1983

A Polynomial-Time Algorithm for the Topological Type of a Real Algebraic Curve

Dennis S. Arnon

Scott McCallum

Report Number:
83-454

1

A POLYNOMIAL-TIME ALGORITHM FOR THE TOPOLOGICAL TYPE OF A REAL ALGEBRAIC CURVE

by

Dennis S. Arnon
Computer Science Department
Purdue University
West Lafayette, Indiana 47907

and

Scott McCallum
Computer Science Department
University of Wisconsin
Madison, Wisconsin 53706

CSD TR-454
Department of Computer Sciences
Purdue University
West Lafayette, Indiana 47907
August 31, 1983

ABSTRACT

Let $f(x,y,z)$ be a homogeneous polynomial with rational coefficients. Let C_f be the real projective curve defined by $f = 0$, and suppose that C_f is nonsingular. It is well known that C_f is essentially a finite collection of disjoint circles, all except possibly one of which lie in the projective plane $\mathbb{R}P^2$ in such a way as to have both an interior (homeomorphic to a disk), and an exterior (homeomorphic to a Möbius strip). These two-sided components of C_f are called *ovals*. The partial order imposed on its ovals by the relation of inclusion specifies the topological type of C_f. We present an algorithm which, given f, determines the ordering of the ovals of C_f. The algorithm constructs a cell complex for $\mathbb{R}P^2$, such that for each oval O of C_f, the closure of each component of complement(O) is a subcomplex. The Euler characteristic χ of a complex is easily computed, and since χ(disk)$\neq\chi$(Möbius strip), any cell can be classified as being inside, on, or outside a particular oval. This essentially determines the ordering of ovals. The maximum computing time of our algorithm is dominated by a polynomial function of the degree of f and the size of its coefficients.

Keywords: polynomial zeros, computer algebra, computational geometry, semi-algebraic geometry, decision procedures, real algebraic geometry, Hilbert's Sixteenth Problem, cylindrical algebraic decomposition.

The second author would like to acknowledge the support of the National Science Foundation, Grant MCS-8009357.

1. Introduction.

Let $f(x,y,z)$ be a homogeneous polynomial with rational coefficients. Let C_f be the real projective curve defined by $f = 0$. It is well known [Wil78a] that if C_f is nonsingular, then it is a compact one-dimensional manifold, and so homeomorphic to a disjoint union of circles. A circle can have either a one-sided or two-sided imbedding in $\mathbb{R}P^2$ (see Section 4); in the latter case it has both an interior (homeomorphic to a disk), and an exterior (homeomorphic to a Möbius strip). The two-sided components of C_f are called *ovals*. If f has even degree, then every component of C_f is an oval; if degree(f) is odd, every component except one is an oval.

Curves C_1 and C_2 have the same *topological type* if there is a homeomorphism $\varphi:\mathbb{R}P^2 \to \mathbb{R}P^2$ which maps C_1 onto C_2. Each oval of a nonsingular curve C_f is either inside or outside any other; the partial ordering of the ovals induced by this inclusion relation, together with the parity of the degree of f, determine the topological type of the curve.

In this paper we present an algorithm which, given $f(x,y,z)$ with rational coefficients, determines whether C_f is nonsingular, and if so, determines the ordering of its ovals. The main step of the algorithm is construction of a cellular decomposition D_f of $\mathbb{R}P^2$ such that every component of C_f is a union of cells of D_f. The following description of D_f is produced: (1) a list of the pairs of adjacent cells (two cells are *adjacent* if their union is connected), and (2) a list of the cells contained in C_f. In the course of constructing D_f we determine if C_f has singularities (see Section 3), and if so, halt.

Assuming C_f is nonsingular, the rest of the algorithm is straightforward. The reflexive transitive closure $\overline{R}$ of the adjacency relation is an equivalence relation; for a subset X of $\mathbb{R}P^2$, let $\overline{R}(X)$ denote its restriction to the cells of D_f which meet X. We construct the equivalence classes of $\overline{R}(C_f)$; (the union of) each class is a component of C_f. Let O be one of these components and K the corresponding class of $\overline{R}(C_f)$. We construct the equivalence classes of $\overline{R}$(complement(O)); (the union of) each class is a component of complement(O). O is an oval if and only if there are two such classes; if there is only one, we do not process O further. Suppose there are two classes K_1 and K_2. Let V be the union of K_1 and let W be the union of K_2. We want to determine which of V and W is the interior (Int(O)) and which is the exterior (Ext(O)) of O. We prove in Section 5 that D_f gives $\mathbb{R}P^2$ the structure of a finite cell complex (see e.g [Gra75a, Hil3,a], or [Mas67a]). Theorem 4.1 of Section 4 establishes that $O \bigcup V = \overline{V}$ and $O \bigcup W = \overline{W}$, hence $K \bigcup K_1$ and $K \bigcup K_2$ give $\overline{V}$ and $\overline{W}$ respectively the structure of subcomplexes of $\mathbb{R}P^2$. We can therefore compute the Euler characteristic χ of each of $\overline{V}$ and $\overline{W}$ using the formula

$$\chi = \alpha_0 - \alpha_1 + \alpha_2$$

where α_i is the number of i-cells (see e.g [Vic73a]). By Theorem 4.1, we have $\overline{\text{Int}(O)}$ homeomorphic to the closed disk and $\overline{\text{Ext}(O)}$ homeomorphic to the closed Möbius strip. Thus $\chi(\overline{\text{Int}(O)}) = 1$ and $\chi(\overline{\text{Ext}(O)}) = 0$. Hence we can determine from χ which of $\overline{V}$ and $\overline{W}$ is $\overline{\text{Int}(O)}$ and which is $\overline{\text{Ext}(O)}$. After making this determination for all ovals of C_f, we know, for any oval, which cells of D_f are inside, which on, and which outside it. From this information the ordering of ovals follows.

The chief tool for constructing D_f is the cylindrical algebraic decomposition (cad) algorithm [Arn82a, Arn82b, Col75a]. We use it to construct a cellular decomposition for an affine plane in $\mathbb{R}P^2$. Then by appropriately partitioning the line at infinity into cells, we extend to a cellular decomposition of $\mathbb{R}P^2$. These steps are described in detail in Section 3. Before applying the cad algorithm, we may possibly perform a linear change of coordinates of $\mathbb{R}P^2$. Section 2 gives the conditions under which we change coordinates, and defines the transformation used. In Section 5 we prove that the cellular decomposition of $\mathbb{R}P^2$ constructed in Section 3 is a complex.

