Proving Mordell-Weil: A Descent in Three Parts

A Senior Thesis Of
Danielle Li
li15@fas.harvard.edu (617) 493-3107

Thesis Advisor: William A. Stein

Submitted to
The Department of Mathematics
in partial fulfillment of the requirements
for the degree of Bachelor of Arts
in the subject of

Mathematics and History and Science

Harvard University
Cambridge, Massachusetts
4 April 2005

Acknowledgements

First and foremost, I am grateful to Professor William Stein "for his guidance and support." Read: thank you for putting up with me, for devoting your full energy into this project even when I did not, and for being a rare professor who is generous with both his extensive knowledge and his limited time.

Many thanks to David Spanagel for advising my work and making me feel welcome in the History of Science department even when I had strayed. Thanks to Barry Mazur and Federica La Nave; I learned a great deal from our conversations. I would have written this on a napkin and submitted in June were it not for Allie Belser and Sveltana Alpert who kept after me when I was late for every deadline.

Thank you to Josh Reyes for helping with translations and listening to me talk about number theory even though you are a mathematical physicist at heart (I would not do the same for you). Thank you to Jennifer Balakrishnan for reading my drafts and keeping me happy with a steady supply of yellow books, to Jonathan Leong for being my study buddy, to Alex Turnbull for making sure I stayed alert, to Amittai Axelrod for catching all those dangling participles, and to Jennifer Sinnott for never once having asked me: "So how's your thesis going?"

Contents

Introductions

In 1659, Pierre de Fermat wrote to Christiaan Huygens claiming to have discovered a "completely unique" method for solving number theoretic problems. Although poorly described in his writings, Fermat used this method, which he called *infinite descent*, to provide a scant, but most likely correct, proof of:

PROPOSITION 0.0.1. *There is no integral right triangle whose area is a square.*

and a most likely incorrect proof of:

THEOREM 0.0.2. *It is impossible for a cube to be written as the sum of two cubes or a fourth power to be written as the sum of two fourth powers or, in general, for any number which is a power greater than the second to be written as a sum of two like powers.*

In the intervening 350 years, the concept of descent, or at least the word, has come up in a variety of contexts. In particular it has, at several points in time, been linked to the problem of finding rational points on elliptic curves. In this thesis, we examine three of these moments:

(1) Pierre de Fermat's proof of Proposition 0.0.1 using infinite descent in the 1650s;
(2) Mordell and Weil's use of descent in proving the finite generatedness of $E(K)$ from the finiteness of $E(K)/mE(K)$ in the 1920s;
(3) Contemporary descent methods used to compute the group of rational points on $E(K)/mE(K)$.

Specifically, it is conjectured that the method of descent in (3) will always be able to tell us the generators and rank of the free group associated to $E(K)$, the group of rational points on an elliptic curve E. Since the torsion subgroup of $E(K)$ has already been shown to be computable, this would significantly further our knowledge of rational points on elliptic curves.

The organization of this thesis is motivated by the various "descents" which appear in our discussions. Instead of starting chronologically with Fermat's infinite descent, we begin with contemporary questions of computation and move backwards toward Fermat's initial insight. Chapter 1 motivates the study of the Mordell-Weil theorem from the perspective of the open problems it created. Specifically, we work toward explaining the descent methods used to compute the Weak-Mordell Weil group, $E(K)/mE(K)$. To do this, we define the Selmer and Tate-Shafarevich groups and interpret their elements as related to covering curves of E. We show that the weak Mordell-Weil, Selmer, and Tate-Shafarevich groups form a short exact sequence and that the Selmer group is computable but the other two not necessarily. Through this approach, the weak Mordell-Weil theorem, that $E(K)/mE(K)$ is finite, follows as a corollary.

Chapter 2 concludes the proof of the Mordell-Weil theorem with Mordell and Weil's descent procedure. We begin by introducing the notion of a height function on $E(K)$ and then proving some of its properties. At the end of the chapter, we use the height to do descent.

Chapter 3 concerns Fermat's method of infinite descent. We reconstruct Fermat's proof of Proposition 0.0.1 through his correspondence and posthumous publications.

The final chapter investigates the extent to which we can view these three descents as part of a single tradition. Finally, the appendix gives background on group cohomology. Throughout, we assume an undergraduate background in algebraic number theory up through the unit theorem and finiteness of the class group, as well as some familiarity with algebraic geometry, valuations, and ramification.

However... first things first:

1. Math

An elliptic curve is a smooth projective genus 1 curve defined over a field K with at least one K-rational point, or, equivalently:[1]

DEFINITION 0.1.1. An *elliptic curve* E is the variety of points defined by the *Weierstrass equation*:

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

with coefficients $a_i \in K$ along with an extra point, $[0:1:0]$, at infinity. If $\mathrm{char}(K) \neq 2, 3$ we can further simplify the Weierstrass equation to the form

$$y^2 = x^3 + ax + b \tag{1}$$

through an appropriate change of variables. As we are working with number fields, we will write an elliptic curve in the form (1) from now on. We denote by $E(K)$ the set of K-rational points on $E(\bar{K})$, which we write simply as E. In general, we use this notation for all curves C.

Above, we have situated the elliptic curve in the affine plane, which is a slight abuse of notation, but nonetheless acceptable because an affine curve has a unique and well-defined projective closure.

1.1. The Group Law.

The Mordell-Weil theorem concerns the group of rational points on E. In order for this to make sense, we must first attach a group structure to E. A suitable composition law can be defined geometrically as follows:[2]

For any two rational points $P, Q \in E(K)$, take the intersection of the chord connecting P and Q with the curve E. This yields another rational point $P * Q \in E(K)$. This procedure is known as the *chord-tangent process*. However, if we just take $P * Q$ to be the composition of P and Q, the resulting law we get does not have an identity element. In order to fix this, we pick a rational base point $\mathcal{O}$ and define $P + Q$ to be the intersection of the chord connecting $P * Q$ and $\mathcal{O}$ with E. If $P = Q$, we use the tangent line at P, instead of the chord connecting P and Q, and proceed in the same manner.

In general, this construction does not rely on the specific mapping of the elliptic curve into the projective plane. That is, the resulting group does not depend on our initial choice of basepoint; the map $P \mapsto P + (\mathcal{O} - \mathcal{O}')$ is easily verified as an isomorphism between $(E, \mathcal{O})$ and $(E, \mathcal{O}')$.

We may also write down an explicit formula for the sum of P and Q. Suppose that we have an elliptic curve E in Weierstrass form with $P = (x_1, y_1)$ and $Q = (x_2, y_2)$. We wish to find the

[1] The equivalence follows from the Riemann-Roch Theorem. See [**15**] Chap. 4, Sec. 1.

[2] For details, see [**37**] Chap. 2.

coordinates for $P * Q = (x_3, y_3)$ because this gives $P + Q = (x_3, -y_3)$. The line joining P and Q is

$$y = \lambda x + \nu, \qquad \lambda = \frac{y_2 - y_1}{x_2 - x_1}, \qquad \nu = y_1 - \lambda x_1 = y_2 - \lambda x_2.$$

To find its intersection with E, we substitute $\lambda x + \nu$ for y in $y^2 = x^3 + ax + b$. The three roots of this cubic are now (x_1, y_1), (x_2, y_2), (x_3, y_3) so that:

$$x^3 - \lambda^2 x^2 + (a - 2\lambda\nu)x + (b - \nu^2) = (x - x_1)(x - x_2)(x - x_3)$$
$$= x^3 - (x_1 + x_2 + x_3)x^2 + (x_1x_2 + x_1x_3 + x_2x_3)x + x_1x_2x_3.$$

Equating the coefficients of the x^2 terms, we obtain:

$$x_3 = \lambda^2 - x_1 - x_2, \qquad y_3 = \lambda x_3 + \nu.$$

This provides a formula for the composition of two points. Moreover, depending on the curve, iteration of this formula can yield an infinite number of distinct points on $E(K)$.

1.2. Maps of curves. Now that we have described the structure of a single elliptic curve, we turn to maps between them. In order to define a map between elliptic curves, which are projective varieties, we first define maps on affine varieties and then define a map of projective varieties to be one obtained by glueing together such maps of affine varieties. As such, we begin with a discussion of maps on affine varieties.

Definition 0.1.2. Given an affine variety V over K, we say that the *coordinate ring of* V is

$$K[V] \cong \frac{K[X]}{I(V/K)}.$$

We call the field of fractions of $K[V]$ the *function field of* V and denote it $K(V)$.

We can think of $K[V]$ as the ring of functions regular on all of X.[3] $K[V]$ has a K-algebra structure (it is an associative ring with the structure of an K-vector space).
Now, let $f : V \to W$ be a regular map (defined as $f = (f_1, \ldots, f_n)$) between two varieties V and W defined over K. For regular map w on W, we associate a map v on V given by $v(x) = w(f(x))$, called the *pullback of* w. As the composition of two regular maps, v will be regular as well. Thus, f induces a map of coordinate rings:

$$f^* : K[W] \to K[V]$$

via the pullback. Moreover, because f is regular, f^* is a K-algebra homomorphism of coordinate rings. In fact:

Proposition 0.1.3. *Let V and W be affine varieties defined over a field K. Then there is a natural bijective mapping of sets*

$$\mathrm{Hom}(V, W) \cong \mathrm{Hom}(K[W], K[V])$$

where the LHS are morphisms of varieties and the RHS are homomorphisms of K-algebras.

Proof. See [**15**] Chap. 1, Prop 3.5. □

Thus, affine varieties are isomorphic if and only if their coordinate rings are.

Although elliptic curves are projective varieties, Proposition 0.1.3 will be useful in Chapter 1. For now, we turn to a discussion of mappings on curves.

[3] See [**14**] Chap. 2, Lem. 2.1

DEFINITION 0.1.4. An *isogeny* is a morphism of curves that preserves the basepoint $\mathcal{O}$.

PROPOSITION 0.1.5. *Let $\phi : C \to C'$ be a nonconstant isogeny of curves. Then ϕ is surjective.*

PROOF. See [**33**] Chap. 1, Sec. 5, Thm 4. □

PROPOSITION 0.1.6. *If, in addition, ϕ is separable, and $deg(\phi) = r$, then*

$$\#(\mathrm{Ker}(\phi)) = r.$$

PROOF. By separability, $\phi^{-1}(p')$ has r elements for all but finitely many elements $p' \in C'$. However, by surjectivity, for any $p', q' \in C'$, we can find $s \in C$ such that $\phi(s) = q' - p'$. But because ϕ is also a homomorphism, this means that there is a bijective correspondence

$$\begin{aligned} \phi^{-1}(p') &\to \phi^{-1}(q') \\ t &\to s + t. \end{aligned}$$

Thus, $\#(\phi^{-1}(p'))$ is independent of our choice of p'. Since K is infinite, we must have $\#(\phi^{-1}(p')) = r$ for all $p' \in C'$. Taking p' trivial, we see that $\#(\mathrm{Ker}(\phi)) = r$. □

In particular, we define the multiplication by $m : E \to E$ map:

$$P \mapsto [m]P = \overbrace{P + \cdots + P}^{m \text{ times}} \qquad \text{for all } P \in E \text{ and } m \geq 0$$

In general, m can be any integer; if it is negative, we define $[m]P = [-m](-P)$. The map $[m]$ is an isogeny as it keeps $\mathcal{O}$ fixed. By Proposition 0.1.5, $[m]$ is surjective for all $m \neq 0$. Further:

PROPOSITION 0.1.7. *Let $\phi : E_1 \to E_2$ be a nonconstant isogeny of degree m. Then there exists a unique isogeny*

$$\phi' : E_2 \to E_1$$

such that $\phi' \circ \phi = [m]$. We call ϕ' the dual isogeny *of ϕ.*

PROOF. See [**36**] Chap 3, Thm 6.8. □

PROPOSITION 0.1.8. *Then, $E[m]$, the m-torsion of E, is finite, and in particular*

$$E[m] \cong \mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/m\mathbb{Z}.$$

PROOF. From Proposition 0.1.7, it follows that $\deg([m]) = m^2$. Since $\mathrm{char}(K) = 0$, the map $[m]$ is separable. Proposition 0.1.6 implies $\#(E[m]) = \deg([m]) = m^2$. Similarly, $E[d]$ has d^2 elements for all d dividing m and we must have $E[d] \subset E[m]$. As $E[m]$ is abelian, the classification of finite, abelian groups shows that the only possibility is $E[m] \cong \mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/m\mathbb{Z}$. □

Finally, since $E[\phi] \subset E[m]$ and has order m, $E[\phi] \cong \mathbb{Z}/m\mathbb{Z}$.

Another important map is the *reduction modulo v* map. Let R be the ring of integers of the field K. For v a nonarchimedean valuation, denote by R_v the valuation ring associated to v. Given the Weierstrass equation $y^2 = x^3 + ax + b$, we can change coordinates so that $a, b \in R_v$. This means that the valuation function will take nonnegative values on the discriminant $\Delta = -16(4a^3 + 27b^2)$ so $v(\Delta)$ is bounded below by 0. Since v is discrete, there are coefficients a and b which we can get from a change of coordinates, for which the valuation of the discriminant is minimal. The resulting form of E is known as a *minimal Weierstrass equation*, which is, in

general not unique.[4]

Under these circumstances, we have a natural reduction map

$$\sim: \mathbb{P}^n(K) \to \mathbb{P}^n(k)$$

where k is the residue field of K and where the map sends an element to its reduction modulo v. This can be restricted to give a map

$$\begin{array}{rcl} E(K) & \to & \tilde{E}(k) \\ y^2 = x^3 + ax + b & \mapsto & y^2 = x^3 + \tilde{a}x + \tilde{b} \qquad \tilde{a}, \tilde{b} \in k. \end{array}$$

Since the reduced curve $\tilde{E}(k)$ is not necessarily nonsingular at v, we make the following definition:

DEFINITION 0.1.9. Let E/K be an elliptic curve and let $\tilde{E}(k)$ be the reduced curve for a minimal Weierstrass equation with respect to v. Then we say that E has *good reduction* modulo v if $\tilde{E}(k)$ is nonsingular. Otherwise, we say that it has *bad reduction*.

Because a curve $y^2 = x^3 + \tilde{a}x + \tilde{b}$ will have a multiple root only if $p \mid \Delta$, it follows that there can only be a finite number of places, v, at which E has a bad reduction.

In general, for every reduction, we denote the group of nonsingular points of $\tilde{E}(k)$ by $\tilde{E}_{ns}(k)$. We also define several subgroups of $E(K)$:

$$E_0(K) = \{P \in E(K) : \tilde{P} \in \tilde{E}_{ns}(k)\}$$

$E_0(K)$ is the set of elements of $E(K)$ that are mapped to nonsingular points via reduction.

$$E_1(K) = \{P \in E(K) : \tilde{P} = \tilde{\mathcal{O}}\}$$

$E_1(K)$ is the kernel of the reduction.

The following result will be useful in the proof of the weak Mordell-Weil theorem:

PROPOSITION 0.1.10. *Let K be a complete field with respect to a discrete normalized valuation v, where $v(\pi) = 1$ if π is a uniformizer. Then, the sequence*

$$0 \to E_1(K) \to E_0(K) \to \tilde{E}_{ns}(k) \to 0$$

is exact.

REMARK 0.1.11. So that this proposition makes sense, note that $\tilde{E}_{ns}(k)$ inherits the group law from $E(K)$ because the reduction map sends lines to lines.

PROOF. Exactness at the left and middle follows from the definition of E_1. Now to show exactness of at the right.

By definition, $E_0(K)$ is the preimage of $\tilde{E_{ns}}(k)$. Note that this only implies the map is well-defined. To show surjectivity, let $f(x, y)$ be our minimal Weierstrass equation for E, and $\tilde{f}(x, y)$ be its image under the projection map. Take $\tilde{P} = (a, b)$ to be any element of $\tilde{E}_{ns}(k)$. Because $\tilde{f}$ is nonsingular, at least one partial derivative of $\tilde{f}(x, y)$ will be nonzero at $\tilde{P}$. We will prove this for the case that $\tilde{f}_x(\tilde{P}) \neq 0$; the other case is completely symmetric.

First, pick any element $y_0 \in R$ that reduces to b, the y-coordinate of $\tilde{P}$. Fixing, y_0, we turn $f(x, y) = 0$ into an equation in x only, $f(x, y_0)$. Now, a is a simple root of the reduction of f because $\tilde{f}_x(a, y_0) \neq 0$, by assumption. Therefore we can apply Hensel's Lemma to $\tilde{f}(x, y_0)$ with

[4]For example, the curve $y^2 = x^3 - x$ has minimal discriminant $\Delta = 64$. However, the isomorphic curve $y^2 = (x+1)^3 - (x+1) = x^3 + 3x^2 + 2x$ obtained via the transformation $x \mapsto x + 1$ also has discriminant $\Delta = 64$. In general, a infinite number of transformations exist that preserve the determinant of the given curve. For details, see [**36**], page 49.

a as a root. Hensel's Lemma says that a can be lifted to an element, x_0, of R such that $\tilde{x}_0 = a$ and $f(x_0, y_0) = 0$. This means that (x_0, y_0) is a point on the elliptic curve which maps to $\tilde{P}$ under reduction and thus the map is surjective. □

2. History

THEOREM 0.2.1 (Finite-Basis). *All the rational points of an elliptic curve over $\mathbb{Q}$ can be obtained via finite iteration of the modified chord-tangent process on a finite generating set in $E(\mathbb{Q})$.*

This theorem was conjectured by Poincaré in his 1901 writings on the chord-tangent process, proven by Mordell in 1922, and generalized to abelian varieties over arbitrary number fields by Weil in 1928.[5] We will prove the following version of the theorem:

THEOREM 0.2.2 (Mordell-Weil). *Let K be any algebraic number field or function field and let E be any elliptic defined over K. Then the group of K-rational points $E(K)$ is finitely generated.*

However, we first consider some historical background.

2.1. The arithmetic structure of elliptic curves.

The preface to Mordell's proof reads:

> *Mathematicians have been familiar with very few questions for so long a period with so little accomplished in the way of general results, as that of finding the rational* [points on nonsingular cubic curves].[6]

In making this statement, however, Mordell elides considerable changes in the way this question was approached. Indeed, while the problem of finding rational points on polynomials predates Diophantus, the problem of understanding the group structure on $E(\mathbb{Q})$ as a way of approaching Diophantine problems was no more than 40 years old at the time Mordell published his proof of the Finite Basis theorem.

Specifically, the study of $E(\mathbb{Q})$ demands the existence of a particularly cohesive way of thinking about points on E. At a minimum, the group law we defined in Section 1 combines knowledge of geometric and arithmetic properties of E to endow $E(\mathbb{Q})$ with an algebraic structure. The formal tools for this approach - namely the coordinate geometry and group theory - were not available until the early 19th century and even then, it took some time to develop a group structure on $E(\mathbb{Q})$.

In contrast, a variation of the chord-tangent process was known, at least algebraically, in antiquity, by Diophantus, who used it systematically in his *Arithmetica*. This method, sometimes referred to as *ascent*, was widely used by Fermat and his contemporaries, who considered it one of the traditional tools in the search of solutions to polynomial equations.[7] Another such tool was Bachet's duplication formula for points on

$$y^2 - x^3 = c \qquad c \in \mathbb{Z}$$

[5]The original papers are, respectively, [**26**], [**25**], and [**43**].

