T-congruence of a pair (A, B) with A arbitrary and B nonsingular symmetric. We first make B into the identity I by applying the Autonne–Takagi factorization and reduce (A, B) to (A', I). Next, we write $A' = C + Z$ with S symmetric and Z skew-symmetric. By a suitable orthogonal similarity transformation (keeping I intact), we put C into the symmetric normal form $S(C)$; we obtain $(S(C) + Z', I)$ with Z' skew-symmetric. Finally, Z' is simplified by using the isotropy group of $S(C)$ with respect to (1) (keeping I, $S(C)$ intact). We add that a reduction of symmetric pairs under T-congruence was considered in [7].

In connection to isotropy groups, we also mention that the so-called linear isotropy representation at $A \in S_n(\mathbb{C})$ is indeed the restriction of (1):

$$\Sigma_A \times T_A \to T_A, \quad (Q, A) \mapsto Q^T A Q, \quad T_A := \{X^T A + A X \mid X = -X^T \in \mathbb{C}^{n \times n}\}, \quad (4)$$

a representation of Σ_A on a complex vector space $T_A \subset S_n(\mathbb{C})$ associated to the tangent space of $\mathrm{Orb}(A) \subset S_n(\mathbb{C})$ at A (see also Section 3). It is closely related to invariant objects of $\mathrm{Orb}(A)$ (see, e.g. [[4, II.Ch.2.3];[8, Ch. X]]). On the other hand, (1) can be seen as a representation of $O_n(\mathbb{C})$; note that the classification of representations of complex classical groups along with their invariants is well understood (see [9]). We shall not consider this matter further here.

2. Preparatory material

In this section, we prepare some preliminary material.

First, let a *block upper triangular Toeplitz* matrix be

$$T(A_0, A_1, \ldots, A_{\beta-1}) = \begin{bmatrix} A_0 & A_1 & A_2 & \cdots & \cdots & A_{\beta-1} \\ 0 & A_0 & A_1 & A_2 & & \vdots \\ \vdots & \ddots & A_0 & A_1 & \ddots & \vdots \\ \vdots & & \ddots & \ddots & \ddots & A_2 \\ \vdots & & & \ddots & \ddots & A_1 \\ 0 & \cdots & \cdots & \cdots & 0 & A_0 \end{bmatrix} \quad (\beta\text{-by-}\beta),$$

$$A_0, A_1, \ldots, A_{\beta-1} \in \mathbb{C}^{m \times n},$$

where $T(A_0, A_1, \ldots, A_{\beta-1}) = [T_{jk}]_{j,k=1}^{\beta}$ with $T_{jk} = 0$ for $j > k$ and $T_{jk} = T_{(j+1)(k+1)}$.

The principal object of our investigation will be (up to similarity) the solutions of the homogeneous linear Sylvester equation $AX - XA = 0$ with a given square matrix A. Thus, we recall the following classical result (see, e.g. [10, Ch. VIII]).

Theorem 2.1: *Let J be of form* (11). *Let us consider the following matrix equation:*

$$JX = XJ. \tag{5}$$

(1) *Assume that $J = \bigoplus_{r=1}^{N} J_r$, in which all summands of J corresponding to the eigenvalue ρ_r are collected together into J_r. Then X is a solution of Equation (5) if and only if it is of the form $X = \bigoplus_{r=1}^{N} X_r$ with $J_r X_r = X_r J_r$.*

(2) *Let $J = \bigoplus_{r=1}^{N} \bigoplus_{j=1}^{m_r} J_{\alpha_j}(\lambda)$ for $\lambda \in \mathbb{C}$ and $\alpha_1 > \alpha_2 > \ldots > \alpha_N$, and let X be partitioned conformally into blocks as J. Then X is a solution of (5) if and only if $X = [X_{rs}]_{r,s=1}^{N}$ is such that every block X_{rs} is further an m_r-by-m_s block matrix with blocks of size $\alpha_r \times \alpha_s$ and of the form*

$$\begin{cases} [0 \ \ T], & \alpha_r < \alpha_s, \\ \begin{bmatrix} T \\ 0 \end{bmatrix}, & \alpha_r > \alpha_s, \\ T, & \alpha_r = \alpha_s, \end{cases} \tag{6}$$

in which T is an b_{rs}-by-b_{rs} upper triangular Toeplitz matrix ($b_{rs} = \min\{\alpha_r, \alpha_s\}$).

For our developments, it is convenient to work with matrices having fewer Toeplitz blocks. In the paper by Lin et al. [11, Sec. 3.1], this was achieved by conjugating with a suitable permutation matrix. Let $e_1, e_2, \ldots, e_{\alpha m}$ be the standard orthonormal basis in $\mathbb{C}^{\alpha m}$. We set a permutation matrix formed by these vectors:

$$\Omega_{\alpha,m} := \begin{bmatrix} e_1 \ e_{\alpha+1} \ \cdots \ e_{(m-1)\alpha+1} \ e_2 \ e_{\alpha+2} \ \cdots \ e_{(m-1)\alpha+2} \ \cdots \ e_\alpha \ e_{2\alpha} \ \cdots \ e_{\alpha m} \end{bmatrix}. \tag{7}$$

Observe that multiplication with $\Omega_{\alpha,m}$ from the right puts the 1 st, the $(\alpha + 1)$ th, …, the $((m - 1)\alpha + 1)$th column together, further the 2 nd, the $(\alpha + 2)$ th, …, the $((m - 1)\alpha + 2)$th column together, and so forth. Similarly, multiplying by $\Omega_{\alpha,m}^T$ from the left collects the 1 st, the $(\alpha + 1)$ th, …, the $((m - 1)\alpha + 1)$th row together, further the 2 nd, the $(\alpha + 2)$ th, …, the $((m - 1)\alpha + 2)$th row together, and so forth.

Suppose $X = [X_{rs}]_{r,s=1}^{N}$ is as in Theorem 2.1 (2.1). Next, fix r, s and let $b = \min\{\alpha_r, \alpha_s\}$. Denote the block of X_{rs} in the jth row and the kth column by

$$(X_{rs})_{jk} = \begin{cases} [0 \ \ T_{jk}], & \alpha_r < \alpha_s, \\ \begin{bmatrix} T_{jk} \\ 0 \end{bmatrix}, & \alpha_r > \alpha_s, \\ T_{jk}, & \alpha_r = \alpha_s, \end{cases} \qquad \begin{aligned} & j \in \{1, \ldots m_r\}, \quad k \in \{1, \ldots m_s\}, \\ & T_{jk} := T(a_0^{jk}, a_1^{jk}, \ldots, a_{b-1}^{jk}). \end{aligned}$$

By setting $A_n := [a_n^{jk}]_{j,k=1}^{m_r,m_s} \in \mathbb{C}^{m_r \times m_r}$ for $n \in \{0, \ldots, b-1\}$ and $\mathcal{T} = T(A_0, \ldots, A_{b-1})$, it gives a rectangular block upper triangular Toeplitz matrix of size $\alpha_r \times \alpha_s$:

$$\Omega_{\alpha_r,m_r}^T X_{rs} \Omega_{\alpha_s,m_s} = \begin{cases} [0 \ \ \mathcal{T}], & \alpha_r < \alpha_s, \\ \begin{bmatrix} \mathcal{T} \\ 0 \end{bmatrix}, & \alpha_r > \alpha_s, \\ \mathcal{T}, & \alpha_r = \alpha_s. \end{cases}$$

By defining

$$\mathcal{X} := \Omega^T X \Omega = [\Omega_{\alpha_r,m_r}^T X_{rs} \Omega_{\alpha_s,m_s}]_{r,s=1}^N, \quad (\Omega := \oplus_{r=1}^N \Omega_{\alpha_r,m_r}), \tag{8}$$

we obtain an $N-$by$-N$ block matrix such that its block $\mathcal{X}_{rs}$ is a rectangular $\alpha_r-$by$-\alpha_s$ block upper triangular Toeplitz matrix with blocks of size $m_r \times m_s$:

$$\mathcal{X} = [\mathcal{X}_{rs}]_{r,s=1}^N, \quad \mathcal{X}_{rs} = \begin{cases} [0 \ \ \mathcal{T}_{rs}], & \alpha_r < \alpha_s, \\ \begin{bmatrix} \mathcal{T}_{rs} \\ 0 \end{bmatrix}, & \alpha_r > \alpha_s, \quad b_{rs} := \min\{\alpha_s, \alpha_r\}, \\ \mathcal{T}_{rs}, & \alpha_r = \alpha_s, \end{cases} \tag{9}$$

in which $\mathcal{T}_{rs} = T(A_0^{rs}, \ldots, A_{b_{rs}-1}^{rs})$ with $A_j^{rs} \in \mathbb{C}^{m_r \times m_r}$ is a block upper triangular Toeplitz matrix of size $b_{rs} \times b_{rs}$.

Example 2.1: $N = 2, \alpha_1 = 3, m_1 = 2, \alpha_2 = 2, m_2 = 3$:

$$\Omega_{3,2}^T \left[\begin{array}{cc|cc|cc} a_1 & b_1 & a_2 & b_2 & a_3 & b_3 \\ 0 & a_1 & 0 & a_2 & 0 & a_3 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ \hline a_4 & b_4 & a_5 & b_5 & a_6 & b_6 \\ 0 & a_4 & 0 & a_5 & 0 & a_6 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array} \right] \qquad \Omega_{2,3} = \left[\begin{array}{ccc|ccc} a_1 & a_2 & a_3 & b_1 & b_2 & b_3 \\ a_4 & a_5 & a_6 & b_4 & b_5 & b_6 \\ 0 & 0 & 0 & a_1 & a_2 & a_3 \\ 0 & 0 & 0 & a_4 & a_5 & a_6 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array} \right].$$

Next, we show that the set of nonsingular matrices of form (9) has a group structure similar to the structure of the group of all nonsingular upper triangular matrices. We use ideas from the proof of a somewhat stronger result for upper unitriangular matrices [[12, Lecture 21];[5, Example 6.49]].

Lemma 2.2: *Let $\mathbb{T}$ be the set of all nonsingular matrices of form (9). Then $\mathbb{T}$ is a subgroup of the group of all nonsingular matrices. Furthermore, $\mathbb{T} = \mathbb{D} \ltimes \mathbb{U}$ is a semidirect product of subgroups, where $\mathbb{D} \subset \mathbb{T}$ contains all nonsingular block-diagonal matrices, and $\mathbb{U} \subset \mathbb{T}$ is a normal subgroup that consists of matrices whose diagonal blocks are block upper triangular Toeplitz matrices with identity as the diagonal block. Further, $\mathbb{U}$ is unipotent of order at most $\alpha_1 - 1$ and it has nilpotency class at most α_1.*

Proof: First, we examine the set $\mathbb{U}$ of all nonsingular matrices of form (9) such that their diagonal blocks are block upper triangular Toeplitz matrices with identities as the diagonal blocks.

For $k \in \{1, \ldots, \alpha_1 - 1\}$, let $\mathfrak{N}_k$ be the set of nonsingular matrices of form (9) with $\mathcal{T}_{rs} = T(0, \ldots, 0, A_k^{rs}, A_{k+1}^{rs}, \ldots, A_{b_{rs}-1}^{rs})$ (i.e. $A_0^{rs} = \ldots = A_{k-1}^{rs} = 0$) for $b_{rs} > k$ and $\mathcal{T}_{rs} = 0$ for $k \geq b_{rs}$, and such that $A_k^{rr} = 0$ for all r. Thus

$$\mathbb{U} - \mathcal{I} =: \mathfrak{N}_0 \supset \mathfrak{N}_1 \supset \cdots \supset \mathfrak{N}_{\alpha_1-1} = \{0\}.$$

Sums and products of rectangular upper triangular Toeplitz matrices of the appropriate size are again rectangular upper triangular Toeplitz matrices. Moreover,

$$\mathfrak{N}_k + \mathfrak{N}_k \subset \mathfrak{N}_k, \quad \mathfrak{N}_0 \mathfrak{N}_k \subset \mathfrak{N}_{k+1}, \quad \mathfrak{N}_k \mathfrak{N}_0 \subset \mathfrak{N}_{k+1}.$$

In particular, $\mathfrak{N}_k^{\alpha_1-k-1} = \{0\}$; thus matrices in $\mathfrak{N}_k$ are nilpotent. For $\mathcal{N} \in \mathfrak{N}_k$, we have

$$(\mathcal{I} + \mathcal{N})^{-1} = \mathcal{I} - \mathcal{N} + \mathcal{N}^2 - \ldots + (-1)^{\alpha_1-k-1} \mathcal{N}^{\alpha_1-k-1}.$$

