CUMC Talk
An Algebraic view on the Arithmetic of Elliptic Curve
Sachin Kumar
University of Waterloo, Faculty of Mathematics

Notation

We will follow the standard Bourbaki notations: $\mathbb{F}$ denotes a field, K denotes a number field, $\mathbb{Q}$ denotes the set of rational numbers and $\mathbb{C}$ denotes the set of complex numbers.

What is it?

Firstly, elliptic curve is not an ellipse! An elliptic curve is an abelian variety of dimension, ie., genus 1.

Definition 0.1. *algebraic variety is the set of solutions of a system of polynomial equations over the* $\mathbb{R}$ *or* $\mathbb{C}$*. Non-irreducible algebraic varieties are called algebraic sets.*

(Abelian variety is a projective algebraic variety that is also an algebraic group, i.e., has a group law that can be defined by regular functions). What makes it non-singular? The group law of an abelian variety is necessarily commutative and the variety is non-singular.

Definition 0.2. *char* $(\mathbb{F}) = p$*, a prime, is the size of the smallest subfield in the field, or 0, if this subfield has* ∞ *size.*

We talk about elliptic curves in characteristic $\neq 2$ or 3. The equation y^2 equal to a cubic in x is called the Weierstrass normal form (affine form). A curve is non-singular, if it has no points at which both the partial derivatives vanish, ie., E does not have any singular points, which are points where the curve fails to be smooth.

Geometrical View

$\mathcal{O}$, point at infinity acts like an identity element, how? since if we take a point $P \in E(\mathbb{F})$, where $\mathbb{F}$ is a field, then geometrically we can draw a line L through P that tends to a point of infinity, $\mathcal{O}$. But the L does not intersect another point at $\mathcal{O}$, so the only point through which L intersects is P, so $P \oplus \mathcal{O} = P$. Why $P \oplus (-P) \neq 0$? Since $0 \in E(F)$, ie., 0 is not a point on the elliptic curve. Process of adding two points in $E/\mathbb{F}$: Start with two points P and Q on $E/\mathbb{F}$. Draw a line L through P and Q. L intersects at a third point on E, namely R. Draw a vertical line through R, which hits another point on E, which is $P \oplus Q$. Process of adding same point in $E/\mathbb{F}$: Start with a point P on $E/\mathbb{F}$. Draw a tangent T on P. T intersects at a second point on E/F, namely R. Draw a vertical line, through R which hits another point on E, which is $P \oplus P = 2P$. So the reflection point $-P$ also exists on the curve, opposite to the point P.

Isogenies

Isogenies are algebraic maps or morphisms between E and E' that preserve the group structure. We ask for functions from E to E', to be algebraic (Each coordinate is a rational function, that are two variables, meaning just quotients of polynomials in 2 variables). Concretely:

$$\phi : E \to E'$$
$$\phi(x, y) = (\phi_x(x, y), \phi_y(x, y))$$
$$\phi_x(x, y) = \frac{f_1(x, y)}{f_2(x, y)}$$
$$\phi_y(x, y) = \frac{g_1(x, y)}{g_2(x, y)}$$

