

# Invariants of $S_4$ and the Shape of Sets of Vectors

Helmer Aslaksen<sup>1</sup>, Shih-Ping Chan<sup>2</sup>, Tor Gulliksen<sup>3</sup>

<sup>1</sup> Department of Mathematics, National University of Singapore, Singapore 0511, Republic of Singapore, e-mail: aslaksen@math.nus.sg

<sup>2</sup> Department of Mathematics, National University of Singapore, Singapore 0511, Republic of Singapore, e-mail: matchansp@math.nus.sg

<sup>3</sup> Department of Mathematics, University of Oslo, P.O. Box 1053, Blindern, 0316 Oslo, Norway, e-mail: torg@math.uio.no

Received October 7, 1993; revised version November 28, 1994

**Abstract.** We study a representation of  $S_n$  that is related to the shape of sets of vectors in  $\mathbb{R}^n$ . We want to determine the invariants of this representation, and obtain a complete description for the case of  $S_4$ .

**Keywords:** Invariants of finite groups, Shape of sets of vectors.

## 1 Introduction

When studying the shape of a set of  $n$  vectors in  $\mathbb{R}^n$ , we are led to consider the following representation of  $S_n$ . Let  $\mathcal{M}$  be the set of symmetric  $n \times n$  matrices over  $\mathbb{R}$  with zeros along the diagonal. We can identify a permutation in  $S_n$  with the corresponding  $n \times n$  permutation matrix, and we get a representation  $\phi: S_n \rightarrow GL(\mathcal{M})$  given by

$$\phi(P)(M) = P^t M P,$$

where  $P \in S_n$  and  $M \in \mathcal{M}$ . This action moves the  $(i, j)$  entry of  $M$  to the  $(P(i), P(j))$  position. If we let  $V$  denote the standard representation of  $S_n$  corresponding to the partition  $(n-1, 1)$  (i.e., we start with the natural  $n$ -dimensional representations, and cancel the 1-dimensional invariant subspace spanned by  $(1, \dots, 1)$ ), then we will see later that

$$\phi = S^2 V - V = \text{ind}_{S_{n-2} \times S_2}^{S_n} \mathbf{1} = (n) \oplus (n-1, 1) \oplus (n-2, 2).$$

We can also think of this as a mapping from  $S_n$  into  $S_{n(n-1)/2}$ .

We would like to determine the invariants of this representation. When  $n \geq 5$ , the structure of the ring of invariants is too complicated, but we can get a completed description for  $n=4$ . In Section 2 we describe the background for the problem, and

in Section 3 we describe the ring of invariants when  $n = 4$ . The first author would like to thank Hanspeter Kraft for helpful suggestions and for patiently answering many questions.

## 2 Geometric Background

The problem studied in this paper stems from our need for a computationally efficient way of discriminating finite sets of vectors in  $\mathbb{R}^m$  according to their “shape” (see Definition 1 below). With respect to certain applications in the field of neural networks, where vectors frequently have components  $\pm 1$ , it is no severe limitation to restrict ourselves to vectors with fixed norm, say normalized to 1.

The idea is now to find algebraic shape-invariants that can be used for the discrimination process. Suppose that we are given two sets of vectors with different shape, but with the same cardinality  $n$ . With a sufficiently rich sequence of invariants at hand, the natural discrimination procedure would be to go through the invariants one by one as long as their values on the two sets coincide. Thus we will only occasionally have to evaluate all the invariants in the sequence. The big problem with this approach, however, is to find an adequate sequence of invariants.

For  $n = 3$  the problem is simple. Indeed, for a given set of unit vectors  $\{v_1, v_2, v_3\}$  it is easily seen that the (unordered) set of three inner products  $\langle v_i, v_j \rangle$  for  $i < j$  determines the shape of the set. So the sequence of invariants can be taken to be the elementary symmetric functions  $s_1, s_2, s_3$  of these inner products.

For  $n = 4$  the problem is much more complicated. The shape of a set of unit vectors  $\{v_1, \dots, v_4\}$  is no longer determined by the elementary symmetric functions of the inner products  $a_{ij} = \langle v_i, v_j \rangle$  for  $i < j$ . However, from the theory in the Section 3, we are able to extract an explicit sequence of 9 invariants (expressible in the  $a_{ij}$ ), capable of discriminating shape.