We show in Section 6 that the computing time of our algorithm is $O(\,p(n,d)\,)$, for some polynomial function p of the degree n of f and the size d of its coefficients. Polotovskii [Pol73a] gave a topological type algorithm for curves of even degree, but did not establish a bound for it. His approach is quite different from ours: he examines the curves $f(x,y,z) + \varepsilon z^n$, ($n = degree(f)$), for various small values of ε. As noted by Fuks [Fuka] and Delzell [Del80a], one could get a topological type algorithm from Tarski's decision procedure for elementary algebra and geometry [Tar51a], but such an algorithm would have an exponential computing time bound. We have recently learned of an independently developed topological type algorithm by P. Gianni and C. Traverso [Gia83a], which has some resemblance to our method, but does not make use of cell complexes.

Section 7 provides an example of our algorithm. Because the time of our method depends almost entirely on the time required by the cad algorithm, and because the cad algorithm has recently been implemented [Arn81a], our algorithm appears to have some practical value. It could be used, for example, to study examples relating to Hilbert's 16th problem

[Wil78a].

Our method existed in rough form in summer 1982. It was first fully presented at a Purdue University Symposium in February, 1983.

2. The change of coordinates.

Assume that f is squarefree; if not, we may replace it with its greatest squarefree divisor h (i.e. the product of its distinct squarefree factors). since $C_f = C_h$. (See [Kal82a], p. 98, or [Col73a] for information on squarefree factorization). The line at infinity in $\mathbb{R}P^2$, written l_∞, consists of all points $[x,y,0]$ in $\mathbb{R}P^2$.

We want C_f to satisfy the following conditions:

(1) C_f has only simple intersections with l_∞ (this will be the case if and only if $f(x,y,0)$ does not have a multiple factor);

(2) C_f does not contain the point $[0,1,0]$.

If C_f does not satisfy these conditions initially, then we will transform $f(x,y,z)$ to a polynomial $E(U,V,W)$ such that E is squarefree and homogeneous of the same degree as f, C_f is non-singular if and only if C_E is non-singular, C_f and C_E have the same topological type, and E satisfies conditions (1) and (2). We shall then assume, by replacing f by E, that conditions (1) and (2) hold for f.

We show now how to obtain E. Let the degree of f be n, and let

$$f(x,y,z) = f_r(x,y,)z^{n-r} + \cdots + f_n(x,y)$$

where $0 \le r \le n$, each $f_i(x,y)$ is homogeneous of degree i, and $f_r(x,y) \ne 0$. The more typical situation is that in which $f_n(x,y) \ne 0$. However let us pause to consider our strategy in the event that $f_n(x,y) = 0$. In this case $z \mid f(x,y,z)$ but $z^2 \nmid f(x,y,z)$ as $f(x,y,z)$ is squarefree. We can therefore write $f(x,y,z) = zf_1(x,y,z)$, where

$$f_1(x,y,z) = f_r(x,y)z^{n-r-1} + \cdots + f_{n-1}(x,y)$$

and $f_{n-1}(x,y) \neq 0$. Thus l_∞ is contained in the curve C_f. Hence if C_{f_1} has any point on l_∞ (that is, if either $f_{n-1}(0,1) = 0$ or $f_{n-1}(1,y)$ has a real root) then C_f is a singular curve, and we report this fact and exit from the algorithm. If C_{f_1} does not meet l_∞ then C_f is non-singular if and only if C_{f_1} is non-singular. Moreover, if C_{f_1} is non-singular, then C_f and C_{f_1} have the same number and arrangement of ovals. Hence we can replace f by f_1; since C_{f_1} does not meet l_∞, conditions (1) and (2) of are trivially satisfied.

Let us assume, then, that $f_n(x,y) \neq 0$. We will now transform $f(x,y,z)$ into $F(X,Y,Z)$, where $F(0,1,0) \neq 0$ (so the point $[0,1,0]$ does not lie on the curve C_F). We know $f_n(x,1) \neq 0$, as otherwise $f_n(x,y) = 0$, a contradiction. Thus there is an integer λ such that $f_n(\lambda,1) \neq 0$. Define $F(X,Y,Z)$ by

$$F(X,Y,Z) = f(X + \lambda Y, Y, Z)$$

Then one has $F(0,1,0) = f(\lambda,1,0) = f_n(\lambda,1) \neq 0$.

Let $G(X,Y) = F(X,Y,1)$ and let $D(X)$ be the discriminant of $G(X,Y)$. Then $D(X) \neq 0$ as $G(X,Y)$ is squarefree. Find an integer α with $D(\alpha) \neq 0$, and consider the following change of variables: $X = W + \alpha U$, $Y = V$, $Z = U$. As $W = X - \alpha Z$, the line $X = \alpha Z$ (i.e. the affine line $X = \alpha$) corresponds to the line $W = 0$ (i.e. the line at ∞ in the U, V, W coordinates). Let

$$E(U,V,W) = F(W + \alpha U,\ V,\ U)$$

Now E is clearly squarefree and homogeneous of the same degree as f. Observe $E(0,1,0) = F(0,1,0) \neq 0$. Now $E(U,V,0) = F(\alpha U, V, U)$, so that $E(1,V,0) = F(\alpha, V, 1) = G(\alpha, V)$, a squarefree polynomial (since $D(\alpha) \neq 0$). Thus $E(U,V,0)$ is squarefree. Hence the curve C_E satisfies conditions (1) and (2). It remains to show that (i) C_f is non-singular if and only if C_E is non-singular; and (ii) C_f and C_E have the same topological type. Let $T(x,y,z) = (z, y, x - \alpha z - \lambda y)$. Then T is an invertible linear transformation of $\mathbb{R}^3$ with inverse given by $T^{-1}(U,V,W) = (W + \alpha U + \lambda V, V, U)$. Note that we have

$$E(U,V,W) = f(T^{-1}(U,V,W)). \tag{2.1}$$

Applying the chain rule for differentiation one finds

$$\begin{pmatrix} E_U \\ E_V \\ E_W \end{pmatrix} = \begin{pmatrix} \alpha & 0 & 1 \\ \lambda & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} f_x \\ f_y \\ f_z \end{pmatrix}. \tag{2.2}$$

Now the matrix on the left hand side of (2.2) is invertible. Hence (2.1) and (2.2) imply that (U,V,W) is a singular point of C_E if and only if $T^{-1}(U,V,W)$ is a singular point of C_f. This proves (i). As T is an invertible linear transformation of $\mathbb{R}^3$, T induces a homeomorphism $\widetilde{T}: \mathbb{R}P^2 \to \mathbb{R}P^2$ given by $\widetilde{T}[x,y,z] = [T(x,y,z)]$. By (4.2) $\widetilde{T}$ carries C_f onto C_E. Thus C_f and C_E have the same topological type: so (ii) is proved.

