Journal of Algebra **213,** 195–224 (1999)
Article ID jabr.1995.6650, available online at http://www.idealibrary.com on IDEAL®

Canonical Forms of Two by Two by Two Matrices

Richard Ehrenborg*

Department of Mathematics, White Hall, Cornell University, Ithaca, New York 14853-7901

Communicated by David Buchsbaum

Received December 10, 1995

DEDICATED TO THE MEMORY OF ROSA Q. HUANG

We determine the canonical forms of two by two by two matrices. In order to do so, we present a collection of covariants that separates the canonical forms. The covariants also fulfill a number of identities. Moreover, we show that these canonical forms and covariants are related to those of the binary cubic and skew-symmetric tensors of step 3 in the exterior algebra $\mathrm{Ext}(U)$, where $\dim(U) = 6$.

1. INTRODUCTION

One of the goals of invariant theory is to find canonical ways to write symmetric tensors and skew-symmetric tensors. The symmetric tensors may be viewed as polynomials and the skew-symmetric tensors as elements of the exterior algebra.

In this paper we raise the same question for general tensors. These general tensors may be seen as multi-dimensional matrices. We will concentrate our efforts on the first nontrivial case, namely the three-dimensional two by two by two matrix. In Section 3 we introduce the notion of invariants and covariants of the two by two by two matrices. In Proposition 3.4 we introduce an invariant Γ. We show by a Gaussian elimination argument that any invariant of two by two by two matrices has the form $s \cdot \Gamma^k$, where s is a constant and k is a nonnegative integer.

*The research for this paper was done at MIT, and the paper was written when the author was a joint postdoctoral fellow between Centre de Recherches Mathématiques at Université de Montréal and LACIM at Université du Québec à Montréal.

0021-8693/99 $30.00

In Section 4 we introduce the umbral notation. This notation helps us to express the covariants of two by two by two matrices. Moreover, we present a list of six covariants, Γ, S, $H^{(1)}$, $H^{(2)}$, $H^{(3)}$, and Id, which help us to differentiate between the canonical forms of two by two by two matrices.

In Section 5 we present the seven different canonical forms for the two by two by two matrices. We also give a table of the six covariants evaluated on these seven canonical forms. As a corollary we know that $pqr + stu$ is a generic canonical form. By using this fact it is easy to prove certain relations that hold between these covariants.

It turns out that these canonical forms and covariants are related to those of a binary cubic; see Section 6. Indeed, the binary cubic may be viewed as a subclass of the two by two by two matrices, namely the symmetric matrices. Thus we find a natural correspondence between the canonical forms and covariants of two by two by two matrices and binary cubics. In Section 7 a similar relation is obtained between two by two by two matrices and elements of step three in the exterior algebra $\mathrm{Ext}(U)$, where $\dim(U) = 6$. This suggests a question for future study: Are the covariants of symmetric tensors, skew-symmetric tensors, and general tensors all interconnected?

I thank Gian-Carlo Rota, who introduced me to the study of invariant theory. I also thank Dan Klain, Gábor Hetyei, and Margaret Readdy for many helpful comments.

2. INVARIANTS AND COVARIANTS OF POLYNOMIALS

We start by recalling some facts about two-dimensional vector spaces. For $\mathbf{a}, \mathbf{b} \in \mathbb{C}^2$ define

$$(\mathbf{a} \mid \mathbf{b}) = a_1 b_1 + a_2 b_2,$$
$$[\mathbf{a}, \mathbf{b}] = a_1 b_2 - a_2 b_1,$$

where $\mathbf{a} = (a_1, a_2)$ and $\mathbf{b} = (b_1, b_2)$. Let ϕ be a linear map from $\mathbb{C}^2$ to itself. Define the determinant of ϕ to be the determinant of its matrix, that is,

$$\det(\phi) = \begin{vmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{vmatrix},$$

where $\phi(\mathbf{e}_1) = (a_{1,1}, a_{1,2})$ and $\phi(\mathbf{e}_2) = (a_{2,1}, a_{2,2})$. Observe that the determinant of the linear map ϕ is independent of the basis of the linear space.

Let ϕ^* be the adjoint map, that is, the linear map defined by $\phi^*(\mathbf{e}_1) = (a_{1,1}, a_{2,1})$ and $\phi^*(\mathbf{e}_2) = (a_{1,2}, a_{2,2})$.

LEMMA 2.1. *Let* $\mathbf{a}, \mathbf{b} \in \mathbb{C}^2$, *and let* $\phi, \psi\colon \mathbb{C}^2 \to \mathbb{C}^2$ *be two linear maps. Then the following are true*:

$$\begin{aligned} (\phi(\mathbf{a}) \mid \mathbf{b}) &= (\mathbf{a} | \phi^*(\mathbf{b})), \\ [\phi(\mathbf{a}), \phi(\mathbf{b})] &= \det(\phi) \cdot [\mathbf{a}, \mathbf{b}], \\ \det(\phi^*) &= \det(\phi), \\ \det(\phi \circ \psi) &= \det(\phi) \cdot \det(\psi), \\ (\phi \circ \psi)^* &= \psi^* \circ \phi^*. \end{aligned}$$

Let y_1, y_2 be variables and let $\operatorname{span}(y_1, y_2)$ be the two-dimensional linear space spanned by y_1 and y_2. Let A be an algebra over the field of complex numbers $\mathbb{C}$, and let $A[y_1, y_2]$ be the polynomial algebra with variables y_1, y_2 and coefficients in A. A linear map $\phi\colon \operatorname{span}(y_1, y_2) \to \operatorname{span}(y_1, y_2)$ extends to a map $\hat{\phi}\colon A[y_1, y_2] \to A[y_1, y_2]$ by substitution. More formally, the action of $\hat{\phi}$ is determined by

$$\hat{\phi} p(y_1, y_2) = p(\phi(y_1), \phi(y_2)),$$

where $p(\mathbf{y}) \in A[y_1, y_2]$. Thus the map $\hat{\phi}$ corresponds to a change of variables. Moreover, note that $\widehat{\phi \circ \psi} = \hat{\phi} \circ \hat{\psi}$.

DEFINITION 2.2. A polynomial map f from a finite-dimensional linear space U to a finite-dimensional linear space V is a function from U to V that can be written in the form

$$f\left(\sum_{i=1}^{n} a_i u_i\right) = \sum_{j=1}^{m} p_j(a_1, \ldots, a_n) v_j,$$

where $p_1, \ldots, p_m$ are polynomials, the sequence $u_1, \ldots, u_n$ is a basis for U, and $v_1, \ldots, v_n$ is a basis for V.

Let V_p be the linear subspace of $\mathbb{C}[\mathbf{y}] = \mathbb{C}[y_1, y_2]$ consisting of all homogeneous elements of degree p.

DEFINITION 2.3. A covariant C of V_p is a polynomial map from V_p to $\mathbb{C}[\mathbf{y}]$, which for all linear maps $\phi\colon \operatorname{span}(y_1, y_2) \to \operatorname{span}(y_1, y_2)$ satisfies the identity

$$C(\hat{\phi} w) = \det(\phi)^g \cdot \hat{\phi} C(w),$$

where the nonnegative integer g is called the index of C.

DEFINITION 2.4. An invariant I of V_p is a covariant of V_p which maps V_p into $\mathbb{C}$.

Since $\hat{\phi}c = c$ for all $c \in \mathbb{C}$ we have that the condition for I being an invariant is

$$I\left(\hat{\phi}w\right) = \det(\phi)^g \cdot I(w),$$

for the index g.

An example of these concepts is the discriminant, Δ, on the space V_2 of binary quadratics. This is a classical invariant and it has index 2.

$$\Delta\left(a_0 y_1^2 + a_1 y_1 y_2 + a_2 y_2^2\right) = a_1^2 - 4a_0 a_2.$$

It can also be described by

$$\Delta(w) = -\begin{vmatrix} \dfrac{\partial^2 w}{\partial y_1^2} & \dfrac{\partial^2 w}{\partial y_1\, \partial y_2} \\ \dfrac{\partial^2 w}{\partial y_2\, \partial y_1} & \dfrac{\partial^2 w}{\partial y_2^2} \end{vmatrix}.$$

Moreover, we define a bracket on $\operatorname{span}(y_1, y_2)$ by $[a_1 y_1 + a_2 y_2, b_1 y_1 + b_2 y_2] = a_1 b_2 - a_2 b_1$. The discriminant can be computed as follows:

$$\Delta(pq) = [p, q]^2.$$

Observe that we also may describe the discriminant by

$$\Delta\left(p^2 + q^2\right) = -4[p, q]^2.$$

3. INVARIANTS AND COVARIANTS OF TWO BY TWO BY TWO MATRICES

We introduce two by two by two matrices as certain homogeneous polynomials in six variables. This has the advantage that row, column, and height operations will correspond to a change of variables. This notion of polynomials representing two by two by two matrices is helpful when we define covariants and invariants of two by two by two matrices.

Let $\mathbf{x}$ denote the sextet of variables $(x_j^{(i)})$ where $i = 1, 2, 3$ and $j = 1, 2$. For $i = 1, 2, 3$ let $\mathbf{x}^{(i)} = (x_1^{(i)}, x_2^{(i)})$. Let $\mathbb{C}[\mathbf{x}]$ denote the algebra of polynomials in these six variables. Let W be the linear subspace of $\mathbb{C}[\mathbf{x}]$ which consists of polynomials that are homogeneous and have degree 1 in $\mathbf{x}^{(i)}$ for all $i = 1, 2, 3$. An element w of W can be written in the form

$$\sum_{i=1}^{2} \sum_{j=1}^{2} \sum_{k=1}^{2} a_{i,j,k} \cdot x_i^{(1)} x_j^{(2)} x_k^{(3)}.$$

Observe that w is isomorphic to the linear space of two by two by two matrices by sending w to the three-dimensional matrix

$$\left(\begin{pmatrix} a_{1,1,1} & a_{1,1,2} \\ a_{1,2,1} & a_{1,2,2} \end{pmatrix}, \begin{pmatrix} a_{2,1,1} & a_{2,1,2} \\ a_{2,2,1} & a_{2,2,2} \end{pmatrix} \right).$$

Let W_i be the linear space spanned by $x_1^{(i)}$ and $x_2^{(i)}$ for $i = 1, 2, 3$. Observe that $W \cong W_1 \otimes W_2 \otimes W_3$. Let $\phi_i \colon W_i \to W_i$ be a linear map. Thus $\hat{\phi}_i$ is a change of variables in the $\mathbf{x}^{(i)}$'s.

