The Mordell-Weil Theorem

Guo Li

January 1, 2025

Contents

1 Introduction

The main purpose of this note is to demonstrate:

Theorem 1.1 (Mordell-Weil Theorem)**.** *Let K be a number field and E an elliptic curve over K. then the group $E(K)$ is finitely generated.*

The proof path is usually divided into two steps:

(i) Weak Mordell-Weil thm: For any $m \geq 2, E(K)/mE(K)$ is a finite abelian group.

(ii) the "infinite descent" argument: Through a general theorem, we see that the key of this part is to define the "height".

The proof of the week Mordell-Weil thm for $\mathbb{Q}$ comes from Chapter 19 of Book [2], which combines Weil's original 1929 paper and Cassels' simplified version. And The rest mainly refers to [1] and [3]. After the proof, we will discuss how to calculate some simple examples of calculating $E(\mathbb{Q})$ (including the rank and torsion part), as well as some related famous results and conjectures.

- Selmer group and Tate–Shafarevich group. The two essential concepts.
- The famous Birch and Swinnerton-Dyer conjecture, the Tate–Shafarevich conjecture, Ogg's conjecture (about torsion part) $\cdots$
- Roth's Theorem (about Diophantine approximation), Faltings's theorem (Mordell conjecture) $\cdots$

2 The Week Mordell-Weil Theorem

K a number field with characteristic 0
$\overline{K}$ the algebraic closure of K
$E(F)$ the group of F-rational points of E
$A[m]$ the m-torsion subgroup of an abelian group A

Let E be an elliptic curve over K with Weierstrass equation $y^2 = f(x) = x^3 + ax + b, a, b \in K$, which we assume to be non-singular (i.e. has three distinct roots in $\overline{K}$, or the discriminant $\Delta = -(4a^3 + 27b^2) \neq 0$). We can define a group law on E by setting the sum of collinear three pionts $P, Q, R \in E(K)$ zero with O the point at infinity.

Remark 2.1 (The group law on E)**.**

We have other ways to define the group structure on E:

(1) There is a one-to-one correspondence

$$\sigma : E \to \mathrm{Pic}^0(E), P \mapsto [(P) - (O)]$$

(By the Riemann–Roch theorem), and this induces an addition law on E.

(2) If E is defined over $\mathbb{C}$, we can the fact that the following categories are equivalent

$$\left(\begin{array}{l} \textit{Objects} : \textit{Elliptic curves } E/\mathbb{C} \\ \textit{Maps} : \textit{Isogenies} \end{array} \right) \leftrightsquigarrow \left(\begin{array}{l} \textit{Objects} : \textit{Lattices (up to homothety) } \Lambda \subset \mathbb{C} \\ \textit{Maps} : \{\alpha \in \mathbb{C} \mid \alpha\Lambda_1 \subset \Lambda_2\} \end{array} \right)$$

to get the group law on E from the natural addition law on $\mathbb{C}/\Lambda$.

2.1 the Week Theorem for $\mathbb{Q}$

Assume $K[x]/(f(x)) = K[\xi]$ with $\xi \equiv x \bmod f(x)$ and denote the group of units of $K[\xi]$ by U. U contains elements $h(x)$ mod $f(x)$ with $h(x)$ coprime to $f(x)$, and has direct product decomposition depending on the decomposition of $f(x)$ over K.

2.1.1 the Weil map

The key point of this section is a map $\phi : E \to U/U^2$ which was defined by Weil in his 1929 paper and simplified by Cassels.

Define ϕ as follows: $\phi(O) := 1 \in U/U^2$. For $P = (\alpha, \beta) \in E(K)$, if $\beta \neq 0$, then $\alpha - \xi \in U$ and define $\phi(P)$ be the image of $\alpha - \xi$ in U/U^2; if $\beta = 0$, then $(x - \alpha)g(x) = f(x)$ and define $\phi(P)$ be the image of $(f'(\alpha), (\alpha - x) \bmod g(x))$ in U/U^2 by consider $K[\xi] = K[x]/(x - \alpha) \oplus K[x]/(g(x))$.

Lemma 2.2. *ϕ is a homomorphism with kernel $2E$.*

Proof. First, we see $\phi(P) = \phi(-P)$ and $\phi(P)^2 = 1$, so $\phi(P + Q) = \phi(P)\phi(Q) \iff \phi(P + Q)\phi(-P)\phi(-Q) = 1$ and we just show that if $A_1 + A_2 + A_3 = O$ on E, then $\phi(A_1)\phi(A_2)\phi(A_3) = 1$. Put $A_k = (x_k, y_k)$, and we can assume none of them be O, they are distinct and $x_1 \neq x_2$. If none of them has order 2, then by the collinearity of A_k, there is nontrivial linear form $cx+d$ with $c, d \in K$ s.t. $f(x)-(cx+d)^2 = (x-x_1)(x-x_2)(x-x_3)$. Just mod $f(x)$ to get the result. If precisely A_1 has order 2, then $\phi(A_1)\phi(A_2)\phi(A_3) = ((f'(\alpha))^2, *)$ with $* \in U^2$ by the former case and mod $\frac{f(x)}{x-x_1}$. The final case is all A_k have order 2 and we see that the three components of $\phi(A_k)$ are $f'(x_1)^2, f'(x_2)^2, f'(x_3)^2$. Hence ϕ is a homomorphism.

