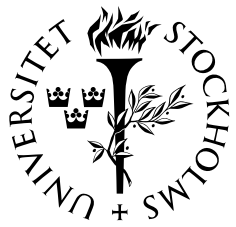

SJÄLVSTÄNDIGA ARBETEN I MATEMATIK

MATEMATISKA INSTITUTIONEN, STOCKHOLMS UNIVERSITET

An introduction to algebraic geometry and Bezout's theorem

av

Elisabeth Bonnevier

2018 - No K24

An introduction to algebraic geometry and Bezout's theorem

Elisabeth Bonnevier

Självständigt arbete i matematik 15 högskolepoäng, grundnivå

Handledare: Gregory Arone

2018

Abstract

The fundamental theorem of algebra tells us the number of roots of a polynomial. As a generalization, Bezout's theorem tells us the number of intersection points between two arbitrary polynomial curves in a plane. The aim of this text is to develop some of the theory of algebraic geometry and prove Bezout's theorem. First, after some initial definitions and propositions we will prove the classical result of Hilbert's nullstellensatz, which describes the relationship between algebraic sets and ideals of a polynomial ring. From that we continue on to define the projective space, to which we extend our previous definitions of algebraic sets and ideals. Also needed for Bezout's theorem is the notion of intersection number, which is a generalization of counting zeros with multiplicities. The properties expected of the intersection number are given and we show that there is only one number which satisfies those properties. Then we have all the theory needed and we will prove Bezout's theorem.

Acknowledgements

I would like to thank my supervisor Gregory Arone for all his valuable help in understanding the theory and in writing the thesis. You always manage to give good concrete examples to illustrate abstract or complicated concepts. I would also like to thank Rikard Bögvad for giving me the idea to study Bezout's theorem and for reviewing my thesis.

Contents

1 Introduction

A classical result in algebra is that the number of roots of a polynomial f of one variable over $\mathbb{C}$ is equal to the degree of f, if we count multiplicities correctly. An equivalent statement is that the number of times that the curves $y = f$ and $y = 0$ intersect in $\mathbb{C}^2$ is equal to the degree of f, if we count multiplicities correctly. This prompts a new question: how many times do two arbitrary curves in $\mathbb{C}^2$ intersect? The answer to this question is called Bezout's theorem, which we aim to prove in this text.

In its weakest form, Bezout's theorem states that two polynomial curves f and g in the plane of an arbitrary field intersect in at most $\deg(f) \cdot \deg(g)$ points. There are three steps which we may take to sharpen the theorem. The first is to take the intersection points in the plane of an *algebraically closed* field. We could have hoped that this would be enough to ensure the existence of at least one intersection point. However, it turns that it is not enough. We also need to extend the plane in order to include *points at infinity*. Then any two curves will intersect at least once.

The third and last step is in analogue to the need for correctly counting multiplicities of zeros of a polynomial of one variable in order to get the full number of roots. We define the *intersection number* of f and g at a point in a natural way as a non-negative integer that states "how many times" f and g intersect there. Then we finally have the full Bezout's theorem, which states that if f and g are curves in the plane of an algebraically closed field and we count points at infinity and intersection multiplicities correctly then there are exactly $\deg(f) \cdot \deg(g)$ intersection points.

2 Foundations

2.1 Basic definitions and concepts

Unless otherwise stated, the results and proofs in this text are based on the book *Algebraic Curves* by William Fulton [2].

Let $\mathbf{K}$ be a field and let $\mathbf{K}[x_1, ..., x_n]$ be the polynomial ring over $\mathbf{K}$. For any polynomial $f \in \mathbf{K}[x_1, ..., x_n]$, let

$$\mathbf{V}(f) := \{\mathbf{x} \in \mathbf{K}^n \mid f(\mathbf{x}) = 0\},$$

i.e. $\mathbf{V}(f)$ is the zero-set of f.

For any set S of polynomials in $\mathbf{K}[x_1, ..., x_n]$ let

$$\mathbf{V}(S) := \bigcap_{f \in S} \mathbf{V}(f),$$

i.e. $\mathbf{V}(S)$ is the set of common zeros of the polynomials in S. When $S = \{f_i\}_{i=1}^k$ is finite we will write $\mathbf{V}(f_1, ..., f_k)$ instead of $\mathbf{V}(\{f_1, ..., f_k\})$.

Definition 2.1. Let $X \subseteq \mathbf{K}^n$, then X is called an **algebraic set** if $X = \mathbf{V}(S)$ for some $S \subseteq \mathbf{K}[x_1, ..., x_n]$.

The main purpose of algebraic geometry is to study algebraic sets. There is a correspondence between algebraic sets of $\mathbf{K}^n$ and ideals in $\mathbf{K}[x_1, ..., x_n]$ so that many properties of algebraic sets can be reduced to properties of ideals, which are often easier to work with.

Example 2.1. Here are some examples of algebraic sets in $\mathbb{R}^2$.

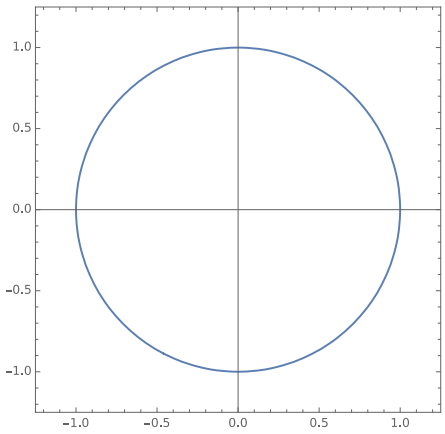

Figure 1: $\mathbf{V}(y^2 + x^2 - 1)$

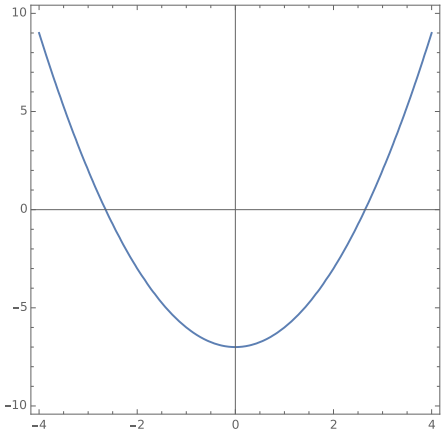

Figure 2: $\mathbf{V}(y - x^2 + 7)$

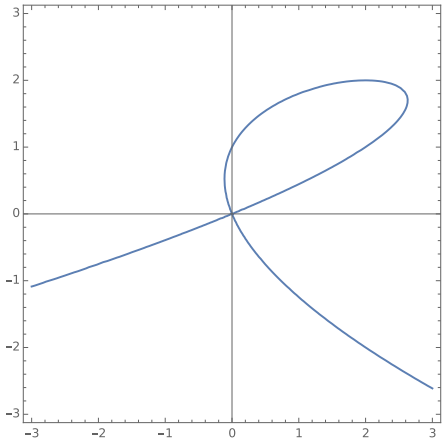

Figure 3: $\mathbf{V}(x^2 + y^3 - y^2 - 2xy)$

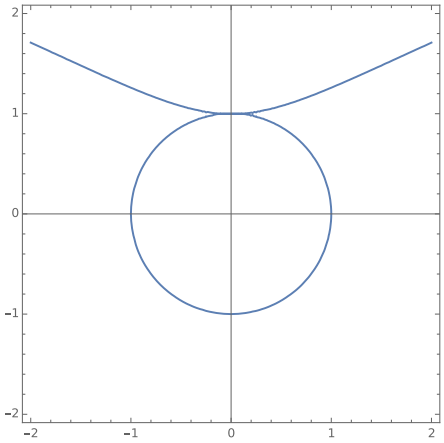

Figure 4: $\mathbf{V}((y^2 + x^2 - 1)(x^2 - y^3 + 1))$

Figure 1 is the unit circle, which is the set of solutions to the equation $y^2+x^2 = 1$. So the set of solutions is also the zero-set of the polynomial $y^2 + x^2 - 1$. Equally for Figure 2 which is the curve $y = x^2 - 7$. The curves in Figure 3 and Figure 4 are more complicated but are nonetheless algebraic sets.

We continue with a proposition on algebraic sets.

Proposition 2.1. *Let I be the ideal in $\mathbf{K}[x_1, ...x_n]$ that is generated by S. Then $\mathbf{V}(S) = \mathbf{V}(I)$.*

Proof. Since $S \subseteq I$ we have the inclusion $\mathbf{V}(I) \subseteq \mathbf{V}(S)$. So what we need to prove is the other inclusion.

Suppose $\mathbf{x} \in \mathbf{V}(S)$. Then $f(\mathbf{x}) = 0$ for all $f \in S$, and thus for any $f_1, ..., f_n \in S$ and for any $g_1, ..., g_n \in \mathbf{K}[x_1, ..., x_n]$ we have $(g_1 f_1 + ... + g_n f_n)(\mathbf{x}) = 0$. So

$\mathbf{x} \in \mathbf{V}(I)$ and thus $\mathbf{V}(S) \subseteq \mathbf{V}(I)$. $\square$

This means that every algebraic set is the zero-set of some ideal in $\mathbf{K}[x_1, ..., x_n]$ and that in order to study the properties of $\mathbf{V}(S)$ it is enough to consider zero-set of the ideal generated by S. What follows are some basic properties of algebraic sets.

Proposition 2.2. *The following properties hold:*

1. *For any two ideals I and J, if $I \subseteq J$ then $\mathbf{V}(I) \supseteq \mathbf{V}(J)$.*
2. *For any collection $\{I_\alpha\}$ of ideals, $\mathbf{V}(\cup_\alpha I_\alpha) = \cap_\alpha \mathbf{V}(I_\alpha)$.*
3. *For any two ideals I and J, $\mathbf{V}(I) \cup \mathbf{V}(J) = \mathbf{V}(\{fg \in \mathbf{K}[x_1, ..., x_n] \,|\, f \in I,\, g \in J\})$.*

Proof.

1. Suppose $I \subseteq J$ and let $\mathbf{x} \in \mathbf{V}(J)$. Then $f(\mathbf{x}) = 0$ for all $f \in J$. In particular $f(\mathbf{x}) = 0$ for all $f \in I$, so $\mathbf{x} \in \mathbf{V}(I)$ and thus $\mathbf{V}(I) \supseteq \mathbf{V}(J)$.

2. We have the following equivalences:

$$\begin{aligned}
\mathbf{x} \in \mathbf{V}(\cup_\alpha I_\alpha) &\Leftrightarrow f(\mathbf{x}) = 0 \text{ for all } f \in \cup_\alpha I_\alpha \\
&\Leftrightarrow f(\mathbf{x}) = 0 \text{ for all } f \in I_\alpha \text{ for all } \alpha \\
&\Leftrightarrow \mathbf{x} \in \mathbf{V}(I_\alpha) \text{ for all } \alpha \\
&\Leftrightarrow \mathbf{x} \in \cap_\alpha \mathbf{V}(I_\alpha).
\end{aligned}$$

3. We have the following equivalences:

$$\begin{aligned}
\mathbf{x} \in \mathbf{V}(I) \cup \mathbf{V}(J) &\Leftrightarrow \mathbf{x} \in \mathbf{V}(I) \text{ or } \mathbf{x} \in \mathbf{V}(J) \\
&\Leftrightarrow f(\mathbf{x}) = 0 \text{ for all } f \in I \text{ or } f(\mathbf{x}) = 0 \text{ for all } f \in J \\
&\Leftrightarrow (fg)(\mathbf{x}) = 0 \text{ for all } f \in I,\, g \in J \\
&\Leftrightarrow \mathbf{x} \in \mathbf{V}(\{fg \in \mathbf{K}[x_1, ..., x_n] \,|\, f \in I,\, g \in J\}),
\end{aligned}$$

   where the left implication of the third equivalence is given by the following argument: suppose $(fg)(\mathbf{x}) = 0$ for all $f \in I$ and all $g \in J$. Then for any given combination fg we have either $f(\mathbf{x}) = 0$ or $g(\mathbf{x}) = 0$. Now suppose that some but not all of the polynomials in I are zero at $\mathbf{x}$ and suppose the same for the polynomials in J. Then take any $f \in I$ and $g \in J$ such that $f(\mathbf{x}) \neq 0$ and $g(\mathbf{x}) \neq 0$ and then we get $(fg)(\mathbf{x}) \neq 0$ which is a contradiction. So either all of I vanishes on $\mathbf{x}$ or all of J vanishes on $\mathbf{x}$.

$\square$

Observe that $\mathbf{V}(x_1 - a_1, ..., x_n - a_n) = \{(a_1, ..., a_n)\}$. So by property 3 in proposition 2.2 any finite set is an algebraic set.

We have seen how we may define sets in $\mathbf{K}^n$ by using ideals in $\mathbf{K}[x_1, ..., x_n]$. In a similar manner we may define ideals in $\mathbf{K}[x_1, ..., x_n]$ using sets in $\mathbf{K}^n$.

For any set $X \subseteq \mathbf{K}^n$ let

$$\mathbf{I}(X) = \{f \in \mathbf{K}[x_1, ...x_n] \mid f(\mathbf{x}) = 0 \; \forall \, \mathbf{x} \in X\},$$

i.e. $\mathbf{I}(X)$ is the set of polynomials which vanish on X.

If f and g vanish on X then so do $f + g$ and hf for any $h \in \mathbf{K}[x_1, ...x_n]$. So $\mathbf{I}(X)$ is an ideal in $\mathbf{K}[x_1, ..., x_n]$. Here are some basic properties of ideals of sets.

Proposition 2.3. *The following properties hold:*

1. *For any two sets X and Y in $\mathbf{K}^n$, if $X \subseteq Y$ then $\mathbf{I}(X) \supseteq \mathbf{I}(Y)$,*
2. *For any two sets X and Y in $\mathbf{K}^n$, $\mathbf{I}(X \cup Y) = \mathbf{I}(X) \cap \mathbf{I}(Y)$.*

Proof.

