Canad. Math. Bull. Vol. **60** (3), 2017 pp. 613–630
http://dx.doi.org/10.4153/CMB-2016-070-8

On the Dimension of the Locus of Determinantal Hypersurfaces

Zinovy Reichstein and Angelo Vistoli

Abstract. The characteristic polynomial $P_A(x_0,\dots,x_r)$ of an r-tuple $A := (A_1,\dots,A_r)$ of $n\times n$-matrices is defined as

$$P_A(x_0,\dots,x_r) := \det(x_0 I + x_1 A_1 + \cdots + x_r A_r).$$

We show that if $r \geqslant 3$ and $A := (A_1,\dots,A_r)$ is an r-tuple of $n\times n$-matrices in general position, then up to conjugacy, there are only finitely many r-tuples $A' := (A'_1,\dots,A'_r)$ such that $p_A = p_{A'}$. Equivalently, the locus of determinantal hypersurfaces of degree n in $\mathbf{P}^r$ is irreducible of dimension $(r-1)n^2+1$.

1 Introduction

Let $r, n \geqslant 2$ be integers and let k be a base field. Assume that $\mathrm{char}(k) = 0$ or $> n$. Given an r-tuple $A := (A_1,\dots,A_r) \in \mathrm{M}_n^r$ of $n\times n$-matrices, we define *the characteristic polynomial* of A as

$$P_A(x_0,\dots,x_r) := \det(x_0 I + x_1 A_1 + \cdots + x_r A_r),$$

where I denotes that $n\times n$ identity matrix. The purpose of this paper is to answer the following question, due to B. Reichstein.

Question 1.1 *For $(A_1,\dots,A_r)$ in general position in M_n^r, are there finitely many or infinitely many conjugacy classes of r-tuples $A' := (A'_1,\dots,A'_r)$ such that $p_A = p_{A'}$?*

To restate this question in geometric terms, consider the following diagram:

(1.1)
$$\begin{tikzcd}
\mathrm{M}_n^r \arrow[d, "\pi"] \arrow[dr, "P"] & & \\
Q_{r,n} = \mathrm{M}_n^r /\!\!/ \mathrm{PGL}_n \arrow[r, "\overline{P}"] & \mathrm{DHyp}_{r,n} \arrow[r, hook] & \mathrm{Hypersurf}_{r,n}\,.
\end{tikzcd}$$

Here

- $\mathrm{Hypersurf}_{r,n} \simeq \mathbb{P}^{\binom{r+n}{n}-1}$ denotes the space of degree n hypersurfaces in $\mathbb{P}^r$;
- $Q_{r,n} := \mathrm{M}_n^r /\!\!/ \mathrm{PGL}_n = \mathrm{Spec}\, k[\mathrm{M}_n^r]^{\mathrm{PGL}_n}$ denotes the categorical quotient space for the conjugation action of PGL_n on r-tuples of $n\times n$-matrices;
- π denotes the natural projection induced by the inclusion $k[\mathrm{M}_n^r]^{\mathrm{PGL}_n} \hookrightarrow k[\mathrm{M}_n^r]$;

Received by the editors May 24, 2016; revised September 26, 2016.
Published electronically November 11, 2016.
Z.R. supported in part by an NSERC Discovery grant.
A.V. supported in part by research funds from the Scuola Normale Superiore.
AMS subject classification: **14M12**, 15A22, 05A10.
Keywords: determinantal hypersurface, matrix invariant, q-binomial coefficient.

- P takes an r-tuple $A = (A_1, \dots, A_r)$ of $n \times n$ matrices to the hypersurface in $\mathbb{P}^r$ cut out by the homogeneous polynomial $P_A(x_0, \dots, x_r)$ of degree n. Hypersurfaces of this form are called *determinantal.*
- $\mathrm{DHyp}_{r,n}$ denotes the closure of the image of P in $\mathrm{Hypersurf}_{r,n}$; this is the *locus of determinantal hypersurfaces* of degree n in $\mathbb{P}^r$.

Question 1.2 *What is the dimension of* $\mathrm{DHyp}_{r,n}$*?*

Questions 1.1 and 1.2 are closely related. Indeed, Question 1.1 asks whether or not fibers of $\overline{P}$ in general position are finite, or equivalently, whether or not

$$\dim(\mathrm{DHyp}_{r,n}) = \dim(Q_{r,n}),$$

where

$$\dim(Q_{r,n}) = \dim(\mathrm{M}_n^r) - \dim(\mathrm{PGL}_n) = (r-1)n^2 + 1.$$

Our main result answers Questions 1.1 and 1.2 for $r \geqslant 3$.

Theorem 1.3 *Assume $r \geqslant 3$. Then the map $\overline{P}$ is generically finite and separable. In particular,* $\dim(\mathrm{DHyp}_{r,n}) = (r-1)n^2 + 1$, *for any $n \geqslant 2$.*

Several remarks are in order.

(1) A classical theorem of G. Frobenius [F1897, §7.1] asserts that the only linear transformations $T\colon \mathrm{M}_n \to \mathrm{M}_n$ preserving the determinant function are of the form $X \mapsto PXQ$ or $X \mapsto PX^tQ$, where X^t denotes the transpose of X, and P and Q are fixed $n \times n$ matrices such that $\det(P)\det(Q) = 1$. (For modern proofs of this theorem, further references, and generalizations, see [Dieu49], [MM59, Theorem 2], [Wat87, Theorem 4.2], [BGL14, Corollary 8.9].) In the case where $r = n^2 - 1$, Frobenius's theorem tells us that the fiber of $\overline{P}$ contains exactly two points corresponding to the conjugacy classes of $(A_1, \dots, A_r)$ and $(A_1^t, \dots, A_r^t)$, where A^t denotes the transpose of A; see Lemma 8.4. In Section 8 we will show that the same is true for any $r \geqslant n^2 - 1$.

(2) In the case where $n = r = 3$, Theorem 1.3 is equivalent to the following assertion: a general hypersurface of degree 3 in $\mathbb{P}^3$ is determinantal. Equivalently, the map $P\colon \mathrm{M}_3^3 \to \mathrm{Hypersurf}_{3,3} \simeq \mathbb{P}^{19}$ is dominant. This result goes back (at least) to H. Grassmann [G1855]; for a modern proof (in arbitrary characteristic), see [Beau00, Corollary 6.4].

(3) In the case where $r = 3$ and $n = 4$, Theorem 1.3 is equivalent to the assertion that determinantal quartic hypersurfaces in $\mathbb{P}^3$ form a codimension 1 locus in $\mathrm{Hypersurf}_{3,4} \simeq \mathbb{P}^{34}$. Over the field of complex numbers this is proved in [Dolg12, Example 4.2.23].

(4) We do not know what the degree of $\overline{P}$ is in general; our proof of Theorem 1.3 sheds no light on this question. As we mentioned above, if $r \geqslant n^2 - 1$, the general fiber of $\overline{P}$ consists of exactly two points corresponding to the conjugacy classes of $(A_1, \dots, A_r)$ and $(A_1^t, \dots, A_r^t)$ (see Theorem 8.2) and thus $\deg(\overline{P}) = 2$. An interesting (and to the best of our knowledge, open) question is whether or not $\deg(\overline{P}) = 2$ for every $n \geqslant 2$ and $r \geqslant 4$. Note however, that this fails for $r = 3$. Indeed, if $r = n = 3$, then $\deg(\overline{P}) = 72$; see [G1855], [Beau00, Corollary 6.4] or [Dolg12, Theorem 9.3.6].

https://doi.org/10.4153/CMB-2016-070-8 Published online by Cambridge University Press

(5) Theorem 1.3 fails for $r = 2$, as long as $n \geqslant 3$. Indeed, in this case,

$$\dim(Q_{2,n}) = n^2 + 1 > \binom{n+2}{2} - 1 = \dim(\mathrm{Hypersurf}_{2,n}),$$

so the fibers of $\overline{P}$ cannot be finite. In fact, this setting has been much studied, both from the theoretical point of view and in connection with applications to control theory. In particular, it is well known that the map $\overline{P}\colon Q_{2,n} \to \mathrm{Hypersurf}_{2,n}$ is dominant, and the points of the fiber of $\overline{P}$ over a general plane curve C of degree n are in a natural bijective correspondence with line bundles of degree $n(n-1)/2$ on C. For details and further references, see [CT79], [Vin86], [Beau00, Section 3], [Dolg12, Section 4.1], [Ne11].

(6) On the other hand, Theorem 1.3 remains true for $r = n = 2$. Indeed, in this case the k-algebra $k[Q_{2,2}] = k[\mathrm{M}_2^2]^{\mathrm{PGL}_2}$ is generated by five algebraically independent elements, $\mathrm{Tr}(A_1)$, $\mathrm{Tr}(A_2)$, $\det(A_1)$, $\det(A_2)$, and $\mathrm{Tr}(A_1A_2)$; see, [P67, Theorem 2.1], [H71, p. 20] or [FHL81, Lemma 1(1)]. One easily checks that these five elements lie in the k-algebra generated by the coefficients of $\det(x_0 I + x_1 A_1 + x_2 A_2)$. We conclude that for $r = n = 2$ the map $\overline{P}\colon \mathrm{M}_2^2 /\!\!/ \mathrm{PGL}_2 \to \mathrm{Hypersurf}_{2,2} \simeq \mathbb{P}^5$ is, in fact, a birational isomorphism, *i.e.,* $\deg(\overline{P}) = 1$. If $r, n \geqslant 2$ but $(n,r) \neq (2,2)$, then $(A_1, \dots, A_r)$ and $(A_1^t, \dots, A_r^t)$ are not conjugate for $(A_1, \dots, A_r) \in \mathrm{M}_n^r$ in general position (see, *e.g.,* [R93, Remark 1 on p. 73]), and hence, $\deg(\overline{P}) \geqslant 2$.