It is clear that $2E \subset \ker \phi$. Consider $P = (\alpha, \beta) \neq O$ but $\phi(P) = 1$ then $\alpha - \xi = (\alpha_1\xi^2 + \alpha_2\xi + \alpha_3)^2 \in K[\xi]$. We see $\alpha_1 \neq 0$ and can get $(e\xi + e')^2 = (\alpha - \xi)(h - \xi)^2$ by using $\xi^3 + a\xi + b = 0$ and choosing suitable $e, e', h \in K$. So $f(x) \mid (ex + e')^2 - (\alpha - x)(h - x)^2$, the latter is monic and then $f(x) - (ex + e')^2 = (\alpha - x)(h - x)^2$. Note that $ex + e' \neq 0$ because E is nonsingular, so the equation tell us there is a point Q with x-coordinate being h satisfies $P = 2Q$. Hence $2E = \ker \phi$. □

2.1.2 the finiteness of $\phi(E)$ and $E/2E$

Now we consider $K = \mathbb{Q}$, $\theta \in \mathbb{C}$ s.t. $f(\theta) = 0$, and let $f(x) = (x - \theta)g(x)$. Then for $P = (\alpha/\beta, w) \neq O$ on $E(\mathbb{Q})$ with α, β are coprime integers, $\alpha - \beta\theta, h_{\alpha,\beta} = g(\alpha/\beta)\beta^2$ are algebraic integers of $\mathbb{Q}(\theta)$. We now consider all in $\mathbb{Q}(\theta)$ and $\mathcal{O}_{\mathbb{Q}(\theta)}$.

By some elmentary calculation, there is

Lemma 2.3. *all rational pionts on $E/\mathbb{Q}$ has form $(\frac{s}{d^2}, \frac{t}{d^3})$ with $s, t, d \in \mathbb{Z}$ and $d > 0, \gcd(s, d) = \gcd(t, d) = 1$.*

Lemma 2.4. *The set of ideals $I(P) := (\alpha - \beta\theta, h_{\alpha,\beta})$ is finite and $(\alpha - \beta\theta) = I(P)C^2$ for some ideal C.*

Proof. By noting that $(x - \theta) \mid [g(x) - g(\theta)]$ and $(x - \theta) \mid [x^2g(\theta) - \theta^2g(x)]$, we get $g(\theta)\beta^2, g(\theta)\alpha^2 \in I(P)$. α, β are coprime, then $I(P) \mid (g(\theta))$, which shows there are only a finite number of possibilities for $I(P)$.

Hence β is a square. But $f(\frac{\alpha}{\beta}) = (\frac{l}{m})^2$ with $\frac{l}{m} \in \mathbb{Q}$. Thus $\beta^3l^2 = m^2(\alpha - \beta\theta)h_{\alpha,\beta} \in (\mathbb{Z})^2$. Hence we must have $(\alpha - \beta\theta) = I(P)C^2$ for some ideal C. □

Finally, we use the facts that $\#\mathrm{Cl}(\mathbb{Q}(\theta)) < \infty$ and the Dirichlet unit theorem to get the following theorem.

Theorem 2.5. $\#(E/2E) < \infty$.

First we show there is a finite set S s.t. for each P above, $\alpha - \beta\theta = u\gamma\tau^2$ for some u in the unit group of $\mathcal{O}_{\mathbb{Q}(\theta)}$, $\gamma \in S$ and $\tau \in \mathbb{Q}(\theta)^*$. Let $C_1, \cdots, C_h$ be representatives for the ideal classes. Thus $(\alpha - \beta\theta) = I(P)C^2 \sim I(P)C_s$ for some s and the latter is then principal, say (γ). Hence $\{(\gamma)\}$ is finite because $\{I(P)\}$ is finite. And there are $\rho, \tau_1 \in \mathcal{O}_{\mathbb{Q}(\theta)}$ s.t. $\rho C = \tau_1 C_s$, which implies $(\rho^2(\alpha - \beta\theta)) = (\gamma\tau_1^2)$, so denote $\tau = \frac{\tau_1}{\rho}$ to get waht we want.

Now it is enough to show $\phi(E)$ is finite. Assume $P \neq O$ with order not 2. If we consider the i^{th} component $K_i := \mathbb{Q}(\theta_i)$ of $\mathbb{Q}[x]/(f(x))$ with assume $f(x) = (x-\theta_1)(x-\theta_2)(x-\theta_3)$ in $\mathbb{C}[x]$, then in $K_i^*/(K_i^*)^2$, $\alpha/\beta - x \sim \frac{1}{\beta}u\gamma \sim u\gamma$. Dirichlet unit theorem says u mod $(K_i^*)^2$ is finite, S is finite, so is $\phi(E)$.

The process of this proof can be summarized as follows. First we consider the Weil map $\phi : E \to U/U^2$ and show that $\ker\phi = 2E$, so tranlate the proof for the finiteness of $E/2(E)$ to prove $\phi(E)$ is finite, which is obtained by using the finiteness of the class group and the Dirichlet unit theorem at last.

2.2 the Week Theorem for number field K

2.2.1 the Kummer pairing: Motivation

We point out that the motivation for this proof comes from the analogy of Kummer's theory on number fields.

Let ζ_m be a primitive m-th root of unity, μ_m be the group of m-th roots and denote $G_{L/K}$ as Galois group $\mathrm{Gal}(L/K)$. Recall that if $\zeta_m \in K$, then we have a long exact sequence with taking $G_{\overline{K}/K}-$cohomology

$$1 \to \mu_m \to K^* \xrightarrow{z \mapsto z^m} K^* \xrightarrow{\delta} H^1(G_{\overline{K}/K}, \mu_m) \to H^1(G_{\overline{K}/K}, \overline{K}^*) \to \cdots$$

By Hilbert's theorem 90, we know that $H^1(G_{\overline{K}/K}, \overline{K}^*) = 0$ and then we have a isomorphism which can explicitly be written as

$$\delta : K^*/(K^*)^m \xrightarrow{\sim} H^1(G_{\overline{K}/K}, \mu_m) \qquad a \mapsto ([\sigma] \mapsto \frac{\alpha^\sigma}{\alpha})$$

where $\alpha \in \overline{K}^*$ is any element s.t. $\alpha^m = a$. And then we have a perfect bilinear pairing:

$$K^*/(K^*)^m \times G_{\overline{K}/K} \to \mu_m \qquad (a, \sigma) \mapsto \frac{\alpha^\sigma}{\alpha}$$

