Universiteit Utrecht

Faculty of Science

The Mordell-Weil theorem for elliptic curves

Bachelor Thesis

Jonathan Grube

Mathematics

Supervisor:

Dr. Marta Pieropan
Utrecht University

June 2023

Contents

1 Introduction

In this thesis we will prove a fundamental result in the study of elliptic curves: the Mordell-Weil theorem. Elliptic curves are a special type of (projective) algebraic variety. There is a very natural way to define a group law on them by constructing specific lines through the points of the variety. What these lines are will be clear when we define the group operation. Elliptic curves have become essential for research in modern mathematics. For example, in 1995 Andrew Wiles proved the famous Fermat's Last Theorem by showing that all elliptic curves over the rationals have the property of being modular. There are also a lot of algorithms that use the group law on elliptic curves, for example in cryptography. The main theorem we will prove is:

Theorem 1.1 (Mordell-Weil)**.** *If K is a finite field extension of $\mathbb{Q}$, then the group $E(K)$ of K-rational points on an elliptic curve is abelian and finitely generated.*

The fundamental theorem of finitely generated abelian groups then tells us that

$$E(K) \cong \mathbb{Z}^r \oplus \mathbb{Z}/a_1\mathbb{Z} \oplus \cdots \oplus \mathbb{Z}/a_n\mathbb{Z},$$

for $r \geq 0$ and $a_1, \ldots, a_n \geq 1$. What exactly this group $E(K)$ is, we will define in section 3.2. In general the torsion part of the group is well understood. There exist algorithms to calculate it and there is a classification showing which groups it can be. On the other hand, the number r is called the *rank* of the group and it is much more mysterious. The curve with the largest know rank has rank 20, but there also exist curves where the rank is not exactly known but at least 28. A very famous conjecture concerning the rank is the Birch and Swinnerton-Dyer conjecture. This is one of the seven millenium problems.

It will be immediate from the construction that the group is abelian, so we will concern ourselves with proving that it is finitely generated. First in section 2 we define what an elliptic curve is and give a geometric interpretation of the group law. After that we give a short introduction in section 3.1 into the Galois theory of infinite extensions. We will use the results we prove here in later sections, such as section 6.1. Next we prove some first results about the group of points on elliptic curve in section 3.2. In section 4 we introduce a method for showing that the group is finitely generated and in section 5 we give a conditional proof for the case $K = \mathbb{Q}$ under the assumption that $E(\mathbb{Q})/mE(\mathbb{Q})$ is finite. In section 6 we prove the aforementioned assumption. In section 6.1 we first reduce the problem to showing that a certain field extension is finite and in section 6.2 we show that the extension is actually finite. For sections 4 to 6 we closely follow the proof by Silverman in [1, section VIII]. Section 3.1 is based on section 7 of [2].

We assume that the reader is familiar with abstract algebra as it is taught at bachelor level. This includes the theory of groups, rings and fields. We also suppose that the reader has been introduced to Galois theory for finite field extensions, as this is our starting point when we introduce Galois theory for infinite extensions. For 2 and 3.2 it is useful to be acquainted with some elementary algebraic geometry. We will also use notions from algebraic number theory, but we will develop these ourselves and they are not prerequisites.

Throughout this thesis, we let K be a fixed field with $\mathrm{char}(K) = 0$. We write $\bar{K}$ for a fixed algebraic closure of K. Any time we mention some algebraic field extension F of K, it will be implicit that it is contained in $\bar{K}$. For two fields F_1, F_2, we write F_2/F_1 for $F_1 \subseteq F_2$ and we consider the extension of F_2 over F_1. When we have a third field, we write $F_3/F_2/F_1$ for $F_1 \subseteq F_2 \subseteq F_3$ and we consider the respective extensions. For a ring R, we define $R^\times$ to be its group of units. If we have two subsets S, T of a commutative ring R, we define ST to be the set $\{\sum_{i=1}^n s_i t_i \mid n \in \mathbb{Z}_{\geq 1}, s_i \in S, t_i \in T \text{ for all } 1 \leq i \leq n\}$. Note that if $T = R$, then ST is the ideal in R generated by S. If $S = \{r\}$, we write rT. We denote the cardinality of a set S by $\#S$. The notation $\mathcal{F}(X, Y)$ denotes the set of all functions from the set X to the set Y. If we work with a equivalence relation, we write $[\cdot]$ for the equivalence classes. This is not to be confused with the notation $[x : y : z]$ for a point of the projective plane $\mathbb{P}^2(F)$ over a field F. For a polynomial $f \in F[X_1, \ldots, X_n]$, we write $V(f)$ for the set of roots of f in F. It will be clear from context if this is supposed to be affine or projective.

2 The group law on elliptic curves

First consider the following homogeneous polynomial in $\bar{K}[X, Y, Z]$:

$$W_{A,B} := -Y^2Z + X^3 + AXZ^2 + BZ^3,$$

where $A, B \in \bar{K}$. We call this the *Weierstrass polynomial*. We write $E(\bar{K})$ for $V(W_{A,B})$, the set of roots of $W_{A,B}$ in $\mathbb{P}^2(\bar{K})$.

Definition 2.1. The curve $E(\bar{K})$ is called an *elliptic curve* if it is non-singular and irreducible.

There are more general ways to define an elliptic curve. See for example [1, section III.3]. When we describe the group law on elliptic curves, it will become obvious why we want the curve to be irreducible and non-singular. First we make the following observation. Roots of $W_{A,B}$ with $Z = 0$ clearly also need $X = 0$. For the Y-coordinate we can choose anything we want in $\bar{K}^\times$. But note that if $X = Z = 0$, then these are all a multiple of each other, so the same point in $\mathbb{P}^2(\bar{K})$. This means that $[0 : 1 : 0]$ is the only point on $E(\bar{K})$ with the last coordinate equal to 0. From this point on, we will call this point O and it will play an important role in the group law of elliptic curves. Namely, it will function as the identity of the group.

We will now describe the group law on elliptic curves, which has a very elegant geometric interpretation. To do this, we will make use of a fundamental theorem in algebraic geometry due to Bézout.

Theorem 2.2 (Bézout)**.** *Let $f, g \in \bar{K}[X, Y, Z]$ be homogeneous without a common irreducible component, of degree $d_1 > 0$ and $d_2 > 0$ respectively. Then*

$$\sum_{P \in \mathbb{P}^2(\bar{K})} I(P, f \cap g) = d_1 d_2.$$

Here $I(P, f \cap g)$ is the intersection number of f and g at P.

Proof. For a proof, see [3, section 5.3]. □

We consider the projective curve $E(\bar{K})$ given by the polynomial $W_{A,B} \in \bar{K}[X, Y, Z]$ and P, Q two points in $E(\bar{K})$. Then the unique line, given by the polynomial $l \in \bar{K}[X, Y, Z]$, trough these points is well-defined if they are distinct. But what if we want to add a point P to itself? In this case we consider the tangent line of $E(\bar{K})$ at P. We will also call this tangent line 'the line through P and P'. Let us assume for now that $W_{A,B}$, and therefore $E(\bar{K})$, is irreducible. So if l and $W_{A,B}$ were to share an irreducible common component, $W_{A,B}$ would divide l. But this is not possible, since l defines a line and so l is of degree one. Therefore $W_{A,B}$ and l do not share a common irreducible component. This means we can use Bézout's theorem to conclude that $E(\bar{K})$ and the line intersect in exactly three points, counted with multiplicity. This means that we can write $V(l) \cap E(\bar{K}) = \{P, Q, R\}$, where R does not need to be distinct from P or Q. For example if $I(P, l \cap W_{A,B}) = 2$, then $P = R$. Since both P and Q are elements of $E(\bar{K})$ and $V(l)$, they are always in the set $V(l) \cap E(\bar{K})$. We will denote this point R by $L(P, Q)$. Now we repeat this by taking the line through $L(P, Q)$ and O. We will denote the point we get from this with $P + Q = L(O, L(P, Q))$. In 1, we see an affine representation of $\mathbb{P}^2$ and the curve given by $W_{-6,3}$ as an example. The claim now is that this operation satisfies the group axioms.

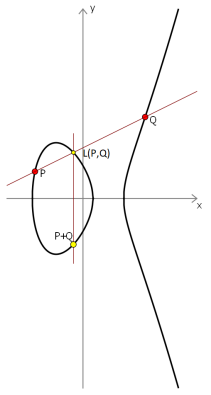

Figure 1: Made with [4]

First note that we assumed that $E(\bar{K})$ is irreducible and implicitly that there is a unique tangent line at every point on $E(\bar{K})$. For this reason we only want to consider $W_{A,B}$ such that $E(\bar{K})$ is non-singular and irreducible.

Theorem 2.3. *If $E(\bar{K})$ is an elliptic curve, the operation on $E(\bar{K})$ defined by $P + Q = L(O, L(P, Q))$ makes $E(\bar{K})$ into an abelian group with O as the identity.*

Proof. Let $E(\bar{K})$ be an elliptic curve given by the Weierstrass polynomial $W_{A,B}$ and $P, Q \in E(\bar{K})$. First we note that the line through P and Q is of course the same as the line through Q and P. Therefore $L(P,Q) = L(Q,P)$ and so $P + Q = Q + P$. This shows that the operation is commutative. Next we will show that it satisfies the group axioms.

We will show that O acts as the identity of the group. By definition we get that $P + O = L(O, L(P,O))$. But $L(P,O)$ lies, by definition, on the line through P and O. Therefore P lies on the line through O and $L(P,O)$ and so $L(O, L(P,O)) = P$.

Next we will show that $L(P,O)$ is the inverse of P. First we have that $P + L(P,O) = L(O, L(P, L(P,O)))$. But O certainly lies on the line through P and $L(P,O)$, so $L(P, L(P,O)) = O$. This shows that $P + L(P,O) = L(O,O)$. Next note that O is a root of both $\frac{\partial W_{A,B}}{\partial X}$ and $\frac{\partial W_{A,B}}{\partial Y}$. Therefore the tangent line of $E(\bar{K})$ at O is given by $Z = O$. But we already noted that O is the only point that is both on this line and the curve. Therefore $I(O, Z \cap W_{A,B}) = 3$ and so $L(O,O) = O$. This means that $P + L(P,O) = O$, so we conclude that $L(P,O)$ is the inverse of P.

The last property that we have to show is associativity. The most straightforward way to do this is to write out explicit formulas for the coordinates of $L(P,Q)$, where $P, Q \in E(\bar{K})$. In this way, one can compute that $P + (Q + R) = P + (Q + R)$ for all $P, Q, R \in E(\bar{K})$. This is unfortunately extremely tedious, since one has to distinguish between multiple cases with different formulas (see theorem 2.8). It is also possible to continue with the geometric arguments we have been using up until now. For a proof of this type, we refer the reader to [3, section 5.6, proposition 4]. □

We would now like to know which Weierstrass polynomials give us an elliptic curve and which don't. It turns out that our two main criteria, smoothness and irreducibility, are actually dependent on each other.

Lemma 2.4. *If the curve $E(\bar{K})$ given by $W_{A,B}$ is non-singular, then it is also irreducible.*

Proof. Suppose $E(\bar{K})$ is reducible. Then $W_{A,B}$ is too, so we write $W_{A,B} = f_1 f_2$ with $f_1, f_2 \in \bar{K}[X,Y,Z]$ homogeneous and non-constant. Using Bézout, the curves given by f_1 and f_2 have to intersect at some point P. But then

$$\frac{\partial W_{A,B}}{\partial X}(P) = \frac{\partial f_1}{\partial X}(P) f_2(P) + \frac{\partial f_2}{\partial X}(P) f_1(P) = 0,$$

since P is a root of both f_1 and f_2. In the same way we see that the other partial derivatives will also be zero. This shows that the point P, which certainly lies in $E(\bar{K})$, is singular. □

So if we can show that $E(\bar{K})$ is smooth, it will also be irreducible. This then makes it into an elliptic curve and we can add points as described above. Now we just have to find a criterion that tells us which combinations of A and B will give us a smooth curve. For this purpose we will introduce a new quantity.

Definition 2.5. We call $\Delta = 4A^3 + 27B^2$ the *discriminant* of $E(\bar{K})$.

The attentive reader might recognise that $-\Delta$ is the discriminant of the polynomial $X^3 + AX + B \in \bar{K}[X]$ (see for example [5, section 14.6]). The next proposition gives us a necessary and sufficient condition to see if a Weierstrass polynomial gives us an elliptic curve.

Proposition 2.6. *The curve $E(\bar{K})$ defined by the the equation $W_{A,B} \in \bar{K}[X,Y,Z]$, has a singular point if and only if $\Delta = 0$.*

Proof. To investigate the point $O = [0:1:0]$, we dehomogenise the equation with respect to Y and get $(W_{A,B})_* = -Z + X^3 + AXZ^2 + BZ^3$. Now note that $\frac{\partial (W_{A,B})_*}{\partial Z}(0,0) = -1$, so O is non-singular. We know that O is the only point on E on the line at infinity with respect to Z. So to check all other points of E, we can dehomogenise with respect to Z. Consider the system of polynomials

$$\begin{cases} (W_{A,B})_* := -Y^2 + X^3 + AX + B \\ \frac{\partial (W_{A,B})_*}{\partial Y} := -2Y \\ \frac{\partial (W_{A,B})_*}{\partial X} := 3X^2 + A. \end{cases}$$

The curve $E(\bar{K})$ has a singular point $[x : y : 1]$ if and only if these polynomials share a common root. Suppose the polynomials have a common root (x, y). From the second equation we get that $y = 0$. The third equation tells us that $A = -3x^2$. Plugging these into the first equation gives $x^3 - 3x^3 + B = 0$, so $2x^3 = B$. Now for the discriminant we get $\Delta = 4(-3x^2)^3 + 27(2x^3)^2 = 0$. For the other direction, suppose that $\Delta = 0$. As we noted above, $-\Delta$ is the discriminant of the cubic $X^3 + AX + B$. By definition of the discriminant of a polynomial, this cubic polynomial has a double root $\alpha \in \bar{K}$. With [5, section 13.5, proposition 33], we see that $\frac{\partial(W_{A,B})_*}{\partial X}$ also has α as a root. Now it is immediate that $(\alpha, 0)$ is a root of every polynomial in the system, so $[\alpha : 0 : 1]$ is a singular point on $E(\bar{K})$. This proves that E has a singular point if and only if $\Delta = 0$. □

Combining lemma 2.4 and proposition 2.6 immediately yields the following.

Corollary 2.7. *The curve $E(\bar{K})$ given by the Weierstrass equation $W_{A,B}$ is an elliptic curve if and only if $4A^3 + 27B^2 \neq 0$.*

From now on we fix the coefficients of the Weierstrass polynomial such that $\Delta \neq 0$. We will now write out the algebraic formulas for the addition of points on an elliptic curve. These are unfortunately not very nice to look at, but we need them for some of our later results (for example theorem 5.2). To check that these are actually correct is just a laborious calculation of constructing lines and we will not present this here. For a proof, see [1, section III.2].

Theorem 2.8 (Formulas for the addition on elliptic curves)**.** *Let $P \in E(\bar{K})$ such that $P \neq O$. Then we can write $P = [x_1 : y_1 : 1]$. The opposite of P is given by $-P = [x_1 : -y_1 : 1]$. Now let $Q \in E(\bar{K})$ such that $Q \neq O, P, -P$. Then we can write $Q = [x_2 : y_2 : 1]$. Now $P + Q = [x_3 : y_3 : 1]$ where*

$$x_3 = \left(\frac{y_2 - y_1}{x_2 - x_1}\right)^2 - (x_1 + x_2),$$
$$y_3 = \frac{y_2 - y_1}{x_2 - x_1}(x_1 - x_3) - y_1.$$

Note that if $x_2 = x_1$, then $y_2 = \pm y_1$ by the shape of $W_{A,B}$. This means that $Q = P$ or $Q = -P$ which gives a contradiction. This shows that it actually makes sense to divide by $x_2 - x_1$. The last case to consider is adding a point to itself. Assume that $P + P \neq O$. Then we write $P + P = [x_4 : y_4 : 1]$, where

$$x_4 = \left(\frac{3x_1^2 + A}{2y_1}\right)^2 - 2x_1,$$
$$y_4 = \frac{3x_1^2 + A}{2y_1}(x_1 - x_3) - y_1.$$

If $y_1 = 0$ then $P = -P$ which is a contradiction. So it makes sense to divide by y_1.

3 Some preliminary results

3.1 Galois theory for infinite (algebraic) field extensions

In this section we will introduce Galois theory for arbitrary algebraic field extensions. Our treatment will be very minimal and is mainly concerned with results that we can directly apply in proofs later on in this thesis. For a more in-depth, but very accessible, treatment of the subject, see [2]. We assume that the reader is familiar with Galois theory for finite field extensions. If not, one can find a nice exposition of this in both [5] and [2]. For these finite field extensions, there are a number of equivalent ways to define what a Galois extension is. See for example [5, section 14.2]. In this section, one definition we find is that an extension is Galois if and only if it is finite, normal and separable. Remember that an algebraic extension F/K is separable if the minimal polynomial of every element in F is *separable* over K i.e., it has no multiple roots. Remember that we assumed that our field is of characteristic zero and therefore perfect. This assumption makes sense for us since we will take K to be an extension of $\mathbb{Q}$ later. An algebraic extension F/K is *normal* if every irreducible polynomial over K, which has a root in F, has all its roots in F. That is, it splits into linear factors over F. We see that the notions of an extension being separable or normal are perfectly well-defined for general algebraic extensions and not just finite ones. So if we relax the condition of being finite to being algebraic in the definition of a Galois extension, we get a definition that coincides with original one in the finite case. This motivates the following definition.

Definition 3.1. A field extension F/K is Galois if it is algebraic, normal and separable.

In this text, we will denote the Galois group of a Galois extension F/K by $\mathrm{Gal}(F/K)$. Just as with finite extensions, this is the collection of all automorphisms of F that fix K. The image of $\alpha \in K$ under the map $\sigma \in \mathrm{Gal}(F/K)$, we will denote by α^σ. We do the same for the image of a set under σ. It is immediate that the extension $\bar{K}/K$ is algebraic and normal. With our assumption that K is perfect, we have that the extension $\bar{K}/K$ is certainly separable. Therefore it is Galois and $\mathrm{Gal}(\bar{K}/K)$ will play a principal part in the rest of this thesis. Note that an arbitrary intersection of normal extension of K is again a normal extension of K and therefore the same holds for Galois extensions. This means that we can define *the Galois closure of an extension* F/K, which we denote by F^c, as the intersection of all Galois extensions of K that contain F. This is therefore the smallest Galois extension of K that contains F.

Lemma 3.2. *Consider the fields F, L such that $F/L/K$. If F/K is Galois, then F/L is also Galois.*

Proof. Suppose F/K is Galois and let $f(x) \in L[x]$ be irreducible. Certainly F/L is also algebraic and separable. Suppose that $f(x)$ has a root $\alpha \in F$, then $f(x)$ is, up to multiplication by a constant, the minimal polynomial of α over L. Let $m(x) \in K[x]$ be the minimal polynomial of α over K. Then $f(x)$ divides $m(x)$ in $L[x]$. Since $\alpha \in F$ and F/K normal, $m(x)$ splits into linear factors over F and therefore $f(x)$ does so too. □

Next we will show a crucial proposition which we will use a number of times throughout this thesis. It can be seen as an analogue to a special case theorem 8 of [5, section 13.1] for infinite extensions. This special case states that if $\alpha, \beta \in \bar{K}$ have the same minimal polynomial over K, then there is an isomorphism $\varphi\colon K(\alpha) \to K(\beta)$ with $\varphi(\alpha) = \beta$ and φ restricted to K is the identity.

Proposition 3.3. *Consider the fields F, L such that $F/L/K$ with F/K Galois. If $\sigma\colon L \to F$ is a homomorphism which fixes K, then there exists a isomorphism $\sigma^F\colon F \to F$ such that $\sigma^F\big|_L = \sigma$.*

Proof. Suppose $\sigma\colon L \to F$ is a homomorphism which fixes K. Let S be the set of pairs (M, σ_M) where $F/M/L$ and $\sigma_M\colon M \to F$ a homomorphism such that $\sigma^M\big|_L = \sigma$. For $(M_1, \sigma^{M_1}), (M_2, \sigma^{M_2}) \in S$, we write $(M_1, \sigma^{M_1}) \leq (M_2, \sigma^{M_2})$ if M_2/M_1 and $\sigma^{M_2}\big|_{M_1} = \sigma^{M_1}$. It is immediate that $(S, \leq)$ is a partially ordered set. We will now use Zorn's lemma to show that S has a maximal element.

