

Chapter 2: Basic Real Algebraic Geometry

Adam Sheffer

April 15, 2015

“Every field has its taboos. In algebraic geometry the taboos are (1) writing a draft that can be followed by anyone but two or three of one’s closest friends, (2) claiming that a result has applications, (3) mentioning the word “combinatorial”, and (4) claiming that algebraic geometry existed before Grothendieck.” / “Indiscrete Thoughts” by Gian-Carlo Rota.

In this chapter, we introduce very basic algebraic geometry over the reals. People with previous background in algebraic geometry would probably be rather bored with definitions of ideals, varieties, etc. However, they may be surprised to learn how nonintuitive things are when working over the reals (especially in the second half of this chapter). Although later on we may consider \mathbb{C} and \mathbb{F}_p as our fields, at this point we only consider \mathbb{R} .

1 Ideals and varieties

Algebraic geometry can be thought of as the study of geometries that arise from algebra (or more specifically, from polynomials). In this section we present varieties, which are the basic geometric object of algebraic geometry, and ideals, which are the main algebraic objects that we study in this course.

The *polynomial ring* $\mathbb{R}[x_1, \dots, x_d]$ is the set of all polynomials in x_1, \dots, x_d with coefficients in \mathbb{R} . Given a (possibly infinite) set of polynomials $f_1, \dots, f_k \in \mathbb{R}[x_1, \dots, x_d]$, the *affine variety* $\mathbf{V}(f_1, \dots, f_k)$ is defined as

$$\mathbf{V}(f_1, \dots, f_k) = \{(a_1, \dots, a_d) \in \mathbb{R}^d : f_i(a_1, \dots, a_d) = 0 \text{ for all } 1 \leq i \leq k\}.$$

The adjective “affine” distinguishes the variety from projective varieties. At this point we only consider affine varieties, and for brevity refer to those simply as varieties.¹ For example, some varieties in \mathbb{R}^3 are a torus, the union of a circle and a line, and a set of 1000 points.

The following is a special case of *Hilbert’s basis theorem* (e.g., see [2, Section 2.5]).

Theorem 1.1. *Every variety can be described by a finite set of polynomials.*

¹To make things even worse, some authors call these objects *algebraic sets*, while using the word variety for what we will refer to as an irreducible variety.

Theorem 1.1 is valid in every field. When working over the reals, we can say something stronger.

Corollary 1.2. *Every variety can be described by a single polynomial.*

Proof. Consider a variety $U \subset \mathbb{R}^d$. By Theorem 1.1, there exist $f_1, \dots, f_k \in \mathbb{R}[x_1, \dots, x_d]$ such that $U = \mathbf{V}(f_1, \dots, f_k)$. We set $f = f_1^2 + f_2^2 + \dots + f_k^2$. Notice that for any point $p \in \mathbb{R}^d$ we have $f(p) = 0$ if and only if $f_1(p) = \dots = f_k(p) = 0$. Thus, we have $U = \mathbf{V}(f)$. \square

We consider some basic properties of varieties.

Claim 1.3. *Let $U, W \subset \mathbb{R}^d$ be two varieties, and let $\tau : \mathbb{R}^d \rightarrow \mathbb{R}^d$ be an invertible linear map (such as a translation, rotation, reflection, stretching, etc.). Then*

- (a) $U \cap W$ is a variety,
- (b) $U \cup W$ is a variety, and
- (c) $\tau(U)$ is a variety.

Proof. Since U and W are varieties, there exist $f_1, \dots, f_k, g_1, \dots, g_m \in \mathbb{R}[x_1, \dots, x_d]$ such that $U = \mathbf{V}(f_1, \dots, f_k)$ and $W = \mathbf{V}(g_1, \dots, g_m)$.² For (a), notice that we have $U \cap W = \mathbf{V}(f_1, \dots, f_k, g_1, \dots, g_m)$. For (b), we have $U \cup W = \mathbf{V}(H)$, where

$$H = \bigcup_{\substack{1 \leq i \leq k \\ 1 \leq j \leq m}} \{f_i \cdot g_j\}.$$

For (c), we write the inverse of τ as $\psi \in (\mathbb{R}[x_1, \dots, x_d])^d$. Then

$$\tau(U) = \mathbf{V}(f_1 \circ \psi, \dots, f_d \circ \psi).$$

\square

At this point it might be instructive to ask what subsets of \mathbb{R}^d are not varieties.

