THE MORDELL-WEIL THEOREM FOR ELLIPTIC CURVES

JEREMY BOOHER

The Mordell-Weil theorem is a fundamental result in the arithmetic of elliptic curves defined over a number field K, describing the structure of their K-valued points. It is proven, for example, in Chapter 8 of Silverman [2]. The proof given here uses the same approach, but aims to be more approachable (the proof in Silverman is spread over 6 sections) and illustrate how in special cases these techniques can be used to effectively calculate the Mordell-Weil group. This is inspired by the presentation in the Cambridge Part III course on elliptic curves taught by Tim Dokchitser.

Theorem 1 (Mordell-Weil)**.** *Let E be an elliptic curve defined over a number field K. The group $E(K)$ is a finitely generated Abelian group.*

We prove the theorem using four main ideas.

(1) Reducing the elliptic curve over good primes to show that the torsion of $E(K)$ is finite.
(2) The weak Mordell-Weil theorem, which states that $E(K)/mE(K)$ is finite.
(3) The theory of heights, which allows us to talk meaningfully about the size of the points on the curve.
(4) A descent argument using representatives for $E(K)/mE(K)$ and the theory of heights to conclude that $E(K)$ is finitely generated.

The first is simple, as we know that the torsion prime to p (for p a good prime) injects into the finite group of points on the elliptic curve over the residue field K_p. The neatest proof of the weak Mordell-Weil theorem uses group cohomology and will be proven first in Section 1. The theory of heights requires some work to develop: the proofs are easiest to understand over $\mathbb{Q}$, so we prove them over $\mathbb{Q}$ and relegate general proofs to an appendix. After developing this theory, the descent argument will be simple to carry out.

1. The Weak Mordell-Weil Theorem

In this section, we will use group cohomology to prove the weak Mordell-Weil theorem.

Theorem 2. *Let E be an elliptic curve defined over a number field K, and $m \geq 2$. Then $E(K)/mE(K)$ is finite.*

The necessary group cohomology is developed from scratch in Section A for those unfamiliar with it.

To prove this, we will first reduce to the case that the number field contains the m-torsion of the elliptic curve. We then use group cohomology (for the absolute Galois group of K) and Kummer theory to construct the Kummer map embedding $E(K)/mE(K)$ into $K^\times/(K^\times)^m \times K^\times/(K^\times)^m$. Finally, we map K into a p-adic field where we have a better understanding of mth powers in order to show the image is finite.

Let E/K be an elliptic curve, G_K the absolute Galois group of K, and $m \geq 2$. Let $L = K(E[m])$ be the field obtained by adjoining the coordinates of the m torsion to K. Note that the mth roots of unity μ_m lie in L as the Weil pairing must be invariant under G_L. It suffices to prove the weak Mordell-Weil theorem for E/L because of the following result.

Proposition 3. *With the notation above, let $E(L)/mE(L)$ be finite. Then $E(K)/mE(K)$ is finite.*

Date: September 4, 2012, updated November 3, 2023.

Proof. Consider the short exact sequence

$$0 \to E[m] \to E(L) \to mE(L) \to 0.$$

Viewing them as $G_{L/K} = \mathrm{Gal}(L/K)$ modules, the long exact sequence in cohomology reads

$$0 \to E(K)[m] \to E(K) \to mE(L) \cap E(K) \to H^1(G_{L/K}, E[m]) \to H^1(G_{L/K}, E(L)).$$

In particular, there is an injection

$$mE(L) \cap E(K)/mE(K) \hookrightarrow H^1(G_{L/K}, E[m]).$$

However, $G_{L/K}$ is finite and so is $E[m]$, thus there are a finite number of crossed homomorphisms in $H^1(G_{L/K}, E[m])$. Thus $(mE(L) \cap E(K))/mE(K)$ is finite. It is the kernel of the map

$$E(K)/mE(K) \to E(L)/mE(L).$$

By assumption $E(L)/mE(L)$ is finite, so $E(K)/mE(K)$ is as well. □

We may now assume that K contains the m torsion of E, in particular the mth roots of unity, and look at a generalization of the above exact sequence to the algebraic closure of K. Start with

$$0 \to E[m] \to E(\overline{K}) \overset{\cdot m}{\to} E(\overline{K}) \to 0$$

and take G_K cohomology. This gives

$$0 \to E(K)[m] \to E(K) \overset{\cdot m}{\to} E(K) \to H^1(G_K, E[m]) \to H^1(G_K, E(\overline{K})) \overset{\cdot m}{\to} \ldots$$

Extracting a short exact sequence in the middle, we get the Kummer sequence

$$0 \to \frac{E(K)}{mE(K)} \to H^1(G_K, E[m]) \to H^1(G_K, E(\overline{K}))[m] \to 0. \tag{1}$$

Little is known about $H^1(G_K, E(\overline{K}))[m]$, but little is required for the proof of the weak Mordell-Weil theorem.

Now as K contains $E[m]$, G_K acts trivially on it so $H^1(G_K, E[m]) = \mathrm{Hom}_{cont}(G_K, E[m])$. Additionally, $E[m] \simeq (\mathbb{Z}/m\mathbb{Z})\,\tau_1 \times (\mathbb{Z}/m\mathbb{Z})\,\tau_2$where τ_1 and τ_2 are a basis for the m-torsion. The maps $\alpha_i : E[m] \to \mu_m$ defined by $\alpha_i(\tau) = e_m(t, \tau_i)$ define an isomorphism $E[m] \to \mu_m \times \mu_m$. Thus

$$\frac{E(K)}{mE(K)} \hookrightarrow H^1(G_K, E[m]) \simeq H^1(G_K, \mu_m) \times H^1(G_K, \mu_m).$$

Kummer's theorem (which is where it is essential the crossed homomorphisms be continuous) says that $H^1(G_K, \mu_m) \simeq K^\times/(K^\times)^m$. Thus it suffices to show that the image of

$$\kappa_K : \frac{E(K)}{mE(K)} \hookrightarrow K^\times/(K^\times)^m \times K^\times/(K^\times)^m$$

is finite. The map $\kappa_K = (\kappa_1, \kappa_2)$ is known as the Kummer map.