where f_1, f_2, g_1, and g_2 are all polynomials. The degree of an isogeny is its degree as an algebraic map. (one of the ways you can get this structure, i.e isogeny, is in group theory. We learnt that when you have a group homomorphism, you have a concept of a kernel, which are points that are annihilated by the homomorphism, and by the first isomorphism theorem, which says that if you know the kernel and you also know that it's a surjective homomorphism, then the homomorphism is uniquely determined by that information, more or less. So there is essentially upto isomorphism, only one surjective group homomorphism for any subgroup of a group, and the idea is that this subgroup should form the kernel). If the group is non-abelian, you need to worry about the normal subgroups.
Let H be any finite subgroup of E, (H must be finite, because these isogenies are $\mathbb{Q}$-maps, so essentially they are polynomials although we are taking quotients, but the idea is that a polynomial can only annihilate finitely many points, unless its the zero polynomial, since a polynomial has only finitely many roots. So what it means is that by the virtue of being algebraic, an isogeny can only having a finite kernel). Vélu formulas reverses these idea, so instead of starting with an isogeny and taking its kernel, which is a finite subgroup, you now start with this finite subgroup and then question, which isogenies can you construct having this finite subgroup as its kernel? and just as is in group theory, where you have this first isomorphism theorem, that tells you in essence that the answer is only one. So here the answer is also only one (for the reverse process), ie., there is exactly one isogeny upto isomorphism, bearing a given kernel, So you might wonder, what about the surjectivity condition? because in the first isomorphism theorem you need the surjectivity, in order for that to work, imagine a polynomial map, ask whether this polynomial is surjective?, well not necessarily, since it can be a constant map (which we dont care, mainly when it is a zero polynomial!), so if I give a non-constant polynomial, is it surjective? You guys may tell, well if we take $f = x^2 + 1$, then it has no solutions in $\mathbb{R}$, yeh sure but if you go to the complex plane, it has a solution, since $\mathbb{C}$ is algebraically closed field, and you can generalize that every polynomial has a root over the algebraic closure of whatever field you are working with, So if you work in the algebraic closure, every polynomial is surjective. So then these isogenies must be surjective and being surjective means that the kernel determines the homomorphism upto isomorphism. Then the map given by $P \mapsto (X, Y)$ where,

$$X = x(P) + \sum_{Q \in H \setminus \{\infty\}} \Big(x(P+Q) - x(Q) \Big)$$
$$Y = y(P) + \sum_{Q \in H \setminus \{\infty\}} \Big(y(P+Q) - y(Q) \Big)$$

is an isogeny ϕ with domain E and kernel H, and the quotient group, E/H denotes the co-domain of ϕ. This co-domain is unique upto isomorphism, and the degree of the isogeny is equal to the order of H or $|H|$. One thing we can prove about isogenies using the geometry of the curve is that whenever we have a finite subgroup of an elliptic curve, essentially you can mod out by that finite subgroup and you can get an isogeny which represents the quotient map. We can take two examples of special cases of vélu's formula, where $H = 2, 3$. So let $E : y^2 = x^3 + ax + b$. Suppose $H = \{\mathcal{O}, P\}$, then $P \oplus P = \mathcal{O}$ (if we go back to the group law, $P \oplus P$ means the secant line through P, ie., the tangent line through P and, the third point, $\mathcal{O}$ exists on this tangent line, which means that this line has to be a vertical line and the curve is symmetric around x-axis so this vertical line has to cross the curve at a point where $y = 0$), so $P = (x_P, 0)$ with $E' : 0 = x_P^3 + ax_P + b$. By vélu's formulas, we get the corresponding isogeny with this kernel,

$$E/H : y^2 = x^3 + (a - 5(3x_P^2 + a))x + (b - 7x_P(3x^2 + a))$$
$$\phi_H(x, y) = \left(x + \frac{3x_P^2 + a}{x - x_P}, y - \frac{y(3x_P^2 + a)}{(x - x_P)^2} \right)$$

and we get $\mathbb{Q}$-functions. Now, if we take an example of isogenies of degree 3 where $H = \{\infty, P, -P\}$. Then $P = (x_P, y_P)$ with $0 = 3x_P^4 + 6ax_P^2 - a^2 + 12bx_P$ and $y_P^2 = x_P^3 + ax_P + b$. We have,

$$E/H : y^2 = x^3 + (a - 10(3x_P^2 + a))x + (b - 28y_P^2 - 14x_P(3x_P^2 + a))$$
$$\phi_H(x, y) = \left(x + \frac{2(3x_P^2 + a)}{x - x_P} + \frac{4y_P^2}{(x - x_P)^2}, y - \frac{8yy_P^2}{(x - x_P)^3} - \frac{2y(3x_P + a)}{(x - x_P)^2} \right)$$