DEFINITION 3.1. A covariant C of W is a polynomial map from W to $\mathbb{C}[\mathbf{x}]$ such that for all linear maps $\phi_i \colon W_i \to W_i$, where $i = 1, 2, 3$, we have that

$$\begin{aligned} C\left(\hat{\phi}_1 \hat{\phi}_2 \hat{\phi}_3 w\right) &= \prod_{i=1}^{3} \det(\phi_i)^{g_i} \cdot \hat{\phi}_1 \hat{\phi}_2 \hat{\phi}_3 C(w) \\ &= \det(\phi_1)^{g_1} \cdot \det(\phi_2)^{g_2} \cdot \det(\phi_3)^{g_3} \cdot \hat{\phi}_1 \hat{\phi}_2 \hat{\phi}_3 C(w), \end{aligned}$$

for some nonnegative integers g_1, g_2, and g_3. We call the triple (g_1, g_2, g_3) the index of the covariant C.

There are two trivial covariants. The first one is $C \equiv c$ where c is a constant. The second covariant is the identity map, that is, $C(w) = w$ for all $w \in W$. Both these covariants have the index $(g_1, g_2, g_3) = (0, 0, 0)$.

LEMMA 3.2. *Let C and C' be two covariants with corresponding indices (g_1, g_2, g_3) and (g_1', g_2', g_3').*

1. *If the two indices are the same, that is, $(g_1, g_2, g_3) = (g_1', g_2', g_3')$, then $C + C'$ is a covariant with the same index.*
2. *$C \cdot C'$ is a covariant with the index $(g_1 + g_1', g_2 + g_2', g_3 + g_3')$.*
3. *Let $\psi_i \colon W_i \to W_i$ be invertible linear maps for $i = 1, 2, 3$. Then*

$$D(w) = \hat{\psi}_1^{-1} \hat{\psi}_2^{-1} \hat{\psi}_3^{-1} C\left(\hat{\psi}_1 \hat{\psi}_2 \hat{\psi}_3 w\right)$$

is also a covariant with index (g_1, g_2, g_3).

Proof. The first two statements follow directly from Definition 3.1. The third statement is proved by the following computation:

$$
\begin{aligned}
&D\left(\hat{\phi}_1\hat{\phi}_2\hat{\phi}_3 w\right)\\
&\quad= \hat{\psi}_1^{-1}\hat{\psi}_2^{-1}\hat{\psi}_3^{-1} C\left(\hat{\psi}_1\hat{\psi}_2\hat{\psi}_3\hat{\phi}_1\hat{\phi}_2\hat{\phi}_3 w\right)\\
&\quad= \hat{\psi}_1^{-1}\hat{\psi}_2^{-1}\hat{\psi}_3^{-1} \prod_{i=1}^{3} \det(\psi\phi)^{g_i} \cdot \hat{\psi}_1\hat{\psi}_2\hat{\psi}_3\hat{\phi}_1\hat{\phi}_2\hat{\phi}_3 C(w)\\
&\quad= \prod_{i=1}^{3} \det(\phi)^{g_i} \cdot \prod_{i=1}^{3} \det(\psi)^{g_i} \cdot \hat{\phi}_1\hat{\phi}_2\hat{\phi}_3 C(w)\\
&\quad= \prod_{i=1}^{3} \det(\phi)^{g_i} \cdot \prod_{i=1}^{3} \det(\psi)^{g_i} \cdot \hat{\phi}_1\hat{\phi}_2\hat{\phi}_3\hat{\psi}_1^{-1}\hat{\psi}_2^{-1}\hat{\psi}_3^{-1}\hat{\psi}_1\hat{\psi}_2\hat{\psi}_3 C(w)\\
&\quad= \prod_{i=1}^{3} \det(\phi)^{g_i} \cdot \hat{\phi}_1\hat{\phi}_2\hat{\phi}_3\hat{\psi}_1^{-1}\hat{\psi}_2^{-1}\hat{\psi}_3^{-1} C\left(\hat{\psi}_1\hat{\psi}_2\hat{\psi}_3 w\right)\\
&\quad= \prod_{i=1}^{3} \det(\phi)^{g_i} \cdot \hat{\phi}_1\hat{\phi}_2\hat{\phi}_3 D(w). \quad \blacksquare
\end{aligned}
$$

DEFINITION 3.3. An invariant I of W is a covariant of W which maps W into $\mathbb{C}$.

Since $\hat{\phi}_1\hat{\phi}_2\hat{\phi}_3 c = c$ for all $c \in \mathbb{C}$ we have that the condition for I being an invariant is

$$
I\left(\hat{\phi}_1\hat{\phi}_2\hat{\phi}_3 w\right) = \prod_{i=1}^{3} \det(\phi_i)^{g_i} \cdot I(w)
$$

for the index (g_1, g_2, g_3).

PROPOSITION 3.4. *All invariants of W are of the form $s \cdot \Gamma^k(w)$, where s is a complex constant, k is a nonnegative integer, and Γ is an invariant defined by*

$$
\begin{aligned}
\Gamma(w) = {}& (a_{1,1,1}a_{2,2,2} + a_{1,2,1}a_{2,1,2} - a_{1,1,2}a_{2,2,1} - a_{1,2,2}a_{2,1,1})^2\\
&- 4(a_{1,1,1}a_{2,1,2} - a_{1,1,2}a_{2,1,1})(a_{1,2,1}a_{2,2,2} - a_{1,2,2}a_{2,2,1}),
\end{aligned}
$$

where

$$w = \sum_{i=1}^{2} \sum_{j=1}^{2} \sum_{k=1}^{2} a_{i,j,k} x_i^{(1)} x_j^{(2)} x_k^{(3)}.$$

The invariant $s \cdot \Gamma^k$ *has the index* $(2k, 2k, 2k)$.

Proof. It is a straightforward calculation to see that Γ is an invariant. By the following rewriting of the definition given for Γ, it is easy to check that $\Gamma(\hat{\phi}_3 w) = \det(\hat{\phi}_3)^2 \cdot \Gamma(w)$.

$$\Gamma(w) = \left(\begin{vmatrix} a_{1,1,1} & a_{2,2,1} \\ a_{1,1,2} & a_{2,2,2} \end{vmatrix} + \begin{vmatrix} a_{1,2,1} & a_{2,1,1} \\ a_{1,2,2} & a_{2,1,2} \end{vmatrix} \right)^2 - 4 \cdot \begin{vmatrix} a_{1,1,1} & a_{2,1,1} \\ a_{1,1,2} & a_{2,1,2} \end{vmatrix} \cdot \begin{vmatrix} a_{1,2,1} & a_{2,2,1} \\ a_{1,2,2} & a_{2,2,2} \end{vmatrix}.$$

The other two conditions $\Gamma(\hat{\phi}_i w) = \det(\hat{\phi}_i)^2 \cdot \Gamma(w)$, $i = 1, 2$, follow by similar reformulations of $\Gamma(w)$. By applying Lemma 3.2 it follows that $s \cdot \Gamma^k$ is an invariant. This argument proves one implication of the proposition.

Let I be an invariant of the space W. Let w be a generic element of W. We compute $I(w)$ using the fact that I is an invariant. This means that we apply different changes of variables to reduce the element w to $w_0 = x_1^{(1)} x_1^{(2)} x_1^{(3)} + x_2^{(1)} x_2^{(2)} x_2^{(3)}$. To make this process easy to follow, we write the elements of W as a two by two by two matrices. Hence the element w_0 corresponds to the matrix

$$\left(\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right).$$

Let us first only apply maps $\hat{\phi}_i$ whose determinant is equal to 1. This corresponds to multiplying one plane in three-dimensional matrix by a scalar and adding it to the parallel plane. We begin with the matrix corresponding to w:

$$\left(\begin{pmatrix} a & b \\ c & d \end{pmatrix}, \begin{pmatrix} e & f \\ g & h \end{pmatrix} \right).$$

$$\left(\begin{pmatrix} a & b \\ c + pa & d + pb \end{pmatrix}, \begin{pmatrix} e & f \\ g + pe & h + pf \end{pmatrix} \right).$$

$$\left(\begin{pmatrix} a + qb & b \\ c + pa + qd + qpb & d + pb \end{pmatrix}, \begin{pmatrix} e + qf & f \\ g + pe + qh + qpf & h + pf \end{pmatrix} \right). \quad (1)$$

Now find complex numbers p and q such that

$$c + pa + qd + qpb = 0,$$
$$g + pe + qh + qpf = 0.$$

This is an equation system of second order, and the solution is given by

$$p = \frac{-(ah + cf - bg - de) \pm \sqrt{D}}{2(af - be)},$$

$$q = \frac{+(ah - cf + bg - de) \mp \sqrt{D}}{2(df - bh)},$$

where

$$D = (ah + cf - bg - de)^2 - 4(af - be)(ch - dg).$$

That is, D is the discriminant of the equation system in p and q. Since we started with a generic element, we may assume that $af - be \neq 0$ and $df - bh \neq 0$, so p and q are well defined. Thus the two by two by two matrix (1) is reduced to

$$\left(\begin{pmatrix} a + qb & b \\ 0 & d + pb \end{pmatrix}, \begin{pmatrix} e + qf & f \\ 0 & h + pf \end{pmatrix} \right).$$

Since we started with a generic element w, we may assume that the two vectors $(a + qb, e + qf)$ and $(d + pb, h + pf)$ are linearly independent, and hence the vector (b, f) is in the span of them. Using this information we may reduce the matrix to

$$\left(\begin{pmatrix} a + qb & 0 \\ 0 & d + pb \end{pmatrix}, \begin{pmatrix} e + qf & 0 \\ 0 & h + pf \end{pmatrix} \right).$$

Continue with the following sequences of operations.

$$\left(\begin{pmatrix} a + qb & 0 \\ 0 & d + pb \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & h + pf - (d + pb)\dfrac{e + qf}{a + qb} \end{pmatrix} \right).$$

$$\left(\begin{pmatrix} a + qb & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & h + pf - (d + pb)\dfrac{e + qf}{a + qb} \end{pmatrix} \right).$$

By dividing the top plane of the matrix with $(a + qb)$ and multiplying the bottom plane by the same value we get

$$\left(\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & (a+qb)(h+pf) - (d+pb)(e+qf) \end{pmatrix}\right).$$

Observe that this last operation corresponds to a linear map with determinant equal to 1. An easy calculation shows that

$$(a + qb)(h + pf) - (d + pb)(e + qf) = \pm\sqrt{D}\,.$$

During all these operations, the invariant I has remained constant. Now we have

$$\begin{aligned} I\left(\begin{pmatrix} a & b \\ c & d \end{pmatrix}, \begin{pmatrix} e & f \\ g & h \end{pmatrix}\right) &= I\left(\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & \pm\sqrt{D} \end{pmatrix}\right) \\ &= \left(\pm\sqrt{D}\right)^{g_1} \cdot I\left(\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\right), \end{aligned}$$

where g_1 is the first coordinate of the index. Let $s = I(w_0) = I(x_1^{(1)}x_1^{(2)}x_1^{(3)} + x_2^{(1)}x_2^{(2)}x_2^{(3)})$. If s is equal to zero then the invariant I vanishes. Assume then that s is nonzero. Since I is a polynomial map, g_1 is an even nonnegative integer. Hence $D = \Gamma(w)$. By symmetry it follows that $g_1 = g_2 = g_3$, which concludes the proof. ∎

4. UMBRAL NOTATION AND THE SIX COVARIANTS

We introduce now umbral notation, which is helpful in describing covariants. We use Greek letters $\alpha, \beta, \dots, \delta$ to denote umbrae. For each of these umbrae there are six variables. Namely, for the umbra α we have the variables $\alpha_1^{(1)}$, $\alpha_2^{(1)}$, $\alpha_1^{(2)}$, $\alpha_2^{(2)}$, $\alpha_1^{(3)}$, and $\alpha_2^{(3)}$. As before we denote the pair $(\alpha_1^{(i)}, \alpha_2^{(i)})$ by $\alpha^{(i)}$.