Hence $\mathbb{U}_k := \mathcal{I} + \mathfrak{N}_k$ is a unipotent group. Taking $\mathcal{I} + \mathcal{N} \in \mathbb{U}_k$ (with $\mathcal{N} \in \mathfrak{N}_k$) and $\mathcal{I} + \mathcal{N}' \in \mathbb{U}$ (with $\mathcal{N}' \in \mathfrak{N}_0$), we get their conjugate and their commutator of the form:

$$(\mathcal{I} + \mathcal{N}')^{-1}(\mathcal{I} + \mathcal{N})(\mathcal{I} + \mathcal{N}') = \mathcal{I} + \left((\mathcal{I} - \mathcal{N}' + (\mathcal{N}')^2 - \ldots)\mathcal{N}(\mathcal{I} + \mathcal{N}')\right) \in \mathbb{U}_k,$$

$$[\mathcal{I} + \mathcal{N}, \mathcal{I} + \mathcal{N}'] = (\mathcal{I} + \mathcal{N})^{-1}(\mathcal{I} + \mathcal{N}')^{-1}(\mathcal{I} + \mathcal{N})(\mathcal{I} + \mathcal{N}')$$

$$= \left((\mathcal{I} - \mathcal{N} + \mathcal{N}^2 - \ldots)(\mathcal{I} - \mathcal{N}' + (\mathcal{N}')^2 - \ldots)\right)$$

$$\times \left((\mathcal{I} + \mathcal{N})(\mathcal{I} + \mathcal{N}')\right)$$

$$= (\mathcal{I} - \mathcal{N} - \mathcal{N}' + \mathcal{M}_1)(\mathcal{I} + \mathcal{N} + \mathcal{N}' + \mathcal{M}_2)$$

$$= \mathcal{I} + \mathcal{M}_3 \in \mathbb{U}_{k+1},$$

where $\mathcal{M}_1, \mathcal{M}_2, \mathcal{M}_3 \in \mathfrak{N}_{k+1}$. Therefore,

$$\mathbb{U} = \mathbb{U}_0 \supset \mathbb{U}_1 \supset \cdots \supset \mathbb{U}_{\alpha_1-1} = \{\mathcal{I}\} \tag{10}$$

is a central series of normal subgroups, i.e. $[\mathbb{U}, \mathbb{U}_j]$ is a commutator group of $\mathbb{U}_{j+1}$.

Any $\mathcal{X} \in \mathbb{T}$ (nonsingular and of form (9)) can be written as $\mathcal{X} = \mathcal{D}\mathcal{U}$, where $\mathcal{U} \in \mathbb{U}$ and $\mathcal{D} \in \mathbb{D}$ is a nonsingular block-diagonal matrix of form (9). For $\mathcal{D}_1, \mathcal{D}_2 \in \mathbb{D}$ and $\mathcal{U}_1, \mathcal{U}_2 \in \mathbb{U}$, we get that $(\mathcal{D}_1\mathcal{U}_1)(\mathcal{D}_2\mathcal{U}_2)^{-1} = \mathcal{D}_1(\mathcal{U}_1\mathcal{U}_2^{-1})\mathcal{D}_2^{-1}$ is of the form (9), thus $\mathbb{T}$ is a group. Next, conjugating $\mathcal{I} + \mathcal{N} \in \mathcal{I} + \mathfrak{N}_0 = \mathbb{U}$ by $\mathcal{D}\mathcal{U}$ gives

$$\mathcal{U}^{-1}\mathcal{D}^{-1}(\mathcal{I} + \mathcal{N})\mathcal{D}\mathcal{U} = \mathcal{I} + \mathcal{U}^{-1}\mathcal{D}^{-1}\mathcal{N}\mathcal{D}\mathcal{U} \in \mathbb{U}.$$

This proves the normality of $\mathbb{U}$ and concludes the proof. ∎

Remark 2.1: It would be interesting to know whether (10) is a lower central sequence or not. Note that the situation seems more involved than in the case of upper unitriangular matrices, in which the commutators of suitably chosen generators are again generators ([12, Proposition 3.31]).

3. Isotropy groups

To be able to compute the isotropy groups under (1), it is important to have simple representatives of orbits. Given a matrix A with its Jordan canonical form:

$$J(A) = \bigoplus_j J_{n_j}(\lambda_j), \quad \lambda_j \in \mathbb{C},$$

$$J_n(\lambda) := \begin{bmatrix} \lambda & 1 & & 0 \\ & \lambda & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda \end{bmatrix}, \quad \lambda \in \mathbb{C} \quad (n\text{–by–}n), \tag{11}$$

the symmetric canonical form under similarity is

$$S(A) = \bigoplus_j K_{n_j}(\lambda_j), \tag{12}$$

in which

$$K_n(\lambda) := \frac{1}{2}\left(\begin{bmatrix} 2\lambda & 1 & & 0 \\ 1 & \ddots & \ddots & \\ & \ddots & \ddots & 1 \\ 0 & & 1 & 2\lambda \end{bmatrix} + i \begin{bmatrix} 0 & & -1 & 0 \\ & \ddots & \ddots & 1 \\ -1 & \ddots & \ddots & \\ 0 & 1 & & 0 \end{bmatrix} \right), \quad \lambda \in \mathbb{C} \quad (n\text{–by–}n). \tag{13}$$

It is uniquely determined up to a permutation of its direct summands; recall that symmetric matrices are similar if and only if they are orthogonally similar (see the textbook by Gantmacher [10, Ch. XI]). For tridiagonal normal forms check [13].

Since $Q^T A Q = A$ is equivalent to $(J(A))X = X(J(A))$ with $J(A) = PAP^{-1}, X = PQP^{-1}$, the following fact on isotropy groups follows immediately from Theorem 2.1 (2.1).

Proposition 3.1: *If $\lambda_1, \ldots, \lambda_k$ are distinct eigenvalues of $A = \bigoplus_{j=1}^k S_j$, where each S_j is a direct sum whose summands are of form (13) and correspond to the eigenvalue λ_j, it then follows that $\Sigma_A = \bigoplus_{j=1}^k \Sigma_{S_j}$. Furthermore, if $S_j = \lambda_j I_{n_j}$ for some index j, then $\Sigma_{S_j} = O_{n_j}(\mathbb{C})$. (We denote the $n \times n$ identity matrix by I_n.)*

It implies that the isotropy groups under (1) are trivial on a complement of a complex analytic subset of codimension 1 in $S_n(\mathbb{C})$ (having nonvanishing discriminants of their characteristic polynomials).

Our major goal is to inspect the isotropy groups of nongeneric matrices (especially non-diagonalizable). We shall see later on that these are related to matrices of form (9) such that the following conditions are satisfied:

(I) For $r, s \in \{1, \ldots, N\}$ the nonzero entries of $\mathcal{X}_{rs}$ with $r > s$ can be taken as free variables, and $(\mathcal{X}_{rr})_{11} = A_0^{rr} \in O_{m_r}(\mathbb{C})$ can be any orthogonal matrix.

(II) For $r \in \{1, \ldots, N\}$ with $\alpha_r \geq 2$ let $Z_{j'}^r = -(Z_{j'}^r)^T$ for $j' \in \{1, \ldots, \alpha_r - 1\}$ be a freely chosen skew-symmetric matrix of size $m_r \times m_r$. The entries of A_j^{rs} for either $r = s$,

$\alpha_r \geq 2, j \in \{1, \ldots, \alpha_r - 1\}$ or $s > r$, $N \geq 2$ are determined uniquely by the entries of all $A_{\widetilde{j}}^{r's'}$ with $\widetilde{j} = 0, s' = r'$ or $s' \leq r', \widetilde{j} \in \{0, \ldots, \alpha_{r'} - 1\}$ (chosen in (I)), by the entries of all $Z_{\widetilde{j}}^{r'}$ with $\widetilde{j} \in \{1, \ldots, j - 1\}$ (if $j \geq 2$) or $\widetilde{j} = j$, $r = r'$, and when $s > r$, $N \geq 2$ also by the entries of $Z_j^{r'}$ for all r'; here $r', s' \in \{1, \ldots, N\}$.

Two significant examples of matrices of form (9) satisfying (I) and (II) are:

Example 3.1: Given skew-symmetric Z_n^r for $r \in \{1, \ldots, N\}$, $n \in \{1, \ldots, \alpha_r - 1\}$, we set

$$\mathcal{W} = \bigoplus_{r=1}^{N} T(I_{m_r}, W_1^r, \ldots, W_{\alpha_r-1}^r),$$

$$W_1^r := \frac{1}{2} Z_1^r, \quad W_{n+1}^r := \frac{1}{2} \left(Z_{n+1}^r - \sum_{j=1}^{n} (W_j^r)^T W_{n-j+1}^r \right), \quad n \geq 1. \tag{14}$$

Example 3.2: The following matrix contains the identity matrix as a principal submatrix, formed by all blocks except those at the pth and the tth columns and rows, while blocks in the pth and the tth columns and rows are as follows:

$$\mathcal{G}_{p,t}^k(F) = [(\mathcal{G}_{p,t}^k(F))_{rs}]_{r,s=1}^N, \quad (\mathcal{G}_{p,t}^k(F))_{rs} = \begin{cases} \begin{bmatrix} 0 & \mathcal{U}_{rs} \end{bmatrix}, & \alpha_r < \alpha_s, \\ \begin{bmatrix} \mathcal{U}_{rs} \\ 0 \end{bmatrix}, & \alpha_r > \alpha_s, & p < t, \\ \mathcal{U}_{rs}, & \alpha_r = \alpha_s, \end{cases} \tag{15}$$

where

$$\mathcal{U}_{rs} = \begin{cases} \oplus_{j=1}^{\alpha_r} I_{m_r}, & r = s, \\ 0, & r \neq s, \end{cases}, \quad \{r, s\} \not\subset \{p, t\}, \quad \mathcal{U}_{rr} = T(I_{m_r}, A_1^{rr}, \ldots, A_{\alpha_r-1}^{rr}), \quad r \in \{p, t\},$$

$$A_j^{pp} = \begin{cases} a_{n-1}(F^T F)^n, & j = n(2k + \alpha - \beta), \\ 0, & \text{otherwise}, \end{cases} \quad a_n = -\frac{1}{2^{2n+1}} \frac{1}{n+1} \binom{2n}{n},$$

$$A_j^{tt} = \begin{cases} a_{n-1}(F F^T)^n, & j = n(2k + \alpha - \beta), \\ 0, & \text{otherwise}, \end{cases}$$

$$\mathcal{U}_{pt} = N_{\alpha_t}^k(F), \quad \mathcal{U}_{tp} = N_{\alpha_t}^k(-F^T), \quad 0 \leq k \leq \alpha_t - 1,$$

in which $N_\beta^k(F)$ is a β−by−β block matrix with $F \in \mathbb{C}^{m_p \times m_t}$ on the kth diagonal above the main diagonal for $k \geq 1$ (on the main diagonal for $k = 0$) and zeros otherwise.

If $N = 3$, $\alpha_1 = 4$, $\alpha_2 = 2$, $\alpha_3 = 1$, $m_1 = 2$, $m_2 = 3$, $m_3 = 1$, $F \in \mathbb{C}^{2 \times 3}$, we have

$$\mathcal{G}_{1,2}^0(F) = \left[\begin{array}{cccc|ccc} I_2 & 0 & -\frac{1}{2}F^T F & 0 & -F^T & 0 & 0 \\ 0 & I_2 & 0 & -\frac{1}{2}F^T F & 0 & -F^T & 0 \\ 0 & 0 & I_2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & I_2 & 0 & 0 & 0 \\ \hline 0 & 0 & F & 0 & I_3 & 0 & 0 \\ 0 & 0 & 0 & F & 0 & I_3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{array} \right].$$

The following theorem is our main result; we prove it in Section 5.

Theorem 3.2: *If $A = \bigoplus_{r=1}^{N}(\bigoplus_{j=1}^{m_r} K_{\alpha_r}(\lambda))$ for $\lambda \in \mathbb{C}$, then its isotropy group Σ_A with respect to (1) has the following properties:*

(1) *It is isomorphic to the subgroup of the group of all invertible matrices of form (9) and such that its elements satisfy conditions (I) and (II). In particular, $\dim(\Sigma_S) = \sum_{r=1}^{N} \alpha_r m_r(\frac{1}{2}(m_r - 1) + \sum_{s=1}^{r-1} m_s)$.*
(2) *It is isomorphic to a semidirect product of groups, i.e. $\Sigma_S \cong \mathbb{O} \ltimes \mathbb{V}$, in which the subgroup $\mathbb{O}$ consists of all matrices of the form $\mathcal{Q} = \bigoplus_{r=1}^{N}(\bigoplus_{j=1}^{\alpha_r} Q_r)$ with $Q_r \in O_{m_j}(\mathbb{C})$, and a unipotent normal subgroup $\mathbb{V}$ (of order at most $\alpha_1 - 1$ and nilpotency class at most α_1) generated by the set of matrices of form (14) and (15).*

We refer to [5, Chs. 6 and 14] for a comprehensive introduction to the theory of nilpotent and unipotent algebraic groups.

Remark 3.1: An algorithm to compute the isotropy groups is provided as (the essential) part of the proof of Theorem 3.2, more precisely, by Lemma 4.1. Due to technical reasons, the lemma is stated and proved in Section 4.

