

TOPOLOGICAL EQUIVALENCE OF REAL BINARY FORMS

DAVID WEINBERG AND DAVE WITTE

(Communicated by Frederick R. Cohen)

ABSTRACT. Necessary and sufficient conditions are given for two real binary homogeneous polynomials to be equivalent under a continuous change of variables, and for the forms to be equivalent under a continuous change of variables that is differentiable (or that is real-analytic) away from the origin.

1. INTRODUCTION

Consider the natural action of $GL(2, \mathbb{R})$ on the space T_n of binary homogeneous polynomials of degree n , induced by the action of $GL(2, \mathbb{R})$ on \mathbb{R}^2 via the variables of the polynomial. (The natural left action of $GL(2, \mathbb{R})$ on the variables induces a right action of $GL(2, \mathbb{R})$ on T_n as follows: if $f \in T_n$, $h \in GL(2, \mathbb{R})$, and $v = \begin{pmatrix} x \\ y \end{pmatrix}$, then $(fh)v = f(hv)$. What this means is that we change variables by h , expand the results, and collect like terms to get a new binary form.) This representation was the subject of intense investigation in the second half of the nineteenth century [GY, E]. The list of orbits in the case $n = 2$ is the simplest case of Sylvester's Law of Inertia [MB, pp. 382–390]. A set of representatives of the orbits in the case $n = 3$ is as follows: x^3 , x^2y , $x(x^2 - y^2)$, $x(x^2 + y^2)$ [G, pp. 261–271], [W]. In the case $n = 4$, the number of orbits is infinite. For a complete enumeration of these orbits see [G, pp. 285–294, 409].

A natural classification of binary homogeneous polynomials can be based on the nature of the roots (if $p(x, y) \in T_n$, let $y = 1$, and then consider the roots). For example, this divides the quartic forms into classes with the following representatives: $\pm x^4$, x^3y , $\pm x^2y^2$, $x^2(x^2 - y^2)$, $\pm x^2(x^2 + y^2)$, $xy(x^2 - y^2)$, $xy(x^2 + y^2)$, $\pm(x^2 + y^2)^2$, $(x^2 + y^2)(x^2 + 4y^2)$. The purpose of this paper is to attempt to furnish some sort of group-theoretical foundation for this classification.

Notation. Let H denote the group of homeomorphisms of the plane.

Received by the editors April 10, 1989 and, in revised form, October 14, 1990.

1980 *Mathematics Subject Classification* (1985 Revision). Primary 57M99, 58C99; Secondary 12D05, 14A25, 14N99, 51H99, 51N25.

Definition. Let $f(v), g(v) \in T_n$, where $v = (x, y)$. The group H determines an equivalence relation on T_n as follows: f is defined to be equivalent to g provided $\exists h \in H$ such that $g(v) = f(h(v))$.

Remark. The traditional study of binary forms is based on the group action of $GL(2, \mathbb{R})$ on T_n , which is described in the first sentence of this paper. The group H does not act on T_n because polynomials are obviously not preserved by an arbitrary continuous change of variables. However, the equivalence relation specified by the definition is well defined.

It turns out that this group-based equivalence relation is coarser than the above-mentioned classification by the nature of the roots. The group-based equivalence relation completely ignores complex roots. Thus, for example, the quartic forms $(x^2 + y^2)^2$ and $(x^2 + y^2)(x^2 + 4y^2)$ are equivalent under H .

Furthermore, the group-based equivalence relation does not notice that two real roots have different multiplicity unless the multiplicities are of opposite parity. Thus, for example, the quartic forms x^3y and $xy(x^2 + y^2)$ are equivalent under H . In order to get a classification that exhibits multiplicities of the real roots, we replace H by the subgroup of homeomorphisms of \mathbb{R}^2 that are differentiable (or even real-analytic) away from the origin. Precise statements can be found in the next section.

The ideas behind the theorems can be briefly summarized. Let $f(x, y) \in T_n$. The zero set of the function $z = f(x, y)$ is a set of straight lines through the origin. A continuous change of variables that is continuously invertible must preserve the number of straight lines in the set. The function f must have constant sign in each sector between these lines. A continuous change of variables must preserve or reverse this sequence of signs up to a rotation. In order to distinguish root multiplicities, we take transformations that preserve the order of vanishing: for this, homeomorphisms of \mathbb{R}^2 that are differentiable away from the origin can be used, or homeomorphisms of \mathbb{R}^2 that are real analytic away from the origin can be used. The details will be provided in the next section.