We define the umbral map U_w, for an element $w \in W$, where we write w as

$$w = \sum_{i=1}^{2}\sum_{j=1}^{2}\sum_{k=1}^{2} a_{i,j,k}\, x_i^{(1)}x_j^{(2)}x_k^{(3)},$$

as follows: Let $\mathbb{C}[\mathbf{x}, \alpha, \dots, \delta]$ be the algebra of polynomials in the variables $x_1^{(1)}, \dots, x_2^{(3)}, \alpha_1^{(1)}, \dots, \delta_2^{(3)}$, and let $\mathbb{C}[\mathbf{x}]$ be the algebra of polynomials in $x_1^{(1)}, \dots, x_2^{(3)}$. The umbral map U_w is a linear map from the algebra $\mathbb{C}[\mathbf{x}, \alpha, \dots, \delta]$ to the algebra $\mathbb{C}[\mathbf{x}]$, which satisfies the following properties:

1. $U_w(p(\mathbf{x}) \cdot q(\alpha) \cdots r(\delta)) = U_w(p(\mathbf{x})) \cdot U_w(q(\alpha)) \cdots U_w(r(\delta))$, where $p(\mathbf{x}) \in \mathbb{C}[\mathbf{x}], q(\alpha) \in \mathbb{C}[\alpha], \ldots, r(\delta) \in \mathbb{C}[\delta]$,

2. $U_w(p(\mathbf{x})) = p(\mathbf{x})$, for all $p(\mathbf{x}) \in \mathbb{C}[\mathbf{x}]$.

3. for each umbra α, we have that $U_w(\alpha_i^{(1)}\alpha_j^{(2)}\alpha_k^{(3)}) = a_{i,j,k}$ and that the umbral map U_w vanishes on any other monomial in the variables belonging to the umbra α.

We may express w in terms of the umbral map U_w:

$$\begin{aligned} &U_w\big((\alpha^{(1)} \mid \mathbf{x}^{(1)})(\alpha^{(2)} \mid \mathbf{x}^{(2)})(\alpha^{(3)} \mid \mathbf{x}^{(3)})\big) \\ &\quad = U_w\left(\sum_{i=1}^{2}\sum_{j=1}^{2}\sum_{k=1}^{2} \alpha_i^{(1)}\alpha_j^{(2)}\alpha_k^{(3)} \cdot x_i^{(1)}x_j^{(2)}x_k^{(3)}\right) \\ &\quad = \sum_{i=1}^{2}\sum_{j=1}^{2}\sum_{k=1}^{2} a_{i,j,k} \cdot x_i^{(1)}x_j^{(2)}x_k^{(3)} = w. \end{aligned}$$

DEFINITION 4.1. A bracket monomial is a product of terms of the form $[\alpha^{(i)}, \beta^{(i)}]$ and $(\alpha^{(i)} \mid \mathbf{x}^{(i)})$.

Observe that a bracket monomial M vanishes when the umbral operator U_w is applied if, for some umbra α in the monomial M, the umbral variable $\alpha^{(i)}$ does not occur exactly once.

PROPOSITION 4.2. *Let M be a bracket monomial with g_i factors of the form $[\alpha^{(i)}, \beta^{(i)}]$ for $i = 1, 2, 3$. Then $U_w(M)$ is a covariant of W with index (g_1, g_2, g_3).*

We call the covariant $U_w(M)$ the associated covariant of the bracket monomial M. Also observe that when there is no factor of the form $(\alpha^{(i)} \mid \mathbf{x}^{(i)})$ in the bracket monomial, then it corresponds to an invariant.

Proof of Proposition 4.2. Let M be a bracket monomial. Assume that

$$M = \prod_{j=1}^{n} \left[\alpha(j)^{(i)}, \beta(j)^{(i)}\right] \cdot \prod_{j=1}^{m} \left(\gamma(j)^{(i)} \mid \mathbf{x}^{(i)}\right),$$

where $\alpha(j)$, $\beta(j)$, and $\gamma(j)$ are umbrae. The claim is that $C(w) = U_w(M)$ is a covariant, where $w = \sum_{i=1}^{2}\sum_{j=1}^{2}\sum_{k=1}^{2} a_{i,j,k} \cdot x_i^{(1)}x_j^{(2)}x_k^{(3)}$.

The element $w \in W$ can be represented umbrally by

$$w = U_w\big((\alpha^{(1)} \mid \mathbf{x}^{(1)})(\alpha^{(2)} \mid \mathbf{x}^{(2)})(\alpha^{(3)} \mid \mathbf{x}^{(3)})\big).$$

Thus $\hat{\phi}_1 \hat{\phi}_2 \hat{\phi}_3 w$ can be represented umbrally by

$$\begin{aligned}\hat{\phi}_1 \hat{\phi}_2 \hat{\phi}_3 w &= U_w\Big(\big(\alpha^{(1)} | \phi_1(\mathbf{x}^{(1)})\big)\big(\alpha^{(2)} | \phi_2(\mathbf{x}^{(2)})\big)\big(\alpha^{(3)} | \phi_3(\mathbf{x}^{(3)})\big)\Big) \\ &= U_w\Big(\big(\phi_1^*(\alpha^{(1)}) \mid \mathbf{x}^{(1)}\big)\big(\phi_2^*(\alpha^{(2)}) \mid \mathbf{x}^{(2)}\big)\big(\phi_3^*(\alpha^{(3)}) \mid \mathbf{x}^{(3)}\big)\Big),\end{aligned}$$

where ϕ_i^* is the adjoint map. To compute the umbral operator of the bracket monomial corresponding to $\hat{\phi}_1 \hat{\phi}_2 \hat{\phi}_3 w$, we replace $\alpha^{(i)}$ with $\phi_i^*(\alpha^{(i)})$ for each umbrae α and $i = 1, 2, 3$. Hence

$$\begin{aligned}C\Big(\hat{\phi}_1 \hat{\phi}_2 \hat{\phi}_3 w\Big) = U_w\Bigg(\prod_{j=1}^{n}\Big[\phi^*\big(\alpha(j)^{(i)}\big), \phi^*\big(\beta(j)^{(i)}\big)\Big] \\ \cdot \prod_{j=1}^{m}\Big(\phi^*\big(\gamma(j)^{(i)}\big) \mid \mathbf{x}^{(i)}\Big)\Bigg).\end{aligned}$$

By Lemma 2.1 we have that

$$\Big[\phi_i^*\big(\alpha(j)^{(i)}\big), \phi_i^*\big(\beta(j)^{(i)}\big)\Big] = \det(\phi_i) \cdot \Big[\alpha(j)^{(i)}, \beta(j)^{(i)}\Big],$$

$$\Big(\phi_i^*\big(\gamma(j)^{(i)}\big) \mid \mathbf{x}^{(i)}\Big) = \Big(\gamma(j)^{(i)} | \phi_i(\mathbf{x}^{(i)})\Big).$$

Recall that g_i is the number of terms in M of the form $[\alpha^{(i)}, \beta^{(i)}]$. Then the expression $C(\hat{\phi}_1 \hat{\phi}_2 \hat{\phi}_3 w)$ is equal to

$$\begin{aligned}C\Big(\hat{\phi}_1 \hat{\phi}_2 \hat{\phi}_3 w\Big) = \prod_{i=1}^{3} \det(\phi_i)^{g_i} \\ \cdot U_w\Bigg(\prod_{j=1}^{n}\Big[\alpha(j)^{(i)}, \beta(j)^{(i)}\Big] \cdot \prod_{j=1}^{m}\Big(\gamma(j)^{(i)} | \phi(\mathbf{x}^{(i)})\Big)\Bigg).\end{aligned}$$

By evaluating the umbral operator we obtain

$$C\Big(\hat{\phi}_1 \hat{\phi}_2 \hat{\phi}_3 w\Big) = \prod_{i=1}^{3} \det(\phi_i)^{g_i} \cdot \hat{\phi}_1 \hat{\phi}_2 \hat{\phi}_3 C(w).$$

Hence we conclude that $C(w) = U_w(M)$ is a covariant. ∎

We now have a tool for expressing covariants of the space W. We give a list of six bracket monomials, $\mathscr{C}$, $\mathscr{S}$, $\mathscr{H}^{(1)}$, $\mathscr{H}^{(2)}$, $\mathscr{H}^{(3)}$, and $\mathscr{I}$, and discuss the associated covariants.

