Pré-Publicações do Departamento de Matemática
Universidade de Coimbra
Preprint Number 16–04

CANONICAL FORMS FOR SIMULTANEOUS SIMILARITY OF PAIRS OF INVOLUTIONS

EDUARDO MARQUES DE SÁ

Abstract: In this paper we determine complete canonical forms for the action of simultaneous similarity on pairs of involutions. This is done in the context of linear operators of a finite dimensional vector space over an arbitrary field. In some pinpointed cases, the field is supposed to have characteristic $\neq 2$. The determination of a canonical form in the simultaneous similarity orbit of a pair (L, R) of involutions uses as ingredient the similarity class of LR, a conspicuous simultaneous similarity invariant; in the course of proof we get, for a given square matrix A, a detailed description of the set $\mathfrak{I}_A$ of involutions L such that $A = LR$ for some involution R; canonical forms are obtained for the action of G_A (the group of invertible matrices commuting with A) acting by similarity on the set $\mathfrak{I}_A$. The set of pairs (L, R) of involutions such that LR lies in a fixed similarity class is a union of a finite number of simultaneous similarity orbits; we determine that number, as well as the number of such orbits assuming the similarity classes of L, R and LR are kept fixed.

Keywords: involutions, simultaneous similarity, canonical forms.
AMS Subject Classification (2010): 15A21, 15A18, 15A23, 15B05.

1. Introduction

The general problem of classifying the orbits of m-tuples of $n \times n$ complex matrices under simultaneous similarity is solved in S. Friedland's renowned paper [9]; in that paper, a set of explicit invariants is given which characterize the orbits, but no canonical form is obtained. Another interesting work on simultaneous similarity is [6] where the more general concept of *PS-equivalence* is proposed, and a near canonical form is proved for the new equivalence. I wasn't able to use that near canonical form to shorten the argument producing the canonical forms in the particular case treated in this paper.

Received February 25, 2016.

This work was partially supported by the Centre for Mathematics of the University of Coimbra – UID/MAT/00324/2013, funded by the Portuguese Government through FCT/MEC and co-funded by the European Regional Development Fund through the Partnership Agreement PT2020.

The determination of canonical forms for the general problem of simultaneous similarity is considered hopeless in the sense of an effective determination of such forms. In the hierarchy of matrix problems, which goes back to the 1970's (as presented in, *e.g.*, [3]), the general simultaneous similarity problem serves as a milestone to the complexity of such problems; it is said to be *wild*, to express its high degree of complexity. The determination of canonical forms for that general problem has been the object of an interesting algorithm due to G.R. Belitskiĭ [2, 20], which, in a restricted sense to be discussed later on, may be viewed as an archetypal model to the methods used below.

If the simultaneous similarity action of GL_n is restricted to a subset of pairs (or k-tuples) of n-square matrices, we may have the chance of getting a feasible solution to the canonical form problem. For example, for pairs of matrices (A, B) such that $AB = BA = 0$, canonical forms have been given in [4]. On the other hand, I.M. Gelfand and V.A. Ponomarev proved that the subproblem with pairs (A, B) such that $AB = BA$ is equivalent to the general problem of simultaneous similarity for k-tuples of matrices [11].

In the present paper we consider pairs of involutions, and give canonical forms for such pairs. In sections 2-3, using a simple localization technique, we split the problem into several easier subproblems, studying pairs (L, R) of involutions such that LR has a characteristic polynomial of simple structure. In section 4 an easy solution is got in case LR has no eigenvalue ± 1; in this case we show GL_n acts transitively on the set of pairs such that LR lies in a fixed similarity orbit. In section 5, by far the lengthier, we treat the case when the characteristic polynomial of LR is a power of $x - 1$, or $x + 1$; then a block diagonal canonical form is obtained, each block being a pair of special involutions called *Pascal matrices*; we also give a canonical form of companion type. For LR in a fixed similarity class, the pairs (L, R) fill in a finite number of simultaneous similarity orbits; in section 6 we determine that number, as well as the number of such orbits assuming the similarity classes of L, R and LR are kept fixed.

Some of our concepts may be given in a semigroup $\mathcal{S}$ with a bilateral identity I. Suppose an element $A \in \mathcal{S}$ is the product of two involutions, say $A = LR$, with $L^2 = R^2 = I$. Then RL is the (bilateral) inverse of A, and the equalities $A^{-1} = LAL = RAR$ show that A is similar to its inverse by an involution

(L or R); conversely, if A is similar to its inverse by an involution, then A is the product of two involutions. All this is easy to check, but in our main object of study, the semigroup $\mathfrak{M}_n$ of the n-square matrices over a field, it is nontrivial and well-known that the similarity of A to A^{-1} implies similarity by an involution. Note that the left involution L in a factorization $A = LR$ also occurs as right factor in another factorization of A into two involutory factors, because LRL is an involution and $A = (LRL)L$.

The following sets will play an important role in the sequel:

$$\mathscr{P}_A = \{X : AXA = X\}, \qquad \mathscr{C}_A = \{X : AX = XA\},$$
$$\mathfrak{I}_A = \{\text{involutions of } \mathscr{P}_A\}, \qquad \mathrm{G}_A = \{\text{units of } \mathscr{C}_A\}.$$

Note that $\mathfrak{I}_A$ is the set of all involutions L $[R]$ which occur in a factorization $A = LR$. From such a factorization we may obtain other factorizations of the same kind: $A = (TLT^{-1})(TRT^{-1})$, for $T \in \mathrm{G}_A$. So G_A acts on $\mathfrak{I}_A$ by similarity. In our concrete setting, the purpose is

To determine canonical forms for the action of G_A *on* $\mathfrak{I}_A$.

As $\mathscr{P}_{TAT^{-1}} = T\mathscr{P}_A T^{-1}$, the displayed problem is closely related *to determining canonical forms for simultaneous similarity of pairs of involutions of* $\mathcal{S}$. A moments thought shows that a solution to one of these problems implies a solution to the other. The reader may enjoy proving the following trivial exercises, for any $A \in \mathcal{S}$:

1) $\mathscr{P}_A\mathscr{P}_A \subseteq \mathscr{C}_A$.
2) $\mathscr{C}_A\mathscr{P}_A = \mathscr{P}_A\mathscr{C}_A = \mathscr{P}_A$.
3) If $P \in \mathscr{P}_A$, where P and A are units, then $\mathscr{P}_A = P\mathscr{C}_A = \mathscr{C}_A P$.
4) For units of $\mathcal{S}$, we have $(X, Y) \approx (X', Y')$ iff $(XY, Y) \approx (X'Y', Y')$, where $\approx$ means simultaneous similarity.

The last exercise shows the way followed below to get a canonical form for a pair (L, R): apply the natural procedure to (LR, R), namely, reduce LR to a similarity normal form, call it N, and then act with G_N on the second term of the pair for a final reduction.

General notations. $\mathbb{F}[x]$ denotes the ring of polynomials over an arbitrary field $\mathbb{F}$, $\mathfrak{M}_n = \mathfrak{M}_n(\mathbb{F})$ the algebra of n-square matrices over $\mathbb{F}$, and GL_n the group of units of $\mathfrak{M}_n$; $\mathscr{S}_n$ is the group of permutations of $\{1, \ldots, n\}$; $|\Sigma|$ is the cardinality of a set Σ; $f[\Sigma]$ denotes $\{f(x) : x \in \Sigma\}$; I_n, or just I, is the $n \times n$ identity matrix; S_n is the $n \times n$ *skew-identity*, resulting from I_n by reversing the rows' order; χ_A denotes the characteristic polynomial of A; the transpose of A is denoted A^t; similarity of matrices and simultaneous similarity of matrix pairs will be denoted by $A \sim B$ and $(A_1, A_2) \approx (B_1, B_2)$, respectively.

2. Split forms

Let φ be a monic polynomial of positive degree. A pair $(G, F) \in \mathfrak{M}_n^2$ of the form $(G_1 \oplus G_2, F_1 \oplus F_2)$ is called a *φ-split form* if G_1 and F_1 are square matrices of the same order, φ is the characteristic polynomial of G_1F_1, and the characteristic polynomial of G_2F_2 has positive degree and is relatively prime with φ. We say that $(L, R) \in \mathfrak{M}_n^2$ *has a φ-splitting* if it is simultaneously similar to a φ-split form. A pair of the kind $(G_1 \oplus G_2, F_1 \oplus F_2)$ will also be denoted by

$$(G_1, F_1) \oplus (G_2, F_2).$$

Lemma 2.1. *Let $(G_1, F_1) \oplus (G_2, F_2)$ and $(\widetilde{G}_1, \widetilde{F}_1) \oplus (\widetilde{G}_2, \widetilde{F}_2)$ be two φ-split forms of (L, R). Then $(G_1, F_1) \approx (\widetilde{G}_1, \widetilde{F}_1)$ and $(G_2, F_2) \approx (\widetilde{G}_2, \widetilde{F}_2)$.*

Proof. There exists a nonsingular matrix T such that $T(G_1 \oplus G_2) = (\widetilde{G}_1 \oplus \widetilde{G}_2)T$ and $T(F_1 \oplus F_2) = (\widetilde{F}_1 \oplus \widetilde{F}_2)T$. Then $T(G_1F_1 \oplus G_2F_2) = (\widetilde{G}_1\widetilde{F}_1 \oplus \widetilde{G}_2\widetilde{F}_2)T$. Partition T accordingly: $T = (T_{ij})_{i,j=1,2}$. Then we have

$$T_{12}G_2F_2 = \widetilde{G}_1\widetilde{F}_1T_{12} \quad \text{and} \quad T_{21}G_1F_1 = \widetilde{G}_2\widetilde{F}_2T_{21}.$$

The characteristic polynomials of G_1F_1 and G_2F_2 are pairwise relatively prime, and are respectively equal to the characteristic polynomials of $\widetilde{G}_1\widetilde{F}_1$ and $\widetilde{G}_2\widetilde{F}_2$. By [10, pp. 215-ff], the displayed conditions imply $T_{12} = 0, T_{21} = 0$; therefore $T = T_{11} \oplus T_{22}$, and we get $\widetilde{F}_i = T_{ii}F_iT_{ii}^{-1}$ and $\widetilde{G}_i = T_{ii}G_iT_{ii}^{-1}$, for $i = 1, 2$, as required. □

While this lemma cares for uniqueness of the direct summands of a φ-splitting, the next one deals with existence.

Lemma 2.2. *Assume that $(L, R) \in \mathfrak{M}_n^2$ satisfies $LR = P_1 \oplus P_2$, where the P_1, P_2 are nonempty and have relatively prime characteristic polynomials φ_1, φ_2. Then (L, R) has a φ_1-splitting iff (L, R) is already a φ_1-split form.*

Proof. There exist L_1, R_1, L_2, R_2 of appropriate sizes, such that

$$(L, L^{-1}(P_1 \oplus P_2)) \approx (L_1, R_1) \oplus (L_2, R_2)),$$
$$L_1R_1 \sim P_1 \quad \text{and} \quad L_2R_2 \sim P_2.$$

Let T be a nonsingular matrix such that

$$L = T(L_1 \oplus L_2)T^{-1}, \text{and} \quad L^{-1}(P_1 \oplus P_2) = T(R_1 \oplus R_2)T^{-1}.$$

We get $(P_1 \oplus P_2)T = T(L_1R_1 \oplus L_2R_2)$, and the argument used in the proof of lemma 2.1 shows that $T = T_{11} \oplus T_{22}$. Therefore

$$(L, R) = (T_{11}L_1T_{11}^{-1}, T_{11}R_1T_{11}^{-1}) \oplus (T_{22}L_2T_{22}^{-1}, T_{22}R_2T_{22}^{-1}).$$

□

Of course a pair $(L, R) \in \mathfrak{M}_n^2$ may have no φ-splitting, for any φ. A trivial case occurs when LR has an irreducible characteristic polynomial. For a less trivial example, take for L a matrix with an irreducible pattern of zeroes, and let D be a diagonal matrix with simple spectrum. By lemma 2.2, $(L, L^{-1}D)$ has no φ-splitting, for any φ.