So given a curve E and a kernel H, you can construct a projection map from $E \to (E \bmod H)$ and there is some elliptic curve structure on $(E \bmod H)$, ie, E modular any finite subgroup group gives an elliptic curve, and the projection map can be interpreted as an isogeny. For every such choice of curve and finite subgroup, there is essentially a unique isogeny (up to isomorphism) that starts from that curve and annihilates that kernel, so $(E \bmod H)$ is essentially unique, so the projection map is also unique. Degree of an isogeny, is defined as a degree of the algebraic maps, in the setting of finite subgroups and kernels, the degree is just the cardinality of the kernel. If interested, we have a set of formula's called the vélu's formulas that gives us equations to construct such separable isogenies. If I have an elliptic curve and a finite subgroup, then I can use vélu's formulas to get an equation for this isogeny. A question may arise, what if the subfield H is infinite, does it still preserve the structure? No. In the context of elliptic curves, the operation of **modular reduction or mod** p refers to reducing the coordinates of points on the curve modulo a prime p. When dealing with a finite subgroup H, the points on the elliptic curve are defined over a finite field. In this case, reducing the coordinates modulo p preserves the finite structure of the subgroup H, and resulting set of points still form an elliptic curve. If H is infinite subgroup, reducing the coordinates modulo p does not preserve the infinite structure of H. The resulting set of points will not satisfy the necessary properties of an elliptic curve. We can conclude that, E mod H gives an elliptic curve $\iff$ H is a finite subgroup.

Rank and Mordell-Weil Theorem

Previously, we told that for an elliptic curve over a field $\mathbb{Q}$, the abelian group of $\mathbb{Q}$-points on E is denoted by $E(\mathbb{Q})$ is called the mordell-Weil group. One question, you might ask is can we always start with some finite set of $\mathbb{Q}$-points so that just by playing with this addition operation, we can produce all the other $\mathbb{Q}$-points? This was one of the very first fundamental question asked about elliptic curves. This was answered by the celebrated theorem of Mordell, which states that the group $E(\mathbb{Q})$ of $\mathbb{Q}$-points on E is a finitely generated abelian group and later, Weil generalized this result for all Abelian varieties on a number field, $A(K)$. In elementary terms, using a finite set of $\mathbb{Q}$-points we can generate all the other $\mathbb{Q}$-points using the addition law, that lie on this elliptic curve E. In terms of group theory, since $E(\mathbb{Q})$ is thus a finitely generated abelian group, we can use addition law to produce all the $\mathbb{Q}$-points from some finite starting set, and the fundamental theorem of abelian group says (FTAG) if you have a finitely generated abelian group, it always looks like the product of cyclic groups. Some no. of infinite cyclic groups which looks like $\mathbb{Z}$, so $E(\mathbb{Q})$ since its finitely generated, by FTAG, $E(\mathbb{Q})$ has to look like some product of copies of infinite cyclic groups to the power of some r (i.e, r copies) times some finite abelian group T, which will also be product of finite cyclic groups.

Theorem 0.3. *If E is an elliptic curve over $\mathbb{Q}$, then*

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

for some $r \in \mathbb{Z}_{\geq 0}$ and T is a finite abelian group, mostly $T = E(\mathbb{Q})^{tors}$.

$E(\mathbb{Q})^{\text{tors}}$ denotes the torsion points of the Mordell-Weil group on $E/\mathbb{Q}$, where torsion points of $E(\mathbb{Q})$ are the points of finite order in the group, ie, points that satisfy,

$$nP = \mathcal{O}$$

for some $n \in \mathbb{Z}_+$. For $E/\mathbb{Q}$, Mazur classified which groups appear as the torsion part of $E(\mathbb{Q})$, mainly the torsion part of $E(\mathbb{Q})$ is isomorphic to one of the following 15 groups, ie,

$$\begin{aligned} \mathbb{Z}/n\mathbb{Z}, &\quad n \in \{1, \ldots, 10, 12\} \\ \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2n\mathbb{Z} &\quad n \in \{1, \ldots, 4\} \end{aligned}$$

