Real Plane Algebraic Curves*

Maria Jesus de la Puente

Departamento de Algebra, Facultad de Matemáticas, Universidad Complutense, 28040 Madrid, Spain

Abstract

We study real algebraic plane curves, at an elementary level, using as little algebra as possible. Both cases, affine and projective, are addressed. A real curve is infinite, finite or empty according to the fact that a minimal polynomial for the curve is indefinite, semi-definite nondefinite or definite. We present a discussion about isolated points. By means of the P operator, these points can be easily identified for curves defined by minimal polynomials of order bigger than one. We also discuss the conditions that a curve must satisfy in order to have a minimal polynomial. Finally, we list the most relevant topological properties of affine and projective, complex and real plane algebraic curves.

1 Introduction

What are the qualitative differences between the following subsets of \mathbb{R}^2 :

$$C_j = \{(x, y) \in \mathbb{R}^2 : f_j(x, y) = 0\},\$$

where $f_1(x, y) = -x^2 + y^2 - x^3$, $f_2(x, y) = x^2 + y^2 - x^3$, $f_3(x, y) = x + y^2 - x^3$, $f_4(x, y) = x^2 + y^4(y-1)^6$, $f_5(x, y) = -x^2 - y^4(y-1)^6 - 1$, and $f_6(x, y) = x^2 + 11y^2 - 10$?

This paper contains results and techniques to prove some properties of these sets, such as

- The sets C_1 , C_2 , C_3 and C_6 are infinite, C_4 consists of two points and C_5 is empty.
- The point (0,0) belongs to C_1 , C_2 , C_3 and C_4 ; it is isolated in C_2 and C_4 but not so in C_1 and C_3 .

^{*}Partially supported by DGICYT PB98-0756C-02-01.

E-mail address: MariaJesus_delaPuente@mat.ucm.es



Figure 1: $-X^2 + Y^2 - X^3 = 0$

• The sets C_4 and C_6 are bounded, while C_1 , C_2 and C_3 are not.

See figures 1, 2 and 3, for C_1 , C_2 and C_3 and note that C_6 is an ellipse.



Figure 2: $X^2 + Y^2 - X^3 = 0$

These and other topological and geometrical properties of the sets C_j partially depend on properties of the polynomials f_j and, very particularly, on whether f_j has a constant sign when evaluated at points of \mathbb{R}^2 . Let us say that f_j is *indefinite* if there exist points x and y in \mathbb{R}^2 such that $f_j(x) < 0 < f_j(y)$ and that f_j is *semi-definite* otherwise. It is easy to check that

- f_j is indefinite for $j \in \{1, 2, 3, 6\}$ since, say $f_j(3, 0) < 0 < f_j(0, 1)$.
- $f_4(x,y) \ge 0$ and $f_5(x,y) < 0$, for all $(x,y) \in \mathbb{R}^2$.



Figure 3: $X + Y^2 - X^3 = 0$

In this paper we study sets like the C_j above, called *real algebraic plane affine curves*. A real algebraic plane affine curve is the zero-set of one nonconstant real polynomial in two variables. Now the first question arises: is it fair to say that C_4 and C_5 are curves?

One can plot real plane curves with the help of a computer, in order to visualize them: there are several computer programs which plot any polynomial equation in two variables. However, to any good observer further questions will come up. Indeed, these computer plots are somewhat unreliable. For instance, one cannot be totally sure that a curve is empty, based on the fact that the plot on the screen looks empty. Also, plots are not very precise near singular points of curves. (In fact, the figures shown in this paper have not been plotted by a computer, merely using their equations, due to the lack of accuracy near singular points.) These facts make it necessary to have some theoretical results about real plane algebraic curves at hand. However, every elementary text on algebraic curves that we know of gives up the study of real curves at a very early stage.

Our favorite book on plane (complex) algebraic curves is [18]. Based on it, we have searched in the literature on real algebraic sets, for statements about real curves analogous to those found in the first chapters of [18]. Sometimes our search has been fruitless but we have been able to provide a result ourselves. The result is this paper, where we present a theory of real plane algebraic curves, at an elementary level. Both cases, affine and projective, are addressed. It is often better to work with projective curves, for global problems, and with affine ones, for local problems. Being a projective curve the closure of an affine one, we know how to draw conclusions for the affine case from the projective case and conversely. For the study of a real curve, affine or projective, it is always useful to examine the curve first over the field of complex numbers.

The material presented here comes from various sources; it includes:

- results already published by different authors and generally well-known. These go from Whitney's example up to the *real Study's lemma*, and again from the *sign change criterion* up to the crucial theorem 25.
- examples, which are either well-known or easy to propose.

• results due to the author. These go from corollary 14 up to lemma 16, lemma 19, the *real projective Study's lemma*, and from corollary 26 up to the end of the second section. Some of these are just real versions of results known for complex curves. The author should also be credited for putting all this material together.

Some true statements about the zero-set, say in the affine complex plane \mathbb{C}^2 , of a complex polynomial $f \in \mathbb{C}[X, Y]$ also hold true for the zero-set in \mathbb{R}^2 of a real polynomial $f \in \mathbb{R}[X, Y]$, as long as f is indefinite; see corollary 9 and theorems 12 and 15. This means that the zero-set of a semi-definite real polynomial may have unexpected properties. Indeed, a definite polynomial has empty zero-set and the converse is also true, by continuity. Now let us think of semi-definite nondefinite polynomials $f \in \mathbb{R}[X, Y]$ and, for the time being, let us consider only polynomials without multiple factors, since they are easier to deal with. In corollary 26 we show that the zero-set of f is finite nonempty.

These results are known, in far more generality, to real algebraic geometers. And only to them, mostly. Our intention has been to present real algebraic plane curves, in elementary terms, to a broader audience, including students. Why do most algebraic geometers like to work over an algebraically closed field, such as C? It is essentially due to the fundamental theorem of algebra and its many consequences.

In this paper, we use as little algebra as possible: just basic properties of polynomials (such as: order and degree, irreducibility, including some irreducibility criteria, unique factorization, greatest common divisor, Gauss's lemma and Euclid's algorithm). We do not mention ideals, although some are behind the curtains, whenever we talk about minimal polynomials. Further, three more tools suffice. First, the *continuity* of polynomials as functions on \mathbb{R}^2 , and consequences of this, such as the *intermediate value theorem*. Second, the *character* (indefinite, semi-definite or definite) of polynomials, and third, a very mild version of the *implicit function theorem*.

We give a number of examples and it is possible to apply the techniques presented here to other instances. Of course, only those of low degree will be feasible to work by hand.

These notes are organized as follows.

The second section starts with the basic properties of affine algebraic sets and, in particular, of curves. After a couple of algebraic lemmas, we come to the properties of real affine curves. Next, we recall some facts of complex projective curves. Only then can we study real curves (affine and projective). This is so because the proof of theorem 25, which is crucial in our presentation, is based on a well-known upper bound for the number of singular points of an irreducible complex projective curve. Latter on, we present a discussion about isolated points. The question of whether a point in a real curve (affine or projective) is isolated can be completely settled down by means of Puiseux expansions, rational over \mathbb{R} , see [5]. We do not go into this topic, because it would make this paper too long. Instead, we will obtain some partial results, using elementary methods. At the end of this section we define, in very elementary algebrogeometric terms, the class \mathcal{F} of all plane real algebraic curves having a *minimal polynomial* and henceforth, admitting a notion of *degree*. These real curves behave like complex curves, in many respects. It may not be easy to determine by hand if a given real algebraic curve $\mathcal{C} = V_{\mathbb{R}}(f)$ belongs to the class \mathcal{F} , for a given polynomial f with rational coefficients, if the

degree of f is big; however this can be decided with a computer. The details of this assertion will be developed elsewhere.

In the third section we have gathered the most relevant topological properties of affine and projective, complex and real plane algebraic curves. This information is well known, though scattered in the literature. For most statements there, we have provided easy examples as well as references for the proofs, which are beyond the scope of this paper.

The bibliography splits into two parts: the first 14 items are the references quoted in this paper, whilst the latter items are some standard updated references on algebraic curves.

These notes owe a lot to [3] and [18].

An earlier, somewhat different version of this paper is [10].

Acknowledgements

We take the opportunity to express our gratitude to our colleagues C. Andradas and J.M. Gamboa, who made very valuable suggestions. We are also thankful to M.E. Alonso, for her interest.

2 Affine and projective curves: algebraic aspects

Let \mathbb{K} denote either \mathbb{C} or \mathbb{R} . Let $n \in \mathbb{N}$. An algebraic set in the affine space \mathbb{K}^n is any set of the form

$$\{(x_1,\ldots,x_n)\in\mathbb{K}^n: f_1(x_1,\ldots,x_n)=0,\ldots,f_s(x_1,\ldots,x_n)=0\}$$

for some $s \in \mathbb{N}$ and some polynomials $f_1, \ldots, f_s \in \mathbb{K}[X_1, \ldots, X_n]$. It is denoted $V_{\mathbb{K}}(f_1, \ldots, f_s)$. It is also called the *set of zeros* (or *zero-set*) of f_1, \ldots, f_s in \mathbb{K}^n . We mostly use n = 2 or 3, in which case we prefer to use variables X, Y or X, Y, Z.

