ON TYPES OF MATRICES AND CENTRALIZERS OF MATRICES AND PERMUTATIONS

JOHN R. BRITNELL AND MARK WILDON

Abstract. It is known that that the centralizer of a matrix over a finite field depends, up to conjugacy, only on the type of the matrix, in the sense defined by J. A. Green. In this paper an analogue of the type invariant is defined that in general captures more information; using this invariant the result on centralizers is extended to arbitrary fields. The converse is also proved: thus two matrices have conjugate centralizers if and only if they have the same generalized type. The paper ends with the analogous results for symmetric and alternating groups.

1. Introduction

The notion of the type of a matrix over a finite field was defined by Green in his influential paper [2] on characters of finite general linear group, generalizing early work of Steinberg [4]. In Green's definition, the type of a matrix is obtained from its cycle type by formally replacing each irreducible polynomial with its degree. In [2, Lemma 2.1] Green showed that two matrices with the same type have isomorphic centralizer algebras. In [1, Theorem 2.7] the authors strengthened this result by proving that the centralizers are in fact conjugate. In this paper we generalize Green's definition of type to matrices over an arbitrary field, and prove the following theorem characterizing all matrices with conjugate centralizers.

Theorem 1.1. *Let K be a field and let X, $Y \in \mathrm{Mat}_n(K)$. The centralizers of X and Y in $\mathrm{Mat}_n(K)$ are conjugate by an element of $\mathrm{GL}_n(K)$ if and only if X and Y have the same generalized type.*

The definition of generalized type given in Section 2 below agrees with Green's for fields with the unique extension property; these include finite fields, and also algebraically closed fields. Thus an immediate corollary of Theorem 1.1 is that two matrices over a finite field have the same type if and only if their centralizers are conjugate. This gives the converse of Theorem 2.7 of [1].

The proof of Theorem 1.1 is given in Sections 4 and 5 below. In Section 4 we prove that two matrices with the same generalized type have conjugate centralizers. We obtain this result as a corollary of Theorem 4.3, which states that two matrices have the same generalized type if and only if their similarity classes contain representatives that are polynomial in one another.

2010 *Mathematics Subject Classification.* Primary 15A27; Secondary 15A21, 12F15, 20B35.

In Section 5 we prove the converse implication of Theorem 1.1, that if two matrices have conjugate centralizers then their generalized types agree. This requires a number of 'recognition' results on centralizers that build on the work in [1]. Some preliminary results needed in both parts of the proof are collected in Section 3.

An aspect of our work to which we would like to direct attention is our method, in the proof of Theorem 4.3, for dealing with a possibly inseparable field extension. This result is a generalization of [1, Theorem 2.6], but the proof of the earlier result depends on the existence of a Jordan–Chevalley decomposition, which can fail when the field is arbitrary. We avoid this problem by means of Lemma 4.2, which offers a dichotomy: if the minimal polynomial of a matrix X is a power of an irreducible polynomial, then either X has a Jordan–Chevalley decomposition, or else X possesses a very strong stability property under polynomial functions.

It is possible to make a similar statement about centralizers in symmetric groups, to the effect that permutations with conjugate centralizers have the same cycle type, except for certain 'edge cases'. It is clear that this result is directly analogous to Theorem 1.1, and since we have not found it in the literature, we have included it here. Section 6 contains this result (Theorem 6.2), and also the corresponding result for centralizers in alternating groups.

It is natural to ask whether the generalized type of a matrix is determined by the unit group of its centralizer. In the case of a matrix X over any field other than $\mathbf{F}_2$, the answer is that its type is indeed so determined; this follows from Theorem 1.1 via the observation that any element of the centralizer algebra of X is a sum of two units. For let $Y \in \mathrm{Cent}(X)$, and consider the primary decomposition of Y; define T to act as the identity on all but the unipotent summand of Y, and as any non-identity, non-zero scalar on that summand; then T and $Y - T$ are both units. Centralizers over the field $\mathbf{F}_2$ are not always generated linearly by their unit groups however, and for instance the centralizers of the two matrices

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$$

are distinct, although each has a trivial unit group.

2. Types and generalized types

Let K be a field, let $n \in \mathbf{N}$, let V be the K-vector space K^n, and let $X \in \mathrm{Mat}_n(K)$; we suppose throughout that matrices act on the right. Let $V = \bigoplus U_i$ be a decomposition of V as a sum of X-invariant subspaces, on each of which the action of X is indecomposable. Let X_i be X restricted to U_i. Then each X_i is a cyclic matrix and the minimum polynomial of X_i is f^t, for some polynomial f irreducible over K, and some positive integer t. For each such irreducible f, let λ_f be the partition obtained by collecting together the values of t arising in this way (counted with multiplicity). Although the decomposition of V is not in general unique, the partitions λ_f are invariants of X and collectively they determine X up to similarity of matrices.

