

VECTOR INVARIANTS OF SYMMETRIC GROUPS

BY

H. E. A. CAMPBELL, I. HUGHES, AND R. D. POLLACK

ABSTRACT. Let M be a free module of rank n over a commutative ring R with unit and let Σ_n denote the symmetric group acting on a fixed basis of M in the usual way. Let M^m denote the direct sum of m copies of M and let S be the symmetric ring of M^m over R . Then each element of Σ_n acts diagonally on M^m and consequently on S ; denote by Σ_n the subgroup of $Gl(M^m)$ so defined. The ring of invariants, S^{Σ_n} , is called the ring of vector invariants by H. Weyl [3, Chapter II, p. 27] when $R = \mathbf{Q}$. In this paper a set of generators valid over any ring R is given. This set of generators is somewhat larger than Weyl's. It is interesting to note that, over the integers, his algebra and S^{Σ_n} have the same Hilbert-Poincaré series, are equal after tensoring with the rationals, and have the same fraction fields, although they are not equal.

0. Introduction. Let M be a free module of rank n over a commutative ring R with unit and suppose z_1, \dots, z_n is a basis for M . Let Σ_n denote the symmetric group of all permutations of the z_i . Let $M^m = M \oplus \dots \oplus M$ (m copies); then M^m is a free module of rank mn . Denote by z_{1i}, \dots, z_{ni} a basis for the i th copy of M in M^m . The symmetric R -algebra S of M^m over R is isomorphic to the polynomial ring $R[z_{11}, \dots, z_{nm}]$. Each element g in Σ_n acts diagonally on M^m , that is, $g(z_{ij}) = z_{g(i)j}$. The reader might find it useful to think of the matrix of indeterminates

$$Z = \begin{pmatrix} z_{11} & z_{12} & \dots & z_{1m} \\ z_{21} & z_{22} & \dots & z_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ z_{n1} & z_{n2} & \dots & z_{nm} \end{pmatrix}.$$

Elements of Σ_n act by permuting the rows of Z . Σ_n acts as a group of degree-preserving R -algebra automorphisms of S . The ring of invariants, S^{Σ_n} , is called the ring of vector invariants of Σ_n by H. Weyl [3, Chapter II, p. 27] who computed a set of generators called 'generalized elementary symmetric functions' when $R = \mathbf{Q}, \mathbf{R}$, or \mathbf{C} —see [3, pp. 36–39]. One can also find a brief discussion of this ring when $R = \mathbf{C}$ in Stanley's expository article [2, example 5.3, pp. 492–493].

In this paper a set of generators valid over any ring R is described (see Theorem 4.1). This set of generators is somewhat larger than Weyl's. It is interesting to note that over the integers his algebra and S^{Σ_n} have the same Hilbert-Poincaré series, are equal after

Received June 26, 1990.

AMS subject classification: 13F20.

©Canadian Mathematical Society 1990.

The first author gratefully acknowledges the support of NSERC.

tensoring with the rationals, and have the same fraction fields, although they are not equal.

The paper is organized as follows. Section 1 sets notations and conventions and discusses a homogeneous basis for the ring of invariants. In section 2, following a suggestion of the referee, a partial order on the sequences comprising the columns of an ‘exponent’ matrix is defined—this provides the key step in the proof of the main theorem. Section 3 discusses a set of exponent matrices which determines the generators for the ring of invariants. Section 4 is devoted to the main theorem (4.1). Finally, in Section 5 Weyl’s algebra is shown not to equal the ring of invariants over \mathbf{Z} , although they are equal over \mathbf{Q} .

We would like to thank D. Richman for detecting errors in an early version of this paper. In addition, Richman has a version of the main theorem, although his set of generators is somewhat larger than ours—see the remark following the proof of 4.1. We would also like to express our gratitude to the referee, who offered many suggestions and improvements, found errors, and even pointed out that we had proved more than we originally claimed.