Looking at the world of elliptic curves, we notice something similar:

$$1 \to E(\overline{K})[m] \to E(\overline{K}) \xrightarrow{P \mapsto mP} E(\overline{K}) \to 0$$

If $E[m] \subset E(K)$ (similar to assume $\zeta_m \in K$), then we also take the $G_{\overline{K}/K}-$cohomology to get an exact sequence

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

The failure of $H^1(G_{\overline{K}/K}, E(\overline{K}))$ to vanish adds a key complexity to the theory. But we still analysis the map $\delta : P \mapsto ([\sigma] \mapsto Q^\sigma - Q)$ with $Q \in E(\overline{K})$ s.t. $mQ = P$. And thus we hope there is a similar bilinear pairing.

Remark 2.6 (Similarity between Curves and Number Fields)**.**
The first is two contravariant category equivalences (there we assume $K = \overline{K}$):

$$\left(\begin{array}{l} \textit{Objects}: \textit{irreducible quasi-projective varieties } V \\ \textit{Maps}: \textit{dominant rational maps } f : V \dashrightarrow W \end{array}\right) \leftrightsquigarrow \left(\begin{array}{l} \textit{Objects}: \textit{finitely generated extensions of} K \\ \textit{Maps}: K\textit{-homomorphisms} \end{array}\right)$$

$$\left(\begin{array}{l} \textit{Objects}: \textit{smooth curves} \\ \textit{Maps}: \textit{nonconstant rational maps} \end{array}\right) \leftrightsquigarrow \left(\begin{array}{l} \textit{Objects}: \textit{finitely generated extensions of } K \textit{ with} \\ \textit{transcendence degree one} \\ \textit{Maps}: K\textit{-homomorphisms} \end{array}\right)$$

There is also a similar version for an important exact sequence in algebraic number theory:

$$1 \longrightarrow \mathcal{O}_K^* \longrightarrow K^* \longrightarrow (\textit{fractional ideals of } K) \longrightarrow \mathrm{Cl}(K) \longrightarrow 1.$$

which says

$$1 \longrightarrow \overline{K}^* \longrightarrow \overline{K}(E)^* \xrightarrow{\mathrm{div}} \mathrm{Div}^0(E) \longrightarrow \mathrm{Pic}^0(E) \longrightarrow 0.$$

This follows that $\mathrm{div}(f) = 0 \iff f \in \overline{K}^*$ *and* $\deg(\mathrm{div}(f)) = 0$.

2.2.2 the Kummer pairing: Construction

Now that the analogy is over, we have two things to do next. One is to show that we can always assume $E[m] \subset E(K)$, and the other is to actually define the bilinear form called Kummer pairing.

$E[m]$ is a finite group and then there is a finite extension L/K s.t. $E[m] \subset E(L)$, so we just need the following lemma.

Remark 2.7 (The structure of m-torsion points)**.**
Actually We have:
(i) If $\mathrm{char}K = 0$ *or* $\mathrm{char}K \nmid m$*, then* $E[m] \cong \mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/m\mathbb{Z}$.
(ii) If $\mathrm{char}K = p$ *then* $E[p^e] \cong \mathbb{Z}/p^e\mathbb{Z}$ *for all* $e \in \mathbb{N}$ *or* $E[p^e] \cong \{O\}$ *for all* $e \in \mathbb{N}$.
See III 6.4 of [1]. If $E/\mathbb{C}$, We can intuitively get the above result by noting $E(\mathbb{C}) \cong \mathbb{C}/\Lambda$ for some lattice $\Lambda \subset \mathbb{C}$, thus $E[m] \cong (\mathbb{C}/\Lambda)[m] \cong \mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/m\mathbb{Z}$. (As for the case $\mathrm{char}(K) = 0$*, we can use the Lefschetz Principle to get similar results.)*

Lemma 2.8. *If L/K is finite and Galois, then $E(L)/mE(L)$ is finite implies $E(K)/mE(K)$ is finite.*

Proof. We have a natural map $E(K)/mE(K) \to E(L)/mE(L)$ with kernel $\Phi = \dfrac{E(K) \cap mE(L)}{mE(K)}$. By using the inflation-restriction sequence, we have the following commutative diagram:

$$\begin{tikzcd}
0 \arrow[r] \arrow[d, hook] & \Phi \arrow[r] \arrow[d, dashed] & E(K)/mE(K) \arrow[r] \arrow[d, hook] & E(L)/mE(L) \arrow[d, hook] \\
0 \arrow[r] & H^1(G_{L/K}, E[m]) \arrow[r, "inf"] & H^1(G_{\overline{K}/K}, E[m]) \arrow[r, "res"] & H^1(G_{\overline{L}/L}, E[m])
\end{tikzcd}$$

The short five lemma shows that the dashed arrow is injective, so Φ is finite because $G_{L/K}$ and $E[m]$ are finite. Now $E(L)/mE(L)$ is finite, so is $E(K)/mE(K)$.

There is another elmentary way to see $\Phi \hookrightarrow \mathrm{Map}(G_{L/K}, E[m])$. By setting $\lambda_P : G_{L/K} \to E[m], \sigma \mapsto Q^\sigma - Q$ with $Q \in E(L), mQ = P$. If λ_P is trivial, then $Q^\sigma = Q$ for all σ, so $Q \in E(K)$ and $P \in mE(K)$, which shows the injectivity. Then also get Φ is finite. □

Now we define the Kummer pairing:

$$\kappa : E(K) \times G_{\overline{K}/K} \to E[m] \qquad (P, \sigma) \mapsto Q^\sigma - Q$$

where $Q \in E(\overline{K})$ s.t. $mQ = P$. Recall that $E(K)/mE(K) \overset{\delta}{\hookrightarrow} H^1(G_{\overline{K}/K}, E[m]) = \mathrm{Hom}(G_{\overline{K}/K}, E[m])$ ("=" follow that $G_{\overline{K}/K}$ trivially act on $E[m] \subset E(K)$), then we have κ is well defined, bilinear and with left kernel $mE(K)$.