Example 4. For computational purposes, we need to understand this map when doing a two descent. The Weil pairing has an explicit description in this case, so the Kummer map takes on a simple form. If the Weierstrauss form over $\overline{K}$ is $y^2 = (x - e_1)(x - e_2)(x - e_3)$, then it turns out the Kummer map sends $P = (x, y) \in E(K)/2E(K)$ to $(x - e_1) \times (x - e_2) \in (K^\times/(K^\times)^2) \times (K^\times/(K^\times)^2)$. Extra care must be taken if P is one of the two-torsion points. If $\kappa_i(P) = x - e_i$ for $i = 1, 2, 3$ and P is not a two-torsion point, note that $\prod_i \kappa_i(P)$ is always a square. If P is a two-torsion point, $x = e_i$ for some i and $\kappa_i(P)$ is undefined as $0 \notin K^\times/(K^\times)^2$. Instead, define $\kappa_i(P)$ so that $\prod_i \kappa_i(P)$ is the identity in $K^\times/(K^\times)^2$.

Now let p be a prime of K not dividing m such that E has good reduction at p. Consider the maximal unramified extension of K_p, denoted by K_p^{ur}.

Lemma 5. *With the notation above, $E(K_p^{ur})/mE(K_p^{ur}) = 0$.*

Proof. Look at the long exact sequence

$$0 \to E_1(K_p^{\mathrm{ur}}) \to E(K_p^{\mathrm{ur}}) \to \tilde{E}_{ns}(k) \to 0$$

where k is the (algebraically closed) residue field of K_p^{ur}. Mapping it to the same sequence via the multiplication by m map and taking the kernel-cokernel exact sequence yields

$$\frac{E_1(K_p^{\mathrm{ur}})}{mE_1(K_p^{\mathrm{ur}})} \to \frac{E(K_p^{\mathrm{ur}})}{mE(K_p^{\mathrm{ur}})} \to \frac{\tilde{E}_{ns}(k)}{m\tilde{E}_{ns}(k)} \to 0.$$

By the theory of formal groups, multiplication by m is an isomorphism on $E_1(K_p^{\mathrm{ur}})$ hence the left term is 0. Since k is algebraically closed, $\frac{\tilde{E}_{ns}(k)}{m\tilde{E}_{ns}(k)} = 0$ as well. The desired conclusion follows. □

Note the following diagram, coming from the inclusion of K into K_p^{ur}, commutes:

$$\begin{tikzcd}
E(K)/mE(K) \arrow[r, "\kappa_K"] \arrow[d] & K^\times/(K^\times)^m \times K^\times/(K^\times)^m \arrow[d] \\
E(K_p^{\mathrm{ur}})/mE(K_p^{\mathrm{ur}}) \arrow[r, "\kappa_{K_p^{\mathrm{ur}}}"] & (K_p^{\mathrm{ur}})^\times/(K_p^{\mathrm{ur}\times})^m \times (K_p^{\mathrm{ur}})^\times/(K_p^{\mathrm{ur}\times})^m
\end{tikzcd}$$

As the lower left group is trivial, the image of $E(K)/mE(K)$ in the lower right is trivial. Therefore the image of the Kummer map satisfies $v_p(\kappa_i(P)) \equiv 0 \mod m$. This holds for any prime p not dividing m or the discriminant of E.

Define $H := \{\alpha \in K^\times/(K^\times)^m : v_p(\alpha) \equiv 0 \mod m \text{ for } p \nmid \Delta(E)m\}$. If we know H is finite, the weak Mordell-Weil theorem will follow as $E(K)/mE(K)$ injects into H.

Proposition 6. *With the notation above, H is a finite group.*

Proof. Let $\mathfrak{p}_1, \ldots, \mathfrak{p}_n$ be the primes of bad reduction and the primes dividing m. Map $H \to (\mathbb{Z}/m\mathbb{Z})^n$ via the map $\alpha \to \{v_{\mathfrak{p}_i}(\alpha) \mod n\}_i$. Let H_0 be the kernel. We will show H_0 is finite, which will imply H is finite.

Now, if α is in the kernel then it is a perfect mth power. Factoring the ideal (α) in $\mathcal{O}_K$, we know that it is of the form $\mathfrak{a}^m$ for some ideal $\mathfrak{a}$. Define a map

$$H_0 \to \mathrm{Cl}(\mathcal{O}_K) \quad \text{sending} \quad \alpha \to \mathfrak{a}.$$

This is well defined as $(K^\times)^m$ is sent to a principal ideal. An big theorem in basic algebraic number theory says that the class group is finite. Therefore it suffices to show the kernel of this map is finite. α is mapped to a principal ideal if and only it is a unit times an mth power. The other important theorem in algebraic number theory, Dirichlet's unit theorem, says the group of units is a finitely generated Abelian group, so $\mathcal{O}_K^\times/(\mathcal{O}_K^\times)^m$ is finite. □

This completes the proof of the weak Mordell-Weil theorem.

2. The Theory of Heights

The next ingredient is a way to meaningfully talk about the size of a point on an elliptic curve. We first define the height of an algebraic number as a measure of the complexity of the number: intuitively $\frac{2}{3}$ is much simpler "arithmetically" than $\frac{200}{301}$, and we will see this reflected in the fact that heights are $\log 3$ and $\log 301$ respectively. Then we extend this to define heights of points on elliptic curves, and establish the parallelogram law and determine how heights interact with the multiplication by m map.

