SINGULARITIES
BANACH CENTER PUBLICATIONS, VOLUME 20
PWN POLISH SCIENTIFIC PUBLISHERS
WARSAW 1988

DEFINING ORBIT SPACES BY INEQUALITIES

CLAUDIO PROCESI

Department of Mathematics, Institute "Guido Castelnuovo", University of Roma, Roma, Italy

and

GERALD SCHWARZ*

Department of Mathematics, Brandeis University, Waltham, MA, U.S.A.

We discuss the authors' recent results [7] concerning the description of orbit spaces of representations of compact Lie groups.

0. Introduction

We begin by discussing two problems:

0.1. Let $p\colon \mathbf{R}^n \to \mathbf{R}^m$ be a polynomial mapping. Then the image $\operatorname{Im} p$ of p is a semialgebraic subset of $\mathbf{R}^m$. How can one find "simply" or "explicitly" the inequalities defining $\operatorname{Im} p$?

0.2. Let K be a compact Lie group and W a real representation space for K. Can one find a nice description of the orbit space W/K?

In Section 1 we will see that Problem 0.2 is a special case of Problem 0.1, and we describe the solution to 0.2. In Section 2 we present the details for the case of a finite group. In Section 3 we discuss some connections with Hilbert's 17th problem.

The contents of this paper approximate the talk given by the second named author at the Semester on Singularities held at the Stefan Banach Center in Warsaw, Spring 1985. He thanks the organizing committee for their invitation and hospitality.

---

* Partially supported by National Science Foundation Grant # MCS 8302575

1. Orbit spaces

1.0. Let W and K be as in 0.2. Then the graded algebra $\mathbf{R}[W]^K$ of K-invariant polynomial functions is finitely generated ([9], p. 274). Let $p_1, \ldots, p_m$ be homogeneous generators of $\mathbf{R}[W]^K$ and let $p = (p_1, \ldots, p_m)$ be the associated mapping from W to $\mathbf{R}^m$. Then p is proper and constant on the orbits of K, hence it induces a homeomorphism of W/K (quotient topology) with $X = \operatorname{Im} p \subseteq \mathbf{R}^m$ ([8]).

Let I denote the ideal of relations of the p_i in $\mathbf{R}[y_1, \ldots, y_m]$, and let Z denote the corresponding algebraic subset of $\mathbf{R}^m$. Then p induces an isomorphism $p^*\colon \mathbf{R}[Z] \to \mathbf{R}[W]^K$, and $X \subseteq Z$. Note that Z is determined by $\mathbf{R}[W]^K$ and our choice of generators, while to describe X we need some extra information.

1.1. Example. Let $K = \{\pm 1\}$ act by multiplication on $W = \mathbf{R}^2$. Then $\mathbf{R}[W]^K$ is generated by polynomials $p_1 = x^2+y^2$, $p_2 = x^2-y^2$ and $p_3 = 2xy$. Their ideal of relations is generated by the single polynomial $y_1^2-y_2^2-y_3^2$. Thus $Z = \{(y_1, y_2, y_3) \in \mathbf{R}^3\colon y_1^2 = y_2^2+y_3^2\}$. Since p_1 is non-negative, we must have that $X \subseteq \{(y_1, y_2, y_3) \in Z\colon y_1 \geqslant 0\}$. We will see that, in fact, there is equality.

1.2. Example. Let $W = \mathbf{R}^2$ and K the group of rotations by angles 0, $2\pi/3$, and $4\pi/3$. Then $\mathbf{R}[W]^K$ is generated by $p_1 = x^2+y^2$, $p_2 = x^3-3xy^2$ and $p_3 = y^3-3x^2y$. Their ideal of relations is generated by $y_1^3-y_2^2-y_3^2$, so $Z = \{(y_1, y_2, y_3)\colon y_1^3 = y_2^2+y_3^2\}$, and one can show that $X = \operatorname{Im} p = Z$ in this case.

