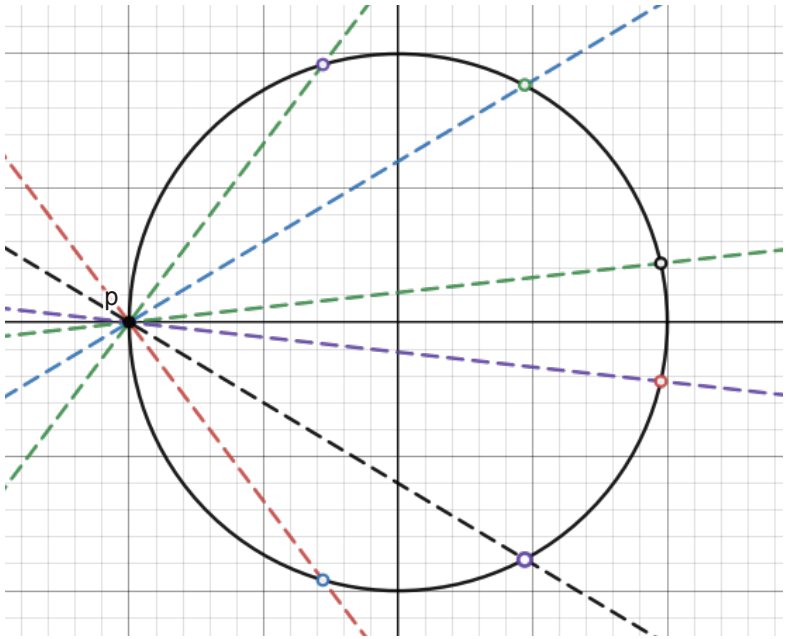
Lecture 3: Elliptic curves

Brian Lehmann
Boston College

Introduction

Suppose that $P(x, y) = 0$ is a polynomial equation with rational coefficients. Guiding question: what are the rational solutions to this equation?

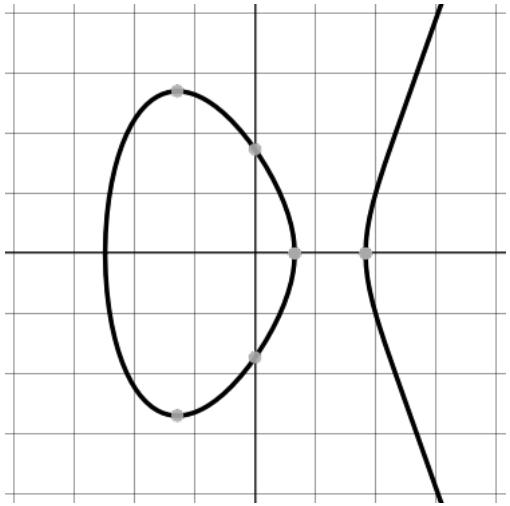
Earlier we studied the case when $P(x, y)$ has degree 2:



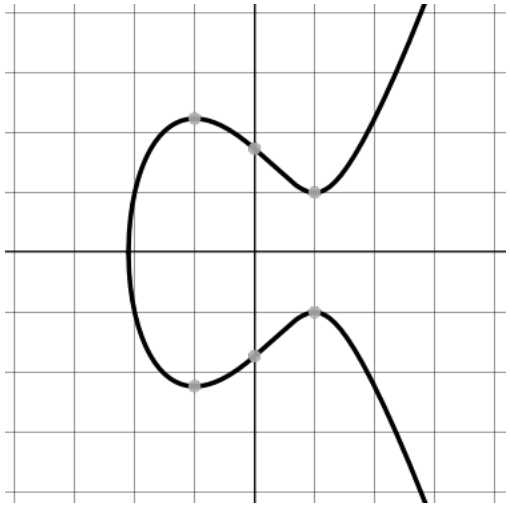
Rational points via projection

Introduction

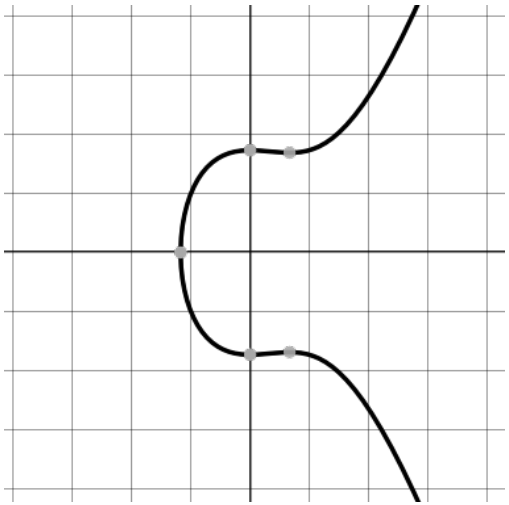
In this lecture, we study the equation $P(x, y) = 0$ when $P(x, y)$ has degree 3.



$y^2 = x^3 - 5x + 3$



$y^2 = x^3 - 3x + 3$



$y^2 = x^3 - x^2 + 3$

Describing cubics

(In this section we will be implicitly working with curves in $\mathbb{P}^2$, not in $\mathbb{R}^2$.)

Definition

Let $P(x, y)$ be a polynomial of degree 3 with rational coefficients. Consider the curve C defined by the equation $P(x, y) = 0$. We say that P is an elliptic curve if:

1. $P(x, y) = 0$ has at least one rational solution.
2. C is "smooth."[1]

[1] We define smoothness later.

Introduction

To make our lives a little easier, we will only work with equations which are in "reduced Weierstrass form":

$$y^2 = x^3 + ax^2 + bx + c$$

Fact

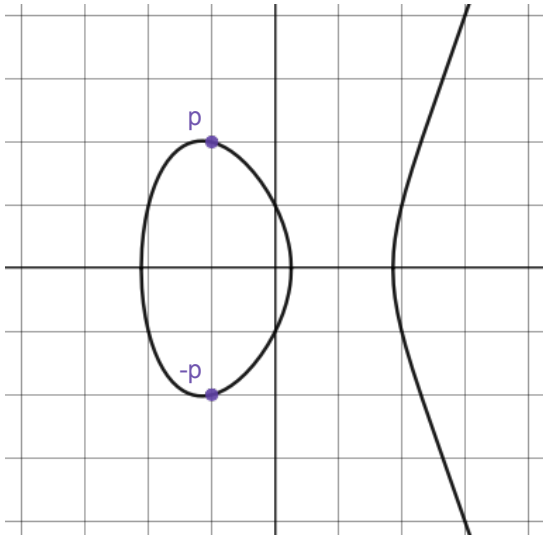
Every elliptic curve can be put in reduced Weierstrass form using a change of coordinates.[2]

[2] More precisely we apply a projective change of coordinates to the compactified curve in $\mathbb{P}^2$. This may have the side effect of moving the rational point "to infinity".

Rational points on cubics

We now turn to studying the rational solutions to a Weierstrass equation $y^2 = x^3 + ax^2 + bx + c$. We will denote the corresponding elliptic curve by C.

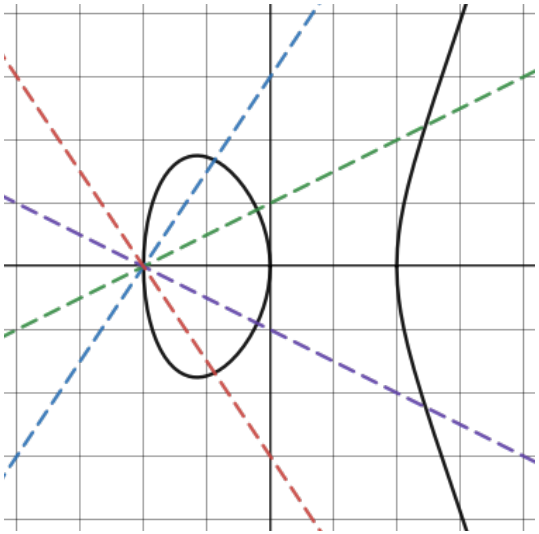
All such curves C have a symmetry: they are unchanged by reflection over the x-axis. In particular, if $p = (x, y)$ is a point on C with rational coordinates, then $-p = (x, -y)$ is also a point on C with rational coordinates.



p and $-p$

Rational points on cubics

Can we use projection from a point to find rational solutions to degree 3 equations?



