Indian Institute of Technology, Kanpur

MTH393-Undergraduate Project

Project Report
Mordell-Weil Theorem

Submitted by
Ajay Prajapati
Roll No- 17817063
Dept. of Mathematics and Statistics
Indian Institute of Technology, Kanpur

Under the guidance of
Dr. Somnath Jha
Dept. of Mathematics and Statistics
Indian Institute of Technology, Kanpur

May 2021

ANNEXURE-II

DECLARATION

I hereby declare that the work presented in the project report entitled "**Mordell-Weil Theorem**" contains my own ideas in my own words. At places, where ideas and words are borrowed from other sources, proper references, as applicable have been cited. To the best of my knowledge this work does not emanate from or resemble other work created by person(s) other than mentioned herein.

Name: Ajay Prajapati

Date: 11-05-2021

ABSTRACT

This report is the culmination of a semester-long reading project on Arithmetic of Elliptic curves under the guidance of Dr. Somnath Jha, IIT Kanpur. The theory of elliptic curves lies in the center of what we understand and what we don't understand. The field is a very rich branch of mathematics and once a famous mathematician said that it is possible to write endlessly on elliptic curves. Yet the field is full deep conjectures and even very fundamental results are unknown. In this report, we look at one of the (proven)fundamental results about elliptic curves, the Mordell-Weil Theorem and ideas and tools that go into its proof.

This report is expository in nature and no new result is being claimed.

Contents

1 Introduction

Elliptic curves are non-singular curves of genus 1. It can be proved that every such curve can be given by a so called Weierstrass equation. Now curves given by Weierstrass equation have an amazing property that their points forms a group. The group law can be described geometrically very easily which consists of drawing lines and taking intersection with curve so it is called chord-tangent law.

Mordell-Weil theorem is one of the fundamental results in the theory of Elliptic curves which says that when an elliptic curve is defined over a number field K, then its group of K-rational points is finitely generated. So there are only finite number of $K-$rational points which generate all other $K-$rational point using the chord tangent law. It was proved by Louis Mordell for $\mathbb{Q}$ and generalized by Andre Weil to general number fields.

In this report we define elliptic curves, prove that every elliptic curve can be given by a Weierstrass equation, define its group law and prove the Mordell-Weil theorem for $\mathbb{Q}$. The most difficult part in proving the theorem is a result called Weak Mordell-Weil theorem which we prove in full generality. From here, with a little more effort, one can prove the Mordell-Weil theorem for arbitrary number fields.

2 Preliminaries

Notation which will be used throughout this note:

K	a perfect field
$\bar{K}$	a fixed algebraic closure of K
$Gal(\bar{K}/K)$	the Galois group of $\bar{K}/K$

Here we state some standard results related to curves (algebraic varieties of dimension 1). *The results of this section will be used to prove the equivalence between elliptic curves and non-singular Weierstrass equations.*(Theorem 3.11). The main result of this section is the *Riemann-Roch theorem* and its consequences

Here we assume familiarity with basic algebraic geometry (Silverman [2008], Chapter 1). Let C be a curve and $P \in C$. Following notation will be used throughout this section:

C/K	a curve defined over K
$\bar{K}(C)$	the function field of C over $\overline{K}$
$K(C)$	the function field of C over K

2.1 Map between curves

Proposition 2.1. *Let C be a curve, $V \subset \mathbb{P}^N$ be a variety, $P \in C$ be a non-singular point and $\phi : C \longrightarrow V$ be a rational map. Then ϕ is regular at P. In particular, if C is a non-singular curve then ϕ is a morphism.*

Theorem 2.2. *Let $\phi : C_1 \to C_2$ be a morphism of curves. Then ϕ is either constant or surjective.*

2.2 Divisors

Let C be a non-singular curve for the rest of this section.

Definition 2.3. The *divisor group of C*, denoted by $div(C)$ is the free abelian group generated by points on C. i.e. $div(C) = \oplus_{P \in C} \mathbb{Z}(P)$.

Definition 2.4. To every $f \in \bar{K}(C)^*$, we associate a divisor to f given by

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

Definition 2.5. A divisor $D \in Div(C)$ is called *principal* if $\exists f \in \bar{K}(C)^*$ s.t. $D = div(f)$. The set of principal divisors are denoted by $Princ(C)$. Two divisors are called *linearly dependent* if their difference is principal. If D_1 and D_2 are linearly dependent then we write $D_1 \sim D_2$.

Definition 2.6. The *divisor class group* (or *Picard group*) of C, denoted by $Pic(C)$ is $Div(C)/Princ(C)$.

Proposition 2.7. *Let C be a non-singular curve and $f \in \bar{K}(C)^*$.*

(a) div(f)=0 $\iff$ $f \in \bar{K}^*$.

(b) deg(div(f))=0.

2.3 Differentials

Definition 2.8. Let C be a curve. The *space of (meromorphic) differential forms* on C, denoted by Ω_C is the $\bar{K}$-vector space generated by symbols of the form dx where $x \in \bar{K}(C)$ subject to relations

(a) $d(x + y) = dx + dy$ for all $x, y \in \bar{K}(C)$,

(b) $d(xy) = xdy + ydx$ for all $x, y \in \bar{K}(C)$,

(c) $da = 0$ for all $a \in \bar{K}$.

Proposition 2.9. *The space Ω_C is a one dimensional $\bar{K}(C)$-vector space.*

Proposition 2.10. *Let $P \in C$, and let $t \in \bar{K}(C)$ be a uniformizer at P.*

(a) For every $\omega \in \Omega_C$, there exists a unique function $g \in \bar{K}(C)$, depending on ω and t, satisfying $\omega = gdt$. We denote g by ω/dt.

(b) Let $f \in \bar{K}(C)$ be regular at P. Then df/dt is also regular at P.

(c) Let $\omega \in \Omega_C$. The quantity $ord_P(\omega/dt)$ is independent of choice of uniformizer t. We call this value the order of ω at P and is denoted by $ord_P(\omega)$.

Definition 2.11. Let $\omega \in \Omega_C$. The *divisor associated to ω* is

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

Definition 2.12. The differential ω is called *regular* or *holomorphic* if $ord_P(\omega) \geq 0$ for all $P \in C$. It is called *non-vanishing* if $ord_P(\omega) \leq 0$ for all $P \in C$.

Remark 2.13. Let $\omega_1, \omega_2 \in \Omega_C$ be non-zero differentials, then the proposition 2.9 implies that $\exists$ a function $f \in \bar{K}(C)^*$ s.t. $\omega_1 = f\omega_2$. Hence $div(\omega_1) \sim div(\omega_2)$ and the following definition makes sense.

Definition 2.14. The *canonical divisor class* on C is the image in $Pic(C)$ of $div(\omega)$ for any non-zero differential $\omega \in \Omega_C$. Any divisor in this divisor class is called a *canonical divisor*.

2.4 The Riemann-Roch Theorem

Definition 2.15. A divisor $D = \sum n_P(P)$ is called *positive* (or *effective*), denoted by $D \succcurlyeq 0$, if $n_P \geq 0$ for every $P \in C$. For $D_1, D_2 \in Div(C)$, we write $D_1 \succcurlyeq D_2$ to indicate $D_1 - D_2 \succcurlyeq 0$.

Example 2.16. Let $f \in \bar{K}(C)^*$ which is regular everywhere except at $P \in C$ where it has pole of order atmost n. These requirements on f can be written as $div(f) \succcurlyeq -n(P)$. Similarly, $div(f) \succcurlyeq (Q) - n(P)$ says that in addition, f has a zero at Q.

Definition 2.17. Let $D \in Div(C)$. We associate to D the set of functions,

$$\mathcal{L}(D) = \{f \in \bar{K}(C)^* : div(f) \succcurlyeq -D\} \bigcup \{0\}$$

It can be proven that $\mathcal{L}(D)$ forms a $\bar{K}$-vector space and its dimension over $\bar{K}$ is denoted by $\ell(D)$.

Proposition 2.18. *Let $D \in Div(C)$.*

(a) If $deg(D) < 0$ then $\mathcal{L}(D) = \{0\}$ and $\ell(D) = 0$.

(b) If $D' \in Div(C)$ s.t. $D' \sim D$. Then $\mathcal{L}(D) \cong \mathcal{L}(D')$ and so $\ell(D) = \ell(D')$.

Proof. (a) If $f \in \mathcal{L}(D)$ with $f \neq 0$, then $div(f) \succcurlyeq -D$. This implies $deg(div(f)) \geq -deg(D)$. Now use 2.7(b) to conclude that $deg(D) \geq 0$.