1. **Preliminaries.** Let I be any $n \times m$ matrix of non-negative integers. Then z^I denotes the monomial $z_1^{i_1} \cdots z_m^{i_m}$ and I is called an exponent matrix. Denote the degree of z^I by $|I|$. Let $O(z^I)$ be the orbit of z^I under Σ_n , that is,

$$O(z^I) = \{ z^J \mid \exists g \in \Sigma_n \text{ with } g(z^I) = z^J \}.$$

Write $g(J) = I$ if $g(z^J) = z^I$ and $O(I) = \{ J \mid \exists g \in \Sigma_n \text{ with } g(J) = I \}$. Note that $J \in O(I)$ if and only if J is obtained from I by some permutation of the rows of I . In particular, if $J \in O(I)$, $|J| = |I|$.

Let $s(I) = \sum_{J \in O(I)} z^J$ and observe that $s(I) \in S^{\Sigma_n}$, for all I .

LEMMA 1.1. $\{s(I) \mid |I| = t\}$ is a basis for $S_t^{\Sigma_n}$, the homogeneous polynomials of degree t invariant under the action of Σ_n .

PROOF. Omitted. ■

2. **A partial order on columns of exponent matrices.** Elements of Σ_n acts on any sequence $\mathbf{a} = (\mathbf{a}_1, \dots, \mathbf{a}_n)$ of n non-negative integers by permuting the entries of \mathbf{a} ; denote the orbit of \mathbf{a} under this action by $O(\mathbf{a})$ and define the stabilizer of \mathbf{a} , $\text{Stab}(\mathbf{a})$, to be $\{g \in \Sigma_n \mid g(\mathbf{a}) = \mathbf{a}\}$, a subgroup of Σ_n . Note that each orbit $O(\mathbf{a})$ contains a unique sequence $\mathbf{b} = (\mathbf{b}_1, \dots, \mathbf{b}_n)$ in descending order $b_1 \geq b_2 \geq \dots \geq b_n$. Given two sequences in descending order $\mathbf{b} = (\mathbf{b}_1, \dots, \mathbf{b}_n)$ and $\mathbf{c} = (\mathbf{c}_1, \dots, \mathbf{c}_n)$ define $\mathbf{b} \preceq \mathbf{c}$ if $\sum_{i=1}^r b_i \leq \sum_{i=1}^r c_i$ for $1 \leq r \leq n$. This is a partial order on the set of descending sequences. Write $O(\mathbf{a}) \preceq O(\mathbf{b})$ if \mathbf{a} and \mathbf{b} are in descending order with $\mathbf{a} \preceq \mathbf{b}$. This is a partial order on the set of orbits.

Sequences are added component by component.

LEMMA 2.1. Suppose \mathbf{a} and \mathbf{b} are two sequences in descending order (so $\mathbf{a} + \mathbf{b}$ is also). If $\mathbf{c} \in O(\mathbf{a})$ and $\mathbf{d} \in O(\mathbf{b})$ have the property that $\mathbf{c} + \mathbf{d}$ is in descending order, then $\mathbf{c} + \mathbf{d} \preceq \mathbf{a} + \mathbf{b}$ with equality if and only if $\mathbf{c} = \mathbf{a}$ and $\mathbf{d} = \mathbf{b}$.

PROOF. Omitted. ■

Denote by I^i the i th column of the $n \times m$ exponent matrix I . Say exponent matrices J and K *cohere* if there exist $g_1, \dots, g_m \in \Sigma_n$ such that $g_i(J^i)$ and $g_i(K^i)$ are in descending order for $i = 1, \dots, m$.

LEMMA 2.2. *Suppose the exponent matrices J and K cohere. Then*

- (1) *If $L \in O(J)$ and $M \in O(K)$ and if $L + M = J + K$, then $L = J$ and $M = K$;*
- (2) *If $g \in \Sigma_n$, then for $1 \leq i \leq m$, $O(J^i + g(K^i)) \preceq O(J^i + K^i)$, with equality if and only if $g \in \text{Stab}(J^i)\text{Stab}(K^i)$.*

PROOF. Choose $g_1, \dots, g_m \in \Sigma_n$ so that $g_i(J^i)$ and $g_i(K^i)$ are in descending order.

(1) Since $g_i(L^i) + g_i(M^i) = g_i(L^i + M^i) = g_i(J^i + K^i) = g_i(J^i) + g_i(K^i)$ is in descending order, $g_i(L^i) + g_i(M^i)$ is in descending order. Applying Lemma 2.1, $g_i(L^i) = g_i(J^i)$ and $g_i(M^i) = g_i(K^i)$. So $L^i = J^i$ and $M^i = K^i$, $1 \leq i \leq m$.