3. Products of two involutions

It is well-known that an invertible matrix $A \in \mathfrak{M}_n$ is the product of two involutions of $\mathfrak{M}_n$ iff A is similar to A^{-1}; and this holds iff any similarity invariant factor of A is self-reciprocal. These results have been proved and improved in several directions, *e.g.*, [21, 7, 12, 1]. The reader may also find in [21, 8, 14, 15] modified results of the same kind, in some cases worked out in the context of orthogonal and symplectic geometry. For the theory of invariant factors and elementary divisors check, *e.g.*, [10, 13].

For $f \in \mathbb{F}[x]$ of degree m such that $f(0) \neq 0$, define the *reciprocal* of f, denoted f^*, as follows: $f^*(x) = x^m f\left(1/x\right)/f(0)$ (check [22, p. 38]). Note that f^* is always a monic polynomial; and $f^{**} = f$ if f is monic. In case $f^* = f$, we say f is *self-reciprocal.* The reciprocal operation preserves products, $(fg)^* = f^*g^*$, and this entails: f is irreducible iff f^* is irreducible. So, if ψ is an irreducible factor of a self-reciprocal polynomial χ, and if ψ is not self-reciprocal, then ψ^* is also an irreducible factor of χ; moreover, ψ^k divides χ if and only if $\left(\psi^*\right)^k$ divides χ; so in the primary factorization of χ, ψ and ψ^* occur with the same power. (*Cf.* [18] for a generalization.) In the sequel we shall consider two kinds of companion matrices of a monic polynomial $f(x) = x^n - c_n x^{n-1} - \cdots - c_2 x - c_1$, namely

$$C(f) = \left[\begin{array}{c|cccc} & 1 & & & \\ & & 1 & & \\ & & & \ddots & \\ & & & & 1 \\ \hline c_1 & c_2 & c_3 & \cdots & c_n \end{array}\right] \quad \text{and} \quad \widehat{C}(f) = \left[\begin{array}{cccc|c} c_n & \cdots & c_3 & c_2 & c_1 \\ \hline 1 & & & & \\ & \ddots & & & \\ & & 1 & & \\ & & & 1 & \end{array}\right] \qquad (1)$$

where non-specified entries are 0. We now give a trimmed proof of the main result of [21, 7, 12].

Theorem 3.1. *For $A \in \mathfrak{M}_n$, the following are pairwise equivalent:*

(a) *A is the product of two involutions of $\mathfrak{M}_n$;*
(b) *A is similar to its inverse by an involution of $\mathfrak{M}_n$;*
(c) *A is similar to its inverse;*
(d) *The similarity invariant polynomials of A are self-reciprocal.*

Proof. $(a) \Rightarrow (b)$ and $(b) \Rightarrow (c)$ are obvious.

$(c) \Rightarrow (d)$. If $f_1|f_2|\ldots|f_s$ are the similarity invariant polynomials of A then $f_1^*|f_2^*|\ldots|f_s^*$ are the similarity invariant polynomials of A^{-1} (*cf.* [10, p. 153]). Therefore (c) implies $(f_1, \ldots, f_s) = (f_1^*, \ldots, f_s^*)$.

$(d) \Rightarrow (a)$. Clearly we only have to consider the case when A is non-derogatory. Without loss of generality we assume $A = C(f)$. Some simple computations show that $C(f)^{-1} = \widehat{C}(f^*)$ and $C(f) = S_n \widehat{C}(f) S_n$, where S_n is the $n \times n$ skew-identity. As S_n is an involution and we are assuming $f = f^*$, we get $C(f)$ as a product of two involutions (*cf.* [12]). □

A *palindrome* is a polynomial $g(x) = \sum_{i=0}^m c_i x^i$ such that $(c_0, \ldots, c_m) = (c_m, \ldots, c_0)$. It is easy to show that

Lemma 3.2. *With the exception of $x+1$ and $x-1$, all irreducible self-reciprocal polynomials are palindromes of even degrees.* □

The cyclotomic polynomials are remarkable examples of palindromes; they are irreducible over the field of rational numbers, and their irreducibility for other fields is a well-studied matter in Galois Theory [17].

Now suppose that $(L, R) \in \mathfrak{M}_n^2$ is a pair of involutions, and χ_A is the characteristic polynomial of $A = LR$. We may factorize χ_A as

$$\chi_A = p_1^{e_1} \cdots p_w^{e_w} (\pi_1^* \pi_1)^{\varepsilon_1} \cdots (\pi_s^* \pi_s)^{\varepsilon_s} \tag{2}$$

where the e_i, ε_j are positive integers, the p_i are the distinct self-reciprocal prime factors of χ_A, and the $2s$ polynomials π_j, π_j^* are the distinct non-self-reciprocal prime monic factors of χ_A. Now let φ_1 be a divisor of χ_A such that φ_1 and $\varphi_2 := \chi/\varphi_1$ are non-unit, relatively prime and self-reciprocal; then $A = W(A_1 \oplus A_2)W^{-1}$, for some $W \in \mathrm{GL}_n$, where $\chi_{A_i} = \varphi_i$. We may write $A_i = L_i R_i$, where L_i, R_i are involutions; we get a φ_1-splitting

$$(L, R) \approx (L_1, R_1) \oplus (L_2, R_2)$$

with the bonus attribute that the four factors on the right are involutions. Then $\mathfrak{I}_A$ and G_A also split accordingly. More precisely:

Lemma 3.3. $\mathfrak{I}_A = W(\mathfrak{I}_{A_1} \oplus \mathfrak{I}_{A_2})W^{-1}$ *and* $\mathrm{G}_A = W(\mathrm{G}_{A_1} \oplus \mathrm{G}_{A_2})W^{-1}$.

Proof. We may assume $W = I$. $M \in \mathfrak{I}_A$ iff $M^2 = I$ and $AMA = A$. Partition M as $\big(M_{ij}\big)_{i,j=1}^2$ according to the split form. Then we get $A_i M_{ij} A_j = M_{ij}$; as the characteristic polynomials of A_1 and A_2 are relatively prime and self-reciprocal, $M_{12} = 0$ and $M_{21} = 0$; the desired splitting of $\mathfrak{I}_A$ follows at once. The case of G_A may be done in the same manner. □

The previous arguments (with the help of lemma 2.1 for uniqueness) allow a simple induction to get

Theorem 3.4. *If $A = LR$, where L, R are involutions, then*

$$(L, R) \approx (L_1, R_1) \oplus \cdots \oplus (L_{w+s}, R_{w+s}), \tag{3}$$

where all terms L_k, R_k are involutions and $L_i R_i$ has characteristic polynomial equal to the i-th power of the factorization (2)*. Moreover, the direct summand pairs (L_i, R_i) are unique up to simultaneous similarity. If $W \in \mathrm{GL}_n$ provides the simultaneous similarity* (3)*, then we have $\mathfrak{I}_A = W\big(\bigoplus_i \mathfrak{I}_{L_i R_i}\big)W^{-1}$ and $\mathrm{G}_A = W\big(\bigoplus_i \mathrm{G}_{L_i R_i}\big)W^{-1}$.* □

This theorem suggests to divide the determination of orbits and canonical forms into cases according to the factorization of χ_A.

4. A has no eigenvalues ±1

In this section we prove the following theorem and then discuss some available canonical forms. In this section, n is necessarily an even number.

Theorem 4.1. *Let L and R be involutions such that 1 and -1 are not eigenvalues of $A = LR$. Then G_A acts transitively on $\mathfrak{I}_A$; that is, $(L, R) \approx (L', R')$ is equivalent to $LR \sim L'R'$, for any involutions L', R'. All elements of $\mathfrak{I}_A$ are similar to S_n, the n-th order skew identity.*

Proof. The argument is split into several cases.

Case 1: *χ_A has no self-reciprocal prime factors.* This corresponds to $w = 0$ in (2). It is easy to check that an elementary divisor canonical form of A may be organized to show that $A \sim B \oplus B^{-1}$, where B and B^{-1} have relatively prime characteristic polynomials. Choose any such B, and let $\Sigma = B \oplus B^{-1}$. The general forms of the elements in G_Σ and $\mathfrak{I}_\Sigma$ are

$$\begin{bmatrix} C & 0 \\ 0 & D \end{bmatrix} \in \mathrm{G}_\Sigma \quad \text{and} \quad \begin{bmatrix} 0 & E^{-1} \\ E & 0 \end{bmatrix} \in \mathfrak{I}_\Sigma,$$

with $C, D, E \in \mathrm{G}_B$. Thus the (similarity) action of G_Σ on $\mathfrak{I}_\Sigma$ is transitive, and so is the action of G_A on $\mathfrak{I}_A$. As $\left[\begin{smallmatrix} 0 & I \\ I & 0 \end{smallmatrix}\right]$ lies in $\mathfrak{I}_\Sigma$, and this matrix is permutationally similar to S_n, all elements of $\mathfrak{I}_A$ are similar to S_n.

Case 2: *χ_A is a power of a self-reciprocal monic prime p of degree > 1.* Denote by $\mathbb{K}$ the splitting field of p. Over $\mathbb{K}$, χ_A has no self-reciprocal prime factors. In the extended field we fall under the jurisdiction of Case 1; thus in the present case our theorem holds over $\mathbb{K}$. According to Theorem 1 of [16], if two pairs of $\mathbb{F}$-matrices are simultaneously similar over an extension $\mathbb{K}$ of $\mathbb{F}$, then they are simultaneously similar over $\mathbb{F}$ as well. So, in Case 2, our theorem holds over $\mathbb{F}$.

The general case follows from appropriate application of theorem 3.4. □

The transitivity assures that any pair of simple structure in the orbit of (L, R) may serve as canonical form. We go back to the proof of theorem 3.1, and denote by $\alpha_1 | \dots | \alpha_u$ the similarity invariant polynomials of LR. Each $C(\alpha_i)$ is the product of the two involutions, S_{a_i} and $S_{a_i} C(\alpha_i)$, where a_i denotes

$\deg \alpha_i$. This gives rise to the first proposed canonical form

$$\begin{aligned}(L,R) \approx & \big(S_{a_1} \oplus \cdots \oplus S_{a_u}, S_{a_1}C(\alpha_1) \oplus \cdots \oplus S_{a_u}C(\alpha_u)\big) = \\ & \big(S_{a_1}, S_{a_1}C(\alpha_1)\big) \oplus \cdots \oplus \big(S_{a_u}, S_{a_u}C(\alpha_u)\big) \end{aligned} \tag{4}$$

Using the factorization (2), split $A = LR$ as $A \sim A^{(1)} \oplus A^{(2)}$, where each elementary divisors of $A^{(1)}$ is a power of a self-reciprocal prime polynomial p_i, and each elementary divisor of $A^{(2)}$ is a power of a non-self-reciprocal prime polynomial in the set $\Pi = \{\pi_1, \pi_1^*, \ldots, \pi_s, \pi_s^*\}$. Then

$$(L,R) \approx (L^{(1)}, R^{(1)}) \oplus (L^{(2)}, R^{(2)}), \tag{5}$$

where $L^{(i)}, R^{(i)}$ are involutions, and $L^{(i)}R^{(i)} = A^{(i)}$. In (5) the $(L^{(i)}, R^{(i)})$ may be replaced by a canonical form of the kind (4); we may also replace $(L^{(1)}, R^{(1)})$ by a form as (4) with the α_i denoting the elementary divisors of $A^{(1)}$; and we may replace $(L^{(2)}, R^{(2)})$ by a canonical form of a different kind that we shall describe under (6).