(b) Let $D = D' + div(g)$, then $\mathcal{L}(D) \longrightarrow \mathcal{L}(D'), f \mapsto fg$ is an isomorphism. □

Example 2.19. Let $K_C = div(\omega)$ be a canonical divisor. If $f \in \mathcal{L}(K_C)$, then

$$div(f) \succcurlyeq -div(\omega) \implies div(f\omega) \succcurlyeq 0$$

So $f\omega$ is holomorphic. Conversely, if $f\omega$ is holomorphic then $f \in \mathcal{L}(K_C)$. Hence

$$\mathcal{L}(K_C) = \{\omega \in \Omega_C : \omega \text{ is holomorphic}\}$$

Now we state one of the fundamental results in algebraic geometry.

Theorem 2.20. ***(Riemann-Roch)*** *Let C be a non-singular curve and K_C be a canonical divisor. Then there is an integer $g \geq 0$ s.t. for every $D \in Div(C)$,*

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

Corollary 2.21. *(a) $\ell(K_C) = g$*

(b) $deg(K_C) = 2g - 2$

(c) If $deg(D) \geq 2g - 2$, *then* $\ell(D) = deg(D) - g + 1$.

Proof. (a) Use theorem 2.20 with $D = 0$. Note $\ell(0) = 1$

(b) Use theorem 2.20 with $D = K_C$.

(c) From (b) we have $deg(K_C - D) < 0$. Now use proposition 2.18(a).

□

3 Geometry of Elliptic curves

In this section, we define *elliptic curves*, our main objects of study and see some of their basic properties.

3.1 Weierstrass equations

Definition 3.1. A *Weierstrass equation* is a homogenous equation of the form

$$Y^2Z + a_1XYZ + a_3YZ^2 = X^3 + a_2X^2Z + a_4XZ^2 + a_6Z^3 \tag{1}$$

where $a_1, ..., a_6 \in \bar{K}$.

Elliptic curves are curves of genus one with a specified base point. We will see later that every such curve has a Weierstrass equation with specified point $O = [0, 1, 0]$ on the line at ∞.

To ease the notation, we generally write dehomogenized equation of E

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

while always remembering that there is an extra point $O = [0, 1, 0]$. If $a_1, ..., a_6 \in K$, we say that E is defined over K.

If $char(K) \neq 2$, then we can simplify equation by completing square. Thus the substitution $y \longmapsto (y - a_1x - a_3)/2$ gives (where $b_2 = a_1^2 + 4a_4$, $b_4 = 2a_4 + a_1a_3$ and $b_6 = a_3^2 + 4a_6$)

$$E : y^2 = 4x^3 + b_2x^2 + 2b_4x + b_6$$

Useful quantities associated to E are

$$\begin{aligned}
b_8 &= a_1^2a_6 + 4a_2a_6 - a_1a_3a_4 + a_2a_3^2 - a_4^2, \\
c_4 &= b_2^2 - 24b^4, \\
c_6 &= -b_2^3 + 36b_2b_4 - 216b_6, \\
\Delta &= -b_2^2b_8 - 8b_4^3 - 27b_6^2 + 9b_2b_4b_6, \\
j &= c_4^3/\Delta \\
\omega &= \frac{dx}{2y + a_1x + a_3} = \frac{dy}{3x^2 + 2a_2x + a_4 - a_1y}
\end{aligned}$$

Definition 3.2. Δ is called the *discriminant* of the Weierstrass equation 1, j is called the *j-invariant* and ω is called *invariant differential* of the elliptic curve.

Let $P = (x_0, y_0)$ be a point satisfying a Weierstrass equation

$$f(x, y) = y^2 + a_1xy + a_3y - x^3 - a_2x^2 - a_4x - a_6$$

and assume that P is singular point of the curve $f(x, y) = 0$. Then

$$\frac{\partial f}{\partial x}(P) = \frac{\partial f}{\partial y}(P) = 0$$

It follows that there are $\alpha, \beta \in \bar{K}$ s.t. the Taylor series expansion of $f(x, y)$ at P

$$f(x, y) - f(x_0, y_0) = ((y - y_0) - \alpha(x - x_0))((y - y_0) - \beta(x - x_0)) - (x - x_0)^3$$

Definition 3.3. With notation as above, the singular point P is a *node* if $\alpha \neq \beta$. In this case, the lines $(y - y_0) = \alpha(x - x_0)$ and $(y - y_0) = \beta(x - x_0)$ are the *tangent lines* at P. If $\alpha = \beta$, then we say P is a *cusp*.

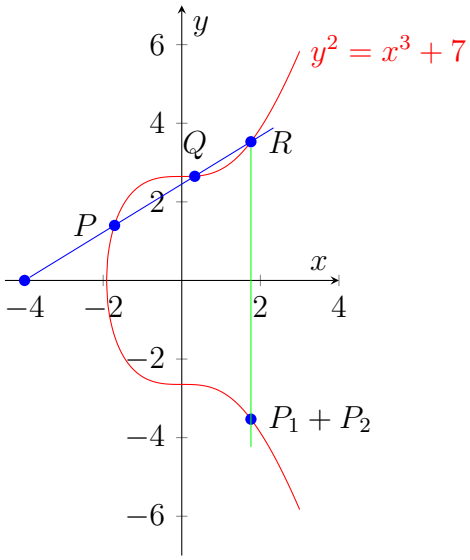
Proposition 3.4. *The curve given by a Weierstrass equation is non-singular* $\iff$ $\Delta = 0$.

Proposition 3.5. *The invariant differential associated to a non-singular Weierstrass equation is holomorphic and non-vanishing. i.e.* $div(\omega) = 0$.

3.2 The group law

Let $E \subset \mathbb{P}^2$ be an elliptic curve given by Weierstrass equation, $P = (x, y) \in E$ and $L \subset \mathbb{P}^2$ be a line. Since equation has degree 3, the line L intersects E at exactly 3 points say P, Q, R. Of course, when L is tangent then P, Q, R need not be distinct. We define a composition law $+$ given by the following:

Definition 3.6. (Composition law) Let $P, Q \in E$, let L be line through P and Q (if $P = Q$ then L is tangent at P) and let R be the third point of intersection of L with E. Let L' be line through R and O which intersects E at a third point. We denote third point by $P + Q$.



Proposition 3.7. *The composition law has following properties:*

(a) $P+O=P$ *for all* $P \in E$.

(b) $P+Q=Q+P$ *for all* $P, Q \in E$.

(c) Let $P \in E$, *then there exists a point on* E, *denoted by* $-P$ *s.t.* $P + (-P) = O$.

(d) Let $P, Q, R \in E$. *Then* $(P + Q) + R = P + (Q + R)$.

In other words, the composition law makes E *into an abelian group with identity* O. *Further if* E *is defined over* K, *then*

$$E(K) = \{(x, y) \in E : x, y \in K\} \cup \{O\}$$

is a subgroup of E.

Proof. All of this is clear except the associativity law which can be verified directly by a tedious calculation using explicit formulas for addition law we give later. □

Remark 3.8. Let $m \in \mathbb{Z}$ be an integer and $P \in E$. Then we let $[0]P = O$,

$$[m]P = \underbrace{P + ... + P}_{m-times}(\text{ if } m > 0) \text{ and } [m]P = \underbrace{-P - ... - P}_{|m|-times}(\text{ if } m < 0)$$

Proposition 3.9. ***(Group law algorithm)*** *Let E be an elliptic curve given by Weierstrass equation (1).*

(a) Let $P_0 = (x_0, y_0)$. Then $-P_0 = (x_0, -y_0 - a_1x_0 - a_3)$.
Next let $P_1 + P_2 = P_3$ with $P_i = (x_i, y_i) \in E$ for $i = 1, 2, 3$

(b) If $x_1 = x_2$ and $y_1 + y_2 + a_1x_2 + a_3 = 0$, then $P_1 + P_2 = O$. Otherwise define λ and ν by the following formulas:

$$\text{If } x_1 \neq x_2, \text{ then } \lambda = \frac{y_2 - y_1}{x_2 - x_1} \text{ and } \nu = \frac{y_1x_2 - y_2x_1}{x_2 - x_1}$$

$$\text{If } x_1 = x_2, \text{ then } \lambda = \frac{3x_1^2 + 2a_2x_1 + a_4 - a_1y_1}{2y_1 + a_1x_1 + a_3} \text{ and } \nu = \frac{-x_1^3 + a_4x_1 + 2a_6 - a_3y_1}{2y_1 + a_1x_1 + a_3}$$

Then $y = \lambda x + \nu$ is the line through P_1 and P_2 or tangent to E if $P_1 = P_2$.

