WHY DO WE NORMALLY DO ALGEBRAIC GEOMETRY OVER AN ALGEBRAICALLY CLOSED FIELD?

Reference: Section 1.1 "Plane curves" in "Basic Algebraic Geometry 1", Shafarevich.

INTRODUCTION

The first goal of this first lecture is to show why normally an algebraic geometer prefers to work over an algebraically closed field rather than an arbitrary field.

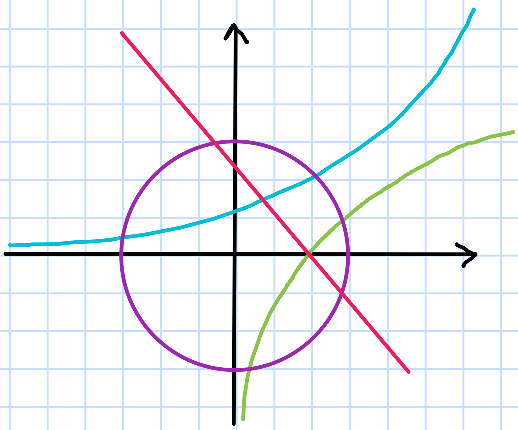
The second goal is to start showing how much the geometry of an algebraic curve is related to the algebra of its "corresponding polynomial".

In order to keep it simple and have a better geometrical intuition, we will work in this lecture in the affine plane.

What is a curve?

The word "curve" automatically suggests us a geometrical object of dimension one.

In the real plane we imagine something like this:



For an algebraic geometer not all the above geometrical objects are curves, i.e. not all curves are algebraic.

Indeed in algebraic geometry we only consider those curves that arise from polynomials, like lines, cicles, parabolas, etc...

Since the algebra of polynomial rings plays a fundamental role in the study of algebraic curves we will start by reviewing some algebraic facts in commutative algebra.

REVIEW: COMMUTATIVE RINGS

FIELDS → examples: $\mathbb{Q}$, $\mathbb{C}$, $\mathbb{Q}_p$, $\mathbb{F}_q$, etc...

$\cap$

ED → "Euclidean domains": integral domain + euclidean function (generalization of the euclidean division in $\mathbb{Z}$)
examples: $\mathbb{Z}$, $k[x]$, etc.

$\cap$

PID → "Principal ideal domain": every ideal is principal.

$\cap$

UFD → "Unique factorization domain": every non-zero non-unit element can be written as a product of irreducible (= prime elements), uniquely up to order and units.
example: $k[x,y]$, which is not a PID since the ideal (x,y) is not principal.

$\cap$

BÉZOUT DOMAINS → For every two elements a Bézout identity holds $\Leftrightarrow$ every finitely generated ideal is principal

$\cap$

GCD DOMAINS → Every two elements have a greatest common divisor

$\cap$

INTEGRAL DOMAINS → If $ab=0 \Rightarrow a=0$ or $b=0$

$\cap$

COMMUTATIVE RINGS → rings commutative with respect to multiplication.

Some additional algebraic preliminaries can be found in Sec. 1.1. of "Algebraic curves: an introduction to algebraic geometry", W. Fulton.

ALGEBRAIC PLANE CURVES

Let us fix, at the moment, an arbitrary field K.

Recall: A field $(F,+,\cdot)$ is a set with two binary operations $+$ and $\cdot$, called generally addition and multiplication

$$+ : F \times F \to F$$

$$\cdot : F \times F \to F$$

such that:

- $+$ is commutative: $\forall\, a,b \in F$, $a+b=b+a$.
- $+$ is associative: $\forall\, a,b,c \in F$, $(a+b)+c = a+(b+c)$.
- $+$ has an identity element 0 such that $a+0=0+a=a$
- $\forall\, a \in F$, $+$ has an inverse $-a$ such that $a+(-a)=(-a)+a=0$

- $\cdot$ is commutative: $\forall\, a,b \in F$, $a\cdot b = b\cdot a$
- $\cdot$ is associative: $\forall\, a,b,c \in F$, $(a\cdot b)\cdot c = a\cdot(b\cdot c)$
- $\cdot$ has an identity element 1 such that $a\cdot 1 = 1\cdot a = a$
- $\cdot$ $\forall\, a \in F$, $a \neq 0$, $\cdot$ has an inverse a^{-1} such that $a\cdot a^{-1} = a^{-1}\cdot a = 1$

- $\forall\, a,b,c \in F$, $a\cdot(b+c) = ab+ac$ (distributivity)

Remark: If F is a field, then $(F,+)$ and $(F\setminus\{0\},\cdot)$ are groups, called the additive and multiplicative groups.