Basic properties of an action of a Lie group on a manifold (check, e.g. [4, Ch. II.1]) imply that orbits of action (1) are immersed homogeneous submanifolds in $S_n(\mathbb{C})$ and the isotropy group Σ_A for any $A \in S_n(\mathbb{C})$ is a Lie subgroup of $O_n(\mathbb{C})$. Moreover, the orbit $\mathrm{Orb}(A)$ is biholomorphic to the quotient $O_n(\mathbb{C})/\Sigma_A$, and the codimension of Σ_A in $O_n(\mathbb{C})$ is equal to the codimension of $\mathrm{Orb}(A) \subset S_n(\mathbb{C})$.

Corollary 3.3: *If $\lambda_1, \ldots, \lambda_k$ are distinct eigenvalues of $S = \bigoplus_{j=1}^{k} S_j$, where each $\mathcal{S}_j$ is a direct sum whose summands are of form (13) and correspond to the eigenvalue λ_j, then $\mathrm{codim}(\mathrm{Orb}(S)) = \sum_{j=1}^{k} \mathrm{codim}(\mathrm{Orb}(S_j))$. Moreover, if $S = \bigoplus_{r=1}^{N}(\bigoplus_{j=1}^{m_r} S_{\alpha_r}(\lambda))$ for $\lambda \in \mathbb{C}$, it then follows that $\mathrm{codim}(\mathrm{Orb}(S)) = \sum_{r=1}^{N} \alpha_r m_r(\frac{1}{2}(m_r + 1) + \sum_{s=1}^{r-1} m_s)$.*

Although the above corollary is an immediate consequence of Theorem 3.2, we also give a direct simple proof, since the tangent space of $\mathrm{Orb}(A)$ at A (see T_A in (4)) is easily computed. Indeed, if $Q(t)$ is a complex-differentiable path of orthogonal matrices with $Q(0) = I$, then

$$\frac{\mathrm{d}}{\mathrm{dt}}\Big|_{t=0}\big((Q(t))^T A Q(t)\big) = (Q'(0))^T A + A Q'(0),$$

and differentiation of $(Q(t))^T Q(t) = I$ at $t = 0$ yields $(Q'(0))^T + Q'(0) = 0$; conversely, for any $X = -X^T$, we have e^{tX} orthogonal with $e^{0 \cdot X} = I$ and $\frac{\mathrm{d}}{\mathrm{dt}}|_{t=0}(e^{tX}) = X$.

The dimension of T_A in (4) is precisely the codimension of the solution space of $X^T A + AX = 0$ with $X = -X^T$ (with respect to the space of skew-symmetric matrices). If J is the

Jordan form of a symmetric matrix A with $A = P^{-1}JP$, we get

$$JY = YJ, \quad Y = PXP^{-1}, \quad X = -X^T.$$

Thus $Y = [Y_{jk}]_{j,k=1}^N$ has rectangular upper triangular Toeplitz blocks Y_{jk} (see Theorem 2.1), and $Y = -PX^TP^{-1} = -P^2Y^TP^{-2}$. Note that (see, e.g. [14, Theorem 4.4.24]):

$$K_\alpha(\lambda) = P_\alpha J_\alpha(\lambda)P_\alpha^{-1}, \quad P_\alpha := \frac{1}{\sqrt{2}}(I_\alpha + iE_\alpha), \quad E_\alpha = \begin{bmatrix} 0 & & 1 \\ & \cdot^{\cdot^{\cdot}} & \\ 1 & & 0 \end{bmatrix} \quad (\alpha-\text{by}-\alpha), \tag{16}$$

in which E_α is the backward identity matrix (with ones on the anti-diagonal); $P_\alpha^2 = E_\alpha$. If A is of form (12), then $Y = -EY^TE$, in which E is a direct sum of backward identity matrices and it is partitioned conformally to Y. We further obtain that all $Y_{jj} = 0$ and $Y_{jk} = \begin{bmatrix} T_{jk} \\ 0 \end{bmatrix}$, $Y_{kj} = [\, 0 \; T_{kj} \,]$ with both T_{jk}, T_{kj} upper triangular Toeplitz and related by $T_{jk} = -T_{kj}$ (see also (25) and (26)). This proves Corollary 3.3.

Note that codim(Orb(A)) in $\mathbb{C}^{n \times n}$ with respect to similarity is equal to the dimension of the set of solutions of $AX - XA = 0$ (e.g. [15, Section 30]), while to get codim(Orb(A)) under T-congruence De Teran and Dopico [16] solved $XA + AX^T = 0$.

4. Certain block upper triangular Toeplitz matrix equation

In this section, we consider a certain block upper triangular Toeplitz matrix equation. Its solution (Lemma 4.1) is the key ingredient in the proof of Theorem 3.2.

Let $\alpha_1 > \alpha_2 > \ldots > \alpha_N$ and $m_1, \ldots, m_N \in \mathbb{N}$. Suppose

$$\mathcal{B} = \bigoplus_{r=1}^N T(B_0^r, B_1^r, \ldots, B_{\alpha_r-1}^r), \quad \mathcal{C} = \bigoplus_{r=1}^N T(C_0^r, C_1^r, \ldots, C_{\alpha_r-1}^r), \quad \mathcal{F} = \bigoplus_{r=1}^N E_{\alpha_r}(I_{m_r}),$$

$$B_0^r, C_0^r \in GL_{m_r}(\mathbb{C}) \cap S_{m_r}(\mathbb{C}), \quad B_1^r, C_1^r, \ldots, B_{\alpha_r-1}^r, C_{\alpha_r-1}^r \in S_{m_r}(\mathbb{C}), \tag{17}$$

where $E_\alpha(I_m) = \begin{bmatrix} 0 & & I_m \\ & \cdot^{\cdot^{\cdot}} & \\ I_m & & 0 \end{bmatrix}$ is an $\alpha-$by$-\alpha$ block matrix with I_m on the anti-diagonal and zero-matrices otherwise. We shall solve a matrix equation

$$\mathcal{C} = \mathcal{F}\mathcal{X}^T\mathcal{F}\mathcal{B}\mathcal{X}, \tag{18}$$

in which $\mathcal{X} = [\mathcal{X}_{rs}]_{r,s=1}^N$ is as in (9).

We first observe a few simple facts.

The calculation

$$(\mathcal{F}\mathcal{X}^T\mathcal{F}\mathcal{B}\mathcal{X})^T = \mathcal{X}^T\mathcal{B}^T\mathcal{F}\mathcal{X}\mathcal{F} = \mathcal{F}\mathcal{F}\mathcal{X}^T\mathcal{F}(\mathcal{F}\mathcal{B}^T\mathcal{F})\mathcal{X}\mathcal{F} = \mathcal{F}(\mathcal{F}\mathcal{X}^T\mathcal{F}\mathcal{B}\mathcal{X})\mathcal{F}$$

shows that for $r \neq s$ we have $(\mathcal{F}\mathcal{X}^T\mathcal{F}\mathcal{B}\mathcal{X})_{rs} = 0$ if and only if $(\mathcal{F}\mathcal{X}^T\mathcal{F}\mathcal{B}\mathcal{X})_{sr} = 0$. When comparing the left-hand side with the right-hand side of (18) blockwise, it thus suffices to observe only blocks in the upper triangular parts of $\mathcal{F}X^T\mathcal{F}\mathcal{B}X$ and $\mathcal{C}$. Since $(\mathcal{F}\mathcal{X}^T\mathcal{F}\mathcal{B}\mathcal{X})_{rs}$

and C_{rs} are rectangular block upper triangular Toeplitz and of the same form for each r, s, it is enough to compare the first rows of these blocks.

The following lemma explains the process of computing solutions of (18). Note that in the proof of Theorem 3.2 we shall consider (18) for $\mathcal{B}$ and $\mathcal{C}$ equal to the identity matrix. However, due to a possible application when computing isotropy groups of actions similar to (1) (see Remark 5.1) and since it makes no serious difference to the proof, we prove a little more general result.

Lemma 4.1: *Let $\mathcal{B}$, $\mathcal{C}$ as in (17) be given. Then the dimension of the space of solutions of (18) that are of the form $\mathcal{X} = [\mathcal{X}_{rs}]_{r,s=1}^{N}$ (partitioned conformally to $\mathcal{B}$, $\mathcal{C}$) with*

$$
\mathcal{X}_{rs} = \begin{cases} [0 \quad \mathcal{T}_{rs}], & \alpha_r < \alpha_s, \\ \begin{bmatrix} \mathcal{T}_{rs} \\ 0 \end{bmatrix}, & \alpha_r > \alpha_s, \\ \mathcal{T}_{rs}, & \alpha_r = \alpha_s, \end{cases} \quad \begin{matrix} \alpha_1 > \alpha_2 > \ldots > \alpha_N, \quad b_{rs} := \min\{\alpha_s, \alpha_r\} \\ \mathcal{T}_{rs} = T\big(A_0^{rs}, A_1^{rs}, \ldots, A_{b_{rs}-1}^{rs}\big), \quad A_j^{rs} \in \mathbb{C}^{m_r \times m_s} \end{matrix} \tag{19}
$$

is $\sum_{r=1}^{N} \alpha_r m_r (\frac{1}{2}(m_r - 1) + \sum_{s=1}^{r-1} m_s)$. In particular, the general solution satisfies the following properties:

(a) *The entries of A_0^{rr} for $r \in \{1, \ldots, N\}$ can be taken so that A_0^{rr} is any solution of the equation $C_0^r = (A_0^{rr})^T B_0^r A_0^{rr}$. If $N \geq 2$ the entries of A_j^{rs} for $r, s \in \{1, \ldots, N\}$ with $r > s$ and $j \in \{0, \ldots, \alpha_r - 1\}$ can be taken as free variables.*

(b) *Assuming (a) and choosing the entries of matrices $Z_j^r = -Z_j^r \in \mathbb{C}^{m_r \times m_r}$ for $r \in \{1, \ldots, N\}$, $\alpha_r - 1 \geq j \geq 1$ as free variables, the remaining entries of $\mathcal{X}$ are computed by the following algorithm:*

$$\Psi_n^{krs} := \sum_{j=0}^{n} \sum_{l=0}^{n-j} (A_j^{kr})^T B_{n-j-l}^k A_l^{ks}$$

for $j = 0 : \alpha_1 - 1$ **do**

    if $r \in \{1, \ldots, N\}$, $1 \leq j \leq \alpha_r - 1$ **then**

$$A_j^{rr} = A_0^{rr} - A_0^{rr}(C_0^r)^{-1}\Big(Z_j^r + \sum_{l=1}^{j-1}\sum_{m=0}^{n-j}(A_l^{rr})^T B_{n-j-m}^k A_m^{kr}$$
$$+ \sum_{k=1}^{r-1} \Psi_{j-\alpha_k+\alpha_r}^{rr} + \sum_{k=r+1}^{N} \Psi_{j-\alpha_r+\alpha_k}^{rr}\Big)$$

    end if

    for $p = 1 : N - 1$ **do**

        if $r \in \{1, \ldots, N\}$, $j \leq \alpha_{r+p} - 1$, $r + p \leq N$ **then**

$$A_j^{r(r+p)} = -A_0^{r(r+p)}(C_0^r)^{-1}\Big(\sum_{l=1}^{j}\sum_{m=0}^{n-j}(A_l^{rr})^T B_{n-j-m}^k A_m^{k(r+p)} + \sum_{k=1}^{r-1} \Psi_{j-\alpha_k+\alpha_r}^{r(r+p)}$$
$$+ \sum_{k=r+1}^{r+p} \Psi_{j}^{r(r+p)} + \sum_{k=r+p+1}^{N} \Psi_{j-\alpha_{r+p}+\alpha_k}^{r(r+p)}\Big),$$

        end if

    end for

end for

    For simplicity, in this algorithm, we define $\sum_{j=l}^{n} a_j = 0$ if $l > n$, and it is understood that the inner loop (i.e. for $p = 1: N-1$) is not performed for $N = 1$.

Furthermore, assume that $\mathcal{B}$ and $\mathcal{C}$ are real. Then the solution $\mathcal{X}$ is real if and only if the following statements hold:

(i) *Matrices B_0^r and C_0^r in (17) have the same inertia for all $r \in \{1, \ldots, N\}$.*

(ii) *Matrices A_j^{rs} with $r > s$, $j \in \{0, \ldots, \alpha_r - 1\}$, $N \geq 2$, matrices A_0^{rr}, and matrices Z_j^r for $1 \leq j \leq \alpha_r - 1$ ($r, s \in \{1, \ldots, N\}$) in (a) and (b) are chosen real.*

For the sake of clarity, we point out the importance of the correct order of calculating the entries of $\mathcal{X}$ in Lemma 4.1. It is essential for the proof of the lemma.

