Talk XII

The Birch and Swinnerton-Dyer Conjecture

H. Liang & F. Naccarato

" I am simply here to open windows for other people and talk with them about what they see and I hope they will do the same for me." — Martin Buber

Today we are moving to a rather fancy topic (finally!) – the Birch and Swinnerton-Dyer (BSD) conjecture, which is an important open problem in the arithmetic theory of elliptic curves. Hence, in contrast to previous talks where systematic theories and rigorous proofs are presented in detail, my partner Francesco Naccarato and I would like to give a short survey on this amazing topic instead and leave the proof as an exercise.[1]

12.1 *Some Historical Notes*

Before we actually get started, I would prefer to say something about the "bigger picture", e.g. history, background, motivations, etc. There may be some notions that you are currently not familiar with, but fear not: we will develop what we need in more detail. Also, you may feel free to skip this section and go back to it whenever you like.

The following discussions borrow (copy) heavily from Keqin Feng[2] [1] (§3.5-§3.6), Silverman & Tate [2] (introduction) and Stewart & Tall [3] (§12.1-§12.4).

One important and fundamental problem in algebraic geometry is to classify algebraic varieties, and in particular, algebraic curves, under birational equivalence. When the genus is ≥ 1, there are in general many equivalence classes of algebraic curves, even of the same genus. However, from a view point of number theory, when studying points on algebraic curves over a number field (or more generally, a *global field*) K , there are only three types of them —

The first one is *rational curves*, i.e., algebraic curves of genus 0, which consist of lines and quadratic curves. A rational curve defined on K has either no points in K (we shall call them K-points) or infinitely many ones.

The second one is curves of genus ≥ 2. In 1923, Louis J. Mordell conjectured that

Conjecture 12.1 (Mordell, 1923) *Any algebraic curve over rational numbers $\mathbb{Q}$ of genus greater than 1 has finitely many rational points.*

which was confirmed by Gerd Faltings in 1983. For instance, the Fermat's last theorem states that there are no non-trivial intergal solutions to the equation

$$x^n + y^n = z^n$$

[1] This is an exercise that worths 1,000,000 $. In fact, the BSD conjecture was chosen as one of the seven Millennium Prize Problems listed by the Clay Mathematics Institute, who has offered a 1,000,000 prize for the first correct proof.

[2] This enlightening book is in Chinese. But as far as I know, there is no English version of it...

for any integer $n > 2$, which is equivalent to that "the only rational points on $F : x^n + y^n - 1 = 0$ are $(0, \pm 1)$ and $(\pm 1, 0)$". Since the genus of this curve is $n(n-1)/2$, at least now we know that there are at most finitely many rational points on it.[1]

The third one is those of genus 1, now known as *elliptic curves*. An elliptic curve over a global field K may have finitely or infinitely many K-points. More interestingly, as we shall see explicitly later, the set of all K-points (together with the point $\mathcal{O}$ at infinity) of an elliptic curves E can be naturally given an operation to make it into an abelian group $E(K)$. Moreover, Mordell proved that this group is finitely generated.[2]

Theorem 12.2 (Mordell-Weil, 1928) *The group $E(K)$ is finitely generated.*

Before you might get too proud of how powerful mathematics is, new questions just pour into our minds: what can we say about this group? How to compute its rank? Can we find a set of generator? What are the possible structures of the torsion part of this group? $\cdots$

More restrictedly, what if we are only interested in integral points on an algebraic curve C? Are there any integral point? If so, are there infinitely many? Or more generally, what if we are concerned with hypersurfaces (i.e., more than two variables) instead of curves?

In fact, some of these questions are fully answered. For instance, Siegel proved in the 1920s that any non-singular cubic equation has only finitely many integer solutions, and in 1970 Baker and Coates gave an explicit upper bound for the largest solution in terms of the coefficients of the polynomials...

Anyway, let us go back to elliptic curves and start our journey through some first examples.

12.2 A Taste of Elliptic Curves

Long before the notion of an elliptic curve was born, people had already been thinking of problems which concerns rational solutions of binary cubic equations.