Choose one element in each element of $\big\{\{\pi_1, \pi_1^*\}, \ldots, \{\pi_s, \pi_s^*\}\big\}$, and call $\varPhi$ the set of chosen elements. Clearly $\{\varPhi, \varPhi^*\}$ is a partition of Π. Let $B_\varPhi$ be any matrix whose elementary divisors are those of A which are powers of elements of $\varPhi$. Obviously $A^{(2)} \sim B_\varPhi \oplus B_{\varPhi^*} \sim B_\varPhi \oplus B_\varPhi^{-1}$. Let $\zeta_1, \ldots, \zeta_k$ be the elementary divisors of $B_\varPhi$. Then we get a canonical form

$$(L^{(2)}, R^{(2)}) \approx \left(\begin{bmatrix} 0 & I \\ I & 0 \end{bmatrix}, \begin{bmatrix} 0 & \bigoplus_{i=1}^k \widehat{C}(\zeta_i) \\ \bigoplus_{i=1}^k C(\zeta_i) & 0 \end{bmatrix} \right). \tag{6}$$

5. χ_A is a power of $x-1$, or $x+1$

So we are left with the prime factors $x \pm 1$, the exceptions of lemma 3.2. Firstly we treat the case when $\chi_A(x) = (x-1)^n$.

Let H_k be the $k \times k$ nilpotent Jordan matrix: $h_{ij} = 1$ if $j = i+1$, and $h_{ij} = 0$ otherwise. Take the Jordan $k \times k$ matrix with eigenvalue 1, $J_k := I_k - H_k$. If A has characteristic polynomial $(x-1)^n$, then A is similar to

$$M = J_{n_1} \oplus \cdots \oplus J_{n_w}, \tag{7}$$

where $(n_1, n_2, \ldots)$ is an integer partition of n, with w positive parts; we are assuming $n_1 \geqslant n_2 \geqslant \ldots$, and $n_1 + \cdots + n_w = n$. Let $\mathscr{C}_M := \{X \in \mathfrak{M}_n : XM = MX\}$, and

$$\mathscr{P}_M := \{X \in \mathfrak{M}_n : MXM = X\}.$$

The set $\mathscr{C}_M$ is a well-known sub-algebra of $\mathfrak{M}_n$, and $\mathscr{P}_M$ is a subspace of $\mathfrak{M}_n$, satisfying some useful, elementary properties like:

$$\mathscr{P}_{M^{-1}} = \mathscr{P}_M, \quad \text{and} \quad M\mathscr{P}_M = \mathscr{P}_M M = \mathscr{P}_M;$$

moreover, if $X \in \mathscr{P}_M$ is nonsingular, then $X^{-1} \in \mathscr{P}_M$. We know $M = LR$, with L and R involutions. Our aim is to find all involutions of $\mathscr{P}_M$.

We say that a $k \times r$ matrix $Y = (y_{ij})$ is a *Pascal matrix* if $J_k Y J_r = Y$; the set of such matrices is denoted by $\mathscr{P}_{kr}$. Clearly $Y \in \mathscr{P}_{kr}$ iff $H_k Y H_r = H_k Y + Y H_r$. Note that $[H_k Y]_{ij} = y_{i+1,j}$, $[Y H_r]_{ij} = y_{i,j-1}$, and $[H_k Y H_r]_{ij} = y_{i+1,j-1}$, with the convention $y_{uv} = 0$ if $u = k+1$ or $v = 0$. So $Y \in \mathscr{P}_{kr}$ iff $y_{i+1,j-1} = y_{i+1,j} + y_{i,j-1}$, a rule similar to Pascal's for the binomial coefficients, namely: two diagonally adjacent entries α, β of Y, as in the diagram

$$\begin{array}{cc} \alpha & * \\ S & \beta \end{array} \qquad \rightsquigarrow S = \alpha + \beta \tag{8}$$

determine the value $\alpha + \beta$ for the entry S just bellow α. It is easy to see, using (8), that Y is an *upper triangle*, more precisely $y_{ij} = 0$, for $j < i + \max\{0, r-k\}$. Note that an upper triangle has one of the following two patterns

$$\boxed{\begin{array}{ccccccccc} &&&& \circledast & * & * & \cdots & * \\ &&&&& * & * & \cdots & * \\ &&&&&& * & \cdots & * \\ &&&&&&& \ddots & \vdots \\ &&&&&&&& * \end{array}} \quad \text{or} \quad \boxed{\begin{array}{ccccc} \circledast & * & * & \cdots & * \\ & * & * & \cdots & * \\ && * & \cdots & * \\ &&& \ddots & \vdots \\ &&&& * \\ &&&& \\ &&&& \\ &&&& \end{array}} \tag{9}$$

according to $k < r$ or $k \geqslant r$, respectively, with zeroes in non specified entries. Let $m := \min\{k, r\}$. Let $c_1, \ldots, c_m$ be the nonzero entries of a generic Y's first row (c_1 is the entry marked as $\circledast$). Applying the rule (8) we get the entries $c_1, -c_1, c_1, -c_1, \ldots$ down the longest diagonal of Y; and all other $*$'s are uniquely determined as well. So $\dim \mathscr{P}_{kr} = \min\{k, r\}$.

We frequently refer to matrices $X \in \mathfrak{M}_n$ related to the direct sum (7); then we shall always assume X partitioned as

$$X = (X_{\sigma\tau})_{\sigma,\tau=1}^{w}, \quad \text{where } X_{\sigma\tau} \text{ is } n_\sigma \times n_\tau, \tag{10}$$

and refer to the $X_{\sigma\tau}$ as the *standard blocks of* X. For example, we may express the condition $Y \in \mathscr{P}_M$ (in terms of Y's standard blocks) as follows $J_\sigma Y_{\sigma\tau} J_\tau = Y_{\sigma\tau}$. Therefore, the arguments above lead to the following theorem.

Theorem 5.1. *$Y \in \mathscr{P}_M$ if and only if all standard blocks of Y are Pascal matrices. The dimension of $\mathscr{P}_M$ is $\sum_{\sigma,\tau=1}^{w} \min\{n_\sigma, n_\tau\}$.* □

The condition $X \in \mathscr{C}_M$, is equivalent to $J_{n_\sigma} X_{\sigma\tau} = X_{\sigma\tau} J_{n_\tau}$; and this in turn is equivalent to $H_{n_\sigma} X_{\sigma\tau} = X_{\sigma\tau} H_{n_\tau}$. The set $\mathscr{C}_M$ is completely determined in [10, pp. 220-*ff*]; the result is expressed here in terms of *Toeplitz matrices* (those satisfying $x_{ij} = x_{i+1,j+1}$): *$X \in \mathscr{C}_M$ if and only if all standard blocks $X_{\sigma\tau}$ are upper triangular Toeplitz matrices. $\mathscr{C}_M$ has the same dimension as $\mathscr{P}_M$.* So the generic $Y \in \mathscr{P}_M$ and $X \in \mathscr{C}_M$ have the same pattern of zeroes.

5.1. Partitioned matrices with upper triangular blocks. The algebra $\mathscr{C}_M$ and the vector space $\mathscr{P}_M$ are contained in the algebra $\mathscr{T}_M$ of all matrices having a partition (10) with upper triangular standard blocks. This algebra may be reduced to a block-upper triangular form by means of a permutational similarity, the permutation being the same which transforms the Jordan canonical form M into its Weyr canonical form (*cf.* [19, 2, 3]). This permutational reduction is the same as done in [20] for the algebra $\mathscr{C}_M$. Later on we use a refined notation. Let $\eta_1 > \eta_2 > \cdots > \eta_v$ be the distinct positive parts of the partition $(n_1, n_2, \dots)$, and denote by m_i the multiplicity of η_i, for $i = 1, \dots, v$. Clearly $n = m_1\eta_1 + \cdots + m_v\eta_v$. Let X_i be the submatrix of X made up of the square standard blocks $X_{\sigma\tau}$ of size $\eta_i \times \eta_i$; note that X_i is a principal square submatrix of X, of order $m_i\eta_i$, composed by m_i^2 standard blocks of X. Then

$$X = \begin{bmatrix} X_1 & * & \cdots & * \\ * & X_2 & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ * & * & \cdots & X_v \end{bmatrix}, \tag{11}$$

where each $*$ represents a composite of non-square standard blocks $X_{\sigma\tau}$ of X, and any square block of X is a block of some X_k. For $k = 1, \dots, v$, select in each block $X_{\sigma\tau}$ of X_k the first entry of the main diagonal of $X_{\sigma\tau}$; the selected entries form an $m_k \times m_k$ principal submatrix of X_k that we denote by $Z_k(X)$, and often shorten to Z_k. The following lemma is proved by simple combinatorics based on the fact that any standard block $X_{\sigma\tau}$ below the X_i's in (11) has a zero first column (check (9)), and if $X_{\sigma\tau}$ is a block of some X_k, the only nonzero entry of the first column of $X_{\sigma\tau}$ is its top-left entry (which is an entry of Z_k).

