Lecture 5: Trichotomy of curves

Brian Lehmann
Boston College

Definition

Trichotomy: a division into three categories.

Today we discuss the famous "trichotomy" of curves: all algebraic curves can be split into three different types.

This decomposition is the same no matter what field we are working in! Complex geometry, number theory, and algebraic geometry all give the same trichotomic classification. We will focus on the case of plane curves $P(x, y) = 0$.

Notation

Suppose $P(x, y)$ is an irreducible polynomial with rational coefficients. The solutions to $P(x, y) = 0$ is called a "curve"; we will denote it by C. (We will might think of C as an object in $\mathbb{R}^2$ or an object in $\mathbb{C}^2$.)

The set of rational solutions is denoted by $C(\mathbb{Q})$.

Definition

We say that C is smooth if the partial derivatives $\frac{\partial P}{\partial x}$ and $\frac{\partial P}{\partial y}$ do not simultaneously vanish for any point in C.

To be clear: we are asking whether there is any point (x, y) for which the three polynomials P, $\frac{\partial P}{\partial x}$, and $\frac{\partial P}{\partial y}$ all vanish simultaneously.

Smoothness should be thought of as a property in $\mathbb{C}^2$: we need to make sure there are no **complex** pairs (x, y) where these polynomials vanish.

Trichotomy of curves

Example

The curve $y - x^2 = 0$ is smooth. The partial derivatives are $\frac{\partial P}{\partial x} = -2x$ and $\frac{\partial P}{\partial y} = 1$. Note that the second polynomial will never vanish (no matter what inputs we choose).

Example

The curve $x^2 + y^2 = 0$ is not smooth. The partial derivatives are $\frac{\partial P}{\partial x} = 2x$ and $\frac{\partial P}{\partial y} = 2y$. These will both vanish at the point $(x, y) = (0, 0)$. Since $(0, 0)$ also lies on C, it is a non-smooth point of C.

Question

Let $P(x, y)$ be a degree d polynomial with rational coefficients that defines a smooth curve C. What is $\chi(C)$?

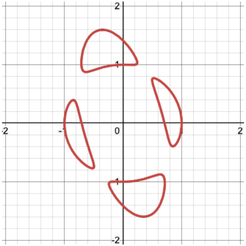
The "right" version of this question requires two improvements:

1. We take a "compactification" of C in projective space.
2. We think of C as an object over the complex numbers, not the reals.

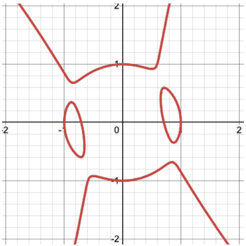
Let's discuss these one at a time.

Compactification

Suppose that $P(x, y) = 0$ defines a curve C in $\mathbb{R}^2$. The curve C will contain several pieces which we will call “ovals”.



Four ovals



Four ovals

Some ovals are bounded, while some escape to infinity.

Compactification

The bounded ovals in C are examples of compact sets.

However, the unbounded ovals are not compact. To fix this issue, we need to add some "points at infinity" to turn them into compact sets. Loosely, we can think of these points at infinity as directions along the horizon.

Compactification

Definition

$\mathbb{P}^2_{\mathbb{R}}$ is equal to $\mathbb{R}^2 \cup (\mathbb{R} \cup \{\infty\})$. We call the new points $(\mathbb{R} \cup \{\infty\})$ the "points at infinity" in projective space.

The point at infinity labeled t corresponds to the points in the horizon "at slope t". More precisely, suppose we start from some point and start walking along the line with slope t (in either direction). If we walk "infinitely far out" along the line, we will eventually reach the horizon at the point labeled t.

Introduction
Geometry of plane curves
Counting ovals
Trichotomy of curves

Compactification rule: suppose we have an unbounded oval.