For E/K, where K is a number field, the idea is non-trivial. In general, the torsion subgroup of $E(K)$ is finite, but its structure can be quite complicated. There has been some recent work on computing the torsion subgroups of elliptic curves over certain number fields, such as multi-quadratic fields (these are number fields of degree $2n$ represented as the composite of n quadratic fields, in other words, its is a field extension of the $\mathbb{Q}$ that can be obtained by adjoining square roots of $\mathbb{Q}$). We can kind of see from this

that, a big part of this group is measured by this r, how many infinite cyclic groups do we have inside E?, that kind of measures the size of $E(\mathbb{Q})$. It is a theorem of Mazur that the group T is bounded by size 16, so it never gets any big so even in a stronger sense it is really this r that's measuring the size of this set of solutions $E(\mathbb{Q})$. Thus r measures how "big" the group $E(\mathbb{Q})$ is, the quantity r is called the arithmetic rank of $E/\mathbb{Q}$. The rank of E is the r that comes out that measures the number of infinite cyclic groups that you're getting inside this group of $\mathbb{Q}$-points. In elementary terms if you're not familiar with group theory, the rank of E just measures the number of points that you need to start with on our elliptic curve so that every other points can be obtained just by connecting pairs and finding the third point of intersection and repeating the procedure. SO what is the minimum number of points that we need to start with to get all the information about all the $\mathbb{Q}$-points on the elliptic curve by just using the addition law, and that's the rank. We know by Mordell's theorem that this number if always finite, you can always start with some finite set of $\mathbb{Q}$-points so that all the $\mathbb{Q}$-points can be found, by connect-the-dots procedure (addition law) and the minimal number of points that you need is essentially the rank. So r is the fundamental invariant that measures how many solutions you have for E. If $r = 0$ corresponds to the case where $E(\mathbb{Q})$ is finite, but if $r \geq 1$, then the number of $\mathbb{Q}$-points is infinite, and then r measures how infinite?, then bigger the tr, the more infinite the set of $\mathbb{Q}$-points. In group theory, the rank of $E/\mathbb{Q}$ is the size of the smallest torsion-free subgroup of the Mordell-Weil group, ie., $E(\mathbb{Q})$. In elementary term, it is the smallest set of independent points that generate all the other points (independent generators) on $E/\mathbb{Q}$. In elementary terms, the rank is the group of independent points on the Elliptic curves, from which all the other points in $E(\mathbb{Q})$ can be generated. Talk about Mordell-Weil group, generalized Mordell's theorem and Falting's theorem for abelian varieties. Mordell-Weill group is not necessarily for $E/\mathbb{Q}$, but its for any abelian group associated to any abelian variety over a number field K, denoted A/K.

Elliptic curves with (CM) Complex Multiplication

Can the rank, $r \in \mathbb{Q}$? Surprisingly Yes!, mostly in the case where the curve belongs to a special class called elliptic curves with complex multiplication (CM), ie., where E has an endomorphism ring, larger than the $\mathbb{Z}$, End $\nsubseteq \mathbb{Z}$. This occurs when the elliptic curve has complex multiplication by an order in an imaginary quadratic field, that is maximal, because the endomorphism ring contains elements beyond those coming from the maximal order. These additional elements can multiply the independent $\mathbb{Q}$-points by the non-integer values, leading to a fractional rank. The endomorphism ring contains $\mathbb{C}$, and it can be isomorphic to an order in an imaginary quadratic field. Let $E(K)$, denote the Elliptic curve over an imaginary quadratic field K,

$$\begin{aligned} \text{rank } E(K) &= \text{ maximal order } + \text{ Non-maximal order} \\ &= r + \frac{s}{d} \\ &\in \mathbb{Q} \end{aligned}$$

where r denotes the regular arithmetic rank, $s \in \mathbb{Z}_{\geq 0}$, d denotes the index of the order in K, with respect to the maximal order.