Suppose that X is a matrix whose characteristic polynomial has the irreducible factors $f_1, \ldots, f_t$, with respective degrees $d_1, \ldots, d_t$, and that the partition invariants corresponding to these polynomials are $\lambda_1, \ldots \lambda_t$ respectively. The *cycle type* of X is the formal product $f_1^{\lambda_1} \cdots f_t^{\lambda_t}$. We say that a matrix over a field K is *primary* if it has cycle type f^λ for some irreducible polynomial f and partition λ. The *type* of X, as defined by Green in [2, page 407] is the formal product $d_1^{\lambda_1} \cdots d_t^{\lambda_t}$.

Green's definition of type makes sense when K is an arbitrary field. However Theorem 2.8 of [1], which states that matrices over a finite field with the same type have conjugate centralizers, would not extend to matrices over arbitrary fields if this definition were in force. To give an instance, let X and Y be the rational companion matrices of the irreducible polynomials $f(x) = x^2 - 2$ and $g(x) = x^2 - 3$. These matrices both have type $2^{(1)}$. Since X and Y are cyclic we have that $\operatorname{Cent} X = \mathbf{Q}\langle X\rangle$ and $\operatorname{Cent} Y = \mathbf{Q}\langle Y\rangle$. But X is not conjugate to a polynomial in Y, since the eigenvalues of X and Y lie in distinct quadratic extensions of $\mathbf{Q}$.

This example, however, suggests a very natural way of extending Green's definition which, as we shall show, allows the theorem we have mentioned to be generalized to infinite fields.

Definition 2.1. *Let K be a field, and let Φ be the set of irreducible polynomials over K. Let f, $g \in \Phi$ and let L be a splitting field for fg. We say that f is* equivalent *to g if whenever $\alpha \in L$ is a root of f there exists a root $\beta \in L$ of g such that $K(\alpha) = K(\beta)$, and vice versa. We denote equivalence by $f \sim g$, and denote the equivalence class of f by $[f]$.*

Since all splitting fields for fg are isomorphic as extensions of K, this definition does not depend on the choice of L.

Definition 2.2. *Let $X \in \operatorname{Mat}_d(K)$ and let Φ_X be the set of irreducible polynomials for which the partition invariant λ_f of X is non-empty. We define the* generalized type *of X to be the formal product*

$$\prod_{f \in \Phi_X} [f]^{\lambda_f}$$

in which the order of terms is unimportant.

We note that if K has the unique extension property (and in particular, if K is finite), then two polynomials are equivalent under $\sim$ if and only if they have the same degree. Our definition of generalized type therefore agrees with Green's in this case.

3. Preliminary results

We require two general results from [1]. For $d \in \mathbf{N}$, and for a partition λ, we write $d\lambda$ for the partition with d parts of size i for every part of size i in λ. For a partition λ we write $N(\lambda)$ for the similarity class of nilpotent matrices of type 1^λ. The dominance order on partitions will be denoted by $\trianglelefteq$.

Proposition 3.1 ([1, Proposition 2.2]). *Let M be a matrix of primary type d^λ. If the cycle type of M is f^λ then $f(M)$ is nilpotent and $f(M) \in N(d\lambda)$.*

Proposition 3.2 ([1, Proposition 2.4]). *Let X be a primary matrix of type d^λ with entries from a field K, and let $h \in K[x]$ be a polynomial. The type of $h(X)$ is e^μ for some e dividing d, and some partition μ such that $e|\mu| = d|\lambda|$ and $e\mu \trianglelefteq d\lambda$.*

We also need the following result giving the dimension of the centralizer of a matrix. If λ is a partition with exactly m_i parts of size i, we define

$$F(\lambda) = \sum_j \sum_k \min(j,k) m_j m_k.$$

Proposition 3.3. *Let K be a field and let $X \in \mathrm{Mat}_n(K)$ have type $d_1^{\lambda_1} \ldots d_t^{\lambda_t}$. Then $\dim_K \mathrm{Cent}\, X = \sum_{i=1}^t d_i F(\lambda_i)$.*

Proof. Let $V = K^n$. Since the subspaces corresponding to the primary decomposition of X are preserved by $\mathrm{Cent}\, X$, we may reduce to the case where X is a primary matrix of cycle type f^λ. Let the degree of f be d.