(7) The fact that $\overline{P}\colon \mathrm{M}_n^r \to \mathrm{Hypersurf}_{r,n}$ is dominant if and only if $r = 2$ or $r = n = 3$ was known to L. E. Dickson; see [Dickson21]. Dickson also noted that the determinantal form

$$\det(A_0 x_0 + \cdots + A_r x_r) \sum_{i_0 + \cdots + i_r = n} a_{i_0,\dots,i_r} x_0^{i_0} \dots x_r^{a_r}$$

"involves no more than $(r-1)n^2 + 2$ parameters", *i.e.,* the transcendence degree of the field generated by the coefficients $a_{i_1}, \dots, a_{i_r}$ over k is $\leqslant (r-1)n^2 + 2$; see [Dickson21, Theorem 6]. Our Theorem 1.3 implies that this bound is, in fact, attained for the generic determinantal form. [1]

Our standing assumption on the base field k is that $\mathrm{char}(k) = 0$ or $> n$. Among other things, this allows us to use Newton's formulas to express the coefficients of the characteristic polynomial of an $n \times n$-matrix X in terms of $\mathrm{Tr}(X), \mathrm{Tr}(X^2), \dots, \mathrm{Tr}(X^n)$. Our main results are of a geometric nature in the sense that in the course of proving them we can replace k by a larger field. In particular, we can usually assume without loss of generality that k is algebraically closed. We do not know to what extent Theorem 1.3 remains valid in the case where $0 < \mathrm{char}(k) \leqslant n$; our argument breaks down in this setting.

A modern approach to the study of determinantal hypersurfaces is based on the fact that a hypersurface $X \subset \mathbb{P}^n$ is determinantal if and only if X carries an Ulrich sheaf of rank 1; see [Beau00] in the case where X is smooth, and [ES03] in general. We have not been able to prove Theorem 1.3 using this approach, even though this may well be possible (one complication is that for $r > 3$ every determinantal hypersurface is singular). The proof we give here is entirely elementary.

[1] The reason for the discrepancy between $(r-1)n^2 + 2$ in Dickson's Theorem 6 and $(r-1)n^2 + 1$ in our Theorem 1.3 is that we take $A_0 = I$. The "extra" parameter in Dickson's setting is $\det(A_0)$.

https://doi.org/10.4153/CMB-2016-070-8 Published online by Cambridge University Press

2 A General Strategy for the Proof of Theorem 1.3

The first step is to reduce Theorem 1.3 to the case where $r = 3$. We will do this in Section 3; then assume that $r = 3$ for the rest of the proof. Clearly,

$$\dim(\mathrm{DHyp}_{3,n}) \leqslant \dim(Q_{3,n}) = 2n^2 + 1, \tag{2.1}$$

since the morphism $\overline{P}\colon Q_{3,n} \to \mathrm{DHyp}_{3,n}$ is dominant, by definition. The following lemma will supply a key ingredient for our proof of Theorem 1.3.

Lemma 2.1 *There exists a triple of $n \times n$ matrices $A = (A_1, A_2, A_3) \in \mathrm{M}_n^3$ such that the differential $dP_{|A}$ of P at A has rank $2n^2 + 1$.*

Once Lemma 2.1 is established, we know that $\operatorname{rank} dP_{|B} \geqslant 2n^2 + 1$ for $B \in \mathrm{M}_n^3$ is general position. Hence, (2.1) is an equality. Moreover, for $B \in \mathrm{M}_n^3$ in general position

$$\operatorname{rank} d\overline{P}_{|\pi(B)} \geqslant \operatorname{rank} dP_{|B} = 2n^2 + 1.$$

Since $\dim(Q_{3,n}) = \dim(\mathrm{DHyp}_{3,n}) = 2n^2 + 1$, we conclude that for $B \in \mathrm{M}_3^r$ in general position, $d\overline{P}_{|\pi(B)}$ is an isomorphism. In other words, $\overline{P}$ is generically finite and separable, as desired.

Our proof of Lemma 2.1 will be structured as follows. In Section 4 we will exhibit a homogeneous system of linear equations cutting out $\operatorname{Ker}(dP_{|A})$ inside the tangent space $T_A(\mathrm{M}_n^3)$ (which we identify with M_n^3) in Section 4. We will do this for any triple $A = (A_1, A_2, A_3) \in \mathrm{M}_n^3$ such that the linear span of A_1, A_2, and A_3 in M_n contains a matrix with distinct eigenvalues; see Lemma 4.1(ii). Our goal will be to prove Lemma 2.1 by showing that $\dim \operatorname{Ker}(dP_{|A}) = n^2 - 1$. The system of linear equations we obtain, cutting out $\operatorname{Ker}(dP_{|A})$ in M_n^3, is rather complicated (in particular, it is badly overdetermined). For this reason we have not been able to compute the dimension of $\operatorname{Ker}(dP_{|A})$ for an arbitrary triple $A = (A_1, A_2, A_3) \in M_n^3$ whose linear span contains a matrix with distinct eigenvalues. However, for the particular triple $A = (A_1, A_2, A_3)$ defined in (5.1), the kernel of $dP_{|A}$ carries a $(\mathbb{Z}/n\mathbb{Z})^2$-grading, *i.e.,* remains invariant under a certain linear action of the finite abelian group $G := (\mathbb{Z}/n\mathbb{Z})^2$ on M_n^3; see Section 6. This will allow us to decompose M_n^3 as a direct sum of n^2 three-dimensional character spaces, and verify that $\operatorname{Ker}(dP_{|A})$ has the desired dimension, $n^2 - 1$, by solving our linear system in each character space. This computation, completing the proof of Lemma 2.1 (and thus of Theorem 1.3), will be carried out in Sections 6 and 7. It relies on properties of q-binomial and trinomial coefficients, which are recalled in Section 5.

3 Reduction to the Case Where $r = 3$

Throughout this section, we will fix $n \geqslant 2$ and denote the map

$$\mathrm{M}_n^r /\!\!/ \mathrm{PGL}_n \to \mathrm{DHyp}_{r,n}$$

in diagram (1.1) by $\overline{P}(r, n)$.

Proposition 3.1 *Assume that $r \geqslant 3$. If the morphism $\overline{P}(r, n)$ is generically finite and separable, then so is $\overline{P}(r + 1, n)$.*

Let $K_{r,n} := k(\mathrm{M}_n^r)^{\mathrm{PGL}_n}$ be the field of rational functions on $\mathrm{M}_n^r /\!\!/ \mathrm{PGL}_n$ and $K'_{r,n}$ be the subfield of $K_{r,n}$ generated by the coefficients of the characteristic polynomial

$$(A_1, \dots, A_r) \longmapsto \det(x_0 I + x_1 A_1 + \dots + x_r A_r).$$

Clearly, $K'_{r,n}$ is the field of rational functions on $\mathrm{DHyp}_{r,n}$ and the inclusion of function fields $P^*: k(\mathrm{DHyp}_{r,n}) \hookrightarrow k(Q_{r,n})$ is the natural inclusion $K'_{r,n} \hookrightarrow K_{r,n}$. Thus, Proposition 3.1 can be restated, in purely algebraic terms, as follows.

Proposition 3.2 *Assume that $r \geqslant 3$. If the field extension $K_{r,n}/K'_{r,n}$ is finite and separable, then so is $K_{r+1,n}/K'_{r+1,n}$.*

The key to our proof of Proposition 3.2 is the following lemma, which asserts that $K_{r,n}$ is generated, as a field extension of k, by functions that depend on at most three of the matrices $A_1, \dots, A_r$.

Lemma 3.3 (C. Procesi) *Assume that $r \geqslant 3$. There are finitely many monomials $M_1, \dots, M_N$ in A_1 and A_2 such that $K_{r,n}$ is generated, as a field extension of k, by the elements $\mathrm{Tr}(M_i)$ and $\mathrm{Tr}(M_i A_j)$, where $i = 1, \dots, N$, and $j = 3, \dots, r$.*

Proof See [P67, Prop. 2.3, p. 255] or [FGG97, Thm. 3.2 and Ex. 3.3(a)]. ∎

Proof of Proposition 3.2 First observe that $K_{r,n} \subset K_{r+1,n}$ and $K'_{r,n} \subset K'_{r+1,n}$ (just set $A_{r+1} = 0$).

By Lemma 3.3, there exist finitely many monomials $M_1, \dots, M_N$ in A_1 and A_2 such that $K_{r+1,n}$ is generated, as a field extension of k, by $\mathrm{Tr}(M_i)$ and $\mathrm{Tr}(M_i A_j)$, where $i = 1, \dots, N$, and $j = 3, \dots, r+1$. It thus suffices to show that each of these elements is algebraic and separable over $K'_{r+1,n}$.

