SO(2) INVARIANTS OF A SET OF 2 \times 2 MATRICES

HELMER ASLAKSEN

Abstract.

We give an alternative proof of a result due to Sibirskii on the polynomial invariants of SO(2, C) (or SO (2, R)) acting on M(2, C) (or M(2, R)) by conjugation. We show that the invariants are given in terms of traces and Pfaffians, and we find a minimal basis which is a minimal complete set of invariants in the real case.

The polynomial invariants of $O(2, \mathbb{C})$ (or $O(2, \mathbb{R})$) acting on $M(2, \mathbb{C})$ (or $M(2, \mathbb{R})$) by conjugation has been studied by Sibirskii [8]. In this paper we study the invariants when restricting to SO (2, \mathbb{C}) (or SO (2, \mathbb{R})). After submitting a first version of this paper, we were informed by Professor Sibirskii that this case had already been studied by him. The results in this paper are equivalent to results in [10, pp. 126–127], but our approach is different. We essentially follow the approach in [8], instead of using the resuls of [9].

Let $\{f_j\}$ be a set of invariants. We will call $\{f_j\}$ a basis if any invariant can be expressed polynomially in the f_j -s. We will call $\{f_j\}$ a functional basis if any invariant can be expressed as a function (not necessarily a plynomial) in the f_j -s. We will call $\{f_j\}$ a complete set of invariants if they separate orbits (i.e., conjugacy classes).

The starting point is the following results from [8], which was later proved independently by Procesi [5].

THEOREM 1. Let $\{A_i\}$ be a set of complex (or real) $n \times n$ matrices. The invariants of the form tr (A_i, A_i^t) , where P is a monomial in the A_i and A_i^t , form a basis for the $O(n, \mathbb{C})$ (or $O(n, \mathbb{R})$) invariants of the A_i . In the real case the invariants also form a complete set of $O(n, \mathbb{R})$ invariants.

Partially supported by a grant from Norwegian Reseach Council for Science and Humanities (NAVF) and NSF grant DMS 87-01609. Received July 3, 1988

The fact that they form a basis was proved for the case of one matrix by Gurevich [2], while the fact that they form a complete set of inariants is due to Pearcy [4]. The problem is now to reduce these trace expressions and find a finite basis. In the 2×2 case, Sibirskii proved the following.

THEOREM 2. Let $\{A_i\}$ be a set of complex (or real) 2×2 matrices. 1. The invariants

> tr A_i , tr A_iA_j ($i \leq j$), tr $A_iA_j^t$ ($i \leq j$), tr $A_iA_jA_j^t$ ($i \neq j$), tr $A_i^tA_jA_k$, tr $A_iA_j^tA_k$, tr $A_iA_jA_k^t$ (i < j < k)

form a minimal basis for the $O(2, \mathbb{C})$ (or $O(2, \mathbb{R})$) invariants of the A_i . The invariants tr $A_i A_j A_k (i < j < k)$ can replace any of the last three types of invariants.

2. The invariants

tr
$$A_i$$
, tr $A_i A_j (i \leq j)$, tr $A_i A_j^t (i \leq j)$, tr $A_i A_j A_j^t (i \neq j)$, tr $A_i A_j A_k (i < j < k)$

form a minimal functional basis of $O(2, \mathbb{C})$ invariants of the A_i . In the real case they also form a minimal complete set of $O(2, \mathbb{R})$ invariants of the A_i .

In the complex case the invariants do not separate orbits, as the following example from [8] shows. Set

$$A = \begin{pmatrix} 2 & i \\ i & 0 \end{pmatrix}.$$
 Then $A^k = \begin{pmatrix} k+1 & ki \\ ki & 1-k \end{pmatrix},$

so tr $A^k = 2$. Hence the invariants do not separate A and I_2 . The reason for this is essentially that $O(2, \mathbb{C})$ is non-compact.