The reader may wonder why we do not transform C_f to a curve which has no intersections with l_∞. It is stated in [Rag06a] that there exist curves which, for any linear change of coordinates, will have points on l_∞.

3. Cellular decomposition of the projective plane.

Our objectives in this section are to define a cellular decomposition D_f of $\mathbb{R}P^2$, such that some subset of D_f is a decomposition for C_f, and to describe how we construct the following information about D_f: (1) a list of the pairs of adjacent cells, and (2) a list of the cells contained in C_f.

Let $g(x,y) = f(x,y,1)$. Using algorithm CADA2 of [Arn82b], we determine a proper g-invariant cylindrical algebraic decomposition (cad) D of $\mathbb{R}^2$. CADA2 produces a list of pairs of (the indices of) adjacent cells in the cad, a list of the (indices of the) cells on which g vanishes (these are exactly the sections of the cad, identifiable by their indices), and sample points for the cells. By exact evaluation of $g_x(x,y)$ and $g_y(x,y)$ at the 0-cell sample points, we determine whether they vanish simultaneously on some 0-cell. If so, we report that C_f is singular and exit from the topological type algorithm. If not, we report that C_f is non-singular and continue. (By Conditions (1) and (2) of Section 2, C_f has no singularities on l_∞).

Recall that $\mathbb{R}P^2$ is the disjoint union of U and l_∞, where U is the image of the affine plane $\mathbb{R}^2$ under the natural embedding $\iota : \langle x,y \rangle \to [x,y,1]$. Thus the images of the cells of D under ι are a cellular decomposition for U.[1] Furthermore the cells of D on which g vanishes are exactly the cells of U contained in C_f. Suppose that there are $k \geq 0$ points of C_f on l_∞. Since $[0,1,0]$ does not lie on C_f (by Section 2), these points can be written $[1,\gamma_1,0],\ldots,[1,\gamma_k,0]$, where $\gamma_1 < \cdots < \gamma_k$ are the real roots of $f(1,y,0)$. By isolating these roots [Col82a], we determine a cellular decomposition of l_∞ consisting of the following elements: the points of C_f on l_∞, the point

[1] for convenience we will not distinguish between a cell c of D and $\iota(c)$.

[0,1,0], and the $k+1$ 1-cells which comprise the remainder of l_∞. We can assign indices to these cells (in the sense of [Arn82b]) in some arbitrary fashion (cf. the example in Section 7). Thus we have defined D_f, we have an index for every cell, and we have (a list of the indices of) the cells which belong to C_f.

The adjacencies within U have been given to us by the cad algorithm. The adjacencies of cells within l_∞ are obvious. The following theorem is the basis for determination of adjacency between a cell of l_∞ and a cell of U (see [Arn82b], Sec. 2 for the definitions of stack, section, and φ-section):

THEOREM 3.1. *Let S and T be (respectively) the "rightmost" and "leftmost" stacks of D. Then*

(i) S has k sections, say $s_1 < \cdots < s_k$, and T has k sections, say $t_1 < \cdots < t_k$;

(ii) if s_i is the graph of the continuous real-valued function φ, and t_i is the graph of the continuous real-valued function ψ, for $1 \le i \le k$, then

$$\lim_{x \to +\infty} \frac{\varphi(x)}{x} = \gamma_i, \text{ and } \lim_{x \to -\infty} \frac{\psi(x)}{x} = \gamma_{k-i+1}.$$

Proof. Let n be the degree of $f(x,y,z)$. By condition (1) of Section 2, each γ_i is a simple root of $f(1,y,0)$. Let $G(X,Y) = f(1,Y,X)$. Then since $f(0,1,0) \neq 0$ by condition (2) of Section 2, $G(X,Y) = g_0 Y^n + g_1(X)Y^{n-1} + \cdots + g_n(X)$, for some constant $g_0 \neq 0$ and some polynomials $g_1(X),\ldots,g_n(X)$. Since $G(0,Y) = f(1,Y,0)$, $G(0,Y)$ has exactly k real roots $\gamma_1 < \cdots < \gamma_k$, each of them simple. Hence by root continuity, there is some $\delta>0$ such that $|X|<\delta$ implies $G(X,Y)$ has exactly k real roots, each of them simple, the i^{th} of which approaches γ_i as $|X| \to 0$. Since $g(x,y) = x^n G(1/x, y/x)$ for nonzero x, $g(x,y)$ has k real roots, each simple, for all sufficiently large positive x. Hence S has k sections. A similar argument shows that T has k sections.

For any x in the interval $(\alpha, +\infty)$ on which φ is defined, $\varphi(x)$ is the i^{th} real root of $g(x,y)$. Hence, for positive x greater than α, $\frac{\varphi(x)}{x}$ is the i-th real root of $G(1/x, Y)$. Hence, as x approaches $+\infty$, $\frac{\varphi(x)}{x}$ approaches γ_i. For any x in the interval $(-\infty, \beta)$ in which ψ is defined, $\psi(x)$ is the i-th real root of $g(x,y)$. Hence, for negative x less than β, $\frac{\psi(x)}{x}$ is the $(k-i+1)$-th real root of $G(1/x, Y)$. Hence, as x approaches $-\infty$, $\frac{\psi(x)}{x}$ approaches γ_{k-i+1}

■

Figure 3 in Section 7 illustrates the theorem. S and T each have four sections. One sees that the asymptotic slope of s_i, namely γ_i, is equal to the asymptotic slope of t_{k-i+1}.

Let $P_i = [1, \gamma_i, 0]$ for $1 \le i \le k$, let $P_0 = P_{k+1} = [0,1,0]$, and let e_i denote the 1-cell in l_∞ between P_i and P_{i+1}, for $0 \le i \le k$. Note that $\overline{e_i} = e_i \cup \{P_i, P_{i+1}\}$. Let R be any stack of D, say over the interval I, with sections $r_0 < r_1 < \cdots < r_l < r_{l+1}$, where $r_0 = I \times \{-\infty\}$ and $r_{l+1} = I \times \{+\infty\}$ are the infinite sections. For $0 \le i \le l$ let $\hat{r}_i$ denote the sector of R lying between r_i and r_{i+1}.