2.1. **Heights over Number Fields.** Let K be a number field, and Σ denote the set of normalized absolute values on K. Define

$$h_K(\alpha) := \sum_{v\in\Sigma} \max(0, \log(|\alpha|_v)).$$

Example 7. Let $K = \mathbb{Q}$. Then let $\alpha = \frac{a}{b}$ for a and b relatively prime, and note $\Sigma = \{p : p \text{ prime}\} \cup \{\infty\}$. Suppose $v_p(a) \geq 0$. Then $\max(1, |\alpha|_p) = \max(1, p^{-v_p(a)}) = 1$. If $v_p(b) \geq 0$, then $\max(1, |\alpha|_p) = \max(1, p^{v_p(b)}) = p^{v_p(b)}$. Additionally, $\max(1, |\alpha|_\infty) = |\frac{a}{b}|$ if $|\frac{a}{b}| > 1$, 1 otherwise. Therefore

$$\prod_v \max(1, |\alpha|_v) = \prod_{p:v_p(b)>0} p^{v_p(b)} \cdot \max(1, \left|\frac{a}{b}\right|) = |b| \max(1, \left|\frac{a}{b}\right|) = \max(|a|, |b|).$$

Therefore we have that

$$h_Q(\frac{a}{b}) = \log \max(|a|, |b|)$$

which certainly measures the size of a rational number.

Here are three essential properties of heights.

Proposition 8. *Let K be a number field.*

1. *For any constant C, the set $\{\alpha \mid h_K(\alpha) < C\}$ is finite.*
2. *Given a map $f : \mathbb{P}^1 \to \mathbb{P}^1$ defined over the field K of degree d, $h_K(f(\alpha)) = \deg(f)h_K(\alpha) + O(1)$, where the constant depends on f but not α.*
3. *We have $h_K(\alpha) = 0$ if and only if $\alpha = 0$ or α is a root of unity in K.*

The full proof of this proposition is somewhat technical and no more enlightening than the proof over $\mathbb{Q}$. It may be found in Section B. Instead, we will prove the special case when $K = \mathbb{Q}$ in which case the height function $h_q(\frac{a}{b}) = \max(|a|, |b|)$ is particularly nice to work with.

The first assertion, that there are finitely many rational numbers $\frac{a}{b}$ with $\max(|a|, |b|) < C$, is obvious.

Now let f be a rational function (with coefficients in $\mathbb{Q}$) of degree d, defining a degree map $\mathbb{P}^1 \to \mathbb{P}^1$. Write

$$f(T) = \frac{g(T)}{h(T)} = \frac{a_dT^d + a_{d-1}T^{d-1} + \ldots + a_0}{b_dT^d + b_{d-1}T^{d-1} + \ldots + b_0}.$$

Let $\alpha = \frac{p}{q}$ with p and q relatively prime. Then

$$f(\alpha) = \frac{G(p,q)}{H(p,q)} = \frac{a_dp^d + a_{d-1}p^{d-1}q + \ldots + a_0q^d}{b_dp^d + b_{d-1}p^{d-1}q + \ldots + a_0q^d}$$

Therefore both the numerator and denominator can be bounded by $(d+1)C\max(|p|, |q|)^d$ where C is the maximum size of all of the coefficients. Therefore taking logarithms

$$h_{\mathbb{Q}}(f(\alpha)) \leq dh_{\mathbb{Q}}(\alpha) + C'.$$

For a lower bound, we need a lemma about homogeneous polynomials which is an application of the Euclidean algorithm. We will assume that the polynomials have integer coefficients: the generalization to other number fields is trivial.

Lemma 9. *Suppose $g(t)$ and $h(t) \in \mathbb{Z}[t]$ are relatively prime polynomials of degree at most d, and $G(x,y)$ and $H(x,y)$ are the corresponding homogeneous degree d polynomials. Then there exist integers d_1, d_2 and polynomials $R, R', S, S' \in \mathbb{Z}[x,y]$ such that*

$$G(x,y)R(x,y) + H(x,y)S(x,y) = x^d d_1 \quad \textit{and} \quad G(x,y)R'(x,y) + H(x,y)S'(x,y) = y^d d_2.$$

Proof. In $\mathbb{Q}[t]$ use the Euclidean algorithm to write $gr' + hs' = 1$. Then homogenize r' and s' (letting t correspond to x) to get two variable polynomials R' and S' so that $GR' + HS' = y^d$. Finally multiply by an integer d_2 to clear the denominators of the coefficients of R' and S'. Doing the same thing with the roles of x and y reversed gives R and S. □

Given these expressions, consider the fraction . Any integer that cancels in $\frac{G(p,q)}{H(p,q)}$ divides both $G(p,q)$ and $H(p,q)$, and hence divides d_1p^d and d_2q^d. As p and q are relatively prime, writing the fraction in lowest terms will at worst divide $\max(|g(p,q)|, |h(p,q)|)$ by constant factor (depending on f, but independent of p and q). But then

$$|d_1p^d| = |G(p,q)R(p,q) + H(p,q)S(p,q)| \leq 2\max(|G(p,q)|, |H(p,q)|)\max(|R(p,q)|, |S(p,q)|).$$

Combining this with a similar expression for d_2q^d, we see that

$$\log(\max(|G(p,q)|, |H(p,q)|)) \geq d\log\max(|p|, |q|) - C$$

for some constant C. Therefore we conclude that $h_{\mathbb{Q}}(f(\alpha)) = dh_{\mathbb{Q}}(\alpha) + O(1)$.

The third statement, that $h_{\mathbb{Q}}(\alpha) = 0$ if and only if α is ±1 or zero, is obvious. □

2.2. Heights on Elliptic Curves.

Let K be a number field, E be an elliptic curve in Weierstrauss form defined over K, and $P = (a,b) \in E(K)$. Define the (naive) height $h_E(P)$ to be $h_K(a)$.[1] This height satisfies similar properties to the height of an algebraic number.

Proposition 10. *Let E be an elliptic curve defined over a number field K.*

1. *For any $C > 0$, the set $\{Q \in E(K) : h_E(Q) < C\}$ is finite.*
2. *For $P \in E(K)$ and $m \in \mathbb{Z}$, $h_E(mP) = m^2h_E(P) + O(1)$, where the constant depends only on E and m.*
3. *For $P, Q \in E(K)$, the parallelogram law holds:*

$$h_E(P+Q) + h_E(P-Q) = 2h_E(P) + 2h_E(Q) + O(1)$$

where the constant depends only on the curve E.