1.3. Example. Let W be the space of $n \times q$ real matrices, and let $K = O(n) = O(n, \mathbf{R})$ act by left multiplication. Then W is just q copies of the standard representation of K on $\mathbf{R}^n$. By classical invariant theory, the K-invariants are generated by the inner products of the various copies of $\mathbf{R}^n$, i.e., of the columns of our $n \times q$ matrices. Thus we can define

$$p\colon W \to \operatorname{Sym}_q, \qquad A \mapsto A^t A,$$

where Sym_q denotes the space of real symmetric $q \times q$ matrices. It is an easy exercise to show that X is the set of all matrices $B \in \operatorname{Sym}_q$ such that:

(1.3.1) $$\operatorname{rank} B \leqslant n.$$

(1.3.2) B is positive semidefinite.

Then Z is defined by (1.3.1), i.e., by the condition that the determinants of all $(n+1) \times (n+1)$ minors of B are zero. The inequalities defining X come from the following:

1.4. *Remark.* Let C be a real symmetric matrix. Then C is positive semidefinite (we write $C \geqslant 0$) if and only if $C_\alpha \geqslant 0$ for all α, where $\{C_\alpha\}$ is the set of determinants of principal (i.e. symmetric) minors of C.

We now show how to find the inequalities describing X in general. The description was, essentially, conjectured by the physicists Abud and Sartori ([1], [2]): Let (,) denote a K-invariant inner product on W as well as the dual inner product on W^*. The differentials dp_i: $W \to W^*$ are K-equivariant, and the functions $w \mapsto (dp_i(w), dp_j(w))$ give an $m \times m$ symmetric matrix valued function $\mathrm{Grad}^\sim(w)$ with entries in $\boldsymbol{R}[W]^K$. There is a unique matrix valued function Grad on Z such that $\mathrm{Grad}^\sim(w) = \mathrm{Grad}(p(w))$ for all $w \in W$.

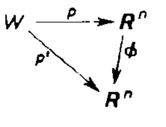
1.5. *Remark.* One may choose orthonormal co-ordinates $x_1, \ldots, x_n$ on W relative to (,). Then $\mathrm{Grad}^\sim(w) = J(w)J(w)^t$ where $J(w)$ is the Jacobian matrix of p at w. This shows that $\mathrm{Grad}^\sim(w) \geqslant 0$, or, in other words, $\mathrm{Grad}(x) \geqslant 0$ for all $x \in X$.

1.6. THEOREM. $X = \{z \in Z\colon \mathrm{Grad}(z) \geqslant 0\}$.

In Example 1.1, the theorem gives inequalities $y_1 \geqslant 0$; $y_1^2 - y_2^2 \geqslant 0$ and $y_1^2 - y_3^2 \geqslant 0$. But the last two inequalities are automatically satisfied on Z (since $y_1^2 = y_2^2 + y_3^2$), so $X = \{y \in Z\colon y_1 \geqslant 0\}$. In 1.2 one similarly gets that the inequality $y_1 \geqslant 0$ defines X, but this is already forced by the equality $y_1^3 = y_2^2 + y_3^2$, hence $X = Z$. (In [7] we show that $X = Z$ if and only if K (assumed acting effectively on W) is a finite group of odd order.) In 1.3 the theorem gives a redundant set of inequalities. One gets the condition in (1.3.2) exactly by applying a variant of Theorem 1.6 (see [7]).

