DIOPHANTINE GEOMETRY WEEK 08 NOTES

XIAORUN WU

Abstract. This week, we will present a proof of Mordell-Weil Theorem for Elliptic Curves. Following the approach as outlined in [1], we first introduce the concepts of arithmetic heights as well as other necessary tools, then we will briefly outline the proof for the weak Mordell-Weil theorem. Finally, we extend the result to the strong version using Fermat's descent argument.

Contents

1. Introduction

The main objective of this week is to give a proof of the strong Mordell-Weil over, namely the finite generation of the group of rational points of an abelian variety defined over a number field.

We first present a brief historical overview of the problem. The theorem of Mordell-Weil was first proved over a special case, namely elliptic curves. In fact, L.J. Mordell proved the finiteness of the rank of the group of rational points on an elliptic curve E defined over Q in his famous paper [5]. The paper focused on the elliptic curve in the form of quartic equation $y^2 = a_0x^4 + \cdots + a_4$, and the proof used its parametrization by means of Jacobi elliptic functions and theta functions.

About twenty years later, Weil generalized the results of elliptic curves over number fields to abelian varieties in his famous thesis [9]. A. Weil made the rather critical observation in [10] that, for elliptic curves, a Weierstrass model rather than the quartic equation used by Mordell would simplify the proof, since it allows the addition and duplication formulas of elliptic functions used by Mordell to be replaced by rational functions on the curve. Since then, this has become the standard elementary approach to the Mordell–Weil theorem for elliptic curves over a field. We will follow the development sequence from Mordell to Weil: namely, we will first give a rather simple-minded proof of the Mordell-Weil Theorem for elliptic curves.

Then using results from Naron-Tate and Galois Cohomology, we will generalize this results to any Abelian variety.

For the rest of the section, we will state the Mordell-Weil theorem on Elliptic curves, outline the main stages of the proofs, and present the relevant theorems. We will then fill in the proofs and details of those theorems over the next sections.

1.1. Mordell-Weil Theorem Over Elliptic Curves.

The statement of the theorem (for elliptic curves) is as follows:

Theorem 1.1 (Mordell-Weil)**.** *For an elliptic curve E defined over a number field K, the elliptic curve group $E(K)$ is finitely generated.*

The basic structure for the proof consists of two stages. In the first stage, we prove a weak version, commonly known as the weak Mordell-Weil Theorem. The statement of the theorem is as follows:

Theorem 1.2 (Weak Mordell-Weil)**.** *For an elliptic curve E over a number field K, the quotient $E(K)/mE(K)$ is finite for all integers $m \geq 2$.*

In the second stage, we use a technique known as Fermat Descent Argument, which goes as following:

Theorem 1.3 (Fermat's Descent Theorem, group formulation)**.** *Let A be an Abelian group. If there exists some $m \geq 2$ and a function $h : A \mapsto \mathbb{R}$ such that:*

(1) *For any $Q \in A$ there exists some constant C_Q such that for $\forall P \in A$, $h(P+Q) \leq 2h(P) + C_Q$;*
(2) *There exists some finite C such that for any $P \in A$, $h(mP) \geq m^2 h(P) - C$;*
(3) *For any $k \in \mathbb{R}$, $h(P) \leq k$ only holds for finitely many $P \in A$;*
(4) *The quotient A/mA is finite*

then the group A is finitely generated.

Remark 1.1. The first two conditions tells us that any point $P \in A$ can be generated by a specific set of coset representatives of A/mA and points of height less than some constant M, with both the representatives and M independent of P; (3) and (4) guarantees the finiteness of the set.

We would leave proves of the three theorems in section 3.

Thus if we could find an appropriate height function $h : E(K) \to \mathbb{R}$, which satisfies the criteria (1)(2)(3) of theorem 1.3, then we would be able to prove theorem 1.1 from theorem 1.2. In the next sections, we will formally construct a height function and show that it indeed satisfy the conditions.

2. Heights over Projective Space and $E(K)$

In this section, we will start out with the definition of absolute value, and then we define height function over projective spaces. From that, we define height over the group $E(K)$.

Definition 2.1 (Absolute value)**.** *An absolute value on a field K is a real valued function $|\cdot|$ on K such that:*

(1) $|x| \geq 0$ *and* $|x| = 0$ *iff* $x = 0$*;*
(2) $|xy| = |x| \cdot |y|$*;*
(3) $|x + y| \leq |x| + |y|$.

and if an absolute value satisfy a stronger condition

(1) $|x+y| \leq \max\{|x|, |y|\}$,

then it is called non-archemedean. An absolute value that satisfy (3) but does not satisfy (4) is called non-archemedean.

and we define the equivalence between absolute values as follows:

Proposition 2.1. *Two absolute values $|\cdot|_1$, $|\cdot|_2$ are equivalent if and only if there is a positive real number s such that*

$$|x|_1 = |x|_2^s$$

for $x \in K$. When this happends, they define the same topology over K.

For proof, see for example [3]. We define place v to be an equivalence class of non-trivial absolute values. By $|\cdot|_v$ we denote an absolute value in the equivalence class determined by the place v. If the field L is an extension of K and v is a place of K, we write $w|v$ for a place w of L if and only if the restriction to K of any representative of w is a representative of v, and say that w extends v and, equivalently, that w lies over v . In this case we denote this by $w|v$.

We also note the following useful lemma:

Lemma 2.2. *Let $x \in K \setminus \{0\}$ and $y \in L \setminus \{0\}$. Then:*

$$\sum_{w|v} \log |x|_w = \log |x|_v$$

The proof of this uses Hensel's lemma, for a complete proof see [1].