It is well known from classical invariant theory that when considering SO (2) invariants we must include certain determinants. We will first observe that these determinants can be expressed in terms of traces and Pfaffians of the A_i . We define the Pfaffian of a (not necessarily skew-symmetric) 2×2 matrix by

$$pf(a_{ij}) = a_{12} - a_{21}$$

It is easy to show that

 $pf(g^t Ag) = \det g \, pf A,$

so the Pfaffian is an SO(2) invariant but not an O(2) invariant. We want to show

THEOREM 3. Let $\{A_i\}$ be a set of complex (or real) 2×2 matrices.

1. The invariants of the form tr $P(A_i, A_i^t)$ and pf $P(A_i, A_i^t)$ where P is a monomial in the A_i and A_i^t , form a basis for the SO(2, C) (or SO(2, R)) invariant of the A_i . In the real case they also form a complete set of SO(2, R) invariants of the A_i .

2. The invariants

(1)

$$A_i, \operatorname{tr} A_i^2, \operatorname{tr} A_i A_j (i < j), \operatorname{pf} A_i \text{ and } \operatorname{pf} A_i A_j (i < j)$$

form a minimal basis (and a minimal functional basis) for the SO (2, C) invariants of the A_i . In the real case they also form a minimal complete set of SO (2, R) invariants of the A_i .

Part 1 will follow from classical invariant theory, using an approach similar to Procesis [5]. Let K denote R or C. We can first reduce the problems to finding the multihomogeneous invariants of order (d_1, \ldots, d_k) , and then reduce further to studying multilinear invariants of $(K^2 \otimes K^2)^{\otimes d}$, $d = \sum_{i=1}^k d_i$. We will use the correspondence

$$u \otimes v \to uv^t$$

between M(2, K) and $K^2 \otimes K^2$ and we can assume that $A_i = u_i \otimes u_i$. The invariants of $(K^2 \otimes \leq^2)^{\otimes d}$ are generated by inner products and determinants, i.e., invariants of the form

$$\phi(x_1 \otimes \ldots \otimes x_{2d}) =$$
$$\langle x_{i_1}, x_{i_2} \rangle \ldots \langle x_{i_{2l-1}}, x_{i_{2l}} \rangle [x_{i_{2l+1}}, x_{i_{2l+2}}] \ldots [x_{i_{2d-1}}, x_{i_{2d}}]$$

where
$$\langle x_i, x_j \rangle = x_i^t x_j$$

and $[x_i, x_i] = \det(x_i, x_i)$.

Here (x_i, x_j) denotes the matrix with columns x_i and x_j . Now we observe that

(2)
$$\langle x_i, x_j \rangle = x_i^t x_j = \operatorname{tr} x_i x_j^t = \operatorname{tr} x_i \otimes x_j$$

 $[x_i, x_j] = \det(x_i, x_j) = \operatorname{pf} x_i x_j^t = \operatorname{pf} x_i \otimes x_j,$

and we claim that all invariants of type (1) can be written in terms of traces and Pfaffians of the A_i . Consider $\langle w_1, w'_2 \rangle \langle w_2, w'_3 \rangle \dots \langle w_l, w'_1 \rangle$ where w_i is either u_j or v_j and

$$w'_i = \begin{cases} u_j & \text{if } w_i = v_j \\ v_j & \text{if } w_i = u_j. \end{cases}$$

Then $(w'_1 \otimes w_1)(w'_2 \otimes w_2) \dots (w'_l \otimes w_l) = w'_1 w_1^t w_2^t w_2^t \dots w'_l w_l^t = \langle w_1, w'_2 \rangle \dots \langle w_{l-1}, w'_l \rangle w'_1 w_l^t$. Taking the trace we get

$$\langle w_1, w_2' \rangle \langle w_2, w_3' \rangle \dots \langle w_l, w_1' \rangle = \operatorname{tr} [(w_1' \otimes w_1)(w_2' \otimes w_2) \dots (w_l' \otimes w_l)].$$

If we instead take the Pfaffian and use (2), we get

$$\langle w_1, w_2' \rangle \dots \langle w_{l-1}, w_l' \rangle [w_1', w_l] = pf[(w_1' \otimes w_1)(w_2' \otimes w_2) \dots (w_l' \otimes w_l)].$$

Since the products involving an even number of determinants are O(2) invariant, and hence expressible in terms of traces, we need only consider invariants of type (1) with one determinant factor. Since $w'_i \otimes w_i = A_j$ or A_j^t , we see that the traces and Pfaffians of $P(A_i, A_i^t)$ generate the ring of SO (2) invariants.

