BULL. AUSTRAL. MATH. SOC. 14H45, 68Q25
VOL. 43 (1991) [37–50]

THE TOPOLOGICAL CONFIGURATION OF A REAL ALGEBRAIC CURVE

TAKIS SAKKALIS

This paper presents an algorithm, motivated by Morse Theory, for the topological configuration of the components of a real algebraic curve $\{f(x, y) = 0\}$. The running time of the algorithm is $O\left(n^{12}(d + \log n)^2 \log n\right)$, where n, d are the degree and maximum coefficient size of $f(x, y)$.

1. INTRODUCTION

Let $f(x, y)$ be a polynomial with integer coefficients, of degree n, $n \geqslant 2$. Let C be the real affine curve defined by $C = \{(x, y) \in \mathbf{R}^2 \mid f(x, y) = 0\}$.

It is known that C consists of at most finitely many connected components.

More precisely, when the curve is real non-singular (Section 2) each unbounded component of it is homeomorphic to a line and each bounded component is homeomorphic to a circle. We will call a bounded component an oval. An oval has a definite interior, homeomorphic to an open disk, and an exterior, homeomorphic to $\mathbf{R}^2$ minus a closed disk. On the other hand, if C is real singular and K is the set of its real singular points (Section 2), then $C - K$ is a differentiable 1-manifold. Therefore, each component of $C - K$ is homeomorphic to a line or a circle. Furthermore, a component of C is either a component of $C - K$ or a disjoint union of components of $C - K$ and a subset of K.

In this paper we present a method for the topological configuration of the components of C. More precisely, in Section 2 we will first decide whether C is non-singular. When this is the case our procedure does the following:

1. It counts the number of components of C, and triangulates each component.
2. It finds the configuration of the components. In particular, given two ovals A and B, we will find their relative position.

The main idea of our method is based on an efficient way of locating the critical points of the projection map $h: (x, y) \to x$; that is, the real solutions of the system $f(x, y) = \partial f/\partial y(x, y) = 0$. For each critical point (x_0, y_0) we construct an isolating

Received 31 January 1990

rational rectangle Γ and moreover we compute (3.3.4) an integer $I(x_0, y_0)$, called the index of (x_0, y_0). Then Proposition 3.3.3 enables us to determine the shape of the curve near (x_0, y_0). We will see that the above local analysis is sufficient for the determination of the configuration of C.

On the other hand, if C is real singular we first locate its real singular points in addition to the regular critical points of h. Then if (x_0, y_0) is a singular point and Γ an isolating rational rectangle of it, we find the local topology of the curve near (x_0, y_0). In particular we decide on the number of components of $C - K$ which go through (x_0, y_0).

In Section 3 we present a brief review of our previous work [**11, 12**], which provides the necessary geometric and algebraic tools pertinent to this work. In addition we include a procedure for finding a linear and a quadratic Morse function on a real non-singular curve. (A Morse function M is a real function which has only non-degenerate critical points.) Although the latter procedure seems to have no resemblance to the main algorithm, it is in fact Morse Theory (in particular Proposition 3.3.2) which inspired our method.

In Section 5 we calculate the computing time of our method. We note that it is polynomial in the degree n and the coefficient size d of f.

We conclude with some illustrative examples of both singular and non-singular curves.

The problem of the configuration of an algebraic curve is not new. It has been considered by several authors, along with similar problems, such as curve tracing and the topological type of a curve. Arnon and McCallum [**3**] described how one can use the cylindrical algebraic decomposition (cad) algorithm to find the topological type of a curve. Arnborg and Feng [**1**], using a different method, were able to decompose a non-singular curve. In addition, Gianni and Traverso [**7**] have a procedure for shape determination of real curves.

It is interesting to note that most of the above procedures use symbolic manipulation methods and exact arithmetic. Our approach however is quite different. It has been motivated by Morse Theory, and all of our computations are done over the ring $\mathbb{Q}[x_1, \ldots, x_k]$, $k \geqslant 1$, using only the basic operations of that ring. This has the advantage that the computing time is of the order $O\left(n^{12}(d + \log n)^2 \log n\right)$, and moreover our procedure can be implemented using any low level language with full rational arithmetic capabilities.

2. Decision procedure for a curve to be singular

In this section we present a procedure which decides whether a curve is singular. It is based on a convenient linear change of coordinates (Lemma 2.2). Most of the key

results mentioned below appear in [13], and therefore we state them without proof.

We begin with some preliminaries. For polynomials $p(x)$, $q(x)$ with coefficients in a unique factorisation domain I, we denote by $R = \mathrm{Res}_x(p, q)$, their resultant with respect to the variable x [13]. Note that $R \in I$. Let $s(x, y)$ be a polynomial over $\mathbf{R}[x, y]$, and let k be its degree. We say that s is regular in y whenever the coefficient of y^k in s is a non-zero constant.

Now consider our polynomial $f(x, y)$. We may suppose that f is square factor free; that is for every $g \in \mathbf{Q}[x, y]$ of positive degree, g^2 does not divide f. Let Σ be the following system of equations:

$$\Sigma : f = 0, \quad \frac{\partial f}{\partial x} = 0, \quad \frac{\partial f}{\partial y} = 0.$$

We observe that Σ has at most finitely many solutions over $\mathbf{C}^2$.

DEFINITION 2.1: We say that C is non-singular over $\mathbf{R}^2$ (respectively over $\mathbf{C}^2$) if there is no point in $\mathbf{R}^2$ (respectively in $\mathbf{C}^2$) which satisfies S. C is singular over $\mathbf{R}^2$ ($\mathbf{C}^2$) if such a point can be found.