It is clear that

$$\mathsf{V}_{\mathbb{K}}(f_1,\ldots,f_s)=\mathsf{V}_{\mathbb{K}}(f_1)\cap\ldots\cap\mathsf{V}_{\mathbb{K}}(f_s).$$

Here are a few basic properties satisfied by algebraic sets. They all are very easy to check.

- 1. $V_{\mathbb{K}}(f) = \mathbb{K}^n$ if and only if f = 0,
- 2. $V_{\mathbb{K}}(c) = \emptyset$, for all $c \in \mathbb{K} \setminus \{0\}$,
- 3. $V_{\mathbb{K}}(f_1f_2) = V_{\mathbb{K}}(f_1) \cup V_{\mathbb{K}}(f_2)$. In particular, if f divides h then $V_{\mathbb{K}}(f) \subseteq V_{\mathbb{K}}(h)$.
- 4. Every finite subset of \mathbb{K}^n is algebraic.

An algebraic set C in \mathbb{K}^n is called *irreducible* if it cannot be decomposed into the union of proper algebraic subsets, i.e., $C = C_1 \cup C_2$ implies $C = C_1$ or $C = C_2$, whenever C_1 and C_2 are algebraic sets in \mathbb{K}^n .

An affine plane algebraic curve over \mathbb{K} is, by definition, the set of zeros in \mathbb{K}^2 of just one nonconstant polynomial $f \in \mathbb{K}[X, Y]$. We say that f vanishes (precisely) on C. A curve $V_{\mathbb{K}}(f)$ is irreducible if it is irreducible as an algebraic set.

This is the naivest definition of curve possible. At a later stage, the definition of curve may be extended by allowing multiplicities. This is needed, for instance, when one studies families of curves, such as linear systems. We do not do it here, for these notes are merely introductory.

Here is a known theorem:

Every complex plane algebraic curve $\mathcal{C} \subset \mathbb{C}^2$ has infinitely many points.

This result is very easy to deduce from the fundamental theorem of algebra, as follows: we intersect the curve C with an infinite family of complex parallel lines and then we notice that there exist points in every intersection, except for possibly finitely many lines.

Here is another known theorem:

Every complex plane algebraic curve $\mathcal{C} \subset \mathbb{C}^2$ has a *minimal* polynomial f.

By this, we mean that every polynomial vanishing precisely on C is a multiple of f. Such an f is unique up to multiplication by nonzero constants and we write $C = V_{\mathbb{C}}(f)$. The degree of f is also called the *degree of* C. A famous result follows:

Bezout's theorem: Two complex plane algebraic curves C and D without common irreducible components, of degrees m and n respectively, intersect in j different points, with $0 < j \le mn$. In fact, if the curves are projective and the intersection points are counted with an adequate multiplicity, then C and D intersect in exactly mn points.

The theorem is named after the French mathematician Étienne Bezout (1739–1783). See [18] p. 25, for the proof.

Now, let us look at some examples over \mathbb{R} .

Example. 1

- 1. This is taken from [14]. For each $c \in \mathbb{K}$, set $f_c = X^2 + 2cXY + Y^2 \in \mathbb{K}[X, Y]$. Then the curve $V_{\mathbb{R}}(f_c) \subseteq \mathbb{R}^2$ is either the point (0, 0), or a line or the union of two lines, according to |c| < 1, |c| = 1 or |c| > 1, respectively. To check it, just notice that $f_c = (X + cY)^2 + (1 - c^2)Y^2$. On the other hand, $V_{\mathbb{C}}(f_c) \subseteq \mathbb{C}^2$ is the union of two lines (possibly equal), for all $c \in \mathbb{C}$. See figure 4, for c = 2.
- 2. The empty set is a real algebraic curve. It is the zero-set of, say, $Y^2 + 1$ or $X^2 + (X-1)^2 \in \mathbb{R}[X, Y]$. However, both $\mathbb{V}_{\mathbb{C}}(Y^2 + 1)$ and $\mathbb{V}_{\mathbb{C}}(X^2 + (X-1)^2)$ are unions of two parallel lines.

We continue with some algebra, on which we base certain arguments; see [13] for more details. Given two polynomials $f, h \in \mathbb{K}[X, Y]$, a greatest common divisor (g.c.d.) of f and h exists in $\mathbb{K}[X, Y]$, by unique factorization. It is unique up to multiplication by nonzero constants. If factorizations of f and h into irreducible factors are known, then the product of the common factors of f and h is a g.c.d. of f and h. This is, of course, well known.

We may think of elements in $\mathbb{K}[X, Y]$ as polynomials in Y having coefficients in $\mathbb{K}[X]$. Concerning g.c.d.'s, we can do even better by working in the bigger ring $\mathbb{K}(X)[Y]$, where the coefficients are rational expressions in X: that is, we allow denominators which are polynomial



Figure 4: $X^2 + 4XY + Y^2 = 0$

expressions in X. Being $\mathbb{K}(X)$ a field, we can use Euclid's algorithm to express the g.c.d. of f and h as a sum lf + mh, for some $l, m \in \mathbb{K}(X)[Y]$. This expression *cannot* be obtained in $\mathbb{K}[X, Y]$, in general.

Back in $\mathbb{K}[X, Y]$, we say that a nonzero polynomial $f = a_0(X) + a_1(X)Y + \cdots + a_d(X)Y^d$ is primitive (in Y) if its coefficients $a_0(X)$, $a_1(X)$, ..., $a_d(X)$ are coprime. For example, if fis a monomial, then f is primitive in Y if and only if f does not depend on X. It is an useful result, easy to prove, that the product of primitive polynomials is primitive. This is known as *Gauss's lemma*.

Algebraic Lemma. 2 Let $f \in \mathbb{K}[X,Y]$ be a polynomial of positive degree. If f equals a product gh with $g,h \in \mathbb{K}(X)[Y]$, then there exist $\tilde{g}, \tilde{h} \in \mathbb{K}[X,Y]$ so that $f = \tilde{g}\tilde{h}$ and the degrees with respect to Y of g and \tilde{g} are equal.

PROOF. Let us look at the degree of f with respect to Y. If $\deg_Y(f) = 0$ then f does not depend on Y and the same holds for g and h, since $\deg_Y(f) = \deg_Y(g) + \deg_Y(h)$. In this case the result follows from unique factorization in $\mathbb{K}[X]$.

Now, if $\deg_Y(f) > 0$, then we write $g = ag^*$, with $a \in \mathbb{K}(X)$ and $g^* \in \mathbb{K}[X, Y]$ primitive in Y. Similarly, we write $h = bh^*$, with $b \in \mathbb{K}(X)$ and $h^* \in \mathbb{K}[X, Y]$ primitive in Y. Further, we write $f = cf^*$, where $c \in \mathbb{K}[X]$ is a g.c.d of the coefficients of f in $\mathbb{K}[X]$ and $f^* \in \mathbb{K}[X, Y]$ is primitive in Y. This expression is *unique* up to multiplication by nonzero constants from \mathbb{K} . By hypothesis we have $cf^* = abg^*h^*$ and Gauss's lemma tells us that g^*h^* is primitive. Then the uniqueness claimed right above yields $ab \in \mathbb{K}[X]$. Thus we may take, say, $\tilde{g} = abg^*$ and $\tilde{h} = h^*$.

Algebraic Corollary. 3 Let $f, h \in \mathbb{K}[X, Y]$ be both of positive degree and coprime. Then there exist $d \in \mathbb{K}[X] \setminus \{0\}$ and $l', m' \in \mathbb{K}[X, Y]$ such that d = l'f + m'h.

PROOF. Assume that f and h are coprime in $\mathbb{K}[X, Y]$. It follows from the algebraic lemma 2 that f and h remain coprime in $\mathbb{K}(X)[Y]$, so that there exist $l, m \in \mathbb{K}(X)[Y]$ with 1 = lf + mh. Removing denominators we get $d \in \mathbb{K}[X] \setminus \{0\}$ such that d = l'f + m'h, with l' = dl and $m' = dm \in \mathbb{K}[X, Y]$.

Algebraic Lemma. 4 Let $f, h \in \mathbb{K}[X, Y]$ be both of positive degree and coprime. Then $\mathbb{V}_{\mathbb{K}}(f) \cap \mathbb{V}_{\mathbb{K}}(h)$ is finite, possibly empty.

PROOF. Assume that f and h are coprime in $\mathbb{K}[X, Y]$. By corollary 3, we get $d \in \mathbb{K}[X] \setminus \{0\}$ such that d = l'f + m'h, with $l', m' \in \mathbb{K}[X, Y]$. If (x_0, y_0) lies in $\mathbb{V}_{\mathbb{K}}(f) \cap \mathbb{V}_{\mathbb{K}}(h)$ then $d(x_0)$ vanishes, i.e., x_0 is a root of the polynomial d, so that only finitely many x_0 qualify. Now reproducing the argument on $\mathbb{K}(Y)[X]$, we prove that only finitely many y_0 qualify. Altogether, only finitely many (x_0, y_0) may lie in $\mathbb{V}_{\mathbb{K}}(f) \cap \mathbb{V}_{\mathbb{K}}(h)$.

Another proof of this algebraic lemma can be given, using resultants.

Now set $\mathbb{K} = \mathbb{R}$. The following theorem doesn't hold if \mathbb{R} is replaced by \mathbb{C} .

