ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

EPFL - SMA Master Semester 2
Under the supervision of : Prof. Dr. Eva Bayer-Fluckiger
Assisted by : Dr. Mathieu Huruguen

The Mordell-Weil Theorem

Semester project presented by

David Ter-Borch Gram Schjoldager Lilienfeldt

April 24, 2017

Contents

Introduction

The study of Diophantine equations, that is, the problem of finding rational solutions to polynomial equations in several variables with integer coefficients, occupies a central place in Mathematics both by its long history, which can be traced back to ancient Greece, and by its apparent simplicity in spite of the complexity of the problems induced by these equations. To cite an example, there is the famous Fermat's Last Theorem proved in 1995 by Andrew Wiles. But even though the problems considered are ancient, Diophantine equations, and especially Diophantine geometry, are rich areas of research still today. The term Diophantine geometry describes the study of Diophantine equations through a combination of techniques from Algebraic Geometry and Algebraic Number Theory.

Consider an equation in two variables of the form $y^2 = P(x)$ where P is a polynomial in one variable. Such a polynomial relation defines a curve C_a in the affine plane. If the polynomial has rational coefficients, then one can ask for rational solutions to the equation $y^2 = P(x)$ with x, y in $\mathbb{Q}$. In other words, one asks for rational points of the curve C_a. Hence the problem becomes geometrical. The natural questions that arise are the following:

(i) *Are there any rational points ?*

(ii) *If yes, how many are there ? Finitely many or infinitely many ?*

One can consider the projective closure C of C_a, which is a plane projective curve. Suppose in this example that the curve C is smooth. Attached to this curve is a number g called the genus of the curve.

If P is a polynomial of degree 2, then the equation $y^2 = P(x)$ defines a conic in the plane and C is a curve of genus $g = 0$. If P is a polynomial of degree 3, then C is a curve of genus $g = 1$ and such a curve is known as an elliptic curve. In the case where P is of degree greater than or equal to 4, the curve C has genus $g \geq 2$.

The case $g = 0$ was proved by Hurwitz and Hilbert [HH] around 1890:

Theorem (Hurwitz-Hilbert)**.** *Let C be a curve of genus zero. If it has a rational point over $\mathbb{Q}$, then it has an infinite number of rational points over $\mathbb{Q}$.*

The case of $g \geq 2$ was conjectured by Mordell [Mor] in 1922 and proved by Faltings [Fal] in 1983.

Theorem (Faltings)**.** *Let C be a curve of genus greater than or equal to 2. Then $C(\mathbb{Q})$ is finite.*

This leaves us with the case when $g = 1$. In this case, the above considered curve C is a smooth plane projective curve of genus 1. It is what is called an elliptic curve. These curves are particular in the sense that the rational points on the curve can be given a group structure via a geometric group law. Around 1901, Poincaré [Poi] conjectured that the rational points of an elliptic curve formed a finitely generated group. This was proved in 1922 by Mordell [Mor] and became known as the Mordell's Theorem.

Theorem (Mordell)**.** *Let $E/\mathbb{Q}$ be an elliptic curve. The group of rational points $E(\mathbb{Q})$ is finitely generated.*

In this paper we are concerned with a slightly more general result:

Theorem (Mordell-Weil Theorem)**.** *Let K be a number field and E/K an elliptic curve. The group of rational points $E(K)$ is finitely generated.*

The aim of this paper is to present the proof of this theorem with the stated goal of being mostly self-contained. Some notions and theories are not developed in full generality but introduced and proved in the special cases we need. When this is the case, it will be indicated and a reference for the more general case will be provided. Most importantly, all notions and results needed to understand the proofs are included except in a few cases where we make use of general theorems whose demonstration would lead us too far away. Of course, there is only a limit to which one can pretend to be self-contained : we make free use of basic tools from Algebraic Number Theory, the theory of elliptic curves as well as finite Galois Theory.

The paper is organized in the following way : after a brief introduction to the theory of elliptic curves, we turn to the proof of the Mordell-Weil Theorem. We start by deriving the result from two theorems and the rest of the proof is concerned with the demonstration of these two results. The first part is purely technical while the second one is more conceptually demanding. Throughout, we closely follow the exposition in [Sil]. We have organized the proof in a way that we find more convenient and have tried to fill in the details as much as possible. Our treatment of the Weak Mordell-Weil Theorem slightly differs in the sense that we have reformulated all results in order to work only in finite extensions.

We have left quite a few appendices concerning discrete valuation rings, ramification theory, completion of fields and formal groups. These have been part of the learning process of the author. The reader who is familiar with these concepts should be able to read through the proof of Mordell-Weil without much effort.

The Mordell-Weil Theorem opens the way to a lot of interesting both solved and unsolved problems. One direction for further work is to ask questions concerning rational points of more general objects than elliptic curves. This question was considered by Weil in 1928 in his thesis. He considered a curve of genus greater than 1 over a number field with jacobian J. His result concerns the group of points $J(K)$ rational over K and states that it is finitely generated. He later generalized this statement to any abelian variety and this theorem is also known as the Mordell-Weil Theorem (see [Se1]).

Another possible direction is to stay focused on elliptic curves. The structure of finitely generated abelian groups being well-known, we can restate the Mordell-Weil Theorem as follows:

Theorem. *Let E be an elliptic curve defined over a number field K. Then*

$$E(K) \cong E_{tors}(K) \oplus \mathbb{Z}^r$$

where r, which is called the rank of the elliptic curve, is some natural number and $E_{tors}(K)$ denotes the torsion subgroup of $E(K)$.

One can ask about the structure of the finite group $E_{\mathrm{tors}}(K)$. The case $K = \mathbb{Q}$ was completely answered in 1977 in a theorem due to Mazur [Maz]. He proved that up to isomorphism, the torsion group $E_{\mathrm{tors}}(\mathbb{Q})$ could either be of the form $\mathbb{Z}/n\mathbb{Z}$ for $1 \leq n \leq 10$ or $n = 12$, or of the form $(\mathbb{Z}/2\mathbb{Z}) \oplus (\mathbb{Z}/2n\mathbb{Z})$ for $1 \leq n \leq 4$. The case of a general number field has been discussed by Merel [Mer].

One can also ask questions concerning the rank of an elliptic curve. This leads to many still unanswered problems. Most famous is the Birch and Swinnerton-Dyer Conjecture which is one of the seven Millennium Prize Problems. It states that for an elliptic curve E defined over K, the order of the zero of the Hasse-Weil L-function $L(E, s)$ of E at $s = 1$ is exactly the rank of E. This is still an open problem. One of the great advances toward a proof of this conjecture is the Gross-Zagier formula [GZ], which relates heights of Heegner points to the derivative of $L(E, s)$ at $s = 1$.

Acknowledgments

I would like to thank Professor Eva Bayer-Fluckiger for giving me the opportunity to do this project under her direction. I would also like to thank Mathieu Huruguen for his guidance and encouragement through every stage of the process. I thank him for his patience and his time and I am very grateful for him sharing his experience and knowledge with me.

1 Elliptic curves

We define and briefly discuss the objects that we are going to be working with, namely elliptic curves. This is a brief introduction to the subject and we leave out most of the proofs.

1.1 Definition

For the definition of the notions in this section we refer to [Har].

Definition 1.1. An elliptic curve (E, O_E) over a field K consists of a non-singular projective curve E over K of genus 1 together with a point $O_E \in E(K)$.

By abuse of language, we will often use the notation E/K to refer to an elliptic curve defined over K without specifying the K-point O_E. Its existence will be implicit.

Because elliptic curves are curves of genus one, one can show that E is isomorphic to its Jacobian $\mathrm{Pic}^0(E)$ and, more precisely, each choice of a point $O_E \in E(K)$ gives a natural isomorphism $E \xrightarrow{\sim} \mathrm{Pic}^0(E)$ ([Sil] ch. 3 § 3 prop. 3.4, p. 61). Since $\mathrm{Pic}^0(E)$ has the structure of an abelian group, this isomorphism provides (E, O_E) with two morphisms of varieties

$$\mu : E \times E \longrightarrow E \quad \text{and} \quad \iota : E \longrightarrow E$$

making E into a group scheme over K. For all extensions L of K, these morphisms make $E(L)$ into an abstract abelian group where μ is the addition, ι is the inverse and O_E the identity.

1.2 Example : Weierstrass curves

Definition 1.2. A Weierstrass curve C defined over a field K is a curve in $\mathbb{P}^2_K$ which is defined by an equation of the following form:

$$Y^2Z + a_1XYZ + a_3YZ^2 = X^3 + a_2X^2Z + a_4XZ^2 + a_6Z^3$$

where all the coefficients a_i belong to K. Such an equation is called a long Weierstrass equation.

Remark 1.3. Let $\bar{K}$ be a fixed algebraic closure of K. If C is such a Weierstrass curve, one checks by setting $Z = 0$ that the only $\bar{K}$-point C that does not belong to the chart $Z = 1$ is the point $[0, 1, 0]$ that we will denote by O_C. For this reason we will often directly refer to C using the dehomogenized equation

$$y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6.$$

We will talk about a point P of $C(\bar{K})$ given by coordinates (x_0, y_0) when we actually mean the point in $\mathbb{P}^2_K(\bar{K})$ with coordinates $[x_0, y_0, 1]$.

Definition 1.4. Consider a long Weierstrass equation:

$$y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6..$$

We introduce the following constants associated to this equation:

$$\begin{array}{rcl} b_2 & = & a_1^2 + 4a_2 \\ b_4 & = & a_1a_3 + 2a_4 \\ b_6 & = & a_3^2 + 4a_6 \\ b_8 & = & a_1^2a_6 - a_1a_3a_4 + 4a_2a_6 + a_2a_3^2 - a_4^2 \\ c_4 & = & b_2^2 - 24b_4 \\ c_6 & = & -b_2^3 + 36b_2b_4 - 216b_6. \end{array}$$

Proposition 1.5. *Let C be a Weierstrass curve over a field K defined by the equation*

$$y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6.$$

If we suppose that the characteristic of K is neither 2 nor 3, then this equation takes the simpler form

$$y^2 = x^3 + Ax + B,$$

where $A = -\frac{c_4}{48}$ and $B = -\frac{c_6}{864}$ and the constants c_4 and c_6 are the ones of Definition 1.4. Such an equation is called a short Weierstrass equation.

Proof. This is simple algebra. □

Definition 1.6. Let C be a Weierstrass curve over a field K given by the equation

$$y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6.$$

To this equation we associate a quantity Δ, called the determinant, defined by:

$$\Delta = -b_2^2 b_8 - 8b_4^3 - 27b_6^2 + 9b_2b_4b_6 = \frac{c_4^3 - c_6^2}{12^3},$$

where the constants are the ones of Definition 1.4. If further the characteristic of K is neither 2 nor 3, the discriminant is equal to $-16(4A^3 + 27B^2)$, where A and B are the constants defined in Proposition 1.5.

Proposition 1.7. *Let C be a Weierstrass curve over K. Then C is non-singular if and only if its discriminant Δ is non-zero.*

Proof. [Sil] ch. 3 prop. 1.4 (a) (i), p. 45. □

Proposition 1.8. *Let C be a Weierstrass curve over K with non-zero discriminant and let $O_C = [0, 1, 0]$. Then (C, O_C) is an elliptic curve over K.*

Proof. [Sil] ch. 3 prop. 3.1 (c), p. 59. □

Proposition 1.9. *Let K be a field and $\bar{K}$ a fixed algebraic closure of K. Let E/K be an elliptic curve over K given by a Weierstrass equation and with base point $O_E = [0, 1, 0]$. The group law induced on $E(\bar{K})$ by $Pic^0(E)$ is given by the following geometrical law: three points of $E(\bar{K})$ sum to zero if and only if they are colinear in $\mathbb{P}^2_K(\bar{K})$ and O_E is the neutral element.*

Proof. [Sil] ch. 3 prop. 3.4, p. 61. □

Using this proposition, one can derive explicit formulas. We illustrate this for the inversion. Let E/K be an elliptic curve given by the Weierstrass equation

$$y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6.$$

Let P_0 be a point of $C(\bar{K})$ with coordinates (x_0, y_0). Let L be the vertical line through P_0 and O_C. It is given by the equation $L : x - x_0 = 0$. Substituting in the equation of C we obtain a polynomial equation of degree two in the variable y :

$$y^2 + y + (-x_0^3 - a_2x_0^2 - a_4x_0 - a_6) = 0.$$

We already know of one solution, namely y_0. Extracting the other solution y_0', we find that

$$y_0' = -y_0 - a_1x_0 - a_3.$$

Whence the inverse $-P_0$ has coordinates given by

$$-P_0 = [x_0, -y_0 - a_1 x_0 - a_3, 1]. \tag{1.9.1}$$

By similar arguments, one derives formulas for addition. This is summed up in the following proposition.

Proposition 1.10. *Let K be a field and $\bar{K}$ a fixed algebraic closure of K. Let E/K be an elliptic curve given by a Weierstrass equation*

$$y^2 + a_1 xy + a_3 y = x^3 + a_2 x^2 + a_4 x + a_6.$$

(a) Let $P_0 = (x_0, y_0)$ be a point of $E(\bar{K})$. Then

$$-P_0 = [x_0, -y_0 - a_1 x_0 - a_3, 1]. \tag{1.10.1}$$

Let $P_1 = (x_1, y_1)$, $P_2 = (x_2, y_2)$ be points of $E(\bar{K})$ and denote by P_3 their sum.

(b) If $x_1 = x_2$ and $y_1 + y_2 + a_1 x_2 + a_3 = 0$, then $P_3 = O_E$. Otherwise define the quantities λ and ν as in the following table:

	λ	ν
$x_1 \neq x_2$	$\frac{y_2 - y_1}{x_2 - x_1}$	$\frac{y_1 x_2 - y_2 x_1}{x_2 - x_1}$
$x_1 = x_2$	$\frac{3x_1^2 + 2a_2 x_1 + a_4 - a_1 y_1}{2y_1 + a_1 x_1 + a_3}$	$\frac{-x_1^3 + a_4 x_1 + 2a_6 - a_3 y_1}{2y_1 + a_1 x_1 + a_3}$

(c) Denote by (x_3, y_3) the coordinates of P_3. Then

$$\begin{cases} x_3 = & \lambda^2 + a_1 \lambda - a_2 - x_1 - x_2 \\ y_3 = & -(\lambda + a_1) x_3 - \nu - a_3 \end{cases} \tag{1.10.2}$$

(d) The duplication formula for $P = (x, y) \in E(\bar{K})$ is given by

$$x(P + P) = \frac{x^4 - b_4 x^2 - 2b_6 - b_8}{4x^3 + b_2 x^2 + 2b_4 x + b_6}, \tag{1.10.3}$$

where the constants b_i are the ones given in Definition 1.4.

Proof. [Sil] ch. 3 prop. 2.3, p. 53-54. □

1.3 Embedding of elliptic curves

The above example concerning elliptic curves defined by Weierstrass equations is not as restrictive as one might initially think. In fact, all elliptic curves can be embedded in the projective plane and can be shown to be isomorphic to a Weierstrass curve. This is made precise in the following proposition.

Proposition 1.11. *Let (E, O_E) be an elliptic curve over a field K. There exists an isomorphism defined over K from (E, O_E) to an elliptic curve $(E', O_{E'})$ in $\mathbb{P}^2_K$ given by a Weierstrass equation and such that O_E is mapped to $O_{E'}$.*

Moreover, any isomorphism between Weierstrass curves defined over K that preserves the base point $[0, 1, 0]$ is of the form

$$[u^2X + r, u^3Y + sX + t, Z].$$

where u, r, s, t belong to K and u is non-zero. We say that (u, r, s, t) is an admissible change of coordinates.

Proof. [Sil] ch. 3 prop. 3.1. p. 59. □

Let E/K be an elliptic curve given by a Weierstrass equation

$$y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6.$$

Applying an admissible change of coordinates (u, r, s, t) gives a new Weierstrass equation with coefficients a'_i given by the following relations:

$$\begin{aligned}
ua'_1 &= a_1 + 2s \\
u^2a'_2 &= a_3 - sa_1 + 3r - s^2 \\
u^3a'_3 &= a_3 + ra_1 + 2t \\
u^4a'_4 &= a_4 - sa_3 + 2ra_2 - (t + rs)a_1 + 3r^2 - 2st \\
u^6a'_6 &= a_6 + ra_4 + r^2a_2 + r^3 - ta_3 - t^2 - rta_1 \\
u^4c'_4 &= c_4 \\
u^6c'_6 &= c_6 \\
u^{12}\Delta' &= \Delta.
\end{aligned} \tag{1.11.1}$$

Remark 1.12. Embedding an elliptic curve (E, O_E) into $\mathbb{P}^2_K$ involves the choice of an isomorphic Weierstrass curve. Since the group law of an elliptic curve is induced by the group law of $\mathrm{Pic}^0(E)$ and the bijection between the curve and the Picard group depends only on the chosen base point, the group laws on different isomorphic Weierstrass curves are preserved by isomorphism if we impose that the base point is $[0, 1, 0]$. It therefore suffices to prove the Mordell-Weil Theorem for elliptic curves in $\mathbb{P}^2_K$ given by a fixed Weierstrass equation. In the rest of this paper, we will always present an elliptic curve given by a fixed Weierstrass equation.

1.4 The m-torsion subgroup

Let K be a field and $\bar{K}$ a fixed algebraic closure of K. Let E/K be an elliptic curve given by a Weierstrass equation

$$Y^2Z + a_1XYZ + a_3YZ^2 = X^3 + a_2X^2Z + a_4XZ^2 + a_6Z^3$$

with coefficients in K. We define the multiplication-by-m map and introduce the m-torsion subgroup of $E(\bar{K})$. The goal of this section is to prove that this subgroup is finite. We start with a general definition concerning abstract groups.

Definition 1.13. Let G be an abstract group. For any natural number m, we define the m-torsion subgroup of G to be

$$G[m] := \{g \in G \mid g^m = 1\}.$$

Definition 1.14. Let m be an integer. We define the multiplication-by-m map $[m] : E(\bar{K}) \longrightarrow E(\bar{K})$ by setting $[0](P) = O_E$ and then for m positive,

$$[m](P) = \underbrace{P + \ldots + P}_{m \text{ times}}.$$

For m negative we set $[m](P) = [-m](-P)$.

Since addition is a morphism, it follows by induction that the multiplication-by-m map is a morphism and it is clear that it is also a homomorphism of abelian groups. Since $E(\bar{K})$, by the above definition applied to $E(\bar{K})$,

$$E(\bar{K}) = \{P \in E(\bar{K}) \mid [m](P) = O_E\}.$$

One can prove finiteness of the m-torsion subgroup of an elliptic curve defined over any field K, but since we are mainly interested in number fields, we will restrict ourselves to the case where K has characteristic zero. The general proof is not conceptually more demanding but this restriction will save us some timely calculations. We state and prove a result concerning morphisms of projective varieties.

Lemma 1.15. *Let ϕ be a morphism of embedded projective varieties between a projective variety V and a curve C. Then ϕ is either constant of surjective.*

Proof. By ([Sha], ch. 1 § 5.1 Theorem 2), the image $\phi(V)$ is a (non-empty) closed subset of C. Decompose $\phi(V)$ into irreducible closed components :

$$\phi(V) = \bigcup_{i=1}^{r} V_i$$

where the V_i's are non-empty closed irreducible subsets of C and none of them contains another. This decomposition exists and is unique. Let P_i be a point of V_i. Then $\{P_i\}$ is a closed subset of C and we have the following chain of closed irreducible sets $\{P_i\} \subset V_i \subset C$. Since the dimension of C is 1, we must either have $V_i = \{P\}$ or $V_i = C$. Suppose that ϕ is not surjective. Then all the V_i's are strictly contained in C and therefore we must have $V_i = \{P_i\}$ and as a consequence $\phi(V) = \{P_1, \ldots, P_r\}$. Since any morphism is continuous and V being irreducible implies that V is connected, we have that $\phi(V)$ is connected and therefore $\phi(V)$ must be a single point. In other words, ϕ is constant. □

Proposition 1.16. *Let E/K be an elliptic curve defined over a field of characteristic zero and given by a Weierstrass equation*

$$Y^2Z + a_1XYZ + a_3YZ^2 = X^3 + a_2X^2Z + a_4XZ^2 + a_6Z^3$$

with coefficients in K. Then the multiplication-by-m map $m : E(\bar{K}) \longrightarrow E(\bar{K})$ is non-constant. As a consequence, it is surjective.

Proof. The last part of this result follows from Lemma 1.15.

We start by proving that [2] is not constant. Consider a point $P = (x, y)$ of $E(\bar{K})$. The duplication formula 1.10.3 states that

$$x([2](P)) = \frac{x^4 - b_4x^2 - 2b_6 - b_8}{4x^3 + b_2x^2 + 2b_4x + b_6}.$$

It is not well-defined if P is a two-torsion point. In order for $[2](P)$ to be equal to O_E it is necessary that

$$4x^3 + b_2x^2 + 2b_4x + b_6 = 0.$$

Thus there are finitely many possible choices for x and as a consequence finitely many points of 2-torsion. Since $E(\bar{K})$ contains infinitely many points, [2] can therefore not be constant.

This reduces the proof to the case where m is odd. Indeed, suppose that we have shown the result for n odd. Then [2] and $[n]$ are surjective. Let m be any positive integer. We may write it as $m = 2^kn$ for some integer k and n odd. Since the composition of surjective maps is surjective and by definition of the multiplication-by-m map $[m] = [2]^k \circ [n]$, we see that $[m]$ is also surjective.

In order to treat the case of an odd integer we search for a non-trivial point of 2-torsion. Dividing $x^4 - b_4x^2 - 2b_6 - b_8$ by $4x^3 + b_2x^2 + 2b_4x + b_6$ we obtain $x/4 - b_2/16$ with a rest

$$R(x) = \left(\frac{b_2^2}{16} - \frac{3b_4}{2}\right)x^2 + \left(\frac{b_2b_4}{8} - \frac{9}{4}b_6\right)x + \left(\frac{b_2b_6}{16} - b_8\right).$$

One can check by computation that $R(x) = 0$ if and only if the discriminant Δ is zero. Since E is non-singular, this is impossible. We conclude that $4x^3 + b_2x^2 + 2b_4x + b_6$ does not divide $x^4 - b_4x^2 - 2b_6 - b_8$ and therefore there must exist an element x_0 in $\bar{K}$ such that the first polynomial vanishes at a higher order than the second at x_0. Let y_0 be any solution in $\bar{K}$ to the second degree polynomial equation

$$y^2 + (a_1x_0 + a_3)y = x_0^3 + a_2x_0^2 + a_4x_0 + a_6$$

and define P_0 to be the point of $E(\bar{K})$ with coordinates (x_0, y_0). This is a non-zero element of 2-torsion.

Let m be an odd integer which we write as $m = 2n + 1$. Then

$$[m](P_0) = [2n](P_0) + P_0 = P_0 \neq \mathcal{O}$$

and consequently $[m]$ is non-constant. □

Corollary 1.17. *Let E/K be an elliptic curve defined over a field of characteristic zero given by a Weierstrass equation*

$$Y^2Z + a_1XYZ + a_3YZ^2 = X^3 + a_2X^2Z + a_4XZ^2 + a_6Z^3$$

with coefficients in K. If m is any positive integer, then $E(\bar{K})[m]$ is finite.

Proof. By Proposition 1.16, the multiplication-by-m map is surjective. By ([Sha], ch. 1 § 6.3 Theorem 7 *(ii)*), there exists a non-empty open set U of $E(\bar{K})$ such that for all points P in U

$$\dim[m]^{-1}(P) = \dim E - \dim E = 0. \tag{1.17.1}$$

In other words, the fiber of $[m]$ over P in U is finite. The torsion subgroup is by definition the fiber of $[m]$ over O_E. Let Q be any point of $E(\bar{K})$. By surjectivity of $[m]$, there exists a point Q' in $E(\bar{K})$ such that $[m](Q') = Q$. Then the fiber over Q becomes

$$[m]^{-1}(Q) = \{P \in E(\bar{K}) \mid [m](P - Q') = O_E\}$$

and thus restricting the translation-by-$-Q'$ map $\tau_{-Q'}$ to $[m]^{-1}(Q)$ gives a morphism of projective varieties between $[m]^{-1}(Q)$ and $[m]^{-1}(O_E)$ which is actually an isomorphism with inverse $\tau_{Q'}$. Thus the dimension of $[m]^{-1}(Q)$ is the same as the dimension of $[m]^{-1}(O_E)$. Since Q is arbitrary, all fibers of $[m]$ have the same dimension which is zero by equation 1.17.1. So all fibers are finite and in particular $E(\bar{K})[m]$ is finite. □

2 Sketch of the proof

We give two results concerning elliptic curves that will be proved respectively in Section 3 and Section 4. Using these results, we derive the Mordell-Weil Theorem using a descent argument.

Theorem 2.1 (Height Theorem). *Let K be a number field and E/K an elliptic curve. There exists a function $h_E : E(K) \longrightarrow \mathbb{R}$, called a height function, that satisfies the following three conditions:*

(i) For any point Q of $E(K)$ there exists a constant C_Q depending only on Q and such that for every P belonging to $E(K)$, $h(P+Q) \leq 2h(P) + C_Q$.

(ii) There exists a natural number m greater or equal to 2 and a constant C_2 such that $h(mP) \geq m^2 h(P) - C_2$ for all P in $E(K)$.

(iii) For any constant C_3, the set of all points P of $E(K)$ for which $h(P)$ is bounded by C_3 is finite.