Example 12.3 (The congruent number problem) *In Ancient Greek, people studied the following geometric problem: which positive integer n is the area of some right triangle all of whose sides are rational numbers? In other words, does there exist $a, b, c \in \mathbb{Q}^+$, such that*

$$a^2 + b^2 = c^2, \quad \frac{1}{2}ab = n \tag{12-1}$$

*holds? If so, we say that n is a **congruent number**. For example, from the triple $(3, 4, 5)$ we know that 6 is a congruent number. For a non-trivial example, in 1225 Fibonacci found that 5 is a congruent number, since we can take*

$$(a, b, c) = \left(\frac{3}{2}, \frac{20}{3}, \frac{41}{6}\right).$$

Also, Fermat prove that $1, 2, 3$ are not congruent numbers, then neither is 4 (n is a congruent number if and only if nm^2 is so, for some $m \geq 1$. Hence, we assme that n is square-free).

In 10th century, Arabs considered the following number theory problem, which turns out to be equivalent to the previous one:

[1] This seemingly innocuous result actually turned out to be an important step towards the proof of Fermat's Last theorem.
[2] Actually, Mordell proved the case when $K = \mathbb{Q}$ and André Weil generalized this result in his doctorial dissertation.

"*Given a positive integer n, does there exist a rational number x, such that*

$$\sqrt{x}, \sqrt{x+n}, \sqrt{x-n}$$

are all rational numbers?"

Indeed, if (a, b, c) is a triple of positive rational numbers such that (12-1) holds, then one can take $x = c^2/4$. Conversely, if x has the property above, then $a = \sqrt{x+n} + \sqrt{x-n}, b = \sqrt{x+n} - \sqrt{x-n}, c = 2\sqrt{x}$ is such that (12-1) holds.

One can go further. If $x = D$ is a solution to the second problem, then $D, D \pm n$ are perfect squares of some positive rational numbers, say A, B and C. Then, the cubic curve $E_n : y^2 = x^3 - n^2x$ has a rational point $(x, y) = (D, ABC)$ where $ABC \neq 0$. Conversely, if $(x, y) = (M, N)$ is a rational point on E_n with $N \neq 0$, then one can check directly that

$$\left(\frac{M^2+n^2}{2N}\right)^2, \quad \text{and} \quad \left(\frac{M^2+n^2}{2N}\right)^2 \pm n$$

are all squares of rational numbers. i.e., n is a congruent number if and only if

*The cubic curve E_n has **a** rational point (x, y) with $y \neq 0$.*

In fact, E_n is an example of an elliptic curve, and in turns out that if we have such a rational point, we can find infinitely many ones on E_n, using the addition operation on E_n. i.e., n is a congruent number if and only if

*The cubic curve E_n has **infinitely many** rational points.*

Now, let us give a formal definition of an elliptic curve and discover the group structure on it. Unless otherwise stated, K is an (arbitrary) field. We write $\mathbb{A}^n_K = \{(a_1, \ldots, a_n) \mid a_i \in K, \forall\, i\}$ to be the *affine n-space* over K. $\mathbb{A}^n_K$ is nothing but K^n, except that we do not care much about its vector space structure. We often omit K if it is understood.

Definition 12.4 ((affine) elliptic curve) *An (affine) elliptic curve $E \subset \mathbb{A}^2_K$ over K is a smooth curve of an equation of the form (the **reduced Weierstrass form**[1])*

$$\boxed{y^2 = x^3 + ax + b} \tag{12-2}$$

for some $a, b \in K$.

Note that we also assume that the curve $y^2 = x^3 + ax + b$ is smooth, which means that if we set $F(x, y) = y^2 - x^3 - ax - b$, then the gradient of F never vanishes. A direct computation gives that

$$\boxed{\Delta := 4a^3 + 27b^2 \neq 0.}$$

Geometrically, F has no singular points. See the following examples[2]:

Example 12.5 *(a) $E = E_1 : y^2 = x^3 - x$ is an elliptic curve. More generally, for every $n \in \mathbb{N}^+$, E_n is an elliptic curve.*

(b) $C : y^2 = x^3 - x^2$ is NOT an elliptic curve since it has a singular point at the origin.

[1] In general, an elliptic curve is defined as $y^2 = a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6$ for some constants $a_1, a_3, a_4, a_6 \in K$, but we shall not bother to deal with these technical details.

[2] The following two figures are bollowed from Gathmann's notes *Plane Algebraic Curves*.