Proof. There are finitely many elements in K with height less than C, hence only finitely many potential first coordinates for points in $E(K)$ with $h_E(Q) < C$. There are at most two second coordinates for such a point, so the first statement is proven.

For the second, recall that the multiplication by m map induces a degree m map on the first coordinates. Then Proposition 8.2 implies the result, as the degree of the multiplication by m map is m^2.

The last part is a consequence of the explicit formulas for the group law on an elliptic curve and gets a bit messy. The details are Theorem VIII.6.2 in Silverman [2]. □

The next step is to introduce the canonical height of points on elliptic curves.

Proposition 11. *There exists a unique function $\hat{h}_E : E(K) \to \mathbb{R}$ such that*

- *We have $|\hat{h}_E(P) - h_E(P)| < C$ for some constant C and all $P \in E(K)$.*
- *For $m \geq 1$ and $P \in E(K)$, $\hat{h}_E(mP) = m^2\hat{h}_E(P)$.*

The function $\hat{h}_E$ is called the canonical height of E. It is close to the naive height function, but some of the error terms vanish.

[1] For those who care, define the height of the point at infinity to be 0. Alternately, map $E \to \mathbb{P}^1$ using the first coordinate map and use a height on projective space.

Proof. For uniqueness, suppose $\hat{h}_E$ and $\hat{h}'_E$ are two functions satisfying these two properties. For a point P, consider $|\hat{h}_E(2^nP) - \hat{h}'_E(2^nP)|$. By the first property, we know it is bounded by some constant C. On the other hand, by the second property it equals $4^n|\hat{h}_E(P) - \hat{h}'_E(P)|$. Letting n go to infinity, we see that $\hat{h}_E(P) = \hat{h}'_E(P)$.

To show it exists, set

$$\hat{h}_E(P) = \lim_{n\to\infty} \frac{1}{4^n} h_E(2^nP).$$

This limit exists: if $a_n = 4^{-n}h_E(2^nP)$, then

$$|a_n - a_m| = \left| \frac{1}{4^n}\left(4^n h_E(P) + O(1)\right) - \frac{1}{4^m}\left(4^m h_E(P) + O(1)\right) \right| \leq 4^{-\min(m,n)}O(1)$$

and hence the a_n form a Cauchy sequence. Therefore $\hat{h}_E(P)$ is well-defined. This also shows that $|\hat{h}_E(P) - h_E(P)|$ is bounded.

Now consider the function $\hat{h}'_E$ sending $Q \to \frac{1}{m^2}\hat{h}_E(mQ)$.

$$|\hat{h}'_E(P) - \hat{h}_E(P)| = \lim_{n\to\infty} 4^{-n}\left|\frac{1}{m^2}h_E(m2^nP) - h_E(2^nP)\right| = \lim_{n\to\infty} 4^{-n}O(1)$$

which goes to zero as n goes to infinity. Thus $\frac{1}{m^2}\hat{h}_E(mP) = \hat{h}_E(P)$. □

There is an analog of Proposition 10 for the canonical height.

Proposition 12. *Let E be an elliptic curve defined over a number field K.*

(1) *For any $C > 0$, the set $\{Q \in E(K) : \hat{h}_E(Q) < C\}$ is finite.*
(2) *For $P \in E(K)$ and $m \in \mathbb{Z}$, $\hat{h}_E(mP) = m^2\hat{h}_E(P)$.*
(3) *For $P, Q \in E(K)$, the parallelogram law holds:*

$$\hat{h}_E(P+Q) + \hat{h}_E(P-Q) = 2\hat{h}_E(P) + 2\hat{h}_E(Q).$$

(4) *$\hat{h}_E(P) \geq 0$ and is equal to 0 if and only if P is a torsion point.*

Proof. The first follows immediately from the corresponding fact for h_E. The second has already been proven. The third follows from the parallelogram law for h:

$$|h_E(2^n(P+Q)) + h_E((2^n(P-Q))) - 2h_E(2^nP) - 2h_E(2^nQ)| \leq C$$

for any n so multiplying by 2^{-n} and taking the limit as n goes to infinity gives the result for canonical heights.

It is clear that $\hat{h}_E$ is non-negative. If $\hat{h}_E(P) = 0$, then all multiples of P have height 0, so the set of multiples must be finite by the first part. Conversely, if P is a torsion point then $(1+m)P = P$ with $m \neq 1$ and hence $(m+1)^2\hat{h}_E(P) = \hat{h}_E(P)$ by the second part. □

This gives enough information about heights to carry out the proof of the Mordell-Weil theorem.

3. The Descent Argument

We will now prove the Mordell-Weil theorem using the weak Mordell-Weil theorem and the theory of heights.

Proof. Given an elliptic curve E defined over a number field K, the weak Mordell-Weil theorem (Theorem 2) says that $E(K)/mE(K)$ is finite for any integer $m \geq 2$. Pick coset representatives $P_1, \ldots, P_n$, and let $C = \max_i \hat{h}_E(P_i)$. Let S be the set $\{Q \in E(K) : \hat{h}_E(Q) \leq C\}$. By Proposition 12, this is finite. We claim that S generates $E(K)$.

Assume this is not the case, and let Q be an element of $E(K)$ not generated by S of smallest canonical height. Now $Q = mR + P_i$ for some P_i, so by Proposition 12

$$m^2\hat{h}_E(R) = \hat{h}_E(mR) = \hat{h}_E(Q - P_i) \leq \hat{h}_E(Q - P_i) + \hat{h}_E(Q + P_i) = 2\hat{h}_E(Q) + 2\hat{h}_E(P_i) < 4\hat{h}(Q)$$

where the last step uses that $\hat{h}(Q) > C \geq \hat{h}_E(P_i)$. This shows that $\hat{h}_E(R) < \hat{h}_E(Q)$, so by hypothesis R lies in the subgroup of $E(K)$ generated by S. As $Q = mR + P_i$, Q must as well, a contradiction. □

4. Calculating the Rational Points on An Elliptic Curve

In general, it is difficult to calculate the rational points on an elliptic curve. We will illustrate a procedure, based on the proof of the weak Mordell-Weil theorem, that often works.

