A PROOF OF THE MORDELL-WEIL THEOREM

HUISHI YU

Abstract. The goal of this expository paper is to give a self-contained proof of the Mordell-Weil theorem.

This paper assumes a familiarity with algebraic number theory and some basic knowledge of algebraic geometry.

Contents

1. Introduction

The origin of number theory goes back to Ancient Greek when people studied Diophantine equations, the solution of polynomial equations in rational numbers. For linear equations, we can solve them via linear algebra. For quadratic equations in two variables, we can use quadratic reciprocity and Hensel's lemma to solve them in every completion of $\mathbb{Q}$. Then piecing together information over every local field helps us get results for the global field $\mathbb{Q}$. The cubic equations in three variables become more complicated. We are interested in a special type of cubic called elliptic curves, on which there is an addition law. This makes the arithmetic of elliptic curves fruitful and interesting.

The main goal of this paper is to provide a self-contained proof of the famous Mordell-Weil theorem. In section 2, we discuss some basic properties of elliptic curves. After that we break the proof of the theorem into two distinct parts. In section 3, we prove the weak Mordell-Weil theorem. In section 4, we prove the

Date: AUGUST 16, 2021.

descent theorem, and then finish the proof of the Mordell theorem by constructing a height function on elliptic curves.

2. Basic Properties of Elliptic Curves

2.1. **Addition law on Elliptic Curves.** Let K be a field, we can define elliptic curves over the field K.

Definition 2.1. An elliptic curve over K is a pair (E, O), written as E/K, where E is a smooth projective curve over K of genus 1 and $O \in E$. The special point O is called the basepoint. Sometimes $E/\bar{K}$ is denoted as E if the field $\bar{K}$ is clear from context.

There is another useful equivalent definition.

Definition 2.1'. An elliptic curve E over K is a subvariety of $\mathbb{P}^2$ defined by a homogenous equation

$$Y^2Z + a_1XYZ + a_3YZ^2 = X^3 + a_2X^2Z + a_4XZ^2 + a_6Z^3$$

with coefficients $a_1, \ldots, a_6 \in K$. The point $O = [0, 1, 0]$ is called the point at infinity.

Such an elliptic curve consists of an affine subvariety of the affine plane $z \neq 0$ and an infinity point $[0, 1, 0]$. The Weierstrass equation of the elliptic curve is defined by

$$E : y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6. \tag{2.2}$$

To see that the above two definitions are equivalent, we need to use the Riemann-Roch theorem (see Theorem A.4). For a detailed proof, see [2, Chap. 3] or [4, Chap. 4]. For $\sigma \in \mathrm{Gal}(\bar{K}/K)$ and $P \in E$, we can define a natural Galois group action, denoted by Q^σ.

The divisor group of an elliptic curve E, denoted $\mathrm{Div}(E)$, is defined by the free abelian group generated by points of E. For a divisor $D \in \mathrm{Div}(E)$ of the form

$$D = \sum_{P \in E} n_P(P), n_P \in \mathbb{Z},$$

the degree of D is defined by

$$\deg D = \sum_{P \in E} n_P.$$

For $P \in E$, since E is a smooth curve, the local ring of E at P, denoted $K(E)_P$, is a discrete valuation ring. The normalized valuation is denoted by v_P. For $f \in K(E)$, the order of f at P is defined as

$$ord_P(f) := v_P(f).$$

If f is a non-zero function on E, that is $f \in K(E)^*$, then

$$div(f) = \sum_{P \in E} ord_P(f)(P)$$

is a divisor on the elliptic curve E. Such a divisor is called a principal divisor. The Picard group of E, denoted Pic(E), is the quotient of Div(E) by the subgroup of principal divisors.

Since $\deg(div(f)) = 0$ (Theorem A.1), the degree map factors through $Pic(E)$. Thus we get a degree map on $Pic(E)$ and $Pic^0(E)$ is defined as the preimages of 0.

Remark 2.3. Those who are familiar with algebraic geometry know that Weil divisors are defined on every variety (see [1, Chap.6]). As a special case, Weil divisors on a smooth curve are defined above. In general, the quotient of the Weil divisor group by the subgroup of principal divisors is called the divisor class group or the Weil class group. The Picard group of a variety is defined quite differently and possibly not isomorphic to its Weil divisor group. However, when dealing with smooth curves these two groups are isomorphic, so it is reasonable to regard the divisor class group as the Picard group.

We now define an addition law on E. The following theorem largely depends on the Riemann-Roch theorem, which is introduced in the Appendix A.

Theorem 2.4. *There is a bijection between sets*

$$\begin{aligned} \sigma : E &\to Pic^0(E) \\ P &\mapsto (P) - (O). \end{aligned}$$

As a consequence, this induces an addition law on E, and the basepoint O is the zero element.

Proof. Let $K(E)$ be the function field of E. We first show that σ is surjective. For every divisor D with degree 0, the Riemann-Roch theorem says that

$$\dim \mathscr{L}(D + (O)) = 1.$$

Let $f \in \mathscr{L}(D + (O))$ be a non-zero function. Then

$$div(f) \geq -D - (O).$$

Since we also have $div(f) = 0$, there is a point $P \in E$ such that

$$div(f) = -D - (O) + (P).$$

Then we have $\sigma(P) = D$.

To show that the map is injective, it suffices to show that if $(P) \sim (Q)$, then $P = Q$. Now let $(P) - (Q) = div(f)$. Consider the degree 1 map

$$[1, f] : E \to \mathbb{P}^2,$$

which is a homeomorphism between E and $\mathbb{P}^2$. This contradicts the assumption that the genus of E is 1. □

We now turn to the study of maps between two elliptic curves. Since elliptic curves are special cases of projective varieties, the map between them should be a morphism between varieties. We also hope that the map preserves the addition law.

Definition 2.5. Let E_1 and E_2 be elliptic curves. An *isogeny* between E_1 and E_2 is a morphism

$$\phi : E_1 \to E_2$$

satisfying $\phi(O) = O$.