(c) $P_3 = P_1 + P_2$ has coordinates $x_3 = \lambda^2 + a_1\lambda - a_2 - x_1 - x_2$ and $y_3 = -(\lambda + a_1)x_3 - \nu - a_3$. In particular, the duplication formula for $P = (x, y) \in E$

$$x([2]P) = \frac{x^4 - b_4x^2 - 2b_6x - b_8}{4x^3 + b_2x^2 + 2b_4x + b_6} \tag{2}$$

where b_2, b_4, b_6, b_8 are the polynomials in the a_i's given in section 2.1.

Proof. The formulas are a just trivial matter of finding slope of line passing through two points, slope of tangent line at a given point and finding coordinates of intersection of a line with the curve. □

3.3 Singular Weierstrass Equations

Definition 3.10. Let E be a curve given by Weierstrass equation. The *non-singular part* of E, denoted by E_{ns} is the set of non-singular points of E.

Here we see that if we remove singular points from the curve then remaining curve E_{ns} is still a group. Recall from definition 3.4 that there are two possibilities for a singular point.

Proposition 3.11. *Let E be a curve given by Weierstrass equation with $\Delta = 0$, so E has a singular point S. The composition law 3.6 makes E_{ns} into an abelian group.*

(a) Suppose E has a node and let $y = \alpha_1x + \beta_1$ and $y = \alpha_2x + \beta_2$ are the distinct tangent lines at S. Then the map

$$E_{ns} \longrightarrow \bar{K}^*, \quad (x, y) \longmapsto \frac{y - \alpha_1x - \beta_1}{y - \alpha_2x - \beta_2}$$

is an isomorphism of abelian groups.

(b) Suppose that E has a cusp and $y = \alpha x + \beta$ be the tangent line at S. Then

$$E_{ns} \longrightarrow (\bar{K}, +), \quad (x, y) \longmapsto \frac{x - x(S)}{y - \alpha x - \beta}$$

is an isomorphism of abelian groups. ($(\bar{K}, +)$ is the additive group of $\bar{K}$)

3.4 Elliptic curves

Definition 3.12. An **elliptic curve** is a pair (E, O), where E is non-singular curve of genus one and $O \in E$. If E is defined over K and $O \in E(K)$ then we write E/K.

It can be proved using Riemann-Roch theorem that every elliptic curve is isomorphic to plane cubic. Precisely

Proposition 3.13. *Let E/K be an elliptic curve. Then*

*(a) There exist functions $x, y \in K(E)$ s.t. the map $\phi : E \to \mathbb{P}^2$, $\phi = [x, y, 1]$ gives an isomorphism of E/K onto a curve given by Weierstrass equation (1) with $a_1, a_2, ..a_6 \in K$ and $\phi(O) = [0, 1, 0]$. The functions x and y are called **Weierstrass coordinates** of E.*

(b) Any two Weierstrass coordinates for E as in (a) are related by linear change of variables of the form

$$X = u^2X' + r, \; Y = u^3Y' + su^2X' + t \quad \text{where } u \in K^* \text{ and } r, s, t \in K \qquad (3)$$

(c) Conversely, every non-singular cubic curve given by Weierstrass equation 1 defines an elliptic curve over K.

Proof. (a) Since $g = 1$, we have $\ell(n(O)) = dim(\mathcal{L}(n(O))) = n$ for all $n \geq 1$. Choose functions x and y such that $\{1, x\}$ and $\{1, x, y\}$ are bases of $\mathcal{L}(2(O))$ and $\mathcal{L}(3(O))$ respectively. Also $\mathcal{L}(6(O))$ has dimension 6 but it contains seven functions

$$1, x, y, x^2, xy, y^2, x^3$$

It follows there is linear relation $A_1 + A_2x + A_3y + A_4x^2 + A_5xy + A_6y^2 + A_7x^3 = 0$. Note that $A_6A_7 \neq 0$, since otherwise every term would have a pole at O of a different order, and so all of the A_j's would vanish. Replacing x and y by $-A_6A_7x$ and $A_6A_7^2y$, respectively, and dividing by $A_6^3A_7^4$, we get a cubic equation in Weierstrass form. This gives a map $\phi : E \longrightarrow \mathbb{P}^2, \quad \phi = [x, y, 1]$
whose image lies in C, the locus described by a Weierstrass equation. Note that $\phi : E \longrightarrow C$ is a morphism (by 2.1) and it is surjective (by 2.2) and $\phi(O) = [0, 1, 0]$. One can prove that ϕ is an isomorphism.

(b) Clearly $\{1, x\}$ and $\{1, x'\}$ are both bases for $\mathcal{L}(2(O))$ and $\{1, x, y\}$ and $\{1, x', y'\}$ are bases for $\mathcal{L}(3(O))$. Thus there are constants $u_1, u_2 \in K^*$ and $r, s_2, t \in K$ s.t.

$$x = u_1x' + r, \quad \text{and} \quad y = u_2y' + s_2x' + t$$

Derive that $u_1^3 = u_2^2$ and let $u = u_2/u_1$ and $s = s_2/u^2$.

(c) Use Riemann-Roch theorem and proposition 3.5.

□

4 The formal group of elliptic curves

In this section, we study "infinitesimal" neighbourhood of E centered at O. To do this, we associate a group to E called its formal group. The main result is theorem 4.11 which will be used to prove theorem 5.6.

4.1 Expansion around O

To study addition law of E close to O, we make change of variables

$$z = -x/y \text{ and } w = -1/y \quad \text{so that} \quad x = z/w \text{ and } y = -1/w \tag{4}$$

The point O is now $(z, w) = (0, 0)$. The Weierstrass equation of E becomes

$$w = z^3 + a_1zw + a_2z^2w + a_3w^2 + a_4zw^2 + a_6w^3 = f(z, w)$$

Now we substitute this equation into itself recursively so as to express w as a power series in z.

$$\begin{aligned} w &= z^3 + (a_1z + a_2z^2)w + (a_3 + a_4z)w^2 + a_6w^3 \\ &= z^3 + (a_1z + a_2z^2)[z^3 + (a_1z + a_2z^2)w + (a_3 + a_4z)w^2 \\ &\quad + a_6w^3] + (a_3 + a_4z)[z^3 + (a_1z + a_2z^2)w + (a_3 + a_4z)w^2 + a_6w^3]^2 \\ &\quad + a_6[z^3 + (a_1z + a_2z^2)w + (a_3 + a_4z)w^2 + a_6w^3]^3 \\ &\quad . \\ &\quad . \\ &\quad . \\ &= z^3(1 + A_1z + A_2z^2 + \text{ůůů}) = w(z) \end{aligned}$$

To describe more precisely the algorithm for producing $w(z)$, define a sequence

$$f_1(z, w) = f(z, w), \quad \text{and} \quad f_{m+1}(z, w) = f_m(z, f(z, w))$$

Then set $w(z) = lim_{m\to\infty} f_m(z, 0)$ provided this limit makes sense.

Proposition 4.1. *(a) The procedure described above gives a power series*

$$w(z) = z^3(1 + A_1z + A_2z^2 + ...) \in \mathbb{Z}[a_1, ..., a_6][\![z]\!]$$

(b) The series w(z) is the unique power series in $\mathbb{Z}[a_1, ..., a_6][\![z]\!]$ *satisfying*

$$w(z) = f(z, w(z))$$

Proof. Both (a) and (b) are special cases of Hensel's lemma. To prove it, apply lemma 4.2 with $R = \mathbb{Z}[a_1, ..., a_6][\![z]\!]$, $I = (z)$, $F(w) = f(z, w) - w$, $a = 0$ and $\alpha = -1$. □

We recall Hensel's lemma from Algebraic Number Theory.

Lemma 4.2. ***(Hensel's Lemma)*** *Let R be a ring that is complete w.r.t. some ideal $I \subset R$, and let $F(w) \in R[w]$ be a polynomial. Suppose that there is an integer $n \geq 1$ and an element $a \in R$ satisfying*

$$F(a) \in I^n \quad \text{and} \quad F'(a) \in R^*$$

Then for any $\alpha \in R$ satisfying $\alpha \equiv F'(a) \pmod{I}$, the sequence

$$w_0 = a \quad \text{and} \quad w_{m+1} = w_m - \frac{F(w_m)}{\alpha}$$

converges to an element $b \in R$ satisfying $F(b) = 0$ and $b \equiv a (mod\ I^n)$. If R is an integral domain, then these conditions determine b uniquely.

