1.1 06/10/24 - Introduction

Example 1.1.1 (Student Examples). Get Desmos to plot the subsets of the plane (over $k = \mathbb{R}$) defined by the vanishing of the following polynomials

(a) $3x + 4y - 7$ (line)
(b) $x^2 + y^2 - 1$ (circle),
(c) $y - x^2$ (parabola),
(d) $y^2 + x^3$ (semicubical parabola/cuspidal cubic),
(e) $y^2 - x^3 - x$ (one-component elliptic curve),
(f) $y^2 - x^3 + x$ (two-component elliptic curve),
(g) $(x^2 + y^2)(x + y - 1)$ (line and point not on it),
(h) $xy - 1$ (hyperbola), and
(i) $x^2 + y^2 + 1$ (empty set).

These are all examples of algebraic curves. Now get Desmos to plot

(a) $y - \sin(1/x)$, and
(b) $y - |x|$.

These are not plane algebraic curves (why?). See also Exercise 2.1.8.

We will fix a field k throughout (see Remark 1.1.17).

Definition 1.1.2. The affine plane over k, denoted $\mathbb{A}^2_k$, is the set of ordered pairs of elements of k, so that

$$\mathbb{A}^2_k := \{(p, q) : p, q \in k\}.$$

If you want, see Remark 1.1.18 for an explanation of why we use $\mathbb{A}^2_k$ to denote the set others sometimes denote by k^2.

Given a function $F : \mathbb{A}^2_k \to k$, we can look at its vanishing locus, denoted variously by

$$F^{-1}(0) = C_F = \mathbb{V}(F) = \mathrm{Z}(F) = \{(p, q) : F(p, q) = 0\}.$$

We will usually stick to the notation C_F.

Remark 1.1.3. More generally, we can look at the level sets $F^{-1}(a)$ for all $a \in k$. Why does this perspective not add anything new?

Any polynomial $f(x, y) \in k[x, y]$ gives rise to a function $F_f : \mathbb{A}^2_k \to k$ by evaluation.

Remark 1.1.4. Why is it important to keep the notions of a polynomial and polynomial function separate? See Exercise 2.2.6.

Definition 1.1.5. An affine plane algebraic curve is the vanishing locus of a polynomial function in the affine plane given by a nonconstant polynomial, i.e. a subset $C \subset \mathbb{A}^2_k$ of the form $C = C_{F_f}$ for some nonconstant polynomial $f(x, y) \in k[x, y]$.

For simplicity, we'll use the notation $C_f := C_{F_f}$. We will sometimes write $C_f(k)$ to denote C_f if we want to emphasize the underlying field. Finally, we will often abbreviate "affine plane algebraic curves" to simply "curves," since we will not have occasion to deal with other kinds of curves, at least initially.

Remark 1.1.6. Our definition is currently a little weird. For instance, with our current definition, for certain fields k, a curve can be

- empty (think $x^2 + y^2 + 1 = 0$ over $\mathbb{R}$),
- a finite collection of points (think $x^2 + y^2 = 0$ over $\mathbb{R}$ and Proposition 1.1.7, or think of what happens when $k = \mathbb{F}_q$ is a finite field),
- and all of $\mathbb{A}^2_k$ (again think of $k = \mathbb{F}_q$ being a finite field).

Neither of these sets seem to be "1-dimensional," which is the elusive notion we are trying to capture. We could either choose to restrict ourselves to working over infinite fields or algebraically closed fields (even in positive characteristic–see Exercise 2.2.8), but this misses a lot of important number theory (see Examples 1.1.11 and 1.1.15). Alternatively, we can accept that our definition is broader than initially intended, and try to study its consequences.

Proposition 1.1.7. Let k be a field. If $C, D \subset \mathbb{A}^2_k$ are curves, then so is $C \cup D$.