Projection from a point

The answer is no: the argument doesn't work in the same way.

Rational points on cubics

For degree 2 polynomials, the key fact we used was:

Theorem

Suppose $ax^2 + bx + c = 0$ is a quadratic equation with rational coefficients. If this equation has one rational solution, then every solution is rational.

The corresponding statement for cubic equations is false! For example, the cubic equation $x^3 - 3x^2 - 2x + 6 = 0$ has solutions $x = 3, \sqrt{2}, -\sqrt{2}$. Only one is rational!

Rational points on cubics

The right analogue in degree 3 is:

Theorem

*Suppose $ax^3 + bx^2 + cx + d = 0$ is a cubic equation with rational coefficients. If this equation has **two** rational solutions, then every solution is rational.*

Proof.

We can factor

$$ax^3 + bx^2 + cx + d = a(x - r_1)(x - r_2)(x - r_3).$$

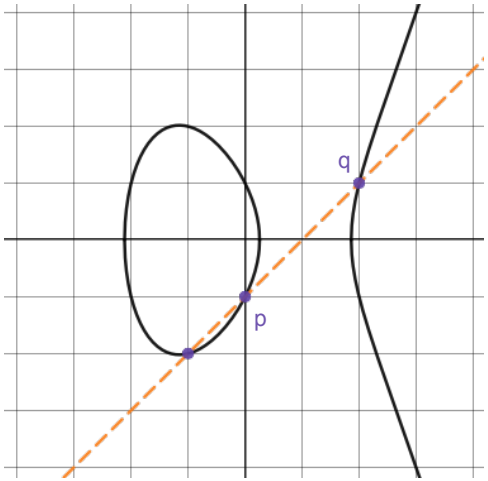
where r_1, r_2, r_3 are the (possibly complex) solutions to the equation. By comparing the coefficients of x^2 on both sides, we see that

$$b = a(-r_1 - r_2 - r_3) \implies r_3 = -r_1 - r_2 - \frac{b}{a}.$$

If two roots r_1, r_2 are rational, then r_3 is as well. □

Rational points on cubics

Geometrically this means the following. Let C be the curve $P(x, y) = 0$ where P has degree 3. Suppose that p_1, p_2 are **two** rational points on C. If we draw the line ℓ between them, the third intersection point with C will also be rational.

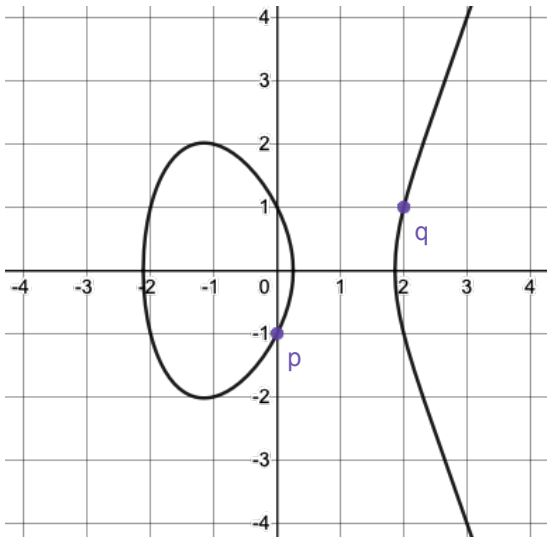


Third intersection point

Example

We can find this third intersection point explicitly by simultaneously solving the equations for ℓ and C.

