INVARIANT THEORY OF SPECIAL ORTHOGONAL GROUPS

HELMER ASLAKSEN, ENG-CHYE TAN AND CHEN-BO ZHU

INVARIANT THEORY OF SPECIAL ORTHOGONAL GROUPS

Helmer Aslaksen, Eng-Chye Tan and Chen-bo Zhu

In this paper we study the action of $SO(n)$ on m-tuples of $n \times n$ matrices by simultaneous conjugation. We show that the polynomial invariants are generated by traces and polarized Pfaffians of skewsymmetric projections. We also discuss the same problem for other classical groups.

1. Special orthogonal groups. Let F be a field of characteristic 0. If A is a skewsymmetric $2k \times 2k$ matrix over F, we denote the Pfaffian of A by $\operatorname{pf} A$. It satisfies $\det A = \operatorname{pf}^2 A$ and $\operatorname{pf}(gAg^t) = \det g \operatorname{pf} A$. For an arbitrary $2k \times 2k$ matrix M, we define $\widetilde{\operatorname{pf}}(M) = \operatorname{pf}(M - M^t)$ to be the Pfaffian of the skewsymmetric projection of M. This is clearly an $SO(2k, F)$ invariant. By abuse of notation we will refer to $\widetilde{\operatorname{pf}}$ as the Pfaffian, too.

Let $W = W(n, m, F)$ be the vector space of m-tuples of $n \times n$ matrices over F on which a group $G \subset GL(n, F)$ acts by simultaneous conjugation. For $G = SO(2, F)$,the invariants $P[W(2, m, F)]^G$ were determined in [1]. They are generated by the invariants $\operatorname{tr} P(A, A^t)$ and $\widetilde{\operatorname{pf}} P(A, A^t)$ where $A \in W(2, m, F)$ and P is noncommutative polynomial.

We will see that for n odd we do not get any more invariants when we restrict $O(n, F)$ to $SO(n, F)$. In the even case we have the following crucial Lemma that indicates why Pfaffian appears.

LEMMA 1. *Let* $x_1, \dots, x_{2k} \in F^{2k}$ *and let* $[x_1, \dots, x_{2k}]$ *denote the determinant of the matrix with columns* $x_1, \dots, x_{2k}$. *Then*

$$[x_1, \dots, x_{2k}] = \widetilde{\operatorname{pf}}(x_1 x_2^t + \dots + x_{2k-1} x_{2k}^t).$$

Proof. Let X be the matrix with columns $x_1, \ldots, x_{2k}$ and set

$$
\tilde{J} = \begin{pmatrix} 0 & 1 & & & \\ -1 & 0 & & & \\ & & \ddots & & \\ & & & 0 & 1 \\ & & & -1 & 0 \end{pmatrix}.
$$

Then

$$
\widetilde{\mathrm{pf}}(x_1 x_2^t + \cdots + x_{2k-1} x_{2k}^t) = \mathrm{pf}(X \tilde{J} X^t) = \det X \, \mathrm{pf} \, \tilde{J} = \det X.
$$

$\square$

When $k > 1$ the Pfaffian will no longer be linear. Instead we will consider the polarized Pfaffian, which we will denote by pl. Thus $\mathrm{pl}(A_1, \ldots, A_k)$ is the coefficient of $t_1 \cdots t_k$ in the expansion of $\widetilde{\mathrm{pf}}(t_1 A_1 + \cdots + t_k A_k)$. It is a symmetric, multilinear function of k matrices and satisfies $\mathrm{pl}(A, \ldots, A) = k! \widetilde{\mathrm{pf}}(A)$.

When proving our main theorem about the invariants of $\mathrm{SO}(2k, F)$ we will use classical invariant theory and consider decomposable matrices of the form $A_i = u_i v_i^t$. Notice that rank $u_i v_i^t \leq 1$. We will need the following simple Lemma about the polarized Pfaffian.