Figure 12.2.1 Elliptic Curve and Nodal Curve

It is said that the discovery of a geometric way of finding a group structure on $E(K)$ may owe to Issac Newton [4]: If we have already known two K-points $P = (x_1, y_1)$ and $Q = (x_2, y_2)$ on a given elliptic curve E (not necessarily different), then the line l passing through P and Q (or the tangent line at P, if $Q = P$) will intersect with E at another K-point $P * Q$ (by Vieta's theorem):[1]

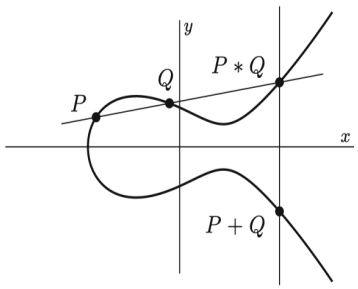

Figure 12.2.2 The Group Structure

But wait, what if l does not? Well, this could happen only when l is parallel to the y-axis. In this case, we define $P * Q$ to be the point $\mathcal{O}$ "at infinity". Since E is symmetric about the x-axis, so we define $P + Q$ as the reflection of $P * Q$ about the x-axis, denoted $\overline{P * Q}$.

You may not be convinced why this "point at infinity" works, but we also have a perfect explanation using projective geometry. Recall that the projective n-space $K\mathbb{P}^n$ over K is defined as the quotient space of $\mathbb{A}_K^{n+1} \setminus \{0\} = \{a_0, \ldots, a_n\}/\sim$, where

$$(a_0, \ldots, a_n) \sim (b_0, \ldots, b_n) \Leftrightarrow \exists\, \lambda \in K^* = K \setminus \{0\},\ \text{s.t. } b_i = \lambda a_i, \forall\, i.$$

Again, we often omit K and simply write $\mathbb{P}^n$ if (lazy) it is understood.

By convention, the equivalence class of $(a_0, \ldots, a_n)$ is denoted $[a_0 : \cdots : a_n]$, the so-called *homogeneous coordinate*. Also, we shall call the points in $K\mathbb{P}^n$ with $a_n = 0$ *infinite points*, and other points *finite points*. Note that there is an one-to-one correspondence between finite points and points in $\mathbb{A}^n$:

$$[a_0 : \cdots : a_n] = [a_0/a_n : \cdots : a_{n-1}/a_n : 1] \leftrightarrow (a_0/a_n, \ldots, a_{n-1}/a_n),$$

we have a natural embedding $i : \mathbb{A}^n \hookrightarrow \mathbb{P}^n$.

The only polynomials living on projective spaces are homogeneous, so in order to embed an affine elliptic curve $\mathbb{A}^2$ into $\mathbb{P}^2$, we need an operation – *homogenization*: If $f(x, y) \in k[x, y]$

[1] The figure below is borrowed from Silverman & Tate [2].

is a polynomial of degree d, then we put

$$\hat{f}(x,y,z) := z^d f(x/z, y/z) \in k[x,y,z],$$

which is indeed a homogeneous polynomial of degree d. Conversely, for any homogeneous polynomial $\hat{g} \in k[x,y,z]$, we can *dehomogenize* it by evaluating at $z = 1$. To sum it up, from any affine elliptic curve E in the reduced Weierstrass form (12-2), we can homogenize it to get a **projective elliptic curve**

$$\boxed{y^2z = x^3 + axz^2 + bz^3.} \tag{12-3}$$

and conversely, we can dehomogenize it to get the affine elliptic curve back. Therefore, we need not bother to distinguish these two. The point is: if $z = 0$, then $x = 0$ and thus $y = 1$, i.e., an (projective) elliptic curve has exactly one infinite point $\mathcal{O} = [0 : 1 : 0]$.

Now, we are ready for the group law on $E(K)$.

Theorem 12.6 *For any elliptic curve $E : y^2z = x^3 + axz^2 + bz^3$, the set of its K-points (including $\mathcal{O}$) $E(K)$ forms an abelian group $(E(K), +)$.*

Proof: (Sketch) It is not hard to see that for any $P \in E(K)$, $P + \mathcal{O} = P$ and $P + \overline{P} = \mathcal{O}$, i.e., $\mathcal{O}$ is the identity element and each element P has an inverse $\overline{P}$. The non-trivial part is to show the associativity of "+", i.e., for any $P, Q, R \in E(K)$,

$$(P + Q) + R = P + (Q + R).$$

As far as I know, there are at least five ways to solve this problem. we would adopt (e) here, and others are listed for interested readers:

(a) ~~Direct computation~~ (annoying painful). By Vieta's theorem one can give an explict formula for the addition law (see Silverman & Tate [2] §1.4).