1. Suppose $X \subseteq Y$ and let $f \in \mathbf{I}(Y)$. Then $f(\mathbf{x}) = 0$ for all $\mathbf{x} \in Y$ and in particular, $f(\mathbf{x}) = 0$ for all $\mathbf{x} \in X$. So $f \in \mathbf{I}(X)$ and thus $\mathbf{I}(X) \supseteq \mathbf{I}(Y)$.
2. We have the following equivalences:

$$\begin{aligned} f \in \mathbf{I}(X \cup Y) &\Leftrightarrow f(\mathbf{x}) = 0 \text{ for all } \mathbf{x} \in X \cup Y \\ &\Leftrightarrow f(\mathbf{x}) = 0 \; \forall \, \mathbf{x} \in X \text{ and } f(\mathbf{x}) = 0 \; \forall \, \mathbf{x} \in Y \\ &\Leftrightarrow f \in \mathbf{I}(X) \text{ and } f \in \mathbf{I}(Y) \\ &\Leftrightarrow f \in \mathbf{I}(X) \cap \mathbf{I}(Y). \end{aligned}$$

□

The maps $\mathbf{V}$ and $\mathbf{I}$ are almost inverses to each other. We see the relation between them in the following proposition.

Proposition 2.4. *The following properties hold for any set $X \subseteq \mathbf{K}^n$ and any set $S \subseteq \mathbf{K}[x_1, ..., x_n]$:*

1. $S \subseteq \mathbf{I}(\mathbf{V}(S))$.
2. $X \subseteq \mathbf{V}(\mathbf{I}(X))$.
3. $\mathbf{V}(\mathbf{I}(\mathbf{V}(S))) = \mathbf{V}(S)$.
4. $\mathbf{I}(\mathbf{V}(\mathbf{I}(X))) = \mathbf{I}(X)$.

Proof.

1. Suppose $f \in S$. Then for all $\mathbf{x} \in \mathbf{V}(S)$ we have $f(\mathbf{x}) = 0$. So $f \in \mathbf{I}(\mathbf{V}(x))$.
2. Suppose $\mathbf{x} \in X$. Then for all $f \in \mathbf{I}(X)$ we have $f(\mathbf{x}) = 0$. So $\mathbf{x} \in \mathbf{V}(\mathbf{I}(X))$.
3. By property 2 we have $\mathbf{V}(S) \subseteq \mathbf{V}(\mathbf{I}(\mathbf{V}(S)))$.

   Let $\mathbf{x} \in \mathbf{V}(\mathbf{I}(\mathbf{V}(S)))$. Then $f(\mathbf{x}) = 0$ for all $f \in \mathbf{I}(\mathbf{V}(S))$. By property 1. $f(\mathbf{x}) = 0$ for all polynomials $f \in S$ i.e. $\mathbf{x} \in \mathbf{V}(S)$. Thus $\mathbf{V}(\mathbf{I}(\mathbf{V}(S))) \subseteq \mathbf{V}(S)$.
4. By property 1 we have $\mathbf{I}(X) \subseteq \mathbf{I}(\mathbf{V}(\mathbf{I}(X)))$.

   Let $f \in \mathbf{I}(\mathbf{V}(\mathbf{I}(X)))$. Then $f(\mathbf{x}) = 0$ for all $\mathbf{x} \in \mathbf{V}(\mathbf{I}(X))$. By property 2 $f(\mathbf{x}) = 0$ for all points $\mathbf{x} \in X$ i.e. $f \in \mathbf{I}(X)$. Thus $\mathbf{I}(\mathbf{V}(\mathbf{I}(X))) \subseteq \mathbf{I}(X)$.

□

If $f \in \mathbf{K}[x_1, ..., x_n]$ is a non-constant polynomial then at least one of $x_1, ...x_n$ appears in f. Say that x_i appears in f. If $n \geq 2$ then we can set all variables except x_i in f to constants in $\mathbf{K}$. Consider $f(a_1, ..., a_{i-1}, x_i, a_{i+1}, ..., a_n)$ for some choice of $a_1, ..., a_{i-1}, a_{i+1}, ..., a_n \in \mathbf{K}$. This is a polynomial in only one variable and so it has at least one root if $\mathbf{K}$ is algebraically closed. This holds for any choice of the constants. So if $\mathbf{K}$ is algebraically closed and therefore in particular infinite then there are infinitely many choices of constants $a_1, ..., a_{i-1}, a_{i+1}, ..., a_n \in \mathbf{K}$, all for which f has at least one root. So $\mathbf{V}(f)$ is infinite.

We will now show the existence of a special set of polynomials that will be useful later.

Proposition 2.5. *Let $\{p_1, ..., p_r\}$ be a finite subset of $\mathbf{K}^n$. Then there exists polynomials $f_1, ..., f_r \in \mathbf{K}[x_1, ..., x_n]$ such that $f_i(p_i) = 1$ and $f_i(p_j) = 0$ for all $i \neq j$.*

Proof. Let $S = \{p_1, ..., p_r\}$ and let $S_i = S \backslash \{p_i\}$. Since $\mathbf{I}(S_i) \nsubseteq \mathbf{I}(\{p_i\})$ we get, by property 2 in proposition 2.3, that $\mathbf{I}(S_i \cup \{p_i\}) = \mathbf{I}(S_i) \cap \mathbf{I}(\{p_i\}) \subsetneq \mathbf{I}(S_i)$. So there is some polynomial g_i such that $g_i \in \mathbf{I}(S_i)$ but $g_i \notin \mathbf{I}(\{p_i\})$, i.e. $g_i(p_j) = 0$ for all $j \neq i$ and $g_i(p_i) = a_i$ for some non-zero $a_i \in \mathbf{K}$. Now set $f_i = \frac{g_i}{a_i}$ for all $1 \leq i \leq r$. Then $f_1, ..., f_r$ is a set of polynomials which satisfies the proposition. □

It is often useful to be able to make a change of coordinates in $\mathbf{K}^n$. But not every change of coordinates will be nice to work with. So let us define a specific type of change of coordinates.

Definition 2.2. Let $T = (T_1, ..., T_n)$ where $T_i \in \mathbf{K}[x_1, ..., x_n]$ is a polynomial of degree 1 for all $1 \leq i \leq n$ such that T is a bijection of $\mathbf{K}^n$. Then T is called an **affine change of coordinates**.

An affine change of coordinates is a combination of a translation and an invertible linear transformation. We will use the notation f^T to mean $f(T_1, ..., T_n)$.

2.2 Hilbert's nullstellensatz

Hilbert's nullstellensatz is a fundamental theorem in algebraic geometry because it gives the correspondence between algebraic sets in $\mathbf{K}^n$ and ideals in $\mathbf{K}[x_1, ..., x_n]$, which provides us with the possibility to use an algebraic framework to determine properties of geometrical objects. We will in this section work our way towards stating and proving Hilbert's nullstellensatz.

The ring $\mathbf{K}[x_1, ..., x_n]$ has a special property that makes the ideals easier to work with. First we begin with a definition.

Definition 2.3. A commutative ring R is called **Noetherian** if every ideal in R is finitely generated.

The following is a basic theorem in algebraic geometry that lies in the background to many other deeper results.

Theorem 2.6 (Hillbert basis theorem)**.** *If R is a Noetherian ring then so is $R[x_1, ..., x_n]$.*

Proof. It is enough to show that the theorem holds for $R[x]$ since $R[x_1, ..., x_{n-1}][x_n]$ is isomorphic to $R[x_1, ..., x_n]$, so the result follows by induction.

Let I be an ideal in $R[x]$. For a polynomial $f \in I$, let the coefficient of the highest power of x be called the leading coefficient of f. Let J be the set of all leading coefficients of the polynomials in I. Then J is an ideal. Since R is Noetherian, J is finitely generated. Let $\{f_i\}_{i=1}^{s}$ be a (finite) set of polynomials in I whose leading coefficients generate J and let N be an integer larger than $\max_{1 \leq i \leq s}(\{\deg(f_i)\})$.

For each $m \leq N$ let J_m be the set of all leading coefficients of the polynomials of I of degree less than or equal to m. Then J_m is an ideal. Let $\{f_{m,j}\}_{j=1}^{s_m}$ be a (finite) set of polynomials of degree less than or equal to m whose leading coefficients generate J_m.

Now consider the ideal I' in $R[x]$ generated by $\{f_i\}_{i=1}^{s}$ and $\{f_{m,j}\}_{j=1}^{s_m}$. It is a finitely generated ideal, so if we can show that $I' = I$ then we are done.

Firstly, since the generators of I' all belong to I, $I' \subseteq I$. So it is enough to show that $I \subseteq I'$.

Suppose that $I \nsubseteq I'$. Let $g \in I$ be a polynomial of lowest degree which is not in I'. If $\deg(g) > N$, then since $g \in I$, the leading coefficient of g belongs to J. Since the leading coefficients of $\{f_i\}_{i=1}^{s}$ generate J we can find some polynomials h_i in $R[x]$ such that the leading coefficient of g is the same as the leading coefficient of $\sum_{i=1}^{s} h_i f_i$ and such that $\deg\left(\sum_{i=1}^{s} h_i f_i\right) = \deg(g)$. Then $\deg(g - \sum_{i=1}^{s} h_i f_i) < \deg(g)$. Since g is a polynomial of lowest degree not in I' and $g - \sum_{i=1}^{s} h_i f_i$ is of lower degree, $g - \sum_{i=1}^{s} h_i f_i \in I'$. But then g is a sum of two polynomials in I' and thus $g \in I'$, which is a contradiction.

If $\deg(g) = m \leq N$ then, again since the leading coefficient of g belongs to J_m and the leading coefficients of $\{f_{m,j}\}_{j=1}^{s_m}$ generate J_m we can find some polynomials $h_j \in R[x]$ such that the leading coefficient of g is the same as the leading coefficient of $\sum_{j=1}^{s_m} h_j f_{m,j}$ and such that $\deg\left(\sum_{j=1}^{s_m} h_j f_{m,j}\right) = \deg(g)$. Then $\deg(g - \sum_{j=1}^{s_m} h_i f_{m,j}) < \deg(g)$. Using the same argument as above we get $g \in I'$, which is again a contradiction.

So $I \subseteq I'$ and thus $I = I'$. □

Since the only two ideals of a field are the zero-ideal and the entire field, which are both finitely generated, $\mathbf{K}$ is a Noetherian ring. So by theorem 2.6, $\mathbf{K}[x_1, ..., x_n]$ is a Noetherian ring.

Proposition 2.7. *Any non-empty collection of ideals in a Noetherian ring has a maximal member.*

Proof. Let R be a Noetherian ring, let $\mathcal{S}$ be a non-empty collection of ideals in R and choose an ideal $I_0 \in \mathcal{S}$. Assume by way of contradiction that $\mathcal{S}$ does not contain a maximal member. Now choose an ideal $I_1 \in \mathcal{S}$ such that $I_0 \subsetneq I_1$. This is possible since otherwise I_0 would be a maximal member of $\mathcal{S}$. Continue choosing ideals in this manner, i.e. such that $I_n \subsetneq I_{n+1}$. This produces an infinite chain of ideals ordered by set inclusion, since if it were finite, then the last ideal would be a maximal member of $\mathcal{S}$. Now consider $I = \bigcup_{n=0}^{\infty} I_n$. This is an ideal and since R is a Noetherian ring, I is generated by some $f_1, ..., f_m \in R$. But then for a large enough integer N, all of $f_1, ..., f_m$ belong to I_N and therefore $I_N = I_{N+1} = ... = I$, so the chain of ideals is finite, which is a contradiction. So $\mathcal{S}$ contains a maximal member. □

Corollary 2.8. *Every proper ideal in a Noetherian ring is contained in a maximal ideal.*

Proof. Let $I \subsetneq \mathbf{K}[x_1, ..., x_n]$ be a proper ideal and let S be the set of proper ideals in $\mathbf{K}[x_1, ..., x_n]$ that contain I. By proposition 2.7, there exists some ideal

J_0 that is a maximal member of S. Now suppose, by way of contradiction, that J_0 is not a maximal ideal. Then there is some proper ideal $J' \subsetneq \mathbf{K}[x_1, ..., x_n]$ such that $J_0 \subsetneq J'$. But then $I \subsetneq J'$ and thus $J' \in S$, which contradicts the maximality of J_0 in S. Hence J_0 is a maximal ideal that contains I. □

We are interested in determining the maximal ideals of $\mathbf{K}[x_1, ..., x_n]$ since they correspond to minimal algebraic sets in $\mathbf{K}^n$. We begin with a lemma.

Lemma 2.9. *For any point $(a_1, ..., a_n) \in \mathbf{K}^n$ the ideal $(x_1 - a_1, ..., x_n - a_n)$ in $\mathbf{K}[x_1, ..., x_n]$ is a maximal ideal.*

Proof. Consider the homomorphism

$$\begin{aligned} \varphi : \mathbf{K}[x_1, ..., x_n] &\longrightarrow \mathbf{K} \\ f(x_1, ..., x_n) &\longmapsto f(a_1, ..., a_n). \end{aligned}$$

This map is surjective since for any element $a \in \mathbf{K}$, the constant polynomial $f(x_1, ..., x_n) = a$ is mapped to a.

Let $I = (x_1 - a_1, ..., x_n - a_n)$. Then clearly $I \subseteq \ker \varphi$. If we can prove the other inclusion then the theorem will follow.

We first note that $x_i \equiv a_i \pmod{I}$ for all $1 \leq i \leq n$. Thus $f(x_1, ..., x_n) \equiv f(a_1, ..., a_n) \pmod{I}$ for all $f \in \mathbf{K}[x_1, ..., x_n]$.

Let $f \in \ker \varphi$, then $f(a_1, ..., a_n) = 0$ which implies

$$f(x_1, ..., x_n) \equiv f(a_1, ..., a_n) \equiv 0 \pmod{I},$$

and thus $f(x_1, ..., x_n) \in I$. So $\ker \varphi \subseteq I$ and therefore $I = \ker \varphi$. This means that $\mathbf{K}[x_1, ..., x_n]/I \cong \mathbf{K}$ and therefore I is a maximal ideal. □

In order to prove Hilbert's nullstellensatz, we first need to prove a theorem called Zariski's lemma and for that we need some definitions.

Definition 2.4. If R is a subring of S, then S is **module-finite** over R if there is a finite subset $X \subseteq S$ such that every element of S can be written as an R-linear combination of the elements of X. Additionally, S is called **ring-finite** over R if $S = R[v_1, ..., v_n]$ for some elements $v_1, ..., v_n \in S$. An element $s \in S$ is called **integral** over R if s is the root of some monic polynomial with coefficients in R.

We begin by noting that if S is module-finite over R, with generators $v_1, ..., v_n$ and T is module-finite over S with generators $w_1, ..., w_m$ then T is module-finite over R with generators $\{v_i w_j\}$. So the property of being module-finite is transitive.