LEMMA 2. *Let* $A_1, \ldots, A_k$ *be of the form* $A_i = u_i v_i^t$. *The polarized Pfaffian is alternating in the* A_i, *i.e.,*

$$
\mathrm{pl}(A_{i_1}, \ldots, A_{i_k}) = 0
$$

if two of the A_i, *are equal. Hence*

$$
\widetilde{\mathrm{pf}}(A_1 + \cdots + A_k) = \mathrm{pl}(A_1, \ldots, A_k).
$$

Proof. Set $A = t_1 A_{i_1} + \cdots + t_k A_{i_k}$. Then $\mathrm{pl}(A_{i_1}, \ldots, A_{i_k})$ is the coefficient of $t_1 \cdots t_k$ in the expansion of $\widetilde{\mathrm{pf}}(A)$. But if the A_{i_k} are not all distinct, the rank of A will be at most $k - 1$. But then the rank of $A - A^t$ will be at most $2k - 2$, so $\det(A - A^t) = 0$. Hence

$$
\widetilde{\mathrm{pf}}(A) = \mathrm{pf}(A - A^t) = 0,
$$

and so $\mathrm{pl}(A_{i_1}, \dots, A_{i_k}) = 0$. Combining this with the multilinearity of pl, we get

$$\widetilde{\mathrm{pf}}(A_1 + \cdots + A_k) = \mathrm{pl}(A_1, \dots, A_k, \cdots, A_1, \dots, A_k,)/k!$$
$$= \mathrm{pl}(A_1, \dots, A_k).$$

$\square$

We can now prove our main theorem.

THEOREM 3. *When $G = \mathrm{SO}(2k+1, F)$, the invariants $P[W(2k+1, m, F)]^G$ are the same as the $O(2k+1, F)$ invariants. When $G = \mathrm{SO}(2k, F)$, the invariants $P[W(2k, m, F)]^G$ are generated by traces and polarized Pfaffians of the A_i and A_i^t, i.e.,*

$$\mathrm{tr}\, P(A, A^t) \quad and \quad \mathrm{pl}\left(P_1(A, A^t), \dots, P_k(A, A^t)\right),$$

where $P, P_1, \dots, P_k$ are noncommutative polynomials.

Proof. The proof will follow from classical invariant theory. We can first reduce the problems to finding the multihomogeneous invariants of order $(d_1, \dots, d_r)$ in m matrices, and then reduce further to studying multilinear invariants of dm matrices where $d = \sum_{i=1}^r d_i$. We will identify $F^n \otimes F^n$ and $M(n, F)$ using

$$u \otimes v \rightarrow uv^t.$$

We can assume that $A_i = u_i \otimes v_i$ (the symbolic method). The invariants of $(F^n \otimes F^n)^{\otimes dm}$ are generated by inner products and determinants, i.e., invariants of the form

$$\phi(x_1 \otimes \cdots \otimes x_{2dm})$$
$$= [x_1, \dots, x_n] \cdots [x_{n(l-1)+1}, \dots, x_{nl}]\langle x_{nl+1}, x_{nl+2}\rangle \cdots \langle x_{2dm-1}, x_{2dm}\rangle$$

where $\langle x_i, x_j\rangle = x_i^t x_j$ and $[x_1, \dots, x_n]$ denotes the determinant of the matrix with columns $x_1, \dots, x_n$.

We first observe that if n is odd, we must have an even number of determinants. But then the whole expression will be an $O(n, F)$ invariant. This proves the first part of the theorem.

We can therefore assume that $n = 2k$. We observe that

$$\langle x_i, x_j\rangle = x_i^t x_j = \mathrm{tr}\, x_i x_j^t = \mathrm{tr}\, x_i \otimes x_j.$$

If w_i is u_i or v_i and w_i' the other one, then a product of the form

$$\langle w_{i_1}, w_{i_2}'\rangle\langle w_{i_2}, w_{i_3}'\rangle \cdots \langle w_{i_r}, w_{i_1}'\rangle$$

is equal to

$$\operatorname{tr} B_{i_1} \cdots B_{i_r},$$

where B_i equals A_i or A_i^t depending on whether w_i is u_i or v_i.

We will now show that ϕ can be written in terms of traces and polarized Pfaffians of the A_i and the A_i^t.