We note that $(L, \sigma) \in S$, so the empty chain has an upper bound. Now let C be a non-empty chain in S. Consider $(M', \sigma^{M'}\colon M' \to F)$, where $M' = \bigcup_{(M,\sigma^M)\in C} M$ and $\sigma^{M'}(x) = \sigma^M(x)$ for any $(M, \sigma^M) \in C$ with

$x \in M$. Clearly $1 \in M'$, so $M' \neq \emptyset$. Let $x, y \in M'$, so $x \in M_x$ and $y \in M_y$ where $(M_x, \sigma^{M_x}), (M_y, \sigma^{M_y}) \in C$. Since C is a chain, assume without loss of generality that $M_y \leq M_x$. Then $x - y \in M_x$, so $x - y \in M'$. Assume without loss of generality that $y \in M'^{\times}$, then $xy^{-1} \in M_x$ so $xy^{-1} \in M'$. By the subfield criterion we can now conclude that M' is a subfield of F and we clearly have that $L \subseteq M'$. Next we show that $\sigma^{M'}$ is well-defined. Let $x \in M'$ and suppose that $x \in M_1$ and $x \in M_2$, where $(M_1, \sigma^{M_1}), (M_2, \sigma^{M_2}) \in C$. Assume without loss of generality that $(M_1, \sigma^{M_1}) \leq (M_2, \sigma^{M_2})$. Then $\sigma^{M_2}\big|_{M_1} = \sigma^{M_1}$, so $\sigma^{M_2}(x) = \sigma^{M_1}(x)$. This means that $\sigma^{M'}$ is well-defined. In the same way we can see that $\sigma^{M'}$ is a homomorphism and it extends σ. This shows that $(M', \sigma^{M'}) \in S$. For each $(M, \sigma^M) \in C$, we now immediately get that $(M, \sigma^M) \leq (M', \sigma^{M'})$. This shows that $(M', \sigma^{M'})$ is an upper bound for C. By Zorn's lemma we can conclude that S has a maximal element. For now we will also denote this maximal element by $(M', \sigma^{M'})$ and we note that by definition, $\sigma^{M'}$ is a field homomorphism from M' into F.

Suppose that $M' \subsetneq F$, so we have $\alpha \in F \setminus M'$. But by assumption F/K is Galois, so by lemma 3.2 F/M' is Galois. Therefore F/M' is algebraic and this means that α has a minimal polynomial $m(x) \in M'[x]$. But $\sigma^{M'}$ still fixes K, so $\mathrm{im}\sigma^{M'} \neq 0$. Therefore $\ker \sigma^{M'} \neq M'$ an so $\ker \sigma^{M'} = 0$. But this means that $\sigma^{M'}$ is injective and an isomorphism if we reduce the codomain to $\sigma^{M'}(M')$. Now by theorem 8 in [5, section 13.1], we can extend $\sigma^{M'}$ to $\sigma^{M'(\alpha)} \colon M'(\alpha) \to F$ since F/K is normal and therefore contains all other roots of $m(x)$. But this contradicts the maximality of $(M', \sigma^{M'})$, so we conclude that $M' = F$. We will from now on denote $\sigma^{M'}$ by σ^F.

Lastly we show that σ^F is an isomorphism. We already noted that σ^F is injective. Let $\alpha \in F$. Then α is algebraic over K so therefore α has a minimal polynomial $m(x) \in K[x]$ of degree $d \geq 1$. Minimal polynomials are irreducible, so by normality F contains all d roots of $m(x)$. These roots are distinct by separability. Since σ^F fixes K, α^{σ^F} is still a root of $m(x)$. Since d is finite and σ^F injective, σ^F is bijective on the set of d roots of $m(x)$. Therefore there exists $\beta \in F$ such that $\beta^{\sigma^F} = \alpha$. This shows that σ^F is surjective and so we conclude that $\sigma^F \colon F \to F$ is an isomorphism, which by definition extends σ. □

Lemma 3.4. *Consider the fields F, L such that $F/L/K$ with F/K Galois. The extension L/K is Galois if and only if for all $\sigma \in \mathrm{Gal}(F/K)$ we have that $L^\sigma \subseteq L$.*

Proof. Suppose first that L/K is Galois. Let $\sigma \in \mathrm{Gal}(F/K)$ and $\alpha \in L$. Since L is algebraic over K, there exists a minimal polynomial $m(x) \in K[x]$ of α. Since σ fixes K, α^σ is also a root of $m(x)$. But $m(x)$ is irreducible and L/K normal, so L contains all roots of $m(x)$. This means that $\alpha^\sigma \in L$, so we conclude that $L^\sigma \subseteq L$ for all $\sigma \in \mathrm{Gal}(F/K)$.

Now suppose that $L^\sigma \subseteq L$ for all $\sigma \in \mathrm{Gal}(F/K)$. We immediately have that L/K is algebraic and separable, so we only need to show normality. Let $f(x) \in K[x]$ be irreducible and suppose that $\alpha \in L$ is a root of $f(x)$. Then $\alpha \in F$ and F/K is normal, so F contains all other roots of $f(x)$. Let β be such a root. We can use theorem 8 in [5, section 13.1] to extend id_K to an isomorphism $\sigma' \colon K(\alpha) \to K(\beta)$, with $\alpha^{\sigma'} = \beta$. We can increase the codomain of σ' to F and use proposition 3.3 to extend σ' to an isomorphism $\sigma \colon F \to F$. Since σ clearly still fixes K, we have that $\sigma \in \mathrm{Gal}(F/K)$. We can then use our assumption to conclude that $\beta = \alpha^{\sigma'} = \alpha^\sigma \in L$. This shows that L/K is normal and therefore Galois. □

Theorem 3.5. *Consider the fields F, L such that $F/L/K$ with both F/K and L/K Galois. Then*

$$\mathrm{Gal}(F/K)/\mathrm{Gal}(F/L) \cong \mathrm{Gal}(L/K).$$

Proof. We consider the map $r \colon \mathrm{Gal}(F/K) \to \mathrm{Gal}(L/K) \colon \sigma \mapsto \sigma|_L$. To see that this map is well-defined, we have to check that its image is actually contained in $\mathrm{Gal}(L/K)$. Let $\sigma \in \mathrm{Gal}(F/K)$. Then $r(\sigma)$ is certainly still a homomorphism and fixes K. Since L/K is Galois, we can use lemma 3.4 to say that $L^\sigma \subseteq L$ and therefore it makes sense to restrict the codomain of $r(\sigma)$ to L. Of course, $r(\sigma)$ is still injective. Now let $\alpha \in L$. We proceed as in the last paragraph of the proof of proposition 3.3. We know that α has a minimal polynomial $m(x) \in K[x]$ of degree $d \geq 1$. Since L/K is Galois, so both normal and separable, L contains all d distinct roots of $m(x)$. We know that $r(\sigma)$ fixes K, so $\alpha^{r(\sigma)}$ is still a root of $m(x)$. But $r(\sigma)$ is still injective,

so it permutes the roots of $m(x)$. But an injection from a finite set to itself is certainly surjective, therefore there exists $\beta \in L$ such that $\beta^{r(\sigma)} = \alpha$. This shows that $r(\sigma)$ is surjective and therefore $r(\sigma) \in \mathrm{Gal}(L/K)$. We conclude that r is well-defined.

By construction, it is immediate that r is a homomorphism. Furthermore, we see that $\sigma \in \ker r$ if and only if $\sigma|_L = id_L$, if and only if σ fixes L, if and only if $\sigma \in \mathrm{Gal}(F/L)$. Note that F/L is Galois by lemma 3.2. Therefore $\ker r = \mathrm{Gal}(F/L)$. Now let $\sigma' \in \mathrm{Gal}(L/K)$. We can expand the codomain to F and use proposition 3.3 to extend σ' to an automorphism σ of F. This still fixes K, so $\sigma \in \mathrm{Gal}(F/K)$. Certainly $r(\sigma) = \sigma'$, so we conclude that r is surjective. Now the first isomorphism theorem tells us that

$$\mathrm{Gal}(F/K)/\mathrm{Gal}(F/L) \cong \mathrm{Gal}(L/K). \qquad \square$$

For a field extension F/K that is Galois and $G \leq \mathrm{Gal}(F/K)$, we write F_G for the fixed field of G.

Theorem 3.6. *Let F/K be some Galois extension, then $F_{\mathrm{Gal}(F/K)} = K$.*

Proof. It is immediate that $K \subseteq F_{\mathrm{Gal}(F/K)}$, since every element of $\mathrm{Gal}(F/K)$ fixes K. Let $\alpha \in F \setminus K$ and $m(x) \in K[x]$ the minimal polynomial of α with degree $d \geq 1$. If $d = 1$, $\alpha \in K$ which is a contradiction and therefore $d \geq 2$. Since the extension is separable and normal, there exists $\beta \in K$ such that $m(\beta) = 0$ and $\beta \neq \alpha$. By theorem 8 of [5, section 13.1], we can extend id_K to an isomorphism $\sigma' \colon K(\alpha) \to K(\beta)$ with $\sigma'(\alpha) = \beta$. We know that $K(\beta) \subseteq F$, so we can use proposition 3.3 to extend σ' to an isomorphism $\sigma \colon F \to F$. This still fixes K, so $\sigma \in \mathrm{Gal}(F/K)$, but also $\alpha^\sigma = \beta \neq \alpha$. This means that $\alpha \notin F_{\mathrm{Gal}(F/K)}$. The contrapositive gives us that if $\alpha \in F_{\mathrm{Gal}(F/K)}$, then $\alpha \in K$. We can now conclude that $F_{\mathrm{Gal}(F/K)} = K$. $\square$

3.2 First results on elliptic curves

From now on we will take $A, B \in K$ and say that the curve $E(\bar{K})$ of $W_{A,B}$ is *defined over* K. We will mainly be interested in a particular kind of subset of $E(\bar{K})$.

Definition 3.7. Let F be a field such that $\bar{K}/F/K$. We define

$$E(F) := \{P \in E(\bar{K}) \mid P = [x : y : 1] \text{ with } x, y \in F\} \cup \{O\}.$$

Note that if $F = \bar{K}$, this definition coincides with our previous definition of $E(\bar{K})$. Therefore there will be no ambiguity if we write $E(\bar{K})$.

Lemma 3.8. *Let F be a field such that $\bar{K}/F/K$. Then $E(F)$ is a subgroup of $E(\bar{K})$.*

Proof. First we note that $O \in E(F)$, so $E(F) \neq \emptyset$. We will show that $E(F)$ is closed under opposites. Suppose that $P \in E(F)$. Since $-O = O$ we can assume that $P \neq O$. Then we know that $P = [x : y : 1]$ and $-P = [x : -y : 1]$ with $x, y \in F$. But then certainly $-y \in F$, so $-P \in E(F)$. Now we only need to show that $E(F)$ is closed under addition. Let $Q \in E(F)$. If $Q = O$ or $Q = -P$, we are done. If $Q = P$ and $P + P = O$ we are also done. If $Q = P$ and $P + P \neq O$ we see in the formula in theorem 2.8 that both coordinates of $P + P$ can be expressed as rational combination of x, y and A. Since F is a field, these rational combinations are again element of F, so $P + P \in E(F)$. Note that here we have used that $A \in K \subseteq F$. If $Q \neq O, P, -P$, write $Q = [x' : y' : 1]$. We can use the other formula from theorem 2.8 to see that both coordinates of $P + Q$ are rational combinations of x, x', y and y'. Since all of these are contained in F, we again have that $P + Q \in E(F)$. We have exhausted all possible cases, so $E(F)$ is closed under addition. We now conclude that $E(F)$ is a subgroup of $E(\bar{K})$. $\square$

For a field F with $\bar{K}/F/K$, we call $E(F)$ the *group of F-rational points*. From now on we take K to be a number field. Number fields are finite field extensions of $\mathbb{Q}$ so, as we have noted before, they are perfect. We are now ready to state the main result of this thesis.

Theorem 3.9 (Mordell-Weil)**.** *The group $E(K)$ is abelian and finitely generated i.e,*

$$E(\mathbb{Q}) \cong \mathbb{Z}^r \oplus \mathbb{Z}/a_1\mathbb{Z} \oplus \cdots \oplus \mathbb{Z}/a_n\mathbb{Z},$$

for $r \geq 0$ and $a_0, \ldots, a_n \geq 1$.

The rest of this thesis is dedicated to proving this fact. We note here that the theorem does not hold for $E(\bar{K}$. In fact, we can show that group is never finitely generated! This will be done in corollary 3.14. We will now show that $E(\mathbb{C})$ also cannot be finitely generated. Here we take $W_{A,B} \in \mathbb{C}[X, Y, Z]$. Consider $z \in \mathbb{C}$. Since $\mathbb{C}$ is algebraically closed, there exist $x, y \in \mathbb{C}$ such that $y^2 = z = x^3 + Ax + B$. Then we see that $[x : y : 1] \in E(\mathbb{C})$. For different values for z, we certainly get different points. Since $\mathbb{C}$ is uncountable, we see that $E(\mathbb{C})$ is uncountable. But a finitely generated abelian group is certainly countable since it is isomorphic to a cartesian product of copies of $\mathbb{Z}$ and some finite sets. This shows that $E(\mathbb{C})$ can never be finitely generated. The Galois group $\mathrm{Gal}(\bar{K}/K)$ has a very natural action on the points, as is described in the next proposition.

Proposition 3.10. *For each $\sigma \in \mathrm{Gal}(\bar{K}/K)$, the map $-^\sigma \colon E(\bar{K}) \to E(\bar{K}) \colon [x : y : z] \mapsto [x^\sigma : y^\sigma : z^\sigma]$ is a group automorphism. (The notation $-^\sigma$ means that we write P^σ for the image of the point P under this map)*

Proof. Let $\sigma \in \mathrm{Gal}(\bar{K}/K)$. First we show that the action is well-defined. To do this, we show that it is does not depend on the representative of the point in projective space and that the new point is again an element of $E(\bar{K})$. Suppose that $[x : y : z] = [x' : y' : z'] \in E(\bar{K})$. Then $x' = \lambda x$, $y' = \lambda y$ and $z' = \lambda z$ for some $\lambda \in \bar{K}^\times$. Then also $\lambda^\sigma \in \bar{K}^\times$. Using this we get that

$$[x' : y' : z']^\sigma = [(\lambda x)^\sigma : (\lambda y)^\sigma : (\lambda z)^\sigma] = [\lambda^\sigma x^\sigma : \lambda^\sigma y^\sigma : \lambda^\sigma z^\sigma] = [x^\sigma : y^\sigma : z^\sigma] = [x : y : z]^\sigma.$$

This means that the action does not depend on the representative. Now we just have to show that $[x : y : z]^\sigma \in E(\bar{K})$. Since $A, B \in K$, they are fixed by $\sigma \in \mathrm{Gal}(\bar{K}/K)$. Using this we get that

$$\begin{aligned} W_{A,B}(x^\sigma, y^\sigma, z^\sigma) =& -(y^\sigma)^2 z^\sigma + (x^\sigma)^3 + Ax^\sigma (z^\sigma)^2 + B(z^\sigma)^3 \\ =& -(y^\sigma)^2 z^\sigma + (x^\sigma)^3 + A^\sigma x^\sigma (z^\sigma)^2 + B^\sigma (z^\sigma)^3 \\ =& (-y^2 z + x^3 + Axz^2 + Bz^3)^\sigma \\ =& 0^\sigma \\ =& 0. \end{aligned}$$

This shows that $[x : y : z]^\sigma \in E(\bar{K})$ and therefore the action is well-defined. Next note that for all $P \in E(\bar{K})$, $(P^\sigma)^{\sigma^{-1}} = P$ so the map is invertible.

Now we just have to show that the map respects the addition on the curve. Let $P, Q \in E(\bar{K})$ and write $P = [x_p : y_p : z_p]$ and $Q = [x_q : y_q : z_q]$, with $x_p, y_p, z_p, x_q, y_q, z_q \in \bar{K}$. Define $\ell \in \bar{K}[X, Y, Z]$ as $aX + bY + cZ$, such that $V(\ell)$ is the line through P and Q. Remember that if $P = Q$, then we have defined the line through P and Q as the tangent line at P. We write ℓ^σ for the polynomial $a^\sigma X + b^\sigma Y + c^\sigma Z \in \bar{K}[X, Y, Z]$. Since $P, Q \in V(\ell)$, we see that $P^\sigma, Q^\sigma \in V(\ell^\sigma)$. Note that if $P^\sigma = Q^\sigma$, then $P = Q$. This shows that $V(\ell^\sigma)$ is well-defined as the line through P^σ and Q^σ. The point $L(P, Q)$ lies on the line through P and Q, so it is a root of ℓ. Then in the same way as before, $L(P, Q)^\sigma$ is a root of ℓ^σ. This shows that $L(P, Q)^\sigma \in V(\ell^\sigma) \cap E(\bar{K}) = \{P^\sigma, Q^\sigma, L(P^\sigma, Q^\sigma)\}$. We now wish to show that $L(P, Q)^\sigma = L(P^\sigma, Q^\sigma)$. For this we consider three cases:

1. Suppose that $P \neq L(P, Q) \neq Q$. Suppose that $L(P, Q)^\sigma = P^\sigma$ or $L(P, Q)^\sigma = Q^\sigma$. By injectivity, then $L(P, Q) = P$ or $L(P, Q) = Q$ which is a contradiction. We conclude that $L(P, Q)^\sigma = L(P^\sigma, Q^\sigma)$.

2. Suppose that $L(P, Q) = P = Q$. Suppose that $L(P^\sigma, Q^\sigma) \neq P^\sigma$, then $L(P^\sigma, Q^\sigma)^{\sigma^{-1}} \neq (P^{\sigma^{-1}})^\sigma = P$. But $L(P^\sigma, Q^\sigma)$ certainly lies on the line through P^σ and Q^σ and therefore $L(P^\sigma, Q^\sigma)^{\sigma^{-1}}$ lies on the line through P and Q. But our first assumption then tells us that $L(P^\sigma, Q^\sigma)^{\sigma^{-1}} = P$, which is a contradiction. Therefore $L(P^\sigma, Q^\sigma) = P^\sigma = L(P, Q)^\sigma$.

3. Suppose that $L(P,Q)$ is equal to either P or Q, but not both. Without loss of generality, assume that $L(P,Q) = P$. Then $L(P,P) = Q$ and therefore $V(\ell)$ is the tangent line to $E(\bar{K})$ at P. In particular we see that
$$V(\ell) = V\left(\frac{\partial W_{A,B}}{\partial X}(P)X + \frac{\partial W_{A,B}}{\partial Y}(P)Y + \frac{\partial W_{A,B}}{\partial Z}(P)Z\right).$$
The polynomial for the tangent line is defined only up to a constant, since all the partial derivatives are homogeneous of degree two. The set of roots is therefore well-defined. Since $E(\bar{K})$ is defined over K, the coefficients of the partial derivatives will certainly be elements of K. For a fixed representative of P, we then get that
$$\frac{\partial W_{A,B}}{\partial X}(P^\sigma) = \left(\frac{\partial W_{A,B}}{\partial X}(P)\right)^\sigma, \frac{\partial W_{A,B}}{\partial Y}(P^\sigma) = \left(\frac{\partial W_{A,B}}{\partial Y}(P)\right)^\sigma \text{ and } \frac{\partial W_{A,B}}{\partial Z}(P^\sigma) = \left(\frac{\partial W_{A,B}}{\partial Z}(P)\right)^\sigma.$$
This then means that
$$V(\ell^\sigma) = V\left(\frac{\partial W_{A,B}}{\partial X}(P^\sigma)X + \frac{\partial W_{A,B}}{\partial Y}(P^\sigma)Y + \frac{\partial W_{A,B}}{\partial Z}(P^\sigma)Z\right),$$
so $V(\ell^\sigma)$ is the tangent line of $E(\bar{K})$ at P^σ. Since P^σ and $Q^\sigma \in V(\ell^\sigma)$, we then see that $L(P^\sigma, Q^\sigma) = P^\sigma = L(P,Q)^\sigma$.