(a)

$$\begin{aligned}\mathscr{C} = -\frac{1}{2} \\ \cdot\big[\alpha^{(1)}, \beta^{(1)}\big]\big[\alpha^{(2)}, \beta^{(2)}\big]\big[\alpha^{(3)}, \gamma^{(3)}\big]\big[\beta^{(3)}, \delta^{(3)}\big]\big[\gamma^{(1)}, \delta^{(1)}\big]\big[\gamma^{(2)}, \delta^{(2)}\big].\end{aligned}$$

The associated covariant is the invariant Γ, described in Proposition 3.4. To see this, first observe that the two invariants have the same index, namely $(2,2,2)$. Hence by Proposition 3.4 we know that they are equal up to a constant. All we need to do is to verify that this constant is actually equal to one. To do this, evaluate both invariants on the element $w_0 = x_1^{(1)}x_1^{(2)}x_1^{(3)} + x_2^{(1)}x_2^{(2)}x_2^{(3)}$. Clearly, $\Gamma(w_0) = 1$. The umbral calculation goes as follows.

$$U_{w_0}(\mathscr{C}) = -\frac{1}{2} \cdot U_{w_0}\Big(\big(\alpha_1^{(1)}\beta_2^{(1)} - \alpha_2^{(1)}\beta_1^{(1)}\big)\big(\alpha_1^{(2)}\beta_2^{(2)} - \alpha_2^{(2)}\beta_1^{(2)}\big) \times \big(\alpha_1^{(3)}\gamma_2^{(3)} - \alpha_2^{(3)}\gamma_1^{(3)}\big). \cdot \big[\beta^{(3)}, \delta^{(3)}\big]\big[\gamma^{(1)}, \delta^{(1)}\big]\big[\gamma^{(2)}, \delta^{(2)}\big]\Big). \tag{2}$$

The only monomials of the umbra α which do not vanish are the two monomials $\alpha_1^{(1)}\alpha_1^{(2)}\alpha_1^{(3)}$ and $\alpha_2^{(1)}\alpha_2^{(2)}\alpha_2^{(3)}$. These instead evaluate to 1. Thus, (2) becomes

$$-\frac{1}{2} \cdot U_{w_0}\Big(\big(\beta_2^{(1)}\beta_2^{(2)}\gamma_2^{(3)} - \beta_1^{(1)}\beta_1^{(2)}\gamma_1^{(3)}\big)\big[\beta^{(3)}, \delta^{(3)}\big]\big[\gamma^{(1)}, \delta^{(1)}\big]\big[\gamma^{(2)}, \delta^{(2)}\big]\Big).$$

By a similar argument on the umbra δ the umbral expressions evaluate to

$$\begin{aligned} U_{w_0}(\mathscr{C}) &= -\frac{1}{2} \cdot U_{w_0}\Big(\big(\beta_2^{(1)}\beta_2^{(2)}\gamma_2^{(3)} - \beta_1^{(1)}\beta_1^{(2)}\gamma_1^{(3)}\big)\big(\beta_1^{(3)}\gamma_1^{(1)}\gamma_1^{(2)} - \beta_2^{(3)}\gamma_2^{(1)}\gamma_2^{(2)}\big)\Big) \\ &= -\frac{1}{2} \cdot U_{w_0}\big(\beta_2^{(1)}\beta_2^{(2)}\gamma_2^{(3)}\beta_1^{(3)}\gamma_1^{(1)}\gamma_1^{(2)} - \beta_2^{(1)}\beta_2^{(2)}\gamma_2^{(3)}\beta_2^{(3)}\gamma_2^{(1)}\gamma_2^{(2)} \\ &\qquad - \beta_1^{(1)}\beta_1^{(2)}\gamma_1^{(3)}\beta_1^{(3)}\gamma_1^{(1)}\gamma_1^{(2)} + \beta_1^{(1)}\beta_1^{(2)}\gamma_1^{(3)}\beta_2^{(3)}\gamma_2^{(1)}\gamma_2^{(2)}\big) \\ &= -\frac{1}{2} \cdot U_{w_0}\big(-\beta_2^{(1)}\beta_2^{(2)}\beta_2^{(3)}\gamma_2^{(1)}\gamma_2^{(2)}\gamma_2^{(3)} - \beta_1^{(1)}\beta_1^{(2)}\beta_1^{(3)}\gamma_1^{(1)}\gamma_1^{(2)}\gamma_1^{(3)}\big) = 1. \end{aligned}$$

(b)

$$\mathscr{S} = -\big[\alpha^{(1)}, \beta^{(1)}\big]\big[\alpha^{(2)}, \beta^{(2)}\big]\big[\alpha^{(3)}, \gamma^{(3)}\big]\big(\beta^{(3)} \mid \mathbf{x}^{(3)}\big)\big(\gamma^{(1)} \mid \mathbf{x}^{(1)}\big)\big(\gamma^{(2)} \mid \mathbf{x}^{(2)}\big).$$

The covariant $U_w(\mathscr{S})$ is denoted by S. Observe that S maps W into itself. The index of this covariant is $(1,1,1)$.

(c)

$$\mathscr{H}^{(1)} = \frac{1}{2}\big[\alpha^{(2)}, \beta^{(2)}\big]\big[\alpha^{(3)}, \beta^{(3)}\big]\big(\alpha^{(1)} \mid \mathbf{x}^{(1)}\big)\big(\beta^{(1)} \mid \mathbf{x}^{(1)}\big).$$

Denote the associated covariant to this bracket monomial by $H^{(1)}$. We call this covariant the Hessian, and it has index $(0,1,1)$. The Hessian may also be described by

$$H^{(1)}(w) = \begin{vmatrix} \dfrac{\partial^2 w}{\partial x_1^{(2)}\,\partial x_1^{(3)}} & \dfrac{\partial^2 w}{\partial x_1^{(2)}\,\partial x_2^{(3)}} \\ \dfrac{\partial^2 w}{\partial x_2^{(2)}\,\partial x_1^{(3)}} & \dfrac{\partial^2 w}{\partial x_2^{(2)}\,\partial x_2^{(3)}} \end{vmatrix}.$$

This may be seen to be true by the following computation.

$$\begin{aligned} &U_w\Big(\big[\alpha^{(2)}, \beta^{(2)}\big]\big[\alpha^{(3)}, \beta^{(3)}\big]\alpha_i^{(1)}\beta_j^{(1)}\Big) \\ &\quad = U_w\Big(\big(\alpha_1^{(2)} \cdot \beta_2^{(2)} - \alpha_2^{(2)} \cdot \beta_1^{(2)}\big) \cdot \big(\alpha_1^{(3)} \cdot \beta_2^{(3)} - \alpha_2^{(3)} \cdot \beta_1^{(3)}\big)\alpha_i^{(1)}\beta_j^{(1)}\Big) \\ &\quad = U_w\Big(\alpha_i^{(1)}\alpha_1^{(2)}\alpha_1^{(3)}\beta_j^{(1)}\beta_2^{(2)}\beta_2^{(3)} - \alpha_i^{(1)}\alpha_1^{(2)}\alpha_2^{(3)}\beta_j^{(1)}\beta_2^{(2)}\beta_1^{(3)} \\ &\qquad\qquad - \alpha_i^{(1)}\alpha_2^{(2)}\alpha_1^{(3)}\beta_j^{(1)}\beta_1^{(2)}\beta_2^{(3)} + \alpha_i^{(1)}\alpha_2^{(2)}\alpha_2^{(3)}\beta_j^{(1)}\beta_1^{(2)}\beta_1^{(3)}\Big) \\ &\quad = a_{i,1,1}a_{j,2,2} - a_{i,1,2}a_{j,2,1} - a_{i,2,1}a_{j,1,2} + a_{i,2,2}a_{j,1,1}. \end{aligned}$$

Hence the associated covariant to the bracket monomial $\mathscr{H}^{(1)}$ is given by

$$\begin{aligned} &\frac{1}{2}\sum_{i=1}^{2}\sum_{j=1}^{2}\left(a_{i,1,1}a_{j,2,2} - a_{i,1,2}a_{j,2,1} - a_{i,2,1}a_{j,1,2} + a_{i,2,2}a_{j,1,1}\right)x_i^{(1)}x_j^{(1)} \\ &\quad = \frac{1}{2}\sum_{i=1}^{2}\sum_{j=1}^{2}\left(\begin{vmatrix} a_{i,1,1}x_i^{(1)} & a_{i,1,2}x_i^{(1)} \\ a_{j,2,1}x_j^{(1)} & a_{j,2,2}x_j^{(1)} \end{vmatrix} + \begin{vmatrix} a_{i,2,2}x_i^{(1)} & a_{i,2,1}x_i^{(1)} \\ a_{j,1,2}x_j^{(1)} & a_{j,1,1}x_j^{(1)} \end{vmatrix}\right) \\ &\quad = \frac{1}{2}\begin{vmatrix} a_{1,1,1}x_1^{(1)} + a_{2,1,1}x_2^{(1)} & a_{1,1,2}x_1^{(1)} + a_{2,1,2}x_2^{(1)} \\ a_{1,2,1}x_1^{(1)} + a_{2,2,1}x_2^{(1)} & a_{1,2,2}x_1^{(1)} + a_{2,2,2}x_2^{(1)} \end{vmatrix} \\ &\qquad + \frac{1}{2}\begin{vmatrix} a_{1,2,2}x_1^{(1)} + a_{2,2,2}x_2^{(1)} & a_{1,2,1}x_1^{(1)} + a_{2,2,1}x_2^{(1)} \\ a_{1,1,2}x_1^{(1)} + a_{2,1,2}x_2^{(1)} & a_{1,1,1}x_1^{(1)} + a_{2,1,1}x_2^{(1)} \end{vmatrix}. \end{aligned}$$

These two determinants are equal. Moreover, they are equal to the determinant given for the Hessian, since

$$\frac{\partial^2 w}{\partial x_i^{(2)}\,\partial x_j^{(3)}} = a_{1,i,j}x_1^{(1)} + a_{2,i,j}x_2^{(1)}.$$

(d)

$$\mathscr{H}^{(2)} = \frac{1}{2}\big[\alpha^{(1)}, \beta^{(1)}\big]\big[\alpha^{(3)}, \beta^{(3)}\big]\big(\alpha^{(2)} \mid \mathbf{x}^{(2)}\big)\big(\beta^{(2)} \mid \mathbf{x}^{(2)}\big).$$

Similarly, this is also a Hessian, but in the second variable. It is denoted by $H^{(2)}$ and has index $(1,0,1)$.

(e)

$$\mathscr{H}^{(3)} = \frac{1}{2}\left[\alpha^{(1)}, \beta^{(1)}\right]\left[\alpha^{(2)}, \beta^{(2)}\right]\left(\alpha^{(3)} \mid \mathbf{x}^{(3)}\right)\left(\beta^{(3)} \mid \mathbf{x}^{(3)}\right).$$

This Hessian is denoted by $H^{(3)}$ and has index $(1,1,0)$.

(f)

$$\mathscr{I} = \left(\alpha^{(1)} \mid \mathbf{x}^{(1)}\right)\left(\alpha^{(2)} \mid \mathbf{x}^{(2)}\right)\left(\alpha^{(3)} \mid \mathbf{x}^{(3)}\right).$$

This is the identity covariant, Id. That is, $\mathrm{Id}(w) = w$.

For all the bracket monomials in examples a–f we have chosen the constant such that the coefficients of the associated covariant will have no common factor.