Theorem. 5 Except for \mathbb{R}^2 itself, every algebraic set in \mathbb{R}^2 is a real algebraic curve.

PROOF. We have $\mathbb{V}_{\mathbb{R}}(f_1, \ldots, f_s) = \mathbb{V}_{\mathbb{R}}(f_1) \cap \ldots \cap \mathbb{V}_{\mathbb{R}}(f_s) = \mathbb{V}_{\mathbb{R}}(f)$, with $f = f_1^2 + \cdots + f_s^2$.

Corollary. 6 Every finite set in \mathbb{R}^2 is a real algebraic curve.

We can easily find the expression of a polynomial vanishing on a given finite set of points $(a_1, b_1), \ldots, (a_n, b_n) \in \mathbb{R}^2$. We choose, as in example 1, $g_c = \prod_{j=1}^n [(X - a_j)^2 + 2c(X - a_j)(Y - b_j) + (Y - b_j)^2]$, for any $c \in \mathbb{R}$ with |c| < 1. Then we have $\mathbb{V}_{\mathbb{R}}(g_c) = \{(a_1, b_1), \ldots, (a_n, b_n)\}$.

Now we introduce the *character* of a real polynomial. Later we only use n = 2 or 3.

Definition. 7 The polynomial $f \in \mathbb{R}[X_1, \ldots, X_n]$ is indefinite if there exist $a, b \in \mathbb{R}^n$ such that f(a) < 0 < f(b). Otherwise, f is positive definite if f(a) > 0, for every $a \in \mathbb{R}^n$, negative definite if f(a) < 0, for every $a \in \mathbb{R}^n$, positive semi-definite if $f(a) \ge 0$, for every $a \in \mathbb{R}^n$ or negative semi-definite if $f(a) \le 0$, for every $a \in \mathbb{R}^n$ or

By the character of a polynomial we mean one of the following three mutually exclusive conditions:

- indefinite,
- semi-definite nondefinite,
- definite.

Note that the character of a polynomial does not change by an affine/projective change of coordinates. In lemma 19 we will see that it (almost) remains unchanged by homogeneization.

Let us say a few words on positive semi-definite polynomials. An easy example of such is $X^2 + Y^2$ or, more generally, any sum of squares of polynomials. Hilbert knew that not every positive semi-definite polynomial is a sum of squares of polynomials, although he had no explicit example. In his famous collection of problems given in the 1900 International Congress of Mathematicians, he proposed the following generalization as the 17^{th} problem:

Is every positive semi-definite polynomial equal to a sum of squares of rational expressions?

The problem was solved by E. Artin in 1927, on the affirmative. The first examples of positive semi-definite polynomials which are not sums of squares of polynomials are due to T.S. Moztkin, and were published in 1967. R. Robinson gave more examples in 1969. These and other examples as well as the references can be found in [11]. Positive semi-definite polynomials which are not sum of squares of polynomials are rare; however, there is a lot of literature on the subject. In this paper, every example of positive semi-definite polynomial is actually a sum of squares of polynomials, for easiness.

First, we will study curves defined by indefinite polynomials.

Lemma. 8 If $f \in \mathbb{R}[X, Y]$ is indefinite then, after an affine change of coordinates, there exists an open interval $\emptyset \neq I \subseteq \mathbb{R}$ such that, for each $u \in I$ there exists $b_u \in \mathbb{R}$ with $(u, b_u) \in V_{\mathbb{R}}(f)$.

PROOF. An affine change of coordinates allows us to assume that there exist a and $b_1 < b_2$ in \mathbb{R} such that $f(a, b_1) < 0 < f(a, b_2)$. By continuity, there exists $\delta > 0$ such that $f(u, b_1) < 0 < f(u, b_2)$, for every u in the open interval $I_{\delta} := (a - \delta, a + \delta)$. By the intermediate value theorem, for each $u \in I_{\delta}$ there exists b_u with $b_1 < b_u < b_2$ such that $f(u, b_u) = 0$, i.e., $(u, b_u) \in V_{\mathbb{R}}(f)$.

Corollary. 9 If $f \in \mathbb{R}[X, Y]$ is indefinite, then $V_{\mathbb{R}}(f)$ is infinite.

So far, indefiniteness looks like a promising property. Here are some sufficient conditions for it.

Lemma. 10 Let $f \in \mathbb{R}[X_1, ..., X_n]$ be of positive degree d. Then f is indefinite in any of the following circumstances:

- 1. if d is odd,
- 2. if there exists $P \in V_{\mathbb{R}}(f)$ such that the order of f at P is odd.

PROOF.

- An affine change of coordinates allows us to assume that 0 ≠ f(0,...,0). For each λ ∈ R, consider the nonzero polynomial g_λ ∈ R[X] given by g_λ(X) := f(X, λX,..., λX). We have deg(g_λ) ≤ deg(f) for every λ ∈ R, with equality for every λ ∈ R \ S, where S ⊆ R is some finite set (indeed, S = {λ : f_(d)(X, λX,..., λX) = 0}, where f = f₍₀₎ + f₍₁₎ + ··· + f_(d), with f_(k) homogeneous of degree k, and f₍₀₎ ≠ 0 ≠ f_(d)). Then g_λ is a polynomial of odd degree in one variable, for each λ ∈ R \ S. Now, it is well known that there exist x_λ, x'_λ ∈ R such that g_λ(x_λ) < 0 < g_λ(x'_λ), and we conclude.
- 2. Similar to part 1.

Lemma. 11 Let $f \in \mathbb{R}[X_1, ..., X_n]$ be of positive degree d. If there exists $P \in V_{\mathbb{R}}(f)$ such that the gradient of f at P does not vanish then f is indefinite.

PROOF. The proof is taken from [2]. Let $P = (p_1, \ldots, p_n)$ and say $\frac{\partial f}{\partial X_1}(P) \neq 0$. Then the function $\mathbb{R} \to \mathbb{R}$ that maps x to $f(x, p_2, \ldots, p_n)$ is strictly monotonous on some neighborhood of p_1 and, by continuity, it must change sign on such a neighborhood, since this function vanishes at p_1 .

Notice that if $f \in \mathbb{R}[X, Y]$ is irreducible then $V_{\mathbb{R}}(f) \cap V_{\mathbb{R}}(f_X) \cap V_{\mathbb{R}}(f_Y)$ is finite. If we further assume that $V_{\mathbb{R}}(f)$ is infinite, then we conclude that f is indefinite, applying the previous lemma.

Here comes a couple of well-known results: first, a sort of converse to property 3 in page 5 and then a consequence of it. The former is named after the German mathematician Eduard Study (1862–1930).

- Study's lemma: If $f, h \in \mathbb{C}[X, Y]$ are both of positive degree such that f is irreducible in $\mathbb{C}[X, Y]$ and $\mathbb{V}_{\mathbb{C}}(f) \subseteq \mathbb{V}_{\mathbb{C}}(h)$, then f divides h.
- *Irreducibility Criterion:* Given $f \in \mathbb{C}[X, Y]$ of positive degree, the affine curve $\mathbb{V}_{\mathbb{C}}(f)$ is irreducible if and only if there exist $g \in \mathbb{C}[X, Y]$ irreducible in $\mathbb{C}[X, Y]$ and $k \in \mathbb{N}$ such that $f = g^k$. In particular, if g is irreducible in $\mathbb{C}[X, Y]$ then $\mathbb{V}_{\mathbb{C}}(g)$ is irreducible.

See [18], pp. 13 and 15 for the easy proofs of these two results. Next, let us prove the following analogs over the real field and notice the indefiniteness requirement on f.

Theorem. 12 (Real Study's lemma). Given $f, h \in \mathbb{R}[X, Y]$ both of positive degree such that f is irreducible in $\mathbb{R}[X, Y]$, indefinite and $\mathbb{V}_{\mathbb{R}}(f) \subseteq \mathbb{V}_{\mathbb{R}}(h)$, then f divides h.

PROOF. This is taken from [8], lemma 6.14. The polynomial f is indefinite, so we may assume that there exist a and $b_1 < b_2$ in \mathbb{R} such that $f(a, b_1) < 0 < f(a, b_2)$. It follows from the algebraic lemma 2 that if f does not divide h in $\mathbb{R}[X, Y]$ then f does not divide h in $\mathbb{R}(X)[Y]$ either and that f remains irreducible in $\mathbb{R}(X)[Y]$. Then f and h must be coprime. By the algebraic corollary 3, there exist $l', m' \in \mathbb{R}[X, Y]$ and $d \in \mathbb{R}[X] \setminus \{0\}$ such that d = l'f + m'h. Moreover, by lemma 8, after an affine change of coordinates, there exists an open interval $\emptyset \neq I \subseteq \mathbb{R}$ such that, for each $u \in I$ there exists $b_u \in \mathbb{R}$ with $(u, b_u) \in V_{\mathbb{R}}(f)$. By hypothesis, it follows that $(u, b_u) \in V_{\mathbb{R}}(h)$ and so d(u) = 0, for every $u \in I$. This implies d = 0, a contradiction.

Actually, the proof just given shows that the following statement holds true.