Given a vector $v \in V$ we say that v has *height* $h \in \mathbf{N}$ if $f(X)^{h-1}v \neq 0$ and $f(X)^h v = 0$. Let $V = \bigoplus_{i=1}^r U_i$ be a direct sum decomposition of V into indecomposable X-invariant subspaces such that the dimension of U_i is equal to the ith part of λ. Let u_i be a cyclic vector generating U_i. If h is a part of λ then the images of the m_h cyclic vectors of height h can be chosen freely from the subspace of V of vectors of height at most h. This subspace has dimension

$$d\big(h \sum_{j \geq h} m_j + \sum_{k<j} k m_k\big).$$

The proposition now follows by a straightforward counting argument. $\square$

As a corollary, we see that the dimension of the centralizer of a matrix depends on the field of definition only through the information captured by its type.

In the special case of nilpotent matrices this proposition is well known. For two equivalent formulations see Propositions 3.1.3 and 3.2.2 in [3]. The first implies that $F(\lambda) = \sum (2i-1)\ell_i$, where ℓ_i is the ith part of λ; the second, which is originally due to Frobenius, gives $F(\lambda) = \sum {\ell'_i}^2$, where ℓ'_i is the ith part of the conjugate partition to λ.

4. Matrices with conjugate centralizers

The aim of the remainder of this section is to prove Theorem 4.3 and hence the 'if' direction of Theorem 1.1.

Proposition 4.1. *Let X be nilpotent of class f^λ, where f has degree d and λ is a partition with at least one part of size greater than 1. Let $r(x)$ be a polynomial. Then $r(X) \in N(d\lambda)$ if and only if $r(x)$ is divisible by $f(x)$ but not by $f(x)^2$.*

Proof. It is clear that $r(X)$ is nilpotent if and only if $f(x)$ divides $r(x)$. Let $r(x) = g(x)f(x)^a$ where $g(x)$ is coprime to $f(x)$. Since $g(X)$ is invertible, and commutes with $f(X)^a$, we see that the dimensions of the kernels of $r(X)^i$ and $f(X)^{ai}$ are the same for all i. Since these dimensions determine the similarity class of X, it follows that $r(X)$ is similar to $f(X)^a$. By Proposition 3.1 we have $f(X) \in N(d\lambda)$. Hence if $a = 1$ then $r(X) \in N(d\lambda)$, while if $a > 1$ then $r(x) \notin N(d\lambda)$, since λ has a part of size greater than 1. □

Let X be a matrix over a field K. Recall that an additive Jordan–Chevalley decomposition of X is a decomposition $X = S + N$, where S and N are matrices over K such that S is semisimple, N is nilpotent, and $SN = NS$. If a Jordan–Chevalley decomposition of X exists then it is unique, and both S and N are polynomial in X. Over a perfect field, every matrix admits a Jordan–Chevalley decomposition, and the proof of [1, Theorem 2.6] (in which the field is finite) relies on this fact. Over an arbitrary field these decompositions do not generally exist; but the following lemma allows us to compensate for their lack.

Lemma 4.2. *Let X be a primary matrix over a field K of cycle type f^λ. Let r be a polynomial over K such that $r(X)$ has class f^μ. If $\mu \neq \lambda$, then X has a Jordan–Chevalley decomposition over K.*

Proof. If all parts of λ are equal to 1, then X is semisimple, and has an obvious Jordan–Chevalley decomposition. So we suppose that λ has a part greater than 1.

Let d be the degree of f. Since $r(X)$ has class f^μ, we see from Proposition 3.1 that $(f \circ r)(X)$ is nilpotent and lies in the similarity class $N(d\mu)$. It follows that f divides $f \circ r$. Let $f \circ r = gf$ for some polynomial g. If g is coprime with f, then by Proposition 4.1 we see that $gf(X)$ is the same nilpotent class as $f(X)$, and so we have $\mu = \lambda$.

Suppose, then, that g is divisible by f, and so $f \circ r = hf^2$ for some polynomial h. Observe that

$$f \circ (r \circ r) = (f \circ r) \circ r = hf^2 \circ r = (h \circ r)(f \circ r)^2 = (h \circ r)h^2 f^4.$$

Similarly, writing $r^{(a)}$ for the a-th power of r under composition, we see that $f \circ r^{(a)}$ is divisible by f^{2^a}. So for sufficiently large a, we have $(f \circ r^{(a)})(X) = 0$.

Let L be a splitting field for f over K. Notice that the polynomial r acts on the roots of f in L by permuting them, since these roots are the eigenvalues of both X and $r(X)$. We may suppose (by increasing a as necessary) that $r^{(a)}$ fixes each root of f. Then certainly $r^{(a)}(X) \neq 0$, and since $f\big(r^{(a)}(X)\big) = 0$, it follows that $S = r^{(a)}(X)$ is a semisimple matrix with minimum polynomial f. But since any eigenvector of X over L is an eigenvector of S with the same eigenvalue, we see that $N = X - S$ must be nilpotent. So we have found a Jordan–Chevalley decomposition $S + N$ for X. □

Theorem 4.3. *Let K be a field, and let X, $Y \in \mathrm{Mat}_d(K)$. Then X and Y have the same generalized type if and only if there exist polynomials p and q such that $p(X)$ is similar to Y and $q(Y)$ is similar to X.*

Proof. This is the generalization to an arbitrary field of [1, Theorem 2.6], and only part of the proof is complicated by the necessity of appealing to Lemma 4.2. We shall therefore present the unaffected parts of the argument very concisely, referring the reader to our earlier paper for a gentler exposition.