[6]See [**25**] pg. 179.

[7]See the discussion in [**42**] pg. 104-112, and [**7**] pg. 31.

which can be applied recursively to yield an infinitude of distinct rational solutions in most cases.[8] These methods were stated purely algebraically; according to Weil, a geometric description of the chord-tangent process first appears a few decades later in one of Newton's manuscripts.[9]

Although the chord-tangent process hints at the possibility of a more structured composition law on $E(\mathbb{Q})$, the use of a group law is not suggested until a century later, with Jacobi's work on Diophantine analysis, which marked the beginnings of a program to bring more theoretical structure to the study of Diophantine equations.[10] Proponents of this reform, in particular Hurwitz, Hilbert, Poincaré, and Levi, felt that traditional approaches to solving Diophantine equations relied too heavily on tricks adapted to specific equations and lacked a general theory.[11]

Whereas the study of Diophantine equations was very curve specific, the emerging field of algebraic geometry sought to study algebraic curves and varieties over closed fields using more general tools. The program hinted at by Jacobi, and developed more explicitly by Hilbert and Hurwitz, was to situate the study of Diophantine equations in the tradition of algebraic geometry by applying its language and tools to the study of curves over arbitrary fields such as $\mathbb{Q}$ and its extensions. Particularly, the idea of a *birational equivalence* between two curves is important. If C and C' are birationally equivalent (we will just say isomorphic), then the rational points of C get mapped to the rational points of C'. Thus, the process of finding rational points on one curve is essentially the same as the problem of finding rational points on a birationally equivalent curve (this is part of the reason we find it so natural to study an elliptic curve in Weierstrass form). One characteristic of a curve invariant under birational equivalence is its genus; from this perspective, the problem of finding solutions to specific Diophantine equations changes into the more general problem of finding rational points on algebraic curves over $\mathbb{Q}$ of a given genus, up to birational equivalence.

In one variable, an algebraic curve defined by a polynomial equation of the form:

$$a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 = 0.$$

To find its rational or integer solutions, we simply apply Gauss's lemma, which says that if p/q is a solution in reduced form, then $p \mid a_0$ and $q \mid a_n$, to get a finite list of possible solutions to check. Algebraic curves in two variables are the first incompletely understood case.

2.2. Nonsingular projective curves in two variables. For such curves of genus 0, the Hasse principle holds so that $C(\mathbb{Q})$ is non-empty if and only if C has a $\mathbb{Q}_p$-rational point at all primes p. Finding these local points is computable because Hensel's Lemma reduces it to a matter of computation over finite rings. If a rational point does exist, then Hilbert and Hurwitz showed that C is birationally equivalent over $\mathbb{Q}$ to a line. Thus, if there exists one rational point, then we can find infinitely many points by parameterizing.[12]

For curves of genus ≥ 2, Mordell conjectured, and in 1983, Faltings proved, the following deep result:

[8]Indeed this will happen for $c \neq 1, -432$ when the original solution, (x, y) does not have $xy = 0$, although Bachet did not prove this. See [**37**], Intro.

[9][**42**] pg. 108.

[10][**17**] pg. 13.

[11][**29**] pgs 64-68.

[12]For the original paper, see [**16**]. Otherwise, see the discussion in [**44**].

THEOREM 0.2.3 (Faltings). *If the genus of C is greater than or equal to two, then $C(\mathbb{Q})$ is finite.*

Faltings' theorem, however, does not give criteria for when rational solutions exist, nor an algorithm for how to find them when they do.

Yet in the case of genus 1, we have even fewer general facts. There may or may not exist a rational solution, and if there does, there may be finite or infinitely many; currently, we have no algorithm for determining which is the case given an arbitrary curve. Thus, Mordell's theorem was a considerable breakthrough for the case of elliptic curves, a subset of such genus 1 curves. Further, unlike the proof of Faltings' theorem, the proof of the Mordell-Weil theorem does suggest a way of computing the generators of $E(K)$ using a "descent" method. However, this method relies on finding local points and, as a result, might not always yield a global rational point on $E(K)$ due to the failure of the Hasse Principle for curves of genus 1.

2.3. Examples. As a bonus, some classic number theory problems can be formulated in terms of elliptic curves.

EXAMPLE 0.2.4. We show that Proposition 0.0.1, which Fermat proved, can be written as a problem involving elliptic curves. First, write the sides, a, b, c, of a Pythagorean triangle, in primitive form: $a = 2pq$, $b = p^2 - q^2$, and $c = p^2 + q^2$ where p, q are mutually prime, $p > q$, and $p - q$ is odd. This characterization forms a general solution (up to permutation and multiplication by a constant) to the equation $a^2 + b^2 = c^2$. The area of this triangle is $A = pq(p - q)(p + q)$. Substituting $x = q/p$ we obtain the equality

$$x - x^3 = At^2 \qquad \text{where } t = 1/p^2.$$

If, further, we assume that A is a square, then this equation can be written as

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

Thus, Fermat's problem becomes showing that this curve has only the trivial rational points, $(0, 0)$, $(1, 0)$, $(-1, 0)$, and the point at infinity.

EXAMPLE 0.2.5. An even older and more difficult problem that can be phrases in terms of elliptic curves is the Congruent Number problem. This problem asks, whether, given an integer n, there exists a right triangle with rational sides that has area n.[13] This problem turns out to be equivalent to the question of whether the Mordell-Weil group associated to the curve

$$E : y^2 = x^3 - n^2x \tag{2}$$

has rank $r \geq 1$. The sketch of the argument is as follows:[14]

If n is congruent, then there must exist $X, Y, Z \in \mathbb{Q}$ such that

$$X^2 + Y^2 = Z^2 \qquad XY = 2n.$$

We can assume without loss of generality that n is a squarefree positive integer. Adding or subtracting four times the second equation from the first, we obtain $(X \pm Y)^2 = Z^2 \pm 4n$.

[13]The problem apparently first appears on a Chinese manuscript around the year 980 and was only resolved in 1989, pending verification of the Birch and Swinnerton-Dyer Conjecture. See Tunnell's proof in [**41**] and Koblitz's discussion throughout [**18**].

[14]for more details, see [**18**] Chap. 1.

Dividing through by 4 gives the two equations

$$\left(\frac{X+Y}{2}\right)^2 = \left(\frac{Z}{2}\right)^2 \qquad \left(\frac{X+Y}{2}\right)^2 = \left(\frac{Z}{2}\right)^2.$$

Multiplying these two equations together gives

$$\left(\frac{X^2-Y^2}{4}\right)^2 = \left(\frac{Z}{2}\right)^4 - n^2,$$

indicating that $u^4 - n^2 = v^4$ has a rational solution. Multiplying through by u^2 gives $u^6 - n^2u^2 = (uv)^2$. Finally, substituting $x = u^2 = (Z/2)^2$ and $y = uv = [(X^2 - Y^2)Z]/8$ yields the elliptic curve

$$E : y^2 = x^3 - n^2x.$$

However, a point on E will not always yield an elliptic curve. Since $x = (Z/2)^2$, two necessary conditions are that $x \in (\mathbb{Q}^+)^2$ and that the denominator of x be divisible by 2. Another necessary condition is that the numerator of x and n be mutually prime. If there was some p, then $p^2 \mid n$, but we assumed n was squarefree. It turns out that these conditions are sufficient as well.

Now consider the curve E given in (2). It clearly contains three roots at $(0,0), (0,\pm n)$ as well as the point at infinity, none of which satisfy all three conditions given above. It turns out that these are the only torsion points on E. Thus, any other point on E, in particular one that yields a rational right triangle with area n, must be a point of infinite order. If such a point exists, then the rank of $E(K) \geq 1$. The converse of this statement also true: if there exists a point with infinite order, we can find a rational right triangle with area n. Thus, the problem of determining whether n is a congruent number is equivalent to determining whether the curve $E : y^2 = x^3 - n^2x$ has positive rank.

Throughout this thesis, we will use the family of curves $E : y^2 = x^3 - n^2x$ to illustrate some of the problems with, and applications of, the descent methods we introduce.

CHAPTER 1

Descent 3: Finiteness and Computation of the Weak Mordell-Weil Group

The Mordell-Weil theorem tells us that the group of K-rational points on an elliptic curve $E(K)$ is finitely generated. By the structure theorem for abelian groups, this means:

$$E(K) \cong E(K)_{\text{tors}} \times \mathbb{Z}^r.$$

Thus, in order to fully characterize $E(K)$, we must compute the torsion, the rank, and then find the generators. However, finding an algorithm that does this in generality remains an open problem. Currently, the torsion subgroup is far better understood than the free group. In particular, the Nagell-Lutz theorem describes individual points in $E(\mathbb{Q})_{\text{tors}}$ and Mazur's theorem classifies the possible isomorphism classes of the torsion subgroup. Further, Nagell-Lutz provides an (inefficient) algorithm for computing $E(\mathbb{Q})_{\text{tors}}$ which has been improved by Doud [**10**]. For exposition on the torsion subgroup, see [**8**] Chap. 1 and [**28**] Sec. 2.

Thus, the primary obstacle to understanding $E(K)$ lies with computing the rank and generators of the free group since the current procedure used is not guaranteed to terminate for all curves E/K. This problem can be somewhat simplified by focusing on what is called the weak Mordell-Weil group, $E(K)/mE(K)$. If we can compute the generators of $E(K)/mE(K)$, then we can give generators for $E(K)$ (see Chapter 2, Section 3).

This chapter is dedicated to understanding the weak Mordell-Weil, Selmer, and Tate-Shafarevich groups, where the latter two are important in computing the former. Section 1 outlines the general framwork used and the remaining sections fill in gaps in the argument and provide examples.

We adopt modern language throughout this chapter and in general ignore the chronology of results. For example, in our exposition, the proof that the weak Mordell-Weil group is finite (which is crucial to the proof of the Mordell-Weil theorem) will follow from the finiteness of the Selmer group, whose study emerged much later. The rationale for this approach is that Weil's generalization of Mordell's theorem over $\mathbb{Q}$ highlights the distinction between the modern and classical approaches in the proof of Theorem 0.2.2. The first part, in which Weil proves that $E(K)/mE(K)$ is finitely generated suggests an algorithm to compute the rank and generators of $E(K)/mE(K)$. Thus, we discuss the older problem of finiteness and the contemporary problem of computation together to emphasize their methodological similarities.

Finally, it should also be noted that in some cases, for example $K = \mathbb{Q}$ and $m = 2$, everything can be described in terms of explicit computations with rational functions; this method is pursued throughout [**37**].

1. A Framework for Computation

Now we turn to the problem of computing the rank and generators of the weak Mordell-Weil group. There is an important conjecture relating the rank of E to its L-series; we give a simplified statement:

CONJECTURE 1.1.1 (Birch and Swinnerton-Dyer). *Let E/K be an elliptic curve of rank r and let $L(E/K;t)$ be its L-series. Then the function $L(E/K;t)$ has a zero of order $r' = r$ at $t = 1$.*

If this conjecture were true, then the problem of computing the rank is reduced to computing r', the order of vanishing at $s = 1$, and assuming $r = r'$. Then, to find generators, we search on E or its associated homogenous spaces until we find r independent points. However, we do not have an algorithm to compute r'. Indeed, no one has been able to prove that we can have $r' \geq 4$ even though elliptic curves of higher rank r have been found.[1] Yet even in these cases, knowing BSD still results in an algorithm for computing $E(K)$: suppose we have found s independent generators. Then, compute $L^s(E/K, 1)$. If BSD is true and $L^s(E/K, 1)$ is zero, then $s < r' = r$ and we should search longer for additional generators. If it is not zero, then we have found enough generators.

However, we will concern ourselves with the second class of methods for determining rank: the method of descent. This procedure is guaranteed to give at least an upper bound on the rank of $E(K)$ and in many cases, this bound will be exact. To formulate how it works, we begin with some cohomology.

Consider an isogeny $\phi : E \to E'$ and let $\mathrm{Gal}(\bar{K}/K) = G_{\bar{K}/K}$. To motivate the steps, we can think of $\phi = m$ and $E' = E$ even though we will often want to work with other isogenies. We have the following short exact sequence of $G_{\bar{K}/K}$-modules:

$$0 \to E[\phi] \to E \xrightarrow{\phi} E' \to 0 \tag{3}$$

where the first map injects because it is simply an inclusion and the second map surjects by Proposition 0.1.5. Making use of cohomology gives the standard long exact sequence:

$$0 \to E(K)[\phi] \to E(K) \xrightarrow{\phi} E'(K) \xrightarrow{\delta} H^1(G_{\bar{K}/K}, E[\phi])$$

$$\to H^1(G_{\bar{K}/K}, E) \xrightarrow{\phi} H^1(G_{\bar{K}/K}, E').$$

Notice that the subgroups of $E[\phi]$ and E that are fixed by all $\sigma \in G_{\bar{K}/K}$ are precisely those defined over the base field K. In the middle of this long exact sequence is the fundamental short exact sequence:

$$0 \to \frac{E'(K)}{\phi E(K)} \xrightarrow{\delta} H^1(G_{\bar{K}/K}, E[\phi]) \to H^1(G_{\bar{K}/K}, E)[\phi] \to 0.$$

To see this, notice that the image of ϕ is $\phi E(K)$ is the kernel of the δ map. Thus, the first map $\frac{E'(K)}{\phi E(K)} \xrightarrow{\delta} H^1(G_{\bar{K}/K}, E[\phi])$ is injective. The next mapping is surjective because, from the long exact sequence, the image of $H^1(G_{\bar{K}/K}, E[\phi]) \to H^1(G_{\bar{K}/K}, E)$ is the kernel of $H^1(G_{\bar{K}/K}, E) \xrightarrow{\phi} H^1(G_{\bar{K}/K}, E[\phi])$, which is just the ϕ-torsion of $H^1(G_{\bar{K}/K}, E)$.

The δ is given by the general theory of group cohomology. Namely, we take $P \in E(K)$ and find Q such that $[\phi]Q = P$. Then the map $\sigma \mapsto Q^\sigma - Q$ is a cocycle and we take $\delta(P)$ to be its cohomology class. Notice that while $Q^\sigma - Q \in E$, Q itself is not necessarily an element of E so

[1] In 2000, an elliptic curve of rank at least 24 was found. See [**23**].

that $Q^\sigma - Q$ is, in general, a nontrivial cocycle.

We also repeat the same steps, replacing K by K_v, to obtain:

$$0 \to \frac{E'(K_v)}{\phi E(K_v)} \xrightarrow{\delta} H^1(G_{\bar{K_v}/K_v}, E[\phi]) \to H^1(G_{\bar{K_v}/K_v}, E(\bar{K_v}))[\phi] \to 0$$

which we join to obtain:

$$0 \to \prod_{v \in M_K} \frac{E'(K_v)}{\phi E(K_v)} \xrightarrow{\delta} \prod_{v \in M_K} H^1(G_{\bar{K_v}/K_v}, E[\phi]) \to \prod_{v \in M_K} H^1(G_{\bar{K_v}/K_v}, E(\bar{K_v}))[\phi] \to 0. \tag{4}$$

Combining the global and local sequences, we have the following commutative diagram:

$$\begin{array}{ccccccccc}
0 & \longrightarrow & \frac{E'(K)}{\phi E(K)} & \xrightarrow{\delta} & H^1(G_{\bar{K}/K}, E[\phi]) & \longrightarrow & H^1(G_{\bar{K}/K}, E)[\phi] & \longrightarrow & 0 \\
 & & \downarrow & & \downarrow & & \downarrow & & \\
0 & \longrightarrow & \prod_{v \in M_K} \frac{E'(K_v)}{\phi E(K_v)} & \xrightarrow{\delta} & \prod_{v \in M_K} H^1(G_{\bar{K_v}/K_v}, E[\phi]) & \longrightarrow & \prod_{v \in M_K} H^1(G_{\bar{K_v}/K_v}, E(\bar{K_v}))[\phi] & \longrightarrow & 0
\end{array}$$

We do this because our goal is to investigate the image of $\frac{E(K)}{\phi E(K)}$ in $H^1(G_{\bar{K}/K}, E[\phi])$, which, by exactness, is equivalent to looking at the kernel of

$$H^1(G_{\bar{K}/K}, E[\phi]) \to H^1(G_{\bar{K}/K}, E)[\phi]. \tag{5}$$

While this problem is in general difficult, the analogous local problem of determining the kernel of

$$H^1(G_{\bar{K_v}/K_v}, E[\phi]) \to H^1(G_{\bar{K_v}/K_v}, E(\bar{K_v}))[\phi] \tag{6}$$

is simpler for reasons we will discuss in Section 4

Because we are ultimately interested in the kernel of the map (5), we define the following peripheral kernels:

Definition 1.1.2. With the isogeny ϕ as above, the *ϕ-Selmer group of E/K* is the subgroup of
$H^1(G_{\bar{K}/K}, E[\phi])$ defined by:

$$S^\phi(E/K) = \ker\Big\{H^1(G_{\bar{K}/K}, E[\phi]) \to \prod_{v \in M_K} H^1(G_{\bar{K_v}/K_v}, E(\bar{K_v}))\Big\}.$$

Definition 1.1.3. With the isogeny ϕ as above, the *Tate-Shafarevich* is the subgroup of $H^1(G_{\bar{K}/K}, E)$ defined by:

$$Ш(E/K) = \ker\Big\{H^1(G_{\bar{K}/K}, E) \to \prod_{v \in M_K} H^1(G_{\bar{K_v}/K_v}, E(\bar{K_v}))\Big\}.$$

Further, the ϕ-torsion subgroup, $Ш(E/K)[\phi]$ is the kernel of the map $H^1(G_{\bar{K}/K}, E)$ into $\prod_{v \in M_K} H^1(G_{\bar{K_v}/K_v}, E(\bar{K_v}))[\phi]$.

From these definitions and the commutative diagram above, we obtain the following short exact sequence:

$$0 \to E'(K)/\phi E(K) \to S^\phi(E/K) \to Ш(E/K)[\phi] \to 0. \tag{7}$$

From here, it is easy to see that the weak Mordell-Weil group injects into the Selmer group with the ϕ-torsion of the Tate-Shafarevich group as the "error". It turns out that the Selmer

group is computable, making it far more tractable than either $E'(K)/\phi E(K)$ or $Ш(E/K)[\phi]$. This suggests the following way of looking for generators of $E'(K)/\phi E(K)$:

(1) Consider an element of $H^1(G_{\bar{K}/K}, E[\phi])$ and check to see if it maps to a trivial element of $\prod_{v\in M_K} H^1(G_{\bar{K_v}/K_v}, E(\bar{K_v}))$. This will tell us if this element is in the ϕ-Selmer group.
(2) Once we know this, we must determine if this element comes from $E'(K)/\phi E(K)$ or if it goes to a non-zero element of the Tate-Shafarevich group. In most cases $Ш(E/K)[\phi]$ is trivial (enough) and exactness tells us that the Selmer group is in fact isomorphic to $E'(K)/\phi E(K)$.
(3) If this is the case, then this means that we can map this element back to $E'(K)/\phi E(K)$. We then see the point in $E'(K)/\phi E(K)$ it maps to can be generated from any of the other points we already have. If not, we have found a new generator.

There are several problems we are omitting: In order to determine the elements of the Selmer group, *a priori* we must check all elements of $H^1(G_{\bar{K}/K}, E[\phi])$ to see which map trivially into $\prod_{v\in M_K} H^1(G_{\bar{K_v}/K_v}, E[\phi])$. First, not only have we not explained how to do this, we don't even know when to stop as $H^1(G_{\bar{K}/K}, E[\phi])$ is infinite for $\phi \neq 1$.