NOTATION

- $K[x,y]$: the ring of polynomials in two variables with coefficients in K.

  example: $f(x,y) = x^2y^2 + \sqrt{2}x - \frac{y}{3}$.

  $f(x,y) \in \mathbb{R}[x,y]$ but $f(x,y) \notin \mathbb{Q}[x,y]$.

- $\mathbb{A}^2(K) = \{(x,y) : x,y \in K\}$: the affine plane with coordinates in K.

An element of $\mathbb{A}^2(K)$ is called a point.

We would like to give the following definition for an (affine) algebraic plane curve:

Def: An algebraic plane curve is a curve $\mathcal{C} \subseteq \mathbb{A}^2(k)$ consisting of the points $(x,y) \in \mathbb{A}^2(k)$ such that

$$f(x,y) = 0$$

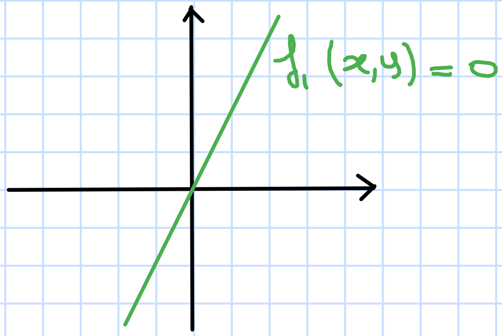
where $f(x,y) \in k[x,y]$ is a nonconstant polynomial.

The degree of the polynomial f is also called the degree of the curve. A curve of degree 1 is a line, a curve of degree 2 is called a conic and a curve of degree 3 a cubic.

Remark: The adjective "plane" follows from the fact that the curve is a subset of the affine plane $\mathbb{A}^2(k)$, while the adjective "algebraic" from the fact that it is defined by a polynomial equation.

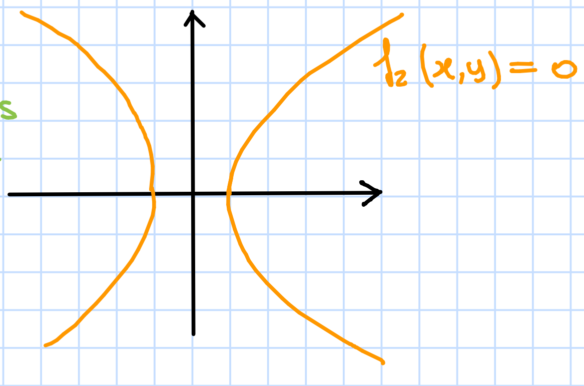
e.g.: $K = \mathbb{R}$

1) If $f_1(x,y) = y - 2x \Rightarrow f_1(x,y) = 0$ is the line which passes through $(0,0)$ and $(1,2)$:

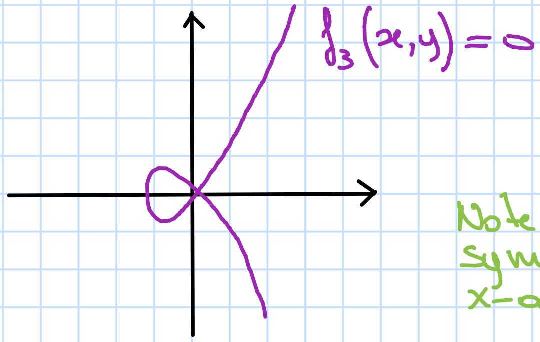


2) If $f_2(x,y) = x^2 - y^2 - 1 \Rightarrow f_2(x,y) = 0$ defines a conic (hyperbola):

Note that the curve is symmetric about the x- and the y-axis.