We will now prove that the invariants separate orbits in the real case. Assume that the traces and Pfaffians agree on two sets of matrices, A_i and B_i . Since the traces separate O(2) orbits, there must be a g in O(2) with $gA_ig^{-1} = B_i$. We want to show that g is in SO (2). If at least on of the A_i is non-symmetric, i.e. $pf A_i \neq 0$, it follows from $pf B_i = pf(gA_ig^{-1}) = \det g pf A_i = \det g pf B_i$ that $\det g = 1$, so g is actually in SO (2). Assume then that the A_i are all symmetric. If there is at least one pair, A_i and A_j , which do not commute, then A_iA_j is not symmetric, and hence $pf B_iB_j = pf(gA_iA_jg^{-1}) = \det g pf A_iA_j = \det g pf B_iB_j$, which implies that g is in SO (2). If they all commute, then they are simultaneously diagonalizable by conjugation with g in O(2), but sine diagonal matrices commute, we can assure that g is in SO (2) by multiplying g with the diagonal matrix diag (1, -1). It then follows that the A_i are SO (2) conjugate and hence the invariants separate SO(2) orbits. This completes the proof of the first part of the Theorem.

We will say that pf $P(A_i, A_i^t)$ is reducible if it can be expressed in terms of traces and Pfaffians of products of fewer matrices. We will write

$$\operatorname{pf} F(A_i, A_i^t) \equiv \operatorname{pf} G(A_i, A_i^t)$$

if pf(F - G) is reducible. In order to prove the second part of the theorem, we first need to reduce expressions of the form $pf P(A_i, A_i^t)$. Let us first state some basic properties of the Pfaffian of 2×2 matrices which follow from a simple calculation.

Lemma 1.

(3)
$$pf X^{t} = -pf X$$

(4) $pf YX = pf X tr Y + tr X pf Y - pf XY$
(5) $pf XY^{t} = pf XY - tr X pf Y.$

Hence

$$pf YX \equiv -pf XY$$
$$pf XY^{t} \equiv pf XY.$$

We see that (3) and (4) are more complicated than the corresponding formulas for the trace, but (5) is a big simplification which will allow us to carry the reduction further than in the case of the trace.

For n = 2 the Cayley-Hamilton Theorem says that

(6)
$$X^{2} - X \operatorname{tr} X + 1/2I [\operatorname{tr} X)^{2} - \operatorname{tr} X^{2}] = 0$$

or in its polarized version

(7)
$$XY + YX - X\operatorname{tr} Y - Y\operatorname{tr} X + I[\operatorname{tr} X\operatorname{tr} Y - \operatorname{tr} XY] = 0.$$

This equation is the fundamental tool for reducing Pfaffian expressions, just as in the case of trace expressions. In fact it is even more powerful in the case of the Pfaffian, since if we take the trace in (6) or (7) everything cancels. In order to get a non-trivial relation, we must first multiply the equations with a matrix $\neq I$ before taking the trace. This is not necessary if we take the Pfaffian. In particular, (4) follows from taking the Pfaffian of (7), and taking the Pfaffian of (6) we get that pf X^2 is reducible. By comparison, tr X^2 is not reducible, but if we first multiply (6) by X and then take the trace, we see that tr X^3 is reducible.