Lemma 2.9. *The right kernel of κ is $G_{\overline{K}/L}$ where $L = K\left(\frac{1}{m}E(K)\right)$, and then the bilinear pairing $E(K)/mE(K) \times G_{L/K} \to E[m]$ is perfect.*

Proof. If $\sigma \in G_{\overline{K}/L}$ then $\kappa(P, \sigma) \equiv O$ since $Q \in E(L)$. Conversely, if $\kappa(P, \sigma) \equiv O$ for all $P \in E(K)$, then $Q^\sigma = Q$ for all $Q \in E(\overline{K})$ s.t. $mQ = P$. So $\sigma \in G_{\overline{K}/L}$. And $G_{\overline{K}/L} = \mathrm{Ker}\left(G_{\overline{K}/K} \to \mathrm{Hom}(E(K), E[m])\right)$ thus normal. Hence L/K is Galois and the latter bilinear pairing is perfect. □

2.2.3 the finiteness of L/K and $E(K)/mE(K)$

Now we can prove $E(K)/mE(K)$ is finite by just showing that $G_{L/K}$ is finite, and this need to analyze the field L.

Denote $M_K = M_K^\infty \cup M_K^0$ as a complete set of inequivalent absolute values on K with the former being archimedean and the latter being non-archimedean. For $v \in M_K^0$, we denote $\mathfrak{m}_v, k_v$ as the maximal ideal and residue field of $\mathcal{O}_{K_v}$.

We discuss the reduction of E. Let $v \in M_K^0$,then E is said to have good (respectively bad) reduction at v if E has good (respectively bad) reduction when considered over the completion K_v, which means the reduction modulo $\mathfrak{m}_v$ of a minimal (means that in this time $v(\Delta)$ is the smallest possible nonnegative integer among all the Weierstrass equation for E) Weierstrass equation for E, denoted by $\widetilde{E}$ is non-singular (respectively singular). We point out two important facts:

(i) E has good reduction at v if and only if $v(\Delta) = 0$, in this case $\widetilde{E}/k_v$ is an elliptic curve (See VII 5.1 of [1]).

(ii) If E has good reduction at v, $v(m) = 0$, then the reduction map $E(K_v)[m] \to \widetilde{E}(k_v)$ is injective (See VII 3.1 of [1]).

We are now ready to analyze the extensionL/K.

Proposition 2.10. *L/K is abelian with $G_{L/K}$ having exponent m. And L/K is unramified outside the finite set $S := \{v \in M_K^0 \mid E \text{ has bad reduction at } v\} \cup \{v \in M_K^0 \mid v(m) \neq 0\} \cup M_K^\infty$.*

Proof. The first part is clear because $G_{L/K} \hookrightarrow \mathrm{Hom}(E(K), E[m])$.

Let $v \notin S$ and we just need show if $Q \in E(\overline{K})$ s.t. $mQ \in E(K)$, then $K' := K(Q)$ is unramified over K at v because the compositum of some unramified fields is also unramified.

And let $v' \in M^0_{K'}$ be the place over v and then E has a good reduction at v' with reduction map $E(K') \to \widetilde{E}(k'_{v'})$. Take an element of the inertia group $I_{v'/v}$ of v'/v, say σ, which acts trivially on $\widetilde{E}(k'_{v'})$, so $\widetilde{Q^\sigma - Q} = \widetilde{O}$ and $m(Q^\sigma - Q) = O$ because of $mQ \in E(K)$. By the injectivity of $E(K')[m] \to \widetilde{E}(k'_{v'})$, we get $Q^\sigma - Q = O$. Hence $K(Q)$ is unramified over K at v' for Q is fixed by $I_{v'/v}$. This holds for every v' over v and every $v \notin S$. Thus $K(Q)/K$ is unramified outside S and the Proposition follows. □

The proof of the weak Mordell - Weil theorem follows that any field extension L/K satisfying the above Proposition is finite, which relies on two facts of algebraic number theory: the finiteness of the class group and the finiteness of the $S-$unit group.

Proposition 2.11. *Let L/K be an abelian extension with exponent m, and is unramified outside a finite set $M_K^\infty \subset S \subset M_K$, then L/K is finite.*

Proof. Firstly, we see that if the Proposition is true for some finite extension K'/K, then it is true for K, then we can assume $\mu_m \subset K$. And we may increase the size of the set S, since the only makes L larger. By adjointing all place corresponding to prime ideals in the decomposition of the representatives $\mathfrak{a}_1, \cdots, \mathfrak{a}_h$ of $\mathrm{Cl}(K)$. Then the $S-$integers ring $\mathcal{O}_{K,S} := \{x \in K \mid x \in \mathcal{O}_{K_v} \text{ for all } v \notin S\}$ is PID. We can also enlarge S to get $v(m) = 0$ for all $v \notin S$, and assume L is maximal under the conditions of the Proposition.

Kummer theory says the maximal abelian extension of exponent m of K is $K(\sqrt[m]{a} \mid a \in K)$ and then L is its subfield. Let $v \notin S$, then $K_v(\sqrt[m]{a})/K_v$ is unramified $\iff$ $\mathrm{ord}_v(a) = \mathrm{ord}_v(X^m) = m \cdot \mathrm{ord}_v(X) \equiv 0 \bmod m$. Because $\mu_m \subset K$, L can be written as $K(\sqrt[m]{a} \mid a \in T_S)$ with $T_S := \{a \in K^*/(K^*)^m \mid \mathrm{ord}_v(a) \equiv 0 \bmod m \text{ for all } v \notin S\}$.