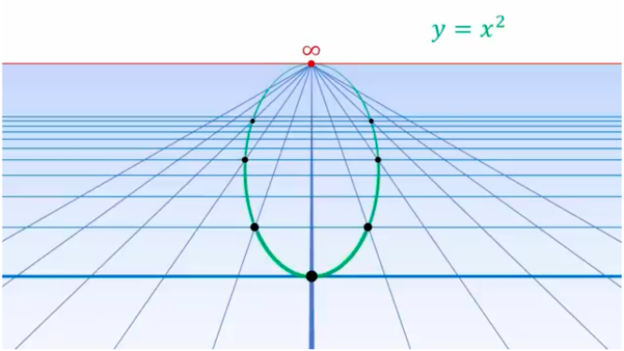
1. It is possible that as we move out to infinity the slopes of the tangent lines to the oval will converge to a value t. In that case we add the point marked t at infinity to our oval.
2. It is possible that as we move out to infinity the slopes of the tangent lines to the oval will converge to two different values. In this case we add the two corresponding points at infinity to our oval.

Sometimes two unbounded ovals have the same points at infinity. In this case we "glue the two ovals" to get a single oval that "stretches out over ∞."

Introduction
Geometry of plane curves
Counting ovals
Trichotomy of curves

Example

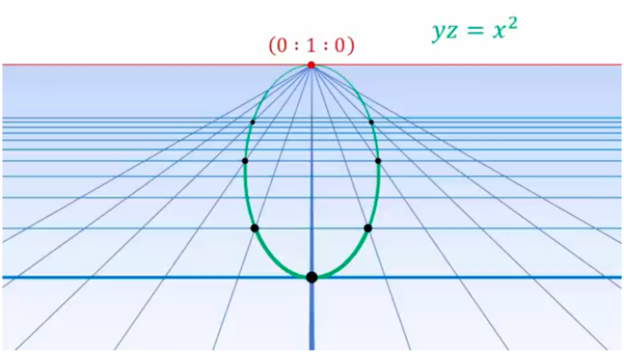
The parabola limits to a single point at infinity. When we add this point, the parabola "looks like" a typical ellipse.



Compactifying the parabola

Example

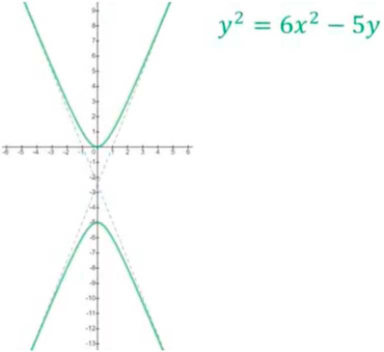
The parabola limits to a single point at infinity. When we add this point, the parabola "looks like" a typical ellipse.



Compactifying the parabola

Example

In addition to ellipses and parabolas, the third type of conic section in the real plane is a hyperbola.

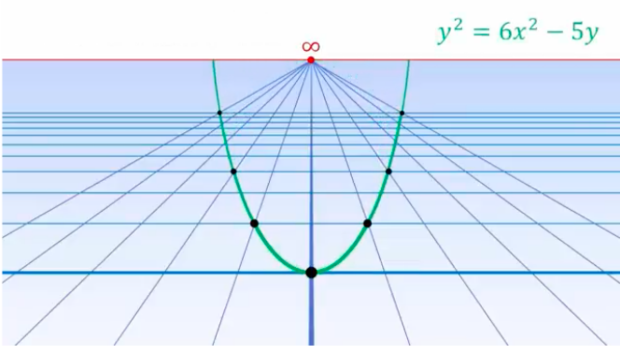


A hyperbola

Compactification

Example

The hyperbola limits to two points at infinity corresponding to the two asymptotes. The two branches of the hyperbola become "glued" together at these infinite points to form a single object.