We have not been able to exhibit an adequate sequence of invariants in the case  $n \geq 5$ .

We will now see how the definition of shape of a set of vectors in  $\mathbb{R}^n$  leads to a certain representation of the symmetric group  $S_n$ .

**Definition 1.** *Let  $K$  and  $L$  be two sets of  $n$  vectors in  $\mathbb{R}^n$ . We say that  $K$  and  $L$  have the same shape if there exists a linear orthogonal transformation  $F$  such that  $F(K) = L$ . Thinking of the elements of the unordered set  $L$  as the columns of an  $m \times n$  matrix, we are lead to the following definition.*

**Definition 2.** *Let  $A$  and  $B$  be real  $m \times n$ -matrices. We will call  $A$  and  $B$  congruent if there exists an orthogonal matrix  $U$  and a permutation matrix  $P$  such that*

$$UAP = B$$

So if the matrices  $A$  and  $B$  correspond to the sets  $K$  and  $L$ , then  $K$  and  $L$  have the same shape if and only if  $A$  and  $B$  are congruent. In order to study the congruence classes of  $m \times n$  matrices, we will prove the following simple lemma.

**Lemma 3.** *Let  $C$  and  $D$  be real matrices of the same size. Then we have  $C^t C = D^t D$  if and only if there exists an orthogonal matrix  $U$  such that  $D = UC$ .*

*Proof.* One of the implications is trivial. For the other, assume that  $C^t C = D^t D$ . Let  $c_1, \dots, c_n$  and  $d_1, \dots, d_n$  be the columns of  $C$  and  $D$ , respectively. Let  $r_1, \dots, r_n$  be real

numbers. By squaring the norm it follows from the assumption that

$$\sum r_i c_i = 0 \Leftrightarrow \sum r_i d_i = 0,$$

from which it follows that the equations  $F(c_i) = d_i$  for  $i = 1 \dots n$  determine a well-defined linear map  $F$  from the column space of  $C$  to that of  $D$ . It is clear that  $F$  is norm preserving, and that it is possible to extend  $F$  to an orthogonal transformation defined on  $\mathbb{R}^n$ .  $\square$

It follows from Lemma 3 that  $A$  and  $B$  are congruent if and only if there exists a permutation matrix  $P$  such that

$$P^t A^t A P = B^t B.$$

Thus the congruence classes for the congruence relation defined above correspond to the orbits of the symmetric group  $S_n$  acting on the set of symmetric positive semidefinite  $n \times n$  matrices, the action of  $P$  on a symmetric matrix  $M$  being  $P^t M P$ . Notice that  $A^t A$  represents the inner products of the vectors in our set.

We assume from now on that the elements of  $L$  have unit length. Thus the corresponding symmetric matrices have ones along the diagonal.

### 3 Invariants

The geometric considerations above give us an action of  $S_n$  on symmetric  $n \times n$  matrices with ones along the diagonal. This action moves the  $(i, j)$  entry to the  $(P(i), P(j))$  position. This fixes the diagonal, so we can restrict ourselves to looking at symmetric  $n \times n$  matrices with zeros along the diagonal in order to get a vector space. This gives a representation  $\phi = S^2 V - V$ , where  $V$  is the standard  $(n - 1)$ -dimensional representation of  $S_n$ , and it follows from [5, 4.19] that

$$\phi = (n) \oplus (n - 1, 1) \oplus (n - 2, 2).$$

But from [5, 4.39] we know that this is equal to  $\text{ind}_{S_{n-2} \times S_2}^{S_n} 1$ . We will write

$$\tilde{V} = \phi_1 = (n) \oplus (n - 1, 1) = k \oplus V \quad \text{and} \quad W = \phi_2 = (n - 2, 2).$$

Hence  $\phi = \tilde{V} \oplus W$ , where  $\tilde{V}$  is the natural  $n$ -dimensional permutation representation.