2. THEOREMS AND PROOFS

Binary homogeneous polynomials are our main interest, but it is convenient to prove our results for more general functions that are continuous and homogeneous. When we need to discuss the order of vanishing, we assume the function is real-analytic. If the order of vanishing is not required, we assume a much weaker property, monotonicity on each side of the zeros.

Definition. Let f be a continuous, real-valued function on the plane. (In what follows, (x, y) will denote rectangular coordinates and (r, θ) will denote polar coordinates in the plane.) We say that f is *homogeneous* if there is some $k \in \mathbb{R}^+$ with $f(cr, \theta) = c^k \cdot f(r, \theta)$ for all $c \in \mathbb{R}^+$.

Definition. Suppose f is a homogeneous, real-analytic function on the plane. Assume first that $f(1, \theta)$ is not identically zero, but has at least one zero. Let $0 \leq \theta_1 < \theta_2 < \dots < \theta_n < 2\pi$ be the zeros of $f(1, \theta)$ in $[0, 2\pi)$.

- Let d_i be the degree (i.e., the order of vanishing) of $f(1, \theta)$ at θ_i ; and
- let ε_i be the symbol $+$ or $-$ depending on whether the sign of $f(1, \theta)$ is positive or negative in the interval (θ_i, θ_{i+1}) .

The *real-analytic zero-set data* for f is the sequence $d_1, \varepsilon_1, d_2, \varepsilon_2, \dots, d_n, \varepsilon_n$. If $f(1, \theta)$ is identically 0, or has no zeros, then we define the *real-analytic zero-set data* for f to be a single symbol 0, $+$, or $-$, as appropriate.

Two homogeneous, real-analytic functions are considered to have the *same* real-analytic zero-set data if the sequence for one of the functions can be obtained by a cyclic permutation and/or a reversal of the sequence for the other.

Remark. If d_i is even, then $\varepsilon_i = \varepsilon_{i-1}$; if d_i is odd, then $\varepsilon_i = -\varepsilon_{i-1}$.

Theorem 2.1. *Suppose f and g are homogeneous, real-analytic functions on the plane. Then the following are equivalent:*

- (1) f and g have the same real-analytic zero-set data;
- (2) there is a homeomorphism h of \mathbb{R}^2 with $f \circ h = g$, such that both h and h^{-1} are real analytic except perhaps at the origin;
- (3) there is a homeomorphism h of \mathbb{R}^2 with $f \circ h = g$, such that both h and h^{-1} are differentiable except perhaps at the origin.

Proof. (2 \Rightarrow 3) Any real-analytic function is differentiable.

(3 \Rightarrow 1) This argument is described in the last paragraph of § 1. Note that a diffeomorphism must preserve the order of vanishing of a real-analytic function (see Lemma 2.3).

(1 \Rightarrow 2) There is no harm in assuming that f and g are not identically zero, so $f(1, \theta)$ and $g(1, \theta)$ have only finitely many zeros. By composing g with a homeomorphism in the θ -variable (real-analytic except at the origin), we may assume $g(1, \theta)$ and $f(1, \theta)$ have the same zeros, have the same degree at each of these zeros, and have the same sign between the zeros (because f and g have the same real-analytic zero-set data). Then $q(\theta) = g(1, \theta)/f(1, \theta)$ is real-analytic in θ , and has no zeros. By composing each of f and g with a homeomorphism (of the form $r^{1/k}$) in the r -variable, we may assume f and g are homogeneous of degree 1. Let $h(r, \theta) = (r \cdot q(\theta), \theta)$. Because q is continuous and nowhere zero, h is a homeomorphism. Because q and $1/q$ are real-analytic, h is a real-analytic diffeomorphism except at the origin. From the definitions of h and q , and the homogeneity of f and g , we have

$$\begin{aligned} f \circ h(r, \theta) &= f(r \cdot q(\theta), \theta) = r \cdot q(\theta) \cdot f(1, \theta) \\ &= r \cdot \frac{g(1, \theta)}{f(1, \theta)} \cdot f(1, \theta) = r \cdot g(1, \theta) = g(r, \theta). \end{aligned}$$

This is what we wanted. \square

Corollary 2.2. *Let f and g be binary homogeneous polynomials. Then the following are equivalent:*

- (1) f and g have the same real-analytic zero-set data;
- (2) there is a homeomorphism h of \mathbb{R}^2 with $f \circ h = g$, such that both h and h^{-1} are real analytic except perhaps at the origin;
- (3) there is a homeomorphism h of \mathbb{R}^2 with $f \circ h = g$, such that both h and h^{-1} are differentiable except perhaps at the origin.