Since the product of two determinants is an $O(2k, F)$ invariant and hence expressible in terms of scalar products, we can assume that we have only one determinant. By reordering the A_i and replacing A_i by A_i^t, i.e., interchanging u_i and v_i, we can assume that

$$\phi = \pm[u_1, v_1, \ldots, u_l, v_l, u_{l+1}, \ldots, u_{2k-l}]$$
$$\langle v_{l+1}, u_{2k-l+1}\rangle\langle v_{2k-l+1}, u_{2k-l+2}\rangle \cdots \langle v_{2k-l+d_1}, v_{l+2}\rangle$$
$$\langle v_{l+3}, u_{2k-l+d_1+1}\rangle \cdots \langle v_{2k-l+d_1+d_2}, v_{l+4}\rangle$$
$$\cdots \langle v_{2k-l-1}, u_{2k-l+d_1+\cdots+d_{k-l-1}+1}\rangle \cdots \langle v_{2k-l+d_1+\cdots+d_{k-l}}, v_{2k-l}\rangle Q.$$

where Q only involves scalar products of u_i and v_i for $i > 2k - l + d_1 + \cdots + d_{k-l}$. But this can be expressed as the trace of a polynomial in A_i and A_i^t for $i > 2k - l + d_1 + \cdots + d_{k-l}$. We can now apply Lemma 2 and we get that

$$(1) \quad [u_1, v_1, \ldots, u_l, v_l, u_{l+1}, \ldots, u_{2k-l}]$$
$$= \widetilde{\operatorname{pf}}(A_1 + \cdots + A_l + u_{l+1}u_{l+2}^t + \cdots + u_{2k-l-1}u_{2k-l}^t)$$
$$= \operatorname{pl}(A_1, \ldots, A_l, u_{l+1}u_{l+2}^t, \ldots, u_{2k-l-1}u_{2k-l}^t).$$

But since the function pl is multilinear, we can combine

$$\langle v_{l+1}, u_{2k-l+1}\rangle \langle v_{2k-l+1}, u_{2k-l+2}\rangle \cdots \langle v_{2k-l+d_1}, v_{l+2}\rangle$$

with $u_{l+1}u_{l+2}^t$ to get

$$u_{l+1}u_{l+2}^t v_{l+1}^t u_{2k-l+1} v_{2k-l+1}^t u_{2k-l+2} \cdots v_{2k-l+d_1}^t v_{l+2}$$
$$= u_{l+1}v_{l+1}^t u_{2k-l+1}v_{2k-l+1}^t u_{2k-l+2} \cdots v_{2k-l+d_1}^t v_{l+2}u_{l+2}^t$$
$$= A_{l+1}A_{2k-l+1} \cdots A_{2k-l+d_1}A_{l+2}^t.$$

Repeating this for each of the last $k - l$ terms in (1), we get that

$$\phi/Q = \mathrm{pl}(A_1, \ldots, A_l, A_{l+1}A_{2k-l+1} \cdots A_{2k-l+d_1}A_{l+2}^t, \ldots,$$
$$A_{2k-l-1}A_{2k-l+d_1+\cdots+d_{k-l-1}+1} \cdots A_{2k-l+d_1+\cdots+d_{k-l}}A_{2k-l}^t).$$

This shows that ϕ is the required form.   □

We now want to derive a similar result for $\mathrm{SO}(p, q)$. The method is similar to the case of $\mathrm{SO}(n, F)$, except that we identify $F^n \otimes F^n$ and $M(n, F)$ using

$$u \otimes v \to u(I_{p,q}v)^t = uv^t I_{p,q},$$

where

$$I_{p,q} = \begin{pmatrix} I_p & 0 \\ 0 & -I_q \end{pmatrix}.$$

The corresponding involution of $M(n, F)$ is

$$A^* = I_{p,q}^{-1}A^t I_{p,q},$$

so that $(u \otimes v)^* = v \otimes u$. When $p + q = 2k$ is even, we also need to change the Pfaffian to the (p, q) Pfaffian. It is defined by

$$\widetilde{\mathrm{pf}}_{p,q}(A) = \mathrm{pf}(I_{p,q}A - A^t I_{p,q}) = \widetilde{\mathrm{pf}}(I_{p,q}A).$$

It is easy to see that the (p, q) Pfaffian is an $\mathrm{SO}(p, q)$ invariant. Moreover, it satisfies the following modified version of Lemma 1

$$\widetilde{\mathrm{pf}}_{p,q}(x_1 \otimes x_2 + \cdots + x_{2k-1} \otimes x_{2k}) = (-1)^q[x_1, \ldots, x_{2k}].$$

We let $\mathrm{pl}_{p,q}$ be the polarization of $\widetilde{\mathrm{pf}}_{p,q}$. We can now derive the following theorem.