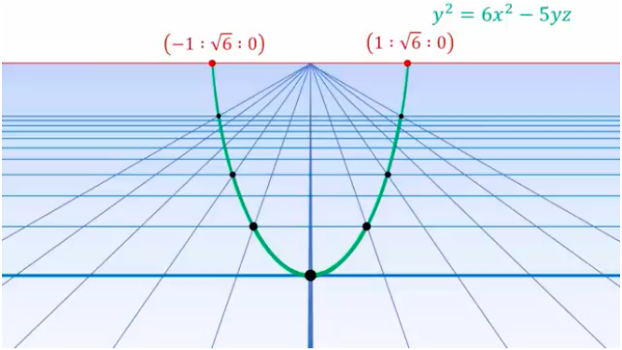


Compactifying the hyperbola

Compactification

Example

The hyperbola limits to two points at infinity corresponding to the two asymptotes. The two branches of the hyperbola become "glued" together at these infinite points to form a single object.



Compactifying the hyperbola

Plane curves

As these examples demonstrate, when we add in the points at infinity we are getting objects that honestly look like ovals (justifying the name).

The other change is to consider the curve $P(x, y) = 0$ over the complex numbers. Unfortunately it is hard to visualize this procedure:

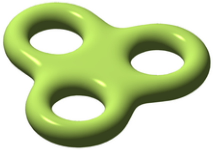
- The spaces $\mathbb{C}^2$ and $\mathbb{P}^2_{\mathbb{C}}$ have four real dimensions.
- The curve complex C will have two real dimensions.

Plane curves

Fact

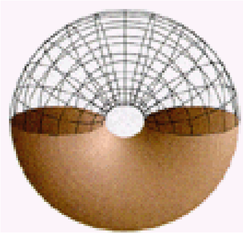
Suppose C is the closure of a plane curve $P(x, y) = 0$ inside of $\mathbb{P}^2_{\mathbb{C}}$. If C is smooth, then the geometric shape underlying C is a compact orientable surface.

To compute $\chi(C)$, all we need to do is count the number of holes!

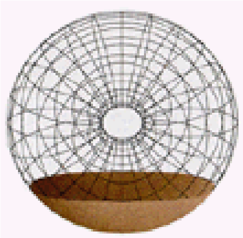


A compact algebraic curve

We will use the following visualization procedure. When we intersect C with the plane $\mathbb{R}^2 \subset \mathbb{C}^2$, we will obtain a "slice" of our surface. If we get lucky, this slice will intersect every hole. (However, it is also possible that our slice will miss some holes.)



A good slice



A bad slice

To estimate $\chi(C)$, we will count the number of ovals for the real curve $P(x, y) = 0$ in $\mathbb{R}^2$. We then guess that the number of holes in C is one less than the number of ovals. Using our formula for $\chi(C)$:

$$\begin{aligned}\chi(C) &= 2 - 2(\#\text{ of holes}) \\ &= 4 - 2(\#\text{ of ovals})\end{aligned}$$

This isn't quite right; there's no guarantee that when taking the slice $\mathbb{R}^2 \subset \mathbb{C}^2$ we actually see every hole in C. This may be a serious problem depending on which curve we picked.

To improve our guess, recall that $\chi(C)$ does not change if we "deform C continuously". In particular, every curve of degree d should have the same Euler characteristic! (The number of real ovals we see will change; the value of $\chi(C)$ will not.)

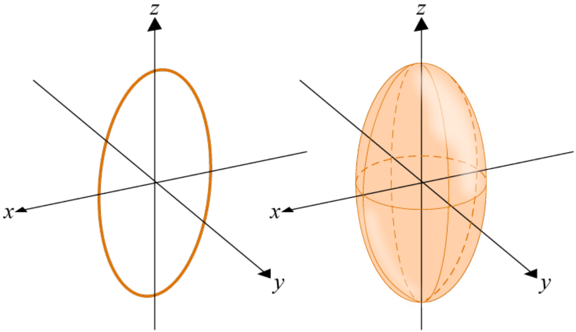
In summary, we pose:

Hypothesis

Suppose that r is the maximal possible number of ovals for a real plane curve of degree d. Then **every** curve of degree d will satisfy $\chi(C) = 4 - 2r$.