Let us start with $\mathrm{Tr}(M_i)$. By definition, $\mathrm{Tr}(M_i) \subset K_{2,n} \subset K_{r,n}$. By our assumption $\mathrm{Tr}(M_i)$ is thus algebraic and separable over $K'_{r,n}$. Since $K'_{r,n} \subset K'_{r+1,n}$, $\mathrm{Tr}(M_i)$ is algebraic and separable over $K'_{r+1,n}$, as desired.

Similarly, $\mathrm{Tr}(M_i A_3) \subset K_{3,n} \subset K_{r,n}$, since $r \geqslant 3$. By our assumption, $\mathrm{Tr}(M_i A_3)$ is algebraic and separable over $K'_{r,n}$. Hence, it is algebraic and separable over $K'_{r+1,n}$. By symmetry $\mathrm{Tr}(M_i A_j)$ is also algebraic and separable over $K'_{r+1,n}$ for every $j = 3, \dots, r+1$, and the proof of Proposition 3.2 is complete. ∎

4 The Kernel of dP

Observe that the image of the map P lies in the affine subspace $\mathbb{A}^{\binom{r+n}{n}-1}$ of $\mathbb{P}^{\binom{r+n}{n}-1} = \mathrm{Hypersurf}_{r,n}$ consisting of hypersurfaces of the form

$$\sum_{i_0+\dots+i_r=n} a_{i_1,\dots,i_r} x_0^{i_0} \dots x_r^{i_r} = 0,$$

where $a_{n,0,\dots,0} \neq 0$ (or equivalently, $a_{n,0,\dots,0} = 1$, after rescaling). Thus, we can view P as a polynomial map between the affine spaces M_n^r and $\mathbb{A}^{\binom{r+n}{n}-1}$. The differential $dP_{|A}$ at a point $A \in \mathrm{M}_n^r$ is a linear map $T_A(\mathrm{M}_n^r) \to T_A(\mathbb{A}^{\binom{r+n}{n}-1})$. We will identify $T_A(\mathrm{M}_n^r)$ with M_n^r and $T_A(\mathbb{A}^{\binom{r+n}{n}-1})$ with $\mathbb{A}^{\binom{r+n}{n}-1}$ in the obvious way.

Given an $n \times n$ matrix X, we will denote the classical adjoint of X by X^{ad}. Recall that X^{ad} is, by definition, the $n \times n$ matrix whose (i,j)-component is $(-1)^{i+j}\det(X_{ji})$, where X_{ji} is the $(n-1)\times(n-1)$ matrix obtained from X by deleting row j and column i. If X is invertible, then $X^{\mathrm{ad}} = \det(X)X^{-1}$.

Lemma 4.1 *Let $A = (A_1, \dots, A_r)$ be an r-tuple of $n \times n$-matrices.*

(i) *The differential $dP_{|A}$ sends $(B_1, \dots, B_r) \in T_A(\mathrm{M}_n^r) \simeq \mathrm{M}_n^r$ to*

$$\mathrm{Tr}\Big((x_0 I + x_1 A_1 + \cdots + x_r A_r)^{\mathrm{ad}} (x_1 B_1 + \cdots + x_r B_r) \Big).$$

(ii) *Suppose some matrix in the linear span of $A_1, \dots, A_r$ has distinct eigenvalues. Then the kernel of $dP_{|A}$ is the space of r-tuples $(B_1, \dots, B_r) \in \mathrm{M}_n^r$ satisfying*

$$\mathrm{Tr}\Big((x_1 A_1 + \cdots + x_r A_r)^d (x_1 B_1 + \cdots + x_r B_r) \Big) = 0$$

for every $d = 0, 1, \dots, n-1$.

In part (ii) we require that for every $d = 0, 1, \dots, n-1$, the left-hand side of the formula should be identically zero as a polynomial in $x_1, \dots, x_r$. This gives rise to a system of linear equations in $(B_1, \dots, B_r) \in \mathrm{M}_n^r$, whose solution space is $\mathrm{Ker}(dP_{|A})$.

Proof (i) Let $Y = (y_{ij})$ and $\Delta Y = (\Delta y_{ij})$ be $n \times n$ matrices. We think of the entries Δy_{ij} as being "small" and of the entries of Y as being constant. We claim that

$$\det(Y + \Delta Y) = \det(Y) + \mathrm{Tr}(Y^{\mathrm{ad}} \Delta Y) + (\text{terms of degree} \geqslant 2 \text{ in } \Delta y_{ij}). \tag{4.1}$$

In the special case where $Y = I$, (4.1) readily follows from the usual expansion of the characteristic polynomial of ΔY:

$$\det(I + \Delta Y) = 1 + \mathrm{Tr}(\Delta Y) + (\text{terms of degree} \geqslant 2 \text{ in } \Delta y_{ij}). \tag{4.2}$$

To prove the claim for arbitrary Y, note that both sides of (4.1) are $n \times n$-matrices, whose entries are polynomials in y_{ij} and Δy_{ij}. Hence, in order to establish (4.1) for an arbitrary $n \times n$ matrix Y, we can assume without loss of generality that Y is non-singular. In this case,

$$\det(Y + \Delta Y) = \det(Y)\det(I + Y^{-1}\Delta Y).$$

Expanding the second factor as in (4.2), we arrive at (4.1). This completes the proof of the claim.

In order to finish the proof of part (i), we will compute the directional derivative of P in the direction of $(B_1, \dots, B_r) \in \mathrm{M}_n^r$. Setting $Y := x_0 I + x_1 A_1 + \cdots + x_r A_r$ and $\Delta Y := (x_1 B_1 + \cdots + x_r B_r)h$, and applying (4.1), we see that

$$\begin{aligned} &P(A_1 + hB_1, \dots A_r + hB_r) \\ &\quad = \det(Y + \Delta Y) = \det(Y + \Delta Y) = \det(Y) + \mathrm{Tr}(Y^{\mathrm{ad}} \Delta Y)h + O(h^2) \\ &\quad = P(A_1, \dots, A_r) + \mathrm{Tr}\Big((x_0 I + x_1 A_1 + \cdots + x_r A_r)^{\mathrm{ad}} (x_1 B_1 + \cdots + x_r B_r) \Big) h + O(h^2). \end{aligned}$$

This shows that the directional derivative of P at A in the direction of B is

$$\mathrm{Tr}\Big((x_0 I + x_1 A_1 + \cdots + x_r A_r)^{\mathrm{ad}} (x_1 B_1 + \cdots + x_r B_r) \Big),$$

and part (i) follows. (Note that in the last computation $h \to 0$, but $x_0, x_1, \dots, x_n$ remain constant throughout.)

https://doi.org/10.4153/CMB-2016-070-8 Published online by Cambridge University Press

(ii) Let A be an $n \times n$ matrix with distinct eigenvalues, over a field K. We claim that $B \in \mathrm{M}_n$ satisfies

(a) $\operatorname{Tr}((x_0 I + A)^{\mathrm{ad}} B) = 0$ for every x_0

if and only if B satisfies

(b) $\operatorname{Tr}(A^d B) = 0$ for every $d = 0, \dots, n-1$.

Once this claim is established, we can deduce part (ii) from part (i) by setting $A := x_1 A_1 + \cdots + x_r A_r$ and $B := x_1 B_1 + \cdots + x_r B_r$ and working over the field $K = k(x_1, \dots, x_r)$.

To prove the claim, we can pass to the algebraic closure of K. By our assumption, A has distinct eigenvalues, and hence, is diagonalizable. We can thus assume without loss of generality that A is the diagonal matrix $\operatorname{diag}(\lambda_1, \dots, \lambda_n)$, where $\lambda_1, \dots, \lambda_n$ are distinct elements of K. Then

$$(tI + A)^{\mathrm{ad}} = \operatorname{diag}\Big(\frac{\Pi(t)}{t+\lambda_1}, \dots, \frac{\Pi(t)}{t+\lambda_n} \Big),$$

where $\Pi(t) = (t+\lambda_1)(t+\lambda_2)\dots(t+\lambda_n) = \det(tI + A)$ and each diagonal entry $\frac{\Pi(t)}{t+\lambda_i}$ is a polynomial of degree $n-1$ in t. Condition (a) now translates to

$$\sum_{i=1}^{n} b_{ii} \frac{\Pi(t)}{t+\lambda_i} = 0,$$

where $b_{11}, \dots, b_{nn}$ are the diagonal entries of B. Setting $t = -\lambda_i$, for $i = 1, \dots, n$, we obtain $b_{11} = b_{22} = \cdots = b_{nn} = 0$. On the other hand, condition (b) translates to

$$\sum_{i=1}^{n} \lambda_i^d b_{ii} = 0,$$

for each $d = 0, 1, \dots, n-1$, which we view as a homogeneous system of n linear equations in n unknowns $b_{11}, \dots, b_{nn}$. The matrix of this system is the Vandermonde matrix

$$\begin{pmatrix} 1 & 1 & \dots & 1 \\ \lambda_1 & \lambda_2 & \dots & \lambda_n \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_1^{n-1} & \lambda_2^{n-1} & \dots & \lambda_n^{n-1} \end{pmatrix}.$$

Since $\lambda_1, \dots, \lambda_n$ are distinct, this Vandermonde matrix is non-singular, and the above system has only the trivial solution, $b_{11} = b_{22} = \cdots = b_{nn} = 0$.