Proof. If $C = C_f$ and $D = C_g$ for $f, g \in k[x, y]$, then $C \cup D = C_{fg}$. ■

Remark 1.1.8. Here we are using that $k[x, y]$ is a ring (how?), and that k is a field (or at least that it is a domain–what happens if k is not even a domain?). We will say more about this when we talk about irreducibility and reducedness of curves.

1.1.1 Motivating Questions

Given a field k and a curve $C \subset \mathbb{A}^2_k$, we can ask several questions about it.

Question 1.1.9. Is $C = \emptyset$?

This is not at all as trivial as it seems. Many number-theoretic questions can be phrased in this language, if we take k to be $\mathbb{Q}$ or a finite field $\mathbb{F}_q$, for instance.

Example 1.1.10. Take $k = \mathbb{Q}$, fix a prime p, and look at the curve C defined by

$$f(x, y) := x^2 + y^2 - p \in \mathbb{Q}[x, y].$$

Then $C = \emptyset$ iff p satisfies a certain congruence condition (which?). See Exercise 2.1.1.

Example 1.1.11. Take $k = \mathbb{F}_p$ to be a finite field of prime order and $a \in k$ to be any element, and look at the curve C defined by

$$f(x, y) = x^2 - a \in \mathbb{F}_p[x, y].$$

Then $C = \emptyset$ iff a is quadratic nonresidue modulo p, i.e. $\left(\frac{a}{p}\right) = -1$.

Remark 1.1.12. For any field k, if $f(x, y) \in k[x, y]$ is a polynomial of x only, then the curve C_f defined by f is a finite (possibly empty) union of "vertical lines". Can you make this precise?

Example 1.1.13. Take $k = \mathbb{Q}$ and $n \geq 1$ to be a positive integer. Let

$$f_n(x, y) := x^n + y^n - 1 \in \mathbb{Q}[x, y],$$

and $C_n := C_{f_n}$ be the curve defined by f_n. Then Fermat's Last Theorem says that

$$C_n(\mathbb{Q}) = \emptyset \Leftrightarrow n > 2.$$

Question 1.1.14. If C is nonempty, what can we say about the locus C? Is it finite or infinite? What can we say about its topology[a]?

[a] What's that?

Example 1.1.15. For instance, if k is finite, what is the cardinality of $C(k)$? Suppose $k = \mathbb{F}_q$ is a finite field, and that C is an elliptic curve[1], e.g. the curve defined by

$$f(x, y) = y^2 - x^3 - x \in \mathbb{F}_q[x, y]$$

when q is not a power of 2. The Hasse Theorem says that, in the above case,

$$(\sqrt{q} - 1)^2 \leq \#C(\mathbb{F}_q) \leq (\sqrt{q} + 1)^2.$$

In particular, we have $\#C(F_q) \sim q$ for all large q. (What does that even mean? Aren't we starting with a fixed q to begin with?) We will not prove this theorem in this course.

Example 1.1.16. If $k = \mathbb{R}$ or $k = \mathbb{C}$, how many pieces (i.e. connected components) does $C(k)$ have? How are they related to each other? See Exercise 2.1.2 for the case when $k = \mathbb{R}$. Another theorem, which will not prove in this course, asserts that if $k = \mathbb{C}$, then any irreducible curve[2] is connected.

1.1.2 Some Unimportant Remarks

Remark 1.1.17. Why did we require k to be a field? What would happen if k were just a ring–does the notion of an affine plane curve over a ring make sense? [Hint: some things make sense, whereas other things like Proposition 1.1.7 break down. See Remark 1.1.8.] Can you see how far you can go till things break down and what you can salvage by adapting definitions?

Remark 1.1.18. As sets, $\mathbb{A}^2_k$ and $k^2 = k \times k$ are identical[3], but $\mathbb{A}^2_k$ does not come equipped with additional structure that k^2 is often (implicitly) interpreted to have: k^2 is often seen (by students who have seen some linear algebra) as a vector space with an additive structure and a distinguished origin, but for us $\mathbb{A}^2_k$ is just a set[4] and, as will become clear when we discuss affine changes of coordinates, there is no distinguished point in $\mathbb{A}^2_k$–all points "look the same". In slightly more grown-up terminology, the affine plane over k is a principal homogenous space or torsor for the (underlying additive group) of the vector space k^2. If you do not understand what this remark means, you can safely ignore it.