Lemma 5.2. *There exists a permutation matrix P, which satisfies*

$$PXP^t = \begin{bmatrix} Z_1 & * & \dots & * & * \\ & Z_2 & \dots & * & * \\ & & \ddots & \vdots & \vdots \\ & & & Z_v & * \\ & & & & \mathfrak{r}(X) \end{bmatrix}, \qquad (12)$$

for any $X \in \mathscr{T}_M$, where $\mathfrak{r}(X)$ is the matrix obtained by deleting all rows and columns of X corresponding to $Z_1, \dots, Z_v$. □

Clearly, as X runs over $\mathscr{P}_M$ $[\mathrm{G}_M]$, $Z_1 \oplus \cdots \oplus Z_v$ fills completely the algebra $\mathfrak{M}_{n_1} \oplus \cdots \oplus \mathfrak{M}_{n_v}$ [resp., the group $\mathrm{GL}_{n_1} \oplus \cdots \oplus \mathrm{GL}_{n_v}$]. For any $X, Y \in \mathscr{T}_M$, $T \in \mathscr{T}_M^*$ and $i \in \{1, \dots, v\}$, we have

$$Z_i(XY) = Z_i(X)Z_i(Y), \quad \text{and} \quad Z_i(TXT^{-1}) = Z_i(T)Z_i(X)Z_i(T)^{-1}.$$

This implies the following

Lemma 5.3. *The similarity classes of $Z_1(X), \dots, Z_v(X)$ are invariant for the action of $\mathscr{T}_M^*$ on $\mathscr{T}_M$. Suppose $X \in \mathscr{T}_M$ $[X \in \mathscr{P}_M]$. For any W_i similar to $Z_i(X)$, $i = 1, \dots, w$, there exists Y in the $\mathscr{T}_M^*$-orbit [resp., G_M -orbit] of X that satisfies $W_1 \oplus \cdots \oplus W_v = Z_1(Y) \oplus \cdots \oplus Z_v(Y)$.* □

Lemma 5.4. *Let Z_{kd} be the m_k-square submatrix of X_k made up of the d-th entry of the main diagonal of each block $X_{\sigma\tau}$ of X_k. Then*

$$\det X = \prod_{k=1}^{v} \prod_{d=1}^{\eta_k} \det Z_{kd}.$$

Proof. The matrix $\mathfrak{r}(X)$ of (11) has a partition $(X'_{\sigma\tau})_{\sigma,\tau=1}^{w}$, where $X'_{\sigma\tau}$ is upper triangular of size $(n_\sigma - 1) \times (n_\tau - 1)$. So the triangular reduction (12) may be inductively refined by a permutational similarity performed on the rows and columns of $\mathfrak{r}(X)$. We convention that matrices of one non-positive size is empty; so some of the $X'_{\sigma\tau}$ may be empty. Note that Z_{k1} is the previous Z_k. Therefore, from (12),

$$\det X = \left(\prod_{k=1}^{v} \det Z_{k1} \right) \det \mathfrak{r}(X),$$

and we get the lemma's formula by induction. □

Corollary 5.5. *In the notation of lemma* 5.4, $\det X = \prod_{k=1}^{v} \det X_k$ *and, if the* $X_{\sigma\tau}$ *are Toeplitz upper triangular standard blocks,*

$$\det X = \prod_{k=1}^{v} \big(\det Z_{k1}\big)^{\eta_k}. \qquad \square$$

Restriction homomorphisms. The method used to get $\mathfrak{r}(X)$ in lemma 5.2 may be slightly extended. For a fixed $e \in \mathbb{N}$, eliminate in each block $X_{\sigma\tau}$ the first e rows and the first e columns; let $\mathfrak{r}(X_{\sigma\tau})$ be the resulting restricted standard blocks, and $\mathfrak{r}(X)$ the resulting restricted matrix. As $\mathfrak{r}(X_{\sigma\tau})$ is upper triangular, $\mathfrak{r}(X)$ lies in $\mathscr{T}_{\mathfrak{r}(M)}$. Note that if $e \geqslant n_k$, then all standard blocks $X_{\sigma\tau}$ are empty for $\min\{\sigma, \tau\} \geqslant k$. It is clear that

$$X_{\sigma\tau} = \left[\begin{array}{cc} E_{\sigma\tau} & * \\ 0 & \mathfrak{r}(X_{\sigma\tau}) \end{array} \right],$$

where $E_{\sigma\tau}$ is an e-square upper triangular matrix; as e does not depend on σ, τ, $\mathfrak{r}(X_{\sigma\tau}X_{\tau\mu}) = \mathfrak{r}(X_{\sigma\tau})\mathfrak{r}(X_{\tau\mu})$. Therefore $\mathfrak{r}(XY) = \mathfrak{r}(X)\mathfrak{r}(Y)$ for $X, Y \in \mathscr{T}_M$. So the *restriction* mapping $\mathfrak{r} : \mathscr{T}_M \to \mathscr{T}_{\mathfrak{r}(M)}$ is an algebra homomorphism, which obviously preserves the Toeplitz property and the Pascal rule (8). From corollary 5.5, if $X \in \mathscr{C}_M$ is nonsingular, then $\mathfrak{r}(X)$ is nonsingular as well (this follows from the fact that $Z_{kr} = Z_{k1}$ or $Z_{kr} = \varnothing$). According to this we have three restrictions of $\mathfrak{r}$:

$$\mathfrak{r}_c : \mathscr{C}_M \to \mathscr{C}_{\mathfrak{r}(M)}, \quad \mathfrak{r}_p : \mathscr{P}_M \to \mathscr{P}_{\mathfrak{r}(M)} \quad \text{and} \quad \mathfrak{r}_g : \mathrm{G}_M \to \mathrm{G}_{\mathfrak{r}(M)}\,.$$

Lemma 5.6. $\mathfrak{r}_c$, $\mathfrak{r}_p$ *and* $\mathfrak{r}_g$ *are epimorphisms of, respectively, algebras, vector spaces and groups.* □

The simple proof is omitted. An example shows the way to do it. In the following diagram, a south-east standard block of size 7 × 11, of Pascal [or Toeplitz] type, whose nonzero entries are denoted by $*$'s, is shown in the process of extension to a matrix of the *same type* (Pascal or Toeplitz), by adding e rows on top and e columns on the left, with $e = 6$:

$$\begin{array}{|cccccc|ccccccccccc|}
\hline
&&&&\diamond&\diamond&\diamond&\diamond&\diamond&\diamond&\diamond&\cdot&\cdot&\cdot&\cdot&\cdot&\cdot\\
&&&&&\diamond&\diamond&\diamond&\diamond&\diamond&\diamond&\diamond&\cdot&\cdot&\cdot&\cdot&\cdot\\
&&&&&&\diamond&\diamond&\diamond&\diamond&\diamond&\diamond&\diamond&\cdot&\cdot&\cdot&\cdot\\
&&&&&&&\diamond&\diamond&\diamond&\diamond&\diamond&\diamond&\diamond&\cdot&\cdot&\cdot\\
&&&&&&&&\diamond&\diamond&\diamond&\diamond&\diamond&\diamond&\diamond&\cdot&\cdot\\
&&&&&&&&&\diamond&\diamond&\diamond&\diamond&\diamond&\diamond&\diamond&\cdot\\
\hline
&&&&&&&&&&\ast&\ast&\ast&\ast&\ast&\ast&\ast\\
&&&&&&&&&&&\ast&\ast&\ast&\ast&\ast&\ast\\
&&&&&&&&&&&&\ast&\ast&\ast&\ast&\ast\\
&&&&&&&&&&&&&\ast&\ast&\ast&\ast\\
&&&&&&&&&&&&&&\ast&\ast&\ast\\
&&&&&&&&&&&&&&&\ast&\ast\\
&&&&&&&&&&&&&&&&\ast\\
\hline
\end{array}\,. \tag{13}$$

The $\diamond$'s represent the entries of the extension which are uniquely determined by the top row of $*$'s, and by the type of extension, namely, by the Pascal rule (8) [or the Toeplitz rule]. The entries denoted by $\cdot$'s may be filled in with some flexibility: we may choose arbitrarily the first row of dots, the remaining dots being uniquely determined by that choice; the dotted entries contribute to the kernel of $\mathfrak{r}_x$, for $x = c, p, g$ (note, *en passant*, that $\ker \mathfrak{r}_x$ is an affine variety of $\mathscr{T}_M$ with dimension $\sum \{\min\{e, n_i, n_j\} : 1 \leqslant i, j \leqslant w\}$).

The similarity actions of G_M and $\mathrm{G}_{\mathfrak{r}(M)}$ are also well related, in the sense that the following diagram

$$\begin{tikzcd}
\mathrm{G}_M \times \mathscr{P}_M \arrow[r] \arrow[d, "\mathfrak{r}_g \times \mathfrak{r}_p"'] & \mathscr{P}_M \arrow[d, "\mathfrak{r}_p"] \\
\mathrm{G}_{\mathfrak{r}(M)} \times \mathscr{P}_{\mathfrak{r}(M)} \arrow[r] & \mathscr{P}_{\mathfrak{r}(M)}
\end{tikzcd} \tag{14}$$

is commutative, where the horizontal arrows denote the group actions.

The whole thing here may be done with the restriction defined as the elimination, in each block $X_{\sigma\tau}$, of the *last* e rows and columns.

5.2. The case of a single Jordan block with eigenvalue 1. We now settle the case when M has only one Jordan block, namely $M = J_n$; we denote H_n and J_n simply by H and J. Then $\mathscr{C}_J$ is $\mathbb{F}[J] = \mathbb{F}[H]$; it is the set of upper-triangular Toeplitz matrices, the sub-algebra of $\mathfrak{M}_n$ generated by H (and by J). The elements of $\mathscr{C}_J$ may be presented as $g(H)$, where $g \in \mathbb{F}[[x]]$, the ring of formal power series, because $g(x)$ will be ultimately chopped off modulo x^n. For example, the inverse of $1 - x$ is $\ell(x) = \sum_{k \geqslant 0} x^k$; so $J^{-1} = \ell(H)$, the upper triangle of all 1's. Note that $\mathscr{P}_J = \mathscr{P}_{nn}$, the set of the n-square Pascal matrices.

Lemma 5.7. *Let $Y \in \mathscr{P}_J$ and $f \in \mathbb{F}[[x]]$. Define $\xi(x) = \frac{x}{x-1}$. We have $\xi(H) = -\sum_{k\geqslant 1} H^k$, $Yf(H) = f(\xi(H))Y$ and $f(H)Y = Yf(\xi(H))$.*

Proof. The formula for $\xi(H)$ is obvious. It is easy to check that $YH = Y - YJ = (I - J^{-1})Y = \xi(H)Y$; so the second formula holds for $f(x) = x$, and the general case is obtained by induction. The third formula follows from the second one, because $\xi(\xi(x)) = x$. □

As explained above, for any $\gamma = (\gamma_1, \dots, \gamma_n) \in \mathbb{F}^n$ there exists one and only one n-square Pascal matrix with γ as first row; this matrix will be denoted $P(\gamma_1, \dots, \gamma_n)$. We single out two special Pascal matrices

$$\mathit{\Delta} = P(1, 0, \dots, 0) \quad \text{and} \quad \mathit{\Gamma} = P(1, 1, \dots, 1), \tag{15}$$

also denoted $\mathit{\Delta}_n$ and $\mathit{\Gamma}_n$ if needed. Computing the entries $\mathit{\Gamma}_{ij}$ of $\mathit{\Gamma}$ by the rule (8) is like building a Pascal triangle; we get $\mathit{\Gamma}_{ij} = (-1)^{i-1}\binom{j-1}{i-1}$, with the convention $\binom{j}{i} = 0$ if $j < i$. Note that any leading or trailing principal submatrix of a Pascal matrix is again Pascal, and that the second row of $\mathit{\Delta}$ is $(0, -1, \dots, -1)$; therefore $\mathit{\Delta}_n = 1 \oplus (-\mathit{\Gamma}_{n-1})$.