Let now u, v be new coordinates so that $x = u + mv$, $y = v$, and consider $g(u, v) = f(u + mv, v)$. We have:

LEMMA 2.2. *We can pick an integer m so that the following two conditions are satisfied:*

1. *$g(u, v)$ is regular in v;*
2. *whenever the points (u_0, v_0), (u_0, v_1) satisfy the equations $g = \partial g/\partial u = 0$, then $v_0 = v_1$.*

PROOF: See [13].

NOTATION. Denote by $f_x = \partial f/\partial x$, $f_y = \partial f/\partial y$, $g_u = \partial g/\partial u$, $g_v = \partial g/\partial v$.

Let now t be an indeterminate. Consider the polynomial $q(u, v, t) = tg_u - g_v$, and let $a(u, t) = \mathrm{Res}_v(g, q)$. Also, without confusion we will call the set $\{(u, v) \in \mathbf{R}^2 \mid g(u, v) = 0\}$ C. We have:

PROPOSITION 2.3. *Let a be as above. Write $a(u, t) = \Sigma_i a_i(u)t^i$, and consider $A(u) = \gcd_i(a_i(u))$. Then*

1. *C is singular over $\mathbf{R}^2$ if and only if $A(u) = 0$ is solvable in $\mathbf{R}$.*
2. *C is singular over $\mathbf{C}^2$ if and only if $A(u) = 0$ is solvable in $\mathbf{C}$.*

PROOF: 1. Suppose $(u_0, v_0) \in \mathbf{R}^2$ is a singular point of C. Then $g(u_0, v_0) = q(u_0, v_0, t) = 0$ for all t. Therefore $a(u_0, t) \equiv 0$, which implies that u_0 is a real root of $A(u)$. Conversely, let $u_0 \in \mathbf{R}$ be a root of $A(u)$. In that case $a(u_0, t) \equiv 0$, and for each

https://doi.org/10.1017/S0004972700028756 Published online by Cambridge University Press

t we can find a corresponding $v_0^t \in \mathbf{C}$ such that $0 = g(u_0, v_0^t) = tg_u(u_0, v_0^t) - g_u(u_0, v_0^t)$. But since g is regular in v, there can be only finitely many values $v_0^t \in \mathbf{C}$ such that $g(u_0, v_0^t) = 0$. One of these, v_0 say, must therefore satisfy $g(u_0, v_0) = tg_u(u_0, v_0) - g_u(u_0, v_0) = 0$ for at least two distinct t's. This implies that (u_0, v_0) is a singular point of C. Moreover if $v_0 \in \mathbf{C} - \mathbf{R}$, then $(u_0, \bar{v}_0)$ is also a singular point. But condition 2 of the above lemma implies that $v_0 = \bar{v}_0$; that is v_0 must be real.

2. The proof of this case is similar to Case 1. □

REMARK 2.4. The proof indicates that the singular points of C are in one-to-one correspondence with the distinct roots of $A(u)$. Moreover, the real (complex) singular points are in one-to-one correspondence with the distinct real (complex) roots of $A(u)$. Furthermore, we note that by invoking a Sturm-type algorithm we can decide whether $A(u)$ has any real roots and therefore whether C is real singular.

A similar procedure can be applied directly to the original polynomial $f(x, y)$ for testing whether C is singular over $\mathbf{C}^2$. We caution that this test fails to give us a definite answer as to whether C is real singular.

By supposing that $f(x, y)$ is regular in y and keeping the same notation as before let $q(x, y, t) = tf_x - f_y$ and $a(x, t) = \mathrm{Res}_y(f, q) = \Sigma_i a_i(x)t^i$, $A(x) = \underset{i}{gcd}(a_i(x))$. Then a similar argument shows

REMARK 2.5. C is singular over $\mathbf{C}^2$ if and only if $A(x) = 0$ is solvable over $\mathbf{C}$.

3. PRELIMINARIES

3.1 A BRIEF REVIEW.

In this section we state some of the key results taken from [**11, 12**], which are essential to this work. In addition we include a procedure for finding a linear and a quadratic Morse function on a real non-singular curve. We conclude with Proposition 3.3.3 which is one of the basic ingredients of our algorithm. We begin with a definition.

DEFINITION 3.1.1: Let $R(x)$ be a rational function, and $[a, b]$, $a < b$ a closed interval such that $R(a)$ and $R(b)$ are finite. The Cauchy index $I_a^b R$ of $R(x)$ over $[a, b]$ is defined as $I_a^b R = N_+ - N_-$, where N_+ and N_- denote the number of points in (a, b) at which $R(x)$ jumps from $-\infty$ to $+\infty$ and from $+\infty$ to $-\infty$ respectively, as x increases from a to b. By convention $I_a^b R = -I_b^a R$.

EXAMPLE. *Let $P(x)$ be a real polynomial and $[a, b]$ a closed interval so that $P(a)P(b) \neq 0$. Then $I_a^b P'/P$ is simply the number of distinct real roots of $P(x)$ in (a, b). In particular $I_{-\infty}^{\infty} P'/P$ is the number of distinct real roots of $P(x)$.*

REMARK. We note that if $R(x) = s(x)/r(x)$, where $s(x), r(x) \in \mathbf{Q}[x]$ and $a, b \in \mathbf{Q}$, then $I_a^b R$ can algorithmically be computed via the modified Euclidean Algorithm [**6,**

https://doi.org/10.1017/S0004972700028756 Published online by Cambridge University Press

11].