Let E be the elliptic curve defined over $\mathbb{Q}$ by the Weierstrauss equation $y^2 = x(x-6)(x+6)$. This is in fact a global minimal model for E, as can be verified by calculating the discriminant ($2^{12}3^6$) and checking the prime 2 directly. We will show that $E(\mathbb{Q}) \simeq \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}$.

There are four obvious torsion points: the point at infinity, $(0,0)$, $(6,0)$, and $(-6,0)$. The last three are obviously of order two. A short search reveals that $(-3,9)$ is a point that is not obviously a torsion point. We will show it has infinite order, and that it generates $E(\mathbb{Q})$ along with the torsion points.

To find the torsion part of $E(\mathbb{Q})$, we will use the fact that the torsion prime to p injects into $E(\mathbb{F}_p)$ provided p is a prime of good reduction. Taking $p = 5$, a direct search shows that $E(\mathbb{F}_5)$ consists of 8 points. Taking $p = 11$, a direct search shows that $E(\mathbb{F}_{11})$ has 12 points. Thus the 4 torsion points found are all the rational torsion points.

This shows that $P = (-3,9)$ is a point of infinite order, so the rank of $E(\mathbb{Q})$ is at least one. We will use the Kummer map to show the rank is exactly one and $(-3,9)$ is a generator. Since there are four two-torsion points, we know that $E(K)/2E(K) \simeq (\mathbb{Z}/2\mathbb{Z})^{r+2}$, where $r \geq 1$ is the rank of E. Now consider the diagram we obtain by looking at $\mathbb{Q}_2$ valued points:

$$\begin{tikzcd}
E(\mathbb{Q})/2E(\mathbb{Q}) \arrow[r, "\kappa_{\mathbb{Q}}"] \arrow[d] & \mathbb{Q}^\times/(\mathbb{Q}^\times)^2 \times \mathbb{Q}^\times/(\mathbb{Q}^\times)^2 \arrow[d] \\
E(\mathbb{Q}_2)/2E(\mathbb{Q}_2) \arrow[r, "\kappa_{\mathbb{Q}_2}"] & \mathbb{Q}_2^\times/({\mathbb{Q}_2}^\times)^2 \times \mathbb{Q}_2^\times/({\mathbb{Q}_2}^\times)^2
\end{tikzcd}$$

In fact, we know more about the image of $E(\mathbb{Q})/2E(\mathbb{Q})$ under $\kappa_{\mathbb{Q}}$. If p is a prime of good reduction, then the proof of Lemma 5 shows both coordinates of the Kummer map have even p-adic valuation. As a rational number is a square if and only if it is positive and every prime divides it an even number of times, this shows the image is contained in

$$\{\pm 1, \pm 2, \pm 3, \pm 6\} \times \{\pm 1, \pm 2, \pm 3, \pm 6\} \subset \mathbb{Q}^\times/(\mathbb{Q}^\times)^2 \times \mathbb{Q}^\times/(\mathbb{Q}^\times)^2.$$

It is a coincidence that in this case, it injects into $\mathbb{Q}_2^\times/({\mathbb{Q}_2}^\times)^2 \times \mathbb{Q}_2^\times/({\mathbb{Q}_2}^\times)^2$: in general there will be a (simple to understand) kernel. So to understand the image of $\kappa_{\mathbb{Q}}$, we can attempt to understand the image of $\kappa_{\mathbb{Q}_2}$. We will soon see that this image is abstractly isomorphic to $(\mathbb{Z}/2\mathbb{Z})^3$. This shows the rank of E is exactly one. To show that $(-3,9)$ is a generator, one simply does a search for points on the elliptic curve with smaller height. This is slightly complicated by the fact one uses canonical height instead of the naive height, but is a standard calculation that can be done for example by SAGE.

We now analyze the image of $\kappa_{\mathbb{Q}}$ by repeatedly using the kernel cokernel exact sequence and the standard short exact sequences for an elliptic curve over a local field. First, we know that $E(\mathbb{Q}_2)/E_0(\mathbb{Q}_2)$ has order at most 4 since the curve has additive reduction at $p = 2$. Furthermore, $(0,0)$ is a singular point on E over $\mathbb{F}_2$, and all three of the non-trivial two torsion points reduce to it. Therefore $E(\mathbb{Q}_2)/E_0(\mathbb{Q}_2) \simeq (\mathbb{Z}/2\mathbb{Z})^2$, with generators $(0,0)$ and $(6,0)$. Now consider the exact sequences

$$\begin{tikzcd}
0 \arrow[r] \arrow[d] & E_0(\mathbb{Q}_2) \arrow[r] \arrow[d, "\cdot 2"'] & E(\mathbb{Q}_2) \arrow[r] \arrow[d, "\cdot 2"'] & (\mathbb{Z}/2\mathbb{Z})^2 \arrow[r] \arrow[d, "\cdot 2"'] & 0 \arrow[d] \\
0 \arrow[r] & E_0(\mathbb{Q}_2) \arrow[r] & E(\mathbb{Q}_2) \arrow[r] & (\mathbb{Z}/2\mathbb{Z})^2 \arrow[r] & 0
\end{tikzcd}$$

Taking the kernel-cokernel exact sequence, we see that

$$0 \to E_0(\mathbb{Q}_2)[2] \to E(\mathbb{Q}_2)[2] \to (\mathbb{Z}/2\mathbb{Z})^2 \xrightarrow{\delta} E_0(\mathbb{Q}_2)/2E_0(\mathbb{Q}_2) \to E(\mathbb{Q}_2)/2E(\mathbb{Q}_2) \to (\mathbb{Z}/2\mathbb{Z})^2 \to 0.$$

The two-torsion reduces to singular points, so $E_0(\mathbb{Q}_2)[2] = 0$. As $E(\mathbb{Q}_2)[2] \simeq (\mathbb{Z}/2\mathbb{Z})^2$, the map δ is the zero map. Thus we have a short exact sequence

$$0 \to E_0(\mathbb{Q}_2)/2E_0(\mathbb{Q}_2) \to E(\mathbb{Q}_2)/2E(\mathbb{Q}_2) \to (\mathbb{Z}/2\mathbb{Z})^2 \to 0. \tag{2}$$

So to understand $E(\mathbb{Q}_2)/2E(\mathbb{Q}_2)$, we must understand $E_0(\mathbb{Q}_2)/2E_0(\mathbb{Q}_2)$.