These questions will become easier once we have a concrete way of thinking about elements of these cohomology groups and specifically the Selmer and Tate-Shafarevich groups.

2. Principal Homogenous Spaces

In this section, we associate to each elliptic curve E/K, a genus one algebraic curve C/K known as a *principal homogenous space of* E/K.

Definition 1.2.1. For E/K an elliptic curve, we define a *principal homogenous space for* E as a smooth curve C/K together with a mapping $i : C \times E \to C$ such that:

(1) $i(p, \mathcal{O}) = p$ for all $p \in C$
(2) $i(i(p, P), Q) = i(p, P + Q)$ for all $p \in C$ and $P, Q \in E$
(3) For all $p, q \in C$, there exists a unique $P \in E$ such that $i(p, P) = q$

More intuitively, the map i acts essentially as addition so that we may write $i(p, P) = p + P$.

In fact, these spaces can be defined over arbitrary abelian groups (e.g., $A(K)$ the group of rational points on abelian variety). More generally, the theory we discuss generalizes to arbitrary abelian vareties even though we only treat the genus 1 case.

Example 1.2.2. Addition $E \times E \to E$ given by $(P, Q) \mapsto P + Q$ turns E into a principal homogeneous space for itself, called the *trivial principal homogeneous space*.

Another way to think of a principal homogenous space is as a twist of E - a genus 1 curve which is isomorphic to E over an extension field of K, but not necessarily over K itself - with slightly more structure. To see this, we will construct a $\bar{K}$ isomorphism from C to E:
Pick any $p_0 \in C$. Then let

$$f : E \to C \qquad f(P) = i(p_0, P) \text{ or } f(P) = p_0 + P. \tag{8}$$

Since E acts simply transitively on C, for each $p \in C$, there exists a unique $P \in E$, such that $f(P) = p$ and so $\deg(f) = 1$. This means that the induced map on function fields $f^* : K(C) \to K(E)$ is an isomorphism . Therefore, f^*, $f^{*-1} : K(E) \to K(C)$. This isomorphism gives a degree 1 rational function $g : C \to E$. But since E is smooth, g is a morphism and thus is actually an isomorphism because it is degree 1.

DEFINITION 1.2.3. We consider two homogenous spaces C and C' equivalent if there exists an isomorphism $\theta : C \to C'$ defined over K which is also compatible with the action of E on C and C', that is, such that for all $p \in C$ and $P \in E$

$$\theta(p + P) = \theta(p) + P$$

where the $+$ operator is defined in context.

The equivalence classes of homogenous spaces for E/K forms a group called the *Weil-Châtelet group for E/K*, denoted $WC(E/K)$. The next proposition characterizes the trivial element of this group as the class of homogenous spaces equivalent to the trivial space defined in Example 1.2.2.

PROPOSITION 1.2.4. *A homogenous space C/K is in the trivial equivalence class if and only if $C(K)$ is non-empty.*

PROOF. If C/K is in the trivial class, then there must be an isomorphism $\theta : E \to C$ which induces an isomorphism $E(K) \to C(K)$. Since $\mathcal{O} \in E(K)$ by definition, it must get mapped to an element of $C(K)$.

Conversely, suppose that $C(K)$ is nonempty and take $p_0 \in C(K)$. We want to show that there exists an equivalence between E/K and C/K. Consider the map $\theta' : E \to C$ defined by

$$\theta'(P) = p_0 + P. \tag{9}$$

This is the same as the map f from (8) so we already know it is a $\bar{K}$-isomorphism. So all we need to do is show that it is actually defined over K:

$$\theta(P)^\sigma = (p_0 + P)^\sigma = p_0^\sigma + P^\sigma = p_0 + P^\sigma = \theta(P^\sigma).$$

This shows that the action of σ is the same before and after the map θ' for all $\sigma \in G_{\bar{K}/K}$ so that θ' must be defined over K. Thus, θ' is an equivalence between E and C. □

Proposition 1.2.4 tells us that in order to determine if a homogenous space is in the trivial class, we need "only" look for a global rational point on it. This classificiation becomes quite useful when applied to our discussion on Selmer and Tate-Shafarevich groups in the previous section. The connection between homogenous spaces and the exact sequence (7) is made possible by the association of the Weil-Châtelet group with the first cohomology group $H^1(G_{\bar{K}/K}, E)$, which we pursue in the next section. Once this association is established, we will have a geometric interpretation of the sequences given in Section 1.

3. The Bijection $WC(E/K) \to H^1(G_{\bar{K}/K}, E)$

In this section, we want to show that there is a bijection between the cohomology classes $[\xi] \in H^1(G_{\bar{K}/K}, E)$ and equivalence classes $[i, C/K] \in WC(E/K)$.

PROPOSITION 1.3.1. *For E/K an elliptic curve, there is a bijection*

$$WC(E/K) \to H^1(G_{\bar{K}/K}, E)$$

given as follows: for C/K a homogenous space, choose any $p_0 \in C$ and we define our bijection

$$[i, C/K] \to [\sigma : p_0^\sigma - p_0] \tag{10}$$

where $p_0^\sigma - p_0$ is the unique element of E defined to satisfy this relation.

The idea behind this proof is essentially to construct two maps. First, we verify that the map above, which takes a homogenous space and associates to it a cocycle class, is indeed well-defined and injective. To show that it is surjective, we must construct another mapping, taking a cocycle class and associating to it a homogenous space. We then show that the homogenous space we get maps back to the cocycle class we started with, under the map (10). To lighten our notation, we will generally drop the equivalence class notation.

3.1. Well-definedness and injectivity. We first show that the map (10) is well-defined and injective.

PROOF. (*Of well-definedness and injectivity*). Notice that $p_0^\sigma - p_0$ satisfies the cocycle condition:

$$p_0^{\sigma\tau} - p_0 = (p_0^{\sigma\tau} - p_0^\tau) + (p_0^\tau - p_0) = (p_0^\sigma - p_0)^\tau - (p_0^\tau - p_0).$$

To show well-definedness, pick some other homogenous space C'/K equivalent to C/K under the map θ and another point $p_0' \in C'$. We want to show that the cocycles $p_0^\sigma - p_0$ and $p_0'^\sigma - p_0'$ are in the same equivalence class of $H^1(G_{\bar{K}/K}, E)$. This is just a matter of computation:

$$\begin{aligned} p_0^\sigma - p_0 &= (\theta[p_0^\sigma + (p_0 - p_0^\sigma)] - \theta(p_0)) + (p_0^\sigma - p_0) \\ &= ([\theta(p_0^\sigma) + (p_0 - p_0^\sigma)] - \theta(p_0)) + (p_0^\sigma - p_0) \qquad \text{since } \theta \text{ is an equivalence} \\ &= \theta(p_0^\sigma) - \theta(p_0) \\ &= (p_0'^\sigma - p_0') + [(\theta(p_0) - p_0')^\sigma - (\theta(p_0) - p_0')]. \end{aligned}$$

For injectivity, suppose that there exist two spaces, C/K and C'/K with points p_0 and p_0', that map to equivalent cocycles. That is, there is some $P_0 \in E$

$$p_0^\sigma - p_0 = (p_0'^\sigma - p_0') + (P_0^\sigma + P_0)$$

for all $\sigma \in G_{\bar{K}/K}$. Let $\theta : C \to C'$ be the map defined by $\theta(p) = p_0' + (p - p_0) + P_0$. We will show that θ is an equivalence. θ is a $\bar{K}$-isomorphism for the same reasons as the map θ' in the proof of Proposition 1.2.4. Further, by definition, θ is also compatible with the action of E. To see that θ is defined over K, notice that:

$$\begin{aligned} \theta(p)^\sigma &= p_0'^\sigma + (p^\sigma - p_0^\sigma) + P_0^\sigma \\ &= p_0' + (p^\sigma - p_0^\sigma) + P_0 + [(p_0'^\sigma - p_0') + P_0^\sigma - P_0 - (p_0^\sigma - p_0)] \\ &= \theta(p^\sigma). \end{aligned}$$

Therefore, θ is an equivalence between C and C'. □

3.2. Surjectivity. In order to show surjectivity, we make use of a general method for algebraic varieties known as *descent of the base field.*[2] Let L be a finite Galois extension of K and let $G = \mathrm{Gal}(L/K)$. Suppose we have some variety V, defined over L. Under certain circumstances, we can find a variety W, defined over K, such that there exists some L-isomorphism, $f : W \to V$. If such a W and f exists we can "descend" from the variety V defined over L to some variety W defined over the base field K.

Descent of the base field tells us how to find W/K given a variety V/L and, for all $\sigma \in G$, L-isomorphisms, $h_\sigma : V \to V^\sigma$ such that

$$h_{\sigma\tau} = h_\sigma^\tau \circ h_\tau, \text{ for all } \sigma, \tau \in G$$

[2]Throughout this section, our reference will be [**31**] Chap. 5, Sec. 4 and [**20**] Chaps. 4 and 5.

where V^σ is the variety obtained from V by letting $\sigma \in G$ act on the coefficients of the charts and glueings that define V.

Before continuing, we discuss how this relates to attempt to find, given $\xi \in H^1(G_{\bar{K}/K}, E)$, a principal homogenous space C/K associated to E. First, although we are working with the infinite extension $\bar{K}/K$, we can still use descent of the base field because any variety V defined over $\bar{K}$ is actually defined over a finite extension of K as V is only allowed a finite number of defining polynomials.

Our elliptic curve E is defined over K, but we embed K in $\bar{K}$, and consider $E/\bar{K}$ instead. Now, take a element $\xi' \in H^1(G_{\bar{K}/K}, E)$ and consider the map

$$\xi : G_{\bar{K}/K} \to \mathrm{Isom}(E) \text{ given by } \sigma \mapsto T_{\xi'_\sigma}$$

where T_{ξ_σ} denotes translation by ξ_σ. This new map inherits the cocycle property from ξ':

$$\xi_{\sigma\tau} = T_{\xi'_{\sigma\tau}} = T_{(\xi'_\sigma)^\tau \circ \xi'_\tau} = (T_{\xi'_\sigma})^\tau \circ T_{\xi'_\tau} = (\xi_\sigma)^\tau \circ \xi_\tau.$$

This construction allows us to embed $H^1(G_{\bar{K}/K}, E)$ inside $H^1(G_{\bar{K}/K}, \mathrm{Isom}(E))$ by sending an element $P \in E$ to T_P, which is an isomorphism of E as a curve defined over $\bar{K}$. We note that translation is not an automorphism of E as an elliptic curve because $\mathcal{O}$ is not preserved; this is why we denote the group $\mathrm{Isom}(E)$ instead of $\mathrm{Aut}(E)$.

Now, the cocycle T_{ξ_σ} is exactly the $\bar{K}$-isomorphism of $E \to E$ (because E is defined over K, $E^\sigma = E$ for all σ) that acts as the input to the process of descent of the base field. To recap, we just consider the general program discussed above, except with

$$V = E \qquad h_\sigma = T_{\xi_\sigma}.$$

Since $E^\sigma = E$, the maps T_{ξ_σ} can be summarized by giving $T_\sigma \in H^1(G_{\bar{K}/K}, \mathrm{Isom}(E))$.

Assuming descent of the base field applies, we use T_{ξ_σ} to obtain a curve C/K with a $\bar{K}$-isomorphism $f : C \to E$. This tells us that C/K is a twist of E. All that remains is to show is that C also has the structure of a principal homogenous space, and that the cohomology class associated to C/K is indeed $\{\xi\}$.

This describes our motivation for considering descent of the base field. Now we prove an algebraic lemma that will help us in the descent procedure.

LEMMA 1.3.2. *Let A be an L-vector space and suppose, for every $\sigma \in G$, we have a σ-linear bijection, $\bar{\sigma} : B \to B$, such that $\overline{\sigma\tau} = \bar{\sigma}\bar{\tau}$. If B denotes the elements of A invariant under this action, then*

$$A = B \otimes_K L.$$

PROOF. Let $r = [L : K]$. We consider the following three objects:

(1) Let C be the endomorphism algebra of L, considered as a K-vector space. Once we choose a basis, we can just think of C as $M_{r\times r}(K)$, the $r \times r$ matrices with entries in K.
(2) There is a sub-algebra of C consisting of the matrices that correspond to left multiplication by elements of L. While these matrices are in general quite complicated, we can identify this sub-algebra with L itself.
(3) Finally, consider the group ring of G over K; that is, the K-linear combinations of elements of G and denote it $K[G]$. Since all elements of G send L to itself, elements of $K[G]$ are endomorphisms of L as a K-vector space and thus we can consider $K[G]$ as another sub-algebra of C.

Define the map

$$\theta : L \otimes_K K[G] \to C$$

via composition. Since automorphisms are independent (there is no nontrivial linear combination of them that is equal to zero), it follows that θ must be injective. Since $\dim(L \otimes_K K[G]) = r^2 = \dim(C)$, θ must be surjective as well, so that θ is, in fact, an isomorphism of K-algebras.

Now we claim that $L \otimes_K K[G]$ is a *simple K-algebra* by showing that $C \cong M_{r\times r}(K)$ is. Recall that a simple algebra is one which cannot be written as the direct sum of two proper sub-algebras. Suppose to the contrary that

$$C = R \oplus S.$$

Then consider the following subspaces of the vector space $V = K^r$:

$$W_R = \{x \in V : Rx = 0\} \qquad W_S = \{x \in V : Sx = 0\}.$$

Both R and S leave W_R and W_S invariant. But this means that the matrices in C must be of the form:

$$\left(\begin{array}{c|c} * & 0 \\ \hline 0 & * \end{array}\right)$$

But this has dimension much less than r^2. Contradiction.

This shows that $L \otimes_K K[G]$ is a simple algebra, which has the property that all modules over it can be expressed as the direct sum of simple modules over $L \otimes_K K[G]$.[3] Further, a simple ring such as $L \otimes_K K[G]$ can have only one simple module, up to isomorphism.[4] We claim that L is this simple $L \otimes_K K[G]$-module.

We define the action of $L \otimes_K K[G]$ on L by

$$c \otimes_K \sigma'(\alpha) = c * \sigma'(\alpha)$$

where $c, \alpha \in L$, $\sigma' \in K[G]$, and $*$ denotes multiplication in L. This definition makes sense because G preserves L so all elements of $K[G]$ will as well, so that $\sigma'(\alpha) \in L$. It is straightforward to verify the properties of a module structure for this action.

L must be simple because we cannot have invariant subspaces for the action of $L \otimes_K K[G]$. Thus, all $L \otimes_K K[G]$-modules are isomorphic to a direct sum of copies of L as a module.

Now notice that A, our original L-vector space, has an action on it via G. Using this action, we extend to an action of $L \otimes_K K[G]$ on A so that A takes on a $L \otimes_K K[G]$-module structure. Thus, we can write:

$$\begin{aligned} A &\cong \bigoplus_{i=1}^{m} L = (K \otimes_K L) \overbrace{\oplus \cdots \oplus}^{m} (K \otimes_K L) \\ &= (K \overbrace{\oplus \cdots \oplus}^{m} K) \otimes_K L = K^{\oplus^m} \otimes_K L \\ &= (K^{\oplus^m} \otimes_K K) \times L = (K^{\oplus^m} \otimes_K L)^G \otimes_K L \\ &= A^G \otimes_K L = B \otimes_K L \end{aligned}$$

□

Now, we are ready to return to our discussion of descent.

[3] See [**20**] Chap. 17, Prop. 4.1.

[4] This is Cor. 4.6 of [**20**].

PROPOSITION 1.3.3. *Descent to the base field is possible when V is the union of affine opens U_i, defined over L, such that:*

$$h_\sigma(U_i) \subseteq U_i^\sigma. \tag{11}$$

PROOF. Condition (11) allows us to restrict to cases where V is an affine variety. Let $L[V]$ be the coordinate ring of V and let

$$L[V^\sigma] = \frac{L[X]^\sigma}{I(V^\sigma)}$$

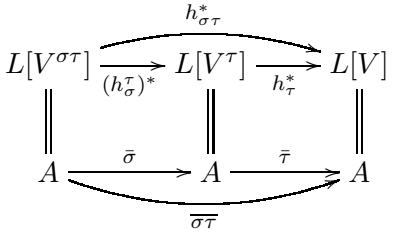
be the coordinate ring of V^σ. Although $L[V]$ and $L[V^\sigma]$ are equal as rings, they have different structures as L-algebras. Specifically, $L[V^\sigma]$ obtains the action of L from $L[V]$ *by means of* σ. That is, given $\alpha \in L$ and $x \in L[V^\sigma]$, we figure out how α acts on x by mapping everything back to $L[V]$, applying the action there, and then mapping back. Thus, the right action in $L[V]$ is $x \mapsto x^\alpha$, whereas in $L[V^\sigma]$, it is $x \mapsto x^{\alpha\sigma}$.

Via the pullback, we obtain from $h_\sigma : V \to V^\sigma$ an isomorphism of L-algebras.

$$h_\sigma^* : L[V^\sigma] \to L[V]$$

where, for $u \in L[V^\sigma]$, $h_\sigma^*(u) = u \circ h_\sigma$.

Let $A = L[V] = L[V^\sigma]$ as a ring and let $\bar{\sigma} : A \to A$ be the automorphism of A obtained from h_σ^* by simply forgetting the L-algebra structure. Then we have the following diagram



where the maps on L-algebras induce maps on the ring structure A. This diagram commutes because h_σ^* satisfies the cocycle condition. Thus, $\overline{\sigma\tau} = \bar{\sigma}\bar{\tau}$ and h_σ induces a group action of G on A via h_σ^*. Thus, it makes sense to consider the subgroup $B \subset A$ of elements invariant under the action of $\bar{\sigma}$. Since $L \subset A$, we can think of B as an A-algebra. By Proposition 5.1 of [**1**] A is integral over B if and only if it is a finite A-module. Since we took $[L : K]$ finite, this is the case. This means that B is a finitely generated K-algebra so that we can associate to B an affine K-variety W having B as its coordinate ring. We claim that W is the affine variety we are looking for. But this follows from Lemma 1.3.2.

Since B, the coordinate ring of W, is isomorphic over $\bar{K}$ to A, the coordinate ring of V, it follows that W and V must be isomorphic over L under some map $f : W \to V$.

□

In fact, we apply σ to the coefficients of f to obtain an isomorphism $f^\sigma : W^\sigma \to V^\sigma$ but since W is a K-variety, $W^\sigma = W$ so that we actually get

$$f^\sigma : W \to V^\sigma.$$

This allows us to decompose the map $h_\sigma : V \to V^\sigma$ into

$$h_\sigma = f^\sigma \circ f^{-1}.$$

Returning to the surjectivity of the map (10), descent tells us that given E and an element $\xi \in H^1(G_{\bar{K}/K}, E)$, we can find C/K such that

$$f^\sigma \circ f^{-1} = \text{translation by } -\xi_\sigma$$

where f is a $\bar{K}$-isomorphism from C to E.