1.7. Let $f(x) = x^n - b_1 x^{n-1} + \ldots + (-1)b_n$ be a real polynomial. When does f have only real roots? We use Theorem 1.6 to recover the classical criterion of Sylvester: Let $K = S_n$ denote the symmetric group which acts as usual on $W = \boldsymbol{R}^n$. Let $\sigma_1, \ldots, \sigma_n$ denote the elementary symmetric functions on the co-ordinates $x_1, \ldots, x_n$ of $\boldsymbol{R}^n$. Then $\boldsymbol{R}[W]^K = \boldsymbol{R}[\sigma_1, \ldots, \sigma_n]$. Let $p = (\sigma_1, \ldots, \sigma_n)$: $W \to \boldsymbol{R}^n$. Then f has real roots $a_1, \ldots, a_n$ if and only if $p(a) = b$, where $a = (a_1, \ldots, a_n)$ and $b = (b_1, \ldots, b_n)$. In other words, f has only real roots if and only if $b \in \operatorname{Im} p$.

We apply Theorem 1.6: Let $\tau_i = \sum_{j=1}^{n} x_j^i$, $i \geqslant 0$. Then $\tau_1, \ldots, \tau_n$ generate $\boldsymbol{R}[W]^K$. Let $p' = (\tau_1, \ldots, \tau_n)$: $W \to \boldsymbol{R}^n$. The classical Newton formulae: $\tau_1 = \sigma_1$, $\tau_2 = \sigma_1^2 - 2\sigma_2$, $\tau_3 = \sigma_1^3 - 3\sigma_1\sigma_3 + 3\sigma_3$, etc. give a polynomial isomorphism ϕ: $\boldsymbol{R}^n \to \boldsymbol{R}^n$ so that the following diagram commutes.



Thus $b \in \operatorname{Im} p$ if and only if $\phi(b) \in \operatorname{Im} p'$. Now $(d\tau_i, d\tau_j) = ij\tau_{i+j-2}$, and by Theorem 1.6 and our remarks above we see that $b \in \operatorname{Im} p$ if and only if $B(b) \geqslant 0$, where $B = (B_{ij})$ and $B_{ij}(\sigma_1, \ldots, \sigma_n) = ij\tau_{i+j-2}$. It does not affect positive semidefiniteness if we replace B_{ij} by $\frac{1}{ij} B_{ij}$ and in this way we arrive at the "Bezoutiant" matrix Bez of Sylvester. We have shown:

1.8. Corollary (Sylvester, see [6])). *Let $f(x) = x^n - b_1 x^{n-1} + \ldots$ $\ldots + (-1)^n b_n$ be a real polynomial. Then f has only real roots if and only if* $\operatorname{Bez}(b) \geqslant 0$.

For $n = 2$, one can compute that

$$\operatorname{Bez}(b) = \begin{bmatrix} 2 & b_1 \\ b_1 & b_1^2 - 2b_2 \end{bmatrix}$$

and for $n = 3$ one has

$$\operatorname{Bez}(b) = \begin{bmatrix} 3 & b_1 & b_1^2 - 2b_2 \\ b_1 & b_1^2 - 2b_2 & b_1^3 - 3b_1 b_2 + 3b_3 \\ b_1^2 - 2b_2 & b_1^3 - 3b_1 b_2 + 3b_3 & b_1^4 - 4b_1^2 b_2 + 2b_2^2 + 4b_1 b_3 \end{bmatrix}.$$

2. Finite groups

We give a proof of Theorem 1.6 in the case that K is finite: Let W, k, p and $X \subseteq Z \subseteq \boldsymbol{R}^m$ be as before. Recall that we have K-invariant inner products $(\ ,\ )$ on W and W^*, and $\widetilde{\operatorname{Grad}}(w) = (dp_i(w), dp_j(w))$. We have a point $z \in Z$ with the property that $\operatorname{Grad}(z) \geqslant 0$, and we want to show that $z \in X$.

Let $V = W \otimes_{\boldsymbol{R}} \boldsymbol{C}$. Our K-invariant inner products extend to K-invariant non-degenerate symmetric bilinear forms on V and V^*, denoted as usual by $(\ ,\ )$. We identify $\boldsymbol{R}[W]^K$ with the elements of $\boldsymbol{C}[V]^K$ which are real on W, and then $p_1, \ldots, p_m$ generate $\boldsymbol{C}[V]^K$. Our mapping $p\colon W \to \boldsymbol{R}^n$ extends to $p\colon V \to \boldsymbol{C}^n$, and the image $p(V)$ lies in the set of complex zeroes $Z_{\boldsymbol{C}}$ of the ideal of relations of the p_i (see 1.0).