We conclude that $L(P,Q)^\sigma = L(P^\sigma, Q^\sigma)$. A direct calculation shows that $O^\sigma = O$. Now we can finish the proof by seeing that
$$(P+Q)^\sigma = L(O, L(P,Q))^\sigma = L(O^\sigma, L(P,Q)^\sigma) = L(O, L(P^\sigma, Q^\sigma)) = P^\sigma + Q^\sigma.$$

□

From this point on, let $m \in \mathbb{Z}_{\geq 2}$. We know that being an abelian group is equivalent to being a $\mathbb{Z}$-module. So for some $P \in E(\bar{K})$, we can write
$$mP = \underbrace{P + \cdots + P}_{m \text{ terms}}.$$
With this notation, we can now define some objects which will be very important to us later on. For a field F such that $\bar{K}/F/K$, we write $mE(F)$ for the subgroup of m-multiples.

Definition 3.11. Consider the homomorphism $m\colon E(\bar{K}) \to E(\bar{K})\colon P \mapsto mP$. We define $E(\bar{K})[m] := \ker m$. We call this the *m-torsion subgroup of $E(\bar{K})$*.

We will show that this subgroup is actually finite, as this will be very important later on. For example in the proof of the first item of lemma 6.8. To do this, we first show a theorem which we will also use again later, namely in 6.4. We will use some key facts from algebraic geometry to do this.

Theorem 3.12. *For all $P \in E(\bar{K})$, there exists $Q \in E(\bar{K})$ such that $P = mQ$.*

Proof. With the formulas from theorem 2.8, it is clear that $[m]\colon P \mapsto mP$ is a rational map from $E(\bar{K})$ to $E(\bar{K})$. Since $E(\bar{K})$ is by assumption smooth, this map must be a morphism ([1, section II.2, proposition 2.1]). But any non-constant morphism between curves is surjective (proposition 6.8 in [6, section II]), so it just remains to show that the map $[m]$ is non-constant. First note that $mO = O$, so if $[m]$ were constant, it would map the whole curve to O. Now we see that for any $n \in \mathbb{Z}_{\geq 1}$, $[mn] = [m] \circ [n]$. Therefore if we prove that $[\pi]$ is non-constant for every prime number $\pi \in \mathbb{Z}_{>1}$, we are done.

First we consider the case $\pi = 2$. Suppose that $P \in E(\bar{K})$ such that $P \neq O$. Then we can write $P = [x : y : 1]$ with $x, y \in \bar{K}$. Now suppose that $2P = O$. Then $P = -P$, so with the formula from theorem 2.8 we get that $y = -y$ and therefore $y = 0$. Then x has to satisfy the equation $X^3 + AX + B = 0$, so this shows that there are at most 3 points of exact order 2. So if $[2]$ were constant, there could lie no more than 4 points on the curve. But for each $y \in \bar{K}$, there exist an $x \in \bar{K}$ such that (x,y) satisfies $W_{A,B}(X,Y,1) = 0$ since $\bar{K}$ is algebraically closed. And since $\bar{K}$ is infinite, this means that $E(\bar{K})$ is infinite. Therefore $[2]$ is not constant.

Now let $x \in \bar{K}$ such that $x^3 + Ax + B = 0$. This x exists since $\bar{K}$ is algebraically closed. As we have noted above the point $P = [x : 0 : 1] \in E(\bar{K})$ and $2P = O$. Suppose that π is a prime greater than 2. Then π is odd, so $\pi P = \frac{\pi-1}{2}(2P) + P = P$. But $P \neq O$, so $[\pi]$ is not constant. We now conclude that $[m]$ is not constant for any $m \geq 2$, so as we have noted before, $[m]$ is surjective. □

Corollary 3.13. *The set $E(\bar{K})[m]$ is a finite subgroup of $E(\bar{K})$.*

Proof. In the proof of theorem 3.12 we have already seen that $E(\bar{K})[2]$ is finite, so supppose that $m \geq 3$. Using the formulas in theorem 2.8 we define the following sequences of rational functions with coefficients in $\bar{K}$:

$$S_2^X(X,Y) := \frac{(3X^2 + A)^2 - 8XY^2}{4Y^2},$$
$$S_2^Y(X,Y) := \frac{4XY^2(3X^2 + A) - 8Y^4 S_2^X - 8Y^4}{8Y^3},$$
and
$$S_{n+1}^X(X,Y) := \frac{(S_n^Y - Y)^2 - (X + S_n^X)((S_n^X - X)^2)^2}{(S_n^X - X)^2},$$
$$S_{n+1}^Y(X,Y) := \frac{X(S_n^Y - Y)(S_n^X - X)^2 - S_n^X(S_n^X - X)^2 - Y(S_n^X - X)^3}{(S_n^X - X)^3},$$

For $n \geq 3$. We have constructed this sequence such that for $P = [x : y : 1] \in E(\bar{K})$, $nP = [S_n^X(x,y) : S_n^y(x,y) : 1]$ if $kP \neq O$ for all $k \in \mathbb{Z}$ with $1 \leq k \leq n$. Let $P = [x : y : 1] \in E(\bar{K})$ and suppose that $mP = O$ and $nP \neq O$ for all $n \in \mathbb{Z}$ such that $1 \leq n < m$. Then $(m-1)P = -P$ and so (x,y) is a solution of $S_{m-1}^X(X,Y) - X = 0$. But if we clear the denominator of $S_{m-1}^X(X,Y)$ we see that (x,y) also has to be a root of some polynomial $h(X,Y) \in K[X,Y]$. Suppose that $W_{A,B}$ divides h^*, where h^* is the homogenisation of h^*, so $[x : y : 1]$ is a root of h^*. Then any $Q = [x' : y' : 1] \in E(\bar{K})$ also is a root of h^*, so in particular $S_{m-1}^X(x',y') = x'$ or the denominator of $S_{m-1}^X(x',y')$ is zero. But then either $mQ = O$ or $nQ = O$ for some $1 \leq n < m$. Then every point on the curve would have order less than or equal to m. But by theorem 3.12, we can choose a prime number $\pi > m$ and $Q' \in E(\bar{K})$ such that $\pi Q' = O$. Since π is prime, Q' has order bigger than m so we get a contradiction. Now we see that $W_{A,B}$ does not divide h^*. Then $W_{A,B}$ and h do not share a common irreducible component, since $W_{A,B}$ is irreducible and now it follows from theorem 2.2 that $V(W_{A,B}) \cap V(h^*)$ is finite. We see that there are only finitely many $Q \in E(\bar{K})$ with $mQ = O$ and $nQ \neq O$ for all $n \in \mathbb{Z}$ such that $1 \leq n < m$. Since this holds for all $m \geq 3$, we can now (technically using induction) conclude that $E(\bar{K})[m]$ is finite. □

We now present a corollary which shows that $E(\bar{K})$ is never finitely generated. This again motivates us to only consider the smaller group $E(K)$.

Corollary 3.14. *$E(\bar{K})$ is not finitely generated.*

Proof. Suppose that $E(\bar{K})$ is finitely generated. Then we know that

$$E(\bar{K}) \cong \mathbb{Z}^r \oplus \mathbb{Z}/a_1\mathbb{Z} \oplus \cdots \oplus \mathbb{Z}/a_n\mathbb{Z},$$

for $r \geq 0$ and $a_1, \ldots, a_n \geq 1$. But then the points with finite order are exactly those that have their first r coordinates equal to zero under a chosen isomorphism. In particular, all points of finite order, have an order dividing $a_1 \ldots a_n$. Let π be a prime number greater than $a_1 \ldots a_n$, so theorem 3.12 now tells us that there exists $Q \in E(\bar{K})$ such that $\pi Q = O$. Since π is prime, it is the order of Q. But then π would divide $a_1 \ldots a_n$, which is a contradiction. We conclude that $E(\bar{K})$ is not finitely generated. □

4 The method of descent

We will now state a criterion to check if an abelian group is finite. Since abelian groups are just $\mathbb{Z}$-modules, we present a present a proof in the more general case.

Theorem 4.1. *Let R be a commutative ring with $1 \neq 0$ and M an R-module. Suppose we have maps $h\colon M \to \mathbb{R}$ and $\varphi\colon R \to \mathbb{R}$ satisfying:*

1. *There exists $x \in \mathbb{R}$, such that for all $m \in M$ there exists $c_m \in \mathbb{R}$, such that $h(m+m') \leq xh(m') + c_m$ for all $m' \in M$.*
2. *There exists $c \in \mathbb{R}$ and $r \in R$, such that M/rM is finite and $h(rm) \geq \varphi(r)h(m) + c$ for all $m \in M$.*
3. *For all $y \in \mathbb{R}$, we have that $\{m \in M \mid h(m) \leq y\}$ is finite.*

Assume furthermore that $0 < x < \varphi(r)$. Then M is finitely generated as an R-module.

Proof. By assumption 2, we can write $M/rM = \{[m_1], \dots, [m_k]\}$. Let $n_0 \in M$, then $n_0 \in [m_{i_1}]$ for exactly one $1 \leq i_1 \leq k$. We can therefore write $n_0 = rn_1 + m_{i_1}$ for some $n_1 \in M$. We can now do the same for n_1 and write $n_1 = rn_2 + m_{i_2}$, for some $n_2 \in M$ and $1 \leq i_2 \leq k$. Nothing is stopping us from iterating this process, say $l \in \mathbb{Z}_{\geq 1}$ times.

We will now show by induction that $n_0 = r^l n_l + \sum_{j=1}^{l} r^{j-1} m_{i_j}$, for all $l \in \mathbb{Z}_{\geq 1}$. The base case $l = 1$ is immediate. Now suppose that it holds for some $l' \in \mathbb{Z}_{\geq 1}$. Then $n_0 = r^{l'} n_{l'} + \sum_{j=1}^{l'} r^{j-1} m_{i_j}$. We write $n_{l'} = rn_{l'+1} + m_{i_{l'+1}}$ and plug this in. This gives us $n_0 = r^{l'}(rn_{l'+1} + m_{i_{l'+1}}) + \sum_{j=1}^{l'} r^{j-1} m_{i_j} = r^{l'+1} n_{l'+1} + \sum_{j=1}^{l'+1} r^{j-1} m_{i_j}$. By induction, it follows that the formula holds for all $l \in \mathbb{Z}_{\geq 1}$. This shows that n_0 is in the submodule generated by $\{n_l, m_1, \dots, m_k\}$.

Since $\varphi(r) > 0$, we can use the second property to say that $h(n_l) \leq \frac{(h(rn_l)-c)}{\varphi(r)}$. With property 1 we have that

$$\frac{h(rn_l) - c}{\varphi(r)} = \frac{h(n_{l-1} - m_{i_l}) - c}{\varphi(r)} \leq \frac{xh(n_{l-1}) + c_l - c}{\varphi(r)},$$

where c_l is the constant in property 1 for $-m_{i_l}$. Define C as the maximum of these constants for all $m \in \{m_1, \dots, m_k\}$. We can then conclude that $h(n_l) \leq \frac{xh(n_{l-1})+C-c}{\varphi(r)}$. Note here that neither C nor c depend on the n_i. In a similar way as before, we can use induction to prove that

$$h(n_l) \leq \left(\frac{x}{\varphi(r)}\right)^l h(n_0) + \frac{C-c}{\varphi(r)} \sum_{i=0}^{l-1} \left(\frac{x}{\varphi(r)}\right)^i = \left(\frac{x}{\varphi(r)}\right)^l h(n_0) + (C-c)\frac{1 - (\frac{x}{\varphi(r)})^l}{\varphi(r) - x}.$$

The last equality follows immediately if we recognise the sum as a geometric series. Since $0 < x < \varphi(r)$, we have that $0 < 1 - (\frac{x}{\varphi(r)})^l < 1$ and $0 < \varphi(r) - x$. We don't know the sign of $C - c$, but we can still say that

$$\left(\frac{x}{\varphi(r)}\right)^l h(n_0) + (C-c)\frac{1 - (\frac{x}{\varphi(r)})^l}{\varphi(r) - x} \leq \left(\frac{x}{\varphi(r)}\right)^l h(n_0) + \frac{|C-c|}{\varphi(r) - x}.$$

Since $0 < \frac{x}{\varphi(r)} < 1$, we can choose l large enough, such that

$$\left(\frac{x}{\varphi(r)}\right)^l h(n_0) + \frac{|C-c|}{\varphi(r) - x} < 1 + \frac{|C-c|}{\varphi(r) - x}.$$

This shows that for l large enough, $n_l \in \left\{ m \in M \;\middle|\; h(m) \leq 1 + \frac{|C-c|}{\varphi(r)-x} \right\} =: S$. With what we said before, it follows that M is generated by $S \cup \{m_1, \dots, m_k\}$. But property 3 tells us that S is finite, so M is finitely generated. □

We will use this theorem in the special case were $R = \mathbb{Z}$. If a function h satisfies the constraints of the theorem, we call it a *height function*.

5 A conditional proof of the Mordell-Weil theorem for the case $K = \mathbb{Q}$

We will now give explicit functions $\varphi\colon \mathbb{Z} \to \mathbb{R}$ and $h\colon E(\mathbb{Q}) \to \mathbb{R}$ and show that they satisfy the constraints of theorem 4.1.

Definition 5.1. First we consider the *height of a rational number*. This is defined by

$$H\colon \mathbb{Q} \to \mathbb{Z}_{>0}\colon \frac{p}{q} \mapsto \frac{\max\{|p|,|q|\}}{\gcd(p,q)}.$$

We use this to define the *height of a point* $P \in E(\mathbb{Q})$ as

$$h\colon E(\mathbb{Q}) \to \mathbb{R}\colon P \mapsto \begin{cases} \log_b(H(x(P))) & \text{if } P \neq O \\ 0 & \text{if } P = O. \end{cases}$$

For the base of the logarithm one can choose any $b \in \mathbb{R}_{>1}$. Note that if $P \neq O$, we have that P is not on the hyperplane at infinity so it has a representative with last coordinate equal to 1. The notation $x(P)$ then means that we take the first coordinate of this representative.

To see that our definition of the height of a rational number makes sense, we have to show that it does not depend on the representative of the number. Let $\frac{p}{q} \in \mathbb{Q}$ with $\frac{p}{q} = \frac{p'}{q'} \in \mathbb{Q}$ such that $\gcd(p',q') = 1$. Then

$$H\left(\frac{p}{q}\right) = \frac{\max\{|p|,|q|\}}{\gcd(p,q)} = \frac{\max\{\gcd(p,q)|p'|, \gcd(p,q)|q'|\}}{\gcd(p,q)} = \max\{|p'|,|q'|\} = H\left(\frac{p'}{q'}\right),$$

since the absolute value respects multiplication and the greatest common divisor is positive. This shows that the height of every representative is the same as the height of the representative in lowest terms, so H is well-defined. The logarithm in the definition of the height of a point in $E(\mathbb{Q})$ might seem a bit arbitrary at this point, but it will certainly be of use. For example at the end of the proof of part 1 of the next theorem, we need the logarithm to transform multiplication into addition. Under the assumption that $E(\mathbb{Q})/2E(\mathbb{Q})$ is finite, we will now prove our main theorem 3.9. As it turns out, proving this assumption will be the hardest part of the Mordell-Weil theorem and we will do so in section 6.

Theorem 5.2 (Mordell-Weil theorem for $\mathbb{Q}$)**.** *Assume $E(\mathbb{Q})/2E(\mathbb{Q})$ is finite. The functions h from 5.1 and $\varphi\colon \mathbb{Z} \to \mathbb{R}\colon n \mapsto n^2$, together with the choices $x = r = 2$, satisfy the three conditions of 4.1. Therefore $E(\mathbb{Q})$ is finitely generated.*

Proof. If we fill in our choices for h, φ, x and r the three conditions of 4.1 we have to verify, become:

1. For all $P \in E(\mathbb{Q})$, there exists $c_P \in \mathbb{R}$, such that $h(P+Q) \leq 2h(Q) + c_P$ for all $Q \in E(\mathbb{Q})$.
2. There exists $c \in \mathbb{R}$, such that $h(2P) \geq 4h(P) + c$ for all $P \in E(\mathbb{Q})$.
3. For all $y \in \mathbb{R}$, we have that $\{P \in E(\mathbb{Q}) \mid h(P) \leq y\}$ is finite.

We will check these one by one:

1. Let $P \in E(\mathbb{Q})$. Let $Q \in E(\mathbb{Q})$. Since the denominator of a rational number is never zero, its height is at least one. The height of a point on $E(\mathbb{Q})$ is therefore always non-negative. So if we make sure that $c_P \geq 0$, then condition 1 holds for $P = O$. If $Q = -P$, the non-negativity of the height function and the condition $c_P \geq 0$ make sure that condition 1 hold. By setting $c_P \geq h(P)$, we have dealt with the case $Q = O$. The condition $c_P \geq h(2P)$ takes care of the case $Q = P$. To conclude, we need to make sure that $c_P \geq \max\{h(P), h(2P)\}$. Note that $c_P \geq 0$ follows immediately from this.

   Now assume that $P \neq O$ and $Q \neq O, P, -P$. We write $P = [\frac{x_1}{x_2} : \frac{y_1}{y_2} : 1]$, with $\gcd(x_1,x_2) = \gcd(y_1,y_2) = 1$ and $x_2, y_2 > 0$. Plugging this into the Weierstrass equation gives

   $$\frac{y_1^2}{y_2^2} = \frac{x_1^3 + Ax_1x_2^2 + Bx_2^3}{x_2^3}.$$

Note that if $\gcd(y_1^2, y_2^2) > 1$, then also $\gcd(y_1, y_2) > 1$. Therefore $\gcd(y_1^2, y_2^2) = 1$. With basic rules for the gcd, we get that $\gcd(x_1^3 + (Ax_1x_2 + Bx_2^2)x_2, x_2) = \gcd(x_1^3, x_2) = 1$ since x_1 and x_2 have no primes in common. But then also $\gcd(x_1^3 + (Ax_1x_2 + Bx_2^2)x_2, x_2^3) = 1$, since we do not introduce any new primes. Now we have shown that two fractions are equal, are both in lowest terms and the denominators have by assumption the same sign. We can then conclude that the denominators are actually equal, so $y_2^2 = x_2^3 = s$. So s is a perfect square, so each prime in its decomposition appears an even number of times. But on the other hand, s is also a perfect cube so the number of times a prime appears in its decomposition is also divisible by three. But two and three are coprime so each prime in its decomposition is also divisible by six. This means that we can write $s = d^6$, so $y_2 = d^3$ and $x_2 = d^2$. Now we have that $P = [\frac{x_1}{d^2} : \frac{y_1}{d^3} : 1]$. Since $Q \neq O$, we can do the exact same thing and get $Q = [\frac{x_3}{e^2} : \frac{y_3}{e^3} : 1]$ with $\gcd(x_3, e^2) = \gcd(y_3, e^3) = 1$ and $e^2, e^3 > 0$. $h(Q) = \max\{|e|^2, |x_3|\}$

With our assumptions, we can use the first formula from theorem 2.8 for the x-coordinate of $P + Q$. We get that

$$\begin{aligned} x(P+Q) &= \left(\frac{y(Q) - y(P)}{x(Q) - x(P)}\right)^2 - x(Q) - x(P) \\ &= \frac{y(Q)^2 - 2y(P)y(Q) + y(P)^2}{(x(Q) - x(P))^2} - \frac{x(Q)(x(Q) - x(P))^2 + x(P)(x(Q) - x(P))^2}{(x(Q) - x(P))^2}. \end{aligned}$$

Now we can use the fact that both P and Q lie on E and therefore satisfy the Weierstrass equation. Substituting $y(P)^2$ and $y(Q)^2$ in the last expression gives

$$\begin{aligned} &\frac{y(Q)^2 - 2y(P)y(Q) + y(P)^2}{(x(Q) - x(P))^2} - \frac{x(Q)(x(Q) - x(P))^2 + x(P)(x(Q) - x(P))^2}{(x(Q) - x(P))^2} \\ =&\frac{x(Q)^3 + Ax(Q) + B - 2y(P)y(Q) + x(P)^3 + Ax(P) + B}{(x(Q) - x(P))^2} \\ &- \frac{\big(x(Q)(x(Q) - x(P))^2 + x(P)(x(Q) - x(P))^2\big)}{x(Q) - x(P))^2} \\ =&\frac{2(B - y(P)y(Q)) + A(x(P) + x(Q)) + x(P)^2x(Q) + x(P)x(Q)^2}{(x(Q) - x(P))^2} \\ =&\frac{2(B - y(P)y(Q)) + (A + x(P)x(Q))(x(P) + x(Q))}{(x(Q) - x(P))^2}. \end{aligned}$$