We show first that if X and Y have the same generalized type then there exists a polynomial p such that $p(X)$ is similar to Y. By an appeal to the Chinese Remainder Theorem, we see that it is enough to prove the result in the case that X a primary matrix of cycle type f^λ for some irreducible polynomial f and some partition λ. By hypothesis there exists an irreducible polynomial g with $f \sim g$, such that Y has cycle type g^λ.

Let α be a root of f in an extension field of K in which f and g split. Since $f \sim g$ there exists a root β of g and polynomials r and s over K such that $r(\alpha) = \beta$ and $s(\beta) = \alpha$. Now if α' is any root of f then, since α is sent to α' by an automorphism of L fixing K, we see that $r(\alpha')$ is a root of g and $s(r(\alpha')) = \alpha'$. It follows that $r(X)$ has class g^μ for some partition μ and $(s \circ r)(X)$ has class f^ν for some partition ν. Since $(s \circ r)(X)$ is polynomial in $r(X)$, it follows from Proposition 3.2 that $\lambda \trianglerighteq \mu \trianglerighteq \nu$.

Suppose that $\lambda = \nu$. Then the classes f^λ and g^λ are polynomial in one another, witnessed by the polynomials r and s.

Suppose, on the other hand, that $\nu \neq \lambda$. Then by Lemma 4.2, the matrix X has a Jordan–Chevalley decomposition $X = S + N$. It is now easy to see that $r(S) + N$ is the Jordan–Chevalley decomposition for some matrix Y' belonging to the class g^π for some partition π. Since both S and N are polynomials in X, we have that $r(S)+N$ is a polynomial in X. Similarly, we see that $(s \circ r)(S) + N$ is polynomial in Y'. Since $s \circ r$ fixes the eigenvalues of X we must have $(s \circ r)(S) = S$, and so X is polynomial in Y'. But now it follows from Proposition 3.2 that $\lambda = \pi$, and so the classes f^λ and g^λ are polynomial in one another in this case too.

Conversely, suppose that $p(X)$ is similar to Y and $q(Y)$ similar to X. Since the number of summands in the primary decomposition of $p(X)$ is at most the number in that of X, we see that the primary decomposition of X and Y have the same number of summands. Let X_f be the summand of X corresponding to the polynomial f, and let Y_g be the summand of Y similar to $p(X_f)$, corresponding to the polynomial g. Since p sends the eigenvalues of X (in a suitable extension field) to eigenvalues of Y, it is clear that $K(\alpha)$ embeds into $K(\beta)$. By symmetry we have $K(\alpha) = K(\beta)$ and so $f \sim g$. Now it follows from Proposition 3.2 that the partition invariants λ_f of X and λ_g of Y are the same. So X and Y have the same type. □

We now obtain one half of Theorem 1.1.

Proof of 'if' direction of Theorem 1.1. By Theorem 4.3 there exist polynomials p and q such that $p(X)$ is similar to Y and $q(Y)$ is similar to X. Now $\operatorname{Cent} X$ is a subalgebra of $\operatorname{Cent} p(X)$ and so $\operatorname{Cent} X$ is conjugate to a subalgebra of $\operatorname{Cent} Y$. Similarly $\operatorname{Cent} Y$ is a subalgebra of $\operatorname{Cent} q(Y)$, and so $\operatorname{Cent} Y$ is conjugate to a subalgebra of $\operatorname{Cent} X$. It follows from considering the dimensions of these subalgebras that $\operatorname{Cent} X = \operatorname{Cent} p(X)$ and that $\operatorname{Cent} Y = \operatorname{Cent} q(Y)$. □

5. Recognizing the generalized type of a matrix from its centralizer

Throughout this section we let K be a field. Let X and Y be matrices in $\mathrm{Mat}_n(K)$ with conjugate centralizer algebras. By replacing Y with an appropriate conjugate, we may assume that in fact $\operatorname{Cent} X$ and $\operatorname{Cent} Y$ are equal. We shall show that X and Y have the same generalized type.

The proof proceeds by a series of reductions. We first prove the result for nilpotent matrices, then for primary matrices, and finally, for general matrices.