Lemma 5.8. *$\mathit{\Delta}$ and $\mathit{\Gamma}$ are involutions, and $J = \mathit{\Gamma}\mathit{\Delta}$.*

Proof. For $\ell(x) = (1-x)^{-1}$, the first row of $\ell(H)$ is $(1, 1, \dots, 1)$, the same as the first row of J^{-1}; thus $\mathit{\Gamma} = \mathit{\Delta}J^{-1}$. As $\mathit{\Delta}_n, \mathit{\Gamma}_n \in \mathscr{P}_{nn}$ and $\mathit{\Delta}_n = \mathit{\Gamma}_n J$, we get: $\mathit{\Delta}_n$ is an involution iff $\mathit{\Gamma}_n$ is an involution. The fact that $\mathit{\Gamma}_n^2 = \mathit{\Delta}_n^2 = I_n$ follows by a simple induction using $\mathit{\Delta}_n = 1 \oplus (-\mathit{\Gamma}_{n-1})$. □

Remarks on characteristic 2. The case when $\mathbb{F}$ has characteristic 2 is going to be left out by the following main reasons. Firstly, in characteristic $\neq 2$, an involution is similar to one of the $n+1$ matrices $-I_k \oplus I_{n-k}$, while in characteristic 2 the Jordan form of an involution is a direct sum of blocks of two types: $J_2 = \left[\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right]$ and $J_1 = [1]$. As a second fact, to find all involutions of $\mathscr{P}_J$ amounts to finding all power series $f(x)$ such that $f(H)\mathit{\Delta}f(H) = \mathit{\Delta}$; from lemma 5.7, this means to characterize those f's which satisfy

$$f(x)f(\xi(x)) = 1.$$

Introducing the expressions $f(x) = \sum_{k\geqslant 0} c_k x^k$ and $\xi(x) = -\sum_{k\geqslant 1} x^k$, we get

$$(c_0 + c_1 x + c_2 x^2 + \dots)\left[c_0 - c_1 \textstyle\sum_{k\geqslant 1} x^k + c_2 \left(\sum_{k\geqslant 1} x^k\right)^2 - \cdots\right] = 1. \tag{16}$$

This gives rise to an infinite sequence of equations involving the c_i's. The first equation is $c_0^2 = 1$ and, when we equate to 0 the coefficient of x^{2k} of the left hand side of (16), $2c_0c_{2k}$ comes out as a quadratic expression in the c_i's for $i < 2k$. So the case of characteristic 2 raises problems of a very different nature from those treated below. □

From now on we assume
$\mathbb{F}$ has characteristic $\neq 2$

Lemma 5.9. *Assume $Y \in \mathscr{P}_J$ is nonsingular. There exists one and only one polynomial $g(x)$ of the form*

$$g(x) = \prod_{odd\ w<n} (1 - \theta_w x^w), \tag{17}$$

(a product of $\lfloor \frac{n}{2} \rfloor$ factors) such that the matrix $Y^* = g(H)^{-1}Yg(H)$ has first row with entries $y^*_{1i} = 0$ for all even i.

Proof. For upper triangular matrices Y consider the similarity transformation $Y \mapsto Y^* := (I - \theta H^w)^{-1}Y(I - \theta H^w)$. Let e_k be the k-th row of I, and $Y_k := e_kY$. Using $e_1H^k = e_{k+1}$, compute the first row of Y^*:

$$\begin{aligned}
e_1Y^* &= e_1(I - \theta H^w)^{-1}Y(I - \theta H^w) \\
&= e_1 \sum_{i\geqslant 0} \theta^i H^{iw} Y(I - \theta H^w) = \sum_{i\geqslant 0} \theta^i Y_{iw+1}(I - \theta H^w) \\
&= \left(Y_1 + \theta Y_{w+1} + \sum_{i\geqslant 2} \theta^i Y_{iw+1} \right) (I - \theta H^w) \\
&= Y_1 - \theta(Y_1H^w - Y_{w+1}) + \left[\sum_{i\geqslant 2} \theta^i Y_{iw+1} - \sum_{i\geqslant 1} \theta^{i+1} Y_{iw+1} H^w \right].
\end{aligned} \tag{18}$$

As Y is upper triangular, the first $w + k - 1$ entries of Y_kH^w are 0; therefore, the row in (18) between brackets has its first $2w$ entries 0. Thus we have: the transformation $Y \mapsto Y^*$ does not change the elements y_{i1} for $1 \leqslant i \leqslant w$, and $y^*_{1,w+1} = y_{1,w+1} - \theta(y_{11} - y_{w+1,w+1})$.

For $Y \in \mathscr{P}_J$, the diagonal elements of Y are $y_{ii} = (-1)^{i+1}y_{11}$; therefore, if w is odd, then $y^*_{w+1,1} = y_{w+1,1} - 2\theta y_{11}$, and we may zero out $y^*_{w+1,1}$ by appropriate choice of θ, without changing the elements of the first row of Y on the left of $y_{1,w+1}$. Taking $w = 1, 3, 5, \ldots$ we may successively eliminate all entries of the first row of Y in even positions.

To show the uniqueness of g note that in (17) there is no repeated exponent w. So we may order the factors $(1-\theta_w x^w)$ by strictly increasing values of w from left to right. We act on Y by similarity with matrices $(I-\theta_1 H), (I-\theta_3 H^3), \dots,$ in this order; to zero out successively $x_{12}, x_{14}, \dots$; in each step, each $\theta_1, \theta_3, \dots$ is uniquely determined. □

Lemma 5.10. *Let $P \in \mathscr{P}_J$ be an involution. The involutions of $\mathscr{P}_J$ are the matrices $\pm f(H)Pf(H)^{-1}$, for $f \in \mathbb{F}[x]$, $f(0) \neq 0$. For an involution $X \in \mathscr{P}_J$, $x_{11} = \pm 1$ and the sign of x_{11} is a complete invariant for the G_J-similarity action on $\mathfrak{I}_J$.*

Proof. Clearly the matrices $\pm f(H)\varDelta f(H)^{-1}$ are involutions. Conversely let us pick any involution $Y \in \mathscr{P}_J$. Then Y has diagonal $(\pm 1, \mp 1, \pm 1, \mp 1, \dots)$. The matrix Y^* of lemma 5.9 is also an involution; therefore, for $k \geqslant 2$, the product of the k-th column of Y^* by the first row is zero; reading this condition entry-wise for successively increasing odd values of k, we get $y^*_{1k} = 0$ for all odd $k \neq 1$. Therefore, the first row of Y^* is $(\pm 1, 0, \dots, 0)$, in other words, $Y^* = \pm\varDelta$. We thus have $Y = \pm g(H)\varDelta g(H)^{-1}$. The value range and invariance of x_{11} are obvious. □

Therefore, we have two G_J-orbits in $\mathfrak{I}_J$, that we denote by $\mathfrak{I}_J^+$ and $\mathfrak{I}_J^-$

$$\mathfrak{I}_J^+ = \{X \in \mathfrak{I}_J : x_{11} = 1\} \quad \text{and} \quad \mathfrak{I}_J^- = \{X \in \mathfrak{I}_J : x_{11} = -1\}.$$

Theorem 5.11. *Suppose $C \sim J$, say $C = TJT^{-1}$. Clearly $\mathfrak{I}_C = T\mathfrak{I}_J T^{-1}$. The action of G_C on $\mathfrak{I}_C$ has two orbits, $\mathfrak{I}_C^+ := T\mathfrak{I}_J^+ T^{-1}$ and $\mathfrak{I}_C^- := T\mathfrak{I}_J^- T^{-1}$. The orbit $\mathfrak{I}_C^+$ consists of the matrices $Y \in \mathfrak{I}_C$ such that $(I-Y)(I-C)^{n-1} = O$.*

Proof. The last assertion is the only one deserving some attention. $X \in \mathfrak{I}_J^+$ iff the first column of $I - X$ is zero; and this holds iff $(I-X)(I-J)^{n-1} = O$, *i.e.*, $(I - TXT^{-1})(I-C)^{n-1} = O$. The theorem follows from $Y \in \mathfrak{I}_C^+$ iff $Y = TXT^{-1}$, with $X \in \mathfrak{I}_J^+$. □

Corollary 5.12. *We have $(\varGamma_n, \varDelta_n) \approx (S_n, S_n C((x-1)^n))$, where $C(f)$ is the companion matrix in* (1)*, and S_n is the skew-identity.*

Proof. We let $C := C((x-1)^n))$, and apply theorem (5.11). Note that $\varGamma_n \in \mathfrak{I}_J^+$. So we only need to prove that $S_n \in \mathfrak{I}_C^+$ (because $T\varGamma T^{-1} \in \mathfrak{I}_C^+$, and all matrices in $\mathfrak{I}_C^+$ are G_C-similar). So we have to show that

$$(I - S_n)(I - C)^{n-1} = O. \tag{19}$$

Note that $I-C$ is nilpotent, non-derogatory and, therefore, $(I-C)^{n-1}$ has rank one. The sum of the last row of $C(f)$ is $1-f(1)$; in case $f(x)=(x-1)^n$ that sum is 1. Hence C is row stochastic, and therefore each row of $I-C$ has sum 0; so $u=[1,1,\dots,1]^t$ is an eigenvector of $I-C$ corresponding to the eigenvalue 0. It is easy to check that, if $e_1=[1,0,\dots,0]^t$, then the vectors

$$e_1,(I-C)e_1,(I-C)^2e_1,\dots,(I-C)^{n-1}e_1$$

are linearly independent, *i.e.*, they form a Jordan chain for $I-C$. As a consequence, $(I-C)^{n-1}e_1$ is proportional to u (in fact, it equals u). This implies that all rows of $(I-C)^{n-1}$ are equal; thus $S_n(I-C)^{n-1}=(I-C)^{n-1}$, and (19) holds, as desired. □

5.3. Back to a direct sum of Jordan blocks with eigenvalue 1. For any involution K, we denote by $\mu(K)$ the multiplicity of 1 as an eigenvalue of K. The *signature* of an involution L of $\mathscr{P}_M$, is the v-tuple

$$\operatorname{sg}(L)=\big(\mu(Z_1(L)),\dots,\mu(Z_v(L))\big). \tag{20}$$

The *signature*$^{\approx}$ of a pair of involutions, (L,R), such that $LR\sim M$, is defined as the signature of any $L'\in\mathscr{P}_M$ similar to L; the notation is $\operatorname{sg}^{\approx}(L,R)$, or just $\operatorname{sg}(L,R)$. In the definition of $\operatorname{sg}(L,R)$ we used lemma 5.3. The invariance of $\operatorname{sg}(R)$ $[\operatorname{sg}^{\approx}(L,R)]$ for G_M-similarity [resp., simultaneous similarity] also follows easily.

For each w-tuple of signs, $\epsilon=(\epsilon_1,\dots,\epsilon_w)\in\{1,-1\}^w$, let L_ϵ and R_ϵ be the matrices

$$L_\epsilon=\epsilon_1\mathit{\Gamma}_{n_1}\oplus\cdots\oplus\epsilon_w\mathit{\Gamma}_{n_w}\quad\text{and}\quad R_\epsilon=\epsilon_1\mathit{\Delta}_{n_1}\oplus\cdots\oplus\epsilon_w\mathit{\Delta}_{n_w}. \tag{21}$$

Clearly $M=L_\epsilon R_\epsilon$ for any ϵ. For $k\in\{1,\dots,v\}$, let U_k be the set of those $i\in\{1,\dots,w\}$ such that $n_i=\eta_k$; we know $|U_k|=m_k$. So $U_1,\dots,U_v$ are disjoint consecutive intervals which cover $\{1,\dots,w\}$. Any permutation $\sigma\in\mathscr{S}_w$, such that $\sigma[U_k]=U_k$ (for all k) will transform ϵ into $\epsilon'=(\epsilon_{\sigma(1)},\dots,\epsilon_{\sigma(w)})$; then $(L_\epsilon,R_\epsilon)\approx(L_{\epsilon'},R_{\epsilon'})$ by an obvious permutational simultaneous similarity. As $\mathit{\Gamma}_{n_i}\sim\mathit{\Delta}_{n_i}$, then L_ϵ and R_ϵ are similar, *i.e.*, $\mu(L_\epsilon)=\mu(R_\epsilon)$.