Proposition 2.10. *If R is a subring of a domain S and $v \in S$ then the following are equivalent:*

1. *v is integral over R.*
2. *$R[v]$ is module-finite over R.*
3. *There is a subring $R' \subseteq S$ that contains $R[v]$ such that R' is module-finite over R.*

Proof.

$1 \Rightarrow 2$**:** Suppose that v is integral over R. Then $v^n = a_{n-1}v^{n-1} + ... + a_0$ for some $a_{n-1}, ..., a_0 \in R$. Then for any positive integer m, v^m is an R-linear combination of $1, v, ..., v^{n-1}$ and therefore $R[v]$ is module-finite over R.

$2 \Rightarrow 3$**:** Set $R' = R[v]$.

$3 \Rightarrow 1$**:** Suppose R' is module-finite over R and let $w_1, ..., w_k$ be generators for R'. Then $vw_i \in R'$ for all $1 \leq i \leq k$ and therefore $vw_i = \sum_{j=1}^{k} a_{ij}w_i$ for some $\{a_{ij}\} \subseteq R$. Then $\sum_{j=1}^{k}(\delta_{ij}v - a_{ij})w_i = 0$, where δ_{ij} is the Kronecker delta. This gives us a system of linear equations in the quotient field of S to which $(w_1, ..., w_k)$ is a non-trivial solution. Thus the determinant of the matrix correspondning to the system is 0. Since v only appears on the main diagonal, the determinant is a monic polynomial in v with coefficients in R and thus v is algebraic over R.

□

Corollary 2.11. *The set of elements of S which are integral over R form a subring of S that contains R.*

Proof. Suppose that a and b are integral over R. Then $R[a]$ is module-finite over R and since $R \subseteq R[a]$, b is integral over $R[a]$. Thus $R[a, b] = R[a][b]$ is module-finite over $R[a]$ and therefore, by transitivity, $R[a, b]$ is module-finite over R. Then $R[a, b]$ is a subring of S that contains $R[ab]$ and since $R[a, b]$ is module-finite over R, this means that ab is integral over R. The same argument holds for $a + b$. □

We are now ready to prove Zariski's lemma, which is needed to prove Hilbert's nullstellensatz.

Theorem 2.12 (Zariski's lemma)**.** *If $\mathbf{L}$ is a field and $\mathbf{K} \subseteq \mathbf{L}$ is a subfield such that $\mathbf{L}$ is ring-finite over $\mathbf{K}$ then $\mathbf{L}$ is module-finite over $\mathbf{K}$.*

Proof. Let $\mathbf{L} = \mathbf{K}[v_1, ..., v_n]$ be a field. We will prove the theorem by induction on n.

Let $n = 1$, then $\mathbf{L} = \mathbf{K}[v_1]$. Assume that $\mathbf{L}$ is not module-finite over $\mathbf{K}$. Then the elements $1, v_1, v_1^2, ...$ are all linearly independent. Hence there is no polynomial $p \in \mathbf{K}[x]$ such that $p(v_1) = 0$. Now consider the evaluation map $\varphi : \mathbf{K}[x] \longrightarrow \mathbf{K}[v_1]$ that sends x to v_1. This is a surjective homomorphism and by the discussion above, $\ker(\varphi) = \{0\}$. So it is an isomorphism. But $\mathbf{K}[x]$ is not a field and therefore $\mathbf{L} = \mathbf{K}[v_1]$ is not a field, which is a contradicition. Thus $\mathbf{L}$ is module-finite over $\mathbf{K}$.

Now assume that the theorem holds for $n-1$ generators. We have

$$\mathbf{L} = \mathbf{K}[v_1, ..., v_n] = \mathbf{K}[v_1][v_2, ..., v_n] \subseteq \mathbf{K}(v_1)[v_2, ..., v_n] \subseteq \mathbf{K}(v_1, ..., v_n) = \mathbf{L},$$

since $\mathbf{L}$ is a field. Thus $\mathbf{L} = \mathbf{K}(v_1)[v_2, ..., v_n]$. By induction, $\mathbf{K}(v_1)[v_2, ..., v_n]$ is module-finite over $\mathbf{K}(v_1)$.

Suppose first that v_1 is not integral over $\mathbf{K}$. We will show that this leads to a contradiction.

Since $\mathbf{K}(v_1)$ is a field, $\mathbf{K}(v_1)[v_2, ..., v_n]$ being module-finite over $\mathbf{K}(v_1)$ is equivalent to saying that $\mathbf{K}(v_1)[v_2, ..., v_n]$ is a finitely generated vector space over $\mathbf{K}(v_1)$. Then all of $1, v_i, v_i^2, ...$ cannot be linearly independent for any $2 \leq i \leq n$ since then we would have infinitely many linearly independent vectors in $\mathbf{K}(v_1)[v_2, ..., v_n]$ which contradicts the fact that $\mathbf{K}(v_1)[v_2, ..., v_n]$ is a finite dimensional vector space over $\mathbf{K}(v_1)$. So for all $2 \leq i \leq n$, there is a smallest n_i such that v_i satisfies some non-trivial equation

$$v_i^{n_i} + a_{i1}v_i^{n_i-1} + ... + a_{in_i} = 0,$$

where $a_{ij} \in \mathbf{K}(v_1)$. Let a be a multiple of the denominators of all the a_{ij}:s. Then

$$(av_i)^{n_i} + aa_{i1}(av_i)^{n_i-1} + ... + a^{n_i}a_{in_i} = 0,$$

which is an equation with coefficients in $\mathbf{K}[v_1]$ and thus av_i is integral over $\mathbf{K}[v_1]$ for all $2 \leq i \leq n$.

Let $z \in \mathbf{K}[v_1, ..., v_n]$ and let $N = \deg(z)$. Then $a^N z$ is a polynomial in $av_1, ..., av_n$ with coefficients in $\mathbf{K}[v_1]$. Since all elements of $\mathbf{K}[v_1]$ are integral over $\mathbf{K}[v_1]$ we get that $a^N z$ is a combination of integral elements over $\mathbf{K}[v_1]$. Since the set of integral elements over $\mathbf{K}[v_1]$ form a ring by corollary 2.11, we conclude that $a^N z$ is integral over $\mathbf{K}[v_1]$.

Let $z = \frac{f}{g}$ be an element of $\mathbf{K}(v_1)$ and therefore also an element of $\mathbf{K}[v_1, ..., v_n]$, where g is non-constant, f and g are relatively prime and a^N and g are relatively

prime. Then $a^N z = \frac{a^N f}{g}$ where $a^N f$ and g are relatively prime. Since $a^N z$ is integral over $\mathbf{K}[v_1]$, $\frac{a^N f}{g}$ satisfies some equation

$$\left(\frac{a^N f}{g}\right)^m + b_1 \left(\frac{a^N f}{g}\right)^{m-1} + ... + b_m = 0,$$

where $b_1, ..., b_m \in \mathbf{K}[v_1]$.

If we multiply this equation with g^m we get

$$(a^N f)^m + b_1 g(a^N f)^{m-1} + ... + b_m g^m = 0,$$

and we see that g divides $(a^N f)^m$. But since $a^N f$ and g are relatively prime and g is non-constant this is a contradiction. Thus v_1 must be integral over $\mathbf{K}$.

Now consider the map $\psi : \mathbf{K}[x] \longrightarrow \mathbf{K}(v_1)$ that sends x to v_1. Since v_1 is integral over $\mathbf{K}$, $\ker(\psi)$ is non-trivial. Furthermore, $\mathbf{K}[x]/\ker(\psi)$ is isomorphic to $\mathbf{K}[v_1]$, which is an integral domain and therefore $\ker(\psi)$ is a prime ideal in $\mathbf{K}[x]$. Since $\mathbf{K}[x]$ is a principal ideal domain, $\ker(\psi)$ is a maximal ideal. Therefore $\mathbf{K}[v_1] \cong \mathbf{K}[x]/\ker(\psi)$ is a field and thus $\mathbf{K}[v_1] = \mathbf{K}(v_1)$.

By proposition 2.10, since v_1 is integral over $\mathbf{K}$, $\mathbf{K}[v_1]$ and therefore $\mathbf{K}(v_1)$ is module-finite over $\mathbf{K}$. By the induction hypothesis, $\mathbf{L} = \mathbf{K}(v_1)[v_2, ..., v_n]$ is module-finite over $\mathbf{K}(v_1)$. So by transitivity, $\mathbf{L}$ is module-finite over $\mathbf{K}$. □

Of special interest is algebraically closed fields, for which we have the following lemma.

Lemma 2.13. *An algebraically closed field* $\mathbf{K}$ *has no module-finite field extensions except itself.*

Proof. Suppose $\mathbf{L}$ is a module-finite field extension of $\mathbf{K}$. Let $\{v_1, ..., v_n\}$ be a set of generators for $\mathbf{L}$. Then for each $1 \leq i \leq n$ there is some m_i such that $1, v_i, ..., v_i^{m_i}$ are linearly dependent. Thus v_i is a root of some polynomial with coefficients in $\mathbf{K}$. Then since $\mathbf{K}$ is algebraically closed, $v_i \in \mathbf{K}$. Thus $\mathbf{L} = \mathbf{K}$. □

In order to prove Hilbert's nullstellensatz we will follow the outline of the proof by Rabinowitsch for which we first need to prove a weaker theorem.

Theorem 2.14 (Weak nullstellensatz)**.** *If* $\mathbf{K}$ *is algebraically closed and* I *is a proper ideal in* $\mathbf{K}[x_1, ..., x_n]$ *then* $\mathbf{V}(I) \neq \emptyset$*.*

Proof. It is enough to prove the theorem for maximal ideals since by corollary 2.8, every proper ideal is contained in a maximal ideal and by proposition 2.2, if $J \subseteq I$ then $\mathbf{V}(J) \supseteq \mathbf{V}(I)$.

So let I be a maximal ideal of $\mathbf{K}[x_1, ..., x_n]$. Then $\mathbf{K}[x_1, ..., x_n]/I$ is a field which is ring-finite over $\mathbf{K}$ and has $\mathbf{K}$ as a subfield. Then by Zariski's lemma (theorem 2.12), $\mathbf{K}[x_1, ..., x_n]/I$ is module-finite over $\mathbf{K}$. Since $\mathbf{K}$ is algebraically closed, lemma 2.13 then implies that $\mathbf{K}[x_1, ..., x_n]/I = \mathbf{K}$.

Then the image of x_i under the quotient map is a_i for some $a_i \in \mathbf{K}$, so $x_i - a_i \in I$ for all $1 \leq i \leq n$. But by lemma 2.9, $(x_1 - a_1, ..., x_n - a_n)$ is a maximal ideal. Therefore we must have $I = (x_1 - a_1, ..., x_n - a_n)$ and thus $\mathbf{V}(I) = \{(a_1, ..., a_n)\} \neq \emptyset$. □

In this proof we see that the maximal ideals of $\mathbf{K}[x_1, ..., x_n]$ correspond exactly to points $(a_1, ..., a_n) \in \mathbf{K}^n$.

We continue with a definition needed in order to state Hilbert's nullstellensatz.

Definition 2.5. If I is an ideal in a commutative ring R then

$$\mathrm{Rad}(I) = \{f \in R \,|\, f^n \in I \text{ for some positive integer n}\}$$

is called the **radical** of I.

We also need a lemma for the proof.

Lemma 2.15. *Let I be an ideal in $\mathbf{K}[x_1, ..., x_n]$. Then $\mathbf{V}(I) = \mathbf{V}(\mathrm{Rad}(I))$.*

Proof. Since $I \subseteq \mathrm{Rad}(I)$ we have the inclusion $\mathbf{V}(\mathrm{Rad}(I)) \subseteq \mathbf{V}(I)$. What we need to show is the inclusion $\mathbf{V}(\mathrm{Rad}(I)) \subseteq \mathbf{V}(I)$.

Suppose $\mathbf{x} \in \mathbf{V}(I)$ and $g \in \mathrm{Rad}(I)$. Then $g^n \in I$ for some positive integer n and thus $g(\mathbf{x})^n = 0$. Since $\mathbf{K}$ is a field this means that $g(\mathbf{x}) = 0$ and therefore $\mathbf{x} \in \mathbf{V}(\mathrm{Rad}(I))$. So $\mathbf{V}(I) \subseteq \mathbf{V}(\mathrm{Rad}(I))$. □

We are now ready to prove Hilbert's nullstellensatz.

Theorem 2.16 (Hilbert's nullstellensatz)**.** *Let I be an ideal in $\mathbf{K}$, where $\mathbf{K}$ is algebraically closed. Then*

$$\mathbf{I}(\mathbf{V}(I)) = \mathrm{Rad}(I).$$

Proof. Let I be an ideal in $\mathbf{K}[x_1, ..., x_n]$. By lemma 2.15 together with property 1 of proposition 2.4 we see that $\mathrm{Rad}(I) \subseteq \mathbf{I}(\mathbf{V}(I))$. Let us now prove the other inclusion.

Let $g \in \mathbf{I}(\mathbf{V}(I))$. Since $\mathbf{K}[x_1, ..., x_n]$ is Noetherian, I is generated by some finite set of polynomials $f_1, ..., f_t \in \mathbf{K}[x_1, ..., x_n]$. Now consider the ideal $J = (f_1, ..., f_t, x_{n+1}g - 1) \subseteq \mathbf{K}[x_1, ..., x_{n+1}]$. Since g vanishes whenever $f_1, ..., f_t$ all

vanish, $\mathbf{V}(J) = \emptyset$. By the weak nullstellensatz this means that $J = \mathbf{K}[x_1, ..., x_{n+1}]$ and thus $1 \in J$. So there are some polynomials $p_1, ..., p_t, q \in \mathbf{K}[x_1, ..., x_{n+1}]$ such that

$$1 = \sum_{k=1}^{t} p_k(x_1, ..., x_{n+1}) f_k(x_1, ..., x_n) + q(x_1, ..., x_{n+1})(x_{n+1}g - 1). \qquad (1)$$

Now let d be the highest power of x_{n+1} in equation (1) and set $y = \frac{1}{x_{n+1}}$. Let D be an integer larger than d and multiply equation (1) by y^D. Then we get

$$y^D = \sum_{k=1}^{t} r_k(x_1, ..., x_n, y) f_k(x_1, ..., x_n) + s(x_1, ..., x_n, y)(g - y), \qquad (2)$$

for some $r_1, ..., r_t, s \in \mathbf{K}[x_1, ..., x_{n+1}]$. By setting $y = g$ in equation (2), since g is a polynomial in $x_1, ..., x_n$ we get

$$g^D = \sum_{k=1}^{t} r'_k(x_1, ..., x_n) f_k(x_1, ..., x_n), \qquad (3)$$

for some $r'_1, ..., r'_t \in \mathbf{K}[x_1, ..., x_n]$ and thus $\mathbf{I}(\mathbf{V}(I)) \subseteq \mathrm{Rad}(I)$. □

So there is a one-to-one correspondence between the algebraic sets in $\mathbf{K}^n$ and the radical ideals in $\mathbf{K}[x_1, ..., x_n]$.

