

Introduction to Arithmetic Geometry

18.782

Andrew V. Sutherland

September 5, 2013

What is arithmetic geometry?

Arithmetic geometry applies the techniques of algebraic geometry to problems in number theory (a.k.a. **arithmetic**).

Algebraic geometry studies systems of polynomial equations (**varieties**):

$$\begin{aligned}f_1(x_1, \dots, x_n) &= 0 \\ &\vdots \\ f_m(x_1, \dots, x_n) &= 0,\end{aligned}$$

typically over algebraically closed fields of characteristic zero (like \mathbb{C}).

In arithmetic geometry we usually work over non-algebraically closed fields (like \mathbb{Q}), and often in fields of non-zero characteristic (like \mathbb{F}_p), and we may even restrict ourselves to rings that are not a field (like \mathbb{Z}).

Diophantine equations

Example (Pythagorean triples – easy)

The equation $x^2 + y^2 = 1$ has infinitely many rational solutions. Each corresponds to an integer solution to $x^2 + y^2 = z^2$.

Example (Fermat's last theorem – hard)

$x^n + y^n = z^n$ has no rational solutions with $xyz \neq 0$ for integer $n > 2$.

Example (Congruent number problem – unsolved)

A **congruent number** n is the integer area of a right triangle with rational sides. For example, 5 is the area of a $(3/2, 20/3, 41/6)$ triangle.

This occurs iff $y^2 = x^3 - n^2x$ has infinitely many rational solutions. Determining when this happens is an open problem (solved if BSD holds).

Hilbert's 10th problem



Is there a process according to which it can be determined in a finite number of operations whether a given Diophantine equation has any integer solutions?

The answer is **no**; this problem is formally undecidable (proved in 1970 by Matiyasevich, building on the work of Davis, Putnam, and Robinson).

It is unknown whether the problem of determining the existence of rational solutions is undecidable or not (it is conjectured to be so).

If we restrict to equations with just 2 variables (plane curves), then this problem is believed to be decidable, but there is no known algorithm for doing so (unless we also restrict the degree to 2).

Algebraic curves

A **curve** is an algebraic variety of dimension 1 (defined over a field k). In n -dimensional affine space k^n , this means we have a system of $n - 1$ polynomials in n variables ($n + 1$ variables in projective space).

For the sake of simplicity we will usually work with **plane curves** ($n = 2$) defined by a single polynomial equation. But we should remember that not all curves are plane curves (the intersection of two quadric surfaces in 3-space, for example), and we may sometimes need to embed our plane curves in a higher dimensional space in order to remove singularities.

While we will often define our curves using affine equations $f(x, y) = 0$, it is more convenient to work with projective curves defined by homogeneous equations $g(x, y, z) = 0$ in which every term has the same degree. Such a g can be obtained by **homogenizing** f (multiplying each term by an appropriate power of z); we then have $f(x, y) = g(x, y, 1)$.

The projective plane

Definition

The **projective plane** $\mathbb{P}^2(k)$ is the set of all nonzero triples $(x, y, z) \in k^3$, modulo the equivalence relation $(x, y, z) \sim (\lambda x, \lambda y, \lambda z)$ for all $\lambda \in k^*$.

The **projective point** $(x : y : z)$ is the equivalence class of (x, y, z) .

The projective plane

Definition

The **projective plane** $\mathbb{P}^2(k)$ is the set of all nonzero triples $(x, y, z) \in k^3$, modulo the equivalence relation $(x, y, z) \sim (\lambda x, \lambda y, \lambda z)$ for all $\lambda \in k^*$.

The **projective point** $(x : y : z)$ is the equivalence class of (x, y, z) .

We may call points of the form $(x : y : 1)$ **affine points**.

They form an affine plane $\mathbb{A}^2(k)$ embedded in $\mathbb{P}^2(k)$.

We may call points of the form $(x : y : 0)$ **points at infinity**.

These consist of the points $(x : 1 : 0)$ and the point $(1 : 0 : 0)$, which form the **line at infinity**, a projective line $\mathbb{P}^1(k)$ embedded in $\mathbb{P}^2(k)$.

(of course this is just one way to partition \mathbb{P}^2 as $\mathbb{A}^2 \cup \mathbb{P}^1$).

Plane projective curves

Definition

A **plane projective curve** C/k is defined by a homogeneous polynomial $f(x, y, z)$ with coefficients in k .¹

For any field K containing k , the **K -rational points** of C_f form the set

$$C(K) = \{(x : y : z) \in \mathbb{P}^2(K) \mid f(x, y, z) = 0\}.$$

A point $P \in C(K)$ is **singular** if $\frac{\partial f}{\partial x}$, $\frac{\partial f}{\partial y}$, $\frac{\partial f}{\partial z}$ all vanish at P .