We now define a map:

$$i : C \times E \to C \qquad \imath(p, P) = f^{-1}(f(p) + P)$$

and show that it gives C the structure of a principal homogenous space for E. To see that i is simply transitive, let $p, q \in C$. Then, $i(p, P) = f^{-1}(f(p) + P) = q$ so that the only choice for P is $P = f(q) - f(p)$. To check that i is defined over K, let $\sigma \in G_{\bar{K}/K}$ and notice:

$$\begin{aligned} i(p, P)^\sigma &= (f^{-1})^\sigma (f^\sigma(p^\sigma) + P^\sigma) \\ &= f^{-1}([f(p^\sigma) + \xi_\sigma + P^\sigma] - \xi_\sigma) \\ &= i(p^\sigma, P^\sigma). \end{aligned}$$

To finish off the proof that $WC(E/K)$ and $H^1(G_{\bar{K}/K}, E)$ are in bijection, using the E and ξ as inputs, we need to show that the C/K we get from descent is mapped back to the cohomology class of ξ under the original map.

To do this, we choose $p_0 \in C$ and consider the cocycle $\sigma \mapsto p_0^\sigma - p_0$. Because this map is well defined, we can take $p_0 = f^{-1}(\mathcal{O})$. Then,

$$\begin{aligned} p_0^\sigma - p_0 &= (f^\sigma)^{-1}(\mathcal{O}) - f^{-1}(\mathcal{O}) \\ &= f^{-1}(\mathcal{O} + \xi_\sigma) - f^{-1}(\mathcal{O}) \\ &= \xi_\sigma \end{aligned}$$

Thus, $WC(E/K)$ and $H^1(G_{\bar{K}/K}, E)$ are in bijective correspondence.

4. A Breather

We now rewrite the diagram on page 12 with the appropriate Weil-Châtelet groups.

$$\begin{array}{ccccccccc} 0 & \longrightarrow & \frac{E'(K)}{\phi E(K)} & \longrightarrow & S^\phi(E/K) & \longrightarrow & Ш(E/K)[\phi] & \longrightarrow & 0 \\ & & \downarrow & & \downarrow & & \downarrow & & \\ 0 & \longrightarrow & \frac{E'(K)}{\phi E(K)} & \xrightarrow{\delta} & H^1(G_{\bar{K}/K}, E[\phi]) & \longrightarrow & WC(E/K)[\phi] & \longrightarrow & 0 \\ & & \downarrow & & \downarrow & \searrow & \downarrow & & \\ 0 & \longrightarrow & \prod_{v \in M_K} \frac{E(K_v)}{\phi E(K_v)} & \xrightarrow{\delta} & \prod_{v \in M_K} H^1(G_{\bar{K}_v/K_v}, E[\phi]) & \longrightarrow & \prod_{v \in M_K} WC(E/K_v)[\phi] & \longrightarrow & 0 \end{array}$$

Now, checking if an element $\alpha \in H^1(G_{\bar{K}/K}, E[\phi])$ is in the Selmer group is equivalent, by Propositions 1.3.1 and 1.2.4, to seeing if the curve associated to the image of α in $WC(E/K)$ has a K_v-rational point for all valuations v. By Hensel's Lemma, this is much easier than the corresponding global problem; indeed it is reducible to checking over a finite ring. For each valuation we check, this is only a finite amount of computation. This indicates that it may be possible to compute the Selmer group.

Some elements of the Selmer group may map to the Tate-Shafarevich group. Concretely, these elements correspond to classes of twists which have rational points everywhere locally, but

none globally. From this perspective, $Ш(E/K)$ has the interesting property that it measures the failure of the Hasse principle.

5. Finiteness of the Selmer Group and the Weak Mordell-Weil Group

To ensure that the Selmer group is computable, we must reduce the number of elements we test for membership in $S^{\phi}(E/K)$ to a finite subset of $H^1(G_{\bar{K}/K}, E[\phi])$. This is formalized as follows:

DEFINITION 1.5.1. Let M be a $G_{\bar{K}/K}$-module and v be a discrete valuation. A cohomology class $\xi \in H^1(G_{\bar{K}/K}, M)$ is *unramified at* v if its image under the restriction map

$$H^1(G_{\bar{K}/K}, M) \to H^1(I_v, M)$$

is trivial, where I_v is the inertia group of v.

THEOREM 1.5.2. *Let $S \subset M_K$ be a finite set of places and let $\phi : E/K \to E'/K$ with $\deg(\phi) = m$. Every element of the Selmer group $S^{\phi}(E/K)$ is unramified outside of S, where S is the set of all archimeadean places, union those places where E has bad reduction, union those v where $v(m) \neq 0$.*

Moreover, the group

$$H^1(G_{\bar{K}/K}, E[\phi]; S) = \{\xi \in H^1(G_{\bar{K}/K}, E[\phi]) : \ \xi \text{ is unramified outside } S\}.$$

is finite for S a finite set of places. The Selmer group is contained in this set and so it must be finite.

COROLLARY 1.5.3 (weak Mordell-Weil). *By the exact sequence (7), we conclude that $E(K)/mE(K)$ is finite.*

PROOF. We can just take $\phi = [m]$ and $E' = E$. □

The outline of the proof of Theorem 1.5.2 is as follows: First, we show that the Selmer group is composed of cocycle classes that are unramified outside of S. Then we show that such a set must be finite. To prove this last statement, we define $L_{K,S}$ to be the maximal abelian extension of K of exponent m which is unramified outside of S and show it is a finite extension.[5] We then relate the set of cocycle classes to $G_{L_{K,S}/K}$.

First, we begin with two lemmas which are used to prove that $L_{K,S}$ is finite.

LEMMA 1.5.4. *Let K be a field of characteristic 0, containing the m-th roots of unity μ_m. Then L, the maximal abelian extension of K of exponent m, is obtained by adjoining m-th roots of all elements of K. That is, $L = K(a^{1/m} : a \in K)$.*

PROOF. Let L be an abelian extension of K of exponent dividing m and let $G = \mathrm{Gal}(L/K)$. Since G is abelian, we can decompose G into the product of cyclic groups

$$G \cong G_1 \times \cdots \times G_k.$$

Let L_i be the fixed field of $G_1 \times \cdots \times G_{i-1} \times G_{i+1} \times \cdots \times G_n$. Then, the Galois group of L_i/K is G_i, which is cyclic of order m' dividing m. It follows that $\xi \in K$ where ξ is a primitive m'-th root of unity. Since $N_{L_i/K}(\xi) = 1$, we can apply the Hilbert Satz 90, which says that $\xi = \sigma(v_i)/v_i$, where σ is a generator of $\mathrm{Gal}(L_i/K)$. Now, $\sigma(v_i^n) = \sigma(v_i)^n = (\xi^{-1}v_i)^n = v_i^n$.

[5]To ease the notation, we do not write $L_{K,S,m}$, but the m is understood.

Thus, v_i^n must be in the base field, K. Because $\sigma^i(v_i) = \xi^{-i}v_i$, the minimum polynomial of v_i must be degree m'. Thus $L_i = K(v_i)$, noting that $v_i^{m'} \in K$. Finally, the maximal abelian extension of K of exponent dividing m is

$$L = K(a^{1/m} : a \in K).$$

□

DEFINITION 1.5.5. The *S-integers* of K are those elements of K which are integral outside of S:

$$R_S = \{a \in K : v(a) \geq 0 \text{ for all } v \in M_K, v \notin S\}.$$

Notice that R_S is larger than the ring of integers R, but smaller than the valuation ring for any one valuation not in S.

Let $P(S)$ denote the set of prime ideals of $R = \mathcal{O}_K$ of the form

$$P_v = \{a \in R : v(a) > 0\} \qquad \text{for } v \in S \text{ nonarchimedean.}$$

Another way to think of R_S is as the ring of elements $a/b \in K$ where $a \in R$ and b is divisible only by prime ideals in $P(S)$.

We know that the valuation ring defined by a discrete valuation is always a PID, but that the ring of integers may not be; in general, neither is R_S. However, by the following lemma, R_S is not "far" from being a PID:

LEMMA 1.5.6. *There exists a finite set of places S such that R_S is a PID.*

PROOF. First we show that the prime ideals $\tilde{\mathfrak{p}} \subset R_S$ are in bijective correspondence with the prime ideals $\mathfrak{p} \subset R$, $\mathfrak{p} \notin P(S)$. The bijection works as follows: let Σ be the set of elements of K contained only in elements of $P(S)$. Then, for $\mathfrak{p} \subset R$ such that $\mathfrak{p} \notin P(S)$, $\Sigma^{-1}\mathfrak{p} = \tilde{\mathfrak{p}}$ is a prime ideal of R_S because $\mathfrak{p}$ does not contain any elements of Σ so that no units are introduced. If, however, $\mathfrak{p} \in P(S)$, then this procedure will yield a unit ideal, which, of course, is not prime. This gives a map from $\mathfrak{p} \notin P(S)$ to $\tilde{\mathfrak{p}} \in R_S$. Mapping the other way, we associate to $\tilde{\mathfrak{p}} \subset R_S$ the ideal $\mathfrak{p} = \tilde{\mathfrak{p}} \cap R$.

Now Let H be the class number of R, and let $I_1, \ldots, I_H$ be representatives of each of the ideal classes of R. We define S as the set of places which consist of all archimedean valuations of K, together with the set of all finite places of K corresponding to prime ideals q of R such that $q \mid I_j$ for some j. Since all the I_j can be factored into a finite product of prime ideals, S must be finite. Now we show that R_S is a PID. Because R_S is a Dedekind ring, it suffices to show that every prime ideal is principal. Every prime ideal of R_S, $\tilde{\mathfrak{p}}$, can be associated with an ideal $\mathfrak{p} \subset R$ where $\mathfrak{p} \notin P(S)$. Since $[p] = [I_j]$ for some j, we write $p = (\alpha)I_j$ for some $\alpha \in K^*$. But all the I_j can be written as $I_j = \mathfrak{q}_1 \cdots \mathfrak{q}_r$ for $\mathfrak{q}_i \in P(S)$. Thus, when we map to R_S, each $\mathfrak{q}_i$ becomes the unit ideal. As a result, $\tilde{I}_j = (1)$ so that $\tilde{\mathfrak{p}} = (\tilde{\alpha})$ is principal. □

Now we show that the extension $L_{K,S}$ is indeed finite.

PROPOSITION 1.5.7. *Let K be any number field, $S \subset M_K$ be a finite set of valuations including the archimedean valuation, and $m \geq$ be an integer. Then $L_{K,S}$, the maximal abelian extension of exponent m unramified outside S is finite.*

PROOF. We can add a finite set of places to S because if $L_{K,S}$ were unramified outside the original set, it will clearly be unramified outside of the new set. Thus, using Lemma 1.5.6, we assume R_S is a PID. We may also enlarge S so that $v(m) = 0$ for all $v \notin S$ because the valuation of any nonzero element of K is equal to 0 for almost all places by the Product Formula. Finally,

reduce to the case when $\mu_m \subset K$. Suppose this proposition were true for some finite extension K' of K where S' is the set of valuations extending those in S. Then $L_{K',S'} = LK'/K'$, the maximal abelian extension of K' of exponent m unramified outside S', is finite. Looking at the resulting tower of field extensions, we see that $L_{K,S}$ must be finite as well. Thus we may always extend K to a K' containing μ_m. Now Lemma 1.5.4 tells us that $L_{K,S} \subset K(a^{1/m} : a \in K)$.

In order for $L_{K,S}$ to be unramified outside S, $a^{1/m} \in L_{K,S}$ if and only if the extension $K_v(a^{1/m})/K_v$ is unramified. We claim that this corresponds precisely to those $a \in K$ such that:

$$m \mid \mathrm{ord}_v(a).$$

An extension M of a complete field K_v is umramified if and only if all the elements of M have integral valuation under the unique extension of the normalized valuation on K_v.Suppose $m \nmid \mathrm{ord}_v(a)$. Then the valuation of $a^{1/m}$ is not integral so that the extension must be ramified. Conversely, suppose that $m \mid \mathrm{ord}_v(a)$. Then we write $a = \pi^{rm}u$ or $a^{1/m} = \pi^r u^{1/m}$ where u is a unit. This shows that we may reduce to the case where $a = u$ is a unit. The discriminant of $K_v(a^{1/m})$ divides $m^m u^{m-1}$.[6] However, since $v(u) = 0$ and $v(m) = 0$, $v \nmid m$ and $v \nmid u$. Thus, v does not divide the discriminant of $K(a^{1/m})$ and therefore the extension is unramified.

Thus, we conclude that

$$L_{K,S} = K(a^{1/m} : a \in K(S, m))$$

where $K(S,m) = \{a \in K^*/(K^*)^m : m \mid \mathrm{ord}_v(a) \text{ for all } v \in M_K, v \notin S\}$. To finish the proof, we show that $K(S,m)$ is finite.

Consider the natural reduction map:

$$R_S^* \to K(S, m). \tag{12}$$

We want to show that it is surjective. This will allow us to use Dirichlet's S-Unit theorem to show that $K(S,m)$ is finite. Take $a \in K^*$, a representative of an element in $K(S,m)$ and consider the ideal $aR_S \subset R_S$. By Lemma 1.5.6, it has a unique factorization into prime ideals $\mathfrak{p} \notin P(S)$. Since $m \mid \mathrm{ord}_v(a)$, the power of each prime ideal in the factorization of aR_S is divisible by m. This means that aR_S is the m-th power of some other ideal in R_S. As R_S is a PID, this ideal must be of the form bR_S for some $b \in K^*$. Hence, $aR_S = b^m R_S$ where

$$a = b^m u \qquad u \in R_S^*.$$

This u is an element of R_S^* that maps onto the element of $K(S,m)$ that a represents.

Because R_S^{*m} is in the kernel of the map (12), the map

$$\frac{R_S^*}{R_S^{*m}} \to K(S, m)$$

is surjective. Dirichlet's S-Unit Theorem says that R_S^* is finitely generated so that $K(S,m)$ must be finite. □

The next two lemmas are directly related to the proof of Theorem 1.5.2.

Lemma 1.5.8. *Let v be a discrete valuation on K and suppose that $v(m) = 0$ and E has good reduction at v. Then the reduction map*

$$E(K)[m] \to \tilde{E}(k_v)$$

is injective.

[6] See [**5**] Chap. 5, Lem. 5.

PROOF. From Proposition 0.1.10, the sequence:

$$0 \to E_1(K_v) \to E_0(K_v) \to \tilde{E}_{ns}(k_v) \to 0$$

is exact. $E_1(K_v)$ is a isomorphic to a formal group and from the general theory of formal groups, this means $E_1(K_v)[m] = 0$.[7] Further, since $\tilde{E}$ is nonsingular, $E_0(K_v) = E(K_v)$ and $\tilde{E}_{ns}(k_v) = \tilde{E}(k_v)$, so we get that $E(K)[m]$ injects into $\tilde{E}(k_v)$. □

LEMMA 1.5.9. *Let everything be as in the statement of Theorem 1.5.2 and additionally, assume that $E[\phi] \subset K$. Then*

$$\mathrm{Hom}(G_{L_{K,S}/K}, E[\phi]) \cong \mathrm{Hom}(G_{\bar{K}/K}, E[\phi]; S).$$

PROOF. Because the action of $G_{L_{K,S}/K}$ and $G_{\bar{K}/K}$ is trivial on $E[\phi] \subset K$, our homomorphism groups are really the same as cocycle groups. Because $L_{K,S}$ is Galois by definition, we may write down the inflation-restriction sequence:

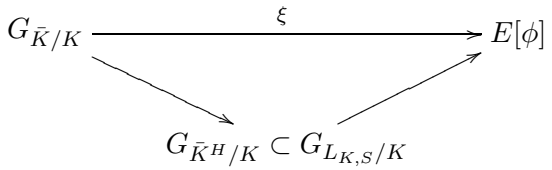
$$0 \to \mathrm{Hom}(G_{L_{K,S}/K}, E[\phi]) \overset{\inf}{\to} \mathrm{Hom}(G_{\bar{K}/K}, E[\phi])$$

where the image actually lands in $\mathrm{Hom}(G_{\bar{K}/K}, E[\phi]; S)$ because $L_{K,S}$ is unramified outside of S. This shows injectivity.

To show surjectivity, we claim that every cocycle of $\mathrm{Hom}(G_{\bar{K}/K}, E[\phi]; S)$ factors through $\mathrm{Hom}(G_{L_{K,S}/K}, E[\phi])$. Consider the sequence

$$0 \to H \to G_{\bar{K}/K} \overset{\xi}{\to} E[\phi]$$

where $\xi \in \mathrm{Hom}(G_{L_{K,S}/K}, E[\phi])$ and H is defined so that the sequence is exact. By Galois theory, the extension $\bar{K}^H/K$ is unramified outside S and has Galois group $G_{\bar{K}/K}/H \subset E[\phi]$. Together, this tells us that $\bar{K}^H$ is an abelian extension of exponent m which is unramified outside S. But $L_{K,S}$ is defined to be the maximal such extension so we must have $\bar{K}^H \subset L_{K,S}$. This shows that we have the diagram:



□

Finally, we prove that the Selmer group is finite:

THEOREM 1.1.5.2 (restatement) *Let $S \subset M_K$ be a finite set of places and let $\phi : E/K \to E'/K$ with $\deg(\phi) = m$. Every element of the Selmer group $S^\phi(E/K)$ is unramified outside of S, where S is the set of all archimeadean places, union those places where E has bad reduction, union those v where $v(m) \neq 0$.*

Moreover, the group

$$H^1(G_{\bar{K}/K}, E[\phi]; S) = \{\xi \in H^1(G_{\bar{K}/K}, E[\phi]) : \ \xi \text{ is unramified outside } S\}.$$

is finite for S a finite set of places. The Selmer group is contained in this set and so it must be finite.

[7] See [**36**] Chap 4, Prop. 3.2b.

PROOF. Consider any $\xi \in S^\phi(E/K)$ and $v \notin S$. We first show that ξ is unramified at v. Since $\xi \in S^\phi(E/K)$, it must map to a trivial cocycle of $H^1(G_v, E(\bar{K}_v))$ so that there exists some $P \in E(\bar{K}_v)$ such that ξ is of the form

$$\xi_\sigma = P^\sigma - P \qquad \text{for all } \sigma \in G_v$$

where $P^\sigma - P \in E[\phi] \subset E[m]$ since $S^\phi(E/K) \subset H^1(G_{\bar{K}/K}, E[\phi])$. Since $I_v \subset G_v$, this relation holds for all elements of the inertia subgroup. But moreover, the inertia group acts trivially on $\tilde{E}_v$ so that

$$\widetilde{P^\sigma - P} = \tilde{P}^\sigma - \tilde{P} = \tilde{\mathcal{O}}.$$

Thus, $P^\sigma - P$ is in the kernel of the reduction map. Lemma 1.5.8 tells us that $E(K)[m] \to \tilde{E}_v$ is injective. Thus, since $P^\sigma - P \in E[m]$ maps to something trivial, it must have been trivial to begin with:

$$P^\sigma - P = \xi_\sigma = 0 \qquad \text{for all } \sigma \in I_v$$

so that v is unramified and $S^\phi(E/K) \subset H^1(G_{\bar{K}/K}, E[\phi]; S)$.