Theorem. 13 (Real Study's lemma –second version–). Let $f, h \in \mathbb{R}[X, Y]$ be polynomials both of positive degree such that f is irreducible in $\mathbb{R}[X, Y]$ and indefinite. Assume that after a certain projective change of coordinates, there exists an open interval $\emptyset \neq I \subseteq \mathbb{R}$ such that, for each $u \in I$ there exists $b_u \in \mathbb{R}$ with $(u, b_u) \in \mathbb{V}_{\mathbb{R}}(f) \cap \mathbb{V}_{\mathbb{R}}(h)$. Then f divides h.

Corollary. 14 Given $f, h \in \mathbb{R}[X, Y]$ both of positive degree such that f is irreducible in $\mathbb{R}[X, Y]$, indefinite and $V_{\mathbb{R}}(f) \subseteq V_{\mathbb{R}}(h)$. If, moreover h is semi-definite, then f^2 divides h.

PROOF. We may assume that h is positive semi-definite. So far, we have that h = fg, for some $g \in \mathbb{R}[X, Y]$. Take δ , I_{δ} , b_1 , b_2 , u and b_u as in the proof of 12, so that $(u, b_u) \in V_{\mathbb{R}}(f)$. Thus $0 = h(u, b_u) = f(u, b_u)g(u, b_u)$. By assumption, $0 \le h(u, b_1) = f(u, b_1)g(u, b_1)$ and being $f(u, b_1)$ strictly negative, it follows that $g(u, b_1)$ must be negative. Similarly, $0 \le h(u, b_2) =$ $f(u, b_2)g(u, b_2)$ and being $f(u, b_2)$ strictly positive, then $g(u, b_1)$ must be positive. Letting b_1 and b_2 both converge to b_u , we conclude that $g(u, b_u)$ must vanish, for every $u \in I_{\delta}$. If we now apply theorem 13 to f and g, we obtain that f divides g.

In [3] proposition 2.3, this corollary has been proved for a linear polynomial f, while the general case is claimed to be proved in a paper never published.

Next, let us find the relationship between the irreducibility of a real polynomial and the irreducibility of the curve it defines.

Theorem. 15 (Real irreducibility condition). If $g \in \mathbb{R}[X, Y]$ is indefinite and irreducible in $\mathbb{R}[X, Y]$ then $V_{\mathbb{R}}(g)$ is irreducible.

PROOF. Suppose that $V_{\mathbb{R}}(g) = C_1 \cup C_2$, with $C_1 = V_{\mathbb{R}}(h_1, \ldots, h_s)$ and $C_2 = V_{\mathbb{R}}(f_1, \ldots, f_t)$ for some $s, t \in \mathbb{N}$ and some $h_1, \ldots, h_s, f_1, \ldots, f_t \in \mathbb{R}[X, Y]$ of positive degrees. We have $C_1 = V_{\mathbb{R}}(h)$ and $C_2 = V_{\mathbb{R}}(f)$, with $h = h_1^2 + \ldots + h_s^2$ and $f = f_1^2 + \ldots + f_t^2$. Then $V_{\mathbb{R}}(g) = V_{\mathbb{R}}(h) \cup V_{\mathbb{R}}(f) = V_{\mathbb{R}}(hf)$ and by the real Study's lemma we get g|hf. By the irreducibility of g, either g|h or g|f holds. Thus, either $V_{\mathbb{R}}(g) \subseteq V_{\mathbb{R}}(h)$ or $V_{\mathbb{R}}(g) \subseteq V_{\mathbb{R}}(f)$ holds, so that $V_{\mathbb{R}}(g) = C_1$ or $V_{\mathbb{R}}(g) = C_2$.

Below we present examples related to the real irreducibility condition. The polynomials appearing there are irreducible, by the following lemma.

Lemma. 16 If $f = X_n^2 + \phi$ where $\phi \in \mathbb{R}[X_1, \dots, X_{n-1}]$ is a positive semi-definite nonzero polynomial, then f is irreducible in $\mathbb{R}[X_1, \dots, X_n]$.

PROOF. Suppose that $X_n^2 + \phi = (X_n + a)(X_n + b) = X_n^2 + X_n(a + b) + ab$, for some $a, b \in \mathbb{R}[X_1, \dots, X_{n-1}]$. Then a = -b and $\phi = -b^2$. Thus ϕ is also negative semi-definite and so $\phi = 0$, a contradiction.

Example. 17

- Take $g = X^2 + Y^2$ and $V_{\mathbb{R}}(g) = \{(0,0)\}$. Thus, the irreducibility of $V_{\mathbb{R}}(g)$ does not imply the indefiniteness of g.
- Take $g = X^2 + Y^2(Y-1)^2 \in \mathbb{R}[X, Y]$, $V_{\mathbb{R}}(g) = \{(0,0), (0,1)\}$. Thus, a polynomial may be semi-definite and irreducible, having reducible zero-set.

The moral of the results proved so far is that real affine curves $V_{\mathbb{R}}(f)$ with f indefinite have similar properties to complex affine curves. Our next question is

What can be said about $V_{\mathbb{R}}(f)$, if $f \in \mathbb{R}[X, Y]$ is semi-definite?

But, before we can answer it, we must examine projective curves.

If $F_1, \ldots, F_s \in \mathbb{K}[X, Y, Z]$ are homogeneous polynomials (also called *forms*) then the set $\{(x : y : z) \in P^2(\mathbb{K}) : F_1(x, y, z) = 0, \ldots, F_s(x, y, z) = 0\}$ is denoted $\mathbb{V}_{\mathbb{K}}(F_1, \ldots, F_s)$. It is, by definition, an *algebraic set* in the projective plane $P^2(\mathbb{K})$. It is also called the *set of zeros of* F_1, \ldots, F_s in $P^2(\mathbb{K})$. The same notation $\mathbb{V}_{\mathbb{K}}$ to denote affine or projective zero-sets should not be confusing.

It is clear why one works only with homogeneous polynomials in the projective case: since both (x : y : z) and $(\lambda x : \lambda y : \lambda z)$ are homogeneous coordinates of just one point in $P^2(\mathbb{K})$ (when x, y, z are not simultaneously zero and $\lambda \neq 0$) then it is necessary that $F(\lambda x, \lambda y, \lambda z)$ vanishes if and only if F(x, y, z) vanishes, which occurs only when F is homogeneous. It is obvious that the irreducible factors and the partial derivatives of a form are forms. We use capital letters to denote forms.

Properties stated at the beginning of this section for affine algebraic sets as well as definitions of projective (irreducible) algebraic sets and of plane projective algebraic curves are analogous.

We can easily prove the projective versions of theorem 5 and corollary 6: except for $P^2(\mathbb{R})$ itself, every algebraic subset in $P^2(\mathbb{R})$ is a real algebraic curve and every finite subset of $P^2(\mathbb{R})$ is a real curve. Also, the empty set is a real algebraic projective curve.

Example. 18

- $V_{\mathbb{R}}(X^2 + Y^2 + Z^2) = \emptyset$, since (0:0:0) is not a point in $P^2(\mathbb{R})$.
- $V_{\mathbb{R}}(X^2 + Y^2) = \{(0:0:1)\}.$

Let us recall a few well-known facts. Given $f \in \mathbb{K}[X, Y]$ of positive degree d, we produce a form $F \in \mathbb{K}[X, Y, Z]$ with $\deg(f) = \deg(F)$ in a unique fashion, as follows: $F = Z^d f(X/Z, Y/Z)$. Equivalently, $F = f_{(0)}Z^d + \cdots + f_{(d-1)}Z + f_{(d)}$, where $f = f_{(0)} + \cdots + f_{(d-1)} + f_{(d)}$ and $f_{(j)}$ is the homogeneous part of f of degree j or is zero. We call F the homogeneization of f. It is clear that $Z \mid /F$. Conversely, if $F \in \mathbb{K}[X, Y, Z]$ is a form of degree d > 0 and neither Z, nor Y, nor X divide F, then we define $f_3 := F(X, Y, 1) \in \mathbb{K}[X, Y]$, $f_2 := F(X, 1, Z) \in \mathbb{K}[X, Z]$ and $f_1 := F(1, Y, Z) \in \mathbb{K}[Y, Z]$, with $\deg(f_j) = \deg(F)$, j = 1, 2, 3. They are called respectively, dehomogeneization of F with respect to Z, Y and X. It is immediate to verify that if $F \in \mathbb{K}[X, Y, Z]$ is a form of degree d > 0 and neither Z, nor Y, nor X divide F, then F is irreducible if and only if there exists $k \in \{1, 2, 3\}$ such that f_k is irreducible.