Let S and T be as in Theorem 3.1. We now consider adjacencies between cells of S and l_∞, and cells of T and l_∞. In general, S and T are distinct stacks of D, however it is possible that $S = T$ is the only stack of D. Suppose $S \ne T$. We show that P_i is a limit point of s_i (and hence that s_i is adjacent to P_i). Let $[x_i, \varphi(x_i), 1]$ be a sequence of points in s_i, with x_i approaching $+\infty$. Then $\lim [x_i, \varphi(x_i), 1] = \lim [1, \frac{\varphi(x_i)}{x_i}, \frac{1}{x_i}] = [1, \gamma_i, 0] = P_i$. It can be shown that

P_i is in fact the unique limit point of s_i on l_∞. Similarly, P_{k-i+1} is the unique limit point of t_i on l_∞. If $S = T$, then $s_i = t_i$ has exactly two limit points P_i and P_{k-i+1} in l_∞.

If $S \neq T$, it is evident that the portion of the boundary of $\hat{s}_i$ contained in l_∞ is $\bar{e}_i$, while the portion of the boundary of $\hat{t}_i$ contained in l_∞ is $\bar{e}_{k-i}$, $0 \leq i \leq k$ (Figure 1) If $S = T$ is the only stack of D, the portion of the boundary of $\hat{s}_i$ contained in l_∞ is $\bar{e}_i \cup \bar{e}_{k-i}$ (Figure 2). A sector of S or T is adjacent to exactly those cells in l_∞ which belong to its boundary.

Now let R be any stack of D besides S and T. Let $r_1 < \cdots < r_l$ be the finite sections of R. Then clearly only $\hat{r}_0$ and $\hat{r}_l$ have limit points on l_∞, and each in fact has the unique limit point P_0 on l_∞. This completes the determination of all adjacencies between cells of l_∞ and cells of U.