We say that ϵ is *adjusted* if each section $(\epsilon_i : i\in U_k)$ is a non-increasing m_k-tuple, *i.e.*,

$$\big(\epsilon_i : i\in U_k\big)=(\underbrace{1,1,\dots,1}_{\mu_k},\underbrace{-1,-1,\dots,-1}_{m_k-\mu_k}),$$

where $0 \leqslant \mu_k \leqslant m_k$. If ϵ is not adjusted, then we may reorder each section $(\epsilon_i : i \in U_k)$ in non-increasing order; the w-tuple ϵ' obtained in this way is called the *adjusted of* ϵ. We thus have proven

Lemma 5.13. *Two pairs (L_ϵ, R_ϵ) and (L_τ, R_τ) are simultaneously similar if and only if they have the same signature.* □

Theorem 5.14. *For an involution $\mathit{\Psi} \in \mathscr{P}_M$, there exists a unique adjusted sign tuple ϵ such that $\mathit{\Psi}$ is G_M-similar to L_ϵ.*

Proof. Partition $\mathit{\Psi}$ as $(\mathit{\Psi}_{rs})_{r,s=1}^{w}$, where each $\mathit{\Psi}_{rs}$ is a Pascal matrix of size $n_r \times n_s$. We keep the notation related to the partition (11), namely n_i, η_k, m_k, and the meaning of $Z_k = Z_k(\mathit{\Psi})$ and $\mathfrak{r}(\mathit{\Psi})$ of (12). In the current case, the diagonal blocks $Z_1, \dots, Z_v, \mathfrak{r}(\mathit{\Psi})$ are involutions. By induction, there exists $W \in \mathrm{G}_{\mathfrak{r}(M)}$ such that

$$W\mathfrak{r}(\mathit{\Psi})W^{-1} = \mathit{\Lambda}_1 \oplus \cdots \oplus \mathit{\Lambda}_w,$$

where each $\mathit{\Lambda}_r$ is a Pascal involution of order $n_r - 1$ for $1 \leqslant r \leqslant w$ (the last m_v $\mathit{\Lambda}_i$'s are empty if $n_w = 1$). As $\mathfrak{r}_g$ is surjective (lemma 5.6) choose $U \in \mathrm{G}_M$ such that $\mathfrak{r}(U) = W$, and define $\mathit{\Omega} := U\mathit{\Psi}U^{-1}$. The commutativity of (14) implies

$$\mathfrak{r}(\mathit{\Omega}) = \mathit{\Lambda}_1 \oplus \cdots \oplus \mathit{\Lambda}_w.$$

In case $n_s > 1$, $\mathfrak{r}(\mathit{\Omega}_{ss})$ is an involution of $\mathscr{P}_{n_s-1,n_s-1}$; so the diagonal entries of $\mathit{\Omega}_{ss} \in \mathscr{P}_{n_s n_s}$ are alternately ±1. Each off-diagonal block $\mathit{\Omega}_{rs}$ satisfies $\mathfrak{r}(\mathit{\Omega}_{rs}) = 0$, in other words: all entries of $\mathit{\Omega}_{rs}$ are zero except, possibly, those in $\ker \mathfrak{r}$-positions (check the comments following (13)), in this case the entry in the top-right position of $\mathit{\Omega}_{rs}$. Let ω_{ij} be the entries of $\mathit{\Omega}$, $1 \leqslant i, j \leqslant n$. For $1 \leqslant r \leqslant w$, let $\alpha_r, \beta_r \in \{1, \dots, n\}$ be defined by the conditions:

$$\begin{array}{ll} \omega_{\alpha_r\alpha_r} & \text{is the \textit{first} diagonal entry of } \mathit{\Omega}_{rr} \\ \omega_{\beta_r\beta_r} & \text{is the \textit{last} diagonal entry of } \mathit{\Omega}_{rr}. \end{array}$$

So the top-right entry of $\mathit{\Omega}_{rs}$ is $\omega_{\alpha_r\beta_s}$. Note that $\alpha_r = \beta_r$ iff $n_r = 1$. Clearly $\mathit{\Omega}$ has a partition like (12), with $\mathfrak{r}(\mathit{\Psi})$ replaced by $\mathfrak{r}(\mathit{\Omega})$.

Case $\eta_v = 1$. In this case, $\mathit{\Omega}_v$, the matrix consisting of the 1-by-1 standard blocks $\mathit{\Omega}_{rs}$, is nothing but Z_v; thus it is an involution of $\mathfrak{M}_{m_v}$, and it may be any such involution. Partition $\mathit{\Omega}$ as

$$\mathit{\Omega} = \begin{bmatrix} \mathit{\Sigma} & K \\ C & \mathit{\Omega}_v \end{bmatrix}. \tag{22}$$

The pattern of zeros of K and C are as follows: all rows of K [columns of C] are zero except (possibly) those of indices α_r [resp., β_r], for r such that $n_r > 1$.

For any $Q \in \mathrm{GL}_{m_v}$, the matrix $\bar{Q} = I_{n-m_v} \oplus Q$ lies in G_M. Choose Q such that $Q\varOmega_v Q^{-1}$ is a diagonal involution; when we transform $\varOmega$ into $\bar{Q}\bar{\varOmega}\bar{Q}^{-1}$, the matrix $\varSigma$ and the patterns of K and C do not change, and $\varOmega_v$ becomes a diagonal involution. So we shall assume that *$\varOmega_v$ is a diagonal involution.* Therefore the diagonal entries of $\varOmega$ are ± 1.

Let E_{pq} be the $n_p \times n_q$ matrix with all entries 0 except the top-right entry which is 1; and let $T_{pq}(x) = I_n + xE_{pq}$. Clearly $T_{pq}(x) \in \mathrm{G}_M$ and $T_{pq}(x)^{-1} = T_{pq}(-x)$. We shall zero out all off diagonal standard blocks $\varOmega_{rs}$ using a sequence of similarity transformations by matrices $T_{pq}(x)$. Let

$$\varOmega' = T_{pq}(x)\varOmega T_{pq}(x)^{-1}.$$

This is the similarity which transforms $\varOmega$ into $\varOmega'$ in two steps:

Step 1: *add to row α_p, the row β_q multiplied by x*, followed by
Step 2: *add to column β_q, the column α_p multiplied by $-x$.*

Pick any nonzero entry of C, say $\omega_{\alpha_p\beta_q}$, such that $\omega_{\alpha_p\alpha_p}\omega_{\beta_q\beta_q} = -1$. Note that $n_q > n_p = 1$. In step 1, the row β_q has a sole nonzero entry, namely $\omega_{\beta_q\beta_q} = \pm 1$; so this step only changes the chosen entry $\omega_{\alpha_p\beta_q}$; and adds to it $x\omega_{\beta_q\beta_q}$. In step 2, some of the top-right entries of the standard blocks of $\varSigma$ may change, but all the rest of $\varSigma$ remains intact; moreover, as $\varOmega_v$ is diagonal, the only element of C which is changed is the chosen entry $\omega_{\alpha_p\beta_q}$, which transforms, after the two steps are made, into

$$\omega'_{\alpha_p\beta_q} = \omega_{\alpha_p\beta_q} + x(\omega_{\beta_q\beta_q} - \omega_{\alpha_p\alpha_p})$$

So, as $\omega_{\beta_q\beta_q} - \omega_{\alpha_p\alpha_p} = \pm 2$, we may zero out $\omega'_{\alpha_p\beta_q}$ by appropriate choice of x.

Now we treat the case $\omega_{\alpha_p\alpha_p}\omega_{\beta_q\beta_q} = 1$. The assumption $\varOmega^2 = I_n$ when read in terms of standard blocks, implies

$$\sum_{k=1}^{w} \varOmega_{pk}\varOmega_{kq} = 0, \quad \text{for } p \neq q. \tag{23}$$

In our case ($n_p = 1$), $\varOmega_{pk}\varOmega_{kq}$ is the last row of $\varOmega_{kq}$ multiplied by $\omega_{\alpha_p\beta_k}$; this is a zero row if $(n_k > 1 \wedge q \neq k)$ or $(n_k = 1 \wedge p \neq k)$; (the last term of this disjunction follows from the fact that $\mathit{\Omega}_v$ is diagonal). So (23) reduces to:

$$\mathit{\Omega}_{pp}\mathit{\Omega}_{pq} + \mathit{\Omega}_{pq}\mathit{\Omega}_{qq} = 0.$$

As $\mathit{\Omega}_{pp} = \omega_{\alpha_p\alpha_p}$ and $\mathit{\Omega}_{pq}\mathit{\Omega}_{qq}$ is the last row of $\mathit{\Omega}_{qq}$ multiplied by $\omega_{\alpha_p\beta_q}$, we have $\omega_{\alpha_p\alpha_p}\omega_{\alpha_p\beta_q} + \omega_{\alpha_p\beta_q}\omega_{\beta_q\beta_q} = 0$, that is $\omega_{\alpha_p\beta_q} = 0$.

We have proven the existence of a similarity action of G_M that zeroes out the matrix C in (22). Of course the same procedure applies to K with minor changes, with similarity transformations that do not change the null matrix C, and zero out K. The conclusion of case $\eta_v = 1$ is that we may assume that $\mathit{\Omega}$ splits as $\mathit{\Omega} = \mathit{\Sigma} \oplus \mathit{\Omega}_v$. We shall assume, without loss of generality, that $\eta_v \geqslant 2$.

Case $\eta_v \geqslant 2$. The elimination of an entry $\omega_{\alpha_p\beta_q}$, with $p \neq q$, for which $\omega_{\alpha_p\alpha_p}\omega_{\beta_q\beta_q} = -1$ is done as before, using an elementary matrix $T_{pq}(x)$; things are easier because the acting row α_p and column β_q each have a sole nonzero entry, namely $\omega_{\alpha_p\alpha_q}$ and $\omega_{\beta_q\beta_q}$. The proof that $\omega_{\beta_p\beta_p}\omega_{\alpha_q\alpha_q} = 1$ implies $\omega_{\beta_p\alpha_q} = 0$ is done as before using the equation (23).

We have proven that $\mathit{\Psi}$ is G_M-similar to a direct sum $\mathit{\Phi}_1 \oplus \cdots \oplus \mathit{\Phi}_w$, where each $\mathit{\Phi}_r$ is an involution of $\mathscr{P}_{n_rn_r}$. By lemma 5.10, $\mathit{\Phi}_r$ is $\mathrm{G}_{J_{n_r}}$-similar to $\epsilon_r\mathit{\Gamma}_{n_r}$, where ϵ_r is the first diagonal entry of $\mathit{\Phi}_r$. So $\mathit{\Psi}$ is G_M-similar to L_ϵ. By permuting the order of the diagonal blocks, if necessary, we may assume ϵ is adjusted. The G_M-invariance of the signature proves uniqueness. □

The work done so far is enough to present a canonical form:

Theorem 5.15. *For any pair of involutions,* (L, R)*, such that* $LR \sim M$*, there exists a unique adjusted sign tuple* ϵ *such that* $(L, R) \approx (L_\epsilon, R_\epsilon)$*.* □

5.4. Direct sum of Jordan blocks with eigenvalue -1.

The whole theory developed for the eigenvalue 1 may be repeated for -1 with minor changes. One way of doing this is to replace M with $-M$; as $\mathscr{P}_{-M} = \mathscr{P}_M$, $\mathscr{C}_{-M} = \mathscr{C}_M$ and $\mathrm{G}_{-M} = \mathrm{G}_M$, the action of G_{-M} on $\mathfrak{I}_{-M}$ is the same as we have seen above.