Theorem 2.2 (Weak Mordell-Weil Theorem). *Let K be a number field and E/K be an elliptic curve defined over K. If m is a natural number greater than 2, then the quotient group $E(K)/mE(K)$ is finite.*

Theorem 2.3 (Mordell-Weil Theorem). *Let E/K be an elliptic curve over a number field. Then $E(K)$ is a finitely generated abelian group.*

Proof. Let $Q_1, \ldots, Q_r$ be representatives of the quotient $E(K)/mE(K)$ which is finite by Theorem 2.2. Let P be a point of $E(K)$. Then there exists P_1 in $E(K)$ and an index i_1 between 1 and r such that $P = mP_1 + Q_{i_1}$. Similarly for P_1, there exists P_2 in $E(K)$ and an index i_2 between 1 and r such that $P_1 = mP_2 + Q_{i_2}$. Proceeding inductively, at the n^{th} stage we obtain a element P_n of $E(K)$ and an index i_n between 1 and r such that $P_{n-1} = mP_n + Q_{i_n}$. We may then write P as a linear combination of the point P_n and the representatives $Q_1, \ldots, Q_r$. Consider $h : E(K) \longrightarrow \mathbb{R}$ a function satisfying the properties of Theorem 2.1. If we can bound $h(P_n)$ by a constant, taking n large if necessary, we will be able to conclude that $E(K)$ is finitely generated by using property (iii) of Theorem 2.1. So our goal is to find a suitable bound. By property (ii), $h(P_n)$ is bounded by $m^{-2}(h(mP_n) + C_2)$. Applying the definition of P_n, this is equal to $m^{-2}(h(P_{n-1} - Q_{i_n}) + C_2)$ and by property (i) we see that this is in turn bounded by $m^{-2}(2h(P_{n-1}) + C_{Q_{i_n}} + C_2)$. We define $C = \max_{1 \leq i \leq r} C_{Q_i}$.

Thus, $h(P_n)$ is bounded by $m^{-2}(2h(P_{n-1})+C+C_2)$. Applying the same result but for $n-1$, we obtain

$$\begin{aligned} h(P_n) \leq m^{-2}(2m^{-2}(2h(P_{n-2}) + C + C_2) + C + C_2) \\ = \left(\frac{2}{m^2}\right)^2 h(P_{n-2}) + (C + C_2)\left(\frac{1}{m^2} + \frac{2}{m^4}\right) \end{aligned}$$

Repeating n times yields

$$h(P_n) \leq \left(\frac{2}{m^2}\right)^n h(P) + \frac{C + C_2}{m^2} \sum_{i=0}^{n-1} \left(\frac{2}{m^2}\right)^i.$$

and this bound is equal to

$$\left(\frac{2}{m^2}\right)^n h(P) + \frac{\left(1 - \left(\frac{2}{m^2}\right)^n\right)(C + C_2)}{m^2 - 2}.$$

Using the fact that m is greater than 2 and taking n large enough so that $(2/m^2)^n h(P) \leq 1$, we finally obtain

$$h(P_n) \leq 1 + \frac{C + C_2}{2} =: C_3.$$

Every point in $E(K)$ is a linear combination of points from the set

$$\{Q_1, \dots, Q_r\} \cup \{P \in E(K) : h(P) \leq C_3\},$$

which is finite by property (iii) of Theorem 2.1. We conclude that $E(K)$ is finitely generated.

□

3 Height Theory

We prove Theorem 2.1. From now on until the end of this chapter, K will denote a number field and M_K will be the set of standard absolute values on K (see Appendix C § 7.2). We first introduce a height function defined on $\mathbb{P}^n_{\mathbb{Q}}(\bar{\mathbb{Q}})$. We will then use this function to define the desired height function on $E(K)$. We closely follow [Sil] ch. 8, p. 224-239.

3.1 Heights in projective space

Definition 3.1. Let P be a point in $\mathbb{P}^n_{\mathbb{Q}}(K)$ with homogeneous coordinates $[x_0, \ldots, x_n]$. We define the height associated to K, which we denote by H_K, by

$$H_K(P) = \prod_{v \in M_K} \max_{0 \leq i \leq n} \{|x_i|_v\}^{n_v},$$

where $n_v = [K_v : \mathbb{Q}_{v|_{\mathbb{Q}}}]$.

Remark 3.2. Let P be a point of $\mathbb{P}^n_{\mathbb{Q}}(K)$ given by homogeneous coordinates $[x_0, \ldots, x_n]$. For each i, there are only finitely many discrete valuations v in M_K for which $v(x_i)$ is non-zero. So there are only finitely many discrete valuations for which $v(x_i)$ is non-zero for at least one i, which is to say that $\max_{0 \leq i \leq n}\{|x_i|_v\}^{n_v}$ is not 1. On the other hand, there are only finitely many archimedean absolute values in M_K. As a consequence, the seemingly infinite product in our definition of the height is actually a finite product so there is no need to worry about convergence.

We show that H_K is well-defined (i.e. it does not depend on the choice of coordinates) and give some first properties of this function.

Proposition 3.3. *Let P be a point of $\mathbb{P}^n(K)$.*

(i) The value $H_K(P)$ does not depend on the chosen homogeneous coordinates of P.

(ii) The height satisfies $H_K(P) \geq 1$.

(iii) Let L/K be a finite extension. Then $H_L(P) = H_K(P)^{[L:\mathbb{Q}]/[K:\mathbb{Q}]}$.

Proof. Let $[x_0, \ldots, x_n]$ be homogeneous coordinates for P and let λ be a non-zero element of K. Then

$$H_K([\lambda x_0, \ldots, \lambda x_n]) = \left(\prod_{v \in M_K} |\lambda|_v^{n_v} \right) H_K([x_0, \ldots, x_n])$$

and by the Product Formula (Appendix C § 7.2 Proposition 7.20) this is equal to $H_K([x_0, \dots, x_n])$ and we have proved (i).

We can always choose homogeneous coordinates for P such that one of the x_i's equals 1. Since $|1|_v = 1$ for all $v \in M_K$, we immediately see that $H_K(P)$ is greater or equal to 1 and thus we have proved (ii).

In order to prove (iii) we compute

$$H_L(P) = \prod_{w \in M_L} \max_{0 \le i \le n} \{|x_i|_w\}^{n_w} = \prod_{v \in M_K} \prod_{\substack{w \in M_L \\ w|v}} \max_{0 \le i \le n} \{|x_i|_w\}^{n_w}$$

Since the x_i's are in K and w restricts to v, we obtain

$$H_L(P) = \prod_{v \in M_K} \prod_{\substack{w \in M_L \\ w|v}} \max_{0 \le i \le n} \{|x_i|_v\}^{n_w} = \prod_{v \in M_K} \max_{0 \le i \le n} \{|x_i|_v\}^{\sum_{w|v} n_w}.$$

Using the Extension Formula (Appendix C § 7.2 Proposition 7.19), this becomes

$$H_L(P) = \prod_{v \in M_K} \max_{0 \le i \le n} \{|x_i|_v\}^{[L:K]n_v} = H_K(P)^{[L:\mathbb{Q}]/[K:\mathbb{Q}]}.$$

□

We introduce a height function that is not relative to a particular number field K.

Definition 3.4. Let P be a point of $\mathbb{P}^n_{\mathbb{Q}}(\bar{\mathbb{Q}})$ and choose a number field K such that P belongs to $\mathbb{P}^n_{\mathbb{Q}}(K)$. We define the absolute height function H to be $H(P) = H_K(P)^{1/[K:\mathbb{Q}]}$.

This function is well-defined, independent of the choice of K by point (iii) of Proposition 3.3 and $H(P)$ is greater or equal to 1.

Our first goal is to prove that if K is a number field, then the set of points in $\mathbb{P}^n(K)$ whose height is bounded by a given constant is finite. In order to prove this we need some properties of the height function. We start by investigating the relation between the height of a polynomial and the heights of its roots.

Notation If x belongs to $\bar{\mathbb{Q}}$, we set $H(x) = H([x, 1])$ and if x belongs to a number field K, we set $H_K(x) = H_K([x, 1])$.

Proposition 3.5. *Let $f(X) = a_0X^d + a_1X^{d-1} + \ldots + a_d$ be a polynomial with coefficients in $\bar{\mathbb{Q}}$ and roots $\alpha_1, \ldots, \alpha_d$ in $\bar{\mathbb{Q}}$. Then*

$$2^{-d} \prod_{j=1}^{d} H(\alpha_j) \le H([a_0, \ldots, a_d]) \le 2^{d-1} \prod_{j=1}^{d} H(\alpha_j).$$

Proof. Multiplying the polynomial by a non-zero scalar does not change the conclusion of the theorem since we are working with homogeneous coordinates and this operation has no effect on the roots. We may therefore suppose that $a_0 = 1$. Let $K = \mathbb{Q}(\alpha_1, \dots, \alpha_d)$. Since the coefficients of the polynomial F can be expressed as polynomials in the roots, K contains them. Define

$$\epsilon(v) = \begin{cases} 2, & \text{if } v \in M_K^\infty \\ 1, & \text{otherwise.} \end{cases}$$

so that the triangle inequality becomes $|x+y|_v \leq \epsilon(v) \max\{|x|_v, |y|_v\}$ whether or not v is archimedean.

For $v \in M_K$, we will prove that

$$\epsilon(v)^{-d} \prod_{j=1}^{d} \max\{|\alpha_j|_v, 1\} \leq \max_{0 \leq i \leq d}\{|a_i|_v\} \leq \epsilon(v)^{d-1} \prod_{j=1}^{d} \max\{|\alpha_j|_v, 1\}. \tag{3.5.1}$$

The result will then follow by elevating to the power n_v, taking the product over all $v \in M_K$ and then the $[K:Q]^{\text{th}}$ root since

$$\prod_{v \in M_K} \epsilon(v)^{n_v(d-1)} = 2^{(d-1)\sum_{v \in M_K^\infty} n_v} = 2^{(d-1)[K:Q]}$$

by the Extension Formula (Appendix C § 7.2 Proposition 7.19.

We will prove the above inequality by induction on the degree d of f. If $d = 1$, then $a_1 = -\alpha_1$ and the inequalities are obvious. Suppose now that the inequalities 3.5.1 hold for polynomials of degree $d-1$ and pick k such that $|\alpha_k|_v$ is largest among the absolute values of the roots. Define the polynomial

$$g(X) = \prod_{\substack{j=1 \\ j \neq k}}^{d} (X - \alpha_j) = X^{d-1} + b_1 X^{d-2} + \ldots + b_{d-1}.$$

Since $f(X) = (X - \alpha_k) g(X)$, by comparing the coefficients we see that

$$a_i = b_i - \alpha_k b_{i-1}$$

for i ranging between 0 and d if we additionally set $b_0 = 1$, $b_{-1} = 0$ and $b_d = 0$. We then see that

$$\max_{0 \leq i \leq d}\{|a_i|_v\} = \max_{0 \leq i \leq d}\{|b_i - \alpha_k b_{i-1}|_v\} \leq \epsilon(v) \max_{0 \leq i \leq d}\{|b_i|_v, |\alpha_k b_{i-1}|_v\}$$

by the triangle inequality. This in turn is bounded by

$$\epsilon(v) \max_{0 \leq i \leq d}\{|b_i|_v\} \max\{|\alpha_k|_v, 1\}.$$

Using the induction hypothesis for the polynomial g, we obtain the inequality

$$\max_{0\leq i\leq d}\{|b_i|_v\} \leq \epsilon(v)^{d-2}\prod_{\substack{j=1\\ j\neq k}}^{d}\max\{|\alpha_j|_v, 1\}$$

and injecting this into the previous bound, we obtain the desired upper bound

$$\max_{0\leq i\leq d}\{|a_i|_v\} \leq \epsilon(v)^{d-1}\prod_{j=1}^{d}\max\{|\alpha_j|_v, 1\}.$$

In order to prove the lower bound, we distinguish between two cases. If $|\alpha_k|_v$ is less than $\epsilon(v)$, then

$$\prod_{j=1}^{d}\max\{|\alpha_j|_v, 1\} \leq \max\{|\alpha_k|_v, 1\}^d \leq \epsilon(v)^d$$

and on the other hand since $a_0 = 1$, we have that $\max\{|a_i|_v\}$ is greater than 1 from which we deduce the lower bound :

$$\epsilon(v)^{-d}\prod_{j=1}^{d}\max\{|\alpha_j|_v, 1\} \leq 1 \leq \max_{0\leq i\leq d}\{|a_i|_v\}.$$

If $|\alpha_k|_v$ is strictly greater than $\epsilon(v)$ we distinguish between the archimedean and non-archimedean cases. Suppose that v is archimedean. Then by the triangle inequality

$$|b_j - \alpha_k b_{j-1}|_v \geq -|b_j|_v + |\alpha_k|_v|b_{j-1}|_v$$

so that

$$|\alpha_k|_v|b_{j-1}|_v \leq \max_{1\leq i\leq d}\{|b_i - \alpha_k b_{i-1}|_v\} + \max_{1\leq i\leq d}\{|b_i|_v\}.$$

This holds for all j ranging from 1 to d, so taking the maximum over all j's, we obtain

$$(|\alpha_k|_v - 1)\max_{1\leq i\leq d}\{|b_i|_v\} \leq \max_{1\leq i\leq d}\{|b_i - \alpha_k b_{i-1}|_v\}.$$

Rewriting the left hand side as $|\alpha_k|_v(1 - |\alpha_k|_v^{-1})\max_{1\leq i\leq d}\{|b_i|_v\}$ and using the fact that $\alpha_k > \epsilon(v) = 2$, we see that

$$\max_{1\leq i\leq d}\{|b_i - \alpha_k b_{i-1}|_v\} \geq \epsilon(v)^{-1}|\alpha_k|_v\max_{1\leq i\leq d}\{|b_i|_v\}.$$

Suppose that v is non-archimedean. If all b_i's have the same absolute value, using the fact that $|\alpha_k|_v > \epsilon(v) = 1$, we see that $|b_i - \alpha_k b_{i-1}|_v = |\alpha_k|_v|b_{i-1}|_v$ for all i. If the b_i's do not all have the same absolute value, let i_0 be such that $|b_{i_0}|_v$ is the maximal absolute value among the b_i's. Then

$$\max_{1\leq i\leq d}\{|b_i - \alpha_k b_{i-1}|_v\} \geq |b_{i_0+1} - \alpha_k b_{i_0}|_v = |\alpha_k|_v|b_{i_0}|_v = |\alpha_k|_v\max_{1\leq i\leq d}\{|b_i|_v\}.$$

So for all v in M_K, we have the inequality

$$\max_{0\leq i\leq d}\{|a_i|_v\} \geq \epsilon(v)^{-1}|\alpha_k|_v \max_{0\leq i\leq d}\{|b_i|_v\}$$

and applying the induction hypothesis ends the proof. □

Notation Let P be a point in $\mathbb{P}^n_{\mathbb{Q}}(\bar{\mathbb{Q}})$ and σ be an element of $\mathrm{Gal}(\bar{\mathbb{Q}}|\mathbb{Q})$. We shall write P^σ to denote the coordinate-wise action of σ on P.

Proposition 3.6. *Let P be a point in $\mathbb{P}^n_{\mathbb{Q}}(\bar{\mathbb{Q}})$ and let σ be an element of the absolute Galois group Gal($\bar{\mathbb{Q}}|\mathbb{Q}$). Then $H(P^\sigma) = H(P)$.*

Proof. Let $[x_0, \ldots, x_n]$ be homogeneous coordinates of P and let K be a number field that contains these coordinates. The restriction of σ to K induces a $\mathbb{Q}$-isomorphism from K to $\sigma(K)$ (K is not necessarily a normal extension). We now prove that it also induces a bijection

$$\begin{array}{ccc} M_K & \overset{\sim}{\longrightarrow} & M_{\sigma(K)} \\ v & \longmapsto & v^\sigma. \end{array}$$

It follows that $|\sigma(x)|_{v^\sigma} = |x|_v$ for all x in K and we can then compute that

$$\begin{aligned} H_K(P^\sigma) &= \prod_{w\in M_{\sigma(K)}} \max_{0\leq i\leq n}\{|\sigma(x_i)|_w\}^{n_w} \\ &= \prod_{v\in M_K} \max_{0\leq i\leq n}\{|\sigma(x_i)|_{v^\sigma}\}^{n_{v^\sigma}} = \prod_{v\in M_K} \max_{0\leq i\leq n}\{|x_i|_v\}^{n_v} = H_K(P) \end{aligned}$$

and since the degrees of K and $\sigma(K)$ over $\mathbb{Q}$ are equal, we obtain equality between $H(P^\sigma)$ and $H(P)$ as desired.

It remains to prove our claim. First, we prove that σ induces a ring isomorphism between $\mathcal{O}_K$ and $\mathcal{O}_{\sigma(K)}$. On one hand, $\sigma(\mathcal{O}_K)$ is contained in $\mathcal{O}_{\sigma(K)}$ since algebraic integers are mapped via isomorphisms to algebraic integers. On the other hand, suppose that x belongs to $\mathcal{O}_{\sigma(K)}$. By surjectivity of σ, there exists an element y in K such that $\sigma(y) = x$. But then y belongs to $\mathcal{O}_K$ since y equals $\sigma^{-1}(x)$ and σ^{-1} is an isomorphism. As a consequence, the prime ideals of $\mathcal{O}_K$ are in bijection with those of $\mathcal{O}_{\sigma(K)}$. In other words, the discrete valuations on K are in bijection with those of $\sigma(K)$.

In order to conclude, we need to show that the archimedean absolute values are also in bijection. Let $|\,|$ be such an absolute value on K. It is induced by an embedding f of K into $\bar{K}$, such that for every x in K, $|x| = |f(x)|_\infty$ where $|\,|_\infty$ is the standard absolute value on $\mathbb{C}$. But then $f\circ\sigma^{-1}$ is an embedding of $\sigma(K)$ into $\bar{K}$ and thus defines an archimedean absolute value on $\sigma(K)$. This defines a map $M_K^\infty \longrightarrow M_{\sigma(K)}^\infty$. Starting from an absolute value on $\sigma(K)$ induced by an embedding f, $f\circ\sigma$ defines an embedding of K and therefore an archimedean absolute value on K. This defines the inverse map and thus we have a bijection. □

Definition 3.7. Let $P = [x_0, \ldots, x_n]$ be a point in $\mathbb{P}^n_{\mathbb{Q}}(\bar{\mathbb{Q}})$. We define the minimal field of definition of P over a number field K to be

$$K(P) := K\left(\frac{x_0}{x_i}, \ldots, \frac{x_n}{x_i}\right) \text{ for any } i \text{ with } x_i \neq 0.$$

Proposition 3.8. *Let $P = [x_0, \ldots, x_n]$ be a point in $\mathbb{P}^n_{\mathbb{Q}}(\bar{\mathbb{Q}})$ and K a number field. Then $K(P)$ is well-defined in the sense that it neither depends on the choice of homogeneous coordinates of P nor on the choice of the index i such that x_i is non-zero.*

Proof. Let $[y_0, \ldots, y_n]$ be another choice of homogeneous coordinates for P. Then there exists a non-zero element λ of $\bar{\mathbb{Q}}$ such that $x_j = \lambda y_j$ for all j ranging from 0 to n. Suppose that x_i is non-zero. Then y_i is non-zero and for all j we have $x_j/x_i = \lambda x_j/(\lambda x_i) = y_j/y_i$. This proves that $K(P)$ does not depend on the choice of coordinates.

We now show that $K(P)$ does not depend on the choice of the index. Suppose that i and j are two indexes such that both x_i and x_j are non-zero. Let k be any other index than i and j. Then

$$\frac{x_k}{x_i} = \frac{x_k}{x_j}\frac{x_j}{x_i} = \frac{x_k}{x_j}\left(\frac{x_i}{x_j}\right)^{-1} \in K\left(\frac{x_0}{x_j}, \ldots, \frac{x_n}{x_j}\right)$$

since x_i/x_j is an element of this field and is therefore invertible in this field. Since this holds for all k, it proves that

$$K\left(\frac{x_0}{x_i}, \ldots, \frac{x_n}{x_i}\right) \subset K\left(\frac{x_0}{x_j}, \ldots, \frac{x_n}{x_j}\right)$$

and the same reasoning shows the other inclusion. We conclude that $K(P)$ is well-defined. □

Theorem 3.9. *Let C and d be constants. Then*

$$A_{C,d} = \{P \in \mathbb{P}^n_{\mathbb{Q}}(\bar{\mathbb{Q}}) \mid H(P) \leq C, [\mathbb{Q}(P) : \mathbb{Q}] \leq d\}$$

is a finite set. In particular, if K is a number field,

$$\{P \in \mathbb{P}^n_{\mathbb{Q}}(K) \mid H_K(P) \leq C\}$$

is a finite set.

Proof. Let P be a point in $\mathbb{P}^n_{\mathbb{Q}}(\bar{\mathbb{Q}})$ given by homogeneous coordinates $[x_0, \ldots, x_n]$. We may suppose that one of the coefficients, say x_j, equals 1. Then we can take $\mathbb{Q}(P)$ to be $\mathbb{Q}(x_0, \ldots, x_n)$ and

$$H_{\mathbb{Q}(P)}(P) = \prod_{v \in M_{\mathbb{Q}(P)}} \max_{0 \leq i \leq n}\{|x_i|_v\}^{n_v} = \prod_{v \in M_{\mathbb{Q}(P)}} \max_{0 \leq i \leq n}\{|x_i|_v, 1\}^{n_v}$$

since $x_j = 1$. This is clearly greater or equal to

$$\max_{0\leq i\leq n}\left(\prod_{v\in M_{\mathbb{Q}(P)}}\max\{|x_i|_v,1\}^{n_v}\right)=\max_{0\leq i\leq n}H_{\mathbb{Q}(P)}(x_i).$$

Thus, if $H(P)$ is less than C then the same is true for $\max_{0\leq i\leq n}H(x_i)$. Moreover, since $\mathbb{Q}(P)$ contains all of the $\mathbb{Q}(x_i)$, we have that $\max_{0\leq i\leq n}[\mathbb{Q}(x_i):\mathbb{Q}]$ is bounded by $[\mathbb{Q}(P):\mathbb{Q}]$. We conclude that $A_{C,d}$ is contained in

$$\{[x_0,\ldots,x_n]\in\mathbb{P}^n(\bar{\mathbb{Q}})\mid\max_{0\leq i\leq n}H(x_i)\leq C,\max_{0\leq i\leq n}[\mathbb{Q}(x_i):\mathbb{Q}]\leq d\}$$

and it therefore suffices to prove that the set

$$\mathcal{C}:=\{x\in\bar{\mathbb{Q}}\mid H(x)\leq C,[\mathbb{Q}(x):\mathbb{Q}]\leq d\}$$

is finite. So we have reduced to the case $n=1$.

Let x be an element of this set and let e be the degree of $\mathbb{Q}(x)$ over $\mathbb{Q}$ so that e is less than or equal to d by assumption. Denote by $\alpha_1,\ldots,\alpha_e$ the conjugates of x in $\bar{\mathbb{Q}}$. We set $\alpha_1=x$. Consider the minimal polynomial F of x over $\mathbb{Q}$:

$$F(X)=X^e+a_1X^{e-1}+\ldots+a_e\in\mathbb{Q}[X].$$

Using Proposition 3.5, we obtain

$$H([1,a_1,\ldots,a_e])\leq 2^{e-1}\prod_{i=1}^{e}H(\alpha_i).$$

By Proposition 3.6, all α_i's have the same height which is the height of x. The above inequality then becomes

$$H([1,a_1,\ldots,a_e])\leq 2^{e-1}H(x)^e\leq(2C)^d$$

since $H(x)$ is bounded by C and e is less than d. Remember that all the coefficients a_i belong to $\mathbb{Q}$. We therefore need to study the height $H_{\mathbb{Q}}$.

Let $P=[y_0,\ldots,y_n]$ be a point in $\mathbb{P}^n_{\mathbb{Q}}(\mathbb{Q})$. Multiplying by the homogeneous coordinates by a/b where b is the greatest common divisor of the numerators of the y_i's and a is the least common multiple of the denominators, we may suppose that the coordinates are integers and their greatest common divisor is 1. Then for every prime p we have $|y_i|_p\leq 1$ for all i's and for at least one i we have $|y_i|_p=1$. Thus, the maximum of $|y_i|_p$ over i equals 1. Therefore the non-archimedean absolute values do not contribute to the product defining $H_{\mathbb{Q}}$ and we get

$$H_{\mathbb{Q}}(P)=\max\{|y_0|_\infty,\ldots,|y_n|_\infty\}$$

where $|\cdot|_\infty$ denotes the archimedean absolute value on $\mathbb{Q}$. For any C, one can bound the size of the set of points in $\mathbb{P}^n_{\mathbb{Q}}(\mathbb{Q})$ whose height is bounded by C by for example $(2C+1)^{n+1}$.

Returning to our initial problem, we recall the inequality

$$H([1, a_1, \ldots, a_e]) \leq (2C)^d.$$

It follows from our discussion that there are only finitely many possibilities for the coefficients $a_1, \ldots, a_e$. This implies that there are only finitely many possibilities for the minimal polynomial of an element of $\mathcal{C}$. But each such polynomial contributes with at most d elements to our set and we therefore conclude that $\mathcal{C}$ is finite, which ends the proof.

Let K be a number field of degree d. If $P = [x_0, \ldots, x_n]$ is an element of $\mathbb{P}^n(K)$ then $\mathbb{Q}(P) = \mathbb{Q}(x_0, \ldots, x_n)$ is contained in K so that its degree over $\mathbb{Q}$ is less than d. Therefore the set $\{P \in \mathbb{P}^n_{\mathbb{Q}}(K) \mid H_K(P) \leq C\}$ is contained in $A_{C,d}$ and is therefore finite. □

The next result will be of use in a short while.

Proposition 3.10. *Let $F : \mathbb{P}^n_{\mathbb{Q}}(\bar{\mathbb{Q}}) \longrightarrow \mathbb{P}^m_{\mathbb{Q}}(\bar{\mathbb{Q}})$ be a morphism of degree d (i.e. a morphism of projective varieties defined by homogeneous polynomials of degree d). There exist positive constants C_1 and C_2 depending only on n, d and F and such that*

$$C_1 H(P)^d \leq H(F(P)) \leq C_2 H(P)^d$$

for all P in $\mathbb{P}^n_{\mathbb{Q}}(\bar{\mathbb{Q}})$.

Proof. Since F is a morphism of degree d, there exist homogeneous polynomials $f_0, \ldots, f_m$ of degree d such that $F = [f_0, \ldots, f_m]$ and the only common zero of these polynomials is zero. Let P be a point of $\mathbb{P}^n_{\mathbb{Q}}(\bar{\mathbb{Q}})$ given by homogeneous coordinates $[x_0, \ldots, x_n]$. Choose K to be a number field that contains the coordinates of P and the coefficients of all the f_i's. Let v be a valuation in M_K. To ease notations, we introduce the following :

$$|P|_v = \max_{1 \leq i \leq n} |x_i|_v \qquad \text{and} \qquad |F(P)|_v = \max_{1 \leq j \leq m} |f_j(P)|_v$$

so that $H_K(P) = \prod_{v \in M_K} |P|_v^{n_v}$ and $H_K(F(P)) = \prod_{v \in M_K} |F(P)|_v^{n_v}$. We also introduce

$$|F|_v = \max\{|a|_v \ : \ a \text{ is a coefficient of some } f_i\}$$

and define $H_K(F) = \prod_{v \in M_K} |F|_v^{n_v}$. Finally, we define

$$\epsilon(v) = \begin{cases} 1, & \text{if } v \in M_K^\infty \\ 0, & \text{otherwise.} \end{cases}$$

This enables us to use the triangle inequality

$$|s_1 + \ldots + s_k|_v \leq k^{\epsilon(v)} \max\{|s_1|_v, \ldots, |s_k|_v\}$$

without distinguishing between whether or not v is archimedean.