5. CANONICAL FORMS OF TWO BY TWO BY TWO MATRICES

PROPOSITION 5.1. *An element $w \in W$ can be written in exactly one of the following seven forms*

$$\begin{aligned}
&pqr + stu,\\
&sqr + ptr + pqu,\\
&pqr + ptu,\\
&pqr + squ,\\
&pqr + str,\\
&pqr,\\
&0,
\end{aligned}$$

where $p, s \in W_1$, $q, t \in W_2$, $r, u \in W_3$, p and s are linearly independent, q and t are linearly independent, and r and u are linearly independent.

We begin proving this proposition with the following two lemmas.

LEMMA 5.2. *An element $w \in W$ can be written in at least one of two forms, $pqr + stu$ and $sqr + ptr + pqu$ where $p, s \in W_1$, $q, t \in W_2$, and $r, u \in W_3$.*

Proof. Observe that these two forms are invariant under changes of variables. To prove that such changes of variables are possible is quite similar to the proof of Proposition 3.4. However, we need to be more careful, since we do not begin with a generic element of W, but rather an arbitrary element of W.

We claim that by a change of variables we can transform any two by two by two matrix to a two by two by two matrix where two adjacent entries are equal to zero. Following the proof of Proposition 3.4, we can do this if $af - be \neq 0$ and $df - bh \neq 0$. Without loss of generality we can assume that $af - be = 0$, since the case $df - bh = 0$ is symmetric. Since $af - be = 0$, the two vectors (a, e) and (b, f) are linearly dependent. Now, by a change of variables we can make one of the vectors vanish, and thus there are two adjacent entries in the matrix that are equal to zero. Hence we can assume that the matrix looks like

$$\left(\begin{pmatrix} a' & b' \\ 0 & d' \end{pmatrix}, \begin{pmatrix} e' & f' \\ 0 & h' \end{pmatrix} \right).$$

If the vectors (a', e') and (d', h') are linearly independent, then by two more changes of variables we can eliminate b' and f'. Thus we have the matrix

$$\left(\begin{pmatrix} a' & 0 \\ 0 & d' \end{pmatrix}, \begin{pmatrix} e' & 0 \\ 0 & h' \end{pmatrix} \right).$$

This matrix corresponds to the element

$$a' x_1^{(1)} x_1^{(2)} x_1^{(3)} + d' x_1^{(1)} x_2^{(2)} x_2^{(3)} + e' x_2^{(1)} x_1^{(2)} x_1^{(3)} + h' x_2^{(1)} x_2^{(2)} x_2^{(3)},$$

which can be written as

$$\left(a' x_1^{(1)} + e' x_2^{(1)}\right) x_1^{(2)} x_1^{(3)} + \left(d' x_1^{(1)} + h' x_2^{(1)}\right) x_2^{(2)} x_2^{(3)}.$$

This is the first desired form.

It remains to consider the case when the vectors (a', e') and (d', h') are linearly dependent. If both are equal to zero, then the matrix is trivially one of the two stated forms. Hence assume that $(a', e') \neq \mathbf{0}$, so we can write $(d', h') = j \cdot (a', e')$. Now the matrix corresponds to the element

$$\begin{aligned} &a' x_1^{(1)} x_1^{(2)} x_1^{(3)} + b' x_1^{(1)} x_1^{(2)} x_2^{(3)} + j \cdot a' x_1^{(1)} x_2^{(2)} x_2^{(3)} \\ &\qquad + e' x_2^{(1)} x_1^{(2)} x_1^{(3)} + f' x_2^{(1)} x_1^{(2)} x_2^{(3)} + j \cdot e' x_2^{(1)} x_2^{(2)} x_2^{(3)}, \end{aligned}$$

which we can write as

$$\begin{aligned} &\left(b' x_1^{(1)} + f' x_2^{(1)}\right) x_1^{(2)} x_2^{(3)} + \left(a' x_1^{(1)} + e' x_2^{(1)}\right)\left(j x_2^{(2)}\right) x_2^{(3)} \\ &\qquad + \left(a' x_1^{(1)} + e' x_2^{(1)}\right) x_1^{(2)} x_1^{(3)}, \end{aligned}$$

which is in the second desired form. ∎

Define a bracket on the linear space W_i by

$$\left[a_1 x_1^{(i)} + a_2 x_2^{(i)}, b_1 x_1^{(i)} + b_2 x_2^{(i)}\right] = a_1 b_2 - a_2 b_1.$$

LEMMA 5.3. *The six different covariants in Section* 4 *evaluated on the seven canonical forms in Proposition* 5.1 *are given by the table*:

	$pqr + stu$	$sqr + ptr + pqu$
Γ	$[p,s]^2[q,t]^2[r,u]^2$	0
S	$[p,s][q,t][r,u](pqr - stu)$	$2[p,s][q,t][r,u]pqr$
$H^{(1)}$	$[q,t][r,u]ps$	$-[q,t][r,u]p^2$
$H^{(2)}$	$[p,s][r,u]qt$	$-[p,s][r,u]q^2$
$H^{(3)}$	$[p,s][q,t]ru$	$-[p,s][q,t]r^2$
Id	$pqr + stu$	$sqr + ptr + pqu$

	$pqr + ptu$	$pqr + squ$	$pqr + str$	pqr	0
Γ	0	0	0	0	0
S	0	0	0	0	0
$H^{(1)}$	$[q,t][r,u]p^2$	0	0	0	0
$H^{(2)}$	0	$[p,s][r,u]q^2$	0	0	0
$H^{(3)}$	0	0	$[p,s][q,t]r^2$	0	0
Id	$pqr + ptu$	$pqr + squ$	$pqr + str$	pqr	0

Proof. To do these six times seven calculations, evaluate the covariant C on the element $w_0 = x_1^{(1)}x_1^{(2)}x_1^{(3)} + x_1^{(1)}x_2^{(2)}x_2^{(3)}$. Then, by a suitable change of variables, one obtains the value $C(pqr + stu)$.

If we know the value of $C(pqr + stu)$, we may use this information to find $C(sqr + ptr + pqu)$. Since the covariant is a polynomial map, and hence continuous, we may apply the following limit. Let $p' = -Rp$, $q' = -Rq$, $r' = -Rr$, $s' = Rp + R^{-2}s$, $t' = Rq + R^{-2}t$, and $u' = Rr + R^{-2}u$. Then we have

$$\begin{aligned}
&\lim_{R\to\infty} p'q'r' + s't'u' \\
&\quad= \lim_{R\to\infty} (-Rp)(-Rq)(-Rr) \\
&\qquad + \left(Rp + R^{-2}s\right)\left(Rq + R^{-2}t\right)\left(Rr + R^{-2}u\right) \\
&\quad= \lim_{R\to\infty} (sqr + ptr + pqu) \\
&\qquad + R^{-3}\cdot(ptu + squ + str) + R^{-6}\cdot stu \\
&\quad= sqr + ptr + pqu
\end{aligned}$$

to find $C(sqr + ptr + pqu)$. For instance, if we have that

$$S(pqr + stu) = [p, s][q, t][r, u](pqr - stu), \tag{3}$$

then we may compute $S(sqr + ptr + pqu)$ by

$$\begin{aligned}
S(sqr + ptr + pqu)& \\
&= \lim_{R \to \infty} S(p'q'r' + s't'u') \\
&= \lim_{R \to \infty} [p', s'][q', t'][r', u'] \cdot (p'q'r' - s't'u') \\
&= \lim_{R \to \infty} [p', s'][q', t'][r', u'] \cdot \big((-Rp)(-Rq)(-Rr) \\
&\qquad\qquad - (Rp + R^{-2}s)(Rq + R^{-2}t)(Rr + R^{-2}u)\big).
\end{aligned}$$

But $[p', s'] = [-Rp, Rp + R^{-2}s] = [-Rp, R^{-2}s] = -R^{-1} \cdot [p, s]$. Hence this limit is equal to

$$\begin{aligned}
&\lim_{R \to \infty} (-R)^{-3} \cdot [p, s][q, t][r, u] \cdot \big(-2 \cdot R^3 \cdot pqr - O(1)\big) \\
&\qquad = 2 \cdot [p, s][q, t][r, u] \cdot pqr.
\end{aligned}$$

Similarly, we apply the four limits

$$\lim_{\epsilon \to 0} pqr + (p + \epsilon s)tu = pqr + ptu,$$

$$\lim_{\epsilon \to 0} pqr + s(q + \epsilon t)u = pqr + squ,$$

$$\lim_{\epsilon \to 0} pqr + st(r + \epsilon u) = pqr + str,$$

$$\lim_{\epsilon \to 0} pqr + (\epsilon s)tu = pqr,$$

to find $C(pqr + ptu)$, $C(pqr + squ)$, $C(pqr + str)$, and $C(pqr)$ for a covariant C. ∎

Now we are able to prove Proposition 5.1.

Proof of Proposition 5.1. By Lemma 5.2 we know that an element $w \in W$ may be written in the forms $pqr + stu$ and $sqr + ptr + pqu$, where $p, s \in W_1$, $q, t \in W_2$, and $r, u \in W_3$.

If any of the pairs $\{p, s\}$, $\{q, t\}$, or $\{r, u\}$ are linearly dependent then it is easy to see that we can reduce the expressions further to one of the following five forms: $pqr + ptu$, $pqr + squ$, $pqr + str$, pqr, or 0. Hence we know that every $w \in W$ can be written in at least one of the seven forms.

It remains to prove that every w can be written in exactly one of the seven forms. Assume that w may be written in two of the forms. Then by

Lemma 5.3 we may find a covariant C which is zero on one of the forms and nonzero on the second. This contradicts the assumption. ∎

COROLLARY 5.4. *The elements w in W that may be written in the form $pqr + stu$, where $p, s \in W_1$, $q, t \in W_2$, $r, u \in W_3$, p and s are linearly independent, q and t are linearly independent, and r and u are linearly independent, form a dense set in W under the Euclidean topology. That is, the form $pqr + stu$ is a generic canonical form.*

Hence, if two continuous functions, with domain W, agree on the elements of the form $pqr + stu$, then these two functions agree on all of W. With this observation Corollaries 5.5 and 5.6 will follow.

COROLLARY 5.5. *The following two algebraic relations hold*:

$$S^2 = \Gamma \cdot \mathrm{Id}^2 - 4 \cdot H^{(1)} \cdot H^{(2)} \cdot H^{(3)},$$
$$\Gamma = \Delta^{(i)} \circ H^{(i)},$$

where $\Delta^{(i)}$ is the discriminant on quadratic polynomials in the two variables $x_1^{(i)}$ and $x_2^{(i)}$.