Lemma 2.3. *Suppose f and g are homogeneous, real-analytic functions on the plane. Assume $\theta = 0$ is an isolated zero of both $f(1, \theta)$ and $g(1, \theta)$. If there is a homeomorphism h of \mathbb{R}^2 with $f \circ h = g$, such that both h and h^{-1} are differentiable except perhaps at the origin, and such that $h(1, 0) = (r_0, 0)$ for some $r_0 \in \mathbb{R}^+$, then $f(1, \theta)$ and $g(1, \theta)$ have the same order of vanishing at $\theta = 0$.*

Proof. Since h is differentiable, and $h(1, 0) = (r_0, 0)$, we have

$$h(1, \theta) = (r_0 + a\theta + o(\theta), b\theta + o(\theta)) \quad \text{as } \theta \rightarrow 0,$$

for some $a, b \in \mathbb{R}$. Because h must map the zero set of f to the zero set of g , we know that h maps the positive x -axis onto itself. Since h is a diffeomorphism, then h must map the unit circle $\{(1, \theta)\}$ to a curve transverse to the positive x -axis, and $b \neq 0$.

Let d be the order of vanishing of $f(1, \theta)$. Then, since $f(r, \theta)$ is homogeneous, we have

$$f(r, \theta) = r^k \cdot f(1, \theta) = r^k (c_1 \theta^d + o(\theta^d)) \quad \text{as } \theta \rightarrow 0,$$

for some nonzero $c_1 \in \mathbb{R}$. If (r, θ) approaches $(r_0, 0)$ along a curve transverse to the positive x -axis, we have $r = r_0 + O(\theta)$. Then $r^k = r_0^k + O(\theta)$, so

$$f(r, \theta) = (r_0^k + O(\theta))(c_1 \theta^d + o(\theta^d)) = c_2 \theta^d + o(\theta^d),$$

for some nonzero $c_2 \in \mathbb{R}$. Then, for $\theta \rightarrow 0$, we have

$$\begin{aligned} g(1, \theta) &= f(h(1, \theta)) = f(r_0 + a\theta + o(\theta), b\theta + o(\theta)) \\ &= c_2 (b\theta)^d + o((b\theta)^d) = c_3 \theta^d + o(\theta^d). \end{aligned}$$

Therefore, the order of vanishing of $g(1, \theta)$ at $\theta = 0$ is d . \square

Definition. Let f be a continuous, homogeneous, real-valued function on the plane. We say that f is *strictly monotonic on each side of its zeros* if, for each zero θ_0 of $f(1, \theta)$, there is an open interval (a, b) containing θ_0 such that $f(1, \theta)$ is either strictly monotonic or identically 0 on each of the two subintervals (a, θ_0) and (θ_0, b) .

Remark. Let f be a homogeneous, real-valued function on the plane. If f is real-analytic, then f is strictly monotonic on each side of its zeros.

Definition. Suppose f and g are two continuous, homogeneous functions on the plane. We say that f and g have the same topological zero-set behavior if there is a homeomorphism h in the θ -variable such that $f \circ h(1, \theta)$ has the same sign (+, -, or 0) as $g(1, \theta)$ for every $\theta \in [0, 2\pi)$.

If $f(1, \theta)$ and $g(1, \theta)$ have only finitely many zeros, we can describe this much more concretely. Let $0 \leq \theta_1 < \theta_2 < \dots < \theta_n < 2\pi$ be the zeros of $f(1, \theta)$ in $[0, 2\pi)$. Let ε_i be the symbol + or - depending on whether the sign of $f(1, \theta)$ is positive or negative in the interval (θ_i, θ_{i+1}) . The topological zero-set data for f is the sequence $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$. Similarly, we define the topological zero-set data for g , and then f and g have the same topological zero-set behavior if the sequence for one of the functions can be obtained by a cyclic permutation and/or a reversal of the sequence for the other.

Theorem 2.4. Suppose f and g are two continuous, homogeneous functions on the plane. If each of f and g is strictly monotonic on each side of its zeros, then the following are equivalent:

- (1) f and g have the same topological zero-set behavior;
- (2) there is a homeomorphism h of \mathbb{R}^2 with $f \circ h = g$.

Proof. (2 \Rightarrow 1) This simple argument is described in the last paragraph of §1.