Everything is reduced to show T_S is finite. Consider a natural map $\varphi : \mathcal{O}^*_{K,S} \to T_S$, for $\alpha \in K^*$ represents an element of T_S, the ideal $\alpha\mathcal{O}_{K,S}$ is a m^{th} of an ideal, say $\beta\mathcal{O}_{K,S}$ by using $\mathcal{O}_{K,S}$ is PID and note that the prime ideals of $\mathcal{O}_{K,S}$ correspond to the valuations $v \notin S$. Then $\alpha = u\beta^m$ for some unit $u \in \mathcal{O}^*_{K,S}$, and then φ is surjective. Moreover it induces a surjection (actually it is an isomorphism) :

$$\mathcal{O}^*_{K,S}/(\mathcal{O}^*_{K,S})^m \twoheadrightarrow T_S$$

Hence the Proposition follows the Dirichlet $S-$unit theorem, which says $\mathcal{O}^*_{K,S}$ is a finite generated abelian group. □

Let's summarize this proof, which is actually similar to the case of $\mathbb{Q}$. By analogy with the classic Kummer theory, we define the Kummer pairing $\kappa : E(K)/mE(K) \times G_{L/K} \to E[m]$ and show it is perfect. Because $E[m]$ is finite, the theorem can be obtained by showing L/K is finite. Hence we analyze the field extension L/K and prove show that such L/K is necessarily finite, which is done by using the finiteness of the class group and the Dirichlet $S-$unit theorem.

2.3 Selmer group and Tate-Shafarevich group

Suppose that we have elliptic curves $E/K, E'/K$ and a nonzero isogeny $\phi : E \to E'$ defined over K (we could take $E = E'$ and $\phi = [m]$). Let $E[\phi]$ denote the kernel of ϕ. Taking Galois cohomology yields the fundamental short exact sequence we need

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

Theorem X.3.6 of [1] says

$$\mathrm{WC}(E/K) \xrightarrow{\cong} H^1(G_{\overline{K}/K}, E)$$

where the Weil–Châtelet group $\mathrm{WC}(E/K)$ is the collection of equivalence classes of homogeneousspaces for E/K. Two homogeneous spaces C/K and C'/K for E/K are equivalent if there is an isomorphism $\theta : C \to C'$ defined over K that is compatible with the action of E on C and C'. The equivalence class containing E/K, acting on itself by translation, is called the trivial class. One can see that C/K is in the trivial class if and only if $C(K)$ is not the empty set.

Now we consider locally. For each $v \in M_K$ we fix an extension of v to $\overline{K}$, which serves to fix an embedding $\overline{K} \subset \overline{K}_v$ and a decomposition group $G_v \subset G_{\overline{K}/K}$. Now G_v acts on $E(\overline{K}_v)$ and $E'(\overline{K}_v)$, similarly, we see the exact sequence:

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

Hence

$$\begin{tikzcd}
0 \arrow[r] & E'(K)/\phi(E(K)) \arrow[r, "\delta"] \arrow[d] & H^1(G_{\overline{K}/K}, E[\phi]) \arrow[r] \arrow[d] & \mathrm{WC}(E/K)[\phi] \arrow[r] \arrow[d] & 0 \\
0 \arrow[r] & \prod_{v \in M_K} E'(K_v)/\phi(E(K_v)) \arrow[r, "\delta"] & \prod_{v \in M_K} H^1(G_v, E[\phi]) \arrow[r] & \prod_{v \in M_K} \mathrm{WC}(E/K_v)[\phi] \arrow[r] & 0
\end{tikzcd}$$

Our ultimate goal is to compute the image of $E'(K)/\phi(E(K))$ in the cohomology group $H^1(G_{\overline{K}/K}, E[\phi])$, i.e. the kernel of the map $H^1(G_{\overline{K}/K}, E[\phi]) \longrightarrow \mathrm{WC}(E/K)[\phi]$. And this last problem is the same as determining whether certain homogeneous spaces possess a K-rational point, which may be a very difficult question to answer. On the other hand, by the same reasoning, the determination of each local kernel

$$\ker\left(H^1(G_v, E[\phi]) \longrightarrow \mathrm{WC}(E/K_v)[\phi]\right)$$

is straightforward, since we have Hensel's lemma. This prompts the following definitions.

Definition 2.12 (Selmer group and Tate-Shafarevich group)**.**
Let $\phi : E/K \to E'/K$ be an isogeny. The ϕ-Selmer group of E/K is the subgroup of $H^1(G_{\overline{K}/K}, E[\phi])$ defined by

$$S^{(\phi)}(E/K) = \ker\left\{ H^1(G_{\overline{K}/K}, E[\phi]) \longrightarrow \prod_{v \in M_K} \mathrm{WC}(E/K_v) \right\}.$$

The Shafarevich–Tate group of E/K is the subgroup of $\mathrm{WC}(E/K)$ *defined by*

$$\text{Ш}(E/K) = \ker\left\{ \mathrm{WC}(E/K) \longrightarrow \prod_{v \in M_K} \mathrm{WC}(E/K_v) \right\}.$$

One see in order to determine whether an element of $\mathrm{WC}(E/K)$ becomes trivial in $\mathrm{WC}(E/K_v)$, we must check whether the associated homogeneous space, which is a curve defined over K, has any points defined over K_v. This last problem is clearly independent of our choice of extension of v to $\overline{K}$, since v itself determines the embedding of K into K_v. Therefore $S^{(\phi)}(E/K)$ and $\text{Ш}(E/K)$ depend only on E and K. Alternatively, one can directly work with cocycles that the cohomological definitions of $S^{(\phi)}$ and Ш do not depend on the extension of the $v \in M_K$ to $\overline{K}$. We leave this verification for the reader.