To do this, look at exact sequences

$$\begin{tikzcd}
0 \arrow[r] \arrow[d] & E_1(\mathbb{Q}_2) \arrow[r] \arrow[d, "\cdot 2"'] & E_0(\mathbb{Q}_2) \arrow[r] \arrow[d, "\cdot 2"'] & \tilde{E}_{ns}(\mathbb{F}_2) \arrow[r] \arrow[d, "\cdot 2"'] & 0 \arrow[d] \\
0 \arrow[r] & E_1(\mathbb{Q}_2) \arrow[r] & E_0(\mathbb{Q}_2) \arrow[r] & \tilde{E}_{ns}(\mathbb{F}_2) \arrow[r] & 0
\end{tikzcd}$$

We can directly see that $\tilde{E}_{ns}(\mathbb{F}_2) = \{\mathcal{O}, (1,1)\} \simeq \mathbb{Z}/2\mathbb{Z}$. Then taking the kernel-cokernel exact sequence, we see

$$0 \to E_1(\mathbb{Q}_2)[2] \to E_0(\mathbb{Q}_2)[2] \to \mathbb{Z}/2\mathbb{Z} \to E_1(\mathbb{Q}_2)/2E_1(\mathbb{Q}_2) \to E_0(\mathbb{Q}_2)/2E_0(\mathbb{Q}_2) \to \mathbb{Z}/2\mathbb{Z} \to 0.$$

Since the two torsion of E reduces to singular points, the first two groups are trivial. Thus we obtain the exact sequence

$$0 \to \mathbb{Z}/2\mathbb{Z} \to E_1(\mathbb{Q}_2)/2E_1(\mathbb{Q}_2) \to E_0(\mathbb{Q}_2)/2E_0(\mathbb{Q}_2) \to \mathbb{Z}/2\mathbb{Z} \to 0. \tag{3}$$

We can understand $E_1(\mathbb{Q}_2)/2E_1(\mathbb{Q}_2)$ using the theory of formal groups. In particular, we know that $E_1(\mathbb{Q}_2) \simeq \hat{E}(2\mathbb{Z}_2)$ and $\hat{E}(4\mathbb{Z}_2) \simeq (\mathbb{Z}_2, +)$ have no two torsion. As $\hat{E}(2\mathbb{Z}_2)/\hat{E}(4\mathbb{Z}_2) \simeq \mathbb{Z}/2\mathbb{Z}$, we have a short exact sequence

$$0 \to \hat{E}(4\mathbb{Z}_2) \to \hat{E}(2\mathbb{Z}_2) \to \mathbb{Z}/2\mathbb{Z} \to 0.$$

Using the multiplication by two map and taking the kernel-cokernel exact sequence as before, we get an exact sequence

$$0 \to 0 \to 0 \to \mathbb{Z}/2\mathbb{Z} \xrightarrow{\delta'} \hat{E}(4\mathbb{Z}_2)/2\hat{E}(4\mathbb{Z}_2) \to \hat{E}(2\mathbb{Z}_2)/2\hat{E}(2\mathbb{Z}_2) \to \mathbb{Z}/2\mathbb{Z} \to 0.$$

As we know that $\hat{E}(4\mathbb{Z}_2)/2\hat{E}(4\mathbb{Z}_2) \simeq \mathbb{Z}_2/2\mathbb{Z}_2 \simeq \mathbb{Z}/2\mathbb{Z}$, δ' is an isomorphism so

$$E_1(\mathbb{Q}_2)/2E_1(\mathbb{Q}_2) \simeq \hat{E}(2\mathbb{Z}_2)/2\hat{E}(2\mathbb{Z}_2) \simeq \mathbb{Z}/2\mathbb{Z}.$$

Using this in (3), we see that $E_0(\mathbb{Q}_2)/2E_0(\mathbb{Q}_2)$ must be isomorphic to $\mathbb{Z}/2\mathbb{Z}$. Using this in (2), this shows that $E(\mathbb{Q}_2)/2E(\mathbb{Q}_2)$ has order 8. Since we know it is two-torsion, we have established that $E(\mathbb{Q}_2)/2E(\mathbb{Q}_2) \simeq (\mathbb{Z}/2\mathbb{Z})^3$.

In conclusion, we have shown that $E(\mathbb{Q}_2)/2E(\mathbb{Q}_2)$, and hence its image under the Kummer map, is isomorphic to $(\mathbb{Z}/2\mathbb{Z})^3$.

Remark 13. This procedure does not always work. It is possible that the image of the $2-$adic Kummer map will be larger than the image of the rational Kummer map, in which case it is impossible to get a tight bound on the rank. In addition this requires that the two-torsion be rational, although this can be overcome. The case when the two-torsion is rational is known as complete two-descent. It is also possible to change from $m = 2$ to $m = 3$, or use the real numbers or a different p-adic field to obtain additional information.