THEOREM 4. *Let* $G = \mathrm{SO}(p, q)$. *Define* $A^* = I_{p,q}^{-1}A^t I_{p,q}$. *When* $p + q$ *is odd, the invariants* $P[W(p + q, m, \mathbb{R})]^G$ *are the same as for* $O(p, q)$. *When* $p + q = 2k$ *is even, the invariants* $P[W(p+q, m, \mathbb{R})]^G$ *are generated by traces and polarized* (p, q) *Pfaffians of the* A_i *and* A_i^*, *i.e.,*

$$\mathrm{tr}\, P(A, A^*) \text{ and } \mathrm{pl}_{p,q}\left(P_1(A, A^*), \ldots, P_k(A, A^*)\right),$$

where $P, P_1, \ldots, P_k$ *are noncommutative polynomials.*

2. Classical groups. We will now discuss the invariants of other classical groups. Let us first define what we mean by a classical group. Let F be either $\mathbb{R}$ or $\mathbb{C}$ and let D be a division algebra over F with involution $a \mapsto a^{\#}$. Let V be a vector space over D and let (,) be a nondegenerate sesquilinear form on V that is Hermitian or skew-Hermitian relative to $^{\#}$. (If the involution is the identity we get bilinear symmetric and skewsymmetric forms.) Let G be the isometry groups. We will call such groups classical isometry groups.

If the involution is either the identity or complex or quaternionic conjugation, we can describe the classical isometry groups by Table 1.

	$\mathbb{R}$	$\mathbb{C}$	$\mathbb{H}$
bilinear and symmetric	$O(p,q)$	$O(n,\mathbb{C})$	–
bilinear and skewsymmetric	$\mathrm{Sp}(k,\mathbb{R})$	$\mathrm{Sp}(k,\mathbb{C})$	–
sesquilinear and Hermitian	–	$U(p,q)$	$\mathrm{Sp}(p,q)$
sesquilinear and skew-Hermitian	–	$U(p,q)$	$\mathrm{Sp}(k,\mathbb{H})$

Table 1

(We are assuming that $n = p + q$ or $n = 2k$ when appropriate.) A "–" means that there are no such forms. Helgason [2] writes $SO^*(2k)$ for the group we have called $\mathrm{Sp}(k,\mathbb{H})$.

The remaining classical groups are the general and special linear groups $GL(n, F)$ and $SL(n, F)$ and the special isometry groups, i.e., the intersections of $SL(n, F)$ and the classical isometry groups.

Sibirskii [4] gave a set of generators of $P[W]^G$ when G equals $GL(n, F)$, $O(n, F)$ or $U(n)$. His proof for the case of $U(n)$ is essentially Weyl's unitary trick, which can easily be adapted to work for $U(p, q)$. These results were also proved independently by Procesi [3], who also solved the problem for $\mathrm{Sp}(k, F)$. His proof for the $O(n, \mathbb{R})$ case can also easily be modified for the case of $O(p, q)$. For the special isometry groups, all the cases are trivial except for the special orthogonal groups.

It therefore remains to discuss the quaternionic groups $\mathrm{Sp}(k, \mathbb{H})$ and $\mathrm{Sp}(p, q)$. We will use the complexification trick to reduce $\mathrm{Sp}(k, \mathbb{H})$ to $SO(2k, \mathbb{C})$ and $\mathrm{Sp}(p, q)$ to a group isomorphic to $\mathrm{Sp}(p + q, \mathbb{C})$.