Next, we show that $H^1(G_{\bar{K}/K}, E[\phi]; S)$ is finite. Since $E[\phi] \subset E[m]$ is finite and $G_{\bar{K}/K}$ acts continuously on $E[\phi]$, we obtain a subgroup of finite index in $G_{\bar{K}/K}$ which fixes all elements of $E[\phi]$. Let's call this $G_{\bar{K}/M}$ for some finite extension M/K. From the inflation/restriction sequence, we obtain the following exact sequence:

$$0 \to H^1(G_{M/K}, E[\phi]) \overset{\text{inf}}{\to} H^1(G_{\bar{K}/K}, E[\phi]) \overset{\text{res}}{\to} H^1(G_{\bar{K}/M}, E[\phi])$$

which means that

$$H^1(G_{\bar{K}/K}, E[\phi])/H^1(G_{M/K}, E[\phi]) \cong H^1(G_{\bar{K}/M}, E[\phi]).$$

However, because M/K is finite, $H^1(G_{M/K}, E[\phi])$ must be as well. Thus, if we show that $H^1(G_{\bar{K}/M}, E[\phi])$ is finite, we automatically get that $H^1(G_{\bar{K}/K}, E[\phi])$ is as well. Thus, we may proceed with $K = M$ and assume the action of $G_{\bar{K}/K}$ on $E[\phi]$ is trivial. This lets us rewrite $H^1(G_{\bar{K}/K}, E[\phi])$ as $\text{Hom}(G_{\bar{K}/K}, E[\phi])$. By Lemma 1.5.9,

$$\text{Hom}(G_{L_{K,S}/K}, E[\phi]) \cong \text{Hom}(G_{\bar{K}/K}, E[\phi]; S).$$

But Proposition 1.5.7 tells us that $L_{K,S}$ is a finite extension so that $\text{Hom}(G_{L_{K,S}/K}, E[\phi])$ is a subset of maps between two finite sets. Therefore, it must be finite, implying that $H^1(G_{\bar{K}/K}, E[\phi]; S)$ is as well. □

REMARK 1.5.10. By definition, $S^\phi(E/K) \subset H^1(G_{\bar{K}/K}, E[\phi]; S)$. Elements of the Selmer group, however, must also be trivial in $H^1(G_{\bar{K}_v/K_v}, E(\bar{K}_v))$ for $v \in S$. Thus, in order to check whether an element of $H^1(G_{\bar{K}/K}, E[\phi]; S)$ is in the Selmer group, we need only check local soluability at the finite number of places $v \in S$. It follows that if $H^1(G_{\bar{K}/K}, E[\phi]; S)$ is computable, then so is the Selmer group.

THEOREM 1.5.11. *$H^1(G_{\bar{K}/K}, E[\phi]; S)$ is computable.*

PROOF. Inflation/restriction gives:

$$0 \to H^1(G_{L_{K,S}/K}, E[\phi]) \overset{\text{inf}}{\to} H^1(G_{\bar{K}/K}, E[\phi]) \overset{\text{res}}{\to} H^1(G_{\bar{K}/L_{K,S}}, E[\phi])$$

We claim that $H^1(G_{\bar{K}/K}, E[\phi]; S) \subset H^1(G_{\bar{K}/K}, E[\phi])$ maps to 0 under the restriction to $H^1(G_{\bar{K}/L_{K,S}}, E[\phi])$. By exactness, this would imply that $H^1(G_{\bar{K}/K}, E[\phi]; S)$ is contained in $H^1(G_{L_{K,S}/K}, E[\phi])$. Since both $G_{L_{K,S}/K}$ and $E[\phi]$ are finite and computable, we can just enumerate the possible maps between them and test to see which are cocycles and it would

follow that $H^1(G_{L_{K,S}/K}, E[\phi])$ is computable.

To show this claim, let M be the field extension obtained by adjoining the coordinates of all elements of $E[\phi]$ to K. Note that we have the tower of fields $K \subset M \subset L_{M,S}$. We obtain the commutative diagram:

$$\begin{array}{ccccc} 0 \longrightarrow \mathrm{Hom}(G_{L_{M,S}/M}, E[\phi]) & \xrightarrow{\inf} & \mathrm{Hom}(G_{\bar{K}/M}, E[\phi]; S) & \xrightarrow{\mathrm{res}} & H^1(G_{\bar{K}/L_{M,S}}, E[\phi]; S) \\ & & \uparrow{\scriptstyle \mathrm{res}} & \nearrow{\scriptstyle \mathrm{res}} & \\ & & H^1(G_{\bar{K}/K}, E[\phi]; S) & & \end{array}$$

where we have replaced H^1 by Hom where appropriate and have written $H^1(G_{L_{M,S}/M}, E[\phi])$ for $H^1(G_{L_{M,S}/M}, E[\phi]; S)$ because $L_{M,S}$ is unramified outside of S. Further, by Lemma 1.5.9, the inflation map is actually an isomorphism. Thus, the restriction of $H^1(G_{\bar{K}/M}, E[\phi]; S)$ to $H^1(G_{\bar{K}/L_{K,S}}, E[\phi]; S) \subset H^1(G_{\bar{K}/L_{M,S}}, E[\phi]; S)$ is the zero map. By commutativity, this means that the restriction $H^1(G_{\bar{K}/K}, E[\phi]; S)$ to $H^1(G_{\bar{K}/L_{K,S}}, E[\phi]; S)$ is as well. □

In proving the finiteness of the Selmer group, we assumed two major results from algebraic number theory: finiteness of the class group of R and Dirichlet's S-unit theorem.

THEOREM 1.5.12 (Finiteness of Class Number). *The ideal class group C_K of a number field K is finite.*

The proof involves finding a uniform bound on the norms of integral ideals in ideal classes of R and then showing that there are only finitely many integral ideals in R within a certain bound. The bound generally used is the Minkowski bound, and it depends only on the number of real and complex embeddings. For a complete proof, see [**39**] Chap. 10, Thm. 4.2.

THEOREM 1.5.13 (Dirichlet's S-unit Theorem).

$$R_S^* \cong \mu_K \times \mathbb{Z}^{|S|-1}.$$

In particular, this implies that R_S^* is finitely generated. The S-unit theorem follows fairly quickly from dirichlet's unit theorem, whose proof involves defining a map of R^* to $\mathbb{R}^m$ and showing that the kernel is finite and that the image is a lattice in a hyperplane of $\mathbb{R}^m$. Thus, the kernel gives the finite cyclic part of the unit group structure and the image gives the free part. A proof is given in [**39**] Chap. 12, Thm. 1.2.

6. Descent for $m = 2$

The previous sections have told us that the process of finding generators for $E(K)/mE(K)$ amounts to taking a curve E, looking for global points on a principal homogenous space C and "descending" from those points on C to ones on E. In practice, the most common algorithms for descent take $m = 2$ and attempt to compute the generators of $E(K)/2E(K)$ (although 3 and 4 descents are becoming more popular).

There are roughly three methods of 2-descent, which work depending on the structure of $E[2]$.

(1) The simplest is the method of *complete 2-descent* which works[8] when

$$E(K)[2] \cong \mu_2 \times \mu_2.$$

[8]By works, we mean that it computes generators of $E(K)/2E(K)$ when $Ш(E/K)[2] = \emptyset$.

(2) Next is the method of *descent via 2-isogeny* which works for any E with $E(K)[2] \neq \emptyset$ by decomposing the multiplication-by-m map into the two isogenies ϕ and ϕ' from Proposition 0.1.7 and then computing $E'(K)/\phi E(K)$ and $E(K)/\phi' E'(K)$.
(3) Finally, general 2-descent works for an arbitrary elliptic curve, but requires more effort as we need to find the Selmer group itself (as opposed to embedding it in a simpler group). This was developed by Birch and Swinnerton-Dyer in [**3**] and generalized by Cassels [**7**] to general number fields.

Although all of these methods may fail, we are at least guaranteed of computing $S^2(E/K)$. By the exact sequence (7) with $\phi = [2]$, all elements of $S^2(K, E)$ must be of order 2 so that

$$\#(S^2(K, E)) = 2^k \qquad \text{for } k \in \mathbb{N}.$$

Since $E(K)/2E(K)$ injects into the 2-Selmer group, its order must be

$$\#(E(K)/2E(K)) = 2^{r+t} \qquad \text{for } r = \text{rank}(E(K)),\ 2^t = \#E(K)[2], \text{ and } k - t \geq r.$$

Thus, computing the rank of the Selmer group will always give an upper bound on the rank of the Mordell-Weil group, $E(K)$.

We will give an example of complete 2-descent. Details on the others can be found in [**36**] Chap.10, Sec. 4 and [**45**] Chapter 7 respectively.

Now, suppose that $E(K)[2] \cong \mu_2 \times \mu_2$. Recall that we have the exact sequence:

$$0 \to E(K)/2E(K) \xrightarrow{\delta} H^1(G_{\bar{K}/K}, E[2])$$

This means that there is an injective homomorphism

$$\Theta : E(K)/2E(K) \hookrightarrow H^1(G_{\bar{K}/K}, \mu_2 \times \mu_2).$$

Our goal then, is to take an element of $H^1(G_{\bar{K}/K}, \mu_2 \times \mu_2)$ and see when it maps back to an element of $E(K)/2E(K)$. In our previous discussion, we saw that it sufficed to consider the injection of $E(K)/2E(K)$ into the *finite* set $H^1(G_{\bar{K}/K}, \mu_2 \times \mu_2; S)$. However, this group is still rather abstract. We now try to make it more concrete.

There is an exact sequence analogous to the sequence (3) on page 11:

$$1 \to \mu_2 \to \bar{K}^* \xrightarrow{[2]} \bar{K}^* \to 1.$$

Cohomology gives the following:

$$1 \to \bar{K}^*/(\bar{K}^*)^2 \xrightarrow{\delta} H^1(G_{\bar{K}/K}, \mu_2) \to H^1(G_{\bar{K}/K}, \bar{K}^*).$$

The Hilbert Satz 90 says that $H^1(G_{\bar{K}/K}, \bar{K}^*) = 0$ so that we get an isomorphism

$$\bar{K}^*/(\bar{K}^*)^2 \cong H^1(G_{\bar{K}/K}, \mu_2).$$

Since the subgroup $K(S, 2) \subset \bar{K}^*/(\bar{K}^*)^2$ corresponds to the finite cohomology group $H^1(G_{\bar{K}/K}, \mu_2; S)$, we can rewrite the injection of $E(K)/2E(K)$ into $H^1(G_{\bar{K}/K}, \mu_2 \times \mu_2)$ as

$$\Theta : E(K)/2E(K) \hookrightarrow K(S, 2) \times K(S, 2).$$

The above discussion applies for arbitrary m. In the case of $m = 2$ we can explicitly write the injection Θ as follows: Given an elliptic curve with $E[2](K) = 4$, we write it as

$$y^2 = (x - e_1)(x - e_2)(x - e_3) \qquad \text{for } e_1, e_2, e_3 \in K.$$

The points $T_i = (e_i, 0)$ for $i = 1, 2, 3$ form the nontrivial 2-torsion. Then

$$\Theta(P) = \Theta(x, y) = \begin{cases} (x - e_1, x - e_2) & \text{if } x \neq T_1 \text{ or } T_2 \\ \left(\frac{e_1 - e_3}{e_1 - e_2}, e_1 - e_2\right) & \text{if } P = T_1 \\ \left(e_2 - e_1, \frac{e_2 - e_3}{e_2 - e_1}\right) & \text{if } P = T_2 \\ (1, 1) & \text{if } P = \mathcal{O} \end{cases}$$

Verifying that this is an injective homomorphism is a matter of computation and we omit the calculations.[9]

Using this injection, we want to examine each of the finitely many elements $(b_1, b_2) \in K(S, 2) \times K(S, 2)$ and see when they can be mapped back to $E(K)/2E(K)$. This is analogous to our discussion of how to take an element of $H^1(G_{\bar{K}/K}, E[\phi]; S)$ and figure out whether it comes from an element of $E(K)/mE(K)$ via the Selmer group. To do this, we first mapped the cocycle to a principal homogenous space. This is what we will do here explicitly.

Consider an arbitrary element, $(b_1, b_2) \in K(S, 2)$. Each representative b of a class in $K(S, 2)$ can be written as $b'z^2$ where $z \in K^*$. If there exists some $P = (x, y) \in E(K)/2E(K)$, which maps to (b_1, b_2), then the following equations must be satisfied:

(1) $y^2 = (x - e_1)(x - e_2)(x - e_3)$ because (x, y) must be on the curve E.
(2) $b_1 z_1^2 = x - e_1$ and $b_2 z_2^2 = x - e_2$ by definition of the map Θ.

Combining these, we obtain

$$b_1 z_1^2 - b_2 z_2^2 = e_2 - e_1, \qquad b_1 z_1^2 - b_1 b_2 z_3^2 = e_3 - e_1$$

where $z_3 = y/b_1 b_2 z_1 z_2$ These equations define another curve associated to E, the principal homogenous space. If there is a rational solution on this curve, then we recover the point $P = (x, y) \in E(K)/2E(K)$ where $(x, y) = (b_1 z_1^2 + e_1, b_1 b_2 z_1 z_2 z_3)$.

Now we consider an explicit example using the curve from Example 0.2.4 of the introduction,

$$E = y^2 = x - x^3 = -x(x + 1)(x - 1)$$

There are three nontrivial torsion points: $T_1 = (e_1, 0)$, $T_2 = (e_2, 0)$, $T_3 = (e_3, 0)$. Thus,

$$E[2] = \{\mathcal{O}, T_1, T_2, T_3\} \cong \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$$

so that we can apply complete 2-decent.

The discriminant of E is 2^6 so $S = \{\infty, 2\}$. Thus, the group $\mathbb{Q}(S, 2) = \langle 2, -1\rangle$, and the injective homomorphism $E(\mathbb{Q})/2E(\mathbb{Q}) \hookrightarrow S^2(E/\mathbb{Q}) \subset H^1(G_{\bar{\mathbb{Q}}/\mathbb{Q}}, E[2]; S)$ can be written as

$$\Theta : E(\mathbb{Q})/2\mathbb{Q} \hookrightarrow \langle -1, 2\rangle^{\oplus^2}$$

Now, using the equations we define above, for $(a, b) \in \langle -1, 2\rangle^{\oplus^2}$, we recover a point of $E(\mathbb{Q})/2E(\mathbb{Q})$ if the following equations have a solution in $\mathbb{R}$ and $\mathbb{Q}_2$.

$$x + 1 = az_1^2, \qquad x = bz_2^2, \qquad x - 1 = abz_3^2.$$

Let's call the curve that these equations define $C_{a,b}$. If $a < 0$ and $C_{a,b}$ has a real point, then the first equation shows that it satisfies $x \leq -1$, the second equation shows that $b < 0$, and the third equation says that $x > 1$, a contradiction. Thus, any element of $\langle -1, 2\rangle^{\oplus^2}$ with a negative first coordinate is ruled out. This leaves $\{(1, 1), (1, 2), (2, -1), (1, -1), (2, 1)\}$. Except for $(1, 2)$, all of these points lead to curves $C_{a,b}$ with solutions in $\mathbb{Q}$ so that they correspond to elements in the Selmer group

$$S^2(E/\mathbb{Q}) = \{(1, 1), (2, -1), (1, -1), (2, 1)\},$$

[9]For more discussion, see [**36**] Chap. 10, Sec. 1.

which has order 4. This shows that $E(\mathbb{Q})/2E(\mathbb{Q})$ surjects onto $S^2(E/\mathbb{Q})$ (because their orders are the same) so that Ш$(E/\mathbb{Q})[2] = 0$ Finally, as $E[2] \subset E(\mathbb{Q})$, we know that $\#(E(\mathbb{Q})/2E(\mathbb{Q}) = 2^{2+r}$. However, it is also true that $\#(E(\mathbb{Q})/2E(\mathbb{Q})) \leq 4$, so $r = 0$.

7. Nontrivial Ш$(E/K)[\phi]$

We have thus far been working with curves for which Ш(E/K) is trivial or has at least trivial ϕ-torsion. This means that once we have found an element of the ϕ-Selmer group, we can immediately conclude that it comes from $E'(K)/\phi E(K)$. This is equivalent to saying that the Hasse Principle holds enough so that finding a homogenous space with K_v-rational points will always give us a K-rational point, which can be mapped down to $E(K)/mE(K)$.

However, the ϕ-torsion of the Tate-Shafarevich group is not always trivial, for example with the curve:

$$E : y^2 = x^3 - 450079^2 x$$

which we encounter when trying to decide if $n = 450079 = 7 \cdot 113 \cdot 569$ is a congruent number. This curve was found by Elkies and more examples are available on his website:
`http://www.math.harvard.edu/ elkies/compnt.html`.

7.1. An example with `mwrank`. In this section, we compute $E(\mathbb{Q})$ for the above curve, using Cremona's `mwrank` program, which implements a descent via 2-isogeny.[10]

In `mwrank`'s notation, this curve is expressed as $[0, 0, 0, -450079^2, 0]$. First, `mwrank` checks if E has any nontrivial 2-torsion points. In this case it does, finding:

```
3 points of order 2:  [0:0:1], [450079:0:1], [-450079:0:1]
```

where the torsion points are written in projective notation. For a given 2-torsion point, T_1, a change of coordinates yields:

$$E : y^2 = x(x^2 + cx + d) \qquad \text{where } T_1 = (0, 0) \text{ and } c, d \in \mathbb{Z}.$$

Each T_1 gives a 2-isogenous curve $E' = E/\langle T \rangle$ with the equation

$$E : y^2 = x(x^2 + c'x + d') \qquad \text{where } c' = -2c \text{ and } d' = c^2 - 4d.$$

Let $\phi' : E \to E'$ be the 2-isogeny and let $\phi : E' \to E$ be its dual (this seems backward, but is consistent with `mwrank`'s notation).

In our example, `mwrank` finds the following 2-isogenous curves:

(1) `[0,0,0,810284424964,0]` or $y^2 = x^3 + 810284424964x$.
(2) `[0,-2700474,0,202571106241,0]` or $y^2 - 2700474y = x^3 + 202571106241x$.
(3) `[0,2700474,0,202571106241,0]` or $y^2 + 2700474y = x^3 + 202571106241x$.

Using each of these curves E', `mwrank` attempts to compute

$$E(\mathbb{Q})/\phi E'(\mathbb{Q}) \qquad \text{and} \qquad E(\mathbb{Q})/\phi E'(\mathbb{Q})$$

and, if successful, uses them to find generators for $E(\mathbb{Q})/2E(\mathbb{Q})$. Although descent via 2-isogeny requires only one 2-isogenous curve, having more is beneficial as the rank of $S^{\phi}(E'/\mathbb{Q})$ may be different for different E'; because the rank of $E(\mathbb{Q})$ is invariant under isogeny this means that we may take the best bounds from among the isogenous curves we test.

`mwrank` first begins with descent on the isogenous curve (1). The first step is to compute the ϕ and ϕ' Selmer groups so that we can determine $S^2(E/\mathbb{Q})$. Although it is theoretically possible to compute $S^2(E/\mathbb{Q})$ directly, `mwrank` works with 2-isogenies instead because they are

[10] Details can be found in [**8**] Chap. 3, Sec. 6 or in [**9**].

degree 2 maps whereas multiplication by 2 is degree 4 and thus increases computation time dramatically.