Lemma 5.1. *If M and N are nilpotent matrices, and $\operatorname{Cent} M = \operatorname{Cent} N$, then M and N are conjugate by an element of $\mathrm{GL}_n(K)$.*

Proof. We use results from Section 3 of [1]. Let $A = \operatorname{Cent} M$. Let the partition associated with M have m_h parts of size h for each $h \in \mathbf{N}$. By Propositions 3.4 and 3.5 of [1], for each h such that $m_h > 0$, the A-module V has a composition factor of dimension m_h which appears with multiplicity h; these are all of the composition factors of V. Thus the similarity class of M can be recovered from a composition series for V. □

Lemma 5.2. *Suppose that X, $Y \in \mathrm{Mat}_n(K)$ have equal centralizers. Then the primary decompositions of V as a $K\langle X\rangle$-module and as an $K\langle Y\rangle$-module have the same subspaces of V of summands.*

Proof. Since X and Y commute we may form the simultaneous primary decomposition

$$V = \bigoplus_{f,g} V_{f,g}, \tag{$\star$}$$

where the direct sum is over pairs of irreducible polynomials in $K[x]$ and $V_{f,g}$ is the maximal subspace of V on which both $f(X)$ and $g(Y)$ have nilpotent restrictions. Suppose that V_{f,g_1} and V_{f,g_2} are both non-trivial, where g_1 and g_2 are distinct irreducible polynomials. Let v generate V_{f,g_1} as a $K\langle f(X)\rangle$-module and let w be a vector in the kernel of the restriction of $f(X)$ to V_{f,g_2}. There is a $K\langle X\rangle$-endomorphism of V that maps v to w. Such an endomorphism corresponds to matrix $Z \in \operatorname{Cent} X$ such that $V_{f,g_1}Z$ intersects non-trivially with V_{f,g_2}. On the other hand, no such Z can belong to $\operatorname{Cent} Y$; this contradicts the assumption that $\operatorname{Cent} X = \operatorname{Cent} Y$.

It follows that the decomposition ($\star$) is simply the primary decomposition of V as a $K\langle X\rangle$-module. The lemma follows by symmetry. □

To complete the proof in the primary case we need the following lemma and proposition describing how the type and centralizer algebra of a matrix change on field extensions. Given a partition λ, let $\lambda \times p$ denote the partition obtained by multiplying all of the parts of λ by p.

Lemma 5.3. *Suppose that K has prime characteristic p. Let $X \in \mathrm{Mat}_n(K)$ be a primary matrix of cycle type f^λ where $f(x^p) \in K[x]$ is an inseparable irreducible polynomial. Let L be an extension field of K containing the pth roots of the coefficients of f, and let $g \in L[x]$ be such that $g(x)^p = f(x^p)$. Then the cycle type of X over L is $g^{\lambda \times p}$.*

Proof. It is sufficient to prove the lemma when X is cyclic and so λ has a single part. Suppose that $\lambda = (h)$. Let $V = K^n$ regarded as a $K\langle X\rangle$-module. Since $V \cong K[x]/(f(x^p)^h)$, there is an isomorphismism of $L\langle X\rangle$-modules

$$V \otimes_K L \cong \frac{K[x]}{\langle f(x^p)^h\rangle} \otimes_K L \cong \frac{L[x]}{\langle f(x^p)^h\rangle} = \frac{L[x]}{\langle g(x)^{hp}\rangle}.$$

Hence $X \otimes 1$ acts as a cyclic matrix on $V \otimes_K L$ with minimal polynomial $g(x)^{hp}$. Therefore $X \otimes 1$ has cycle type $g^{(hp)}$, as required. □

Proposition 5.4. *Let $X \in \mathrm{Mat}_n(K)$ be a primary matrix of cycle type f^λ and let L be a splitting field for f. Under the isomorphism between $\mathrm{Mat}_n(L)$ and $\mathrm{Mat}_n(K) \otimes L$, the image of $\mathrm{Cent}_{\mathrm{Mat}_n(L)} X$ is $\mathrm{Cent}_{\mathrm{Mat}_n(K)} X \otimes 1$. Moreover if f has distinct roots $\alpha_1, \ldots, \alpha_d$ in L, where each root of f has multiplicity p^a, then the cycle type of X, regarded as an element of $\mathrm{Mat}_n(L)$, is*

$$(x-\alpha_1)^\lambda \ldots (x-\alpha_d)^\lambda$$

if f is separable, and

$$(x-\alpha_1)^{\lambda\times p^a} \ldots (x-\alpha_d)^{\lambda\times p^a}$$

if f is inseparable and each root of f in L has multiplicity p^a.

Proof. Clearly $\mathrm{Cent}_{\mathrm{Mat}_n(K)} X \otimes L$ is isomorphic to a subalgebra of $\mathrm{Cent}_{\mathrm{Mat}_n(L)} X$. We shall prove that the dimensions are the same, and at the same time establish the other claims in the proposition.