With this, we now introduce the product formula.

Definition 2.3 (Product Formula)**.** *Let K be a field and M_K be a set of non-trivial inequivalent absolute values on K such that the set*

$$\{|\cdot|_v \in M_K| \; |x|_v \neq 1\},$$

is finite for any $x \in K\backslash\{0\}$. We identify the elements of M_K with the corresponding places and say that M_K satisfies the product formula if

$$\prod_{v \in M_K} |x|_v = 1$$

for any $x \in K \setminus \{0\}$, or equivalently for $x \neq 0$, $\sum_{v\in M_K} \log |x|_v = 0$.

With this property, now we may formally define height on the projective space. We denote by $\overline{\mathbb{Q}}$ a choice of an algebraic closure of $\mathbb{Q}$. Let us consider the projective space $\mathbb{P}^n_{\overline{\mathbb{Q}}}$ with standard global homogeneous coordinates $x = (x_0 : x1 : \cdots : x_n)$. Let $P \in \mathbb{P}^n_{\overline{\mathbb{Q}}}$. We now define a function, called height, on algebraic points of $\mathbb{P}^n_{\overline{\mathbb{Q}}}$, which may be considered as a measure of the "algebraic complication" needed to describe P. This is a fundamental notion at the basis of diophantine geometry.

In particular, we note the following proposition (again check [1] for a complete proof):

Proposition 2.2. *If K is a number field, and let M_K be the associated set of places and normalized absolute values, obtained from the above construction applied to the extension $K/\mathbb{Q}$. Then M_K satisfies product formula.*

Definition 2.4. *Height Let P be a point of $\mathbb{P}^n_{\overline{\mathbb{Q}}}$ represented by a homogeneous nonzero vector $\mathbf{x}$ with coordinates in a number field K. Then $h : \mathbb{P}^n_{\overline{\mathbb{Q}}} \mapsto \mathbb{R}$ is defined as*

$$h(\mathbf{x}) = \sum_{v \in M_K} \max_j \log |x_j|_v$$

As a convention, we also denote $H(\mathbf{x}) = e^{h(\mathbf{x})}$. In particular, this choice is extremely powerful, in that we can show the following two lemmas:

Lemma 2.5. *$h(\mathbf{x})$ is independent of the choice of coordinates.*

Proof. Let L be another number field containing the coordinates $x_0, \cdots, x_n$ of $\mathbf{x}$. We can assume that $K \subset L$. Then

$$\sum_{w \in M_K} \max_j \log |x_j|_w = \sum_{v \in M_K} \sum_{w|v} \max_j \log |x_j|_w,$$

then the claim follows directly from lemma 2.2. □

Lemma 2.6. *$h(\mathbf{x})$ is independent of the choice of coordinates.*

Proof. Let $\mathbf{y}$ be another coordinate vector. By the preceding lemma, we may assume that $x_0, \cdots, x_n, y_0, \cdots, y_n \in K$. There is $\lambda \in K$, $\lambda \neq 0$, with $y = \lambda x$, hence

$$h(\mathbf{y}) = \sum_{v \in M_K} \max_j \log |y_j|_v = \sum_{v \in M_K} \log |\lambda|_v + \sum_{v \in M_K} \max_j \log |x_j|_v.$$

Thus by proposition 2.2, we have $h(\mathbf{y}) = h(\mathbf{x})$. □

Since the definition of the height above, we can define the absolute height, or more commonly known, Weil Height, as follows:

Definition 2.7. *Weil Height Let $P \in K*$ be a point, then*

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

Again, the above result shows that the definition of the height is independent of choice of K.

Finally, to define height over an elliptic curve (and more generally, for a projective variety over $\overline{\mathbb{Q}}$), we have the following:

Definition 2.8. *Weil Height Let X be a projective variety over $\overline{\mathbb{Q}}$. Let $\varphi : X \mapsto \mathbb{P}^n_{\overline{\mathbb{Q}}}$ be a morphism over $\overline{\mathbb{Q}}$. The Weil height of $P \in X(\overline{\mathbb{Q}})$ relative to φ is defined by $h_\varphi(P) := h \circ \varphi(P)$, with h the usual height on $\mathbb{P}^n_{\overline{\mathbb{Q}}}$. In particular, the height of a point over an elliptic curve $E(K)$ with regards to a nonconstant function $f \in K(E)$ is defined as $h_f(P) = \log H(f(P))$.*

For the last thing of this chapter, let us extend the definition height function for polynomials.

Definition 2.9. *The height of a polynomial $f(t_1, \cdots, t_n) = \sum_{j_1, \cdots, j_n} a_{j_1 \cdots j_n} t_1^{j_1} \cdots t_n^{j_n} = \sum_j a_j \boldsymbol{t}^{\boldsymbol{j}}$ with coefficients in a number field K is the quantity $h(f) = \sum_{v \in M_K} \log |f|_v$, where*

$$|f|_v = \max_j |a_j|_v.$$

is the Gauss Norm. Now let $F = [f_1, f_2, \cdots, f_n]$. If we label the coefficents of each f_i as a_{ij} then define

$$|F|_v = \max_{i,j}(|a_ij|_v).$$

And similarly $H_K(F)$ is

$$H_K(F) = \prod_{v \in M_K} |F|_v.$$

With this we've completed our definition of the height function. We now need to show that indeed, this height function would satisfy the conditions of theorem 1.3. The next section would introduce some famous lemmas which we would use to verify those conditions.