Recall first that by (a) (when $N \geq 2$) all entries of the blocks below the main diagonal of $\mathcal{X} = [\mathcal{X}_{rs}]_{r,s=1}^N$ can be chosen freely. Next, we compute the diagonal entries of the blocks in the upper triangular part of $\mathcal{X}$. We first obtain the diagonal entries of the main diagonal blocks $\mathcal{X}_{rr}$ for $r \in \{1, \ldots, N\}$ (see (a) again). Secondly, step $j = 0$, $p = 1$ (if $N \geq 2$) of the algorithm in (b) yields the diagonal entries of the first upper off-diagonal blocks of $\mathcal{X}$ (i.e. $(\mathcal{X}_{r(r+1)})_{11} = A_0^{r(r+1)}$). Further, step $j = 0, p = 2$ gives the diagonal entries of the second upper off-diagonal blocks of $\mathcal{X}$ (i.e. $(\mathcal{X}_{r(r+2)})_{11} = A_0^{r(r+2)}$), step $j = 0, p = 3$ gives the diagonal entries of the third upper off-diagonal blocks of $\mathcal{X}$ (i.e. $(\mathcal{X}_{r(r+3)})_{11} = A_0^{r(r+3)}$), and so forth. In the same fashion, the step for fixed $j \in \{1, \ldots, \alpha_1 - 1\}$, $p \in \{0, \ldots, N\}$ yields the entries on the jth upper off-diagonals of the pth upper off-diagonal blocks of $\mathcal{X}$, i.e. $(\mathcal{X}_{r(r+p)})_{1(j+1)} = A_{j+1}^{r(r+p)}$ with $r + p \leq N$, provided that $j \leq \alpha_{r+p} - 1$. Finally, when $j = \alpha_1 - 1$, $p = 0$, we compute $(\mathcal{X}_{11})_{1\alpha_1} = A_{\alpha_1-1}^{11}$. Note that we add $\frac{1}{2} m_r(m_r - 1)$ free variables when calculating each entry $A_j^{rr} \in \mathbb{C}^{m_r \times m_r}$.

Proof of Lemma 4.1: The idea is to write Equation (18) entrywise as a system of several simpler matrix equations and then consider them in an appropriate order.

First, we analyse the right-hand side of Equation (18) for $\mathcal{B}, \mathcal{F}$ of form (17) and $\mathcal{X} = [\mathcal{X}_{rs}]_{r,s=1}^N$ with blocks as in (19). To simplify the notation we set $\mathcal{Y} := \mathcal{B}\mathcal{X}$ and $\widetilde{\mathcal{X}} := \mathcal{F}X^T\mathcal{F}$. The entries in the jth column and in the first row of $(\widetilde{\mathcal{X}}\mathcal{Y})_{rs}$ are obtained by multiplying the first rows of the blocks $\widetilde{\mathcal{X}}_{r1}, \ldots, \widetilde{\mathcal{X}}_{rN}$ with the jth columns of the blocks $(\mathcal{Y})_{1s}, \ldots, (\mathcal{Y})_{Ns}$, respectively, and then adding them:

$$((\widetilde{\mathcal{X}}\mathcal{Y})_{rs})_{1j} = \sum_{k=1}^N (\widetilde{\mathcal{X}}_{rk})_{(1)}(\mathcal{Y}_{ks})^{(j)}, \quad r,s \in \{1, \ldots, N\}, \quad j \in \{1, \ldots, \alpha_s\}. \tag{20}$$

As mentioned in the discussion at the beginning of this section, it suffices to analyse the upper triangular blocks of $\widetilde{\mathcal{X}}\mathcal{Y}$:

$$((\widetilde{\mathcal{X}}\mathcal{Y})_{r(r+p)})_{1j} = \sum_{k=1}^N (\widetilde{\mathcal{X}}_{rk})_{(1)}(\mathcal{Y}_{k(r+p)})^{(j)}, \quad 1 \leq j \leq \alpha_{r+p}, \quad 0 \leq p \leq N - r.$$

When $N = 1$ (hence $r = 1, p = 0$), we have

$$((\widetilde{\mathcal{X}}\mathcal{Y})_{11})_{1j} = (\widetilde{\mathcal{X}}_{11})_{(1)}((\mathcal{Y})_{11})^{(j)}, \tag{21}$$

while for $N \geq 2$ we obtain

$$((\widetilde{\mathcal{X}}\mathcal{Y})_{1(1+p)})_{1j} = (\widetilde{\mathcal{X}}_{11})_{(1)}(\mathcal{Y}_{1(1+p)})^{(j)} + \sum_{k=2}^N (\widetilde{\mathcal{X}}_{1k})_{(1)}(\mathcal{Y}_{k(1+p)})^{(j)} \quad (r = 1), \tag{22}$$

$$((\widetilde{\mathcal{X}}\mathcal{Y})_{r(r+p)})_{1j} = (\widetilde{\mathcal{X}}_{rr})_{(1)}(\mathcal{Y}_{r(r+p)})^{(j)} + \sum_{k=r+1}^{N}(\widetilde{\mathcal{X}}_{rk})_{(1)}(\mathcal{Y}_{k(r+p)})^{(j)}$$

$$+ \sum_{k=1}^{r-1}(\widetilde{\mathcal{X}}_{rk})_{(1)}(\mathcal{Y}_{k(r+p)})^{(j)}, \quad 1 < r < N, \tag{23}$$

$$((\widetilde{\mathcal{X}}\mathcal{Y})_{NN})_{1j} = (\widetilde{\mathcal{X}}_{NN})_{(1)}(\mathcal{Y}_{NN})^{(j)} + \sum_{k=1}^{N-1}(\widetilde{\mathcal{X}}_{Nk})_{(1)}(\mathcal{Y}_{kN})^{(j)} \quad (r = N). \tag{24}$$

For any $r, k \in \{1, \dots, N\}$, we get

$$E_{a_{rk}}(I_{m_k})\big(T(A_0, A_1, \dots, A_{b_{rk}-1})\big)^T E_{a_{rk}}(I_{m_r}) = T(A_0^T, A_1^T, \dots, A_{b_{rk}-1}^T), \tag{25}$$

and it further implies

$$\widetilde{\mathcal{X}}_{rk} = E_{\alpha_r}(I_{m_r})\mathcal{X}_{kr}^T E_{\alpha_k}(I_{m_k}) = \begin{cases} \begin{bmatrix} \widetilde{\mathcal{T}}_{rk} \\ 0 \end{bmatrix}, & \alpha_r > \alpha_k, \\ \begin{bmatrix} 0 & \widetilde{\mathcal{T}}_{rk} \end{bmatrix}, & \alpha_r < \alpha_k, \\ \widetilde{\mathcal{T}}_{rk}, & \alpha_r = \alpha_k, \end{cases} \quad \widetilde{\mathcal{T}}_{rk} = T\big((A_0^{kr})^T, \dots, (A_{b_{kr}}^{kr})^T\big),$$

$$(\widetilde{\mathcal{X}}_{rk})_{(1)} = \begin{cases} \begin{bmatrix} (A_0^{kr})^T & (A_1^{kr})^T & \dots & (A_{a_k-1}^{kr})^T \end{bmatrix}, & \alpha_k \le \alpha_r, \\ \begin{bmatrix} 0 & \dots & 0 & (A_0^{kr})^T & \dots & (A_{\alpha_r-1}^{kr})^T \end{bmatrix}, & \alpha_k > \alpha_r. \end{cases} \tag{26}$$

We define $\Phi_n^{ks} := \sum_{j=0}^n B_{n-j}^k A_j^{ks}$ and observe that

$$\mathcal{Y}_{ks} = T\big(B_0^k, B_1^k, \dots, B_{\alpha_k-1}^k\big) \begin{cases} \begin{bmatrix} \mathcal{T}_{ks} \\ 0 \end{bmatrix}, & \alpha_k > \alpha_s \\ \begin{bmatrix} 0 & \mathcal{T}_{ks} \end{bmatrix}, & \alpha_k < \alpha_s \\ \mathcal{T}_{ks}, & \alpha_k = \alpha_s \end{cases} = \begin{cases} \begin{bmatrix} \mathcal{S}_{ks} \\ 0 \end{bmatrix}, & \alpha_k > \alpha_s, \\ \begin{bmatrix} 0 & \mathcal{S}_{ks} \end{bmatrix}, & \alpha_k < \alpha_s, \\ \mathcal{S}_{ks}, & \alpha_k = \alpha_s, \end{cases}$$

$$\mathcal{S}_{ks} = T\big(B_0^k, B_1^k, \dots, B_{b_{ks}-1}^k\big) T\big(A_0^{ks}, A_1^{ks}, \dots, A_{b_{ks}-1}^{ks}\big) = T\big(\Phi_0^{ks}, \Phi_1^{ks}, \dots, \Phi_{b_{ks}-1}^{ks}\big). \tag{27}$$

We begin with the calculation of matrices A_0^{rr} for $r \in \{1, \dots, N\}$. Since

$$(\widetilde{\mathcal{X}}_{rk})_{(1)} = \begin{cases} \begin{bmatrix} (A_0^{kr})^T & * & \dots & * \end{bmatrix}, & k \ge r, \\ \begin{bmatrix} 0 & * & \dots & * \end{bmatrix}, & k < r, \end{cases} \qquad (\mathcal{Y}_{kr})^{(1)} = \begin{cases} \begin{bmatrix} B_0^k A_0^{kr} \\ 0 \\ \vdots \\ 0 \end{bmatrix}, & k \le r, \\ 0, & k > r, \end{cases}$$

we deduce from (20) for $r = s, j = 1$ that

$$((\widetilde{\mathcal{X}}\mathcal{Y})_{rr})_{11} = \sum_{k=1}^{N}(\widetilde{\mathcal{X}}_{rk})_{(1)}((\mathcal{Y})_{kr})^{(1)} = (\widetilde{\mathcal{X}}_{rr})_{(1)}(\mathcal{Y}_{rr})^{(1)} = (A_0^{rr})^T B_0^r A_0^{rr}, \quad r \in \{1, \dots, N\}.$$

Together with $(\mathcal{C}_{rr})_{11} = C_0^r$ Equation (18) yields an equation that gives A_0^{rr}:

$$C_0^r = (A_0^{rr})^T B_0^r A_0^{rr}, \quad r \in \{1, \ldots, N\}; \tag{28}$$

it adds $\sum_{r=1}^{N} \frac{m_r(m_r-1)}{2}$ to the dimension of the solution. Next, if $N \geq 2$, we fix arbitrarily the blocks below the main diagonal of $[\mathcal{X}_{rs}]_{r,s=1}^{N}$ (the blocks above the main diagonal of $[\widetilde{\mathcal{X}}_{rs}]_{r,s=1}^{N}$), with $\sum_{r=1}^{N} \sum_{s=1}^{r-1} \alpha_r m_r m_s$ entries altogether. This corresponds to (a).

Proceed with the key step in the proof: an inductive procedure that enables us to compute the remaining entries (the algorithm in (b)). We fix $r \in \{1, \ldots, N\}, p \in \{0, \ldots, N-r\}$ and $j \leq \alpha_r - 1$, but not $p = j = 0$. Assuming that we have already determined the matrices $A_{\widetilde{j}}^{r's'}$ (with $1 \leq r', s' \leq N$) for

$$j \geq 1, \ \ \widetilde{j} \in \{0, \ldots, j-1\}, \ \ s' \geq r' \quad \text{or} \quad p \geq 1, \ \ \widetilde{j} = j, \ \ r' \leq s' \leq r' + p - 1$$

$$\text{or} \quad s' \leq r', \ \ \widetilde{j} \in \{0, \ldots, b_{r's'} - 1\}, \quad N \geq 2, \tag{29}$$

we shall compute $A_j^{r(r+p)}$ (step j, p of the algorithm in (b)). Essentially, we shall solve the equation $(\mathcal{C}_{r(r+p)})_{1j} = ((\widetilde{\mathcal{X}}\mathcal{Y})_{r(r+p)})_{1j}$ (see (18)). By a careful analysis of the structures of $(\widetilde{\mathcal{X}}_{rk})_{(1)}$ and $(\mathcal{Y}_{k(r+p)})^{(j)}$ in formulas (22), (23), and (24), we shall reduce this equation to a simple linear matrix equation in $A_j^{r(r+p)}$ (and possibly in $(A_j^{r(r+p)})^T$) with coefficients depending only on $A_{\widetilde{j}}^{r's'}$ for (29).