Theorem 2.6. *Let ϕ be a an isogeny between E_1 and E_2. Then*

$$\phi(P + Q) = \phi(P) + \phi(Q).$$

Proof. Let the map ϕ_* be defined as

$$\begin{aligned} \mathrm{Pic}^0(E_1) &\to \mathrm{Pic}^0(E_2) \\ \sum_{P\in E_1} n_P(P) &\mapsto \sum_{P\in E_1} n_P(\phi(P)). \end{aligned}$$

The above map is a group homomorphism. Since $\phi(O) = O$, we have the following commutative diagram

$$\begin{array}{ccc} E_1 & \xrightarrow{\cong} & \mathrm{Pic}^0(E_1) \\ \downarrow{\scriptstyle \phi} & & \downarrow{\scriptstyle \phi_*} \\ E_2 & \xrightarrow{\cong} & \mathrm{Pic}^0(E_2). \end{array} \tag{2.7}$$

Thus ϕ is also a group homomorphism. □

For every $m \in \mathbb{Z}$, we define an isogeny

$$[m] : E \to E.$$

If $m > 0$, then $[m]P := m(P) = P + \cdots + P(m \text{ terms})$. If $m = 0$, then $[0]P := O$. If $m < 0$, then $[m]P := [-m](-P)$.

Then for every $m \in \mathbb{Z}, m \geq 2$, the m-torsion subgroup of E, denoted by $E[m]$, is the kernel of $[m]$.

2.2. **The Structure of $E[m]$.** In this subsection, we will decide the group structure of $E[m]$. This part is not necessary for the proof of the Mordell-Weil theorem; however, we will still discuss it since it is somehow related to the Lubin-Tate theory.

First note that $E[m]$ is a finite group. Suppose that E is defined by a Weierstrass equation. Then the addition law on E is given by rational functions. In this way, we can find points in $E[m]$ by solving the polynomial equations

$$\begin{aligned} f(x, y) &= 0 \\ g(x, y) &= 0. \end{aligned}$$

We can show that there are finitely many common roots via the resultant of f and g.

To decide the group structure of $E[m]$, we first need to calculate its order. From now on

$$\phi : E_1 \to E_2$$

is a non-constant isogeny of degree m. We see from (2.7) that ϕ is associated to ϕ_* between Picard groups. Recall that there is another natural homomorphism between Picard groups

$$\begin{aligned} \phi^* : \mathrm{Pic}^0(E_2) &\to \mathrm{Pic}^0(E_1) \\ \sum_{Q\in E_2} n_Q(Q) &\mapsto \sum_{Q\in E_2} n_Q \sum_{P\in\phi^{-1}(Q)} e_\phi(P)(P). \end{aligned}$$

So there is another diagram

$$\begin{array}{ccc} E_2 & \xrightarrow[\sigma_2]{\cong} & \mathrm{Pic}^0(E_2) \\ \downarrow{\scriptstyle \hat{\phi}} & & \downarrow{\scriptstyle \phi^*} \\ E_1 & \xrightarrow[\sigma_2]{\cong} & \mathrm{Pic}^0(E_1), \end{array} \tag{2.8}$$

where the map $\hat{\phi}$ is defined so as to make the above diagram commute. Since ϕ^* is the dual of ϕ_*, it is natural to say that $\hat{\phi}$ is the dual of ϕ. However, it is not obvious that $\hat{\phi}$ is an isogeny.

Theorem 2.9. *(a) There exists a unique isogeny*

$$\psi : E_2 \to E_1$$

satisfying

$$\psi \circ \phi = [m].$$

(b) The map ψ is equal to $\hat{\phi}$.

See [2, Theorem 6.1 of Chap. 3] for a detailed proof. We call $\hat{\phi}$ the dual isogeny of ϕ.

Theorem 2.10. *For any $m \in \mathbb{Z}$, we have*

$$\widehat{[m]} = [m].$$

Proof. We first prove an enhanced proposition. If ϕ and ψ are isogenies from E_1 to E_2, then

$$\widehat{\phi + \psi} = \hat{\phi} + \hat{\psi}.$$

To make the proof more clearly, we now let x, y and z, w be Weierstrass coordinates for E_1 and E_2 respectively. For a given $P = (x_1, y_1) \in E_1$, Theorem 2.4 shows that

$$D = ((\phi + \psi)(P)) - (\phi(P)) - (\psi(P)) + (O) \tag{2.11}$$

is a principle divisor of E_2. Let $F_P \in K(E_2)$ satisfying

$$div(F_P) = D.$$

By the calculation in [4, Chap. 5], we can find a rational function $f(x, y, z, w)$ such that

$$F_{(x,y)} = f(x, y, \cdot, \cdot).$$

In this way, for a given $Q = (z_1, w_1) \in E_2$, we get a function $G_{(z_1,w_1)} = f(\cdot, \cdot, z_1, w_1) \in K(E_1)$. $P = (x_1, y_1) \in E_1$ is a zero point of $G_{(z_1,w_1)}$ if and only if $Q = (z_1, w_1) \in E_2$ is a zero point of $F_{(x_1,y_1)}$. Then (2.11) shows that

$$\begin{aligned} (\phi + \psi)(P) &= Q \\ \mathrm{ord}_P(G_Q) &= e_{\phi+\psi}(P) \end{aligned}$$

For the same reason, $P = (x_1, y_1) \in E_1$ is a pole of $G_{(z_1,w_1)}$ if and only if

$$\begin{aligned} \phi(P) &= Q, \mathrm{ord}_P(G_Q) = e_\phi(P); \textit{or} \\ \psi(P) &= Q, \mathrm{ord}_P(G_Q) = e_\psi(P). \end{aligned}$$

Thus we have

$$\begin{aligned} div(G_Q) &= \sum_{P\in(\phi+\psi)^{-1}(Q)} \mathrm{ord}_P(G_Q)(P) \\ &\quad - \sum_{P\in(\phi)^{-1}(Q)} \mathrm{ord}_P(G_Q)(P) \\ &\quad - \sum_{P\in(\psi)^{-1}(Q)} \mathrm{ord}_P(G_Q)(P) \\ &= (\phi+\psi)^*((Q)) - \phi^*((Q)) - \psi^*((Q)). \end{aligned}$$

Thus for every $Q \in E_2$, (2.8) yields

$$(\widehat{\phi+\psi})(Q) - \hat{\phi}(Q) - \hat{\psi}(Q) = \sigma_1^{-1}(div(G(Q)) - div(G(O))) = O, \tag{2.12}$$

which completes the proof that $\widehat{\phi+\psi} = \hat{\phi} + \hat{\psi}$.