We now turn to the proof and start with the upper bound. Let I be the set of vectors $\mathbf{k} = (k_0, \ldots, k_n)$ of natural numbers that satisfy $\sum_{i=0}^{n} k_i = d$. One sees that the size of the set I is the number of monomials of degree d in $n+1$ variables which is $C = \binom{n+d}{n}$. For v in M_K, we have

$$|f_i(P)|_v = \left| \sum_{\mathbf{k} \in I} \lambda_{i,\mathbf{k}} x_0^{k_0} \ldots x_n^{k_n} \right|_v$$

where the $\lambda_{i,\mathbf{k}}$'s are the coefficients of f_i (some of them may be zero). Using the triangle inequality, this expression is bounded by

$$C^{\epsilon(v)} \max_{\mathbf{k} \in \mathbf{I}} \{|\lambda_{i,\mathbf{k}}|_v |x_0|_v^{k_0} \ldots |x_n|_v^{k_n}\}.$$

We bound $|\lambda_{i,\mathbf{k}}|_v$ by $|F|_v$ and we bound each $|x_i|_v$ by $|P|_v$. Since the k_j's sum to d, we get the following bound

$$|f_i(P)|_v \leq C^{\epsilon(v)} |F|_v |P|_v^d.$$

This inequality holds for all i and the right hand side is independent of i. Taking the maximum over i, we obtain

$$|F(P)|_v \leq C^{\epsilon(v)} |F|_v |P|_v^d.$$

Taking the n_v^{th} power and taking the product over all v in M_K, we obtain

$$H_K(F(P)) \leq C^{\sum_{v \in M_K} n_v \epsilon(v)} H_K(F) H_K(P)^d.$$

Note that by definition of ϵ,

$$\sum_{v \in M_K} \epsilon(v) n_v = \sum_{v \in M_K^\infty} n_v = [K : \mathbb{Q}],$$

by the Extension Formula (Appendix C § 7.2 Proposition 7.19). So taking the $[K : \mathbb{Q}]^{\text{th}}$ root, we finally get

$$H(F(P)) \leq C H(F) H(P)^d.$$

Taking C_2 to be $CH(F)$ which only depends on m, n, d and F, we have proved the upper bound.

We now prove the lower bound. Let I_F be the ideal in $\bar{\mathbb{Q}}[X_0, \ldots, X_n]$ generated by the polynomials $f_0, \ldots, f_m$. The ideal I_F is a homogeneous ideal and since F is a morphism, we have that the projective algebraic set $V_p(I_F)$ generated by I_F is empty. By the homogeneous Nullstellensatz, there exists a natural number N depending possibly on F, n and m and such that $I_{irr}^N \subset I_F$ where I_{irr} is the irrelevant ideal $(X_0, \ldots, X_n)$. Thus for each i ranging from 0 to n, there exist polynomials $g_{ij} \in \mathbb{Q}[X_0, \ldots, X_n]$ for $j = 0, \ldots, m$ such that

$$X_i^N = \sum_{j=0}^{m} g_{ij} f_j.$$

By taking a finite extension of K if necessary, we may suppose that the coefficients of the g_{ij}'s lie in K. Since the left hand side is homogeneous of degree N and the f_i's are homogeneous of degree d, we may suppose that all the g_{ij}'s are homogeneous of degree $N - d$ since non-homogeneous terms will eventually cancel out. Imitating our notations from before, we set

$$|G|_v = \max\{|a|_v \ : \ a \text{ is a coefficient of some } g_{ij}\}$$

and $H_K(G) = \prod_{v \in M_K} |G|_v^{n_v}$. Since $P = [x_0, \ldots, x_n]$, the above formula implies that

$$|x_i|_v^N = \left| \sum_{j=0}^{m} g_{ij}(P) f_j(P) \right|_v \leq m^{\epsilon(v)} \max_{0 \leq j \leq m} \{|g_{ij}(P) f_j(P)|_v\}$$

by the triangle inequality. This quantity is bounded by

$$m^{\epsilon(v)} \max_{0 \leq j \leq m} \{|g_{ij}(P)|_v\} \max_{0 \leq j \leq m} \{|f_j(P)|_v\} = m^{\epsilon(v)} \max_{0 \leq j \leq m} \{|g_{ij}(P)|_v\} |F(P)|_v.$$

Taking the maximum over i, we obtain

$$|P|_v^N \leq m^{\epsilon(v)} \max_{i,j} \{|g_{ij}(P)|_v\} |F(P)|_v.$$

Since the g_{ij}'s are homogeneous of degree $N - d$, using the same argument as for the upper bound we proved before, we obtain

$$|g_{ij}(P)|_v \leq \binom{n+N-d}{n}^{\epsilon(v)} |G|_v |P|_v^{N-d}$$

and we can take the maximum over i and j which does not affect the right hand side. Define C_3 to be the constant $m \binom{n+N-d}{n}$. Then regrouping the above expressions, we have shown that

$$|P|_v^N \leq C_3^{\epsilon(v)} |G|_v |P|_v^{N-d} |F(P)|_v.$$

Multiplying this expression by $|P|_v^{d-N}$ yields

$$|P|_v^d \leq C_3^{\epsilon(v)} |G|_v |F(P)|_v.$$

Elevating to the power n_v, taking the product over all v in M_K and taking the $[K : Q]^{\text{th}}$ root, we obtain $H(P)^d \leq C_3 H(G) H(F(P))$. Letting C_1 be $(C_3 H(G))^{-1}$ yields

$$H(F(P)) \geq C_1 H(P)^d.$$

Note that C_3 depends on m, n, N and d. But N and $H(G)$ can be bounded in terms of m, n, d and $H(F)$ and therefore C_1 only depends on m, n, d and F so we have proved the desired result. □

3.2 Heights on elliptic curves

We now define a height on elliptic curves and show that it satisfies the three properties of Theorem 2.1. Throughout this section, E/K will denote an elliptic curve defined over a number field K, given by a Weierstrass equation

$$E : Y^2Z = X^3 + AXZ^2 + BZ^3$$

and with base point $O_E = [0, 1, 0]$. Consider the morphism of projective varieties $x : E \longrightarrow \mathbb{P}^1_K$ defined by

$$P \longmapsto \begin{cases} [x_0, 1], & \text{if } P = [x_0, y_0, 1] \\ [1, 0], & \text{if } P = O_E. \end{cases}$$

To see that this is indeed a morphism, note that $[x, 1]$ is simply the projection on the first coordinate restricted to the open subset $E \setminus \{O_E\}$ so this is a morphism and it extends in a unique way to the whole curve E. We now show that it in fact extends to $[1, 0]$ when evaluated at O_E. First, note that if $P = [x_0, y_0, z_0]$ and z is non-zero, then $x(P) = [x_0/z_0, 1] = [x_0, z_0]$. Dehomogenizing with respect to Y yields the equation $z = x^3 + Axz^2 + Bz^3$ and $O_E = (0, 0)$ belongs to this chart. Rewriting this equation as $z(1 - Bz^2) = x(x^2 + Az^2)$ we see that $[x_0, z_0] = [1 - Bz_0^2, x_0^2 + Az_0^2]$ which is well-defined at O_E so that $x(O_E) = [1, 0]$. We conclude that x is indeed a morphism.

Moreover, it is a non-constant morphism and by Proposition 1.15 it is surjective. In view of the additive structure on E, we introduce the following definition.

Definition 3.11. We define the logarithmic height function on $\mathbb{P}^n(\bar{\mathbb{Q}})$ as follows :

$$\begin{aligned} h : \ \mathbb{P}^n_{\mathbb{Q}}(\bar{\mathbb{Q}}) &\longrightarrow \mathbb{R} \\ P &\longmapsto \log(H(P)). \end{aligned}$$

By Proposition 3.3 *(ii)*, we have that $h(P)$ is non-negative. We define the height on an elliptic curve to be

$$\begin{aligned} h_E : \ E(\bar{\mathbb{Q}}) &\longrightarrow \mathbb{R} \\ P &\longmapsto h(x(P)). \end{aligned}$$

We start by proving that property *(iii)* of Theorem 2.1 is satisfied.

Proposition 3.12. *Let C be a positive constant. Then the set of points P of $E(K)$ whose height is bounded by C is finite.*

Proof. Clearly, x maps points of $E(K)$ to points of $\mathbb{P}^1_K(K)$. Moreover, for any fixed point $[a, b]$ of $\mathbb{P}^1_K(K)$, there are only finitely many points of $E(K)$ that are mapped to $[a, b]$. If $b = 0$, then the preimage consists only of O_E. Otherwise, we suppose $b = 1$. We are then looking at the fiber of x at $[a, 1]$. Elements of this fiber are points $[a, y, 1]$ which satisfy the Weierstrass equation for E. Having fixed a, this gives a polynomial equation in y of degree 2 which has at most two roots in K. So x is a finite-to-one map from the set in question to the set

$$\{P \in \mathbb{P}^1_K(K) \mid H(P) \leq e^C\}$$

which is finite by Theorem 3.9. It follows that the set in question is finite. □

In order to prove the remaining properties (i) and (ii) of Theorem 2.1, we need the following theorem.

Theorem 3.13. *For all P and Q in $E(\bar{K})$ we have*

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

where the constant involved in the big-O notation are independent of the points considered.

Proof. Recall that E is given by the Weierstrass equation

$$E : Y^2Z = X^3 + AXZ^2 + BZ^3$$

where A and B are in K.

First, we consider the case where either P or Q is the origin. Recall that $x(O_E) = [1, 0]$ and therefore $H(x(O_E)) = 1$ which brings us to $h_E(O_E) = 0$. If P is the origin then the left hand side is exactly $2h_E(P)$ so we get a real equality without any constant involved. Using the fact that $h_E(P) = h_E(-P)$ since $x(P) = x(-P)$ be the inversion formula 1.10.1, we see that the case where Q is O_E is similar.

From now on we suppose that neither P nor Q is the origin. We set the following notation

$$\begin{aligned} x(P) &= [x_1, 1], \qquad & x(P + Q) &= [x_3, 1], \\ x(Q) &= [x_3, 1], \qquad & x(P - Q) &= [x_4, 1]. \end{aligned}$$

Notice that we may have "blow-ups" in the cases where Q equals ±P in which case we assign an infinite value to x_3 or x_4. Using Proposition 1.10 and doing some computations yields

$$\begin{aligned} x_3 + x_4 &= \tfrac{2(x_1+x_2)(A+x_1x_2)+4B}{(x_1+x_2)^2-4x_1x_2}, \\ x_3x_4 &= \tfrac{(x_1x_2-A)^2-4B(x_1+x_2)}{(x_1+x_2)^2-4x_1x_2}. \end{aligned} \tag{3.13.1}$$

Define the following composition of morphisms :

$$\begin{array}{rcccl} \sigma : & E(\bar{K}) \times E(\bar{K}) & \longrightarrow & \mathbb{P}^1_K(\bar{K}) \times \mathbb{P}^1_K(\bar{K}) & \longrightarrow & \mathbb{P}^2_K(\bar{K}) \\ & (P_1, P_2) & \longmapsto & (x(P_1), x(P_2)) & & \\ & & & ([\alpha_1, \beta_1], [\alpha_2, \beta_2]) & \longmapsto & [\beta_1\beta_2, \alpha_1\beta_2 + \alpha_2\beta_1, \alpha_1\alpha_2]. \end{array}$$

We also define the morphism

$$\begin{array}{rccl} G : & E(\bar{K}) \times E(\bar{K}) & \longrightarrow & E(\bar{K}) \times E(\bar{K}) \\ & (P_1, P_2) & \longmapsto & (P_1 + P_2, P_1 - P_2). \end{array}$$

This is a morphism since addition is a morphism. Finally, we define a map g from $\mathbb{P}^2_K(\bar{K})$ to $\mathbb{P}^2_K(\mathbb{K})$ by

$$g([t, u, v]) = [u^2 - 4tv, 2u(At + v) + 4Bt^2, (v - At)^2 - 4Btu].$$

With these definitions and the relations in equation 3.13.1, one checks that the following diagram commutes :

$$\begin{tikzcd} E(\bar{K}) \times E(\bar{K}) \arrow[r, "G"] \arrow[d, "\sigma"'] & E(\bar{K}) \times E(\bar{K}) \arrow[d, "\sigma"] \\ \mathbb{P}^2_K(\bar{K}) \arrow[r, "g"'] & \mathbb{P}^2_K(\bar{K}) \end{tikzcd}$$

We will now prove that g is a morphism. It is given by three homogeneous polynomials of degree 2 in t, u and v so all we need to do it to check that these polynomials do not have a common zero. In other words, the equation $g([t, u, v]) = 0$ must not have any other solution than $(0, 0, 0)$. If $t = 0$, then the equation becomes

$$u^2 = 0, \qquad 2uv = 0, \qquad v^2 = 0$$

so that $u = v = 0$ is the only possible solution. Thus we may suppose t is non-zero and define a new variable $z = u/2t$. Then the equation $u^2 - 4tv = 0$ becomes $z^2 = v/t$. The two remaining equations are

$$2u(At + v) + 4Bt^2 = 0 \quad \text{and} \quad (v - At)^2 - 4Btu = 0.$$

Dividing by t^2 and rewriting in terms of the new variable z, we obtain

$$\psi(z) = 4z^3 + 4Az + 4B = 0$$
$$\phi(z) = z^4 - 2Az^2 - 8Bz + A^2 = 0.$$

We need to check that these two polynomials have no common roots. Computing the gcd of ϕ and ψ we find $4(4A^3 + 27B^2)$ which is $-\Delta/4$, where Δ is the discriminant of the Weierstrass equation. Since E is non-singular, the discriminant is non-zero and therefore the two polynomials cannot have any common roots. This proves that g is indeed a morphism.

The map g is a morphism of degree 2 from $\mathbb{P}^2_K(\bar{K})$ to $\mathbb{P}^2_K(\bar{K})$. Using Proposition 3.10, there exist constants C_1 and C_2 such that for all $S \in \mathbb{P}^2_K(\bar{K})$,

$$C_1 H(S)^2 \leq H(g(S)) \leq C_2 H(S)^2$$

and the constants are independent of S. Taking the logarithm, we see that

$$\log C_1 + 2h(S) \leq h(g(S)) \leq 2h(S) + \log C_2.$$

In other words, $h(g(S)) = 2h(S) + O(1)$ and this holds for all S in $\mathbb{P}^2_K(\bar{K})$.

Using the fact that the above diagram commutes, we see that

$$h(\sigma(P+Q, P-Q)) = h(\sigma \circ G(P,Q)) = h(g \circ \sigma(P,Q)) = 2h(\sigma(P,Q)) + O(1).$$

It remains to prove that if P_1 and P_2 are two point of E, then

$$h(\sigma(P_1, P_2)) = h_E(P_1) + h_E(P_2) + O(1).$$

Applying this to both sides of the above equality ends the proof. If P_1 and P_2 are equal to the origin, then $\sigma(P_1, P_2) = [1, 0, 0]$ so that $h_E(\sigma(P_1, P_2)) = 0$. On the other hand, $h_E(\mathcal{O}) = 0$ so we have equality. If P_1 is the origin and P_2 is not, write $x(P_2) = [\alpha, 1]$. Then $\sigma(P_1, P_2) = [0, 1, \alpha]$ so that $h_E(\sigma(P_1, P_2)) = h(x(P_2)) = h_E(P_2)$ and we have the desired equality since the height of the origin is zero.

Finally, suppose neither P_1 nor P_2 is the origin and write $x(P_1) = [\alpha_1, 1]$ and $x(P_2) = [\alpha_2, 1]$. Then $\sigma(P_1, P_2) = [1, \alpha_1 + \alpha_2, \alpha_1\alpha_2]$. Consider the polynomial

$$F(X) = X^2 + (\alpha_1 + \alpha_2)X + \alpha_1\alpha_2 = (X + \alpha_1)(X + \alpha_2).$$

Applying Proposition 3.5 to F, we see that

$$2^{-2} H(\alpha_1) H(\alpha_2) \leq H([1, \alpha_1 + \alpha_2, \alpha_1\alpha_2]) \leq 2H(\alpha_1)H(\alpha_2).$$

Taking the logarithm, this becomes

$$-\log 4 + h(\alpha_1) + h(\alpha_2) \leq h(\sigma(P_1, P_2)) \leq \log 2 + h(\alpha_1) + h(\alpha_2).$$

As a consequence, $h(\sigma(P_1, P_2)) = h_E(P_1) + h_E(P_2) + O(1)$ and we this finishes the proof. □

Corollary 3.14. *Let E/K be an elliptic curve over a number field.*

(i) For any point Q of $E(\bar{K})$ we have

$$h_E(P+Q) \leq 2h(P) + O(1) \quad \textit{for all } P \in E(\bar{K}),$$

where the constant depends only on Q.

(ii) Let m be any integer. Then

$$h([m]P) = m^2 h_E(P) + O(1) \quad \textit{for all } P \in E(\bar{K}),$$

where the constant depends only on m.

Combined with Proposition 3.12 this proves Theorem 2.1.

Proof. To prove (i), we use the fact that the height on E is always non-negative and the result from Theorem 3.13 in order to obtain

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

Bringing $2h_E(Q)$ into the big O yields the desired result.

To prove (ii), note that since $x(P) = x(-P)$, we only need to prove this for m non-negative. The cases $m = 0$ and $m = 1$ are trivial and involve no constants. We proceed by induction. For m greater than 2, suppose that the result is true for n less than m. Using Theorem 3.13 with P and $[m-1]P$, we get

$$\begin{aligned}
h_E([m]P) &= h_E([m-1]P + P) \\
&= -h_E([m-1]P - P) + 2h_E([m-1]P) + 2h_E(P) + O(1) \\
&= (-(m-2)^2 + 2(m-1)^2 + 2)h_E(P) + O(1) \quad \text{(by induction)} \\
&= m^2 h_E(P) + O(1)
\end{aligned}$$

and this completes the proof of (ii). □

4 The Weak Mordell-Weil Theorem

We prove Theorem 2.2. This constitutes the technical heart of the proof of Mordell-Weil. The proof is articulated in several steps. We start by reducing the problem to the one of proving finiteness of a Galois extension. This first reduction is made via the Kummer pairing. We then study some general field theory and this will reduce the problem to proving a ramification property. For this, we will need some technology from the theory of elliptic curves which involves reduction of elliptic curves and the formal group of an elliptic curve.

4.1 The Kummer pairing

We start with a lemma that we will need in order to define the Kummer pairing.

Lemma 4.1. *Let K be a number field, E/K an elliptic curve. Suppose L/K is a finite Galois extension such that $E(L)/mE(L)$ is finite. Then $E(K)/mE(K)$ is finite.*

Proof. The inclusion of $E(K)$ in $E(L)$ provides us with a natural homomorphism of groups $\psi : E(K)/mE(K) \longrightarrow E(L)/mE(L)$. Denoting by Φ the kernel of ψ, we have the following short exact sequence

$$0 \longrightarrow \Phi \longrightarrow E(K)/mE(K) \longrightarrow \mathrm{Im}(\psi) \longrightarrow 0.$$

By assumption, the image of ψ is finite so it suffices to prove that the kernel is finite in order to have that $E(K)/mE(K)$ is finite.

We have reduced the argument to proving that Φ is finite. We will do this by embedding it into $\mathrm{Map}(\mathrm{Gal}(L|K), E(\bar{K})[m])$, which is finite since both $\mathrm{Gal}(L|K)$ and $E(\bar{K})[m]$ are finite, the latter being Corollary 1.17.

First, note that the kernel Φ is equal to $(E(K) \cap mE(L))/mE(K)$. For each element $P + mE(K)$ of Φ, there exists a point Q_P in $E(L)$ such that P equals mQ_P. We define the following map

$$\begin{array}{rccc} \lambda_P : & \mathrm{Gal}(L|K) & \longrightarrow & E(\bar{K})[m] \\ & \sigma & \longmapsto & Q_P^\sigma - Q_P. \end{array}$$

To see that this is well-defined, we need to verify that the image of λ_P is indeed in $E(\bar{K})[m]$. Letting σ be an element of the Galois group, we see that $m\lambda_P(\sigma)$ is equal to $mQ_P^\sigma - mQ_P$. Since addition is a morphism defined over K, $(P+Q)^\sigma$ is equal to $P^\sigma + Q^\sigma$. In particular, mP^σ is equal to $(mP)^\sigma$. Consequently, $m\lambda_P(\sigma)$ is equal to $(mQ_P)^\sigma - mQ_P$ which, by choice of Q_P, is $P^\sigma - P$. Since P is an element of $E(K)$, the action of the Galois group on P is trivial. Whence P^σ equals P and this proves that the image of λ_P is indeed in $E(\bar{K})[m]$.

We have now constructed a map

$$\begin{array}{rccl}\lambda : & \Phi & \longrightarrow & \mathrm{Map}(\mathrm{Gal}(L|K), E[m]) \\ & P & \longmapsto & \lambda_P .\end{array}$$

To see that this is an injection, let P and P' be two elements of the kernel Φ such that the maps λ_P and $\lambda_{P'}$ coincide. Then for every element σ of the Galois group, we have equality between $Q_P^\sigma - Q_P$ and $Q_{P'}^\sigma - Q_{P'}$ which is equivalent to equality between $(Q_P - Q_{P'})^\sigma$ and $Q_P - Q_{P'}$. Whence $Q_P - Q_{P'}$ is an element of $E(K)$. But then $P - P' = m(Q_P - Q_{P'})$ is an element of $mE(K)$ so that P equals P' in Φ. □

Corollary 4.2. *Let K be a number field, E/K an elliptic curve. In order to prove that $E(K)/mE(K)$ is finite, we may suppose that $E(K)$ contains $E(\bar{K})[m]$.*

Proof. Let L/K be a finite Galois extension such that $E(L)$ contains $E(\bar{K})[m]$, which is possible since $E(\bar{K})[m]$ is finite by Corollary 1.17. By Lemma 4.1, in order to prove that $E(K)/mE(K)$ is finite it suffices to show that $E(L)/mE(L)$ and thus we may directly work with L instead of K. □

Proposition-Definition 4.3 (Kummer Pairing). *Let K be a number field and E/K an elliptic curve. Suppose that $E(\bar{K})[m]$ is contained in $E(K)$. Consider the following map, which is called the Kummer pairing,*

$$\begin{array}{rccl}\kappa : & E(K) \times \mathit{Gal}(\bar{K}|K) & \longrightarrow & E(\bar{K})[m] \\ & (P, \sigma) & \longmapsto & Q^\sigma - Q,\end{array}$$

where $Q \in E(\bar{K})$ is a point such that $mQ = P$. This is a well-defined map.

Proof. First, the existence of such an element Q of $E(\bar{K})$ is guaranteed by Proposition 1.16.

Secondly, we need to verify that the image of κ is indeed an element of the m-torsion subgroup of $E(\bar{K})$. The proof of this is similar to the one in the proof of Lemma 4.1.

Finally, we need to show that the definition of κ is independent of our choice of Q. So let Q and Q' be two points in $E(\bar{K})$ such that $mQ = P = mQ'$. Then $m(Q - Q')$ equals zero. This implies that $Q - Q'$ is an element of $E(\bar{K})[m]$. So $Q = Q' + T$ for some T in $E(\bar{K})[m]$. But then

$$Q^\sigma - Q = Q'^\sigma - Q' + T^\sigma - T = Q'^\sigma - Q'$$

because T^σ equals T since T is in particular an element of $E(K)$ by our assumption that $E(K)$ contains $E(\bar{K})[m]$. We conclude that the pairing is well-defined. □

Definition 4.4. Let K be a number field and E/K and elliptic curve. For any point Q in $[m]^{-1}(E(K))$, i.e. any Q in $E(\bar{K})$ such that mQ belongs to $mE(K)$, we denote by $K_G(Q)$ the compositum of the minimal fields of definition $K(Q^\sigma)$ (Definition 3.7), where σ is an element of the absolute Galois group $\mathrm{Gal}(\bar{K}|K)$. The fields $K_G(Q)$ are finite Galois extensions of K. We further define $K_{m,E}$ to be the compositum of all the $K_G(Q)$'s where Q belongs to $[m]^{-1}(E(K))$.

Proposition 4.5. *Let K be a number field and E/K an elliptic curve. Suppose that $E(\bar{K})[m]$ is contained in $E(K)$. The Kummer pairing enjoys the following properties :*

(i) It is bilinear.

(ii) We have the equality

$$mE(K) = \{P \in E(K) \mid \kappa(P,\sigma) = O_E,\ \forall \sigma \in \mathit{Gal}(\bar{K}|K)\}.$$

(iii) We have the equality

$$\mathit{Gal}(\bar{K}|K_{m,E}) = \{\sigma \in \mathit{Gal}(\bar{K}|K) \mid \kappa(P,\sigma) = O_E,\ \forall P \in E(K)\}.$$

Proof. (i) For linearity in the first variable, let P and P' be elements of $E(K)$. Then for all σ belonging to $\mathrm{Gal}(\bar{K}|K)$,

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

For linearity in the second variable, let σ and τ belong to $\mathrm{Gal}(\bar{K}|K)$. Then for all P in $E(K)$,

$$\kappa(P,\sigma\tau) = Q^{\sigma\tau} - Q = Q^{\sigma\tau} - Q^\tau + Q^\tau - Q = (Q^\sigma - Q)^\tau + \kappa(P,\tau).$$

But $Q^\sigma - Q$ belongs to $E[m]$ so in particular to $E(K)$. Hence, it is fixed by τ. Therefore we obtain the sum of $\kappa(P,\sigma)$ and $\kappa(P,\tau)$ as desired.