3. Mahler Measure, Northcott Theorem, Gauss's & Gelfond's Lemma

In this section, we will introduce Mahler Measure, which would be essential in the proof of Northcott Theorem. This will in turn be used later to verify condition (3) in Theorem 1.3. We also introduce Gelfond and Gauss Lemma, which would be relevant in verifying condition (1) and (2).

We first prove for finite places **Gauss's Lemma**:

Lemma 3.1. *If v is not archimedean, then $|fg|_v = |f|_v|g|_v$.*

Proof. The inequality $|fg|_v \leq |f|_v|g|_v$ is immediate because v is not archimedean. Let us assume first that $f(t)$ and $g(t)$ are polynomials in one variable t . We denote by c_j the coefficient

$$\sum_{j=k+l} a_k b_l$$

of $f(t)g(t)$. Without loss of generality, we can assume that $|f|_v = 1, |g|_v = 1$.

Suppose $|fg|_v < 1$. Let j be the smallest index with $|a_j|_v = 1$. Since $|c_j|_v < 1$ and $|a_k|_v < 1$ for $k < j$, we get $|b_0|_v < 1$. Now we apply the above formula for the coefficient c_{j+l} and conclude $|b_l|_v < 1$ by induction. This contradiction proves the lemma in the one-variable case. For several variables, let d be an integer larger than the degree of fg. The Kronecker substitution

$$x_j = t^{d^{j-1}}(j = 1, \cdots, n)$$

reduces the problem to the one-variable case. □

We now introduce Mahler Measure, which we will use to prove Northcott's Theorem

Definition 3.2 (Mahler Measure)**.** *We define Mahler Measure of a $n-variable$ polynomial f as:*

$$M(f) := \exp\left(\int_{\mathbb{T}^n} \log|f(e^{i\theta_1}, \cdots, e^{i\theta_n})|d\mu_1 \cdots d\mu_n\right),$$

where we have abbreviated $\mathbb{T}$ for the unit circle $\{e^{i\theta}|0 \leq \theta < 2\pi\}$ equipped with the standard measure $d\mu = (1/2\pi)d\theta$.

The advantage of Mahler Measure is two-fold: first, it permits a multiplicativity property: it's not hard to show that $M(fg) = M(f) \cdot M(g)$. Moreover, it permits an upper and lower bound by ℓ_∞ norm, which is a very important gradient in Northcott's theorem:

Lemma 3.3. *Let $\alpha \in \overline{\mathbb{Q}}$ and let $f(t) := a_d t^d + \cdots + a_0$ be the minimal polynomial of α over $\mathbb{Z}$, where $d = \deg(\alpha)$ and denote $\alpha_1, \alpha_2, \cdots \alpha_d$ roots of f, then $M(f) \leq \ell_1(f)$. Moreover*

$$\binom{d}{\lceil d/2 \rceil}^{-1} \ell_\infty(f) \leq M(f) \leq \ell_2(f) \leq (d+1)^{1/2} \ell_\infty(f).$$

Proof. The first inequality is obvious from the definition of $M(f)$ and the pointwise bound $|f(e^{i\theta})| \leq \ell_1(f)$ on $\mathbb{T}$. Next by convexity, we get

$$M(f) \leq \left(\int_{\mathbb{T}^n} |f(e^{i\theta})|^2 d\mu \right)^{1/2}.$$

Now by Parseval's identity, the right-hand-side equals

$$\ell_2(f) = \left(\sum_{j=0}^{d} |a_j|^2 \right)^{1/2} \leq (d+1)^{1/2} \ell_\infty(f).$$

Finally, we remark that

$$|\frac{a_{d-r}}{a_d}| = | \sum_{j_1 < \cdots < j_r} \alpha_{j_1} \cdots \alpha_{j_r}|,$$

hence

$$|a_{d-r}| \leq \binom{d}{r} |a_d| \prod_{j=1}^{d} \max(1, |\alpha_j|).$$

Finally by Jensen's formula ($\log M(f) = \log |a_d| + \sum_{j=1}^{d} \log^+ |\alpha_j|$,for proof see [1]), we have

$$|a_{d-r}| \leq \binom{d}{r} M(f).$$

□

We also note another property of Mahler's measure:

Proposition 3.1 (Property of Mahler's Measure)**.** *Let $\alpha \in \overline{\mathbb{Q}}$ and let f be the minimal polynomial of α over $\mathbb{Z}$. Then*

$$\log M(f) = \deg(\alpha) h(\alpha).$$

In particular

$$\log |N_{\mathbb{Q}(\alpha)/\mathbb{Q}}(\alpha)| \leq \deg(\alpha) h(\alpha).$$

The following consequence is known as Northcott's theorem:

Theorem 3.4 (Northcott)**.** *There are only finitely many algebraic numbers of bounded degree and bounded height.*

Proof. Let α be algebraic of degree d and height $h(\alpha) \leq \log H$. Let $f(t) = a_d t^d + \cdots + a_0$ be the minimal polynomial of α over $\mathbb{Z}$. By Proposition 3.1, we have $M(f) \leq H_d$. Also, Lemma 3.3 shows that $\max |a_i| \leq 2^d M(f)$. Therefore, the coefficients of f are bounded by $(2H)^d$. Since there are $d+1$ integer coefficients for each f, they give rise to not more than $(2\lceil (2H)^d \rceil + 1)^{d+1}$ distinct polynomials f. Since each f has d roots, the number of algebraic integers of degree d and height at most H is at most $d(2\lceil (2H)d \rceil + 1)^{d+1} \leq (5H)^{d^2+d}$. □