Consider the elliptic curve $y^2 = x^3 - 4x + 1$. Two examples of rational points on this curve are $p = (0, -1)$ and $q = (2, 1)$.



p and q on C

Example

The line between p and q has equation $y = x - 1$. The third intersection point can be found by solving the equations

$$y = x - 1$$
$$y^2 = x^3 - 4x + 1$$

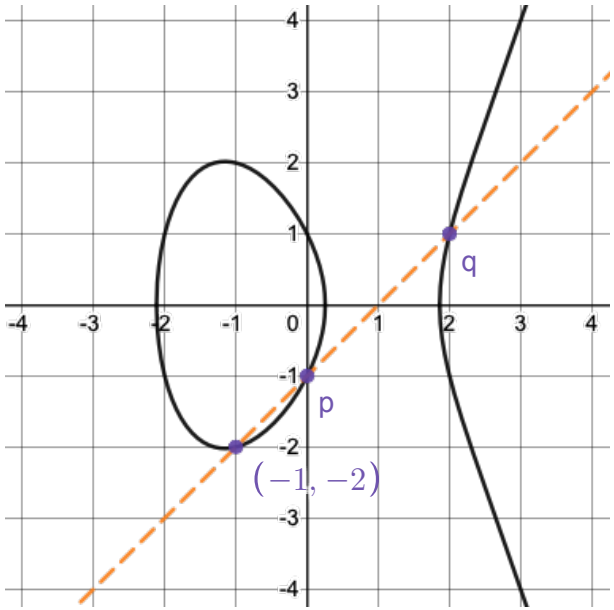
Substituting the top equation into the bottom and simplifying, we see that $x^3 - x^2 - 2x = 0$. Note that we already know two solutions to this equation, i.e. $x = 0$ and $x = 2$. We can find the third by factoring:

$$x^3 - x^2 - 2x = (x - 0)(x - 2)(x - ??).$$

Comparing the coefficient of x^2 shows that the last root is -1.

Example

This tells us that -1 is the x-coordinate of the third intersection point. Plugging in to the equation $y = x - 1$, we see that the last point is $(-1, -2)$.



Third intersection point

Rational points on cubics

We formalize this important operation:

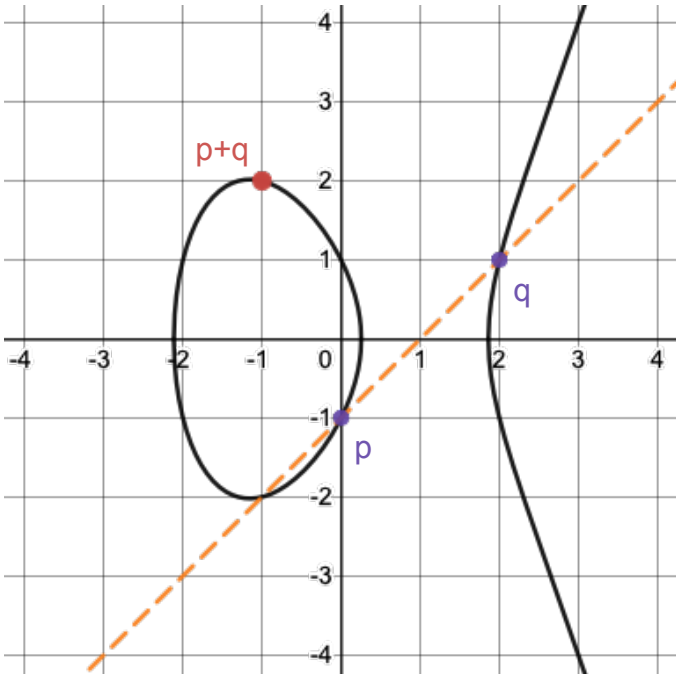
Definition

Given two rational points $p_1, p_2 \in C$, let ℓ denote the line through p_1, p_2 and let p_3 denote the third intersection point in $\ell \cap C$ (if such a point exists). We define

$$p_1 + p_2 = -p_3.$$

Understanding this operation is the key to unlocking the structure of rational points on C.

Rational points on cubics



Picturing $p + q$

Aside

It is natural to wonder: why do we define $p_1 + p_2 = -p_3$ instead of $p_1 + p_2 = p_3$?