Example

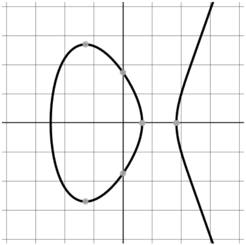
Suppose $P(x, y)$ has degree 2. Then $\chi(C) = 2$.



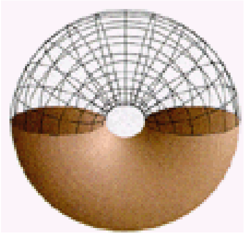
An elliptic slice of a surface with $\chi(S) = 2$

Example

Suppose $P(x, y)$ has degree 3. Then $\chi(C) = 0$.



An elliptic curve



A slice of a torus

Theorem (Harnack's Theorem, 1876)

Let $P(x, y)$ be an irreducible degree d polynomial with rational coefficients. Let $C \subset \mathbb{R}^2$ denote the solution set to $P(x, y) = 0$. If k denotes the number of ovals making up C, then

$$k \leq \frac{(d-1)(d-2)}{2} + 1.$$

Furthermore, some polynomials P of degree d will achieve this equality.

The proof passes through some results that are interesting in their own right.

Theorem

Let d be a positive integer and set $N = \frac{(d+2)(d+1)}{2} - 1$. Suppose we fix N different points $\{p_i\}_{i=1}^{N}$ in $\mathbb{R}^2$. There is a non-zero degree d polynomial $P(x, y)$ that vanishes at every point p_i.

For example, the following table shows that we can find a line through 2 points, a conic through 5 points, etc.

d	1	2	3	4	5
N	2	5	9	14	20

Proof.

A polynomial $P(x,y)$ of degree d is determined by $\frac{(d+2)(d+1)}{2}$ coefficients. Let's write down the general form of such a polynomial. If we denote the coefficient of $x^k y^\ell$ by a_{kl}, then

$$P(x,y) = \sum_{\substack{k+\ell \leq d \\ k,\ell \text{ non-neg}}} a_{k\ell} x^k y^\ell$$
$$= a_{d0}x^d + a_{(d-1)1}x^{d-1}y + \ldots + a_{10}x + a_{01}y + a_{00}$$

We also write write $p_i = (x_i, y_i)$ for $1 \leq i \leq N$. Note that each x_i and each y_i should be interpreted as a constant: it is given to us by the problem!

Proof.

Finding a suitable polynomial $P(x, y)$ is the same as finding coefficients satisfying

$$\begin{aligned} a_{d0}x_1^d + a_{(d-1)1}x_1^{d-1}y_1 + \ldots + a_{10}x_1 + a_{01}y_1 + a_{00} &= 0 \\ a_{d0}x_2^d + a_{(d-1)1}x_2^{d-1}y_2 + \ldots + a_{10}x_2 + a_{01}y_2 + a_{00} &= 0 \\ \vdots \qquad \vdots \qquad \vdots \qquad \vdots & \\ a_{d0}x_N^d + a_{(d-1)1}x_N^{d-1}y_N + \ldots + a_{10}x_N + a_{01}y_N + a_{00} &= 0 \end{aligned}$$

For emphasis: the x_is and y_is are constant!

This is a linear system in the variables $a_{k\ell}$ with N unknowns and $N - 1$ equations. This means that the solution set has dimension ≥ 1. In particular, there is always a non-zero solution. □

Theorem (Weak Bezout's Theorem)

Suppose that $P(x, y)$ and $Q(x, y)$ are two polynomials which do not share any factors. Let d and e denote the degrees of P and Q respectively. The number of points simultaneously satisfying

$$P(x, y) = 0$$
$$Q(x, y) = 0$$

is at most $d \cdot e$ (counted with multiplicity).

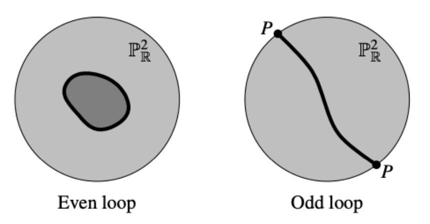
In fact, if we take the closures of $P(x, y) = 0$ and $Q(x, y) = 0$ in $\mathbb{P}^2$ then the number of simultaneous solutions is exactly $d \cdot e$ (counted with multiplicity).