4.2 Formal group of Elliptic curves

Using the power series $w(z)$, we derive *Laurent series* for x and y

$$x(z) = \frac{z}{w(z)} = \frac{1}{z^2} - \frac{a_1}{z} - a_2 - a_3 z - (a_4 + a_1 a_3) z^2 - ...,$$
$$y(z) = -\frac{1}{w(z)} = -\frac{1}{z^3} + \frac{a_1}{z^2} + \frac{a_2}{z} + a_3 + (a_4 + a_1 a_3) z - ...$$

The pair $(x(z), y(z))$ provides a formal solution to the Weierstrass equation

$$E : y^2 + a_1 xy + a_3 y = x^3 + a_2 x^2 + a_4 x + a_6 \tag{5}$$

Now let K be a complete local field with the ring of integers R and the maximal ideal $\mathcal{M}$. Suppose E is defined over K with $a_1, ..., a_6 \in R$. Then the power series $x(z)$ and $y(z)$ will converge for any $z \in \mathcal{M}$. Thus we get an injective map

$$\mathcal{M} \longrightarrow E(K), \quad z \longrightarrow (x(z), y(z))$$

(This map is injective because it has an inverse $(x, y) \longmapsto -x/y$). Let z_1, z_2 be independent indeterminates and let $w_1 = w(z_1)$ and $w_2 = w(z_2)$. In the (z, w)-plane, the slope of line connecting (z_1, w_1) and (z_2, w_2) has slope

$$\lambda = \lambda(z_1, z_2) = \frac{w_2 - w_1}{z_2 - z_1} = \sum_{n=3}^{\infty} A_{n-3} \frac{z_2^n - z_1^n}{z_2 - z_1} \in \mathbb{Z}[a_1, ..., a_6][\![z_1, z_2]\!]$$

Letting $\nu = \nu(z_1, z_2) = w_1 - \lambda z_1$ and substituting $w = \lambda z - \nu$ into equation 4, we get a cubic with roots z_1, z_2 and z_3(say). Then

$$z_3 = z_3(z_1, z_2) = -z_1 - z_2 + \frac{a_1\lambda + a_3\lambda^2 - a_2 y - 2a_4\lambda\nu - 3a_6\lambda^2\nu}{1 + a_2\lambda + a_4\lambda^2 + a_6\lambda^3}$$

Letting $w_3 = \lambda(z_1, z_2) z_3(z_1, z_2) + \nu(z_1, z_2)$,. Using uniqueness condition of proposition 4.1(b), we get that $w_3 = w(z_3)$. i.e. we can compute the w coordinate of $-(z_1, w_1) - (z_2, w_2)$ using the power series w.

In the (x, y)-plane, the inverse of (x, y) is $(x, -y - a_1 x - a_3)$. Remembering that $z = -x/y$, we can compute $z-$coordinate of inverse of (z, w)

$$i(z) = \frac{x(z)}{y(z) + a_1 x(z) + a_3} = \frac{z^{-2} - a_1^z - 1 - ...}{-z^{-3} + 2a_1 z^{-2} + ...}$$

and similarly $w-$coordinate of inverse of (x, y) is $w(i(z))$. This gives **formal addition law** $F(z_1, z_2) = i(z_3(z_1, z_2))$.

4.3 Formal Groups and their properties

Definition 4.3. Let R be a ring. A *(one parameter commutative)* formal group $\mathcal{F}$ over R is a power series $F(X,Y) \in R[\![z]\!]$ with the following properties:

(a) $F(X,Y) = X + Y +$ (degree $\geq$2 terms)

(b) $F(X, F(Y,Z)) = F(F(X,Y), Z)$ (associativity)

(c) $F(X,Y) = F(Y,X)$ (commutativity)

(d) There is a unique power series $i(T) \in R[\![z]\!]$ s.t. $F(T, i(T)) = 0$

(e) $F(X,0) = X, F(0,Y) = Y$

We call $F(X,Y)$ the *formal group law* of $\mathcal{F}$.

Definition 4.4. Let $(\mathcal{F}, F)$ and $(\mathcal{G}, G)$ be formal groups defined over R. A *homomorphism from $\mathcal{F}$ to $\mathcal{G}$ defined over R* is a power series $f \in R[\![z]\!]$ that satisfies $f(F(X,Y)) = G(f(X), f(Y))$.

Isomorphim of formal groups are defined in the obvious way.

Definition 4.5. Let E be an elliptic curve given by Weierstrass equation with coefficients in R. The formal group associated to E is given by $F(z_1, z_2)$ described in previous section and is denoted by $\hat{E}$.

Example 4.6. Let $(\mathcal{F}, F)$ be a formal group. Define homomorphisms $[m] : \mathcal{F} \longrightarrow \mathcal{F}$ inductively for $m \in \mathbb{Z}$:

$$[0](T) = 0, \quad [m+1](T) = F([m](T), T), \quad [m-1](T) = F([m](T), i(T))$$

Proposition 4.7. *(a) $[m](T) = mT +$(higher order terms)*

(b) If $m \in R^*$, then $[m] : \mathcal{F} \longrightarrow \mathcal{F}$ is an isomorphism.

Proof. a) follows from induction on m and noting that $0 = F(T, i(T)) = T + i(T) + .. \implies i(T) = -T + ...$ (b) follows from following lemma. □

Lemma 4.8. *Let $a \in R^*$ and $f(T) \in R[\![T]\!]$ be a power series of the form $f(T) = aT +$(higher-order terms). Then there is a unique power series $g(T) \in R[\![T]\!]$ satisfying $f(g(T)) = T$. The series g also satisfies $g(f(T)) = T$.*

Proof. Construct a sequence of polynomials $g_n(T) \in R[T]$ inductively satisfying

$$f(g_n(T)) \equiv T \ (mod\ T^{n+1}) \quad \text{and} \quad g_{n+1}(T) = g_n(T) \ (mod\ T^{n+1})$$

Then the limit $g(T) = lim\ g_n(T)$ exists in $R[\![T]\!]$ and satisfies $f(g(T)) = T$. To start the induction, let $g_1(T) = a^{-1}T$. To prove $g(f(T)) = T$, apply above to $g(T) = a^{-1}T + ...$ to get $h(T) \in R[\![T]\!]$ s.t. $g(h(T)) = T$. Now

$$g(f(T)) = g(f(g(h(T)))) = g(f \circ g(h(T))) = g(h(T)) = T$$

To prove uniqueness, let $G(T) \in R[\![T]\!]$ be another power series satisfying $f(G(T)) = T$. Then

$$g(T) = g(f(G(T))) = g \circ f(G(T)) = G(T)$$

□

Now we come to our main theorem of this section. We fix the following notation:

R	a complete local ring
$\mathcal{M}$	the maximal ideal of R
k	$=R/\mathcal{M}$, the residue field of R
$\mathcal{F}$	a formal group defined over R, with formal group law $F(X,Y)$

Definition 4.9. The *group associated to* $\mathcal{F}/R$, denoted by $\mathcal{F}(\mathcal{M})$, is the set $\mathcal{M}$ endowed with group operations $x \oplus_{\mathcal{F}} y = F(x,y)$ and $\ominus_{\mathcal{F}} x = i(x)$ for all $x, y \in \mathcal{M}$.

Example 4.10. Let $\hat{E}$ be the formal group associated to an elliptic curve E/K where $K = Q(R)$(field of fractions). As noted in section 4.2, we have injective map

$$\mathcal{M} \longrightarrow E(K), \qquad z \longmapsto (x(z) = z/w(z), y(z) = -1/w(z))$$

The construction of the power series for $\hat{E}$ imply that this map is a homomorphism from $\hat{E}(\mathcal{M})$ to $E(K)$.

Theorem 4.11. *Let p=char(k)(where p can be 0). Then every torsion element in* $\mathcal{F}(\mathcal{M})$ *has order that is power of p.*

Proof. Multiplying an arbitrary torsion element by an appropriate power of p, it suffices to prove that there are no non-zero torsion elements of order prime to p. So let $m \geq 1$ with $(p,m) = 1$. Since R is local, this means $m \in R^*$. So $[m]$ is an automorphism of the formal group $\mathcal{F}$ by proposition 4.7(b). Now let $x \in \mathcal{F}(\mathcal{M})$ such that $[m](x) = 0 \implies x = 0$. □

5 Elliptic curves over local fields

Here we prove one key indredient (theorem 5.6) which is used in the proof of the Weak Mordell-Weil theorem.