2.1. Lemma. (1) $p(V) = Z_{\boldsymbol{C}}$.

(2) *The fibers of p are (set-theoretically) the orbits of K.*

Proof. Let $\varrho\colon \boldsymbol{C}[V] \to \boldsymbol{C}[V]^K$ be the Reynold's operator (averaging over the group). Let $z \in Z$, let I_z be the corresponding maximal ideal of $\boldsymbol{C}[V]^K$ and set $J_z = I_z \boldsymbol{C}[V]$. Then $\varrho(J_z) = I_z$, so J_z is a proper ideal of $\boldsymbol{C}[V]$ and $p(x) = z$ for any zero x of J_z. Thus (1) holds, and another averaging over the group argument shows that $\boldsymbol{C}[V]^K$ separates distinct K-orbits, proving (2). □

It follows from Lemma 2.1(1) that:

2.2. *There is a point $v \in V$ such that $p(v) = z$.*

Write $v = w_1 + iw_2$ where $w_1, w_2 \in W$. Then $\bar{v} = w_1 - iw_2$. Assume the following:

2.3. PROPOSITION. *Define $\lambda \in V^*$ by $\lambda(x) = (x, iw_2)$, $x \in V$. Then there are $a_1, \ldots, a_m \in \boldsymbol{R}$ such that $\lambda = \sum a_i \, dp_i(v)$.*

Proof of Theorem 1.6. Consider the value of (λ, λ) where λ is as above. On the one hand

$$(\lambda, \lambda) = \left(\sum a_i \, dp_i(v), \sum a_j \, dp_j(v)\right) = \sum a_i a_j \operatorname{Grad}(p(v))_{ij} \geqslant 0$$

since $p(v) = z$ and $\operatorname{Grad}(z) \geqslant 0$. On the other hand

$$(\lambda, \lambda) = (iw_2, iw_2) = -(w_2, w_2) \leqslant 0.$$

Hence $(w_2, w_2) = 0$ and $v = w_1 \in W$. Hence $z = p(w_1) \in X$. □

We now establish Proposition 2.3: Since p is real on W, it follows that $p(\bar{v}) = \overline{p(v)} = z$. Hence, by Lemma 2.1(2).

2.4. *There is a $k_0 \in K$ such that $k_0 v = \bar{v}$.*
Set

$$\Delta(v) = (V^*)^{K_v},$$

where K_v is the isotropy group of K at v, and set

$$D(v) = \{df(v) \colon f \in \boldsymbol{C}[V]^K\}.$$

2.5. *Remarks.* (1) If $f \in \boldsymbol{C}[V]^K$, then

$$df(v) = d(f \circ k)(v) = df(kv) \circ k$$

for all $k \in K$, hence $df(v)$ is K_v-invariant. Thus $D(v) \subseteq \Delta(v)$.

(2) Since the p_i generate $\boldsymbol{C}[V]^K$, the complex span of the $dp_i(v)$ is $D(v)$.

2.6. LEMMA. $D(v) = \Delta(v)$.

We establish Lemma 2.6 below. Now set

$$\Delta_{\boldsymbol{R}}(v) = \{\mu \in \Delta(v) \colon \mu \circ k_0 = \bar{\mu}\},$$

where $\bar{\mu}(x) = \overline{\mu(\bar{x})}$. Note that each $dp_i(v)$ is $\Delta_{\boldsymbol{R}}(v)$, since

$$dp_i(v) \circ k_0 = dp_i(k_0^{-1} v) = dp_i(\bar{v}) = \overline{dp_i(v)}.$$

Now by Lemma 2.6, each μ in $\Delta(v)$ is a sum $\sum a_i \, dp_i(v)$, and if $\mu \in \Delta_{\boldsymbol{R}}(v)$ one easily sees, using our computation above, that one may assume that the a_i are real. Hence