Corollary 2.17. *Let I be an ideal in $\mathbf{K}[x_1, ..., x_n]$. Then $\mathbf{V}(I)$ is a finite set if and only if $\mathbf{K}[x_1, ..., x_n]/I$ is a finite dimensional $\mathbf{K}$-vector space. If this is the case then*

$$|\,\mathbf{V}(I)| \leq \dim_{\mathbf{K}} \mathbf{K}[x_1, ..., x_n]/I.$$

Proof. Let I be an ideal of $\mathbf{K}[x_1, ..., x_n]$ and assume that $\mathbf{V}(I)$ is finite. Let $\mathbf{V}(I) = \{p_1, ..., p_s\}$, where $p_i = (a_{i1}, ..., a_{in})$, and let $f_i = \prod_{j=1}^{s}(x_i - a_{ji})$ for all $1 \leq i \leq n$. Then $f_i \in \mathbf{I}(\mathbf{V}(I))$ and thus, by Hilbert's nullstellensatz, $f_i^{m_i} \in I$ for some positive integer m_i for all $1 \leq i \leq n$. Let $N = \max(\{m_1, ..., m_s\})$, then $f_i^N \in I$ for all $1 \leq i \leq n$. This means that $\overline{f_i}^N = 0$ in $\mathbf{K}[x_1, ..., x_n]/I$ and since $\overline{f_i} = \prod_{j=1}^{s}(\overline{x_i} - \overline{a_{ji}})$, this implies that $\overline{x_i}^{sN}$ can be written as a linear combination of $\overline{1}, \overline{x_i}, ..., \overline{x_i}^{sN-1}$ with coefficients in $\mathbf{K}$. This holds for all $1 \leq i \leq n$, and therefore all elements in $\mathbf{K}[x_1, ..., x_n]/I$ can be written as a linear combination of the set $\{x_1^{r_1} \cdots x_n^{r_n} \,|\, r_j < sN\}$, with coefficients in $\mathbf{K}$. So $\dim_{\mathbf{K}} \mathbf{K}[x_1, ..., x_n]/I$ is finite.

Now assume that $\dim_{\mathbf{K}} \mathbf{K}[x_1, ..., x_n]/I$ is finite. Let $p_1, ..., p_s \in \mathbf{V}(I)$. We can by proposition 2.5 choose polynomials $f_1, ..., f_s \in \mathbf{K}[x_1, ..., x_n]$ such that $f_i(p_i) = 1$ and $f_i(p_j) = 0$ for all $i \neq j$. Let $\sum_{i=1}^{s} \lambda_i \overline{f_i}$ be a linear combination of $\overline{f_1}, ..., \overline{f_s}$

with coefficients in $\mathbf{K}$ and assume that $\sum_{i=1}^{s} \lambda_i \overline{f_i} = 0$. Then $\sum_{i=1}^{s} \lambda_i \overline{f_i} \in I$ and so $\lambda_j = \sum_{i=1}^{s} \lambda_i f_i(p_j) = 0$ for all $1 \leq j \leq s$. So the polynomials $\overline{f_1}, ..., \overline{f_s}$ are linearly independent in $\mathbf{K}[x_1, ..., x_n]/I$ and therefore $s \leq \dim_{\mathbf{K}} \mathbf{K}[x_1, ..., x_n]/I$, which is finite. □

2.3 Algebraic varieties

We say that an algebraic set is **irreducible** if it is not the union of two proper algebraic subsets, otherwise it is called **reducible**.

Definition 2.6. An irreducible algebraic set is called an **algebraic variety**.

There is a one-to-one correspondence between algebraic varieties in $\mathbf{K}^n$ and prime ideals in $\mathbf{K}[x_1, ..., x_n]$, which follows by the next two theorems.

Theorem 2.18. *An algebraic set V is irreducible if and only if $\mathbf{I}(V)$ is a prime ideal.*

Proof. Suppose $\mathbf{I}(V)$ is not prime. Then there are some $f, g \in \mathbf{K}[x_1, ...x_n]$ such that $fg \in \mathbf{I}(V)$ but $f \notin \mathbf{I}(V)$ and $g \notin \mathbf{I}(V)$. Since $fg \in \mathbf{I}(V)$, on any point $\mathbf{x} \in V$, at least one of f and g vanish. So $V = (V \cap \mathbf{V}(f)) \cup (V \cap \mathbf{V}(g))$. But since neither f nor g vanish on all of V the sets $V \cap \mathbf{V}(f)$ and $V \cap \mathbf{V}(g)$ are proper subsets of V. So V is reducible.

Suppose V is reducible, i.e. $V = V_1 \cup V_2$ where $V_1 \subsetneq V$ and $V_2 \subsetneq V$ are algebraic sets. Then $\mathbf{I}(V_1) \supsetneq \mathbf{I}(V)$ and $\mathbf{I}(V_2) \supsetneq \mathbf{I}(V)$, for suppose for example that $\mathbf{I}(V_1) = \mathbf{I}(V)$. Then $V_1 = \mathbf{V}(\mathbf{I}(V_1)) = \mathbf{V}(\mathbf{I}(V)) = V$, which is a contradiction. The same result holds for V_2. So take any $f \in \mathbf{I}(V_1) \setminus \mathbf{I}(V)$ and $g \in \mathbf{I}(V_2) \setminus \mathbf{I}(V)$. Then $fg \in \mathbf{I}(V)$ but $f \notin \mathbf{I}(V)$ and $g \notin \mathbf{I}(V)$. So $\mathbf{I}(V)$ is not prime. □

Proposition 2.19. *Prime ideals are radical.*

Proof. Let I be a prime ideal. We have the inclusion $I \subseteq \mathrm{Rad}(I)$, which holds for any ideal. We need to show $\mathrm{Rad}(I) \subseteq I$.

Suppose $x \in \mathrm{Rad}(I)$, i.e. that $x^n \in I$ for some positive integer n. We want to show that this implies that $x \in I$. We proceed by induction on n. If $n = 1$ there is nothing to prove. Now assume that $x^{n-1} \in I \Rightarrow x \in I$. Since I is a prime ideal, if $x^n \in I$ then either $x \in I$ or $x^{n-1} \in I$. But by the induction hypothesis, if $x^{n-1} \in I$ then $x \in I$. So $\mathrm{Rad}(I) \subseteq I$ and thus $I = \mathrm{Rad}(I)$. □

This, together with Hilbert's nullstellensatz, shows that the algebraic varieties of $\mathbf{K}^n$ are in a one-to-one correspondence with the prime ideals in $\mathbf{K}[x_1, ..., x_n]$.

A very useful fact in algebraic geometry is that much like any integer can be decomposed as a unique product of prime numbers, an algebraic set can be decomposed as a unique union of algebraic varieties.

Theorem 2.20. *Let V be an algebraic set in $\mathbf{K}^n$. Then there are unique irreducible algebraic sets $V_1, ..., V_m$ such that $V = V_1 \cup ... \cup V_m$ and $V_i \nsubseteq V_j$ for all $i \neq j$.*

Proof. We begin by showing the existence of the decomposition.

Let $\mathcal{S}$ be the collection of algebraic sets which are not a finite union of irreducible algebraic sets and assume by way of contradiction that $\mathcal{S} \neq \emptyset$. It follows by proposition 2.7 and the inclusion reversion between ideals and algebraic sets that any non-empty collection of algebraic sets has a minimal member. Let therefore V be a minimal member of $\mathcal{S}$. Since $V \in \mathcal{S}$ it is in particular not irreducible and therefore there are some algebraic sets $V_1 \subsetneq V$ and $V_2 \subsetneq V$ such that $V = V_1 \cup V_2$. But since V is a minimal member of $\mathcal{S}$ neither V_1 nor V_2 belong to $\mathcal{S}$. So both V_1 and V_2 are finite unions of algebraic varieties and therefore also V is a finite union of algebraic varieties, which is a contradiction. So $\mathcal{S} = \emptyset$ and thus all algebraic sets can be decomposed as finite unions of irreducible algebraic sets. We now show uniqueness of the decomposition.

Let $V = V_1 \cup ... \cup V_n$ be a decomposition of V into a finite union of algebraic varieties. We may assume, without loss of generality, that $V_i \nsubseteq V_j$ for all $i \neq j$ for if $V_i \subseteq V_j$ then $V_i \cup V_j = V_j$ so it does not add anything to the union. Now assume that also $V = W_1 \cup ... \cup W_m$ for some algebraic varieties $W_1, ..., W_m$. Then for any $1 \leq i \leq n$ we have $V_i = V_i \cap V = (V_i \cap W_1) \cup ... \cup (V_i \cap W_m)$. Since V_i is irreducible we must have $V_i = V_i \cap W_j$ for some $1 \leq j \leq m$. But then $V_i \subseteq W_j$. By a similar argument we have $W_j \subseteq V_k$ for some $1 \leq k \leq n$. But then $V_i \subseteq W_j \subseteq V_k$, so we must have $i = k$ and thus $V_i = W_j$. This holds for all $1 \leq i \leq n$, and thus the decomposition is unique. □

So the study of algebraic sets is often reduced to the study of algebraic varieties.

Since the ideal $\mathbf{I}(V)$ of an algebraic variety $V \subseteq \mathbf{K}^n$ is prime, $\mathbf{K}[x_1, ..., x_n]/\mathbf{I}(V)$ is an integral domain. This special ring is called the **coordinate ring** of V and will be very useful. It is also a natural concept since two different polynomials may define the same polynomial function on a variety. For example, the polynomials $f(x, y) = x$ and $g(x, y) = x + y^2 + x^2 - 1$ define the same function on the unit circle since $y^2 + x^2 - 1 = 0$ there. So the coordinate ring consists of equivalence classes of polynomials which define the same function on V. We will often denote the coordinate ring of V by $\Gamma(V)$.

Proposition 2.21. *Let f and g be polynomials in $\mathbf{K}[x, y]$ with no common factor. Then $\mathbf{V}(f) \cap \mathbf{V}(g) = \mathbf{V}(f, g)$ is a finite set of points in $\mathbf{K}^2$.*

Proof. Since $\mathbf{K}[x]$ is a principal ideal domain, and thus a unique factorization domain, by Gauss lemma on polynomials (for a proof, see [1, pp. 303-304]), any element that is irreducible in $\mathbf{K}[x][y]$ is also irreducible in $\mathbf{K}(x)[y]$. So if f and g have no common factor in $\mathbf{K}[x][y]$ then they have no common factor in $\mathbf{K}(x)[y]$ either. Since $\mathbf{K}(x)$ is a field, $\mathbf{K}(x)[y]$ is a PID and thus $(f, g) = (h)$ where $h \in \mathbf{K}(x)[y]$ is the greatest common divisor of f and g. But since f and g have no common factor, $\gcd(f, g) = 1$. So $af + bg = 1$.

Let $d \in \mathbf{K}[x]$ be a multiple of the denominators of a and b. Then $da \in \mathbf{K}[x, y]$ and $db \in \mathbf{K}[x, y]$. Then $daf + dbg = d$, which is in $\mathbf{K}[x]$. If $(x_0, y_0) \in \mathbf{V}(f, g)$ then $d(x_0) = 0$. But since d is a polynomial of one variable it has only finitely many zeros and therefore there are only finitely many x-coordinates in $\mathbf{V}(f, g)$. By the same argument there are only finitely many y-coordinates in $\mathbf{V}(f, g)$ and therefore $\mathbf{V}(f, g)$ is a finite set of points. □

3 Projective space

3.1 Definition and examples

Projective space is an extension of $\mathbf{K}^n$ such that we also include points at infinity. We begin with a motivating example.

Example 3.1. Consider the lines L_1 defined by $y = x$ and L_2 defined by $y = \alpha x + 1$.

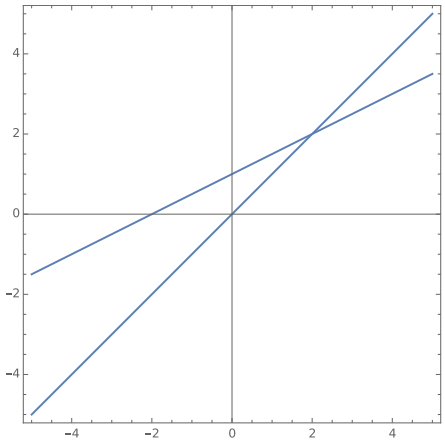

Figure 5: $\alpha = \frac{1}{2}$

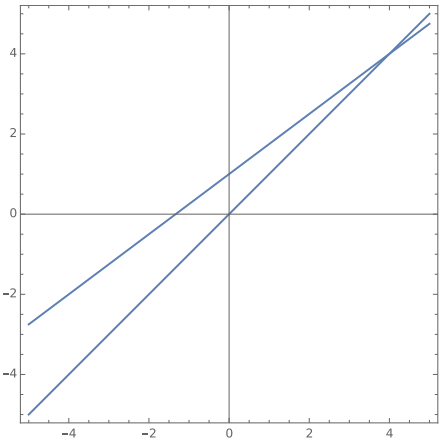

Figure 6: $\alpha = \frac{3}{4}$

If $\alpha = \frac{1}{2}$ then the lines intersect at $(2, 2)$, as seen in Figure 5. If $\alpha = \frac{3}{4}$ then they intersect at $(4, 4)$, as seen in Figure 6. More generally, if we let α approach 1 then the intersection point between the lines moves farther and farther away in the direction of the vector $(1, 1)^T$. So we would like to say that when $\alpha = 1$, i.e. when L_1 and L_2 are parallel, then L_1 and L_2 intersect "at infinity". This can be achieved in a mathematically sound way by using projective space.