In summary, for $A = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$ both (a) and (b) are equivalent to $b_{11} = b_{22} = \cdots = b_{nn} = 0$. Hence, (a) and (b) are equivalent to each other. This completes the proof of the claim and thus of Lemma 4.1(ii). ∎

5 Skew-commuting Matrices and q-binomial Coefficients

Recall that we are working over a base field k of characteristic 0 or $> n$. For the sake of proving Theorem 1.3, we can assume without loss of generality that k is algebraically closed. In particular, we can assume that k contains a primitive n-th root of unity, which we will denote by q. We will also assume that $r = 3$; see Proposition 3.1(i). For the remainder of the proof of Theorem 1.3, we will set

(5.1)
$$A_1 := \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & q & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & q^{n-1} \end{pmatrix}, \quad A_2 := \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 1 & 0 & 0 & \dots & 1 \end{pmatrix}, \quad \text{and} \quad A_3 := A_1 A_2.$$

It is easy to see that

$$A_2 A_1 = q A_1 A_2 \quad \text{and} \quad A_1^n = A_2^n = I,$$

where, as usual, I denotes that $n \times n$-identity matrix. Hence, conjugation by A_1 commutes with conjugation by A_2; we will denote these commuting linear operators by Conj_{A_1} and $\mathrm{Conj}_{A_2}: \mathrm{M}_n \to \mathrm{M}_n$, respectively. They generate a subgroup of $\mathrm{GL}(\mathrm{M}_n)$ isomorphic to $(\mathbb{Z}/n\mathbb{Z})^2$. One readily checks that

$$\mathrm{Conj}_{A_1}(A_1^{e_1} A_2^{e_2}) = q^{-e_2} A_1^{e_1} A_2^{e_2} \quad \text{and} \quad \mathrm{Conj}_{A_2}(A_1^{e_1} A_2^{e_2}) = q^{e_1} A_1^{e_1} A_2^{e_2}.$$

In particular,

$$\mathrm{Tr}(A_1^{e_1} A_2^{e_2}) = \begin{cases} n & \text{if } e_1 \equiv e_2 \equiv 0 \pmod{n}, \\ 0 & \text{otherwise.} \end{cases} \tag{5.2}$$

Letting e_1 and e_2 range over $\mathbb{Z}/n\mathbb{Z}$, we see that each of the n^2 one-dimensional subspaces $\mathrm{Span}_k(A_1^{e_1} A_2^{e_2})$ is a character space for the abelian group

$$\langle \mathrm{Conj}_{A_1}, \mathrm{Conj}_{A_2} \rangle \simeq (\mathbb{Z}/n\mathbb{Z})^2.$$

Since these spaces have distinct associated characters, the matrices $A_1^{e_1} A_2^{e_2}$ form a k-basis of M_n, as e_1 and e_2 range over $\mathbb{Z}/n\mathbb{Z}$. In the sequel it will often be more convenient for us to work in this basis than in the standard basis of M_n, consisting of elementary matrices.

We now recall that the *q-factorial* $[d]_q!$ of an integer $d \geqslant 0$ is given by

$$[d]_q! := [1]_q [2]_q \cdots [d]_q,$$

where $[a]_q := (1 - q^a)/(1-q) = 1 + q + \cdots + q^{a-1}$. In particular, $[0]_q! = 1$. (Recall that we are assuming that $n \geqslant 2$ throughout, and thus $q \neq 1$.) If a and b are non-negative integers and $a + b = d \leqslant n - 1$, then

$$\binom{d}{a, b}_q := \frac{[d]_q!}{[a]_q! [b]_q!} \tag{5.3}$$

is called a *q-binomial coefficient*. If $a < 0$ or $b < 0$, we set

$$\binom{d}{a, b}_q := 0.$$

Similarly, if $a + b + c = d \leqslant n - 1$, then

$$\binom{d}{a, b, c}_q := \begin{cases} \dfrac{[d]_q!}{[a]_q! [b]_q! [c]_q!} & \text{if } a, b, c \geqslant 0, \\ 0 & \text{otherwise.} \end{cases} \tag{5.4}$$

is called a *q-trinomial coefficient*. This terminology is justified by parts (i) and (ii) of the following lemma. Part (iii) will play an important role in the sequel.

Lemma 5.1 *Assume $d = 0, \dots, n-1$.*

(i) *Let X and Y be matrices such that $XY = qYX$. Then*

$$(X+Y)^d = \sum_{a+b=d} \binom{d}{a,b}_q X^a Y^b.$$

(ii) *Let A_1 and A_2 be as in (5.1). Then*

$$(x_1A_1 + x_2A_2 + x_3A_1A_2)^d = \sum_{a+b+c=d} q^{\frac{c(c-1)}{2}} \binom{d}{a,b,c}_q x_1^a x_2^b x_3^c A_1^{a+c} A_2^{b+c}.$$

(iii) *For any $e_1, e_2 \in \mathbb{Z}/n\mathbb{Z}$,*

$$\operatorname{Tr}\Big((x_1A_1 + x_2A_2 + x_3A_1A_2)^d A_1^{e_1} A_2^{e_2} \Big) = n \sum_{a,b,c} q^{e_1(b+c)+\frac{c(c-1)}{2}} \binom{d}{a,b,c}_q x_1^a x_2^b x_3^c,$$

where the sum ranges over triples of non-negative integers (a,b,c), subject to the following conditions: $a+b+c=d$, $a+c+e_1 \equiv 0 \pmod{n}$, and $b+c+e_2 \equiv 0 \pmod{n}$.

Proof The binomial formula in part (i) was proved by M. P. Schützenberger [Sch53]; for a detailed discussion of this formula and further references, see [HMS04].

(ii) We apply part (i) twice. First we set $X = x_1A_1 + x_3A_1A_2$ and $Y := x_2A_2$ to obtain

$$(x_1A_1 + x_2A_2 + x_3A_1A_2)^d = \sum_{i+j=d} \binom{d}{i,j}_q (x_1A_1 + x_3A_1A_2)^i x_2^j A_2^j. \tag{5.5}$$

Next we apply part (i) with $X := x_1A_1$ and $Y := x_3A_1A_2$:

$$(x_1A_1 + x_3A_1A_2)^i = \sum_{a+c=i} \binom{i}{a,c}_q x_1^a x_3^c A_1^a (A_1A_2)^c. \tag{5.6}$$

Substituting (5.6) into (5.5), setting $i := a+c$ and $b := j$, and using the identities

$$\binom{d}{a,b,c}_q = \binom{d}{i,b}_q \binom{i}{a,c}_q, \tag{5.7}$$

$$(A_1A_2)^c = q^{\frac{c(c-1)}{2}} A_1^c A_2^c, \tag{5.8}$$

we obtain the formula in part (ii). Note that (5.7) is an immediate consequence of the definitions (5.3) and (5.4), and (5.8) follows from $A_2A_1 = qA_1A_2$.

To deduce part (iii) from part (ii), multiply both sides of (ii) by $A_1^{e_1}A_2^{e_2}$, rewrite $A_2^{b+c}A_1^{e_1}$ as $q^{e_1(b+c)}A_1^{e_1}A_2^{b+c}$, and take the trace on both sides. The desired equality now follows from (5.2). ∎

For future reference we record a simple identity involving q-trinomial coefficients.

Lemma 5.2 *Suppose α, β, and γ are integers, $0 \leqslant \alpha, \beta, \gamma \leqslant n-1$, and $1 \leqslant \alpha+\beta+\gamma \leqslant n$. Set $d := \alpha + \beta + \gamma - 1$. Then*

$$\left(\binom{d}{\alpha - 1, \beta, \gamma}_q : \binom{d}{\alpha, \beta - 1, \gamma}_q : \binom{d}{\alpha, \beta, \gamma - 1}_q \right) = (1 - q^\alpha : 1 - q^\beta : 1 - q^\gamma)$$

as points in the projective plane $\mathbb{P}^2$.

Proof If $\alpha, \beta, \gamma > 0$, the lemma is obtained by multiplying each of the numbers

$$\binom{d}{\alpha - 1, \beta, \gamma}_q, \quad \binom{d}{\alpha, \beta - 1, \gamma}_q, \quad \text{and} \quad \binom{d}{\alpha, \beta, \gamma - 1}_q$$

by the non-zero scalar $(1 - q)\frac{[\alpha]_q![\beta]_q![\gamma]_q!}{[d]_q!} \in k$. If one of the integers α, β, γ is 0, say, $\alpha = 0$, then

$$\binom{d}{\alpha - 1, \beta, \gamma}_q = 1 - q^\alpha = 0,$$

and the lemma follows. ■

6 A Grading of $\operatorname{Ker}(dP_{|A})$

Let A_1, A_2 and $A_3 = A_1 A_2$ be as in (5.1). Let $V := \operatorname{Ker}(dP_{|A}) \subset \mathrm{M}_n^3$, where the map $P\colon \mathrm{M}_n^3 \to \mathrm{Hypersurf}_{3,n}$ is defined in the introduction. Since A_1 has distinct eigenvalues, Lemma 4.1(ii) tells us that $V \subset \mathrm{M}_n^3$ consists of triples (B_1, B_2, B_3) satisfying

$$\operatorname{Tr}\big((x_1 A_1 + x_2 A_2 + x_3 A_1 A_2)^d (x_1 B_1 + x_2 B_2 + x_3 B_3) \big) = 0$$

for $d = 0, 1, \dots, n - 1$. Here, the left-hand side is required to be zero as a polynomial in x_1, x_2, x_3, for every $d = 0, 1, \dots, n - 1$.