Now consider polynomials $p(x, y)$, $q(x, y)$ over $\mathbf{Q}[x, y]$, with no common factors. Let $F = (p, q)$ and consider a rational rectangle $\Gamma = [a, b] \times [c, d]$, $a < b$, $c < d$ so that no zero of F lies on its boundary $\partial\Gamma$, and $p \cdot q \neq 0$ at its vertices. Set

$$R_3 = \frac{q(x, c)}{p(x, c)},\ R_2 = \frac{q(b, y)}{p(b, y)},\ R_4 = \frac{q(x, d)}{p(x, d)},\ R_1 = \frac{q(a, y)}{p(a, y)},\ \text{and}$$
$$I_\Gamma F = I_a^b R_3 + I_c^d R_2 + I_b^a R_4 + I_d^c R_1.$$

Also, consider the Gauss Map $G = F/\|F\|$, $G : \partial\Gamma \to S^1$, where S^1 is the unit circle and both $\partial\Gamma$ and S^1 carry the counterclockwise orientation. Finally, let d be the degree of G. We have:

PROPOSITION 3.1.2. [11]. *For G, F, Γ, d as above, $d = -I_\Gamma F/2$.*

Let $J = \partial p/\partial x\ \partial q/\partial y - \partial q/\partial x\ \partial p/\partial y$ be the Jacobian determinant of F, and $z_0 = (x_0, y_0) \in \mathbf{R}^2$ be a zero of F. We say that z_0 is non-degenerate if $J(z_0) \neq 0$. Suppose that all zeros of F, which lie in the interior, Int Γ, of Γ, are non-degenerate. Then the above proposition yields the following:

COROLLARY 3.1.3. *Under the above considerations,*

$$\sum_{\substack{F(z_0)=(0,0)\\ z_0 \in \operatorname{Int}\Gamma}} sign(J(z_0)) = -\frac{1}{2} I_\Gamma F.$$

We now proceed with a result concerning signs of algebraic numbers. Let $g(x)$, $G(x) \in \mathbf{Q}[x]$, $[a, b]$, $a < b$ a rational interval isolating a real root x_0 of $g(x)$. We may assume that x_0 is a simple root of $g(x)$ and $G(a)G(b) \neq 0$. Our aim is to determine the sign $(G(x_0))$. First, consider $D = \gcd(g, G)$. Then for $x \in \mathbf{R}$ set

$$V_\infty(x) = \begin{cases} 1 & \text{if } g(x) < 0 \\ 0 & \text{otherwise} \end{cases}, \qquad V_0(x) = \begin{cases} 1 & \text{if } g(x)G(x) > 0 \\ 0 & \text{otherwise} \end{cases},$$

and let I be the following integer,

$$I = V_\infty(a) - V_0(a) - I_a^b \frac{g}{G} + V_0(b) - V_\infty(b).$$

We have:

COROLLARY 3.1.4.

(i) $G(x_0) = 0$ *if and only if* $D(a)D(b) < 0$.

(ii) *If* $D(a) \cdot D(b) \geqslant 0$ *then* $G(x_0) > 0$ $(G(x_0) < 0)$ *if and only if* $I \neq 0$ $(I = 0)$ *respectively.*

https://doi.org/10.1017/S0004972700028756 Published online by Cambridge University Press

PROOF: (ii) Let F be the vector field defined by $F = (g(x), y - G(x))$, and let M be a positive integer so that $\sup_{a \leqslant x \leqslant b} |G(x)| < M$. Also, consider the rectangle $\Gamma = [a, b] \times [0, M]$. First, we observe that $z_0 = (x_0, G(x_0))$ is the only zero of F within the region $a < x < b$. Further, z_0 is non-degenerate since x_0 is a simple root of $g(x)$. It is now easy to see that $I = -I_\Gamma F$. Therefore, as Corollary 3.1.3 shows, $z_0 \in \operatorname{Int}\Gamma$ if and only if $I \neq 0$. □

We close this paragraph with the following. Let $p(x, y)$, $q(x, y)$ be as before and assume that p and q are regular in y. Consider a rational polynomial $g(x)$ and let (a, b) be a rational interval isolating a real root x_0 of $g(x)$. By invoking the idea of negative polynomial remainder sequences [12], a notion similar to the Euclidean Algorithm, and using Corollary 3.1.4 we can do the following:

1. We can locate and count the real roots of $p(x_0, y)$.
2. We can count and locate the common real roots of $p(x_0, y)$ and $q(x_0, y)$.

3.2. THE CONSTRUCTION OF TWO MORSE FUNCTIONS.

For this paragraph only, we shall assume that we are given a real non-singular curve C defined by $C = \{f(x, y) = 0\}$, with the polynomial $f(x, y)$ satisfying conditions 1, 2 of Lemma 2.1.2. That is f is regular in y, and every vertical line $x = x_0$ contains at most one solution of the system $f = f_x = 0$.

Denote by $f_{xx} = \partial^2 f/\partial x^2$, $f_{xy} = \partial^2 f/\partial x \partial y$, $f_{yy} = \partial^2 f/\partial y^2$. Let t be an indeterminate and consider $q = tf_x - f_y$, $a = \operatorname{Res}_y(f, q)$ and $\mathcal{L} = x + ty$. We will first give a sufficient polynomial condition on t so that $h = \mathcal{L} \mid C$ has only non-degenerate critical points.