C is **smooth** (or **nonsingular**) if there are no singular points in $C(\bar{k})$.

C is (geometrically) **irreducible** if $f(x, y, z)$ does not factor over \bar{k} .

We will always work with curves that are irreducible.

We will usually work with curves that are smooth.

¹Fine print: up to scalar equivalence and with no repeated factors in $\bar{k}[x, y, z]$.

Examples of plane projective curves over \mathbb{Q}

affine equation	$f(x, y, z)$	points at ∞
$y = mx + b$	$y - mx - bz$	$(1 : m : 0)$
$x^2 - y^2 = 1$	$x^2 - y^2 - z^2$	$(1 : 1 : 0), (1, -1, 0)$
$y^2 = x^3 + Ax + B$	$y^2z - x^3 - Axz^2 - Bz^3$	$(0 : 1 : 0)$
$x^4 + y^4 = 1$	$x^4 + y^4 - z^4$	none
$x^2 + y^2 = 1 - x^2y^2$	$x^2z^2 + y^2z^2 - z^4 + x^2y^2$	$(1 : 0 : 0), (0 : 1 : 0)$

These curves are all irreducible.

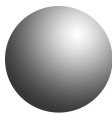
The first four are smooth (provided that $4A^3 + 27B^2 \neq 0$).

The last curve is singular (both points at infinity are singular).

Genus

Over \mathbb{C} , an irreducible smooth projective curve is a connected compact \mathbb{C} -manifold of dimension 1 (a compact Riemann surface).

Every such manifold is topologically equivalent to a sphere with handles. The number of handles is the (topological) **genus**.



genus 0



genus 1



genus 2



genus 3

Later in the course we will give an algebraic definition of genus that works over any field (and agrees with the topological genus over \mathbb{C}).

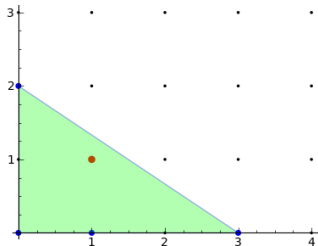
The genus is the single most important invariant of a curve.

Newton polygon

Definition

The **Newton polygon** of a polynomial $f(x, y) = \sum a_{ij}x^i y^j$ is the convex hull of the set $\{(i, j) : a_{ij} \neq 0\}$ in \mathbb{R}^2 .

An easy way to compute the genus of a (sufficiently general) irreducible curve defined by an affine equation $f(x, y) = 0$ is to count the integer lattice points in the interior of its Newton polygon:²



$$y^2 = x^3 + Ax + B.$$

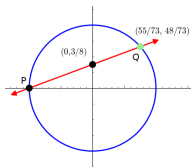
²See <http://arxiv.org/abs/1304.4997> for more on Newton polygons.

Birational equivalence

A **rational map** between two varieties is a map defined by rational functions (which need not be defined at every point). For example, the map that sends the point $(0, t)$ on the vertical line $x = 0$ to the point

$$\left(\frac{1 - t^2}{1 + t^2}, \frac{2t}{1 + t^2} \right)$$

on the unit circle $x^2 + y^2 = 1$ is a rational map of plane curves.



The inverse map sends (x, y) on the unit circle to $(0, y/(x + 1))$. Varieties related by rational maps ϕ and φ such that both $\phi \circ \varphi$ and $\varphi \circ \phi$ are the identity at all points where both are defined are **birationally equivalent**.

The genus is invariant under birational equivalence (as are most things). We will give a finer notion of equivalence (isomorphism) later in the course.

Curves of genus 0

As may be seen by their Newton polytopes, plane curves C of genus 0 are (in general) either **conics** (degree 2 curves), or of the form $y = g(x)$.

In the latter case we can immediately parameterize the points of $C(k)$: the affine points are all of the form $(t, g(t))$, with $t \in k$.

Curves of genus 0

As may be seen by their Newton polytopes, plane curves C of genus 0 are (in general) either **conics** (degree 2 curves), or of the form $y = g(x)$.

In the latter case we can immediately parameterize the points of $C(k)$: the affine points are all of the form $(t, g(t))$, with $t \in k$.