Since we already have $\widehat{[1]} = [1]$, by induction we have $\widehat{[m]} = [m]$. □

Finally, we get the following theorem:

Theorem 2.13. *Let E be an elliptic curve and m be a non-zero integer.*
(a) $\deg[m]=m^2$.
(b) If char(K)=0 or char(K) does not divide m, then

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

(c) If char(K)=p, then either

$$\begin{aligned} E[p^e] &\cong \{O\} \text{ for all } e \in \mathbb{N}; \text{or} \\ E[p^e] &\cong \mathbb{Z}/p^e\mathbb{Z} \text{ for all } e \in \mathbb{N}. \end{aligned}$$

Proof. (a) Let $d = \deg[m]$. Then Theorem 2.10 yields

$$[d] = \widehat{[m]} \circ [m] = [m] \circ [m] = [m^2].$$

Thus $d = m^2$.
(b) Take a non-zero differential $\omega \in \Omega_E$. Since

$$[m]^*\omega = m\omega \neq 0,$$

$[m]$ is separable (see[Proposition A.3]). Thus

$$\#E[m] = \#[m]^{-1}(O) = \deg[m] = m^2.$$

For every $d|m$, we have $\#E[m][d] = \#E[d] = d^2$. Let $E[m] = \mathbb{Z}/d_1\mathbb{Z} \oplus \cdots \oplus \mathbb{Z}/d_n\mathbb{Z}$. Then we have

$$d_1^n = \#E[d_1] = d_1^2,$$

which shows that $n = 2$. Moreover we have

$$d_1 d_2 = \#E[d_2] = d_2^2,$$

which shows that $d_1 = d_2 = m$. This completes the proof.
(c) Assume that E is defined by Weierstrass equation f. Let $f^{(p)}$ be the equation obtained from f by raising each coefficient of f to the p^{th} degree. Let $E^{(p)}$ be the elliptic curve defined by $f^{(p)}$. Then we have the p^{th} Frobenius morphism

$$\begin{aligned} \phi : E &\to E^{(p)} \\ (x, y) &\mapsto (x^p, y^p). \end{aligned}$$

Since K is perfect, Proposition A.2(b) shows that $\deg \phi = p$. While $p^{2e} = \deg[p^e] \geq \#E[p^e]$, we have

$$\begin{aligned}
\#E[p^e] &= \deg_s[p^e] && \text{Proposition A.2}(a)\\
&= \deg_s[p]^e \\
&= (\deg_s(\hat{\phi} \circ \phi))^e && \text{Theorem 2.9}\\
&= (\deg_s \hat{\phi})^e \\
&= \begin{cases} 1 & \text{if } \hat{\phi} \text{ is inseparable} \\ p^e & \text{otherwise.} \end{cases}
\end{aligned}$$

For $e = 1$, the only abelian group of order p is $\mathbb{Z}/p\mathbb{Z}$ and the proof is done. We now prove by induction on e. If the proof is done for $e-1$, let $E[p^{e-1}]$ be generated by a point $P_n \in E[p^{e-1}]$. The map $[p]$ is surjective, so there exists a point $Q \in [p]^{-1}(P_n)$. Then $Q + P_1$ is of degree p^e, therefore generates $E[p^e]$. As a conclusion, $E[p^{e-1}] \cong \mathbb{Z}/p^e\mathbb{Z}$. □

Remark 2.14. The result of Theorem 2.13(c) is analogous to some part of Lubin-Tate theory. This is not surprising since purely inseparable is somehow analogous to totally ramified, and both of them involve looking at the structure of torsion points. To be more precise, let K be a local field with a uniformizer ϖ. Let $\bar{K}$ be its algebraic closure. Then we can use Lubin-Tate formal group f to construct a $\mathcal{O}_K$-action on $\Lambda_f = \mathfrak{m}_{\mathcal{O}_{\bar{K}}}$. An important fact is that $\Lambda_f[\varpi^n] \cong \mathcal{O}_K/\varpi^n$ as $\mathcal{O}_K$ module. As a consequence, $K(\Lambda_f[\varpi^n])/K$ is totally ramified of degree $\#(\mathcal{O}_K/\varpi^n)^*$.

2.3. **Reduction of Elliptic Curves.** Let K be a number field. The following notations will be used through out this paper.

Let $S_\infty := \sum_\infty / \sim$, where elements of $\sum_\infty$ are embeddings $\sigma : K \to \mathbb{C}$ and $\sigma \sim \sigma'$ if $\sigma' = \bar{\sigma}$. An element in S_∞ is called an archimedean place or an infinite place. Each archimedean place associates to a normalized absolute value

$$|\cdot|_\sigma : K \to \mathbb{R}_{\geq 0}, x \mapsto |\sigma(x)|.$$

Let $\mathfrak{p}$ be a prime ideal of $\mathcal{O}_K$. Then we can define the discrete valuation

$$v_\mathfrak{p} : K \to \mathbb{Z} \cup +\infty$$

such that $(x) = \mathfrak{p}^{v_\mathfrak{p}(x)}\mathfrak{a}$ with $\mathfrak{a}$ a fractional ideal that does not have $\mathfrak{p}$-factor and $v_\mathfrak{p}(0) = +\infty$. Such a valuation is called a non-archimedean place or a finite place. Each non-archimedean place associates to a normalized absolute value

$$|\cdot|_\mathfrak{p} : K \to \mathbb{R}_{\geq 0}, x \mapsto N(\mathfrak{p})^{-v_\mathfrak{p}(x)},$$

where $N(\mathfrak{p}) = |\mathcal{O}/\mathfrak{p}|$. Let $K_\mathfrak{p}$ be the completion of K with respect to $|\cdot|_\mathfrak{p}$ and $k_\mathfrak{p}$ the residue field of the local field $K_\mathfrak{p}$.