(ii) Let P belong to $mE(K)$. Let Q be an element of $E(K)$ such that mQ equals P. Then for any σ in $\mathrm{Gal}(\bar{K}|K)$,

$$\kappa(P,\sigma) = Q^\sigma - Q = O_E$$

since Q is fixed by σ. This proves that $mE(K)$ is contained in the set in question. On the other hand, if $\kappa(P,\sigma)$ is zero for all σ, then $Q^\sigma = Q$ for all σ. In other words, Q is fixed by all elements of the absolute Galois group, so that Q is an element of $E(K)$ and P is in turn an element of $mE(K)$. This proves the remaining inclusion.

(iii) Let σ belong $\mathrm{Gal}(\bar{K}|K_{m,E})$. If P is in $E(K)$ and Q is an element of $[m]^{-1}(E(K))$ such that $mQ = P$, then Q belongs to $K_{m,E}$. In particular, Q is a $K_{m,E}$-rational point of E. Therefore Q is fixed by σ and $\kappa(P, \sigma) = O_E$. On the other hand, if $\kappa(P, \sigma)$ is zero for all K-rational points P of E, then σ fixes all elements of $m^{-1}E(\bar{K})$. This implies that σ fixes all elements of $E(K_{m,E})$ and must therefore be an element of $\mathrm{Gal}(\bar{K}|K_{m,E})$. □

Corollary 4.6. *Let K be a number field and E an elliptic curve defined over K such that $E(\bar{K})[m]$ is contained in $E(K)$. Then $E(K)/mE(K)$ is finite if and only if $K_{m,E}/K$ is finite.*

Proof. By Proposition 4.5 *(iii)*, $\mathrm{Gal}(\bar{K}|K_{m,E})$ is the kernel of the homomorphism

$$\begin{array}{ccc} \mathrm{Gal}(\bar{K}|K) & \longrightarrow & \mathrm{Hom}(E(K), E(\bar{K})[m]) \\ \sigma & \longmapsto & \kappa(\,\cdot\,, \sigma) \end{array}$$

which implies that $\mathrm{Gal}(\bar{K}|K_{m,E})$ is a normal subgroup of $\mathrm{Gal}(\bar{K}|K)$. As a consequence, $K_{m,E}$ is a Galois extension of K and we have an isomorphism between $\mathrm{Gal}(\bar{K}|K)/\mathrm{Gal}(\bar{K}|K_{m,E})$ and $\mathrm{Gal}(K_{m,E}|K)$. It follows that the Kummer pairing induces a bilinear pairing

$$E(K)/mE(K) \times \mathrm{Gal}(K_{m,E}|K) \longrightarrow E(\bar{K})[m]$$

for which the two following homomorphisms are injective :

$$\begin{array}{rcl} E(K)/mE(K) & \longrightarrow & \mathrm{Hom}(\mathrm{Gal}(K_{m,E}|K), E(\bar{K})[m]) \\ \mathrm{Gal}(K_{m,E}|K) & \longrightarrow & \mathrm{Hom}(E(K)/mE(K), E(\bar{K})[m]). \end{array}$$

As a consequence, $E(K)/mE(K)$ is finite if and only if $\mathrm{Gal}(K_{m,E}|K)$ is finite. □

4.2 Some general field theory

Having reduced the problem of proving of the Weak Mordell-Weil Theorem 2.2 to the one of proving finiteness of the field extension $K_{m,E}/K$, we turn to some general field theory. The goal here is to prove the following theorem.

Theorem 4.7. *Let K be a number field, m a natural number greater than 2 and S a finite set of places in M_K containing the archimedean places M_K^∞. Then there exists a finite abelian field extension $K_{m,S}^{ab}$ of K such that all finite Galois extensions of K that are abelian of exponent m and unramified outside S are contained in $K_{m,S}^{ab}$.*

Proposition 4.8. *Let K be a number field, E an elliptic curve defined over K and m a natural number greater than 2. Then for all Q in $[m]^{-1}(E(K))$, the extension $K_G(Q)/K$ is abelian of exponent m.*

Proof. By Corollary 4.6, we have an injective homomorphism of groups

$$\begin{array}{ccc} \mathrm{Gal}(K_{m,E}|K) & \longrightarrow & \mathrm{Hom}(E(K), E(\bar{K})[m]) \\ \sigma & \longmapsto & \kappa(\,\cdot\,,\sigma). \end{array}$$

Thus $\mathrm{Gal}(K_{m,E}|K)$ identifies with a subgroup of $\mathrm{Hom}(E(K), E(\bar{K})[m])$ which is abelian since both $E(K)$ and $E(\bar{K})[m]$ are abelian. Moreover, let σ be any element of the Galois group of $K_{m,E}/K$. The image of σ^m via the above homomorphism is the homomorphism $m\kappa(\,\cdot\,,\sigma)$ which is zero since its target is $E(\bar{K})[m]$. Thus σ has order dividing m.

Let Q be in $[m]^{-1}(E(K))$. Since $K_G(Q)/K$ is a Galois extension, its Galois group is isomorphic to a quotient of the Galois group of $K_{m,E}/K$. It is therefore abelian and of exponent m. Since Q is arbitrary this proves the desired assertion. □

Corollary 4.9. *Let K be a number field and E an elliptic curve defined over K such that $E(\bar{K})[m]$ is contained in $E(K)$. Suppose that there exists a finite set S of places in M_K containing M_K^∞ such that for all Q in $[m]^{-1}(E(K))$, the field $K_G(Q)$ is unramified outside S. Then $K_{m,E}/K$ is a finite extension.*

Proof. By Proposition 4.8 and Theorem 4.7, there exists a finite abelian field extension $K^{ab}_{m,S}$ of K that contains $K_G(Q)$ for all Q be in $[m]^{-1}(E(K))$. It therefore contains the compositum of these fields which is $K_{m,E}$. In particular, $K_{m,E}$ is a finite extension of K. □

Quite amazingly, the key argument in the proof of the Weak Mordell-Weil Theorem comes from general field theory. Once we have proved Theorem 4.7, all that remains is to find a set of places S and show that the extensions $K_G(Q)/K$ are unramified outside this set. The rest of this section is concerned with the proof of Theorem 4.7. The question of finding a suitable set S is left for the sections to come. We need a few simplifying lemmas.

Lemma 4.10. *Let K be a number field, m a natural number greater than 2 and S a finite set of places in M_K containing the archimedean places M_K^∞. Let $\tilde{K}$ be a finite Galois extension of K, and let $\tilde{S}$ be the set of places in $M_{\tilde{K}}$ that lie above the places in S. Suppose that there exists a field $\tilde{K}^{ab}_{m,\tilde{S}}$ satisfying the conclusion of Theorem 4.7 for the field $\tilde{K}$ and the set $\tilde{S}$. Then $\tilde{K}^{ab}_{m,\tilde{S}}$ also satisfies the conclusion of Theorem 4.7 for the field K and the set S. We may therefore take $K^{ab}_{m,S}$ to be $\tilde{K}^{ab}_{m,\tilde{S}}$.*

Proof. By assumption, the field $\tilde{K}^{ab}_{m,\tilde{S}}$ is a finite extension of $\tilde{K}$ that contains all finite abelian extensions of exponent m of $\tilde{K}$ that are unramified outside $\tilde{S}$. Let L be a finite abelian extension of exponent m of K that is unramified outside S. Consider the compositum $\tilde{L}$ of L and $\tilde{K}$ in a fixed algebraic closure $\bar{K}$ of K. We claim that $\tilde{L}$ is a finite abelian extension of exponent m of $\tilde{K}$ that is unramified outside $\tilde{S}$. From this, it follows that $\tilde{L}$ is contained in $\tilde{K}^{ab}_{m,\tilde{S}}$ and in particular L is contained in $\tilde{K}^{ab}_{m,\tilde{S}}$. Since this holds for all such L, we may conclude that $\tilde{K}^{ab}_{m,\tilde{S}}$ satisfies the conclusion of Theorem 4.7 for the field K and the set S.

It remains to prove our claim. First, being the compositum of two finite Galois extensions, $\tilde{L}$ is necessarily a finite Galois extension of $\tilde{K}$. Consider the restriction map

$$\begin{array}{rccc} \phi: & \mathrm{Gal}(\tilde{L}|\tilde{K}) & \longrightarrow & \mathrm{Gal}(L|K) \\ & \sigma & \longmapsto & \sigma|_L. \end{array}$$

Since L/K is Galois, $\sigma|_L(L) = L$ and therefore $\sigma|_L$ is an automorphism of L. Since σ fixes $\tilde{K}$ and K is contained in $\tilde{K}$ and L, $\sigma|_L$ fixes K. Therefore $\sigma|_L$ is an element of $\mathrm{Gal}(L|K)$ and the above map is well-defined. Moreover, the restriction ϕ is clearly a homomorphism of groups. We now prove that it is an injective homomorphism of groups : suppose σ is an element of $\mathrm{Gal}(\tilde{L}|\tilde{K})$ such that $\sigma|_L$ is the identity. In other words, σ fixes L. But it also fixes $\tilde{K}$ and must therefore be the identity on the compositum of L and $\tilde{K}$, which is $\tilde{L}$. So we have proved that σ is the identity and this proves injectivity. Since the above map is an injective homomorphism of groups, $\mathrm{Gal}(\tilde{L}|\tilde{K})$ is isomorphic to a subgroup of $\mathrm{Gal}(L|K)$. Since $\mathrm{Gal}(L|K)$ is an abelian group of exponent m, the same is true for all its subgroups and we conclude that $\mathrm{Gal}(\tilde{L}|\tilde{K})$ is abelian of exponent m. We have shown that $\tilde{L}$ is an abelian extension of exponent m of $\tilde{K}$.

We now prove that $\tilde{L}/\tilde{K}$ is unramified outside $\tilde{S}$. We will do this using the theory of inertia groups (Appendix B § 6.3). Let $\tilde{v}$ be a place outside $\tilde{S}$ and let $\tilde{\mathfrak{p}}$ be the corresponding prime ideal of $\mathcal{O}_{\tilde{K}}$. Let $\tilde{\mathfrak{q}}$ be any prime ideal of $\mathcal{O}_{\tilde{L}}$ that lies above $\tilde{\mathfrak{p}}$. Define $\mathfrak{p} = \tilde{\mathfrak{p}} \cap \mathcal{O}_K$ and $\mathfrak{q} = \tilde{\mathfrak{q}} \cap \mathcal{O}_L$, which are prime ideals of $\mathcal{O}_K$ and $\mathcal{O}_L$ respectively and $\mathfrak{q}$ lies above $\mathfrak{p}$. The prime ideal $\mathfrak{p}$ corresponds to a place v in M_K which lies below $\tilde{v}$. By definition of $\tilde{S}$, v does not belong to S. Since L/K is unramified outside S, the inertia group $I_{\mathfrak{q}/\mathfrak{p}}$ is trivial (Appendix B § Corollary 6.27). In order to prove that $\tilde{L}/\tilde{K}$ is unramified at $\tilde{\mathfrak{p}}$, it suffices to prove that the inertia group $I_{\tilde{\mathfrak{q}}/\tilde{\mathfrak{p}}}$ is trivial (Appendix B § Corollary 6.27). We will prove this by showing that the restriction of the above homomorphism ϕ to $I_{\tilde{\mathfrak{q}}/\tilde{\mathfrak{p}}}$ is an injective homomorphism onto $I_{\mathfrak{q}/\mathfrak{p}} = \{1\}$ and therefore $I_{\tilde{\mathfrak{q}}/\tilde{\mathfrak{p}}} = \{1\}$.

We now prove that $\phi(I_{\tilde{\mathfrak{q}}/\tilde{\mathfrak{p}}})$ is contained in $I_{\mathfrak{q}/\mathfrak{p}}$. Let therefore σ be an element of $I_{\tilde{\mathfrak{q}}/\tilde{\mathfrak{p}}}$. We first prove that $\sigma|_L(\mathfrak{q}) = \mathfrak{q}$. In other words we prove that $\sigma|_L$ is in the decomposition group of $\mathfrak{q}$. Since $\mathfrak{q}$ is contained in $\tilde{\mathfrak{q}}$ and $\sigma(\tilde{\mathfrak{q}}) = \tilde{\mathfrak{q}}$, we have that $\sigma|_L(\mathfrak{q})$ is contained in $\tilde{\mathfrak{q}}$. Moreover, since automorphisms map algebraic integers to algebraic integers, we also have that $\sigma|_L(\mathfrak{q})$ is contained in $\mathcal{O}_L$, hence in the intersection $\tilde{\mathfrak{q}} \cap \mathcal{O}_L$ which is $\mathfrak{q}$. Applying the same argument to σ^{-1} which is also in the inertia group $I_{\tilde{\mathfrak{q}}/\tilde{\mathfrak{p}}}$, we get that $\sigma^{-1}|_L(\mathfrak{q})$ is contained in $\mathfrak{q}$ from what we conclude that $\mathfrak{q}$ is contained in $\sigma|_L(\mathfrak{q})$. Thus $\sigma|_L(\mathfrak{q}) = \mathfrak{q}$.

Next, we prove that $\sigma|_L(x) - x$ belongs to $\mathfrak{q}$ for all x in $\mathcal{O}_L$. Let x be an element of $\mathcal{O}_L$. Since $\mathcal{O}_L$ is contained in $\mathcal{O}_{\tilde{L}}$ we know by assumption on σ that $\sigma|_L(x) - x$ belong to $\tilde{\mathfrak{q}}$. But automorphisms map algebraic integers to algebraic integers and thus $\sigma|_L(x)$ belong to $\mathcal{O}_L$. Since this is a ring, $\sigma|_L(x) - x$ also belongs to $\mathcal{O}_L$, hence it belongs to the intersection $\tilde{\mathfrak{q}} \cap \mathcal{O}_L$ which is $\mathfrak{q}$ by definition. This holds for all x in $\mathcal{O}_L$ so we have prove that $\sigma|_L$ belongs to $I_{\mathfrak{q}/\mathfrak{p}}$. We therefore have a well-defined homomorphism

$$\begin{array}{ccc} I_{\tilde{\mathfrak{q}}/\tilde{\mathfrak{p}}} & \longrightarrow & I_{\mathfrak{q}/\mathfrak{p}} \\ \sigma & \longmapsto & \sigma|_L. \end{array}$$

which is injective since ϕ is injective. We conclude that $I_{\tilde{\mathfrak{q}}/\tilde{\mathfrak{p}}}$ is isomorphic to a subgroup of $I_{\mathfrak{q}/\mathfrak{p}}$. But this is the trivial group, so $\tilde{L}/\tilde{K}$ is unramified at $\tilde{v}$. This is true for all $\tilde{v}$ that is outside $\tilde{S}$ and therefore $\tilde{L}/\tilde{K}$ is unramified outside $\tilde{S}$. □

Corollary 4.11. *Let K be a number field, m a natural number greater than 2 and S a finite set of places in M_K containing the archimedean places M_K^∞. If Theorem 4.7 holds in the case of a field that contains the group of m^{th} roots of unity μ_m, then it also holds for K and S.*

Proof. Since μ_m is a finite group, we may choose $\tilde{K}$ to be a finite Galois extension of K that contains μ_m. Take $\tilde{S}$ to be the set of places above the ones in S. By assumption, there exists a field $\tilde{K}^{ab}_{m,\tilde{S}}$ satisfying the conclusion of Theorem 4.7 for the field $\tilde{K}$ and the set $\tilde{S}$. By Lemma 4.10, Theorem 4.7 is also satisfied for K and S. □

Being able to suppose that K contains the group μ_m allows us to use general results from Kummer Theory. We prove the next three results in a general setting.

Theorem 4.12. *Let F be a field an m a natural number such that the characteristic of F does not divide m. Suppose that F contains the group of m^{th} roots of unity μ_m. Let L/F be a cyclic extension of degree m. Then $L = F(\beta)$ where β is an element of L and a root of $X^m - a$ for some a in F.*

Proof. Let σ be a generator of the Galois group of L over F. Choose ζ to be a primitive m^{th} root of unity. By assumption ζ belongs to F and so does its inverse ζ^{-1} which is therefore fixed by σ. Thus

$$N_F^L(\zeta^{-1}) = \zeta^{-1}\sigma(\zeta^1)\sigma^2(\zeta^{-1})\dots\sigma^{m-1}(\zeta^{-1}) = \zeta^{-m} = 1.$$

By Hilbert's Theorem 90 ([Hil] § 54 Theorem 90 p. 104), there exists β in L such that ζ^{-1} equals $\beta\sigma(\beta)^{-1}$. In other words, $\sigma(\beta)$ equals $\zeta\beta$. Since σ is a generator of the Galois group, we see that the Galois conjugates of β are the $\zeta^i\beta$'s where i ranges from 0 to $m-1$ and these are all distinct because of our assumption concerning the characteristic of F. Furthermore, $\sigma(\beta^m)$ equals $\zeta^m\beta^m$ which is β^m. Thus β^m is fixed by all elements of $\mathrm{Gal}(L|F)$ and therefore belongs to K. Taking a to be β^n, β is a root of $X^m - a$ and this is the minimal polynomial of β over K. □

Proposition 4.13. *Let F be a field an m a natural number such that the characteristic of F does not divide m. Suppose that F contains the group of m^{th} roots of unity μ_m. Any field extension of the form $F(\sqrt[m]{a_i} \mid 1 \leq i \leq r)$ for some natural number r and elements a_i in F is abelian of exponent m.*

Proof. Since F contains μ_m it is clear that all such extensions are Galois. Each Galois morphism is uniquely determined by the images of the $\sqrt[m]{a_i}$'s. Fixing ζ a primitive m-root of unity, the image of $\sqrt[m]{a_i}$ is given by $\zeta^{d_i}\sqrt[m]{a_i}$ for some non-negative integer d_i less than m. The set of such integers $\{d_i : 1 \leq i \leq r\}$ uniquely determines a Galois morphism. Consider two morphisms σ and τ which are respectively determined by the sets $\{d_i\}$ and $\{d'_i\}$. Then

$$\tau(\sigma(\sqrt[m]{a_i})) = \tau(\zeta^{d_i}\sqrt[m]{a_i}) = \zeta^{d_i}\tau(\sqrt[m]{a_i}) = \zeta^{d_i+d'_i}\sqrt[m]{a_i} = \sigma(\tau(\sqrt[m]{a_i}))$$

by symmetry of the expression in d_i and d'_i. This holds for each i and it follows that the extension is abelian. Keeping the same notation for σ and composing it m times with itself, we obtain

$$\sigma^m(\sqrt[m]{a_i})) = (\zeta^m)^{d_i}\sqrt[m]{a_i} = \sqrt[m]{a_i}$$

and since this holds for all i and all σ, this shows that the extension is of exponent m. □

Theorem 4.14. *Let F be a field an m a natural number such that the characteristic of F does not divide m. Suppose that F contains the group of m^{th} roots of unity μ_m. Every finite abelian extension of F which has exponent m is of the form $F(\sqrt[m]{a_i} \mid 1 \leq i \leq r)$ for some natural number r and elements a_i in F.*

Proof. Let L be any abelian extension of exponent m of F. Let n denote the degree of the extension L/F. The Galois group of L over F is a finite abelian group of exponent m and thus, by the Structure Theorem of finite abelian groups ([Sam] ch. 1 § 5 Corollary 1 and 2 p. 22), there exist natural numbers $d_1, \ldots, d_r$ such that d_i divides d_j for every j greater than i, d_r divides m and

$$\mathrm{Gal}(L|F) \cong \mathbb{Z}/d_1\mathbb{Z} \oplus \ldots \oplus \mathbb{Z}/d_r\mathbb{Z}.$$

For i ranging between 1 and r, define H_i to be the quotient of $\mathrm{Gal}(L|F)$ by the subgroup corresponding to $\mathbb{Z}/d_i\mathbb{Z}$ via the above isomorphism. It follows that

$$H_i \cong \mathbb{Z}/d_1\mathbb{Z} \oplus \ldots \oplus \widehat{\mathbb{Z}/d_i\mathbb{Z}} \oplus \ldots \oplus \mathbb{Z}/d_r\mathbb{Z}$$

where the hat indicates that the term is not in the sum. This is a normal subgroup of $\mathrm{Gal}(L|F)$ and we denote by L_i the subfield of elements of L which are fixed by H_i. This is a Galois extension of F with Galois group over F isomorphic to $\mathbb{Z}/d_i\mathbb{Z}$. It is thus a cyclic extension of F of degree d_i. Since d_i divides m and F contains μ_m, it also contains μ_{d_i}. By Theorem 4.12, $L_i = F(\sqrt[d_i]{a_i})$ for some a_i in F. Since d_i divides m, we may define b_i to be a_i^{m/d_i} which is in F. Then $L_i = F(\sqrt[m]{b_i})$. All these fields are distinct and they do not contain one another by construction.

The extension L contains L_i as a sub-extension for each i. Thus L contains the compositum of these fields which is the field $M := F(\sqrt[m]{b_i} \mid 1 \leq i \leq r)$. Since L/F is a Galois extension, L/M is also a Galois extension. The Galois group $\mathrm{Gal}(L|M)$ consists exactly of the elements of $\mathrm{Gal}(L|F)$ that fix M. Consider an element σ of $\mathrm{Gal}(L|M)$. It fixes M if and only if it fixes L_i for all i since M is the compositum of these fields. Thus $\sigma|_{L_i}$ is the identity for all i. But the Galois group $\mathrm{Gal}(L_i|F)$ is isomorphic to $\mathbb{Z}/d_i\mathbb{Z}$ and thus σ becomes trivial in this group for all i. But

$$\mathrm{Gal}(L|F) \cong \mathbb{Z}/d_1\mathbb{Z} \oplus \ldots \oplus \mathbb{Z}/d_r\mathbb{Z}$$

and therefore σ is the identity in $\mathrm{Gal}(L|F)$. This proves that the Galois group $\mathrm{Gal}(L|M)$ is trivial and as a consequence

$$L = M = F(\sqrt[m]{b_i} \ : \ 1 \leq i \leq r),$$

which concludes the proof. □

We now return to the proof of Theorem 4.7. Let K, m and S as in the assumptions of this theorem. As we have seen, we may supose that K contains μ_m. Since K is a number field, its characteristic is zero and we may apply Theorem 4.14 to K which implies that all finite abelian extension of exponent m of K are contained in $K(\sqrt[m]{a} \mid a \in K)$. We will now exploit the ramification assumption we made and it will become clear why this assumption is crucial.

Definition 4.15. Let K be a number field, m a natural number greater than 2 and S a finite set of places in M_K containing the archimedean places M_K^∞. We define the ring of S-integers of K to be

$$\mathcal{O}_{K,S} = \{a \in K \mid v(a) \geq 0 \text{ for all } v \in M_K \setminus S\} = \bigcap_{v \in M_K \setminus S} \mathcal{O}_{K,\mathfrak{p}_v},$$

where $\mathfrak{p}_v$ denotes the prime ideal of $\mathcal{O}_K$ corresponding to v.

We will briefly study the structure of $\mathcal{O}_{K,S}$ and show that it can be realized as a localization of $\mathcal{O}_K$.

Proposition 4.16. *Let K be a number field, m a natural number greater than 2 and S a finite set of places in M_K containing the archimedean places M_K^∞.*

- *(i) There exists a non-zero element a of $\mathcal{O}_K$ such that the finite set of non-archimedean places v for which $v(1/a)$ is negative is exactly $S \setminus M_K^\infty$.*
- *(ii) Let a be a non-zero element of $\mathcal{O}_K$ such that the finite set of non-archimedean places v for which $v(1/a)$ is negative is exactly $S \setminus M_K^\infty$. Then $\mathcal{O}_{K,S} = \mathcal{O}_K[1/a]$.*

Proof. In order to prove (i), let $h(K)$ be the ideal class number of K. It is finite (Theorem 2, ch.4, § 3, p.58 [Sam], ch.4 § 3 Theorem 2 p.58) and for any prime ideal $\mathfrak{p}$ of $\mathcal{O}_K$, the ideal $\mathfrak{p}^{h(K)}$ is principal. Hence the finite product $\prod_{v\in S\setminus M_K^\infty} \mathfrak{p}_v^{h(K)}$ is a principal ideal of $\mathcal{O}_K$. Let a_S be a generator of this ideal. This is a non-zero element of $\mathcal{O}_K$ for which $v(1/a_S)$ is negative for exactly the discrete valuations in S.

To prove (ii), let a be a non-zero element of $\mathcal{O}_K$ such that $v(1/a)$ is negative for all v in $S \setminus M_K^\infty$ and $v(1/a)$ is non-negative for all v outside S. Then $1/a$ belongs to $\mathcal{O}_{K,S}$. Since $\mathcal{O}_K$ is contained in $\mathcal{O}_{K,S}$, we get that $\mathcal{O}_K[1/a]$ is contained in $\mathcal{O}_{K,S}$. We want this to be an equality. Let therefore x be an element of $\mathcal{O}_{K,S}$. Since a belongs to $\mathcal{O}_K$, the element $a^n x$ belongs to $\mathcal{O}_{K,S}$ for all positive n and thus $v(a^n x)$ is positive for all v outside S. Let v be in $S \setminus M_K^\infty$. Then $v(a^n x) = nv(a) + v(x)$. But $v(a)$ is positive for such a v, so letting n_v be a large enough number, we get that $v(a^{n_v} x)$ is positive. Since S is finite, we may take N to be the maximum of all n_v where the maximum is taken over all v in $S \setminus M_K^\infty$. Then $v(a^N x)$ is non-negative for all discrete valuations in M_K and thus $a^N x$ belongs to $\mathcal{O}_K$ so that x belongs to $\mathcal{O}_K[1/a]$. This prove that $\mathcal{O}_K[1/a]$ contains $\mathcal{O}_{K,S}$ and ends the proof of (i). □

Lemma 4.17. *Let A be a commutative ring and let S be a multiplicative subset of A that does not contain 0. Then the prime ideals of $S^{-1}A$ are exactly $\mathfrak{p}S^{-1}A$ where $\mathfrak{p}$ is a prime ideal of A such that $\mathfrak{p} \cap A$ is empty.*

Proof. Let $\mathfrak{p}$ be a prime ideal of A such that $\mathfrak{p} \cap A$ is empty. Let $\frac{x}{s}$ and $\frac{y}{s'}$ be two elements of $S^{-1}A$ such that $\frac{xy}{ss'}$ belongs to $\mathfrak{p}S^{-1}A$. Then xy belongs to $\mathfrak{p}$ and therefore either x or y is in $\mathfrak{p}$, say x. But then $\frac{x}{s}$ belongs to $\mathfrak{p}S^{-1}A$. This proves that $\mathfrak{p}S^{-1}A$ is indeed a prime ideal of $S^{-1}A$.