In the case of a plane conic $ax^2 + bxy + cy^2 + dx + ey + f = 0$ over a field whose characteristic is not 2, after a linear change of variables (complete the square at most 3 times) we obtain an affine equation of the form

$$ax^2 + by^2 = 1.$$

If this equation has a rational solution (x_0, y_0) , then the substitution $x = (1 - x_0u)/u$ and $y = (t - y_0u)/u$ yields

$$\begin{aligned} a(1 - 2x_0u + x_0^2u^2) + b(t^2 - 2y_0tu + y_0^2u^2) &= u^2 \\ a + bt^2 - (2ax_0 + 2by_0t)u &= 0, \end{aligned}$$

and the parameterization $(t, g(t))$, where $g(t) = \frac{a+bt^2}{2ax_0+2by_0t} \in k(t)$.

Rational points in genus 0

Thus a curve of genus 0 either has no rational points, or it has infinitely many. Another way to see this is to note that any line with rational slope that passes through a rational point on a degree-2 curve intersects the curve in another rational point (this gives a slope parameterization).

Rational points in genus 0

Thus a curve of genus 0 either has no rational points, or it has infinitely many. Another way to see this is to note that any line with rational slope that passes through a rational point on a degree-2 curve intersects the curve in another rational point (this gives a slope parameterization).

Example

$x^2 + y^2 = 3$ has no rational solutions.

$x^2 + y^2 = 2$ has infinitely many rational solutions.

The question of integer solutions is slightly more delicate.

For example, $x^2 + y^2 = 2$ has only four integer solutions $(\pm 1, \pm 1)$. But $x^2 - 2y^2 = 1$ has infinitely many integer solutions (x_n, y_n) , where $x_n + y_n\sqrt{2} = (3 + 2\sqrt{2})^n$.

The first few are $(3, 2)$, $(17, 12)$, $(577, 408)$, \dots

Curves of genus 1 – elliptic curves

Let E/k be a smooth irreducible genus 1 curve with a rational point P . Then E is called an **elliptic curve** (which is not an ellipse!).

Note: not every smooth irreducible genus 1 curve has a rational point. Over \mathbb{Q} , the curve defined by $x^3 + 2y^3 = 4$ is such a case.

Every elliptic curve can be put in **Weierstrass form**:



$$y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6,$$

where P corresponds to the point $(0 : 1 : 0)$ at infinity.

If k has characteristic different from 2 or 3 we can even put C in the form

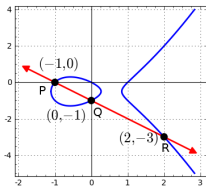
$$y^2 = x^3 + Ax + B.$$

Rational points in genus 1

Now let E be an elliptic curve over \mathbb{Q} defined by a Weierstrass equation.

If P is a rational point and ℓ is a line through P with rational slope, it is not necessarily true that ℓ intersects E in another rational point.

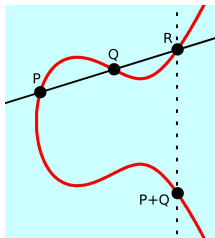
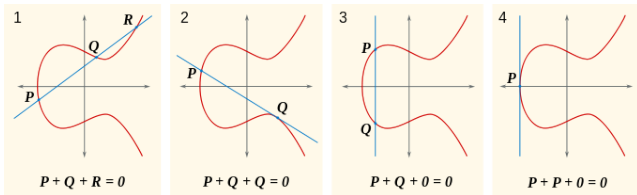
However, if P and Q are two rational points on E , then the line \overline{PQ} intersects E in a third rational point R (this follows from Bezout's theorem and a little algebra). This allows us to generate new rational points from old ones.



Even better, it allows us to define a group operation on $E(\mathbb{Q})$, or on $E(k)$ for any elliptic curve E defined over any field k .

The elliptic curve group law

Three points on a line sum to zero, the point at infinity.



The elliptic curve group law

With addition defined as above, the set $E(k)$ becomes an abelian group.

- ▶ The point $(0 : 1 : 0)$ at infinity is the identity element 0 .
- ▶ The inverse of $P = (x : y : z)$ is the point $-P = (x : -y : z)$.
- ▶ Commutativity is obvious: $P + Q = Q + P$.
- ▶ Associativity is not so obvious: $P + (Q + R) = (P + Q) + R$.

The computation of $P + Q = R$ is purely algebraic.

The coordinates of R are rational functions of the coordinates of P and Q , and can be defined over any field.

Rational points on elliptic curves

The group $E(\mathbb{Q})$ may be finite or infinite, but in every case it is **finitely generated**.

Theorem (Mordell 1922)

The group $E(\mathbb{Q})$ is a finitely generated abelian group. Thus

$$E(\mathbb{Q}) \simeq T \oplus \mathbb{Z}^r,$$

*where the **torsion subgroup** T is a finite abelian group corresponding to the elements of $E(\mathbb{Q})$ with finite order, and r is the **rank** of $E(\mathbb{Q})$.*

It may happen (and often does) that $r = 0$ and T is the trivial group. In this case the only element of $E(\mathbb{Q})$ is the point at infinity.