(2) Choose $h \in \Sigma_n$ so that $h(J^i + g(K^i))$ is in descending order. But $h(J^i + g(K^i)) = h(J^i) + hg(K^i) \preceq g_i(J^i) + g_i(K^i)$ by Lemma 2.1. Hence the first statement.

If $h(J^i) + hg(K^i) = g_i(J^i) + g_i(K^i)$ then, by Lemma 2.1, $h(J^i) = g_i(J^i)$ and $hg(K^i) = g_i(K^i)$. Thus $g_i^{-1}hg \in \text{Stab}(K^i)$ and so $g \in \text{Stab}(J^i)\text{Stab}(K^i)$. ■

The following lemma provides the key step in the proof of the main theorem (4.1).

LEMMA 2.3. *Suppose J and K cohere and $J + K = I$. Then*

$$s(J)s(K) = s(I) + \sum_r s(M_r),$$

where $|M_r| = |I|$, $s(M_r) \neq s(I)$ and $O(M_r^i) \preceq O(I^i)$, $1 \leq i \leq m$, for each r . If K is concentrated in its j th column (i. e. $K^i = \mathbf{0}$, for $i \neq j$) and $\text{Stab}(J^j) \subseteq \text{Stab}(K^j)$ then $O(M_r^i) \prec O(I^i)$, for each r .

PROOF. By Lemma 1.1 $s(J)s(K) = ts(I) + \sum_r s(M_r)$ for $t \in R$ and $s(M_r) \neq s(I)$ for all r . Monomials which occur in the expansion of $s(J)s(K)$ have exponent matrices $g(J) + h(K)$ for some g and h in Σ_n . By Lemma 2.2 (1) and since $J + K = I$, $t = 1$.

Now $M_r = g(J) + h(K)$ for some $g, h \in \Sigma_n$ (in particular $|M_r| = |I|$) and $s(M_r) \neq s(I)$ so that $g(J) \neq h(J)$ and $g(K) \neq h(K)$. Consider $g^{-1}M_r = J + g^{-1}h(K) \in O(M_r)$ with i th column $J^i + g^{-1}h(K^i)$. By Lemma 2.2 (2) $O(M_r^i) = O(J^i + g^{-1}h(K^i)) \preceq O(J^i + K^i) = O(I^i)$, $1 \leq i \leq m$, so the first part of the lemma is proved.

Let K be concentrated in its j th column, so that $J^i + g^{-1}h(K^i) = J^i$ is the i th column of $g^{-1}(M_r)$ and I alike, for $i \neq j$. If $O(J^j + g^{-1}h(K^j)) = O(J^j + K^j) = O(I^j)$ then, by Lemma 2.2 (2) $g^{-1}h = g'h'$ for $g' \in \text{Stab}(J^j)$, $h' \in \text{Stab}(K^j)$. So $J^j + g^{-1}h(K^j) = J^j + g'h'(K^j) = J^j + g'(K^j) = J^j + K^j = I^j$, since $\text{Stab}(J^j) \subset \text{Stab}(K^j)$. Thus $g^{-1}(M_r) = I$, a contradiction. ■

3. A set of exponent matrices. View each sequence

$$\mathbf{a} = (\mathbf{a}_1, \dots, \mathbf{a}_n)$$

of non-negative integers as a function (without changing notation) $\mathbf{a} : \{\mathbf{1}, \dots, \mathbf{n}\} \rightarrow \mathbf{N}$, with $\mathbf{a}(\mathbf{k}) = \mathbf{a}_{\mathbf{k}}$. Here \mathbf{N} denotes the set of non-negative integers. Define $\text{Ker}(\mathbf{a}) = \{\mathbf{k} \mid \mathbf{a}(\mathbf{k}) = \mathbf{0}\}$.

Let Ω be the set of exponent matrices satisfying $I = \mathbf{0}$, or both of

- (1) the image of each I^i is an interval in \mathbf{N} ; and
- (2) $\{ \text{Ker}(I^i) \mid 1 \leq i \leq m \}$ has no minimum element.