Remark 1.1.19. Regarding the different choices of the field k: it's often easiest to plot curves over $k = \mathbb{R}$, but plots can also be made over other fields such as $k = \mathbb{C}$ (using some ingenuity and imagination–how?) or $k = \mathbb{F}_q$ (this may be a silly, uninstructive plot, but not always!). We will see throughout the course that it is, in fact, easier to work with curves over $k = \mathbb{C}$ than over $k = \mathbb{R}$ (why do you think this might be?). However, curves over other fields are equally important:

(a) Fields such as $k = \mathbb{Q}, \mathbb{F}_p$ (or finite extensions and completions of these–such as $k = \mathbb{Q}_p$) show up a lot in solving number-theoretic questions. See Examples 1.1.10, 1.1.11 and 1.1.13.

[1] We will define this notion formally later.

[2] Now, what's that?

[3] Only according to our definition! There are other accepted definitions of $\mathbb{A}^2_k$, such as $\mathbb{A}^2_k = \operatorname{Spec} k[x, y]$, for which this is no longer the case. You don't have to wrorry too much about this right now.

[4] Later on in your studies, it can, and will, be given the structure of a topological space, and in fact a locally ringed space (even affine scheme).

(b) Another case of interest is when $k = K(t)$ for some other field k. When $K = \mathbb{F}_q$ is a finite field, working with curves over $k = \mathbb{F}_q(t)$ is known as a the "function field analog" of the theory of curves. Many important questions which are unsolved in the "usual case" have been solved in the function field case (such as the Riemann Hypothesis), and this provides (one strand of) evidence for the Riemann Hypothesis.

(c) In (b), when we take $K = \mathbb{C}$, so that we are looking at curves over $k = \mathbb{C}(t)$, we are *really* looking at one-parameter families of curves that fit together into an algebraic surface. For instance, elliptic curves over $\mathbb{C}(t)$ often give rise to elliptic K3 surfaces. This perspective is very helpful in the study of higher-dimensional algebraic varieties as well.

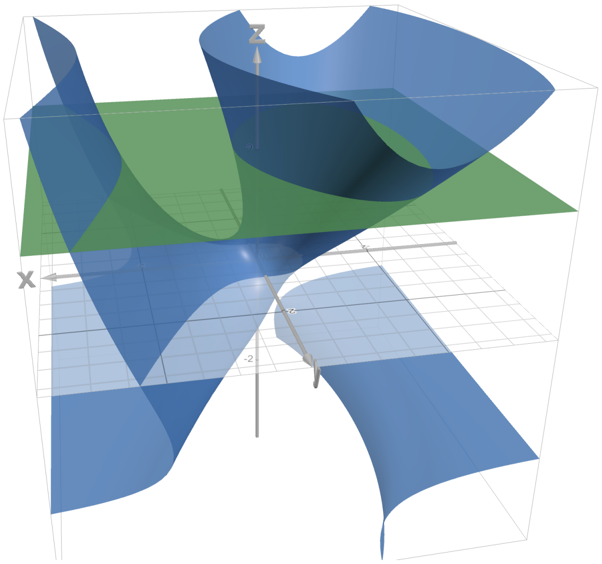

Figure 1.1: The elliptic curve over $k = \mathbb{C}(z)$ defined by $y^2 = x^3 - 3zx + (z^3+1)(z+2)^{-1}$ in blue, along with its hyperplane section at $z = 2$, which is the elliptic curve $y^2 = x^3 - 6x + 9/4$. Picture made with Desmos 3D.

Therefore, it is helpful to have the flexibility to work over arbitrary fields from the beginning.