Let then (x, y) be a critical point of h. Then at (x, y) we have: $1 = \lambda f_x$, $t = \lambda f_y$, $f = 0$, $\lambda \in \mathbf{R}$. Further, (x, y) is non-degenerate if and only if $Q(x, y) \neq 0$, where $Q = f_x^2 f_{yy} + f_y^2 f_{xx} - 2f_x f_y f_{xy}$ [10]. Therefore, if we can eliminate x and y from the system $f = Q = tf_x - f_y = 0$, that will be our condition. A first step in this direction is the following fact:

PROPOSITION 3.2.1. *The polynomials $f(x, y)$ and $Q(x, y)$ have no common factors of positive degree if and only if $f(x, y)$ is linear factor free over $\mathbf{C}[x, y]$.*

PROOF: For a real polynomial $\phi(x, y)$, let us denote by B_ϕ the determinant of the bordered Hessian matrix of ϕ,

$$B_\phi = \det \begin{pmatrix} 0 & \nabla\phi \\ \nabla^t\phi & H(\phi) \end{pmatrix}.$$

Now let $\ell(x, y)$ be a linear factor of $f(x, y)$. Write $f = \ell \cdot g$, $g \in \mathbf{C}[x, y]$. Then a computation shows that $B_f = g^3 \cdot B_\ell + \ell \cdot g_0$, where $g_0 \in \mathbf{C}[x, y]$. But since $B_\ell \equiv 0$ we get that ℓ is a common factor of f and Q (note that $Q = -B_f$).

https://doi.org/10.1017/S0004972700028756 Published online by Cambridge University Press

Conversely, let $\ell(x, y)$ be an irreducible factor of f and Q over $\mathbf{C}[x, y]$. Since f is square free, we can find a point (x_0, y_0) so that $\ell(x_0, y_0) = 0$ and $\partial\ell/\partial y(x_0, y_0) \neq 0$. Using the Implicit Function Theorem we find that $\{\ell(x, y) = 0\}$ is the graph of $y = \psi(x)$ near (x_0, y_0). But the fact that ℓ is a common factor of f and Q implies that $\psi''(x) \equiv 0$, near x_0. Hence $\{\ell = 0\}$ is a straight line near (x_0, y_0). The latter fact implies that ℓ is linear, using the principle of analytic continuation. □

Now let $b(x, y) = \gcd(f, Q)$, and consider $f^* = f/b$, $Q^* = Q/b$, $B^*(x) = \mathrm{Res}_y(f^*, Q^*)$, $a^*(x, t) = \mathrm{Res}_y(f^*, q, y)$, where $q = tf_x - f_y$. Let also $d(x) = gcd(a^*(x, t), B^*(x))$. Then $d(x)$ has no real roots as Proposition 2.3 shows. In addition, consider $a_* = a^*/d$, $B_* = B^*/d$ and finally $\gamma(t) = \mathrm{Res}_x(a_*, B_*)$. We have:

PROPOSITION 3.2.2. *For all real t satisfying $\gamma(t) \neq 0$, $a(x, t) \not\equiv 0$, h has only non-degenerate critical points.*

PROOF: First we note that $\gamma(t) \not\equiv 0$, since a_* and B_* have no common factors. Further we observe that no critical point of h lies on a real line factor of f. Indeed, if L is such a line factor, and (λ_1, λ_2) a non-zero normal vector of L, then h has critical points on l if and only if $t\lambda_1 - \lambda_2 = 0$ on L. Equivalently $tf_x - f_y = 0$. But the latter is a contradiction to $a(x, t)$ being non-zero. Let now (x_0, y_0) be a critical point of h. If $Q(x_0, y_0) = 0$, that would imply $a_*(x_0, t) = B_*(x_0) = 0$, a contradiction to $\gamma(t) \neq 0$. □

We now proceed with the construction of a quadratic Morse function. Let s be another indeterminant and consider the distance function $T(x, y, s, t) = (x - s)^2 + (y - t)^2$. We shal give a sufficient polynomial condition on s and t so that $\Delta = T \mid C$ has only non-degenerate critical points. First, we recall a well-known result of Morse which, roughly speaking, says that there are many Δ's which are Morse functions.

THEOREM 3.2.3. [8]. *For almost all points $(s, t) \in \mathbf{R}^2$, Δ is a Morse function.*

In fact, the following proposition, also due to Morse, gives an explicit characterisation of points (s, t) so that Δ has degenerate critical points.

PROPOSITION 3.2.4. [8] *For (s, t) Δ as above Δ is not a Morse function if and only if (s, t) is a focal point of C.*

Let now (x, y) be a critical point of Δ. Then we note that (x, y) is non-degenerate if and only if

$$p(x, y, s, t) = (x - s)f_y - (y - t)f_x = 0, \quad f(x, y) = 0, \text{ and}$$
$$S(x, y, s, t) = -\|\nabla f\|^2 - 2(x - s)f_y f_{xy} + (x - s)f_x f_{yy} + (y - t)f_y f_{xx} \neq 0. \text{[10]}$$

Define $M(x, y, s) = -\|\nabla f\|^2 f_x + (x - \mu)Q(x, y)$ and let $N(x, s) = \mathrm{Res}_y(f, M)$. The following is one of the key results in constructing a Morse function Δ.

LEMMA 3.2.5. $N(x, s) \not\equiv 0$.

PROOF: We argue by contradiction. Let then $d(x, y)$ be a common factor of f and M of positive degree. Using Proposition 3.2.1 we establish that d consists of linear factors of f and also that

$$d \mid \|\nabla f\|^2 . \tag{1}$$

We consider the following two cases:

(i) $\ell(x, y) = ax + by + c$, $a, b, c \in \mathbf{R}$ is a real linear factor of d. We then write $f = \ell \cdot F$ and observe $\|\nabla f\|^2 = F^2 \cdot \|\nabla \ell\|^2$, whenever $\ell = 0$. But since $\|\nabla \ell\|^2 = a^2 + b^2 \neq 0$, we get a contradiction to (1).