A good way to view $\text{Ш}(E/K)$ is as the group of homogeneous spaces for E/K that possess a K_v-rational point for every $v \in M_K$, which measures the extent to which Hasse's Principle fails on elliptic curves. We have the following theorem.

Theorem 2.13. *Let $\phi : E/K \to E'/K$ be an isogeny of elliptic curves defined over K, then*

(1) There is an exact sequence

$$0 \longrightarrow E'(K)/\phi(E(K)) \longrightarrow S^{(\phi)}(E/K) \longrightarrow \text{Ш}(E/K)[\phi] \longrightarrow 0.$$

(2) The Selmer group $S^{(\phi)}(E/K)$ is finite.

Proof. (1) is by definition and the proof of (2) is similar to Proposition 2.11. One can see Theorem X.4.2 of [1] for details. □

The famous Tate–Shafarevich conjecture states that the Tate–Shafarevich group is finite. Karl Rubin proved this for some elliptic curves of rank at most 1 with complex multiplication. It is known that the Tate–Shafarevich group is a torsion group, thus the conjecture is equivalent to stating that the group is finitely generated.

3 The Descent Argument

Theorem 3.1. *(Descent Theorem)*

Let A be an abelian group. Suppose that there exists a (height) function $h : A \longrightarrow \mathbb{R}$ with the following three properties:

(i) Let $Q \in A$. There is a constant $C_1(Q)$, depending on A and Q, s.t. $h(P+Q) \leq 2h(P) + C_1(Q)$ for all $P \in A$.

(ii) There are an integer $m \geq 2$ and a constant C_2, depending on A, s.t. $h(mP) \geq m^2 h(P) - C_2$ for all $P \in A$.

(iii) 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 (ii), the quotient group A/mA is finite. Then A is finitely generated.

Proof. Let $Q_1, \cdots, Q_r$ be a set of representatives for the cosets of A/mA in A. For any $P \in A$, we write $P = mP_1 + Q_{i_1}$ with $P_1 \in A$ and $i_1 \in \{1, \cdots, r\}$ and do same thing for P_1 to get P_2, Q_{i_2}, which gives a sequence $\{P_n\}$. And we have

$$h(P_n) \leq \frac{1}{m^2}(h(mP_n) + C_2) \leq \frac{1}{m^2}(2h(P_{n-1}) + \widetilde{C_1} + C_2)$$

where $\widetilde{C_1} = \max_{Q \in \{-Q_1, \cdots, -Q_r\}}\{C_1(Q)\}$. Since $m \geq 2$,

$$h(P_n) \leq \left(\frac{2}{m^2}\right)^n h(P) + \sum_{k=1}^{n} \frac{2^{k-1}}{m^{2k}}(\widetilde{C_1} + C_2) \leq \frac{1}{2^n} h(P) + (\widetilde{C_1} + C_2)$$

If n is sufficiently large, then $h(P_n) \leq 1 + \left(\widetilde{C} + C_2\right)$. Hence A is generated by $\{Q_1, \cdots, Q_r\} \cup \{P \mid h(P) \leq 1 + \widetilde{C_1} + C_2\}$, which is a finite set. □

With this general theorem, all that remains is to define the appropriate height function h. And we still divide into two parts: for $\mathbb{Q}$ and for general number field K.

3.1 Height on $\mathbb{Q}$

In this subsection, we assume $E/\mathbb{Q}$ be an elliptic curve with Weierstrass equation $y^2 = f(x) = x^3+ax+b, a, b \in \mathbb{Z}$.We define $H(\frac{m}{n} = \max\{|m|, |n|\})$ for $m, n \in \mathbb{Z}$, $H(P) = H(x(P))$ for $P \in E(\mathbb{Q}), H(O) = 1$ and $x(P)$ is the x-coordinate of P. Lastly, define $h(P) = \log H(P)$. Now we show h is we want and satisfies the three properties in the Descent Theorem with notation same as the theorem.

It is clear that h satisfies (iii).

Lemma 3.2. *h satisfies (i)*

Proof. Write $Q = (x_0, y_0) = (\frac{s}{d^2}, \frac{t}{d^3})$ as the form in Lemma 2.3 and for $P = (x, y) \notin \{-Q_0, O\}$, we have

$$H(P+Q) = H(x(P+Q)) = H\left(-(x+x_0+a) - \frac{(y-y_0)^2}{(x-x_0)^2}\right)$$

Replace y^2 by $x^3 + ax + b$, $H(P + Q)$ has form

$$H\left(\frac{Ay + Bx^2 + Cx + D}{Ex^2 + Fx + G}\right) = H\left(\frac{Atd + Bs^2 + Csd^2 + Dd^4}{Es^2 + Fsd^2 + Gd^4}\right)$$

where $A, \cdots, G$ are integers and depends on $Q, E/\mathbb{Q}$. Note that $d \le H(P)^{\frac{1}{2}}, s \le H(P), t \le CH(P)^{\frac{3}{2}}$ where C depends on $E/\mathbb{Q}$.