Finally, we will introduce Gelfond's lemma:

Lemma 3.5. *Let $f_1, \cdots, f_m$ be complex polynomials in n variables and set $f := f_1 \cdots f_m$. Then*

$$2^{-d} \prod_{j=1}^{m} \ell_\infty(f_j) \leq \ell_\infty(f) \leq 2^d \prod_{j=1}^{m} \ell_\infty(f_j).$$

The proof follows from Lemma 3.3. The readers are welcome to verify the results by first proving the following statement:

Lemma 3.6. *Let $f(t_1, \cdots t_n)$ be a polynomial with complex coefficients and partial degrees $d_1, \cdots, d_n$, then*

$$\prod_{j=1}^{n} (d_j + 1)^{-1/2} M(f) \leq \ell_\infty(f) \leq \prod_{j=1}^{n} \binom{d_j}{\lceil d_j/2 \rceil} M(f).$$

A direct consequence from Gauss's Lemma and Mahler's Lemma is the following:

Lemma 3.7. *Let $f_1, f_2, \cdots, f_m$ be polynomial in n variables with coefficients in $\overline{\mathbb{Q}}$ and let d be the sum of the partial degrees of $f := f_1 \cdots f_m$. Then*

$$-d \log 2 + \sum_{j=1}^{m} h(f_j) \leq h(f) \leq d \log 2 + \sum_{j=1}^{m} h(f_j).$$

4. Boundedness of Height Function

With all the necessary background from the last chapters, we will verify that indeed the height function defined in the previous section would satisfy the three condtions. For the first two conditions, we will devise a purely elementary approach to prove the theorem. Indeed, for the extension of this result to more general abelian variety, we need more technical tools (e.g. Naron Tate Height, Galois Cohomology), which we will be discussing more in chapter 5 onwards.

The outline of this section is also straightforward: we will verify each of the conditions in theorem 1.3 sequentially. The verification of first two condition follows from a purely algebraic approach, while the verification of condition (3) we will introduce Northcott theorem as our important tool.

To verify the first two conditions, we will first introduce the following proposition:

Proposition 4.1. *Fix $m, n > 0$. For any morphism $F : \mathbb{P}^n(K) \mapsto \mathbb{P}^m(K)$ of degree d, there exist constants C_1 and C_2 such that for any $P \in P^n(\overline{\mathbb{Q}})$, there exists constant C_1, C_2 such that $C_1 H(P)^d \leq H(F(P)) \leq C_2 H(P)^d$.*

Remark 4.1. This result allows both an upper bound and an lower bound of the Weil heights we defined in the last section, which we will see later.

Proof. Define $P = [x_0, x_1, \cdots, x_n]$ and $F = [f_0, f_1, f_2, \cdots, f_m]$ where the f_is are homogeneous without common zero. If for place v we define

$$|P|_v = \max_i(|x_i|_v).$$

And set $\iota(v)$ to be a constant such that $\iota(v) = 1$ if v is Archimedean, and 0 otherwise. We may simplify the triangle inequality as $|a_1 + \cdots + a_n|_v = n^{\iota(v)} \max(|a_i|_v)$.

Since the absolute value of every term in any f_i must be less than or equal to $|F|_v |P|_v^d$, the triangle inequality gives us

$$|F(P)|_v \leq \binom{n+d}{d}^{\iota v} |F|_v |P|_v^d,$$

where we note the $\binom{n+d}{d}$ comes from the maximum possible number of terms in f_i

Multiplying across all places $v \in M_K$ gives

$$H_K(F(P)) \leq \binom{n+d}{d}[K:Q]H_K(F)H_K(P)^d H(F(P)) \leq \binom{n+d}{d}H(F)H(P)^d,$$

which is the desired upper bound.

For the lower bound, we first note that because the f_i is homogeneous of degree d and share no nontrivial zeroes, the set upon which all vanish in the affine space $\mathbb{A}^{n+1}(\mathbb{Q})$ is only $(0, 0, \cdots, 0)$.

Thus, we can apply the Nullstellensatz to yield that in the space $\overline{\mathbb{Q}}[X_0, X_1, \cdots, X_n]$ the ideal generated by the f_is must contain powers of every polynomial which vanishes at $(0, 0, \cdots, 0)$, including the polynomials $X_0, X_1, \cdots .X_N$. This means that for some integer $n \geq 1$, each X_i^n can be written as:

$$X_i^n = \sum_{j=0}^{m} g_{ij} f_j$$

with $g_{ij} \in \overline{\mathbb{Q}}[X_0, X_1, \cdots, X_n]$ is homogeneous of degree $n - d$. If we associate the absolute value and heights $|G|_v$ and $H_K(G)$ the same way as we defined the notation for F, taking the maximum across all coefficients of g_{ij}s, then by triangle inequality, there exists some constant $\mathcal{E}_1$ such that

$$|x_i|_v^n \leq \mathcal{E}_1^{\iota(v)} \max_{0 \leq j \leq m} |g_{ij}(P)|_v |F(P)|_v,$$

so that

$$|P|_v^n \leq \mathcal{E}_1^{\iota(v)} \max_{0 \leq j \leq m} |g_{ij}(P)|_v |F(P)|_v,$$