Let I be a prime ideal of $S^{-1}A$. Then $I \cap A$ is a prime ideal of A, say $\mathfrak{p}$. Clearly, I contains $\mathfrak{p}S^{-1}A$. Thus, if $\mathfrak{p} \cap S$ was non-empty, then $I = A$. Therefore $\mathfrak{p} \cap S$ is empty. Let x be an element of I. There exists s an element of S such that sx belongs to A and therefore also to $\mathfrak{p}$. It follows that x belongs to $\mathfrak{p}S^{-1}A$ and we have proved the equality $I = \mathfrak{p}S^{-1}A$. □

Corollary 4.18. *Let K be a number field, m a natural number greater than 2 and S a finite set of places in M_K containing the archimedean places M_K^∞. The prime ideals of $\mathcal{O}_{K,S}$ are exactly the ideals $\mathfrak{p}\mathcal{O}_{K,S}$ where $\mathfrak{p}$ is a prime ideal of $\mathcal{O}_K$ that corresponds to a place v outside S.*

Proof. By Proposition 4.16, $\mathcal{O}_{K,S} = S^{-1}\mathcal{O}_K$ where S is the multiplicative set of power of a_S where a_S is the generator of the ideal $\prod_{v \in S \setminus M_K^\infty} \mathfrak{p}_v^{h(K)}$. By Lemma 4.17, the prime ideals of $\mathcal{O}_{K,S}$ are exactly the ideals $\mathfrak{p}\mathcal{O}_{K,S}$ where $\mathfrak{p}$ is a prime ideal of $\mathcal{O}_K$ for which $\mathfrak{p} \cap S$ is empty. By definition of a_S, an ideal $\mathfrak{p}$ satisfies this intersection property if and only if it corresponds to a place v outside S. □

Proposition 4.19. *Let K be a number field, m a natural number greater than 2 and S a finite set of places in M_K containing the archimedean places M_K^∞. Then $\mathcal{O}_{K,S}$ is a Dedekind domain.*

Proof. The ring $\mathcal{O}_{K,S}$ contains the Dedekind domain $\mathcal{O}_K$. The fact that $\mathcal{O}_{K,S}$ is of dimension 1 follows from the preceding lemma and the fact that $\mathcal{O}_K$ is of dimension 1.

The fact that $\mathcal{O}_{K,S}$ is integrally closed follows from the fact that is is the intersection of the localizations $\mathcal{O}_{K,\mathfrak{p}}$ which are all integrally closed by Proposition 6.8 (Appendix B § 6.1).

Finally, $\mathcal{O}_{K,S}$ is Noetherian since by Proposition 4.16 it is the localization of the Noetherian ring $\mathcal{O}_K$. □

Proposition 4.20. *Let K be a number field, m a natural number greater than 2 and S a finite set of places in M_K containing the archimedean places M_K^∞. Let K'/K be a finite extension of K and let S' be the set of places in $M_{K'}$ that lie above the ones in S. This is a finite set of places that does not contain $M_{K'}^\infty$ and the integral closure of $\mathcal{O}_{K,S}$ in K' is $\mathcal{O}_{K',S'}$.*

Proof. It is clear that S' is a finite set of places that does not contain $M_{K'}^\infty$. Let a be a non-zero element of $\mathcal{O}_K$ such that the finite set of non-archimedean places v for which $v(1/a)$ is negative is exactly $S \setminus M_K^\infty$. By Proposition 4.16, such an element exists and $\mathcal{O}_{K,S} = \mathcal{O}_K[1/a]$.

Let v' be a place of $S' \setminus M_{K'}^{\infty}$. It lies above a place v in $S \setminus M_K^{\infty}$. Thus there exists a positive constant c such that $v'(1/a) = cv(1/a)$. In particular, $v'(1/a)$ is negative. Now let v' be a place in $M_{K'}$ outside S'. For the same reason, $v'(1/a)$ is non-negative. Since $\mathcal{O}_K$ is contained in $\mathcal{O}_{K'}$, by Proposition 4.16, $\mathcal{O}_{K',S'} = \mathcal{O}_{K'}[1/a]$. By Proposition 6.10 (Appendix B § 6.1), $\mathcal{O}_{K'}[1/a]$ is the integral closure of $\mathcal{O}_K[1/a]$ in K'. □

Proposition 4.21. *Let K be a number field, m a natural number greater than 2 and S a finite set of places in M_K containing the archimedean places M_K^{∞}. Let K'/K be a finite extension of K and let S' be the set of places in $M_{K'}$ that lie above the ones in S. Then K'/K is unramified outside S if and only if $\mathcal{O}_{K,S}$ does not ramify in $\mathcal{O}_{K',S'}$.*

Proof. From Proposition 4.20, $\mathcal{O}_{K',S'}$ is the integral closure of $\mathcal{O}_{K,S}$ and the prime ideals of $\mathcal{O}_{K,S}$ are the ideals $\mathfrak{p}\mathcal{O}_{K,S}$ where $\mathfrak{p}$ is a prime ideal of $\mathcal{O}_K$ that corresponds to a place outside S. Similarly, the prime ideals of $\mathcal{O}_{K',S'}$ are the ideals $\mathfrak{q}\mathcal{O}_{K',S'}$ where $\mathfrak{q}$ is a prime ideal of $\mathcal{O}_{K'}$ that corresponds to a place outside S'. It is clear that $\mathfrak{q}$ divides $\mathfrak{p}$ if and only if $\mathfrak{q}\mathcal{O}_{K',S'}$ divides $\mathfrak{p}\mathcal{O}_{K,S}$. Moreover, the localization of $\mathcal{O}_{K,S}$ away from $\mathfrak{p}\mathcal{O}_{K,S}$ is equal to the localization of $\mathcal{O}_K$ away from $\mathfrak{p}$. Since ramification properties can be studied through localization away from the given prime ideal, the result is now clear. □

Lemma 4.22. *Let K be a number field, m a natural number greater than 2 and S a finite set of places in M_K containing the archimedean places M_K^{∞}. Let S' be another finite set of places that contains S. If the conclusion of Theorem 4.7 holds for K and S', then it also holds for K and S.*

Proof. Let $K_{m,S'}^{ab}$ be a finite abelian field extension of K that contains all finite abelian extensions of exponent m of K that are unramified outside S'. Since S' is larger than S, being unramified outside S implies being unramified outside S' and we may therefore take $K_{m,S}^{ab}$ to be $K_{m,S'}^{ab}$. □

Proposition 4.23. *Let K be a number field, m a natural number greater than 2 and S a finite set of places in M_K containing the archimedean places M_K^{∞}. There exists a finite set of places S' containing S such that $\mathcal{O}_{K,S'}$ is principal.*

Proof. By finiteness of the ideal class group (Theorem 2, ch.4, § 3, p.58 [Sam], ch.4 § 3 Theorem 2 p.58) we may choose integral ideals $I_1, \dots, I_r$ that represent the ideal classes of K. We may take these to be ideals of $\mathcal{O}_K$. We decompose these ideals into products of prime ideals. Then we define S' to be S together with the places corresponding to these prime ideals. To see that $\mathcal{O}_{K,S'}$ is indeed principal, consider an ideal I of $\mathcal{O}_{K,S'}$. In particular it is an $\mathcal{O}_{K,S'}$-module and can be seen as an $\mathcal{O}_K$-module since $\mathcal{O}_{K,S'}$ contains $\mathcal{O}_K$ and as such it is a fractional ideal in K. Thus there exists f in K^* and some j such that $I = fI_j$. Note that this is an equality of $\mathcal{O}_K$-modules. As $\mathcal{O}_{K,S'}$-modules, we get $I = fI_j\mathcal{O}_{K,S'}$. To see that I is principal, it suffices to prove that $I_j\mathcal{O}_{K,S'}$ is principal. Decomposing I_j into prime ideals, it suffices to see that $\mathfrak{p}\mathcal{O}_{K,S'}$ is principal where $\mathfrak{p}$ is a prime ideal dividing I_j. Let x be an element of this $\mathfrak{p}\mathcal{O}_{K,S'}$ and let v be the place corresponding to $\mathfrak{p}$. Then $v'(x) = 0$ for all other discrete valuations in M_K. In particular, $v'(x)$ is zero for all discrete valuations outside S' and thus x is invertible in $\mathcal{O}_{K,S}$. Hence, $\mathfrak{p}\mathcal{O}_{K,S'} = \mathcal{O}_{K,S'}$. This proves that $\mathcal{O}_{K,S'}$ is a principal ideal domain. □

We have reduced the initial problem to the case where K contains the group μ_m and $\mathcal{O}_{K,S}$ is principal. We complete this section with this final result.

Theorem 4.24. *Let K be a number field, m a natural number greater than 2 and S a finite set of places in M_K containing the archimedean places M_K^∞. Suppose that K contains μ_m and that $\mathcal{O}_{K,S}$ is principal. Then there exists a finite abelian field extension $K_{m,S}^{ab}$ of K such that all finite Galois extensions of K that are abelian of exponent m and unramified outside S are contained in $K_{m,S}^{ab}$.*

Proof. From Theorem 4.14, we know that all such extensions are contained in $K(\sqrt[m]{a}|a \in K)$ but this did not use the ramification assumption. Note that it suffices to adjoin the m^{th} roots of elements in $\mathcal{O}_K$ since K is the fraction field of $\mathcal{O}_K$. Also note that it suffices to adjoin one representative of each class in $\mathcal{O}_K/(\mathcal{O}_K)^m$. We claim that it actually suffices to adjoin one representative of each class in $\mathcal{O}_{K,S}^*/(\mathcal{O}_{K,S}^*)^m$. By the S-unit Theorem ([Coh], ch. 3 Theorem 3.3 p. 102), $\mathcal{O}_{K,S}^*$ is a finitely generated group, hence $\mathcal{O}_{K,S}^*/(\mathcal{O}_{K,S}^*)^m$ is finite and we may take $K_{m,S}^{ab}$ to be $K(\sqrt[m]{a} \mid a \in \mathcal{O}_{K,S}^*/(\mathcal{O}_{K,S}^*)^m)$.

We now prove the claim. Let K'/K be a finite abelian extension of exponent m that is unramified outside S. By Theorem 4.14, $K' = K(\sqrt[m]{a_i} \mid 1 \leq i \leq r)$. As already mentioned, me may take the a_i's to be elements of $\mathcal{O}_K$. Let a be an element of $\mathcal{O}_K$ such that there exists b in K' such that $b^m = a$. Let S' be the set of places in $M_{K'}$ that lie above the ones in S. From Proposition 4.21, being unramified outside S is equivalent to $\mathcal{O}_{K,S}$ not ramifying in $\mathcal{O}_{K',S'}$. Consider the ideal of $\mathcal{O}_{K,S}$ generated by the element a. Since $\mathcal{O}_{K,S}$ and $\mathcal{O}_{K',S'}$ are Dedekind, we may write

$$a\mathcal{O}_{K,S} = \mathfrak{p}_1^{f_1} \dots \mathfrak{p}_r^{f_r} \quad \text{and} \quad b\mathcal{O}_{K',S'} = \mathfrak{q}_1^{e_1} \dots \mathfrak{q}_s^{e_s}$$

where the $\mathfrak{p}_i$'s are prime ideals of $\mathcal{O}_{K,S}$ and $\mathfrak{q}_i$'s are prime ideals of $\mathcal{O}_{K',S'}$. Consider the ideal of $\mathcal{O}_{K',S'}$ generated by a. Since K'/K is unramified outside S and $b^m = a$, we must have for all i that $f_i = me_{j_i}$ for some $1 \leq j_i \leq s$. Thus

$$a\mathcal{O}_{K,S} = \mathfrak{p}_1^{me_{j_1}} \dots \mathfrak{p}_r^{me_{jr}}.$$

Consider the ideal $\mathfrak{p}_1^{e_{j_1}} \dots \mathfrak{p}_r^{e_{jr}}$ in $\mathcal{O}_{K,S}$. Since $\mathcal{O}_{K,S}$ is principal, there exists, $c \in \mathcal{O}_{K,S}$ such that $\mathfrak{p}_1^{e_{j_1}} \dots \mathfrak{p}_r^{e_{jr}} = c\mathcal{O}_{K,S}$. Thus $a = c^m u$ for some unit u in $\mathcal{O}_{K,S}^*$ and it follows that $K(\sqrt[m]{a}) = K(\sqrt[m]{u})$. Thus $K' = K(\sqrt[m]{a_i} \mid 1 \leq i \leq r, a_i \in \mathcal{O}_{K,S}^*/(\mathcal{O}_{K,S}^*)^m)$. In particular, K' is contained in the field $K(\sqrt[m]{a} \mid a \in \mathcal{O}_{K,S}^*/(\mathcal{O}_{K,S}^*)^m)$ which is a finite extension of K and abelian by Proposition 4.13. This is true for all K', so the proof is complete. □

4.3 Reduction modulo π

By Corollaries 4.6 and 4.9, the proof of the Weak Mordell-Weil Theorem 2.2 has been reduced to finding a finite set S of places in M_K that contains the archimedean absolute values and proving that the fields $K_G(Q)$ are unramified outside this set. We will need some technical tools from the general theory of elliptic curves in order to complete the proof.

From now until the end of the next section we will use the following notation unless otherwise stated:

K a field that is complete with respect to a normalized discrete valuation v

R the valuation ring of v

$\mathcal{M}$ the maximal ideal of R

π a uniformizing parameter for v

ord_v the normalized valuation in $[v]$

k the residue field $k = R/\pi R$.

Proposition 4.25. *Let E/K be an elliptic curve. The following statements hold:*

(i) In the class of Weierstrass curves that are isomorphic to E, there exists a Weierstrass curve for which $v(\Delta) \geq 0$ is minimal subject to the constraint that all the coefficients of its Weierstrass equation are elements of R.

(ii) A Weierstrass equation as described in (i) is unique up to the change of coordinates

$$x' = u^2 x + r \qquad\qquad y' = u^3 y + sx + t$$

where u is a unit of R and r, s, t are elements of R.

(iii) If one starts with an equation with coefficients in R, then any admissible change of coordinates used to produce a Weierstrass equation as in (i) is subject to the condition that u, r, s, t belong to R.

Proof. (i) We claim that we can always choose a representative Weierstrass curve with equation

$$y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6$$

in the class of Weierstrass curves that are isomorphic to E that has all coefficients a_i in R. To see this, start with any Weierstrass curve isomorphic to E given by an equation

$$y^2 + a'_1xy + a'_3y = x^3 + a'_2x^2 + a'_4x + a'_6$$

with coefficients in K. Let u be an element of K^* and apply the admissible change of variables $(x', y') = (u^2x, u^3y)$. We obtain an isomorphic Weierstrass curve given by the equation:

$$y^2 + ua'_1xy + u^3a'_3y = x^3 + u^2a'_2x^2 + u^4a'_4x + u^6a'_6.$$

So choosing u to be a sufficiently large power of π yields new coefficients which are all elements of R. This proves the claim.

Now consider any Weierstrass curve that is isomorphic to E and that is given by an equation

$$y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6$$

with coefficients in R. By Definitions 1.6 and 1.4, the discriminant Δ of this equation is a polynomial in the coefficients a_i. Since R is a ring, this implies that Δ is an element of R. Thus $v(\Delta)$ is positive. But the valuation v is discrete, hence among the Weierstrass curves isomorphic to E given by equations with coefficients in R, there exists at least one such that $v(\Delta)$ is minimal. This proves (i).

(ii) Suppose we have started with a Weierstrass curve as in (i) and apply an admissible change of coordinates to its equation giving an equation for a new Weierstrass curve that also satisfies (i). By the formulas 1.11.1, $u^{12}\Delta' = \Delta$, so if this new equation is to be as in (i), then necessarily $v(\Delta')$ must be equal to $v(\Delta)$ which in turn implies that $v(u)$ is 0 and as a consequence u is a unit of R. The new coefficients a'_i must be elements of R. Using for example the first equation of 1.11.1, $ua'_1 = a_1 + 2s$, we must have that s belongs to R. Using the other equations, one can deduce that r and t also must belong to R.

(iii) Suppose we start from any Weierstrass curve given by an equation with coefficients in R and perform a change of variables resulting in an equation for a Weierstrass curve satisfying (i). Then $v(\Delta')$ is less than or equal to $v(\Delta)$. Meanwhile, $v(\Delta)$ equals $12v(u) + v(\Delta')$ so that $v(u)$ is non-negative and therefore u belongs to R. As in (ii), one can use the remaining equations of 1.11.1 to show that r, s, t also belong to R. □

Definition 4.26. Let E/K be an elliptic curve. An equation for a Weierstrass curve that is isomorphic to E and that satisfies Proposition 4.25 (i) is called a minimal Weierstrass equation for E/K.

Definition 4.27. The ring R comes equipped with a reduction-modulo-π map

$$\begin{array}{ccc} R & \longrightarrow & k \\ r & \longmapsto & \tilde{r}. \end{array}$$

Let E/K be an elliptic curve given by a minimal Weierstrass equation

$$y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6.$$

The curve $\tilde{E}/k$ defined in $\mathbb{P}^2_k$ by the equation

$$y^2 + \tilde{a}_1xy + \tilde{a}_3y = x^3 + \tilde{a}_2x^2 + \tilde{a}_4x + \tilde{a}_6$$

is called the reduced curve modulo π of E. It is unique up to isomorphism.

Remark 4.28. To see why the concept of minimal Weierstrass equation is important, we give the following example: take K to be $\mathbb{Q}_5$, v to be the 5-adic valuation and consider the two Weierstrass curves given by

$$y_1^2 = x_1^3 + 1 \qquad y_2^2 = x_2^3 + 5^6.$$

These two Weierstrass curves are isomorphic by the change of variables $(5^{-2}x, 5^{-3}y)$. Computing the discriminant gives $\Delta_1 = -432$ and $\Delta_2 = -432.5^{12}$. Reducing the two equations gives two Weierstrass curves defined in $\mathbb{P}^2_{\mathbb{F}_5}$ but only the first one is smooth so they cannot be isomorphic.

Requiring the Weierstrass equation to be minimal ensures that the reduced curve is unique up to isomorphism. But it does not necessarily imply that the class of isomorphic curves we obtain is non-singular.

Definition 4.29. Let E/K be the elliptic curve given by the Weierstrass equation

$$y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6.$$

Denote by $\tilde{E}/k$ the reduced curve associated to this equation. If the reduced curve is non-singular, then we say that E/K has good reduction at v. Otherwise, we say that E/K has bad reduction at v. It makes sense to talk about the reduction of the curve E/K since different choices of a minimal Weierstrass equation give rise to isomorphic reduced curves and non-singularity is preserved via isomorphisms.

Definition 4.30. Let $P = [x, y, z]$ be a point in $\mathbb{P}^2_K(K)$. Multiplying by a suitable power of π, we may find coordinates $[x_0, y_0, z_0]$ for P where x_0, y_0, z_0 all belong to R and at least one of them is a unit in R. We define the reduction of P to be the point $\tilde{P} = [\tilde{x}_0, \tilde{y}_0, \tilde{z}_0]$ which is a point in $\mathbb{P}^2(k)$. This gives a map

$$\begin{array}{ccc} \mathbb{P}^2(K) & \longrightarrow & \mathbb{P}^2(k) \\ P & \longmapsto & \tilde{P} \end{array}$$

which we will call the reduction-modulo-π map. Let E/K be an elliptic curve in $\mathbb{P}^2_K$. The above map restricts to a map from $E(K)$ to $\tilde{E}(k)$ which is called the reduction-modulo-π map on E.

Definition 4.31. Let E/K be an elliptic curve in $\mathbb{P}^2_K$. We define $E_1(K)$ to be the following set:

$$E_1(K) := \{P \in E(K) \mid \tilde{P} = O_{\tilde{E}}\}.$$

Proposition 4.32. *Let E/K be an elliptic curve in $\mathbb{P}^2_K$ given by a Weierstrass equation*

$$y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6$$

with coefficients in R. We have the following equality:

$$E_1(K) = \{[x, y, 1] \in E(K) \mid v(x) < 0\} \cup \{O_E\}$$

and if $[x, y, 1]$ is an element of $E_1(K)$, then $2v(y) = 3v(x)$.

Proof. Suppose P is a point of $E_1(K)$ with coordinates $[x_0, y_0, z_0]$ which all belong to R and at least one of the coordinates is a unit. If z_0 is zero, then P is O_E. If this is not the case, then necessarily z_0 belongs to (π) since $\tilde{z}_0$ must be zero. But one of the other two coordinates is a unit, so dividing by z_0, the coordinates of P become $[x, y, 1]$ where $x = x_0/z_0$ and $y = y_0/z_0$. Whence x or y has valuation equal to $-v(z_0)$ which is negative. On the other hand, if P has coordinates $[x, y, 1]$ where either x or y has negative valuation, then multiplying by $z_0 = \pi^{\max(|v(x)|,|v(y)|)}$ we get new coordinates $[x_0, y_0, z_0]$ that are all in R and either x_0 or y_0 is a unit. Then $\tilde{z}_0$ is zero, so $\tilde{P} = O_{\tilde{E}}$. We conclude that

$$E_1(K) = \{[x, y, 1] \in E(K) \mid v(x) < 0 \text{ or } v(y) < 0\} \cup \{O_E\}\}.$$

The result now follows by proving that $2v(y) = 3v(x)$ for all $[x, y, 1]$ in $E_1(K)$.

Using the Weierstrass equation

$$y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6,$$

we take the valuation of both sides. Using the multiplicative property (*ii*) and the non-archimedean property (*iii*) of discrete valuations (see Appendix B § 6.1 Definition 6.2), $v(a_2x^2+a_4x+a_6)$ is greater than $\min(v(a_2)+2v(x), v(a_4)+v(x), v(a_6))$. Since all coefficients a_i are elements of R, their valuations are non-negative. Thus, $v(a_2)+2v(x)$ is greater or equal to $2v(x)$ which in turn is strictly greater than $3v(x)$ since $v(x)$ is less or equal -1. The same argument shows that $v(a_4)+v(x)$ is strictly greater than $3v(x)$. As a consequence, $v(a_2x^2+a_4x+a_6)$ is strictly greater than $3v(x)$ so $v(x^3+a_2x^2+a_4x+a_6)$ is equal to $\min(3v(x), v(a_2x^2+a_4x+a_6))$ which is $3v(x)$. We have shown that the valuation of the right hand side is equal to $3v(x)$. Suppose by contradiction that the valuation of y is positive. Using the same properties of the valuation as above, the valuation of the left hand side is greater than $\min(2v(y), v(a_1y)+v(x), v(a_3)+v(y))$. But, $2v(y)$ and $v(a_3)+v(y)$ are both non-negative and $v(a_1y)+v(x)$ is equal to $\min(v(x), v(a_1y))$ which is $v(x)$. So the valuation on the left hand side is greater than the valuation of x. Putting everything together we see that $v(x)$ is less than or equal to $3v(x)$ which is a contradiction. We conclude that the valuation of y is negative.

Suppose next that the valuation of x is non-negative. This would imply that

$$v(x^3+a_2x^2+a_4x+a_6)$$

is non-negative. Suppose by contradiction that the valuation of y is negative. Then the valuation of the left hand side would be equal to $2v(y)$ which is negative and this would be a contradiction. We conclude that the valuation of y is non-negative.

Now, suppose that the valuation of both x and y is negative. Then we have already seen that the valuation of the right hand side is $3v(x)$. Using the same properties as above, we see that the valuation of the left hand side is $\min(2v(y), v(x))$. If this is $v(x)$ we would have a contradiction, so it necessarily equals $2v(y)$. Finally, we conclude that in this case we have the equality $2v(y) = 3v(x)$. □

Proposition 4.33. *Let E/K be an elliptic curve in $\mathbb{P}^2_K$ and suppose that E/K has good reduction at v. Then the reduction-modulo-π map is a homomorphism of abelian groups.*

Proof. The proof is not particularly enlightening and quite long. We have therefore placed it in Appendix A. □

Remark 4.34. Let E/K be an elliptic curve in $\mathbb{P}^2_K$. If the reduced curve $\tilde{E}/k$ is non-singular, then $E_1(K)$ is exactly the kernel of the reduction-modulo-π homomorphism.

Proposition 4.35. *Let E/K be an elliptic curve in $\mathbb{P}^2_K$ and let m be a natural number that is relatively prime to the characteristic of K. Suppose that $E_1(K)$ has no non-trivial m-torsion points and that E/K has good reduction at v. Then the reduction-modulo-π homomorphism gives an injective homomorphism $E(K)[m] \hookrightarrow \tilde{E}(k)$.*

Proof. Consider the homomorphism $E(K)[m] \hookrightarrow E(K) \twoheadrightarrow E(K)/E_1(K)$ which has kernel equal to $E(K)[m] \cap E_1(K)$. By the assumption on $E_1(K)$, this intersection is trivial, so the above homomorphism is injective. Since $E_1(K)$ is by definition the kernel of the reduction-modulo-π map, we have an injective homomorphism $E(K)/E_1(K) \hookrightarrow \tilde{E}(k)$ and composing the two homomorphisms above yields an injective homomorphism from $E(K)[m]$ to $\tilde{E}(k)$. □

4.4 Formal group of an elliptic curve

The goal in this section is to prove the following:

Proposition 4.36. *Let E/K be an elliptic curve in $\mathbb{P}^2_K$ and let m be a natural number that is relatively prime to the characteristic of the residual field k. Then $E_1(K)$ does not contain any non-trivial m-torsion points.*

As a consequence of Proposition 4.35 and 4.36, we will have proved the following result:

Proposition 4.37. *Let E/K be an elliptic curve in $\mathbb{P}^2_K$ and let m be a natural number that is relatively prime to the characteristic of the residual field k. Suppose that E/K has good reduction at v. Then the reduction-modulo-π homomorphism gives an injective homomorphism*

$$E(K)[m] \hookrightarrow \tilde{E}(k).$$

As we will see in the next section, this result is a key argument in the proof of the Weak Mordell-Weil Theorem 2.2. The strategy to the proof of Proposition 4.36 is to construct the formal group $\hat{E}(\mathcal{M})$ of the elliptic curve E. This formal group contains no non-trivial m-torsion points (Appendix D Proposition 8.9). Then we will prove that $\hat{E}(\mathcal{M})$ is isomorphic as a group to $E_1(K)$.