2.7. *$\Delta_{\mathbf{R}}(v)$ is the real span of the $dp_i(v)$.*

Proof of Proposition 2.3. Let $f(x) = \frac{1}{2}(x, x)$, $x \in V$. Then f is real on W and $\lambda_1 := df(v) \in \Delta_{\mathbf{R}}(v)$, where $\lambda_1(x) = (x, v)$. Define $\lambda_2 \in V^*$ by $\lambda_2(x) = (x, \bar{v})$. Using 2.4 and the fact that $K_v = K_{\bar{v}} = K_{w_1} \cap K_{w_2}$, one easily establishes that $\lambda_2 \in \Delta_{\mathbf{R}}(v)$. Hence $\lambda = \frac{1}{2}(\lambda_1 - \lambda_2) \in \Delta_{\mathbf{R}}(v)$, where $\lambda(x) = (x, iw_2)$. □

Proof of Lemma 2.6. Let B_v be a small ball containing v so that, for any $k \in K$, either $kB_v = B_v$ or $B_v \cap kB_v = \emptyset$. Let $\mathscr{H}(U)$ denote the holomorphic functions on U, for U an open subset of V. Then $\mathscr{H}(KB_v)^K \simeq \mathscr{H}(B_v)^{K_v}$. If $\mu \in (V^*)^{K_v}$, then the function $f(x) := \mu(x-v)$, $x \in B_v$, lies in $\mathscr{H}(B_v)^{K_v}$ and has differential μ at v. Now $\mathbf{C}[V]$ is dense in $\mathscr{H}(B_v)$, hence $\mathbf{C}[V]^K$ is dense in $\mathscr{H}(KB_v)^K$, and it follows that $\Delta(v) = D(v)$. □

2.8. *Remark.* In case K is not finite, one has to consider the action of the complexification $K_{\mathbf{C}}$ of K on V. Not all orbits of $K_{\mathbf{C}}$ are closed, which presents complications. The new ingredients needed for the proof of Theorem 1.6 are Luna's slice theorem [5] (to prove the appropriate analogue of Lemma 2.6) and some results of Kempf and Ness [4] (to establish 2.4).

3. Hilbert's seventeenth problem

We give some applications of Theorem 1.6 to a version of Hilbert's 17th problem: The solution to Hilbert's 17th reads as follows:

3.1. THEOREM. *Let $f \in \mathbf{R}(x_1, \ldots, x_n)$ be positive,* i.e., *f is non-negative wherever it is defined. Then there are $g_1, \ldots, g_d \in \mathbf{R}(x_1, \ldots, x_n)$ such that $f = g_1^2 + \ldots + g_d^2$.*

Let K and W be as in Introduction. Does Theorem 3.1 remain true if we replace $\mathbf{R}(W)$ by $\mathbf{R}(W)^K$? The answer is:

3.2. EXAMPLE. Let $K = \{\pm 1\}$ act by multiplication on $W = \mathbf{R}$. Then $\mathbf{R}(W)^K$ consists of rational functions of x^2, and $f(x) = x^2$ is positive. If $f(x) = g_1(x^2)^2 + \ldots + g_d(x^2)^2$, then $x = g_1(x)^2 + \ldots + g_d(x)^2$, a contradiction. However, one can show that f is, in some sense, the only problem. In other words, if $g(x) \in \mathbf{R}(W)^K$ is positive, then

$$g(x) = g_0(x^2) + g_1(x^2) f(x),$$

where g_0 and g_1 are sums of squares.

3.3. Let F be a subfield of $\mathbf{R}(x_1, \ldots, x_n)$. We say that F *has property* (H) if there are positive elements $h_1, \ldots, h_q \in F$ such that every $f \in F$ which is positive can be written in the form

$$f = \sum s_i h_i \tag{3.3.1}$$

where the s_i are sums of squares in F.