Let $\mathbf{K}$ be a field. The set of lines in $\mathbf{K}^{n+1}$ passing through the origin is called **projective n-space over K** and is denoted by $\mathbf{P}^n(\mathbf{K})$. Since any point different from $(0, ..., 0)$ in $\mathbf{K}^{n+1}$ defines a unique line passing through the oirgin and two points lie on the same line if one is a multiple of the other this means that $\mathbf{P}^n(\mathbf{K})$ consists of equivalence classes of points in $\mathbf{K}^{n+1} \setminus (0, ..., 0)$ under the equivalence relation $(a_1, ..., a_{n+1}) \sim (b_1, ..., b_{n+1})$ if and only if there is some $\lambda \neq 0$ in $\mathbf{K}$ such that $(a_1, ..., a_{n+1}) = (\lambda b_1, ..., \lambda b_{n+1})$. If $(a_1, ..., a_{n+1})$ is a representative of an equivalence class in $\mathbf{P}^n(\mathbf{K})$ we will denote the class by $[a_1, ..., a_{n+1}]$, using square brackets in order to distinguish it from a point in $\mathbf{K}^{n+1}$. Observe that if $a_{n+1} \neq 0$ then $[a_1, ..., a_n, a_{n+1}] = [\frac{a_1}{a_{n+1}}, ..., \frac{a_n}{a_{n+1}}, 1]$. So any point in $\mathbf{P}^n(\mathbf{K})$ with non-zero x_{n+1}-coordinate has a unique representative on the form $(a_1, ..., a_n, 1)$. Thus any point $[a_1, ..., a_{n+1}] \in \mathbf{P}^n(\mathbf{K})$ with $a_{n+1} \neq 0$

can be uniquely identified with a point in $\mathbf{K}^n$ by dividing all coordinates by a_{n+1}, i.e. $\{[a_1, ..., a_{n+1}] \in \mathbf{P}^n(\mathbf{K}) \mid a_{n+1} \neq 0\} = \mathbf{K}^n$. So $\mathbf{K}^n$ is a subset of $\mathbf{P}^n(\mathbf{K})$.

Now consider the points in $\mathbf{P}^n(\mathbf{K})$ with x_{n+1}-coordinate equal to zero. These points make up a copy of $\mathbf{P}^{n-1}(\mathbf{K})$, i.e. $\{[a_1, ..., a_{n+1}] \in \mathbf{P}^n(\mathbf{K}) \mid a_{n+1} = 0\} = \mathbf{P}^{n-1}(\mathbf{K})$. So $\mathbf{P}^n(\mathbf{K}) = \mathbf{K}^n \cup \mathbf{P}^{n-1}(\mathbf{K})$, where we may sometimes call the elements in $\mathbf{P}^{n-1}(\mathbf{K})$ the "points at infinity".

Example 3.2. Consider for example $\mathbf{P}^2(\mathbf{K})$. By the above discussion, $\mathbf{P}^2(\mathbf{K}) = \mathbf{K}^2 \cup \mathbf{P}^1(\mathbf{K})$, where $\mathbf{P}^1(\mathbf{K})$ consists of the lines in $\mathbf{K}^2$ passing through the origin. These lines determine directions in $\mathbf{K}^2$. So $\mathbf{P}^2(\mathbf{K})$ may be seen as $\mathbf{K}^2$ together with points "at infinity" corresponding to every direction in $\mathbf{K}^2$. The set $\mathbf{P}^1(\mathbf{K}) \subseteq \mathbf{P}^2(\mathbf{K})$ is sometimes called the line at infinity.

3.2 Projective algebraic sets

Notice that the value of a polynomial $f \in \mathbf{K}[x_1, ..., x_{n+1}]$ is not in general well defined on the points in $\mathbf{P}^n(\mathbf{K})$ but will depend on the representative. However, if all terms in f are of the same degree d then $f(\lambda a_1, ..., \lambda a_{n+1}) = \lambda^d f(a_1, ..., a_{n+1})$ for all $\lambda \neq 0$ in $\mathbf{K}$. Such a polynomial is called a **homogeneous polynomial**. Therefore if f is homogeneous and $f(a_1, ..., a_{n+1}) = 0$ then $f(\lambda a_1, ..., a_{n+1}) = 0$ for all non-zero $\lambda \in \mathbf{K}$. It is therefore a well-defined notion to talk about the zero-set of a homogeneous polynomial in $\mathbf{P}^n(\mathbf{K})$.

If S is a collection of homogeneous polynomials in $\mathbf{K}[x_1, ..., x_{n+1}]$ then we let $\mathbf{V}(S)$ be the zero-set of S in $\mathbf{P}^n(\mathbf{K})$. In analogy to our definitions in $\mathbf{K}^n$, any set $X \subseteq \mathbf{P}^n(\mathbf{K})$ which is the zero-set of some collection of homogeneous polynomials is called a **projective algebraic set** and if X is irreducible then it is called a **projective variety**. We also define $\mathbf{I}(X)$ to be the set of polynomials in $\mathbf{K}[x_1, ..., x_{n+1}]$ which vanish on X and call it the ideal of X.

We will often need to go back and forth between $\mathbf{K}^n$ and $\mathbf{P}^n(\mathbf{K})$. In order to do this we need to be able to homogenize polynomials in $\mathbf{K}[x_1, ..., x_n]$ such that the zero-sets of the corresponding homogeneous polynomials are well-defined and contain the zero-sets of the original polynomials in $\mathbf{K}^n$.

Definition 3.1. A polynomial $f \in \mathbf{K}[x_1, ..., x_n]$ is called a **form** of degree d if all terms of f are of degree d.

Let $f = f_0 + ... + f_d$ be a polynomial in $\mathbf{K}[x_1, ..., x_n]$ where f_i is a form of degree i. Then we define

$$f^* := x_{n+1}^d f_0 + x_{n+1}^{d-1} f_1 + ... + f_d.$$

So f^* is a homogeneous polynomial in $\mathbf{K}[x_1, ..., x_{n+1}]$ with the property that $f^*(a_1, ..., a_n, 1) = f(a_1, ..., a_n)$. So if $f(a_1, ..., a_n) = 0$ then $f^*(a_1, ..., a_n, 1) = 0$, i.e. $\mathbf{V}(f) \subseteq \mathbf{V}(f^*)$. On the other hand, let $f \in \mathbf{K}[x_1, ..., x_{n+1}]$ be a homogeneous polynomial. Then we define

$$f_* = f(x_1, ..., x_n, 1).$$

So f_* is a polynomial in $\mathbf{K}[x_1, ..., x_n]$ that is not necessarily homogeneous.

Note that for any $f \in \mathbf{K}[x_1, ..., x_n]$ we have $(f^*)_* = f$ and for any $f \in \mathbf{K}[x_1, ..., x_{n+1}]$ we have $(f_*)^* = x_{n+1}^r f$ where x_{n+1}^r is the highest power of x_{n+1} that divides f.

Example 3.3. Consider again the lines L_1 and L_2 of Example 3.1 which are defined by the polynomials $f(x, y) = y - x$ and $g(x, y) = y - \alpha x - 1$ in $\mathbf{K}[x, y]$. We get $f^*(x, y, z) = f(x, y)$ since f was already homogeneous. But g is not a homogeneous polynomial so we get $g^*(x, y, z) = y - \alpha x - z$. We now see that if $\alpha = 1$ then $[1, 1, 0] \in \mathbf{P}^2(\mathbf{K})$ is an intersection point of f^* and g^*. This is the "point at infinity" in the direction of the vector $(1, 1)^T$ in the plane.

For two arbitrary parallel lines defined by $f(x, y) = ay + bx + c_1$ and $g(x, y) = ay + bx + c_2$ with $c_1 \neq c_2$ we get $f^*(x, y, z) = ay + bx + c_1 z$ and $g^*(x, y, z) = ay+bx+c_2 z$. If we now set $z = 0$ the constant difference between the polynomials vanishes and what is left is two identical polynomials. Thus $[-a, b, 0] \in \mathbf{P}^2(\mathbf{K})$ is an intersection point of the polynomials. So we have indeed extended $\mathbf{K}^2$ such that two parallel lines intersect at infinity and thus such that any two lines intersect once.

Example 3.4. Consider the curves defined by the polynomials $f(x, y) = y - x^2$ and $g(x, y) = x$, i.e. the parabola and the y-axis. These curves intersect at $(0, 0)$ in $\mathbf{K}^2$. If we homogenize the equations we get $f^*(x, y, z) = yz - x^2$ and $g(x, y, z) = x$. Then the curves also intersect when $x = z = 0$, i.e. on the point $[0, 1, 0] \in \mathbf{P}^2$. This can in $\mathbf{K}^2$ be seen as the point at infinity in the direction of the y-axis. The result then is reasonable since the slope of the parabola increases so that it comes closer and closer to being parallel to the y-axis and parallel lines intersect at infinity in their common direction.

3.3 Homogeneous ideals

We now continue with ideals corresponding to projective algebraic sets. An ideal $I \subseteq \mathbf{K}[x_1, ..., x_{n+1}]$ is called a **homogeneous ideal** if whenever $f = f_0 + ... + f_d$ is in I, where f_i is a form of degree i, then $f_i \in I$ for all $1 \leq i \leq d$.

Proposition 3.1. $\mathbf{I}(X)$ *is a homogeneous ideal for any set* $X \subseteq \mathbf{P}^n(\mathbf{K})$*, where* $\mathbf{K}$ *is infinite.*

Proof. Suppose $f = f_0 + ... + f_d \in \mathbf{I}(X)$. Let $(a_1, ..., a_{n+1})$ be a representative of a point in X. Then

$$f(\lambda a_1, ..., \lambda a_{n+1}) = f_0(a_1, ..., a_{n+1}) + \lambda f_1(a_1, ..., a_{n+1}) + ... + \lambda^d f_d(a_1, ..., a_{n+1})) = 0$$

for all $\lambda \in \mathbf{K} \setminus \{0\}$. So $f(\lambda a_1, ..., \lambda a_{n+1})$ is a polynomial in λ which is 0 for all non-zero values of λ and therefore it must be the zero polynomial, i.e. $f_0(a_1, ..., a_{n+1}) = ... = f_d(a_1, ..., a_{n+1}) = 0$. □

Similarly to the situation in $\mathbf{K}^n$, V is a projective variety if and only if $\mathbf{I}(V)$ is a homogeneous prime ideal in $\mathbf{K}[x_1, ..., x_{n+1}]$.

Now let I be a homogeneous ideal in $\mathbf{K}[x_1, ..., x_{n+1}]$. For an element $\overline{f} \in \mathbf{K}[x_1, ..., x_{n+1}]/I$ we say that $\overline{f}$ is a form of degree d if there is some $f \in \mathbf{K}[x_1, ..., x_{n+1}]$ which is a form of degree d such that the residue of f is $\overline{f}$.

Proposition 3.2. *Let* I *be a homogeneous ideal in* $\mathbf{K}[x_1, ..., x_{n+1}]$*. Any element* $\overline{f} \in \mathbf{K}[x_1, ..., x_{n+1}]/I$ *can be written uniquely as a sum* $\overline{f} = \overline{f}_0 + ... + \overline{f}_d$ *where* $\overline{f}_i$ *is a form degree* i*.*

Proof. Let $f \in \mathbf{K}[x_1, ...x_{n+1}]$ be such that the residue of f in $\mathbf{K}[x_1, ..., x_{n+1}]/I$ is $\overline{f}$. If $f = f_0 + ... + f_d$ where f_i is a form of degree i then $\overline{f} = \overline{f}_0 + ... + \overline{f}_d$ where $\overline{f}_i$ is a form of degree i.

Now suppose that $\overline{f} = \overline{g}_0 + ... + \overline{g}_s$ where g_j is a form of degree j. If $g_0, ..., g_s$ are polynomials whose residues are $\overline{g_0}, ..., \overline{g_s}$ respectively then $f_0 + ... + f_d - g_0 - ... - g_s \in I$. Since I is homogeneous, each term $f_i - g_i$ is in I and thus $\overline{f}_i = \overline{g}_i$. □

3.4 Projective change of coordinates

Just as for $\mathbf{K}^n$ we will sometimes need to make a change of coordinates on $\mathbf{P}^n(\mathbf{K})$. Observe that if $T : \mathbf{K}^{n+1} \longrightarrow \mathbf{K}^{n+1}$ is an invertible linear transformation then it takes lines through the origin to lines through the origin. So it takes points in $\mathbf{P}^n(\mathbf{K})$ to points in $\mathbf{P}^n(\mathbf{K})$ and we say that T is a **projective change of coordinates**.

Proposition 3.3. *Let* $\mathbf{K}$ *be an infinite field. For any finite set of points* $p_1, ..., p_n \in \mathbf{P}^2(\mathbf{K})$ *there is a line* L *which does not pass through any of the points.*

Proof. Let $S = \{p_1, ..., p_n\}$ be a finite set of points in $\mathbf{P}^2(\mathbf{K})$. Since $\mathbf{K}$ is infinite, $\mathbf{P}^2(\mathbf{K})$ is infinite and therefore $\mathbf{P}^2(\mathbf{K}) - S \neq \emptyset$. So choose some point $p \in \mathbf{P}^2(\mathbf{K}) - S$. Then p and p_i are lines in $\mathbf{K}^3$ passing through the origin for any $1 \leq i \leq n$. Since the lines intersect (at the origin), they define a unique plane in $\mathbf{K}^3$ passing through the origin. The planes in $\mathbf{K}^3$ passing through the origin define distinct lines in $\mathbf{P}^2(\mathbf{K})$. So there is a unique line, let us denote it L_i, in $\mathbf{P}^2(\mathbf{K})$ passing through both p and p_i.

The lines in $\mathbf{P}^2(\mathbf{K})$ passing through the point p correspond uniquely to the planes in $\mathbf{K}^3$ containing the line p, of which there are infinitely many. Let M be the set of lines in $\mathbf{P}^2(\mathbf{K})$ passing through p. Then $M - \{L_1, ..., L_n\} \neq \emptyset$. So there is some line L that passes through p but not any of $p_1, ..., p_n$. □

Proposition 3.4. *For any line L in $\mathbf{P}^2(\mathbf{K})$ there is a projective change of coordinates such that L is mapped to the line at infinity.*

Proof. If L is the line at infinity, the projective change of coordinates is simply the identity map and we are done. Suppose L is not the line at infinity. Then L is defined by the equation $ax + by + cz = 0$ for some $a, b, c \in \mathbf{K}$, where at least one of a and b are non-zero.

If $b \neq 0$, let T be the change of coordinates given by

$$\begin{cases} x' = z \\ y' = x \\ z' = ax + by + cz. \end{cases}$$

The determinant of the corresponding matrix is b which is non-zero, so T is invertible and therefore a projective change of coordinates.