Following notation will be used throughout this section unless otherwise stated:

K	a local field, complete w.r.t a discrete valuation v
R	$=\{x \in K : v(x) \geq 0\}$, the ring of integers of K
R^*	$=\{x \in K : v(x) = 0\}$, the unit group of R
$\mathcal{M}$	$=\{x \in K : v(x) > 0\}$, the maximal ideal of R
π	a uniformizer for R *i.e.* $\mathcal{M} = \pi R$
k	$=R/\mathcal{M}$, the residue field of R

Let E/K be an elliptic curve and let

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

be a Weierstrass equation for E/K. The substitution $(x, y) \mapsto (u^{-2}x, u^{-3}y)$ leads to a new equation in which a_i is replaced by $u^i a_i$. So choosing u to be divisible by large power of π, we can assume that all $a_i \in R$. So we have $v(\Delta) \geq 0$. Since v is discrete, among all such equations, we can choose one which minimises $v(\Delta)$.

Definition 5.1. Let E/K be an elliptic curve. A Weierstrass equation for E is called *minimal (Weierstrass) equation for* E if $v(\Delta)$ is minimized subject to the condition that $a_1, ..., a_6 \in R$. The minimal value of $v(\Delta)$ is called *valuation of minimal discriminant* of E at v.

Remark 5.2. Value of $v(\Delta)$ is changed only by a multiple of 12, so we can say that

$$a_i \in R \text{ and } v(\Delta) < 12 \implies \text{ the equation is minimal}$$

If $char(K) \neq 2, 3$ then converse also holds.

Proposition 5.3. *A minimal Weierstrass equation is unique upto change*

$$x = u^2x' + r, \text{ and } y = u^3y' + u^2sx' + t$$

of coordinates where $u \in R^*$ *and* $r, s, t \in R$.

Proof. The only non-trivial part in above proposition is why $u \in R^*$? This is clear since $12v(u) + v(\Delta') = v(\Delta)$ and $v(\Delta') = v(\Delta)$ implies $v(u) = 0$. □

5.1 Reduction modulo π and torsion points

The natural reduction map $R \longrightarrow k = R/\pi R$ is denoted by $t \mapsto \tilde{t}$. Having chosen a minimal Weierstrass equation for E/K, we reduce its coefficients modulo π to obtain a (possibly singular) curve over k, called *reduction of E modulo* π

$$\tilde{E} : y^2 + \tilde{a_1}xy + \tilde{a_3}y = x^3 + \tilde{a_2}x^2 + \tilde{a_4}x + \tilde{a_6}$$

Let $P \in E(K)$, homogenous coordinates $P = [x_0, y_0, z_0]$ with $x_0, y_0, z_0 \in R$ and atleast one of them in R^*. Then $\tilde{P} = [\tilde{x_0}, \tilde{y_0}, \tilde{z_0}] \in \tilde{E}(k)$. This defines a *reduction map*

$$E(K) \longrightarrow \tilde{E}(k), \quad P \longmapsto \tilde{P}$$

Define two subsets of $E(K)$ as follows:

$E_0(K) := \{P \in E(K) : \tilde{P} \in \tilde{E}_{ns}(k)\} = \{\text{Points in } E(K) \text{with non-singular reduction}\}$,

$E_1(K) := \{P \in E(K) : \tilde{P} = \tilde{O}\} = ker(\text{reduction map})$

From proposition 5.3, the above sets do not depend upon the minimal Weierstrass equation chosen.

Proposition 5.4. *There is an exact sequence of abelian groups*

$$0 \longrightarrow E_1(K) \hookrightarrow E_0(K) \xrightarrow{mod\ \pi} \tilde{E}_{ns}(k) \longrightarrow 0$$

Proof. The non-trivial things to prove in this proposition are **1)**Surjectivity of the reduction map, **2)**Reduction map is a homomorphism and **3)**$E_0(K)$ is a subgroup of $E(K)$.

Surjectivity is shown using Hensel's lemma and completeness of K: Suppose $f(x, y) = y^2 + a_1xy + a_3y - x^3 - a_2x^2 - a_4x - a_6$ be minimal Weierstrass equation and let $\tilde{P} = (\tilde{\alpha}, \tilde{\beta}) \in \tilde{E}_{ns}(k)$. WLOG assume $\frac{\partial \tilde{f}}{\partial x}(\tilde{P}) \neq 0$. Choose any $x_0 \in R$ with $\tilde{x_0} = \tilde{\alpha}$. Then $f(x_0, y) = 0$ and modulo π, it has simple root $\tilde{\beta}$. Now use Hensel's lemma to lift this root to $y_0 \in R$ s.t. $\tilde{y_0} = \tilde{\beta}$ and $f(x_0, y_0) = 0$.

To prove 2), Let $P_1, P_2 \in E_0(K)$ and $P_3 \in E(K)$ be points satisfying $P_1 + P_2 + P_3 = O$. Thus $\exists$ a line L that intersects E at the three points P_1, P_2, P_3, counted with appropriate multiplicities. The idea is to prove that $\tilde{L}$ intersections $\tilde{E}$ at $\tilde{P_1}, \tilde{P_2}, \tilde{P_3}$ with correct multiplicities from which it follows that $P_3 \in E_0(K)$ and $\tilde{P_1} + \tilde{P_2} + \tilde{P_3} = \tilde{O}$ but there are many cases to consider:

(a) $\tilde{P_1}, \tilde{P_2}, \tilde{P_3}$ are distinct

(b) $P_1 \neq P_2$ but $\tilde{P_1} = \tilde{P_2}$

(c) $P_1 = P_2 \neq P_3$ and $\tilde{P_1} = \tilde{P_2} \neq \tilde{P_3}$, etc.

which can be proven case by case. □

Proposition 5.5. *Let E/K be given by a minimal Weierstrass equation, let $\hat{E}/K$ be the formal group associated to E and, let $w(z) \in R[\![t]\!]$ be the power series from section 4.2. Then the map is an isomorphism of groups.*

$$\hat{E}(\mathcal{M}) \longrightarrow E_1(K), \quad z \longmapsto \left(\frac{z}{w(z)}, -\frac{1}{w(z)}\right)$$

Proof. In example, we obtained an injective map $\hat{E}(\mathcal{M}) \to E(K)$ given by above mapping. Its image is in $E_1(K)$ since $v(-1/w(z)) = -3v(z) < 0$. It is surjective since we have a well defined map $E_1(K) \to \hat{E}(\mathcal{M})$ given by $(x, y) \mapsto -x/y$ whose composition with above map is identity on $\hat{E}(\mathcal{M})$. □

Theorem 5.6. *Let E/K be an elliptic curve and $m \geq 1$ be an integer such that $(m, char(k)) = 1$*

(a) The subgroup $E_1(K)$ has no non-trivial points of order m.

(b) Assume further that the reduced curve $\tilde{E}/k$ is non-singular. Then the reduction map $E(K)[m] \longrightarrow \tilde{E}(k)$ is injective.

Proof. (a) is clear from theorem 4.11 and proposition 5.5. (b) is clear from the exact sequence in proposition 5.4 and part (a) of this theorem($\tilde{E}_{ns}(k) = \tilde{E}(k)$). □

From section 3.1, the Weierstrass equation of the reduced curve $\tilde{E}$ is one of the three types. We classify E according to these possibilites.

Definition 5.7. (a) E has *good (or stable reduction)* if $\tilde{E}$ is non-singular.

(b) E has *multiplicative (or semi-stable reduction)* if $\tilde{E}$ has a node.

(c) E has *additive (or unstable reduction)* if $\tilde{E}$ has a cusp.
In cases (*b*) and (*c*), we say that E has *bad* reduction.

We also recall the definition of an unramified extension of K and inertia group of Galois extensions of a local field. These definitions will be useful in next section.

Theorem 5.8. *([Neukirch, 2008], II.4.8) Let K'/K be a finite extension of degree n. Then there is a unique extension w of valuation v where $w = \frac{1}{n} v \circ N_{K'/K}$. Morover, K' is complete w.r.t. the valuation w.*

Definition 5.9. Let K'/K be a finite extension of local fields. We say that K'/K is *unramified* if the residue field extension k'/k is seperable and $[K' : K] = [k' : k]$.