The euclidean topology on \mathbb{K}^3 induces a topology on $P^2(\mathbb{K})$ which makes it a connected, compact space. All three mappings $j_3 : \mathbb{K}^2 \to P^2(\mathbb{K})$ with $j_3(x, y) = (x : y : 1), j_2 : \mathbb{K}^2 \to P^2(\mathbb{K})$ with $j_2(x, z) = (x : 1 : z)$ and $j_1 : \mathbb{K}^2 \to P^2(\mathbb{K})$ with $j_1(y, z) = (1 : y : z)$ are injective and have an open dense image, denoted $\mathbb{K}_{X,Y}$, $\mathbb{K}_{X,Z}$ and $\mathbb{K}_{Y,Z}$ respectively. We have

$$P^{2}(\mathbb{K}) = \mathbb{K}_{X,Y} \cup \mathbb{K}_{X,Z} \cup \mathbb{K}_{Y,Z}.$$

If we identify $V_{\mathbb{K}}(f_k)$ with its image by j_k , then for each form $F \in \mathbb{K}[X, Y, Z]$ of positive degree we have

$$\mathsf{V}_{\mathbb{K}}(F) = \mathsf{V}_{\mathbb{K}}(f_3) \cup \mathsf{V}_{\mathbb{K}}(f_2) \cup \mathsf{V}_{\mathbb{K}}(f_1)$$

This is an open covering of $V_{\mathbb{K}}(F)$ and $V_{\mathbb{K}}(f_3) = V_{\mathbb{K}}(F) \cap \{(x:y:z): z \neq 0\}, V_{\mathbb{K}}(f_2) = V_{\mathbb{K}}(F) \cap \{(x:y:z): y \neq 0\}$ and $V_{\mathbb{K}}(f_1) = V_{\mathbb{K}}(F) \cap \{(x:y:z): x \neq 0\}$.

Set $\mathbb{K} = \mathbb{R}$. Using continuity arguments and bearing in mind that $\mathbb{R}^2 = \mathbb{R}_{X,Y}$ is dense in $P^2(\mathbb{R})$, we can easily prove the following lemma, the part 2 of which is mentioned in [11] p. 252.

Lemma. 19 Let $f \in \mathbb{R}[X, Y]$ be of positive degree and $F \in \mathbb{R}[X, Y, Z]$ be its homogeneization. *Then*

- 1. f is indefinite if and only if F is indefinite,
- 2. f is semi-definite if and only if F is semi-definite,
- 3. if f is definite then F is semi-definite, and
- 4. if F is definite then f is definite.

Example. 20 $f = X^2 + 1$ is definite but $F = X^2 + Z^2$ is semi-definite nondefinite.

Combining parts 1 and 2 of lemma 19 with the real Study's lemma and its corollary, we obtain the following.

Theorem. 21 (Real Projective Study's lemma). Given forms $F, H \in \mathbb{R}[X, Y, Z]$ both of positive degree such that F is irreducible in $\mathbb{R}[X, Y, Z]$ and indefinite. Assume that after a certain projective change of coordinates, there exists an open interval $\emptyset \neq I \subseteq \mathbb{R}$ such that, for each $u \in I$ there exists $b_u \in \mathbb{R}^2$ with $(u, b_u) \in \mathbb{V}_{\mathbb{R}}(F) \cap \mathbb{V}_{\mathbb{R}}(H)$. Then F divides H. If, moreover H is semi-definite, then F^2 divides H.

Now, let us recall a very familiar concept. If the degree of $f \in \mathbb{C}[X_1, \ldots, X_n]$ is positive and f does not have multiple irreducible factors, then a point P in $\mathbb{V}_{\mathbb{C}}(f)$ is *singular* if, by definition, the gradient of f at P vanishes. We denote the set of singular points of $\mathcal{C} = \mathbb{V}_{\mathbb{R}}(f)$ by Sing \mathcal{C} . The following is well-known.

Bound for the set of singular points of a plane complex projective curve: An irreducible form $F \in \mathbb{C}[X, Y, Z]$ of positive degree d satisfies

$$|\operatorname{Sing} \mathtt{V}_{\mathbb{C}}(F)| \leq \frac{(d-1)(d-2)}{2}.$$

See [18] p. 40 for a proof. This bound obviously remains true for projective real curves as well as for affine real or complex ones.

Next we obtain a result which is well-known to real algebraic geometers. In the literature one can find different versions of it.

Theorem. 22 (Sign Change Criterion). Let $F \in \mathbb{R}[X, Y, Z]$ be a form of positive degree without multiple irreducible factors. Then F is indefinite if and only if there exists a nonsingular point $P \in V_{\mathbb{R}}(F)$.

PROOF. Let $F = F_1 \cdots F_s$ be a factorization of F, with the F_j 's irreducible and pairwise coprime. Then

$$\operatorname{grad}(F) = \sum_{j=1}^{s} F_1 \cdots F_{j-1} F_{j+1} \cdots F_s \operatorname{grad}(F_j).$$

If a point $P \in V_{\mathbb{R}}(F)$ is nonsingular, then P belongs to $V_{\mathbb{R}}(F_{j_0})$ for just one j_0 ; say $P \in V_{\mathbb{R}}(F_1)$ and $P \notin V_{\mathbb{R}}(F_j)$ for $j \neq 1$. Then $\operatorname{grad}_P(F) = F_2(P) \cdots F_s(P) \operatorname{grad}_P(F_1)$ does not vanish. Conversely, if $P \in V_{\mathbb{R}}(F_1)$ and $P \notin V_{\mathbb{R}}(F_j)$ for $j \neq 1$, and if $\operatorname{grad}_P(F) = F_2(P) \cdots F_s(P) \operatorname{grad}_P(F_1)$ does not vanish, then P is nonsingular.

It follows that one implication has already been proved in lemma 11. Now suppose that the form F is indefinite. Then at least one irreducible factor of F, say F_1 , is indefinite. Thus $V_{\mathbb{R}}(F_1)$ is an infinite set, and all of its points but possibly a finite number are nonsingular. Therefore $V_{\mathbb{R}}(F)$ contains a nonsingular point.

Here is real/complex type result on irreducibility, which we will use in the proof of theorem 24.

Lemma. 23 Suppose that $f \in \mathbb{R}[X_1, \ldots, X_n]$ is irreducible in $\mathbb{R}[X_1, \ldots, X_n]$. Then f is reducible in $\mathbb{C}[X_1, \ldots, X_n]$ if and only if either f or -f is a sum of two squares in $\mathbb{R}[X_1, \ldots, X_n]$.

PROOF. This is taken from [3]. If $f = r_1^2 + r_2^2$ with $r_1, r_2 \in \mathbb{R}[X_1, \ldots, X_n]$ then $f = (r_1 + ir_2)(r_1 - ir_2)$, where $i = \sqrt{-1}$. Here, neither factor can be constant, since if one of them belonged to \mathbb{C} then the other one would do too, and f itself would be constant, contradicting the irreducibility of f.

Now assume that f factors nontrivially: $f = (r_1+ir_2)(s_1+is_2)$, with $r_j, s_j \in \mathbb{R}[X_1, \ldots, X_n]$, and j = 1, 2. Taking complex conjugates we get $f^2 = f\overline{f} = (r_1^2 + r_2^2)(s_1^2 + s_2^2)$. By unique factorization in $\mathbb{R}[X_1, \ldots, X_n]$ we have $f = a(r_1^2 + r_2^2)$ for some $a \in \mathbb{R} \setminus \{0\}$. Since either a or -a is a square in \mathbb{R} , the result follows.

Following [3], we consider the function

$$lpha(d):=\max\left\{rac{d^2}{4},rac{(d-1)(d-2)}{2}
ight\}$$

for $d \in \mathbb{N}$. Clearly $\alpha(d) = \frac{d^2}{4}$, if $d \leq 5$ and $\alpha(d) = \frac{(d-1)(d-2)}{2}$, if $d \geq 6$. Taking derivatives, it is easy to check that $\frac{\alpha(d)}{d}$ is monotonically increasing on N. Thus,

$$\frac{\alpha(d_j)}{d_j} \leq \frac{\alpha(d_1+d_2)}{d_1+d_2}$$

for j = 1, 2 so that

$$lpha(d_1) + lpha(d_2) \leq \left(rac{d_1}{d_1 + d_2} + rac{d_2}{d_1 + d_2}
ight) lpha(d_1 + d_2) = lpha(d_1 + d_2).$$

This property of α is used in the proof of theorem 25.

Theorem. 24 Let $F \in \mathbb{R}[X, Y, Z]$ be a semi-definite irreducible form of positive degree d. Then $|V_{\mathbb{R}}(F)| \leq \alpha(d)$.

PROOF. This is taken from [3]. First, assume that F is reducible in $\mathbb{C}[X, Y, Z]$. Then, lemma 23 gives $\pm F = R_1^2 + R_2^2$, for some forms $R_1, R_2 \in \mathbb{R}[X, Y, Z]$ both necessarily of degree $\frac{d}{2}$. Moreover, R_1, R_2 are coprime, since F is irreducible in $\mathbb{R}[X, Y, Z]$. Now, by Bezout's theorem, the set $\mathbb{V}_{\mathbb{R}}(F) = \mathbb{V}_{\mathbb{R}}(R_1) \cap \mathbb{V}_{\mathbb{R}}(R_2)$ contains at most $(\frac{d}{2})^2$ different points.

Suppose now that F remains irreducible in $\mathbb{C}[X, Y, Z]$. By lemma 11, each point of $\mathbb{V}_{\mathbb{R}}(F)$ is singular in $\mathbb{V}_{\mathbb{C}}(F)$ and therefore $|\mathbb{V}_{\mathbb{R}}(F)| \leq \frac{(d-1)(d-2)}{2}$.

Theorem. 25 For any semi-definite form $F \in \mathbb{R}[X, Y, Z]$ of degree $d \ge 2$, the following are equivalent:

- 1. $|\mathbf{V}_{\mathbb{R}}(F)| > \alpha(d)$,
- 2. $V_{\mathbb{R}}(F)$ is infinite,
- 3. F is divisible by the square of some indefinite form.