(1 \Rightarrow 2) By composing g with a homeomorphism in the θ -variable, we may assume $g(1, \theta)$ and $f(1, \theta)$ have the same zeros, and have the same sign between the zeros (because f and g have the same topological zero-set behavior). By composing each of f and g with a homeomorphism (of the form $r^{1/k}$) in the r -variable, we may assume f and g are homogeneous of degree 1. Because each of f and g is strictly monotonic on each side of its zeros, Lemma 2.5 shows that, by composing f and g with a homeomorphism in the θ variable, we may assume

$$f(1, \theta) = g(1, \theta) = \begin{cases} \pm(\theta - \theta_0) & \text{or} \\ 0, \end{cases}$$

whenever θ is near a zero θ_0 of $f(1, \theta)$. Let $q(\theta) = g(1, \theta)/f(1, \theta)$, setting $q = 1$ at the zeros of $f(1, \theta)$. Then q is obviously continuous away from the zeros of $f(1, \theta)$. It is also continuous at the zeros, because $f(1, \theta) = g(1, \theta)$ whenever θ is near one of these zeros.

Let $h(r, \theta) = (r \cdot q(\theta), \theta)$. Because q is continuous and nowhere zero, h is a homeomorphism. We have

$$\begin{aligned} f \circ h(r, \theta) &= f(r \cdot q(\theta), \theta) = r \cdot q(\theta) \cdot f(1, \theta) \\ &= r \cdot \frac{g(1, \theta)}{f(1, \theta)} \cdot f(1, \theta) = r \cdot g(1, \theta) = g(r, \theta). \end{aligned}$$

This is what we wanted. \square

Lemma 2.5. Let f be a strictly monotonic, continuous function on an interval $[a, b]$, with $f(a) = 0$. Then there is a homeomorphism h of $[a, b]$, fixing the endpoints, and some $\varepsilon > 0$ such that $f(h(x)) = \pm(x - a)$ for $a < x < a + \varepsilon$. \square

Corollary 2.6. *Let f and g be binary, homogeneous polynomials. Then the following are equivalent:*

- (1) *f and g have the same topological zero-set behavior;*
- (2) *there is a homeomorphism h of \mathbb{R}^2 with $f \circ h = g$.*

Remark. In Theorem 2.4, it is not necessary to assume that the monotonicity of f and g is strict. By homogeneity, if f is monotonic on each side of its zeros, we could compose with a homeomorphism in the r -variable to transform f to a function that is strictly monotonic on each side of its zeros. The monotonicity condition could be weakened even further, but it cannot be entirely eliminated. For example, if

$$f(1, \theta) = \theta^2 + \theta \left(1 + \sin \frac{1}{\theta} \right) \quad \text{for } \theta \text{ near } 0,$$

then $\theta = 0$ is an isolated zero of $f(1, \theta)$, but $f(1, \theta)$ oscillates so badly as θ approaches 0 that f is not equivalent by a homeomorphism to any homogeneous function that is monotonic on each side of its zeros.

Question. As we pointed out in the remark following the definition of equivalence in the space T_n , the group H does not give a group action on the space T_n . However, one can ask the following interesting question. Can the equivalence classes determined in the two theorems be realized in a natural way as the orbits of some group action, or even as the orbits of some group representation?

ACKNOWLEDGMENT

We wish to thank the referee for comments that improved and clarified the results of this paper. Most of this work was done while we were at the Mathematical Sciences Research Institute in Berkeley, CA. We gratefully acknowledge the hospitality of MSRI.

REFERENCES

- [E] E. B. Elliott, *An introduction to the algebra of quantics*, 2nd ed., Oxford Univ. Press, Oxford, 1913.
- [G] G. Gurevich, *Foundations of the theory of algebraic invariants*, Noordhoff, Groningen, 1964.
- [GY] J. H. Grace and A. Young, *The algebra of invariants*, Cambridge Univ. Press, Cambridge, 1903.
- [MB] S. Mac Lane and G. Birkhoff, *Algebra*, MacMillan, London, 1967.
- [W] D. Weinberg, *Canonical forms for symmetric tensors*, *Linear Algebra Appl.* **57** (1984), 271–282.

DEPARTMENT OF MATHEMATICS, TEXAS TECH UNIVERSITY, LUBBOCK, TEXAS 79409

DEPARTMENT OF MATHEMATICS, WILLIAMS COLLEGE, WILLIAMSTOWN, MASSACHUSETTS 01267
E-mail address: dwitte@williams.edu