An element in the set $S_\infty \cup \{v_\mathfrak{p} : \mathfrak{p} \text{ is a prime ideal of } \mathcal{O}_K\}$ is called a place of K. We can prove that the above norms are all the possible norm under equivalence (see [5, Chap. 7]).

Let L/K be a finite Galois extension and w, v places of L and K respectively. If $w|_K = v$, then we say w divides v or w is a place over v. Now let w and v be finite places which are associated to $\mathfrak{P}$ and $\mathfrak{p}$ respectively. We call w an unramified place over v if $\mathfrak{p} = \mathfrak{P}\mathfrak{a}$, where $\mathfrak{a}$ is an ideal of $\mathcal{O}_L$ that does not have $\mathfrak{P}$-factor. If every place of L over v is unramified, we say that L/K is unramified at v.

We now state some facts. For a detailed introduction, see [5, Chap. 8].

Proposition 2.15. *Let L/K be a finite Galois extension and K a number field. Suppose that v is a finite place of K and that w is a place of L over v.*
(a) $Gal(L_w/K_v)$ is isomorphic to the decomposition group D_w. Thus it can be regarded as a subgroup of $Gal(L/K)$.
(b) There is a surjective map

$$Gal(L_w/K_v) \to Gal(l_w/k_v).$$

The kernel of the map, written as $I_{w/v}$, is called the inertia group of w over v.
(c) If in addition L/K is abelian, then the inertia group $I_{w/v}$ does not depend on the choice of w via group isomorphisms. L/K is unramified at v if and only if the inertia group over v is trivial.

If K is a local field and E/K is an elliptic curve. Let $\mathfrak{m}_K$ be its maximal ideal, which is generated by an uniformizer π. Suppose that E/K is defined by a Weierstrass equation. By substituting coordinates, we can find a Weierstrass equation with coefficients in $\mathcal{O}_K$ and its discriminant, denoted Δ, satisfies that $v(\Delta)$ is the smallest possible nonnegative integer among all the Weierstrass equation for E. Such equation is called the minimal Weierstrass equation for E/K.

After choosing a minimal Weierstrass equation for E/K, we reduce its coefficients modulo π to obtain a curve over the residue field k of $\mathcal{O}_K$. That curve, written as $\tilde{E}$, is called a reduction of E. The natural map

$$E(K) \to \tilde{E}(k)$$

is called a reduction map. If $\tilde{E}$ is a smooth curve, then we says the reduction is good. We call the reduction bad otherwise.

Here is an important theorem which will be used in the proof of the weak Mordell theorem.

Theorem 2.16. *Suppose that $\tilde{E}/k$ is nonsingular, and that m is a positive integer relatively prime to $char(k)$. Then*

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

is injective. For the proof, see [2, Chap. 7].

3. Weak Mordell-Weil Theorem

From now on, our goal is to prove the famous Mordell-Weil theorem.

Theorem 3.1. *(Mordell-Weil) Let K be a number field and E/K be an elliptic curve. Then $E(K)$ is a finitely-generated abelian group.*

If Theorem 3.1 has been proved, then the structure theorem for finitely-generated abelian group shows that

$$E(K) \cong \mathbb{Z}^r \oplus \mathbb{Z}/d_1\mathbb{Z} \dots \mathbb{Z}/d_n\mathbb{Z},$$

with $r \in \mathbb{Z}_{\geq 0}$ and $d_i | d_{i+1}$ for $1 \leq i < n$. Then we can get a weaker version of Theorem 3.1.

Theorem 3.2. *(Weak Mordell-Weil) Let E/K be an elliptic curve. Then for any $m \geq 2$, $E(K)/mE(K)$ is a finite abelian group.*

In this section, we will directly prove the weak Mordell-Weil theorem.

3.1. **Kummer pairing.** $E(K)/mE(K)$ is analogous to $K^*/(K^*)^m$ which appears in Kummer thoery. Suppose that $\zeta_n \in K$ (where ζ_n is a primitive n-th root of unity). Let μ_n be the group of n-th roots. We recall that in number theory there is a short exact sequence

$$1 \longrightarrow \mu_n \longrightarrow K^* \xrightarrow{m} K^* \longrightarrow 1.$$

The associated long exact sequence for Galois cohomology yields

$$\begin{array}{c} 1 \longrightarrow \mu_n \longrightarrow K^* \xrightarrow{m} K^* \longrightarrow \\ \longrightarrow H^1(G_{\bar{K}/K}, \mu_n) \longrightarrow H^1(G_{\bar{K}/K}, K^*) \longrightarrow \cdots \end{array}$$

Here we denote $Gal(\bar{K}/K)$ as $G_{\bar{K}/K}$. Hilbert's famous "Hilbert 90" asserts that $H^1(G_{\bar{K}/K}, K^*) = 1$ and we get

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

Similarly, we do the same thing to $E(K)$. The short exact sequence

$$0 \longrightarrow E[m] \longrightarrow E(\bar{K}) \xrightarrow{[m]} E(\bar{K}) \longrightarrow 0$$

yields a long exact sequence

$$\begin{array}{c} 0 \longrightarrow E(K)[m] \longrightarrow E(K) \xrightarrow{[m]} E(K) \xrightarrow{\delta} \\ \longrightarrow H^1(G_{\bar{K}/K}, E[m]) \longrightarrow H^1(G_{\bar{K}/K}, E(\bar{K})) \xrightarrow{[m]} H^1(G_{\bar{K}/K}, E(\bar{K})). \end{array}$$

We extract the following sequence

$$0 \longrightarrow E(K)/mE(K) \xrightarrow{\delta} H^1(G_{\bar{K}/K}, E[m]) \longrightarrow H^1(G_{\bar{K}/K}, E(\bar{K}))[m] \longrightarrow 0.$$

It is likely that $E[m] \not\subseteq E(K)$, in which case elements of H^1 cannot be written explicitly. However, we have proved in section 2.2 that $E[m]$ is a finite group, so there is a finite field extension L/K such that $E[m] \subseteq E(L)$.