The set $\{ \text{Ker}(I^i) \mid 1 \leq i \leq m \}$ is partially ordered by inclusion; a minimum element of this set is a set $\text{Ker}(I^i)$ such that $\text{Ker}(I^i) \subseteq \text{Ker}(I^j)$, for all $j, 1 \leq j \leq m$. For example, if I^i has all entries positive, then $\text{Ker}(I^i) = \emptyset$ and so $\text{Ker}(I^i)$ is a minimum element of $\{ \text{Ker}(I^i) \mid 1 \leq i \leq m \}$.

Note that $J \in O(I)$ and $I \in \Omega$ implies $J \in \Omega$.

In fact, it isn't difficult to see that $0 \neq I \in \Omega$ if and only if there is an element J in $O(I)$ of the form

$$J = \begin{pmatrix} E \\ \mathbf{0} \end{pmatrix},$$

where E is $l \times m$ (and has no zero rows), $1 \leq l \leq n$ and where each column of E viewed as a function has image an interval in \mathbf{N} of the form $[0, \dots, k]$ for $k \leq l - 1$.

4. The theorem. Let K be the exponent matrix which is concentrated in its j th column where $K^j = (1, \dots, 1, 0, \dots, 0)$ (i ones) and denote by σ_{ij} the orbit polynomial $s(K)$. Note that σ_{ij} is the i th elementary symmetric function in the variables z_{1j}, \dots, z_{nj} , so that, for example, $\sigma_{1j} = z_{1j} + z_{2j} + \dots + z_{nj}$, and $\sigma_{nj} = z_{1j}z_{2j} \dots z_{nj}$. Set $B = R[\sigma_{ij} \mid 1 \leq i \leq n, 1 \leq j \leq m] \subset S^{\Sigma_n} \subset S = R[z_{ij}]$; then B is the R -subalgebra of polynomials left invariant under the group $(\Sigma_n)^m$.

The reader might find it helpful to recall the matrix of indeterminates

$$Z = \begin{pmatrix} z_{11} & z_{12} & \dots & z_{1m} \\ z_{21} & z_{22} & \dots & z_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ z_{n1} & z_{n2} & \dots & z_{nm} \end{pmatrix}.$$

Elements of Σ_n act as permutations of the rows of Z , while elements of $(\Sigma_n)^m$ act by permuting the entries of each column of Z independently.

In any event, B is a polynomial algebra over which S and hence S^{Σ_n} is integral. Consequently, S^{Σ_n} is finitely generated as a B -module; in this situation the σ_{ij} are said to be a homogeneous system of parameters for S^{Σ_n} .

Let A be the B -module generated by $\{s(I) \mid I \in \Omega\}$ so that $B \subset A \subseteq S^{\Sigma_n}$.

THEOREM 4.1. $A = S^{\Sigma_n}$.

PROOF. Since $\{s(I)\}$ is a basis for S^{Σ_n} (by Lemma 1.1) it suffices to show $s(I) \in A$, for all $I \notin \Omega$.

For each I , let $\overline{O}(I)$ be the $(\Sigma_n)^m$ -orbit of I . Partially order these orbits by $\overline{O}(I) \leq \overline{O}(J)$ if $O(I^i) \leq O(J^i)$ for all i . The proof proceeds by induction on the ordering of the $(\Sigma_n)^m$ -orbits. The general step of the induction is the same as the first so suppose that $s(J) \in A$ for all J with $\overline{O}(J) < \overline{O}(I)$. Define $l(I)$ to be the number of non-zero rows in I and suppose $l(I) = l$. Observe that $J \in O(I)$ implies $l(J) = l(I)$.

CASE (1): For some j , the image of \bar{I}^j is not an interval in \mathbf{N} .

Replacing I by an element of $O(I)$, if necessary, assume that $\bar{I}^j = (c_1, \dots, c_l, 0, \dots, 0)$ with $c_1 \geq c_2 \geq \dots \geq c_l$ and this column is not the image of an interval. Let q be the least integer for which $c_q - c_{q+1} \geq 2$, $1 \leq q \leq l$.