Unfortunately we don't have time to prove this.

Proof. (Proof of Harnack's Theorem)

We will only prove the upper bound on the number of ovals k. Suppose for a contradiction that $P(x, y)$ has degree d and the curve C defined by $P(x, y) = 0$ has k ovals where $k \geq \frac{(d-1)(d-2)}{2} + 2$.

One can divide the ovals composing C into two types – even and odd – depending on whether or not they divide the plane into two pieces.

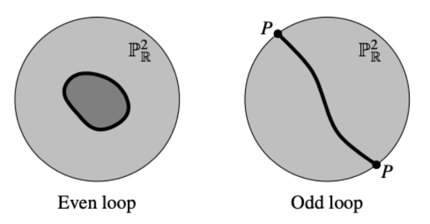


Introduction
Geometry of plane curves
Counting ovals
Trichotomy of curves

Proof. (Proof of Harnack's Theorem)

At most one oval composing C can be odd.

(Aside: we have $\pi_1(\mathbb{P}^2_{\mathbb{R}}) \cong \mathbb{Z}/2\mathbb{Z}$. An oval is even/odd depending on whether it represents 0 or 1 in the fundamental group. If C contains an odd loop, then the complement of this loop is simply connected so that every other oval in C is even.)



Proof. (Proof of Harnack's Theorem)

Next we construct a new curve $Q(x, y) = 0$ of degree $d - 2$ in the following way. Choose $\frac{(d-1)(d-2)}{2} + 1$ different even ovals $O_1, O_2, O_3, \ldots$ and choose 1 point on each of them. We also choose some other oval (even or odd) and choose $d - 3$ points on it. Altogether the number of points we have chosen is

$$\frac{(d-1)(d-2)}{2} + 1 + (d-3) = \frac{d(d-1)}{2} - 1$$

Our earlier theorem guarantees that there is a polynomial $Q(x, y)$ of degree $d - 2$ that vanishes at all of these points. We denote the curve $Q(x, y) = 0$ by D.

Proof. (Proof of Harnack's Theorem)

Now let's count how many times C and D intersect. D meets each oval O_j at a point p and either:

- D is tangent to O_j at p. In this case the multiplicity of intersection of D and C at p is ≥ 2.
- D enters the interior of O_j at p. In this case D must also leave O_j, adding at least one more point of intersection.

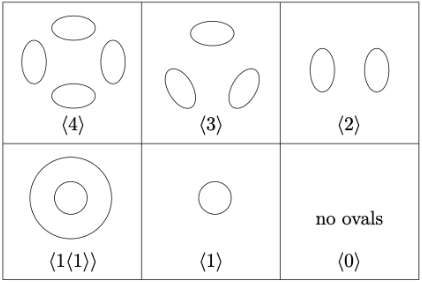
Either way, each O_j contributes at least 2 points to the intersection. This means that the total number of intersection points of C and D is at least

$$2 \cdot \left(\frac{(d-1)(d-2)}{2} + 1 \right) + (d-3) = d^2 - 2d + 1$$

This contradicts the upper bound $(\deg P) \cdot (\deg Q) = d(d-2)$ from Bezout's Theorem.

Question (Hilbert's 16th problem)

Suppose C is a smooth plane curve of degree d. What are the possible "arrangements" of the ovals in C? (Open for degree ≥ 8.)



Oval arrangements for degree 4

The following statement summarizes our work in the previous section.

Corollary

Let $P(x, y)$ be an irreducible degree d polynomial with rational coefficients. Let C be the closure of the curve $P(x, y) = 0$ in $\mathbb{P}^2_{\mathbb{C}}$. If C is smooth, then

$$\begin{aligned}\chi(C) &= 4 - 2 \cdot \left(\frac{(d-1)(d-2)}{2} + 1\right) \\ &= d(3-d).\end{aligned}$$

The first few values are:

degree	1	2	3	4	5	6
$\chi(C)$	2	2	0	-4	-10	-18

Euler characteristic

Remember our Guiding Principle: the curvature of C should determine the geometry/arithmetic of the curve.