We will now assume that  $n = 4$  and set  $G = S_4$ . The elementary symmetric functions  $s_1, s_2, s_3, s_4$  in  $x_1, x_2, x_3, x_4$  form a system of parameters for  $\mathbb{R}[\tilde{V}]^G$ . To find a system of parameters for  $\mathbb{R}[W]$ , we use the fact that  $W$  is the standard 2-dimensional representation of  $S_3$  lifted to  $S_4$ , so there is a system of parameters  $k_2, k_3$  with  $\text{deg } k_2 = 2$  and  $\text{deg } k_3 = 3$ . This gives a system of parameters for  $\mathbb{R}[\tilde{V} \oplus W]^G$ . We can compute the multigraded Poincaré series corresponding to  $\mathbb{R}[\tilde{V} \oplus W]^G$  using Molien's Theorem. We have

$$\begin{aligned} P(s, t) &= \sum_{i, j \geq 0} \dim \mathbb{R}[\tilde{V} \oplus W]_{i, j}^G s^i t^j = \frac{1}{|G|} \sum_{g \in G} \frac{1}{\det(I - s\phi_1(g)) \det(I - t\phi_2(g))} \\ &= \frac{1 + s^2 t + s^2 t^2 + s^4 t + s^4 t^4 + s^6 t^3}{(1 - s)(1 - s^2)(1 - s^3)(1 - s^4)(1 - t^2)(1 - t^3)}. \end{aligned}$$

The form of the denominator agrees with the system of parameters we found above.

To explain the terms appearing in the numerator, we can argue as follows. We have

$$\mathbb{R}[\tilde{V} \oplus W]^G = (S\tilde{V}^* \otimes SW^*)^G = \bigoplus_{(p,q)} (S^p \tilde{V}^* \otimes S^q W^*)^G = \bigoplus_{(p,q)} \text{Hom}_G(S^p \tilde{V}, S^q W),$$

since  $W^* = W$ . Let  $I$  be the ideal of  $\mathbb{R}[\tilde{V}]$  generated by  $s_1, s_2, s_3, s_4$ . Then  $\mathbb{R}[\tilde{V}] \simeq \mathbb{R}[\tilde{V}]/I \otimes \mathbb{R}[\tilde{V}]^G$  as  $G$ -modules. But  $\mathbb{R}[\tilde{V}]/I$  is equivalent to the regular representation of  $S_4$  (see [3], [2, ch. 5, 5.2, Th. 2] or [11, Prop. 4.9]), and it follows from [10] (see also [6]) that we have the following  $G$ -module decomposition of  $\mathbb{R}[\tilde{V}]/I$ . (We use the notation  $\text{sgn} = (1, 1, 1, 1)$  and  $V' = \otimes \text{sgn}$ .)

total degree	0	1	2	3	4	5	6
representation	$k$	$V$	$V \oplus W$	$V \oplus V'$	$W \oplus V'$	$V'$	$\text{sgn}$

We have a similar  $G$ -module decomposition of  $\mathbb{R}[W]/(k_2, k_3)$ .

total degree	0	1	2	3
representation	$k$	$W$	$W$	$\text{sgn}$

It now follows that  $\text{Hom}_G(S^p \tilde{V}, S^q W)$  is nontrivial if and only if

$$(p, q) \in \{(0, 0), (2, 1), (2, 2), (4, 1), (4, 2), (6, 3)\}.$$

This corresponds to the terms in the numerator of the Poincaré series. This proves that there exists invariants  $f_i$  for  $i = 1, \dots, 6$  with degrees corresponding to the terms in the numerator of the Poincaré series, so that if we set  $A = \mathbb{R}[s_1, s_2, s_3, s_4, k_2, k_3]$ , then  $\mathbb{R}[\tilde{V} \oplus W]^G = \bigoplus f_i A$ .

When giving concrete expressions for the invariants, we find it easier to work with the entries  $a_{ij}$  of the symmetric matrix. Let  $s_i$  denote the elementary symmetric functions in the  $a_{ij}$ . We also have the following invariants.

$$\begin{aligned} g_2 &= a_{12}a_{34} + a_{13}a_{24} + a_{14}a_{23}, \\ g_3 &= a_{12}a_{23}a_{31} + a_{12}a_{24}a_{41} + a_{13}a_{34}a_{41} + a_{23}a_{34}a_{42}, \\ g_4 &= a_{12}a_{23}a_{34}a_{41} + a_{12}a_{24}a_{43}a_{31} + a_{13}a_{32}a_{24}a_{41}, \\ h_3 &= a_{12}a_{13}a_{14} + a_{12}a_{23}a_{24} + a_{23}a_{13}a_{34} + a_{24}a_{34}a_{14}. \end{aligned}$$

We set

$$t_2 = s_2 - g_2, \quad t_3 = s_3 - g_3 - h_3, \quad t_4 = s_4 - g_4.$$

We then have the following theorem.

