Chapter 6

Integral points on elliptic curves

Let K be a number field, let $\mathcal{O}_K$ be its ring of integers, and let $S \subseteq M_K$ be a finite set which contains M_K^∞.

Let $\mathcal{O}_{K,S}$ denote the ring of S-integers, *i.e.* $\mathcal{O}_{K,S} = \{x \in K : |x|_v \leq 1 \text{ for all } v \notin S\}$.

The goal of this chapter is to prove the theorem of Siegel.

Theorem 6.0.1. *The equation $Y^2 = X^3 + aX + b$, with $a, b \in \mathcal{O}_{K,S}$ such that $4a^3 + 27b^2 \neq 0$, has only finitely many solutions in $\mathcal{O}_{K,S}^2$.*

A fancier and perhaps more intrinsic way of this theorem is: Each elliptic curve (E, O) defined over K has only finitely many $\mathcal{O}_{K,S}$-points with respect to the divisor $\{O\}$.

6.1 Background on elliptic curves

Let us start with an abstract definition of elliptic curves and then explain how to link it with the equation in Theorem 6.0.1.

Definition 6.1.1. *An elliptic curve is a smooth projective curve E of genus one together with a prescribed point $O \in E$. We say that the elliptic curve is defined over K if E is defined over K as a curve and $O \in E(K)$.*

Usually the point O is well understood, so we simply call E an elliptic curve. Indeed, by theory of curves of genus 1, over the field of complex numbers $\mathbb{C}$, we have $E(\mathbb{C}) \simeq \mathbb{C}/\Lambda$ for a lattice $\Lambda \subseteq \mathbb{C}$ and O is the image of $0 \in \mathbb{C}$ under the natural uniformization $\mathbb{C} \to \mathbb{C}/\Lambda$.

The group $(\mathbb{C}, +; 0)$ induces a natural (abelian) group structure on E, with O being the identity element.

For E/K, set

$$E(K) := \{P \in E(\mathbb{C}) : \sigma(P) = P \text{ for all } \sigma \in \mathrm{Aut}(\mathbb{C}/K)\}.$$

It is not hard to check that $O \in E(K)$ and that $E(K)$ is an abelian group. The following Mordell–Weil theorem is of fundamental importance in the theory of elliptic curves.

Theorem 6.1.2 (Mordell–Weil Theorem)**.** *The abelian group $E(K)$ is finitely generated.*

In this course, we only need a weak version of this theorem, namely

Theorem 6.1.3 (Weak Mordell–Weil Theorem)**.** *For each positive integer $m \in \mathbb{Z}$, the group $E(K)/mE(K)$ is finite.*

Next we turn to a more concrete description of elliptic curves. Using the Weierstraß $\wp$-function, one can prove the following proposition. It can also be obtained as an application of the Riemann–Roch Theorem.

Proposition 6.1.4. *Let E be an elliptic curve defined over K.*

(i) There exist functions $x, y \in K(E)$ such that the map

$$\phi\colon E \to \mathbb{P}^2, \quad P \mapsto [x(P) : y(P) : 1]$$

gives an isomorphism of E onto a curve in $\mathbb{P}^2$ such that $\phi(O) = [0 : 1 : 0]$ and that $\phi(E \setminus \{O\})$ is given by

$$Y^2 = X^3 + aX + b, \tag{6.1.1}$$

in the affine coordinate $[X : Y : 1]$, for some $a, b \in K$. Moreover $4a^3 + 27b^2 \neq 0$.

(ii) Conversely, each curve defined by an equation in the form (6.1.1) such that $4a^3 + 27b^2 \neq 0$ is an elliptic curve defined over K with $O = [0 : 1 : 0]$.

(iii) Given E, the choice of the equation (6.1.1) is not unique. But any two such choices are related by a change of variables of the form $(X, Y) \mapsto (u^2X, u^3Y)$ for some $u \in K^*$.

The equation (6.1.1) is called the *Weierstraß form* of E. We sometimes use the following notation

$$E/K: \quad Y^2 = X^3 + aX + b$$

to mean that E is an elliptic curve defined over K in its *Weierstraß form*. By Proposition 6.1.4.(iii), the Weierstraß form of E is not unique.

The group law on E under the Weierstraß form can be made explicit. Here we only need the following observation. For $P = [x : y : 1] \in E$, the negation $-P = [x : -y : 1]$.

Given an elliptic curve E/K, Theorem 6.0.1 says that any Weierstraß form has only finitely many solutions in $\mathcal{O}_{K,S}^2$. However, one can show that two different Weierstraß forms of E/K may not have the same number of solutions in $\mathcal{O}_{K,S}^2$, and by varying the Weierstraß form the number of solutions in $\mathcal{O}_{K,S}^2$ may not have an upper bound.[1]

6.2 Link with Roth's Theorem

We will prove Theorem 6.0.1 as an application of Roth's Theorem. Let us repeat the statement of Roth's Theorem here. Here we need a version which is more general than Theorem 3.1.4 but less general than Theorem 3.1.5.