Definition 5.10. Let K'/K be finite Galois extension and $\sigma \in Gal(K'/K)$, then define $\bar{\sigma} : k' \longrightarrow k'$, $x \bmod \mathfrak{p}' \longmapsto \sigma(x) \bmod \mathfrak{p}'$. Then $\bar{\sigma} \in Gal(k'/k)$. So we have map $Gal(K'/K) \longrightarrow Gal(k'/k)$, $\sigma \longmapsto \bar{\sigma}$. The kernel of this map is called *inertia group of K'/K*, denoted by $I_{K'/K}$.

6 Mordell-Weil Theorem

In this section, we complete the proof of the Mordell-Weil theorem for elliptic curves over $\mathbb{Q}$. First we prove the Weak Mordell-Weil theorem in full generality (over number fields). Then we see descent theorem. Finally we define height function on $E(\mathbb{Q})$ and prove it has required properties thereby completing the proof of Mordell-Weil theorem using descent theorem.

We fix the notation which will be used throughout this section:

- K a number field
- M_K a complete set of inequivalent absolute values on K
- M_K^∞ the archimedian absolute values in M_K
- M_K^0 the non-archimedian absolute values in M_K
- $v(x) := -log|x|_v$, for absolute value $v \in M_K$
- R is the ring of integers of K, equals $\{x \in K : v(x) \geq 0 \; \forall v \in M_K^0\}$
- K_v is the completion of K at $v \in M_K$

6.1 The Weak Mordell-Weil Theorem

In this section, we give the proof of the following theorem.

Theorem 6.1. ***(Weak Mordell-Weil theorem)*** *Let $m \geq 2$ be an integer. Then $E(K)/mE(K)$ is a finite group.*

For the rest of the section, let E/K and m are as in the theorem 5.1. The following lemma says that it is sufficient to prove theorem 5.1 under additional assumption that $E[m] \subset E(K)$ (which we assume onwards).

Lemma 6.2. *Let L/K be a finite Galois extension. Then*

$$E(L)/mE(L) \text{ finite} \implies E(L)/mE(L) \text{ finite}$$

Proof. The inclusion $E(K) \to E(L)$ induces a natural map with kernel Φ

$$E(K)/mE(K) \longrightarrow E(L)/mE(L), \quad \Phi = \frac{E(K) \cap mE(L)}{mE(K)}$$

For each $P \; mod(mE(K)) \in \Phi$, choose $Q_P \in E(L)$ satisfying $[m]Q_P = P$. Then define a map of sets

$$\lambda_P : Gal(L/K) \longrightarrow E[m], \quad \sigma \longmapsto \sigma(Q_P) - Q_P$$

It is well defined because $[m](\sigma(Q_P) - Q_P) = \sigma([m]Q_P) - [m]Q_P = \sigma(P) - P = O$. Now suppose $P, P' \in E(K) \cap mE(L)$ satisfy $\lambda_P = \lambda_{P'}$. Then

$$\sigma(Q_P) - Q_P = \sigma(Q_{P'}) - Q_{P'} \implies \sigma(Q_P - Q_{P'}) = Q_P - Q_{P'}$$

Hence $Q_P - Q_{P'} \in E(K) \implies P - P' = [m](Q_P - Q_{P'}) \in mE(K)$. This proves that we have bijection of sets

$$\Phi \longrightarrow Hom_{Sets}(Gal(L/K), E[m]), \quad P \longmapsto \lambda_P$$

Therefore Φ is finite and this implies $E(K)/mE(K)$ is finite. □

Next we define the Kummer pairing which reduces the problem of showing finiteness of $E(K)/mE(K)$ to showing finiteness of the field extension L/K where $L = K([m]^{-1}E(K))$ is the compositum of all fields $K(Q)$ as Q ranges over all points in $E(\bar{K})$ satisfying $[m]Q \in E(K)$.

Definition 6.3. The **Kummer pairing** $\kappa : E(K) \times Gal(\bar{K}/K) \to \mu_m$ is defined as: For $P \in E(K)$, choose any $Q \in E(\bar{K})$ satisfying $[m]Q = P$. Then

$$\kappa(P, \sigma) := \sigma(Q) - Q$$

The following are basic properties of Kummer pairing. Although these properties can be proven directly, we remark that all these propeties follows immediately from properties of group cohomology.

Proposition 6.4. *a) The Kummer Pairing is well defined and is bilinear.*
b) The kernel of the Kummer pairing on the left is $mE(K)$ and on the right is $Gal(\bar{K}/L)$ where $L = K([m]^{-1}E(K))$ hence induces a perfect bilinear pairing

$$E(K)/mE(K) \times Gal(L/K) \to \mu_m$$

Proof. a) Checking that the Kummer pairing is well-defined is routine exercise (remembering that $E[m] \subset E(K)$). Bilinearity in first component is obvious. For bilinearity in second component, we have

$$\kappa(P, \sigma\tau) = \sigma\tau(Q) - Q = \sigma(\tau(Q) - Q) + (\tau(Q) - Q) = \sigma(\kappa(P, \tau)) + \kappa(P, \sigma)$$

and $\sigma(\kappa(P, \tau)) = \kappa(P, \tau)$ because $E[m] \subset E(K)$.

b) The kernels on the left and right part are obvious. The induced pairing is perfect is also obvious once we see that the extension L/K is Galois. This follows from the fact that the set $[m]^{-1}E(K)$ is $Gal(\bar{K}/K)$ invariant. Hence L/K is normal extension.(it is seperable because $char(K) = 0$) □

From perfect pairing, we see that $E(K)/mE(K)$ finite $\iff$ $Gal(L/K)$ finite (*proof-* Suppose $E(K)/mE(K)$ is infinite given by $\{a_i\}_{i=1}^{\infty}$ and $Gal(L/K)$ is finite given by $\{\sigma_i\}_{i=1}^{n}$. Since range of pairing is finite, we can extract out a subsequence on which σ_1 is constant. Applying this process n times, we get a subsequence $\{a_{i_j}\}_{j=1}^{\infty}$ on which every σ_i is constant. Hence $\sigma_i(a_{i_1} - a_{i_2}) = 1$ for all $i = 1, ..., n$).

Now we look at some definitions and some facts from algebraic number theory.

Definition 6.5. Let E/K be an elliptic curve and $v \in M_K^0$ be a discrete valuation. We say E has *good*(resp.*bad) reduction at* v if E has good(resp. bad) reduction when considered over K_v. We denote the reduced curve over the residue field k_v by $\tilde{E}/k_v$.

Remark 6.6. Taking a Weierstrass equation for E/K, $y^2+a_1xy+a_3y = x^3+a_2x^2+a_4x+a_6$, with discriminant Δ. Then for all but finitely many $v \in M_K^0$, we have

$$v(a_i) \geq 0 \text{ for all } i = 1, ..., 6 \text{ and } v(\Delta) = 0$$

For any v satisfying these conditions, $\tilde{E}/k_v$ is non-singular. This shows that E has *good reduction at all but finitely many* $v \in M_K^0$.

We now give some standard facts from algebraic number theory. Recall the definition of unramified extension of a local field and of inertia group of finite Galois extensions of a local field from section 5.1.

Theorem 6.7. *([Neukirch, 2008], II.7.2) Let L/K and K'/K be two extensions inside an algebraic closure $\bar{K}/K$ and let $v \in M_K^0$ be a discrete valuation and $v' \in M_{K'}^0$ be a extension of v to K' then*

$$L/K \text{ unramified at } v \implies LK'/K' \text{ unramified at } v'$$

Each subextension of an unramified extension at v is again unramified at v.

Corollary 6.8. *The composite of two unramified extensions at v is again unramified at v.*

Now continuing our proof of theorem 6.9. The following are the properties of the extension L/K.

Proposition 6.9. *a) The extension L/K is abelian and has exponent m. i.e. every element in $Gal(L/K)$ has order dividing m.*
b)Let $S = \{v \in M_K^0 : E \text{ has bad reduction at } v\} \bigcup \{v \in M_k^0 : v(m) \neq 0\} \bigcup M_K^\infty$
Then L/K is unramified outside S.

Proof. (a) This follows immediately from 6.1 which implies that there is an injection

$$Gal(L/K) \longrightarrow Hom(E(K), E[m]), \quad \sigma \longmapsto \kappa(\cdot, \sigma)$$

Note that every element of $Hom(E(\bar{K}), E[m])$ has order dividing m.

(b) Let $v \in M_K \backslash S$ and let $Q \in E(\bar{K})$ s.t. $[m]Q \in E(K)$ and let K' be the normal closure of $K(Q)$. By corollary 6.8 the composite of two unramified extensions is again unramified, it is sufficient to prove that K'/K is unramified. Let $v' \in M_{K'}$ be a place lying above v and $k'_{v'}/k_v$ be corresponding extensions of residue fields. Since $v \notin S$, E has good reduction at v hence good reduction at v' as well. Thus we have

$$E(K') \longrightarrow \tilde{E}(k'_{v'})$$

the usual reduction map.