On the recent rank conjectures

There many natural fundamental questions that one can ask about the rank, but unfortunately they are still unsolved! For example, What is the maximum that the rank of an elliptic curve can be? No one has any idea, we dont even know whether the rank can be bounded. Does a maximum even exist? we dont know, the rank may go all the way up to infinity or stop at 30, I say 30 because the current record for the largest rank ever founded is 28 by Noam Elkies in 2006. We can ask more statistical questions, what is the expected size (ie., average size) of the rank? Do most curves have rank 0 or 1? (Can one prove that even say 1% of all elliptic curves have rank 0 or 1)? Is there any algorithm to determine the rank of an elliptic curve, that will provably terminate? The Birch-Swinnerton Dyer conjectures exactly addresses the last question.

As in often the case in number theory, to try and solve equations over the $\mathbb{Z}$ or $\mathbb{Q}$, the key idea is to look

at solutions in mod p, ie., $\mathbb{F}_p$. So, you look at the normal form of E, reduce it mod p and then you count the number of solutions in $\mathbb{Z}/p\mathbb{Z}$, in other words just in that finite field.

$$y^2 \equiv x^3 + ax + b \pmod{p} \tag{0.1}$$

So in 1960, BSD did some computations on the ranks of elliptic curves and they also did computations on the number of solutions mod p on these elliptic curves, so what are the maximum number of solutions mod p to an elliptic curve?, remember it is a two variable equation so x and y will each have about p possibilities, so the maximum number of solutions is p^2 and so they will loop over all the p^2 possibilities over x and y and they would count the number of solutions. Actually, the correct order of magnitude of how many points you will actually have on this curve, one expects that modulo p have about p points, and not p^2 points. Why is that? once you plug in a value of $x \pmod{p}$, then you can solve for y, y^2 equal this number mod p, that will have at most two solutions, but how often do you expect, when you plug in a value mod p for (0.1), you don't expect it to be a square all the time, it will be a square about half the time, because half the numbers mod p are squares. So when you plug in a number mod p for x, half the time you will get a square and then y will take one of those two square roots as values, so half the time that $x^3 + ax + b$ will be a square and half the time when it is a square y will have two solutions, so when you plug in the p possible values for x, about half of them will be a square and then you will two solutions for y, so the expected number of solutions is p, cause about $\frac{p}{2}$ values of x will give you a square, but for each of those you have two solutions. So you have about p solutions for (x, y) mod p on this curve. So you expect about p solutions. We denote N_p as the number of points mod p on this curve, so you expect N_p to be about p and therefore you expect $\frac{N_p}{p}$ to be about 1, most of the time. BSD observed that if E has lots of $\mathbb{Q}$-points, in other words if this rank r is very large, then if you take those $\mathbb{Q}$ numbers and reduce them mod p and you can do it since mod p is a finite field, so you can divide, by reducing it you will get a point mod p on the elliptic curve and if there are lots $\mathbb{Q}$-points and reduce them mod p, you should be getting lots and lots of points on the elliptic curve mod p and so some how N_p should be pushed a little higher than p in those cases. BSD hypothesized that if the rank of E is large, then on average one should notice E having more than p points modulo p, the expected number of p points in that case. If the rank is 0, then N_p should be about p, but if there lots of $\mathbb{Q}$-points and the rank is large, then N_p tends to be a little higher than p, as you vary across primes p. So maybe you should be able to pick up this large rank by noticing the points mod p for lots of primes p. This hypothesis led to this following conjecture, that

Conjecture 0.4 (Birch-Swinnerton Dyer)**.** *Let E be an elliptic curve, let r be its rank, and let N_p denote the number of points on $E \pmod{p}$. Then,*

$$\prod_{p \leq X} \frac{N_p}{p} \sim c \cdot (\log X)^r$$

where X is some large number.