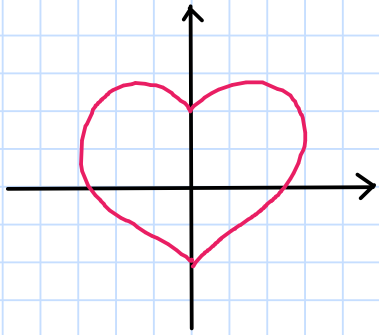


3) If $f_3(x,y) = y^2 - x^2 - x^3 \Rightarrow f_3(x,y) = 0$ defines a cubic:



Note that the curve is symmetric about the x-axis

4) If $f_4(x,y) = (x^2 + y^2 - 1)^3 - x^2 y^3$, then



Note that the curve is symmetric about the y-axis

Recall that $K[x,y]$ is a unique factorization domain (UFD).

This means that every polynomial $f \in K[x,y]$ has a unique factorization (up to constant multiples and the ordering of the factors

$$(*) \quad f = f_1^{e_1} \cdots f_r^{e_r},$$

where f_i are irreducible polynomials and $f_i \neq \alpha f_j \ \forall \alpha \in K$ if $i \neq j$.

Recall: In an integral domain R, an element $a \in R$ is said to be IRREDUCIBLE if it is not a unit or zero and

$$a = b \cdot c, \quad b, c \in R \Rightarrow b \text{ or } c \text{ is a unit (= invertible element)}$$

example: $f(x,y) = x^2 + y^2 \in \mathbb{Q}[x,y]$ is irreducible in $\mathbb{Q}[x,y]$

while $g(x,y) = x^2 - y^2 = (x+y)(x-y) \in \mathbb{Q}[x,y]$ is not irreducible

If now $f(x,y)$ has a decomposition like $(*)$ into irreducible factors, then:

$$f(\alpha,\beta)=0 \iff \exists\, i\in\{1,\dots,r\} \text{ such that } f_i(\alpha,\beta)=0,$$

where $(\alpha,\beta)\in \mathbb{A}^2(K)$. In other words, (α,β) is a point of the curve $\mathcal{C}: f(x,y)=0$ if and only if (α,β) is a point of one of the curves $\mathcal{C}_i: f_i(x,y)=0$. We can write:

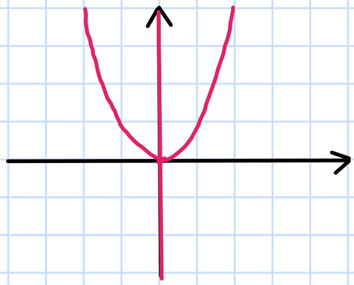
$$\mathcal{C} = \mathcal{C}_1 \cup \dots \cup \mathcal{C}_r.$$

A curve is irreducible if it is defined by an irreducible polynomial.

The decomposition $\mathcal{C}=\mathcal{C}_1\cup\dots\cup\mathcal{C}_r$ is called a decomposition of X into irreducible components.

example: $K=\mathbb{R}$. $f(x,y)=xy-x^3=x(y-x^2)$.

The curve $\mathcal{C}: f(x,y)=0$ is reducible, since it is the union of $\mathcal{C}_1: x=0$, $\mathcal{C}_2: y=x^2$.



In certain cases the notions just introduced turn out not to be well defined, or to differ widly from our intuition...

e.g.: If $K=\mathbb{R}$:

- we should call the point $(0,0)$ a "curve" since it is defined by the equation $x^2+y^2=0$.
- Moreover this "curve" should have "degree" 2, but also any other even number, since the same point $(0,0)$ is also defined by the equation $x^{2n}+y^{2n}=0$.
- the curve is irreducible if we take its equation to be $x^2+y^2=0$, but reducible if we take it to be $x^6+y^6=0$.

This is why in algebraic geometry one prefers to work with algebraically closed fields.

Recall: A field F is algebraically closed if it contains a root for every non constant polynomial in $F[x]$.

e.g.: $\mathbb{R}$ is not algebraically closed since the polynomial x^2+1 has no roots in $\mathbb{R}$.

Examples of algebraically closed field are $\mathbb{C}$, $\overline{\mathbb{Q}}$ (the set of algebraic numbers), $\overline{\mathbb{F}_q}$, ...

note $\overline{\mathbb{Q}} \subsetneq \mathbb{C}$ since any transcendental number does not belong to $\overline{\mathbb{Q}}$. (e.g.: $e, \pi \notin \overline{\mathbb{Q}}$)

this is a consequence of the fundamental theorem of algebra

<u>Remark</u>: Every field is contained in an algebraically closed field and the smallest (with respect to inclusions) algebraically closed field that contains a field K is called the algebraic closure of K and denoted $\overline{K}$.

$$K = \overline{K} \iff K \text{ is algebraically closed.}$$

<u>Proposition</u>: An algebraically closed field K is infinite.