Let $I_{v'/v}$ be the inertia group and $\sigma \in I_{v'/v}$. By definition σ acts trivially on $\tilde{E}(k_v)$, we have $\widetilde{\sigma(Q) - Q} = \sigma(\tilde{Q}) - \tilde{Q} = \tilde{O}$. Also $[m](\sigma(Q) - Q) = \sigma([m]Q) - [m]Q = O$. By theorem 5.6(b), $\sigma(Q) - Q = O$. Since σ was arbitrary, the inertia group $I_{v'/v}$ acts trivially on all generators of K'/K (generators are conjugates of point Q). Therefore K'/K is a unramified extension at v. □

We recall some results from algebraic number theory, Galois theory and Kummer Theory which are used in the proof of the proposition 6.15.

Theorem 6.10. *([Neukirch, 2008], I.6.3) The ideal class group Cl_K of K is finite.*

Let $S \subset M_K$ be a finite subset of places of K which includes all the archimedian places i.e. $M_K^\infty \subset S$.

Definition 6.11. Then the *ring of S-integers*, denoted by R_S is

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

The units in this ring are called the *S-units*.

Theorem 6.12. *([Neukirch, 2008], I.11.7)The group of S-units is finitely generated.*

Proposition 6.13. *([Dummit and Foote, 2003], Proposition 14.19) Let L/K be a Galois extension and K'/K be any extension (not necessarily finite), then LK'/K' is a Galois extension, with Galois group*

$$Gal(LK'/K') \cong Gal(L/L \cap K')$$

Theorem 6.14. ***(Kummer)****Let K be a field contaning μ_m. Then its maximal abelian extension of exponent m is the field $K(a^{1/m} : a \in K)$.*

The next proposition says that any extension satisfying properties a) and b) of proposition 6.9 is necessarily a finite extension thereby completing the proof of the Weak Mordell-Weil theorem.

Proposition 6.15. *Let K be any number field and $S \subset M_K$ be finite set of places containing M_K^∞, and let $m \geq 2$ be an integer. Let L/K be maximal abelian extension of K which is unramified outside S. Then L/K is a finite extension.*

Proof. **Claim:** WLOG, we can assume that $\mu_m \subset K$.

Suppose the above proposition is true for some finite extension K' of K, where S' is the set of places of K' lying over S. Since by 6.13 and 6.7, LK'/K' is an abelian exponent of exponent m and unramified outside S', so it is a finite extension. Again by 6.13, $L/L \cap K'$ is a finite extension and $L \cap K'/K$ is finite since K'/K was finite. Therefore L/K is a finite extension.

Claim: WLOG, we can assume that R_S is a principal ideal domain.

We can also increase the size of S since this only has the effect of making L larger. Now choose integral ideals $\mathfrak{a}_1, \ldots, \mathfrak{a}_h$ representating the ideal classes of K and adjoin to S the valuations $v_\mathfrak{p}$ for primes $\mathfrak{p} | \mathfrak{a}_1 \ldots \mathfrak{a}_h$. Then R_S is a PID. To see this, note that R_S is the localization of R w.r.t. the multiplicative set

$$\left(\bigcup_{\mathfrak{p} \in S \setminus M_K^\infty} \mathfrak{p} \right) \setminus \{0\}$$

Let A be any ring, then ideals of $T^{-1}A$ is in one to one correspondence with ideals of A which do not meet S and if $\mathfrak{a} \cap T \neq 0$, then extension of $\mathfrak{a}$ in $T^{-1}A$ is the whole ring. Now it is clear that class group of R_S is trivial.

We also enlarge S so that $v(m) = 0$ for $v \notin S$. By 6.14, L is the largest subfield of $K(a^{1/m} : a \in K)$ which is unramified outside S. Now let $v \in M_K \setminus S$ and consider

the equation $X^m - a$ over K_v. Since $v(m) = 0$ and $disc(X^m - a) = \pm m^m a^{m-1}$, by [Lang, 1994], II.7, we have that $K_v(a^{1/m})/K_v$ is unramified $\iff m|ord_v(a)$.

It is clear that when we adjoin m^{th} roots, it is sufficent to take only one representative for each class in $K^*/(K^*)^m$. So if we let

$$T_S = \{a \in K^*/(K^*)^m : m|ord_v(a) \text{ for all } v \in M_K \backslash S\}$$

Then $L = K(a^{1/m} : a \in T_S)$. To complete the proof, it suffices to prove that T_S is finite. Consider the natural map $R_S^* \longrightarrow T$. **Claim:** This map is surjective.

Let $a \in T_S$ and let $aR_S = \mathfrak{p}_1^{e_1} \dots \mathfrak{p}_r^{e_r}$. Since $m|ord_v(a)$ for all $v \notin S$ and prime ideals of R_S correspond to the valuations $v \notin R_S$, we see that $m|e_i$ and aR_S is m^{th} power of an ideal in R_S. Since R_S is PID, we can find $b \in K^*$ s.t. $aR_S = b^m R_S$. So there is some $u \in R_S^*$ s.t. $a = ub^m$. So u and a give same elements T_S and prove the surjectivity. Also, kernel clearly contains $(R_S^*)^m$ so we have

$$R_S^*/(R_S^*)^m \twoheadrightarrow T_S$$

By theorem 6.12, $R_S^*/(R_S^*)^m$ is finite. Hence T_S is finite. □

6.2 The Descent Theorem

It is easy to see that finiteness of $E(K)/mE(K)$ does not imply finite generation of $E(K)$. For example, $\mathbb{R}/m\mathbb{R} = 0$ for every integer $m \geq 1$ but $\mathbb{R}$ is certainly not f.g. abelian group. The problem is that in $\mathbb{R}$, arbitrary large elements are divisible by m. We will show that this is not the case with $E(K)$ by defining an appropriate notion "size" on $E(K)$ and then showing $[m]$ increases the "size".

In this section we axiomatize the situation and describe the type of size (or height) function needed to prove that an abelian group is finitely generated.

Theorem 6.16. ***(Descent Theorem)*** *Let A be an abelian group. Suppose there exists a (height) function $h : A \to \mathbb{R}$ satisfying*

(a) For every $Q \in A$, $\exists\, C_1$ only depending on A and E s.t.

$$h(P + Q) \leq 2h(P) + C_1 \text{ for all } P \in A$$

(b) There is an integer $m \geq 2$ and a constant C_2 depending only on A s.t.

$$h(mP) \geq m^2 h(P) - C_2 \; \forall P \in A$$

(c) For every constant C_3, the set $\{P \in A : h(P) \leq C_3\}$ is finite.

Suppose further that for the integer m in (b), the group A/mA is finite then A is finitely generated abelian group.

Proof. Let $Q_1, Q_2, ..., Q_r$ be finitely many cosets representatives of A/mA. Let P be an arbitrary element in A, then there exists indices $i_1, i_2, ..., i_m$ and elements $P_1, P_2, ..., P_n \in A$ such that

$$P = Q_{i_1} + mP_1$$
$$P_1 = Q_{i_2} + mP_2$$

.
.
.

$$P_{n-1} = Q_{i_n} + mP_n$$

For each index j, we have

$$\begin{aligned} h(P_j) &\leq \frac{1}{m^2}(h(2P_j) + C_2) \\ &= \frac{1}{m^2}(h(P_{j-1} - Qi_j) + C_2) \\ &\leq \frac{1}{m^2}(2h(P_{j-1}) + C_1' + C_2) \end{aligned}$$

where C_1' is the maximum of constants from (i) for $Q \in \{-Q_1, ..., -Q_r\}$. Using this inequality repeatedly, we get

$$\begin{aligned} h(P_n) &\leq \left(\frac{1}{m^2}\right)^n h(P) + \frac{1}{m^2}\left(1 + \frac{2}{m^2} + ... + \left(\frac{2}{m^2}\right)^{n-1}\right)(C_1' + C_2) \\ &< \left(\frac{1}{m^2}\right)^n h(P) + \frac{C_1' + C_2}{m^2 - 2} \\ &\leq \frac{1}{2^n}h(P) + \frac{1}{2}(C_1' + C_2) \quad \text{since } m \geq 2 \end{aligned}$$

So for n sufficiently large, we have $h(P_n) \leq 1 + \frac{1}{2}(C_1' + C_2)$
Since P is a linear combination of P_n and $Q_1, ..., Q_r$,

$$P = m^n P_n + \sum_{j=1}^{n} m^{j-1} Q_{i_j}$$

Hence, the set

$$\{Q_1, Q_2, ..., Q_r\} \cup \{P \in A : h(P) \leq 1 + \frac{1}{2}(C_1' + C_2)\}$$

generates A which is finite by (c). □

6.3 Mordell-Weil Theorem over $\mathbb{Q}$

In this section, we complete the proof of the Mordell-Weil Theorem over $\mathbb{Q}$:

Theorem 6.17. *Let $E/\mathbb{Q}$ be an elliptic curve. Then $E(\mathbb{Q})$ is finitely generated.*

In order to apply the theorem 6.16, we need to define a height function on $E(\mathbb{Q})$ and prove that it has the required properties.