Suppose first of all that f is separable. Then f factors as $(x-\alpha_1)\ldots(x-\alpha_d)$ in $L[x]$. Since the α_i are conjugate by automorphisms of L fixing K, there is a partition μ such that, over L, the cycle type of X is $(x-\alpha_1)^\mu \ldots (x-\alpha_d)^\mu$. Therefore $f(X)$, regarded as a matrix over L, lies in the similarity class $N(d\mu)$. But by Proposition 3.1, we have $f(X) \in N(d\lambda)$, and so $\lambda = \mu$. Proposition 3.3 now implies that

$$\dim_L \mathrm{Cent}_{\mathrm{Mat}_n(L)} X = dF(\lambda) = \dim_K \mathrm{Cent}_{\mathrm{Mat}_n(K)} X.$$

Now suppose that f is inseparable. Let K have prime characteristic p and suppose that f factors as $(x-\alpha_1)^{p^a} \ldots (x-\alpha_d)^{p^a}$ where $a \geq 1$ and the α_i are distinct. Let $g(x) = (x-\alpha_1)\ldots(x-\alpha_d)$. Lemma 5.3 implies that the cycle type of X over the field extension of K generated by the coefficients of g is $g^{\lambda\times p^a}$. Since g is separable, it now follows that the cycle type of X over L is $(x-\alpha_1)^{\lambda\times p^a} \ldots (x-\alpha_d)^{\lambda\times p^a}$. Proposition 3.3 implies that

$$\dim_L \mathrm{Cent}_{\mathrm{Mat}_n(L)} X = dF(\lambda \times p^a) = dp^a F(\lambda) = \dim_K \mathrm{Cent}_{\mathrm{Mat}_n(K)} X,$$

again as required. □

Proposition 5.5. *Let f and g be irreducible polynomials over K. Let X, $Y \in \mathrm{Mat}_n(K)$ have cycle types f^λ and g^μ respectively, and suppose that* $\mathrm{Cent}\, X = \mathrm{Cent}\, Y$. *Then $f \sim g$ and $\lambda = \mu$.*

Proof. We shall work over a splitting field L for the product fg. By the first part of Proposition 5.4 the centralizers of X and Y in $\mathrm{Mat}_n(L)$ are equal.

Let f have distinct roots $\alpha_1, \ldots, \alpha_c$ and let g has distinct roots $\beta_1, \ldots, \beta_d$ in L. By Proposition 5.4 if K has characteristic zero then the cycle types of X and Y over L are respectively

$$(x - \alpha_1)^\lambda \cdots (x - \alpha_c)^\lambda,$$
$$(x - \beta_1)^\mu \cdots (x - \beta_d)^\mu,$$

while if K has prime characteristic p then there exists $a, b \in \mathbf{N}_0$ such that the cycle types are respectively

$$(x - \alpha_1)^{\lambda \times p^a} \cdots (x - \alpha_c)^{\lambda \times p^a},$$
$$(x - \beta_1)^{\mu \times p^b} \cdots (x - \beta_d)^{\mu \times p^b}.$$

Since X and Y have the same centralizer over L, it follows from Lemma 5.2 that their primary decompositions have the same number of summands, and so we have $c = d$ in both cases. Furthermore, the primary decompositions of X and Y over L have the same subspaces as summands. Let this decomposition be $\bigoplus V_i$ where X has the eigenvalue α_i and Y the eigenvalue β_i on V_i. Let X_i and Y_i denote the restrictions of X and Y to V_i, respectively. Then it is clear that

$$\mathrm{Cent}\, X = \bigoplus_i \mathrm{Cent}\, X_i, \quad \mathrm{Cent}\, Y = \bigoplus_i \mathrm{Cent}\, Y_i,$$

and since $\mathrm{Cent}\, X = \mathrm{Cent}\, Y$ it follows that $\mathrm{Cent}\, X_i = \mathrm{Cent}\, Y_i$ for all i. But $\mathrm{Cent}\, X_i = \mathrm{Cent}(X_i - \alpha_i I)$ and $\mathrm{Cent}\, Y_i = \mathrm{Cent}(Y_i - \beta_i I)$, and so the nilpotent matrices $X - \alpha_i I$ and $Y - \beta_i I$ have the same centralizer; by Lemma 5.1 they must be conjugate. In the separable case $X - \alpha_i I$ has the partition λ and $Y - \beta_i I$ has the partition μ, and so we have $\lambda = \mu$, as required. In the inseparable case $X - \alpha_i I$ has the partition $\lambda \times p^a$ and $Y - \beta_i I$ has the partition $\mu \times p^b$. Since $c = d$ the partitions λ and μ are partitions of the same number. Hence we have $a = b$ and so $\lambda = \mu$, as required.