PROOF. This proof appears in [3]. It is trivial that part 2 implies 1. It follows from corollary 9 and lemma 19 that part 3 implies 2.

Now, we prove that part 1 implies 3, by induction on d. We may assume that F is positive semi-definite. For d = 2 we have $\alpha(2) = 1$ so that F is a positive semi-definite quadratic form in three variables having, at least, two real zeros. A projective change of variables allows us to write $F(X, Y, Z) = X^2$, and the conclusion follows.

Suppose now $|\mathbb{V}_{\mathbb{R}}(F)| > \alpha(d)$, for $d \ge 2$. By theorem 24, F is reducible in $\mathbb{R}[X, Y, Z]$. Write $F = F_1 \cdots F_s$ with $s \ge 2$ and $F_j \in \mathbb{R}[X, Y, Z]$ irreducible for each j (here we do not assume that F_i and F_j are coprime, if $i \ne j$). Set $d_j = \deg(F_j)$. First assume, in addition, that every F_j is semi-definite. There must exist j_0 such that $|\mathbb{V}_{\mathbb{R}}(F_{j_0})| > \alpha(d_{j_0})$; otherwise we would have

$$|\mathtt{V}_{\mathbb{R}}(F)| \leq \sum_{j=1}^{\circ} |\mathtt{V}_{\mathbb{R}}(F_j)| \leq \sum_{j=1}^{\circ} \alpha(d_j) \leq \alpha(\sum_{j=1}^{\circ} d_j) = \alpha(d),$$

a contradiction. Say $|V_{\mathbb{R}}(F_1)| > \alpha(d_1)$. Since $d_1 < d$, then the conclusion follows by induction, in this case. If on the other hand, F_{j_0} is indefinite for some j_0 then $V_{\mathbb{R}}(F_{j_0})$ is infinite, by corollary 9 and lemma 19. The conclusion follows now from the real projective Study's lemma.

The following result needs no further proof.

Corollary. 26 If $F \in \mathbb{R}[X, Y, Z]$ is a form of positive degree without multiple irreducible factors, (in particular, if F is irreducible) then

- $V_{\mathbb{R}}(F)$ is infinite if and only if F is indefinite,
- $V_{\mathbb{R}}(F)$ is finite nonempty if and only if F is semi-definite nondefinite, and

• $V_{\mathbb{R}}(F)$ is empty if and only if F is definite.

Moreover, in case that F is semi-definite nondefinite, then $|V_{\mathbb{R}}(F)| \leq \alpha(d)$, where d is the degree of F.

Of course, there are affine versions of the last four results.

Let $F \in \mathbb{R}[X, Y, Z]$ be a form of positive degree and

$$F = cF_1^{k_1} \cdots F_s^{k_s} F_{s+1}^{k_{s+1}} \cdots F_t^{k_t} F_{t+1}^{k_{t+1}} \cdots F_r^{k_r} \tag{(*)}$$

be the decomposition of F into coprime irreducible factors, with $c \in \mathbb{R} \setminus \{0\}, 0 \le s \le t \le r \in \mathbb{N}$, and $k_j \in \mathbb{N}$. Suppose that F_j is indefinite, for all $j \le s$, F_j is semi-definite nondefinite, for all j with $s + 1 \le j \le t$ and F_j is definite, for all j with $t + 1 \le j \le r$. Clearly

$$\mathsf{V}_{\mathbb{R}}(F) = \mathsf{V}_{\mathbb{R}}(F_1) \cup \cdots \cup \mathsf{V}_{\mathbb{R}}(F_s) \cup \mathsf{V}_{\mathbb{R}}(F_{s+1}) \cup \cdots \cup \mathsf{V}_{\mathbb{R}}(F_t)$$

and $V_{\mathbb{R}}(F_{s+1}) \cup \cdots \cup V_{\mathbb{R}}(F_t)$ is a finite set. The irreducible definite factors of F are irrelevant concerning the set $V_{\mathbb{R}}(F)$, of course.

Definition. 27 With the notations above, we define the irreducible components of the curve $V_{\mathbb{R}}(F)$. These are of two sorts:

- 1. each $V_{\mathbb{R}}(F_i)$, with $j \leq s$ and
- 2. each point in the set $\cup_{j=s+1}^{t} V_{\mathbb{R}}(F_j) \setminus \bigcup_{j=1}^{s} V_{\mathbb{R}}(F_j)$.

Clearly, each irreducible component of type 2 is an isolated point in $V_{\mathbb{R}}(F)$, and each one of type 1 is an algebraic irreducible infinite curve. We exclude in type 2 any point in the set $\bigcup_{j=1}^{s} V_{\mathbb{R}}(F_j)$ due to the fact that, for us, a curve is just a set of points. Note that the gradient of F vanishes at any excluded point P, although P might be regular in the curve $\bigcup_{j=1}^{s} V_{\mathbb{R}}(F_j)$.

In connection to this, it is high time to say that a more general theory of real plane curves should be elaborated. In this theory, curves would not be just sets because the irreducible components of them would carry multiplicites. If this were the case, then the set $\bigcup_{j=s+1}^{t} V_{\mathbb{R}}(F_j) \setminus \bigcup_{j=1}^{s} V_{\mathbb{R}}(F_j)$ in part 2 above should be replaced by $\bigcup_{j=s+1}^{t} V_{\mathbb{R}}(F_j)$.

For affine curves, a similar definition can be given.

Example. 28 Consider the curve $V_{\mathbb{R}}(f)$, with $f = f_1^{\alpha} f_2^{\beta} f_3^{\gamma} f_4^{\delta}$, where $f_1 = Y + X^5 - 2XY - 2$, $f_2 = X^2 + Y^2 - X^3$, $f_3 = -X^2 - (Y - 2)^4 (Y + 4)^6$ and $f_4 = (X + Y)^2 + 3$ and any α, β, γ and $\delta \in \mathbb{N}$. Here,

- f_1 and f_2 are indefinite, by lemma 10.
- f_1 is irreducible, because it is linear in Y; and f_2 is irreducible, since it is the sum of two coprime forms of consecutive degrees.
- f_3 is negative semi-definite nondefinite. It is irreducible, by lemma 16.
- f_4 is positive definite, and is irreducible again by lemma 16.

Thus c = 1, s = 2, t = 3 and r = 4. The irreducible components of $V_{\mathbb{R}}(f)$ are the following:

- $V_{\mathbb{R}}(f_1)$ and $V_{\mathbb{R}}(f_2)$ are components of type 1.
- The point (0, -4) is a component of type 2. It is isolated in $V_{\mathbb{R}}(f)$.

The point (0,2) in $V_{\mathbb{R}}(f_3)$ is not an irreducible component of type 2, since it also belongs to $V_{\mathbb{R}}(f_1)$.

The point (0,0) is singular and isolated in $V_{\mathbb{R}}(f)$. However, it is not an irreducible component of $V_{\mathbb{R}}(f)$ because it belongs to the irreducible set $V_{\mathbb{R}}(f_2)$.

This example shows that not every isolated point in a real algebraic curve is an irreducible component of it.

The question of whether a point in a real curve (affine or projective) is isolated is a local one; therefore it is suitable to address it just in the affine setting.

Let $\emptyset \neq C \subset \mathbb{R}^2$ be an algebraic curve and consider a point $P \in C$. If P is nonsingular, then P is nonisolated in C, by the implicit function theorem. If P is singular, then P may or may not be isolated in C. For instance, the point (0,0) in the curve $V_{\mathbb{R}}(X^4 - 2X^2Y - Y^3)$ is singular nonisolated. Actually, much more is true in this example, for it can be proved that Y is analytic near 0, as an implicit function of X, see [9] p. 12. This curve looks like the parabola $V_{\mathbb{R}}(X^2 - 2Y)$, near (0,0); see figure 5. On the other hand, in example 28 we have found that (0,2) and (0,0) are singular isolated points in the curve $V_{\mathbb{R}}(f)$.



Figure 5: $X^4 - 2X^2Y - Y^3 = 0$

Set $C = V_{\mathbb{R}}(f) \subset \mathbb{R}^2$, for some nonconstant polynomial $f \in \mathbb{R}[X, Y]$. Consider a singular point $P \in C$ and suppose that $C \setminus \{P\}$ is nonempty. The point P is isolated in C if and only if the minimum of the set

$$\{\operatorname{dist}(P,Q): Q \in \mathcal{C} \setminus \{P\}\}\$$

exists and is positive, where the distance considered is euclidean. We may assume that the coordinates of P are (0,0). Then note that if $Q \in C \setminus \{P\}$ is a local extreme for the function $\operatorname{dist}(P, \cdot)$, the point Q must belong to the curve $V_{\mathbb{R}}(P(f))$, where P(f) is defined as $Y\frac{\partial f}{\partial X}$ –

 $X\frac{\partial f}{\partial Y}$. This is so because either the vector $\operatorname{grad}_Q(f)$ is zero or it is nonzero and proportional to the vector from P = (0, 0) to Q. Therefore, we may consider the minimum of the set

 $\{\operatorname{dist}(P,Q): Q \in \mathcal{C} \cap \mathsf{V}_{\mathbb{R}}(\mathsf{P}(f)) \setminus \{P\}\}.$

Before we proceed, here are a few properties of the P operator.