Definition 6.18. Let $t \in \mathbb{Q}$ and write $t = p/q$ as a fraction in lowest terms. Then *height of t*, denoted by $H(t)$ is defined by $H(t) = max\{|p|, |q|\}$.

Definition 6.19. The *logarithmic height* on $E(\mathbb{Q})$, relative to given Weierstrass equation, is given by

$$h_x : E \longrightarrow \mathbb{R}, \quad h_x(P) = \begin{cases} logH(x(P)) & \text{if } P \neq O \\ 0 & \text{if } P = O \end{cases}$$

Lemma 6.20. *Let $E/\mathbb{Q}$ be an Elliptic curve given by a Weierstrass equation*

$$E : y^2 = x^3 + Ax + B \quad \text{with } A, B \in \mathbb{Z}$$

(a) For every $P_0 \in C(\mathbb{Q})$, $\exists\, C_1$ only depending on P_0, A and B s.t.

$$h(P + P_0) \leq 2h(P) + C_1 \text{ for all } P \in C(\mathbb{Q})$$

(b) There is a constant C_2 that depends only on A and B s.t.

$$h([2]P) \geq 4h(P) - C_2 \quad \text{for all } P \in C(\mathbb{Q})$$

(c) For every constant C_3, the set $\{P \in A : h(P) \leq C_3\}$ is finite.

Proof. (a) Assume that $C_1 > max\{h_x(P_0), h_x([2]P_0)\}$ so that a) is true when $P_0 = O$ or $P \in \{O, \pm P_0\}$. In other cases, write

$$P = (x, y) = \left(\frac{a}{d^2}, \frac{b}{d^3}\right) \text{ and } P_0 = (x_0, y_0) = \left(\frac{a_0}{d_0^2}, \frac{b_0}{d_0^3}\right)$$

where all fractions are in lowest terms. The addition law says that

$$\begin{aligned} x(P + P_0) &= \left(\frac{y - y_0}{x - x_0}\right)^2 - x - x_0 \\ &= \frac{(xx_0 + A)(x + x_0) + 2B - yy_0}{(x - x_0)^2} \\ &= \frac{(aa_0 + Ad^2d_0^2)(ad_0^2 + a_0d^2) + 2Bd^4d_0^4 - 2bdb_0d_0}{(ad_0^2 - a_0dr)^2} \end{aligned}$$

In computing the height of a rational number, cancellation between numerator and denominator can only decrease the height, so we have

$$H(x(P + P_0)) \leq C_1' max\{|a|^2, |d|^4, |bd|\} \tag{6}$$

where C_1' depends on A, B, a_0, b_0, d_0. Now $P \in C$, so we have

$$b^2 = a^3 + Aad^4 + Bd^6 \implies |b| \leq C_1\text{"}max\{|a|^{3/2}, |d|^3\} \tag{7}$$

Combining this with above result, we get

$$H(x(P + P_0)) \leq C_1 max\{|a|^2, |d|^4\} = C_1 H(x(P))^2$$

Now taking logarithm completes the proof.

(b) Choose C_2 to satisfy

$$C_2 \geq 4max\{h_x(T) : T \in E(\mathbb{Q})[2]\}$$

so we may assume that $[2]P \neq O$. Now let $P = (x, y)$, we have

$$x([2]P) = \frac{x^4 - 2Ax^2 - 8Bx + A^2}{4x^3 + 4Ax + 4B}$$

Now define the homogenous polynomials

$$F(X,Y) = X^4 - 2AX^2Z^2 - 8BXZ^3 + A^2Z^4,$$
$$G(X,Z) = 4X^3Z + 4AXZ^3 + 4BZ^4$$

If we write $x(P) = a/b$ in lowest terms, then $x([2]P) = F(a,b)/G(a,b)$.

In contrast to proof of (a), We are now to looking to lower bound the height on doubling which means we have to bound how much cancellation occurs in numerator and denominator of $x([2]P)$.

To do this, we use the fact that $F(X,1)$ and $G(X,1)$ are relatively prime polynomials so they generate unit ideal in $\mathbb{Q}[X]$. This implies that identities of following sort exitst. (remember that $\Delta \neq 0$)

Lemma 6.21. *Let $\Delta = 4A^3 + 27B^2$, and define polynomials*

$$F(X,Y) = X^4 - 2AX^2Z^2 - 8BXZ^3 + A^2Z^4,$$
$$G(X,Z) = 4X^3Z + 4AXZ^3 + 4BZ^4,$$
$$f_1(X,Z) = 12X^2Z + 16AZ^3,$$
$$g_1(X,Z) = 3X^3 - 5AXZ^2 - 27BZ^3,$$
$$f_2(X,Z) = 4\Delta X^3 - 4A^2BX^2Z + 4A(3A^3 + 22B^2)XZ^2 + 12B(A^3 + 8B^2)Z^3,$$
$$g_2(X,Z) = A^2BX^2 + A(5A^3 + 32B^2)X^2Z + 2B(13AB^3 + 96B^2)XZ^2 - 3A^2(A^3 + 8B^2)Z^3.$$

Then the following identities hold in $\mathbb{Z}[A,B,X,Y]$:

$$f_1(X,Z)F(X,Z) - g_1(X,Z)G(X,Z) = 4\Delta Z^7 \tag{8}$$
$$f_2(X,Z)F(X,Z) - g_2(X,Z)G(X,Z) = 4\Delta X^7 \tag{9}$$

Proof. This can be verified directly via a tedious calculation. The identities can be obtained using euclidean algorithm or using theory of resultants. □

Now let $g = gcd(F(a,b), G(a,b))$. From equations 8 and 9, we get $g|4\Delta$. In particular, $|g| \leq 4|\Delta|$ and we get

$$H([2]P) \geq \frac{max\{|F(a,b)|, |G(a,b)|\}}{4|\Delta|}$$

On the other hand, the same identities gives

$$|4\Delta b^7| \leq 2max\{|f_1(a,b)|, |g_1(a,b)|\}max\{|F(a,b)|, |G(a,b)|\}$$
$$|4\Delta a^7| \leq 2max\{|f_2(a,b)|, |g_2(a,b)|\}max\{|F(a,b)|, |G(a,b)|\}$$

Looking at the expressions for f_1, f_2, g_1, g_2 in Lemma 6.21, we get

$$max\{|f_1(a,b)|, |g_1(a,b)|, |f_2(a,b)|, |g_2(a,b)|\} \leq Cmax\{|a|^3, |b|^3\}$$

where C is a constant depending on A and B. Combining last three inequalities

$$max\{|4\Delta a^7|, |4\Delta b^7|\} \leq 2Cmax\{|a|^3, |b|^3\}max\{|F(a,b)|, |G(a,b)|\}$$

Cancelling $max\{|a|^3, |b|^3\}$ from both sides and using $H(x(P)) = max\{|a|, |b|\}$,

$$\frac{max\{|F(a,b)|, |G(a,b)|\}}{4|\Delta|} \geq (2C)^{-1} max\{|a|^4, |b|^4\}$$

(c) This is clear since for any constant C, the set

$$\{t \in \mathbb{Q} : H(P) \leq C\}$$

is finite since it can have atmost $(2C+1)^2$ elements and for any value of x, there can be atmost 2 values of y. Hence the given set is finite. □

The Mordell-Weil theorem over $\mathbb{Q}$ follows using lemma 6.20 and theorem 6.16,

References

D. S. Dummit and R. S. Foote. *Abstract Algebra.* Wiley, 2003.

S. Lang. *Algebraic Number Theory.* Springer, 1994.

J. Neukirch. *Algebraic Number Theory.* Springer, 2008.

J. H. Silverman. *Arithmetic of Elliptic curves, 2nd edition.* Springer, 2008.