It remains to show that $f \sim g$. For this we shall work over the original field K. Take $r \in \mathbf{N}$ such that $f(X)^{r-1} \neq 0$ and $f(X)^r = 0$. The action of X on $\mathrm{im}\, f(X)^{r-1}$ is semisimple, since it acts as a direct sum of copies of the irreducible companion matrix C of f. The X-endomorphisms of this subspace form a full matrix algebra with coefficients in $K\langle C \rangle$. The centre of this algebra consists of the diagonal matrices with coefficients in $K\langle C \rangle$. Therefore $\mathrm{Cent}_{\mathrm{Mat}_n(K)}\, X$ determines $K\langle C \rangle$. Hence we have $K\langle C \rangle = K\langle D \rangle$ where D is the companion matrix for g. It follows that if α is an eigenvalue of C then there is a polynomial $s \in K[x]$ such that $s(D)$ has α as an eigenvalue. But the eigenvalues of $S(D)$ are $\{s(\beta_1), ..., s(\beta_d)\}$ so $K(\alpha) = K(\beta_j)$ for some j. Therefore $f \sim g$. □

We are now ready to prove the other half of Theorem 1.1

Proof of 'only if' direction of Theorem 1.1. By Lemma 5.2 the primary decompositions of X and Y are the same. Let

$$V = \bigoplus_{i=1}^{t} V_i,$$

where for each i there exist irreducible polynomials f_i and g_i such that $f_1, \ldots, f_t$ are distinct, $g_1, \ldots, g_t$ are distinct, and both $f_i(X)$ and $g_i(Y)$ are nilpotent on their restriction to V_i. Now by Lemma 5.2, it follows that $\operatorname{Cent} X_i = \operatorname{Cent} Y_i$, where X_i and Y_i are the restrictions of X and Y to V_i. But then it follows from Proposition 5.5 that $f_i \sim g_i$ and that the partitions associated with these polynomials are equal. Therefore the generalized types of X and Y are the same. □

6. Centralizers in symmetric and alternating groups

Theorem 1.1 is analogous to a result for symmetric groups, which, since we have been unable to find it in the literature, we record here. Let g, h be elements of the symmetric group S_n of all permutations of $\{1, \ldots, n\}$. We write $g = v_1 \cdots v_n$, where v_i is the product of the cycles of g of length i. Similarly, we write $h = w_1 \cdots w_n$.

Definition 6.1.
(1) *If there exists k such that $w_i = v_i^k$, then we say that g and h are* locally equivalent *at i.*
(2) *We say that g and h are* equivalent *if they are locally equivalent at i for all $i \in \{1, 2, \ldots, n\}$.*
(3) *If $S \subseteq \{1, \ldots, n\}$ and if g and h are locally equivalent at all $i \notin S$, but not locally equivalent at $i \in S$, then we say that there is a* local variation *at S.*

Theorem 6.2. *Let g and h be elements of S_n whose centralizers in S_n are equal. Either g and h are equivalent, or there is a local variation at $\{1, 2\}$ described by one of the following statements:*
(1) *v_1v_2 is conjugate to (12) and w_1w_2 is simultaneously conjugate to $(1)(2)$, or vice versa.*
(2) *v_1v_2 is conjugate to $(12)(3)(4)$ and w_1w_2 is simultaneously conjugate to $(1)(2)(34)$, or vice versa.*

Proof. Let X_i be the support of v_i. Then

$$\operatorname{Cent}_{S_n}(g) \cong \bigoplus_i \operatorname{Cent}_{\operatorname{Sym}(X_i)}(v_i).$$

If g and h are locally equivalent at i, then the support of w_i is X_i, and clearly $\operatorname{Cent}_{\operatorname{Sym}(X_i)}(w_i) = \operatorname{Cent}_{\operatorname{Sym}(X_i)}(v_i)$. It follows easily that if g and h are equivalent, then their centralizers are equal.

For the converse, let G be the centralizer of g in S_n. Let $\alpha \in \{1 \ldots n\}$ be a point in X_i. Note that G permutes the orbits of g of length i transitively, as blocks for its action. Thus the orbit α^G is equal to X_i. Let G_α be the stabilizer of α in G. It is not hard to show that that G_α acts transitively on the points in the cycles of length i not containing α, and fixes the points lying in the same g-cycle as α. Thus the set F_α of fixed points of G_α consists precisely of the i points lying in the same g-cycle as α, *except* when $i = 1$ and g has exactly two fixed points. Therefore, when we attempt to reconstruct the orbits of g from the permutation action of G, the ambiguities arise precisely from the local variations in the statement of the theorem. Furthermore since (12) and $(1)(2)$ have the same centralizer in S_2, and since (12) and (34) have the same centralizer in S_4, there is no possibility of resolving these ambiguities.

