The Mordell-Weil theorem

Caleb Ji

These are my notes on the Mordell-Weil theorem for elliptic curves. Their main purpose is to remind me of how the arguments go rather than to provide full details of a proof. They are drawn from Booher [1], Milne [2], and Silverman [4], where one may find complete details. If I ever get to it, I also plan to add an outline of the proof for abelian varieties, following Serre [3].

Contents

1 Some background material

1.1 Kummer theory

Given a field K containing n distinct nth roots of unity, Kummer theory describes the abelian extensions of K with exponent dividing n. Thus it may be regarded as a part of class field theory. The importance of Kummer theory to the Mordell-Weil theorem is the main statement that

$$H_c^1(G_K, \mu_n) = \mathrm{Hom}_c(G_K, \mu_n) \cong K^*/(K^*)^n. \tag{1}$$

This appears when we consider the Galois action on $E[n]$, the n-torsion points of an elliptic curve, and use it as part of a long exact sequence of cohomology to show that the group $E(K)/nE(K)$ is finite (precisely the statement of weak Mordell-Weil). The goal of this section is to develop Kummer theory and prove 1, placing it in the context of class field theory.

Remark. The subscript H_c^i (or Hom_c) denotes continuous cohomology. Like many authors, we will drop this subscript for convenience. When working with a profinite group like G_K, one would generally want continuous maps anyways. For an exposition of continuous cohomology, we refer the reader to [5].

Cohomological proof of 1

Fix a field K and take any n such that $\mathrm{char}(K) \nmid n$. Let μ_n be the group of nth roots of unity contained in K, which must be distinct by the assumption on the characteristic. Let $G_K = \mathrm{Gal}(\overline{K}/K)$. Then we have an exact sequence of discrete G_K-modules

$$1 \to \mu_n \to K^* \xrightarrow{\bullet^n} K^* \to 1.$$

The associated long exact sequence of (continuous!) cohomology reads

$$1 \to \mu_n \to K^* \xrightarrow{\bullet^n} K^* \to H^1(G_K, \mu_n) \xrightarrow{i^1} H^1(G_K, K^*) \xrightarrow{\bullet^n} \cdots$$

From this we deduce that $K^*/(K^*)^n \cong \ker i^1$. Moreover, because G_K acts trivially on these modules, the first cohomology groups are just continuous homomorphisms.

We recall that by Hilbert's Theorem 90, $H^1(\mathrm{Gal}(L/K), K^*) = 0$ when L/K is a finite Galois extension. In our scenario, we have

$$G_K = \varprojlim_{[L:K] \text{ finite}} L/K \Rightarrow H^n(G_K, K^*) = \varinjlim_{L} H^n(L/K, K^*) = 0,$$

where the passage to the direct limit takes some technical checking.

With this, we obtain

$$H^1(G_K, \mu_n) = \mathrm{Hom}(G_K, \mu_n) \cong K^*/(K^*)^n,$$

as desired.

1.2 The Selmer and Tate-Shafarevich groups

Let K be a number field and let K_v be its completion at a finite place v. For convenience, we will denote $E(\overline{F})$ by simply E where appropriate. Then the natural homomorphism $\mathrm{Gal}(G_{K_v}) \to \mathrm{Gal}(G_K)$ induces a map $H^1(G_K, E) \to H^1(G_{K_v}, E)$, and similarly when E is replaced by $E[n]$. We can make sense of these maps using group cohomology.

Let E be an elliptic curve defined over K. It is not hard to show that $\cdot nE(\overline{K}) \to E(\overline{K})$ is surjective. We therefore have an exact sequence

$$0 \to E[n] \to E(\overline{K}) \xrightarrow{\cdot n} E(\overline{K}) \to 0.$$

First, we assume that $E[n] \subset K$. This leads to the long exact sequence

$$0 \to E[n] \to E(K) \xrightarrow{\cdot n} E(K) \to H^1(G_K, E[n]) \to H^1(G_K, E(\overline{K})) \xrightarrow{\cdot n} \cdots.$$

We extract from this the short exact sequence

$$0 \to E(K)/nE(K) \xrightarrow{f} H^1(G_K, E[n]) \to H^1(G_K, E(\overline{K}))[n] \to 0.$$

Replacing K with K_v and taking the maps between them, we obtain the following commutative diagram.

$$\begin{array}{ccccccccc}
0 & \longrightarrow & E(K)/nE(K) & \longrightarrow & H^1(G_K, E[n]) & \longrightarrow & H^1(G_{K_v}, E)[n] & \longrightarrow & 0 \\
 & & \downarrow & & \downarrow & & \downarrow & & \\
0 & \longrightarrow & E(K_v)/nE(K_v) & \longrightarrow & H^1(G_{K_v}, E[n]) & \longrightarrow & H^1(G_{K_v}, E)[n] & \longrightarrow & 0
\end{array}$$

Now we can define the Selmer group and the Tate-Shafarevich group.