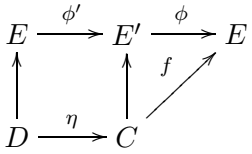
```
First step, determining 1st descent Selmer groups
After first local descent, rank bound = 5
```
$\texttt{rk}(S^{\phi}(E')) = 4$
$\texttt{rk}(S^{\phi'}(E)) = 3$

This process, known as the *first descent*, tells us the principle homogenous curves that have rational points everywhere locally. In the case of 2-descent, they are quartics of the form

$$C_{d_1} : v^2 = d_1u^4 + cu^2 + d_2 \qquad \text{where } d_1d_2 = d \text{ for } d_1, d_2 \in \mathbb{Z} \text{ and } d_1 \text{ is squarefree.}$$

We know from Section 4 that if any of these C_{d_1} has a global rational point, then that point maps down to a rational point on E. This is done by using the method described in Section 1. However, we may not be able to find a global rational point on C_{d_1} for each d_1 either because no such point exists (e.g. the element of the Selmer group maps to a nontrivial element of the Tate-Shafarevich group) or the rational point is so large that it is beyond our search space. In either of these cases, we only get an upper bound.

To refine this bound, `mwrank` performs a *second descent*. Currently, we have curves C which are ϕ-coverings of E of degree 2. We want to extend C to a 2-covering D so that we have a commutative diagram:



A rational point on D maps via η to a rational point on C, which then maps via f to one on E. One benefit of this approach is that a rational point on D is likely to have smaller "size" than one on C because $\deg(\eta) \geq 2$.[11] Another is that if all the curves D associated to C do not have points everywhere locally, then $C \in S^{\phi}(E'/\mathbb{Q})$ must not be in the image of $S^2(E/\mathbb{Q})$ and so must represent a nontrivial element of Ш$(E'/\mathbb{Q})$. This is illustrated in the following exact, commutative diagram:[12]

$$\begin{array}{ccccccccc}
0 & \longrightarrow & B & \longrightarrow & B & \longrightarrow & 0 & & \\
 & & \downarrow & & \downarrow & & \downarrow & & \\
0 & \longrightarrow & E'(\mathbb{Q})/\phi' E(\mathbb{Q}) & \longrightarrow & S^{\phi'}(E/\mathbb{Q}) & \longrightarrow & Ш(E/\mathbb{Q})[\phi] & \longrightarrow & 0 \\
 & & \downarrow{\scriptstyle \phi} & & \downarrow & & \downarrow & & \\
0 & \longrightarrow & E(\mathbb{Q})/mE(\mathbb{Q}) & \longrightarrow & S^{m}(E/\mathbb{Q}) & \longrightarrow & Ш(E/\mathbb{Q})[m] & \longrightarrow & 0 \\
 & & \downarrow & & \downarrow{\scriptstyle \phi'} & & \downarrow{\scriptstyle \phi'} & & \\
0 & \longrightarrow & E(\mathbb{Q})/\phi E'(\mathbb{Q}) & \longrightarrow & S^{\phi}(E'/\mathbb{Q}) & \longrightarrow & Ш(E'/\mathbb{Q})[\phi] & \longrightarrow & 0 \\
 & & \downarrow & & \downarrow & & \downarrow & & \\
 & & 0 & & C & & C & &
\end{array}$$

[11] The notion of size we are referring to is formalized by the height function. We will discuss this in Chapter 2.

[12] See [**9**] pg. 5.

where B and C are defined to be the appropriate kernels and cokernels, respectively.

Homogenous spaces C of E which are everywhere locally trivial represent elements of $S^{\phi}(E'/\mathbb{Q})$ and if there exists D which is also everywhere locally trivial, D is in the preimage of an element of $S^{\phi}(E'/\mathbb{Q})$ in $S^{m}(E/\mathbb{Q})$.

From the center vertical sequence, knowing the ranks of B, $S^{\phi'}(E/\mathbb{Q})$, and $\phi'(S^{m}(E/\mathbb{Q}))$ would allow us to determine the rank of $S^{2}(E/\mathbb{Q})$.

In our first descent, `mwrank` already computed the rank of $S^{\phi'}(E/\mathbb{Q})$, which turned out to be 3. In the second descent, `mwrank` outputs:

```
 Second step, determining 2nd descent Selmer groups
After second local descent, rank bound = 5
rk(φ'(S^2(E))) = 4
rk(φ(S^2(E'))) = 3
rk(S^2(E)) = 7
rk(S^2(E')) = 6
```

Here, `mwrank` is simultaneously computing the rank of $S^{2}(E/\mathbb{Q})$ and $S^{2}(E'/\mathbb{Q})$, which is in the center of an analogous diagram. In this case, $B = 0$, so that by exactness, the rank of $S^{2}(E/\mathbb{Q})$ is equal to the rank of the kernel of the center ϕ' map plus its image which is $7 = 3 + 4$, as noted in the `mwrank` output. Similarly, we obtain the rank of $S^{2}(E'/\mathbb{Q})$ in the same way. Note that in that case, however, $4 + 3 \neq 6$ so that B' must have rank 1 in the diagram we did not show. The rank of the Selmer group gives us a bound on the rank of $E(\mathbb{Q})$, which is invariant under isogeny. In this case, we get a bound of 5.

When we do the same thing with the 2 other isogenous curves, we get better bounds:

```
 Using 2-isogeny number 2
Using 2-isogenous curve [0,-2700474,0,202571106241,0]
First step, determining 1st descent Selmer groups
After first local descent, rank bound = 5
rk(S^φ(E')) = 4
rk(S^φ'(E)) = 3
Second step, determining 2nd descent Selmer groups
After second local descent, rank bound = 3
rk(φ'(S^2(E))) = 4
rk(φ(S^2(E'))) = 1
rk(S^2(E)) = 7
rk(S^2(E')) = 4  Using 2-isogeny number 3
Using 2-isogenous curve [0,2700474,0,202571106241,0]
First step, determining 1st descent Selmer groups
After first local descent, rank bound = 5
rk(S^φ(E')) = 4
rk(S^φ'(E)) = 3
Second step, determining 2nd descent Selmer groups
After second local descent, rank bound = 3
rk(φ'(S^2(E))) = 4
rk(φ(S^2(E'))) = 1
rk(S^2(E)) = 7
rk(S^2(E')) = 4
```

These isogenies give an upper bound on the rank of $E(\mathbb{Q})$ of 3. Since the rank of the Selmer group was 7, this tells us that the 2-torsion of the Tate-Shafarevich group is of order at least 4. To find generators of $E(K)$, we could just test elements of $S^2(E/\mathbb{Q})$ for rational points, but again, working with a degree 4 map would take more time than working with our 2-isogenies. Thus, we look for points on $E(\mathbb{Q})/\phi E'(\mathbb{Q})$ and $E'(\mathbb{Q})/\phi' E(\mathbb{Q})$ instead. We then use the exact sequence:

$$0 \to \frac{E'(K)[\phi]}{\phi'(E(K)[2])} \to \frac{E'(K)}{\phi'(E(K))} \xrightarrow{\phi} \frac{E(K)}{2E(K)} \to \frac{E(K)}{\phi(E'(K))} \to 0$$

to arrive at generators of $E(K)/2E(K)$. `mwrank` outputs:

```
I. Points on E mod phi(E')
Point [53767647152352:117078959809329057:11239424],
height = 15.1743661746344
Point [8070327225:26829571070240:729], height = 15.5585820191474
II. Points on phi(E') mod 2E
Point [-3770887884120:-15492435634205439:314432000],
height = 13.0562446525709
```

From these, `mwrank` computes the following generators of $E(\mathbb{Q})/2E(\mathbb{Q})$:

```
Generator 1 is [53767647152352:117078959809329057:11239424]; height
15.1743661746344
Generator 2 is [8070327225:26829571070240:729]; height 15.5585820191474
Generator 3 is [-3770887884120:-15492435634205439:314432000]; height
13.0562446525709
```

Determining $E(\mathbb{Q})/2E(\mathbb{Q})$ is important for two reasons. First, $\text{rank}(E(\mathbb{Q})/2E(\mathbb{Q})) = r + t$ so that knowing $r + t$ allows us to find $\text{rank}(E(\mathbb{Q})) = r$, since t is computable. This tells us how many independent points of $E(\mathbb{Q})$ we are looking for in our search. Second, as we will see in Chapter 2, Section 3, the coset representatives of $2E(\mathbb{Q})$ in $E(\mathbb{Q})$ will play an important role in bounding the search area for points on $E(\mathbb{Q})$. We will continue this example then.

7.2. General strategies. We generalize the second descent for $m = 2$ to an arbitrary m^n=descent as follows:

For any integer $n \geq 1$, we obtain the following commutative diagram:

$$\begin{array}{ccccccccc}
0 & \longrightarrow & \frac{E(K)}{m^n E(K)} & \longrightarrow & S^{m^n}(E/K) & \longrightarrow & Ш(E/K)[m^n] & \longrightarrow & 0 \\
 & & \downarrow & & \downarrow & & \downarrow{\scriptstyle m^{n-1}} & & \\
0 & \longrightarrow & \frac{E(K)}{mE(K)} & \longrightarrow & S^{m}(E/K) & \longrightarrow & Ш(E/K)[m] & \longrightarrow & 0
\end{array}$$

From this diagram we obtain another short exact sequence:

$$0 \to E(K)/mE(K) \to S^{m,n}(E/K) \to m^{n-1}Ш(E/K)[m^n] \to 0$$

where $S^{m,n}(E/K)$ is the image of $S^{m^n}(E/K)$ in $S^m(E/K)$. The following proposition formalizes a strategy we can use to compute $E(K)/mE(K)$ when Ш$(E/K)[m]$ is nontrivial.

Proposition 1.7.1. *For the above set up, we have*

$$E(K)/mE(K) \subseteq \cap_n S^{m,n}(E/K),$$

and we have equality if and only if the Tate-Shafarevich group contains no element divisible by all powers of m.

PROOF. (*Of Proposition 1.7.1*) First, $E(K)/mE(K) \subseteq S^m(E/K)$ so it follows by definition that $E(K)/mE(K) \subseteq \cap_n S^{m,n}(E/K)$. To show the reverse containment, let $\alpha \in \cap_n S^{m,n}(E/K)$. By definition, for each n, there exists some $\alpha_n \in S^{m^n}(E/K)$ which maps to α. Now suppose β_n is the image of α_n in $W(E/K)[m^n]$. Then $m^{n-1}\beta_n = \beta_1$ for all n which implies that β_1 is divisible by all powers of m. If Ш(E/K) contains no such nontrivial elements, then α must be in the kernel so that by exactness, α is in the image of $E(K)/mE(K)$ in $S^m(E/K)$. □

The problem, however, lies with the m^n-torsion of the Tate-Shafarevich group which may indeed contain such an element. In this case, the strategy we outlined may not terminate - that is, there is no way to conclude that the α from the proof is in the kernel of 5 and thus comes from the Weak Mordell-Weil group. This leaves open the possibility that α is associated with a nontrivial element of the Tate-Shafarevich group instead.

CHAPTER 2

Descent 2: From $E(K)/mE(K)$ to $E(K)$

1. Heights on an Elliptic Curve

The second part of the proof of the Mordell-Weil theorem involves using the finiteness of $E(K)/mE(K)$ to show that $E(K)$ is finitely generated. To do this, we first define a notion of size on the points of $E(K)$ and show that any point $P \in E(K)$ is generated by a set of elements of common bounded size. We begin by defining the notion of a height on $E(\mathbb{Q})$ and then generalize its properties to an arbitrary number field.

EXAMPLE 2.1.1. A height function on $E(\mathbb{Q})$ is relatively straightforward to define. First define

$$H(m/n) = \max\{|m|, |n|\}$$

where $|\cdot|$ is the standard absolute value and m/n is in reduced form. This can be almost trivially extended to multiple dimensions by defining the height of a point as the height of its first coordinate. Thus, given a point $P = (x, y) \in E(\mathbb{Q})$, define $H(P) = H(x)$. This definition is sufficient to complete the proof of the Mordell-Weil theorem over $\mathbb{Q}$.

To extend this definition to K, consider the set M_K of valuations on K. For each $v \in M_K$, define the *local degree at* v to be $n_v = [K_v : \mathbb{Q}_v]$. We define the (naive) height as follows:

DEFINITION 2.1.2. For a point $P = [x_0 : \cdots : x_N]$ in K projective space with homogenous coordinates, the *naive height of* P *relative to* K is:

$$H_K(P) = \prod_{v \in M_K} \max\{v(x_0), \ldots, v(x_N)\}^{n_v}.$$

The naive height satisfies some nice properties:

PROPOSITION 2.1.3. .

(1) *The height function is independent of our choice of homogenous coordinates.*
(2) $H_K(P) \geq 1$.
(3) *The height of points are invariant under Galois automorphisms. That is, for* $P \in \mathbb{P}^N(\bar{K})$ *and* $\sigma \in G_{\bar{K}/K}$, $H_L(P) = H_L(P^\sigma)$ *for all finite extensions* L/K.
(4) *For* L/K *a finite extension,*

$$H_L(P) = H_K(P)^{[L:K]}.$$

PROOF. (1) Another choice of coordinates must have the form $[cx_0 : \cdots : cx_N]$. The height of this point is:

$$\begin{aligned} H_K(P) &= \prod_{v \in M_K} \max\{v(cx_0), \ldots, v(cx_N)\}^{n_v} = \prod_{v \in M_K} v(c)^{n_v} \max\{v(x_0), \ldots, v(x_N)\}^{n_v} \\ &= \prod_{v \in M_K} v(c)^{n_v} \prod_{v \in M_K} \max\{v(x_0), \ldots, v(x_N)\}^{n_v} \end{aligned}$$

By the Product Formula, $\prod_{v\in M_K} v(c)^{n_v} = 1$, so we get the same height back.

(2) Using our result in (1), we can simply pick homogenous coordinates such that one coordinate is 1. Then $H_K(P) \geq 1$.

(3) Let L/K be the finite field extension that contains P. Then σ gives an isomorphism from L to L^σ, a bijective mapping from M_L to M_{L^σ}, and thus an isomorphism from L_v to $L^\sigma_{v^\sigma}$ such that $v^\sigma(x^\sigma) = v(x)$. Thus we compute:

$$\begin{aligned} H_{L^\sigma}(P^\sigma) &= \prod_{w\in M_{L^\sigma}} \max\{w(x_i^\sigma)\}^{n_w} = \prod_{v\in M_L} \max\{v^\sigma(x_i^\sigma)\}^{n_{v^\sigma}} \\ &= \prod_{v\in M_L} \max\{(x_i)\}^{n_v} = H_L(P) \end{aligned}$$

(4) This follows from a simple computation:

$$\begin{aligned} H_L(P) &= \prod_{w\in M_L} \max\{w(x_i)\}^{n_w} = \prod_{v\in M_K} \prod_{w:v\in M_L} \max\{v(x_i)\}^{n_w} \text{ since } x_i \in K \\ &= \prod_{v\in M_K} \max\{v(x_i)\}^{[L:K]n_v} = H_K(P)^{[L:K]} \end{aligned}$$

□

The height we just defined is dependent on our choice of K. We also define an *absolute naive height*,

$$H(P) = H_K(P)^{1/[K:\mathbb{Q}]}$$

where K is any finite extension that contains P. Note that this definition does not depend on our choice of K. To see this, suppose K and L both contain P and $[L:K] \geq 1$. Then:

$$\begin{aligned} H(P) &= H_L(P)^{1/[L:\mathbb{Q}]} \\ &= H_K(P)^{[L:K]/[L:\mathbb{Q}]} \text{ by Proposition 2.1.3 (4).} \\ &= H_K(P)^{1/[K:\mathbb{Q}]} \end{aligned}$$

Now we would like to extend our general concept of height to height on an elliptic curve. In the case of $\mathbb{Q}$, we defined a height on $E(\mathbb{Q})$ by taking the height of the first coordinate. We follow the same procedure here.

Consider the map $f : E \to \mathbb{P}^1$ defined by $f(P) = [1, 0]$ if P is a pole and $f(P) = [f(P), 1]$ if not. Now we define the height of a point P on E as

$$H_f(P) = H(f(P)).$$

We also define the logarithmic height as $h(P) = \log H(P)$. From Proposition 2.1.3, it follows that $h(P) \geq 0$ for all P. For the remainder of this thesis, we take $f = x$, the function that picks out the first coordinate and write $h_x(P)$.

First, we state some technical properties of heights. The proofs of these propositions are basic, but rather dry, so the reader is referred to [**36**] Chap. 8 for proofs.

Proposition 2.1.4. *Let*

$$F : \mathbb{P}^N \to \mathbb{P}^M$$

be a morphism of degree d. Then there exist constants C_1 and C_2, which depend on F, such that for all points $P \in \mathbb{P}^N(\bar{\mathbb{Q}})$,

$$C_1 H(P)^d \leq H(F(P)) \leq C_2 H(P)^d.$$

This is equivalent to saying that there exists some constant C^* such that

$$h(F(P)) = dh(P) + C^*$$

in terms of the additive height.

If F is a one dimensional polynomial morphism, then:

PROPOSITION 2.1.5. *For*

$$f(T) = T^d + a_{d-1}T^{d-1} + \cdots + a_0 = (T - \alpha_1)\cdots(T - \alpha_d) \in \bar{\mathbb{Q}},$$

we have

$$2^d \prod_{j=1}^{d} H(\alpha_j) \leq H([a_0, \ldots, a_d]) \leq 2^{d-1} \prod_{j=1}^{d} H(\alpha_j).$$

2. Two Lemmas, Two Corollaries

Now we prove the specific results that prepare us for Mordell and Weil's descent procedure.

LEMMA 2.2.1. *Let E/K be an elliptic curve. Then, for any constant C, the set*

$$\{P \in E(K) : h_x(P) \leq C\}$$

is finite.

PROOF. x gives a 1-1 map of the set $\{P \in E(K) : h_x(P) \leq C\}$ to the set

$$\mathcal{A} = \{\alpha \in \mathbb{P}^1(K) : H(\alpha) \leq e^C\}$$

by definition. Thus, it suffices to prove this set is finite.

Let $e = [\mathbb{Q}(\alpha) : \mathbb{Q}]$. We know that e is finite and moreover, $e \leq d$ where $d = [K : \mathbb{Q}]$. Let

$$g(T) = (T - \alpha_1)\cdots(T - \alpha_e) = T^e + a_{e-1}T^{e-1} + \cdots + a_1$$

be the minimal polynomial of α over $\mathbb{Q}$ where $\alpha \in \mathcal{A}$. By Proposition 2.1.5,

$$H([1, a_1, \ldots, a_e]) \leq 2^{e-1} \prod_{j=1}^{e} H(\alpha_j)$$

Moreover, because height is invariant under Galois automorphisms, we have

$$\begin{aligned} H([1, a_1, \ldots, a_e]) &\leq 2^{e-1} H(\alpha)^e \\ &\leq (2e^C)^d \end{aligned}$$

because we assumed $H(\alpha) \leq e^C$ and $e \leq d$. Since the a_i are in $\mathbb{Q}$, there can only be finitely many a_i such that $H([1, a_1, \ldots, a_e])$ will be less that a given constant. As such, there must only be finitely many minimal polynomials $g(T)$ of α's in the set $\{\alpha \in \mathbb{P}^1(K) : H(\alpha) \leq e^C\}$. Since each has at most d roots, this is equivalent to saying that the set of such α is finite. □

LEMMA 2.2.2. *Let E/K be an elliptic curve. Then, for all $P, Q \in E$,*

$$h_x(P + Q) + h_x(P - Q) = 2h_x(P) + 2h_x(Q) + C$$

where C is a constant that depends on E, but is independent of the choice of P or Q.