Following the strategy outlined in Section 2, in order to complete the proof of Theorem 1.3 (or equivalently, of Lemma 2.1), it suffices to show that $\dim(V) = n^2 - 1$.

Lemma 6.1 *V is invariant under the linear action of the finite abelian group $(\mathbb{Z}/n\mathbb{Z})^2 = \langle \tau, \sigma \rangle$ on M_n^3 given by*

$$\sigma\colon (B_1, B_2, B_3) \longmapsto \big(\operatorname{Conj}_{A_1}(B_1), q \operatorname{Conj}_{A_1}(B_2), q \operatorname{Conj}_{A_1}(B_3) \big),$$
$$\tau\colon (B_1, B_2, B_3) \longmapsto \big(q^{-1} \operatorname{Conj}_{A_2}(B_1), \operatorname{Conj}_{A_2}(B_2), q^{-1} \operatorname{Conj}_{A_2}(B_3) \big).$$

Proof Suppose $(B_1, B_2, B_3) \in V$, *i.e.*,

$$f_{B_1, B_2, B_3, d}(x_1, x_2, x_3) := \operatorname{Tr}\big((x_1 A_1 + x_2 A_2 + x_3 A_1 A_2)^d (x_1 B_1 + x_2 B_2 + x_3 B_3) \big) = 0$$

for every $d = 0, \dots, n-1$. Here, $f_{B_1, B_2, B_3, d}$ is a polynomial in x_1, x_2, x_3 with coefficients in k, and $f_{B_1, B_2, B_3, d}(x_1, x_2, x_3) = 0$ means that $f_{B_1, B_2, B_3, d}$ is the zero polynomial, *i.e.*, every coefficient vanishes. Let

$$(C_1, C_2, C_3) := \sigma(B_1, B_2, B_3) = \big(\operatorname{Conj}_{A_1}(B_1), q \operatorname{Conj}_{A_1}(B_2), q \operatorname{Conj}_{A_1}(B_3) \big),$$

as above. To prove that V is invariant under σ, we need to show that $(C_1, C_2, C_3) \in V$, i.e., $f_{C_1, C_2, C_3, d}$ is identically 0 for every $d = 0, 1, \dots, n - 1$. Keeping in mind that

$$A_1 := \operatorname{Conj} A_1(A_1), \quad A_2 := q \operatorname{Conj}_{A_1}(A_2), \quad \text{and} \quad A_1 A_2 := q \operatorname{Conj}_{A_1}(A_1 A_2),$$

we see that

$$\begin{aligned} 0 &= f_{B_1,B_2,B_3,d}(x_1,x_2,x_3) \\ &= \operatorname{Tr}\big(\operatorname{Conj}_{A_1}(x_1A_1+x_2A_2+x_3A_1A_2)^d(x_1B_1+x_2B_2+x_3B_3)\big) \\ &= \operatorname{Tr}\big((x_1A_1+x_2q^{-1}A_2+x_3q^{-1}A_1A_2)^d(x_1C_1+x_2q^{-1}C_2+x_3q^{-1}C_3)\big) \\ &= f_{C_1,C_2,C_3,d}(x_1,q^{-1}x_2,q^{-1}x_3). \end{aligned}$$

This shows that $f_{C_1,C_2,C_3,d}(x_1,q^{-1}x_2,q^{-1}x_3)$ is identically zero as a polynomial in x_1, x_2, x_3. Hence, so is $f_{C_1,C_2,C_3,d}(x_1,x_2,x_3)$, as desired.

A similar argument shows that V is invariant under τ. (Here we conjugate by A_2, rather than A_1.) This completes the proof of Lemma 6.1. ∎

Since we are working over an algebraically closed base field k and $\operatorname{char}(k)=0$ or $\operatorname{char}(k)>n$, Lemma 6.1 tells us that V is a direct sum of character spaces for the action of $(\mathbb{Z}/n\mathbb{Z})^2$ on M_n^3. There are n^2 character spaces, each of dimension 3 (one for each character of $(\mathbb{Z}/n\mathbb{Z})^2$). They are defined as follows:

$$W_{e_1,e_2} := \big\{(t_1A_1^{e_1+1}A_2^{e_2}, t_2A_1^{e_1}A_2^{e_2+1}, t_3A_1^{e_1+1}A_2^{e_2+1}) \mid t_1,t_2,t_3\in k\big\},$$

where $(e_1,e_2)\in(\mathbb{Z}/n\mathbb{Z})^2$. Here, σ multiplies every vector in W_{e_1,e_2} by q^{-e_2} and τ by q^{e_1}. In other words, $(\mathbb{Z}/n\mathbb{Z})^2$ acts on W_{e_1,e_2} by the character

$$\chi\colon \sigma^a\tau^b \longmapsto q^{-e_2a+e_1b}.$$

In summary, $V=\bigoplus_{e_1,e_2=0}^{n-1}V_{e_1,e_2}$, where $V_{e_1,e_2}:=V\cap W_{e_1,e_2}$. Recall that our goal is to show that $\dim(V)=n^2-1$. Thus, in order to prove Theorem 1.3, it suffices to establish the following proposition.

Proposition 6.2

(i) $V_{0,0}=(0)$.

(ii) $\dim(V_{e_1,e_2})=1$ *for any* $(0,0)\neq(e_1,e_2)\in(\mathbb{Z}/n\mathbb{Z})^2$.

Proposition 6.2 will be proved in the next section.

Remark 6.3 If X and Y are $n\times n$-matrices, then clearly $\operatorname{Tr}(X^d[X,Y])=0$ for every $d\geqslant 0$. Setting $X=x_1A_1+x_2A_2+x_3A_1A_2$, $Y=A_1^{e_1}A_2^{e_2}$, and thus

$$[X,Y]=x_1(1-q^{e_2})A_1^{e_1+1}A_2+x_2(q^{e_1}-1)A_1^{e_1}A_2^{e_2+1}+x_3(q^{e_1}-q^{e_2})A_1^{e_1+1}A_2^{e_2+1},$$

we see that the triple

$$(B_1,B_2,B_3)=\big((1-q^{e_2})A_1^{e_1+1}A_2^{e_2},(q^{e_1}-1)A_1^{e_1}A_2^{e_2+1},(q^{e_1}-q^{e_2})A_1^{e_1+1}A_2^{e_2+1}\big)$$

lies in V_{e_1,e_2}. Here, $(B_1,B_2,B_3)=(0,0,0)$ if $(e_1,e_2)=(0,0)$ in $(\mathbb{Z}/n\mathbb{Z})^2$ and $(B_1,B_2,B_3)\neq(0,0,0)$ otherwise. Proposition 6.2 tells us that, in fact, (B_1,B_2,B_3) spans V_{e_1,e_2} for every $(e_1,e_2)\in(\mathbb{Z}/n\mathbb{Z})^2$.

https://doi.org/10.4153/CMB-2016-070-8 Published online by Cambridge University Press

7 Conclusion of the Proof of Theorem 1.3

It remains to prove Proposition 6.2. Given $t_1, t_2, t_3 \in k$, recall that an element

$$w := (t_1 A_1^{e_1+1} A_2^{e_2}, t_2 A_1^{e_1} A_2^{e_2+1}, t_3 A_1^{e_1+1} A_2^{e_2+1})$$

of W_{e_1,e_2} lies in V_{e_1,e_2} if and only if

$$\operatorname{Tr}\big((x_1 A_1 + x_2 A_2 + x_3 A_1 A_2)^d (t_1 x_1 A_1^{e_1+1} A_2^{e_2} + t_2 x_2 A_1^{e_1} A_2^{e_2+1} + t_3 x_3 A_1^{e_1+1} A_2^{e_2+1})\big)$$

is identically 0 as a polynomial in x_1, x_2, x_3, for every $d = 0, \dots, n-1$. Rewriting this polynomial as

$$\begin{aligned} &t_1 x_1 \operatorname{Tr}\big((x_1 A_1 + x_2 A_2 + x_3 A_1 A_2)^d A_1^{e_1+1} A_2^{e_2}\big) \\ &\quad + t_2 x_2 \operatorname{Tr}\big((x_1 A_1 + x_2 A_2 + x_3 A_1 A_2)^d A_1^{e_1} A_2^{e_2+1}\big) \\ &\quad + t_3 x_3 \operatorname{Tr}\big((x_1 A_1 + x_2 A_2 + x_3 A_1 A_2)^d A_1^{e_1+1} A_2^{e_2+1}\big) \end{aligned}$$

and applying Lemma 5.1(iii) to each term, we obtain

$$\begin{aligned} (7.1)\qquad & t_1 \sum_{(a,b,c)} n q^{(e_1+1)(b+c)+\frac{c(c-1)}{2}} \binom{d}{a,b,c}_q x_1^{a+1} x_2^b x_3^c \\ &+ t_2 \sum_{(a',b',c')} n q^{e_1(b'+c')+\frac{c'(c'-1)}{2}} \binom{d}{a',b',c'}_q x_1^{a'+1} x_2^{b'} x_3^{c'} \\ &+ t_3 \sum_{(a'',b'',c'')} n q^{(e_1+1)(b''+c'')+\frac{c''(c''-1)}{2}} \binom{d}{a'',b'',c''} x_1^{a''} x_2^{b''} x_3^{c''+1} = 0, \end{aligned}$$

where the sums are taken over triples of non-negative integers (a,b,c), (a',b',c'), and (a'',b'',c'') satisfying

$$\begin{array}{ccc} a+b+c=d, & a'+b'+c'=d, & a''+b''+c''=d, \\ a+c+e_1+1 \equiv 0 \pmod n, & a'+c'+e_1 \equiv 0 \pmod n, & a''+c''+e_1+1 \equiv 0 \pmod n, \\ b+c+e_2 \equiv 0 \pmod n, & b'+c'+e_2+1 \equiv 0 \pmod n, & b''+c''+e_2+1 \equiv 0 \pmod n. \end{array}$$

The expression on the left hand side of (7.1) is a homogeneous polynomial in x_1, x_2, x_3 of degree $d+1$. Our element $w = (t_1 A_1^{e+1} A_2^{e_2}, t_2 A_1^{e_1} A_2^{e_2+1}, t_3 A_1^{e_1+1} A_2^{e_2+1})$ of W_{e_1,e_2} lies in V_{e_1,e_2} if and only if this polynomial is identically zero.