The second equality follows if we expand the squared terms in the numerator of the second expression and cancel accordingly. For the last equality we just collect like terms. Now we can plug in our representations for the coordinates of P and Q and multiply both numerator and denominator by d^4e^4 to get a simple fraction. This gives

$$\begin{aligned} &\frac{2(B - y(P)y(Q)) + (A + x(P)x(Q))(x(P) + x(Q))}{(x(Q) - x(P))^2} \\ =&\frac{2(Bd^4e^4 - y_1y_3de) + (Ad^2e^2 + x_1x_3)(x_1e^2 + x_3d^2)}{(x_3d^2 - x_1e^2)^2}. \end{aligned}$$

Remember that this is an expression for the $x(P+Q)$. For $\frac{p}{q} \in \mathbb{Q}$, we can directly use the definition of the height of a rational number to see that

$$\max\{|p|, |q|\} \geq H\left(\frac{p}{q}\right).$$

We can now conclude that

$$H(x(P+Q)) \leq \max\{\big|2(Bd^4e^4 - y_1y_3de) + (Ad^2e^2 + x_1x_3)(x_1e^2 + x_3d^2)\big|, \big|(x_3d^2 - x_1e^2)^2\big|\}.$$

We will now derive an upper bound for this maximum. To do this we will use some basic properties of the absolute value and maximum. First, by repeated use of the triangle equality, we get that

$$\begin{aligned}&\max\{\left|2(Bd^4e^4 - y_1y_3de) + (Ad^2e^2 + x_1x_3)(x_1e^2 + x_3d^2)\right|, \left|(x_3d^2 - x_1e^2)^2\right|\}\\ \leq &\max\{2|Bd^4e^4| + 2|y_1y_3de| + |Ax_1d^2e^4| + |Ax_3d^4e^2| + |x_1^2x_3e^2| + |x_1x_3^2d^2|,\\ &|x_3^2d^4| + 2|x_1x_3d^2e^2| + |x_1^2e^4|\}.\end{aligned}$$

Note that $\sum_{i=1}^n |a_i| \leq \max_{1\leq i\leq n}\{n|a_i|\}$. This tells us that

$$\begin{aligned}&\max\{2|Bd^4e^4| + 2|y_1y_3de| + |Ax_1d^2e^4| + |Ax_3d^4e^2| + |x_1^2x_3e^2| + |x_1x_3^2d^2|,\\ &|x_3^2d^4| + 2|x_1x_3d^2e^2| + |x_1^2e^4|\}\\ \leq &\max\{12|Bd^4||e|^4, 12|y_1d||y_3e|, 6|Ax_1d^2||e|^4, 6|Ad^4||x_3e^2|, 6|x_1|^2|x_3e^2|,\\ &6|x_1d^2||x_3|^2, 3|d|^4|x_3|^2, 6|x_1d^2||x_3e^2|, 3|x_1|^2|e|^4\}.\end{aligned}$$

We have written each of the terms we take the maximum over in the form $c_1|w_1||w_2|$. Here $c_1 \in \mathbb{Z}_{>0}$, w_1 is a product of elements from $\{A, B, x_1, y_1, d\}$ and w_2 is a product of elements from $\{x_3, y_3, e\}$. Next, consider $c'_P = \max\{12(|B|+1)|d^4|, 12|y_1d|, 6(|A|+1)|x_1d^2|, 6(|A|+1)|d^4|, 6|x_1|^2, 6|x_1d^2|, 3|d|^4, 6|x_1d^2|, 3|x_1|^2\}$. The $+1$ are there to ensure that $|A|+1, |B|+1 \geq 1$. Note that then for each element in the maximum, $c_1|w_1| \leq c'_P$. The important part here is that c'_P depends only on P and not on Q. We can now conclude

$$\begin{aligned}&\max\{12|Bd^4||e|^4, 12|y_1d||y_3e|, 6|Ax_1d^2||e|^4, 6|Ad^4||x_3e^2|, 6|x_1|^2|x_3e^2|,\\ &6|x_1d^2||x_3|^2, 3|d|^4|x_3|^2, 6|x_1d^2||x_3e^2|, 3|x_1|^2|e|^4\}\\ \leq &\max\{c'_P|e|^4, c'_P|y_3e|, c'_P|e|^4, c'_P|x_3e^2|, c'_P|x_3e^2|, c'_P|x_3|^2, c'_P|x_3|^2, c'_P|x_3e^2|, c'_P|e|^4\}\\ = &c'_P \max\{|e|^4, |y_3e|, |x_3e^2|, |x_3|^2\}.\end{aligned}$$

Since $\prod_{i=1}^n |a_i| \leq \max_{1\leq i\leq n}\{|a_i|^n\}$, we can note that $|x_3e^2| \leq \max\{|x_3|^2, |e|^4\}$. Putting all of this together tells us that $H(x(P+Q)) \leq c'_P \max\{|e|^4, |y_3e|, |x_3|^2\}$. Since $h(Q) = \log_b\left(\max\{|e|^2, |x_3|\}\right)$, this is almost what we want. We just need to give an upper bound the term $|y_3e|$. To do this, we notice that the point Q satisfies the Weierstrass equation. This means that

$$\frac{y_3^2}{e^6} = \frac{x_3^3}{e^6} + A\frac{x_3}{e^2} + B$$

and therefore

$$y_3^2 = x_3^3 + Ax_3e^4 + Be^6.$$

We can use the same tricks as above to obtain

$$|x_3^3 + Ax_3e^4 + Be^6| \leq |x_3|^3 + |A||x_3e^4| + |B||e|^6 \leq 3\max\{|x_3|^3, |A||x_3e^4|, |B||e|^6\}.$$

The absolute values around $x_3^3 + Ax_3e^4 + Be^6$ are not doing anything, since by definition this expression is equal to $y_3^2 \geq 0$. This means that $y_3^2 \leq 3\max\{|x_3|^3, |A||x_3e^4|, |B||e|^6\}$ and so

$$|y_3| = \sqrt{y_3^2} \leq \sqrt{3}\sqrt{\max\{|x_3|^3, |A||x_3e^4|, |B||e|^6\}} = \sqrt{3}\max\{|x_3|^{3/2}, \sqrt{|A|}\sqrt{|x_3|}|e|^2, \sqrt{|B|}|e|^3\}.$$

Now we recognise $\sqrt{|x_3|}|e|^2$ as the product of three things, so we can use the same product formula as above to get $\sqrt{|x_3|}|e|^2 \leq \max\{|x_3|^{3/2}, |e|^3\}$. We have $1, \sqrt{|A|}, \sqrt{|B|} \leq \sqrt{(|A|+1)(|B|+1)}$ and so $|y_3| \leq \sqrt{3|AB|}\max\{|x_3|^{3/2}, |e|^3\}$. Now it is immediate that $|y_3e| \leq \sqrt{3|AB|}\max\{|x_3^{3/2}e|, |e|^4\}$. But we can use the same trick for products again by recognising $|x_3^{3/2}e|$ as a product of four things and so $|x_3^{3/2}e| \leq \max\{|x_3|^2, |e|^4\}$. This means that $|y_3e| \leq \sqrt{3|AB|}\max\{|x_3|^2, |e|^4\}$. We can combine this with our previous bound for $H(x(P+Q))$ to get

$$H(x(P+Q)) \leq c''_P \max\{|e|^4, |x_3|^2\},$$

where $c''_P = c'_P\sqrt{3|AB|}$. Now we are ready to give the desired inequality for this case:

$$\begin{aligned} h(P+Q) &= \log_b\left(H(x(P+Q))\right) \leq \log_b\left(c''_P \max\{|e|^4, |x_3|^2\}\right) \\ &= \log_b(c''_P) + 2\log_b\left(\max\{|e|^2, |x_3|\}\right) \\ &= \log_b(c''_P) + 2h(Q). \end{aligned}$$

For each possible Q we have found a constant that works. We now define $c_P = \max\{\log_b(c''_P), h(P), h(2P)\}$ and, with the help of our previous discussions we immediately conclude that, for all $Q \in E(\mathbb{Q})$,

$$h(P+Q) \leq 2h(Q) + c_P.$$

2. We will start in a manner similar to the proof of the first property and take care of the edge cases first. We know that $P + P = O$ if and only if $P \in E(\mathbb{Q})[2]$. But this set is finite so it makes sense to demand that $-c \geq \max\{4h(P) \mid P \in E(\mathbb{Q}) \text{ such that } 2P = O\}$. If c satisfies this condition, then $h(2P) = 0 \geq 4h(P) + c$ for every $P \in E(\mathbb{Q})$ such that $2P = O$.

   Now suppose that $P = [x : y : 1] \in E(\mathbb{Q})$ with $P + P \neq O$. We use the formula in theorem 2.8 and the Weierstrass equation to get that

$$\begin{aligned} x(2P) &= \left(\frac{3x^2 + A}{2y}\right)^2 - 2x \\ &= \frac{9x^4 + 6Ax^2 + A^2}{4(x^3 + Ax + B)} - \frac{8x(x^3 + Ax + B)}{4(x^3 + Ax + B)} \\ &= \frac{x^4 - 2Ax^2 - 8Bx + A^2}{4(x^3 + Ax + B)}. \end{aligned}$$

   We have that $x(2P) \in \mathbb{Q}$, so we can also write $x(P) = \frac{p}{q}$ with $\gcd(p, q) = 1$. Next we define the following polynomials over $\mathbb{Z}$ to make upcoming calculations more readable:

$$\begin{aligned} N(X_1, X_2) &:= X_1^4 - 2AX_1^2X_2^2 - 8BX_1X_2^3 + A^2X_2 \\ D(X_1, X_2) &:= 4(X_1^3X_2 + AX_1X_2^3 + BX_2^4) \\ n_q(X_1, X_2) &:= 12X_1^2X_2 + 16AX_2^3 \\ d_q(X_1, X_2) &:= -3X_1^3 + 5AX_1X_2^2 + 27BX_2^3 \\ n_p(X_1, X_2) &:= 4\left((4A^3 + 27B^2)X_1^3 - A^2BX_1^2X_2 + A(3A^3 + 22B^2)X_1X_2^2 + 3B(A^3 + 8B^2)X_2^3\right) \\ d_p(X_1, X_2) &:= A^2BX_1^3 + A(5A^3 + 32B^2)X_1^2X_2 + 2B(13A^3 + 96B^2)X_1X_2^2 - 3A^2(A^3 + 8B^2)X_2^3. \end{aligned}$$

   Note that

$$x(2P) = \frac{(\frac{p}{q})^4 - 2A(\frac{p}{q})^2 - 8B\frac{p}{q} + A^2}{4((\frac{p}{q})^3 + A\frac{p}{q} + B)} = \frac{p^4 - 2Ap^2q^2 - 8Bpq^3 + A^2q^4}{4(p^3q + Apq^3 + Bq^4)} = \frac{N(p,q)}{D(p,q)}.$$

   Then by the definition of the height of a rational number, we have that

$$\max\{|N(p,q)|, |D(p,q)|\} = \gcd(N(p,q), D(p,q))H(x(2P)).$$

   A direct, but very tedious, calculation will show that

$$\begin{aligned} n_q(p,q)N(p,q) + d_q(p,q)D(p,q) &= 16A^3q^7 + 108B^2q^7 = 4\Delta q^7 \text{ and} \\ n_p(p,q)N(p,q) + d_p(p,q)D(p,q) &= 16A^3q^7 + 108B^2q^7 = 4\Delta p^7. \end{aligned}$$

   We will write $M_1 = \max\{|n_q(p,q)|, |d_q(p,q)|, |n_p(p,q)|, |d_p(p,q)|\}$ and $M_2 = \max\{|N(p,q)|, |D(p,q)|\}$ for the sake of compactness. From here we can use the triangle inequality and the identity $\sum_{i=1}^n |a_i| \leq \max_{1\leq i\leq n}\{n|a_i|\}$ like in the proof of property 1, to get that

$$4|\Delta q^7| \leq |n_q(p,q)||N(p,q)| + |d_q(p,q)||D(p,q)| = M_1(|N(p,q)| + |D(p,q)|) \leq 2M_1M_2,$$

and in the same way that also $4|\Delta p^7| \leq 2M_1M_2$. Therefore $2|\Delta| \max\{|p|^7, |q|^7\} \leq M_1M_2$.

Looking back at the equations for $4\Delta q^7$ and $4\Delta p^7$, we see that $\gcd(F(p,q), G(p,q))$ divides the left hand side of both equations, so it divides both $4\Delta q^7$ and $4\Delta p^7$. But we assumed p and q to be coprime, so p^7 and q^7 are are also coprime and therefore M_2 divides 4Δ. This then means that

$$\frac{M_2}{4|\Delta|} = \frac{\gcd(N(p,q), D(p,q))}{4|\Delta|} H\left(x(2P)\right) \leq H\left(x(2P)\right),$$

since $\frac{\gcd(N(p,q),D(p,q))}{4|\Delta|} \in \mathbb{Z}_{\geq 1}$. Summarising everything gives

$$\max\{|p|^7, |q|^7\} \leq 2M_1 H(x(2P)).$$

Next we will give an upper bound for M_1. We use the same tricks as in the proof of part one to write

$$\begin{aligned}
M_1 &= \max\{|n_q(p,q)|, |d_q(p,q)|, |n_p(p,q)|, |d_p(p,q)|\} \\
&= \max\{|12p^2q + 16Aq^3|, |-3p^3 + 5Apq^2 + 27Bq^3|, \\
&\qquad |4\left((4A^3 + 27B^2)p^3 - A^2Bp^2q + A(3A^3 + 22B^2)pq^2 + 3B(A^3 + 8B^2)q^3\right)|, \\
&\qquad |A^2Bp^3 + A(5A^3 + 32B^2)p^2q + 2B(13A^3 + 96B^2)pq^2 - 3A^2(A^3 + 8B^2)q^3|\} \\
&\leq \max\{12|p|^2|q|, 16|A||q|^3, 3|p|^3, 5|A||p||q|^2, 27|B||q|^3, \\
&\qquad 4|4A^3 + 27B^2||p|^3, 4|A^2B||p|^2|q|, |A(3A^3 + 22B^2)||p||q|^2, 12|B(A^3 + 8B^2)||q|^3, \\
&\qquad |A^2B||p|^3, |A(5A^3 + 32B^2)||p|^2|q|, 2|B(13A^3 + 96B^2)||p||q|^2, 3|A^2(A^3 + 8B^2)||q|^3\} \\
&\leq c' \max\{|p|^3, |q|^3\},
\end{aligned}$$

where

$$\begin{aligned}
c' = \max\{&12, 16|A|, 3, 5|A|, 27|B|, 4|4A^3 + 27B^2|, 4|A^2B|, |A(3A^3 + 22B^2)|, 12|B(A^3 + 8B^2)|, \\
&|A^2B|, |A(5A^3 + 32B^2)|, 2|B(13A^3 + 96B^2)|, 3|A^2(A^3 + 8B^2)|\}.
\end{aligned}$$

We could simplify the expression for c' further, but this will not change anything for the rest of the proof. The important part is that the constant c' only depends on the curve itself and not on the point P. Putting everything together gives us that

$$(\max\{|p|, |q|\})^7 \leq 2c'(\max\{|p|, |q|\})^3 H(x(2P)).$$

Now we see that

$$h(2P) = \log_b(H(x(2P))^2) \geq \log_b\left(\frac{(\max\{|p|, |q|\})^4}{2c'}\right) = 4h(P) - \log_b(2c').$$

So if we choose $-c = \max\{\log_b(2c'), \max\{4h(P) \mid P \in E(\mathbb{Q}) \text{ such that } 2P = O\}\}$, we see that $h(2P) \geq 4h(P) + c$ for all $P \in E(\mathbb{Q})$.

3. If $C < 1$, then there are no rational numbers with height at most C. Suppose $C \geq 1$. Then there are no more than $C+1$ non-negative integers less than or equal to C. For a rational number to have a height less than or equal to C, both denominator and numerator can be no more than C in absolute value. There are at most $(C+1)^2$ such combinations with positive numbers, so there are at most $(C+1)^2$ positive rational numbers with height at most C. This is because for both the numerator and denominator we have $C+1$ choices in principle. Of course in practice a lot of the combinations would give the same rational number, so this bound could be much tighter. For the claim this is not of importance, since we only care about finiteness. In total, there are then no more than $2(C+1)^2$ rational numbers with height at most C. For a given X, at most two Y exist such that $(X, Y, 1)$ is a root of the Weierstrass polynomial. This means that for $C > 0$, we have $\#\{P \in E(\mathbb{Q}) \mid h(P) \leq C\} \leq 2\#\{r \in \mathbb{Q} \mid H(r) \leq b^C\} + 1$. This plus one comes from the point O. So in conclusion $\#\{P \in E(\mathbb{Q}) \mid h(P) \leq C\} \leq C\} \leq 4(b^C + 1)^2 + 1$. If $C < 0$, then $\#\{P \in E(\mathbb{Q}) \mid h(P)\} = 0$ and if $C = 0$, then $\#\{P \in E(\mathbb{Q}) \mid h(P) \leq C\} \leq 5$. This is because the only possible points P with $h(P) = 0$ are $P = O$ or have $x(P) = 0$ or $x(P) = 1$. There are again no more than two points with the same x-coordinate. In any case, the set $\{P \in E(\mathbb{Q}) \mid h(P) \leq C\}$ is finite for all $C \in \mathbb{R}$. □

This concludes the (conditional) proof of the Mordell-Weil theorem (3.9) for the special case that $K = \mathbb{Q}$. Constructing a height function for general number fields and showing that it has the right properties, is a bit more involved. One reason for this, is the relatively complicated algebraic structure of number fields. The rational numbers can be thought of as the fraction field of a unique factorisation domain, namely $\mathbb{Z}$. This is not always the case for number fields. But for our height function on $\mathbb{Q}$, we implicitly made use of this. Unfortunately we do not have the time to work this out in this thesis. A good reference is [1, section VIII.6]. The fundamental theorem of finitely generated abelian groups now states that

$$E(\mathbb{Q}) \cong \mathbb{Z}^r \oplus \mathbb{Z}/a_1\mathbb{Z} \oplus \cdots \oplus \mathbb{Z}/a_n\mathbb{Z},$$

for $r \geq 0$ and $a_0, \ldots, a_n \geq 1$ ([5, section 5.2, theorem 3]).

6 The weak Mordell-Weil theorem

In this section we will prove the assumption of theorem 5.2 in the general setting of a number field. This is often called the weak Mordell-Weil theorem.

Theorem 6.1 (weak Mordell-Weil)**.** *The group $E(K)/mE(K)$ is finite.*

Note that for the case $K = \mathbb{Q}$, we technically only needed to prove this for $m = 2$.

6.1 Reduction of the problem

In this section we show that the proof of the weak Mordell-Weil theorem can be reduced to the proof of the finiteness of a certain field extension L/K. To do this we will first show that it is enough to proof the Mordell-Weil theorem under the assumption that $E(\bar{K})[m] \subseteq E(K)$. We will now state a lemma which help us achieve this and which we can use again later.

Lemma 6.2. *Let $P \in E(\bar{K})$ and suppose that $P^\sigma = P$ for all $\sigma \in \mathrm{Gal}(\bar{K}/K)$. Then $P \in E(K)$.*

Proof. Certainly $O \in E(K)$, so suppose that $P \neq O$. Then we can write $P = [x : y : 1]$ and in particular $P^\sigma = [x^\sigma : y^\sigma : 1]$ for all $\sigma \in \mathrm{Gal}(\bar{K}/K)$. Since we have written both P and P^σ with their last coordinate equal to 1, we can conclude that $x^\sigma = x$ and $y^\sigma = y$ for all $\sigma \in \mathrm{Gal}(\bar{K}/K)$. Now by theorem 3.6, we can conclude that $x, y \in K$ and therefore $P \in E(K)$. □

Proposition 6.3. *For a Galois extension F/K, if $E(F)/mE(F)$ is finite, then $E(K)/mE(K)$ is finite.*

Proof. We have the natural inclusion $\iota\colon E(K) \to E(F)$. Consider the map

$$\varphi\colon E(K)/mE(K) \to E(F)/mE(F)\colon [P] \mapsto [\iota(P)].$$

First we note that both these quotients make sense since both $E(K)$ and $E(F)$ are abelian. To show that this is well-defined, let $P, Q \in E(K)$ and suppose $[P] = [Q] \in E(K)/mE(K)$. Then $P - Q \in mE(K)$ so we can write $P - Q = mR$, where $R \in E(K)$. Using the fact that ι is a group homomorphism we get that

$$\iota(P) - \iota(Q) = \iota(P - Q) = \iota(mR) = m\iota(R).$$

But $\iota(R) \in E(F)$, so $\iota(P) - \iota(Q) \in mE(F)$. Therefore $[\iota(P)] = [\iota(Q)] \in E(F)/mE(F)$ and so φ is well-defined. It is immediate that φ is also a group homomorphism.