PROOF. The lemma is clear in the case of P or Q equals $\mathcal{O}$. Now, consider $P = (x_1, y_1)$ and $Q = (x_2, y_2)$. Let:

$$x(P) = [x_1, 1], \qquad x(Q) = [x_2, 1],$$
$$x(P+Q) = [x_3, 1], \qquad x(P-Q) = [x_4, 1].$$

Using the addition formula from Section 1 of the introduction, we find that:

$$x_3 + x_4 = \frac{2(x_1+x_2)(a+x_1x_2)+4b}{(x_1+x_2)^2 - 4x_1x_2},$$

$$x_3x_4 = \frac{(x_1x_2-a)^2 - 4b(x_1+x_2)}{(x_1+x_2)^2 - 4x_1x_2}.$$

Now define a map $g : \mathbb{P}^2 \to \mathbb{P}^2$ by

$$g([t,u,v]) = [u^2 - 4tv, 2u(at+v) + 4bt^2, (v-at)^2 - 4btu].$$

We combine g with several other maps to form the diagram:

$$\begin{array}{ccc} E \times E & \xrightarrow{G} & E \times E \\ \downarrow\phi & & \downarrow\phi \\ \mathbb{P}^1 \times \mathbb{P}^1 & & \mathbb{P}^1 \times \mathbb{P}^1 \\ \downarrow\psi & & \downarrow\psi \\ \mathbb{P}^2 & \xrightarrow{g} & \mathbb{P}^2 \end{array}$$

$g(P,Q) = (P+Q, P-Q)$, $\phi(P,Q) = (x(P), x(Q))$, $\psi([x_1,y_1],[x_2,y_2]) = [y_1y_2, x_1y_2+x_2y_1, x_1x_2]$.

For ease of notation, we write $\psi \circ \phi = \theta$. This diagram turns out to be commutative. To verify this, we need only plug in and check; going down then right, we obtain:

$$g(\theta(P,Q)) = g(\psi([x_1,1],[x_2,1])) = g(1, x_1+x_2, x_1x_2) =$$
$$[(x_1+x_2)^2 - 4x_1x_2, 2(x_1+x_2)(a+x_1x_2)+4b, (x_1x_2-a)^2 - 4b(x_1+x_2)]$$

Going right then down gives

$$\theta(G(P,Q)) = \theta(P+Q, P-Q) = \psi([x_3,1],[x_4,1]) = [1, x_3+x_4, x_3x_4]$$

which is just the previous result scaled by a factor of $(x_1+x_2)^2 - 4x_1x_2$, which does not matter in projective space.

We claim that g is a morphism. By Proposition 2.1.4, this means that

$$h(\theta(P+Q, P-Q)) = h(g(\theta(P,Q))) = 2h(\theta(P,Q)) + C \tag{13}$$

because g is degree two. Further, assuming that for all $R_1, R_2 \in E$, there is an identity

$$h(\theta(R_1,R_2)) = h_x(R_1) + h_x(R_2) + C, \tag{14}$$

we apply it to both sides of Equation 13 to obtain the desired result.

Thus it remains to prove that g is a morphism and that Equation (14) holds. To show that g is a morphism, we must show that $u^2 - 4tv$, $2u(at+v) + 4bt^2$, and $(v-at)^2 - 4btu$ have no common nontrivial zeros. Suppose then, that $g[t,u,v] = [0,0,0]$. If $t = 0$, we can solve to see that u and v must be zero as well. Thus, we assume that $t \neq 0$ and define $w = u/2t$. Our first equation can be written as

$$u^2 - 4tv = w^2v/t = 0.$$

It suffices to show that the remaining two equations (in terms of w):

$$m(w) = 4w(a+w^2) + 4b = 4w^3 + 4aw + 4b = 0, \text{ and}$$

$$n(w) = (w^2 - a)^2 - 8bw = w^4 - 2aw^2 - 8bw + a^2 = 0$$

have no common zeros. It is computationally straightforward to verify that

$$(12X^2 + 16a)n(X) - (3X^3 - 5aX - 27b)m(X) = 4(4a^3 + 27b^2)$$

holds. This cannot be zero because the discriminant of the Weierstrass equation is nonzero by definition of an elliptic curve. Thus there can be no value of w for which $m(w) = n(w) = 0$ because otherwise the above equation would be zero. This proves that g is a morphism.

We now prove Equation (14) and this concludes the proof. In the case that R_1 or R_2 equals the origin, Equation (14) clearly holds. So now we assume that neither does and write

$$x(R_1) = [a_1, 1] \text{ and } x(R_2) = [a_2, 1].$$

Thus,

$$h(\theta(R_1, R_2)) = h([1, a_1 + a_2, a_1a_2]) \text{ and } h_x(R_1) + h_x(R_2) = h(a_1) + h(a_2).$$

Applying Proposition 2.1.5 to $f(T) = T^2 + (a_1 + a_2)T + a_1a_2$ gives

$$h(a_1) + h(a_2) + C_1 \leq h([1, a_1 + a_2, a_1a_2]) \leq h(a_1) + h(a_2) + C_1.$$

This implies that

$$h_x(R_1) + h_x(R_2) + C = h(\sigma(R_1, R_2))$$

which proves Equation (14) and thus this lemma. □

Corollary 2.2.3. *Fix an arbitrary $Q \in E$. For all $P \in E$,*

$$h_x(P + Q) \leq 2h_x(P) + C_1$$

where C_1 depends on E and Q, but not on P.

Proof. We already know that

$$h_x(P + Q) = 2h_x(P) + 2h_x(Q) - h_x(P - Q) + C.$$

Since $h_x(P - Q) \geq 0$ by Proposition 2.1.3, we have

$$h_x(P + Q) \leq 2h_x(P) + C_1$$

where $C_1 = 2h_x(Q) + C$. □

Corollary 2.2.4. *Fix an arbitrary $m \in \mathbb{Z}$. For all $P \in E(\bar{K})$,*

$$h_x([m]P) = m^2h_x(P) + C_2$$

where C_2 depends on E and m, but not on P.

Proof. Since x is an even function, we need only consider $m \geq 0$. We proceed by induction. For $m = 0$ or 1, we can read the result right off the lemma. Assuming true for $m-1$ and $m-2$, we plug in $Q = [m-1]P$ in the lemma and get:

$$\begin{aligned} h_x([m]P) &= 2h_x(P) + 2h_x([m-1]P) - h_x([m-2]P) + C \\ &= 2h_x(P) + 2(m-1)^2)h_x(P) - (m-2)^2h_x(P) + C \\ &= m^2h_x(P) + C \end{aligned}$$

□

This tells us that the naive height $h(P)$ is almost a quadratic form. Although only need the naive height to prove the Mordell-Weil theorem, we modify $h(P)$ to get nicer properties.

DEFINITION 2.2.5. The *canonical, or Neron-Tate, height* on E/K is the function

$$\hat{h} : E \to \mathbb{R}$$

given by

$$\hat{h}(P) = \lim_{N\to\infty} 4^{-N} h_x(2^N P).$$

It follows quickly from the lemmas we have proven that the canonical height satisfies the following relevant properties:

PROPOSITION 2.2.6. *Let E/K be an elliptic curve and $\hat{h}$ the canonical height on E. Then:*

(1) *For any constant C, the set*

$$\{P \in E(K) : \hat{h}(P) \leq C\}$$

is finite.

(2) *For all $P, Q \in E$,*

$$\hat{h}(P+Q) + \hat{h}(P-Q) = 2\hat{h}(P) + 2\hat{h}(Q).$$

(3) *For all $P \in E$,*

$$\hat{h}([m]P) = m^2 \hat{h}(P).$$

PROOF. (i) This follows directly from Lemma 2.2.1.

(ii) From Lemma 2.2.2,

$$h_x(P+Q) + h_x(P-Q) = 2h_x(P) + 2h_x(Q) + C.$$

Replacing P and Q by $[2^N]P$ and $[2^N]Q$ and multiplying through by 4^{-N}, we obtain:

$$4^{-N} h_x([2^N](P+Q)) + 4^{-N} h_x([2^N](P-Q)) = 4^{-N} 2h_x([2^N]P) + 4^{-N} 2h_x([2^N]Q) + 4^{-N} C.$$

Taking the limit, the constant disappears and we obtain

$$\hat{h}(P+Q) + \hat{h}(P-Q) = 2\hat{h}(P) + 2\hat{h}(Q).$$

(iii) From Corollary 2.2.4,

$$h_x([m]P) = m^2 h_x(P) + C_2.$$

Replace P by $[2^N]P$, multiply through by 4^{-N}, and take the limit $N \to \infty$. □

Moreover, we can bound the difference between the naive and canonical heights. There is a procedure due to Siksek [**35**] that gives a good lower bound for $\hat{h}(P) - h_x(P)$.

3. The Descent

Using the properties of the canonical height discussed in the previous section, we now finish the proof of the Mordell-Weil theorem. We had from Chapter 1 that the weak Mordell-Weil group $E(K)/mE(K)$ is finite. We use descent to show that we need only add a finite number of elements to the finite set of coset representatives for $mE(K)$ in $E(K)$ in order to obtain a generating set for $E(K)$.

THEOREM 2.3.1. *Let $B > 0$ be a constant such that the set*

$$S = \{P \in E(K) : \hat{h}(P) \leq B\}$$

contains a set of coset representatives of $E(K)/mE(K)$. The S generates $E(K)$.

PROOF. Suppose not. Then there exists $Q \in E(K) - \langle S \rangle$ with $\hat{h}(Q)$ minimal. This exists because $\hat{h}$ takes on a discrete set of values. By assumption, there exists $P \in S$ and $R \in E(K)$ such that $Q = P + mR$. Since $Q \notin A$, R must not be either. Thus $\hat{h}(R) \geq \hat{h}(Q)$ by construction. Now:

$$\begin{aligned}
\hat{h}(P) &= \frac{1}{2}\big(\hat{h}(Q+P) + \hat{h}(Q-P)\big) - \hat{h}(Q) \\
&\geq \frac{1}{2}\hat{h}(mR) - \hat{h}(Q) \\
&= \frac{m^2}{2}\hat{h}(R) - \hat{h}(Q) \\
&\geq \hat{h}(Q) \qquad \text{since } \hat{h}(R) \geq \hat{h}(Q) \text{ and } m \geq 2 \\
&> B
\end{aligned}$$

But this contradicts the fact that $P \in S$. □

As a result, if we find coset representatives for $E(K)/mE(K)$, Theorem 2.3.1 tells us that the set of generators for $E(K)$ is contained in the set of points bounded by the maximum canonical height of the coset representatives. In practice, it is often easy (as in the case of $K = \mathbb{Q}$) to bound the naive height of a set of generators and from there, to find points on $E(\mathbb{Q})$ with bounded x-coordinate using Siksek's bound on the difference between the naive and canonical heights.[1]

Now we return to our `mwrank` calculation on the curve:

$$E : y^2 = x^3 - 450079^2 x$$

from Chapter 1. `mwrank` sorts the coset representatives into height order in order to get a bound on the heights of points it should search for.

```
Computing full set of 8 coset representatives for 2E(Q) in E(Q) (modulo
torsion), and sorting into height order....done.
```

and eventually we find the following generators of $E(\mathbb{Q})$:

```
Generator 1 is [53767647152352:117078959809329057:11239424];
height 15.1743661746344
Generator 2 is [8070327225:26829571070240:729]; height 15.5585820191474
Generator 3 is [-3770887884120:-15492435634205439:314432000];
height 13.0562446525709
```

Because there is no saturation, these are the same generators as for the weak Mordell-Weil group.

[1] For details, see [**38**] Chap. 3, Sec. 4.

CHAPTER 3

Descent 1: Fermat's Infinite Descent

We have seen several descent methods in the previous chapters. The term "descent" originates with Fermat and was used to describe a method of finding roots of equations intended to contrast with Diophantus' method of "ascent."[1] Even though Fermat was interested in the study of Diophantine equations, he was critical of the fact that number theory at his time still relied so heavily on such ancient methods. Thus, Fermat's research and correspondence were largely directed at developing and sustaining a new program of research in arithmetic rooted in classical problems, but equipped with modern methods. To arouse interest in new proof methods, Fermat corresponded with many prominent mathematicians and delegated challenge problems and research tasks among them, as if he were coordinating a research effort.[2] For example, in a letter to Pierre de Carcavi, but addressed to Christiaan Huygens, Fermat instructed Huygens to tell another mathematician, Bernard Frenicle, to prove that "no cube can be the sum of two cubes" and expressed his desire that Pascal and Roberval continue research into the "mystery of my [Fermat's] method...at my indication."[3]

This particular letter stands out in Fermat's correspondence because it has been stressed as one of the few instances in which Fermat actually attempted to provide some indication of his proof methods.[4] Written in 1659, less than six years before his death, Fermat's letter, entitled the "Relation of New Discoveries in the Science of Numbers," has been regarded as Fermat's "swan song as a number theorist" where he outlined the methods and philosophy of the arithmetic program he hoped would be continued after his death.[5] In it, Fermat claimed to have discovered a "unique method for arriving at [the solution to difficult propositions]."[6] As an example, he "proved" Proposition 0.0.1: that there is no integral right triangle whose area is a square.[7] However, in his argument, Fermat gave only a rudimentary outline of descent:

> *...if one has* [such a] *triangle, then one will have another triangle smaller than that one which has the same property. If one has a second less than the first, which has the same property, one has, by the same reasoning, a third, less than the second, with the same property, and finally, a fourth, a fifth, etc., descending to infinity...One concludes that it is impossible to have a right* [integral] *triangle with area a square.*[8]

[1] For a discussion of Diophantus' use of ascent, see [**42**] Chap 1., Sec. 10. An interesting observation is that Diophantus never applies it recursively.

[2] Fermat's number theory correspondents, include Descartes, Wallis, Mersenne, Pascal, and Roberval.

[3] [**13**] Vol. 2, pg 433.

[4] See [**22**] pg. 350.

[5] [**22**] pg. 288.

[6] [**13**] Vol 2, pg. 431.

[7] This problem was sent as a challenge to Sainte-Croix via Mersenne as early as 1636, indicating that Fermat was at least partially aware of its impossibility and perhaps even the proof as early as then. See [**13**] Vol. 2, pg. 65.

[8] [**13**] Vol 2, pg. 431-2.

Although Fermat was hesitant to share his insights in his correspondence with Huygens and others, he did provide more detailed arguments in his personal notes and marginalia, which were published posthumously by his son in 1672. In these notes, it turns out, Fermat did provide a somewhat complete proof of Proposition 0.0.1:

> *If the area of the triangle were a square, then there would be two quadro-quadrates* [fourth-powers] *of which the difference would be a square; it follows that one would also have two squares of which the sum and the difference would be squares. Consequently, one would have a square number, the sum of the square, and of the double of the square with the condition that the sum of the two squares who serve to compose it is also a square. But if the square number is the sum of a square and of the double of the square, its root is also the sum of a square and double of a square, which I can prove without difficulty. One will conclude that this root is the sum of two legs of a right triangle, of which one of the squares will form the base and the double of the other square will form the altitude.*
>
> *This right triangle will therefore be formed by two square numbers, of which the sum and the difference will be squares. But we will prove that the sum of two squares is smaller than that of the first two, of which one also supposes that the sum and difference are squares. Therefore, if one gives two squares whose sum and difference are squares, one gives likewise integers, two squares, with the same property and whose sum is smaller.*
>
> *By the same reasoning, one will have, then, another sum smaller than that deduced from the first, and continuing indefinitely one will always find smaller and smaller whole numbers satisfying the same conditions. But this is impossible, for having given a whole number, it is not possible to have an infinity of whole numbers which are smaller.*[9]

In more modern language and notation, Fermat's proof says:

PROOF. *(Of Proposition 0.0.1)* Any right triangle gives a primitive right triangle with sides $(2pq, p^2 - q^2, p^2 + q^2)$ where p, q are mutually prime, $p > q$, and $p - q$ is odd. If the area of the triangle, $A = pq(p - q)(p + q)$, is a square and all the factors are mutually prime, then each factor must itself be a square. Thus we can write

$$x^2 = p \qquad y^2 = q \qquad u^2 = p + q \qquad v^2 = p - q$$

where u, v are odd and mutually prime. This gives $x^4 - y^4 = p^2 - q^2 = z^2$, where $z = uv$, ("*If the area of the triangle were a square, then there would be two quadroquadrates* [fourth-powers] *of which the difference would be a square.*") We factor the last equation to obtain:

$$(x^2 - y^2)(x^2 + y^2) = z^2.$$

Since x and y are mutually prime, it follows that

$$x^2 - y^2 = r^2 \qquad x^2 + y^2 = s^2$$

("*it follows that one would also have two squares of which the sum and the difference would be squares*"). Rearranging, we have $x^2 = r^2 + y^2$ and so substituting, gives $s^2 = r^2 + 2y^2$ ("*Consequently, one would have a square number, the sum of the square, and of the double of the square with the condition that the sum of the two squares who serve to compose it is also a*

[9]Translation is the author's own, based off [**13**] Vol. 1.

square").

However, this means that s must have the form $s = k^2 + 2m^2$ ("*But if the square number is the sum of a square and of the double of the square, its root is also the sum of a square and double of a square, which I can prove without difficulty*"). With this, we can write:

$$s^2 = (k^2 + 2m^2)^2 = (k^2 - 2m^2)^2 + 2(4k^2m^2) = r^2 + 2y^2 \tag{15}$$

$$s^2 = (k^2 + 2m^2)^2 = k^4 + 4m^4 + 4k^2m^2 = x^2 + y^2 \tag{16}$$

Using Equation (15), we write

$$y^2 = 4k^2m^2$$

and from (16), plugging in $y^2 = 4k^2m^2$ gives us

$$x^2 = k^4 + 4m^4.$$

This last equation tells us that k^2 and $2m^2$ are the sides of a right triangle with hypotenuse x and area k^2m^2, a square ("*One will conclude that this root is the sum of two legs of a right triangle, of which one of the squares will form the base and the double of the other square will form the altitude*"). Therefore, we can apply the same argument as above to the new triangle and derive a continual series of right triangles with areas a square. But the hypotenuse of our original triangle was $p^2+q^2 = x^2+y^2$ which is greater than x since we are working with integers. Iterating this method, we obtain ever smaller integral right triangles, but this contradicts the well-ordering principle. □

Although Fermat's proof does not deal with elliptic curves explicitly, we can consider this problem in the context of elliptic curves using the transformation in Example 0.2.4. From this perspective, Proposition 0.0.1 says:

COROLLARY 3.0.2. *Since all nontrivial points on the curve*

$$E : y^2 = -x^3 + x = x(1 - x)(1 + x)$$

represent right triangles with square areas, the set of rational points on E consists only of $(0, 0)$, $(1, 0)$, $(-1, 0)$, and the point at infinity.

This argument shows that the group $E(K)$ contains no nontrivial solutions. In particular, the elements of $E(K)$ that do exist are all torsion on the curve, implying that $E(K)$ has rank 0. Thus, we recover the same result about the rank of E from our example of 2-descent in Chap. 1 Sec. 6.