Proof. By Corollary 5.4 we may prove the first identity by observing:

$$\begin{aligned} S(pqr + stu)^2 &\\ &= ([p,s][q,t][r,u](pqr - stu))^2 \\ &= ([p,s][q,t][r,u](pqr + stu))^2 \\ &\quad - 4[p,s]^2[q,t]^2[r,u]^2 psqtru \\ &= \Gamma(pqr + stu) \cdot \mathrm{Id}(pqr + stu)^2 \\ &\quad - 4 \cdot H^{(1)}(pqr + stu) \cdot H^{(2)}(pqr + stu) \cdot H^{(3)}(pqr + stu). \end{aligned}$$

We only prove the second identity when $i = 1$. By the same argument it is enough to see that

$$\begin{aligned} \Gamma(pqr + stu) &= [p,s]^2[q,t]^2[r,u]^2 \\ &= [[q,t][r,u] \cdot p, s]^2 \\ &= \Delta^{(1)}([q,t][r,u] \cdot p \cdot s) \\ &= \Delta^{(1)}(H^{(1)}(pqr + stu)). \quad ∎ \end{aligned}$$

Recall that the covariant S maps W into W. Thus we can consider the composition of a covariant C with the covariant S.

COROLLARY 5.6. *The following three identities hold*:

$$\begin{aligned} \Gamma \circ S &= \Gamma^3, \\ S \circ S &= -\Gamma^2 \cdot \mathrm{Id}, \\ H^{(i)} \circ S &= -\Gamma \cdot H^{(i)}. \end{aligned}$$

Proof. We will only give the proof of the first identity. By Corollary 5.4 it is enough to evaluate $\Gamma(S(w))$ when w is on the form $pqr + stu$. Let $p' = [p,s][q,t][r,u]p$ and $s' = -[p,s][q,t][r,u]s$. Then we have

$$\begin{aligned} &\Gamma(S(pqr + stu)) \\ &\quad = \Gamma([p,s][q,t][r,u] \cdot (pqr - stu)) \\ &\quad = \Gamma(p'qr + s'tu) \\ &\quad = [p',s']^2[q,t]^2[r,u]^2 \\ &\quad = [[p,s][q,t][r,u]p, -[p,s][q,t][r,u]s]^2[q,t]^2[r,u]^2 \\ &\quad = [p,s]^6[q,t]^6[r,u]^6 \\ &\quad = \Gamma(pqr + stu)^3. \quad \blacksquare \end{aligned}$$

We conclude our discussion by presenting two partially ordered sets. One poset is defined on the covariants of the linear space W and the other poset on the canonical forms of W. The partial orders are described in the following definition.

DEFINITION 5.7. 1. For two covariants C and C' we say that $C \leq C'$ if $C'(w) = 0$ implies $C(w) = 0$ for all $w \in W$.

2. For two forms f and f' we say that $f \leq f'$ if

$$\overline{\{w \in W : w \textit{ is of the form } f\}} \subseteq \overline{\{w \in W : w \textit{ is of the form } f'\}},$$

where the closure is the topological closure in the Euclidean topology on W.

In Fig. 1 the two partial orders are presented. Observe that 1 and 0 are constant covariants and $\varnothing$ is the form in which no element can be written.

6. RELATIONS WITH THE BINARY CUBIC

In this section we show how binary cubics relate to two by two by two matrices by studying their covariants and their canonical forms. Readers are also referred to [8].

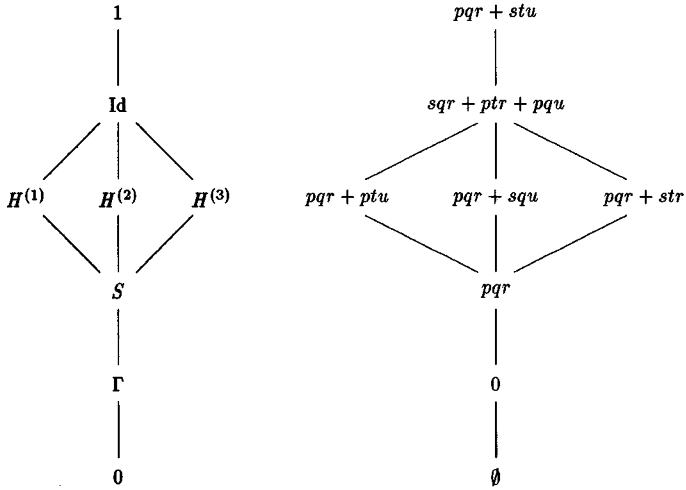

FIG. 1. The two partial orders on canonical forms and covariants of two by two by two matrices.

Recall that V_3 is the linear subspace of $\mathbb{C}[\mathbf{y}] = \mathbb{C}[y_1, y_2]$ consisting of all homogeneous elements of degree 3. An element p of $\mathbb{C}[\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \mathbf{x}^{(3)}]$ is called *symmetric* if

$$p\left(\mathbf{x}^{(\sigma(1))}, \mathbf{x}^{(\sigma(2))}, \mathbf{x}^{(\sigma(3))}\right) = p\left(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \mathbf{x}^{(3)}\right),$$

for all permutations σ of the set $\{1, 2, 3\}$. Recall that W is isomorphic to the linear space of two by two by two matrices, see Section 3. The symmetric two by two by two matrices are of the form

$$\left(\begin{pmatrix} a_0 & a_1 \\ a_1 & a_2 \end{pmatrix}, \begin{pmatrix} a_1 & a_2 \\ a_2 & a_3 \end{pmatrix} \right).$$

Define the algebra homomorphism $\Omega\colon \mathbb{C}[\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \mathbf{x}^{(3)}] \to \mathbb{C}[\mathbf{y}]$ by $\Omega(x_j^{(i)}) = y_j$. Moreover, define a linear map $\omega\colon V_3 \to W$ by

$$\omega\left(a_0 y_1^3 + 3a_1 y_1^2 y_2 + 3a_2 y_1 y_2^2 + a_3 y_2^3\right) = \left(\begin{pmatrix} a_0 & a_1 \\ a_1 & a_2 \end{pmatrix}, \begin{pmatrix} a_1 & a_2 \\ a_2 & a_3 \end{pmatrix} \right).$$

Note that for all $w \in V_3$ we have that $\Omega(\omega(w)) = w$.

LEMMA 6.1. *If C is a covariant of W with index (g_1, g_2, g_3), then the map $\Omega \circ C \circ \omega$ is a covariant of V_3 with index $g_1 + g_2 + g_3$.*

Proof. For $\phi(y_j) = a_{j,1}y_1 + a_{j,2}y_2$, define $\phi_i\colon W_i \to W_i$ by $\phi_i(x_j^{(i)}) = a_{j,1}x_1^{(i)} + a_{j,2}x_2^{(i)}$. Directly we have that $\det(\phi_i) = \det(\phi)$. Moreover, we have the two following identities:

$$\Omega\left(\hat{\phi}_1\hat{\phi}_2\hat{\phi}_3 p(\mathbf{x})\right) = \hat{\phi}\Omega(p(\mathbf{x})),$$

$$\omega\left(\hat{\phi}w\right) = \hat{\phi}_1\hat{\phi}_2\hat{\phi}_3\omega(w).$$

The proofs of both of them are straightforward calculations.

Consider now the map $\Omega \circ C \circ \omega$.

$$\begin{aligned}
\left(\Omega \circ C \circ \omega \circ \hat{\phi}\right)(w) &= \left(\Omega \circ C \circ \left(\hat{\phi}_1\hat{\phi}_2\hat{\phi}_3\right)\right)(\omega(w)) \\
&= \Omega\left(\prod_{i=1}^{3}\det(\phi_i)^{g_i} \cdot \hat{\phi}_1\hat{\phi}_2\hat{\phi}_3 C(\omega(w))\right) \\
&= \prod_{i=1}^{3}\det(\phi_i)^{g_i} \cdot \Omega\left(\hat{\phi}_1\hat{\phi}_2\hat{\phi}_3 C(\omega(w))\right) \\
&= \prod_{i=1}^{3}\det(\phi)^{g_i} \cdot \hat{\phi}\Omega(C(\omega(w))) \\
&= \det(\phi)^{g_1+g_2+g_3} \cdot \hat{\phi}(\Omega \circ C \circ \omega)(w).
\end{aligned}$$

Thus $\Omega \circ C \circ \omega$ is a covariant of V_3 of index $g_1 + g_2 + g_3$. ∎

Let us now see to what the covariants Γ, S, $H^{(1)}$, $H^{(2)}$, $H^{(3)}$, and Id of two by two by two matrices correspond for binary cubics. Let $w = a_0y_1^3 + 3a_1y_1^2y_2 + 3a_2y_1y_2^2 + a_3y_2^3$.

a. The covariant given by $\Omega \circ \Gamma \circ \omega$ is called the discriminant. It is computed by

$$\Delta(w) = (a_0a_3 - a_1a_2)^2 - 4\left(a_0a_2 - a_1^2\right)\left(a_1a_3 - a_2^2\right).$$

b. The covariant $\Omega \circ S \circ \omega$ of binary cubics is called the Jacobian and is denoted by T.

c–e. For the three Hessians we have that

$$\Omega \circ H^{(1)} \circ \omega = \Omega \circ H^{(2)} \circ \omega = \Omega \circ H^{(3)} \circ \omega.$$

Denote this covariant by H. It is also described by

$$H(w) = \frac{1}{36} \cdot \begin{vmatrix} \dfrac{\partial^2 w}{\partial y_1^2} & \dfrac{\partial^2 w}{\partial y_1 \, \partial y_2} \\ \dfrac{\partial^2 w}{\partial y_2 \, \partial y_1} & \dfrac{\partial^2 w}{\partial y_2^2} \end{vmatrix}.$$

f. Clearly $\Omega \circ \mathrm{Id} \circ \omega = \mathrm{Id}$, the identity map on V_3.

A binary cubic can be written in one of the following forms: $p^3 + q^3$, $3p^2q$, p^3, and 0, where $p, q \in \mathrm{span}(y_1, y_2)$ and p and q are linearly independent.

We conclude the discussion about covariants and canonical forms of the binary cubic by presenting the following table.

	$p^3 + q^3$	$3p^2q$	p^3	0
Δ	$[p,q]^6$	0	0	0
T	$[p,q]^3(p^3 - q^3)$	$2[p,q]^3p^3$	0	0
H	$[p,q]^2pq$	$-[p,q]^2p^2$	0	0
Id	$p^3 + q^3$	$3p^2q$	p^3	0

The Fundamental Theorem of Algebra says that a binary cubic, w, may be written as the product of three linear forms, s, t, and u. We present here the covariants evaluated on the binary cubic stu.