Suppose $P \in E(K)$ with $[P] \in \ker\varphi$. Then $[\iota(P)] \in mE(F)$, so we can choose a $Q \in E(F)$ such that $mQ = P$. We can then consider the map $\kappa_P\colon \mathrm{Gal}(F/K) \to E(\bar{K})[m]\colon \sigma \mapsto Q^\sigma - Q$. We know from proposition 3.10 that the action of and element the Galois group is a group homomorphism and therefore a $\mathbb{Z}$-module homomorphism. This means that the Galois action also respects multiplication by m and so

$$m(Q^\sigma - Q) = mQ^\sigma - mQ = (mQ)^\sigma - P = P^\sigma - P = O.$$

This last equality we get from the fact that $P \in E(K)$. If $P \neq O$ we can write $P = [x : y : 1]$ with $x, y \in K$, so σ acts trivially on P since it leaves the field K fixed. In the same way, we immediately see that O also is fixed by σ. This shows that the image of this map is truly contained in $E(\bar{K})[m]$ and our definition makes sense. Here we have to note that for each $P \in E(K)$ we need to fix the point Q before defining the map, since in general Q need not be unique in having the property $mQ = P$. Next, suppose that $P_1, P_2 \in E(K)$ such that $[P_1], [P_2] \in \ker \varphi$ and $\kappa_{P_1} = \kappa_{P_2}$. We will show that in the quotient $E(K)/mE(K)$, the points P_1 and P_2 coincide. By definition, we have that

$$Q_1^\sigma - Q_1 = Q_2^\sigma - Q_2,$$

for all $\sigma \in \mathrm{Gal}(F/K)$. Here $mQ_1 = P_1$ and $mQ_2 = P_2$. We can now use the fact that the Galois action respects the addition again, to see that

$$(Q_1 - Q_2)^\sigma = Q_1 - Q_2,$$

for all $\sigma \in \mathrm{Gal}(F/K)$. Lemma 6.2 then tells us that $Q_1 - Q_2 \in E(K)$. We note that $P_1 - P_2 = mQ_1 - mQ_2 = m(Q_1 - Q_2)$ and therefore $P_1 - P_2 \in mE(K)$. This means that $[P_1] = [P_2] \in E(K)/mE(K)$. This shows that each class in $\ker \varphi$ gives rise to distinct functions from $\mathrm{Gal}(F/K)$ to $E(\bar{K})[m]$. Note that we have not shown that distinct points in $E(K)$ give rise to the same function if they coincide in $E(K)/mE(K)$ and therefore a class in the kernel of this quotient could potentially give rise to multiple functions. But this is not of importance to us since we now do know that each class gives rise to at least one function and distinct classes give distinct functions. In particular we have that $\#\ker \varphi \leq \#\mathcal{F}(\mathrm{Gal}(F/K), E(\bar{K})[m])$.

From corollary 3.13 we know that $E(\bar{K})[m]$ is finite. By assumption, we have that F/K is finite so $\mathrm{Gal}(F/K)$ is finite. There are only finitely many maps between two finite sets, so we conclude that $\mathcal{F}(\mathrm{Gal}(F/K), E(\bar{K})[m])$, and therefore $\ker \varphi$, is finite. Our other assumption tells us that $E(F)/mE(F)$, and therefore $\mathrm{im}\ \varphi$, is finite. But then certainly $E(K)/mE(K)$ finite, since its cardinality is the product of the cardinalities of the kernel and the image. □

By corollary 3.13, we can write $E(\bar{K})[m] = \{O, [x_1 : y_1 : 1], \ldots, [x_n, y_n : 1]\}$. We write ζ_m for a primitive m-th root of unity and define $K' = K(x_1, y_1, \ldots, x_n, y_n, \zeta_m)$. All the elements we are adjoining are algebraic over K, so this is a finite algebraic extension of K. But remember that K is a number field and therefore perfect. This means every algebraic extension of K, in particular K', is separable. Now we see that it makes sense to talk about the Galois closure K'^c of K'. We immediately see that $E(\bar{K})[m] \leq E(K'^c)$ and that K'^c contains all m-th roots of unity. By the previous proposition, it is enough to prove the weak Mordell-Weil theorem for $E(K'^c)$. This is again a number field and throughout this section we will just denote it by K. But now note that $E(\bar{K})[m] \leq E(K)$. This will be crucial in the following definition and the rest of this section. We will use the fact that K contains the all roots of unity later in this section.

Definition 6.4. We define the *Kummer pairing* $\kappa\colon E(K) \times \mathrm{Gal}(\bar{K}/K) \to E(\bar{K})[m]$ by $\kappa(P, \sigma) = Q^\sigma - Q$, where $Q \in E(\bar{K})$ such that $mQ = P$.

It is not at all obvious that this gives us a well-defined mapping. First we note that by theorem 3.12 such a Q exists. We now have to prove two things to show that the map is well-defined. The first is that the Kummer pairing does not depend on which point Q we choose. Lastly we have to show that κ actually takes values in $E(\bar{K})[m]$.

Let $P \in E(K)$ and $\sigma \in \mathrm{Gal}(\bar{K}/K)$. Now suppose that $P = mQ = mQ'$, for $Q, Q' \in E(\bar{K})$. Since the action of the Galois group respects the addition on the curve, we have that

$$Q'^\sigma - Q' = Q^\sigma - Q + (Q' - Q)^\sigma - (Q' - Q).$$

Now notice that $m(Q' - Q) = mQ' - mQ = P - P = O$ and therefore $Q' - Q \in E(\bar{K})[m]$. But we have assumed that $E(\bar{K})[m] \subseteq E(K)$, so $Q' - Q$ is fixed by σ. This means that $(Q' - Q)^\sigma - (Q' - Q) = O$, so we conclude that $Q'^\sigma - Q' = Q^\sigma - Q$. This shows that the value of κ does not depend on our choice of Q. Now it only remains to show that κ actually takes values in $E(\bar{K})[m]$. But we have already seen in the proof of

proposition 6.3 that $m\kappa(P,\sigma) = m(Q^\sigma - Q) = O$, so $\kappa(P,\sigma) \in E(\bar{K})[m]$.

The Kummer pairing looks very similar to the maps κ_P we defined in the proof of proposition 6.3. But note that in proposition 6.3 we had to choose a $Q \in E(\bar{K})$ with $mQ = P$, for each $P \in E(K)$. Here we did not have to do that because of the fact that $E(\bar{K})[m] \subseteq E(K)$. We will now show a few properties of the Kummer pairing. It turns out that the property $E(\bar{K})[m] \leq E(K)$ will be crucial for showing these properties. To do this, we will first introduce some notation. We write $\prod_{i\in I} F_i$ for the *compositum* of the fields F_i. This is defined as the smallest field containing all the F_i. We will always regard the F_i as subfields of some common superfield F. In our case, we will choose $F = \bar{K}$. Since an arbitrary intersection of subfields is again a field, the smallest field containing all F_i is actually well-defined.

Definition 6.5. For a point $P = [x : y : 1] \in E(\bar{K})$, we write $K(P)$ for *the minimal field of definition of P over K* which is defined as the field $K(x, y)$. Furthermore, we define $K(O) = K$.

Proposition 6.6. *1. The Kummer pairing respects the group operation in both coordinates.*

2. We have

$$\bigcap_{\sigma\in \mathrm{Gal}(\bar{K}/K)} \ker \kappa(-,\sigma) = mE(K).$$

3. We have

$$\bigcap_{P\in E(K)} \ker \kappa(P,-) = \mathrm{Gal}(\bar{K}/L).$$

Here

$$L = \prod_{\substack{P\in E(\bar{K})\\ mP\in E(K)}} K(P).$$

Proof. 1. Let $\sigma \in \mathrm{Gal}(\bar{K}/K)$ and $P_1, P_2 \in E(K)$ and $Q_1, Q_2 \in E(\bar{K})$, such that $mQ_1 = P_1$ and $mQ_2 = P_2$. Then

$$\kappa(P_1,\sigma) + \kappa(P_2,\sigma) = Q_1^\sigma - Q_1 + Q_2^\sigma - Q_2 = (Q_1+Q_2)^\sigma - (Q_1+Q_2) = \kappa(P_1+P_2,\sigma),$$

since $m(Q_1 + Q_2) = P_1 + P_2$. Next, let $\sigma' \in \mathrm{Gal}(\bar{K}/K)$. Since we have that $E(\bar{K})[m] \subseteq E(K)$, we know that $(P_1^\sigma - P_1)^{\sigma'} = P_1^\sigma - P_1$. Therefore we get

$$\kappa(P_1,\sigma\sigma') = (P_1^\sigma)^{\sigma'} - P_1 = (P_1^\sigma - P_1)^{\sigma'} + P_1^{\sigma'} - P_1 = P_1^\sigma - P_1 + P_1^{\sigma'} - P_1 = \kappa(P_1,\sigma) + \kappa(p_1,\sigma').$$

This shows that κ is respects the group operation in both coordinates. Note that this means that for each $\sigma \in \mathrm{Gal}(\bar{K}/K)$ and $P \in E(K)$, the maps $\kappa(-,\sigma)$ and $\kappa(P,-)$ are group homomorphisms. Now we see that the kernel notation in 2 and 3 is actually warranted.

2. First suppose that $P \in \bigcap_{\sigma\in\mathrm{Gal}(\bar{K}/K)} \ker\kappa(-,\sigma)$. Then $\kappa(P,\sigma) = O$ for all $\sigma \in \mathrm{Gal}(\bar{K}/K)$, so we have $Q \in E(\bar{K})$ with $mQ = P$ that is invariant under all $\sigma \in \mathrm{Gal}(\bar{K}/K)$. We can again use lemma 6.2 to conclude that $Q \in E(K)$.

   On the other hand, suppose that $P \in mE(K)$. Then we can write $P = mQ$ with $Q \in E(K)$. This Q is certainly fixed by every $\sigma \in \mathrm{Gal}(\bar{K}/K)$, so $P \in \ker\kappa(-,\sigma)$ for every $\sigma \in \mathrm{Gal}(\bar{K}/K)$ and therefore $P \in \bigcap_{\sigma\in\mathrm{Gal}(\bar{K}/K)} \ker\kappa(-,\sigma)$. This shows that $\bigcap_{\sigma\in\mathrm{Gal}(\bar{K}/K)} \ker\kappa(-,\sigma) = mE(K)$

3. First of all, we note that $\bar{K}/L$ is Galois by lemma 3.2. Let $\sigma \in \mathrm{Gal}(\bar{K}/L)$ and $P \in E(K)$. We can write $P = mQ$ with $Q \in E(\bar{K})$. But certainly $Q \in E(K(Q))$, so then $Q \in E(L)$ since $K(Q) \subseteq L$. This means that σ fixes Q and therefore $\sigma \in \bigcap_{P\in E(K)} \ker\kappa(P,-)$. Next let $\sigma \in \bigcap_{P\in E(K)} \ker\kappa(P,-)$. Then $Q^\sigma = Q$ for all $Q \in E(\bar{K})$ such that $mQ \in E(K)$. But it is immediate that the set of all elements fixed by σ is a subfield of $\bar{K}$. We now know that this field contains all $Q \in E(\bar{K})$ such that $mQ \in E(K)$, so by minimality it also contains L. This shows that σ fixes L, so $\sigma \in \mathrm{Gal}(\bar{K}/L)$. We conclude that $\bigcap_{P\in E(K)} \ker\kappa(P,-) = \mathrm{Gal}(\bar{K}/L)$. □

Kernels of group homomorphisms are normal subgroups, so part 3 of 6.6 tells us that $\mathrm{Gal}(\bar{K}/L)$ can be written as the intersection of normal subgroups of $\mathrm{Gal}(\bar{K}/K)$. We may then conclude that $\mathrm{Gal}(\bar{K}/L)$ is normal in $\mathrm{Gal}(\bar{K}/K)$. Since $E(K)$ is abelian, $mE(K)$ is also normal in $E(K)$. Now consider the map

$$\kappa' \colon E(K)/mE(K) \times \mathrm{Gal}(\bar{K}/K)/\mathrm{Gal}(\bar{K}/L) \to E(\bar{K})[m] \colon ([P],[\sigma]) \mapsto \kappa(P,\sigma).$$

To show that κ' is well-defined, let $P_1, P_2 \in E(K)$ with $[P_1] = [P_2]$ and $\sigma_1, \sigma_2 \in \mathrm{Gal}(\bar{K}/K)$ with $[\sigma_1] = [\sigma_2]$. This means that we have $P_1 = P_2 + Q$, where $Q \in mE(K)$, and $\sigma_1 = \sigma_2\tau$, where $\tau \in \mathrm{Gal}(\bar{K}/L)$. Using this and all of 6.6, we get that

$$\begin{aligned}\kappa'([P_1],[\sigma_1]) &= \kappa(P_1,\sigma_1) = \kappa(P_2+Q,\sigma_2\tau) = \kappa(P_2,\sigma_2) + \kappa(P_2,\tau) + \kappa(Q,\sigma_2) + \kappa(Q,\tau) \\ &= \kappa(P_2,\sigma_2) = \kappa'([P_2],[\sigma_2]).\end{aligned}$$

So κ' is well-defined and it is immediate that it respects the group operation in both coordinates.

Lemma 6.7. *The extension L/K is Galois and*

$$\mathrm{Gal}(\bar{K}/K)/\mathrm{Gal}(\bar{K}/L) \cong \mathrm{Gal}(L/K).$$

.

Proof. First we define $S' = \{Q \in E(\bar{K}) \mid mQ \in E(K)\}$. Using this, we define

$$S = \{\alpha \in \bar{K} \mid \text{there exists } Q \in S' \text{ such that } Q = [\alpha : \beta : 1] \text{ or } Q = [\beta : \alpha : 1] \text{ for some } \beta \in K\}.$$

We defined L to be minimal with respect to the property of containing the elements of S, so $L = K(S)$. That is, L is the minimal field extension of K generated by all the elements of S. Let $\alpha \in S$ and $\sigma \in \mathrm{Gal}(\bar{K}/K)$. Without loss of generality, there exists $Q \in E(\bar{K})$ and $\beta \in \bar{K}$, such that $Q = [\alpha : \beta : 1]$ and $mQ \in E(K)$. Note that then

$$mQ^\sigma = (mQ)^\sigma = mQ \in E(K),$$

But then the first coordinate of Q^σ, which is α^σ, is an element of S. This shows that $\sigma(S) \subseteq S$. Since σ fixes K, it follows that $\sigma(K[S]) \subseteq K[S]$. Next we note that certainly $K[S] \subseteq L$ and therefore $\mathrm{Frac}(K[S]) \subseteq L$. Here $\mathrm{Frac}(K[S])$ denotes the fraction field of $K[S]$. The reverse inclusion is immediate since L is the minimal field extension of K containing S, and therefore it contains $K[S]$. Let $\alpha \in L$. Then we can write

$$\alpha = \frac{\beta}{\gamma},$$

where $\beta, \gamma \in K[S]$. But we already showed that $\sigma(K[S]) \subseteq K[S]$, so $\beta^\sigma, \gamma^\sigma \in K[S]$. Now we see that

$$\alpha^\sigma = \frac{\beta^\sigma}{\gamma^\sigma} \in \mathrm{Frac}(K[S]) = L.$$

Finally, we use lemma 3.4 to conclude that L/K is Galois. Theorem 3.5 then immediately gives the desired isomorphism. □

This tells us that we can regard our function κ' as having $E(K)/mE(K) \times \mathrm{Gal}(L/K)$ as its domain. In the next lemma we will use κ' to reduce proving the finiteness of $E(K)/mE(K)$, to proving the finiteness of the extension L/K. We will also use κ' to show two properties of $\mathrm{Gal}(L/K)$ that will be useful in the next section.

Lemma 6.8. *1. If the extension L/K is finite, then $E(K)/mE(K)$ is finite.*

2. The group $\mathrm{Gal}(L/K)$ is abelian.

3. We have that $\sigma^m = 1$ for all $\sigma \in \mathrm{Gal}(L/K)$.

Proof. 1. Consider the map

$$\varphi_1 \colon E(K)/mE(K) \to \operatorname{Hom}\big(\operatorname{Gal}(L/K), E(\bar{K})[m]\big) : [P] \mapsto \kappa'([P], -).$$

Since κ' respects the group operation in the second coordinate, the $\kappa'([P], -)$ are actual group homomorphisms. Suppose that $\varphi_1([P_1]) = \varphi_1([P_2])$, for $P_1, P_2 \in E(K)$. Then $\kappa'([P_1], \sigma) = \kappa'([P_2], \sigma)$ for all $\sigma \in \operatorname{Gal}(L/K)$ and therefore $\kappa(P_1, \sigma) = \kappa(P_2, \sigma)$ for all $\sigma \in \operatorname{Gal}(\bar{K}/K)$ with the identification from lemma 6.7. But $\kappa(P_2, \sigma) + \kappa(-P_2, \sigma) = O$, so $-\kappa(P_2, \sigma) = \kappa(-P_2, \sigma)$. Now we conclude that $P_1 - P_2 \in \bigcap_{\sigma \in \operatorname{Gal}(\bar{K}/K)} \ker(-, \sigma)$, so by property 2 of proposition 6.6 we have $P_1 - P_2 \in mE(K)$. Therefore $[P_1] = [P_2] \in E(K)/mE(K)$.

This shows that φ_1 is injective, so the cardinality of its domain is not bigger than the cardinality of its codomain. Next we will use the same trick as in proposition 6.3. Suppose that L/K is finite, so $\operatorname{Gal}(L/K)$ is also finite. We also know from corollary 3.13 that $E(\bar{K})[m]$ is finite, so there are finitely many maps between them. But this means that the codomain of φ_1 is finite and therefore its domain, $E(K)/mE(K)$, is finite.

2. Consider the map

$$\varphi_2 \colon \operatorname{Gal}(L/K) \to \operatorname{Hom}\big(E(K)/mE(K), E(\bar{K})[m]\big) : \sigma \mapsto \kappa'(-, \sigma).$$

Since κ' respects the group operation in the first coordinate, the $\kappa'(-, \sigma)$ are actual group homomorphisms. We have a natural group structure on the set $\operatorname{Hom}\big(E(K)/mE(K), E(\bar{K})[m]\big)$ by adding pointwise. Since both $E(K)/mE(K)$ and $E(\bar{K})[m]$ are abelian, it is immediate that this group is abelian. Since κ' respects the group operation in the second coordinate, φ_2 is a group homomorphism. We will show that φ_2 is also injective.

Suppose that $\varphi_2(\sigma_1) = \varphi_2(\sigma_2)$ for some $\sigma_1, \sigma_2 \in \operatorname{Gal}(L/K)$. Then $\kappa'([P], \sigma_1) = \kappa'([P], \sigma_2)$ for all $[P] \in E(K)/mE(K)$, so $\kappa(P, \sigma_1) = \kappa(P, \sigma_2)$ for all $P \in E(K)$. Here we regard σ_1 and σ_2 as elements of $\operatorname{Gal}(\bar{K}/K)$. Then in the same way as above, $\sigma_1\sigma_2^{-1} \in \bigcap_{P \in E(K)} \ker \kappa(P, -)$. By property 3 of proposition 6.6 we see that $\sigma_1\sigma_2^{-1} \in \operatorname{Gal}(\bar{K}/L)$. But we know from theorem 3.5 that the isomorphism in lemma 6.7 comes from function restriction. Therefore $\sigma_1 = \sigma_2$ as elements of $\operatorname{Gal}(L/K)$, so φ_2 is injective. By the first isomorphism theorem, the group $\operatorname{Gal}(L/K)$ is isomorphic to its image under φ_2. But this is a subgroup of an abelian group, so again abelian. We conclude that $\operatorname{Gal}(L/K)$ is abelian. Since the operation of a Galois group is composition, we will not switch to an additive notation for $\operatorname{Gal}(L/K)$ but will keep writing it multiplicatively.