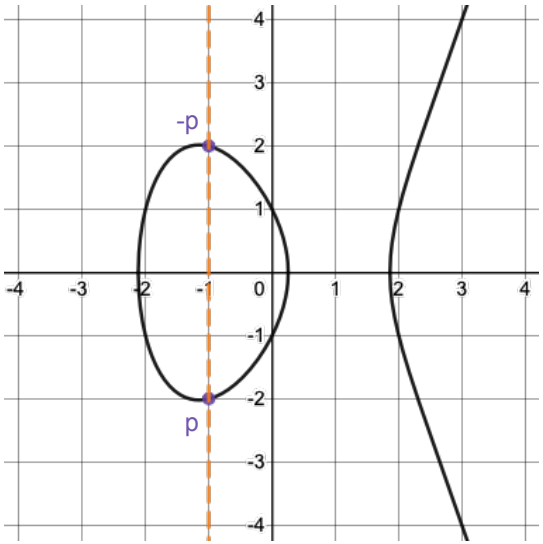
The reason is that it gives a better behaved operation. For example, it is only with our convention that $+$ becomes associative. In particular, the set of rational points on C has the structure of an abelian group under $+$.

To complete our description of $+$, there are a few "extra cases" we need to consider separately.

Rational points on cubics

Suppose we choose p and $-p$ for our two points. In this case the line ℓ is vertical and only meets C twice. We declare that the third meeting point is "at infinity":

$$p + (-p) = \infty$$



p plus $-p$

Rational points on cubics

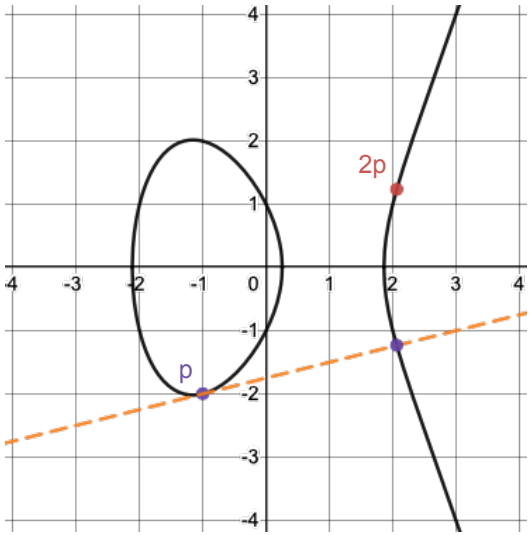
Since we have now added ∞ to our set of points, we also need to introduce rules for combining points with ∞:

$$p + \infty = p = \infty + p \qquad\qquad \infty + \infty = \infty$$

Note that ∞ behaves like the "identity": it doesn't change anything.

Rational points on cubics

The second special case is to consider what happens when we combine a point p with itself. We define $2p = p + p$ by taking the tangent line to C at p and computing the negative of the third point of intersection.



Tangent line to p

We think of the tangent line ℓ as "meeting C twice at p".

Rational points on cubics

Consider again the elliptic curve $y^2 = x^3 - 4x + 1$. The tangent line to the point $p = (-1, -2)$ has equation $y = \frac{1}{4}x - \frac{7}{4}$. Substituting into our cubic and simplifying, we get

$$x^3 - \frac{1}{16}x^2 - \frac{25}{8}x - \frac{33}{16} = 0$$

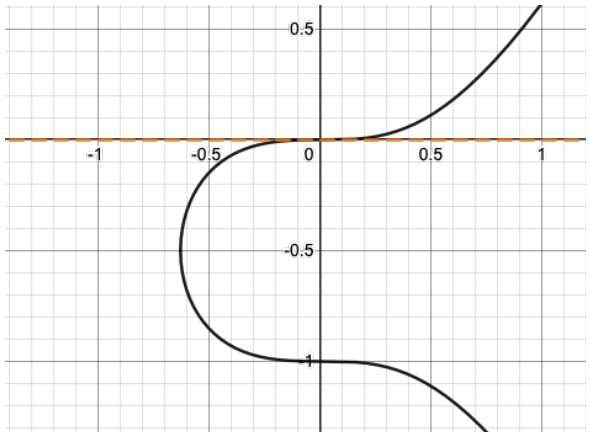
This polynomial has a double root at $x = -1$ (since our line is tangent to the curve at $(-1, -2)$). Using our factoring trick, we see the other intersection point has coordinates $(\frac{33}{16}, -\frac{79}{64})$. Thus

$$2p = \left(\frac{33}{16}, \frac{79}{64}\right).$$

Rational points on cubics

It is also possible for the line through p and q to be tangent to C at one of the points, say at p. This means that p is the third point of intersection, so $p + q = -p$.