(ii) Suppose $\ell = a + ib$ is a complex linear factor of d, $a = a_1 x + a_2 y + a_3$, $b = b_1 x + b_2 y + b_3$, $a_i, b_i \in \mathbf{R}$. Consider $\|\nabla \ell\|^2 = a_1^2 + a_2^2 - b_1^2 - b_2^2 + 2i(a_1 b_1 + a_2 b_2)$. If $\|\nabla \ell\|^2$ is not zero then again this contradicts (1). Now assume $\|\nabla \ell\|^2 = 0$. The latter implies that the vectors $a_1 + ib_1$ and $a_2 + ib_2$ are perpendicular and have equal lengths. In particular, we get $a_1 b_2 - a_2 b_1 \neq 0$. Since d is a real polynomial, then $\bar{\ell} = a - ib$ is also a factor of d. Further, since ℓ and $\bar{\ell}$ have no common factors, we conclude that $\ell \cdot \bar{\ell}$ divides d. But the latter implies that C must have a singular point, namely the common point of the real lines $a = 0$ and $b = 0$, a contradiction to C being real non-singular. □

Consider now $r(x) = \mathrm{Res}_y(f, f_x)$ and $K(x, s, t) = \mathrm{Res}_y(f, p)$. Using Lemma 2.2 one can easily show the following:

REMARK 3.2.6. If $s_0 \in \mathbf{R}$ so that $r(s_0) \neq 0$ and $N(x, s_0) \not\equiv 0$, then there is no real x_0 such that $K(x_0, s_0, t) \equiv 0$.

Finally if $c(x) = \gcd(K(x, s, t), N(x, s))$, $k = K/c$, $\eta = N/c$ and $\Gamma(s, t) = \mathrm{Res}_x(k, \eta)$ we have:

PROPOSITION 3.2.7. *Let $r(x)$, $N(x, s)$, $\Gamma(s, t)$ be as above. Then for any pair (s_0, t_0) of reals satisfying $r(s_0) \neq 0$, $N(x, s_0) \not\equiv 0$ and $\Gamma(s_0, t_0) \neq 0$, $\Delta = (x - s_0)^2 + (y - t_0)^2 \mid C$ has only non-degenerate critical points.*

PROOF: First we note that $\Gamma(s, t) \not\equiv 0$. Secondly, let (x_0, y_0) be a critical point of Δ. We observe that $f(x_0, y_0) = p(x_0, y_0, s_0, t_0) = 0$ and $M(x_0, y_0, s_0) \neq 0$, since $K(x_0, s_0, t) \not\equiv 0$. Now since $r(s_0) \neq 0$, we note that $f_x(x_0, y_0) \neq 0$. Finally we have $M(x_0, y_0, s_0) = (1/f_x(x_0, y_0)) \cdot S(x_0, y_0, s_0, t_0)$ which in turn says that (x_0, y_0) is non-degenerate. □

3.3. AN APPLICATION.

Throughout this paragraph we assume that we are given a real non-singular curve $C = \{f(x, y) = 0\}$. Let $\Delta = (x - s)^2 + (y - t)^2$ be a Morse function on C, $s, t \in \mathbf{Q}$.

https://doi.org/10.1017/S0004972700028756 Published online by Cambridge University Press

Consider $p(x, y) = (y - t)f_x - (x - s)f_y$ and denote (f, p) by G. Let Γ be a rational rectangle so that all zeros of G are inside Γ. Then we have:

PROPOSITION 3.3.1. *The number of unbounded components of C is equal to $-I_\Gamma G/2$. In particular, if $I_\Gamma G = 0$, then the curve is either empty or consists only of ovals.*

PROOF: Let $z_0 = (x_0, y_0)$ be a real zero of G. Then z_0 is also a critical point of Δ, and conversely. Let $i(z_0)$ denote the Morse index of Δ at z_0, and let J be the Jacobian determinant of G. Then a calculation shows that $\operatorname{sign}(J(z_0)) = (-1)^{i(z_0)}$. Recall (Morse's Lemma) that $\Sigma_{z_0}(-1)^{i(z_0)} = \chi(C)$, where $\chi(C)$ denotes the Euler characteristic of C. The proof now follows since $-I_\Gamma G/2 = \Sigma_{z_0} \operatorname{sign}(J(z_0)) = \Sigma_{z_0}(-1)^{i(z_0)} = \chi(C) =$ the number of unbounded components of C. □

Next, let us denote by h the restriction of the projection map $(x, y) \to x$ on the curve. Suppose that $z_0 = (x_0, y_0)$ is a non-degenerate critical point of h. Consider the vector field $F = (f, f_y)$ and let Γ be a proper rational rectangle isolating z_0. Further let G_1, G_2 denote the graphs of $x - x_0 = (y - y_0)^2$, $x - x_0 = -(y - y_0)^2$ respectively. We have:

PROPOSITION 3.3.2. *Let F, Γ, G_1, G_2 be as above. Then $I_\Gamma F = -2, 2$ if and only if near z_0, C looks like G_1, G_2.*

PROOF: Using the Implicit Function Theorem we can find a differentiable function $x = \phi(y)$ so that $f(\phi(y), y) = 0$ and $\phi'(y_0) = 0$, $\phi''(y_0) \neq 0$ near (x_0, y_0). Therefore (x_0, y_0) is a local minimum (maximum) of h if $\phi''(y_0) > 0$ ($\phi''(y_0) < 0$) respectively, and thus the curve looks like either G_1 or G_2. Moreover, if $i(z_0)$ denotes the Morse index of h at z_0 we see that $i(z_0) = 0$ or 1 according to whether z_0 is a local minimum or maximum. The proof now follows since $-I_\Gamma F/2 = (-1)^{i(z_0)}$. □

Finally, let us consider a critical point $z_0 = (x_0, y_0)$, not necessarily non-degenerate, of h and let G^1, G^2, G^3 denote the graphs of

$$x - x_0 = (y - y_0)^{2k_1}, \quad x - x_0 = -(y - y_0)^{2k_2}, \quad x - x_0 \pm (y - y_0)^{2k_3+1}, \quad k_1, k_2, k_3 \in Z^+.$$

The following proposition provides the basis for the local topology of the curve near z_0.