Conclusion

At the end of his 1659 letter to Huygens, Fermat writes:

> *Such is the summary of my thoughts on the subject of numbers....and perhaps posterity will thank me for having shown that the ancients did not know everything, and this relation may come to be regarded by my successors as a "passing of the torch."*[10]

While Fermat wished for Huygens and others to immediately take up the problems he proposed, he would actually have to wait almost a century, for Euler, to have such a successor.[11] From Euler, however, Fermat's number theory became the focus of sustained mathematical research and his method of infinite descent was cited frequently in the 19th century mathematical literature.[12] Thus, it is not surprising that Mordell credits his use of descent in the proof of the Finite Basis to Fermat's method of infinite descent:

> *I shall now prove that if* $[E(K)]$ *has an infinite number of* [elements]*, then the method of infinite descent applies, that is to say, all the solutions can be expressed rationally in terms of a finite number by means of the classic method.*

In his proof, Mordell assumes that $E(\mathbb{Q})$ is not finitely generated and supposes the existence of a smallest point on E satisfying some conditions. He then uses descent to find an even smaller such point, thus deriving a contradiction. In comparison, Fermat's proof of Proposition 0.0.1, when viewed in the context of elliptic curves, takes a point on $E : y^2 = -x^3 + x$, which represents a right triangle with a square area, and derives from it another point on E with smaller coordinates. Repeating this process indefinitely, the areas of these triangles form an infinite sequence of decreasing positive integers, contradicting the well-ordering principle. Although Fermat's proof seems slightly different, the fact that the areas of his triangles are natural numbers means that he could have every supposed the existence of a smallest such triangle and then descended to a contradiction in the same way Mordell does using height functions.

Not surprisingly, the influence of Fermat on Mordell's descent procedure is commonly acknowledged. Indeed, not only was this relationship explicitly acknowledged by Mordell in his original proof, it has also been preserved in contemporary textbooks as well as the the secondary literature; in writing this thesis, almost every text I encountered that included a proof or discussion of the Mordell-Weil theorem also included a discussion of Mordell's use of infinite descent.[13] Indeed, the influence of Fermat on Mordell seems quite natural after reading all these accounts; Cassels, writing in 1986, even adds that Mordell's use of infinite descent "must have been Pavlovian."[14]

[10] [**13**] Vol. 2. pg 436.
[11] [**42**] pg. 120.
[12] [**7**] pg 31.
[13] For example, [**4**], [**7**], [**17**], [**27**], [**42**].
[14] [**7**] pg 37.

However, there are fewer methodological similarities between the computational descents of Chapter 1 and Fermat's infinite descent. Here, it seems that the term "descent," when applied to computing $E(K)$, refers not to a relationship with Fermat, but to the process of "descending" from rational points on homogenous spaces C down to possible generators of $E(K)$. Indeed, while, in computing $E(K)$, "descent" is accomplished by using the algebraic structure on and between E and C to associate points on E with points of smaller height on C, descent for Fermat limits itself to finding smaller points on the same curve without any explicit reliance on the algebraic structure of the curve.

Yet, despite the dissimilarities in method, many mathematicians nonetheless claim a deep relationship between computational descent and Fermat's infinite descent. For example, the complete 2-descent we performed in Chapter 1, Section 6 shows that the curve $E : y^2 = x - x^3$ has rank 0. This is equivalent, by the transformation in Example 0.2.4, to Fermat's proof that no integral right triangle can have area a square. More generally, Poonen writes: "the process of computing the Selmer group and using it to bound $E(K)/mE(K)$ is known as descent because as a very special case it includes Fermats infinite descent method for solving Diophantine equations."[15] Further, Cassels comments that Fermat's infinite descent is "almost completely subsumed in the general theory that arose of the Mordell-Weil theorem."[16]

These remarks indicate that the three descents we discussed can be reconciled into a single cohesive intellectual method aimed at studying the rational points on elliptic curves, and more generally, on abelian varieties.[17] In one way, this represents a remarkable feat, one that supports the claim that mathematic is one of the few truly cumulative disciplines; using modern tools such as the theory of cohomology, we have been able to integrate Fermat's method of finding smaller integral points on curves with a much broader theory of arithmetic geometry equipped with the theoretical tools that Jacobi, Hilbert, Hurwitz, and Poincaré found so lacking in the classical study of Diophantine equations.

In another way, however, the intellectual cohesion of the descents we discuss is only an interpretation made in hindsight, formulated by applying modern techniques - from group laws to cohomology - to problems that were originally conceived of much differently; Fermat, after all, did not have a group notion, nor even a well developed coordinate geometry. As such, the fact that we think of modern computational descent as a generalization of Fermat's classical infinite descent does not necessarily imply that Fermat's infinite descent influenced the development of computational descent (though it very well might have); it could also be indicative of the contextual amnesia that accompanies progress toward more powerful tools and theories. From this perspective, the case of descent indicates that both progress and amnesia play a role in preserving the cumulative nature of the move from classical Diophantine analysis to modern arithmetic geometry.

[15] [**27**] pg. 8.

[16] [**7**] pg. 32.

[17] And schemes as well? I lack the background to say.

APPENDIX A

Cohomology of Groups

Chapter 1 makes frequent use of tools from Galois cohomology. In this section, we provide a quick introduction, assuming results from basic commutative algebra. For more background on commutative algebra, see [**1**], [**24**], or [**11**]. The discussion in this section is adopted from [**36**] Appendix B and [**5**] Chap 4.

We begin by defining the cohomology of finite groups. Let G be a finite group and M an abelian group on which G acts by $m \mapsto m^\sigma$. We say that M is a G-module if the action of G on M satisfies:

$$m^1 = m \qquad (m + m')^\sigma = m^\sigma + m'^\sigma \qquad (m^\sigma)^\tau = m^{\sigma\tau}$$

for all $m, m' \in M$ and $\sigma, \tau \in G$. We can think of the action of G on M as giving a map $G \to \mathrm{Aut}(M)$, the group of group automorphisms of M.

Example A.0.3. If L is a Galois extension of K, and $G = \mathrm{Gal}(L/K)$, then both L (as an additive abelian group) and $E(L)$ are G-modules. K is, trivially, as well.

Now, given two G-modules, M and N, we consider maps between them.

Definition A.0.4. A map $\phi : M \to N$ is a *G-module homomorphism* if

$$\phi(m^\sigma) = \phi(m)^\sigma$$

for all $m \in M$ and $\sigma \in G$.

Now we are ready to define the 0^{th} and 1^{st} cohomology groups, $H^0(G, M)$ and $H^1(G, M)$, respectively. This is all we will need for the proof of Mordell-Weil and our discussion of its computation.

Definition A.0.5. We define the 0^{th} *cohomology group* of the G-module M, denoted by $H^0(G, M)$, as:

$$H^0(G, M) = \{m \in M : m^\sigma = m \text{ for all } \sigma \in G\}$$

This is the largest submodule of M fixed by G. For example, if G were to act trivially on M as it does in the case of $G = \mathrm{Gal}(L/K)$ and $M = K$ above, then $H^0(G, M)$ would be all of M.

Continuing Example A.0.3, we have $H^0(G, L) = K$ and $H^0(G, E(L)) = E(K)$.

Associated to M as a G-module are several other objects: The *group of 1-cochains* is the collection of maps (not necessarily homomorphisms) $f : G \to M$. The *group of 1-cocycles*, also known as the group of *crossed-homomorphisms*, are those maps $f : G \to M$ such that

$$f_{\sigma\tau} = (f_\sigma)^\tau + f_\tau$$

where f_σ denotes the element of M obtained by applying the map f to $\sigma \in G$ and $(f_\sigma)^\tau$ denotes the action of τ on the element of M given by f_σ. We denote this group $Z^1(G, M)$.

For any $m \in M$, we construct a 1-cocycle by setting $f_\sigma = m^\sigma - m \in M$ for all $\sigma \in G$. To verify, notice that

$$f_{\sigma\tau} = m^{\sigma\tau} - m = (m^\sigma - m)^\tau + (m^\tau - m) = (f_\sigma)^\tau + f_\tau.$$

These are an important subgroup of $Z^1(G, M)$ known as the *1-coboundaries* or *principal crossed homomorphisms* and denoted $B^1(G, M)$. It is straightforward to verify that both $Z^1(G, M)$ and $B^1(G, M)$ are indeed groups.

DEFINITION A.0.6. We define the 1^{st} *cohomology group* of the G-module M to be:

$$H^1(G, M) = Z^1(G, M)/B^1(G, M)$$

$H^1(G, M)$ is the group of 1-cocycles modulo the equivalence relation that two cocycles are equivalent if their difference is of the form of a coboundary.

Continuing Example A.0.3, $G = \mathrm{Gal}(L/K)$ acts trivially on K so that the 1-cocycles are just the standard homorphisms and the 1-coboundaries are identically zero. Thus, $H^1(G, K) = \mathrm{Hom}(G, K)$. Finding $H^1(G, L)$ and $H^1(G, E(L)$ is more involved.

Now, suppose we have a short exact sequence of G-modules (recall that in an exact sequence, the image of the preceding map is the kernel of the next):

$$0 \to M \xrightarrow{\phi} N \xrightarrow{\psi} P \to 0$$

Restricting to H^0, the G-invariant submodules, preserves most of this exact sequence:

$$0 \to H^0(G, M) \to^{\phi} H^0(G, N) \to^{\psi} H^0(G, P)$$

but ψ may no longer be surjective because $H^0(G, P)$ may be quite large. Although we no longer have a short exact sequence, the move to cohomology groups gives rise to a long exact sequence as the next proposition describes.

PROPOSITION A.0.7. *For any short exact sequence of G-modules*

$$0 \to M \xrightarrow{\phi} N \xrightarrow{\psi} P \to 0,$$

there is a long exact sequence

$$0 \to H^0(G, M) \to H^0(G, N) \to H^0(G, P) \xrightarrow{\delta} H^1(G, M) \to H^1(G, N) \to H^1(G, P)$$

PROOF. We have already demonstrated how the short exact sequence can give rise to the initial sequence of 0^{th}-cohomology groups. The same logic holds for the sequences involving 1^{st} cohomology groups. To complete the proof, we define the bridging map δ as follows: Let $p \in H^0(G, P)$. There exists an $n \in N$ mapping to p by the surjectivity of the original map. Further, $n^\sigma - n \in M$ for all $\sigma \in G$. To see this, notice that $\psi(n)^\sigma = \psi(n)$ for all σ which we rewrite as $\psi(n^\sigma - n) \in \mathrm{Ker}(\psi)$. But since the map is exact, any element of $\mathrm{Ker}(\psi)$ is in the image of ϕ so we can think of $n^\sigma - n \in M$ since it comes from some element of M. The map $n_\sigma = \sigma \mapsto n^\sigma - n$ is a cocycle from $G \to M$ so that we define $\delta(p) = n_\sigma$. This map is well defined because if we chose a different n' in the preimage of p, the two maps would differ by a coboundary. □

Now, consider H a subgroup of G. Automatically, any G-module M can be thought of as an H-module by simply restricting to H. Similarly, cocycles from $G \to M$ can be restricted to H. Under these restrictions, cocycles and coboundaries are mapped to themselves. Putting this together, we get a restriction homomorphism

$$\mathrm{Res} : H^1(G, M) \to H^1(H, M)$$

of first cohomology groups.

Suppose H is a normal subgroup of G so that the quotient group G/H is defined. Then the submodule $H^0(H, M)$ consisting of elements fixed by H has the structure of a G/H-module since elements of H act as the identity. A cocycle $f : G/H \to H^0(H, M)$ gives rise to a cocycle from G to M by composition:

$$G \to G/H \to H^0(H, M) \subset M$$

This gives the *inflation homomorphism*:

$$\text{Inf} : H^1(G/H, M) \to H^1(G, M)$$

The sequence

$$H^1(G/H, M) \overset{\text{inf}}{\to} H^1(G, M) \overset{\text{res}}{\to} H^1(H, M) \to 0$$

is exact.

Now we turn to the case where our group G is a Galois group. Specifically, let K be a perfect field, one in which all algebraic extensions are separable, (of most interest to us, when K is a number field) and let $\bar{K}$ denote the algebraic closure of K.

Let $G = \text{Gal}(\bar{K}/K)$. G is equipped with a topology where the basis of open sets about 1 are the subgroups which fix intermediate fields L where $[L : K]$ is finite. Two elements of G are considered close if they agree on larger and larger fields. This is known as the *Krull Topology*.

We say that a G-module is *discrete* if, for any $\sigma \in G$, the action on M given by $\sigma : G \times M \to M$ is continuous for the discrete topology on M and the Krull topology on G. Since the any set is open in the discrete topology, this is equivalent to requiring that the stabilizer of M have finite index in $G_{\bar{K}/K}$ under the action of $G_{\bar{K}/K}$ on M.

For example, $\bar{K}$ is a discrete G-module because for all $x \in \bar{K}$, the extension $K(x)/K$ is finite so that the stabilizer of x will have finite index. The same logic shows that E is a discrete G-module.

We now refine the concepts introduced in the last section with $G = G_{\bar{K}/K}$ in mind. The 0^{th} cohomology group is defined exactly as it was before. For cocycles however, we need the added condition that the map $f : G \to M$ be continuous with respect to the Krull and discrete topologies, respectively. In this case, it is equivalent to saying that the preimage of every point in M contains a subgroup of finite index in G. Similarly, we must restrict the coboundaries, $\sigma \to m^\sigma - m$ for all $\sigma \in G$, to those which are continuous. However, the condition of continuity is implied by the discreteness of M as a G-module immediately from the definition.

We define the 1^{st} cohomology group once again as the quotient group of continuous cocycles mod the coboundaries. Using these definitions, the results from the finite case follow.

Bibliography

[1] Atiyah, M. and I.G. Macdonald. *Introduction to Commutative Algebra.* Reading MA: Addision-Wesley, 1969.

[2] Balakrishan, J. "4-descent on the elliptic curve $y^2 = x^3 + 7823$," *Harvard College, Freshman Seminar paper* (2003).

[3] Birch, B.J. and H.P.F. Swinnerton-Dyer. "Notes on elliptic curves," *J. Reine Angew. Math.* 212 (1963), 7-25.

[4] Calinger, R. *A Contextual History of Mathematics.* Prentice Hall: New Jersey, 1999.

[5] Cassels, J. W. S. and A. Frölich. Eds. *Algebraic Number Theory.* Academic Press: New York, 1967.

[6] Cassels, J. W. S. *Lectures on Elliptic Curves.* Cambridge: Cambridge University Press, 1991.

[7] Cassels, J. W. S. "Mordell's finite basis theorem revisited," *Math. Proc. Camb. Phil. Soc.* 100 (1986), 31-41.

[8] Cremona, J. *Algorthims for Modular Elliptic Curves.* 2nd Ed. Cambridge: Cambridge University Press, 1997.

[9] Cremona, J. "Higher descents on elliptic curves," Preprint. (1997).

[10] Doud, D. "A procedure to calculate torsion of elliptic curves over $\mathbb{Q}$," *Manuscripta Mathematica.* 95 (1998), 463-469.

[11] Eisenbud, D. *Commutative Algebra with a View Toward Algebraic Geometry.* New York: Springer-Verlag, 1995.

[12] Faltings, G. "Endlichkeitsätze für abelsche varietäten über Zahlkörpern," *Invent. Math.* 73 (1983), 549576.

[13] Fermat, P. *Oeuvres de Pierre Fermat.* Paul Tannery, Trans. Paris: Gauthier-Villars et fils, 1891.

[14] Harris, Joe. *Algebraic Geometry: A First Course.* New York: Springer Verlag, 1992.

[15] Hartshorne, Robin. *Algebriac Geometry.* New York: Springer Verlag, 1977.

[16] Hilbert, D. and A. Hurwitz. "Über die diophantischen gleichungen vom geschlecht null," *Acta Mathematica.* 14 (1890), 217-224.

[17] Husemöller, D. *Elliptic Curves.* 2nd Ed. New York: Springer Verlag, 2004.

[18] Koblitz, N. *Introduction to Elliptic Curves and Modular Forms.* New York: Springer-Verlag, 1993.

[19] Lang, S. and A. Neron. "Rational points of abelian varieties over function fields," *Amer. J. Math.* 81 (1959), 95-118.

[20] Lang, S. *Algebra.* New York: Springer-Verlag, 1999.

[21] Lang, Serge and John Tate. "Principal homogenous spaces over abelian varieties," *Amer. J. Math.* 80 (1958), 659-684.

[22] Mahoney, M. *The Mathematical Career of Pierre de Fermat.* Princeton: Princeton University Press, 1994.

[23] Martin, R. and W. McMillen. "An elliptic curve over $\mathbb{Q}$ with rank at least 24." *Number Theory Listserver.* (2000).

[24] Matsumura, H. *Commutative Algebra.* Reading, MA: Benjamin/Cumming, 1980.

[25] Mordell, L. J. "On the rational solutions of the indeterminate equations of the third and fourth degrees," *Proc. Cam. Phil. Soc.* 21 (1922), 179-182.

[26] Poincaré, H. "Sur les propriétiés arithmétiques des courbes algébriques," *J. Math. Pures Appl.* 5 (1901), 161-233.

[27] Poonen, B. "The Selmer group, the Shafarevich-Tate group, and the weak Mordell-Weil Ttheorem." Preprint.

[28] Rubin, K. and A. Silverberg. "Ranks of elliptic curves," *Bull. Amer. Math. Soc.* 39 (2002), 455-474.

[29] Schoof, R. and N. Schappacher. "Beppo Levi and the arithmetic of elliptic curves," *The Mathematical Intelligencer.* 18 (1996), 5769.

[30] Serre, J. P. *Lectures on the Mordell-Weil Theorem.* Martin Brown, Trans. Vieweg: Wiesbaden, 1997.

[31] Serre, J. P. *Algebraic Groups and Class Fields.* Springer-Verlag: New York.

[32] Serre, J. P. *Local Fields.* Springer Verlag: New York, 1979.

[33] Shafarevich, I. G. *Basic Algebriac Geometry: Varieties in Projective Space.* Vol 1. Springer-Verlag: New York, 1972.

[34] Siegel, C. "Über einige Anwendungen diophantischer Approximationen," (1929) in *Collected Works*. Springer-Verlag: New York, 1966.
[35] Siksek, S. "Infinite descent on elliptic curves," *Rocky Mountain Journal of Math.* 44 (1995), 1501-1538.
[36] Silverman, J. *The Arithmetic of Elliptic Curves*. Springer Verlag: New York, 1986.
[37] Silverman, J. and J. Tate. *Rational Points on Elliptic Curves*. Springer Verlag: New York, 1992.
[38] Smart, N. P. *The Algorithmic Resolution of Diophantine Equations*. Cambridge: Cambridge University Press, 1998.
[39] Stein, W. *Algebriac Number Theory*. Math 129 Course Notes. Cambridge, MA: Harvard University, 2003.
[40] Stoll, M. "Explicit 4-descent on an elliptic curve." Online. Available: http://www.faculty.iu-bremen.de/stoll/
papers/4-descent.pdf.
[41] Tunnell, J. "A classical Diophantine problem and modular forms of weight 3/2," *Inventiones math.* 72 (1983), 323-334.
[42] Weil, A. *Number Theory: an approach through history*. New York: Birkhauser, 1983.
[43] Weil, A. "Sur un theoreme de Mordell." *Bull. Sci. Math.* 54 (1929), 182-191.
[44] Wiles, A. "The Birch and Swinnerton-Dyer Conjecture," Preprint. Online. Available: www.fen.bilkent.edu.tr/
franz/ta/birchsd.pdf.
[45] S. Schmitt and H.G. Zimmer. *Elliptic Curves: A computational approach*. Berlin: de Gruyter, 2003.