So in elementary terms, they conjectured that if you take $\frac{N_p}{p}$ which tends to be very close to 1 and multiply $\frac{N_p}{p}$ over all primes $\leq X$, then as X grows this should also grow as c times $\log X$ to the rank. This is exactly what we were saying before that if the rank is large, then some how these $\frac{N_p}{p}$ should be maybe a little bit bigger than 1, so that when you take the product it grows as $\log X$, and if you take the product over all primes leX it actually grows in a way that depends on the rank r. If $r = 0$, then the product $\frac{N_p}{p}$ for all $p \leq X$ as X gets large, they would remain bounded, in other word $\frac{N_p}{p} = 1$, otherwise $\frac{N_p}{p}$ grows in a away that depends on the r. This is the *weak form* of the conjecture.
Now, what is c? BSD gave an explicit expression for c in terms of E. There is a modern formulation for the BSD conjecture, is in terms of L-function of the elliptic curve. Let E be an elliptic curve, and let N_p denote the number of points on $E \pmod{p}$. We said that N_p tends to be about p. So we can measure the deviation from that expected value of p or $p+1$ by setting $a_p = p+1-N_p$, now a_p tends to be fairly small, and then you can define what's called the incomplete L-function of E (incomplete, since we omit the Euler

factor's for primes $p \mid 2\Delta$) by, taking the product over all primes that do not divide the discriminant,

$$L(E,s) = \prod_{p \nmid 2\Delta} \frac{1}{1 - a_p p^{-s} + p^{1-2s}}$$

The L-function just depends on the count of the number of points mod p on the curve. You think of it as function of $s \in \mathbb{C}$, you can find that the product converges for all $\Re(s) > \frac{3}{2}$.

Conjecture 0.5 (Hasse)**.** *$L(E,s)$ should have a holomorphic continuation as a function of s to the entire complex plane.*

The conjecture states that $L(E,s)$ should have a analytic continuation to the entire complex plane, and if it did then it makes sense to talk about the value and order of vanishing of this holomorphic function of s at $s = 1$. When you plug in $s = 1$, well then you notice the partial products of this function are exactly the reciprocals of,

$$\prod_{p \leq X} \frac{N_p}{p}$$

and that might have to do something with the BSD's original conjecture. They reformulated it in terms of values at $s = 1$, since they recently knew about the analytic continuation. Guess how!

Conjecture 0.6 (Mordern Birch-Swinnerton Dyer)**.** *The rank of E is equal to $ord_{s=1}L(E,s)$.*

where, $\text{ord}_{s=1}L(E,s)$ is the order of vanishing of the L-function at $s = 1$. In other words, if you take the taylor expansion around $s = 1$, The first taylor coefficient will look like, $c(s-1)^r$ and all the other terms will be higher order terms. Formally, the Taylor expansion of $L(C,s)$ at $s = 1$ has the form,

$$L(C,s) = c(s-1)^r + \text{higher order terms}$$

(ie., order of vanishing) with $c \neq 0$ and $r = \text{rank}(E(\mathbb{Q}))$. Remember, now it looks like the reciprocal, so the rate of going to infinity has been replaced by the rate of going to 0, the order to which it is vanishing is measuring the growth of $\frac{N_p}{p}$, for $p \leq X$. So the conjecture just says that the, rank of E is equal to the order of vanishing at $s = 1$ of a particular analytic function. In particular this conjecture asserts that,

$$L(C,1) = 0 \iff |E(\mathbb{Q})| = \infty$$

One of the great contributions of Wiles's modularity theorem, which proved for semistable elliptic curves ($E/\mathbb{Q}$ is semistable at prime q if it is isomorphic to an $E'/\mathbb{Q}$ which modulo q either is non singular or has a singular point with two distinct tangent direction. An $E/\mathbb{Q}$ is called semistable if it is semistable at every prime, so it order words the curve has good reduction at all but finitely many primes, and at the bad primes, the curve has only nodes or cusps as singularities) and rest was proved by Richard Taylor, Brian Conrad and Fred Diamond, is that it allows the modern formulation of BSD conjecture to make sense,

Theorem 0.7 (Wiles, 1995)**.** *For any E, the L-function $L(E,s)$ has a holomorphic (analytic) continuation to the entire complex plane.*