Lemma 3.3. *Let L/K be a finite Galois field extension. If $E(L)/mE(L)$ is finite, then $E(K)/mE(K)$ is finite.*

Proof. We notice the following commutative diagram

$$\begin{array}{ccccc} 0 & \longrightarrow & K & \longrightarrow & E(K)/mE(K) \longrightarrow E(L)/mE(L) \\ & & \downarrow & & \downarrow \quad\quad \downarrow \\ 0 & \longrightarrow & H^1(G_{L/K}, E[m]) & \xrightarrow{\text{Inf}} & H^1(G_{\bar{K}/K}, E[m]) \longrightarrow H^1(G_{\bar{K}/L}, E[m]). \end{array}$$

where the second row is the inflation-restriction exact sequence. The short 5-lemma yields the first injection. Since $H^1(G_{L/K}, E[m])$ is a finite group, K is also a finite group. So the first row of the diagram gives the desired result. □

Now suppose that $E[m] \subseteq E(K)$. Then we have

$$H^1(G_{\bar{K}/K}, E[m]) = Hom(G_{\bar{K}/K}, E[m]).$$

The explicit description of δ shows that

$$\delta : E(K)/mE(K) \to Hom(G_{\bar{K}}/K), E[m]) \tag{3.4}$$

$$P \mapsto (\sigma \mapsto Q^\sigma - Q), \text{ where } Q \text{ satisfies } mQ = P. \tag{3.5}$$

Now, consider the pairing

$$\kappa : E(K)/mE(K) \times G_{\bar{K}/K} \to E[m].$$

(3.4) shows that the left kernel of the pairing is trivial. If $\tau \in G_{\bar{K}/K}$ is in the right kernel, then for any $Q \in E$ satisfying $mQ \in E(K)$ we have $Q^\sigma = Q$. The converse is also true. So the right kernel is $G_{\bar{K}/L}$, where L is the composition of all fields $K(Q)$ over all $Q \in E(K)$ such that $mQ \in E(K)$.

By abuse of notation, we still use κ to denote the perfect pairing

$$\kappa : E(K)/mE(K) \times G_{L/K} \to E[m].$$

This is called the Kummer pairing.

3.2. **The Extension L/K is Finite.** In the previous subsection, we saw that $E(K)/mE(K) \cong \mathrm{Hom}(G_{L/K}, E[m])$. Our goal is to show that $E(K)/mE(K)$ is a finite group, so it suffices to prove that L/K is a finite field extension.

Lemma 3.6. *(a) L/K is an abelian extension of exponent m. (i.e $G_{L/K}$ is an abelian group such that the order of each element divides m.)*
(b) Let

$$S := S_\infty \cup \{E(K) \text{ has bad reduction at } v\} \cup \{v(m) \neq 0\}.$$

be a finite set of places of K. Then L/K is unramified outside S.

Proof. (a) Let $\sigma \in G_{L/K}$. Then

$$\kappa(P, \sigma^m) = m \cdot \kappa(P, \sigma) = 0$$

for every $P \in E(K)$. The fact that the Kummer paring is perfect yields $\sigma^m = 0$.
(b) Let v be a place of K outside S. Recall that if L_1/K and L_2/K are both unramified over v, then L_1L_2/K is unramified over v. It suffices to prove that $K(Q)/K$ is unramified over v, where $Q \in E(\bar{K})$ satisfying $mQ \in E(K)$.

Let $L' = K(Q)$. w is a place of L' over v. Proposition 2.15 says it suffices to prove that $I_{w/v}$ is trivial. If $\tau \in I_{w/v}$, then τ acts trivially on $\tilde{E}(k_v)$. Thus

$$\widetilde{Q^\tau - Q} = \tilde{O}. \tag{3.7}$$

$Q^\tau - Q$ is a point in E satisfying

$$m(Q^\tau - Q) = (mQ)^\tau - mQ = 0.$$

The last equation follows from $mQ \in E(K)$ and that τ can be regarded as an element in $G_{L'/K}$. Hence

$$Q^\tau - Q \in E[m] \subseteq E(K)[m] \subseteq E(K_v)[m].$$

Now, Theorem 2.16 together with (3.7) shows that $Q^\tau - Q = O$, i.e. τ acts trivially on L'. □

Lemma 3.8. *Let S be a finite set of places of K. If L/K is abelian of exponent m and L/K is unramified outside S, then L/K is finite.*

Proof. We can first reduce to the case where $\zeta_m \in K$ (where ζ_m is a primitive $m-$th root of unity) because $L(\zeta_m)/K(\zeta_m)$ is finite induces L/K is finite.

Increasing the set S would only make the field L become larger, so we can enlarge the set S. Let the class group of K be $C_K = \{\mathfrak{a}_1, \cdots, \mathfrak{a}_n\}$. Here we use the finiteness of class number (see [5, Chap. 4]). For any $\mathfrak{p}$ which satisfies that there exists some $\mathfrak{a}_i \subseteq \mathfrak{p}$, add the place corresponding to $\mathfrak{p}$ into the set S. Then the ring

$$\mathcal{O}_{K,S} := \{x \in K : x \in \mathcal{O}_{K_v} \text{ for every } v \notin S\},$$

which is called the ring of S-integers in K, is a principal ideal domain. In fact, if $\mathfrak{a}$ is an ideal of $\mathcal{O}_{K,S}$, then there is an ideal of $\mathcal{O}_K$ called $\mathfrak{b}$ such that

$$\mathfrak{a} = \mathcal{O}_{K,S} \cdot \mathfrak{b}.$$

Let

$$\mathfrak{b} = \alpha \cdot \prod_{i=1}^{n} \mathfrak{a}_i^{e_i}$$

with $e_i \in \mathbb{Z}$ and $\alpha \in K^*$. Then since we have enlarged the set S, we have

$$\prod_{i=1}^{n} \mathfrak{a}_i^{e_i} \mathcal{O}_{K,S} = \mathcal{O}_{K,S}.$$

Thus $\mathfrak{a} = \alpha \mathcal{O}_{K,S}$.