For the sake of clarity, we set the notation ($n \in \mathbb{Z}, k, r, s \in \{1, \ldots, N\}$):

$$\Psi_n^{krs} := \begin{cases} \begin{bmatrix} (A_0^{kr})^T & (A_1^{kr})^T & \cdots & (A_n^{rr})^T \end{bmatrix} \begin{bmatrix} \Phi_n^{ks} \\ \vdots \\ \Phi_0^{ks} \end{bmatrix}, & n \geq 0 \\ 0, & n < 0 \end{cases} = \begin{cases} \sum_{j=0}^{n} (A_j^{kr})^T \Phi_{n-j}^{ks}, & n \geq 0, \\ 0, & n < 0. \end{cases} \tag{30}$$

Note that

$$\Psi_n^{krs} = \sum_{j=0}^{n} (\Phi_j^{kr})^T A_{n-j}^{ks} = \sum_{j=0}^{n} \sum_{l=0}^{j} (A_l^{kr})^T (B_{j-l}^r)^T A_{n-j}^{ks} = \sum_{l=0}^{n} \sum_{j=l}^{n} (A_l^{kr})^T (B_{j-l}^r)^T A_{n-j}^{ks}$$

$$= \sum_{l=0}^{n} \sum_{j'=0}^{n-l} (A_l^{kr})^T (B_{j'}^r)^T A_{n-l-j'}^{ks} = \sum_{l=0}^{n} (A_l^{kr})^T \Phi_{n-l}^{ks} = (\Psi_n^{ksr})^T, \quad n \geq 0. \tag{31}$$

Since $(\widetilde{\mathcal{X}}_{rr})_{(1)} = \begin{bmatrix} (A_0^{rr})^T & \cdots & (A_{\alpha_r-1}^{rr})^T \end{bmatrix}$ and

$$(\mathcal{Y}_{rr})^{(\alpha_r-1)} = \begin{bmatrix} \Phi_{\alpha_r-1}^{rr} \\ \vdots \\ \Phi_0^{rr} \end{bmatrix}, \quad (\mathcal{Y}_{r(r+p)})^{(j+1)} = \begin{bmatrix} \Phi_j^{r(r+p)} \\ \vdots \\ \Phi_0^{r(r+p)} \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad j < \alpha_s - 1 \quad \text{or} \ p \geq 1,$$

the first term of (21), (22), (23), and (24) is equal to

$$(\widetilde{\mathcal{X}}_{rr})_{(1)}(\mathcal{Y})_{rr}^{(j+1)} = \Psi_j^{rrr} = (A_0^{rr})^T B_0^r A_j^{rr} + (A_j^{rr})^T B_0^r A_0^{rr} + \Xi(j,r,0), \quad (p=0)$$

$$(\widetilde{\mathcal{X}}_{rr})_{(1)}(\mathcal{Y})_{r(r+p)}^{(j+1)} = \Psi_j^{rr(r+p)} = (A_0^{rr})^T B_0^r A_j^{r(r+p)} + \Xi(j,r,p), \qquad p \geq 1,$$

$$\Xi(j,r,p) := \begin{cases} \sum_{l=1}^{j-1} A_l^{rr} \Phi_{j-l}^{rr}, & j \geq 1, \ p = 0, \\ \sum_{l=1}^{j} A_l^{r} \Phi_{j-l}^{r(r+p)}, & j \geq 0, \ p \geq 1. \end{cases} \tag{32}$$

(For simplicity, we have defined $\sum_{l=1}^{n} a_l = 0$ for $n < l$.)

When $N \geq 2$ the second term in (22) and (23) for $j+1$ instead of j consists of summands $(\widetilde{\mathcal{X}}_{rk})_{(1)}(\mathcal{Y}_{k(r+p)})^{(j+1)}$ with $k \geq r+1$, and such that $(\widetilde{\mathcal{X}}_{rk})_{(1)} = \left[(A_0^{kr})^T \quad \cdots \quad (A_{\alpha_r-1}^{kr})^T \right]$,

$$(\mathcal{Y}_{(r+p)(r+p)})^{(j+1)} = \begin{bmatrix} \Phi_j^{r(r+p)} \\ \vdots \\ \Phi_0^{r(r+p)} \end{bmatrix}, \quad (\mathcal{Y}_{k(r+p)})^{(j+1)} = \begin{cases} \begin{bmatrix} \Phi_j^{r(r+p)} \\ \vdots \\ \Phi_0^{r(r+p)} \\ 0 \\ \vdots \\ 0 \end{bmatrix}, & r+p > k, \\ \begin{bmatrix} \Phi_{j-\alpha_{r+p}-\alpha_k}^{r(r+p)} \\ \vdots \\ \Phi_0^{r(r+p)} \\ 0 \\ \vdots \\ 0 \end{bmatrix}, & k > r+p. \end{cases}$$

Hence for $N \geq r+1 \geq 2$:

$$\Theta(j,r,p) := \sum_{k=r+1}^{N} (\widetilde{\mathcal{X}}_{rk})_{(1)}(\mathcal{Y}_{k(r+p)})^{(j+1)}$$

$$= \begin{cases} \sum_{k=r+1}^{N} \Psi_{j-\alpha_r+\alpha_k}^{krr}, & j \geq 1, \quad p = 0, \\ \sum_{k=r+1}^{r+p} \Psi_j^{kr(r+p)} + \sum_{k=r+p+1}^{N} \Psi_{j-\alpha_{r+p}+\alpha_k}^{kr(r+p)}, & j \geq 0, \quad p \geq 1. \end{cases} \tag{33}$$

(For simplicity, we defined $\sum_{k=r+p+1}^{N} \Psi_{j-\alpha_{r+p}-\alpha_k}^{r(r+p)} = 0$ for $r+p+1 > N$.)

Finally, the third term in (23) and the second term in (24) (with $N \geq 2$) consist of summands which are products of matrices

$$(\widetilde{\mathcal{X}}_{rk})_{(1)} = \begin{bmatrix} 0 & \cdots & 0 & (A_0^{kr})^T & \cdots & (A_{b_{kr}}^{kr})^T \end{bmatrix},$$

$$(\mathcal{Y}_{k(r+p)})^{(j+1)} = \begin{bmatrix} \Phi_j^{k(r+p)} \\ \cdots \\ \Phi_0^{k(r+p)} \\ 0 \\ \cdots \\ 0 \end{bmatrix}, \quad 1 \le k \le r-1,$$

hence

$$\Lambda(j,r,p) := \sum_{k=1}^{r-1} (\widetilde{\mathcal{X}}_{rk})_{(1)} (\mathcal{Y}_{k(r+p)})^{(j+1)} = \sum_{k=1}^{r-1} \Psi_{j-\alpha_k+\alpha_r}^{kr(r+p)}. \tag{34}$$

We extend Ξ, Θ and Λ by 0:

$$\widetilde{\Xi}(j,r,p) = \begin{cases} \Xi(j,r,p), & j \ge 2, p \ge 0, \\ 0, & \text{otherwise}, \end{cases} \qquad \widetilde{\Theta}(j,r,p) = \begin{cases} \Theta(j,r,p), & N \ge r+1 \ge 2, \\ 0, & \text{otherwise} \end{cases}$$

$$\widetilde{\Lambda}(j,r,p) = \begin{cases} \Lambda(j,r,p), & N \ge r \ge 2, \\ 0, & \text{otherwise}, \end{cases}$$

and define

$$D_j^{r(r+p)} := \widetilde{\Xi}(j,r,p) + \widetilde{\Theta}(j,r,p) + \widetilde{\Lambda}(j,r,p). \tag{35}$$

The equation $(\mathcal{C}_{r(r+p)})_{1j} = ((\widetilde{\mathcal{X}}\mathcal{Y})_{r(r+p)})_{1j}$ combined with (21), (22), (23), (24) and with (32), (33), (34), (35) yields

$$\begin{aligned} (A_0^{rr})^T B_0^r A_j^{r(r+p)} &= -D_j^{r(r+p)}, \quad p \ge 1, \\ (A_0^{rr})^T B_0^r A_j^{rr} + (A_j^{rr})^T B_0^r A_0^{rr} &= C_j^r - D_j^{rr}, \quad p = 0. \end{aligned} \tag{36}$$

Moreover, from (31) it follows that Ψ_n^{krs} and $r = s$ are symmetric, thus $\Xi(j,r,0)$, $\Theta(j,r,0)$, $\Lambda(j,r,0)$ (and hence $C_j^r - D_j^{rr}$) are symmetric, too.

To get $A_j^{r(r+p)}$ for $p \ge 1$, we solve a simple equation of the form $A^T X = B$ with given nonsingular A and arbitrary B, while to get A_j^{rr} we solve the equation of the form $A^T X + X^T A = B$ with known nonsingular A and symmetric B; the solution is $X = \frac{1}{2}(A^T)^{-1}B + (A^T)^{-1}Z$ with Z skew-symmetric. We have $A = (A_0^{rr})^T B_0^r$ with $(A^T)^{-1} = ((A_0^{rr})^T B_0^r)^{-1} = A_0^r(C_0^r)^{-1}$ (see (28)) and $B = C_j^r - D_j^{rr}$ (for $p = 0$) or $B = -D_j^{r(r+p)}$ for $p \ge 1$. By recalling the definition of $D_j^{r(r+p)}$ in (35), we deduce the algorithm in (b).

Furthermore, $\Xi(j,r,p)$, $\Theta(j,r,p)$, $\Lambda(j,r,p)$ (thus also $D_j^{r(r+p)}$ and $A_j^{r(r+p)}$) depend on the entries of $A_{\tilde{j}}^{r's'}$ with (29). It is thus straightforward to see that the algorithm in (b) allows us to compute each entry from the entries that are already known.

If B_0^r, G_0^r are real, then by Sylvester's theorem Equation (28) has a real solution A_0^{rr} precisely when B_0^r, G_0^r are of the same inertia. The equivalence at the end of the lemma is then apparent. ∎

Remark 4.1: (1) The equation in (a) is of the form $C = X^T BX$ with given nonsingular symmetric matrices B, C. By the Autonne-Takagi factorization (see, e.g. [14, Corollary

4.4.4]) $B = R^T IR$, $C = S^T IS$ for some nonsingular R, S and the identity matrix I. The above equation thus reduces to $I = Y^T Y$ with $Y = RXS^{-1}$. When B and C are real with the same inertia matrix $\widetilde{I}$, i.e. $B = R^T \widetilde{I} R$ and $C = S^T \widetilde{I} S$ for some real orthogonal R and S, we get $\widetilde{I} = Y^T \widetilde{I} Y$ with $Y = RXS^{-1}$ (real pseudo-orthogonal).

(2) One could consider the Equation (18) even when the diagonal blocks of $\mathcal{B}, \mathcal{C}$ are non-singular. In this more general setting, the equation $C = A^T BA$ is more involved, while the solution of the equation $A^T X + X^T A = B$ is known (see [17]).

Example 4.1: We solve (18) for $\mathcal{F} = E_4(I) \oplus E_2(I) \oplus I$, $\mathcal{B} = \mathcal{C} = \mathcal{I} := I_4(I) \oplus I_2(I) \oplus I$. Set

$$\mathcal{Y} = \left[\begin{array}{cccc|ccc} A_1 & B_1 & C_1 & D_1 & H_1 & G_1 & J_1 \\ 0 & A_1 & B_1 & C_1 & 0 & H_1 & 0 \\ 0 & 0 & A_1 & B_1 & 0 & 0 & 0 \\ 0 & 0 & 0 & A_1 & 0 & 0 & 0 \\ \hline 0 & 0 & N_1 & P_1 & A_3 & B_3 & J_3 \\ 0 & 0 & 0 & N_1 & 0 & A_3 & 0 \\ 0 & 0 & 0 & R_1 & 0 & R_3 & A_4 \end{array}\right].$$

We compute

$$\mathcal{F}\mathcal{Y}^T\mathcal{F}\mathcal{Y} = \widetilde{\mathcal{Y}}\mathcal{Y} = \left[\begin{array}{cccc|ccc} A_1^T & B_1^T & C_1^T & D_1^T & N_1^T & P_1^T & R_1^T \\ 0 & A_1^T & B_1^T & C_1^T & 0 & N_1^T & 0 \\ 0 & 0 & A_1^T & B_1^T & 0 & 0 & 0 \\ 0 & 0 & 0 & A_1^T & 0 & 0 & 0 \\ \hline 0 & 0 & H_1^T & G_1^T & A_3^T & B_3^T & R_3^T \\ 0 & 0 & 0 & H_1^T & 0 & A_3 & 0 \\ 0 & 0 & 0 & J_1^T & 0 & J_3^T & A_4 \end{array}\right]$$

$$\times \left[\begin{array}{cccc|ccc} A_1 & B_1 & C_1 & D_1 & H_1 & G_1 & J_1 \\ 0 & A_1 & B_1 & C_1 & 0 & H_1 & 0 \\ 0 & 0 & A_1 & B_1 & 0 & 0 & 0 \\ 0 & 0 & 0 & A_1 & 0 & 0 & 0 \\ \hline 0 & 0 & N_1 & P_1 & A_3 & B_3 & J_3 \\ 0 & 0 & 0 & N_1 & 0 & A_3 & 0 \\ 0 & 0 & 0 & R_1 & 0 & R_3 & A_4 \end{array}\right]$$

$$= \left[\begin{array}{cccc} A_1^T A_1 & A_1^T B_1 + B_1^T A_1 & \begin{array}{c} A_1^T C_1 + C_1^T A_1 \\ +B_1^T B_1 + N_1^T N_1 \end{array} & * \\ & A_1^T A_1 & A_1^T B_1 + B_1^T A_1 & \begin{array}{c} A_1^T C_1 + C_1^T A_1 \\ +B_1^T B_1 + N_1^T N_1 \end{array} \\ & & A_1^T A_1 & \begin{array}{c} A_1^T B_1 + B_1^T A_1 \\ A_1^T A_1 \end{array} \\ \hline & & & \end{array}\right]$$

$$\left[\begin{array}{ccc|c}
A_1^T H_1 + N_1^T A_3 & \begin{array}{c} A_1^T G_1 + B_1^T H_1 + N_1^T B_3 \\ + P_1^T A_3 + R_1^T R_3 \end{array} & N_1^T J_3 + R_1^T A_4 \\ & & + A_1^T J_1 \\
0 & A_1^T H_1 + N_1^T A_3 & 0 \\
\\
0 & 0 & 0 \\
0 & 0 & 0 \\
\hline
A_3^T A_3 & A_3^T B_3 + B_3^T A_3 + R_3^T R_3 & A_3^T J_3 + R_3^T A_4 \\
& A_3^T A_3 & 0 \\
\hline
& & A_4^T A_4
\end{array} \right]$$