(b) Geometric approach. The proof mainly relies on two important facts:
(1) (**Bézout's theorem**) Assume C_1, C_2 are projective curves with no common components, then over the algebraic closure of K, we have

$$|C_1 \cap C_2| = \deg(C_1)\deg(C_2).$$

(2) (**Cayley-Bacharach theorem**) If cubic projective curves C_1 and C_2 intersect in 9 different points, then any cubic curve pass through 8 of them must pass through the last one.

Provided these two facts, the figure below[1] shows that $P * (Q + R) = (P + Q) * R$ since the three vertial lines and the three horizontal lines are cubics passing through eight points of E, and both intersect E with exactly 9 points.

(c) Algebraic approach. The point is: there is a natural abelian group Pic^0_E of the given elliptic curve E, and a bijection $E(K) \to \mathrm{Pic}^0_E$, so we can pullback the group structure on Pic^0_E to $E(K)$, and this is precisely $(E(K), +)$.

(d) Complex approach. The point is: any elliptic curve over $\mathbb{C}$ is biholomorphic via the Weierstrass map to a complex torus $\mathbb{T}$. But $\mathbb{T}$ inherits a group structure from $\mathbb{C}$ via the quotient map, and one can show that the pushforward of the group structure on a torus via the Weierstrass map is again exactly the one we just defined.

[1] This figure is borrowed from professor Joachim Rosenthal.

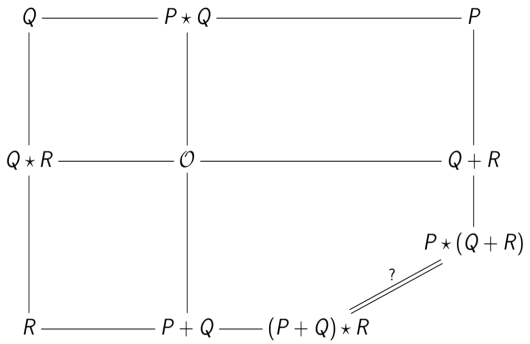

Figure 12.2.3 Proving the Associativity of "+"

(e) ~~Believe this is true.~~

□

As I mentioned in the first section, a crucial fact is that $E(K)$ is finitely generated (theorem 12.2). In other words,

$$E(K) \cong E(K)_t \oplus E(K)_f,$$

where $E(K)_t$ is the finite group consisting of all points of finite order, called the **torsion group** of E, and $E(K)_f$ is free of rank $r = r(E)$, namely $E(K)_f \cong \mathbb{Z}^r$. We say that r is the rank of the elliptic curve. Again, we shall admit this fact since it is rather difficult.

Example 12.7 (Torsion group of $E_n(\mathbb{Q})$) *For the family of elliptic curves $E_n : y^2 - x^3 + n^2 x = 0$, $n \in \mathbb{N}^+$, the torsion group*

$$E_n(\mathbb{Q})_t = \{(0,0), (n,0), (-n,0), \mathcal{O}\} \cong \mathbb{Z}/2\mathbb{Z} \oplus \mathbb{Z}/2\mathbb{Z}.$$

Hence, if we find one rational point with $y \neq 0$, then we can find infinitely many ones.

Proof: (Sketch[1]) The key idea is that we have a **reduction mod p map**: for any elliptic curve $E : y^2 = x^3 + ax + b$ with $a, b \in \mathbb{Z}$, consider

$$E_p : y^2 = x^3 + \overline{a}x + \overline{b},$$

where $\bar{\cdot}$ is the image of $\cdot$ under the map $\mathbb{Z} \to \mathbb{Z}/p\mathbb{Z}$. This allows us to reduce the given elliptic curve over $\mathbb{Q}$ to one over the finite field $\mathbb{F}_p$, while it is much easier to compute the torsion group of the latter.

Note that for E_p to be an elliptic curve, we must check that $\overline{\Delta} = 4\overline{a}^3 + 27\overline{b}^2 \neq \overline{0}$, i.e., $p \nmid \Delta$. Also, $p \nmid 2$ since otherwise one cannot write an elliptic curve in the reduced Weierstrass form (always singular). We say E has **good reduction at** p if $p \nmid 2\Delta$, and otherwise we say E has **bad reduction at** p. When we talk about reduction mod p, we shall generally assume that we have good reduction at p.