To make the conditions the vanishing of this polynomial imposes on t_1, t_2, t_3 more explicit, let us examine the coefficient of $x_1^\alpha x_2^\beta x_3^\gamma$ (with $d+1 = \alpha+\beta+\gamma$). This coefficient is zero unless α, β, and γ are chosen so that

$$(7.2)\qquad \alpha+\beta+\gamma \leqslant n, \quad \alpha+\gamma+e_1 \equiv 0 \pmod n, \quad \beta+\gamma+e_2 \equiv 0 \pmod n.$$

On the other hand, if α, β and γ satisfy conditions (7.2), then setting

$$\begin{gathered} d := \alpha+\beta+\gamma-1 \\ a = \alpha-1, \quad b = \beta, \quad c = \gamma \\ a' = \alpha, \quad b' = \beta-1, \quad c = \gamma \\ a'' = \alpha, \quad b'' = \beta, \quad c'' = \gamma-1, \end{gathered}$$

https://doi.org/10.4153/CMB-2016-070-8 Published online by Cambridge University Press

we see that the coefficient of $x_1^{\alpha}x_2^{\beta}x_3^{\gamma}$ is

$$t_1 n q^{(e_1+1)(\beta+\gamma)+\frac{\gamma(\gamma-1)}{2}} \binom{d}{\alpha-1,\beta,\gamma}_q + t_2 n q^{e_1(\beta-1+\gamma)+\frac{\gamma(\gamma-1)}{2}} \binom{d}{\alpha,\beta-1,\gamma}_q \\ + t_3 n q^{(e_1+1)(\beta+\gamma-1)+\frac{(\gamma-2)(\gamma-1)}{2}} \binom{d}{\alpha,\beta,\gamma-1}_q .$$

Equating this coefficient to 0 and dividing through by $nq^{e_1(\beta+\gamma)+\frac{\gamma(\gamma-1)}{2}}$, we obtain

$$t_1 q^{\beta+\gamma}\binom{d}{\alpha-1,\beta,\gamma}_q + t_2 q^{-e_1}\binom{d}{\alpha,\beta-1,\gamma}_q + t_3 q^{\beta-e_1}\binom{d}{\alpha,\beta,\gamma-1}_q = 0. \tag{7.3}$$

In summary, $w = (t_1A_1^{e+1}A_2^{e_2}, t_2A_1^{e_1}A_2^{e_2+1}, t_3A_1^{e_1+1}A_2^{e_2+1})$ lies in V_{e_1,e_2} if and only if (7.3) holds for every α,β,γ satisfying conditions (7.2).

Proof of Proposition 6.2(i) Our goal is to show that $w = (t_1A_1, t_2A_2, t_3A_1A_2)$ lies in $V_{0,0}$ if and only if $t_1 = t_2 = t_3 = 0$. Note that here $e_1 = e_2 = 0$, and $(\alpha,\beta,\gamma) = (n,0,0)$, $(0,n,0)$, $(0,0,n)$ satisfy conditions (7.2). Substituting $(\alpha,\beta,\gamma) = (n,0,0)$ into (7.3), and remembering that $\binom{d}{a,b,c}_q = 0$ whenever a, b or c is < 0, we obtain

$$t_1\binom{n-1}{n-1,0,0}_q = 0,$$

or equivalently, $t_1 = 0$. Similarly, setting $(\alpha,\beta,\gamma) = (0,n,0)$ yields $t_2 = 0$, and setting $(\alpha,\beta,\gamma) = (0,0,n)$ yields $t_3 = 0$. This proves part (i). ■

Proof of Proposition 6.2(ii) Here, $(e_1,e_2) \neq (0,0)$, and we can use Lemma 5.2 to simplify formula (7.3) as follows:

$$t_1q^{\beta+\gamma}(1-q^{\alpha}) + t_2q^{-e_1}(1-q^{\beta}) + t_3q^{\beta-e_1}(1-q^{\gamma}) = 0.$$

Using (7.2), we can rewrite this in a more symmetric way, as

$$t_1(q^{-e_2}-q^{d+1}) + t_2(q^{-e_1}-q^{d+1}) + t_3(q^{d+1}-q^{-e_1-e_2}) = 0, \tag{7.4}$$

where $d+1 = \alpha+\beta+\gamma$, as before.

Claim Suppose $e_1, e_2 = 0,\dots,n-1$ and $(e_1,e_2) \neq (0,0)$. Then there exist triples of non-negative integers, $(\alpha_1,\beta_1,\gamma_1)$ and $(\alpha_2,\beta_2,\gamma_2)$ satisfying conditions (7.2) such that $d_1 \not\equiv d_2 \pmod{n}$. Here, $d_1 = \alpha_1+\beta_1+\gamma_1-1$ and $d_2 = \alpha_2+\beta_2+\gamma_2-1$.

We will now deduce Proposition 6.2(ii) from this claim. The proof of the claim will be deferred to the end of this section. Assuming the claim is established, formula (7.4) tells us that if $(t_1A_1^{e_1+1}A_2^{e_2}, t_2A_1^{e_1}A_2^{e_2+1}, t_3A_1^{e_1+1}A_2^{e_2+1})$ lies in V_{e_1,e_2}, then t_1, t_2, and t_3 satisfy the linear equations

$$\begin{aligned} t_1(q^{-e_2}-q^{d_1+1}) + t_2(q^{-e_1}-q^{d_1+1}) + t_3(q^{d_1+1}-q^{-e_1-e_2}) &= 0,\\ t_1(q^{-e_2}-q^{d_2+1}) + t_2(q^{-e_1}-q^{d_2+1}) + t_3(q^{d_2+1}-q^{-e_1-e_2}) &= 0. \end{aligned} \tag{7.5}$$

The matrix of this system

$$\begin{pmatrix} q^{e_2}-q^{d_1+1} & q^{-e_1}-q^{d_1+1} & q^{d_1+1}-q^{-e_1-e_2} \\ q^{e_2}-q^{d_2+1} & q^{-e_1}-q^{d_2+1} & q^{d_2+1}-q^{-e_1-e_2} \end{pmatrix}$$

https://doi.org/10.4153/CMB-2016-070-8 Published online by Cambridge University Press

is easily seen to have rank 2. Indeed, the determinants of the 2×2 minors are

$$\pm (q^{d_1+1} - q^{d_2+1})(q^{-e_2} - q^{-e_1}),$$
$$\pm (q^{d_1+1} - q^{d_2+1})(q^{-e_1-e_2} - q^{-e_1}),$$
$$\pm (q^{d_1+1} - q^{d_2+1})(q^{-e_1-e_2} - q^{-e_1}).$$

Since $q^{d_1+1} \neq q^{d_2+1}$, all three of these determinants can only be zero if $q^{-e_1} = q^{-e_2} = q^{-e_1-e_2}$, or equivalently, $e_1 \equiv e_2 \equiv e_1 + e_2 \pmod n$, *i.e.*, $(e_1, e_2) = (0,0) \pmod n$, contradicting our assumption that $(e_1, e_2) \neq (0,0)$. We conclude that the solution space to system (7.5) is of dimension $\leqslant 1$ and consequently, $\dim(V_{e_1,e_2}) \leqslant 1$. On the other hand, by Remark 6.3, $\dim(V_{e_1,e_2}) \geqslant 1$. This shows that $\dim(V_{e_1,e_2}) = 1$, thus completing the proof of Proposition 6.2(ii). ∎

We now turn to the proof of the claim. The statement of the claim is clearly symmetric with respect to e_1 and e_2. That is, if the triples

$$(\alpha_1, \beta_1, \gamma_1) \text{ and } (\alpha_2, \beta_2, \gamma_2)$$

satisfy the claim for (e_1, e_2), then the triples $(\beta_1, \alpha_1, \gamma_1)$, $(\beta_2, \alpha_2, \gamma_2)$ will satisfy the claim for (e_2, e_1). Thus, for the purpose of proving this claim, we can assume without loss of generality that $0 \leqslant e_2 \leqslant e_1 \leqslant n-1$.