PROPOSITION 3.3.3. *Let F, Γ, G^1, G^2, G^3 be as above. Then $I_\Gamma F = -2, 2, 0$ if and only if near z_0, C looks like G^1, G^2, G^3 respectively.*

PROOF: Let $x = \phi(y)$ be as in Proposition 3.3.2. Then since f is regular in y there exists an integer k, $k > 0$ so that $\phi^{(j)}(y_0) = 0$ for $1 \leqslant j < k$ and $\phi^{(k)}(y_0) \neq 0$. Therefore if k is even, C looks like G^1 or G^2 according to whether $\phi^{(k)}(y_0)$ is positive or negative, while if k is odd, it looks like G^3. Now we note that $\partial f/\partial y = -\partial f/\partial x \, d\phi/dy$ near z_0. Let Γ_0 be a rectangle containing z_0 so that $\partial f/\partial x \neq 0$ on Γ_0 and Γ_0 is

contained in Γ. Consider the vector field $F_0 = (\phi - x, d\phi/dy)$. It is easy to see that $I_\Gamma F = I_{\Gamma_0} F_0$. A direct calculation shows that $I_{\Gamma_0} F_0 = -2, 2, 0$ according to whether the graph of $x = \phi(y)$ looks like G^1, G^2, G^3. That completes the proof. □

We close this section with a definition.

DEFINITION 3.3.4: Let $z \in C$. We define an integer $I(z)$, called the index of z, as follows:

$$I(z) = \begin{cases} 0 & \text{if } z \text{ is a regular point of } h \\ -\frac{1}{2} I_\Gamma F & \text{if } z \text{ is a critical point of } h. \end{cases}$$

4. THE ALGORITHM

4.1 THE NON-SINGULAR CASE.

Let $\alpha < \beta$ be two consecutive critical values of h, and let k be the number of real roots of $f(\gamma, y)$, where $\alpha < \gamma < \beta$. We first have:

PROPOSITION 4.1.1. [9]. *There exist real continuous functions* $r_1(x), r_2(x), \ldots, r_k(x)$ *over* $[\alpha, \beta]$ *so that if* $\gamma \in (\alpha, \beta)$, $r_1(\gamma) < r_2(\gamma) < \ldots < r_k(\gamma)$ *and* $r_j(\gamma)$ *are the real roots of* $f(\gamma, y)$, *for* $j = 1, \ldots, k$.

It is apparent from the above proposition that $h^{-1}[\alpha, \beta]$ is the union of the graphs of $r_j(x)$ over $[\alpha, \beta]$. Now let $z \in h^{-1}(\alpha)$. According to whether $I(z) = 1, -1, 0$ the point z is the left endpoint of $2, 0, 1$ of the graphs of the $r_j(x)$'s. The same holds for each $z \in h^{-1}(\beta)$ with the numbers $1, -1$ interchanged. Therefore, in order to find the configuration of C it is enough to do the following:

(1) Determine the number of common points, along with their order, of a critical line $x = \delta$ with C and decide which points have index $1, -1$ or 0.

(2) Decide how the graphs of $r_j(x)$'s are glued together at a critical point of non-zero index.

Consider $p(x) = \mathrm{Res}_y(f, f_y)$, $q(y) = \mathrm{Res}_x(f, f_y)$, $v(x) = \mathrm{Res}_y(f, f_x)$. We observe that a critical value δ of h is a real root of $p(x)$. Let then x_0 be a real root of $p(x)$ and let (a, b) be a rational isolating interval of x_0 so that x_0 is the only real root of $p(x) \cdot v(x)$ in $[a, b]$. Moreover, let $[c, d]$ be a rational isolating interval of a real root y_0 of $q(y)$, and let $\Gamma = [a, b] \times [c, d]$, $F = (f, f_y)$. We now compute $I_\Gamma F$. If $I_\Gamma F = 0$, we note that either $(x_0, y_0) \notin C$ or (x_0, y_0) is a point of index 0. If $I_\Gamma F \neq 0$, we consider the following two cases:

A. $I_\Gamma F = -2$. Then (x_0, y_0) is a critical point of h of index 1. Let $y_1 < y_2 < \ldots < y_k$ be the real roots of $f(b, y)$. Let the unique i, $1 \leqslant i \leqslant k - 1$, be such that $r_i(x_0) =$

https://doi.org/10.1017/S0004972700028756 Published online by Cambridge University Press

$r_{i+1}(x_0) = y_0$. Observe, from the choice of b, that $r_i(x)$ is a decreasing function over $[x_0, b]$, while $r_{i+1}(x)$ is increasing. Thus $y_i < y_0 < y_{i+1}$.

Conversely, let i, $1 \leqslant i \leqslant k-1$, be such that $y_i < y_0 < y_{i+1}$.

Consider $r_i(x)$, $r_{i+1}(x)$. Then $r_i(b) = y_i$, $r_{i+1}(b) = y_{i+1}$, and the above argument shows $r_i(x_0) = r_{i+1}(x_0) = y_0$.

B. $I_\Gamma F = 2$. In this case, using a similar argument, we conclude that (x_0, y_0) is the right endpoint of the graphs of $r_j(x)$, $r_{j+1}(x)$, where $r_j(a)$, $r_{j+1}(a)$ are real roots of $f(a, y)$ and $r_j(a) < y_0 < r_{j+1}(a)$, for a unique j.