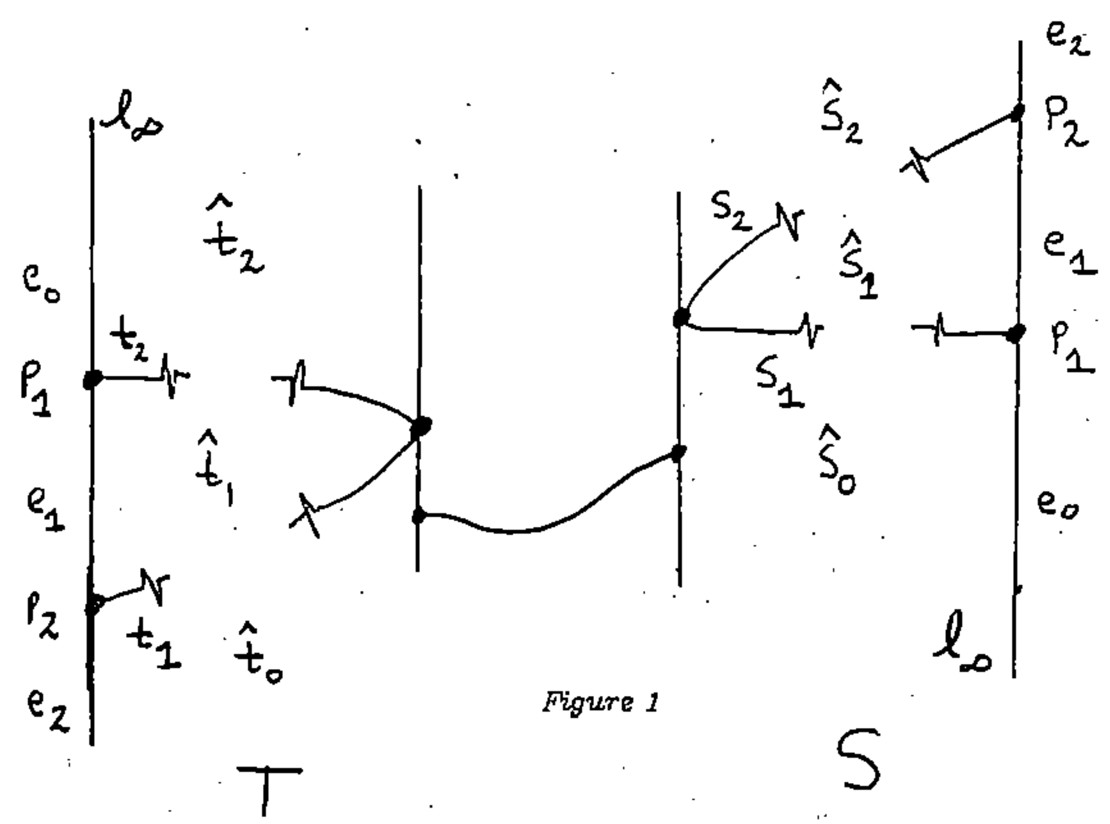

Figure 1

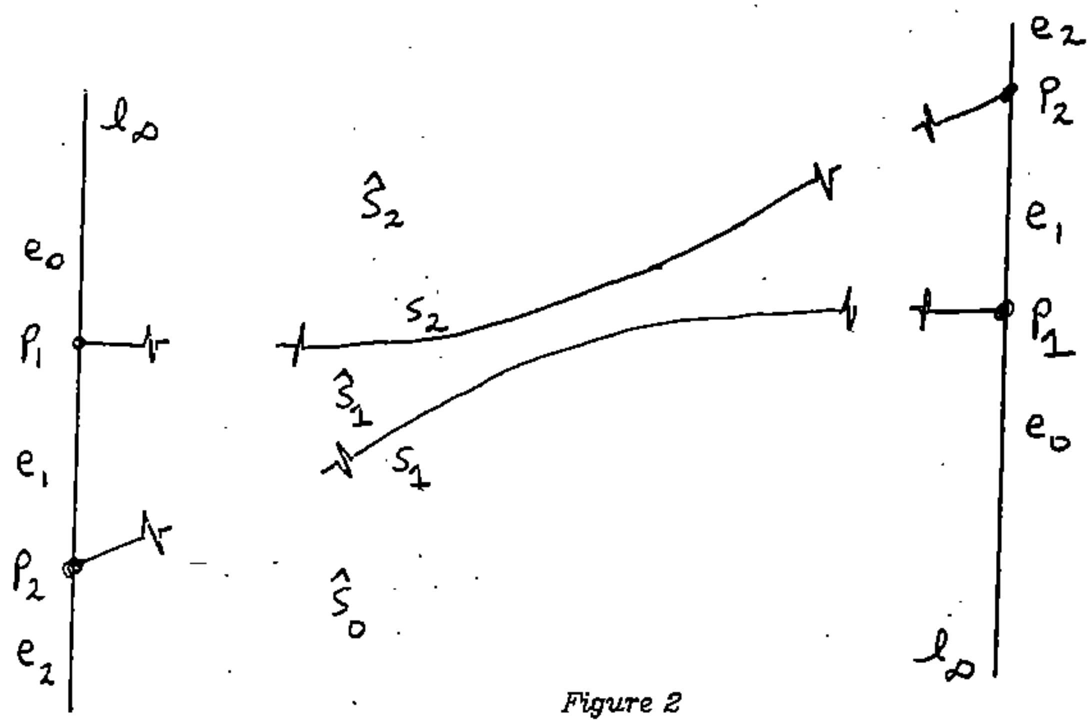

Figure 2

4. Simple closed curves in the projective plane.

In this section we characterize the two possible imbeddings of a circle in $\mathbb{R}P^2$. A *simple closed curve* is a topological space homeomorphic to S^1, i.e. a space which is essentially a circle. In the following, M^2 denotes the closed Möbius band, X denotes $\mathbb{R}P^2$, and "$\approx$" denotes homeomorphism.

THEOREM 4.1. *Let C be a simple closed curve in X. Then either:*

(i) $X \setminus C$ has exactly two connected components V and W with common boundary C (i.e. $C = \bar{V} \setminus V = \overline{W} \setminus W$), such that, after interchanging V and W if necessary, $\bar{V} \approx B^2$ and $\overline{W} \approx M^2$; or

(ii) $X \setminus C = V$ is connected, $V \approx U^2$, C is the boundary of V, and $\bar{V} = X$.

Proof. We make use of the fact (see [Mun75a], Sec. 8-7) that S^2 is a covering space of X, with covering map $\pi : S^2 \to X$ given by $\pi(x,y,z) = [x,y,z]$. It can be shown using the path lifting property (Lemma 4.1 in Ch. 8 of

[Mun75a], that the subset $\pi^{-1}(C)$ of S^2 consists of either one or two disjoint simple closed curves. In the latter case, let C_1 and C_2 be the two disjoint simple closed curves comprising $\pi^{-1}(C)$. By the Jordan curve theorem and the Schoenflies theorem (Sec. 8-13 of [Mun75a]), C_1 and C_2 separate the sphere into three components V_1, V_2 and W_1, with $\bar{V}_1 \approx \bar{V}_2 \approx B^2$, and $\bar{W}_1 \approx$ a closed annulus. Moreover, C_1 is the boundary of V_1, C_2 is the boundary of V_2, and $C_1 \cup C_2$ is the boundary of W_1. Using these facts it can be shown that $\pi(V_1) = \pi(V_2)$, and that (i) holds, with $V = \pi(V_1)$ and $W = \pi(W_1)$. In the former case, let C_1 be the simple closed curve $\pi^{-1}(C)$. By the Jordan curve theorem, C_1 separates the sphere into two components V_1 and V_2, of which C_1 is the common boundary. Moreover, by the Schoenflies theorem, $\bar{V}_1$ and $\bar{V}_2$ are each homeomorphic to B^2. One can show that $\pi(V_1) = \pi(V_2)$ and that (ii) holds, with $V = \pi(V_1)$ ▪

5. Cell complex structure for the projective plane.

We prove that the cellular decomposition D_f of $\mathbb{R}P^2$ defined in Section 3 gives $\mathbb{R}P^2$ the structure of a finite cell complex (and hence, the structure of a finite CW-complex). For convenience, we review the definition of a complex. Let B^n denote the n-dimensional closed unit ball in $\mathbb{R}^n$, U^n the n-dimensional open unit ball in $\mathbb{R}^n$, S^{n-1} the $(n-1)$-sphere in $\mathbb{R}^n$, I^n the closed n-cube in $\mathbb{R}^n$. A *finite cell complex* X is a Hausdorff space which is the union of finitely many disjoint open cells e^i_α ($\alpha \in A$) such that to each e^i_α there corresponds a continuous map $\chi_\alpha : B^i \to X$ for which $\chi_\alpha(S^{i-1}) \subseteq X^{i-1}$ (where X^{i-1}, called the *(i-1)-skeleton* of X, is the union of all cells of dimension $\leq i-1$) and $\chi_\alpha|_{U^i}$ is a homeomorphism from U^i onto e^i_α. The map χ_α is called a *characteristic map* for e^i_α. For more information on cell complexes the reader can consult any of the following texts: [Hil3,a], Chapter 7 [Mas67a], Chapter 7 [Gra75a], Section 14.

The proof of the following theorem involves some of the basic notions of cylindrical algebraic decompositions (e.g. sector, section), for which the reader may wish to consult Section 2 of [Arn82a].

THEOREM 5.1. *Every cell of D_f has a characteristic map. D_f thus gives $\mathbb{R}P^2$ the structure of a finite cell complex.*

Proof. Let $X = \mathbb{R}P^2$. Let us adopt the convention that $[x, +\infty, z] = [x, -\infty, z] = [0,1,0]$ for x & z finite. Note that, as B^i is homeomorphic to I^i under a map carrying S^{i-1} to $\partial(I^i)$, it suffices to give characteristic maps from I^i into X.

Characteristic maps for 0-cells are trivial. Let e^1 be a 1-cell of D_f. e^1 is contained either in l_∞ or in the affine plane U (cf. Section 3). In the latter case, it is a cell of the cad that we constructed in Section 2, and hence is either a sector over a point $x = \alpha$, α finite (case 1) or a section over an interval (α,β), $-\infty \leq \alpha < \beta \leq +\infty$ (case 2). If e^1 is contained in l_∞, then $e^1 = \{[1,y,0] \in X : \gamma < y < \delta\}$, some γ, δ, $-\infty \leq \gamma < \delta \leq +\infty$. Let $\sigma:[0,1] \rightarrow [\gamma,\delta]$ be a homeomorphism such that $\sigma(0) = \gamma$, $\sigma(1) = \delta$. Define $\chi:[0,1] \rightarrow X$ by $\chi(s) = [1,\sigma(s),0]$. Clearly χ is a characteristic map for e^1. In case (1) $e^1 = \{[\alpha,y,1] \in X : \gamma < y < \delta\}$, $-\infty \leq \gamma < \delta \leq +\infty$. Let $\sigma:[0,1] \rightarrow [\gamma,\delta]$ be a homeomorphism such that $\sigma(0) = \gamma$, $\sigma(1) = \delta$. Define $\chi:[0,1] \rightarrow X$ by $\chi(s) = [\alpha,\sigma(s),1]$. Clearly χ is a characteristic map for e^1. In case (2) $e^1 = \{[x,y,1] \in X : \alpha < x < \beta, y = \varphi(x)\}$, where $-\infty \leq \alpha < \beta \leq +\infty$ and φ is a continuous function from (α,β) into $\mathbb{R}$. Let $\sigma : [0,1] \rightarrow [\alpha,\beta]$ be a homeomorphism with $\sigma(0) = \alpha$, $\sigma(1) = \beta$. Define $\chi : (0,1) \rightarrow X$ by