If $b = 0$, let T be the change of coordinates given by

$$\begin{cases} x' = z \\ y' = y \\ z' = ax + by + cz. \end{cases}$$

The determinant of the corresponding matrix is $-a$ which is non-zero, so T is invertible and therefore a projective change of coordinates.

For any of the two changes of coordinates we get

$$L = \{[x, y, z] \in \mathbf{P}^2(\mathbf{K}) \,|\, ax + by + cz = 0\} = \{[x', y', z'] \in \mathbf{P}^2(\mathbf{K}) \,|\, z' = 0\},$$

which is the line at infinity. □

4 Intersection number

4.1 Properties of ideals

The intersection number is a generalization of counting the multiplicity of a root of a polynomial of one variable, which can be seen as counting the intersection multiplicity between the polynomial and the curve $y = 0$. But instead we are counting the multiplicity of the intersection of two arbitrary polynomials of arbitrary dimension at a given point. First we begin with some properties of ideals.

Proposition 4.1. *Let I and J be ideals in a ring R such that $I \subseteq J$. Then*

$$\begin{aligned} \varphi : R/I &\longrightarrow R/J \\ \overline{a} &\longmapsto \overline{a}, \end{aligned}$$

where the residue is calculated in each respective ring, is a well-defined surjective ring homomorphism.

Proof. Suppose $\overline{a} = \overline{b}$ in R/I. Then $a - b \in I \subseteq J$, and thus $\overline{a} = \overline{b}$ in R/J. So φ is well-defined.

Furthermore,

$$\begin{aligned} &\varphi(\overline{1}) = \overline{1}, \\ &\varphi(\overline{a} + \overline{b}) = \varphi(\overline{a+b}) = \overline{a+b} = \overline{a} + \overline{b} = \varphi(\overline{a}) + \varphi(\overline{b}) \quad \text{and} \\ &\varphi(\overline{a}\overline{b}) = \varphi(\overline{ab}) = \overline{ab} = \overline{a}\overline{b} = \varphi(\overline{a})\varphi(\overline{b}). \end{aligned}$$

So φ is a homomorphism that is clearly surjective. □

Definition 4.1. Let I and J be ideals of some ring R. If $I + J = R$ then I and J are called **comaximal**.

Proposition 4.2. *If I and J are comaximal ideals of a commutative ring then $IJ = I \cap J$.*

Proof. Clearly $IJ \subseteq I$ and $IJ \subseteq J$, so $IJ \subseteq I \cap J$.

Since I and J are comaximal there are some elements $a \in I$ and $b \in J$ such that $1 = a + b$. Let $c \in I \cap J$, then $c = ac + bc \in IJ$. So $I \cap J \subseteq IJ$ and thus $IJ = I \cap J$. □

Proposition 4.3. *Let $I, J, I_1, ..., I_N$ be ideals in a commutative ring R. The following properties hold:*

1. $(I_1 \cdot ... \cdot I_N)^n = I_1^n \cdot ... \cdot I_N^n$.
2. *If I and J are comaximal then so are I^m and J^n for all positive integers m and n.*
3. *If I_i and $\bigcap_{i \neq j} I_j$ are comaximal for all $1 \leq i \leq N$, then*

$$I_1^n \cap ... \cap I_N^n = (I_1 \cdot ... \cdot I_N)^n = (I_1 \cap ... \cap I_N)^n.$$

Proof.

1. This follows directly from the fact that R is commutative.
2. It is enough to show that I and J^n are comaximal for any positive integer n. The result then follows by interchanging I and J.

   If I and J are comaximal then there are some elements $a \in I$ and $b \in J$ such that $a + b = 1$. But then

$$1 = (a+b)^n = a \sum_{k=1}^{n} \left(\binom{n}{k} a^{k-1} b^{n-k} \right) + b^n \ \in I + J^n.$$

   Therefore $I + J^n = R$ and thus I and J^n are comaximal.
3. If I_i and $\bigcap_{i \neq j} I_j$ are comaximal for all $1 \leq i \leq N$ then I_i is comaximal with any intersection of ideals in $\{I_1, ..., I_N\} \backslash I_i$, for all $1 \leq i \leq N$. So by proposition 4.2:

$$(I_1 \cap ... \cap I_N)^n = (I_1 \cdot (I_2 \cap ... \cap I_N))^n = (I_1 I_2 \cdot (I_3 \cap ... \cap I_N))^n = ... = (I_1 \cdot ... \cdot I_N)^n.$$

   By 1 and 2 together with the fact that I_i and I_j are comaximal for all $i \neq j$ and again using proposition 4.2 we then get

$$(I_1 \cdot ... \cdot I_N)^n = I_1^n \cdot ... \cdot I_N^n = I_1^n \cap ... \cap I_N^n.$$

□

Proposition 4.4. *Let I be an ideal in a ring R such that* $\mathrm{Rad}(I)$ *is finitely generated. Then* $(\mathrm{Rad}(I))^n \subseteq I$ *for some positive integer n.*

Proof. Let $\mathrm{Rad}(I)$ be generated by $\{a_1, ..., a_r\}$. Then $(\mathrm{Rad}(I))^n$ is generated by $\{a_1^{i_1} \cdot ... \cdot a_r^{i_r} \mid \sum_{k=1}^{r} i_k = n\}$.

For all $1 \leq i \leq r$ we have $a_i^{m_i} \in I$ for some positive integer m_i. Let $m = \max(\{m_1, ..., m_N\})$. Then $a_i^m \in I$ for all $1 \leq i \leq r$. Now let $n = r(m-1)+1$. Then for any element of the form $a_1^{i_1} \cdot ... \cdot a_r^{i_r}$ where $\sum_{k=1}^{r} i_k = n$ there is at least one exponent i_k such that $i_k \geq m$. Therefore $a_k^{i_k} \in I$ and thus $a_1^{i_1} \cdot ... \cdot a_r^{i_r} \in I$. Then since all the generators of $(\mathrm{Rad}(I))^n$ belong to I, we have $(\mathrm{Rad}(I))^n \subseteq I$. □

Proposition 4.5. *Let I and J be ideals in $\mathbf{K}[x_1, ..., x_n]$, where $\mathbf{K}$ is algebraically closed. Then I and J are comaximal if and only if $\mathbf{V}(I) \cap \mathbf{V}(J) = \emptyset$.*

Proof. Firstly, note that $\mathbf{V}(I) \cap \mathbf{V}(J) = \mathbf{V}(I+J)$. Then by the weak nullstellensatz, $\mathbf{V}(I+J) = \emptyset$ if and only if $I + J = \mathbf{K}[x_1, ..., x_n]$, i.e. if and only if I and J are comaximal. □

4.2 The local ring

Let V be an affine variety and let $\Gamma(V)$ be the coordinate ring of V. Since $\Gamma(V)$ is a domain, we may form its quotient field. Let the quotient field of $\Gamma(V)$ be denoted by $\mathbf{K}(V)$. This ring consists of rational polynomial functions on V. In analogy to the rational functions on $\mathbb{R}$ we say that an element $\frac{f}{g} \in \mathbf{K}(V)$ is undefined at a point $p \in \mathbf{K}^n$ if $g(p) = 0$ and defined at p if $g(p) \neq 0$. This allows us to form another ring, called the **local ring** of V at p, which consists of the set of rational functions in $\mathbf{K}(V)$ which are defined at p and it is denoted by $\mathcal{O}_p(V)$. This ring will be used when defining the intersection number between two algebraic varieties.

Theorem 4.6. *Let I be an ideal in $\mathbf{K}[x_1, ..., x_n]$ such that $\mathbf{V}(I) = \{p_1, ..., p_N\}$ is finite. Then*

$$\mathbf{K}[x_1, ..., x_n]/I \cong \prod_{i=1}^{N} \mathcal{O}_{p_i}(\mathbf{K}^n)/I\,\mathcal{O}_{p_i}(\mathbf{K}^n).$$

Proof. Let $R = \mathbf{K}[x_1, ..., x_n]/I$ and $R_i = \mathcal{O}_{p_i}(\mathbf{K}^n)/I\,\mathcal{O}_{p_i}(\mathbf{K}^n)$. Let $I_i = \mathbf{I}(\{p_i\})$ for all $1 \leq i \leq N$. Since the maximal ideals of $\mathbf{K}[x_1, ..., x_n]$ correspond to points in $\mathbf{K}^n$ this means that $\{I_1, ..., I_N\}$ are the distinct maximal ideals which contain I.

For each i, there is a homomorphism $\varphi_i : R \longrightarrow R_i$. This follows from the fact that we have the following sequence of homomorphisms:

$$\mathbf{K}[x_1, ..., x_n] \xhookrightarrow{\psi_i} \mathcal{O}_{p_i}(\mathbf{K}^n) \xrightarrow{\pi_i} \mathcal{O}_{p_i}(\mathbf{K}^n)/I\,\mathcal{O}_{p_i}(\mathbf{K}^n),$$

where ψ_i is the inclusion homomorphism and π_i is the quotient homomorphism. The kernel of $\pi_i \circ \psi_i$ is exactly I, so there is a homomorphism $\varphi_i : \mathbf{K}[x_1, ..., x_n]/I \longrightarrow \mathcal{O}_{p_i}(\mathbf{K}^n)/I\,\mathcal{O}_{p_i}(\mathbf{K}^n)$. This is the homomorphism that sends an equivalence class $\overline{f}$ in R to the corresponding equivalence class $\overline{f}$ in R_i.

The homomorphisms φ_i define a homomorphism $\varphi : R \to \prod_{i=1}^N R_i$ by setting $\varphi(\overline{f}) = (\varphi_1(\overline{f}), ..., \varphi_N(\overline{f}))$. We will show that φ is an isomorphism.

First, we note that $\mathrm{Rad}(I) = \mathbf{I}(\{p_1, ..., p_N\}) = \bigcap_{i=1}^N I_i$ by Hilbert's nullstellensatz (theorem 2.16). By proposition 4.4 we have $(\bigcap_{i=1}^N I_i)^d \subseteq I$ for some positive integer d. Then, by proposition 4.5, I_i and $\bigcap_{i\neq j} I_j$ are comaximal, and therefore, by proposition 4.3, $\bigcap_{i=1}^N I_i^d = (\bigcap_{i=1}^N I_i)^d \subseteq I$.

Now, by proposition 2.5, let $f_1, ..., f_N \in \mathbf{K}[x_1, ..., x_n]$ be polynomials such that $f_i(p_i) = 1$ and $f_i(p_j) = 0$ for all $i \neq j$ and set $e_i = 1-(1-f_i^d)^d$. If we expand the parentheses in e_i then we see that the resulting sum has no constant term and every term contains some power of f_i^d. So $e_i = f_i^d q_i$ for some $q_i \in \mathbf{K}[x_1, ..., x_n]$ and thus, since $f_i \in I_j$ for all $i \neq j$, we have $e_i \in I_j^d$ for all $i \neq j$. Furthermore,

$$1 - \sum_{i=1}^N e_i = (1 - e_j) - \sum_{i\neq j} e_i,$$

where $(1 - e_j) \in I_j^d$ and $\sum_{i\neq j} e_i \in I_j^d$. So $1 - \sum_{i=1}^N e_i \in \bigcap_{i=1}^N I_i^d \subseteq I$.

Now let $\overline{e}_i$ be the residue of e_i in R. Then $\sum_{i=1}^N \overline{e}_i = 1$ and $e_i(1 - e_i) \in I$, so $\overline{e}_i^2 = \overline{e}_i$. Furthermore, $e_i e_j \in I$ and thus $\overline{e}_i\overline{e}_j = 0$ for all $i \neq j$.

Let $g_i \in \mathbf{K}[x_1, ..., x_n]$ be a polynomial such that $g_i(p_i) \neq 0$. Without loss of generality we may assume that $g_i(p_i) = 1$. Set $h_i = 1 - g_i$, then

$$g_i(e_i + h_i e_i + ... + h_i^{d-1} e_i) = (1 - h_i)(e_i + h_i e_i + ... + h_i^{d-1} e_i) = e_i - h_i^d e_i.$$

Since $h_i(p_i) = 0$ we have $h_i \in I_i$ and thus $h_i^d \in I_i^d$. This together with the fact that $e_i \in I_j^d$ for all $j \neq i$ means that $h_i^d e_i \in \bigcap_{i=1}^N I_i \subseteq I$. Thus, if we set $t_i = e_i + h_i e_i + ... + h_i^{d-1} e_i$, then $\overline{g}_i\overline{t}_i = \overline{e}_i$.

Now let $\overline{f} \in R$ be such that $\varphi(\overline{f}) = 0$. Then $\varphi_i(\overline{f}) = 0$ for all $1 \leq i \leq N$. So $f \in I\,\mathcal{O}_{p_i}(\mathbf{K}^n)$ which implies that there is some $g_i \in \mathbf{K}[x_1, ..., x_n]$ such that $g_i(p_i) \neq 0$ and $g_i f \in I$. Let t_i be defined as above, then

$$\overline{f} = \sum_{i=1}^N \overline{e}_i\overline{f} = \sum_{i=1}^N \overline{t}_i\overline{g}_i\overline{f} = 0,$$

so φ is injective.

Let $r = (\overline{\frac{a_1}{b_1}}, ..., \overline{\frac{a_N}{b_N}}) \in \prod_{i=1}^{N} R_i$. Since $e_i(p_i) = 1$, we have $\frac{1}{e_i} \in \mathcal{O}_{p_i}(\mathbf{K}^n)$. Thus $\varphi_i(\overline{e_i}) \cdot \overline{\left(\frac{1}{e_i}\right)} = \overline{e_i} \cdot \overline{\left(\frac{1}{e_i}\right)} = 1$, so $\varphi_i(\overline{e_i})$ is a unit in R_i. Therefore

$$\varphi_i(\overline{e_i})\varphi_i(\overline{e_j}) = \varphi_i(\overline{e_i e_j}) = \varphi_i(0) = 0$$

implies that $\varphi_i(\overline{e_j}) = 0$ for all $j \neq i$. This then gives

$$\varphi_i(\overline{e_i}) = \varphi_i\left(\sum_{i=1}^{N} \overline{e_i}\right) = \varphi_i(1) = 1.$$

For all $1 \leq i \leq N$ we have $b_i(p_i) \neq 0$, and thus, with $\overline{t_i}$ defined similarly as above, we have $\overline{t_i}\overline{b_i} = \overline{e_i}$. Then

$$\frac{\overline{a_i}}{\overline{b_i}} = \frac{\overline{a_i}\overline{t_i}}{\overline{e_i}} = \frac{\overline{a_i}\overline{t_i}\overline{e_i}}{\overline{e_i}^2} = \frac{\overline{a_i}\overline{t_i}\overline{e_i}}{\overline{e_i}} = \overline{a_i}\overline{t_i},$$

and thus $\varphi_i\left(\sum_{j=1}^{N} \overline{a_j}\overline{t_j}\overline{e_j}\right) = \varphi_i(\overline{a_i}\overline{t_i}) = \overline{a_i}\overline{t_i} = \frac{\overline{a_i}}{\overline{b_i}}$. So $\varphi\left(\sum_{i=1}^{N} \overline{a_i}\overline{t_i}\overline{e_i}\right) = r$ and therefore φ is surjective. □

One immediate corollary is as follows.