Let G be a real connected Lie group that acts rationally on a real vector space V. We can find a complex connected group $G_{\mathbb{C}}$

such that the Lie algebra of $G_\mathbb{C}$ is $\mathfrak{g}_\mathbb{C} = \mathfrak{g} \oplus i\mathfrak{g}$ where $\mathfrak{g}$ is the Lie algebra of G and with $G \subset G_\mathbb{C}$. Then both G and $G_\mathbb{C}$ act on $V_\mathbb{C} = V \otimes_\mathbb{R} \mathbb{C} = V \oplus iV$, and for $A = X + iY \in V_\mathbb{C}$ with $X, Y \in V$ we define $A^* = X - iY$. We prove the following, which is a generalization of Lemma 1 in [**4**].

LEMMA 5. *Consider $f(A, B) \in P(V_\mathbb{C} \oplus V_\mathbb{C})$. Then $f(A, B)$ is $G_\mathbb{C}$ invariant if and only if $f(A, A^*)$ is G invariant.*

Proof. The lemma follows easily from the following two facts.
(i) $f(A, B)$ is G invariant if and only if it is $G_\mathbb{C}$ invariant.
(ii) If $f(A, A^*) = 0$ for all $A \in V_\mathbb{C}$, then $f(A, B) = 0$ for all $A, B \in V_\mathbb{C}$.

To prove (i), we fix arbitrary $A, B \in V_\mathbb{C}$ and consider the function

$$F(g) = f(g \cdot A, g \cdot B) - f(A, B), \quad g \in G_\mathbb{C}.$$

We can write $F(g)$ as $h(g)/\det^j g$ for polynomial function h and some $j \in \mathbb{N}$. We now let

$$H(X) = h(\exp X), \quad X \in \mathfrak{g}_\mathbb{C}.$$

We see that H is a power series that vanishes on $\mathfrak{g}$ and hence also on $\mathfrak{g}_\mathbb{C}$. Since $G_\mathbb{C}$ is connected it follows that $h = 0$ and so $F = 0$. This proves (i).

To prove (ii), we write $A = X + iY$ with $X, Y \in V$, and set

$$\tilde{f}(X, Y) = f(A, A^*) = f(X + iY, X - iY) = 0.$$

Since a polynomial that vanishes on real variables also vanishes on complex variables, the above remains true for $X, Y \in V_\mathbb{C}$. We can now substitute $X = (A + B)/2$ and $Y = (A - B)/2i$ to get $f(A, B) = 0$. This proves (ii), and the lemma follows easily.   □

Conversely, suppose that we start with a complex connected Lie group $G_\mathbb{C} \subset GL(n, \mathbb{C})$ and an involution $*$ of $M(n, \mathbb{C})$ such that $(iI_n)^* = -iI_n$. We then obtain a real from G of $G_\mathbb{C}$ by setting $G = G_\mathbb{C} \cap \{gg^* = I_n\}$, and we get the following.

COROLLARY 6. *Let $W = W(n, m, \mathbb{C})$ and consider $f(A, B) \in P(W \oplus W)$. Then $f(A, B)$ is $G_\mathbb{C}$ invariant if and only if $f(A, A^*)$*

is G invariant.

We would like to remark that Corollary 6 allows us to state the known results [**3**, **4**] in a uniform way.

	$O(p,q)$	$O(n,\mathbb{C})$	$Sp(k,\mathbb{R})$	$Sp(k,\mathbb{C})$	$U(p,q)$
K	$I_{p,q}$	I_n	J_k	J_k	$I_{p,q}$

Table 2

THEOREM 7. *Let G be a classical isometry group that is not quaternionic, i.e., $G = \{(g^{\#})^t K g = K\}$ where K is given in Table 2. The ring of invariants $P[W]^G$ is generated by invariants of the form $\operatorname{tr} P(A, A^*)$, where P is a noncommutative polynomial and $A^* = K^{-1}(A^{\#})^t K$.*

When discussing quaternionic groups, we will use the decomposition $\mathbb{H}^n = \mathbb{C}^n \oplus j\mathbb{C}^n$ to identify $\mathbb{H}^n$ with $\mathbb{C}^{2n}$. This gives a corresponding identification of $M(n, m, \mathbb{H})$ with $M(2n, m, \mathbb{C})$. Then

$$Sp(k, \mathbb{H}) = SO(2k, \mathbb{C}) \cap \{g \in GL(2k, \mathbb{C}) \mid \bar{g}^t J_k g = J_k\},$$

where

$$J_k = \begin{pmatrix} 0 & -I_k \\ I_k & 0 \end{pmatrix}.$$

We can apply Corollary 6 to get the following result for $Sp(k, \mathbb{H})$.