<u>Proof</u>: Indeed, if it was finite, i.e. $K = \{a_1, \dots, a_n\}$ then the polynomial:

$$f(x) = (x-a_1)(x-a_2)\cdots(x-a_n)+1 \quad \in K[x]$$

would have no roots in K, since $f(a_i) = 1 \ \forall a_i \in K$.

<u>Corollary</u>: Assume that K is algebraically closed and let $f(x,y) \in K[x,y]$. Then the curve $\mathcal{C}: f(x,y)=0$ has infinitely many points.

<u>Proof</u>: $\forall \beta \in K$ the polynomial $f(x,\beta)=0$ has at least a root in K (since K is algebraically closed).

Since K is infinite, the conclusion follows.

We want to prove that if K is algebraically closed, for each irreducible curve we can define an unique irreducible polynomial (up to a constant multiple).

We will use the following lemma, which holds for an arbitrary field.

Lemma: Let K be an arbitrary field, $f \in K[x,y]$ an irreducible polynomial and $g \in K[x,y]$ an arbitrary polynomial.
If f does not divide g then the system of equations:

$$(**) \quad \begin{cases} f(x,y)=0 \\ g(x,y)=0 \end{cases}$$

has only a finite number of solutions, i.e. the curves described by f and g have finitely many intersections.

Proof: Suppose the degree of f in x is positive.

We can see f and g as elements of $K(y)[x]$, that is as polynomials in one variable x, whose coefficients are rational function of y. Unlike $K[x,y]$, the ring $K(y)[x]$ is an Euclidean domain (thus a Bézout domain).

We need the following fact:

FACT: Let R be a GCD domain and F its field of fractions. If a nonconstant polynomial is irreducible in $R[X]$ then it is irreducible in $F[X]$.

generalization of Gauss's lemma

By the previous fact, since f is irreducible in $K[x,y] = K[y][x]$, then f is irreducible in $K(y)[x]$

$K(y) = \mathrm{Frac}(K[y])$

Note also that f does not divide g in $K(y)[x]$.

Indeed if $f \mid g$ then:

$$g(x,y) = f(x,y) \cdot h, \quad \text{with } h \in K(y)[x]$$

$$\Downarrow \quad a(y) \in K[y] \text{ common denominator of } h.$$

$$a(y)g(x,y) = f(x,y) \cdot \underbrace{h \cdot a(y)}_{\in K[x,y]}$$

Then, $f(x,y) \mid a(y)g(x,y)$ $\xRightarrow{(f \text{ irr.} \Rightarrow f \text{ prime}) + f \nmid a(y)}$ $f(x,y) \mid g(x,y)$ ↯

This implies $\gcd(f,g)=1$. Hence there exist two polynomials $\tilde{u}, \tilde{v} \in K(y)[x]$ such that:

$$f(x,y)\,\tilde{u} + g(x,y)\,\tilde{v} = 1 \qquad \text{(Bézout identity)}$$

$$\Big\Downarrow \quad b(y) \in K[y] \text{ common denominator of } f\tilde{u}+g\tilde{v}$$

$$f(x,y)\,\underbrace{b(y)\tilde{u}}_{\in K[x,y]} + g(x,y)\,\underbrace{b(y)\,\tilde{v}}_{\in K[x,y]} = b(y)$$

Now, if $(\alpha,\beta) \in K\times K$ is a common solution of the system $(**)$, i.e. $f(\alpha,\beta) = g(\alpha,\beta)=0$, then from the last equation we obtain that

$$b(\beta)=0,$$

i.e. β is a root of the polynomial $b(y)$.

Thus we have finitely many possible values for the second coordinate β.

For each such value, the first coordinate α is a root of $f(x,\beta)=0$. The polynomial $f(x,\beta)$ is not identically 0, since otherwise $f(x,y)$ would be divisible by $y-\beta$ (while f is irreducible) and hence there is only a finite number of possibilities for the first coordinate α. The lemma is proved.

$\square$

Now assume that K is algebraically closed. Let $f(x,y) \in K[x,y]$ be an irreducible polynomial and let us consider the curve $f(x,y)=0$ (which has infinitely many points since K is algebraically closed).