This is a consequence of the Modularity theorem. This theorem allows us to know that the L-function is defined at $s = 1$, which makes our modern formulation make sense. And the reason that the L-function has an analytic continuation is basically because its is modular, and we know that modular forms are analytic in the entire complex plane and therefore, the L-function is also analytic through a technique of Mellin transform. So because modular functions are defined over the entire complex plane, so are the L-functions of these elliptic curves, if the elliptic curves are modular. So now, we know that the L-function is defined at $s = 1$ and so the modern formulation of BSD conjecture intuitively makes sense. Let E be a elliptic curve over $\mathbb{Q}$, and let r be it's rank. By (0.7), there exists an $r' \in \mathbb{Z}$ such that the Taylor expansion of an analytic function $L(E,s)$ at $s = 1$ is of the form $a(s-1)^{r'}$+ higher order terms, where $a \neq 0$. What does analytic mean? it just means that you have a power series expansion at that point. This r' is exactly what it means to talk about the order of the vanishing at $s = 1$ and the quantity r' is called the analytic rank of $E/\mathbb{Q}$. It is the exponent of the first term in the Taylor series at $s = 1$, where we now know it's known to be defined by the modularity theorem.

Conjecture 0.8 (Birch, Swinnerton-Dyer)**.** *Let r be the algebraic rank and r' be the analytic rank of E, then $r = r'$.*

In an overview, rank is defined in terms of order of the group of solutions of $\mathbb{Q}$-points on this elliptic curve, that is totally algebraic quantity; and r' is the order of vanishing of this analytic function that is defined in terms of the point counts mod p, you make a generating function called the L-function and look at the vanishing, which is totally analytic construction that's called r'; and BSD says that $r = r'$.
Quite a few is known about the BSD conjecture! The first theoretical evidence was actually in Wiles's PhD thesis with John Coates,

Theorem 0.9 (Coates and Wiles (1977))**.** *If E is an elliptic curve of the form $y^2 = x^3 + Ax$ or $y^2 = x^3 + B$, and if $r' = 0$, then $r = 0$.*

Thus there are finitely many $\mathbb{Q}$-points for the curves of the following form. The result actually applied to a more general class of elliptic curves, not just those two forms, although those two kinds were sided for the most, because those were part of the original BSD computations. More generally, the Coates-Wiles Theorem applies to any elliptic curve having complex multiplication by the ring of integers of an imaginary quadratic field of class number one, ie., the endomorphism ring of E is larger than the $\mathbb{Z}$. Where $y^2 = x^3 + Ax$ has complex multiplication inside $\mathbb{Q}[i]$ and $y^2 = x^3 + B$ has CM inside $\mathbb{Q}[\sqrt[3]{-1}]$, ie., cube root of unity. This was the first indication, that progress could be made relating the BSD. Why did they try only for this special class of elliptic curves? because these were the initial curves, on which BSD did computations on. The results of Coates-Wiles was extended to general elliptic curves and also for the case of $r' = 1$,

Theorem 0.10 (Gross, Zagier, and Kolyvagin (1989))**.** *Let E be an* ***modular*** *elliptic curve.*

(1) *If $r' = 0$, then $r = 0$.*
(2) *If $r' = 1$, then $r = 1$.*

For this, an assumption was made that E is modular, which was not known at that point. Now let's think! What about the converse, ie., if $r = 0$, then $r' = 0$? Not yet known. Inspired by and taking further these ideas of Coates-Wiles, Gross-Zagier-Kolyvagin, as well as Mazur-Wiles, and Wiles on the Iwasawa Theory, some kind of converse was proved, although it had lots of restrictions on E,

Theorem 0.11 (Skinner, Urban, and Zhang (2013))**.** *Let E be an elliptic curve that satisfies some further technical conditions (for some prime $p \geq 5$, E has p-Selmer rank 0 or 1, E had good ordinary or multiplicative reduction at p, etc...).*

(1) *If $r = 0$, then $r' = 0$.*
(2) *If $r = 1$, then $r' = 1$.*

The height of an elliptic curve. With all these technical hypothesis, do any elliptic curves satisfy these conditions?
We may write any $E/\mathbb{Q}$ in the form, $E_{A,B} : y^2 = x^3 + Ax + B$, where $A, B \in \mathbb{Z}$.