Claim 1.4. *The set $X = \{(x, x) : x \in \mathbb{R}, x \neq 0\} \subset \mathbb{R}^2$ is not a variety.*

Proof. Assume for contradiction that there exist $f_1, \dots, f_k \in \mathbb{R}[x_1, x_2]$ such that $X = \mathbf{V}(f_1, \dots, f_k)$. For every $1 \leq i \leq k$, we set $g_i(t) = f_i(t, t)$ and notice that $g_i \in \mathbb{R}[t]$. Since by definition $g_i(t)$ vanishes on every $t \neq 1$, we have that $g_i(t) = 0$. This in turn implies that $f_i(1, 1) = 0$. Since this holds for every $1 \leq i \leq k$, we get a contradiction to $(1, 1) \notin X$. \square

Similarly, a line segment and half a circle are not varieties.

We say that a set U' is a *subvariety* of a variety U if $U' \subseteq U$ and U' is a variety. We say that U' is a *proper subvariety* if $U' \neq \emptyset, U$. A variety U is *reducible* if there exist two proper subvarieties $U', U'' \subset U$ such that $U = U' \cup U''$. Otherwise, U is *irreducible*. For example, the union of the two axes $\mathbf{V}(x, y) \subset \mathbb{R}^2$ is reducible since $\mathbf{V}(x, y) = \mathbf{V}(x) \cup \mathbf{V}(y)$.

²By Corollary 1.2, it suffices to use a single polynomial for each variety. We present this slightly less elegant proof since it applies in every field.

Every variety U can be decomposed into irreducible subvarieties U_1, \dots, U_k such that $U = \bigcup_{i=1}^k U_i$. After removing any U_i that is a proper subvariety of another U_j , we obtain a unique decomposition of U . The subvarieties of this decomposition are said to be the *irreducible components of U* (or components, for brevity).

Ideals. A subset $J \subseteq \mathbb{R}[x_1, \dots, x_d]$ is an *ideal* if it satisfies:

- $0 \in J$.
- If $f, g \in J$ then $f + g \in J$.
- If $f \in J$ and $h \in \mathbb{R}[x_1, \dots, x_d]$, then $f \cdot h \in J$.

As a first example of an ideal, consider a polynomial $f \in \mathbb{R}[x_1, \dots, x_d]$ and notice that $\{f \cdot h : h \in \mathbb{R}[x_1, \dots, x_d]\}$ is an ideal. More generally, given $f_1, \dots, f_k \in \mathbb{R}[x_1, \dots, x_d]$, the set

$$\langle f_1, \dots, f_k \rangle = \left\{ \sum_{i=1}^k f_i \cdot h_i : h_1, \dots, h_k \in \mathbb{R}[x_1, \dots, x_d] \right\}$$

is an ideal.³ We say that this ideal is generated by f_1, \dots, f_k . We also say that $\{f_1, \dots, f_k\}$ is a *basis* of this ideal.

We are specifically interested in ideals of varieties. Given a variety $U \subset \mathbb{R}^d$, the ideal of U is

$$\mathbf{I}(U) = \{f \in \mathbb{R}[x_1, \dots, x_d] : f(a) = 0 \text{ for every } a \in U\}.$$

It can be easily verified that $\mathbf{I}(U)$ satisfies the three requirements in the definition of an ideal. This seems to lead to the following question: Given $f_1, \dots, f_k \in \mathbb{R}[x_1, \dots, x_d]$, is it always the case that $\langle f_1, \dots, f_k \rangle = \mathbf{I}(\mathbf{V}(f_1, \dots, f_k))$?

Claim 1.5. *Given $f_1, \dots, f_k \in \mathbb{R}[x_1, \dots, x_d]$, we have $\langle f_1, \dots, f_k \rangle \subseteq \mathbf{I}(\mathbf{V}(f_1, \dots, f_k))$ although equality need not occur.*

Proof. To see that the containment relation holds, we set $U = \mathbf{V}(f_1, \dots, f_k)$. If $g \in \langle f_1, \dots, f_k \rangle$ then by definition g vanishes on every point of U , and is thus in $\mathbf{I}(\mathbf{V}(f_1, \dots, f_k))$.