The group $E(\mathbb{Q})$

The torsion subgroup T of $E(\mathbb{Q})$ is well understood.

Theorem (Mazur 1977)

The torsion subgroup of $E(\mathbb{Q})$ is isomorphic to one of the following:

$$\mathbb{Z}/n\mathbb{Z} \quad \text{or} \quad \mathbb{Z}/2\mathbb{Z} \oplus \mathbb{Z}/2m\mathbb{Z},$$

where $n \in \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 12\}$ and $m \in \{1, 2, 3, 4\}$.



Barry Mazur (and some other guy)

The ranks of elliptic curves over \mathbb{Q}

The rank r of $E(\mathbb{Q})$ is **not** well understood.

Here are some of the things we do not know about r :

1. Is there an algorithm that is guaranteed to correctly compute r ?
2. Which values of r can occur?
3. How often does each possible value of r occur, on average?
4. Is there an upper limit, or can r be arbitrarily large?

The ranks of elliptic curves over \mathbb{Q}

The rank r of $E(\mathbb{Q})$ is **not** well understood.

Here are some of the things we do not know about r :

1. Is there an algorithm that is guaranteed to correctly compute r ?
2. Which values of r can occur?
3. How often does each possible value of r occur, on average?
4. Is there an upper limit, or can r be arbitrarily large?

We do know a few things about r . We can compute r in most cases where r is small. When r is large often the best we can do is a lower bound; the largest example is a curve with $r \geq 28$ due to Elkies (2006).



Noam Elkies

The ranks of elliptic curves over \mathbb{Q}

The most significant thing we know about r is a bound on its average value over all elliptic curves (suitably ordered).

The following result is very recent and is still being improved.

Theorem (Bhargava, Shankar 2010-2012)

The average rank of all elliptic curves over \mathbb{Q} is less than 1.



Manjul Bhargava



Arul Shankar

It is believed that the average rank is exactly $1/2$.

The group $E(\mathbb{F}_p)$

Over a finite field \mathbb{F}_p , the group $E(\mathbb{F}_p)$ is necessarily finite.

On average, the size of the group is $p + 1$, but it varies, depending on E .

The following theorem of Hasse was originally conjectured by Emil Artin.

Theorem (Hasse 1933)

The cardinality of $E(\mathbb{F}_p)$ satisfies $\#E(\mathbb{F}_p) = p + 1 - t$, with $|t| \leq 2\sqrt{p}$.



Emil Artin



Helmut Hasse

The group $E(\mathbb{F}_p)$

Over a finite field \mathbb{F}_p , the group $E(\mathbb{F}_p)$ is necessarily finite.

On average, the size of the group is $p + 1$, but it varies, depending on E .

The following theorem of Hasse was originally conjectured by Emil Artin.

Theorem (Hasse 1933)

The cardinality of $E(\mathbb{F}_p)$ satisfies $\#E(\mathbb{F}_p) = p + 1 - t$, with $|t| \leq 2\sqrt{p}$.



Emil Artin



Helmut Hasse

The fact that $E(\mathbb{F}_p)$ is a group whose size is not fixed by p is unique to genus 1 curves. This is the basis of many useful applications.

For curves C of genus $g = 0$, we always have $\#C(\mathbb{F}_p) = p + 1$.

For curves C of genus $g > 1$, the set $C(\mathbb{F}_p)$ does not form a group.

(But as we shall see, there is way to associate an abelian group to C .)

Reducing elliptic curves over \mathbb{Q} modulo p

Let E/\mathbb{Q} be an elliptic curve defined by $y^2 = x^3 + Ax + B$, and let p be a prime that does not divide the **discriminant** $\Delta(E) = -16(4A^3 + 27B^2)$.

The elliptic curve E is then said to have **good reduction** at p .

If we reduce A and B modulo p , we obtain an elliptic curve $\bar{E} = E \bmod p$ defined over the finite field $\mathbb{F}_p \simeq \mathbb{Z}/p\mathbb{Z}$.

Thus from a single curve E/\mathbb{Q} we get an infinite family of curves \bar{E} , one for each prime p where E has good reduction.

Now we may ask, how does $\#\bar{E}(\mathbb{F}_p)$ vary with p ?

The Sato-Tate conjecture

We know $\#\bar{E}(\mathbb{F}_p) = p + 1 - a_p$ for some integer a_p with $|a_p| \leq 2\sqrt{p}$.
So let $x_p = a_p/\sqrt{p}$. Then x_p is a real number in the interval $[-2, 2]$.