We shall assume that we are not in this exceptional case. Suppose that g has j cycles of length i. We have seen that the set X_i is determined by the permutation action of G. We observe that G contains an element which acts as a full cycle c on X_i. Let g_i be the restriction of g to X_i. Since the centralizer of c in $\mathrm{Sym}(X_i)$ is the cyclic group $\langle c\rangle$, we see that $g_i = c^m$ for some m. Since c has order ij, it is clear that $m = jk$ for some k coprime with i.

Now if h is another permutation whose centralizer in S_n is G, and if h_i is the restriction of h to X_i, then we must similarly have that $h_i = c^{j\ell}$, where ℓ is coprime with i. Now since k and ℓ are invertible modulo i, we have $g_i = h_i^{k/\ell}$ and $h_i = g_i^{\ell/k}$. So g and h are locally equivalent at i as required. □

An obvious consequence of Theorem 6.2 is that if two elements x and y of S_n have centralizers which are isomorphic as permutation groups, then either x is conjugate to y, or else there is a unique transposition t such that t centralizes y and x is conjugate to ty. We remark that this conclusion does not hold if the centralizers of x and y are isomorphic merely as abstract groups. As an example, suppose that $n = 2k\ell + k + \ell - 1$, where k and ℓ are greater than 1, and such that k, ℓ, $2k - 1$ and $2\ell - 1$ are pairwise coprime. Let x and y be permutations such that x has cycles of lengths k, $2\ell - 1$ and $\ell(2k - 1)$, and y has cycles of lengths ℓ, $2k - 1$ and $k(2\ell - 1)$. Then x and y have no cycle lengths in common, but each has a centralizer that is cyclic of order $k\ell(2k - 1)(2\ell - 1)$.

Finally, it is worthwhile to state the analogous result to Theorem 6.2 for the alternating groups A_n. We shall not prove it here; the proof follows similar lines to that of Theorem 6.2, but is complicated slightly by the fact that centralizer G of an element g in A_n is not in general a direct product of permutation groups on the sets X_i, though it has index at most 2 in such a product: in fact the restriction of G to the set X_i acts either as $\mathrm{Cent}_{\mathrm{Alt}(X_i)}(v_i)$ or as $\mathrm{Cent}_{\mathrm{Sym}(X_i)}(v_i)$, depending on whether the cycles of g of length other than i have distinct odd lengths.

Theorem 6.3. *Let g and h be elements of A_n whose centralizers in A_n are equal. Then either g and h are equivalent, or one of the following statements is true.*

(1) *There is a local variation at $\{1, 2\}$, with v_1v_2 conjugate to $(12)(3)(4)$ and w_1w_2 simultaneously conjugate to $(1)(2)(34)$.*

(2) *There is a local variation at* $\{2\}$*, with* v_2 *being conjugate to* $(12)(34)$ *and* w_2 *simultaneously conjugate to* $(13)(24)$*. Elsewhere, each of* g *and* h *has only odd cycles of distinct lengths.*

(3) *There is a local variation at* $\{1, 3\}$*, with* $v_1 v_3$ *conjugate to* (123) *and* $w_1 w_3$ *simultaneously conjugate to* $(1)(2)(3)$*. Elsewhere, each of* g *and* h *has only odd cycles of distinct lengths.*

(4) *For some odd integer* m *there is a local variation at* $\{m\}$*, with* v_m *and* w_m *each having exactly two cycles. Each cycle of* w_m *is a power of a cycle of* v_m*, but the two exponents, taken modulo* i*, are distinct. Elsewhere, each of* g *and* h *has only odd cycles of distinct lengths.*

References

[1] John R. Britnell and Mark Wildon, 'On types and classes of commuting matrices over finite fields', *J. London Math. Soc.* 83 (2011) 470–492.

[2] J. A. Green, 'The characters of the finite general linear groups', *Trans. Amer. Math. Soc.* 80 (1955) 402–447.

[3] Kevin C. O'Meara, John Clark, and Charles I. Vinsonhaler, 'Advanced topics in linear algebra', Oxford University Press, Oxford, 2011.

[4] R. Steinberg, 'A geometric approach to the representations of the full linear group over a Galois field', *Trans. Amer. Math. Soc.* 71 (1951) 274–282.

Department of Mathematics, Imperial College London, London, SW7 2AZ
E-mail address: `j.britnell@imperial.ac.uk`

Department of Mathematics, Royal Holloway, University of London, Egham, Surrey TW20 0EX, United Kingdom
E-mail address: `mark.wildon@rhul.ac.uk`