Case 1: $e_2 \geqslant 1$. Here, the triples

$$(\alpha_1, \beta_1, \gamma_1) = (0, e_1 - e_2, n - e_1) \quad \text{and} \quad (\alpha, \beta, \gamma) = (1, e_1 - e_2 + 1, n - e_1 - 1)$$

satisfy conditions (7.2) and yield distinct sums $d_1 + 1 = \alpha_1 + \beta_1 + \gamma_1 = n - e_2$ and $d_2 + 1 = \alpha_2 + \beta_2 + \gamma_2 = n - e_2 + 1$. Note that $d_2 + 1 \leqslant n$, because we are assuming that $e_2 \geqslant 1$.

Case 2: $e_2 = 0$ but $1 \leqslant e_1 \leqslant n-1$. Set $(\alpha_1, \beta_1, \gamma_1) = (0, e_1, n - e_1)$ as in Case 1, and $(\alpha_2, \beta_2, \gamma_2) = (n - e_1, 0, 0)$. Then $d_1 + 1 = n$ and $d_2 + 1 = n - e_1$ are, once again, distinct modulo n. This completes the proof of the claim, and hence, of Proposition 6.2 and of Theorem 1.3. ∎

8 The Case Where $r \geqslant n^2 - 1$

Let $K_{r,n} \coloneqq k(\mathrm{M}_n^r)^{\mathrm{PGL}_n}$ be the field of matrix invariants and let $K'_{r,n}$ be the subfield generated by the coefficients of the generalized characteristic polynomial

$$(A_1, \dots, A_r) \longmapsto \det(x_0 I + x_1 A_1 + \cdots + x_r A_r),$$

as in Section 3. Recall that $K_{r,n}$ is the field of rational functions on $\mathrm{M}_n^r /\!\!/ \mathrm{PGL}_n$ and $K'_{r,n}$ is the field of rational functions on $\mathrm{DHyp}_{r,n}$.

By abuse of notation we will denote by t the transposition map $\mathrm{M}_n \to \mathrm{M}_n$ as well as the maps it induces on M_n^r (by applying t to each component), $\mathrm{M}_n^r /\!\!/ \mathrm{PGL}_n$, and their function fields. For example,

$$t\big(\mathrm{Tr}(A_1 A_2 A_3)\big) \coloneqq \mathrm{Tr}(A_1^t A_2^t A_3^t) = \mathrm{Tr}(A_3 A_2 A_1).$$

Since $\det(x_0 I + x_1 A_1 + \cdots + x_r A_r) = \det(x_0 I + x_1 A_1^t + \cdots + x_r A_r^t)$, we have

$$K'_{r,n} \subset K^t_{r,n}. \tag{8.1}$$

https://doi.org/10.4153/CMB-2016-070-8 Published online by Cambridge University Press

Our standing assumption that the base field k is algebraically closed of characteristic 0 or $> n$ remains in force.

Lemma 8.1 *Assume that $r \geqslant 2$, $n \geqslant 2$, and $(r,n) \neq (2,2)$. Then the following assertions are equivalent.*

(i) *The general fiber of $\overline{P}\colon \mathrm{M}_n^r /\!/ \mathrm{PGL}_n \to \mathrm{DHyp}_{r,n}$ consists of exactly two points corresponding to the conjugacy classes of $(A_1,\dots,A_r)$ and $(A_1^t,\dots,A_r^t)$.*

(ii) $[K_{r,n} : K'_{r,n}] = 2$.

(iii) $K'_{r,n} = K^t_{r,n}$.

Proof (i) $\Rightarrow$ (ii). Theorem 1.3 tells us that $K_{n,r}/K'_{n,r}$ is a finite separable extension. Thus the general fiber of $\overline{P}$ consists of exactly $[K_{r,n} : K'_{r,n}]$ points.

(ii) $\Leftrightarrow$ (iii). Under our assumptions on r and n, t is an automorphism of $K_{r,n}$ of order 2. Thus $[K_{r,n} : K^t_{r,n}] = 2$. In view of (8.1), $[K_{r,n} : K'_{r,n}] \geqslant 2$, and equality holds if and only if $K'_{r,n} = K^t_{r,n}$.

(iii) $\Rightarrow$ (i). If (iii) holds, then a general fiber of $\overline{P}$ has exactly two elements. If such a fiber contains a point representing A, it also contains a point representing A^t. For $A \in \mathrm{M}_n^r$ in general position, these points are distinct (here we are using the assumption that $(r,n) \neq (2,2)$!), so there cannot be any others. ∎

Our goal now is show that in the case where $r \geqslant n^2 - 1$, Theorem 1.3 can be strengthened as follows.

Theorem 8.2 *The equivalent conditions of Lemma 8.1 hold if $r \geqslant n^2 - 1$, for any $n \geqslant 2$.*

The rest of this section will be devoted to proving Theorem 8.2. We proceed in three steps. (1) Lemma 8.3 settles the case where $n = 2$; (2) Lemma 8.4 settles the case where $r = n^2 - 1$; (3) Proposition 8.5 supplies the induction step, showing that if the equivalent conditions of Lemma 8.1 hold for some parameters r and n, then they also hold for $r+1$ and n, provided that $r, n \geqslant 3$.

Lemma 8.3 *Assume that $r \geqslant 2$. Then*

(i) $K'_{r,2} = k\big(\mathrm{Tr}(A_i), \mathrm{Tr}(A_iA_j) \mid i,j = 1,\dots,r\big)$.

(ii) $K'_{r,2} = K^t_{r,2}$.

Proof (i) Recall that $K'_{r,2}$ is generated over k by the coefficients of

$$\det(x_0 I + x_1 A_1 + \cdots + x_r A_r),$$

where I is the 2×2 identity matrix. Setting

$$X := x_0 I + x_1 A_1 + \cdots + x_r A_r$$

and using the formula $\det(X) = \frac{1}{2}(\mathrm{Tr}(X)^2 - \mathrm{Tr}(X^2))$, we see that $K'_{r,2}$ is generated over $k(\mathrm{Tr}(A_i) | i = 1,\dots,r)$ by the coefficients of $\mathrm{Tr}(X^2)$, and part (i) follows. (ii) Let V be the 3-dimensional subspace of trace zero 2×2 matrices, equipped with the non-degenerate quadratic form $q(A,B) = \mathrm{Tr}(AB)$. Then the representation $\mathrm{PGL}_2 \to \mathrm{GL}(V)$ given by the conjugation action is an isomorphism between PGL_2 and $\mathrm{SO}(V) \simeq \mathrm{SO}_3$. The transposition map $t\colon V \to V$ also preserves the trace form; the subgroup G of $\mathrm{GL}(V)$ generated by $\mathrm{SO}(V)$ and t is easily seen to be the full orthogonal group $\mathrm{O}(V)$. Now observe that by definition, $K^t_{r,2} = k(\mathrm{M}_2^r)^G$. Let us identify M_2 with $V_0 \oplus V$, via the isomorphism

$$A \to \Big(\operatorname{Tr}(A), A - \frac{1}{2}\operatorname{Tr}(A)\Big).$$

Here, V_0 denotes the 1-dimensional trivial representation of G. This identifies $K^t_{r,2}$ with the field of $\mathrm{O}(V)$-invariants of $V_0^r \oplus V^r$. The First Fundamental Theorem of classical invariant theory tells us that this field of invariants is generated by $k(V_0^r)$ and the functions

$$(t_1,\dots,t_r,v_1,\dots,v_r) \longmapsto q(v_i,v_j),$$

where $t_1,\dots,t_r \in V_0$, $v_1,\dots,v_r \in V$; see, *e.g.*, [dCP, Theorem 5.7]. Remembering our identification between M_n and $V_0 \oplus V$, we readily translate this into

$$K^t_{r,2} = k\big(\operatorname{Tr}(A_i), \operatorname{Tr}(A_iA_j) \mid i,j = 1,\dots,r\big).$$

The desired equality, $K'_{r,2} = K^t_{r,2}$ now follows from part (i). ∎

Lemma 8.4 *Let $r = n^2 - 1$ and assume that $I, A_1,\dots,A_r$ span M_n as a k-vector space. If*

$$\det(x_0 I + x_1 A_1 + \cdots + x_r A_r) = \det(x_0 I + x_1 B_1 + \cdots + x_r B_r)$$

for some $B = (B_1,\dots,B_r) \in \mathrm{M}_n^r$, then B is conjugate to A or B is conjugate to A^t.