PROPOSITION 6.2. *We have that*

$$\Delta(stu) = -\frac{1}{27} \cdot [s,t]^2[t,u]^2[u,s]^2,$$

$$H(stu) = \frac{1}{9} \cdot e_2([t,u]s, [u,s]t, [s,t]u),$$

$$T(stu) = \frac{1}{27} \cdot ([t,u]s - [u,s]t) \cdot ([u,s]t - [s,t]u) \cdot ([s,t]u - [t,u]s),$$

where e_2 is the elementary symmetric function of degree two, that is, $e_2(x, y, z) = xy + xz + yz$.

Proof. We know that the factorization is unique up to order of the factors and multiplication by scalars. It is easy to see that the three expressions above are symmetric in s, t, and u. Moreover, if the expression holds for s, t, and u, they also hold for $\alpha \cdot s$, $\beta \cdot t$, and $\gamma \cdot u$, where

$\alpha \cdot \beta \cdot \gamma = 1$. Hence it is enough to prove these expressions for one factorization of a binary cubic. Let $\omega = \exp(2\pi i/3)$, that is, a third root of unity. Since $p^3 + q^3 = (p + q) \cdot (\omega \cdot p + \omega^2 \cdot q) \cdot (\omega^2 \cdot p + \omega \cdot q)$ is a generic canonical form for the binary cubic, it is enough to prove our identities for this factorization.

Let $s = p + q$, $t = \omega \cdot p + \omega^2 \cdot q$, and $u = \omega^2 \cdot p + \omega \cdot q$. Then it is easy to calculate that $[s, t] = [t, u] = [u, s] = (\omega^2 - \omega) \cdot [p, q]$. Since $\omega^2 - \omega = -\sqrt{3} \cdot i$, we obtain

$$-\frac{1}{27} \cdot [s, t]^2 [t, u]^2 [u, s]^2 = -\frac{1}{27} \cdot \left(-\sqrt{3} \cdot i\right)^6 \cdot [p, q]^6 = [p, q]^6,$$

which proves the first identity. It is easy to see that

$$\begin{aligned}(s \cdot t + t \cdot u + u \cdot s) &= \big((1 + \omega + \omega^2) \cdot p^2 + 3 \cdot (\omega + \omega^2) \cdot pq \\ &\qquad + (1 + \omega + \omega^2) \cdot q^2\big) \\ &= -3 \cdot pq.\end{aligned}$$

The second identity follows now by

$$\begin{aligned}&\frac{1}{9} \cdot e_2([t, u]s, [u, s]t, [s, t]u) \\ &\quad = \frac{1}{9} \cdot \left(-\sqrt{3} \cdot i\right)^2 \cdot [p, q]^2 \cdot (s \cdot t + t \cdot u + u \cdot s) \\ &\quad = -\frac{1}{3} \cdot [p, q]^2 \cdot (-3) \cdot pq = [p, q]^2 \cdot pq.\end{aligned}$$

Let $\tau = \exp(\pi i/6)$, that is, a twelfth root of unity. Then we calculate that

$$\begin{aligned} s - t &= \sqrt{3} \cdot \left(\tau^{11} \cdot p + \tau \cdot q\right), \\ t - u &= \sqrt{3} \cdot \left(\tau^{3} \cdot p + \tau^{9} \cdot q\right), \\ u - s &= \sqrt{3} \cdot \left(\tau^{7} \cdot p + \tau^{5} \cdot q\right). \end{aligned}$$

Multiplying these three identities together, we obtain

$$\begin{aligned}&(s - t) \cdot (t - u) \cdot (u - s) \\ &\quad = \left(\sqrt{3}\right)^3 \cdot \tau^9 \cdot \left(p + \tau^2 \cdot q\right) \cdot \left(p + \tau^6 \cdot q\right) \cdot \left(p + \tau^{10} \cdot q\right) \\ &\quad = -\sqrt{27} \cdot i \cdot \left(p^3 - q^3\right).\end{aligned}$$

Now

$$\frac{1}{27} \cdot ([t,u]s - [u,s]t) \cdot ([u,s]t - [s,t]u) \cdot ([s,t]u - [t,u]s)$$
$$= \frac{1}{27} \cdot (-\sqrt{3} \cdot i)^3 \cdot [p,q]^3 \cdot (s - t) \cdot (t - u) \cdot (u - s)$$
$$= \frac{1}{\sqrt{27}} \cdot i \cdot [p,q]^3 \cdot (-\sqrt{27} \cdot i) \cdot (p^3 - q^3) = [p,q]^3 \cdot (p^3 - q^3),$$

and the third identity is proved. ∎

By Vandermonde identity the Jacobian T may also be written as

$$T(stu) = \frac{1}{27} \cdot \det \begin{pmatrix} 1 & 1 & 1 \\ [t,u] \cdot s & [u,s] \cdot t & [s,t] \cdot u \\ [t,u]^2 \cdot s^2 & [u,s]^2 \cdot t^2 & [s,t]^2 \cdot u^2 \end{pmatrix}.$$

7. RELATIONS WITH SKEW-TENSORS OF STEP THREE

Let U be a six-dimensional linear space spanned by the six variables $x_1^{(1)}, \ldots, x_2^{(3)}$. That is, $U \cong W_1 \oplus W_2 \oplus W_3$. In this section we consider the linear space $\mathrm{Ext}_3(U)$, consisting of a skew-symmetric tensor of step 3, and see how the skew-symmetric tensor relates to two by two by two matrices.

PROPOSITION 7.1 (Grosshans, Rota, and Stein [5]). *An element $w \in \mathrm{Ext}_3(U)$, where $\dim(U) = 6$, can be written in exactly one of the five forms*

$$p \wedge q \wedge r + s \wedge t \wedge u,$$
$$s \wedge q \wedge r + p \wedge t \wedge r + p \wedge q \wedge u,$$
$$p \wedge q \wedge r + p \wedge t \wedge u,$$
$$p \wedge q \wedge r,$$
$$0,$$

where $p, q, r, s, t, u \in U$, and the elements p, q, r, s, t, and u are linearly independent. Moreover, a generic element $w \in \mathrm{Ext}_3(U)$ can be written in the first form, namely $p \wedge q \wedge r + s \wedge t \wedge u$.

For a proof see [5, pp. 69–72]. Grosshans *et al.* also present three covariants, C_1, C_2, and C_3, and show how they relate to the above canonical forms. Their relations are that C_1 vanishes on the two last forms, C_2 vanishes on the three last forms, and C_3 is only nonzero on the first form. We will do a more explicit calculation in the spirit of Lemma 5.3. To do so, we multiply these three covariants with appropriate constants to make them easier to handle.

Consider the following covariants:

$$D_1 = -C_1, \qquad D_2 = -\frac{1}{3} \cdot C_2, \qquad \text{and} \qquad D_3 = -\frac{1}{6} \cdot C_3.$$

Observe that

$$D_1 : \mathrm{Ext}_3(U) \to \mathrm{Ext}_5(U) \otimes U,$$

$$D_2 : \mathrm{Ext}_3(U) \to \mathrm{Ext}_6(U) \otimes \mathrm{Ext}_3(U) \cong \mathbb{C} \otimes \mathrm{Ext}_3(U) \cong \mathrm{Ext}_3(U),$$

$$D_3 : \mathrm{Ext}_3(U) \to \mathrm{Ext}_6(U) \otimes \mathrm{Ext}_6(U) \cong \mathbb{C} \otimes \mathbb{C} \cong \mathbb{C}.$$

The covariant D_1 has index 0, and the covariant D_2 has index 1. Last, D_3 is an invariant with index 2.

LEMMA 7.2. *The following table contains the relations which hold between the canonical forms of* $\mathrm{Ext}_3(U)$ *and the covariants of* $\mathrm{Ext}_3(U)$. (*For brevity we have suppressed the wedge product between the vectors*.)

	$pqr + stu$	$sqr + ptr + pqu$
D_3	$[p, q, r, s, t, u]^2$	0
D_2	$[p, q, r, s, t, u](pqr - stu)$	$2[p, q, r, s, t, u]pqr$
D_1	$(psqtr) \otimes u$	$-2 \cdot (psqtr) \otimes r$
	$+(psqtu) \otimes r$	$-2 \cdot (psruq) \otimes q$
	$+(psruq) \otimes t$	$-2 \cdot (qtrup) \otimes p$
	$+(psrut) \otimes q$	
	$+(qtrup) \otimes s$	
	$+(qtrus) \otimes p$	
Id	$pqr + stu$	$sqr + ptr + pqu$

	$pqr + ptu$	pqr	0
D_3	0	0	0
D_2	0	0	0
D_1	$2 \cdot (qtrup) \otimes p$	0	0
Id	$pqr + ptu$	pqr	0

The table is computed in the same manner as was the corresponding table for two by two by two matrices in Lemma 5.3. First compute the covariants on the element $e_1 \wedge e_2 \wedge e_3 + e_4 \wedge e_5 \wedge e_6$, where $e_1, \ldots, e_6$ is a basis for U. Now by changing the basis and the continuity argument, the other entries are easily computed.

Similarly to Corollary 5.6 we have the next corollary.

COROLLARY 7.3. *The following two identities hold*:

$$D_3 \circ D_2 = D_3^3,$$
$$D_2 \circ D_2 = -D_3^2 \cdot \mathrm{Id}.$$

Proof. Since $p \wedge q \wedge r + s \wedge t \wedge u$ is a generic canonical form, see Proposition 7.1, it is enough to consider elements on this form. For brevity we suppress the wedge product between the vectors. Let $p' = [p,q,r,s,t,u] \cdot p$ and $s' = -[p,q,r,s,t,u] \cdot s$. Then we have

$$\begin{aligned}
D_3(D_2(pqr + stu)) &= D_3([p,q,r,s,t,u] \cdot (pqr - stu)) \\
&= D_3(p'qr + s'tu) \\
&= [p',q,r,s',t,u] \\
&= [[p,q,r,s,t,u] \cdot p, q, r, -[p,q,r,s,t,u] \cdot s, t, u]^2 \\
&= [p,q,r,s,t,u]^6 \\
&= D_3(pqr + stu)^3. \\
D_2(D_2(pqr + stu)) &= D_2([p,q,r,s,t,u] \cdot (pqr - stu)) \\
&= D_2(p'qr + s'tu) \\
&= [p',q,r,s',t,u] \cdot (p'qr - s'tu) \\
&= [[p,q,r,s,t,u] \cdot p, q, r, -[p,q,r,s,t,u] \cdot s, t, u] \cdot \\
&\qquad \cdot([p,q,r,s,t,u] \cdot pqr + [p,q,r,s,t,u] \cdot stu) \\
&= -[p,q,r,s,t,u]^4 \cdot (pqr + stu) \\
&= -D_3(pqr + stu)^2 \cdot (pqr + stu). \quad \blacksquare
\end{aligned}$$

Define the linear map $\xi\colon W \to \mathrm{Ext}_3(U)$ by

$$\xi(p \otimes q \otimes r) = p \wedge q \wedge r,$$

where $p \in W_1$, $q \in W_2$, and $r \in W_3$. Observe that this map is injective. Hence we can define the inverse map Ξ from the image of ξ to $W_1 \otimes W_2 \otimes W_3$ by

$$\Xi(p \wedge q \wedge r) = p \otimes q \otimes r,$$

where $p \in W_1$, $q \in W_2$, and $r \in W_3$. Note that Ξ is linear. Also, we have that $\Xi(\xi(w)) = w$ for all $w \in W \cong W_1 \otimes W_2 \otimes W_3$.