To see that equality does not always hold, we set $f = x^2 + y^2$. We then have $\mathbf{V}(f) = \{(0, 0)\} \subset \mathbb{R}^2$, even though $x \in \mathbf{I}(\mathbf{V}(f))$ and $x \notin \langle f \rangle$. \square

When defining a variety U , it is often useful to use a basis of $\mathbf{I}(U)$ rather than an arbitrary set of polynomials that define U .

We conclude this section by inspecting another connection between ideals and varieties.

Claim 1.6. *Let $U, W \subset \mathbb{R}^d$ be varieties. Then*

- (a) $U \subset W$ if and only if $\mathbf{I}(W) \subset \mathbf{I}(U)$.
- (b) $U = W$ if and only if $\mathbf{I}(W) = \mathbf{I}(U)$.

³There are many different notations for an ideal generated by a set of polynomials. We use $\langle \cdot \rangle$ following the notation of [2].

Proof. First assume that $U \subset W$ and let $f \in \mathbf{I}(W)$. That is, the polynomial f vanishes on every point of W . Since $U \subset W$, then f vanishes on every point of U , which implies $f \in \mathbf{I}(U)$. To see that this containment is proper, notice that there must exist polynomials that vanish on U but not on W (otherwise we would have $U = W$).

Next, assume that $\mathbf{I}(W) \subset \mathbf{I}(U)$ and let $p \in U$. Every polynomial of $\mathbf{I}(U)$ vanishes on the point p , which in turn implies that every polynomial of $\mathbf{I}(W)$ vanishes on p . Thus, we have $p \in W$.

Part (b) is proved in a similar manner. □

2 Dimension

Consider an irreducible variety $U \subset \mathbb{R}^d$. One intuitive definition of the *dimension* d' of U , denoted $\dim U$, is the maximum integer such that there exist

$$U_0 \subset U_1 \subset \cdots \subset U_{d'} = U,$$

where all of the subsets are proper and all of the sets U_i are irreducible varieties. This definition immediately implies that the intersection of two distinct irreducible varieties of dimension d' is of dimension smaller than d' .

If $U \subset \mathbb{R}^d$ is a reducible variety with irreducible components U_1, \dots, U_k , then the dimension of U is $\max_i \{\dim U_i\}$.

A *curve* is a variety with all of its components of dimension one. A *k-flat* is a translation of a k -dimensional linear space. A *hypersurface* in \mathbb{R}^d is a variety with all of its components of dimension $d - 1$. Similarly, a hyperplane in \mathbb{R}^d is a $(d - 1)$ -flat, a hypersphere is a $(d - 1)$ -dimensional sphere, etc. We present the following claim without proof.

Claim 2.1. *For every hypersurface $U \subset \mathbb{R}^d$ there exists a polynomial $f \in \mathbb{R}[x_1, \dots, x_d]$ such that $\langle f \rangle = \mathbf{I}(U)$.*

Given an irreducible variety $U \subset \mathbb{R}^d$ of dimension d' , one might expect U to “look” like a d' -dimensional set in a small neighborhood around any point $p \in U$. To see that this is not the case, consider the cubic curve $\mathbf{V}(y^2 - x^3 + x^2)$. Even though this is an irreducible variety, it consists of a one-dimensional curve together with the origin (see the left part of Figure 1). If we remove the origin, we obtain a set that is not a variety. A similar example is the Whitney umbrella $\mathbf{V}(x^2 - y^2z) \subset \mathbb{R}^3$, which consists of a two-dimensional surface together with the z -axis (see the right part of Figure 1). As before, if we remove the line we obtain a set that is not a variety.

3 Singular points

Let $U \subset \mathbb{R}^d$ be a variety and let $p \in U$. Intuitively (and with some exceptions), p is a *singular point* of U if the tangent to U at p is not well defined (e.g., see Figure 2), if p is contained in more than one irreducible component of U , or if the tangent at p is

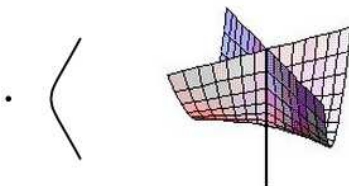


Figure 1: A cubic curve in \mathbb{R}^2 and the Whitney umbrella.

of a smaller dimension than the dimension of an irreducible component that contains p (for example, the line of the Whitney umbrella consists of singular points). A point of U that is not singular is said to be a *regular* point of U .