If one takes an elliptic curve E/K in $\mathbb{P}^2_K$ given by a Weierstrass equation, then the tangent line to $E(\bar{K})$ at O_E is defined by the equation $Z = 0$. This leads to the intuitive idea that we should be able to parametrize the points of the curve E locally around O_E using the

x-coordinate as a parameter. As it happens, this can be done using formal power series and this will lead us to the definition of the formal group of the elliptic curve.

Parametrization Let E/K be the elliptic curve in $\mathbb{P}^2_K$ given by the Weierstrass equation

$$Y^2Z + a_1XYZ + a_3YZ^2 - X^3 - a_2X^2Z - a_4XZ^2 - a_6Z^3 = 0.$$

As mentioned above, the idea is to study the curve locally around O_E which is given by homogeneous coordinates $[0, 1, 0]$. We dehomogenizing the above equation with respect to Y, placing ourselves in the affine chart $Y = 1$ where O_E is given by the coordinates $(0, 0)$. In this chart, the Weierstrass equation becomes

$$z + a_1xz + a_3z^2 - x^3 - a_2x^2z - a_4xz^2 - a_6z^3 = 0.$$

Define $g(x, z)$ to be the polynomial $x^3 - a_1xz + a_2x^2z - a_3z^2 + a_4xz^2 + a_6z^3$. Then the Weierstrass equation is

$$z = g(x, z). \tag{4.37.1}$$

Consider the ring of regular functions $A := A(E_Y) = K[x, z]/(g(x, z) - z)$ and localize it with respect to the origin. Denote the localization $A_{\mathcal{O}}$. This is a local ring with maximal ideal $\mathfrak{m}_{\mathcal{O}}A$ where $\mathfrak{m}_{\mathcal{O}}$ is the ideal of $K[x, z]$ of polynomials that vanish at the origin. Take the completion $\hat{A}_{\mathcal{O}}$ of $A_{\mathcal{O}}$ with respect to its maximal ideal. Since E is smooth at the origin, we know by the Cohen Structure Theorem ([Har] ch.1, Theorem 5.5A p. 34) that $\hat{A}_{\mathcal{O}} = K[[\pi]]$ for some uniformizing parameter π at the origin.

Once can show that x is a uniformizing parameter at the origin. We express z as a formal power series in x by iteratively replacing z by $g(x, z)$ in equation 4.37.1. Formally, we recursively define a sequence

$$g_1(x, z) = g(x, z) \qquad g_{m+1}(x, z) = g_m(x, g(x, z))$$

and then define $w_m(x) = g_m(x, 0)$ which belongs to $\mathbb{Z}[a_1, \ldots, a_6][x]$. Explicitly, the two first terms of this sequence are

$$w_1(x) = x^3, \qquad w_2(x) = x^3 - a_1x^4 + a_2x^5 - a_3x^6 + a_4x^7 + a_6x^9.$$

Proposition 4.38. *The sequence $w_m(x)$ converges in $\mathbb{Z}[a_1, \ldots, a_6][[x]]$ and*

$$z(x) := \lim_{m\to+\infty} w_m(x) = x^3(1 + \ldots) \in \mathbb{Z}[a_1, \ldots, a_6][[x]]$$

is the unique element satisfying $z(x) = g(x, z(x))$.

Proof. We apply Hensel's Lemma ([Sil], ch. 4 Lemma 1.2 p. 117) to the ring $\mathbb{Z}[a_1, \ldots, a_6][[x]]$ which is complete with respect to its maximal ideal (x). Consider the polynomial $F(X)$ in the variable X defined by $X - g(x, X)$. We check that the assumptions of the lemma hold. Since $F(0)$ is equal to $-x^3$ it is an element of (x^3). Computing the derivative of F,

$$F'(X) = 1 + a_1x - a_2x^2 + 2a_3X - 2a_4xX - 3a_6X^2$$

we see that $F'(0)$ is $1 - a_1x - a_2x^2$ which is not an element of (x). This implies that $F'(0)$ is a unit in $\mathbb{Z}[a_1, \ldots, a_6][[x]]$ since this is a local ring with maximal ideal (x). Hensel's Lemma implies that there exists an element $z(x)$ of (x^3) that satisfies $F(z(x)) = 0 = z(x) - g(x, z(x))$. Since $\mathbb{Z}[a_1, \ldots, a_6][[x]]$ is integral it also implies the uniqueness of $z(x)$. □

This proposition tells us that $(x, z(x))$ is a solution to the equation defining E_Y. Note that for each positive integer n, $z(x)$ is equal to $w_n(x)$ modulo (x^{n+3}) and $w_{n+1}(x) = w_n(x) + A(x)x^{n+3}$ for some polynomial A. Computing the first terms of $z(x)$, we get

$$\begin{aligned} z(x) = x^3 - a_1x^4 + (a_1^2 + a_2)x^5 - (a_1^3 + 2a_1a_2 + a_3)x^6 \\ + (a_1^4 + 3a_1^2a_2 + 3a_1a_3 + a_2^2 + a_4)x^7 + (h.o.t.) \end{aligned}$$

From now on we will denote A_i the coefficients of $z(x)$ so that

$$z(x) = \sum_{i=3}^{\infty} A_i x^i \in \mathbb{Z}[a_1, \ldots, a_6][[x]].$$

Furthermore, $z(x)$ is zero if and only x is zero since $\mathcal{O}$ is the unique point of E with last coordinate zero.

Formal addition law We now turn to constructing the power series formally giving the addition law on E. Let x_1 and x_2 be two independent variables. From Proposition 4.38, we know that for $i = 1, 2$ there exists a unique $z_i := z(x_i)$ belonging to $\mathbb{Z}[a_1, \ldots, a_6][[x_i]]$ such that $P_i := (x_i, z_i)$ is a point of E_Y. Let L be the line in the plane $Y = 1$ through P_1 and P_2. This line has equation $L : z = \lambda x + v$ with

$$\lambda = \frac{z_2 - z_1}{x_1 - x_2} = \sum_{i=3}^{\infty} A_i \frac{x_2^i - x_1^i}{x_2 - x_1} \in \mathbb{Z}[a_1, \ldots, a_6][[z_1, z_2]].$$

and $v = z_1 - \lambda x_1$. We are looking for the third intersection of L with E. Substituting $z = \lambda x + v$ in equation 4.37.1 gives a polynomial equation of degree three in the variable x:

$$\begin{aligned} t(x) := x^3(1 + a_6\lambda^3 + a_2\lambda + a_4\lambda^2) + x^2(-a_1\lambda + a_2v - a_3\lambda^2 + 2a_4\lambda v + 3a_6\lambda^2 v) \\ + x(-a_1v - 2a_3\lambda v + a_4v^2 + 3a_6\lambda v^2 - \lambda) + (v^2 + v^3 - v) = 0. \end{aligned}$$

We already know two roots, namely x_1 and x_2. Let a, b, c and d be the coefficients of t so that $t(x) = ax^3 + bx^2 + cx + d$. Let x_3 be the third root. We compute x_3 by solving $t(x) = a(x - x_1)(x - x_2)(x - x_3)$. Identifying the coefficients of x^2 on either side we see that $x_1 + x_2 + x_3 = -b/a$. Thus,

$$x_3 = -x_1 - x_2 - \frac{-a_1\lambda + a_2 v - a_3\lambda^2 + 2a_4\lambda v + 3a_6\lambda^2 v}{1 + a_6\lambda^3 + a_2\lambda + a_4\lambda^2}.$$

Using the formal equality $1/(1 - X) = \sum_{k\geq 0} X^k$, we see that $x_3(x_1, x_2)$ is an element of $\mathbb{Z}[a_1, \ldots, a_6][[x_1, x_2]]$. Defining

$$z_3(x_1, x_2) = \lambda(x_1, x_2)x_3(x_1, x_2) + v(x_1, x_2),$$

we have found the third point of intersection $P_3 = (x_3(x_1, x_2), z_3(x_1, x_2))$. Thus $g(x_3, z_3) = z_3$ and by Proposition 4.38, $z_3 = z(x_3)$. So z_3 can be computed using only x_3 so this variable suffices to describe P_3. Now, by construction, the points P_1, P_2 and P_3 are colinear and therefore by the properties of the group law on E, they sum to zero. The inversion formula we derived in chapter one says the following : $-[x, y, z] = [x, -y - a_1x - a_3z, z]$. If we want to invert a point $[x, 1, z]$ of E in the plane $Y = 1$, then the formula becomes

$$-[x, 1, z] = \left[\frac{-x}{1 + a_1x + a_3z}, 1, \frac{-z}{1 + a_1x + a_3z}\right].$$

We define the inverse of x to be $x(-[x, 1, z(x)])$. Explicitly,

$$\iota(x) = \frac{-x}{1 + a_1x + a_3z(x)}$$

which is an element of $\mathbb{Z}[a_1, \ldots, a_6][[x]]$ (this can easily be checked by using the formula for geometric series).

Definition 4.39. We define the formal group operation $F \in \mathbb{Z}[a_1, \ldots, a_6][[x_1, x_2]]$ as follows:

$$F(x_1, x_2) = \iota(x_3(x_1, x_2)).$$

Proposition 4.40. *The above defined power series F is a formal group law (Appendix D Definition 8.1).*

Proof. By construction, the sum of P_1 and P_2 equals $(F(x_1, x_2), w(F(x_1, x_2)))$ and therefore F immediately inherits properties $(ii) - (v)$ of Definition 8.1 from the geometric group law on E. Using the formulas for x_3 and $\iota(x_3)$, we see that

$$F(x_1, x_2) = -x_3(1 + (h.o.t.)) = x_1 + x_2 + (h.o.t.)$$

so F satisfies property (i), making F into a formal group over E. □

Definition 4.41. Let K be local field complete with respect to a discrete valuation v and $\mathcal{M}$ be the maximal ideal of its valuation ring R. Let E/K be an elliptic curve in $\mathbb{P}^2_K$ defined by a given minimal Weierstrass equation. The formal group of E, denoted $\hat{E}(\mathcal{M})$, is the set $\mathcal{M}$ with the group law induced by the formal group law $F \in R[[X, Y]]$ from Definition 4.39.

Remark 4.42. Note that since E is given by a minimal equation for E, the coefficients a_i are all elements of R and thus $\mathbb{Z}[a_1, \ldots, a_6]$ is included in R. So it makes sense to talk about the formal group F defined over R.

Proposition 4.43. *Let E/K be an elliptic curve in $\mathbb{P}^2_K$ and let m be a natural number that is relatively prime to the characteristic of k. Then $\hat{E}(\mathcal{M})$ does not contain any non-trivial m-torsion points.*

Proof. This is general result concerning formal groups, see Appendix D Proposition 8.9. □

The next result concludes this section and proves Proposition 4.36 by our considerations from the beginning of this section.

Proposition 4.44. *Let E/K be an elliptic curve embedded into $\mathbb{P}^2(\bar{K})$. With the same notations as in the previous section, the map*

$$\begin{array}{cccc} \phi: & \hat{E}(\mathcal{M}) & \longrightarrow & E_1(K) \\ & x & \longmapsto & [x, 1, z(x)] \end{array}$$

is an isomorphism of groups. In particular, if m is a natural number that is relatively prime to the characteristic of k, then $E_1(K)$ does not contain any non-trivial m-torsion points.

Proof. To see that this is a well-defined map, recall from our constructions in the last section that $\phi(x)$ is a point of E. Furthermore, $\phi(x)$ has coordinates in R by completeness (as discussed in Appendix D when defining groups associated to formal groups) so in particular it is a K-rational point of E. It remains to prove that it reduces to $\mathcal{O}$. If x is zero, then $\phi(x)$ is $\mathcal{O}$ so there is nothing to prove. If x is non-zero then $z(x)$ is also non-zero. We rewrite the coordinates as $[x/z(x), 1/z(x), 1]$. Now, recall from Remark 4.32 that it suffices to prove that $x/z(x)$ or $1/z(x)$ has negative valuation. Since x belongs to $\mathcal{M}$, we know that the valuation of x is positive. Using the formula

$$z(x) = x^3 \left(1 + \sum_{n \geq 1} A_{n+3} z^n \right) \in R[[z]]$$

we see that the valuation of $1/z(x)$ is equal to $-v(z(x))$ which is less than or equal to $-3v(z)$ which in turn is negative. As a consequence, $\phi(z)$ belongs to $E_1(K)$ and we conclude that ϕ is well-defined.

To see that it is a homomorphism of groups, let x and y belong to $\hat{E}(\mathcal{M})$. If either one of these two elements, say x, is zero, then

$$\phi(x \oplus_F y) = \phi(y) = \mathcal{O} + \phi(y) = \phi(x) + \phi(y).$$

Otherwise, $\phi(x \oplus_F y)$ is the point $[x \oplus_F y, 1, z(x \oplus_F y)]$. By definition of the group law induced by F, this is $[F(x, y), 1, z(F(x, y))]$. This is, by construction of F, equal to the sum of $[x, 1, z(x)]$ and $[y, 1, z(y)]$.

In order to prove that this is a bijection, we provide the inverse

$$\psi : E_1(K) \longrightarrow \hat{E}(\mathcal{M})$$

that sends a point $[x, y, z]$ of $E_1(K)$ to x/y. Either z is zero and then the point is $\mathcal{O}$, or else z is non-zero in which case we rewrite the coordinates as $[x/z, y/z, 1]$. We know from Remark 4.32 that the valuations of both x/z and y/z are negative. Thus neither x or y can be zero and we can divide by y. Moreover, using the same remark as before, we know that twice the valuation of y/z is three times the valuation of x/z. So there must exist a positive integer r such that

$$2v(y/z) = 3v(x/z) = -6r.$$

But then the valuation of x/y is exactly r which is positive. We conclude that x/y does indeed belong to $\mathcal{M}$ and therefore the map ψ is well-defined.

Now, let x be an element of $\hat{E}(\mathcal{M})$. Then

$$\psi(\phi(x)) = \psi([x, 1, z(x)]) = x.$$

Let $[x, y, z]$ be a point of $E_1(K)$. Then

$$\phi(\psi([x, y, z])) = \phi(x/y) = [x/y, 1, z(x/y)].$$

Recall that $w(x/y)$ is the unique element of R such that $[x/y, 1, w(x/y)]$ belongs to E by Proposition 4.38. On the other hand, $[x/y, 1, z/y]$ already belongs to E. As a consequence, $w(x/y)$ is equal to z/y and $\phi(\psi([x, y, z]))$ is the point $[x, y, z]$. This ends the proof and this section. □

4.5 Non-ramification of the extensions $K_G(Q)$

We are now ready to conclude the proof of the Weak Mordell-Weil Theorem 2.2. In this section, K denotes a number field. The fields $K_{m,E}$ and $K_G(Q)$ are the ones from Definition 4.4. By Corollaries 4.6 and 4.9, in order to conclude the proof, we must find a finite set S of places in M_K that contains M_K^∞ and such that all the extensions $K_G(Q)$ are unramified outside S.

We are now working on global fields in contrast to the last two sections where we were concerned with local fields. We therefore start by adjusting some definitions and restating Proposition 4.36 in this new setting. The definition of a minimal Weierstrass equation with respect to a discrete valuation v in M_K is the same as Definition 4.26.

Definition 4.45. Let v be a discrete valuation in M_K with a uniformizing parameter π. The discrete valuation $\hat{v}$ on the completion K_v also has π as a uniformizing parameter (Appendix C § 7.1). Let E/K be the elliptic curve in P_K^2 defined by the minimal Weierstrass equation (with respect to v)

$$y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6.$$

We define the reduced curve $\tilde{E}_v/k_v$ modulo π of E/K with respect to v to be the reduced curve modulo π of E/K_v in $\mathbb{P}_{K_v}$ with respect to $\hat{v}$ in the sense of Definition 4.27.

Definition 4.46. Let E/K be an elliptic curve defined over a number field K and v be a discrete valuation in M_K. Let $\hat{v}$ be the valuation on K_v that extends v. We will say that E/K has good (resp. bad) reduction at v if the elliptic curve E/K_v has good (resp. bad) reduction at $\hat{v}$ in the sense of Definition 4.29.

Proposition 4.47. *Let E/K be the elliptic curve given by the Weierstrass equation*

$$y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6$$

where the coefficients belong to K. Then E/K has good reduction at v for all but finitely many discrete valuations v's in M_K.

Proof. The coefficients a_i's and the discriminant Δ all belong to K. Looking at the denominators of the a_i's and Δ, there can only be a finite number of prime ideals of $\mathcal{O}_K$ that contain each of them and again only a finite number of prime ideals that contain the numerator of Δ. Thus, for all but finitely many valuations v in M_K^0, we have

$$v(a_i) \geq 0 \text{ for } i = 1, \ldots, 6 \text{ and } v(\Delta) = 0.$$

For all valuations satisfying this, the equation is already minimal and since $v(\Delta) = 0$, the reduction of E at such a v is good. There can therefore only be finitely many places v in M_K^0 at which the reduction of E is bad. □

We reformulate Proposition 4.36 in the case of a number field.

Proposition 4.48. *Let E/K be an elliptic curve in $\mathbb{P}_K^2$ and v a discrete valuation in M_K such that E/K has good reduction at v. Let m be a natural number such that $v(m) = 0$. Then we have an injective homomorphism*

$$E(K)[m] \hookrightarrow E(K_v)[m] \hookrightarrow \tilde{E}(k_v).$$

Proof. The condition $v(m) = 0$ is just a reformulation of the fact that m is relatively prime tho the characteristic of k_v. The result is now immediate from Proposition 4.36. □

The next result concludes the proof of the Weak Mordell-Weil Theorem 2.2.

Proposition 4.49. *Let E/K be an elliptic curve in $\mathbb{P}_K^2$. Let m be a natural number greater than 2 and define*

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

This is a finite set of place in M_K and for all Q in $[m]^{-1}(E(K))$, the extension $K_G(Q)$ is unramified outside S.

Proof. We start by proving that S is a finite set of places. The set S is union of three set: the first one is finite by Proposition 4.47, the second is finite since the integer m can only belong to finitely many prime ideals and the third is M_K^∞ which is finite.

Let Q be an element of $[m]^{-1}(E(K))$. We now prove that $K_G(Q)$ is unramified outside S. From Definition 4.4, $K_G(Q)$ is the compositum of the fields $K(Q^\sigma)$ where σ is an element of the absolute Galois group $\mathrm{Gal}(\bar{K}|K)$ and $K_G(Q)$ is therefore a finite Galois extension of K. We will prove the result using the theory of inertia groups (Appendix B § 6.3).

Let v be a place in $M_K \setminus S$ and let $\mathfrak{p}$ be the corresponding ideal of $\mathcal{O}_K$. Consider any prime ideal $\mathfrak{q}$ of $\mathcal{O}_{K_G(Q)}$ that divides $\mathfrak{p}$ and let w be the corresponding valuation in $M_{K_G(Q)}$. Denote by $\tilde{k}_w$ the residual field of $K_G(Q)$ at w. Consider a minimal Weierstrass equation of E over the completion K_v and let Δ be the associated discriminant. The fact that v is outside S implies that E has good reduction at v and therefore $v(\Delta)$ is zero. Since w extends v, we also have that $w(\Delta)$ is zero and therefore this equation is also a minimal Weierstrass equation of E over $K_G(Q)_w$ and $E/K_G(Q)_w$ has good reduction at w. By Proposition 4.33, the reduction map

$$E(K_G(Q)) \longrightarrow \tilde{E}_w(\tilde{k}_w)$$

is a homomorphism. Let σ be an element of the inertia group $I_{\mathfrak{q}/\mathfrak{p}}$. By Definition 6.25, the action of σ on $\tilde{k}_w$ and therefore on $\tilde{E}_w(\tilde{k}_w)$ is trivial. Let τ be an element of $\mathrm{Gal}(\bar{K})|K))$. Using the fact that the reduction is a homomorphism, we get that:

$$\widetilde{(Q^\tau)^\sigma - Q^\tau} = (\widetilde{Q^\tau})^\sigma - \widetilde{Q^\tau} = \tilde{\mathcal{O}}.$$

On the other hand, $m((Q^\tau)^\sigma - Q^\tau)$ is equal to $(mQ^\tau)^\sigma - mQ^\tau$. The fact that mQ^τ belongs to $E(K)$ implies that this is zero and therefore $(Q^\tau)^\sigma - Q^\tau$ belongs to $E(K_G(Q))[m]$ which injects into $\tilde{E}(\tilde{k}_w)$ by Proposition 4.48 since $w(m) = 0$. Therefore we must have that $(Q^\tau)^\sigma = Q^\tau$ and this equality holds for all τ in $\mathrm{Gal}(\bar{K}|K)$. In other words, σ fixes all the conjugates of Q and since $K_G(Q)$ is the compositum of the conjugate fields of $K(Q)$, σ must be the identity. Since this argument is valid for all σ in $I_{\mathfrak{q}/\mathfrak{p}}$, it proves that the inertia group is trivial and $K_G(Q)/K$ is unramified at v. This holds for all v that does not belong to S and therefore $K_G(Q)/K$ is unramified outside S. □

5 Appendix A - Reduction-modulo-π is a homomorphism of groups

Let K be a field that is complete with respect to a discrete valuation v. We will use the notation introduced in Section 4.3. We prove that the reduction-modulo-π map on E/K is a homomorphism of groups in the case where $\tilde{E}/k$ is non-singular. We will need to treat several different cases but first we prove two useful lemmas.

Definition 5.1. For any line L in $\mathbb{P}^2_K(\bar{K})$ defined over K, it is always possible to find an equation

$$L : aX + bY + cZ = 0$$

where the coefficients a, b and c are all in R and at least one of them is a unit (it suffices to multiply the equation by a well-chosen power of π). We define the reduced line $\tilde{L}$ to be the line in $\mathbb{P}^2(k)$ defined by the equation

$$\tilde{L} : \tilde{a}X + \tilde{b}Y + \tilde{c}Z = 0.$$

Clearly, if a point P belongs to L, then the reduced point $\tilde{P}$ belongs to the reduced curve $\tilde{L}$.

Lemma 5.2. *Let E/K be the elliptic curve in $\mathbb{P}^2_K$ given by the minimal Weierstrass equation*

$$Y^2Z + a_1XYZ + a_3YZ^2 - X^3 - a_2X^2Z - a_4XZ^2 - a_6Z^3.$$

and suppose that the reduced curve $\tilde{E}/k$ is non-singular. Let P and Q be two distinct K-rational points of E that reduce to the same point and let L be the line through P and Q. Then the reduced line $\tilde{L}$ is tangent to $\tilde{E}$ at $\tilde{P}$.

Proof. Suppose first that $\tilde{P}$ is not $O_{\tilde{E}}$. Then P has coordinates $[\alpha, \beta, 1]$ with α and β belonging to R (see Proposition 4.32). Since Q reduces to the same point as P, this point necessarily has coordinates of the form $[\alpha + \mu, \beta + \nu, 1]$ where μ and ν both belong to (π). Since neither P, Q nor $\tilde{P}$ is O_E, we dehomogenize the Weierstrass equation with respect to the variable Z and place ourselves in the affine chart $Z = 1$. The equation of the elliptic curve in this chart is

$$y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6.$$

We define $f(x, y) = y^2 + a_1xy + a_3y - (x^3 + a_2x^2 + a_4x + a_6)$. Since $\tilde{P}$ is a non-singular point of $\tilde{E}(k)$, either $\partial_x\tilde{f}(\tilde{P})$ or $\partial_y\tilde{f}(\tilde{P})$ is non-zero.

We start by considering the case where $\partial_y \tilde{f}(\tilde{P})$ is non-zero. We compute the Taylor expansion of $f(x, y)$ around the point P and evaluate it at Q :

$$f(Q) = f(P) + \partial_x f(P)\mu + \partial_y f(P)\nu - (3\alpha - a_2)\mu^2 + a_1\mu\nu + \nu^2 - \mu^3.$$

Since $f(P)$ and $f(Q)$ are zero, we obtain

$$\partial_y f(P)\nu = -\partial_x f(P)\mu + (3\alpha - a_2)\mu^2 - a_1\mu\nu - \nu^2 + \mu^3.$$

By assumption, $\partial_y f(P)$ is a unit in R. Thus the valuation of ν is the same as the one of $\partial_y f(P)\nu$. Using the above equality, this valuation is the same as

$$v(\partial_x f(P)\mu + (3\alpha - a_2)\mu^2 + a_1\mu\nu + \nu^2 + \mu^3).$$

But this is greater than the minimum of the valuations of each term of the sum. Since $v(\partial_x f(P))$ belongs to R, the valuation of $\partial_x f(P)\mu$ is necessarily greater than the one of μ. The same is true for the valuation of $(3\alpha - a_2)\mu^2$, $a_1\mu\nu$ and μ^3. Finally, it is impossible for the minimum to be the valuation of ν^2 since this would lead to the contradiction that $v(\nu)$ is greater than $2v(\nu)$. So the minimum is necessarily greater than the valuation of μ. To sum up, we have shown that $v(\nu)$ is larger than $v(\mu)$. As a consequence, ν/μ belongs to R. Therefore we may divide the Taylor expansion by μ and reduce it modulo π. From this we obtain

$$\partial_x \tilde{f}(\tilde{P}) + \partial_y \tilde{f}(\tilde{P})\widetilde{\nu/\mu} \equiv 0 \mod \pi.$$

As a consequence, the slope of the tangent line to $\tilde{E}$ at $\tilde{P}$ is given by

$$-\frac{\partial_x \tilde{f}(\tilde{P})}{\partial_y \tilde{f}(\tilde{P})} = \widetilde{\nu/\mu}.$$

On the other hand, the equation of L is given by

$$L : y - \beta = \frac{\nu}{\mu}(x - \alpha).$$

So the reduced line $\tilde{L}$ is the line through $\tilde{P}$ with slope $\widetilde{\nu/\mu}$. Thus, it is the tangent line to $\tilde{E}$ at $\tilde{P}$.

Now consider the case where $\partial_x \tilde{f}(\tilde{P})$ is non-zero and $\partial_y \tilde{f}(\tilde{P})$ is zero. Then the tangent line to $\tilde{E}$ at $\tilde{P}$ is the vertical line $x - \tilde{\alpha} = 0$. Looking again at the Taylor expansion, we see that

$$\partial_x f(P)\mu - (3\alpha - a_2)\mu^2 + a_1\mu\nu + \nu^2 - \mu^3 = 0$$

and by reasoning in the same way as before, one can easily see that the valuation of μ is greater than the one of ν, whence μ/ν belongs to R. Dividing by ν and reducing modulo π, we obtain

$$\partial_x \tilde{f}(\tilde{P})\widetilde{\mu/\nu} \equiv 0 \mod \pi$$

and therefore μ/ν is in (π). Since the line L is given by the equation

$$L : \mu/\nu(y - \beta) = x - \alpha,$$

the equation of the reduced line $\tilde{L}$ is $x - \tilde{\alpha} = 0$ so it is the tangent line to $\tilde{E}$ at $\tilde{P}$.