**Theorem 4.** *We can write*

$$\mathbb{R}[\tilde{V} \oplus W] = \bigoplus_{i=1}^6 f_i \mathbb{R}[y_1, \dots, y_6],$$

where the  $y_i$  are  $s_1, g_2, t_2, g_3, h_3, g_4$  and the  $f_i$  are  $1, t_3, t_4, s_5, t_3^2, t_3^3$ .

*Proof.* Some simple computations show that we can replace the system of parameters  $s_1, s_2, s_3, s_4, k_2, k_3$  described above by the system of parameters given in the theorem, and that the listed  $f_i$  correspond to the invariants generating the nontrivial  $\text{Hom}_G(S^p \tilde{V}, S^q W)$ .  $\square$

We have also studied the syzygies. Using the computer algebra system Macaulay [1] and the methods described in [12], we were able to determine all the basic syzygies. Unfortunately, the syzygies are quite complicated, so we have left them out.

It follows from the theorem that the algebra of invariants is generated by the 9 invariants  $s_1, g_2, t_2, g_3, h_3, g_4, t_3, t_4$  and  $s_5$ . So the congruence class of an  $m \times 4$  matrix  $A$  is determined by the values of these 9 invariants on  $A^t A$ .

We would finally like to make some comments about the case  $n=5$ . In that case the elementary symmetric functions  $s_1, \dots, s_5$  in 5 variables would be a system of parameters for  $(5) \oplus (4, 1)$ , while a system of parameters for  $(3, 2)$  consisting of 5 invariants has been given by Dixmier in [4]. But when we computed the Poincaré series, we got 720 terms in the numerator, so it seems difficult to give a simple description of the ring of invariants.

## References

1. Bayer, D., Stillman, M.: Macaulay: A system for computation in algebraic geometry and commutative algebra
2. Bourbaki, N.: Groupes et algèbres de Lie, Chap. 4–6, Hermann, Paris, 1975
3. Chevalley, C.: Invariants of finite groups generated by reflections. *Am. J. Math.* **77**, 778–782 (1955)
4. Dixmier, J.: Sur les invariants du group symétrique dans certains représentations, II. In: Malliavin, M.-P. (ed.) *Topics in Invariant Theory*, Séminaire P. Dubreil et M.-P. Malliavin 1989–1990, Lecture Notes in Math., vol. 1478, pp 1–34. Berlin, Heidelberg, New York: Springer 1991
5. Fulton, W., Harris, J.: *Representation Theory*, Graduate Text in Mathematics, vol. 129. Berlin, Heidelberg, New York: Springer 1991
6. Garsia, A.M., Procesi, C.: On certain graded  $S_n$ -modules and the  $q$ -Kostka polynomial. *Adv. Math.* **94**, 82–138 (1992)
7. Grove, L.C., McShane, J.M.: Polynomial invariants of finite groups *Algebras Groups Geom.* **10**, 1–12 (1993)
8. Jacobson, N.: *Basic Algebra I*. Oxford: W.H. Freeman 1985
9. Kemper, G.: The package for calculating rings of invariant, Maple share library, Waterloo Maple Software, Maple V Release 3, Waterloo, Canada, 1994
10. Kraft, H.: Conjugacy classes and Weyl group representations. *Tableaux de Young et foncteur de Schur en algèbre et géométrie* (Torun, 1980), Asterisque 87–88, Société mathématique de France, 1981, pp. 191–205
11. Stanley, R.P.: Invariants of finite groups and their applications to combinatorics. *Bull. Am. Math. Soc. (N.S.)* **1**, 475–511 (1979)
12. Stillman, M.: Methods for computing in algebraic geometry and commutative algebra. *Acta Appl. Math.* **21**, 77–103 (1990)
13. Sturmfels, B.: *Algorithms in invariant theory*. Wien, Springer 1993