3. Let $\sigma \in \operatorname{Gal}(L/K)$, $P \in E(K)$ and φ_2 as above. We know from the proof of the second property that φ_2 is a group homomorphism, so

$$\varphi_2(\sigma^m)([P]) = (m\varphi_2(\sigma))([P]) = m\varphi_2(\sigma)([P]) = m\kappa'([P], \sigma) = O.$$

The last equality holds because, by definition, $\kappa'([P], \sigma) \in E(\bar{K})[m]$. This shows that $\varphi_2(\sigma^m) = 0 \in \operatorname{Hom}\big(E(K)/mE(K), E(\bar{K})[m]\big)$. But we also know from the proof of property 2 that φ_2 is injective, so σ^m is the identity in $\operatorname{Gal}(L/K)$. □

These properties will be very useful in light of the main theorem of Kummer theory, which we will state here without proof. Now we also see why we assumed that K contains a primitive m-th root of unity (and therefore all m-th roots of unity).

Theorem 6.9. *Let F be field that contains a primitive m-th root of unity. Then $F(\{\sqrt[m]{a} \mid a \in F\})$ is the maximal extension F' of F such that $Gal(F'/F)$ is abelian and for all $\sigma \in Gal(F'/F)$, $\sigma^m = id_{F'}$.*

Proof. For a proof, see [5, section 17.3]. □

6.2 Finiteness of the Extension L/K

As we noted before, we have reduced the problem to proving the finiteness of the field extension L/K. This section will be dedicated to proving that the extension is finite. To do this, we will need results from algebraic number theory and general commutative algebra. The proof of some these results are beyond the scope of this thesis, but we will provide references to them. To state these results, we first introduce some new objects.

Definition 6.10. A *(discrete) valuation* on a integral domain R is a function $v\colon R \to \mathbb{Z} \cup \{\infty\}$, such that for all $\alpha, \beta \in R$

1. $v(\alpha\beta) = v(\alpha) + v(\beta)$
2. $v(\alpha + \beta) \geq \min\{v(\alpha), v(\beta)\}$
3. $v(\alpha) = \infty$ if and only if $\alpha = 0$.

It is implicit here that the symbol ∞ is bigger than all integers and that $\infty + a = \infty$ and $\infty + \infty = \infty$, for all $a \in \mathbb{Z}$. We always have the trivial valuation v with $v(r) = 0$ for all $r \in R \setminus \{0\}$, but we will not consider this any further in this thesis.

Lemma 6.11. *1. For a field F and a set S of valuations on F, the set*

$$R = \{a \in F \mid v(a) \geq 0 \text{ for all } v \in S\}$$

is a subring of F. It is also an integral domain.

2. If $S = \{v\}$, then F is the quotient field of R and the subset $M = \{a \in F \mid v(a) > 0\}$ is the unique maximal ideal of R and M is principal. We call R a discrete valuation ring.

Proof. 1. Note that $0 \in R$, so $R \neq \emptyset$. Suppose that $a, b \in R$ and let $v \in S$. First note that $v(a) = v(1) + v(a)$, so $v(1) = 0$. Let $c \in F^\times$, then we see that $v(c) + v(c^{-1}) = v(1) = 0$, so $v(c^{-1}) = -v(c)$. The units of R will then be the elements that are zero under all valuations in S, since the valuation of the inverse of the other elements of R will be less than zero for some valuation in S and so it will not be contained in R. Note that $2v(-1) = v((-1)^2) = 0$, so $v(-1) = 0$. Next we have that $v(-a) = v(-1) + v(a) = v(a) \geq 0$, so R is closed under opposites. Also $v(a + b) = \min\{v(a), v(b)\} \geq 0$ and $v(ab) = v(a) + v(b) \geq 0$, so R is closed under addition and multiplication. We now conclude that R is a subring of K. It is immediate that R is an integral domain since it is commutative and every zero-divisor would also be a zero-divisor of F, which is impossible.

2. First we show that F is the quotient field of R. Let $a \in F^\times$ and suppose that $a \notin R$, so $v(a) < 0$. There exists $x \in R$ with $v(x) > 0$, since v is non-trivial. Then $v(ax^{-v(a)}) = v(a)(1 - v(x)) \geq 0$, so $ax^{-v(a)} \in R$. But certainly $x^{-v(a)} \in R$. This shows that we can multiply any element of F with an element of R to get an element of R and therefore F is the qotient field of R. Now we show that M is actually an ideal of R. Let $a, b \in M$, so $v(a), v(b) > 0$. We have that $v(a - b) \geq \min\{v(a), v(-b)\}$, but we know from the proof of item 1 that $v(-b) = v(b)$. Then $v(a - b) \geq \min\{v(a), v(b)\} > 0$, so $a - b \in M$. Suppose that $r \in R$, so $v(r) \geq 0$. We immediately see that $v(ra) = v(r) + v(a) > 0$ and therefore $ra \in M$. We again have that $0 \in M$, so $M \neq \emptyset$ and therefore we can conclude that M is indeed an ideal of R. Now suppose that N is an ideal of R that contains M and an element $r \in R \setminus M$. Then $v(r) = 0$ and, as we saw before, this means that r is a unit. But any ideal of R that contains a unit is equal to R, so there are no proper ideals of R that contain M. This shows that M is a maximal ideal. Note that any other maximal ideal can also not contain any units, so it is a subset of M. This is a contradiction, so M is the unique maximal ideal of R.

Now it only remains to show that M is principal. Let $\pi \in M$ such that $v(\pi)$ is minimal. Let $a \in M$ and note that $v(\pi^{-k}a) = v(a) - kv(\pi)$ by induction, for all $k \in \mathbb{Z}_{\geq 0}$. Therefore $v(\pi^{-k}a)$ is decreasing as k increases. Then there exists a fixed minimal $k \in \mathbb{Z}_{\geq 0}$ such that $v(\pi^{-(k+1)}a) \leq 0$, so $v(\pi^{-k}a) > 0$.

If $v(\pi^{-k}a) < v(\pi)$, then this is a contradiction with the minimality of $v(\pi)$. If $v(\pi^{-k}a) > v(\pi)$, then $v(\pi^{-(k+1)}a) > 0$ which is again a contradiction. Therefore $v(\pi^{-k}a) = v(\pi)$ and so $v(\pi^{-(k+1)}a) = 0$. But then $\pi^{-(k+1)}a$ must be equal to a unit $u \in R$. so we conclude that $u\pi^{k+1} = a$. This shows that $a \in \pi R$ and therefore $M \subseteq \pi R$. The other inclusion is immediate since $\pi \in M$. This shows that $M = \pi R$. □

Take F, v, R, M, π to be as in the proof of item 2 of the last proposition. We call π a *uniformising parameter* and say that $R/\pi R$ is the *residue field of* R. Note that the valuation v determines R and therefore determines $R/\pi R$. If F_2 and F_1 are fields such that $F_1 \subseteq F_2$ and v' and v are valuations of F_2 and F_1 respectively such that v' restricted to F_1 is v, we say that v' *is an extension of* v and write v'/v.

Definition 6.12. Define the *ring of integers of* K as

$$\mathcal{O}_K = \{a \in K \mid \text{there exists } f(X) \in \mathbb{Z}[X] \text{ such that } f(X) \text{ is monic and } f(a) = 0\}.$$

We say that $\mathcal{O}_K$ is the *integral closure of* $\mathbb{Z}$ *in* K.

We know that the integral closure of a ring in another ring is actually a subring (see for example [3, section 1.9] or [5, section 15.3, corollary 24]). Note that every $n \in \mathbb{Z}$ is a root of $X - n \in \mathbb{Z}[X]$, so $\mathbb{Z} \subseteq \mathcal{O}_K$. The idea behind this construction is to generalise the notion of integers in some sense. The following lemma makes it clear why we consider this to be a generalisation.

Lemma 6.13. *1. We have that* $\mathcal{O}_{\mathbb{Q}} = \mathbb{Z}$.

2. The fraction field of $\mathcal{O}_K$ *is* K.

Proof. 1. Suppose that $x \in \mathbb{Q}$ is a root of $f(X) \in \mathbb{Z}[X]$ and $f(X)$ is monic and irreducible. A direct corollary of Gauss' lemma ([5, section 9.3, corollary 6]) tells us that $f(X)$ is also irreducible over $\mathbb{Q}$, so it is the minimal polynomial of x. But then the degree of $f(X)$ is equal to one, so x is equal to the constant term which is an integer by definition. Note that if $f(X)$ is not irreducible, we can factor it and repeat the argument on the irreducible factors that have x as a root. This shows that $\mathcal{O}_{\mathbb{Q}} = \mathbb{Z}$.

2. Let $a \in \mathcal{O}_K$. The extension $K/\mathbb{Q}$ is algebraic so a is the root of some monic $f(X) \in \mathbb{Q}[X]$. By multiplying $f(X)$ by all the denominators of the coefficients, we get a polynomial with integer coefficients which we write as $\sum_{i=0}^{n} c_i X^i$. Remember that a is still a root of this, so

$$0 = c_n^{n-1}(c_n a^n + \cdots + c_1 a + c_0) = (c_n a)^n + \ldots c_n^{n-2} c_1 (c_n a) + c_n^{n-1} c_0.$$

This shows that $c_n a \in \mathcal{O}_K$. Next note that $c_n \in \mathbb{Z} \setminus \{0\} \subseteq K^\times$, so $a = \frac{c_n a}{c_n}$. Since $c_n \in \mathcal{O}_K$, a can be written as a fraction of elements of $\mathcal{O}_K$ and therefore K is the fraction field of $\mathcal{O}_K$. □

So the ring of integers of $\mathbb{Q}$ is $\mathbb{Z}$ and general rings of integers also preserve the property that their fraction fields are the whole field K again. Note that we have actually shown something stronger. We can assume that any element of K has a denominator in $\mathbb{Z}$. Unfortunately, there are also some properties of $\mathbb{Z}$ that $\mathcal{O}_K$ does not have. A very important one is that $\mathbb{Z}$ is a unique factorisation domain (UFD). We used this implicitly in the proof of the first item of the lemma. Depending on K, $\mathcal{O}_K$ might or might not have this property. For example, one can check that $\mathbb{Z}[\sqrt{-5}]$ is the ring of integers of $\mathbb{Q}(\sqrt{-5})$ and that $2 \cdot 3 = (1+\sqrt{-5})(1-\sqrt{-5})$. One can also check that $2, 3, (1+\sqrt{-5})(1+\sqrt{-5})$ and $(1-\sqrt{-5})$ are all irreducible in $\mathbb{Z}[\sqrt{-5}]$, so $\mathbb{Z}[\sqrt{-5}]$ does not have unique factorisation. But as it turns out, ideals can be factored in an analogous way! This is made precise in the next theorem.

Theorem 6.14. *Let* I *be a proper ideal of* $\mathcal{O}_K$. *Then there exist prime ideals* $\mathfrak{p}_1, \ldots, \mathfrak{p}_n$ *of* $\mathcal{O}_K$ *such that* $I = \mathfrak{p}_1 \ldots \mathfrak{p}_n$. *These prime ideals are unique up to their order.*

Proof. For a proof, see [5, section 16.3, corollary 16]. □

More generally, if an integral domain R has the unique factorisation of theorem 6.14, we say that R is a *Dedekind domain*. For two ideals $I, J \subseteq R$, we say that I divides J if there exists an ideal I' such that $II' = J$. We write $I \mid J$.

Proposition 6.15. *Let R be a Dedekind domain and I, J be two ideals. Then $I \subseteq J$, if and only if $J \mid I$.*

Proof. For a proof, see [5, section 16.3, proposition 17]. □

Theorem 6.14 is in some sense a generalisation of the factorisation into prime numbers in $\mathbb{Z}$. To see this, note that for $a, b \in \mathbb{Z}$ we have that $(a\mathbb{Z})(b\mathbb{Z}) = (ab)\mathbb{Z}$, so $n\mathbb{Z} = (p_1\mathbb{Z})^{e_1} \dots (p_k\mathbb{Z})^{e_k}$ where $p_1^{e_1} \dots p_k^{e_k}$ is the prime factorisation of n. But we know that $p \mapsto p\mathbb{Z}$ is a bijection between the prime numbers (up to units of course) and of proper prime ideals of $\mathbb{Z}$, so this actually gives us the factorisation of $n\mathbb{Z}$. We will now introduce a very special family of valuations on the number field K. To do this we first define a family of valuations on $\mathcal{O}_K$.

Proposition 6.16. *Let $\mathfrak{p}$ be a proper prime ideal of $\mathcal{O}_K$. For $x \in \mathcal{O}_K \setminus \{0\}$ define $v_\mathfrak{p}(x)$ as the largest integer n such that $\mathfrak{p}^n \mid x\mathcal{O}_K$ and $v_\mathfrak{p}(0) = \infty$. Here we define $\mathfrak{p}^0 = \mathcal{O}_K$. This makes $v_\mathfrak{p}$ into a discrete valuation.*

Proof. First we note that because of theorem 6.14 this function is well-defined. We now have to check that the three conditions in definition 6.10 hold. The first one holds by definition. Let $a, b \in \mathcal{O}_K$. Note that if a or b is zero, the conditions hold trivially so assume that $a \neq 0$ and $b \neq 0$.

Write $a\mathcal{O}_K = \mathfrak{p}^{a_1}\mathfrak{p}_1^{e_1} \dots \mathfrak{p}_n^{e_n}$ and $b\mathcal{O}_K = \mathfrak{p}^{b_1}\mathfrak{p}_1^{f_1} \dots \mathfrak{p}_n^{f_n}$, for some $n \in \mathbb{Z}_{>0}$ and $e_i, f_i \geq 0$ for all $1 \leq i \leq n$, for the factorisation of $a\mathcal{O}_K$ and $b\mathcal{O}_K$. Note that if a or b is a unit, this is still fine since then all exponents will be zero. We see that $ab\mathcal{O}_K \subseteq (a\mathcal{O}_K)(b\mathcal{O}_K)$. Every element of $(a\mathcal{O}_K)(b\mathcal{O}_K)$ can be written in the form $\sum_l ar_lbr'_l = ab\sum_l r_lr'_l$ and therefore the reverse inclusion also holds. We can now write $(ab)\mathcal{O}_K = (a\mathcal{O}_K)(b\mathcal{O}_K) = \mathfrak{p}^{a_1+b_1}\mathfrak{p}_1^{e_1+f_1} \dots \mathfrak{p}_n^{e_n+f_n}$. This shows that $v_\mathfrak{p}(ab) = a_1 + b_1 = v_\mathfrak{p}(a) + v_\mathfrak{p}(b)$.

Write k for $\min\{v_\mathfrak{p}(a), v_\mathfrak{p}(b)\}$. We certainly have that $\mathfrak{p}^k \mid a\mathcal{O}_K$ and $\mathfrak{p}^k \mid b\mathcal{O}_K$, so by proposition 6.15 $a\mathcal{O}_K, b\mathcal{O}_K \subseteq \mathfrak{p}^k$. This then means that $a, b \in \mathfrak{p}^k$ and therefore $a+b \in \mathfrak{p}^k$. Now $(a+b)\mathcal{O}_K \subseteq \mathfrak{p}^k$, so if $a+b \neq 0$ we can now use the same proposition as before to conclude that $\mathfrak{p}^k \mid (a+b)\mathcal{O}_K$. Then $v_\mathfrak{p}(a+b) \geq k$. Note that if $a + b = 0$, the inequality follows immediately. This shows that $v_\mathfrak{p}$ satisfies all three conditions of a discrete valuation. □

Note that for a proper prime ideal $\mathfrak{p}$ of $\mathcal{O}_K$, we can extend the valuation $v_\mathfrak{p}$ to K by using item 2 of lemma 6.13. Let $x \in K^\times$ and write $x = \frac{a}{b}$ with $a, b \in \mathcal{O}_K \setminus \{0\}$. Then we define $v_\mathfrak{p}(x) = v_\mathfrak{p}(a) - v_\mathfrak{p}(b)$. To see that this is well defined, consider $\alpha, \beta \in \mathcal{O}_K \setminus \{0\}$ such that $\frac{\alpha}{\beta} = \frac{a}{b}$. Then $\alpha b = a\beta$ and therefore $v_\mathfrak{p}(\alpha) - v_\mathfrak{p}(\beta) = v_\mathfrak{p}(a) - v_\mathfrak{p}(b)$ so $v_\mathfrak{p}$ is well-defined on K. That this extension still has the properties of a valuation is an easy check from the definitions. In the case of $K = \mathbb{Q}$, we can write every element as $p^n\frac{a}{b}$, where p is a prime that divides neither a nor b and $n \in \mathbb{Z}$. Then $v_{p\mathbb{Z}}(p^n\frac{a}{b}) = n$. This is called the *p-adic valuation*.

An absolute value $|\cdot|$ on a field K is called *non-archimedean* if $|\alpha + \beta| \leq \max\{|\alpha|, |\beta|\}$. It is immediate that if v is a valuation on K, then $\lambda^{v(-)}$ defines a non-archimedean absolute value on K. Here $0 < \lambda < 1$. We don't consider the trivial absolute value, with $|a| = 1$ for all $a \in K^\times$. It is known that, up to equivalence, all non-archimedean absolute values on K are of the form $\lambda^{v_\mathfrak{p}(-)}$ for some proper prime ideal $\mathfrak{p}$ of $\mathcal{O}_K$. The archimedean absolute values correspond to real and complex conjugate embeddings of K (see for example [7, section 7, page 111]). Remember that two absolute values $|\cdot|$ and $|\cdot|'$ are equivalent if their exist some $\alpha \in \mathbb{R}_{>0}$ such that $|\cdot|^\alpha = |\cdot|'$. From now on we will use the following notation. We write M_K^0 for the set of non-archimedean absolute values on K given by $\lambda^{v_\mathfrak{p}(-)}$, where $v_\mathfrak{p}$ is a proper prime ideal K and $\lambda \in \mathbb{R}$ such that $0 < \lambda < 1$. We write M_K^∞ for a complete set of inequivalent archimedean absolute values on K, which is finite ([7, section 4, page 70]), and M_K for their union. We write $K_{v_\mathfrak{p}}$ for the completion of K with respect to the absolute value $\lambda^{v_\mathfrak{p}(-)} \in M_K^0$, where $\mathfrak{p}$ is a proper prime ideal of $\mathcal{O}_K$. This construction is completely analogous to the construction of $\mathbb{R}$ from $\mathbb{Q}$, by considering equivalence classes of Cauchy sequences. From now on we will consider M_K^0 as a set of valuations by the correspondence from before.

Lemma 6.17. $\mathcal{O}_K = \{a \in K \mid v_\mathfrak{p}(a) \geq 0 \text{ for all } v_\mathfrak{p} \in M_K^0\}$

Proof. For $a \in \mathcal{O}_K$, it is clear by the definition that $v_\mathfrak{p}(a) \geq 0$ for all $v_\mathfrak{p} \in M_K^0$. Now let $a \in K$ and suppose that $v_\mathfrak{p}(a) \geq 0$ for all proper prime ideals $\mathfrak{p}$ of $\mathcal{O}_K$. We know from the second item of lemma 6.13 that $a = \frac{r_1}{r_2}$, where $r_1, r_2 \in \mathcal{O}_K$ and $r_2 \neq 0$. If $r_1 = 0$, then $a = 0$ and we are done. So assume that $r_1 \neq 0$. We see that $v_\mathfrak{p}(r_1) \geq v_\mathfrak{p}(r_2)$ for all $v_\mathfrak{p} \in M_K^0$. In particular, we can write

$$r_1\mathcal{O}_K = (r_2\mathcal{O}_K) \prod_{\mathfrak{p} \text{ prime}} \mathfrak{p}^{v_\mathfrak{p}(r_1) - v_\mathfrak{p}(r_2)},$$

where this product is well-defined since if $v_\mathfrak{p}(r_1) - v_\mathfrak{p}(r_2) \neq 0$, then $\mathfrak{p}$ appears in the (finite) prime ideal decomposition of r_1. We denote this product by I. From the definitions for ideal multiplication it follows that $(r_2\mathcal{O}_K)I = r_2 I$. This shows that $r_1\mathcal{O}_K = r_2 I$ and therefore $r_1 = r_2 i$ for some $i \in I$. But I is an ideal in $\mathcal{O}_K$ so we have shown that $a \in \mathcal{O}_K$. We conclude that for some $a \in K$, $a \in \mathcal{O}_K$ if and only if $v_\mathfrak{p}(a) \geq 0$ for all $v_\mathfrak{p} \in M_K^0$. □

Definition 6.18. Let R be an integral domain with fraction field F. Suppose that $S \subseteq R$ such that $0 \notin S$, $1 \in S$, and if $a, b \in S$, then $ab \in S$. We say that S is *multiplicative*. For such a multiplicative subset of R, we define $S^{-1}R = \{\frac{\alpha}{\beta} \in F \mid \beta \in S\}$ and call this the *localisation of R at S*.