It turns out that the reduction mod p map induces a homomorphism of groups

$$E(\mathbb{Q})_t \to E_p(\mathbb{F}_p)$$

given by $P \mapsto \overline{P}$. This map is injective[2] for most p, so the order of $E(\mathbb{Q})_t$ divides the order of

[1] For a detailed proof, see for instance Koblitz [4] §1.9.

[2] Indeed, if we consider the subgroup $E[n] = \{P \in E(K) \mid nP = 0\}$ of $E(\mathbb{Q})_t$, one can show that this group homomorphism is an isomorphism whenever $(n, p) = 1$.

$E_p(\mathbb{F}_p)$ for such p. This gives us an efficient way to find $E(\mathbb{Q})_t$.

□

Let me end this section by giving some further historical remarks:

(a) If we are working over an algebraically closed field K, then whenever $\chi(K) = p = 0$ or $p \nmid n$, the subgroup of torsion points $E[n]$ of order at most n of an elliptic cuvre E is

$$E[n] \cong \mathbb{Z}/n\mathbb{Z} \oplus \mathbb{Z}/n\mathbb{Z}.$$

Moreover, there is a *Weil pairing* $w : E[n] \times E[n] \mapsto \mu_n$ (a "well-behaved" bilinear form on $E[n]$ taking value in the n-th roots of unity μ_n), which may lead one to deeper areas in number theory and cryptography.

(b) There are several natural questions concerning the group structure of $E(K)$ over a fixed number field K:

(1) (Torsion group) Are there only finitely many possible structures of $E(K)_t$?

(2) (Rank) For any $N \in \mathbb{N}^+$, does there exist an elliptic curve over K with rank greater than N.

(3) (Rank distribution conjecture) The "probability"[1] of $r(E) = 0$ and $r(E) = 1$ are both 50%, and that of $r(E) \geq 2$ is 0%.

In 1978, Barry Mazur confirmed (1) for the case $K = \mathbb{Q}$. $E(\mathbb{Q})_t$ can and only can be one of the following groups:

$$\mathbb{Z}_n, \quad n = 1, \ldots, 9, 10, 12 \quad \text{or } \mathbb{Z}_{2n} \oplus \mathbb{Z}_2, \quad n = 1, 2, 3, 4.$$

Later in 1997, (1) was completely solved for number fields [1]; For (2), Tate-Shafarevich proved in 1967 that the unboundedness of the rank over functions fields. The case of number fields is still unknown: the current record over $\mathbb{Q}$ is kept by Elkies, he found an elliptic curve with at least 28 independent $\mathbb{Q}$-points in 2006 [5]; As for (3), it is not even known whether the average rank has an upper bound until the recent work of Bhargava-Shankar in 2010. They showed that the "limsup" of the average rank is bounded, and if the average of rank exists, it is less than 0.99 [5]. Also, the record now is 0.885, also by Bhargava-Shankar [6].

12.3 When Elliptic Curves Meet L-Functions

Miraculously, the rank of an elliptic curve E, which we do not understand well, is (we guess) related to the analytic properties of its L-functions $L(E, s)$.

To motivate this, let us first assign a zeta function on a projective algebraic curve C over a finite field $\mathbb{F}_q$: For any positive integer n, let $N_n(C)$ denote all the $\mathbb{F}_{q^n}$-points of C, we define

$$\zeta(C, s) := \exp\left(\sum_{n \geq 1} \frac{N_n(C)}{n} T^n\right), \tag{12-4}$$

where $T = q^{-s}$, $s \in \mathbb{C}$. Two of the famous Weil conjectures (baby version)[2] (proved by

[1] Here, the "probability" stands for the average in the sense $\lim_{X\to\infty}(\sum_{E\in\mathscr{E}_{<X}} \chi_P)/(\sum_{E\in\mathscr{E}_{<X}} 1)$, where $\mathscr{E}$ is the set of all elliptic curves with integral coefficients, and $\mathscr{E}_{<X}$ is its subset consisting of all such elliptic curves of height ($H = \max\{|4A^3|, 27B^2\}$) less than X, and χ_P is the characteristic function of a given property, say having rank 1.

[2] Indeed, Weil came up with four conjectures (including functional equation, compatibility with complex topology) for n-dimensional algebraic varieties in general.

Grothendieck, Deligne, etc.) state that[1]:

(a) (rationality) $\zeta(C,s)$ is a rational function of T. More precisely, $\zeta(C,s)$ can be written of the form

$$\zeta(C,s) = \frac{P_1(T)}{P_0(T)P_2(T)}, \tag{12-5}$$

where $P_0(T) = 1-T$, $P_2(T) = 1-qT$, and $P_1(T) = \prod_j (1-a_jT)$ over $\mathbb{C}$ for some a_j's.