Now applying the triangle inequality on each $|g_{ij}(P)|_v$, there exists some constant $\mathcal{E}_2$ such that

$$g_{ij}(P)|_v \leq \mathcal{E}_2^{\iota(v)} |G|_v |P|_v^{n-d},$$

so that

$$|P|_v^n \leq \mathcal{E}_1^{\iota(v)} \mathcal{E}_2^{\iota(v)} |G|_v |P|_v^{n-d} |F(P)|_v,$$

$$|P|_v^d \leq \mathcal{E}_1^{\iota(v)} \mathcal{E}_2^{\iota(v)} |G|_v |F(P)|_v,$$

Hence we have

$$H_K(P)^d \leq (E_1 \mathcal{E}_2)^{[K:\mathbb{Q}]} H_K(G) H_K(F(P)),$$

which is precisely

$$H(P)^d \leq (E_1 \mathcal{E}_2) H(G) H(F(P)).$$

So setting $C_1 = \frac{1}{(E_1 \mathcal{E}_2) H(G)}$ completes the lower bound. □

Proposition 4.1 allows the following lemma, which gives a condition for bounding:

Lemma 4.1. *If f and g are even functions, then $h_f \deg(g) = h_g \deg(f) + O(1)$.*

Proof. Firstly, note that elliptic curve that the even functions in $K(E)$ are exactly the functions of $K(X)$. This allows us to find a rational function $r(X) \in K(X)$ with $r \circ x = f$. It follows from Weil Height's definition that $h_f(P) = \log H(f(P)) = \log H(r(x(P))$. We can take the logarithm of both sides in Proposition 4.1 to yield:

$$\log H(x(P))(\deg r) + \log C_1 \leq \log H(r(x(P)) \leq log H(x(P))(\deg r) + \log C_2$$

where $\deg r$ is the degree of r, defined as the degree of numerator minus the degree of demoninator. Treating C_1, C_2 as constants,

$$\log H(r(x(P)) = (\deg r) \log H(x(P)) + O(1) = (\deg r) h_x(P) + O(1).$$

Since $\deg x = 2$, we have that $\deg f = 2 \deg r$ and thus

$$2h_f(P) = (\deg f) h_x(P) + O(1)$$

and likewise

$$2h_g(P) = (\deg g) h_x(P) + O(1)$$

hence

$$h_f(P)(\deg g) = \frac{1}{2}(\deg f)(\deg g) h_x(P) + O(1) = h_g(P)(\deg f).$$

□

Finally, we will follow the treatment on [7] and [8] to prove the following lemma, which would allow us to verify the first two conditions:

Lemma 4.2. *For $P, Q \in E(K)$ and $f : E \mapsto \mathbb{P}^1_K$ even,*

$$h_f(P+Q) + h_f(P-Q) = 2h_f(P) + 2h_f(Q) + O(1),$$

where we note that the added constant $O(1)$ is independent of the choice of P and Q.

Proof. From the previous lemma we have $2h_f = (\deg f)h_x + O(1)$. Hence substitue this back to the equation in the lemma, we see that it suffice to prove h_x also satisfy the equation.

Without loss of generality, we assum $P, Q \neq O$. (Otherwise if $P = O$ and $Q = O$ gives us the trivial case with h_x), , setting $x(P) = [x_1, 1]$, $x(Q) = [x_2, 1]$, $x(P+Q) = [x_3, 1]$, and $x(P-Q) = [x_4, 1]$.

We also write out E in Weierstrass form:

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

Then the Addition formula gives us:

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

$$x_3x_4 = \frac{(x_1x_2 - A)^2 - 4B(x_1+x_2)}{(x_1+x_2)^2 - 4(x_1x_2)},$$

This admits a map to $[1, x_3+x_4, x_3x_4]$ in terms of $[1, x_1+x_2, x_1x_2]$; which is given by

$$g : \mathbb{P}^2 \mapsto \mathbb{P}^2, g(a,b,c) = \left(b^2 - 4ac, 2b(Aa + c) + 4Ba^2, (c - Aa)^2 - 4Bba\right).$$

This, when combining with the maps $\sigma : E \times E \to \mathbb{P}^2$, $(P, Q) \to [1, x(P) + x(Q), x(P)x(Q)]$ and $G : E \times E \to E \times E$, $G(P,Q) = (P+Q, P-Q)$ clearly gives us that $\sigma \circ g = G \circ \sigma$.

We next show that g is a proper morphism on $\mathbb{P}^2$ by showing that $g([a,b,c]) \neq [0,0,0]$ for any $[a,b,c] \neq [0,0,0]$. If $a = 0$, then we have $g([a,b,c]) = [b^2, 2bc, c^2]$, meaning that there are no zeroes for which $a = 0$ but $b \neq 0$ or $c \neq 0$.

Thus, we set $a = 1$, giving

$$g([a,b,c]) = [b^2 - 4c, 2b(A+c) + 4B, (c-A)^2 - 4Bb].$$

If we want all three polynomials to vanish, then we must have $c = \frac{1}{4}b^2$. This gives us the other polynomials $\frac{1}{2}b^3 + 2Ab + 4B$ and $\frac{1}{16}b^4 - \frac{1}{2}Ab^2 - 4Bb + A^2$. Therefore

$$(24x^2 + 128A)(\frac{1}{16}x^4 - \frac{1}{2}Ax^2 - 4Bx + A^2) - (3x^3 - 20Ax - 216B)$$
$$(\frac{1}{2}x^3 + 2Ax + 4B) = 32(4A^3 + 27B^2).$$

This means that both polynomials vanishing on any x would indicate $4A3 + 27Bz^2 = 0$; which would imply E is singular; since an initial assumption is that E is not singular, this does not occur and therefore g is a morphism.