$\chi(s) = [\sigma(s), \varphi(\sigma(s)), 1]$. Clearly χ is a homeomorphism from $(0,1)$ onto e^1. Now χ has a continuous extension to $[0,1]$: if α is finite, set $\chi(0) = [\alpha, \gamma, 1]$, where $\gamma = \lim_{x \to \alpha +} \varphi(x)$; if α is infinite ($\alpha = -\infty$), set $\chi(0) = [1, \gamma, 0]$, where $\gamma = \lim_{x \to -\infty} \frac{\varphi(x)}{x}$ (this limit exists and is finite by Theorem 3.1); define $\chi(1)$ similarly. Hence χ is a characteristic map for e^1.

Let e^2 be a 2-cell of D_f. Then $e^2 = \{[x, y, 1] \in X : \alpha < x < \beta, \varphi(x) < y < \psi(x)\}$, where $-\infty \le \alpha < \beta \le +\infty$, and φ & ψ are continuous functions on (α, β), with $\varphi < \psi$ (possibly $\varphi \equiv -\infty$ or $\psi \equiv +\infty$). Let $\sigma : [0,1] \to [\alpha, \beta]$ be a homeomorphism such that $\sigma(0) = \alpha$, $\sigma(1) = \beta$. There are three cases to consider.

Case I. φ and ψ both finite. For $0 < s < 1$ and $0 \le t \le 1$ set $\chi(s,t) = [\sigma(s), \tau(s,t), 1]$ where $\tau(s,t) = \varphi(\sigma(s)) + t(\psi(\sigma(s)) - \varphi(\sigma(s)))$. Clearly χ maps $(0,1)x(0,1)$ homeomorphically onto e^2. Now χ has a continuous extension to I^2 : if α is finite, set $\chi(0,t) = [\alpha, \gamma + t(\delta - \gamma), 1]$ where $\gamma = \lim_{x \to \alpha +} \varphi(x)$ & $\delta = \lim_{x \to \alpha +} \psi(x)$; if α is infinite $(\alpha = -\infty)$, set $\chi(0,t) = [1, \gamma + t(\delta - \gamma), 0]$, where $\gamma = \lim_{x \to -\infty} \frac{\varphi(x)}{x}$ and $\delta = \lim_{x \to -\infty} \frac{\psi(x)}{x}$; define $\chi(1,t)$ similarly.

Case II. φ and ψ both infinite (i.e. $\varphi \equiv -\infty$ & $\psi \equiv +\infty$). Suppose first that either α or β is finite. After a change of coordinates (of the type discussed in Section 2), we may assume that both α and β are finite. We can define a characteristic map $\chi : I^2 \to X$ for e^2 by $\chi(s,t) = [\sigma(s), \tau(t), 1]$, where τ is a homeomorphism from $[0,1]$ onto $[-\infty, +\infty]$. Suppose, on the other hand, that both α and β are infinite (that is, $\alpha = -\infty$ & $\beta = +\infty$). Notice that in the case $\bar{e}^2 = X$. The closed disk B^2 can be mapped into X as follows:

$\chi(x,y) = [x,y,\sqrt{1-x^2-y^2}]$, (for $x^2+y^2 \le 1$). The map χ is a characteristic map for e^2.

Case III. either φ or ψ, not both, infinite. Say φ is finite & $\psi \equiv +\infty$. If either α or β is finite then one can reduce as in Case II to the case in which both α and β are finite, and easily write down a characteristic map for e^2. Suppose, then, that α and β are both infinite. Let D^2 be the closed semi-circular region $\{(x,y)\in\mathbb{R}^2: x^2+y^2 \le 1, y\ge 0\}$. Define a map from the portion of D^2 in which $x^2+y^2<1$ into X by

$$\chi(x,y) = \left[\frac{x}{\sqrt{1-x^2-y^2}}, \frac{y}{\sqrt{1-x^2-y^2}} + \varphi\left(\frac{x}{\sqrt{1-x^2-y^2}}\right), 1\right]$$

Now χ maps the interior of D^2 homeomorphically onto e^2. But χ has a continuous extension to the whole of D^2: if $x^2+y^2 = 1$ & $x<0$ then set $\chi(x,y) = [1, y/x+\gamma, 0]$, where $\gamma = \lim_{x\to-\infty} \frac{\varphi(x)}{x}$; if $x^2+y^2 = 1$ & $x>0$ then set $\chi(x,y) = [1, y/x+\gamma', 0]$, where $\gamma' = \lim_{x\to+\infty} \frac{\varphi(x)}{x}$; finally, set $\chi(0,1) = [0,1,0]$. As D^2 is homeomorphic to B^2 under a map carrying ∂D^2 onto S^1, we may regard χ as an characteristic map for e^2.

6. Computing time analysis.

We show that the maximum computing time of our algorithm is dominated by a polynomial function of n (the degree of $f(x,y,z)$) and $d = \log \hat{d}$, where $\hat{d}$ is the sum of the absolute values of the numerators and denominators of all rational coefficients of f, i.e. the norm of f.

The steps of the algorithm to implement Section 2 are: carry out two linear changes of coordinates, compute the discriminant of a bivariate polynomial of degree n, and isolate the real roots of this discriminant. The times for these operations are polynomial in n and d [Col71a, Col82a] When we are done with these steps,we will have some (possibly new) $f(x,y,z)$ of degree n; let e denote its norm. log e is bounded by a polynomial function of n and d.

Let $g(x,y) = f(x,y,1)$. Collins [Col75a] established that the time for construction of a g-invariant cad of E^2 is polynomial in n and log e. The cad algorithm in [Arn82a, Arn82b] is slightly different from that which Collins analyzed, and in addition constructs adjacencies, but its computing time is also so bounded. Furthermore, there are at most $O(n^3)$ cells in the cad constructed by the algorithm, as we now show: since [0,1,0] is not on C_f, $f(x,y,z) = cy^n$ + (terms of lower degree in y), for some rational number c, so $g(x,y) = cy^n$ + (terms of lower degree in y), hence $PROJ(g)$ = discriminant(g), and *degree*(discriminant(g)) $= O(n^2)$. (See [Arn82a] for the definition of PROJ). The evaluation of g_x and g_y at 0-cell sample points of the cad takes polynomial time.