Appendix A. Group Cohomology

Let G be a profinite group and A, B, and C be continuous G-modules.[2] Given enough algebraic machinery, group cohomology is easy to construct. Look at the functor $\mathrm{Hom}_{\text{G-mod}}(\mathbb{Z}G, -)$ from the category of continuous G-modules to the category of groups. It is left exact, so if

$$0 \to A \to B \to C \to 0 \tag{4}$$

then we know that

$$0 \to \mathrm{Hom}(\mathbb{Z}, A) \to \mathrm{Hom}(\mathbb{Z}, B) \to \mathrm{Hom}(\mathbb{Z}, C).$$

It is not right exact. However, as it can be viewed as the category of modules over $\mathbb{Z}G$ there are enough injectives so there are right derived functors that extend the sequence.[3] Note that as $\mathbb{Z}$ has the trivial G action $\mathrm{Hom}(\mathbb{Z}, M) = M^G$, the elements of M fixed by G. We write the long exact sequence of derived functors as

$$0 \to A^G \to B^G \to C^G \to H^1(G, A) \to H^1(G, B) \to H^1(G, C) \to H^2(G, A) \to \dots$$

However, as we need a specific interpretation of $H^1(G, -)$, it is simplest just to explicitly define them directly and ignore the general theory.

Definition 14. For a continuous G-module M, a *crossed homomorphism* is a continuous map $\psi : G \to M$ such that $\psi(gh) = \psi(g) + \psi(h)^g$. A *principal crossed homomorphism* is a map of the form $g \to m^g - m$ for some $m \in M$.

We define $H^0(G, M) = M^G$ and $H^1(G, M)$ to be continuous crossed homomorphisms from G to M module the principal crossed homomorphisms.

Given a map $\varphi : A \to B$, there are obvious maps $H^0(G, A) \to H^0(G, B)$ and $H^1(G, A) \to H^1(G, B)$ given by composing with φ. Additionally, given a short exact sequence (4), we can also define a map $\delta : C^G \to H^1(G, A)$ as follows: let $c \in C^G$. Pick a $b \in B$ that maps to c by exactness. Note that $b^g - b$ maps to $0 \in C$, so $b^g - b \in A$. Now define $\delta(c)$ to be the crossed morphism ψ where $\psi(g) = b^g - b \in A$.

Lemma 15. *With the notation above $\delta : C^G \to H^1(G, A)$ is a well defined map.*

Proof. Note that ψ is a continuous map as the action of G on B is continuous. It is a crossed homomorphism as $\psi(gh) = b^{gh} - b = b^{gh} - b^g + b^g - b = \psi(g) + \psi(h)^g$.

Suppose we pick $b' \in B$ which also maps to $c \in C^G$. Then $b - b'$ maps to 0 in C, so there is an $a \in A$ such that $a = b - b'$. Thus the maps ψ and ψ' satisfy

$$\psi(g) - \psi'(g) = b^g - b - (b')^g + b' = a^g - a.$$

This is a principal crossed homomorphism. □

Proposition 16. *With the notation above we have a long exact sequence*

$$0 \to A^G \to B^G \to C^G \xrightarrow{\delta} H^1(G, A) \to H^1(G, B) \to H^1(G, C).$$

Proof. To show it is exact, simply unwind the definitions. For example, suppose $c \in C^G$ maps to a principal crossed homomorphism. This means there exist a and b such that $b^g - b = a^g - a$. But then $(b - a)^g - (b - a) = 0$ for all g, so $(b - a) \in B^G$. Buts its image is c, so the sequence is exact at C^G. The others are equally simple. □

[2] If G is a finite group, it is a profinite group when given the discrete topology, and every G-module is continuous automatically. The case when G is the Galois group of an infinite extension is relevant for us later.

[3] For more details, especially about profinite groups as opposed to the more basic finite group case, see [1].

In the context we need, G will be a (possibly infinite) Galois group of L over K with the usual profinite topology. The G-modules will be discrete. This means that a G module is continuous if for all $m \in M$ and $g \in \mathrm{Stab}_G(m)$, there is a finite extension K_1 of K such that $g \in \mathrm{Gal}(L/K_1)$ and the group stabilizes m. Continuity for a crossed homomorphism means that for every element g of the kernel, there is a finite extension K_1 so that $\mathrm{Gal}(L/K_1)$ contains g and lies in the kernel.

Appendix B. General Proof of Proposition 8

We begin by proving the first assertion, that $\{\alpha | h_K(\alpha) < C\}$ is finite. Assume K is Galois by passing to the normal closure and renormalizing, and consider the action of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ on K. Now

$$\begin{aligned} h_k(\alpha^\sigma) &= \sum_{v \in \Sigma} \max(0, \log(|\alpha^\sigma|_v)) \\ &= \sum_{v \in \Sigma} \max(0, \log(|\alpha|_v)) \end{aligned}$$

as the absolute values are permuted by σ.

Let $f(T) = T^d + c_{d-1}T^{d-1} + \ldots + c_0 = (T - \alpha_1)(T - \alpha_2) \ldots (T - \alpha_d)$ be a polynomial with coefficients in K. We will prove by induction on d that

$$\max_{0 \le i \le d} \{|c_i|_v\} \le C_d \prod_{i=0}^{d} \max\{|\alpha_i|_v, 1\}. \tag{5}$$

For $d = 1$, it is clear. Assuming the result for $d - 1$, pick k with $0 \le k \le d$ and $|\alpha_k|_v$ maximal and define

$$g(T) = (T - \alpha_1) \ldots (\widehat{T - \alpha_k}) \ldots (T - \alpha_d) = T^{d-1} + b_{d-2}T^{d-2} + \ldots + b_0$$

by omitting the $(T - \alpha_k)$ term. Comparing coefficients shows $c_i = b_i - \alpha_k b_{i-1}$. For notational convenience, set $b_{-1} = b_d = 0$ so there are no problems with the leading and constant terms. But then

$$\begin{aligned} \max_{0 \le i \le d} \{|c_i|_v\} &= \max_{0 \le i \le d} |b_i - \alpha_k b_{i-1}|_v \\ &\le 2 \max_{0 \le i \le d} |b_i|_v \max\{|\alpha_k|_v, 1\} \\ &\le C_d \prod_{j=1}^{d} \max\{|\alpha_i|_v, 1\} \end{aligned}$$

using the inductive hypothesis. This completes the proof of (5). If we multiply this over all absolute values $v \in \Sigma$ and taking logarithms, we get that

$$\sum_{v \in \Sigma} \max_{0 \le i \le d} \log |c_i|_v \le C_d + \sum_{j=1}^{d} h_K(\alpha_j). \tag{6}$$