We now turn to the case where $\tilde{P}$ is $O_{\tilde{E}}$. The tangent line to $\tilde{E}$ at this point is given by the equation $Z = 0$ since the gradient of $\tilde{F}$ evaluated at $O_{\tilde{E}}$ is $(0, 0, 1)$. Since P and Q are both in $E_1(K)$, neither of them can have Y-coordinate equal to zero as a consequence of Proposition 4.32. We place ourselves in the affine chart $Y = 1$ by dehomogenizing the Weierstrass equation with respect to the variable Y. The equation of the elliptic curve in this chart is

$$z + a_1xz + a_3z^2 - x^3 - a_2x^2z - a_4xz^2 - a_6z^3 = 0.$$

We define $g(x, z) = z + a_1xz + a_3z^2 - x^3 - a_2x^2z - a_4xz^2 - a_6z^3$. First, suppose that P is O_E and the coordinates of Q are $[\alpha, 1, \beta]$. From Proposition 4.32, the valuation of β is positive and strictly greater than the one of α. The line through P and Q in this plane is given by $L : z = \frac{\beta}{\alpha}x$. Note that α is non-zero and therefore the division makes. In fact, suppose α is zero. Then β satisfies $\beta(1 + a_3\beta) = 0$. But β is in (π), so the only possibility is $\beta = 0$. But then P and Q coincide and this contradicts our assumption that P and Q are distinct. Since the valuation of β is strictly larger than the one of α, β/α belongs to (π) and therefore reduces to zero modulo π. As a consequence, the reduced line $\tilde{L}$ is indeed the tangent line to $\tilde{E}$ at $O_{\tilde{E}}$.

Now, suppose none of the two points P and Q are O_E. Let $[\alpha_1, 1, \beta_1]$ and $[\alpha_2, 1, \beta_2]$ be the respective coordinates of P and Q. Necessarily, the valuation of β_i must be positive and strictly greater than the one of α_i (see Proposition 4.32). The line through P and Q in the chart $Y = 1$ is given by

$$L : z - \beta_1 = \frac{\beta_2 - \beta_1}{\alpha_2 - \alpha_1}x - \frac{\beta_2 - \beta_1}{\alpha_2 - \alpha_1}\alpha_1.$$

Suppose the valuations of the α_i's are distinct and without loss of generality that the minimum of the two is the one of α_1. As already noted, β_1 is an element of (π), so it reduces to zero modulo π. Moreover, the valuation of $\beta_2 - \beta_1$ is greater than $\min(v(\beta_1), v(\beta_2)$. But the valuation of β_i is strictly greater than the one of α_i. So this minimum is strictly greater than $\min(v(\alpha_1), v(\alpha_2))$ which is the valuation of α_1 which in turn is the valuation of the sum of the α_i's. Hence, $(\beta_2 - \beta_1)/(\alpha_2 - \alpha_1)$ is an element of (π) and reduces to zero modulo π. Finally, the valuation of the last term is

$$v(\alpha_1) + v(\beta_1 + \beta_2) - v(\alpha_1 + \alpha_2) = v(\beta_1 + \beta_2) > 0$$

so this term also reduces to zero and therefore the reduced line is indeed the tangent line to $\tilde{E}$ at $O_{\tilde{E}}$. □

Lemma 5.3. *Let E/K be the elliptic curve in $\mathbb{P}^2_K$ given by the minimal Weierstrass equation*

$$Y^2Z + a_1XYZ + a_3YZ^2 - X^3 - a_2X^2Z - a_4XZ^2 - a_6Z^3 = 0$$

and suppose that the reduced curve $\tilde{E}/k$ is non-singular. If P is a point of $E(K)$ and L is the tangent line to E at P, then the reduced line $\tilde{L}$ is tangent to $\tilde{E}$ at $\tilde{P}$.

Proof. If P is O_E, then L is defined by the equation $Z = 0$ which reduces to the same equation. Thus $\tilde{L}$ is tangent to $\tilde{E}$ at $O_{\tilde{E}}$.

Suppose that P belongs to $E_1(K) \setminus \{O_E\}$. Let the coordinates of P be $[\alpha, \beta, \gamma]$ where all coordinates are in R and at least one of then is a unit of R. By Proposition 4.32, we must have that the valuation of γ is strictly greater than the ones of α and β so in particular it is positive. Note that the partial derivatives of F evaluated at P all lie in R. The equation for L is given by

$$L : \partial_x F(P)(X - \alpha) + \partial_y F(P)(Y - \beta) + \partial_z F(P)(Z - \gamma) = 0.$$

Reducing modulo π, we see that $\partial_x \tilde{F}(\tilde{P})$ and $\partial_y \tilde{F}(\tilde{P})$ are both zero whereas $\partial_z \tilde{F}(\tilde{P})$ equals 1 since P reduces to the origin of $\tilde{E}$. Finally, γ reduces to zero and therefore the equation of the reduced line $\tilde{L}$ is given by $Z = 0$ which is the tangent line to $\tilde{E}$ at $O_{\tilde{E}}$.

Finally, suppose P is not an element of $E_1(K)$. We work directly in the affine chart $Z = 1$. Let $[\alpha, \beta, 1]$ be the coordinates of P. Then the valuations of α and β must be non-negative, and therefore the coordinates of the reduced point $\tilde{P}$ are $[\tilde{\alpha}, \tilde{\beta}, 1]$. The equation of L in the chart $Z = 1$ is given by

$$L : \partial_x F(P)(X - \alpha) + \partial_y F(P)(Y - \beta) = 0$$

which reduces to

$$\tilde{L} : \partial_x \tilde{F}(\tilde{P})(X - \tilde{\alpha}) + \partial_y \tilde{F}(\tilde{P})(Y - \tilde{\beta}) = 0$$

which is the tangent line to $\tilde{E}$ at $\tilde{P}$ in $Z = 1$. □

Proposition 5.4. *Let E/K be the elliptic curve in $\mathbb{P}^2_K$ given by the Weierstrass equation*

$$Y^2Z + a_1XYZ + a_3YZ^2 - X^3 - a_2X^2Z - a_4XZ^2 - a_6Z^3 = 0$$

and suppose $\tilde{E}/k$ is non-singular. Then the reduction-modulo-π map is a homomorphism of abelian groups.

Proof. We will show that if P_1, P_2 and P_3 are points of $E(K)$ that sum to O_E, then necessarily the reduced points also sum to $O_{\tilde{E}}$. There are six cases to consider.

(1) Suppose all three points are distinct and the reduced points are also distinct. Let L be the line through P_1, P_2 and P_3. Then the reduced points all belong to the reduced line $\tilde{L}$ and since they are distinct, there can be no more intersection points of $\tilde{E}$ with $\tilde{L}$ by Bézout's Theorem ([Ful] ch.5 § 3 p. 57). Hence, the reduced points sum to zero.

(2) Suppose all three points are distinct and that P_1 and P_2 reduce to the same point that is distinct from $\tilde{P}_3$. Let L be the line through P_1, P_2 and P_3. By Lemma 5.2, the reduced line $\tilde{L}$ is tangent to $\tilde{E}$ at $\tilde{P}_1$. The reduced point $\tilde{P}_3$ also belongs to $\tilde{L}$, so the intersection with $\tilde{E}$ has multiplicity three and therefore $2\tilde{P}_1 + \tilde{P}_3$ is zero by Bézout's Theorem ([Ful] ch.5 § 3 p. 57). Since $\tilde{P}_1$ equals $\tilde{P}_2$, this is the same as to say that $\tilde{P}_1, \tilde{P}_2$ and $\tilde{P}_3$ sum to zero.

(3) Suppose all three points are distinct and that they all reduce to the same point $\tilde{P}_1$ with coordinates $[\tilde{\alpha}, \tilde{\beta}, \tilde{\gamma}]$. We know from Lemma 5.2 that the reduced line $\tilde{L}$ is tangent to $\tilde{E}$ at $\tilde{P}_1$. Now suppose, there is a point $\tilde{P}$ distinct from $\tilde{P}_1$ that belongs to the intersection of $\tilde{L}$ with $\tilde{E}$. The intersection of L with E is determined by a homogeneous polynomial f in two variables, say X and Y (the other cases are done similarly). Let $\tilde{f}$ denote the reduced polynomial. Then the intersection $\tilde{L} \cap \tilde{E}$ is given by the equation $\tilde{f} = 0$. Let $[\tilde{x}_0, \tilde{y}_0, \tilde{z}_0]$ be the coordinates of the new point $\tilde{P}$. Consider points x_0 and y_0 in R that reduce to $\tilde{x}_0$ and $\tilde{y}_0$ respectively. Since $\tilde{E}$ is smooth, either $\partial_x \tilde{f}(\tilde{x}_0, \tilde{y}_0)$ or $\partial_y \tilde{f}(\tilde{x}_0, \tilde{y}_0)$ is non-zero. Suppose without loss of generality that the latter is true. Consider the polynomial $h(Y) = f(x_0, Y)$. Since $\tilde{f}(\tilde{x}_0, \tilde{y}_0)$ equals zero, $h(y_0)$ belongs to (π). The ring R is complete with respect to the ideal (π). We just saw that $h(y_0)$ belongs to (π). Moreover, $h'(y_0)$ is a unit of R by our assumption on the partial derivative of $\tilde{f}$. By Hensel's Lemma (Appendix C § 7.1 Lemma 7.9), there exists a unique element b of R such that $h(b) = f(x_0, b)$ is zero. Then the point P defined by the coordinates x_0 and b is a point in the intersection of L and E which reduces to $\tilde{P}$, so it is necessarily distinct from P_1, P_2 and P_3. This contradicts Bézout's Theorem ([Ful] ch.5 § 3 p. 57). Consequently, $3\tilde{P}_3 = O_{\tilde{E}}$.

(4) Suppose that P_1 and P_2 coincide but are distinct from P_3 which reduces to a point distinct from $\tilde{P}_1$. By Lemma 5.3, the reduced line $\tilde{L}$ is tangent to $\tilde{P}_1$. Furthermore, $\tilde{P}_3$ belongs to $\tilde{L}$ and is distinct from $\tilde{P}_1$. Counting intersection multiplicities, we see that $2\tilde{P}_1 + \tilde{P}_3$ equals zero.

(5) Suppose that P_1 and P_2 coincide but are distinct from P_3 which in this case reduces to $\tilde{P}_1$. Playing the same game as before where we used Hensel's Lemma (Appendix C § 7.1 Lemma 7.9), we see that necessarily $3\tilde{P}_1$ is zero.

(6) Suppose that P_1, P_2 and P_3 all coincide. By a similar argument using Hensel's Lemma (Apendix C § 7.1 Lemma 7.9), we see that $3\tilde{P}_1$ is zero. □

6 Appendix B - Discrete valuations and ramification

We establish the link between the classical theory of ramification and discrete valuations associated to a Galois extension. We essentially follow Chapter 1 of [Se2] and Chapter 6 of [Sam]. We then discuss completion of number fields, where we follow Chapter 2 of [Se2] and Chapter 2 of [Lan].

6.1 Discrete valuation rings

In what follows we will suppose that all rings are domains.

Definition 6.1. A ring R is said to be a discrete valuation ring if it is a Noetherian local domain and its maximal ideal $\mathfrak{m}$ is principal.

Definition 6.2. A valuation on a field K is a map

$$v : K \longrightarrow \mathbb{R} \cup \{+\infty\}$$

that enjoys the following properties:

(i) For each element x of K, $v(x) = +\infty$ if and only if $x = 0$.

(ii) The application $v : K^* \longrightarrow \mathbb{Z}$ is a homomorphism of groups.

(iii) For all elements x and y in K, $v(x + y) \geq \min\{v(x), v(y)\}$.

Definition 6.3. A valuation v on K is said to be discrete if $v(K^*) = s\mathbb{Z}$ for some real positive number s. If $s = 1$, we say that v is normalized.

Definition 6.4. Two valuations v and w on K are said to be equivalent if there exists a positive number s such that $v = sw$. This is clearly an equivalence relation. An equivalence class of valuations is called a place of K and is denoted by brackets.

Notation Let v be a discrete valuation on K with $v(K^*) = s\mathbb{Z}$. Dividing v by s gives an equivalent normalized valuation. So for all places there exists a normalized representative that we will denote by ord_v or simply ord when there is no confusion in the notations.

Proposition 6.5. *Let R be a domain and K be its fraction field. The following statements are equivalent:*

(i) $R = v^{-1}(\mathbb{R}_{\geq 0})$ for some discrete valuation v. This is called the valuation ring of v.

(ii) R is a discrete valuation ring.

(iii) The exists a uniformizing parameter of R. In other words, there exists an element π of R such that for all non-zero elements r of R there exists a unique unit u and a positive integer n such that $r = \pi^n u$.

Proof. We start by proving that (i) implies (ii). Suppose that $R = v^{-1}(\mathbb{R}_{\geq 0})$ for some discrete valuation v on K and let $s > 0$ be such that $v(K*) = s\mathbb{Z}$.

In order to prove that R is a local ring consider the set $\mathfrak{m} = v^{-1}(\mathbb{Z}_{>0})$. We claim that this is an ideal of R. Let x be an element of $\mathfrak{m}$ so that $v(x) > 0$. Let y be an element of R so that $v(y) \geq 0$. Then $v(xy) = v(x) + v(y) > 0$ and thus xy belongs to $\mathfrak{m}$. This proves our claim.

We now prove that $\mathfrak{m}$ is a maximal ideal. Let x be an element of $R \setminus \mathfrak{m}$. In other words, x is an element of R for which $v(x) = 0$. But then $v(x^{-1}) = 0$ and thus x^{-1} belongs to R. Consequently, x is a unit in R. Thus $\mathfrak{m}$ is an ideal of R that contains all of R but the units. Let I be any ideal of R that contains $\mathfrak{m}$ strictly. Then there exists an element x in $I \setminus \mathfrak{m}$. This is a unit and therefore I contains 1, so that $I = R$. This proves that $\mathfrak{m}$ is indeed a maximal ideal.

We now prove that $\mathfrak{m}$ is the unique maximal ideal. Let $\mathfrak{m}'$ be another maximal ideal. It cannot be contained in $\mathfrak{m}$ since otherwise we would have $\mathfrak{m} = R$ which is a contradiction. Thus there is a point x in $\mathfrak{m}' \setminus \mathfrak{m}$. But then x is a unit and $\mathfrak{m}' = R$. We conclude that $\mathfrak{m}$ is the unique maximal ideal. We have then proved that R is a local ring. It remains to show that R is Noetherian and that $\mathfrak{m}$ is principal.

We prove that $\mathfrak{m}$ is principal. We may pick an element π of $\mathfrak{m}$ such that $v(\pi) = s$ since $v(K*) = s\mathbb{Z}$. Let x be a non-zero element of $\mathfrak{m}$ and let $sn = v(x)$. Write $x = \pi^n \frac{x}{\pi^n}$. Then $v(\frac{x}{\pi^n}) = v(x) - nv(\pi) = 0$ so that $\frac{x}{\pi^n}$ is a unit in R. This proves that x belong to the ideal generated by π. Since x is an arbitrary element of $\mathfrak{m}$, we conclude that $\mathfrak{m}$ is contained in (π) and therefore $\mathfrak{m} = (\pi)$.

We prove that R is Noetherian. Let I be any proper ideal of R. Being proper implies that I does not contain any units and must therefore be contained in $\mathfrak{m}$. In particular, all elements of I have positive valuation and since the valuation is discrete, there must exist a minimum value to the set $\{v(y) \mid y \in I\}$. Let sn be this number. We claim that $I = (\pi^n)$. First of all, let x be an element of I for which $v(x) = sn$. Then as in the previous paragraph, we can write $x = \pi^n u$ where u is a unit in R. Thus π^n belongs to I and we have the inclusion $(\pi^n) \subset I$. For the other inclusion, let z be any element of I and let $sm = v(z)$. By definition of n, we have that m is greater or equal to n. As in the previous paragraph, we can write $z = \pi^m u$ where u is a unit. We get that $z = \pi^n \pi^{m-n} u$ and $v(\pi^{m-n}u) = s(m-n) \geq 0$ so $\pi^{m-n}u$ belongs to R. Thus z belong to (π^n). Since z is an arbitrary element of I, we have proved that $I \subset (\pi^n)$. We conclude that $I = (\pi^n)$. We have prove that the only ideals of R are the ones of the form (π^n) with n a natural number. In particular, R is Noetherian.

We conclude that R is a discrete valuation ring.

We prove that (ii) implies (iii). Let π be a generator of the maximal ideal $\mathfrak{m}$ of R. We will prove that π is a uniformizing parameter of R.

Let r be an element of R. We prove that we can write $r = \pi^n u$ for some n and some unit u. If r does not belong to (π), then it is necessarily a unit so there is nothing to prove. If r does belong to (π), then there exists r_1 in R such that $r = \pi r_1$. If r_1 does not belong to $\mathfrak{m}$ it is a unit and we are done. Other wise $r_1 = \pi r_2$ for some r_2 in R. By contradiction suppose all r_i's belong to $\mathfrak{m}$. Then we have an increasing sequence of ideals

$$(r) \subset (r_1) \subset (r_2) \subset \ldots (r_i) \subset \ldots$$

and since R is Noetherian this sequence must become stationary. That is, for some integer n, $(r_k) = (r_n)$ for all k greater than n. Hence $r_n = \pi r_{n+1}$ and $r_{n+1} = ar_n$ for some element a of R. It follows that r_{n+1} equals $a\pi r_{n+1}$ so that $r_{n+1}(1 - a\pi)$ is zero. Since R is a domain and $1 - a\pi$ does not belong to $\mathfrak{m}$ (since 1 is not in $\mathfrak{m}$), we must have $r_{n+1} = 0$ and therefore all r_i's are zero which is a contradiction.

We prove the uniqueness of the equality $r = \pi^n u$. Suppose that $r = \pi^n u = \pi^m s$ where u and s are both units and suppose without loss of generality that n is greater than m. Then $\pi^m(\pi^{n-m}u - s)$ is zero and since R is a domain, $\pi^{n-m}u = s$ which does not belong to $\mathfrak{m}$. Thus $n = m$ and consequently $u = s$.

We prove that (iii) implies (i). Let π be a uniformizing parameter of R and consider the map

$$\text{ord} : K \longrightarrow \mathbb{Z} \cup \{+\infty\}$$

defined by sending $r/s = (\pi^a r_0)/(\pi^b s_0)$ (where r_0 and s_0 are units) to $a - b$ and 0 to $+\infty$. One can check that this is a discrete valuation and that $R = \text{ord}^{-1}(\mathbb{Z} \geq 0)$. □

Corollary 6.6. *Let K be a field with two discrete valuations v and w. Then v and w determine the same place if and only if their valuation rings are equal.*

Proof. The direct implication is clear.

For the converse, denote by R_v and R_w the respective valuation rings of v and w and $\mathfrak{m}_v$ and $\mathfrak{m}_w$ the corresponding maximal ideals. Suppose that $R_v = R_w$. Then also $R_v^* = R_w^*$

and $\mathfrak{m}_v = \mathfrak{m}_w$. By Proposition 6.5, there exists π a uniformizing parameter such that $\mathfrak{m}_v = \pi R_v = \pi R_w = \mathfrak{m}_w$. Since π generates the maximal ideals, π must have the minimal valuation among the elements in $\mathfrak{m}_v = \mathfrak{m}_w$. Set $v(\pi) = s$ and $w(\pi) = s'$. Since the valuations are discrete, this gives $v(K^*) = s\mathbb{Z}$ and $w(K^*) = s'\mathbb{Z}$. Clearly, for all x in $R_v^* = R_w^*$, $v(x) = w(x) = 0$. Let x be a non-zero element in $K \setminus R_v^* = K \setminus R_w^*$. By Proposition 6.5 (iii) and the fact that K is the fraction field of $R_v = R_w$, we can write x uniquely as $x = \pi^n u$ for some $n \in \mathbb{Z}$ and u in $R_v^* = R_w^*$. So, $v(x) = ns$ and $w(x) = ns'$ and this gives $w(x) = \frac{s}{s'}v(x)$. This equality holds for all x and therefore v and w are equivalent. □

The goal of this section is to prove the following: if A is a Dedekind domain ([Sam] ch. 3 § 3 p. 49 Definition 1), then for any non-zero prime ideal $\mathfrak{p}$ of A, the local ring $A_\mathfrak{p}$ is a discrete valuation ring. This is important in the sense that the fields we are interested in are number fields that the ring of algebraic integers $\mathcal{O}_K$ of a number field K is a Dedekind domain ([Sam] ch. 3 § 3 p. 49 Theorem 1 since $\mathcal{O}_K$ is the integral closure of $\mathbb{Z}$ and $\mathbb{Z}$ is a Dedekind domain).

Lemma 6.7. *Let R be a discrete valuation ring with valuation v defined on the fraction field K of R. Let $x_1, \ldots, x_n$ be elements of K such that $v(x_i) > v(x_1)$ for all i larger than 2. Then the sum of the x_i's is non-zero.*

Proof. Dividing all x_i's by x_1, we may assume that $x_1 = 1$ and therefore $v(x_i) > 0$ for all i. In other words, each x_i for i greater than 2 belongs to $\mathfrak{m}$. But then the sum of the x_i's does not belong to $\mathfrak{m}$ since x_1 is not an element of the maximal ideal. In particular, the sum cannot be equal to zero. □

Proposition 6.8. *Let R be a Noetherian domain and K is fraction field. In order for R to be a discrete valuation ring it is necessary and sufficient that it satisfies the following two conditions :*

(i) R is integrally closed.

(ii) R possesses one and only one non-zero prime ideal.

Proof. We prove the "necessary direction". Let R be a discrete valuation ring. By Proposition 6.5, there exists a discrete valuation v such that $R = v^{-1}(\mathbb{Z}_{\geq 0})$ and we may pick a uniformizing parameter π of R. By Corollary 6.6, we may take v to be ord_v. Then the ideals of R are exactly the ideals (π^n) where n is a natural number (see proof of Proposition 6.5). Thus the maximal ideal (π) is the unique prime ideal.

We now prove that R satisfies (i). Let x be an element of K that satisfies an equation of integral dependence $x^n + a_{n-1}x^{n-1} + \ldots + a_0 = 0$ where the coefficients a_i are elements of R. Suppose by contradiction that x does not belong to R. Then $v(x) = -m$ for some positive integer m. Then the first term of the above equation has valuation equal to $-nm$ while for each i between 0 and $n-1$, $v(a_{n-i}x^i) = v(a_{n-1}) - im$ which is greater than $-im$ and therefore strictly greater than $-nm$. By Lemma 6.7, the sum cannot be zero which is a contradiction. Thus $x \in R$ and R is integrally closed.

We prove the "sufficient" direction. Suppose that R is a Noetherian domain that satisfies (i) and (ii). The second property implies that R is a local ring since R is commutative and must therefore possess a maximal ideal which is in particular prime, hence unique.

It remains to prove that the maximal ideal $\mathfrak{m}$ of R is principal. Let $\mathfrak{m}'$ be set of all elements x in K such that $x\mathfrak{m}$ is contained in R. It is an R-submodule of K. If y is a non-zero element of $\mathfrak{m}$, then $\mathfrak{m}'$ is contained in $y^{-1}R$ and since R is Noetherian, $\mathfrak{m}'$ is a finitely generated R-module. Since $\mathfrak{m}'$ contains R, the product $\mathfrak{m}.\mathfrak{m}'$ contains $\mathfrak{m}$. By definition of $\mathfrak{m}'$, $\mathfrak{m}.\mathfrak{m}'$ is contained in R. But $\mathfrak{m}.\mathfrak{m}'$ is an ideal of R hence by maximality of $\mathfrak{m}$, we necessarily have that $\mathfrak{m}.\mathfrak{m}'$ is either equal to $\mathfrak{m}$ or R.

We prove that properties (i) and (ii) imply that $\mathfrak{m}.\mathfrak{m}'$ equals R. By contradiction, suppose the product equals $\mathfrak{m}$. Let x be an element of $\mathfrak{m}'$. Then $x\mathfrak{m}$ is contained in $\mathfrak{m}$. Thus $x(x\mathfrak{m})$ is contained in $x\mathfrak{m}$ which is in $\mathfrak{m}$. Repeating this, we see that $x^n\mathfrak{m}$ is contained in $\mathfrak{m}$ for every non-negative integer n. Let $\mathfrak{a_n}$ be R-submodule of K generated by the elements $1, x, \ldots, x^n$. This defines an increasing sequence of modules all contained in the finitely generated module $\mathfrak{m}'$. Since R is Noetherian this sequence becomes stationary. Hence for some large enough n, x^n belongs to $\mathfrak{a_{n-1}}$ and thus has an equation of integral dependence over R. Thus x belongs to the integral closure of R which is R by property (i). As a consequence $\mathfrak{m}' = R$.

We will now see that property (ii) makes this into a contradiction. Let x be any element of $\mathfrak{m}$ and let S be the multiplicative set of powers x^n of x with n a non-negative integer. Let $R_x = S^{-1}R$ be the ring of elements y/x^n with y in R. Suppose by contradiction that R_x is not K. Then it is not a field and thus possesses a proper prime ideal $\mathfrak{p}$. Since x is a unit in R_x it cannot belong to $\mathfrak{p}$. Hence $\mathfrak{p} \cap R$ is not equal to $\mathfrak{m}$. If y/x^n is an element of $\mathfrak{p}$, then y is an element of the intersection $\mathfrak{p} \cap R$. In particular, this intersection is non empty. Thus $\mathfrak{p} \cap R$ is a proper prime ideal of R that is different from $\mathfrak{m}$. This is not possible by property (ii). So we may conclude that R_x equals K. As a consequence, any element of K may be written as y/x^n for some y in R and some non-negative integer n. Let r be an element of R. Write $1/r = y/x^n$ so that x^n equals ry which belongs to the ideal of R generated by r. Let $x_1, \ldots, x_m$ be generators of $\mathfrak{m}$. Then for each i there exists an integer n_i such that $x_i^{n_i}$ belongs to rR. Define n to be the maximum of all the n_i's so that x_i^n belongs to rR for all i. For any positive integer N, the ideal $\mathfrak{m}^N$ is generated by all monomials of degree N in the x_i's. Taking N to be mn, in each of these monomials there will always be one individual power that exceeds n so the whole monomial will belong to rR. Thus for this choice of N, we see that $\mathfrak{m}^N$ is contained in rR. Take r to be an element of $\mathfrak{m}$. There must be a smallest integer N for which $\mathfrak{m}^N$ is contained in rR. Take y to be an element of $\mathfrak{m}^{N-1} - \mathfrak{m}^N$. Then y does not belong to rR but $y\mathfrak{m}$ is contained in rR. Then y/r belongs to $\mathfrak{m}'$ by definition of the latter but it is not an element of R. So $\mathfrak{m}'$ is not equal to R, which is a direct contradiction.