By comparing the diagonal of the diagonal blocks of the left-hand side and the right-hand side of $\widetilde{\mathcal{Y}}\mathcal{Y} = \mathcal{I}$, we deduce that $A_1, \ldots, A_4$ are any orthogonal matrices. Next, we choose N_1, P_1, R_1, R_3 arbitrarily. The diagonal blocks on the first upper diagonal yield equations $A_1^T H_1 + N_1^T A_3 = 0$ and $A_3^T J_3 + R_3^T A_4 = 0$, which further implies $H_1 = -A_1 N_1^T A_3$, $J_3 = -A_3 R_3^T A_4$; note that $(A_1^T)^{-1} = A_1$, $(A_3^T)^{-1} = A_3$. The last upper diagonal gives $N_1^T J_3 + A_1^T J_1 + R_1^T A_4 = 0$, thus $J_1 = A_1(N_1^T A_3 R_3^T A_4 - R_1^T A_4)$.

By inspecting the first upper diagonal of the main diagonal blocks in $\widetilde{\mathcal{Y}}\mathcal{Y} = \mathcal{I}$, we obtain $A_1^T B_1 + B_1^T A_1 = 0$ and $A_3^T B_3 + B_3^T A_3 + R_3^T R_3 = 0$, so we deduce B_1, B_3. Further, $A_1^T G_1 + B_1^T H_1 + N_1^T B_3 + P_1^T A_3 + R_1^T R_3 = 0$ (observe the first upper diagonal of the first upper diagonal) yields G_1.

The third and the fourth upper diagonal block of the first principal diagonal block give $A_1^T C_1 + C_1^T A_1 + B_1^T B_1 + N_1^T N_1 = 0$, $A_1^T D_1 + B_1^T C_1 + C_1^T B_1 + D_1^T A_1 + N_1^T P_1 + P_1^T N_1 + R_1^T R_1 = 0$ (see *), therefore C_1, D_1 follow, respectively.

The solutions of Equation (18) with a block diagonal matrix $\mathcal{C} = \mathcal{B}$ form a group with relatively simple generators. Recall that $\mathbb{U}$ is the set of matrices of form (9) with identities on the diagonals of the main diagonal blocks (Lemma 2.2).

Lemma 4.2: *The set $\mathbb{X}_{\mathcal{B}}$ of solutions of Equation (18) for $\mathcal{C} = \mathcal{B} = \oplus_{r=1}^N (\oplus_{j=1}^{\alpha_r} B_r)$ with $B_r \in GL_{m_r}(\mathbb{C}) \cap S_{m_r}(\mathbb{C})$ is a semidirect product $\mathbb{X}_{\mathcal{B}} = \mathbb{O}_{\mathcal{B}} \ltimes \mathbb{V}_{\mathcal{B}}$, in which the group $\mathbb{O}_{\mathcal{B}}$ consists of all matrices of the form $Q = \oplus_{r=1}^N (\oplus_{j=1}^{\alpha_r} Q_r)$ with $Q_r \in \mathbb{C}^{m_r \times m_r}$ such that $B_r = Q_r^T B_r Q_r$, and $\mathbb{V}_{\mathcal{B}} := \mathbb{U} \cap \mathbb{X}_{\mathcal{B}}$ (hence unipotent of order at most $\alpha_1 - 1$ and in nilpotency class at most α_1). Moreover, $\mathbb{V}_{\mathcal{B}}$ is generated by matrices of the following two forms:*

$$\mathcal{V} = \bigoplus_{r=1}^N T(I_{m_r}, V_1^r, \ldots, V_{\alpha_r-1}^r),$$

$$V_1^r := \frac{1}{2}(B_r)^{-1} Z_1^r, \quad V_{n+1}^r := \frac{1}{2}(B_r)^{-1} \left(Z_{n+1}^r - \sum_{j=1}^n (V_j^r)^T B_r V_{n-j+1}^r \right), \tag{37}$$

$$Z_n^r = -Z_n^r, \quad n \geq 1,$$

and

$$\mathcal{H}^k_{p,t}(F) = [(\mathcal{H}^k_{p,t}(F))_{rs}]^N_{r,s=1}, \quad p < t, \quad (\mathcal{H}^k_{p,t}(F))_{rs} = \begin{cases} [0 \quad \mathcal{U}_{rs}], & \alpha_r < \alpha_s, \\ \begin{bmatrix} \mathcal{U}_{rs} \\ 0 \end{bmatrix}, & \alpha_r > \alpha_s, \\ \mathcal{U}_{rs}, & \alpha_r = \alpha_s, \end{cases}$$

$$\mathcal{U}_{rs} = \begin{cases} \oplus^{\alpha_r}_{j=1} I_{m_r}, & r = s \\ 0, & r \neq s \end{cases}, \quad \{r,s\} \subset \{p,t\}, \quad \mathcal{U}_{rr} = T(I_{m_r}, A^r_1, \ldots, A^{rr}_{\alpha_r - 1}), \quad r \in \{p, t\},$$

$$A^{pp}_j = \begin{cases} a_{n-1} B^{-1}_p (F^T B_t F B^{-1}_p)^n B_p, & j = n(2k + \alpha - \beta) \\ 0, & \text{otherwise} \end{cases}, \quad a_n = -\frac{1}{2^{2n+1}} \frac{1}{n+1} \binom{2n}{n},$$

$$A^{tt}_j = \begin{cases} a_{n-1} B^{-1}_t (B_t F B^{-1}_p F^T)^n B_t, & j = n(2k + \alpha - \beta) \\ 0, & \text{otherwise} \end{cases}, \quad F \in \mathbb{C}^{m_p \times m_t},$$

$$\mathcal{U}_{pt} = N^k_{\alpha_t}(F), \quad \mathcal{U}_{tp} = N^k_{\alpha_t}(-B^{-1}_p F^T B_t).$$

$$(38)$$

Proof: For any $\mathcal{X}_1, \mathcal{X}_2 \in \mathbb{X}_\mathcal{B}$, we have

$$\mathcal{F}(\mathcal{X}_1 \mathcal{X}^{-1}_2)^T \mathcal{F} \mathcal{B}(\mathcal{X}_1 \mathcal{X}^{-1}_2) = \mathcal{F}(\mathcal{X}^{-1}_2)^T \mathcal{F} \mathcal{F} \mathcal{X}^T_1 \mathcal{F} \mathcal{B} \mathcal{X}_1 \mathcal{X}^{-1}_2 = \mathcal{F}(\mathcal{X}^{-1}_2)^T \mathcal{F} \mathcal{B} \mathcal{X}^{-1}_2$$
$$= \mathcal{F}(\mathcal{X}^{-1}_2)^T \mathcal{F} \mathcal{B}(\mathcal{B}^{-1} \mathcal{F} \mathcal{X}^T_2 \mathcal{F} \mathcal{B}) = \mathcal{B}.$$

Thus $\mathcal{X}_1 \mathcal{X}^{-1}_2 \in \mathbb{X}_\mathcal{B}$, so $\mathbb{X}_\mathcal{B}$ is a group.

We describe the structure of $\mathbb{X}_\mathcal{B}$. Lemma 4.1 for $\mathcal{C} = \mathcal{B} = \oplus^N_{r=1} (\oplus^{\alpha_r}_{j=1} B^r_0)$ implies that $\mathcal{X} \in \mathbb{X}_\mathcal{B}$ is of form (9) such that its main diagonal blocks have Q^r_0 on the diagonal (satisfying $B^r_0 = (Q^r_0)^T B^r_0 Q^r_0$). Therefore $\mathcal{X}$ can be written as $\mathcal{X} = \mathcal{Q} \mathcal{Y}$ with $\mathcal{Q} \in \mathbb{O}_\mathcal{B}$ and $\mathcal{Y} \in \mathbb{U}$. Clearly $\mathbb{O}_\mathcal{B} \subset \mathbb{X}_\mathcal{B}$ (hence $\mathcal{Y} \in \mathbb{X}_\mathcal{B}$), thus $\mathbb{X}_\mathcal{B} = \mathbb{O}_\mathcal{B} \ltimes \mathbb{V}_\mathcal{B}$, where $\mathbb{V}_\mathcal{B} = \mathbb{X}_\mathcal{B} \cap \mathbb{U}$. Since $\mathbb{V}_\mathcal{B}$ is a subgroup of $\mathbb{U}$, it is a normal subgroup in $\mathbb{X}_\mathcal{B}$, unipotent of order at most $\alpha_1 - 1$, and nilpotent of class at most α_1 (see Lemma 2.2).

Next, we find matrices in $\mathbb{X}_\mathcal{B}$ that are of a simple form. First, set

$$\mathcal{D}^k_{\alpha,\beta} = \begin{bmatrix} \Delta_{11} & \Delta_{12} \\ \Delta_{21} & \Delta_{22} \end{bmatrix}, \quad \alpha > \beta, \quad 0 \leq k \leq \beta - 1,$$

$$\Delta_{11} = T(I_{m_1}, A_1, \ldots, A_{\alpha-1}), \quad \Delta_{22} = T(I_{m_2}, D_1, \ldots, D_{\beta-1}), \quad (39)$$

$$\Delta_{21} = \begin{bmatrix} 0 & N^k_\beta(F) \end{bmatrix}, \quad \Delta_{12} = \begin{bmatrix} T(G_0, G_1, \ldots, G_{\beta-1}) \\ 0 \end{bmatrix},$$

in which $N^k_\beta(F)$ is a β−by−β block matrix with $F \in \mathbb{C}^{m_1 \times m_2}$ on the kth diagonal above the main diagonal and zeros otherwise, $A_j \in \mathbb{C}^{m_1 \times m_1}$, $D_j \in \mathbb{C}^{m_2 \times m_2}$, and $G_j \in \mathbb{C}^{m_2 \times m_1}$ for all j. Further, suppose $\mathcal{D}^k_{\alpha,\beta}$ is a solution of the matrix equation

$$\mathcal{B}_{\alpha,\beta} = \mathcal{F}_{\alpha,\beta}(\mathcal{D}^k_{\alpha,\beta})^T \mathcal{F}_{\alpha,\beta} \mathcal{B}_{\alpha,\beta} \mathcal{D}^k_{\alpha,\beta}, \quad \mathcal{B}_{\alpha,\beta} = I_\alpha(B) \oplus I_\beta(C),$$

$$\mathcal{F}_{\alpha,\beta} = E_\alpha(I_{m_1}) \oplus E_\beta(I_{m_2}),$$

where $B \in \mathbb{C}^{m_1 \times m_1}$ and $C \in \mathbb{C}^{m_2 \times m_2}$. Blockwise we have

$$I_\alpha(B) = T(I_{m_1}, A_1^T, \ldots, A_{\alpha-1}^T)I_\alpha(B)T(I_{m_1}, A_1, \ldots, A_{\alpha-1})$$
$$+ \begin{bmatrix} N_\beta^k(F^T) \\ 0 \end{bmatrix} I_\beta(C) \begin{bmatrix} 0 & N_\beta^k(F) \end{bmatrix}, \tag{40}$$

$$0 = T(I_{m_1}, A_1^T, \ldots, A_{\alpha-1}^T)I_\alpha(B) \begin{bmatrix} T(G_0, G_1, \ldots, G_{\beta-1}) \\ 0 \end{bmatrix}$$
$$+ \begin{bmatrix} N_\beta^k(F^T) \\ 0 \end{bmatrix} I_\beta(C)T(I_{m_2}, D_1, \ldots, D_{\beta-1}), \tag{41}$$

$$I_\beta(C) = T(I_{m_2}, D_1^T, \ldots, D_{\beta-1}^T)I_\beta(C)T(I_{m_2}, D_1, \ldots, D_{\beta-1})$$
$$+ \begin{bmatrix} 0 & T(G_0^T, G_1^T, \ldots, G_{\beta-1}^T) \end{bmatrix} I_\alpha(B) \begin{bmatrix} T(G_0, G_1, \ldots, G_{\beta-1}) \\ 0 \end{bmatrix}. \tag{42}$$

To determine $\mathcal{D}_{\alpha,\beta}^k$ we basically follow the algorithm in Lemma 4.1(b).