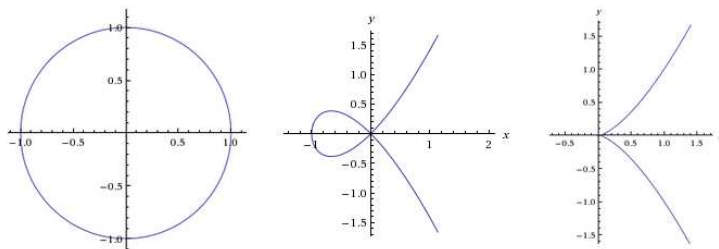


Figure 2: Every point of the circle has a well defined tangent line. In the other two curves the tangent is not well defined at the origin. These curves are $\mathbf{V}(y^2 - x^3 - x^2)$ and $\mathbf{V}(x^3 - y^2)$.

For a more rigorous definition, we begin with the special case of hypersurfaces. Given a polynomial $f \in \mathbb{R}[x_1, \dots, x_d]$, the *gradient* of f is

$$\nabla f = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \dots, \frac{\partial f}{\partial x_d} \right).$$

Let $U \subset \mathbb{R}^d$ be a hypersurface. By Claim 2.1 there exists $f \in \mathbb{R}[x_1, \dots, x_d]$ such that $\mathbf{I}(U) = \langle f \rangle$. Then $p \in U$ is a singular point if and only if $\nabla f(p) = 0$ (that is, the gradient is a vector of zeros). We denote the set of singular points of U as U_{sing} .

To see why we require that $\mathbf{I}(U) = \langle f \rangle$, consider the following example. Let $U \subset \mathbb{R}^2$ be the x -axis. It can be easily verified that U has no singular points. Notice that $U = \mathbf{V}(y^2)$, and that the gradient $\nabla y^2 = (0, 2y)$ is zero for every point of U . Thus, by using y^2 instead of y , we get the false impression that every point of U is singular.

A polynomial is said to be *square-free* if in its factorization into irreducible factors, no factor has a multiplicity larger than one. The problem with the above example is that we took an f that is not square-free. It is not enough to take a square-free f such that $\mathbf{I}(U) = \langle f \rangle$. For example, consider again the case where $U \subset \mathbb{R}^2$ is the x -axis. Notice that $U = \mathbf{V}(y(x^2 + y^2))$, and that the gradient $\nabla y(x^2 + y^2) = (2xy, x^2 + 3y^2)$ is zero at the origin.

A useful property of square-free polynomials is that they do not have common factors with any of their first partial derivatives. Indeed, consider a square-free $f \in \mathbb{R}[x_1, \dots, x_d]$, and let $g \in \mathbb{R}[x_1, \dots, x_d]$ be an irreducible factor of f . That is, $f = g \cdot h$

for some $h \in \mathbb{R}[x_1, \dots, x_d]$ that does not have g as a factor. We then have

$$\frac{\partial f}{\partial x_i} = \frac{\partial g}{\partial x_i} \cdot h + g \cdot \frac{\partial h}{\partial x_i}.$$

Since the second summand has g as a factor but not the first summand, this expression does not have g as a factor.

Claim 3.1. *Every hypersurface $U \subset \mathbb{R}^d$ contains a regular point.*

Proof. By Claim 2.1 there exists $f \in \mathbb{R}[x_1, \dots, x_d]$ such that $\mathbf{I}(U) = \langle f \rangle$. Notice that such f must be square-free, and thus have no common factors with $f_1 = \frac{\partial f}{\partial x_1}$. If every point of U is singular then f_1 is in $\langle f \rangle$, which is a contradiction. \square

As we will see in Theorem 4.1, a much stronger property holds — the set of singular points is a lower-dimensional variety.

To define singular points of varieties that are not hypersurfaces, we require a generalization of the gradient. Given $f_1, \dots, f_k \in \mathbb{R}[x_1, \dots, x_d]$, the Jacobian matrix of f_1, \dots, f_k is

$$\mathbf{J}_{f_1, \dots, f_k} = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_d} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_d} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial f_k}{\partial x_1} & \frac{\partial f_k}{\partial x_2} & \cdots & \frac{\partial f_k}{\partial x_d} \end{pmatrix}$$

Let $U \subset \mathbb{R}^d$ be a variety of dimension d' , and consider $f_1, \dots, f_k \in \mathbb{R}[x_1, \dots, x_d]$ such that $\mathbf{I}(U) = \langle f_1, \dots, f_k \rangle$. Then $p \in U$ is a singular point if and only if $\text{rank}(\mathbf{J}_{f_1, \dots, f_k}(p)) < d - d'$. Recall that $\text{rank}(\mathbf{J}_{f_1, \dots, f_k}(p)) + \dim \ker(\mathbf{J}_{f_1, \dots, f_k}(p)) = d$. Moreover, the tangent to U at p is orthogonal to ∇f_i , for every $1 \leq i \leq k$. Thus, we can intuitively think of $d - \text{rank}(\mathbf{J}_{f_1, \dots, f_k}(p))$ as the dimension of the tangent of U at p .