Hence $H(P + Q) \le \widetilde{C_1}(Q)H(P)^2$ and $h(P + Q) \le 2h(P) + C_1(Q)$ with $C_1(Q) = \log \widetilde{C_1}(Q)$. □

Lemma 3.3. *h satisfies (ii)*

Proof. It is fine to ignore these finite pionts satisfying $2P = O$. Take $P = (x, y) \in E/\mathbb{Q}$, and let $x = \frac{c}{d}$ as $\gcd(c, d) = 1$.

$$x(2P) = \frac{x^4 - 2ax^2 - 8bx + a^2}{4x^3 + 4ax + 4b} = \frac{c^4 - 2ac^2d^2 - 8bcd^3 + a^2d^4}{4c^3d + 4acd^3 + 4bd^4} =: \frac{F}{G}$$

A complicated calculations (See VIII 4.3 of [1]) shows there are f_1, g_1, f_2, g_2 depending on c, d and $\max\{f_1, f_2, g_1, g_2\} \le C(E/\mathbb{Q}) \max\{|c|^3, |d|^3\}$ s.t.

$$f_1F - g_1G = 4\Delta d^7, f_2F - g_2G = 4\Delta c^7$$

Hence $\gcd(F, G) \mid 4\Delta$ and then

$$H(2P) \ge \frac{\max\{|F|, |G|\}}{|4\Delta|} \ge \frac{\max\{|c|^4, |d|^4\}}{2C(E/\mathbb{Q})} \ge \frac{H(P)^4}{2C(E/\mathbb{Q})}$$

Take logarithm to get $h(2P) \ge 4h(P) - \log 2C(E/\mathbb{Q})$. □

Hence we can apply the Descent Theorem to get The Mordell-Weil Theorem for $\mathbb{Q}$.

3.2 Height on number field K

Let $P = [x_0 : \cdots : x_n] \in \mathbb{P}^n(K)$ and define

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

where $n_v := [K_v : \mathbb{Q}_v]$ and H_K is well defined by $\prod_{v \in M_K} |\lambda|_v^{n_v} = |\lambda|_K = 1$ for $\lambda \in K^*$.

If L/K is a finite extension, then $H_L(P) = H_K(P)^{[L:K]}$ by the fact that $\sum_{w|v} n_w = [L : K]n_v$. Hence for $P \in \mathbb{P}^n(\overline{\mathbb{Q}})$, $H(P) := H_K(P)^{\frac{1}{[K:\mathbb{Q}]}}$ is well defined. If $x \in K$, we also define $H_K(x) = H_K([x : 1])$.

Proposition 3.4. *Let C, d be constants. Then*

$$\{P \in \mathbb{P}^n(\overline{\mathbb{Q}}) : H(P) \leq C, [\mathbb{Q}(P) : \mathbb{Q}] \leq d\}$$

is finite and therefore $\{P \in \mathbb{P}^n(K) : H_K(P) \leq C\}$ is also finite.

Proof. Note that if $P = [x_0 : \cdots : x_n] \in \mathbb{P}^n(K)$,then $H_{\mathbb{Q}(P)}(P) \geq \max_{0 \leq i \leq n} H_{\mathbb{Q}(p)}(x_i)$. Thus it is enough to show $S := \{x \in \overline{\mathbb{Q}} \mid H(x) \leq C, [\mathbb{Q}(x) : \mathbb{Q}] \leq d\}$ is finite.

Let $[\mathbb{Q}(x) : \mathbb{Q}] = e$ and $x_1 = x, x_2, \cdots, x_n$ are conjugates of x over $\mathbb{Q}$. Denote the minimal polynomial of x over $\mathbb{Q}$ as $f(X) = \prod_{i=1}^{e}(X - x_i) = X^e + a_1 X^{e-1} + \cdots + a_e$. Consider $P \in \mathbb{P}^n(\overline{\mathbb{Q}})$ and $\sigma \in G_{\overline{\mathbb{Q}}/\mathbb{Q}}$. Then for field $K/\mathbb{Q}$ which may not be Galois, but always σ induces that $K \xrightarrow{\sim} K^\sigma, M_K \xrightarrow{\sim} M_{K^\sigma}, K_v \xrightarrow{\sim} K^\sigma_{v^\sigma}$, which shows $H_{K^\sigma}(P^\sigma) = H_K(P)$, and then $H(P) = H(P^\sigma)$. In our context, we have $H(x_j) = H(x)$.

We want to show S is finite, which can be done if we have consistent upper control for $H([1, a_1, \cdots, a_e])$. Actually,

$$H([1, a_1, \cdots, a_e]) \underbrace{\leq}_{(*)} 2^{e-1} \prod_{j=1}^{e} H(x_j) = 2^{e-1} H(x)^{e-1} \leq (2C)^d$$

Now we start to prove $(*)$. It suffices to prove that for every $v \in M_K$, we have $\max_{1 \leq i \leq e}\{|a_i|_v\} \leq 2^{e-1} \prod_{j=1}^{e} \max\{|x_j|, 1\}$. For $e = 1$, it is clear. Assume we know the result the degree less than $e - 1$, choose $|x_k|_v = \max_j\{|x_j|_v\}$ and denote $g(X) = f(X)/(X - x_k) = X^{e-1} + b_1 X^{e-2} + \cdots + b_{e-1}$. Thus $a_i = b_i - x_k b_{i-1}$(Set $b_{-1} = b_e = 0$). Hence

$$\max_{1 \leq i \leq e}\{|a_i|_v\} \leq 2 \max_{1 \leq i \leq e}\{|b_i|_v, |x_k b_{i-1}|_v\} \leq 2 \max_{1 \leq i \leq e}\{|b_i|_v\} \max\{|x_k|, 1\} \leq 2^{e-1} \prod_{j=1}^{e} \max\{|x_j|, 1\}$$

where the last step is by induction. The proof is done. □

Remark 3.5 (Schanuel' Theorem)**.** *Let K be a number field of degree d, with r_1 real embeddings and r_2 pairs of non-real embeddings. Let Δ be the absolute value of discriminant of $K/\mathbb{Q}$, w be the number of roots of unity in K, $r = r_1 + r_2 - 1$ be the rank of $\mathcal{O}_K^\times$, and R be the regulator. Then*

$$\#\{x \in \mathbb{P}_{N-1} \mid H(x) \leq X\} =: N_X = \frac{h}{\zeta_K(n)} c_n X^{nd} + \begin{cases} O(X \log X), & \textit{if } d = 1, n = 2, \\ O(X^{nd-1}), & \textit{otherwise} \end{cases}$$

where $c_n = \left(\frac{2^{r_1}(2\pi)^{r_2}}{\sqrt{\Delta}}\right)^n \frac{R}{w} n^r$.