Let $v \in M_K$ and let $\alpha_v \in K_v$ be K-algebraic, *i.e.* $\alpha_v \in K_v$ is a root of a polynomial with coefficients in K. Then for each $\epsilon > 0$,

$$\log |\alpha_v - \beta|_v > -(2 + \epsilon)h(\beta) \tag{6.2.1}$$

for all but finitely many $\beta \in K$. In the case of $K = \mathbb{Q}$ and $v = \infty$, this is precisely Theorem 3.1.4.

Let $E/K: \quad Y^2 = X^3 + aX + b$ be an elliptic curve defined over K in its Weierstraß form. We need to understand the analogous inequality of (6.2.1) for E. The right hand side is *height*, and the left hand side is a suitable *distance*.

[1] In fact, such an upper bound exists if and only if $E(K)$ has rank 0.

6.2.1 Height on E

Define the finite morphism

$$f\colon E \to \mathbb{P}^1, \quad \begin{cases} [x:y:1] \mapsto [x:1] \\ [0:1:0] \mapsto [1:0]. \end{cases} \tag{6.2.2}$$

If we set $[1:0] \in \mathbb{P}^1$ to be ∞, then $f([x:y:1]) = x$ and $f(O) = \infty$.

Set

$$L = f^*\mathcal{O}_{\mathbb{P}^1}(1). \tag{6.2.3}$$

Since f is a finite morphism, we have that L is ample on E.

Lemma 6.2.1. *Let $[-1]\colon E \to E$ be the multiplication-by--1 map on E. Then*

$$L \simeq [-1]^*L.$$

In particular, we have the Néron–Tate height $\hat{h}_L\colon E(\overline{\mathbb{Q}}) \to \mathbb{R}_{\geq 0}$ from (5.3.2).

Proof. For each $P = [x:y:1] \in E$, we have $-P = [x:-y:1]$. Thus $f(P) = f(-P)$ for all $P \in E$. Thus $f = f \circ [-1]$. Hence $L = f^*\mathcal{O}_{\mathbb{P}^1}(1) = (f \circ [-1])^*\mathcal{O}_{\mathbb{P}^1}(1) = [-1]^*f^*\mathcal{O}_{\mathbb{P}^1}(1) = [-1]^*L$. □

6.2.2 Distance function on E

Let $Q \in E$. Take t_Q to be a local uniformizer at Q, *i.e.* $t_Q \in K(E)$ such that Q is a zero of t_Q of order 1. Such a t_Q exists; for example if $Q = [x_0 : y_0 : 1]$ with $y_Q \neq 0$, then we can take $t_Q = x - x_0$.

Definition 6.2.2. *Let $v \in M_K$. Let $P \in E(K_v)$. The v-adic distance between P and Q with respect to t_Q is defined to be*

$$d_v(P, t_Q) := \min\{|t_Q(P)|_v, 1\}.$$

If P is a pole of t_Q, then we naturally set $d_v(P, t_Q) = 1$.

Notice that $d_v(P, t_Q)$ depends on the choice of the local uniformizer t_Q at Q. However, for our purpose we only need to understand this distance in the limit process with $P \in E(K_v)$ approaches Q in the v-adic topology.

Lemma 6.2.3. *Let $t'_Q \in K(E)$ be such that Q is a zero of t'_Q of order $e \geq 1$. Then we have*

$$\lim_{P \in E(K_v),\ P \xrightarrow{v} Q} \frac{\log \min\{|t'_Q(P)|_v^{1/e}, 1\}}{\log d_v(P, t_Q)} = 1.$$

Proof. Let $\phi := t'_Q / t_Q^e \in K(E)$. Then Q is neither a zero nor a pole of ϕ. Hence when P is sufficiently close to Q, $|\phi(P)|_v$ is bounded away from 0 and ∞. Thus

$$\lim_{P \in E(K_v),\ P \xrightarrow{v} Q} \frac{\log \min\{|t'_Q(P)|_v^{1/e}, 1\}}{\log d_v(P, t_Q)} = 1 + \lim_{P \in E(K_v),\ P \xrightarrow{v} Q} \frac{\log |\phi(P)|_v^{1/e}}{\log d_v(P, t_Q)} = 1.$$

We are done. □

Because of this lemma, we will write

$$d_v(P,Q)$$

for $d_v(P,t_Q)$ when $P \in E(K_v)$ approaches Q in the v-adic topology.

The following result is then a corollary of Roth's Theorem.