For simultaneous similarity, we have $(L, R) \approx (X, Y)$ iff $(L, -R) \approx (X, -Y)$ [iff $(-L, R) \approx (-X, Y)$]. So theorem 5.15 has the obvious consequence

Corollary 5.16. *For any pair of involutions,* (L, R)*, such that* $LR \sim -M$*, there exist unique adjusted sign tuples* ϵ *and* τ*, such that* $(L, R) \approx (L_\epsilon, -R_\epsilon)$ *and* $(L, R) \approx (-L_\tau, R_\tau)$*.* □

Note that in this corollary τ [ϵ] is the adjusted of $-\epsilon$ [resp., $-\tau$].

Review of the main results of this section

For the action of G_M on $\mathfrak{I}_M$, the signature is a complete invariant, and the L_ϵ for adjusted ϵ's form a set of canonical forms.

For the simultaneous similarity action of GL_n on pairs of involutions (L, R) such that $\chi_{LR} = (x - \lambda)^n$, $\lambda = \pm 1$, the Segre characteristic $(n_1, \dots, n_w)$ of LR and the signature$^{\approx}$ form a complete system of invariants. The Segre characteristic produces the v-tuple of multiplicities, $(m_1, \dots, m_v)$, and the concept of adjusted ϵ. In case $\lambda = 1$ [$\lambda = -1$], the pairs (L_ϵ, R_ϵ) [resp., $(L_\epsilon, -R_\epsilon)$], for adjusted ϵ's form a set of canonical forms.

In the case $\lambda = 1$, we may present the canonical form (L_ϵ, R_ϵ) as

$$(L_\epsilon, R_\epsilon) = \epsilon_1(\mathit{\Gamma}_{n_1}, \mathit{\Delta}_{n_1}) \oplus \cdots \oplus \epsilon_w(\mathit{\Gamma}_{n_w}, \mathit{\Delta}_{n_w}). \tag{24}$$

Let us define V_k by

$$V_k := S_k C((x-1)^k).$$

According to corollary 5.12, we have $(\mathit{\Gamma}_k, \mathit{\Delta}_k) \approx (S_k, V_k)$. In this way, we may replace the canonical pair (L_ϵ, R_ϵ) by a sparser one, (S_ϵ, V_ϵ), defined by

$$(S_\epsilon, V_\epsilon) = \epsilon_1(S_{n_1}, V_{n_1}) \oplus \cdots \oplus \epsilon_w(S_{n_w}, V_{n_w}) \tag{25}$$

In the case $\lambda = -1$, one has the sparser canonical form $(S_\epsilon, -V_\epsilon)$.

The following picture shows the matrix $L_\epsilon = \mathit{\Gamma}_7 \oplus \mathit{\Gamma}_6 \oplus \mathit{\Gamma}_4 \oplus \mathit{\Gamma}_3 \oplus \mathit{\Gamma}_2 \oplus \mathit{\Gamma}_1$, where $\epsilon = (1, 1, 1, 1, 1, 1)$; the block diagonal and the Pascal structure of each block are well visible. On the right we show $PL_\epsilon P^t$, where P is the permutation matrix that transforms the Jordan form $M = J_7 \oplus J_6 \oplus J_4 \oplus J_3 \oplus J_2 \oplus J_1$, into the corresponding Weyr normal form [19]. Although the right hand side matrix shows some regularity in the distribution of the binomial coefficients, the new pattern is a lot scattered and difficult to decipher. This matrix P has been used by G. R. Belitskiĭ to reduce the algebra $\mathscr{C}_M$ (in fact a generalized form of it) to a block triangular form [2, 20]; this has a great advantage in working in Belitskiĭ's much more general framework; in the present reduced case, the option for a Jordan normal form M offers nicer pattern readability of canonical forms like (24).

$$\left[\begin{array}{*{23}{r}}
1&1&1&1&1&1&1\\
&-1&-2&-3&-4&-5&-6\\
&&1&3&6&10&15\\
&&&-1&-4&-10&-20\\
&&&&1&5&15\\
&&&&&-1&-6\\
&&&&&&1\\
&&&&&&&1&1&1&1&1&1\\
&&&&&&&&-1&-2&-3&-4&-5\\
&&&&&&&&&1&3&6&10\\
&&&&&&&&&&-1&-4&-10\\
&&&&&&&&&&&1&5\\
&&&&&&&&&&&&-1\\
&&&&&&&&&&&&&1&1&1&1\\
&&&&&&&&&&&&&&-1&-2&-3\\
&&&&&&&&&&&&&&&1&3\\
&&&&&&&&&&&&&&&&-1\\
&&&&&&&&&&&&&&&&&1&1&1\\
&&&&&&&&&&&&&&&&&&-1&-2\\
&&&&&&&&&&&&&&&&&&&1\\
&&&&&&&&&&&&&&&&&&&&1&1\\
&&&&&&&&&&&&&&&&&&&&&-1\\
&&&&&&&&&&&&&&&&&&&&&&1
\end{array}\right]
\qquad
\left[\begin{array}{*{23}{r}}
1&&&&&&1&&&&&1&&&&1&&&1&&1&&1\\
&1&&&&&&1&&&&&1&&&&1&&&1&&1\\
&&1&&&&&&1&&&&&1&&&&1\\
&&&1&&&&&&1&&&&&1\\
&&&&1&&&&&&1\\
&&&&&1\\
&&&&&&-1&&&&&-2&&&&-3&&&-4&&-5&&-6\\
&&&&&&&-1&&&&&-2&&&&-3&&&-4&&-5\\
&&&&&&&&-1&&&&&-2&&&&-3\\
&&&&&&&&&-1&&&&&-2\\
&&&&&&&&&&-1\\
&&&&&&&&&&&1&&&&3&&&6&&10&&15\\
&&&&&&&&&&&&1&&&&3&&&6&&10\\
&&&&&&&&&&&&&1&&&&3\\
&&&&&&&&&&&&&&1\\
&&&&&&&&&&&&&&&-1&&&-4&&-10&&-20\\
&&&&&&&&&&&&&&&&-1&&&-4&&-10\\
&&&&&&&&&&&&&&&&&-1\\
&&&&&&&&&&&&&&&&&&1&&5&&15\\
&&&&&&&&&&&&&&&&&&&1&&5\\
&&&&&&&&&&&&&&&&&&&&-1&&-6\\
&&&&&&&&&&&&&&&&&&&&&-1\\
&&&&&&&&&&&&&&&&&&&&&&1
\end{array}\right]$$

6. Concluding remarks

6.1. General canonical forms. We may use theorem 3.4 to glue together, as a direct sum, the canonical forms obtained in sections 4 and 5, namely those in (4)-(5)-(6) and (24)-(25). For a given pair (L, R) of involutions, the similarity class of $A = LR$ is a simultaneous similarity invariant. Decompose the characteristic polynomial of A as

$$\chi_A = p_1^{e_1} \cdots p_r^{e_r} \, (\pi_1^*\pi_1)^{\varepsilon_1} \cdots (\pi_s^*\pi_s)^{\varepsilon_s} (x-1)^{n^+} (x+1)^{n^-}, \tag{26}$$

where the p_i are the distinct self-reciprocal prime factors of χ_A, $\deg p_i \geqslant 2$, and the $2s$ polynomials π_j, π_j^* are the distinct non-self-reciprocal prime monic factors of χ_A; and n^+ $[n^-]$ is the algebraic multiplicity of 1 [resp., -1] as an eigenvalue of LR. We let $n_1 \geqslant \cdots \geqslant n_w$ $[\overline{n}_1 \geqslant \cdots \geqslant \overline{n}_{\overline{w}}]$ be the sizes of the Jordan blocks of A with eigenvalue 1 [resp., -1]; clearly $n_1 + \cdots + n_w = n^+$ and $\overline{n}_1 + \cdots + \overline{n}_{\overline{w}} = n^-$; we assume there are v $[\overline{v}]$ distinct n_i's [resp., $\overline{n}_j$'s], and the corresponding multiplicities are denoted by $m_1, \dots, m_v$ [resp., $\overline{m}_1, \dots, \overline{m}_{\overline{v}}$].

Then we have canonical forms expressed as

$$\begin{aligned}(L, R) \approx &\underbrace{(L^{(1)}, R^{(1)})}_{n^{(1)}} \oplus \underbrace{(L^{(2)}, R^{(2)})}_{n^{(2)}} \oplus \underbrace{(L_\epsilon, R_\epsilon)}_{n^+} \oplus \underbrace{(L_\delta, -R_\delta)}_{n^-} \\ \approx &(L^{(1)}, R^{(1)}) \oplus (L^{(2)}, R^{(2)}) \oplus (S_\epsilon, V_\epsilon) \oplus (S_\delta, -V_\delta).\end{aligned} \tag{27}$$

Under the braces, the orders of the pairs are shown, with notations referring to (5) and (26), in particular $n^{(1)}$ and $n^{(2)}$ are the degrees of $p_1^{e_1} \cdots p_r^{e_r}$ and

$(\pi_1^*\pi_1)^{\varepsilon_1} \cdots (\pi_s^*\pi_s)^{\varepsilon_s}$, respectively. We have several choices for the canonical summands $(L^{(i)}, R^{(i)})$ as described in section 4, namely (4) and (6). The third summand (L_ϵ, R_ϵ) in (27) is like (24). The summand $(L_\delta, -R_\delta)$ in (27) is a direct sum

$$(L_\delta, -R_\delta) = \delta_1(\mathit{\Gamma}_{\overline{n}_1}, -\mathit{\Delta}_{\overline{n}_1}) \oplus \cdots \oplus \delta_{\overline{w}}(\mathit{\Gamma}_{\overline{n}_{\overline{w}}}, -\mathit{\Delta}_{\overline{n}_{\overline{w}}}),$$

where δ is a sign $\overline{w}$-tuple, adjusted with respect to $\overline{m}_1, \ldots, \overline{m}_{\overline{v}}$.

6.2. Orbit counting. Fix a matrix A as in the previous subsection. The simultaneous similarity orbit of a pair of involutions (L, R) such that $LR \sim A$ depends only on the choice of the sign tuples ϵ and δ in the canonical form (27). The number of distinct choices of adjusted ϵ and δ is

$$\prod_{i=1}^{v}(1 + m_i) \prod_{j=1}^{\overline{v}}(1 + \overline{m}_j).$$

Therefore, this is the number of simultaneous similarity orbits whose union is [generated by] the set of pairs of involutions (L, R) such that $LR \sim A$ [resp., such that $LR = A$].

If two involutions $L, R \in \mathfrak{M}_n$ are not similar, then -1 is an eigenvalue of LR. This follows easily from the fact that the eigenspace corresponding to the eigenvalue 1 of one of the involutions has a nontrivial intersection with the eigenspace corresponding to -1 of the other involution. The canonical form developed above yields a much stronger result: a complete characterization of the possible similarity classes of involutions L and R, such that $LR \in \mathcal{A}$ (or $LR = A$).