We define $h : \mathbb{P}^n(K) \to \mathbb{R}, P \mapsto \log H(P)$. Any nonconstant function $f \in K(E)$ induces a (surjective) morphism

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

Now we define the true "height" map we want for E/K relative to f as

$$h_f : E(\overline{K}) \to \mathbb{R} \qquad P \mapsto h(f(P))$$

By Prop 3.4, h_f satisfies (iii) in the Descent Theorem. We now show h_f satisfies (i) and (ii) by a general Proposition.

Proposition 3.6. *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 $O(1)$ depends on E/K and f.

We first reduce f to x, then check the relationship between coordinates by straight calculations (See VIII 6.2 of [1] for a detailed proof).

Corollary 3.7. *h_f satisfies (i),(ii) in the Descent Theorem.*

Proof. The above Proposition yields

$$h_f(P+Q) \le 2h_f(P) + 2h_f(Q) + O(1)$$
$$h_f(2P) = h_f(P+P) = 4h_f(P) + O(1)$$

□

Hence we can apply the Descent Theorem to get The Mordell-Weil Theorem for K.

4 About $E(\mathbb{Q})$

4.1 Finite order points and integer points

We see some results in Remark 2.7: *The structure of m-torsion points*, and in this subsection, we introduce a famous theorem which was first proven (independently) by Elisabeth Lutz and Trygve Nagell in the 1930s, and tell us how to find all of the rational points of finite order.

Theorem 4.1 (The Nagell-Lutz Theorem).
For the elliptic curve $y^2 = x^3 + ax^2 + bx + c$, with a, b, and c integers and having (nonzero) discriminant function $D = -4a^3c + a^2b^2 + 18abc - 4b^3 - 27c^2$, let $P = (X, Y)$ be a rational point of finite order greater than 1. Then
(1) X and Y are integers;
(2) Either Y divides D or $Y = 0$ and P has order 2.

Proof. The second part is easy to prove, if one note P and $2P$ have integer coordinates and $D = r(x)f(x) + s(x)f'(x)$ for some $r(x), s(x) \in \mathbb{Z}[x]$. The first part is proved by showing the denominators of X and Y are not divisible by any prime numbers. One can see the Chapter 2 of [3] for a detailed proof. □

Moerover, we can say more about the Integer points on cubic curves. A interesting number is 1729, which is the smallest number expressible as a sum of two cubes in two different ways:

$$1729 = 9^3 + 10^3 = 1^3 + 12^3$$

And we have a Theorem of Thue says

Theorem 4.2 (Thue's Theorem)**.**

Let a, b, c be non-zero integers, Then the equation $ax^3 + by^3 = c$ has only finitely many solutions in integers x, y.

The core part of the proof of Thue's theorem lies in the theory of Diophantine Approximation, which concentrates on rational approximations to irrational quantities. We briefly introduce some important results of this theory.

- (**Dirichlet's Theorem**) For any irrational number α, there are infinitely many rational numbers $\frac{p}{q}$ such that $|\alpha - \frac{p}{q}| < \frac{1}{q^2}$.
- (**Roth's Theorem, 1955**) For every $\varepsilon > 0$, every number field K of degree d has approximation exponent $\tau(d) = 2 + \varepsilon$. More specifically, Let $\alpha \in \overline{K}$, let $d = [K(\alpha) : K]$, and let $v \in M_K$ be an absolute value on K that has been extended to $K(\alpha)$ in some fashion. Then for any constant C there exist only finitely many $x \in K$ satisfying the inequality
$$|x - \alpha|_v < CH_K(x)^{-\tau(d)}.$$
- (**Siegel's Theorem**) Let C be a non-singular cubic curve given by an equation $F(x, y) = 0$ with integer coefficients. Then C has only finitely many points with integer coordinates.

4.2 The free part and the torsion part

We know $E(\mathbb{Q}) \cong \mathbb{Z}^r \oplus$ A finite Abel group G, and we have some famous results and conjectures around this:

- (**Mazur torsion theorem**) G is isomorphic to one of the following groups: $\mathbb{Z}/n\mathbb{Z}$ for $1 \leq n \leq 10$ or $n = 12$, and $\mathbb{Z}/2\mathbb{Z} \oplus \mathbb{Z}/n\mathbb{Z}$ for $n = 2, 4, 6, 8$.
- (**Faltings' Theorem**) A nonsingular algebraic curve of genus greater than 1 over the field $\mathbb{Q}$ has only finitely many rational points.
- (**Ogg's Conjecture**) The order of the torsion group of an abelian variety over a number field is bounded in terms of the dimension of the variety and the degree of the number field.
- (**Conjecture**) There exist elliptic curves of arbitrarily large rank over a fixed algebraic number field.
- (**Birch and Swinnerton-Dyer Conjecture, a Millennium Prize Problem**) The rank of $E(\mathbb{Q})$ is equal to the order of vanishing of the L-function $L(E, s)$ at $s = 1$.

Referrences

[1] J. H. Silverman, *The Arithmetic of Elliptic Curves* (2nd ed.). Graduate Texts in Mathematics, volume 106. Springer, 2009.

[2] K. Ireland & M. Rosen, (1990). *A classical introduction to modern number theory* (2nd ed.). Springer-Verlag.

[3] J. H. Silverman, & J. Tate, (2015). *Rational points on elliptic curves* (2nd ed.). Springer.

[4] J. H. Silverman, *Advanced Topics in the Arithmetic of Elliptic Curves.* Graduate Texts in Mathematics, volume 151. Springer, 1994.

[5] 加藤和也, 黑川信重, 斋藤毅著; 胥鸣伟, 印林生译. 数论 I,Fermat 的梦想和类域论 [M]. 北京: 高等教育出版社, 2009.6.