Proposition 6.19. *Let R be an integral domain with fraction field F and S a multiplicative subset of R. Then $S^{-1}R$ is a subring of F. Furthermore, the map $\mathfrak{p} \mapsto \mathfrak{p} \cap R$ is bijection between the prime ideals of $S^{-1}R$ and the prime ideals of R that are disjoint from S, with $\mathfrak{p} \mapsto \mathfrak{p}S^{-1}R$ as the inverse.*

Proof. Let $x, y \in S^{-1}R$, so we can write $x = \frac{a}{b}$ and $y = \frac{c}{d}$ with $a, c \in R$ and $b, d \in S$. Then $x - y = \frac{ad-bc}{bd} \in S^{-1}R$, since S is multiplicative and R a ring. It is also immediate that $xy = \frac{ac}{bd} \in S^{-1}R$. Since R and S are non-empty, so is $S^{-1}R$. This shows that $S^{-1}R$ is a subring of F. Now we will show that if $\mathfrak{p}$ is a prime ideal of $S^{-1}R$, then $\mathfrak{p} \cap R$ is indeed an prime ideal of R that is disjoint from S, and if $\mathfrak{p}$ is a prime ideal of R that is disjoint from S, then $\mathfrak{p}S^{-1}R$ is a prime ideal of $S^{-1}R$.

Let $\mathfrak{p}$ be a prime ideal of $S^{-1}R$. The inclusion $\iota\colon R \to S^{-1}R$ is a ring homomorphism, so $\iota^{-1}(\mathfrak{p}) = \mathfrak{p} \cap R$ is a prime ideal of R. Suppose that $\mathfrak{p} \cap R$ contains $s \in S$. Then $s \in \mathfrak{p}$ and $\frac{1}{s} \in S^{-1}R$, so $1 \in \mathfrak{p}$. Therefore $\mathfrak{p} = S^{-1}R$ which is a contradiction with the fact that it is prime, so $S \cap \mathfrak{p} \cap R = \emptyset$. Next suppose that $\mathfrak{p}$ is a prime ideal of R disjoint from S and let $\alpha\beta \in \mathfrak{p}S^{-1}R$, where $\alpha, \beta \in S^{-1}R$. So $\alpha = \frac{a}{a'}$ and $\beta = \frac{b}{b'}$ with $a, b \in R$ and $a', b' \in S$ and we can write

$$\frac{ab}{a'b'} = \sum_{i=1}^{n} p_i \frac{c_i}{c_i'}.$$

Here for all $1 \leq i \leq n$, $p_i \in \mathfrak{p}$ $c_i \in R$ and $c_i' \in S$. Then we get that

$$abc_1' \dots c_n' = \sum_{i=1}^{n} p_i c_i r_i,$$

with $r_i \in R$ for all $1 \leq i \leq n$. The right hand side of this equation is clearly contained in $\mathfrak{p}$ and therefore the left hand side is too. But $c_1' \dots c_n'$ is an element of S and can therefore not be contained in $\mathfrak{p}$. Since $\mathfrak{p}$ is prime, $ab \in \mathfrak{p}$. Assume now without loss of generality that $a \in \mathfrak{p}$. Since $\frac{1}{a'} \in S^{-1}R$, we see that $\alpha \in \mathfrak{p}S^{-1}R$. This shows that $\mathfrak{p}S^{-1}R$ is prime.

Let J be an ideal in $S^{-1}R$. We will show that $(J \cap R)S^{-1}R = J$. We have that $(J \cap R)S^{-1}R \subseteq J$ because J is an ideal that contains $J \cap R$ and $(J \cap R)S^{-1}R$ is the smallest ideal that contains $J \cap R$. Now let $\alpha \in J$, so we can write $\alpha = \frac{a}{c}$ where $a \in R$ and $c \in S$. Since $c \in S^{-1}R$, $c\alpha \in J$ because J is an ideal. But we also assumed that $c\alpha = a \in R$, so $a \in J \cap R$. We have that $\frac{1}{c} \in S^{-1}R$, so we conclude that $\frac{a}{c} \in (J \cap R)S^{-1}R$. The inclusion $\iota\colon R \to S^{-1}R$ is a homomorphism, so $\iota^{-1}(J) = J \cap R$ is an ideal of R. In particular, we see that every ideal of $S^{-1}R$ can be written as $IS^{-1}R$ for some ideal I of R. For a prime ideal $\mathfrak{p} \in S^{-1}R$, we then have that $(\mathfrak{p} \cap R)S^{-1}R = \mathfrak{p}$.

For the other direction, suppose that $\mathfrak{p}$ is a prime ideal of R that is disjoint from S. It is immediate that $\mathfrak{p} \subseteq (\mathfrak{p}S^{-1}R) \cap R$. Now suppose that $a \in (\mathfrak{p}S^{-1}R) \cap R$. Then, similar to before, we can write $a = \sum_{i=1}^{n} p_i \frac{c_i}{c_i'}$ and therefore $a = \frac{p}{c_1' \dots c_n'}$, with $p \in \mathfrak{p}$ and $c = c_1' \dots c_n' \in S$. Then either $c \in \mathfrak{p}$ or $\frac{p}{c} \in \mathfrak{p}$. But $c \in S$, so $\frac{p}{c} = a \in \mathfrak{p}$. This shows the reverse inclusion so we conclude that $\mathfrak{p} = (\mathfrak{p}S^{-1}R) \cap R$. Now we have proved that the map $\mathfrak{p} \mapsto \mathfrak{p} \cap R$ is a bijection between the prime ideals of $S^{-1}R$ and the prime ideals of R that are disjoint with S, with $\mathfrak{p} \mapsto \mathfrak{p}S^{-1}R$ as the inverse. □

Definition 6.20. A *fractional ideal* of K is a $\mathcal{O}_K$-submodule I of K, such that there exists $r \in \mathcal{O}_K \setminus \{0\}$ with $rI \subseteq \mathcal{O}_K$.

It is immediate that a fractional ideal of K is an ideal of $\mathcal{O}_K$ if and only if it is contained in $\mathcal{O}_K$. To avoid confusion, we will always say "fractional ideal" when we talk about a fractional ideal and not shorten it to just "ideal".

Proposition 6.21. *Define I_K as the set of non-zero fractional ideals of K. Then I_K is an abelian group under the operation $IJ := \{\sum_{k=1}^{n} i_k j_k \mid n \in \mathbb{Z}_{\geq 1}$ and $i_k \in I, j_k \in J$ for all $1 \leq k \leq n\}$, and $P_K := \{a\mathcal{O}_K \mid a \in K^\times\}$ is a subgroup.*

Proof. Let $I, J \in I_K$, so we have $r_1, r_2 \in \mathcal{O}_K \setminus \{0\}$ with $r_1 I, r_2 J \subseteq \mathcal{O}_K$. It is immediate that the set IJ is again a $\mathcal{O}_K$-submodule of K and that $r_1 r_2 IJ \subseteq \mathcal{O}_K$. This shows that $IJ \in I_K$. It is also clear the operation is abelian. We will now show that the group axioms hold. In the same way as with regular ideals, one can check that the multiplication of fractional ideals is associative. Clearly $\mathcal{O}_K \in I_K$ and $\mathcal{O}_K I = I$, so $\mathcal{O}_K$ will be the identity of the group.
It just remains to show that every fractional ideal has an inverse. Let $I \in I_K$. Then there exists $d \in \mathcal{O}_K \setminus \{0\}$ such that $dI \subseteq \mathcal{O}_K$. In particular, this means that dI is a non-zero ideal of $\mathcal{O}_K$. Now let $a \in I \setminus \{0\}$, so we have that $da \in \mathcal{O}_K$. Then $(da)\mathcal{O}_K \subseteq dI$, so by proposition 6.15 there exists a non-zero ideal J of $\mathcal{O}_K$ such that $(dI)J = (da)\mathcal{O}_K$. It follows from the respective definitions that $(dI)J = d(IJ)$ and for any $b, c \in K$ that $b(cI) = (bc)I$. It then follows that $a^{-1}(IJ) = \mathcal{O}_K$. It again follows immediately from the respective definitions that $a^{-1}(IJ) = I(a^{-1}J)$. Since J is a non-zero ideal of $\mathcal{O}_K$, it is clear that $a^{-1}J$ is a non-zero $\mathcal{O}_K$-submodule of K. Furthermore, we can write $a^{-1} = \frac{r_1}{r_2}$ where $r_1, r_2 \in \mathcal{O}_K \setminus \{0\}$, so it follows that $r_2(a^{-1}J) \subseteq \mathcal{O}_K$. This shows that $a^{-1}J \in I_K$. In conclusion, we have shown that $I(a^{-1}J) = \mathcal{O}_K$ where $a^{-1}J \in I_K$, and therefore I has an inverse.

Now we show that P_K is a subgroup of I_K. Let $a, b \in K^\times$. In the same way as for ideals, we see that $(a\mathcal{O}_K)(b\mathcal{O}_K) = (ab)\mathcal{O}_K$. We note that $ab\mathcal{O}_K \neq 0\mathcal{O}_K$, so P_K is closed under multiplication. We now see that $(a\mathcal{O}_K)(a^{-1}\mathcal{O}_K) = \mathcal{O}_K$, so P_K is also closed under inverses. Since $\mathcal{O}_K \in P_K$, we can conclude that $P_K \leq I_K$. □

The quotient $Cl(K) = I_K/P_K$ is called the *ideal class group* of K and is a very important object in algebraic number theory. Note that in the proof we have only used that $\mathcal{O}_K$ is a Dedekind domain, so one can also define the ideal class group for Dedekind domains and their fraction fields. We only used this property to show that every element in I_K has an inverse, so it is not hard to check that the set of non-zero invertible fractional ideals forms a group for every integral domain. The principal fractional ideals will again be a subgroup of this, so in this way we can also define the ideal class group of an integral domain. Another object that is of interest to us, is the *ring of S-integers* $R_S = \{\alpha \in K \mid v(\alpha) \geq 0$ for all $v \in M_K \setminus S\}$. Here we always assume that $S \subseteq M_K$ such that S is finite and $M_K^\infty \subseteq S$. We will now state two important facts about these objects, which we will need later on.

Theorem 6.22. *1. The ideal class group of a number field is finite.*

2. The group of units of the ring of S-integers is finitely generated.

Proof. 1. For a proof, see [7, section 4, theorem 4.4].

2. For a proof, see [7, section 5, theorem 5.11]. □

We can use immediately use the finiteness of the ideal class group to prove the following result.

Proposition 6.23. *Let $S \subseteq M_K$ be finite with $M_K^\infty \subseteq S$ and $Cl(K) = \{[I_1], \dots [I_n]\}$ for some $n \in \mathbb{Z}_{\geq 1}$ with $I_i \subseteq \mathcal{O}_K$ for all $1 \leq i \leq n$. Let P be the set of valuations that correspond to the prime ideals that divide $\prod_{i=1}^n I_i$. If $P \subseteq S$, then R_S is a principal ideal domain (PID).*

Proof. Note that if I is a fractional ideal of $\mathcal{O}_K$ such that $rI \subseteq \mathcal{O}_K$ with $r \in \mathcal{O}_K$, then $(r\mathcal{O}_K)I = rI$ by an argument similar to that used in proposition 6.21 for the closure of P_K under multiplication. In particular $rI \in [I]$ and therefore we can always choose our representatives of the ideal class group to be ideals. Now note that $\mathcal{O}_K \subseteq R_S$, by lemma 6.17. Define $C = \bigcap_{\mathfrak{p} \in C'} (\mathcal{O}_K \setminus \mathfrak{p})$, where C' is the set of proper prime ideals of $\mathcal{O}_K$ such that their associated valuations are not element of S. We directly see that $C' \neq \emptyset$, so $0 \notin C$. All the prime ideals in C' are proper, so $1 \in C$. We also see that if $a, b \in C$ then they lie outside all the primes in C', so ab also lies outside all the primes in C' by the definition of a prime ideal. This then shows that C is multiplicative.

We will show that $C^{-1}\mathcal{O}_K = R_S$. First suppose that $x \in C^{-1}\mathcal{O}_K$. Then we can write $x = \frac{\alpha}{\beta}$ with $\alpha, \beta \in \mathcal{O}_K$ and $\beta \in C$. Let $\mathfrak{p}$ be a prime ideal of $\mathcal{O}_K$ such that the associated valuation is not an element of S. Then $\mathfrak{p} \in C'$, so $\beta \notin \mathfrak{p}$. This means that $v_\mathfrak{p}(\beta) = 0$, so $v_\mathfrak{p}(\frac{\alpha}{\beta}) = v_\mathfrak{p}(\alpha) \geq 0$, since $\alpha \in \mathcal{O}_K$. This shows that $x \in R_S$. For the reverse inclusion, suppose that $x \in R_S$. If $\mathfrak{p}$ is a prime ideal of $\mathcal{O}_K$ such that $v_\mathfrak{p}(x) < 0$, then certainly $v_\mathfrak{p} \in S$. Let $\mathfrak{p}$ be such an ideal. Since the ideal class group is finite, we have that $[\mathfrak{p}]^{\#Cl(K)} = [\mathcal{O}_K]$. This then means that $\mathfrak{p}^{\#Cl(K)} = a_\mathfrak{p}\mathcal{O}_K$ for some $a_\mathfrak{p} \in K^\times$. But $\mathfrak{p}^{\#Cl(K)} \subseteq \mathfrak{p} \subseteq \mathcal{O}_K$ and $a_\mathfrak{p} \in a_\mathfrak{p}\mathcal{O}_K$, so $a_\mathfrak{p} \in \mathcal{O}_K$. Define

$$a = \prod_{\substack{\mathfrak{p} \text{ prime} \\ v_\mathfrak{p}(x) < 0}} a_\mathfrak{p}^{-v_\mathfrak{p}(x)},$$

so we see that

$$a\mathcal{O}_K = \prod_{\substack{\mathfrak{p} \text{ prime} \\ v_\mathfrak{p}(x) < 0}} (a_\mathfrak{p}\mathcal{O}_K)^{-v_\mathfrak{p}(x)} = \prod_{\substack{\mathfrak{p} \text{ prime} \\ v_\mathfrak{p}(x) < 0}} \mathfrak{p}^{-v_\mathfrak{p}(x)\#Cl(K)}.$$

Since S is finite, this product is also finite and well-defined. It is immediate that $a \in \mathcal{O}_K$. Suppose now that $a \in \mathfrak{p}$, with $\mathfrak{p} \in C'$. Then certainly $v_\mathfrak{p} \notin S$, but also $a\mathcal{O}_K \subseteq \mathfrak{p}$, so with proposition 6.15 $\mathfrak{p}$ divides $a\mathcal{O}_K$. But this is a contradiction with the fact that the factorisation of $a\mathcal{O}_K$ into prime ideals is unique, since we have a factorisation where for every prime divisor $\mathfrak{p}'$, $v_{\mathfrak{p}'} \in S$. This shows that $a \notin \mathfrak{p}$ and therefore we conclude that $a \in C$. Let $\mathfrak{p}$ be a prime ideal of $\mathcal{O}_K$. If $v_\mathfrak{p}(x) \geq 0$, then $v_\mathfrak{p}(ax) = v_\mathfrak{p}(a) + v_\mathfrak{p}(x) \geq 0$ since $a \in \mathcal{O}_K$. If $v_\mathfrak{p}(x) < 0$, then $v_\mathfrak{p}(a) = -v_\mathfrak{p}(x)\#Cl(K)$. Therefore $v_\mathfrak{p}(ax) = v_\mathfrak{p}(a) + v_\mathfrak{p}(x) = -v_\mathfrak{p}(x)(\#Cl(K) - 1) \geq 0$. This shows that $v_\mathfrak{p}(ax) \geq 0$ for all prime ideals $\mathfrak{p}$, so we can conclude that $ax \in \mathcal{O}_K$. Since $a\mathcal{O}_K$ can be written as a product of prime ideals, we certainly have that $a \neq 0$. Now we see that $x = \frac{ax}{a}$ and we conclude that $x \in c^{-1}\mathcal{O}_K$. This shows that $C^{-1}\mathcal{O}_K = R_S$.

We are now ready to show that every ideal is principal. Let J be an ideal of R_S. From the proof of proposition 6.19, we know that there exists an ideal I of $\mathcal{O}_K$, namely $J \cap \mathcal{O}_K$, such that $J = IR_S$. Then $I \in [I_i]$ for some $1 \leq i \leq n$. In particular then $I = rI_i$ for some $r \in K^\times$. Therefore we see that $J = IR_S = (rI_i)R_S = r(I_iR_S)$. This last equality is again immediate after writing out the definitions. We will now show that $I_iR_S = R_S$. Since $I_i \subseteq \mathcal{O}_K$, we get that $I_iR_S \subseteq R_S$. In the same way as earlier, $[I_i]^{\#Cl(K)} = [\mathcal{O}_K]$ so $I_i^{\#Cl(K)} = x\mathcal{O}_K$ for some $x \in k^\times$. But then $x \in I_i^{\#Cl(K)} \subseteq I_i \subseteq \mathcal{O}_K \subseteq R_S$. Now assume that $P \subseteq S$. Suppose that there is some valuation $v_\mathfrak{p} \notin S$ such that $v_\mathfrak{p}(x) > 0$. But then $\mathfrak{p}$ divides $x\mathcal{O}_K$ by the definition of the valuation. Now $v_\mathfrak{p} \in P \subseteq S$, but this is a contradiction. Therefore $v_\mathfrak{p}(x) = 0$ for all $v_\mathfrak{p} \notin S$. This shows $x \in R_S^\times$. But we already know that $x \in I_i$, so I_iR_S is an ideal that contains a unit. We conclude that $I_iR_S = R_S$ and therefore $J = rR_S$. This shows that every ideal of R_S is principal so R_S is a PID. □

We will show one more property of the extension L/K, and after that we have gathered all the tools we need to prove the weak Mordell-Weil theorem.

Definition 6.24. Take F, v, R, M, π as in the proof of lemma 6.11. For a field extension F' with F'/F Galois and valuation v' that extends v, define R', M', π' analogously. Then the *inertia group of F'/F with respect*

to v'/v, is defined to be $I_{v'/v} = \{\sigma \in \mathrm{Gal}(F'_{v'}/F_v) \mid x^\sigma \in [x] \in R'_{v'}/\pi' R'_{v'} \text{ for all } x \in R'_{v'}\}$. If the inertia group is trivial we say that the extension is *unramified at* v'/v.

Note that for $\sigma, \tau \in I_{v'/v}$, certainly $\sigma\tau \in I_{v'/v}$. Also suppose that $x \in R'_{v'}$. Then $x^\sigma \in [x] \in R'_{v'}/\pi' R'_{v'}$ and therefore $x^\sigma = x + \pi' r'$ for some $r' \in R'_{v'}$. This means that $x^{\sigma^{-1}} = x - \pi' r'$, so $x^{\sigma^{-1}} \in I_{v'/v}$. Since $id_{F'} \in I_{v'/v}$, we can conclude that $I_{v'/v}$ is a subgroup of $\mathrm{Gal}(F'/F)$.