Last time, we saw that the Euler characteristic measures the "total curvature". Our trichotomy is determined by whether this "total curvature is positive, zero, or negative. For plane curves, this turns into a trichotomy of degrees:

Trichotomy type	spherical	flat	hyperbolic
Euler characteristic	> 0	$= 0$	< 0
Degree of plane curve	$1, 2$	3	≥ 4

Rational points

Let's see how the trichotomy of curves interacts with rational points. Suppose that $P(x, y)$ is a polynomial with rational coefficients and the equation $P(x, y) = 0$ defines a smooth curve C.

- Degree 1: C always has infinitely many rational points.
- Degree 2: if C has a rational point, then $C(\mathbb{Q})$ can be analyzed projection away from a point.
- Degree 3: the set $C(\mathbb{Q})$ has a finite generating set under the operation $+$.

In both cases the key to our construction was intersecting C with lines through rational points. What about higher degrees?

Higher degrees

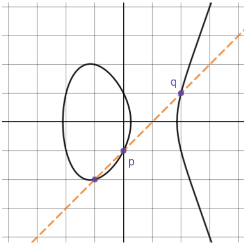
The analogue of our Key Observation in arbitrary degree is:

Theorem

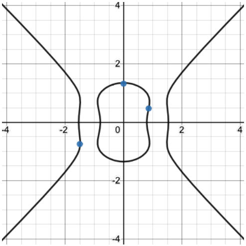
Suppose $a_n x^n + a_{n-1}x^{n-1} + \ldots + a_0 = 0$ is a degree n equation with rational coefficients. If this equation has $n-1$ rational solutions, then every solution is rational.

However, this is not so useful when $n > 3$. The issue comes from geometry!

If we have one or two points in the plane, we can always find a line through them. But if we have ≥ 3 points in the plane, they will "usually" not be collinear. So "usually" there is no way to generate new points by looking at lines through rational points.



2 points: always collinear



3 points: rarely collinear

Of course one can imagine trying to generate new rational points using some other method besides lines. However, a celebrated theorem of Faltings dashes our hopes:

Theorem (Faltings's theorem)

Let $P(x, y)$ be a polynomial of degree ≥ 4 with rational coefficients such that the equation $P(x, y) = 0$ defines a smooth curve C. Then $C(\mathbb{Q})$ is finite.

The following table summarizes the situation for polynomials in two variables:

degree	$1, 2$	3	≥ 4
number of rational points	infinite (if ≥ 1)	could be finite or infinite	finite

This is exactly the same trichotomy we noticed earlier!

For perspective, here is a more complete table describing the "trichotomy" for smooth projective curves over $\mathbb{Q}$.

Trichotomy type	spherical	flat	hyperbolic
genus	0 ($\mathbb{P}^1$)	1 (elliptic)	≥ 2
Euler characteristic / $\mathbb{C}$	> 0	$= 0$	< 0
Universal cover / $\mathbb{C}$	$\mathbb{P}^1$	$\mathbb{C}$	$\mathbb{H}$
Automorphisms / $\mathbb{C}$	PGL_2	$\approx$ itself	finite
number of rational points	infinite (if ≥ 1)	thin set	finite

Question

Is there a “trichotomy” of algebraic varieties in dimension > 1?

There are many famous conjectures in this direction. But very little has been proved!

Recently I have studied the behavior of rational points for higher dimensional varieties with “spherical” curvature, mainly over function fields of complex curves.

Exercises

Exercises:

1. Plane curves: study the relationship between affine curves in $\mathbb{R}^2$ and projective curves in $\mathbb{P}^2$.

Images made using Desmos and taken from Wikipedia, Bill Shillito, Thomas Banchoff, Andreas Gathmann, Viatcheslav Kharlamov and Oleg Viro