What is the distribution of x_p as p varies? Is it uniform?

<http://math.mit.edu/~drew>

The Sato-Tate conjecture

We know $\#\bar{E}(\mathbb{F}_p) = p + 1 - a_p$ for some integer a_p with $|a_p| \leq 2\sqrt{p}$.
So let $x_p = a_p/\sqrt{p}$. Then x_p is a real number in the interval $[-2, 2]$.

What is the distribution of x_p as p varies? Is it uniform?

<http://math.mit.edu/~drew>

The Sato-Tate conjecture, open for nearly 50 years, was recently proven.

Theorem (Taylor et al., 2006 and 2008)

*Let E/\mathbb{Q} be an elliptic curve without complex multiplication.
Then the x_p have a semi-circular distribution.*



Mikio Sato



Richard Taylor



John Tate

The Birch and Swinnerton-Dyer conjecture

There is believed to be a relationship between the infinite sequence of integers a_p associated to an elliptic curve E/\mathbb{Q} and the rank r .

The **L -function** $L_E(s)$ of an elliptic curve E/\mathbb{Q} is a function of a complex variable s that “encodes” the infinite sequence of integers a_p .

For the “bad” primes that divide $\Delta(E)$, one defines a_p to be 0, 1, or -1 , depending on the type of singularity E has when reduced mod p .

$$L_E(s) = \prod_{\text{bad } p} (1 - a_p p^{-s})^{-1} \prod_{\text{good } p} (1 - a_p p^{-s} + p^{1-2s})^{-1} = \sum_{n=0}^{\infty} a_n n^{-s}$$

The Birch and Swinnerton-Dyer conjecture

Based on extensive computer experiments (back in the 1960s!), Bryan Birch and Peter Swinnerton-Dyer made the following conjecture.

Conjecture (Birch and Swinnerton-Dyer)

Let E/\mathbb{Q} be an elliptic curve with rank r . Then

$$L_E(s) = (s - 1)^r g(s),$$

for some complex analytic function $g(s)$ with $g(1) \neq 0, \infty$. In other words, r is equal to the **order of vanishing** of $L_E(s)$ at 1.



Bryan Birch



EDSAC



Sir Peter Swinnerton-Dyer

They subsequently made a much more precise conjecture about $L_E(s)$, but there is already a \$1,000,000 bounty on the conjecture above.

Fermat's Last Theorem

Theorem (Wiles et al. 1995)

$x^n + y^n = z^n$ has no rational solutions with $xyz \neq 0$ and integer $n > 2$.

We may assume that n prime and x and y are relatively prime integers.

Suppose $a^n + b^n = c^n$ with $n > 3$ (Euler addressed the case $n = 3$).

Consider the elliptic curve E/\mathbb{Q} defined by

$$y^2 = x(x - a^n)(x - b^n).$$

Serre and Ribet proved that E is **not modular**.

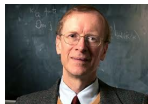
Wiles (with help from Taylor) proved that every **semistable** elliptic curve, including E , is **modular**. So no solution $a^n + b^n = c^n$ can possibly exist.



J.-P. Serre



Ken Ribet



Sir Andrew Wiles



Richard Taylor

Curves of genus greater than 1

The most important thing to know about rational points on curves of genus greater than 1 is that there are only finitely many.

Theorem (Faltings 1983)

Let C/\mathbb{Q} be an irreducible curve of genus greater than 1. Then the cardinality of $C(\mathbb{Q})$ is finite.

This result was originally conjectured by Mordell in 1922 and later generalized to **number fields** (finite extensions of \mathbb{Q}). Faltings actually proved the generalization to number fields.

Faltings' theorem is notably **ineffective**. Even though we know the cardinality of $C(\mathbb{Q})$ is finite, we have no upper bound on its size.



Louis Mordell



Gerd Faltings

Jacobians of curves

As noted earlier, for curves C of genus $g > 1$, there is no natural group operation. However, it is possible to define a g -dimensional variety $\text{Jac}(C)$, the **Jacobian** of C , whose points admit a commutative group operation; $\text{Jac}(C)$ is an example of an **abelian variety**.

An elliptic curve E is isomorphic to $\text{Jac}(E)$ (as abelian varieties).

We will restrict our attention to Jacobians of **hyperelliptic curves**. Over fields of characteristic different from 2, hyperelliptic curves of genus g can be defined by an affine equation of the form

$$y^2 = f(x),$$

where $f(x)$ is a polynomial of degree $2g + 1$ or $2g + 2$.

The case $g = 1$ corresponds to an elliptic curve.

Every curve of genus 2 is a hyperelliptic curve (not true in higher genus).