Let N_b, N_a denote the number of real roots of $f(b, y)$, $f(a, y)$. Further, let $N^1_{x_0}$, $N^{-1}_{x_0}$ be the number of critical points of the form (x_0, y) of index 1, -1 respectively. Then we observe that the number of common points of the line $x = x_0$ with C is equal to $N_b - N^1_{x_0} = N_a - N^{-1}_{x_0}$. Moreover, we may order those points from the respective order of the roots of $f(b, y)$ and $f(a, y)$.

We are now ready to triangulate C. Let α, β, k, $r_j(x)$ be as in Proposition 4.1.1. Define continuous functions $S_1(x) < S_2(x) < \ldots < S_k(x)$ over $[\alpha, \beta]$ so that $S_j(a) = r_j(\alpha)$, $S_j(\beta) = r_j(\beta)$ and $S_j(j) = r_j(\gamma)$, $\gamma \in \mathbb{Q}$, $\alpha < \gamma < \beta$ and $dS_j/dx = 0$ on $(\alpha, \gamma) \cup (\gamma, \beta)$, $j = 1, \ldots, k$. Evidently the graphs of S_j's triangulate $h^{-1}[\alpha, \beta]$. Let now $x_1 < x_2 < \cdots < x_m$ be the real roots of $p(x)$ and consider integers $w_0 < x_1$, $w_{m+1} > x_m$. If $m \geqslant 1$, we triangulate C by repeating the above procedure over the intervals $[w_0, x_1]$, $[x_i, x_{i+1}]$, $[x_m, w_{m+1}]$, $i = 1, \ldots, m-1$. On the other hand, if $m = 0$, we triangulate C using the intervals $[-1, 0]$, $[0, 1]$. We summarise in the following:

PROPOSITION 4.1.2. *Using the above procedure C has been triangulated. Every oval becomes a closed finite polygon and every unbounded component a "broken" line.*

Finally, let A be an oval of C, which is identified with a simple closed polygon. Consider $q \in \mathbb{R}^2$ and let L^s_q be a semi-line starting at q and intersecting A transversely. Further, let A_q be the number of common points of A and L^s_q. We observe:

(i) q is inside $A \Leftrightarrow A_1 \equiv 1 (\mathrm{mod}\ 2)$;

(ii) q is outside $A \Leftrightarrow A_1 \equiv 0 (\mathrm{mod}\ 2)$.

Noting that there is always a vertical rational line which intersects A transversely, it is now apparent how to decide the relative position of two ovals.

4.2. THE SINGULAR CASE.

Let x_0 be a real root of $p(x)$ and let (a, b) be a rational isolating interval of x_0. Then by invoking Remark 3.1.5 we can isolate the common points of C and the line $x = x_0$. Let then (x_0, y_0) be such a point and $\Gamma = [a, b] \times [c, d]$ an isolating rectangle of (x_0, y_0). Since f is regular in y we may shrink the integval $[a, b]$ if necessary, so

https://doi.org/10.1017/S0004972700028756 Published online by Cambridge University Press

that C has no common points with the line segments $c \times [a, b]$ and $d \times [a, b]$. Now we consder the polynomial $f(b, y)$ and let k_b be the number of its real roots inside (c, d). Then we observe that (x_0, y_0) is the left endpoint of exactly k_b graphs of the $r_j(x)$'s. Similarly, if k_a is the number of real roots of $f(a, y)$, (x_0, y_0) is the right endpoint of k_a graphs of the $r_j(x)$'s. Finally, we note that if $k_a = k_b = 0$ then (x_0, y_0) is an isolated singular point of C.

5. The computing time

In this section we calculate the computation time of our method in the case of a non-singular curve. We begin with some well-known notions and results.

Let $k \in \mathbf{Z}$, $p/q \in \mathbf{Q}$, $(p, q) = 1$. We define the size of k, p/q to be $\log k$ and $\log p + \log q$ respectively. Let d be the maximum coefficient size of $f(x, y)$ and n its degree. Then f_x, f_y have degree $n - 1$ and coefficient size $O(\log n + d)$. If γ is a rational number of size d_1, $f(x, \gamma)$, $f(\gamma, y)$ have degree n and coefficient size $nd_1 + d$. Resultants of f, f_y, f_x with respect to x or y have degree $O(n^2)$ and coefficient size $O(n(d + \log n))$. Evaluating a univariate polynomial of degree m and coefficient size δ at a point of size δ_1 takes time $O(m(\delta + \delta_1))$. Evaluation of its Sturm Sequence takes time $O\left(m^2(\delta + \log m + \delta_1)\right)$ and its roots can be isolated in time $O\left(m^4(\log m + \delta)^2 \log m\right)$. The total size of the endpoints of the separation intervals for the roots is $O(m(\log m + \delta))$. [4, 5, 9].

Recall that $p(x) = \mathrm{Res}_y(f, f_y)$, $q(y) = \mathrm{Res}_x(f, f_y)$, $v(x) = \mathrm{Res}_y(f, f_x)$. Note that $p(x) \cdot v(x)$, $q(y)$ have degree $O(n^2)$ and coefficient size $O(n(\log n + d))$. Hence, the time required to find rational isolating intervals $[a, b]$, $[c, d]$ for the roots of $p(x) \cdot v(x)$, $q(y)$ is $O[(n^2)^4 (\log(n^2) + n(\log n + d))^2 \log n] = O\left(n^{10}(d + \log n)^2 \log n\right)$. The total size of the endpoints of these intervals is $O\left(n^3(\log n + d)\right)$. Now let δ_a be the size of a, and consider the polynomial $f(a, y)$. It has degree n and coefficient size $O(n\delta_a + d)$. Therefore its roots can be isolated in time $O\left(n^4(\log n + n\delta_a + d)^2 \log n\right)$. Since $\Sigma_a \delta_a = O\left(n^3(d + \log n)\right)$ we observe that the roots of all $f(a, y)$ can be isolated in time $O[\Sigma\left(n^4(\log n + n\delta_a + d)^2 \log n\right)] = O\left(n^{12}(d + \log n)^2 \log n\right)$. Finally if $\Gamma = [a, b] \times [c, d]$, $F = (f, f_y)$, $I_\Gamma F$ can be computed in time $O\left(n^6(d + \log n)\right)$. We summarise in the following:

Proposition 5.1. *Let n, d be the degree and the maximum coefficient size of a polynomial $f(x, y)$, defining a real non-singular curve $C = \{f = 0\}$. We can find the configuration of the components of C in time $O\left(n^{12}(d + \log n)^2 \log n\right)$.*

6. Illustrative examples

The SCRATCHPAD II computer algebra system was used for the following exam-

https://doi.org/10.1017/S0004972700028756 Published online by Cambridge University Press

ples.

EXAMPLE 1: Consider the degree four curve defined by

$$f(x, y) = -y^4 + 4xy^3 + (-6x^2 + 8)y^2 + (4x^3 - 16x)y + 7 = 0.$$

It is verified that $f(x, y)$ satisfies conditions 1, 2 of Lemma 2.2 and moreover, the curve is real non-singular. Furthermore, it consists of two ovals exterior to each other, and two unbounded components. Its triangulation is shown in Figure 1.

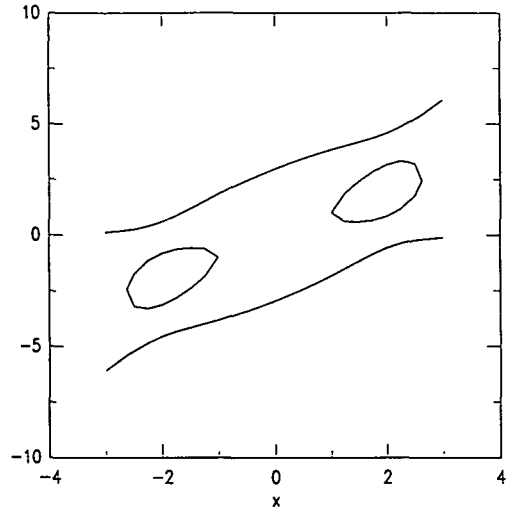

Figure 1

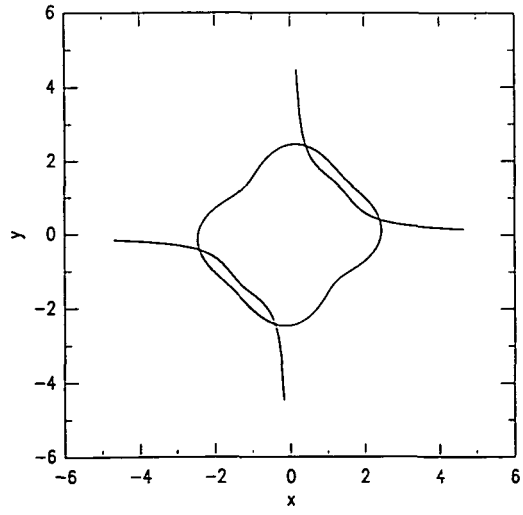

Figure 2

https://doi.org/10.1017/S0004972700028756 Published online by Cambridge University Press

EXAMPLE 2: Let the degree six curve C be defined by

$$C = \{3(x^5y + xy^5) + 10x^3y^3 - 2(x^4 + y^4) - 12x^2y^2 - 23(x^3y + xy^3) + 11(x^2 + y^2) + 34xy + 6 = 0\}.$$

A calculation shows that C is real singular, and its configuration is shown in Figure 2.

REFERENCES

[1] S. Arnborg and H. Feng, 'Algebraic decomposition of regular curves', *J. Symbolic Comput.* **5** (1988), 131–140.

[2] D.S. Arnon, 'Topologically reliable display of algebraic curves', *Computer Graphics* **17** (1983), 219–227.

[3] D.S. Arnon and S. McCallum, 'A polynomial-time algorithm for the topological type of a real algebraic curve', *J. Symbolic Comput.* **5** (1988), 213–236.

[4] G.E. Collins and R. Loos, 'Real zeros of polynomials', *Comput. Suppl.* **4** (1982), 83–94.

[5] J.H. Davenport, 'Computer algebra for cylindrical algebraic decomposition', TRITA-NA-8511, (1985), The Royal Institute of Technology.

[6] F.R. Gantmacher, *The theory of matrices*, Vols. I, II (Chelsea, 1960).

[7] P. Gianni and C. Traverso, 'Shape determination for real curves and surfaces', Manuscript (1983).

[8] J.W. Milnor, *Morse theory* **51** (Annals of Math Studies, Princeton University Press).

[9] D. Prill, 'On approximation and incidence in cylindrical algebraic decompositions', *SIAM J. Comput.* **15** (1986), 972–993.

[10] T. Sakkalis, *An algorithmic application of Morse theory to real algebraic geometry*, Ph.D. Dissertation (University of Rochester, 1986).

[11] T. Sakkalis, 'The Euclidean algorithm and the degree of the Gauss map', *SIAM J. Comput.* **19** (1990), 538-543.

[12] T. Sakkalis, 'On the zeros of a polynomial vector field', *IBM TR, RC 13303* (1987).

[13] T. Sakkalis and R. Farouki, 'Singular points of algebraic curves', *J. Symbolic Comput.* **9** (1990), 405–421.

Department of Mathematical Sciences
Oakland University
Rochester MI 48309-4401 United States of America