We now have that $H(\sigma(P+Q, P-Q)) = H(g(\sigma(P,Q)))$. Since g is a morphism of degree 2, Proposition 4.1 gives us that

$$C_1 H(\sigma(P,Q))^2 \leq H(g(\sigma(P,Q))) \leq C_2(\sigma(P,Q))^2,$$

or

$$\log H(g\sigma(P,Q)) = 2\log H(\sigma(P,Q)) + O(1).$$

Therefore if we apply Lemma 3.7 to the polynomial:

$$f(T) = (T + x_1)(T + x_2) = T^2 + (x_1 + x_2)T + x_1x_2$$

to get

$$\frac{1}{4}H(x_1)H(x_2) \leq H([1, x_1 + x_2, x_1x_2]) \leq 4H(x_1)H(x_2).$$

Taking the logarithm gives

$$\log H(\sigma(P,Q)) = \log H([1, x1 + x2, x1x2])$$
$$= \log H(x_1)H(x_2) + O(1) = h_x(P) + h_x(Q) + O(1),$$

and doing the same with x_3 and x_4 in the polynomial gives the corresponding result for $P+Q$ and $P-Q$. Thus we have:

$$h_x(P+Q) + h_x(P-Q) = \log H(\sigma(P+Q, P-Q)) + O(1)$$
$$= 2\log H(\sigma(P,Q)) + O(1) = 2h_x(P) + 2h_x(Q) + O(1),$$

as desired. □

We are now able to verify the first and second condition of Fermat's Descent Theorem:

Lemma 4.3 (First Condition of Fermat Descent Theorem)**.** *For any $Q \in E(K)$ there exists some constant C such that for any $P \in E(K)$, $h_f(P+Q) \leq 2hf(P)+C$.*

Proof. This follows immediately, as we can set the $C \geq 2h_f(Q) + O(1)$ for the $O(1)$ in Lemma 4.2, and then $h_f(P+Q)+h_f(P-Q) \leq 2h_f(P)+C$ and $h_f(P-Q) \geq 0$. □

Lemma 4.4 (Second Condition of Fermat Descent Theorem)**.** *For any integer $m \geq 2$, and any $P \in E(K)$, $h_f([m]P) = m^2 h_f(P) + O(1)$ where $[m]P$ is the sum $P + P + \cdots + P$ for m times.*

Proof. This is trivial for $m = 0, 1$ so we induct for m given the result on $m-1$ and $m-2$. From Lemma 4.2 we then have that

$$h_f([m]P) + h_f([m-2]P) = 2h_f([m-1]P) + 2h_f(P) + O(1),$$

so that

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

as desired. □

The final condition we need to verify is the third condition. This will also be relatively straightforward. We state a special case of Northcott's Lemma we will be using here:

Lemma 4.5 (Northcott. Special Case)**.** *The set of $P \in \mathbb{P}^n(\overline{\mathbb{Q}})$ with $H(P) \leq C$ and $[\mathbb{Q}(P) : \mathbb{Q}] \leq D$ for some constants C and D is finite.*

From this we will see that the third condition is immediate:

Lemma 4.6 (Third condition for Descent Theorem)**.** *For any number field K and constant c, the set of points of $\mathbb{P}^n(K)$ with height less than or equal to c is finite. In particular for an elliptic curve E/K and nonconstant function $f \in K(E)$, the set of points P for which $h_f(P) \leq C$ is finite.*

Proof. The first half of the statement is immediate. Since for any $P \in \mathbb{P}^n(K)$, the degree $[\mathbb{Q}(P) : \mathbb{Q}] \leq [K : \mathbb{Q}]$, we apply Northcott's theorem with $d = [K : \mathbb{Q}]$ for this result.

For the second half, we note for points on $E(K)$ for which $h_f(P) \leq C$ we have

$$\log H(f(P)) \leq C \text{ or equivalently } \; H(f(P)) \leq e^C.$$

Setting $c = e^C$ in the above result tells us that there are finitely many points $f(P) \in \mathbb{P}^n(K)$ where this occurs. Since the function f is in $K(E)$ and is nonconstant, the preimage of each point in $\mathbb{P}^n(K)$ will be finite, meaning that the union of the preimages of each $f(P)$ with $H(f(P)) \leq e^C$ is finite as desired. □

We are now in good shape to put all things together for Mordell-Weil Theorem Over Elliptic Curves.

5. Putting it altogether: Mordell-Weil Theorem of $E(K)$

So now to prove Theorem 1.1, we are just missing two things:

(1) Proof of Theorem 1.2 (weak Mordell-Weil);
(2) Proof of Theorem 1.3 (Fermat's descent).

The proof of (1) is quite involved: in the next section, we would outline the important lemmas . For a detailed proof, the readers may check [8] or [1]. The proof of (2) is in fact straightforward, which we will write up in detail.

5.1. **Proof of weak Mordell-Weil.**

We first reformulate Theorem 1.2 into two corollaries:

Corollary 5.1. *Let E be an elliptic curve over a number field K with 2- torsion also defined over K. Then $E(K)/2E(K)$ is finite.*

Corollary 5.2. *Let A be an abelian variety defined over a field K and let L be a finite separable extension of K. Let m be a positive integer and suppose that $A(L)/mA(L)$ is a finite group. Then $A(K)/mA(K)$ is a finite group.*

To prove corollary 5.1 would require the following input from commutative algebra:

Proposition 5.1. *Let K be a number field. Then we can find a finite set of places S of K such that for any finite set of places $T \in M_K$ with $T \supset S$, the ring $O_{T,K}$ is a principal ideal domain and hence a unique factorization domain.*

For proof of this see [4] and [6].