Definition 1.1 (Selmer group)**.** *The Selmer group, denoted $S^{(n)}(E/K)$ is defined by*

$$S^{(n)}(E/K) = \ker\left(H^1(G_K, E[n]) \to \prod_v H^1(G_{K_v}, E)\right).$$

Definition 1.2 (Tate-Shafarevich group)**.** *The Tate-Shafarevich group, denoted* $Ш(E/K)$*, is defined by*

$$Ш(E/K) = \ker\left(H^1(G_K, E) \to \prod_v H^1(G_{K_v}, E)\right).$$

Note that the Selmer group can be characterizd as the subgroup of $H^1(G_K, E[n])$ that comes from an element of $E(K_v)$ for every v. Thus it should not be too surprising that by diagram chasing, we obtain the following exact sequence.

$$0 \to E(K) \to nE(K) \to S^{(n)}(E/K) \to Ш(E/K)[n] \to 0. \tag{2}$$

2 The Mordell-Weil theorem for elliptic curves

The proof of the Mordell-Weil theorem, both in the case of elliptic curves and in the general case, proceeds in two steps. First, we prove the weak Mordell-Weil theorem, which states that $E(K)/nE(K)$ is finite for any positive integer n. Then we use the theory of heights to finish.

2.1 Weak Mordell-Weil

In this section, we sketch the proof of the following theorem. Full details may be found in Milne [2]; Silverman [4] gives a different proof. We will prove a generalization of this theorem in this following section.

Theorem 2.1. *[weak Mordell-Weil] Let E be an elliptic curve defined over a number field K. For any positive integer n, the group $E(K)/nE(K)$ is finite.*

We would like to use the exact sequence 2, which states that $E(K)/nE(K)$ embeds into the Selmer group $S^{(n)}(E/K)$. However, recall that this was all based on the assumption that K contains $E[n]$. However, it is not hard to see that we can make this assumption, in which case one can show that we also have $\mu_n \in K$. Indeed, let $L = E(K)[E[n]]$. By considering the long exact sequence of $G_{L/K}$ associated to the short exact sequence

$$0 \to E[n] \to E(L) \to mE(L) \to 0,$$

we readily obtain the following proposition.

Proposition 2.2. *If $E(L)/nE(L)$ is finite, then $E(K)/nE(K)$ is also finite.*

To prove Theorem 2.1, it now suffices to show that the Selmer group $S^{(n)}(E/K)$ is finite.

Sketch of proof of finiteness of $S^{(n)}(E/K)$. We let S be the finite set of places dividing n and where E has bad reduction. (We are only working with finite places.) The Selmer group consists of the elements ξ of $H^1(G_K, E[n])$ which go to 0 in $H^1(G_{K_v}, E)$. For $v \notin S$, this is equivalent to ξ being unramified at v. Viewing ξ as an element of $K^*/(K^*)^n$, one shows that this is equivalent to $n|\nu_v(\alpha)$ for v outside S. Call the subgroup satisfying this H.

We obtain a map $H \to (\mathbb{Z}/n\mathbb{Z})^{|S|}$ defined by $\alpha \mapsto \{\nu_{\mathfrak{p}_i}(\alpha) \pmod n\}$ as $\mathfrak{p}_i$ ranges over the primes in S. We must show that the kernel H_0 is finite. Since the kernel consists of mth powers, we can define another map $H_0 \to \mathrm{Cl}(\mathcal{O}_K)$ sending α to $\mathfrak{a}$, where $\mathfrak{a}^n = (\alpha)$. Since $\mathrm{Cl}(\mathcal{O}_K)$ is finite, it suffices to show that the kernel of this map is finite. This kernel consists of a unit of $\mathcal{O}_K$ multiplied by an nth power, and we are modding out the nth powers too. Thus the kernel may be identified with a subset of $\mathcal{O}_K^*/(\mathcal{O}_K^*)^n$, which is finite by Dirichlet's unit theorem. □

2.2 Heights

Here we sketch the argument using heights from which the Mordell-Weil theorem is deduced from the weak version. The idea is to use an infinite descent argument based on the 'height' of points.