Theorem 6.1. *The locus of $(\mu(L), \mu(R))$ for pairs of involutions (L, R) such that $LR = A$ (or $LR \sim A$), is given by*

$$\begin{aligned} |\mu(L) + \mu(R) - n| &\leqslant n_o, & |\mu(L) - \mu(R)| &\leqslant \overline{n}_o, \\ \mu(L) + \mu(R) - n &\equiv_2 n_o, & \mu(L) - \mu(R) &\equiv_2 \overline{n}_o, \end{aligned} \tag{28}$$

where $\equiv_2$ denotes congruence modulo 2.

Proof. Let $\overline{\mu}(L)$ be the multiplicity of -1 as eigenvalue of L. Clearly $\mu(L) + \overline{\mu}(L) = n$. According to theorem 4.1, and using the notation of (27), we have

$\mu(L^{(k)}) - \overline{\mu}(L^{(k)}) = \mu(R^{(k)}) - \overline{\mu}(R^{(k)}) = 0$, $k = 1, 2$. It is easily seen that

$$\mu(\epsilon_i \mathit{\Gamma}_{n_i}) - \overline{\mu}(\epsilon_i \mathit{\Gamma}_{n_i}) = \begin{cases} 0 & \text{if } n_i \text{ is even} \\ \epsilon_i & \text{if } n_i \text{ is odd.} \end{cases}$$

$$\mu(\delta_j \mathit{\Gamma}_{\overline{n}_j}) - \overline{\mu}(\delta_j \mathit{\Gamma}_{\overline{n}_j}) = \begin{cases} 0 & \text{if } \overline{n}_j \text{ is even} \\ \delta_j & \text{if } \overline{n}_j \text{ is odd,} \end{cases}$$

for $1 \leqslant i \leqslant w$ and $1 \leqslant j \leqslant \overline{w}$, and the same is true with $\mathit{\Gamma}$ replaced by $\mathit{\Delta}$. Define $E = \sum\{\epsilon_i : \textit{odd}\ n_i\}$ and $D = \sum\{\delta_j : \textit{odd}\ \overline{n}_j\}$. The possible values for E and D are characterized by

$$|E| \leqslant n_o,\ E \equiv_2 n_o, \qquad |D| \leqslant \overline{n}_o,\ D \equiv_2 \overline{n}_o. \tag{29}$$

From (27) we get

$$\begin{aligned} \mu(L) - \overline{\mu}(L) &= E + D \\ \mu(R) - \overline{\mu}(R) &= E - D \\ \mu(L) + \overline{\mu}(L) &= \mu(R) + \overline{\mu}(R) = n \end{aligned}$$

Eliminating $\overline{\mu}(L)$ and $\overline{\mu}(R)$, we get $\mu(L)+\mu(R)-n = E$ and $\mu(L)-\mu(R) = D$; these conditions together with (29) characterize the locus of $(\mu(L), \mu(R))$, and are equivalent to (28). □

For the next counting, the v-tuple of multiplicities $\boldsymbol{m} = (m_1, \ldots, m_v)$ will be split into two, according to the parities of the sizes of the Jordan blocks. We let $\boldsymbol{m}_o = (m_{o1}, \ldots, m_{ov_o})$ and $\boldsymbol{m}_e = (m_{e1}, \ldots, m_{ev_e})$, where $\boldsymbol{m}_o$ $[\boldsymbol{m}_e]$ is the sub-tuple of $\boldsymbol{m}$ of the multiplicities of Jordan blocks of A with eigenvalue 1, of odd [resp., even] orders. The $\overline{v}$-tuple of multiplicities $\overline{\boldsymbol{m}} = (\overline{m}_1, \ldots, \overline{m}_{\overline{v}})$ (corresponding to Jordan blocks of A with eigenvalue -1) is split according to the same criteria into two: $\overline{\boldsymbol{m}}_o = (\overline{m}_{o1}, \ldots, \overline{m}_{o\overline{v}_o})$ and $\overline{\boldsymbol{m}}_e = (\overline{m}_{e1}, \ldots, \overline{m}_{e\overline{v}_e})$. Clearly $m_{o1} + \cdots + m_{ov_o} = n_o$ and $\overline{m}_{o1} + \cdots + \overline{m}_{o\overline{v}_o} = \overline{n}_o$.

For any r-tuple $\boldsymbol{q} = (q_1, \ldots, q_r)$, and any nonnegative integer S, define $N(\boldsymbol{q}, S)$ as the number of nonnegative integer solutions to the equation $x_1 + \cdots + x_r = S$, subject to the restrictions $x_i \leqslant q_i$, for $i = 1, \ldots, r$. It is not difficult to prove the following closed formula for this number:

$$N(\boldsymbol{q}, S) = \sum_{k=0}^{r} (-1)^k \sum_{1 \leqslant i_1 < \cdots < i_k \leqslant r} \binom{r + S - q_{i_1} - \cdots - q_{i_k} - k - 1}{r - 1}. \tag{30}$$

To get this one may use the inclusion-exclusion principle, as in [5, p. 138], or traditional generating functions techniques.

Theorem 6.2. *We are given a pair of integers,* (ℓ, r) *in the locus determined by theorem* 6.1*. The set of all pairs of involutions* (L, R) *such that* $LR \sim A$*,* $\mu(L) = \ell$ *and* $\mu(R) = r$ *is a union of*

$$N(\boldsymbol{m}_o, \tfrac{\ell+r-n+n_o}{2}) N(\overline{\boldsymbol{m}}_o, \tfrac{\ell-r+\overline{n}_o}{2}) \prod_{i=1}^{v_e}(1+m_{ei}) \prod_{j=1}^{\overline{v}_e}(1+\overline{m}_{ej}). \tag{31}$$

simultaneous similarity orbits.

Proof. The conditions (28) imply that the numbers

$$p = \tfrac{\ell+r-n+n_o}{2} \quad \text{and} \quad \overline{p} = \tfrac{\ell-r+\overline{n}_o}{2}$$

are integers such that $0 \leqslant p \leqslant n_o$ and $0 \leqslant \overline{p} \leqslant \overline{n}_o$. A closer look at the proof of theorem 6.1 shows that $E = p - (n_o - p)$ and $D = \overline{p} - (\overline{n}_o - \overline{p})$; therefore, to get $\mu(L) = \ell$ and $\mu(R) = r$, it is necessary and sufficient that p $[\overline{p}]$ be the number of positive ϵ_i's [resp., δ_j's] among the n_o [resp., $\overline{n}_o$] Jordan blocks of odd sizes with eigenvalue 1 [resp., -1].

To get all distinct (non simultaneously similar) canonical forms (27) under inspection, we are supposed

a) To assign plus signs to exactly p $[\overline{p}]$ among the n_o [resp., $\overline{n}_o$] Jordan blocks of odd sizes with eigenvalue 1 [resp., -1], in all possible *adjusted* ways. The number of ways of doing this is $N(\boldsymbol{m}_o, p)$ [resp., $N(\overline{\boldsymbol{m}}_o, \overline{p})$].

b) To assign any signs to the Jordan blocks of even sizes with eigenvalue 1 or -1, in all possible *adjusted* ways. The number of distinct ways of doing this is $\prod_{i=1}^{v_e}(1+m_{ei}) \prod_{j=1}^{\overline{v}_e}(1+\overline{m}_{ej})$.

Therefore, as these assignments are independent of each other, the number of simultaneous similarity classes described in the theorem is the product of the numbers in *a*)-*b*) above. Thus we get the formula (31). □

References

[1] C.S. Ballantine, Some involutory similarities, *Linear and Mult. Algebra*, 3(1975), pp. 19-23.

[2] G. R. Belitskiĭ, Normal forms in matrix spaces, *Integral Equat. Operator Theory*, 38(2000), pp. 251-283.

[3] G. R. Belitskiĭ and V.V. Sergeichuk, Complexity of matrix problems, *Linear Algebra Appl.*, 361(2003), pp. 203-222.

[4] V. M. Bondarenko, T. G. Gerasimova and V. V. Sergeichuk, Pairs of mutually annihilating operators, *Linear Algebra Appl.*, 430(2009), pp. 86-105.

[5] C. A. Charalambides, *Enumerative combinatorics*, Chapmam & Hall/CRC, Boca Raton, 2002.

[6] J. Dias da Silva and T. J. Laffey, Simultaneous similarity of matrices and related questions, *Linear Algebra Appl.*, 291(1999), pp. 167-184.

[7] D.Z. Djokovic, Product of Two Involutions, *Archiv Math. (Basel)*, 18(1967), pp. 582-584.

[8] D.Z. Djokovic, The product of two involutions in the unitary group of a hermitian form, *Indiana Univ. Math. J.*, 21(1971/1972), pp. 449456.

[9] S. Friedland, Simultaneous Similarity of Matrices, *Advances in Mathematics*, 50(1983), pp. 189-265

[10] F.R. Gantmacher, *The Theory of Matrices*, Chelsea Publishing Company, New York 1960.

[11] I.M. Gelfand and V.A. Ponomarev, Remarks on the classification of a pair of commuting linear transformations in a finite dimensional vector space, *Funct. Anal. Appl.*, 3(1969), pp. 325-326.

[12] F. Hoffman and E.C. Paige, Products of two involutions in the general linear group, *Indiana Univ. Math. J.*, 20(1970/1971), pp. 1017-1020.

[13] B. Hartley and T.O. Hawkes, *Rings, Modules and Linear Algebra*, UP Cambrodge, Chapman & Hall, London, 1970.

[14] F. Knüppel and K. Nielsen, On products of two involutions in the orthogonal group of a vector space, *Linear Algebra Appl.*, 94(1987), pp. 209216.

[15] T.J. Laffey, Factorization of matrices, involving symmetric matrices and involutions, in F. Uhlig and R. Grone (ed.s), *Current Trends in Matrix Theory*, North Holland 1987, pp. 175-198.

[16] C.S. Pazzis, Invariance of simultaneous similarity and equivalence of matrices under extension of the ground field, *Linear Algebra and its Applications*, 433(2010), pp. 618-624.

[17] S. Roman, *Field theory*, Graduate Texts in Mathematics, 158, Springer-Verlag, New York, 1995.

[18] E.M. Sá, Multiple Roots of Diagonal Multiples of a Square Matrix, *Discrete Mathematics*, 36(1981), pp. 57-67.

[19] H. Shapiro, The Weyr Characteristic, *Amer. Math. Monthly*, 106(1999), pp. 919-929.

[20] V.V. Sergeichuk, Canonical matrices for linear matrix problems, *Linear Algebra Appl.*, 317(2000), pp. 53-102.

[21] M.J. Wonenburger, Transformations which are Products of Two Involutions, *J. Math. Mech.* 16(1966), pp. 327-338.

[22] H. Zassenhaus, On a normal form of the orthogonal transformation, I, II, III, *Canad. Math. Bull.*, 1(1958), pp. 31-39, 101-111, 183-191.

Eduardo Marques de Sá
CMUC, Department of Mathematics, University of Coimbra, Apartado 3008, EC Santa Cruz 3001-501 Coimbra, Portugal

E-mail address: emsa@mat.uc.pt