We also observe that E may be viewed as a plane curve in $\mathbb{P}^2_K$, given in standard affine coordinates by

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

for some $a_i \in K$. Replacing y by $y - \frac{1}{2}(a_1x + a_3)$ (which is allowed because $char(K) \neq 2$), we may assume that $a_1 = a_3 = 0$. Therefore, after this simplification, the affine part of E has equation

$$y^2 = (x - \alpha_1)(x - \alpha_2)(x - \alpha_3)$$

with $\alpha_i \in K, i = 1, 2, 3$.

with this observation we would be able to prove the following proposition:

Proposition 5.2. *The group $E[2]$ of 2-torsion points of E consists of the identity element O and the points $(\alpha_i, 0)$, $i = 1, 2, 3$, of order 2.*

In particular, this will allow us to prove the following:

Lemma 5.1. *The map $\varphi : E(K) \mapsto (K^\times/K^{\times 2})^3$ is a group homomorphism with kernel $2E(K)$.*

whic will in turn allow us to prove:

Proposition 5.3. *Let R be a unique factorization domain with quotient field K. Assume that $char(K) \neq 2$ and that the group of units $R^\times$ in R is finitely generated. Let E be the elliptic curve given by*

$$y^2 = (x - \alpha_1)(x - \alpha_2)(x - \alpha_3),$$

where $\alpha_1, \alpha_2, \alpha_3$ are distinct elements of R. Then there exists a constant C such that $|E(K)/2E(K)| \leq C$.

So putting all these together we will be able to prove Corollary 5.1:

Proof of Corollary 5.1. The ring of integers O_K is not necessarily a unique factorization domain. However, by Proposition 5.1, we can find a finite set of places S in K such that for any finite set of places $T \in M_K$ with $T \supset S$, the ring R of T-integers in K is a unique factorization domain. By Dirichlet's unit theorem, its group of units $R^\times$ is finitely generated. Now by Proposition 5.2 and the observation, we have that E is K-isomorphic to an elliptic curve of the form required in Proposition 5.3, because we can always enlarge the ring R so as to ensure that every $\alpha_i \in R$. The result now follows from Proposition 5.3. □

Now with Corollary 5.1, we introduce an additional Proposition, which would allow us to prove Corollary 5.2:

Proposition 5.4. *Let $n \in \mathbb{Z} \setminus \{0\}$. Then $[n]$ is a finite flat surjective morphism of degree $n^{2\dim(A)}$. The separable degree of $[n]$ equals the number of points of any fibre. If char$(K) \nmid n$, then $[n]$ is an étale morphism and*

$$A[n] \cong (\mathbb{Z}/n\mathbb{Z})^{2\dim(A)}.$$

If $p =$ char(K) divides n, then $[n]$ is not separable.

Proof of Corollary 5.2. Let $d = [L : K]$ and let δ be a positive integer such that we can write $d = d_0 d_1$ with $d_0 | m^{\delta-1}$ and $\gcd(d_1, m) = 1$. Let also E be a set of representatives of $A(L)/mA(L)$ in $A(L)$.

The group $A(L)/m^\delta A(L)$ is again a finite group, since a set of representatives for it is contained in the finite set

$$\mathcal{E}(\delta) = \mathcal{E} + m\mathcal{E} + \cdots + m^{\delta-1}\mathcal{E}.$$

Let F be the Galois closure of L over K, let $G = \mathrm{Gal}(F/K)$, let H be the subgroup of G of index d fixing L, and denote by R a full set of representatives for the left-cosets of H in G.

Let $x \in A(K) \subset A(L)$. Then we have $x - m^\delta y \in E(\delta)$ for some $y \in A(L)$. We apply the automorphisms $\sigma \in R$ to this equation and deduce

$$dx - m^\delta z \in \mathcal{E}'(\delta),$$

where $z := \sum_\sigma \sigma y$ and

$$\mathcal{E}'(\delta) := \left(\sum_{\sigma \in R} \sigma\right) \mathcal{E}(\delta).$$

Clearly, $z \in A(K)$ because any element $\tau \in \mathrm{Gal}(F/K)$ permutes the left-cosets of H. Since d_0 divides $m\delta - 1$, we may divide by d_0 , getting

$$d_1 x - m(m^{\delta-1}/d_0)z \in A(K) \cap \frac{1}{d_0}\mathcal{E}'(\delta),$$

and $A(K) \cap \frac{1}{d_0}\mathcal{E}'(\delta)$ is still a finite set use Proposition above. Finally, since d_1 and m are coprime, by Bézout the euclidean algorithm produces integers u and v such that $d_1 u - mv = 1$. After multiplication by u, it follows that

$$x - m((m^{\delta-1}/d_0)uz - vx) \in A(K) \cap \frac{u}{d_0}\mathcal{E}'(\delta),$$

which completes the proof □

Finally, we conclude the proof of Mordell-Weil Theorem by proving Fermat's Descent Theorem:

5.2. **Proof of Fermat's Descent Theorem.** Recall that Fermat's Descent Theorem states the following:

Theorem 5.2. *Let A be an Abelian group. If there exists some $m \geq 2$ and a function $h : A \mapsto \mathbb{R}$ such that:*

1. *For any $Q \in A$ there exists some constant C_Q such that for $\forall P \in A$, $h(P + Q) \leq 2h(P) + C_Q$;*
2. *There exists some finite C such that for any $P \in A$, $h(mP) \geq m^2 h(P) - C$;*
3. *For any $k \in \mathbb{R}$, $h(P) \leq k$ only holds for finitely many $P \in A$;*
4. *The quotient A/mA is finite*

then the group A is finitely generated.