Now since $\mu_m \subset K$, Kummer theory says that the maximum abelian extension of exponent m over K is $K(\sqrt[m]{a} : a \in K)$. For $a \in K^*$, if $\sqrt[m]{a} \in L$, then $K(\sqrt[m]{a})/K$ is unramified outside S. That is, $K_v(\sqrt[m]{a})/K_v$ is unramified for every $v \notin S$. Then we have

$$ord_v(a) = ord_v(x^m) = m \cdot ord_v(x) \equiv 0 \mod m.$$

The second equation follows from the unramification of $K_v(\sqrt[m]{a})/K_v$.

Now it remains to prove that

$$T_S := \{a \in K^*/(K^*)^m : ord_v(a) \equiv 0 \mod m \text{ for every } v \notin S\}$$

is a finite set. If it were proved, then $L \in K(\sqrt[m]{a} : a \in T_S)$ would yield the finiteness of L/K. We notice that there is a natural map

$$\phi : \mathcal{O}_{K,S}^* \to T_S$$

that factors through $(O)_{K,S}^*/(\mathcal{O}_{K,S}^*)^m$. Dirichlet S-units theorem [see 3, Chap. 5] says that $\mathcal{O}_{K,S}^*$ is finitely generated, so it remains to prove that ϕ is surjective.

For $a \in T_S$, let $a = \frac{b}{c}$ with $b, c \in \mathcal{O}_{K,S}$ and $\phi(b), \phi(c) \in T_S$. Then $c\mathcal{O}_{K,S}$ is the m^{th}-power of an ideal of $\mathcal{O}_{K,S}$. Since $\mathcal{O}_{K,S}$ is principal ideal domain, there exists $c_1 \in \mathcal{O}_{K,S}$ satisfying

$$c\mathcal{O}_{K,S} = c_1^m \mathcal{O}_{K,S}.$$

For the same reason

$$b\mathcal{O}_{K,S} = b_1^m \mathcal{O}_{K,S}, b_1 \in \mathcal{O}_{K,S}.$$

Thus

$$ac_1^m \mathcal{O}_{K,S} = b\mathcal{O}_{K,S} = b_1^m \mathcal{O}_{K,S}.$$

Then $a = (b_1 c_1^{-1})^m \cdot u$ for some $u \in \mathcal{O}_{K,S}^*$, and we have $\phi(u) = a$. □

Hence, Lemma 3.6 and Lemma 3.8 show that L/K is finite extension.

4. Proof of the Mordell-Weil Theorem

4.1. **Descent.**

Theorem 4.1. *(Descent Theorem) Let A be an abelian group. A height function is a map*

$$h : A \to \mathbb{R}$$

satisfying the following three properties:
(1)For each $Q \in A$, there is a constant $C_1(A, Q)$ depending on A and Q, such that

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

for any $P \in A$.
(3)There is an integer $m \geq 2$ and a constant $C_2(A)$ only depending on A, such that

$$h(mP) \geq m^2 h(P) - C_2.$$

(3)For every constant C_3, the set

$$\{P \in A : h(P) \leq C_3\}$$

is finite. Now if A/mA is a finite group (the m here coincides with that in the second property), then A is finitely generated.

Proof. Let

$$A/mA = \{\bar{Q_1}, \cdots, \bar{Q_r}\}, Q_i \in A.$$

For $P \in A$, let

$$\begin{aligned} P &= mP_1 + Q_{i_1} \\ P_1 &= mP_2 + Q_{i_2} \\ &\cdots \\ P_{n-1} &= mP_n + Q_{i_n}. \end{aligned}$$

Then for $2 \leq j \leq n-1$,

$$\begin{aligned} h(P_j) &\leq \frac{1}{m^2}(C_2 + h(mP_j)) \\ &= \frac{1}{m^2}(C_2 + h(P_{j-1} - Q_{i_j})) \\ &\leq \frac{1}{m^2}(C_2 + 2h(P_{j-1}) + C_1) \\ &= \frac{2}{m^2}h(P_{j-1}) + C, C \text{ is a constant} \\ &\leq \frac{1}{2}h(P_{j-1}) + C. \end{aligned}$$

$\frac{1}{2^n}h(P) < 1$ for a sufficiently large n. Then P is a linear combination of $Q_1, \cdots, Q_r$ and P_n. Since $h(P_n) \leq 1 + 2C$, the third property yields that the number of such P_n is finite. Thus A is generated by $\{Q_1, \cdots, Q_r\} \cup \{P : h(P) \leq 1 + 2C\}$. □

Remark 4.2. The procedure of the proof is kind of like an iteration of Euclid division, and the second property of height function plays the role in contracting the height of the remainder.

Take $m = 2$, since the Weak Modell-Weil theorem have shown that $E(K)/2E(K)$ is finite, our last task is to construct a height function for $E(K)$.

4.2. **Heights on elliptic curves.** Recall that we have defined normalized absolute values on K in Section 1.3. The set of standard absolute values on K, written as M_K, consists of all absolute values on K whose restriction on Q is a normalized absolute value.

Here is a well known theorem that we will used.

Theorem 4.3. *(Product Formula) Let K be a number field. Then for any $x \in K^*$, we have*

$$|x|_K := \prod_{v:\in M_K} |x|_v^{n_v} = 1.$$

In fact, $|x|_K$ coincides with the norm defined on I_K, the group of idèles.

We now define the height of a point in projective spaces.