We first simplify the notation by defining $A_0 := I_{m_1}$ and $D := I_{m_2}$. By comparing the first row of the left-hand and the right-hand side of (40), we get

$$0 = \sum_{j=0}^{n} A_j^T B A_{n-j}, \quad 1 \le n \ne 2k + \alpha - \beta, \quad 0 = \sum_{j=0}^{2k+\alpha-\beta} A_j^T B A_{2k+\alpha-\beta-j} + F^T C F.$$

If $\alpha - \beta + k \ge 2$ we choose $A_1 = \ldots = A_{2k+\alpha-\beta-1} = 0$ to satisfy the first equation for $1 \le n \le 2k + \alpha - \beta - 1$. The second equation then yields $-F^T C F = A_{2k+\alpha-\beta}^T B + B A_{2k+\alpha-\beta}$ (in the case $\alpha - \beta + k = 1$ as well) and we take $A_{2k+\alpha-\beta} = -\frac{1}{2}B^{-1}F^T C F$. The first equation for $2k + \alpha - \beta + 1 \le n \le 2(2k + \alpha - \beta)$ further reduces to

$$0 = A_j^T B + B A_j^T, \quad 2k + \alpha - \beta + 1 \le j \le 2(2k + \alpha - \beta) - 1 \quad (\text{if } \alpha - \beta + k \ge 2),$$
$$0 = A_{2(2k+\alpha-\beta)}^T B + A_{2k+\alpha-\beta}^T B A_{2k+\alpha-\beta} + B A_{2(2k+\alpha-\beta)}.$$

Hence we can take $A_j = 0$ for $2k + \alpha - \beta + 1 \le j \le 2(2k + \alpha - \beta) - 1$ (if $\alpha - \beta + k \ge 2$) and $A_{2(2k+\alpha-\beta)} = -\frac{1}{8}(B^{-1}F^T C F)^2$. By continuing in this manner, we obtain

$$A_j = \begin{cases} a_{n-1}(B^{-1}F^T C F)^n, & j = n(2k + \alpha - \beta), \ n \in \mathbb{N}, \\ 0, & \text{otherwise,} \end{cases} \tag{43}$$

where $a_0 = -\frac{1}{2}$ and $a_n = -\frac{1}{2}\sum_{j=0}^{n-1} a_j a_{n-j-1}$ for $n \in \mathbb{N}$. The generating function associated with the sequence a_n is $f(t) := \sum_{j=0}^{\infty} a_j t^j$. Observe that $f(t) = -\frac{1}{2}t(f(t))^2 - \frac{1}{2}$, thus $f(t) = -\frac{1}{t}(1 + (1-t)^{\frac{1}{2}})$ and we obtain $a_n = -\frac{1}{2^{2n+1}}\frac{1}{n+1}\binom{2n}{n}$. For the basic theory of generating functions, see, e.g. [18, Chapter 2].

We now compare the entries in the first row of the left-hand and right-hand side of (41) and get the following equations:

$$0 = \sum_{j=0}^{n} A_j^T B G_{n-j}, \quad 0 \leq n \leq k-1 \quad (\text{if} k \geq 1),$$

$$0 = \sum_{j=0}^{k} A_j^T B G_{k-j} + F^T C, \tag{44}$$

$$0 = \sum_{j=0}^{n} A_j^T B G_{n-j} + F^T C D_{n-k}, \quad n \geq k+1,$$

The first two equations immediately imply

$$G_0 = \ldots = G_{k-1} = 0 \quad (\text{if} k \geq 1), \quad G_k = -B^{-1} F^T C. \tag{45}$$

By comparing the entries in the first row of the left-hand and right-hand side of (42), we obtain

$$0 = \sum_{j=0}^{n} D_j^T C D_{n-j} + \sum_{j=\alpha-\beta}^{n} G_{j-(\alpha-\beta)}^T B G_{n-1}, \quad n \geq 1. \tag{46}$$

Using (45) we deduce that the second summand on the right-hand side of (46) vanishes for $1 \leq n \leq \alpha - \beta + 2k - 1, \alpha - \beta + k \geq 2$, thus

$$0 = \sum_{j=0}^{n} D_j^T C D_{n-j}, \quad 1 \leq n \leq \alpha - \beta + 2k - 1 \quad (\text{if} \alpha - \beta + k \geq 2),$$

$$0 = \sum_{j=0}^{2k+\alpha-\beta} D_j^T C D_{2k+\alpha-\beta-j} + G_k^T B G_k.$$

Therefore, we choose

$$D_1 = \ldots = D_{2k+\alpha-\beta-1} = 0 \quad (\text{if} \alpha - \beta + k \geq 2),$$

$$D_{2k+\alpha-\beta} = -\frac{1}{2} C^{-1} G_k^T B G_k = -\frac{1}{2} F B^{-1} F^T C. \tag{47}$$

Using (43) and (47) for $\alpha - \beta + k \geq 2$, the last equation of (44) reduces to $0 = B G_n$ for $\alpha - \beta + 2k - 1 \geq n \geq k+1$; hence

$$G_{k+1} = \ldots = G_{2k+\alpha-\beta-1} = 0 \quad (\text{if } \alpha - \beta + k \geq 2). \tag{48}$$

Further, we apply (43), (45), (47), (48) to the last equation of (44) for $n = \alpha - \beta + 2k$. If $k \geq 1$ we obtain $B G_{2k+\alpha-\beta} = 0$, while for $\alpha - \beta \geq 2, k = 0$ we get

$$0 = B G_{\alpha-\beta} + A_{\alpha-\beta}^T B G_0 + F^T C D_{\alpha-\beta} = B G_{\alpha-\beta} - \frac{1}{2} F^T C F G_0 - \frac{1}{2} F^T G_0^T B G_0 = B G_{\alpha-\beta}.$$

Similarly for $k = 0, \alpha - \beta = 1$ we deduce $B G_1 = 0$. In any case, we have

$$G_{2k+\alpha-\beta} = 0. \tag{49}$$

If $\alpha - \beta + k \geq 2$ we use (45), (47), (48) and (49) to see that the second summand on the right-hand side of (46) for $\alpha - \beta + 2k + 1 \leq n \leq 2(\alpha - \beta) + 3k$ vanishes, while the first summand is equal to $D_n^T C + C D_n^T$; thus,

$$0 = D_n^T C + C D_n^T, \quad \alpha - \beta + 2k + 1 \leq n \leq 2(\alpha - \beta) + 3k.$$

We take

$$D_n = 0, \quad \alpha - \beta + 2k + 1 \leq n \leq 2(\alpha - \beta) + 3k \quad (\text{if } \alpha - \beta + k \geq 2). \tag{50}$$

Using (43), (47), (50), the third equation of (44) for $\alpha - \beta + 2k + 1 \leq n \leq 2(\alpha - \beta) + 3k$ reduces to $BG_n = 0$; it is clear for $n \neq \alpha - \beta + 3k$, while for $n = \alpha - \beta + 3k$

$$0 = BG_{\alpha-\beta+3k} + A_{\alpha-\beta+2k}^T BG_k + F^T C D_{\alpha-\beta+2k},$$
$$0 = BG_{\alpha-\beta+3k} + \tfrac{1}{2} F^T C F B^{-1} F^T C - \tfrac{1}{2} F^T C F B^{-1} F^T C = BG_{\alpha-\beta+3k}. \tag{51}$$

It yields

$$G_n = 0, \quad \alpha - \beta + 2k + 1 \leq n \leq 2(\alpha - \beta) + 3k. \tag{52}$$

Equation (46) for (45),(47),(48),(52),(50) then gives

$$0 = C D_j + D_j^T C, \quad 2(\alpha - \beta) + 3k + 1 \leq j \leq 2(\alpha - \beta + 2k) - 1,$$
$$0 = C D_{2(\alpha-\beta+2k)} + D_{\alpha-\beta+2k}^T C D_{\alpha-\beta+2k} + D_{2(\alpha-\beta+2k)}^T C.$$

We take

$$D_n = 0, \quad 2(\alpha - \beta) + 3k + 1 \leq n \leq 2(\alpha - \beta + 2k) - 1,$$
$$D_{2(\alpha-\beta+2k)} = -\frac{1}{2} D_{\alpha-\beta+2k}^T C D_{\alpha-\beta+2k} = -\frac{1}{8}(C^{-1} G_k^T BG_k)^2. \tag{53}$$

From (44) for (43), (47), (50), we further deduce

$$G_n = 0, \quad 2(\alpha - \beta) + 3k + 1 \leq n \leq 2(\alpha - \beta + 2k). \tag{54}$$

If $\alpha - \beta = 1$, $k = 0$ then (46) yields $0 = C D_2 + (D_1)^T C D_1 + D_2^T C$ and we choose $D_2 = -\frac{1}{8}(C^{-1} G_0^T BG_0)^2$. Further, similarly as in (51), we apply (44) to get $G_2 = 0$ from $0 = BG_2 + A_1^T BG_0 + F^T C D_1$. Thus (53) and (54) are valid in this case as well.

By continuing this process, we eventually obtain

$$G_j = \begin{cases} B^{-1} F^T C, & j = k, \\ 0, & \text{otherwise,} \end{cases}$$
$$D_j = \begin{cases} a_{n-1}(FB^{-1} F^T C)^n, & j = n(2k + \alpha - \beta), \\ 0, & \text{otherwise,} \end{cases} \quad a_n = -\frac{1}{2^{2n+1}} \frac{1}{n+1} \binom{2n}{n}. \tag{55}$$

We easily compute

$$(\mathcal{D}_{\alpha,\beta}^k(F))^{-1} = \mathcal{B}_{\alpha,\beta}^{-1} \mathcal{F}_{\alpha,\beta}(\mathcal{D}_{\alpha,\beta}^k(F)) \mathcal{F}_{\alpha,\beta} \mathcal{B}_{\alpha,\beta} = \begin{bmatrix} \Delta'_{11} & \Delta'_{12} \\ \Delta'_{21} & \Delta'_{22} \end{bmatrix}$$

with

$$\Delta'_{11} = T(I_m, A'_1, \ldots, A'_{\alpha-1}), \quad A'_j = \begin{cases} a_{n-1} B^{-1}(F^T C F B^{-1})^n B, & j = n(2k + \alpha - \beta), \\ 0, & \text{otherwise,} \end{cases}$$

$$\Delta'_{22} = T(I_n, D'_1, \ldots, D'_{\beta-1}), \quad D'_j = \begin{cases} a_{n-1}C^{-1}(CFB^{-1}F^T)^n C, & j = n(2k+\alpha-\beta), \\ 0, & \text{otherwise,} \end{cases}$$

$$\Delta'_{21} = \begin{bmatrix} 0 & N^k_\beta(F) \end{bmatrix}, \quad \Delta'_{12} = \begin{bmatrix} N^k_\beta(-B^{-1}F^TC) \\ 0 \end{bmatrix}, \quad a_n = -\frac{1}{2^{2n+1}}\frac{1}{n+1}\binom{2n}{n}.$$

Set $\mathcal{K}^k_{p,t}(F) \in \mathbb{V}_\mathcal{B}$ to be an N−by−N block matrix such that its principal submatrix formed by blocks in the pth and the tth columns and rows is equal to $\mathcal{D}^k_{\alpha_p,\alpha_t}(F)$, while the submatrix formed by all other blocks is the identity matrix. Clearly $\mathcal{H}^k_{p,t}(F) := (\mathcal{K}^k_{p,t}(F))^{-1}$ is of the same form as $\mathcal{K}^k_{p,t}(F)$, only with $(\mathcal{D}^k_{\alpha_p,\alpha_t}(F))^{-1}$ as a principal submatrix formed by blocks in the pth and the tth columns and rows.