Let K be the matrix concentrated in its j th column with $K^j = (1, \dots, 1, 0, \dots, 0)$ (q ones) so that $s(K) = \sigma_{qj}$ is one of the generators for B . Set $J = I - K$. The columns of J are those of I except for $J^j = (a_1, \dots, a_l, 0, \dots, 0)$ where $a_i = c_i - 1$, $1 \leq i \leq q$, and $a_i = c_i$, $q < i \leq l$. Clearly $\bar{O}(J) < \bar{O}(I)$, so $s(J) \in A$.

It is easy to verify that J and K cohere. Further, $\text{Stab}(K^j) \supset \text{Stab}(J^j)$ since J^j is in descending order and $a_q > a_{q+1}$. And so by Lemma 2.3

$$s(J)s(K) = s(I) + \sum_r s(M_r),$$

where $O(M_r^j) < O(\bar{I}^j)$. But then $\bar{O}(M_r) < \bar{O}(I)$ and so by induction $s(M_r) \in A$ for each r . Thus $s(I) \in A$ in case (1).

CASE (2): For some j , $\text{Ker}(\bar{I}^j)$ is a minimum.

By case (1) each column of I may be assumed to be the image of an interval in \mathbf{N} . Replacing I by some element of $O(I)$, if necessary, and using the minimality of $\text{Ker}(\bar{I}^j)$ assume that $\bar{I}^j = (c_1, \dots, c_l, 0, \dots, 0)$, with $l \geq c_1 \geq c_2 \geq \dots \geq c_l = 1$ and $c_i - 1 \leq c_{i+1}$ for $1 \leq i \leq l - 1$.

Let K be the matrix concentrated in its j th column with $K^j = (1, \dots, 1, 0, \dots, 0)$ (l ones) so that $s(K) = \sigma_{lj}$ is one of the generators of B . Set $J = I - K$. The columns of J are those of I except that $J^j = (a_1, \dots, a_l, 0, \dots, 0)$ where $a_i = c_i - 1$. Note that $s(J) \in A$ since $\bar{O}(J) < \bar{O}(I)$.

It is trivial to verify that J and K cohere. By Lemma 2.3

$$s(J)s(K) = s(I) + \sum_r s(M_r),$$

where $\bar{O}(M_r) \leq \bar{O}(I)$, and $M_r \notin O(I)$. Those terms $s(M_r)$ with $\bar{O}(M_r) < \bar{O}(I)$ are already in A by induction.

So suppose $\bar{O}(M_r) = \bar{O}(I)$ and recall $M_r = g(J) + h(K)$ for some $g, h \in \Sigma_n$. Consider the matrix $M \in O(M_r)$ defined by

$$M = h^{-1}(M_r) = h^{-1}g(J) + K.$$

Notice $l(M) \geq l(I)$.

Suppose $l(M) = l(I)$. Then the first l rows of M are non-zero, since the first l rows of K are non-zero. It follows that

$$h^{-1}g(J) = \begin{pmatrix} E \\ \mathbf{0} \end{pmatrix},$$

where E is $l \times m$ (E may have zero rows). Observe that $h^{-1}g$ permutes the nonzero rows of J so that they all lie among the first l rows of the result—let h' be any element of Σ_l that acts on the non-zero rows of J in the same way. Then $h^{-1}g(J) = h'(J)$. Now $h' \in \Sigma_l \subset \text{Stab}(K^j)$ so $M = h'(J) + K = h'(J + K) = h'(I)$, contradicting $M_r \notin O(I)$.

Therefore $l(M) > l(I)$. If $l(I) = n$ no such terms $s(M_r)$ exist, and $s(I) \in A$. Suppose $l(I) < n$. Consider the matrix $N \in O(M_r)$ defined by

$$N = g^{-1}h(M_r) = J + g^{-1}h(K),$$

with $n \geq l(N) > l(I)$. No column N^i can have $\text{Ker}(N^i)$ a minimum for $i \neq j$ because $g^{-1}h \notin \text{Stab}(K^j)$ implies N^j has a non-zero entry below row l . Finally $O(N^j) = O(I^j)$ implies that the entries in N^j are some permutation of those of I^j . In particular, N^j has $n - l$ zeros. If all the zero entries in N^j occur in zero rows of N it follows that $l(N) = l(I)$ a contradiction. Hence some zero entry of N^j occurs in a nonzero row of N so that $\text{Ker}(N^j)$ is not a minimum. In this case N and hence $M_r \in \Omega$ so $s(M_r) \in A$. Consequently, $s(I) \in A$ as required. ■