Corollary 4.7. *If I is an ideal of $\mathbf{K}[x_1, ..., x_n]$ with finite zero-set $\mathbf{V}(I) = \{p_1, ..., p_N\}$, then*

$$\dim_{\mathbf{K}}(\mathbf{K}[x_1, ..., x_n]/I) = \sum_{i=1}^{N} \dim_{\mathbf{K}}(\mathcal{O}_{p_i}(\mathbf{K}^n)/I\,\mathcal{O}_{p_i}(\mathbf{K}^n)).$$

4.3 Intersection number

We are now ready to define the intersection number in $\mathbf{K}^2$. We begin with some examples and properties that the intersection number should have.

When two curves intersect at a point at which their tangents are different, i.e. they "cross" each other, it is natural to say that the intersection number between the curves is 1 there.

Example 4.1. For example, we would like to say that the intersection number of the curves $y = x$ and $y = -x$ at $(0, 0)$ is 1.

Example 4.2. For the curve in Figure 3, it is natural to expect that the intersection number at $(0, 0)$ with the x-axis be 2, since the curve "crosses" the x-axis twice there.

Furthermore, since

$$\begin{cases} f(x,y)=0 \\ g(x,y)=0 \end{cases} \quad \text{and} \quad \begin{cases} f(x,y)=0 \\ g(x,y)+f(x,y)h(x,y)=0 \end{cases}$$

are equivalent systems of equations for all $h \in \mathbf{K}[x,y]$, it is natural to expect that the intersection number between f and g at a given point should be the same as the intersection number between f and $g+fh$ at that point, for any h.

More specifically, if we denote the intersection number between f and g at a point p by $I_p(f,g)$, it is natural that the intersection number should have the following properties:

1. If f and g do not have a common factor that passes through p then $I_p(f,g)$ is a non-negative integer and $I_p(f,g)=\infty$ otherwise.
2. $I_p(f,g)=0$ if and only if $p \notin \mathbf{V}(f)\cap\mathbf{V}(g)$ and $I_p(f,g)$ depends only on the factors of f and g which pass through p.
3. $I_p(f,g)=I_p(g,f)$.
4. For any affine change of coordinates T, if $T(q)=p$ then $I_q(f^T,g^T)=I_p(f,g)$.
5. $I_{(0,0)}(x,y)=1$.
6. $I_p(f,gh)=I_p(f,g)+I_p(f,h)$.
7. $I_p(f,g)=I_p(f,g+fh)$ for any $h\in\mathbf{K}[x,y]$.

There is only one number which satisfies the above properties, and it is given by

$$I_p(f,g)=\dim_{\mathbf{K}}(\mathcal{O}_p(\mathbf{K}^2)/(f,g)).$$

Theorem 4.8. *The intersection number defined above is the only number that satisfies properties 1-7.*

Proof.

1. If f and g do not have a common factor then by proposition 2.21, $\mathbf{V}(f,g)$ is finite and therefore by corollary 4.7 and corollary 2.17,

$$\dim_{\mathbf{K}}(\mathbf{K}[x_1,...,x_n]/(f,g))=\sum_{p\in\mathbf{V}(f,g)}\dim_{\mathbf{K}}(\mathcal{O}_{p_i}(\mathbf{K}^2)/I\,\mathcal{O}_{p_i}(\mathbf{K}^2))=\sum_{p\in\mathbf{V}(f,g)}I_p(f,g)$$

is finite, and thus all of $I_p(f,g)$ are finite.

If f and g have a common factor h then $(f,g) \subseteq (h)$. Then, by proposition 4.1, we have a surjective homomorphism

$$\varphi : \mathcal{O}_p(\mathbf{K}^2)/(f,g) \longrightarrow \mathcal{O}_p(\mathbf{K}^2)/(h).$$

Thus $\dim_{\mathbf{K}}(\mathcal{O}_p(\mathbf{K}^2)/(f,g)) \geq \dim_{\mathbf{K}}(\mathcal{O}_p(\mathbf{K}^2)/(h))$.

Let $V = \mathbf{V}(h)$ and consider the homomorphism $\psi : \mathcal{O}_p(\mathbf{K}^2) \longrightarrow \mathcal{O}_p(V)$ defined by $\psi(\frac{f}{g}) = \frac{\overline{f}}{\overline{g}}$. This is a surjective homomorphism with kernel $\ker(\psi) = (h)$. Thus $\mathcal{O}_p(\mathbf{K}^2)/(h) \cong \mathcal{O}_p(V)$, so $\dim_{\mathbf{K}}(\mathcal{O}_p(\mathbf{K}^2)/(h)) = \dim_{\mathbf{K}}(\mathcal{O}_p(V))$. But $\Gamma(V) \subseteq \mathcal{O}_p(V)$, so $\dim_{\mathbf{K}}(\Gamma(V)) \leq \dim_{\mathbf{K}}(\mathcal{O}_p(V))$ and by proposition 2.17, $\dim_{\mathbf{K}}(\Gamma(V)) = \infty$ since $\mathbf{V}(h)$ is infinite. Therefore $\dim_{\mathbf{K}}(\mathcal{O}_p(\mathbf{K}^2)/(f,g)) = \infty$.

2. If $f(p) \neq 0$ then $\frac{1}{f} \in \mathcal{O}_p(\mathbf{K}^2)$, so f is a unit and thus $(f,g) = \mathcal{O}_p(\mathbf{K}^2)$. Then $\dim_{\mathbf{K}}(\mathcal{O}_p(\mathbf{K}^2)/(f,g)) = \dim_{\mathbf{K}}(1) = 0$. The same argument holds for g.

   On the other hand, if $f(p) = g(p) = 0$ then for any $h \in (f,g)$, also $h(p) = 0$. So all elements in (f,g) are non-units and thus $(f,g) \subsetneq \mathcal{O}_p(\mathbf{K}^2)$ and therefore $\dim_{\mathbf{K}}(\mathcal{O}_p(\mathbf{K}^2)/(f,g)) > 0$.

   Suppose, without loss of generality, that $f = f_1 \cdot f_2$ where $f_1(p) \neq 0$ and $f_2(p) = 0$. Then $\frac{1}{f_1} \in \mathcal{O}_p(\mathbf{K}^2)$, so f_1 is a unit. Hence $(f_2, g) = (f_1 f_2, g) = (f,g)$.

3. This follows from the fact that $(f,g) = (g,f)$.

4. Let T be an affine change of coordinates such that $T(p) = q$. Then we get a homomorphism

   $$\widetilde{T} : \mathcal{O}_q(\mathbf{K}^n) \longrightarrow \mathcal{O}_p(\mathbf{K}^n)$$

   by setting $\widetilde{T}\left(\frac{f}{g}\right) = \frac{f(T_1,\ldots,T_n)}{g(T_1,\ldots,T_n)}$. Suppose $g(q) \neq 0$, then $\widetilde{T}(g)(p) = g(T(p)) = g(q) \neq 0$. So the image of $\widetilde{T}$ lies in $\mathcal{O}_p(\mathbf{K}^n)$.

   Now suppose that $\widetilde{T}\left(\frac{f}{g}\right) = 0$, then $\frac{f}{g} \circ T = 0$. Since T is invertible, this implies that $\frac{f}{g} = 0$. So $\widetilde{T}$ is injective.

   Furthermore, suppose that $\frac{f}{g} \in \mathcal{O}_p(\mathbf{K}^n)$. Then, again since T is invertible, there is some $\frac{f'}{g'} \in \mathcal{O}_q(\mathbf{K}^n)$ such that $\frac{f'}{g'} \circ T = \frac{f}{g}$. i.e. $\widetilde{T}\left(\frac{f'}{g'}\right) = \frac{f}{g}$. So $\widetilde{T}$ is surjective and is thus an isomorphism. Therefore

   $$\dim_{\mathbf{K}}(\mathcal{O}_q(\mathbf{K}^2)/(f^T, g^T)) = \dim_{\mathbf{K}}(\mathcal{O}_p(\mathbf{K}^2)/(f,g)).$$

5. It is clear that $\mathcal{O}_p(\mathbf{K}^2)/(x,y) \cong \mathbf{K}$, so $\dim_{\mathbf{K}}(\mathcal{O}_p(\mathbf{K}^2)/(x,y) = 1$.

6. The property clearly holds if f and gh have common factors. So we may assume that f and gh do not have any common factors. Then define the following $\mathbf{K}$-linear map:

$$\begin{aligned} \psi : \mathcal{O}_p(\mathbf{K}^2)/(f,h) &\longrightarrow \mathcal{O}_p(\mathbf{K}^2)/(f,gh) \\ \overline{z} &\longmapsto \overline{gz}. \end{aligned}$$

   Now suppose that $\psi(\overline{z}) = 0$. Then $gz = uf + vgh$ for some $u, v \in \mathcal{O}_p(\mathbf{K}^2)$. Let $s \in \mathbf{K}[x,y]$ be a multiple of the denominators of u, v and z such that $s(p) \neq 0$. Then $g(sz - svh) = suf$ in $\mathbf{K}[x,y]$. Since f and g have no common factors, f must divide $sz - svh$. So $sz - svh = df$ for some $d \in \mathbf{K}[x,y]$. Thus $sz = df + svh$ and hence $z = \frac{d}{s}f + vh$. So $\overline{z} = 0$ and therefore ψ is injective.

   Since $(f,gh) \subseteq (f,g)$, by proposition 4.1 there is a surjective homomorphism φ from $\mathcal{O}_p(\mathbf{K}^2)/(f,gh)$ to $\mathcal{O}_p(\mathbf{K}^2)/(f,g)$. So we have the following sequence of homomorphisms:

$$\mathcal{O}_p(\mathbf{K}^2)/(f,h) \xrightarrow{\psi} \mathcal{O}_p(\mathbf{K}^2)/(f,gh) \xrightarrow{\varphi} \mathcal{O}_p(\mathbf{K}^2)/(f,g).$$

   For any $\overline{gz} \in \mathrm{im}(\psi)$, we have $gz \in (f,g)$ and hence $\overline{gz} \in \ker(\varphi)$. So $\mathrm{im}(\psi) \subseteq \ker(\varphi)$.

   Furthermore, for any $\overline{a} \in \ker(\varphi)$ we have $a = bf + cg$ for some $b, c \in \mathcal{O}_p(\mathbf{K}^2)$ and hence $\overline{a} = \overline{cg} \in \mathrm{im}(\psi)$. So $\ker(\varphi) \subseteq \mathrm{im}(\psi)$ and thus $\mathrm{im}(\psi) = \ker(\varphi)$ and the sequence is exact. Then, by the rank-nullity theorem,

$$\dim_{\mathbf{K}}(\mathcal{O}_p(\mathbf{K}^2)/(f,gh)) = \dim_{\mathbf{K}}(\mathcal{O}_p(\mathbf{K}^2)/(f,g)) + \dim_{\mathbf{K}}(\mathcal{O}_p(\mathbf{K}^2)/(f,h)).$$

7. This follows from the fact that $(f,g) = (f, g + fh)$ for any $h \in \mathbf{K}[x,y]$.

In order to show that the intersection number is unique it is enough to show that it can be calculated using only properties 1-7.

We may assume that $p = (0,0)$, by property 4. So in particular both f and g have constant term 0. Furthermore, we may assume that f and g do not have a common factor that passes through p, since otherwise property 1 gives the intersection number. We will use induction on the value n of the intersection number. The situation when $n = 0$ falls under property 2 and thus can be calculated using only properties 1-7. So assume by induction that the intersection number can be computed for all non-negative integers smaller than n. Consider

$f(x,0)$ and $g(x,0)$ in $\mathbf{K}[x]$ and let $r = \deg(f(x,0))$ and $s = \deg(g(x,0))$. By property 4 we may assume, without loss of generality, that $r \leq s$.

If $r = 0$ then y divides f and thus $f = yh$ for some $h \in \mathbf{K}[x,y]$. Then $I_p(f,g) = I_p(y,g) + I_p(h,g)$. Since y divides $g(x,y) - g(x,0)$, by property 7 we have $I_p(y,g) = I_p(y,g(x,0))$. Let x^m be the largest power of x which divides $g(x,0)$. Then $\left(\frac{g(x,0)}{x^m}\right)(0,0) \neq 0$, and hence by property 6 and 2 we have $I_p(y,g(x,0)) = I_p(y,x^m) = m$. Since g has no constant term, $m > 0$ and thus, $I_p(h,g) < n$ and can therefore be computed, by our induction hypothesis.

Suppose $r > 0$. By property 6 and 2, we may multiply f and g by constants so that they are monic. Let $h = g - x^{s-r}f$, then by property 7, $I_p(f,g) = I_p(f,h)$ where $\deg(h(x,0)) < s$. This process can be repeated finitely many times, possibly interchanging f and h, until we get a pair of polynomials a and b such that one of $\deg(a(x,0))$ and $\deg(b(x,0))$ is 0, and $I_p(f,g) = I_p(a,b)$. Then the result follows by the above discussion. □

We will need to use intersection numbers in $\mathbf{P}^2$, so we need to generalize some earlier notions in order to do this.

If f defines a curve in $\mathbf{P}^2$, i.e. f is a homogeneous polynomial of some degree d, then $\frac{f(x,y,z)}{z^d} = f_*(\frac{x}{z}, \frac{y}{z})$. So on the subset $\mathbf{K}^2$ of $\mathbf{P}^2$, i.e. on the set of elements of $\mathbf{P}^2$ with representative with z-coordinate equal to 1, the equation becomes $f(x,y,1) = f_*(x,y)$, which is exactly our previous definition of f_*. So we may therefore extend f_* in the following way: If $p_1, ..., p_r$ is a finite set of points in $\mathbf{P}^2$ then by proposition 3.3 we may find a line L which does not pass through any of the points. Then, if d is the degree of f we define $f_* = \frac{f}{L^d}$. By the previous discussion this coincides with our earlier definition when L is the line at infinity, i.e. when $L = z$.