It is well known (see for example [1] or [3]) that the polarized Cayley-Hamilton Theorem implies that any product of three 2×2 matrices is reducible. That is, XYZ can be written as a linear combination of matrix products with fewer factors and coefficients expressible in terms of traces. Writing

$$\operatorname{tr}(X, Y) = \operatorname{tr} XY - \operatorname{tr} X\operatorname{tr} Y,$$

we have

2X YZ = X(YZ + ZY) + (XY + YX)Z - [Y(XZ) + (XZ)Y] =XY tr Z + XZ tr Y + X tr (Y, Z) + XZ tr Y + YZ tr X + Z tr (X, Y)- XZ tr Y - Y tr (XZ) - I tr (XZ, Y) = XY tr Z + XZ tr Y+ YZ tr X + X tr (Y, Z) - Y tr (XZ) + Z tr (X, Y) - I tr (XZ, Y).

More generally, it follows from the work of Procesi [5] and Razmyslov [7] that the product of $n^2 - 1$ matrices of order *n* is reducible. For 3×3 matrices the product of 6 matrices is reducible [1].

Consider an irreducible expression of the form pf $P(A_i, A_i^t)$ where P is a monomial. It follows from (8) that P can have at most 2 factors and (6) shows that there can be no squares. Using (4) and (5) we can assume that there are no A_i^t -s and that the A_i -s are in the order of increasing *i*-s. This leaves us with pf A_i and pf A_iA_j (i < j).

This implies that the traces listed in part 1 of Theorem 2 together with pf A_i and pf $A_iA_j(i < j)$ form a basis for the SO(2) invariants. This basis is not minimal, however, since by multiplying two Pfaffians we get an O(2) invariant which is expressible in terms of the traces. A simple argument gives the following relations.

Lemma 2.

(9)
$$pf X pf Y = tr X Y^t - tr X Y$$

(10)
$$pf XY pf Z = -tr XYZ + tr XYZ^{t}$$

We will also use the following relation from [8].

(11)
$$\operatorname{tr} X YZ = \operatorname{tr} X YZ^{t} + \operatorname{tr} X Y^{t}Z + \operatorname{tr} X^{t} YZ$$
$$-\operatorname{tr} X \operatorname{tr} YZ^{t} - \operatorname{tr} Y \operatorname{tr} ZX^{t} - \operatorname{tr} Z \operatorname{tr} XY^{t} + \operatorname{tr} X \operatorname{tr} Y \operatorname{tr} Z.$$

Combining (10) and (11) we get

(12)
$$2 \operatorname{tr} X Y = \operatorname{tr} X \operatorname{tr} Y Z^{t} + \operatorname{tr} Y \operatorname{tr} Z X^{t} + \operatorname{tr} Z \operatorname{tr} X Y t$$
$$-\operatorname{pf} X \operatorname{pf} Y Z - \operatorname{pf} Y \operatorname{pf} Z X - \operatorname{pf} Z \operatorname{pf} X Y - \operatorname{tr} X \operatorname{tr} Y \operatorname{tr} Z$$

Assume now that we know tr A_i , tr A_i^2 , tr $A_iA_j(i < j)$, pf A_i and pf $A_iA_j(i < j)$. We can then determine tr $A_iA_j^t$ from (9), tr $A_iA_jA_k(1 < j < k)$ from (12), tr $A_i^tA_jA_k$, tr $A_iA_kA_k^t(i < j < k)$ from (10), and setting $A_k = A_j$ in (10) we get tr $A_iA_iA_i^t(i \neq j)$. This proves that the above traces and Pfaffians form a basis.

We will now prove that this basis is minimal. We will do this by giving examples of sets of matrices, $\{A_i\}$ and $\{B_i\}$, for which only one of the types of invariants differ. This will imply that this basis is also a minimal functional basis and a minimal complete set of invariants in the real case. The number of matrices in these examples is not significant. We can always add more matrices by setting $A_i = B_i = I_2$ or 0.

1. $A_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, B_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$

Here tr $A_1 \neq$ tr B_1 but the other invariants agree.