Lemma. 29

- *I*. P(a) = 0, for $a \in \mathbb{R}$.
- 2. p(aX + bY) = -bX + aY, for $a, b \in \mathbb{R}$.
- 3. $P(aX^2 + bXY + cY^2) = -bX^2 + 2(a c)XY + bY^2$, for $a, b, c \in \mathbb{R}$. Moreover, $P(aX^2 + bXY + cY^2)$ is reducible in $\mathbb{R}[X, Y]$.
- 4. If $f \in \mathbb{R}[X, Y]$ is homogeneous and irreducible, then $f \nmid p(f)$.
- 5. If $f \in \mathbb{R}[X, Y]$ is homogeneous, then 0 = P(f) if and only if 0 = P(g), for every irreducible factor g of f.
- 6. If $f \in \mathbb{R}[X, Y]$ is homogeneous of odd degree, then $0 \neq p(f)$.
- 7. Let $f = f_{(0)} + f_{(1)} + \cdots + f_{(d)} \in \mathbb{R}[X, Y]$ be a polynomial of degree $d \ge 0$. The homogeneous part of degree j of $\mathbb{P}(f)$ is $\mathbb{P}(f_{(j)})$, for $j = 0, 1, \dots, d$.

PROOF. Parts 1 and 2 are trivial. Part 3 follows from the fact that the discriminant $4(a - c)^2 + 4b^2$ is nonnegative. Part 4 follows from 1, 2 and 3. Parts 5 and 7 are easy and part 6 follows from 2, 3 and 5.

Corollary. 30 Let $f = f_{(0)} + f_{(1)} + \cdots + f_{(d)} \in \mathbb{R}[X, Y]$ be a polynomial of degree $d \ge 0$. Then 0 = P(f) if and only if $0 = f_{(j)}$, for j odd, and $f_{(j)} = a_j(X^2 + Y^2)^{j/2}$, with $a_j \in \mathbb{R}$, for j even, with $0 \le j \le d$. In this case, $V_{\mathbb{R}}(f)$ is the union of the origin and finitely many circles centered at the origin, if d > 0.

PROOF. Apply parts 7, 6, 5 and 3 to obtain the former assertion. For the latter one, note that if $0 = f_{(j)}$, for j odd, and $f_{(j)} = a_j(X^2 + Y^2)^{j/2}$, for j even, then there exists $g \in \mathbb{R}[Z]$ of order bigger than zero such that $f(X, Y) = g(X^2 + Y^2)$.

Lemma. 31 If $f \in \mathbb{R}[X, Y]$ has order bigger than 1 then P(f) is reducible. If, moreover, f is irreducible, then f and P(f) are coprime.

PROOF. We write $f = \sum_{j+k=2}^{d} a_{jk} X^j Y^k$, with $d \ge 2$. The affine change of coordinates X' + Y' = X and X' - Y' = Y allows us to assume that $a_{jk} = 0$ if j = 1 or k = 1, perhaps with the exception of $a_{21} \ne 0$ or $a_{12} \ne 0$. One more change of coordinates makes $a_{12} = 0$ too. Now an easy computation shows that X divides P(f). Then the second assertion follows, using the irreducibility of f and the fact that $\deg(P(f)) \le \deg(f)$.

We return to the discussion above, where we had a curve $C = V_{\mathbb{R}}(f)$, a singular point $(0,0) = P \in C$ such that $C \setminus \{P\} \neq \emptyset$ and we wanted to know whether P is isolated in C. It is enough to consider the case that f is irreducible. If the order of f is bigger than 1, then f and P(f) are coprime, by lemma 31. By Bezout's theorem, the set $C \cap V_{\mathbb{R}}(P(f)) \setminus \{P\}$ is finite, say $\{Q_1, \ldots, Q_s\}$, for some $s \ge 0$. Then P is isolated in C if and only if the function dist (P, \cdot) attains a local minimum at Q_j , for some $1 \le j \le s$.

The auxiliary polynomial P(f) already appears in Seidenberg's method for deciding whether a real curve is empty, see [7], p. 311. The problem of finding algorithms for a given polynomial $f \in \mathbb{R}[X, Y]$, in order to decide whether $V_{\mathbb{R}}(f)$ is empty or not, and further to determine the topology of $V_{\mathbb{R}}(f)$, is still a matter of research, see [4].

Next we present a simple characterization of bounded affine curves. The proof is a straightforward topology exercise.

Lemma. 32 Let $C \subset \mathbb{R}^2$ be an algebraic curve and $\overline{C} \subset P^2(\mathbb{R})$ be the projective closure of C. Then C is compact if and only if each point in $\overline{C} \setminus C$ is isolated in \overline{C} .

Let us discuss now the issue of which conditions must a projective curve $V_{\mathbb{R}}(F)$ satisfy in order to have a minimal form. We assume that (*) is a factorization of F.

- 1. The empty curve does not have a minimal form. Indeed, this follows from considering the forms $X^2 + Y^2 + Z^2$ and $(X + Z)^2 + Y^2 + Z^2$, both vanishing at no point in $P^2(\mathbb{R})$, both irreducible, by lemma 16, and coprime.
- 2. If $V_{\mathbb{R}}(F) \neq \emptyset$ and moreover $V_{\mathbb{R}}(F)$ has irreducible components of type 2, then $V_{\mathbb{R}}(F)$ does not have a minimal form. Indeed, we may assume that every component of type 2 lie outside the line at infinity Z = 0. If $(a_1 : b_1 : 1), \ldots, (a_n : b_n : 1)$ are all the irreducible components of type 2 of $V_{\mathbb{R}}(F)$ then, taking G_c and $G_{c'}$ the homogeneizations of g_c and $g_{c'}$, as in the comment after corollary 6, we have that $H_c = F_1 \cdots F_s G_c$ and $H_{c'} = F_1 \cdots F_s G_{c'}$ are forms vanishing on $V_{\mathbb{R}}(F)$, for any |c| < 1, |c'| < 1, and $|c| \neq |c'|$. Now $F_1 \cdots F_s$ is a g.c.d. of H_c and $H_{c'}$, and it fails to vanish precisely on $(a_j : b_j : 1)$, for every $j = 1, \ldots, n$.

A particular case occurs when 0 = s < t. A more particular case occurs when F is irreducible semi-definite nondefinite.

3. If $V_{\mathbb{R}}(F) \neq \emptyset$ and, in addition, $V_{\mathbb{R}}(F)$ has no irreducible components of type 2, then $F_1 \cdots F_s$ is a minimal form for $V_{\mathbb{R}}(F)$.

A particular case occurs when $1 \le s = t$. A more particular case occurs when F is irreducible and indefinite.

Definition. 33 The class \mathcal{F}_p of real projective algebraic curves which have a minimal form consists of all nonempty curves having no irreducible components of type 2.

Assume that (*) is a factorization of F. If $V_{\mathbb{R}}(F) \in \mathcal{F}_p$ then the degree of $V_{\mathbb{R}}(F)$ is $\sum_{j=1}^{s} \deg(F_j)$, by definition.

For affine curves, the class \mathcal{F}_a is defined similarly. We set

$$\mathcal{F}=\mathcal{F}_a\cup\mathcal{F}_p.$$

Next we list a few properties of forms and of projective curves. The affine case is similarly stated and proved.

Lemma. 34 Let $\mathcal{C} \subset P^2(\mathbb{R})$ be an algebraic curve. Then

- the fact that C belongs to \mathcal{F}_p (if true),
- the degree of C, for $C \in \mathcal{F}_p$, and
- the irreducibility of C, for $C \in \mathcal{F}_p$ (if true),

do not change by a projective change of coordinates.

PROOF. Let $\phi: P^2(\mathbb{R}) \to P^2(\mathbb{R})$ be a projective transformation and $(a_{i,j}) \in \mathcal{M}_{3\times 3}(\mathbb{R})$ be a regular matrix representing ϕ^{-1} with respect to a given reference. Let $F \in \mathbb{R}[X, Y, Z]$ be a given form of positive degree. We have $F(X, Y, Z) = F(a_{11}X' + a_{12}Y' + a_{13}Z', a_{21}X' + a_{22}Y' + a_{23}Z', a_{31}X' + a_{32}Y' + a_{33}Z')$. Define $\tilde{F}(X', Y', Z')$ using the latter expression. It is easy to verify that:

- F is irreducible if and only if \widetilde{F} is irreducible.
- F is indefinite (resp. semi-definite nondefinite) (resp. definite) if and only if \widetilde{F} is indefinite (resp. semi-definite nondefinite) (resp. definite).
- $\deg(F) = \deg(\widetilde{F}).$

Moreover, if (*) is the decomposition of F into coprime irreducible factors, with $c \in \mathbb{R} \setminus \{0\}$, $r \in \mathbb{N}$, $k_j \in \mathbb{N}$ and $0 \le s \le t \le r$ as above, then

$$\widetilde{F} = c \ G_1^{k_1} \cdots G_s^{k_s} G_{s+1}^{k_{s+1}} \cdots G_t^{k_t} G_{t+1}^{k_{t+1}} \cdots G_r^{k_r}$$

is the corresponding decomposition for \widetilde{F} , with $\widetilde{F}_j = G_j$. It follows that $\cup_{j=s+1}^t V_{\mathbb{R}}(F_j) \subseteq \cup_{j=1}^s V_{\mathbb{R}}(F_j)$ if and only if $\cup_{j=s+1}^t V_{\mathbb{R}}(\widetilde{F}_j) \subseteq \cup_{j=1}^s V_{\mathbb{R}}(\widetilde{F}_j)$. Therefore $V_{\mathbb{R}}(F)$ has irreducible components of type 2 if and only if $V_{\mathbb{R}}(\widetilde{F})$ has. Thus, $V_{\mathbb{R}}(F)$ belongs to \mathcal{F}_p if and only if $V_{\mathbb{R}}(\widetilde{F})$ does. The result follows easily.