Heights on $\mathbb{P}^n_K$ and elliptic curves

Definition 2.3. *Let K be a number field and take a point $x = [x_0 : x_1 : \cdots : x_n] \in \mathbb{P}^n(K)$. Then we define the height $H_K(x)$ to be*

$$H_K(x) \coloneqq \prod_v \max_i(|x_i|_v).$$

Here, we are taking the product over all places v, and the valuations are normalized so that the product formula holds. Note that in the case of $K = \mathbb{Q}$, this recovers the definition of taking the maximum absolute value of $x_0, \ldots, x_n$ provided the x_i are integers with no nontrivial common divisor. We may also define the absolute height for $x \in \mathbb{P}^n(\overline{\mathbb{Q}})$ by

$$H(P) \coloneqq H_K(P)^{1/[K:\mathbb{Q}]}.$$

Finally, it will be useful to define the logarithmic height $h(P) = \log H(P)$. If f is a function to $\mathbb{P}^1$, then we define

$$h_f(P) \coloneqq \log H(f(P)).$$

The height function $h_f(P)$ is especially useful when f is an even function applied to an elliptic curve; e.g., taking the x-coordinate. In this case, one proves that

$$h_f(P+Q) + h_f(P-Q) = 2h_f(P) + 2h_f(Q) + O(1), \tag{3}$$

where $O(1)$ depends only on the elliptic curve E and the function f. In other words, by the parallelogram law, h_f is close to giving a quadratic form on $E(K)$. In fact, consider defining

$$\hat{h}(P) = \frac{1}{\deg(f)} \lim_{n\to\infty} 4^{-n} h_f(2^n P).$$

This is known as the canonical, or Néron-Tate height. It gives a canonical quadratic form

$$\langle P, Q\rangle \coloneqq \hat{h}(P+Q) - \hat{h}(P) - \hat{h}(Q).$$

As important as the Néron-Tate height is, it is not strictly necessary for proving the Mordell-Weil theorem, which can be done by using $h_f(P)$.

The descent argument

The Mordell-Weil theorem will directly follow from the weak version and a descent argument if we show the height function $h_f(P)$ satisfies the following properties.

1. For some constant c_1 depending on E, f, and Q, we have

   $$h_f(P+Q) \leq 2h_f(P) + c_1$$

   for $P, Q \in E(K)$.

2. For some constant c_1 depending on E and f, we have

   $$h_f(2P) \geq 4h_f(P) - c_2$$

   for $P \in E(K)$.

3. For all constants c_3, there are finitely many $P \in E(K)$ with $h_f(P) \leq c_3$.

Indeed, assume these results and use the weak Mordell-Weil theorem to pick finite coset representatives $\{P_1, \ldots, P_n\}$ for $E(K)/2E(K)$. Then for any $P \in E(K)$, properties 1 and 2 allow us to write P as a linear combination of $P_1, \ldots, P_n$ and another point with height bounded by a constant independent of P. Then property 3 implies we can take finitely many points to cover all points with such height, which together with $P_1, \ldots, P_n$ must generate $E(K)$.

In order to prove properties 1 and 2, one uses equation 3. The fact that h_f acts similar to a quadratic form with respect to elliptic curve addition, combined with some computation, yields these properties.

Property 3 follows from Northcott's finiteness theorem, which states that there are finitely many points of $\mathbb{P}^n(\overline{\mathbb{Q}})$ with bounded height and degree. To prove this, one first reduces to the case of $n = 1$ via a simple bound. Then from the fact that the height of a point is the same as that of its Galois conjugates, one shows a relation between the height of a point x and the coefficients of its minimal polynomial. This shows that there are only finitely many polynomials that can give rise to points with the prescribed heights and degrees, as desired.

References

[1] Jeremy Booher, *The Mordell-Weil Theorem for Elliptic Curves*. Expository article. https://www.math.canterbury.ac.nz/~j.booher/expos/mordellweil.pdf

[2] James Milne. *Elliptic Curves*. BookSurge Publishers, 2006.

[3] Jean-Pierre Serre. *Lectures on the Mordell-Weil theorem*. Friedr. Vieweg & Sohn Verlagsgesellschaft mbH, Braunschweig, 1989.

[4] Joseph Silverman. *The Arithmetic of Elliptic Curves*. Publishedby Springer Science+Business Media, LLC. 2nd edition, 2009.

[5] Andrew Sutherland. Number Theory II. Course notes. https://math.mit.edu/classes/18.786/LectureNotes30.pdf