Proof. We start by selecting the representatives of each coset of mA in A. Since A/mA is finite, there are finitely many cosets and thus this gives a finite set $Q_1, Q_2, \cdots, Q_n$.

For any point P, we denote the representative of the coset containing P as Q_P.

Now, taking any starting point $P_0 \in A$, we can write that $P_0 - QP_0 \in mA$, or $P_0 = QP_0 + mP_1$ for some $P_1 \in A$. We can repeat this process an arbitrary number of times as well, each time taking $P_k = QP_k + mP_{k+1}$.

Our aim now is to show that there exists some constant M for which, given any P_0, there is eventually a value of k for which $h(P_k) \leq M$. Condition (2) gives us that

$$m^2 h(P_k) \leq h(mP_k) + C$$

or

$$h(P_k) \leq \frac{1}{m^2}(h(mP_k) + C),$$

which, plugging in the equation relating P_{k-1} to P_k gives

$$h(P_k) \leq \frac{1}{m^2}(h(P_{k-1} - Q_{P_{k-1}}) + C).$$

Setting

$$C_{\max} = \max_{1 \leq i \leq n} C_{-Q_i}$$

allows us to in turn apply (1) which tells us that

$$h(P_{k-1} - QP_{k-1}) \leq 2h(P_{k-1}) + C_{-QP_{k-1}} \leq 2h(P_{k-1}) + C_{\max}$$

and thus

$$h(P_k) \leq \frac{1}{m^2}(2h(P_{k-1}) + C_{\max} + C).$$

Since $m \geq 2$, we have that

$$h(P_k) \leq \frac{1}{2}h(P_{k-1}) + \frac{1}{4}(C_{\max} + C),$$

hence

$$h(P_k) - \frac{1}{2}(C_{\max} + C) \leq \frac{1}{2}(h(P_{k-1}) - \frac{1}{2}(C_{\max} + C)).$$

Thus for any P_0 there must eventually be some value of K for which $h(P_k) - \frac{1}{2}(C_{\max} + C) \leq 1$, or $h(P_k) \leq 1 + \frac{1}{2}(C_{\max} + C)$, giving us our constant M. Thus, we have that any $P_0 \in A$ can be written as a sum using only elements of height $\leq 1 + \frac{1}{2}(C_{\max} + C)$ (which, since C and $C_{\max}$ are independent of P_0, is a finite set by (3)) and the coset representatives Q_i (a finite set by (4)). Thus, we have a finite set of generators from which every element of A may be produced. □

Hence Fermat Descent, combined with Weak Mordell-Weil Theorem, allows us to complete the proof of Mordell-Weil Theorem for Elliptic Curves.

References

[1] Bombieri, Enrico, and Walter Gubler. 2006. *Heights in diophantine geometry.* Cambridge: Cambridge University Press. http://public.eblib.com/choice/publicfullrecord.aspx?p=244411

[2] Robin Hartshorne, *Algebraic Geometry* Graduate Texts in Mathematics, Vol. **52** Chapter 1, pp. 54, Springer Science & Business Media, 2013

[3] N. Jacobson, *Basic Algebra II.* First edition. W. H. Freeman & Co., San Francisco 1980. xix+666 pp. Second edition. W. H. Freeman & Company, New York 1989. xviii+686 pp.

[4] S. Lang, *Algebraic Number Theory.* Second edition. Graduate Texts in Mathematics 110. Springer-Verlag, New York 1994. xiv+357 pp.

[5] L.J. Mordell, *On the rational solutions of the indeterminate equations of the third and fourth degrees.* Proc. Cambridge Philos. Soc. **21** (1922), 179–192.

[6] W. Narkiewicz, *Elementary and Analytic Theory of Algebraic Numbers.* Second Edition. PWN– Polish Scientific Publishers and Springer-Verlag, Warszawa 1990. xiv+746 pp.

[7] Jean-Pierre Serre, *Lectures on the Mordell-Weil theorem*, third ed., Aspects of Mathematics, Friedr. Vieweg & Sohn, Braunschweig, 1997, Translated from the French and edited by Martin Brown from notes by Michel Waldschmidt, With a foreword by Brown and Serre. MR 1757192

[8] Joseph H. Silverman, *The arithmetic of elliptic curves*, second ed., Graduate Texts in Mathematics, vol. 106, Springer, Dordrecht, 2009. MR 2514094

[9] A. Weil, L'arithm´etique sur les courbes alg´ebriques, *Acta Math.* **52** (1929), 281–315. Also OEuvres Scientifiques – Collected Papers. Vol. I, Corrected Second Printing, Springer-Verlag, New York–Heidelberg–Berlin 1980, 11–45.

[10] A. Weil, Sur un th´eor‘eme de Mordell, *Bull. Sc. Math.* (2) **54** (1929), 182–191. Also OEuvres Scientifiques–Collected Papers. Vol. I, Corrected Second Printing, Springer-Verlag, New York– Heidelberg–Berlin 1980, 47–56.

Email address: `xiaorunw@math.columbia.edu`