Definition 4.4. Let $P \in \mathbb{P}^N(K)$ be a point with homogeneous coordinates

$$P = [x_0, \cdots, x_N].$$

The height of P is defined as

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

This is well-defined because of the following reasons.
(1)If $\lambda \in K^*$, then

$$\prod_{v\in M_K} \max\{|\lambda x_0|_v, \cdots, |\lambda x_N|_v\}^{n_v} = |\lambda|_K \cdot H_K(P) = H_K(P)$$

Thus $H_K(P)$ is independent of the choice of coordinates.
(2)There is $0 \le j \le N$ such that

$$|x_j|_v = \max\{|\lambda x_0|_v, \cdots, |\lambda x_N|_v\}^{n_v}$$

for all but finitely many $v \in M_K$. So the infinite product is convergent.

Remark 4.5. H_K is a generalization of a natural way of defining height on Q. To be more precise, take $P \in \mathbb{P}^N(Q)$ and let $[x_0, \cdots, x_N]$ be coordinates of P satisfying $X_i \in \mathbb{Z}$ and $\gcd(x_0, \cdots, x_N) = 1$. Then $H_{\mathbb{Q}}(P) = \max\{|x_0|, \cdots, |x_N|\}$

Proposition 4.6. *Let $P \in \mathbb{P}^N(K)$, and L/K is a finite extension. Then*

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

Consequently, let $P \in \mathbb{P}^N(\bar{Q})$ and suppose that $P \in \mathbb{P}^N(K)$ for some finite field extension K/Q. Then

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

is well-defined.

Proof.

$$\begin{aligned} H_L(P) &= \prod_{w\in M_L} \max_{1\le i\le N}\{|x_i|_w\}^{n_w} \\ &= \prod_{v\in M_K} \prod_{w\in M_L, w|v} \max_{1\le i\le N}\{|x_i|_v\}^{n_w} \\ &= \prod_{v\in M_K} (\max_{1\le i\le N}\{|x_i|_v\})^{\sum_{w\in M_L, w|v} n_w} \\ &= H(P)^{[L:K]}. \end{aligned}$$

The last equation follows from the fact that

$$\sum_{w\in M_L, w|v} n_w = [L:K]$$

for every $v \in M_K$. □

Notation 4.7. For $x \in K$, $H_K(x) := H_K([x,1])$.

Proposition 4.8. *(a) Let $\sigma \in G_{K/Q}$, then*

$$H_K(\sigma(x)) = H_K(x).$$

(b) Let

$$f(T) = a_0T^d + a_1T^{d-1} + \cdots + a_d = (T-\alpha_1)\cdots(T-\alpha_d)$$

be a polynomial in $\bar{Q}(T)$. Then

$$H([1, a_1, \cdots, a_d]) \leq 2^{d-1}\prod_{j=1}^{d} H(\alpha_j).$$

That means if we have a polynomial, then the height of coefficients is controlled by the height of roots.

Proof. (a) This follows from the fact that

$$|\sigma(x)|_v = |x|_v.$$

(b) It suffices to prove that for every $v \in M_K$, we have

$$\max_{0\leq i\leq N}\{|a_i|_v\} \leq 2^{d-1}\prod_{j=2}^{d}\max\{|a_j|_v, 1\}.$$

We prove by induction on d. For $d = 1$, it is obvious. Assume that we know the result for polynomials with degree less than $d-1$. Without loss of generality, let

$$|\alpha_1|_v = \max_{0\leq j\leq d}\{|a_j|_v\}.$$

Let

$$g(T) = \prod_{i=2}^{d}(T-\alpha_i) = b_0T^{d-1} + b_1T^{d-2} + \cdots + b_{d-1}.$$

We have $a_i = b_i - \alpha_1 b_{i-1}$ (where $b_d = b_{-1} = 0$).

$$\begin{aligned}\max_{0\leq i\leq d}\{|a_i|_v\} &\leq 2\max_{0\leq i\leq d}\{|b_i|, |\alpha_1 b_{i-1}|_v\}\\ &= 2\max_{0\leq i\leq d-1}\{|b_i|_v\}\cdot\max\{|\alpha_1|_v, 1\}\\ &= 2^{d-1}H(\alpha_1)\cdot\prod_{j=2}^{d}H(\alpha_j).\end{aligned}$$

The last equation follows from the induction hypothesis. □

From its definition, we expect the height behaves multiplicatively. However, to satisfy the first two properties, it is more reasonable to expect the height function behaves additively. So we define a map

$$h : \mathbb{P}^N(\bar{Q}) \to \mathbb{R}, h(P) = \log H(P).$$

Every non-constant function $f \in K(E)$ induces a surjective morphism

$$\begin{aligned} f : E &\to \mathbb{P}^1 \\ P &\mapsto \begin{cases} [1,0] & \text{if } P \text{ is a pole of } f \\ [f(P),1] & \text{otherwise.} \end{cases} \end{aligned}$$

By abuse of notation, we still denote it as f.

We define the height function on E (relative to f) as

$$\begin{aligned} h_f : E(\bar{K}) &\to \mathbb{R} \\ P &\mapsto h(f(P)). \end{aligned}$$

It remains to prove that h_f is a height function. The third property is the easiest to verify.

Theorem 4.9. *Let C be a constant . Then the set*

$$\{P \in \mathbb{P}^N(K) : H_k(P) \leq C\}$$

contains only finitely many points.

Proof. Let $P \in \mathbb{P}^n(K)$. Take a homogenous coordinates $[x_0, \cdots, x_N]$ of P with some $x_j = 1$. Then we have

$$\begin{aligned} H_K(P) &= \prod_{v \in M_K} \max\{|x_0|_v, \cdots, |x_N|_v\}^{n_v} \\ &\geq \max_{0 \leq i \leq N} \prod_{v \in M_K} \max\{|x_i|_v, 1\}^{n_v} \\ &= \max_{0 \leq i \leq N} H_K(x_i). \end{aligned}$$

So it suffices to prove that the set

$$S = \{x \in K : H_K(x) \leq C\}$$

contains finitely many elements.