As a counterexample to the intuition given in the beginning of this section, consider $f = y^3 + 2x^2y - x^4 \in \mathbb{R}[x, y]$. The variety $U = \mathbf{V}(f)$ is depicted in Figure 3. Notice that U is an irreducible variety that does not intersect itself and has a tangent line at every point. Nonetheless, it can be easily verified that $(0, 0)$ is a singular point of U . (For a discussion of this phenomenon, see for example [1, Section 3.3].)

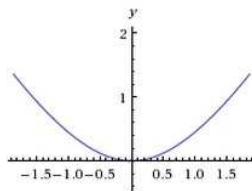


Figure 3: The variety $\mathbf{V}(y^3 + 2x^2y - x^4) \subset \mathbb{R}^2$.

4 Degrees

We will sometimes discuss a *generic* object. When we say that a generic point of \mathbb{R}^d has property X , we mean that the set of points in \mathbb{R}^d that do not have property X are contained in a proper subvariety (i.e., have measure zero). For example, given a plane $h \subset \mathbb{R}^3$, a generic point of \mathbb{R}^3 is not contained in h . We can talk about other generic objects in a similar manner. For example, describing a line in \mathbb{R}^3 requires four parameters, so we can think of the lines of \mathbb{R}^3 as the points of a four-dimensional space.⁴ Thus, we can say that a generic line in \mathbb{R}^3 intersects h in a single point.

Defining the *degree* of a real variety is somewhat problematic. In the complex space there is a well defined notion of degree, with many equivalent definitions. For example, a variety $U \subset \mathbb{C}^d$ of dimension d' has degree k if and only if it intersects a generic $(d - d')$ -flat of \mathbb{C}^d in exactly k points. To see that this does not make sense over the reals, notice that a circle in \mathbb{R}^2 does not intersect a generic line at all (although intuitively we expect a circle to have degree 2).

By corollary 1.2, every real variety is the zero set of a single polynomial, and it might seem tempting to consider the minimum degree of such a polynomial. Unfortunately this definition also has weird consequences. For example, if $U \subset \mathbb{R}^d$ consists of a single point, then it has degree 2.

This has led to a situation where several non-equivalent definitions of degree are being used in \mathbb{R}^d , and some works simply avoid defining this notion. In this course we sometimes consider the degree as being “constant”, and then obtain the same asymptotic results for any reasonable definition of degree. For our purposes, we define the degree of a variety $U \subset \mathbb{R}^d$ as

$$\min_{\substack{f_1, \dots, f_k \in \mathbb{R}[x_1, \dots, x_d] \\ \mathbf{v}(f_1, \dots, f_k) = U}} \max_{1 \leq i \leq k} \deg f_i.$$

That is, the degree of U is the minimum integer D such that U can be defined by a set of polynomials of degree at most D . We might introduce other definitions of degree later on.

Theorem 4.1. *Let $U \subset \mathbb{R}^d$ be a variety of degree k and dimension d' . Then U_{sing} is a variety of dimension smaller than d' and of degree $O_{k,d}(1)$.⁵*

Proof. By definition, there exist polynomials f_1, \dots, f_ℓ of degree at most k such that $\langle f_1, \dots, f_\ell \rangle = I(U)$. We have

$$U_{\text{sing}} = \left\{ p \in U : \text{rank}(\mathbf{J}_{f_1, \dots, f_k}(p)) < d - d' \right\}.$$

Notice that $\mathbf{J}_{f_1, \dots, f_k}(p)$ is of rank smaller than $d - d'$ if and only if every $(d - d') \times (d - d')$ minor of $\mathbf{J}_{f_1, \dots, f_k}(p)$ is zero. Such a minor is an equation of degree at

⁴We actually need this four-dimensional space to be projective, but hopefully the idea itself is clear without the rigorous details.