(b) (Riemann hypothesis) All the zeros of $P_1(T)$ satisfy $|a_j| = q^{1/2}$.

Let us compute a concrete example first.

Example 12.8 (Zeta function of a projective line) *If C is a projective line L, then $N_n(L) = q^n + 1$ (one may think of L as an affine line first and then add a point at infinity on it), so*

$$\begin{aligned} Z(L,s) &= \exp\left(\sum_{n\geq 1} \frac{q^n+1}{n} T^n\right) \\ &= \exp\left(\sum_{n\geq 1} \frac{(qT)^n}{n} + \sum_{n\geq 1} \frac{T^n}{n}\right) \\ &= \exp\left(-\log(1-qT) - \log(1-T)\right) \\ &= \frac{1}{(1-T)(1-qT)}. \end{aligned}$$

This is exactly the zeta function of the function field $\mathbb{F}_q(x)$ (rational functions on L), and this leads one to another rather interesting story - arithmetic of function fields!

Now back to our case, taking reduction mod p of an elliptic curve $E: y^2 = x^3 + ax + b$ ($a, b \in \mathbb{Z}$), we obtain an elliptic curve E_p over $\mathbb{F}_p$, and then we can consider its zeta function $\zeta(E_p, T)$. Thanks to Weil, it is given by

$$\zeta(E_p,T) = \exp\left(\sum_{m\geq 1} \frac{|E_p(\mathbb{F}_p)|}{m} T^m\right) = \frac{P_1(T)}{P_0(T)P_2(T)} = \frac{1-a_pT+pT^2}{(1-T)(1-pT)},$$

where $T = p^{-s}$ and $a_p = p + 1 - |E_p(\mathbb{F}_p)|$, $|a_p| \leq 2\sqrt{p}$ since $P_1(T) = (1-\omega_pT)(1-\overline{\omega}_pT)$, $|\omega_p| = \sqrt{q}$. Analogous to the Euler factors of the Riemann's zeta function, we define the **local L-factor** of E to be

$$L_p(E,s) := \frac{1}{P_1(T)} = \frac{1}{P_1(p^{-s})} = \frac{1}{1-a_pp^{-s}+p^{1-2s}}. \tag{12-6}$$

As you may guess, the **L-function of E** is defined to be the product of all local L-factors[2]:

$$L(E,s) := \prod_{p\nmid 2\Delta} L_p(E,s) = \prod_{p\nmid 2\Delta} \frac{1}{1-a_pp^{-s}+p^{1-2s}}. \tag{12-7}$$

[1] It is said that Weil was somewhat motivated by Guass's *Disquisitiones Arithmeticae*. Often, Gauss's results are concrete but surprisingly deeper than they seem to be at first glance, huh?

[2] It is sometimes called the *incomplete L-function of E* since we are omitting "bad" primes. But this is not a big deal since there are only finitely many "bad" primes.

Since $|a_p| \leq 2\sqrt{p}$, one can show that $L(E, s)$ converges absolutely for $\mathrm{Re}(s) > 3/2$. (For proof, see for instance Hüsemoller [7] chapter §16.2.)

Reference

[1] Feng, Keqin. *A Brief History of Algebraic Number Theory*. Harbin Institute of Technology Press, 2002.

[2] Silverman, Joseph H., and John Torrence Tate. *Rational points on elliptic curves*. Vol. 9. New York: Springer-Verlag, 1992.

[3] Stewart, Ian, and David Tall. *Algebraic number theory and Fermat's last theorem*, 4th ed. AK Peters/CRC Press, 2001.

[4] Koblitz, Neal I. *Introduction to elliptic curves and modular forms*. Vol. 97. Springer Science & Business Media, 2012.

[5] Li, Chao. *What is the Birch and Swinnerton-Dyer conjecture?* (expository notes online) http://www.math.columbia.edu/ chaoli/docs/BSD.html

[6] Bhargava, Manjul, and Arul Shankar. *The average size of the 5-Selmer group of elliptic curves is 6, and the average rank is less than 1.* arXiv preprint arXiv:1312.7859 (2013).

[7] Husemöller, Dale. *Elliptic Curves*, 2nd ed. Vol 111. Springer Springer-Verlag New York, Inc, 2004.