We determine how many points C_f has on l_∞ by isolating the real roots of $f_n(1,y)$. This takes time polynomial in n and log e [Col82a]. There are at most n such roots. Hence D_f has $O(n^3)$ cells.

The sections of D_f are precisely the cells of D_f on which f vanishes, so we can determine whether f vanishes on a cell by examining its cell index (constant time per cell; $O(n^3)$ cells). There are $O(\binom{n^3}{2}) = O(n^6)$ adjacencies. Hence in time $O(n^6)$ we can determine the equivalence classes of $\overline{R}(C_f)$.

For each component of C_f, we can find the equivalence classes of $\overline{R}$(complement(O)) in time $O(n^6)$ For each component which is an oval, we can compute the Euler characteristics of D_1 and D_2 by merely scanning the lists of their $O(n^5)$ cells and calculating their dimensions (from the cell indices, see [Arn82a], Section 4) in constant (or at most $O(\log n)$) time. Since, by Harnack's Theorem [Wil78a], C_f has $O(n^2)$ components, the total time for step 5 is $O(n^8)$.

We have thus shown that the total time required by our algorithm is $O(p(n, d))$, for some polynomial function $p(n, d)$.

7. Example.

We now do an example of our algorithm. Let the input polynomial be:

$$f(x,y,z) = y^4 - 2xy^3 - x^2y^2 + 2x^3y + y^2z^2 + x^2z^2 - z^4.$$

f is irreducible, hence squarefree. $f(x,y,0)$ has no multiple factors, and $[0,1,0]$ does not lie on C_f, so we need not change coordinates.

Let $g(x,y) = f(x,y,1)$. A proper g-invariant cad D of $\mathbb{R}^2$ looks as follows:

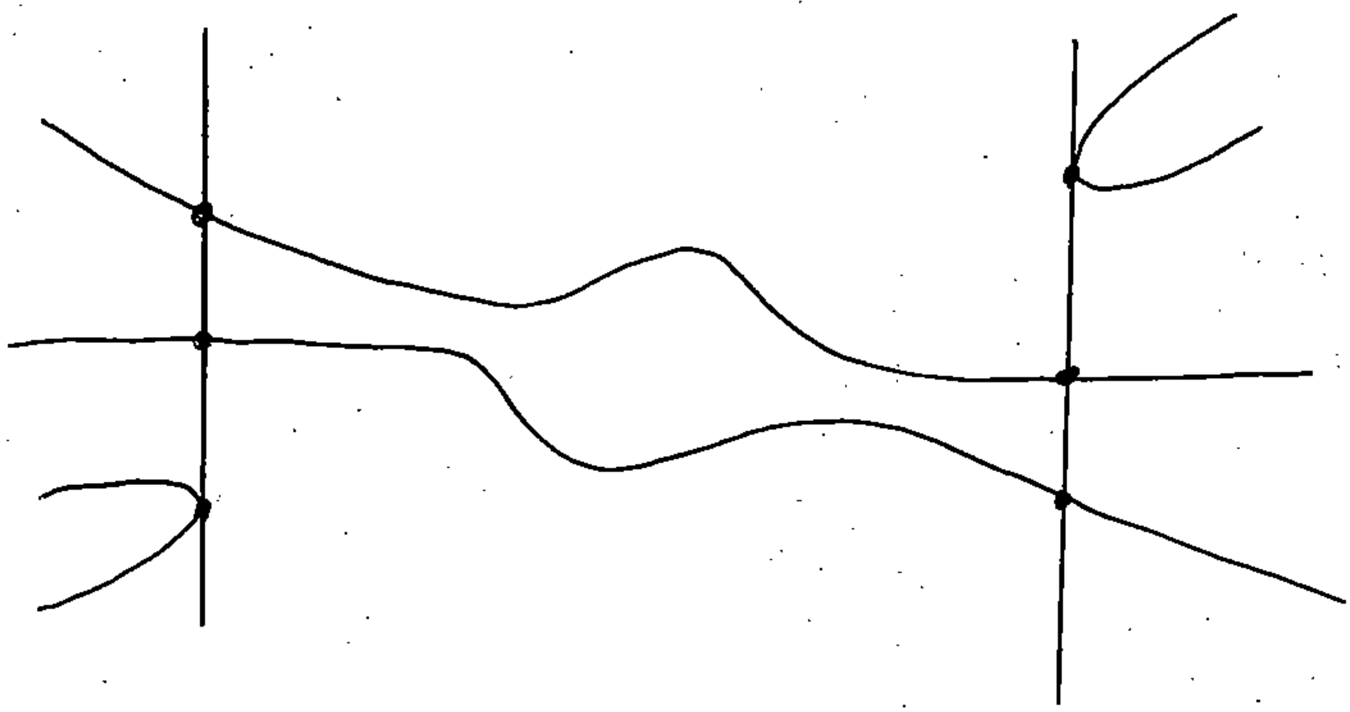

Figure 3

The indices of these cells are:

(1,9)				(5,9)
(1,8)				(5,8)
(1,7)	(2,7)		(4,7)	(5,7)
(1,6)	(2,6)		(4,6)	(5,6)
(1,5)	(2,5)	(3,5)	(4,5)	(5,5)
(1,4)	(2,4)	(3,4)	(4,4)	(5,4)
(1,3)	(2,3)	(3,3)	(4,3)	(5,3)
(1,2)	(2,2)	(3,2)	(4,2)	(5,2)
(1,1)	(2,1)	(3,1)	(4,1)	(5,1)

We find that $g_x(x,y)$ and $g_y(x,y)$ do not vanish simultaneously at any 0-cell of D, so C_f is nonsingular.

Continuing, we find that

$$f(x,y,0) = y(y-x)(y+x)(y-2x),$$

and so C_f has the four points [1,0,0], [1,1,0], [1,-1,0], and [1,2,0] on l_∞. Thus D_f consists of the (imbeddings in $\mathbb{R}P^2$ of the) cells of D, the four just-listed points of l_∞ plus [0,1,0], and the remaining 1-cells that make up l_∞. Let us use the convention that (0,0) is the cell index of [0,1,0], $(0,2i)$ is the cell index of the cell in l_∞ corresponding to the i^{th} real root of $f(1,y,0)$, and the 1-cells in l_∞ have the naturally induced indices consistent with these 0-cell indices. Thus the indices of the cells in D_f which make up l_∞ are:

(0,0), (0,1), (0,2), (0,3), (0,4), (0,5), (0,6), (0,7), (0,8), (0,9)

We find that C_f has two components, composed of the following collections of cells (these are the equivalence classes of $\overline{R}(C_f)$):

{ (1,2), (2,2), (1,4), (0,6), (5,8), (4,6), (5,6), (0,6) }

{ (1,6), (2,4), (3,2), (4,2), (5,2), (0,2), (1,8), (2,6), (3,4), (4,4), (5,4), (0,4) }

Since f has even degree, both are ovals.