Corollary 6.2.4. *Let L be the ample line bundle on E defined by (6.2.3). Then*

$$\liminf_{P\in E(K),\ P\xrightarrow{v}Q} \frac{\log d_v(P,Q)}{\hat{h}_L(P)} \geq -2.$$

Proof. Recall the morphism $f\colon E \to \mathbb{P}^1$ given by (6.2.2). By Height Machine ((i) and (iii) of Proposition 5.1.3), we have $\hat{h}_L(P) = \hat{h}_{f^*\mathcal{O}_{\mathbb{P}^1}(1)}(P) = h_{\mathcal{O}_{\mathbb{P}^1}(1)}(f(P)) + O(1) = h(f(P)) + O(1)$ where $h\colon \mathbb{P}^1(\overline{\mathbb{Q}}) \to \mathbb{R}$ is the Weil height.

The function $f - f(Q)$ is an element in $K(E)$ which vanishes at Q, whose vanishing order we denote by $e \geq 1$. Then by Lemma 6.2.3, we can take

$$d_v(P,Q) = \min\{|f(P) - f(Q)|_v^{1/e}, 1\}.$$

Thus

$$\liminf_{P\in E(K),\ P\xrightarrow{v}Q} \frac{\log d_v(P,Q)}{\hat{h}_L(P)} = \frac{1}{e} \liminf_{P\in E(K),\ P\xrightarrow{v}Q} \frac{\log |f(P) - f(Q)|_v}{h(f(P))}.$$

By (6.2.1) with $\beta = f(P)$ and $\alpha_v = f(Q)$, we then have

$$\liminf_{P\in E(K),\ P\xrightarrow{v}Q} \frac{\log |f(P) - f(Q)|_v}{h(f(P))} \geq -(2+\epsilon)$$

for any $\epsilon > 0$. Hence we are done. □

6.3 Conclusion of Siegel's Theorem

Now we are ready to prove Theorem 6.0.1, the main theorem of this chapter.

Let $E/K : \quad Y^2 = X^3 + aX + b$ be an elliptic curve. Let L be the ample line bundle on E defined by (6.2.3). For each $v \in M_K$, use the distance function defined above Corollary 6.2.4.

Theorem 6.3.1 (Siegel)**.** *Assume $\#E(K) = \infty$. Fix $Q \in E(K)$ and let $v \in M_K$. Then*

$$\lim_{P\in E(K),\ \hat{h}_L(P)\to\infty} \frac{\log d_v(P,Q)}{\hat{h}_L(P)} = 0.$$

6.3.1 Proof of Theorem 6.3.1 implying Theorem 6.0.1

For each $P \in E(K) \setminus \{O\}$, denote by $[x(P) : y(P) : 1]$ its coordinates. Then by definition of L and the Height Machine ((i) and (iii) of Proposition 5.1.3), we have

$$\hat{h}_L(P) = \hat{h}_{f^*\mathcal{O}_{\mathbb{P}^1}(1)}(P) = h_{\mathcal{O}_{\mathbb{P}^1}(1)}(f(P)) + O(1) = h([x(P) : 1]) + O(1)$$

for each $P \in E(K) \setminus \{O\}$, where $O(1)$ is a bounded function. Here the last step is the definition of $f\colon E \to \mathbb{P}^1$ (6.2.2). Thus by definition of the Weil height, we have

$$\hat{h}_L(P) = \frac{1}{[K:\mathbb{Q}]} \sum_{v \in M_K} \log^+ \|x(P)\|_v + O(1).$$

If $x(P) \in \mathcal{O}_{K,S}$, then $\|x(P)\|_v = |x(P)|_v^{[K_v:\mathbb{Q}_p]} \leq 1$ for all $v \notin S$. Notice that $[K_v : \mathbb{Q}_p] \leq [K : \mathbb{Q}]$. Thus we have

$$x(P) \in \mathcal{O}_{K,S} \Rightarrow \hat{h}_L(P) = \frac{1}{[K:\mathbb{Q}]} \sum_{v \in S} \log \|x(P)\|_v + O(1) \leq \sum_{v \in S} \log |x(P)|_v + O(1). \tag{6.3.1}$$

We claim that there are only finitely many $P \in E(K) \setminus \{O\}$ with $x(P) \in \mathcal{O}_{K,S}$. Notice that this suffices to conclude for Theorem 6.0.1. Assume otherwise, then there is a sequence of distinct points $P_1, P_2, \ldots \in E(K)$ with $x(P_n) \in \mathcal{O}_{K,S}$ for each n. By Northcott property (Proposition 5.1.3.(v)), $\hat{h}_L(P_n) \to \infty$. By (6.3.1), up to taking a subsequence there exists $v \in S$ with

$$\hat{h}_L(P_n) \leq \#S \cdot \log |x(P_n)|_v \quad \text{for all } n. \tag{6.3.2}$$