Corollary. 35

- 1. Every nonsingular real algebraic curve belongs to \mathcal{F} .
- 2. Every real algebraic curve in \mathcal{F} is infinite.
- 3. The minimal polynomial of a member of \mathcal{F} is a product of $s \ge 1$ irreducible indefinite coprime polynomials.
- 4. $C \in \mathcal{F}_a$ if and only if $\overline{C} \in \mathcal{F}_p$, for any algebraic curve $C \subset \mathbb{R}^2$ and its projective closure $\overline{C} \subset P^2(\mathbb{R})$.
- 5. An irreducible curve not in \mathcal{F} consists of a single point.

3 Affine and projective curves: topological aspects

In this section we compare some elementary topological properties of plane algebraic curves, affine and projective, complex and real. We include easy examples and references where the missing proofs can be found. The techniques used in these proofs come from topology, complex analysis, real algebra, etc.

Ambient spaces

For \mathbb{C} : \mathbb{C}^2 and $P^2(\mathbb{C})$ are connected, orientable \mathbb{R} -topological manifolds of dimension 4, and $P^2(\mathbb{C})$ is compact.

For \mathbb{R} : \mathbb{R}^2 and $P^2(\mathbb{R})$ are connected \mathbb{R} -topological manifolds of dimension 2. Moreover, \mathbb{R}^2 is orientable and $P^2(\mathbb{R})$ is nonorientable and compact; see [6].

Affine or projective algebraic curves

Curves are closed nowhere dense sets, i.e., they have empty interior. Real curves, their complements and each connected component of the former and of the latter are real semi-algebraic sets (see [1] or [2]). It is known that these sets are connected if and only if they are pathwise connected; see [2] p. 47.

For C: As an \mathbb{R} -topological manifold, every nonsingular curve has dimension 2. If a curve is singular, then the set of its regular points is a dense 2-dimensional manifold and in the neighborhood of a singular point of multiplicity m > 1, the curve is homeomorphic to the union of m discs, the centers of which are all identified to one point; see [12] chapter 7. The complement of a curve is connected, since it has codimension 2.

For \mathbb{R} : Every nonsingular curve has dimension 1, as an \mathbb{R} -topological manifold. If a curve C is singular, the subset \mathcal{D} of regular points of C is either empty or a manifold of dimension 1, not necessarily dense in C. \mathcal{D} is empty if and only if C is finite, if and only if C has dimension 0. \mathcal{D} is not dense in C if and only if C has isolated points. Here the notion of *local dimension of a real curve at a point* appears naturally; see [2] p. 53. For instance, the cubic $V_{\mathbb{R}}(X^2 + Y^2 - X^3)$ has dimension 1, but the local dimension at the (isolated) point (0, 0) is 0. The local dimension is 1 at any other point in this cubic. See figure 2.

The complement of a curve may or may not be connected. It is connected if the curve is finite (since it is a set of codimension 2).

For every point x in a curve C and every sufficiently small open ball U with center at x, the set $U \setminus \{x\}$ is homeomorphic to the union of an even number of open segments; these segments are called *half-branches of C at the point x*; see [2] p. 232. For instance, a closed segment or the letter T cannot possibly be homeomorphic to a real algebraic curve.

There exist curves C having points P which are singular, nevertheless in a neighborhood of P the curve C is differentiable. For instance, take the point (0,0) in the curve $V_{\mathbb{R}}(Y^q - X^{q+1})$, with $q \ge 3$ odd. Here, Y is an implicit function of X of class C^1 . See figure 6, with q = 3. Another example is the point (0,0) in the curve $V_{\mathbb{R}}(X^4 - 2X^2Y - Y^3)$, which has already been mentioned.

Affine algebraic curves

For C: They are unbounded sets; it easily follows by intersecting with lines arbitrary distant from the origin. Their connected components are unions of irreducible components; in particular, every irreducible curve is connected.



For \mathbb{R} : They may or may not be bounded. There exist disconnected irreducible curves, such as the hyperbola $V_{\mathbb{R}}(X^2 - Y^2 - 1)$, or the cubics $V_{\mathbb{R}}(X^2 + Y^2 - X^3)$ and $V_{\mathbb{R}}(X + Y^2 - X^3)$. Curves may have isolated points and each of them may or may not be an irreducible component. This has been discussed in p. 16.

Projective algebraic curves For C: Curves are compact sets and, by Bezout's theorem, they are connected. For R: Curves are compact sets too. Also, we have

Harnack's theorem: The number of connected components of a nonsingular projective curve of degree $d \ge 2$ is less than or equal to $\frac{(d-1)(d-2)}{2} + 1$.

See [1] p. 246 or [2] p. 286 for a proof. This is named after the Prussian mathematician Carl Gustav Axel Harnack (1851-1888). One can wonder how the connected components of a curve are arranged. Actually, this is what the first part of Hilbert's 16^{th} problem asks for:

Give a description of the topological types of embeddings of nonsingular algebraic curves in $P^2(\mathbb{R})$.

This problem remains open.

References

- R. Benedetti, J.J. Risler, Real algebraic and semi-algebraic sets, Actualités Mathématiques, (Hermann, Paris, 1990).
- [2] J. Bochnak, M. Coste, M.F. Roy, *Real algebraic geometry*, Ergebnisse der Mathematik und ihrer Grenzgebiete, (Springer, Berlin, 1998).
- [3] M.D. Choi, T.S. Lam, B. Reznick, Real zeros of Positive Semi-definite Forms, I, Math. Z. 171 (1980) 1-26.

- [4] F. Cucker, L. González-Vega, F. Roselló, On algorithms for real algebraic plane curves in Effective methods in algebraic geometry, T. Mora, C. Traverso eds., (Birkhäuser, Boston, 1991) 63-87.
- [5] D. Duval, Rational Puiseux expansions, Compositio Math. 70 (1989) 119-154.
- [6] M.J. Greenberg, J.R. Harper, Algebraic topology. A first course, MLNS 58, (Benjamin/Cummings, Menlo Park, 1981).
- [7] N. Jacobson, Basic Algebra I, (Freeman, San Francisco, 1974).
- [8] T.Y. Lam, An introduction to real algebra, Rocky Mount. J. of Math. 14, n.4 (1984) 767-813.
- [9] J.W. Milnor, Singular points of complex hypersurfaces, Annals of Math. Studies n. 61, (Princeton U. Press, New Jersey, 1968).
- [10] M.J. de la Puente, Curvas algebraicas reales planas, Contribuciones Matemáticas, Homenaje a J. Arregui (Editorial Complutense S.A., Madrid, 2000).
- [11] B. Reznick, Some concrete aspects of Hilbert's 17th problem, Cont. Math. 253, AMS, (2000), 251-272.
- [12] I.R. Shafarevich, Basic algebraic geometry, (2 vols.) (Springer, Berlin, 1994).
- [13] D. Sharpe, *Rings and Factorization*, (Cambridge Univ. Press, 1987).
- [14] H. Whitney, *The elementary structure of real algebraic varieties*, Ann. of Math. 66 (1957) 545-556.
- [15] R. Bix, Conics and Cubics: a Concrete Introduction to Algebraic Curves, (Springer, Berlin, 1998).
- [16] E. Brieskorn y H. Knörrer, Ebene Algebraische Kurven, (Birkhäuser, Boston, 1981), English translation (Birkhäuser, Boston, 1986).
- [17] A. Chenciner, *Courbes Algebriques Planes*, (Publications mathematiques de l' Université Paris VII, 1978).
- [18] G. Fischer, Ebene Algebraische Kurven, (Vieweg-Studium, Brawnschweig, 1994, English translation AMS 2001).
- [19] W. Fulton, Algebraic Curves: an Introduction to Algebraic Geomentry, (Benjamin, New York, 1969).
- [20] C.G. Gibson, *Elementary Geometry of Algebraic Curves: An Undergraduate Introduction*, (Cambridge Univ. Press, 1998).
- [21] K. Kendig, Elementary Algebraic Geometry, (Springer, Berlin, 1977).
- [22] F. Kirwan, Complex Algebraic Curves, (Cambridge Univ. Press, 1992).

- [23] M. Reid, Undergraduate Algebraic Geometry, (Cambridge Univ. Press, 1988).
- [24] A. Seidenberg, *Elements of the Theory of Algebraic Curves*, (Addison-Wesley, Reading, Mass. 1968).
- [25] I. Vainsencher, Introduçao às curvas algébricas planas, (IMPA, Rio de Janeiro, 1996).
- [26] R.J. Walker, Algebraic Curves, (Springer, Berlin, 1978).