Definition 0.12 (Height)**.** *The size of the coefficients of the defining equation is the height of E.*

You can just list all elliptic curves $E/\mathbb{Q}$, in increasing size of A and B, ie. order of increasing height.

Definition 0.13 (Naive Height)**.** *If $E = E_{A,B}$, then*

$$H(E_{A,B}) := \max\{4|A|^3, 27B^2\}$$

where, we are just taking the maximum of the two terms of Δ of E. We can ask statistical questions relating to the rank and to the probability that BSD is satisfied for these elliptic curves, what proportions of curves satisfies the conditions laid out in (0.11)? The following statistical questions are being ask recently, Do most elliptic curves have small rank?

Conjecture 0.14. 100% *of elliptic curves has rank* 0 *or* 1*.*

100% of E does not mean **all** E, it just means 100%!, there could be 0% of E with higher rank, in a density sense, just like primes are 0% of $\mathbb{Z}$, even though there are infinitely many primes. From recent works, we now know that **most** elliptic curves have rank 0 or 1.

Theorem 0.15 (Bhargava and Shankar (2013))**.** *At least* 83% *of all elliptic curves have rank* 0 *or* 1*.*

Corollary 0.16. *A positive proportion (*$< 1\%$*) of elliptic curves satisfy the BSD conjecture.*

This led to ask, what proportion do all currents results (incl. Dokchitser's brothers result on quotients modulo squares) actually allow us to prove BSD?

Theorem 0.17 (Bhargava, Skinner and Wei Zhang)**.** *The BSD conjecture is true for more than* 66% *of all elliptic curves.*

What remains to be done? Everything we have talked about so far has been about curves of rank 0 or 1, which are conjectured to be 100% of elliptic curves. The technical conditions in the theorem of Skinner, Urban, and Zhang must still be removed; once this is accomplished, this would likely mean that we understand BSD for 100% for elliptic curves! However, its is the remaining 0% of curves, having a rank at least 2, creates the complexity. While rare, there are infinitely many elliptic curves having the rank at least 2. There are now many beautiful extensions of the BSD conjecture, such as Bloch-Kato conjecture, which extends the BSD beyond cubic equations; there is also p-adic analogues of the BSD, that is instead of looking at the L-function as a complex analytic function, you can look at it as a p-adic analytic function and their corresponding conjectures, like p-adic version of the Gross-Zagier formula (Bertolini-Darmon-Prasanna), and there are generalizations over number fields other than $\mathbb{Q}$ (Shou-wu Zhang for proving Gross-Zagier totally real number fields, ...).

Thank you!

Due to some time constraints, I couldn't fill in topics about how $E/\mathbb{Q}$ played a major role in the proof of Fermat's Last Theorem and the applied field of Elliptic curve cryptography, in general, curve-based cryptography; other algebraic-based cryptography protocols.
Other algebraic structures of E, such as Shafarevich-Tate group, Ш and Selmer group; and their implication to BSD.

Book Recommendations
For undergraduates who want to start learning about the abstract theory of elliptic curves and its applications, I woulf really recommend two Springer UTM books, Rational points on Elliptic curves by Joseph Silverman and John Tate; and An introduction to mathematical cryptography by Joseph Silverman.

References

1. Elliptic curve isogenies Based Public Key Cryptography assumptions by David Jao.
2. The problem Landscape of SIDH by David Jao.
3. What is the Birch-Swinnerton Dyer Conjecture, and what is known about it? by Manjul Bhargava
4. Introduction to Birch-Swinnerton Dyer Conjecture by Richard E. Borcherds.
5. Arithmetic of Elliptic curves by John Tate.
5. Birch-Swinnerton Dyer conjecture by Andrew Wiles