The function x has a pole of order 2 at the point $O \in E(K)$. Therefore we may take

$$d_v(P_n, O) = \min\{|x(P_n)|_v^{-1/2}, 1\}$$

by Lemma 6.2.3. Thus (6.3.2) implies

$$\frac{-\log d_v(P_n, O)}{\hat{h}_L(P_n)} \geq \frac{1}{2\#S}.$$

However, as $\hat{h}_L(P_n) \to \infty$, this contradicts Theorem 6.3.1 with $Q = O$. Hence we obtain a contradiction. Therefore there are only finitely many $P \in E(K) \setminus \{O\}$ with $x(P) \in \mathcal{O}_{K,S}$. Hence we are done.

6.3.2 Proof of Theorem 6.3.1

Choose a sequence of distinct points $P_n \in E(K)$ such that

$$\lim_{n\to\infty} \frac{\log d_v(P_n, Q)}{\hat{h}_L(P_n)} = \liminf_{P \in E(K),\ \hat{h}_L(P) \to \infty} \frac{\log d_v(P, Q)}{\hat{h}_L(P)}.$$

Call this limit L. Since $d_v(P_n, Q) \leq 1$ by definition of d_v and $\hat{h}_L(P_n) \geq 0$ by Proposition 5.3.2.(i), we have $L \leq 0$.

Now it remains to show $L \geq 0$.

Let $m \in \mathbb{Z}$ be a positive integer. By the weak Mordell–Weil Theorem (Theorem 6.1.3), $E(K)/mE(K)$ is finite. Thus there exists $R \in E(K)$ such that $mE(K) + R$ (which is a coset of $mE(K)$ in $E(K)$) contains infinitely many points in the sequence $\{P_n\}$. Replacing the sequence by this subsequence, we may assume $P_n \in mE(K) + R$ for all n. Write for each n

$$P_n = [m]P_n' + R$$

with $P_n' \in E(K)$. Then Proposition 5.3.1 yields

$$m^2 \hat{h}_L(P_n') = \hat{h}_L([m]P_n') = \hat{h}_L(P_n - R).$$

By Proposition 5.3.3 and the Polarization Identity, we have $\hat{h}_L(P_n+R)+\hat{h}_L(P_n-R) = 2\hat{h}_L(P_n)+2\hat{h}_L(R)$. Therefore by Proposition 5.3.2.(i), we have

$$m^2\hat{h}_L(P'_n) \leq 2\hat{h}_L(P_n) + 2\hat{h}_L(R). \tag{6.3.3}$$

Next we work with the v-adic distance. Notice that up to taking a subsequence, we may assume $P_n \xrightarrow{v} Q$; otherwise $\log d_v(P_n, Q)$ is bounded and clearly $L = 0$. Hence $[m]P'_n \xrightarrow{v} Q - R$, and therefore at least one of the m^2 possible m-th roots of $Q - R$ is an accumulation point of the sequence $\{P'_n\}$. Again by taking a subsequence, there exists $Q' \in E(\overline{\mathbb{Q}})$ such that

$$P'_n \xrightarrow{v} Q' \quad \text{and} \quad Q = [m]Q' + R.$$

Up to replacing K by a finite extension and by replacing v with a place above, we may assume $Q' \in E(K)$.

Now we need a result from Algebraic Geometry: the morphism $\varphi\colon E \to E$, $P \mapsto [m]P + R$, is everywhere unramified. We claim that

$$\lim_{n\to\infty} \frac{\log d_v(P_n, Q)}{\log d_v(P'_n, Q')} = 1. \tag{6.3.4}$$

Assuming (6.3.4). Then (6.3.3) yields

$$L = \lim_{n\to\infty} \frac{\log d_v(P_n, Q)}{\hat{h}_L(P_n)} \geq \lim_{n\to\infty} \frac{\log d_v(P'_n, Q')}{\frac{1}{2}m^2\hat{h}_L(P'_n) - \hat{h}_L(R)}.$$

Now we apply Corollary 6.2.4 to the right hand side of this inequality. Then we obtain

$$L \geq -4/m^2.$$

This is true for any positive integer m. Therefore $L \geq 0$. We are done.

Now it remains to prove (6.3.4). Let $t_Q \in K(E)$ be such that Q is a zero of t_Q of order 1. Then Q' is a zero of the rational function $t_Q \circ \varphi \in K(E)$. Moreover, since φ is unramified at Q', the order of Q' is 1 (as a zero of $t_Q \circ \varphi$). Therefore we can take $d_v(P'_n, Q') = \min\{|t_Q(\varphi(P'_n))|_v, 1\} = \min\{|t_Q(P_n)|_v, 1\}$, which is precisely $d_v(P_n, Q)$. Hence (6.3.4) holds true.