3.4. THEOREM. *Let K and W be as in* 0.2. *Then $\boldsymbol{R}(W)^K$ has property* (H).

Procesi [6] established Theorem 3.4 and found the polynomials h_i of 3.3 in case $K = S_n$ acting standardly on $\boldsymbol{R}^n$. Bochnak and Efroymson [3] first conjectured Theorem 3.4. We now show how to obtain Theorem 3.4 from Theorem 1.6.

Let P be a closed semialgebraic subset of $\boldsymbol{R}^m$. We assume that the Zariski closure T of P is irreducible. We say that P is *elementary* if there are $f_1, \ldots, f_d \in \boldsymbol{R}[T]$ such that $P = \{t \in T: f_i(t) \geqslant 0, i = 1, \ldots, d\}$. We say that P is *quasi-elementary* if there is an algebraic subset Y of T such that $\dim Y < \dim T$ and $P \cup Y$ is elementary.

3.5. PROPOSITION ([3]). *Let P and T be as above.*

(1) *If P is quasi-elementary, choose $f_1, \ldots, f_d \in \boldsymbol{R}[T]$ so that $\{t \in T: f_i(t) \geqslant 0, i = 1, \ldots, d\} = P \cup Y$, where Y is algebraic and* $\dim Y < \dim T$. *Then every $f \in \boldsymbol{R}(T)$ which is positive on P can be written in the form* (3.3.1), *where the h_i are all possible products $f_{i_1} \ldots f_{i_r}$, $1 \leqslant i_1 < \ldots < i_r \leqslant d$, $0 \leqslant r \leqslant d$.*

(2) *If every $f \in \boldsymbol{R}(T)$ which is positive on P can be written in the form* (3.3.1) *for some h_i, then P is quasi-elementary.*

Proof of Theorem 3.4. Let X, Z and Grad be as in Theorem 1.6. Then one can show that Z is irreducible, that Z is the Zariski closure of X and that p^* induces an isomorphism of $\boldsymbol{R}(Z)$ with $\boldsymbol{R}(W)^K$. Let $f_1, \ldots, f_d$ be the determinants of the principal minors of Grad. Then $X = \{z \in Z: f_i(z) \geqslant 0, i = 1, \ldots, d\}$. Hence X is elementary, and Theorem 3.4 follows. □

If we drop the assumption that K is compact, then $\boldsymbol{R}(W)^K$ may fail to have property (H) ([7]). The problem is that the corresponding orbit space $X \subseteq Z$ may fail to be quasi-elementary!

References

[1] M. Abud and G. Sartori, *The geometry of orbit-space and natural minima of Higgs potentials*, Phys. Lett. B 104 (1981), 147–152.

[2] —, —, *The geometry of spontaneous symmetry breaking*, Ann. Physics 150 (1983), 307–372.

[3] J. Bochnak and G. Efroymson, *Real algebraic geometry and the 17th Hilbert problem*, Math. Ann. 251 (1980), 213–241.

[4] G. Kempf and L. Ness, *The length of vectors in representation spaces*, in *Algebraic Geometry*, Lecture Notes in Math. 732, Springer-Verlag, New York 1979, 233–243.

[5] D. Luna, *Slices étales*, Bull. Soc. Math. France 33 (1973), 81–105.

[6] C. Procesi, *Positive symmetric functions*, Adv. in Math. 29 (1978), 219–225.

[7] C. Procesi and G. Schwarz, *Inequalities defining orbit spaces*, Invent. Math. 81 (1985), 539–554.

[8] G. Schwarz, *Smooth functions invariant under the action of a compact Lie group*, Topology 14 (1975), 63–68.

[9] H. Weyl, *The Classical Groups*, 2nd ed., Princeton University Press, Princeton 1946.

Presented to the semester
Singularities
15 February–15 June, 1985