Consider the first component of C_f above. We get two equivalence classes K_1 and K_2 for $\overline{R}$(complement(O)):

{ (1,3), (0,7), (5,7) }

{ (0,0), (1,9), (5,9), (0,9), (1,8), (1,7), (2,7), (4,7),
(1,6), (2,6), (1,5), (2,5), (3,5), (4,5), (5,5), (0,5),
(2,4), (3,4), (4,4), (5,4), (0,4), (2,3), (3,3), (4,3),
(5,3), (0,3), (3,2), (4,2), (5,2), (0,2),
(1,1), (2,1), (3,1), (4,1), (5,1), (0,1) }

Note that the dimension of a cell is equal to the sum of the parities (even=0, odd=1) of the components of its cell index, e.g. (1,9) is a 2-cell, (4,4) is a 0-cell. Thus we see that the Euler characteristic of the complex consisting of the cells of K_1 together with the cells of the oval is $\chi = 4 - 5 + 2 = 1$. For the complex consisting of the cells of K_2 plus the cells of the oval we have $\chi = 11 - 22 + 11 = 0$. Hence the first cluster is the interior, and the second the exterior, of this oval.

Now consider the second oval of C_f. Again we get two classes K_1 and K_2 for $\bar{R}$(complement(O)):

{ (1,7), (2,5), (3,3), (4,3), (5,3), (0,3) }

{ (0,0), (1,9), (5,9), (0,9), (5,8), (0,8), (2,7), (4,7), (5,7), (0,7),
(4,6), (5,6), (0,6), (1,5), (3,5), (4,5), (5,5), (0,5), (1,4),
(1,3), (2,3), (1,2), (2,2), (1,1), (2,1), (3,1), (4,1), (5,1), (0,1) }

The Euler characteristic of the complex consisting of K_1 plus the cells of the oval is $\chi = 6 - 9 + 3 = 0$. For the complex consisting of K_2 plus the cells of the oval we have $\chi = 11 - 20 + 10 = 1$. Hence the first cluster is the exterior, and the second the interior, of this oval.

We now see by inspection that the cells comprising the first oval occur among the cells comprising the interior of the second oval. Equivalently, the cells comprising the second oval occur among the cells comprising the exterior of the first oval. Hence the topological type of C_f may be specified by saying that it consists of two ovals, one inside the other.

8. Acknowledgements

We are indebted to G. Brumfiel for the observation that Euler characteristic suffices to distinguish the interior of an oval from its exterior (we had originally envisioned a full homology calculation). The second author would like to acknowledge helpful and inspiring conversations on the subject of this paper with the following people: G. Collins, E. Fadell, T.-C. Kuo, E. Mansfield.

References

Arn81a.
DS Arnon, "Algorithms for the geometry of semi-algebraic sets," Technical Report #436, Computer Science Dept., University of Wisconsin, Madison, Wisconsin(1981). (Ph.D. thesis)

Arn82a.
DS Arnon, GE Collins, and S McCallum, "Cylindrical algebraic decomposition I: the basic algorithm," Technical Report CSD TR-427, Computer Science Dept., Purdue University(December, 1982).

Arn82b.
DS Arnon, GE Collins, and S McCallum, "Cylindrical algebraic decomposition II: an adjacency algorithm for the plane," Technical Report CSD TR-428, Computer Science Dept., Purdue University(December, 1982).

Col71a.
GE Collins, "The calculation of multivariate polynomial resultants," *J. Assoc. Comp. Mach.* **18**, pp. 515-532 (1971).

Col73a.
GE Collins, "Computer algebra of polynomials and rational functions," *Amer. Math. Monthly* **80**, pp. 725-755 (1973).

Col75a.
GE Collins, "Quantifier elimination for real closed fields by cylindrical algebraic decomposition," pp. 134-163 in *Proceedings of the Second GI Conference on Automata and Formal Languages*, Lecture notes in Computer Science, 33, Springer-Verlag, Berlin(1975).

Col82a.
GE Collins and RGK Loos, "Real zeros of polynomials," *Computing, Supplementum 4: Computer Algebra - Symbolic and Algebraic Computation,*, pp. 83-94 Springer-Verlag, , (1982).

Del80a.
CN Delzell, , private communication(1980).

Fuka.
D. Fuks, "Review of Polotovskii, G.M., An algorithm for determining the topological type of a structurally stable plane curve of even degree," *Math. Reviews* **58**, Review #28000

Gia83a.
P Gianni and C Traverso, *Shape determination for real curves and surfaces*, manuscript(1983).

Gra75a.
B Gray, *Homotopy theory - an introduction to algebraic topology*, Academic Press, New York(1975).

Hil3,a.
PJ Hilton, *An introduction to homotopy theory*, Cambridge University Press(1966). (Cambridge Tracts in Mathematics and Mathematical Physics, No. 43)

Kal82a.
E Kaltofen, "Polynomial factorization," *Computing, Supplementum 4: Computer Algebra - Symbolic and Algebraic Computation*,, pp. 95-114 Springer-Verlag. , (1982).

Mas67a.
WS Massey, *Algebraic topology - an introduction*, Harcourt, Brace, and World, New York(1967).

Mun75a.
J Munkres, *Topology - a first course*, Prentice-Hall, Englewood Cliffs(1975).

Pol73a.
GM Polotovskii, "Algorithm for determining the topological type of a rough plane algebraic curve of even degree," *Gor'kov Gos. Univ. Ucen. Zap. Vyp.*, **187**, pp. 143-187 (1973). (Russian)

Rag06a.
V Ragsdale, "On the arrangement of the real branches of plane algebraic curves," *Amer. J. Math.* **28**, pp. 377-404 (1906).

Tar51a.
A Tarski, *A decision method for elementary algebra and geometry*, University of California Press(1951). (second revised edition)

Vic73a.
JW Vick, *Homology theory: an introduction to algebraic topology*, Academic Press, New York(1973).

Wil78a.
G Wilson, "Hilbert's Sixteenth Problem," *Topology* **17**, pp. 53-73 (1978).