If L' is another line that does not pass through any of $p_1, ..., p_r$ then $\frac{f}{(L')^d} = (\frac{L}{L'})^d f_*$, where $(\frac{L}{L'})^d$ is a unit in every $\mathcal{O}_{p_i}(\mathbf{P}^2)$. Therefore, any ideal in $\mathcal{O}_{p_i}(\mathbf{P}^2)$ where f_* is one of the generators does not depend on the choice of L for any $1 \leq i \leq r$. So for any two projective curves f and g and point $p \in \mathbf{P}^2$ we may define $I_p(f,g) = \dim_{\mathbf{K}}(\mathcal{O}_p(\mathbf{P}^2)/(f_*, g_*))$, which is then independent on how f_* and g_* are formed. Furthermore, it satisfies properties 1-7 of intersection numbers but with the change that T in property 4 should be a projective change of coordinates.

5 Bezout's theorem

5.1 Projective plane curves

Before we prove Bezout's theorem we must show some properties of projective plane curves, which are zero-sets in $\mathbf{P}^2(\mathbf{K})$ of one polynomial in $\mathbf{K}[x, y, z]$.

Proposition 5.1. *Let f and g be homogeneous polynomials in $\mathbf{K}[x, y, z]$ which have no common factor. Then $\mathbf{V}(f) \cap \mathbf{V}(g) \subseteq \mathbf{P}^2(\mathbf{K})$ is finite.*

Proof. We begin by observing that f_* and g_* have no common factor. Indeed, suppose $h \in \mathbf{K}[x, y]$ is a common factor of f_* and g_*. Then $f_* = ha$ and $g_* = hb$ for some $a, b \in \mathbf{K}[x, y]$. Now let z^s and z^r be the highest powers of z dividing f and g respectively. Then $f = z^s(f_*)^* = z^s h^* a^*$ and $g = z^r(g_*)^* = z^r h^* b^*$. But then h^* is a common factor of f and g, which is a contradiction.

Since f_* and g_* are two polynomials in $\mathbf{K}[x, y]$ with no common factors, by proposition 2.21 they have a finite number of intersection points in $\mathbf{K}^2 \subseteq \mathbf{P}^2(\mathbf{K})$.

We now consider intersection points at infinity. Since f and g have no common factors at least one of them is not a multiple of z. We may, without loss of generality, assume that f is not a multiple of z. Then $f_0(x, y) := f(x, y, 0)$ is a non-zero polynomial in $\mathbf{K}[x, y]$.

For any choice of y, $(1, y, 0)$ is a representative of a distinct point in $\mathbf{P}^2(\mathbf{K})$. If $f_0(1, y)$ is constant then f_0 is a polynomial in only x and therefore has finitely many zeros. If $f_0(1, y)$ is a proper polynomial in y, then it has finitely many zeros. So f_0 has only finitely many zeros of the form $[1, y, 0]$.

In both cases f_0 may additionally have a zero at $[0, 1, 0] \in \mathbf{P}^2(\mathbf{K})$. But in total f_0 will only have a finite number of zeros. This means that f has only a finite number of zeros at infinity and therefore f and g only have a finite number of intersection points at infinity. Thus f and g have only a finite number of intersection points in all of $\mathbf{P}^2(\mathbf{K})$. □

Let $f, g \in \mathbf{K}[x, y, z]$ and let $\Gamma := \mathbf{K}[x, y, z]/(f, g)$ and let Γ_d be the $\mathbf{K}$-vector space of all forms of degree d in Γ.

Lemma 5.2. *If f and g are non-zero polynomials in $\mathbf{K}[x, y, z]$ which have no common factor then*

$$\dim_{\mathbf{K}} \Gamma_d = \deg(f) \cdot \deg(g)$$

for all $d \geq \deg(f) + \deg(g)$.

Proof. Let $R := \mathbf{K}[x, y, z]$ and let $\pi : R \longrightarrow \Gamma$ be the quotient map. Define $\varphi : R \times R \longrightarrow R$ by $\varphi(a, b) = af + bg$. Then it is clear that $\mathrm{im}(\varphi) = \ker(\pi)$.

Let $\psi : R \longrightarrow R \times R$ be defined by $\psi(c) = (cg, -cf)$. Then clearly $\mathrm{im}(\psi) \subseteq \ker(\varphi)$. Now suppose $(a, b) \in \ker(\varphi)$, i.e. $af + bg = 0$. Then $af = -bg$ and since f and g do not have any common factor we must have $f \mid b$ and $g \mid a$. So $a = c_1 g$ and $b = c_2 f$ for some $c_1, c_2 \in \mathbf{K}$. We then have

$$c_1 gf = -c_2 fg \quad \Leftrightarrow \quad c_1 = -c_2.$$

So if we let $c = c_1$ then $a = cg$ and $b = -cf$ and thus $\ker(\varphi) \subseteq \mathrm{im}(\psi)$, so $\mathrm{im}(\psi) = \ker(\varphi)$.

Furthermore, suppose $c \in \ker(\psi)$. Then $cg = 0 = -cf$. So if none of f and g are 0 then $c = 0$. So ψ is injective. Thus

$$0 \longrightarrow R \xrightarrow{\psi} R \times R \xrightarrow{\varphi} R \xrightarrow{\pi} \Gamma \longrightarrow 0$$

is an exact sequence.

Let $\deg(f) = m$ and $\deg(g) = n$ and let the $\mathbf{K}$-vector space of forms in R of degree d be denoted by R_d. Then, if $d \geq m + n$ we may restrict the above maps to get the following sequence:

$$0 \longrightarrow R_{d-m-n} \xrightarrow{\psi} R_{d-m} \times R_{d-n} \xrightarrow{\varphi} R_d \xrightarrow{\pi} \Gamma_d \longrightarrow 0.$$

This sequence is still exact. Thus, by the rank-nullity theorem, we get

$$\dim_{\mathbf{K}}(\Gamma_d) = \dim_{\mathbf{K}}(R_d) - \dim_{\mathbf{K}}(R_{d-m} \times R_{d-n}) + \dim_{\mathbf{K}}(R_{d-m-n}).$$

Observe that there are $\frac{(d+1)(d+2)}{2}$ monomials of degree d in R. So $\dim_{\mathbf{K}}(R_d) = \frac{(d+1)(d+2)}{2}$. We then get, after some calculations, that $\dim_{\mathbf{K}}(\Gamma_d) = mn$. □

5.2 Bezout's theorem

We are now finally ready to state and prove Bezout's theorem.

Theorem 5.3 (Bezout's theorem)**.** *Let* $\mathbf{K}$ *be an algebraically closed field and let* f *and* g *be non-zero polynomials in* $\mathbf{K}[x, y, z]$ *which have no common factor. Then*

$$\sum_{p \in \mathbf{P}^2} I_p(f, g) = \deg(f) \cdot \deg(g).$$

Proof. Since, by proposition 5.1, the number of intersection points of f and g in $\mathbf{P}^2(\mathbf{K})$ is finite, we may, by propositions 3.3 and 3.4 find a projective, and thus linear, change of coordinates such that none of the intersection points lie on the line at infinity. Since the intersection number is unchanged under a linear, and thus affine change of coordinates we may assume, without loss of generality, that none of the intersection points of f and g lie on the line at infinity. By lemma 5.2 it is enough to show that $\sum_{p\in\mathbf{P}^2} I_p(f,g) = \dim_{\mathbf{K}} \Gamma_d$ for some $d \geq \deg(f) + \deg(g)$.

Let $\Gamma_* := \mathbf{K}[x,y]/(f_*, g_*)$. Since the intersection points of f and g all lie in $\mathbf{K}^2$ we get

$$\sum_{p\in\mathbf{P}^2} I_p(f,g) = \sum_{p\in\mathbf{K}^2} I_p(f_*,g_*) = \dim_{\mathbf{K}}(\Gamma_*),$$

where the last equality follows from corollary 4.7. We will now show that $\dim_{\mathbf{K}}(\Gamma_*) = \dim_{\mathbf{K}}(\Gamma_d)$ for some $d \geq \deg(f) + \deg(g)$, from which then the theorem follows.

Consider the map $\alpha : \Gamma \longrightarrow \Gamma$ defined by $\alpha(\overline{h}) = \overline{zh}$. Suppose that $\overline{h} \in \ker(\alpha)$. Then $zh = af + bg$ for some $a, b \in \mathbf{K}[x,y,z]$. For any $u \in \mathbf{K}[x,y,z]$ let u_0 denote $u(x,y,0)$. Then we get $a_0 f_0 + b_0 g_0 = 0$.

Suppose $r(x,y) \in \mathbf{K}[x,y]$ is a common factor of f_0 and g_0. Then $f_0 = ra$ and $g_0 = rb$ for some $a, b \in \mathbf{K}[x,y]$. Since r has infinitely many zeros we can choose some $(x_0, y_0) \neq (0,0)$ which is a zero of r. Then $f(x_0, y_0, 0) = f_0(x_0, y_0) = 0$ and $g(x_0, y_0, 0) = g_0(x_0, y_0) = 0$ which contradicts the assumption that f and g have no intersection points at infinity. Therefore, f_0 and g_0 have no common factors.

Since $a_0 f_0 = -b_0 g_0$ we must have $a_0 = -cg_0$ and $b_0 = cf_0$ for some $c \in \mathbf{K}[x,y]$. Set $a_1 = a + cg$ and $b_1 = b - cf$. Then $(a_1)_0 = 0 = (b_1)_0$ which implies that $a_1 = za'$ and $b_1 = zb'$ for some $a', b' \in \mathbf{K}[x,y,z]$. Since $a_1 f + b_1 g = af + bf$ we get $zh = za'f + zb'g$ and thus $h = a'f + b'g$, so $\overline{h} = 0$ and α is injective. Let $\deg(f) = m$ and $\deg(g) = n$. Then in particular, if $d \geq m + n$ then α restricts to an isomorphism between Γ_d and Γ_{d+1} since an injective linear map between two vector spaces of the same dimension is an isomorphism.

Now let $d \geq m + n$ and let $a_1, ..., a_{mn} \in \mathbf{K}[x,y,z]$ be forms of degree d whose residues $\overline{a_1}, ..., \overline{a_{mn}}$ form a basis for Γ_d, by lemma 5.2. Let $\overline{a_{1*}}, ..., \overline{a_{mn*}}$ be the residues of $a_{1*}, ..., a_{mn*}$ in Γ_*.

Since α is an isomorphism between Γ_d and Γ_{d+1} for all $d \geq m + n$, the residues $\overline{z^r a_1}, ..., \overline{z^r a_{mn}}$ form a basis of Γ_{d+r} for all $r \geq 0$. Now let $\overline{h} \in \Gamma_*$ be the residue of $h \in \mathbf{K}[x,y]$. Then for any $r \geq 0$ such that $r \geq \deg(h) - d$ there is some N

such that $z^N h^*$ is a form of degree $d+r$. Then $\overline{z^N h^*}$ in Γ is a linear combination of $\overline{z^r a_1}, ..., \overline{z^r a_{mn}}$ in Γ_{d+r} and therefore

$$z^N h^* = \sum_{i=1}^{mn} \lambda_i z^r a_i + bf + cg$$

for some $\lambda_1, ..., \lambda_{mn} \in \mathbf{K}$ and $b, c \in \mathbf{K}[x, y, z]$. But then

$$h = (z^N h^*)_* = \sum_{i=1}^{mn} \lambda_i a_{i*} + b_* f_* + c_* g_*$$

and therefore $\overline{h} = \sum_{i=1}^{mn} \lambda_i \overline{a_{i*}}$. So Γ_* is generated by $\overline{a_{1*}}, ..., \overline{a_{mn*}}$.

Suppose $\sum_{i=1}^{mn} \lambda_i \overline{a_{i*}} = 0$. Then $\sum_{i=1}^{mn} \lambda_i a_{i*} = bf_* + cg_*$ for some $b, c \in \mathbf{K}[x, y]$. If we let $s = \deg(bf_*)$, $r = \deg(cg_*)$ and $t = \min(s, r)$ then $z^t(bf_* + cg_*)^* = z^r(bf_*)^* + z^s(cg_*)^* = z^r b^*(f_*)^* + z^s c^*(g_*)^*$. Let z^{u_f} and z^{u_g} be the highest powers of z dividing f and g respectively. Then

$$z^{t+u_f+u_g}(bf_*+cg_*)^* = z^{r+u_g} b^* \cdot z^{u_f}(f_*)^* + z^{s+u_f} c^* \cdot z^{u_g}(g_*)^* = z^{r+u_g} b^* f + z^{r+u_f} c^* g.$$

Let $d' = \max(\deg(a_{1*}), ..., \deg(a_{mn*}))$. Then, since $a_1, ..., a_{mn}$ are all forms of degree d we get

$$z^{d-d'} \left(\sum_{i=1}^{mn} \lambda_i a_{i*} \right)^* = \sum_{i=1}^{mn} \lambda_i a_i.$$

So if we let $t' = \max(t + u_f + u_g, d - d')$ then

$$z^{t'-(d-d')} \sum_{i=1}^{mn} \lambda_i a_i = z^{s'} b^* f + z^{r'} c^* g$$

for some s' and r'. This then means that $\sum_{i=1}^{mn} \lambda_i \overline{z^{t'-(d-d')} a_i} = 0$ in $\Gamma_{d+t'-(d-d')} = \Gamma_{t'+d'}$ and since $\overline{z^{t'-(d-d')} a_1}, ..., \overline{z^{t'-(d-d')} a_{mn}}$ form a basis of $\Gamma_{t'+d'}$, this implies that $\lambda_1 = ... = \lambda_{mn} = 0$. So $\overline{a_{1*}}, ..., \overline{a_{mn*}}$ are linearly independent. Therefore $\overline{a_{1*}}, ..., \overline{a_{mn*}}$ is a basis of Γ_* and hence $\dim_{\mathbf{K}}(\Gamma_*) = \dim_{\mathbf{K}}(\Gamma_d) = mn$. $\square$

References

[1] D.S. Dummit and R.M. Foote. *Abstract Algebra*. Wiley, 2004.

[2] W. Fulton. *Algebraic curves: an introduction to algebraic geometry*. Advanced book classics. Addison-Wesley Pub. Co., Advanced Book Program, 1989.