THEOREM 8. *Let $G = Sp(k, \mathbb{H})$. Define $A^* = J_k^{-1} \bar{A}^t J_k$. The invariants $P[W(2k, m, \mathbb{C})]^G$ are generated by traces and polarized Pfaffians of A_i, A_i^t, A_i^* and $(A_i^*)^t$, i.e.,*

$$\operatorname{tr} P(A, A^t, A^*, (A^*)^t) \quad and$$

$$\operatorname{pl}\left(P_1(A, A^t, A^*, (A^*)^t), \dots, P_k(A, A^t, A^*, (A^*)^t)\right),$$

where $P, P_1, \dots, P_k$ are noncommutative polynomials.

Since $Sp(p, q)$ is not a subgroup of $Sp(p + q, \mathbb{C})$ in the standard realization, we cannot set $G_{\mathbb{C}} = Sp(p+q, \mathbb{C})$ in Corollary 6. Instead we will introduce a group isomorphic to $Sp(p + q, \mathbb{C})$ that contains $Sp(p, q)$. Let

$$J_{p,q} = \begin{pmatrix} 0 & -I_{p,q} \\ I_{p,q} & 0 \end{pmatrix} \quad and \quad K_{p,q} = \begin{pmatrix} I_{p,q} & 0 \\ 0 & I_{p,q} \end{pmatrix}.$$

We define
$$\mathrm{Sp}(p, q, \mathbb{C}) = \{g \in GL(2p + 2q, \mathbb{C}) \mid g^t J_{p,q} g = J_{p,q}\} \cong \mathrm{Sp}(p + q, \mathbb{C}).$$

If we set
$$A^* = K_{p,q}^{-1} \bar{A}^t K_{p,q},$$

we see that
$$\mathrm{Sp}(p, q) = \mathrm{Sp}(p, q, \mathbb{C}) \cap \{gg^* = I_{2p+2q}\}.$$

We can apply Corollary 6, using $G_{\mathbb{C}} = \mathrm{Sp}(p, q, \mathbb{C})$. Since the invariants for this group can be determined by the method used for $\mathrm{Sp}(p + q, \mathbb{C})$ in [**3**], we can now derive the following theorem.

THEOREM 9. *Let* $G = \mathrm{Sp}(p, q)$. *Define* $A^* = K_{p,q}^{-1} \bar{A}^t K_{p,q}$ *and* $A^{\#} = J_{p,q}^{-1} A^t J_{p,q}$. *The invariants* $P[W(2p + 2q, m, \mathbb{C})]^G$ *are generated by invariants of the form* $\mathrm{tr}\, P(A, A^{\#}, A^*, (A^{\#})^*)$, *where* P *is a noncommutative polynomial.*

We would finally like to thank the referee for helpful suggestions.

REFERENCES

[1]   H. Aslaksen, SO(2) *invariants of a set of* 2×2 *matrices*, Math. Scand., **65** (1989), 59-66.

[2]   S. Helgason, *Differential Geometry, Lie Groups, and Symmetric Spaces*, Academic Press, New York (1978).

[3]   C. Procesi, *The invariant theory of* $n \times n$ *matrices*, Adv. in Math., **19** (1976), 306-381.

[4]   K.S. Sibirskii, *Algebraic invariants for a set of matrices*, Siberian Math. J., **9** (1968), 115-124.

References added in proof:

H. Aslaksen, E.C. Tan and C. Zhu, *Generators and relations of invariants of* 2×2 *matrices*, Comm. Algebra, **22** (1994), 1821-1832.

————, Quivers and the invariant theory of Levi subgroups, J. Funct. Anal., **120** (1994), 163-187.

NATIONAL UNIVERSITY OF SINGAPORE
SINGAPORE 0511, REPUBLIC OF SINGAPORE
E-mail addresses: aslaksen@math.nus.sg
                    mattanec@math.nus.sg
                    matzhucb@math.nus.sg

CONTENTS

PACIFIC JOURNAL OF MATHEMATICS