Recall that A and B are the coefficients of the Weierstrass polynomial. Assume now that $A, B \in \mathcal{O}_K$. With this assumption, lemma 6.17 implies that $A, B \in \{a \in K \mid v(a) \geq 0\} =: R_v$, for all $v \in M_K^0$. We write $\tilde{K}$ for $R_v/\pi R_v$. Now let $r\colon R_v \to \tilde{K}$ be the standard projection. Then for a given $v \in M_K^0$, we can define a new polynomial $W_{v,A,B} = -Y^2Z + X^3 + r(A)XZ^2 + r(B)Z^3 \in \tilde{K}[X, Y, Z]$. If the associated projective curve $\tilde{E}(\tilde{K}) := V(W_{v,A,B})$ is non-singular, we see that is an elliptic curve. Note that if $v(\Delta) = 0$, then $[\Delta] \neq [0] \in \tilde{K}$, so the curve $\tilde{E}(\tilde{K})$ is non-singular. We say that $E(\bar{K})$ *has good reduction at* v. If the new curve becomes singular, we say that $E(\bar{K})$ *has bad reduction at* v. Recall that $\Delta = 4A^3 + 27B^2$. Then $v(\Delta) \neq 0$, only when the prime ideal in $\mathcal{O}_K$ that corresponds to v appears in the prime decomposition of $\Delta\mathcal{O}_K$. But this clearly happens only for finitely many prime ideals, so the set $\{v \in M_K^0 \mid E(\bar{K}) \text{ has bad reduction at } v\}$ is finite. In the same way we can argue that $\{v \in M_K^0 \mid v(m) \neq 0\}$ is finite. Now define the finite set

$$\mathcal{S} = \{v \in M_K^0 \mid E(\bar{K}) \text{ has bad reduction at } v\} \cup \{v \in M_K^0 \mid v(m) \neq 0\} \cup M_K^\infty.$$

In the final proof of the weak Mordell-Weil theorem, the idea will be to enlarge $\mathcal{S}$ so that it satisfies to properties of the set S in proposition 6.23. Then we can make use of proposition 6.23. Let $v \in M_K^0$ and suppose that $P = [x : y : z] \in E(K)$. Since one of x, y or z has to be non-zero, $\min\{v(x), v(y), v(z)\}$ exists. Without loss of generality, assume that $\min\{v(x), v(y), v(z)\} = v(z)$. Then $v(xz^{-1}), v(yz^{-1}) \geq 0$ and $P = [xz^{-1} : yz^{-1} : 1]$. This shows that P always has a representative where all coordinates have a non-negative valuation and one has a valuation of exactly zero. Putting a point $P \in E(K)$ in this form and then applying the reduction map on each coordinate defines a new map from $E(K)$ to $\mathbb{P}^2(\tilde{K})$, which we also denote by r. To show that this map is well-defined, suppose that $[x : y : z]$ and $[\lambda x : \lambda y : \lambda z]$ are two such representatives for P, for some $\lambda \in K^\times$. Assume without loss of generality that $v(z) = 0$. Then $v(\lambda z) = v(\lambda)$, but we also see that $v(\lambda x), v(\lambda y) \geq v(\lambda)$. This shows that $v(\lambda z) = \min\{v(\lambda x), v(\lambda y), v(\lambda z)\}$ and therefore $v(\lambda z) = 0$. This shows that $v(\lambda) = 0$, so $\lambda \notin R_v$. This means that $\lambda \in \tilde{K}^\times$. But this means that $[r(\lambda x) : r(\lambda y) : r(\lambda z)] = [r(\lambda)r(x) : r(\lambda)r(y) : r(\lambda)r(z)] = [r(x) : r(y) : r(z)]$ and therefore r is well-defined.

It is immediate that if $P \in E(K)$, then $r(P) \in \tilde{E}(\tilde{K})$. In fact, it is possible to show that r is a group homomorphism from $E(K)$ to $\tilde{E}(\tilde{K})$. For a proof of this, see proposition 2.1 in [1, section VII.2]. The idea is similar to that of the proof of proposition 3.10. We are now ready to prove the following proposition.

Proposition 6.25. *For any* $v \in M_K \setminus \mathcal{S}$*, the extension* L/K *is unramified.*

Proof. Let $v \in M_K \setminus \mathcal{S}$ and v' and extension to L. If L/K were not unramified, then there exists $\sigma \in \mathrm{Gal}(L_v/K_v)$ such that $x^\sigma \in [x] \in \tilde{L_v}$ for all $x \in L_v$ with $v(x) \geq 0$, and $\sigma \neq id_L$. But from the proof of theorem 3.5 we know that σ can be thought of the extension of some automorphism of some intermediate field of L_v and K_v, which is Galois over K_v. So if we choose this to be a finite extension that contains x, we see that is enough to prove the statement for a this extension. So we can choose a finite set of $Q \in E(\bar{K})$, with $mQ \in E(K)$ such that (the Galois closure of) their compositum contains x. Now corollary 7.3 in [8, section 7] tells us that it is enough to prove the statement for the extension $K(Q)/K$, where $Q \in E(\bar{K})$ and $mQ \in E(K)$.

Let $Q \in E(\bar{K})$ with $mQ \in E(K)$, v' be an extension of v to $K(Q)$ and $\sigma \in I_{v'/v}$. Now we see that

$$r(Q^\sigma - Q) = r(Q^\sigma) - r(Q),$$

since r is a group homomorphism. Suppose that $Q = [x : y : z]$ where $x, y, z \in K(Q)$, $v(x), v(y), v(z) \geq 0$ and without loss of generality $v(z) = 0$. Since $\sigma \in I_{v'/v}$ and $[z^\sigma] = [z] \neq [0]$, we see that also $v(z^\sigma) = 0$. It is immediate that $v(x^\sigma), v(y^\sigma) \geq 0$. Therefore

$$r(Q^\sigma) = [[x^\sigma] : [y^\sigma] : [z^\sigma]] = [[x] : [y] : [z]] = r(Q),$$

and so $r(Q^\sigma - Q) = r(O)$ We have also already seen in the proof of 6.3 that $m(Q^\sigma - Q) = O$. Now we can use proposition 1.4 in [1, section VIII.1] tells us that r is injective on the m-torsion kernel, so $Q^\sigma - Q = O$. Note that we have assumed that $v \notin \mathcal{S}$, so $E(\bar{K})$ has good reduction at v. Since $K(O) = K$ and the extension K/K is certainly unramified, we can assume that we can write $Q = [x : y : 1]$ where $x, y \in K(Q)$. Then $Q^\sigma = [x^\sigma : y^\sigma : 1]$, so $x^\sigma = x$ and $y^\sigma = y$. But by definition $K(Q) = K(x, y)$ and therefore we see that σ all of $K(Q)$. This means that $\sigma = id_{K(Q)}$. We conclude that $K(Q)/K$ is unramified. □

Now we are finally ready to prove that L/K is finite and therefore the weak Mordell-Weil theorem. We will use most of our previous results and finish of the proof by using item 2 of theorem 6.22.

Theorem 6.26 (weak Mordell-Weil)**.** *The group $E(K)/mE(K)$ is finite.*

Proof. If S is some subset of M_K that contains $\mathcal{S}$, then the maximal extension of K with properties 2 and 3 of lemma 6.8 that is unramified outside of S, say L' certainly contains L. Therefore if L' is finite, L is as well. We already saw in proposition 6.23 that we can take ideal representatives of each element in $Cl(K)$. We add the valuations corresponding to all the prime ideals that divide one of these representatives to the set $\mathcal{S}$ and call this set S. Since $Cl(K)$ and $\mathcal{S}$ are finite, S is as well. We have done this so we can make use of proposition 6.23 later in this proof. We now just have to prove that L' is finite. By definition, L' is the largest subfield of $K(\{\sqrt[m]{a} \mid a \in K\})$ that is unramified outside of S.

For an abelian group G (written multiplicatively), we write G^m for the subgroup of elements that are a m-th power. We define

$$A_S = \{[a] \in K^\times/(K^\times)^m \mid v(a) \equiv 0 \pmod m \text{ for all } v \in M_K \setminus S\}.$$

Note that if $[a] = [b] \in K^\times/(K^\times)^m$, then $a = bc^m$ for some $c \in K^\times$. In particular $v(a) = v(b) + mv(c)$, so $v(a) \equiv v(b) \pmod m$ and so A_S is well-defined. If $[a], [b] \in A_S$, then $v(ab) = v(a) + v(b) \equiv 0 \pmod m$ and $v(a^{-1}) = -v(a) \equiv 0 \pmod m$. Certainly $A_S \neq \emptyset$, since $[1] \in A_S$ and therefore A_S is a subgroup of $K^\times/(K^\times)^m$. If $a \in R_S^\times$, then certainly $v(a) \equiv 0 \pmod m$ for all $v \in M_K \setminus S$, so $[a] \in A_S$. This shows that the map $\pi\colon R_S^\times \to A_S\colon a \mapsto [a]$ is well-defined. It is immediate that π is a homomorphism. Note that also $K_v(\sqrt[m]{a}) = K_v(c\sqrt[m]{b}) = K_v(\sqrt[m]{b})$. This means that we only need to adjoin one representative for each class of $K^\times/(K^\times)^m$ to K, to get the field $K(\{\sqrt[m]{a} \mid a \in K\})$.

The set of roots of $X^n - a$ is $\left\{\zeta_n^k \sqrt[n]{a} \mid k \in \mathbb{Z} \text{ such that } 1 \leq k \leq n\right\}$. Recall that ζ_n is a primitive n-th root of unity. Then the discriminant of $X^n - a$ is, up to a sign, equal to

$$\prod_{k \neq l} (\zeta_n^k \sqrt[n]{a} - \zeta_n^l \sqrt[n]{a}),$$

for $1 \leq k, l \leq n$. There are $n^2 - n$ terms in this product, so it is equal to $a^{n-1} \prod_{k \neq l} (\zeta_n^k - \zeta_n^l)$. Now we can make use of the fact that the multiplicative group of n-th roots of unity is isomorphic to $\mathbb{Z}/n\mathbb{Z}$ and see that this product is equal to

$$a^{n-1} \prod_{k=1}^{n} \prod_{l=1}^{n-1} \left(\zeta_n^k - \zeta_n^{k+l}\right) = a^{n-1} \prod_{k=1}^{n} \left(\zeta_n^k \prod_{l=1}^{n-1} \left(1 - \zeta_n^l\right)\right).$$

Now note that as l ranges from 1 to $n-1$, ζ_n^l becomes all the n-th roots of unity except for 1. In this way we see that, as polynomials, $\prod_{l=1}^{n-1}(X - \zeta_n^l) = \frac{X^n - 1}{X - 1} = \sum_{l=0}^{n-1} X^l$. But then we immediately get that

$$a^{n-1} \prod_{k=1}^{n} \left(\zeta_n^k \prod_{l=1}^{n-1} \left(1 - \zeta_n^l\right)\right) = a^{n-1} \prod_{k=1}^{n} \left(\zeta_n^k n\right).$$

But, up to a n-th root of unity, this last expression is equal to $a^{n-1}n^n$. For any valuation $v \in M_K^0$, we see that $nv(\zeta_n^k) = v(1) = 0$, so $v(\mathrm{disc}(X^n - a)) = v(a^{n-1}n^n) = (n-1)v(a) + nv(n)$. Here $\mathrm{disc}(X^n - a)$ denotes

the discriminant of $X^n - a$.

Let $a \in K^\times$, so we can write $a = \frac{r}{n}$ for $r \in \mathcal{O}_K$ and $n \in \mathbb{Z}$ by the proof of lemma 6.13. But then $an^m \in \mathcal{O}_K$ and $an^m \in [a] \in K^\times/(K^\times)^m$, so for each class of $K^\times/(K^\times)^m$ we can choose a representative in $\mathcal{O}_K$. Let a be such a representative and t a uniformising parameter for the valuation $v \in M_K^0 \setminus S$. Then with the proof of item 2 of 6.11, we can write $a = ut^{nm+r}$, where $v(u) = 0$, $n, r \in \mathbb{Z}_{\geq 0}$ and $r < m$. We see that $v(a) = nmv(t) + v(t^r)$,. But clearly $[t^r] = [a] \in K^\times/(K^\times)^m$, so t^r and a generate the same extension of K. Now suppose that the extension $K(\sqrt[n]{t^r})/K$ is unramified. Then the valuation of the discriminant of the minimal polynomial is zero $(\dots)$. But the minimal polynomial is certainly $X^{m/d} - \sqrt[d]{t^r}$, where $d = \max\{n \in \mathbb{Z}_{>0} \mid n \text{ divides } m \text{ and } a^{1/n} \in K\}$ and we know that $\operatorname{disc}(X^{m/d} - \sqrt[d]{t^r}) = (m/d)^{m/d}(\sqrt[d]{t^r})^{m/d-1}$. Therefore $v(\operatorname{disc}(X^{m/d} - \sqrt[d]{t^r})) = m/dv(m/d) + (m/d-1)v(\sqrt[d]{t^r})$. But We see that $(m/d\mathcal{O}_K)(d\mathcal{O}_K) = m\mathcal{O}_K$, so by the unique factorisation into prime ideals we get that $v(m/d) = 0$. If $d \neq 1$, we can conclude that $v(\sqrt[d]{t^r}) = 0$. Then $v(t^r) = dv(\sqrt[d]{t^r}) = 0$. Note that if $d = m$, then the extension is trivial and we do not need to consider it. This means that $v(a) = nmv(t) \equiv 0 \pmod m$.

So if we adjoin one element of each class of A_S, we get a field that contains L'. Now if we show that A_S is finite, L'/K will be finite and we are done.

We will show that the map π surjective. To do this we first show that an element of $K^\times/(K^\times)^m$ has a representative in R_S. Now let $[a] \in K^\times/(K^\times)^m$ and write $a = \frac{b}{c}$ with $b, c \in \mathcal{O}_K \setminus \{0\}$. There are only finitely many proper prime ideals that divide $b\mathcal{O}_K$ or $c\mathcal{O}_K$ and therefore there are only finitely many $v \in M_K \setminus S$ such that $v(b) \neq 0$ or $v(c) \neq 0$. Call the finite collection of these valuations N. Then $\{v \in M_K \setminus S \mid v(a) \neq 0\} \subseteq N$. If there is no valuation $v \in M_K \setminus S$ such that $v(a) < 0$, $a \in R_S$ and we are done. So suppose there exist such valuations and let v be the one where $v(a)$ is the smallest. This is well-defined since N is finite. Then $v(a(a^{-1})^m) = v(a) - mv(a) > 0$, since $m \geq 2$. Note that then all other valuations also make $a(a^{-1})^m$ positive and therefore R_S. Furthermore it is clear that $[a] = [a(a^{-1})^m] \in K^\times/(K^\times)^m$.

Now we are ready to prove that π is surjective. Take C as in the proof of proposition 6.23 and recall that $R_S = C^{-1}\mathcal{O}_K$. Now let $\mathcal{A} \in A_S$ and choose $a \in \mathcal{A}$ such that $a \in R_S$. We can then write $a = \frac{r}{s}$, where $r \in \mathcal{O}_K$ and $s \in C$. Since $\frac{1}{s}$ is a unit in R_S, we see that $aR_S = rR_S$. By writing out the definitions, we see that $rR_S = (r\mathcal{O}_K)R_S$. Write

$$r\mathcal{O}_K = \mathfrak{p}_1^{v_{\mathfrak{p}_1}(r)} \dots \mathfrak{p}_n^{v_{\mathfrak{p}_n}(r)}$$

for the unique factorisation of $r\mathcal{O}_K$ into prime ideals. For two ideals I and J in $\mathcal{O}_K$, we can again write out the definitions and see that $(IR_S)(JR_S) = (IJ)R_S$. Using this repeatedly gives

$$(r\mathcal{O}_K)R_S = (\mathfrak{p}_1R_S)^{v_{\mathfrak{p}_1}(r)} \dots (\mathfrak{p}_nR_S)^{v_{\mathfrak{p}_n}(r)}.$$

Let $1 \leq i \leq n$ and suppose that there exists $c \in \mathfrak{p}_i \cap C$. Since $\frac{1}{c} \in R_S$, we see that $1 \in \mathfrak{p}_iR_S$ and therefore $\mathfrak{p}_iR_S = R_S$. But for an ideal I of R_S, we have that $IR_S = I$. This means that we can write (after renaming)

$$(r\mathcal{O}_K)R_S = (\mathfrak{p}_1R_S)^{v_{\mathfrak{p}_1}(r)} \dots (\mathfrak{p}_kR_S)^{v_{\mathfrak{p}_k}(r)},$$

where $k \leq n$ and $\mathfrak{p}_i \cap C = \emptyset$ for all $1 \leq i \leq k$. For all $1 \leq i \leq k$, we see that $v_{\mathfrak{p}_i}(a) = v_{\mathfrak{p}_i}(r) - v_{\mathfrak{p}_i}(s)$. Since $s \in C$, $s \notin \mathfrak{p}_i$ and therefore $s\mathcal{O}_K \not\subseteq \mathfrak{p}_i$. With proposition 6.15 we get that $\mathfrak{p}_i$ does not divide $s\mathcal{O}_K$ and therefore $v(s) = 0$. In particular $v_{\mathfrak{p}_i}(a) = v_{\mathfrak{p}_i}(r)$. Since $\mathfrak{p}_i \cap C = \emptyset$, $v_{\mathfrak{p}_i} \in M_K \setminus S$. But we also know that $[a] \in A_S$ and $a \in R_S$, so $v_{\mathfrak{p}_i}(r) = mn_i$ where $n_i \in \mathbb{Z}_{\geq 0}$. Putting all of this together, we get that

$$\begin{aligned} aR_S &= rR_S \\ &= (\mathfrak{p}_1R_S)^{v_{\mathfrak{p}_1}(r)} \dots (\mathfrak{p}_kR_S)^{v_{\mathfrak{p}_k}(r)} \\ &= ((\mathfrak{p}_1R_S)^{n_1})^m \dots ((\mathfrak{p}_kR_S)^{n_k})^m \\ &= ((\mathfrak{p}_1^{n_1} \dots \mathfrak{p}_k^{n_k})R_S)^m, \end{aligned}$$

where $n_1, \dots, n_k \in \mathbb{Z}_{\geq 0}$. At this point we can use proposition 6.23 to conclude that $(\mathfrak{p}_1^{n_1} \dots \mathfrak{p}_k^{n_k})R_S = bR_S$ for some $b \in R_S$. Now we see that $aR_S = (bR_S)^m = b^mR_S$. Since a and b^m generate the same ideal of R_S,

we can write $a = ub^m$ where $u \in R_S^\times$. But this means that $[a] = [u] \in A_S$, so $\pi(u) = [a]$. This shows that π is surjective.

It is clear that $(R_S^\times)^m \subseteq (K^\times)^m$, so $R_S^\times$ is contained in the kernel of π. Therefore the map $R_S^\times/(R_S^\times)^m \to A_S \colon [a] \mapsto \pi(a)$ is well-defined and of course still surjective. A set of generators of $R_S^\times$ will still be a set of generators of $R_S^\times/(R_S^\times)^m$. Item 2 from theorem 6.22 tells us that we can choose such a set to be finite. Since we took the quotient by $(R_S^\times)^m$, all these generators will have a finite order, namely an integer dividing m. But then it is immediate that $R_S^\times/(R_S^\times)^m$ is finite. Since the quotient map was surjective, we can conclude that A_S is finite. As we mentioned above, we now see that L'/K is finite and therefore L/K is finite. Now with item 1 of lemma 6.8, we conclude that $E(K)/mE(K)$ is finite. This concludes the proof of the weak Mordell-Weil theorem.

□

References

[1] Joseph H. Silverman. *The arithmetic of elliptic curves*. Second. Vol. 106. Graduate Texts in Mathematics. Springer, Dordrecht, 2009, pp. xx+513. ISBN: 978-0-387-09493-9. DOI: 10.1007/978-0-387-09494-6. URL: https://doi.org/10.1007/978-0-387-09494-6.

[2] James S. Milne. *Fields and Galois Theory (v5.10)*. Available at www.jmilne.org/math/. 2022.

[3] William Fulton. *Algebraic curves*. Advanced Book Classics. An introduction to algebraic geometry, Notes written with the collaboration of Richard Weiss, Reprint of 1969 original. Addison-Wesley Publishing Company, Advanced Book Program, Redwood City, CA, 1989, pp. xxii+226. ISBN: 0-201-51010-3.

[4] Stefan Kebekus. *Elliptic Curve Plotter*. June 18, 2023. URL: https://kebekus.gitlab.io/ellipticcurve/.

[5] David S. Dummit and Richard M. Foote. *Abstract algebra*. Third. John Wiley & Sons, Inc., Hoboken, NJ, 2004, pp. xii+932. ISBN: 0-471-43334-9.

[6] Robin Hartshorne. *Algebraic Geometry*. First. Springer Science+Business Media, Inc., 1997, pp. xvi+496. ISBN: 978-1-4757-3849-0.

[7] James S. Milne. *Algebraic Number Theory (v3.08)*. Available at www.jmilne.org/math/. 2020.

[8] Jürgen Neukirch. *Algebraic number theory*. Vol. 322. Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]. Translated from the 1992 German original and with a note by Norbert Schappacher, With a foreword by G. Harder. Springer-Verlag, Berlin, 1999, pp. xviii+571. ISBN: 3-540-65399-6. DOI: 10.1007/978-3-662-03983-0. URL: https://doi.org/10.1007/978-3-662-03983-0.