We now show that $\{\alpha : h_K(\alpha) < C\}$ is finite. For any such α, consider the monic minimal polynomial of α over $\mathbb{Q}$, whose coefficients are rational numbers. All of the roots have the same height as α, so (6) shows that

$$\sum_{v \in \Sigma} \max_{0 \le i \le d} \log |c_i|_v = O(C^d)$$

But as $c_i \in \mathbb{Q}$, $\sum_{v \in \Sigma} \log |c_i|_v = 0$, so there are only a finite number of c_i with $\log |c_i|_v < O(C^d)$ for all v. Therefore there only finitely many choices for the minimal polynomial of α. Therefore there are a finite number of α with bounded height.

Suppose that $\alpha = \frac{p}{q} \in K^\times$. Note that $\max(|\alpha|_v, 1) = \max(\frac{|p|_v}{|q|_v}, 1) = \max(|p_v|, |q|_v)|q|_v^{-1}$. Then

$$\prod_{v\in\Sigma} \max(|\alpha|_v, 1) = \prod_{v\in\Sigma} \max(|p_v|, |q_v|)|q|_v^{-1}.$$

However, we know that $\prod_{v\in\Sigma} |x|_v = 1$ since the absolute values are normalized to make this true. Thus taking logarithms

$$h_K(\alpha) = \sum_{v\in\Sigma} \log(\max(|p_v|, |q_v|)).$$

We now establish the second statement for a function of degree n

$$f(T) = \frac{a_nT^n + a_{n-1}T^{n-1} + \ldots + a_0}{b_nT^n + b_{n-1}T^{n-1} + \ldots + b_0}.$$

Write $\alpha = \frac{p}{q}$ with $p, q \in \mathcal{O}_K$, so

$$f(\alpha) = \frac{a_np^n + a_{n-1}p^{n-1}q + \ldots + a_0q^n}{b_np^n + b_{n-1}p^{n-1}q + \ldots + a_0q^n} = \frac{A(p,q)}{B(p,q)}.$$

Then we know

$$h_K(f(\alpha)) = \sum_{v\in\Sigma} \log(\max(|A(p,q)|_v, |B(p,q))|_v).$$

By the ultrametric inequality, $|A(p,q)|_v \le C\max(|p|_v, |q|_v)^n$, where the constant C depends on the valuation of the coefficients of the polynomial A. For the finite number of infinite places, use the triangle inequality instead, which requires the constant C to increase based on the degree n but doesn't change the conclusion. The same bound holds for $B(p,q)$. Thus

$$h_K(f(\alpha)) \le \sum_{v\in\Sigma} \log C_v \max(|p|_v, |q|_v)^n$$

Note that $\prod_{v\in\Sigma} C_v$ is finite as for all but a finite set of places the absolute value of all of the coefficients is 1, which implies $C_V = 1$. Thus

$$h_K(f(\alpha)) \le nh_K(\alpha) + C_f$$

where n is the degree of f.

To prove the lower bound, note that we may take A and B to be coprime as polynomials. They are homogeneous, so viewing them as being relatively prime polynomials in $K[X]$ we can write $Ar + Bs = p^m$ and $Ar' + Bs' = q^m$ for $r, s, r', s' \in K[p, q]$. (The number m is the maximum of the degree of $Ar + Bs$ viewed as a polynomial in q with the degree of $Ar' + Bs'$ viewed as a polynomial in p.)

Write $|R|_v$ for the maximum of $|c|_v$ over c a coefficient of r, r', s, and s. Note $|R|_v$ is independent of α. Now evaluating the identities of polynomials on (p, q) gives that

$$\begin{aligned} |p|_v^m = |A(p,q)r(p,q) + B(p,q)s(p,q)| &\le C_v \max\{A(p,q)r(p,q), B(p,q)s(p,q)\} \\ &\le C_v \max\{r(p,q), s(p,q)\} \max\{A(p,q), B(p,q)\} \end{aligned}$$

where $C_v = 1$ if the ultrametric inequality holds and $C_v = 2$ for the finite number of exceptions. There is a similar expression for q. Thus

$$\max\{|p|_v, |q|_v\}^m \le C_v \max\{A(p,q), B(p,q)\} \left(\max\{r(p,q), s(p,q)\} + \max\{r'(p,q), s'(p,q)\}\right)$$

However, as the polynomials r, r', s, and s' are homogeneous of degree $m - n$,

$$|r(p,q)|_v \le D|R|_v \max(|p|_v, |q|_v)^{m-n}.$$

where D depends only on whether v is archimedean and the degree. It is 1 unless v is an infinite place. A similar statement holds for r', s, and s'. Substituting and multiplying by $\max(|p|_v, |q|_v)^{n-m}$ gives that

$$\max\{|p|_v, |q|_v\}^n \leq C_v D |R|_v \max\{A(p,q), B(p,q)\}.$$

Taking logarithms and summing over Σ, noting that $C_v D$ is almost always 1, we obtain

$$h_K(f(\alpha)) \geq n h_K(\alpha) + C'.$$

Thus $h_K(f(\alpha)) = n h_K(\alpha) + O(1)$, which is the second part of the proposition.

For the final part of the proposition about heights, it is clear that zero and roots of unity have height zero. Otherwise, if the height of α is 0, consider the powers of α, all of which have height 0. By the first part, this must be a finite set, so α is a root of unity (or zero). □

References

1. Brian Conrad, *Math 210b profinite group cohomology*, `http://math.stanford.edu/~conrad/210BPage/handouts/profcohom.pdf`.
2. J.H. Silverman, *The arithmetic of elliptic curves*, Graduate Texts in Mathematics, Springer, 2009.