We use the inductive procedure of multiplying $\mathcal{Y} \in \mathbb{V}_\mathcal{B}$ by matrices of the form $\mathcal{K}^k_{p,t}(F)$ for the appropriate p, t, k, F. To describe the inductive step, suppose that during the process we have a matrix that by a slight abuse of notation is still called $\mathcal{Y}$, and such that the blocks under the main diagonal in the first $p-1$ columns vanish (i.e. $\mathcal{Y}_{rs}$ vanishes for $p, r > s$), and the first $\alpha_p - \alpha_{p+1} + k$ columns of $\mathcal{Y}_{rp}$ for $r > p$ vanish. Let t be the largest index such that $(\mathcal{Y}_{tp})_{1(\alpha_p-\alpha_{p+1}+k+1)} \neq 0$, i.e. $(\mathcal{Y}_{tp})_{(1)} = \begin{bmatrix} 0 & \cdots & 0 & R^{tp}_{k-\alpha_{p+1}+\alpha_t} & \cdots & R^{tp}_{\alpha_t-1} \end{bmatrix}$ with all $R^{tp}_j \in \mathbb{C}^{m_t \times m_p}$ and $R^{tp}_{k-\alpha_{p+1}+\alpha_t} \neq 0$. We multiply $\mathcal{Y}$ with $\mathcal{K}^k_{p,t}(-R^{tp}_{k-\alpha_{p+1}+\alpha_t})$ to get $\mathcal{Y}'$ that is of the same form as $\mathcal{Y}$, and $(\mathcal{Y}'_{tp})^{(k+\alpha_t-\alpha_{p+1}+1)} = 0$. It is apparent for $\mathcal{Y}'_{rs}$ with $r, p > s$ or $t > r > s = p$, while for $r \geq t, s = p$ we have

$$(\mathcal{Y}'_{tp})_{(1)} = \begin{bmatrix} 0 & \cdots & 0 & R^{tp}_{k-\alpha_{p+1}+\alpha_t} & \cdots & R^{tp}_{\alpha_t-1} \end{bmatrix} T(I_{m_p}, A^{pp}_1, A^{pp}_1, \ldots, A^{pp}_{\alpha_p-1})$$
$$+ T(I_{m_t}, A^{tt}_1, \ldots, A^{tt}_{\alpha_t-1}) \begin{bmatrix} 0 & N^{k+\alpha_t-\alpha_{p+1}}_{\alpha_p-\alpha_t}(-R^{tp}_{k-(\alpha_{p+1}-\alpha_t)}) \end{bmatrix}$$
$$= \begin{bmatrix} 0 & \cdots & 0 & S^{tp}_{k-\alpha_{p+1}+\alpha_t+1} & \cdots & R^{tp}_{\alpha_t-1} \end{bmatrix},$$
$$(\mathcal{Y}'_{rp})_{(1)} = \begin{bmatrix} 0 & \cdots & 0 & R^{rp}_{k-\alpha_{p+1}+\alpha_t} & \cdots & R^{rp}_{\alpha_t-1} \end{bmatrix} T(I_{m_p}, \ldots, A^{pp}_{\alpha_p-1})$$
$$+ \begin{bmatrix} 0 & * & \cdots & * \end{bmatrix} \begin{bmatrix} 0 & N^{k+\alpha_t-\alpha_{p+1}}_{\alpha_p-\alpha_t}(-R^{tp}_{k-(\alpha_{p+1}-\alpha_t)}) \end{bmatrix}$$
$$= \begin{bmatrix} 0 & \cdots & 0 & S^{rp}_{k-\alpha_{p+1}+\alpha_t+1} & \cdots & R^{rp}_{\alpha_t-1} \end{bmatrix}, \quad r > t,$$

for some $S^{sp}_j \in \mathbb{C}^{m_s \times m_p}$ with $s \in \{r, t\}$. This process of choosing the appropriate $\{p_j, t_j, k_j, F_j\}^n_{j=1}$ eventually yields a block upper triangular matrix $\mathcal{V}$ of form (9), and such that the blocks on the main diagonal are block upper triangular Toeplitz with identities on the diagonals:

$$\mathcal{X} = \mathcal{Q}\mathcal{Y} = \mathcal{Q}\mathcal{V}\prod_{j=1}^n (\mathcal{K}^{k_j}_{p_j,t_j}(F_j))^{-1} = \mathcal{Q}\mathcal{V}\prod_{j=1}^n \mathcal{H}^{k_j}_{p_j,t_j}(F_j).$$

The inverse of a nonsingular block upper triangular matrix is again block upper triangular, hence $\mathcal{V}^{-1}$ is block upper triangular. On the other hand, $\mathcal{V}$ is a solution of Equation (18), so $\mathcal{V}^{-1} = \mathcal{B}^{-1}\mathcal{F}\mathcal{V}\mathcal{F}\mathcal{B}$ which is a block lower triangular matrix. Hence

$\mathcal{V} = \oplus_{r=1}^{N} T(I_{m_r}, V_1^r, \ldots, V_{\alpha_r-1}^r)$; the algorithm that provides the solution of (18) (see Lemma 4.1) yields equations that give (37):

$$(B_0^r)^T V_1^r + (V_1^r)^T B_0^r = 0,$$

$$(B_0^r)^T V_{n+1}^r + (V_{n+1}^r)^T B_0^r = -\sum_{j=1}^{n} (V_j^r)^T B_0^r V_{n+1-j}^r, \quad n \geq 1.$$

This concludes the proof of the lemma. ∎

5. Proof of Theorem 3.2

Proof of Theorem 3.2: Given a symmetric matrix A, we need to solve the equation

$$AQ = QA, \tag{56}$$

where Q is an orthogonal matrix and

$$A = \bigoplus_{r=1}^{N} \left(\bigoplus_{j=1}^{m_r} K_{\alpha_r}(\lambda) \right), \quad \lambda \in \mathbb{C}.$$

We shall first use Theorem 2.1 to solve (56) on Q. Taking into account that Q satisfies $Q^T Q = I$ (I is the identity matrix), it will yield a certain matrix equation and further restricting the form of Q; at this point, Lemmas 4.1 and 4.2 will be applied.

We have

$$A = P^{-1} J P, \quad J = \bigoplus_{r=1}^{N} \left(\bigoplus_{j=1}^{m_r} J_{\alpha_r}(\lambda) \right), \quad P = \bigoplus_{r=1}^{N} \left(\bigoplus_{j=1}^{m_r} P_{\alpha_j} \right),$$

where P_{α_j} is defined in (16). Equation (56) thus transforms to

$$JX = XJ, \quad X = PQP^{-1}.$$

From Theorem 2.1 (2.1), we obtain that $X = [X_{rs}]_{r,s=1}^{N}$, where X_{rs} is an m_r−by−m_s block matrix whose blocks of size $\alpha_r \times \alpha_s$ are of the form

$$\begin{cases} [0 \ \ T], & \alpha_r < \alpha_s, \\ \begin{bmatrix} T \\ 0 \end{bmatrix}, & \alpha_r > \alpha_s, \\ T, & \alpha_r = \alpha_s, \end{cases} \tag{57}$$

where $T \in \mathbb{C}^{m \times m}$, $m = \min\{\alpha_r, \alpha_s\}$ is a complex upper triangular Toeplitz matrix.

Since $P_\alpha = P_\alpha^T, P_\alpha^{-1} = \overline{P}_\alpha, P_\alpha^2 = -\overline{P'}_\alpha^2 = -P_\alpha^{-2} = iE_\alpha$, we deduce $P^2 = \overline{P}^2 = -P^{-2} = iE$, where $E := \oplus_{r=1}^N (\oplus_{j=1}^{m_r}(E_{\alpha_r}))$. Therefore, $I = Q^T Q$ is equivalent to

$$I = (P^T X^T (P^{-1})^T)(P^{-1} XP),$$

$$I = P(P^T X^T (P^{-1})^T)(P^{-1} XP)P^{-1},$$

$$I = P^2 X^T P^{-2} X, \tag{58}$$

$$I = iEX^T(-iE)X,$$

$$I = EX^T EX.$$

The trick of the proof is to transform this equation by conjugating the matrices with the permutation matrix $\Omega = \oplus_{r=1}^N \Omega_{\alpha_r,m_r}$ from (8):

$$I = (\Omega^T E\Omega)(\Omega^T X^T \Omega)(\Omega^T E\Omega)(\Omega^T X\Omega),$$

$$I = \mathcal{F}\mathcal{X}^T \mathcal{F}\mathcal{X}, \tag{59}$$

where $\mathcal{F} := \Omega^T E\Omega = \oplus_{j=1}^N E_{\alpha_r}(E_{m_r})$ and $\mathcal{X} := \Omega^T X\Omega$ (see (8)) are of form (9) with block rectangular upper triangular Toeplitz blocks.

We conclude the proof by applying Lemmas 4.1 and 4.2 to Equation (59). Lemma 4.1 for $\mathcal{B} = \mathcal{C} = I$ clearly implies that $\mathcal{X}$ satisfies the conditions (I), (II), while Lemma 4.2 for $\mathcal{B} = I$ yields that $\mathcal{X} \in \mathbb{O} \rtimes \mathbb{V}$; note that $\mathbb{V}$ is generated by matrices of forms (14) and (15) which coincide with (37) and (38) for $\mathcal{B} = I$, respectively. ∎

Remark 5.1: (1) The Equation (58) is very similar to the equation [6, Equation 2.12] which was partly studied by the author when proving the uniqueness of Hong's normal form under orthogonal *congruence. Using the solution of the Equation (18) (Lemma 4.1) and providing a somewhat more detailed analysis, the problem of computing the isotropy groups under orthogonal *congruence on Hermitian matrices will be addressed in the subsequent paper.

(2) Applying the same general approach as in this paper, the isotropy groups under orthogonal similarity on skew-symmetric or orthogonal matrices are described by matrix equations which involve an important difference in comparison to Equations (58) and (59). However, we expect that by developing some special techniques, similar results can be obtained in these cases as well.

(3) A referee of the paper has pointed out to the author that the isotropy groups under congruence $\{Q \in GL_n(\mathbb{Z}) \mid Q^T G_\Delta Q = G_\Delta\}$ at symmetric Gram matrices G_Δ of positive edge-bipartite graphs Δ with $n \geq 1$ vertices were observed in [19,20]. They are playing a key role in the Coxeter spectral analysis of edge-bipartite graphs and positive definite quasi-Cartan matrices recently studied by many authors, see [21–23].

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

References

[1] Hua LK. Geometries of matrices I. Generalizations of van Staudt's theorem. Trans Amer Math Soc. 1945;57:441–481.

[2] Hua LK. Geometries of matrices III. Fundamental theorems in the geometries of symmetric matrices. Trans Amer Math Soc. 1947;61:229–255.

[3] Hua LK. Geometries of symmetric matrices over any field with characteristic other than two. Ann Math. 1949;50:8–31.

[4] Gorbatsevich VV, Onischuk AL, Vinberg EB. Lie groups and lie algebras I. Berlin: Springer-Verlag; 1993. (Encyclopaedia of mathematical sciences; vol. 20).

[5] Milne JS. Algebraic groups: the theory of group schemes of finite type over a field. Cambridge: Cambridge University Press; 2017.

[6] Starčič T. Hong's canonical form of a Hermitian matrix with respect to orthogonal ∗congruence. Linear Algebra Appl. 2021;630:241–251.

[7] Hong YP, Horn RA, Johnson CR. On the reduction of pairs of Hermitian or symmetric matrices to diagonal form by congruence. Linear Algebra Appl. 1986;72:213–226.

[8] Kobayashi S, Nomizu K. Foundations of differential geometry, volume II. New York: Interscience Publishers; 1969.

[9] Weyl H. The classical groups, their invariants and representations. Princeton: Princeton University Press; 1946.

[10] Gantmacher FR. The theory of matrices. New York (NY): Chelsea Publishing Company; 1959.

[11] Lin WW, Mehrmann V, Xu H. Canonical forms for Hamiltonian and symplectic matrices and pencils. Linear Algebra Appl. 1999;302–303:469–533.

[12] Climenhaga V, Katok A. From groups to geometry and back. Providence: American Mathematical Society; 2017.

[13] Đoković DŽ, Zhao K. Tridiagonal normal forms for orthogonal similarity classes of symmetric matrices. Linear Algebra Appl. 2004;384:77–84.

[14] Horn RA, Johnson CR. Matrix analysis. Cambridge: Cambridge University Press; 1990.

[15] Arnold VI. Geometric methods in the theory of ordinary differential equations. New York (NY): Springer-Verlag; 1988.

[16] De Terán F, Dopico FM. The solution of the equation $XA + AX^T = 0$ and its application to the theory of orbits. Linear Algebra Appl. 2011;434:44–67.

[17] Braden HW. The equations $A^T X \pm X^T A = B$. SIAM J Matrix Anal Appl. 1998;20:295–302.

[18] Riordan J. An introduction to combinatorial analysis. Princeton: Princeton University Press; 2014.

[19] Simson D. A Coxeter-Gram classification of simply laced edge-bipartite graphs. SIAM J Discrete Math. 2013;27:827–854.

[20] Simson D. Algorithms determining matrix morsifications, Weyl orbits, Coxeter polynomials and mesh geometries of root orbits for the Dynkin diagrams. Fund Inform. 2013;123:447–490.

[21] Makuracki B, Mroz A. Quadratic algorithm to compute the Dynkin type of a positive definite Quasi-Cartan matrix. Math Comp. 2021;90:389–412.

[22] Perez C, Rivera D. Polynomial-time classification of skew symmetrizable matrices with a positive-definite quasi-Cartan companion. Fund Inform. 2021;181:313–337.

[23] Simson D. A computational technique in Coxeter spectral study of symmetrizable integer Cartan matrices. Linear Algebra Appl. 2020;586:190–238.