2.
$$A_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \ B_1 = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}.$$

Here tr $A_1^2 \neq$ tr B_1^2 but the other invariants agree.

3.
$$A_1 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \ A_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \ B_1 = \begin{pmatrix} 0 & 0 \\ -1 & 0 \end{pmatrix}, \ B_2 = A_2$$

Here tr $A_1A_2 \neq$ tr B_1B_2 but the other invariants agree.

4. If we pick two non-symmetric matrices which are conjugate in O(2) but not in SO(2), we have that pf $A_1 \neq pf B_1$ but the other invariants agree. As an example, take

$$A_1 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \ B_1 = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}.$$

5. In general $pf A_1 A_2 \neq pf A_2 A_1$, and if A_2 is invertible, we can set $B_1 = A_2 A_1 A_2^{-1}$, $B_2 = A_2$. Then $pf B_1 B_2 = pf A_2 A_1 A_2^{-1} A_2 = pf A_2 A_1 \neq pf A_1 A_2$ but the other invariants agree. As an example, take

$$A_1 = \begin{pmatrix} 0 & 1 \\ 2 & 0 \end{pmatrix}, \ A_2 = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}.$$

This completes the proof of part 2 of Theorem 3.

We would like to make some additional comments. Since SO (2) is a smaller group we get more invariants, but we can find a basis with a smaller number of invariants. The reason for this is simply that the Pfaffians give us more invariants of degree one and two, which simplifies the theory considerably. In particular, we see that none of the invariants in the basis for the SO (2) invariants involves more than 2 matrices, while in the orthogonal case we need the invariants tr $A_i A_j A_k (i < j < k)$. Hence the study of $C[mM(2, C)]^{SO(2, C)}$ (or $R[mM(2, R)]^{SO(2, R)}$) reduces to the study of $C[2M(2, C)]^{SO(2, C)}$ (or $R[2M(2, R)]^{SO(2, R)}$).

Let us try to explain the reason for this difference. We were able to delete tr $A_i A_j A_k$ because of (12), but there is no similar formula expressing tr XYZ in terms of traces of one or two factors, as the following example from [8] shows:

$$A_{1} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \ A_{2} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \ A_{3} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$
$$B_{1} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \ B_{2} = A_{2}, \ B_{3} = A_{3}.$$

Here tr $A_1A_2A_3 \neq$ tr $B_1B_2B_3$, but all traces involving one or two factors agree. If we take the trace in (8) everything cancels, and we must first multiply by $U \neq I$ to see that the trace of four 2 × 2 matrices is expressible in terms of traces of three factors or less. The closest we can come to expressing tr XYZ in terms of traces of one or two factors is the following equation from [8], which shows that tr XYZ satisfies a quadratic equation with coefficients expressible in terms of traces of one or two factors.

13)

$$4(\operatorname{tr} X Y Z)^{2} - 4a_{1} \operatorname{tr} X Y Z + a_{2} = 0,$$

$$a_{1} = \{\operatorname{tr} X \operatorname{tr} Y Z\} - \operatorname{tr} X \operatorname{tr} Y \operatorname{tr} Z,$$

$$a_{2} = 2 \{\operatorname{tr} X Y (\operatorname{tr} X Y - \operatorname{tr} X \operatorname{tr} Y) (\operatorname{tr}^{2} Z - \operatorname{tr} Z^{2}) + 4 \operatorname{tr} X Y \operatorname{tr} Y Z \operatorname{tr} Z X + 2 \operatorname{tr} X^{2} \operatorname{tr} Y^{2} \operatorname{tr} Z^{2} + \operatorname{tr}^{2} X \operatorname{tr}^{2} Y \operatorname{tr}^{2} Z - \{\operatorname{tr} X^{2} \operatorname{tr} Y^{2} \operatorname{tr}^{2} Z\}.$$

Here $\{ \}$ denotes the sum of the terms obtained by cyclic permutation of X, Y and Z. The equation is stated without proof in [8], but it follows from clever manipulations of (7).