Let $d = [K : Q]$. Suppose that $x \in S$ and that $x = x_1, \cdots, x_e$ are conjugates of x, $e \leq d$. Then consider the polynomial

$$f(T) = \prod_{i=1}^{e} (T - x_i) = T^d + a_1 T^{d-1} + \cdots + a_e$$

which has coefficients in $\mathbb{Q}$. Proposition 4.8 then yields

$$\begin{aligned} H([1, a_1, \cdots, a_e]) &\leq 2^{d-1} \prod_{i=1}^{d} H_K(x_i) \\ &= 2^{d-1} H_K(x)^e \\ &\leq 2^{d-1} C^d, \end{aligned}$$

which means that there is a polynomial $f(T) \in \mathbb{Z}[T]$ with coefficients less or equal to $2^{d-1}C^d$ and with degree less or equal to d, satisfying $f(x) = 0$. The number of such non-constant polynomials is finite, which yields the finiteness of S. □

We now explain why we say that the height function h behaves additively.

Theorem 4.10. *For all $P, Q \in E(K)$, we have*

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

where the remainder $O(1)$ only depends on f and E.

Assume that E is defined by a Weierstrass equation, we first reduce to the case when f is equal to x. Then we check the relationship between coordinates by some straight calculations. For the detailed proof, see [2, Theorem 6.2 of Chap. 8].

Corollary 4.11. *h_f satisties the first and the second properties.*

Proof. For the first one, Theorem 4.10 yields

$$h_f(P+Q) \leq 2h_f(P) + 2h_f(Q) + O(1) = 2h_f(P) + C_1.$$

For the second one, Theorem 4.10 yields

$$h_f([2]P) = h_f(P+P) = 2h_f(P) + 2h_f(P) + O(1) = 4h_f(P) + C_2.$$

□

We have now completed the proof of the Mordell-Weil thoerem.

Appendix A. Basic results for Algebraic Geometry

We recall some fundamental results from algebraic geometry. For a detailed introduction, see[1,Chap. 1]. In this section, the field K is always algebraically closed and the curve is always smooth.

Theorem A.1. *Let C be a curve. For any $f \in K(C)$, $ord(div(f)) = 0$.*

For the proof, see [1, Chap.2, Corollary 6.10].

Let C_1, C_2 be two K-projective curves and ϕ a nonconstant morphism from C_1 to C_2. The function field of C_1(resp. C_2) is denoted as $K(C_1)$ (resp. $K(C_2)$).

Let $Q = \phi(P)$. If $t \in K(C_2)$ is an uniformizer of $K(C_2)_Q$, then

$$e_\phi(P) := ord_P(\phi^*(t))$$

is called the ramification index of ϕ at P.

Proposition A.2. *(a)$K(C_1)/\phi^*(K(C_1))$ is a finite extension, the degree of the extension, denoted by $\deg \phi$, is called the degree of morphism ϕ. The separable and inseparable degree of the extension are denoted by $\deg_s \phi$ and $\deg_i \phi$ respectively.*
(b)For every $Q \in C_2$,

$$\sum_{P \in \phi^{-1}(Q)} e_\phi(P) = \deg \phi.$$

(c)If furthermore, ϕ is a isogeny between elliptic curves E_1 and E_2, then for every $Q \in E_2$

$$\#\phi^{-1}(Q) = \deg_s \phi.$$

For a curve C, the space of differential forms on C is denoted by Ω_C. And we now have the following result.

Proposition A.3. *ϕ is separable if and only if the map*

$$(\phi)^* : \Omega_{C_2} \to \Omega_{C_2}$$

is non-zero.

Let $P \in C$ and $\omega \in \Omega_C$. If $t \in K(C)$ is an uniformizer of $K(C)_P$, then there exists a unique function $g \in K(C)$ such that $\omega = gdt$ at P. Then we define $ord_P(\omega) = ord_P(g)$. It can be proved that $ord_P(\omega)$ is independent of the choice of t. The divisor associated to ω is

$$div(\omega) = \sum_{P \in C} ord_P(\omega)(P).$$

For any non-zero $\omega \in \Omega_C$, $div(\omega)$ is called a canonical divisor, denoted by K_C.

A divisor $D = \sum n_P(P)$ is positive, denoted by $D \geq 0$ if $n_P \geq 0$ for every $P \in C$. If D_1, D_2 are divisors such that $D_1 - D_2$ is positive, then we write as $D_1 \geq D_2$.

Given a divisor D, we now define a set of functions

$$\mathscr{L}(D) = \{f \in K(C)^* : div(f) + D \geq 0\} \cup \{0\}.$$

This is a finite-dimensional K-vector space, and we use $l(D)$ to denote its dimension. We now state the famous Riemann-Roch theorem.

Theorem A.4. *(Riemann-Roch) Let C be a smooth curve and K_C a canonical divisor on C. The integer g is the genus of C. Then for every divisor $D \in Div(C)$,*

$$l(D) - l(K_C - D) = \deg D - g + 1.$$

Acknowledgments

I would like to thank my mentor Eric for meeting with me weekly and providing guidance throughout the summer. He introduced me various topics that he thought I might be interested in, which largely broaden my horizon.

Also, I want to give a lot of thanks to Professor May and Yueshi Hou for helping me revising this paper. Though the first draft of the paper is hard to read because of my difficulties in English, they read it thoroughly and pointed out its mistakes and imperfections, which helped me a lot to improve it. Finally, I want to thank Professor May for organizing the REU program and giving me the opportunity to participate in it.

References

[1] Robin Hartshorne. *Algebraic geometry*. Springer, 2013.
[2] Silverman, Joseph H. *The Arithmetic of Elliptic Curves*. Springer, 2009.
[3] S. Lang. *Algebraic Number Theory*, Springer, 1994.
[4] Griffiths, Phillip A. *Introduction to Algebraic Curves*. American Mathematical Society, 1989.
[5] Milne, James S. Algebraic Number Theory, https://www.jmilne.org/math/CourseNotes/ANT.pdf
[6] Aleksander Horawa. Mini course: elliptic curves, http://www-personal.umich.edu/~ahorawa/ec_mini_course_new.pdf