Proof Let $T\colon \mathrm{M}_n \to \mathrm{M}_n$ be the linear transformation taking I to I and A_i to B_i for every $i = 1,\dots,r$. By our assumption, T preserves the determinant function. By a theorem of Frobenius, there exist $P, Q \in \mathrm{M}_n$ such that $\det(P)\det(Q) = 1$ and $T(X) = CXD$; see the references in Remark (1) in the introduction. Since $T(I) = I$, we have $C = D^{-1}$, and the lemma follows. ∎

Proposition 8.5 *Assume that $r, n \geqslant 3$. If $K'_r = K^t_{r,n}$, then $K'_{r+1} = K^t_{r+1,n}$.*

Proof This proposition is in the same spirit as Proposition 3.2, and we will use a more elaborate version of the same argument. Once again, a key ingredient will be supplied by Lemma 3.3, which asserts that there exist finitely many monomials $M_1,\dots,M_N$ in A_1 and A_2 such that $K_{r,n}$ is generated, as a field extension of k, by the elements $\operatorname{Tr}(M_i)$ and $\operatorname{Tr}(M_iA_j)$, where $i = 1,\dots,N$, and $j = 3,\dots,r$. To simplify the notation, set

$$\begin{aligned}
s_i &:= \operatorname{Tr}(M_i) + \operatorname{Tr}(M_i)^t,\\
\Delta_i &:= \operatorname{Tr}(M_i) - \operatorname{Tr}(M_i)^t,\\
s_{i,j} &:= \operatorname{Tr}(M_iA_j) + \operatorname{Tr}(A_jM_i)^t,\\
\Delta_{i,j} &:= \operatorname{Tr}(M_iA_j) - \operatorname{Tr}(A_jM_i)^t.
\end{aligned}$$

We will also need a non-zero element $f \in K_{2,n}$ with the property that $t(f) = -f$. Such an element exists for every $n \geqslant 3$; for example, we can take

$$f(A_1,A_2) := \operatorname{Tr}(A_1A_2A_1^2A_2^2) - \operatorname{Tr}(A_2^2A_1^2A_2A_1).$$

For this choice of f, the equality $t(f) = -f$ is clear; the computation on [R93, p. 72] shows that $f \neq 0$. (Note that here we are using the assumption that $n \geqslant 3$. For $n = 2$, f cannot exist because t acts trivially on $K_{2,n}$, and our argument below breaks down. This is the reason we handled the case where $n = 2$ separately, in Lemma 8.3.) Now

$$\begin{aligned} K^t_{r+1,n} &= k\big(\operatorname{Tr}(M_i), \operatorname{Tr}(M_i A_j) \mid i = 1, \dots, N, j = 3, \dots, r+1\big) \\ &= k(s_i, \Delta_i, s_{ij}, \Delta_{ij} \mid i = 1, \dots, N, j = 3, \dots, r+1)^t \\ &= k(s_i, \Delta_i f, s_{ij}, \Delta_{ij} f, f \mid i = 1, \dots, N, j = 3, \dots, r+1)^t. \end{aligned}$$

The elements s_i, $\Delta_i f$, s_{ij}, $\Delta_{ij} f$ are all fixed by t, while $t(f) = -f$. Thus,

$$K^t_{r+1,n} = k(s_i, \Delta_i f, s_{ij}, \Delta_{ij} f, f^2). \tag{8.2}$$

Clearly, $K'_{r+1,n} \subset K^t_{r+1,n}$. To prove equality, it suffices to show that each of the generators $s_i, \Delta_i f, s_{ij}, \Delta_{ij} f$, and f^2 lie in $K'_{r+1,n}$.

Note that s_i, $\Delta_i f$, and f^2 lie in $K^t_{2,n}$, and s_{i3} and $\Delta_{i3} f$ lie in $K^t_{3,n}$. Since $r \geqslant 3$, these elements all lie in $K^t_{r,n}$. By our assumption, $K^t_{r,n} = K'_{r,n} \subset K'_{r+1,n}$. Hence, each of the generators $f^2, s_i, \Delta_i f, s_{i3}, \Delta_{i3} f$ lie in $K'_{r+1,n}$. By symmetry, s_{ij} and $\Delta_{ij} f$ also lie in $K'_{r+1,n}$, for any $j = 3, \dots, r+1$. We conclude that $f^2, s_i, \Delta_i f, s_{ij}, \Delta_{ij} f$ all lie in $K'_{r+1,n}$. By (8.2), $K^t_{r+1,n} = K'_{r+1,n}$, as desired. ∎

Acknowledgments We would like to thank Boris Reichstein for bringing Question 1.1 to our attention. We are also grateful to Arnaud Beauville for helpful comments, and to the referees for calling our attention to [ES03] and encouraging us to include a proof of Theorem 8.2 in this paper. We are in debt to Marian Aprodu for his interest in our work and his very useful observations.

References

[Beau00] A. Beauville, *Determinantal hypersurfaces.* Michigan Math. J. **48**(2000), 39–64. http://dx.doi.org/10.1307/mmj/1030132707

[BGL14] H. Bermudez, S. Garibaldi, and V. Larsen, *Linear preservers and representations with a 1-dimensional ring of invariants.* Trans. Amer. Math. Soc. **366**(2014), no. 9, 4755–4780. http://dx.doi.org/10.1090/S0002-9947-2014-06081-9

[CT79] R. J. Cook and A. D. Thomas, *Line bundles and homogeneous matrices.* Quart. J. Math. Oxford Ser. (2) **30**(1979), no. 120, 423–429. http://dx.doi.org/10.1093/qmath/30.4.423

[dCP] C. de Concini and C. Procesi, *A characteristic free approach to invariant theory*, Advances in Math. **21**(1976), no. 3, 330–354. http://dx.doi.org/10.1016/S0001-8708(76)80003-5

[Dieu49] J. Dieudonné, *Sur une généralisation du groupe orthogonal à quatre variables.* Arch. Math. **1**(1949), 282–287. http://dx.doi.org/10.1007/BF02038756

[Dickson21] L. E. Dickson, *Determination of all general homogeneous polynomials expressible as determinants with linear elements.* Trans. Amer. Math. Soc. **22**(1921), no. 2, 167–179. http://dx.doi.org/10.1090/S0002-9947-1921-1501168-0

[Dolg12] I. V. Dolgachev, *Classical algebraic geometry*. Cambridge University Press, Cambridge, 2012. http://dx.doi.org/10.1017/CBO9781139084437

[ES03] D. Eisenbud, F.-O. Schreyer, and J. Weyman, *Resultants and Chow forms via exterior syzygies.* J. Amer. Math. Soc. **16**(2003), no. 3, 537–579. http://dx.doi.org/10.1090/S0894-0347-03-00423-5

[FHL81] E. Formanek, P. Halpin, and W. C. W. Li, *The Poincaré series of the ring of* 2×2 *generic matrices.* J. Algebra **69**(1981), no. 1, 105–112. http://dx.doi.org/10.1016/0021-8693(81)90130-7

[FGG97] A. Freedman, R. N. Gupta, and R. M. Guralnick, *Shirshov's theorem and representations of semigroups.* Pacific J. Math. **1997**, Special Issue, 159–176. http://dx.doi.org/10.2140/pjm.1997.181.159

[F1897] G. Frobenius, *Über die Darstellung der endlichen Gruppen durch lineare Substitutionen.* Berlin Sitzungsber, 1897, 994–1015.

[G1855] H. Grassmann, *Die stereometrischen Gleichungen dritten Grades, und die dadurch erzeugten Oberflächen.* J. Reine Angew. Math. **49**(1855), 47–65. http://dx.doi.org/10.1515/crll.1855.49.47

[H71] I. N. Herstein, *Notes from a ring theory conference*. Conference Board of the Mathematical Sciences Regional Conference Series in Mathematics, 9, American Mathematical Society, Providence, RI, 1971.

[HMS04] O. Holtz, V. Mehrmann, and H. Schneider, *Potter, Wielandt, and Drazin on the matrix equation $AB = \omega BA$: new answers to old questions.* Amer. Math. Monthly **111**(2004), no. 8, 655–667. http://dx.doi.org/10.2307/4145039

[MM59] M. Marcus and B. N. Moyls, *Linear transformations on algebras of matrices.* Canad. J. Math. **11**(1959), 61–66. http://dx.doi.org/10.4153/CJM-1959-008-0

[Ne11] Yu. A. Neretin, *Spectral data for a pair of matrices of order 3 and the action of the group* $\mathrm{GL}(2, \mathbb{Z})$. (Russian), Izv. Ross. Akad. Nauk Ser. Mat. **75**(2011), no. 5, 93–102; translation in Izv. Math. **75**(2011), no. 5, 959–969. http://dx.doi.org/10.4213/im4127

[P67] C. Procesi, *Non-commutative affine rings.* Atti Accad. Naz. Lincei Mem. Cl. Sci. Fis. Mat. Natur. Sez. I (8) **8**(1967), 237–255.

[R93] Z. Reichstein, *On automorphisms of matrix invariants induced from the trace ring.* Linear Algebra Appl. **193**(1993), 51–74. http://dx.doi.org/10.1016/0024-3795(93)90271-O

[Sch53] M. P. Schützenberger, *Une interprétation de certaines solutions de l'équation fonctionnelle: $F(x + y) = F(x)F(y)$.* C. R. Acad. Sci. Paris **236**(1953), 352–353.

[Vin86] V. Vinnikov, *Determinantal representations of algebraic curves.* In: Linear algebra in signals, systems, and control (Boston, MA, 1986), SIAM, Philadelphia, PA, pp. 73–99.

[Wat87] W. C. Waterhouse, *Automorphisms of* $\det(X_{ij})$*: the group scheme approach.* Adv. in Math. **65**(1987), no. 2, 171–203. http://dx.doi.org/10.1016/0001-8708(87)90021-1

Department of Mathematics, University of British Columbia, Vancouver, BC V6T 1Z2
e-mail: reichst@math.ubc.ca

Scuola Normale Superiore, Piazza dei Cavalieri 7, 56126 Pisa, Italy
e-mail: angelo.vistoli@gmail.com