We may define a bracket on the linear space U so that for $p_1, p_2 \in W_1$, $q_1, q_2 \in W_2$, and $r_1, r_2 \in W_3$ we have

$$[p_1, p_2, q_1, q_2, r_1, r_2] = [p_1, p_2] \cdot [q_1, q_2] \cdot [r_1, r_2],$$

where the brackets of step 2 are the brackets on the three linear spaces W_1, W_2, and W_3. We can now show how covariants on W and $\mathrm{Ext}_3(U)$ are related.

LEMMA 7.4. *The invariant Γ on W is related to the invariant D_3 by the identity*

$$D_3 \circ \xi = \Gamma.$$

Similarly the covariant S on W is related to D_2 by

$$\Xi \circ D_2 \circ \xi = S.$$

Proof. By Corollary 5.4 it is enough to prove these two identities on elements of W in the form $p \otimes q \otimes r + s \otimes t \otimes u$, where $p, s \in W_1$, $q, t \in W_2$, and $r, u \in W_3$. Thus we have

$$\begin{aligned} D_3(\xi(p \otimes q \otimes r + s \otimes t \otimes u)) &= D_3(p \wedge q \wedge r + s \wedge t \wedge u) \\ &= [p, q, r, s, t, u]^2 \\ &= [p, q]^2 \cdot [r, s]^2 \cdot [t, u]^2 \\ &= \Gamma(p \otimes q \otimes r + s \otimes t \otimes u). \end{aligned}$$

$$\begin{aligned} &\Xi(D_2(\xi(p \otimes q \otimes r + s \otimes t \otimes u)) \\ &\quad = \Xi(D_2(p \wedge q \wedge r + s \wedge t \wedge u)) \\ &\quad = \Xi([p, q, r, s, t, u] \cdot (p \wedge q \wedge r - s \wedge t \wedge u)) \\ &\quad = [p, q, r, s, t, u] \cdot (p \otimes q \otimes r - s \otimes t \otimes u) \\ &\quad = [p, q] \cdot [r, s] \cdot [t, u] \cdot (p \otimes q \otimes r - s \otimes t \otimes u) \\ &\quad = S(p \otimes q \otimes r + s \otimes t \otimes u). \quad \blacksquare \end{aligned}$$

Last, we present an explicit polynomial expression for the invariant D_3 on the space $\mathrm{Ext}_3(U)$. This formula may be viewed as a generalization of the expression given for the invariant Γ in Proposition 3.4. Let $e_1, \ldots, e_6$ be a basis for U such that $[e_1, \ldots, e_6] = 1$. Define for $1 \le i < j < k \le 6$

the element $e_{\{i,j,k\}} = e_i \wedge e_j \wedge e_k$. Observe that these elements form a basis for $\mathrm{Ext}_3(U)$.

PROPOSITION 7.5. *The invariant D_3 on the element $v = \sum_{1 \le i < j < k \le 6} a_{\{i,j,k\}} e_{\{i,j,k\}}$ is given by*

$$\begin{aligned} D_3(v) = &\sum_{\mathbf{i} \in L_1} a^2_{\{i_1,i_2,i_3\}} \cdot a^2_{\{i_4,i_5,i_6\}} \\ &- 2 \cdot \sum_{\mathbf{i} \in L_2} a_{\{i_1,i_3,i_4\}} \cdot a_{\{i_1,i_5,i_6\}} \cdot a_{\{i_2,i_3,i_4\}} \cdot a_{\{i_2,i_5,i_6\}} \\ &+ 4 \cdot \sum_{\mathbf{i} \in L_3} \big(a_{\{i_1,i_3,i_5\}} \cdot a_{\{i_1,i_4,i_6\}} \cdot a_{\{i_2,i_3,i_6\}} \cdot a_{\{i_2,i_4,i_5\}} \\ &\qquad + a_{\{i_2,i_4,i_6\}} \cdot a_{\{i_2,i_3,i_5\}} \cdot a_{\{i_1,i_4,i_5\}} \cdot a_{\{i_1,i_3,i_6\}} \big), \end{aligned}$$

where

$$L_1 = \{\mathbf{i} : i_1 < i_2 < i_3, i_4 < i_5 < i_6, i_1 < i_4, \text{ and } i_1, \ldots, i_6 \text{ are distinct}\},$$

$$L_2 = \{\mathbf{i} : i_1 < i_2, i_3 < i_4, i_5 < i_6, i_3 < i_5, \text{ and } i_1, \ldots, i_6 \text{ are distinct}\},$$

$$L_3 = \{\mathbf{i} : i_1 < i_3 < i_5, i_1 < i_2, i_3 < i_4, i_5 < i_6, \text{ and } i_1, \ldots, i_6 \text{ are distinct}\}.$$

8. CONCLUDING REMARKS

The definition of covariants and invariants can be easily extended to the linear space $V_1 \otimes \cdots \otimes V_d$, where the dimension of V_i is n_i. In this setting, is it possible to find a complete set of canonical forms for $V_1 \otimes \cdots \otimes V_d$, that is, the space of d-dimensional $n_1 \times \cdots \times n_d$ matrices? Are we able to find the corresponding set of covariants which distinguishes these canonical forms one from another? For the case when $d = 3$, this question is mentioned in [6, pp. 98–100].

In Proposition 3.4 and Lemma 5.2 a higher dimensional Gaussian elimination was used. Would a similar technique work for a d-dimensional $n_1 \times \cdots \times n_d$ matrix? Which multidimensional matrix would play the role of the identity matrix? The following result may give a hint about the answer.

PROPOSITION 8.1 (Ehrenborg [3]). *If we can place k rooks in a d-dimensional $n_1 \times \cdots \times n_d$ chess board such that every position is attacked by at least one rook* (*or there is a rook in that position*), *then a generic element w in $V_1 \otimes \cdots \otimes V_d$ may be written in the form*

$$\sum_{i=1}^{k} v_{i,1} \otimes \cdots \otimes v_{i,d},$$

where $v_{i,j} \in V_j$ and $\dim(V_j) = n_j$.

The problem of such placements of rooks on a chess board also goes by the name *"The Football Pool Problem,"* see [10, 7]. We call a placement *perfect* when each position is attacked by exactly one rook. (We may view this as an error correcting code, which corrects one error.) It seems that these perfect placements could play the role of multidimensional identity matrices. The next nontrivial case occurs when $d = 4$ and $(n_1, n_2, n_3, n_4) = (3, 3, 3, 3)$.

The umbral map generalizes and will be useful for expressing invariants and covariants of multidimensional matrices. But is it possible to express every covariant of multidimensional matrices in terms of the umbral map?

Does the set of six covariants presented in Section 4 generate all other covariants of two by two by two matrices? That is, given a covariant C, can it be written as a polynomial in these six covariants? The corresponding result for the binary cubic is proven in [8]. Moreover, is there a deeper reason why the two posets presented at the end of Section 5 are isomorphic?

The idea that there are relations of covariants of different spaces, see Sections 6 and 7, also applies to the two most classical invariants, namely the determinant and the Pfaffian. Let W be a vector space over the complex numbers of dimension $2n$. Consider the space $\mathrm{Ext}_2(W)$ consisting of skew-symmetric tensors of step 2. The Pfaffian Pf is an invariant, with the property that $\mathrm{Pf}(w) = 0$ if an only if w may be written in the form $\sum_{i=1}^{n-1} u_i \wedge v_i$, see [2]. Let U and V be n-dimensional spaces such that $U \oplus V = W$. We may view the linear space $U \otimes V$ as the linear space of n by n matrices. Define the map ρ: $U \otimes V \to \mathrm{Ext}_2(W)$ by $\rho(u \otimes v) = u \wedge v$ and extend ρ by linearity. Now it is easy to see that the composition $\mathrm{Pf} \circ \rho$ is the classical determinant on the space $U \otimes V$. Hence, given the determinant, we could lift it up and define the Pfaffian on a larger linear space. Hopefully, this idea of lifting covariants from one space to another may be used in further research about canonical forms and covariants.

REFERENCES

1. R. Ehrenborg and G.-C. Rota, Apolarity and canonical forms for homogeneous polynomials, *European J. Combin.* **14** (1993), 157–181.
2. R. Ehrenborg, G. Hetyei, and G.-C. Rota, "Lectures on Polynomials," to appear, 1995.
3. R. Ehrenborg, On apolarity and generic canonical forms, *J. Algebra* **213** (1999), 167–194.
4. J. H. Grace and A. Young, "The Algebra of Invariants," Chelsea, New York, 1903.
5. F. D. Grosshans, G.-C. Rota, and J. A. Stein, "Invariant Theory and Superalgebras," Regional Conference Series in Mathematics, Vol. 69, American Mathematical Society, Providence, RI, 1987.
6. J. Harris, "Algebraic Geometry, a First Course," Springer-Verlag, New York, 1993.

7. H. J. L. Kamps and J. H. van Lint, The football pool problem for 5 matches, *J. Combin. Theory Ser. A* **3** (1967), 315–325.
8. J. P. S. Kung and G.-C. Rota, The invariant theory of binary forms, *Bull. Amer. Math. Soc.* **10** (1984), 27–85.
9. E. Lasker, Zur Theorie der kanonische Formen, *Math. Ann.* **58** (1904), 434–440.
10. G. Losey, Note on a theorem of Zaremba, *J. Combin. Theory Ser. A* **6** (1969), 208–209.
11. E. K. Wakeford, On canonical forms, *Proc. London Math. Soc.* **2** (1919), 403–410.