If there was another polynomial $g(x,y) \in K[x,y]$ describing the same curve, then the system

$$\begin{cases} f(x,y)=0 \\ g(x,y)=0 \end{cases}$$

would have infinitely many solutions, and by the previous lemma, f divides g.

Note: this is not true if K is not algebraically closed. Indeed

$$f(x,y) = x^2+y^2 \in \mathbb{R}[x,y] \quad \text{and} \quad g(x,y) = x^4+y^4 \in \mathbb{R}[x,y]$$

describe the same curve $\mathcal{C} = \{(0,0)\}$, but $f \nmid g$ and $g \nmid f$.

If now also g is irreducible, then $g = \alpha f$, where $\alpha \in K$.

This shows that an irreducible polynomial $f(x,y)$ is uniquely determined, up to a constant multiple, by the curve $f(x,y)=0$.

Analogously, it is easy to prove that a non irreducible curve is uniquely defined by a polynomial whose factorization into irreducible components has no multiple factors (up to a constant multiple).

The notion of the degree of a curve and of irreducible curve is then well defined when the field is algebraically closed.

Another reason why one would prefer algebraically closed fields comes when one considers the number of points of intersection of curves.

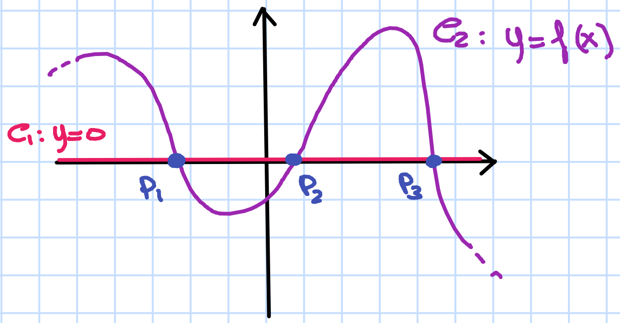
We are already familiar with the following example.

Example: Let us consider the following two curves:

$$\mathcal{C}_1 : y = 0 \quad , \quad \mathcal{C}_2 : y - f(x) = 0, \quad f(x) \in K[x].$$

How many intersections do $\mathcal{C}_1$ and $\mathcal{C}_2$ have?

$$\# \, \mathcal{C}_1 \cap \mathcal{C}_2 = \, ?$$



This is equivalent to count the number of solution of the system:

$$\begin{cases} y = 0 \\ y = f(x) \end{cases}$$

or the number of roots of the polynomial $f(x)$.

From algebra, we know that this number is bounded by the degree of f and the bound is always attained when K is algebraically closed:

If $K = \overline{K}$ then

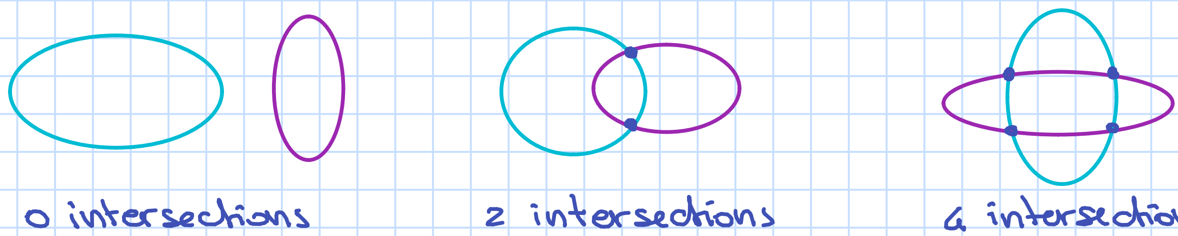
$$\# \, \mathcal{C}_1 \cap \mathcal{C}_2 = \deg(f) = \deg(\mathcal{C}_1) \cdot \deg(\mathcal{C}_2)$$

A generalization of this theorem is the Bézout's theorem, which, without any assumption, sounds like this:

"The number of points of intersection (with multiplicity) of two distinct irreducible curves equals the product of their degree".

One of the necessary assumptions for Bézout's theorem is that K is algebraically closed.

Indeed, by Bézout's theorem two ellipses in the plane should have exacty 4 intersections, but this does not always hold in the real plane:



We will see that the fact of K being algebraically closed is not the only assumption that Bézout's theorem requires to be true. We will come back to this later...