⁵By $O_{k,d}(1)$, we mean $O(1)$ where the constant in the $O()$ -notation depends on k and d .

most $k(d - d')$, so U_{sing} is a variety that is defined by polynomials of degree at most $k(d - d') = O_{d,k}(1)$.

To prove that U_{sing} is of dimension smaller than d' , we first consider the case where U is an irreducible hypersurface. By Claim 2.1, U contains a regular point. Since U_{sing} is a subvariety of an irreducible variety and is not identical to U , it must have a smaller dimension. One can also show that every lower dimensional variety must contain a regular point⁶. Thus, the assertion of the theorem holds for every irreducible variety.

It remains to prove that U_{sing} is of dimension smaller than d' when U is reducible. By the above, the set of points that are singular points of a specific component are of dimension smaller than d' , so it remains to consider points that are singular due to an intersection of at least two components. This is straightforward, since the intersection of two distinct irreducible varieties of dimension d' must be of a smaller dimension. \square

5 Degrees in \mathbb{R}^2

Before concluding this chapter, we consider a couple of useful bounds in \mathbb{R}^2 . First, notice that the degree of a curve in \mathbb{R}^2 coincides with the minimum degree of a polynomial that defines this curve. We begin with a real planar version of Bézout's theorem.

Theorem 5.1 (Bézout's theorem). *Let f and g be two polynomials in $\mathbb{R}[x, y]$ of degrees k_f and k_g , respectively. If $\mathbf{V}(f)$ and $\mathbf{V}(g)$ do not have a common one-dimensional component, then $\mathbf{V}(f) \cap \mathbf{V}(g)$ consists of at most $k_f \cdot k_g$ points.*

As simple examples, notice that two lines indeed intersect in at most one point and that two ellipses (which are of degree 2) intersect in at most four points. For more details about Bézout's theorem, including a proof, see for example [3, Section 14.4]. We now consider the maximum number of connected components that a variety can have.

Theorem 5.2 (Harnack's curve theorem). *Let $f \in \mathbb{R}[x, y]$ be a polynomial of degree k . Then the number of connected components of $\mathbf{V}(f)$ is $O(k^2)$.⁷*

Proof. We may assume that f is square-free, since removing repeated factors does not change the corresponding variety. Every bounded connected component of $\mathbf{V}(f)$ has at least two extreme points in the x -direction (that is, its leftmost and rightmost points). Such a point $p \in \mathbf{V}(f)$ satisfies $f(p) = \frac{\partial f}{\partial y}(p) = 0$. Since f is square-free, it has no common components with $\frac{\partial f}{\partial y}$. Thus, by Bézout's theorem $\mathbf{V}(f) \cap \mathbf{V}\left(\frac{\partial f}{\partial y}\right)$ has at most $k(k - 1)$ points. This in turn implies that the number of bounded connected components of $\mathbf{V}(f)$ is at most $k(k - 1)/2$.

⁶I still need to add a reference for this. Please tell me if you have a good one over the reals (or a simple proof).

⁷Harnack's exact bound is $1 + \binom{k-1}{2}$, although the following proof implies a slightly worse bound.

To bound the number of unbounded connected components of $\mathbf{V}(f)$, we perform a generic rotation of \mathbb{R}^2 around the origin and consider a sufficiently large constant c , so that only the unbounded connected components intersect the lines $\mathbf{V}(x - c)$ and $\mathbf{V}(x + c)$. By Bézout's theorem, $\mathbf{V}(f)$ intersects each of those lines in at most k points, so there are at most $2k$ unbounded connected components. \square

References

- [1] J. Bochnak, M. Coste, and M. Roy, *Real Algebraic Geometry*, Springer-Verlag, Berlin, 1998.
- [2] D. Cox, J. Little and D. O'Shea, *Ideals, Varieties, and Algorithms: An Introduction to Computational Algebraic Geometry and Commutative Algebra*, 3rd edition, Springer-Verlag, Heidelberg, 2007.
- [3] C. G. Gibson, *Elementary geometry of algebraic curves: an undergraduate introduction*, Cambridge University Press, 1998.
- [4] J. Harris, *Algebraic geometry: a first course*, Springer, New York, 1992.
- [5] R. Hartshorne, *Algebraic Geometry*, Springer-Verlag, New York. 1983.