It is even possible that when we substitute the equation for the line into the equation for the elliptic curve we will get a cubic with a triple root. Geometrically, this means that the line ℓ is tangent to C at an inflection point p. We agree that in this case the line ℓ "meets C three times at p" so that $p + p = -p$.



Triple tangent line

Mordell's Theorem

The $+$ operation allows us to generate new rational points from old ones.

Definition

Let $\{p_i\}$ be a subset of the rational points on C. We call $\{p_i\}$ a generating set if every rational point on C can be obtained by combining the p_i and their negatives $-p_i$ under the operation $+$.

Thus generating sets "capture" all the information about rational points on C.

Mordell's Theorem

We are now prepared to state the most important result about rational points on elliptic curves.

Theorem (Mordell's Theorem)

Let C be an elliptic curve. Then the set of rational points on C has a finite generating set $\{p_i\}_{i=1}^{r}$.

A finite amount of information captures all the rational points on C!

Mordell's Theorem

The key tool in the proof is heights of rational points. Remember, if we write $(x, y) = (\frac{a}{c}, \frac{b}{c})$ where $\{a, b, c\}$ are relatively prime integers then $H(x, y) = \max\{|a|, |b|, |c|\}$.

- Step 1: if we fix a positive number T, then only finitely many points $p \in C$ satisfy $H(p) \leq T$.
- Step 2: there is some positive number T such that any point $p \in C$ with height $> T$ can be written as a sum of points with smaller height.

Mordell's Theorem

Theorem

If we fix a positive number T, then only finitely many points $p \in C$ satisfy $H(p) \leq T$.

Proof.

Every such point has the form $(\frac{a}{c}, \frac{b}{c})$ where $|a|, |b|, |c| \leq T$. In particular, there are only finitely many choices for a, b, c and thus only finitely many points. □

Mordell's Theorem

Theorem

There is some positive number T such that any point $p \in C$ with height $> T$ can be written as a sum of points with smaller height.

There is a truly marvelous demonstration of the second step that this slide is too narrow to contain.

Proof.

See "Rational points on elliptic curves" by Silverman and Tate. □

Order

To obtain a better understanding of Mordell's Theorem, it is helpful to split rational points into two types. For a positive integer k we will use the shorthand

$$kp = \underbrace{p + p + \ldots + p}_{\text{k times}}$$

Definition

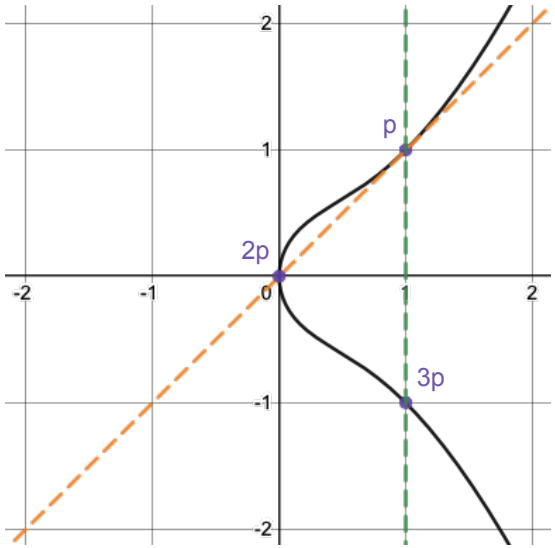
Suppose that p is a rational point on C. Consider the sequence of points $p, 2p, 3p, \ldots$. We say:

- p has finite order if ∞ appears on this list, or equivalently, if this list repeats itself. In this case, the first place that ∞ appears is called the order of p.
- p has infinite order if ∞ does not appear on this list, or equivalently, if there are no repetitions in this list.