If we specialize to the case of m = 1, we get the 3 invariants tr A, tr A^2 , and pf A. Since the orbit space M(2, R)/SO(2) has dimension 3, this is the 'right' number of invariants.

If we consider m = 2, we get the 8 invariants tr A, tr A^2 , pf A, tr B, tr B^2 , pf B, tr AB and pf AB. The dimension of 2M(2, R)/SO(2) is 7, however, so we have too many invariants. It is important to bear in mind that when we say that a complete set of invariants is minimal, we only mean that by deleting any of the invariants, the set will no longer be complete. This does not rule out the possibility that there could be a different minimal complete set of invariants with a smaller set of invariants. The question as to whether it is possible to find a complete set of invariants of 2M(2, R)/SO(2) consisting of 7 invariants in thus open.

If we add the restriction that $2 \operatorname{tr} A^2 + \operatorname{pf}^2 A - \operatorname{tr}^2 A \neq 0$, we can find such a set. We can use $\operatorname{tr} A$, $\operatorname{tr} A^2$, and $\operatorname{pf} A$ to determine A, and $\operatorname{tr} B$, $\operatorname{tr} AB$, and $\operatorname{pf} AB$ give a linear system of equations for the entries of B with determinant equal to $2 \operatorname{tr} A^2 + \operatorname{pf}^2 A - \operatorname{tr}^2 A$, so we can determine B without using $\operatorname{tr} B^2$.

REFERENCES

- J. Dubnov and V. Ivanov, Sur l'abaissement du degré des polynômes en affineurs, Dokl. Akad. Nauk SSSR 41 (1943), no. 3, 95–98.
- 2. G. B. Gurevich, Foundations of the Theory of Algebraic Invariants, P. Noordhoff, Groningen, 1964.
- 3. L. Le Bruyn, Trace rings of generic 2 by 2 matrices, Mem. Amer. Math. Soc. 66 (1987), no. 363.
- 4. C. Pearcy, A complete set of unitary invariants for operators generating finite W*-algebras of type I, Pacific J. Math. 12 (1962), 1405–1416.
- 5. C. Procesi, The invariant theory of $n \times n$ matrices, Adv. in Math. 19 (1976), 306-381.
- 6. C. Procesi, *Trace identities and standard diagrams*, in *Ring Theory* (Proceedings of the 1978 Antwerp Conference), ed. F. van Oystaeyen (Lecture notes in pure and applied mathematics 51), pp. 191–218, Marcel Dekker, New York and Basel, 1979.
- Ju. P. Razmyslov, Trace identities of full matrix algebras over a field of characteristic zero, Math. USSR-Izv. 8 (1974), no. 4, 727–760.
- 8. K. S. Sibirskii, Algebraic invariants for a set of matrices, Siberian Math. J. 9 (1968), no. 1, 115-124.
- K. S. Sibirskii, Invariants of linear representations of the group of rotations of the plane, and their applications to the qualitative theory of differential equations, Differential Equations 2 (1966), no. 6, 384–392; no. 7, 472–477.
- K. S. Sibirskii, Orthogonal invariants of the system of 2 × 2 matrices, Mat. Issled. 2 (1967), vyp. 4, 124–135 (Russian).
- 11. K. S. Sibirskii, Algebraic Invariants of Differential Equations and Matrices, Izdat. "Shtiintsa" Kishinev, 1976 (Russian).
- 12. K. S. Sibirskii, Introduction to the Algebraic Theory of Invariants of Differential Equations, Izdat. "Shtiintsa" Kishinev, 1982 (Russian); Manchester University Press, 1988 (English).

DEPARTMENT OF MATHEMATICS UNIVERSITY OF CALIFORNIA BERKELEY CA 94720 USA CURRENT ADDRESS: DEPARTMENT OF MATHEMATICS NATIONAL UNIVERSITY OF SINGAPORE SINGAPORE OS11 REPUBLIC OF SINGAPORE