5. **Remark.** D. Richman (unpublished) can show that S^{Σ_n} is generated as a B -module by $\{s(I) \mid \text{no entry of } I \text{ is larger than } n - 1\}$. Indeed $\prod_{i=1}^n (t - z_{ij}) = \sum_{i=0}^n (-1)^i \sigma_{ij} t^{n-i}$ (with $\sigma_{0j} = 1$) yielding Newton's formula (set $t = z_{ij}$) $z_{ij}^n = \sum_{i=0}^n (-1)^{i+1} \sigma_{ij} z_{ij}^{n-i}$; inductively, one obtains formulae for z_{ij}^r when $r > n$. Intuitively, these formulae allow one to replace large entries in exponent matrices by smaller entries, using elements of B .

6. **An example.** H. Weyl's "generalized elementary symmetric functions" do not in general generate S^{Σ_n} over \mathbf{Z} . Let Σ_2 act on $M = \mathbf{Z} \oplus \mathbf{Z}$ with basis $\{x, y\}$ by permutations of x and y . Denote a basis for M^3 by $\{x_1, y_1, x_2, y_2, x_3, y_3\}$ with $1 \neq g \in \Sigma_2$ acting by $g(x_i) = y_i, g(y_i) = x_i, i = 1, 2, 3$. Then $S = S(M^3) \cong \mathbf{Z}[x_i, y_i]$. Let B be the subalgebra generated by the elementary symmetric functions. That is, if e_{ij} is the 2×3 matrix which is zero everywhere except for a one in the (i, j) -position, B is generated by $s(e_{1i}) = x_i + y_i$ and $s(e_{1i} + e_{2i}) = x_i y_i, i = 1, 2, 3$.

It follows from the discussion in R. P. Stanley's paper (in particular, see [2, formula 14, page 492 and the paragraph preceeding proposition 5.4]) that $(S \otimes_{\mathbf{Z}} \mathbf{Q})^{S_2}$ is free as a module over $B \otimes_{\mathbf{Z}} \mathbf{Q}$ with basis $\{1, s(e_{1i} + e_{2j}) = x_i y_j + x_j y_i, 1 \leq i < j \leq 3\}$. Weyl's algebra $C \subset S^{\Sigma_2}$ is generated by B and $s(e_{1i} + e_{2j}) = x_i y_j + x_j y_i, 1 \leq i < j \leq 3$ (see [3, pp. 36–39]).

Now consider $f = s(I)$ for $I = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix}$. Over $\mathbf{Q}, f = s(e_{11})s(e_{12})s(e_{13}) - \frac{1}{2}(s(e_{11})s(e_{12} + e_{23}) + s(e_{12})s(e_{11} + e_{23})s(e_{13})s(e_{11} + e_{22}))$, and this expression is unique in the free basis over $B \otimes_{\mathbf{Z}} \mathbf{Q}$. So $2f \in C$ while $f \notin C$. It also follows that A is not free as a B -module.

It is interesting and not difficult to see that in general C and $A = S^{\Sigma_n}$ have the same Poincaré-Hilbert series, $C \otimes_{\mathbf{Z}} \mathbf{Q} = S^{\mathbf{S}^n} \otimes_{\mathbf{Z}} \mathbf{Q}$, and have the same fraction fields, but $C \neq S^{\Sigma_n}$ for $m > 2$ and $n > 1$.

REFERENCES

1. H. E. A. Campbell, I. Hughes, R. D. Pollack, *Rings of invariants and p -Sylow subgroups*, submitted to *Canad. Math. Bull.*
2. R. P. Stanley, *Invariants of finite groups and their applications to combinatorics*, *Bull. A. M. S.* **1** (3) (1979), 475–511.
3. H. Weyl, *Classical Groups*, Princeton University Press, Princeton.

Mathematics and Statistics Department
Queen's University
Kingston, Ontario K7L 3N6