Example

Consider the point $p = (1, 1)$ on the curve $y^2 = x^3 - x^2 + x$. We compute the order of p by writing down $p, 2p, 3p, 4p, \ldots$ until we reach ∞. We claim that p has order 4:

p	$2p$	$3p$	$4p$
$(1, 1)$	$(0, 0)$	$(1, -1)$	∞



A point of order 4

Example

Explanation:

p	$2p$	$3p$	$4p$
$(1,1)$	$(0,0)$	$(1,-1)$	∞

$2p$ is obtained by taking the tangent line $y = x$ at p, intersecting it with C, and taking the **negative** of the third intersection point. Thus we find $2p = (0,0)$.

$3p$ is obtained by taking $p + 2p$. We already know the line connecting p and $2p$: it is $y = x$. Since this line is tangent to the curve at p – i.e. "meets p twice" – the third point of intersection is p again. Thus $3p = -p = (1,-1)$.

$4p$ is $p + (-p) = \infty$.

Identifying finite order generators

We now address rational points of finite order and infinite order separately. Our first result allows us to find all points with finite order:

Theorem (Nagell-Lutz)

Let C be a smooth cubic curve with equation $y^2 = x^3 + ax^2 + bx + c$ where a, b, c are integers. Then every rational point (x, y) of finite order on C satisfies the following properties:

1. *Both x and y are integers.*
2. *The coordinate y is either 0 or a divisor of the discriminant*

$$D = -4a^3c + a^2b^2 + 18abc - 4b^3 - 27c^2.$$

Note that there are only finitely many points satisfying these conditions! (The requirement that a, b, c be integers can always be satisfied after rescaling the coefficients and the variables.)

Identifying infinite order generators

If C has at least one rational point of infinite order, then it will have infinitely many rational points of infinite order. To get a sense of "how many" rational points of infinite order there are, we make the following definition:

Definition

Let C be an elliptic curve. The rank of C is the minimal number of points of infinite order we need to include in a generating set.

Note that the rank is always finite by Mordell's Theorem.

Identifying infinite order generators

Unfortunately we do not have an algorithm for computing the rank of an elliptic curve! Although one can compute it in many examples, no one has been able to write down a definitive process for computing the rank in a finite amount of time.

Some of the most famous conjectures in mathematics are centered around the problem of computing ranks of elliptic curves. Let's look at a few of these conjectures.

Identifying infinite order generators

Question

Is the set of ranks of elliptic curves unbounded?

Currently, the highest rank of an elliptic curve that has been computed exactly is 20, although Elkies has shown that there are elliptic curves of rank ≥ 28.

Theorem (Bhargava-Shankar)

The average[3] *rank of an elliptic curve is bounded above by* $\frac{7}{6}$.

[3] Since there are infinitely many elliptic curves, we need to interpret "average" in an appropriate sense.

Identifying infinite order generators

The most famous conjecture about the rank of an elliptic curve is one of the Millennium Prize Problems:

Conjecture (Birch–Swinnerton-Dyer Conjecture)

The rank of an elliptic curve C is the order of the zero of the Hasse-Weil L-function $L(C, s)$ at the point $s = 1$.

The Hasse-Weil L-function is a function on a complex variable s that measures the "average number" of rational points on the reduction of C modulo a prime p. Just like the Riemann zeta function, the Hasse-Weil L-function is defined via analytic continuation from a certain region inside of $\mathbb{C}$ (namely $\mathrm{Re}(s) > \frac{3}{2}$). One consequence of Wiles' proof of Fermat's Last Theorem (and subsequent work of Breuil, Conrad, Diamond, Taylor) is that $L(C, s)$ can be defined holomorphically over all of $\mathbb{C}$.

One can see the BSD Conjecture as another touchstone between different areas of mathematics!

Exercises

Exercises:

1. Nagell-Lutz Theorem: practice working with points of finite order on elliptic curves.

Images made using Desmos.