As a conclusion, we necessarily have that $\mathfrak{m}.\mathfrak{m}'$ equals R. So there exist elements $x_1, \ldots, x_n$ of $\mathfrak{m}$ and $y_1, \ldots, y_n$ of $\mathfrak{m}'$ such that $\sum x_i y_i = 1$. All products $x_i y_i$ belong to A and since 1 is not in $\mathfrak{m}$ there is at least one product $x_i y_i$ which is not in $\mathfrak{m}$. Rename x_i and y_1 respectively x and y. The product xy is a unit of R say u, so $(xu^{-1})y = 1$. Replace xu^{-1} with π. This is an element of $\mathfrak{m}$. Now let z be any element of $\mathfrak{m}$. Then $z = \pi(yz)$ belongs to $\pi\mathfrak{m}$. We conclude that $\mathfrak{m} = \pi R$ and thus $\mathfrak{m}$ is principal. □

Corollary 6.9. *Let R be a Dedekind domain and K its field of fractions. If $\mathfrak{p}$ is a non-zero prime ideal of R then the localization $R_\mathfrak{p}$ is a discrete valuation ring.*

Proof. The localization $R_\mathfrak{p}$ is an integral local ring with maximal ideal $\mathfrak{p}R_\mathfrak{p}$. In order to see that it is Noetherian, let I be an ideal of $R_\mathfrak{p}$. Then $I' = I \cap R$ is an ideal of R which is Noetherian so there exist elements $x_1, \ldots, x_n$ of R that generate I'. Let x be an element of I. Then x is the quotient a/b of an element in R by an element in $R - \mathfrak{p}$. Since a equals xb, we see that a is an element of I' so there exist elements $a_1, \ldots, a_n$ of R such that $a = \sum a_i x_i$. But then $x = \sum_i^n \frac{a_i}{b} x_i$ and the a_i/b's are all in $R_\mathfrak{p}$. Consequently, I is generated by the x_i's. This proves that $R_\mathfrak{p}$ is Noetherian.

By the previous proposition, it remains to check that $R_\mathfrak{p}$ is integrally closed. Note that the field of fractions of $R_\mathfrak{p}$ is K. Let x be an element of K that satisfies an equation of integral dependence over $R_\mathfrak{p}$:

$$x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0 = 0$$

where the a_i's all belong to $R_\mathfrak{p}$. All the a_i's can be written as a fraction b_i/c where b_i belongs to R and c belongs to $R - \mathfrak{p}$. Multiplying the above equation by c^n, we obtain

$$(cx)^n + b_{n-1}(cx)^{n-1} + \ldots + b_1 c^{n-2}(cx) + b_0 c^{n-1} = 0$$

which is an equation of integral dependence for cx over R. Since R is integrally closed, we obtain that cx belongs to R and thus x belongs to $R_\mathfrak{p}$. □

Notation We shall denote by $v_\mathfrak{p}$ the normalized valuation coming from $R_\mathfrak{p}$.

The proof of the above result almost immediately proves the more general:

Proposition 6.10. *Let R be a subring of a field K and let S be a multiplicative subset of R that does not contain zero. In order for an element of K to be integral over $S^{-1}R$ it is necessary and sufficient for it to be of the form r/s where r is integral over R and s belongs to S. In other words, localization at S commutes with the integral closure.*

Proof. In order to see that this is necessary, let x be an element of K that satisfies an equation of integral dependence over $S^{-1}R$:

$$x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0 = 0$$

where the a_i's all belong to $S^{-1}R$. All the a_i's can be written as a fraction r_i/s where r_i belongs to R and s belongs to S. Multiplying the above equation by s^n, we obtain

$$(sx)^n + r_{n-1}(sx)^{n-1} + \ldots + r_1 s^{n-2}(cx) + r_0 s^{n-1} = 0$$

which is an equation of integral dependence for sx over R. Since R is integrally closed, we obtain that sx belongs to R and thus x can be written as r/s where r is integral over R.

In order to see that this is sufficient, let x be of the form r/s with r integral over R and s in S. Then r satisfies an equation of integral dependence over R:

$$r^n + a_{n-1}r^{n-1} + \ldots + a_1 r + a_0 = 0$$

where the a_i's belong to R. Dividing by s^n yields a equation of integral dependence for r/s over $S^{-1}R$, so that r/s is integral over $S^{-1}R$. □

6.2 Extensions and ramification

We recall some definitions and basic facts about ramification. Throughout this part, K is a field of characteristic zero. Let A be a Dedekind domain whose field of fractions is K. Let K' be a finite extension of degree n of K and let A' denote the integral closure of A in K'. Since we supposed that the characteristic of K is zero, A' is a finitely generated A-module (Theorem 1, § 2.7 in [Sam] ch. 2 § 7 Theorem 1 p. 40) and is therefore Noetherian. One can show that in characteristic zero, A' is actually a Dedekind domain ([Se2], ch. 1 § 4 prop. 9). It is well known that any ideal of a Dedekind domain factors uniquely as a product of powers of prime ideals ([Sam] ch. 3 § 4).

Definition 6.11. Let $\mathfrak{p}$ be a prime ideal of A. If $\mathfrak{q}$ is a prime ideal of A' such that $\mathfrak{q} \cap A = \mathfrak{p}$, we say that $\mathfrak{q}$ divides $\mathfrak{p}$ or lies above $\mathfrak{p}$ and write $\mathfrak{q}|\mathfrak{p}$.

Proposition-Definition 6.12. *Let $\mathfrak{p}$ be a prime ideal in A. A prime ideal in A' divides $\mathfrak{p}$ if and only if it appears in the decomposition of $\mathfrak{p}A'$ in prime ideals of A'. As a consequence, $\mathfrak{p}$ decomposes in A' as follows:*

$$\mathfrak{p}A' = \prod_{\mathfrak{q}|\mathfrak{p}} \mathfrak{q}^{e_\mathfrak{q}}.$$

The positive integer $e_\mathfrak{q}$ is called the ramification index of $\mathfrak{q}$ at $\mathfrak{p}$. If there exists a prime ideal $\mathfrak{q}|\mathfrak{p}$ with ramification index $e_\mathfrak{q}$ strictly greater than one, then the extension K'/K is said to be ramified at $\mathfrak{p}$. Otherwise, the extension is said to be unramified at $\mathfrak{p}$. The extension is said to be unramified if it is unramified at every prime ideal of A.

Proof. This is clear from Definition 6.11. □

Definition 6.13. With the above notations, the field $A'/\mathfrak{q}$ contains $A/\mathfrak{p}$ as a subfield and since A' is finitely generated over A this is a finite extension. The degree $f_\mathfrak{q}$ of this extension is called the residual degree of $\mathfrak{q}$ at $\mathfrak{p}$.

Proposition 6.14. *Let K be a field of characteristic zero and K' be a finite field extension of K of degree n. Let A be a Dedekind domain whose field of fraction is K and let A' be the integral closure of A in K'. For any prime ideal $\mathfrak{p}$ of A, we have that:*

$$n = \sum_{\mathfrak{q}|\mathfrak{p}} e_\mathfrak{q} f_\mathfrak{q}.$$

Proof. [Se2] ch. 1 § 4 prop. 10. □

Lemma 6.15. *Any discrete valuation ring B is a maximal subring of its field of fractions.*

Proof. Let L be the field of fractions of B. Let C be a subring of L containing B and suppose it is not contained in B. Let v be the normalized discrete valuation representing the place coming from B. Pick x an element of $C \setminus B$. Then $v(x)$ is some negative integer. Let y be any element of L. Necessarily, there exists a positive integer n such that $v(y)$ is greater or equal to $nv(x)$. But then y/x^n belongs to B and in particular it is an element of C. Writing $y = \frac{y}{x^n}x^n$, we see that y belongs to C and therefore $C = L$. Thus B is a maximal subring of L. □

Definition 6.16. Let K and L be two fields and let v and w be two discrete valuations defined respectively on K and L. Suppose that L contains K. We say that w extends v or lies above v and write $w|v$ if there exists a positive constant c such that $w(x) = cv(x)$ for all x in K. In we want to specify the constant we say that w extends v with index c. In more symbolic notation, w extends v if $[w|_K] = [v]$. If w does extend v, then clearly all the valuations in $[w]$ extend the ones in $[v]$. We say that $[w]$ extends $[v]$ and write $[w]|[v]$.

Since A and A' are Dedekind domains, by Corollary 6.9 their localizations away from prime ideals are discrete valuation rings.

Proposition 6.17. *With the above notations, the following statements hold:*

(i) Let $\mathfrak{p}$ be a prime ideal of A and $\mathfrak{q}$ a prime ideal of A' that lies above $\mathfrak{p}$. Then the valuation $v_{\mathfrak{q}}$ extends the valuation $v_{\mathfrak{p}}$ with index $e_{\mathfrak{q}}$.

(ii) Let v be a normalized valuation on K' that extends the valuation $v_{\mathfrak{p}}$ with index e. Then there exists a prime ideal $\mathfrak{q}|\mathfrak{p}$ of A' such that $v = v_{\mathfrak{q}}$ and $e = e_{\mathfrak{q}}$.

Proof. We prove (i). From Proposition 6.14, if x belongs to K, then $v_{\mathfrak{q}}(x) = e_{\mathfrak{q}}v_{\mathfrak{p}}(x)$. This proves (i).

We prove (ii). Let R be the ring of integers of v and $\mathfrak{m}$ be its maximal ideal. By Proposition 6.8, the ring R is integrally closed. Moreover, its field of fractions is K' and R contains A. Indeed, let x be an element of A. We have the inclusions $A \subset A_{\mathfrak{p}} = v_{\mathfrak{p}}^{-1}(\mathbb{Z}_{\geq 0})$. Thus $v_{\mathfrak{p}}(x)$ is non-negative and therefore $v(x)$ is also non-negative, hence x belongs to R. It follows that R contains A' since this is by definition the integral closure of A in K'. We may therefore set $\mathfrak{q} = \mathfrak{m} \cap A'$. This is a prime ideal of A' and the intersection of $\mathfrak{q}$ with A is the intersection of $\mathfrak{m}$ with A which is $\mathfrak{p}$. Thus $\mathfrak{q}$ divides $\mathfrak{p}$. Furthermore, R contains the discrete valuation ring $A'_{\mathfrak{q}}$. From Lemma 6.15, we have that $R = A'_{\mathfrak{q}} = v_{\mathfrak{q}}^{-1}(\mathbb{Z}_{\geq 0})$. Thus $v_{\mathfrak{q}}$ and v belong to the same place. Since both are normalized, we have $v = v_{\mathfrak{p}'}$ and it follows that $e = e_{\mathfrak{p}'}$ since by (i), $v_{\mathfrak{q}}$ extends $v_{\mathfrak{p}}$ with index $e_{\mathfrak{q}}$. □

Remark 6.18. As a consequence, the places in K' that lie above $[v_{\mathfrak{p}}]$ are exactly the places $[v_{\mathfrak{q}}]$ where $\mathfrak{q}$ is a prime ideal of A' that lies above $\mathfrak{p}$. In other words, the prime ideals $\mathfrak{q}$ of A' that lie above $\mathfrak{p}$ are in bijection with the places in K' that lie above $[v_{\mathfrak{p}}]$.

6.3 Galois extensions and inertia groups

Let A be a Dedekind domain with field of fractions K and K'/K a finite Galois extension of degree n. Throughout this section, we suppose that K is a field of characteristic zero since we are interested in number fields. Let G denote the Galois group of K' over K and A' denote the integral closure of A in K'.

Lemma 6.19. *Let R be a ring and consider a finite set of prime ideals $\mathfrak{p}_1, \ldots, \mathfrak{p}_\mathfrak{q}$. Let $\mathfrak{q}$ be an ideal of R such that none of the $\mathfrak{p}_\mathfrak{i}$'s contain $\mathfrak{q}$. Then there exists an element b of $\mathfrak{q}$ that does not belong to any of the $\mathfrak{p}_\mathfrak{i}$'s.*

Proof. We may of course assume that $\mathfrak{p}_\mathfrak{i}$ is not contained in $\mathfrak{p}_\mathfrak{j}$ when i is different from j. Let x_{ij} be an element of $\mathfrak{p}_\mathfrak{j} - \mathfrak{p}_\mathfrak{i}$ for i and j different. Since none of the $\mathfrak{p}_\mathfrak{i}$'s contain $\mathfrak{q}$, for each i we may pick an element a_i of $\mathfrak{q} - \mathfrak{p}_\mathfrak{i}$. Define $b_i = a_i \prod_{i \neq j} x_{ij}$. This is an element that belongs to $\mathfrak{q}$ since a_i belongs to $\mathfrak{q}$. The element b_i also belongs to $\mathfrak{p}_\mathfrak{j}$ for every j not equal to i since x_{ij} belongs to $\mathfrak{p}_\mathfrak{j}$ for every j not equal to i. But it does not belong to $\mathfrak{p}_\mathfrak{i}$ since neither a_i nor any of the x_{ij}'s belong to $\mathfrak{p}_\mathfrak{i}$. Now define $b = b_1 + \ldots + b_q$. It is an element of $\mathfrak{q}$ that belongs to none of the $\mathfrak{p}_\mathfrak{i}$'s. □

Lemma 6.20. *The ring A' is stable under the action of G. In other words, we have the equality*

$$\sigma(A') = A' \textit{ for all elements } \sigma \textit{ of } G.$$

Proof. Let σ be an element of G and let x be in A'. Then there exists a monic polynomial P with coefficients in A that vanishes at x. Since the coefficients are in A, we have the equality $P^\sigma = P$. Therefore, we get that $P(\sigma(x)) = \sigma(P(x)) = 0$. As a consequence, $\sigma(A')$ is contained in A' and this inclusion holfds for all σ in G. In particular, $\sigma^{-1}(A')$ is contained in A'. But then $A' = \sigma(\sigma^{-1}(A'))$ which is included in $\sigma(A')$. We conclude that, $\sigma(A') = A'$ for all elements σ of G. □

Let σ be an element of G and $\mathfrak{q}$ be a prime ideal of A' that lies above a fixed prime ideal $\mathfrak{p}$ of A. Then $\sigma(\mathfrak{q})$ is a prime ideal of A' that divides $\mathfrak{p}$. This is a consequence of the fact that σ is an automorphism that fixes K and thus also $\mathfrak{p}$. Moreover, $\mathfrak{p}'$ and $\sigma(\mathfrak{p}')$ must have the same ramification index. A consequence is that the Galois group G acts on the set of prime ideals $\mathfrak{q}$ of A' that lie above $\mathfrak{p}$.

Proposition 6.21. *Let $\mathfrak{p}$ be a prime ideal in A. The action of the Galois group G on the set of primes ideals $\mathfrak{q}$ that lie above $\mathfrak{p}$ is transitive.*

Proof. Let $\mathfrak{q}|\mathfrak{p}$ be a prime ideal of A' and suppose by contradiction that there is a prime ideal $\mathfrak{q}'|\mathfrak{p}$ that is not in the orbit of $\mathfrak{q}$. Then none of the $\sigma(\mathfrak{q})$'s contain $\mathfrak{q}'$. By Lemma 6.19, there is an element x of $\mathfrak{q}'$ that does not belong to any of the $\sigma(\mathfrak{q})$. Consider the norm of x associated to the extension K'/K : $N(x) = \prod_{\sigma \in G} \sigma(x)$ which is an element of A. By Lemma 6.20, each $\sigma(x)$ is in A' and therefore the norm of x belongs to $\mathfrak{q}'$ since x belongs to $\mathfrak{q}'$. Thus the norm belongs to the intersection $A \cap \mathfrak{q}'$ which is the ideal $\mathfrak{p}$. But x does not belong to $\sigma^{-1}(\mathfrak{q})$ hence $\sigma(x)$ does not belong to $\mathfrak{q}$. Since this is a prime ideal, we obtain that $N(x)$ is not an element of $\mathfrak{q}$. This contradicts the fact that $N(x)$ belongs to $\mathfrak{p}$ and $\mathfrak{q}$ lies above $\mathfrak{p}$. □

Corollary 6.22. *Let $\mathfrak{p}$ be a prime ideal in A and let $\mathfrak{q}$ be a prime ideal in A' that lies above $\mathfrak{p}$. Then $\mathfrak{p}$ decomposes in A' as follows:*

$$\mathfrak{p}A' = \left(\prod_{\sigma \in G} \sigma(\mathfrak{p}') \right)^e,$$

where e is the ramification index of $\mathfrak{q}$.

Proof. Immediate from Proposition-Definition 6.12 and Proposition 6.21. □

Definition 6.23. Let $\mathfrak{p}$ be a prime ideal in A and let $\mathfrak{q}$ be a prime ideal of A' that lies above $\mathfrak{p}$. We define the decomposition group $D_\mathfrak{q}$ associated to $\mathfrak{q}$ to be the subgroup of G consisting of elements that fix $\mathfrak{q}$. Explicitly,

$$D_\mathfrak{q} = \{\sigma \in G \mid \sigma(\mathfrak{q}) = \mathfrak{q}\}.$$

Proposition 6.24. *Let $\mathfrak{p}$ be a prime ideal in A and let $\mathfrak{q}$ be a prime ideal of A' that lies above $\mathfrak{p}$. Let e and f denote the ramification index and the residual degree of $\mathfrak{q}$. Then the cardinality of $D_\mathfrak{q}$ is equal to ef.*

Proof. Consider the map of sets from the quotient $G/D_\mathfrak{q}$ to the set $\mathcal{S} = \{\sigma(\mathfrak{q}) \mid \sigma \in G\}$ that sends $[\sigma]$ to $\sigma(\mathfrak{q})$. This is a bijection of sets. Let g be the size of $\mathcal{S}$. Then the cardinality of $G/D_\mathfrak{q}$ is g. By Corollary 6.22 and Proposition 6.14, $[K' : K] = gef$. But the size of G is exactly $[K' : K]$ so we conclude that the cardinality of $D_\mathfrak{q}$ is ef. □

For any element σ of the decomposition group $D_\mathfrak{q}$ we have $\sigma(\mathfrak{q}) = \mathfrak{q}$ and $\sigma(A') = A'$ by Lemma 6.20. We therefore have a well-defined map

$$\begin{array}{rccc} \bar{\sigma} : & A'/\mathfrak{q} & \longrightarrow & A'/\mathfrak{q} \\ & x + \mathfrak{q} & \longmapsto & \sigma(x) + \mathfrak{q}. \end{array}$$

One realizes that this defines an $A/\mathfrak{p}$-automorphism and therefore $\bar{\sigma}$ belongs to the Galois group of $A'/\mathfrak{q}$ over $A/\mathfrak{p}$. We therefore get a well-defined homomorphism of groups

$$\begin{array}{rccc} \Phi: & D_{\mathfrak{q}} & \longrightarrow & \mathrm{Gal}((A'/\mathfrak{q})|(A/\mathfrak{p})) \\ & \sigma & \longmapsto & \bar{\sigma}. \end{array} \tag{6.24.1}$$

Definition 6.25. Let $\mathfrak{p}$ be a prime ideal in A and let $\mathfrak{q}$ be a prime ideal of A' that lies above $\mathfrak{p}$. We define the inertia group $I_{\mathfrak{q}}$ associated to $\mathfrak{q}$ to be kernel of Φ (6.24.1). Explicitly,

$$I_{\mathfrak{q}} = \{\sigma \in D_{\mathfrak{q}} \mid \sigma(x) - x \in \mathfrak{q},\, \forall x \in A'\}$$

and this is a normal subgroup of the decomposition group.

Proposition 6.26. *Let $\mathfrak{p}$ be a prime ideal in A and let $\mathfrak{q}$ be a prime ideal of A' that lies above $\mathfrak{p}$. Suppose that $A/\mathfrak{p}$ is either finite or of characteristic zero. The field $A'/\mathfrak{q}$ is a Galois extension of $A/\mathfrak{p}$ of finite degree f and Φ (6.24.1) is surjective.*

Proof. To ease the notations we replace $D_{\mathfrak{q}}$ simply by D. Let K_D be the subfield of K' that is fixed by D. Then D is the Galois group of the extension K'/K_D. Let A_D be the integral closure of A in K_D and $\mathfrak{p}_{\mathfrak{D}}$ be the intersection of $\mathfrak{q}$ with A_D. Note that $\mathfrak{p}_{\mathfrak{D}}$ is a prime ideal of A_D and that $\mathfrak{q}$ divides it. Thus $\mathfrak{q}$ appears in the decomposition of $\mathfrak{p}_{\mathfrak{D}}A'$. By Proposition 6.21, the only other prime ideals of A' that appear in the decomposition of $\mathfrak{p}_{\mathfrak{D}}$ are conjugates of $\mathfrak{q}$ via D. By definition of the decomposition group, all these conjugates are $\mathfrak{q}$ itself. We conclude that $\mathfrak{p}_{\mathfrak{D}}A' = \mathfrak{q}^{e'}$, where e is the ramification index of $\mathfrak{q}$ at $\mathfrak{p}_{\mathfrak{D}}$. Let f' the residual degree of $\mathfrak{q}$ at $\mathfrak{p}_{\mathfrak{D}}$. Then $ef = e'f'$ since the degree of the extension K'/K_D equals the cardinality of D which we saw is ef. Since we have the inclusions

$$A/\mathfrak{p} \subset A_D/\mathfrak{p}_{\mathfrak{D}} \subset A'/\mathfrak{q},$$

we must have that f' is less than or equal to f. Since $\mathfrak{p}$ is contained in $\mathfrak{p}_{\mathfrak{D}}$ by definition of the latter, we see that $\mathfrak{p}A'$ is contained in $\mathfrak{p}_{\mathfrak{D}}A'$ which is equal to $\mathfrak{q}^{e'}$. Thus e' is less than or equal to e. The equality $ef = e'f'$ then forces the equalities $f = f'$ and $e = e'$. We conclude that

$$A_D/\mathfrak{p}_{\mathfrak{D}} = A/\mathfrak{p}. \tag{6.26.1}$$

We now come to the proof that $A'/\mathfrak{q}$ is a Galois extension of degree f of $A/\mathfrak{p}$. Since $A/\mathfrak{p}$ is either finite or of characteristic zero, the extension is automatically separable and we only need to check that the extension is normal. Let $\bar{x}$ be a primitive element of $A'/\mathfrak{q}$ over $A/\mathfrak{p}$. In other words, $A'/\mathfrak{q} = (A/\mathfrak{p})(\bar{x})$. Pick x an element of A' that reduced modulo

$\mathfrak{q}$ equals $\bar{x}$. Recall that A' is the integral closure of A in K. Thus there exists a monic polynomial with coefficients in A that vanishes at x. In particular, this is a polynomial with coefficients in A_D. Let $P(X) = X^r + a_{r-1}X^{r-1} + \ldots + a_0$ be the minimal polynomial of x over A_D. Its other roots are given by the conjugates of x via the Galois group D. Let $\tilde{P}$ be the reduced polynomial modulo $\mathfrak{p}_{\mathfrak{D}}$. The coefficients of this new polynomial are elements of $A_D/\mathfrak{p}_{\mathfrak{D}}$. But by (6.26.1), they are actually in $A/\mathfrak{p}$. If σ is an element of D, then $\Phi(\sigma) = \bar{\sigma}$ is an $A/\mathfrak{p}$-automorphism of $A'/\mathfrak{q}$ and thus fixes the coefficients of $\tilde{P}$. Hence the roots of $\tilde{P}$ are the elements $\bar{\sigma}(\bar{x})$ where σ belongs to D. Since all these roots are in $A'/\mathfrak{q}$, we deduce that $A'/\mathfrak{q}$ is the splitting field of $\tilde{P}$ and as a consequence it is a normal extension of $A/\mathfrak{p}$ and its degree is by definition f.

Since the minimal polynomial of $\bar{x}$ over $A/\mathfrak{p}$ must divide $\tilde{P}$, its roots are of the form $\bar{\sigma}(\bar{x})$. Let τ be any element of the Galois group of $A'/\mathfrak{q}$ over $A/\mathfrak{p}$. It is completely determined by the image of $\bar{x}$ and must send $\bar{x}$ to a root of its minimal polynomial. So there exists σ in D such that $\tau(\bar{x}) = \bar{\sigma}(\bar{x})$ and therefore $\tau = \bar{\sigma}$ and this proves the surjectivity. □

Corollary 6.27. *Let $\mathfrak{p}$ be a prime ideal in A and let $\mathfrak{q}$ be a prime ideal of A' that lies above $\mathfrak{p}$. Suppose that $A/\mathfrak{p}$ is either finite or of characteristic zero. Then the cardinality of the inertia group $I_{\mathfrak{q}}$ is e, the ramification index of $\mathfrak{q}$ at $\mathfrak{p}$. As a consequence, K' is unramified at $\mathfrak{p}$ if and only if for all prime ideal $\mathfrak{q}$ of A' that lies above $\mathfrak{p}$, the inertia group $I_{\mathfrak{q}}$ is trivial.*

Proof. As an immediate consequence of Proposition 6.26, the quotient group $D_{\mathfrak{q}}/I_{\mathfrak{q}}$ is isomorphic to the Galois group of $A'/\mathfrak{q}$ over $A/\mathfrak{p}$ which has cardinality f by definition. Thus the cardinality of $I_{\mathfrak{q}}$ is equal to e. □

Remark 6.28. Let K be a number field and K' a finite Galois extension of K. We can apply the results of this section in this case by taking $A = \mathcal{O}_K$ and $A' = \mathcal{O}_{K'}$. If $\mathfrak{p}$ is a prime ideal of $\mathcal{O}_K$, then the field $\mathcal{O}_K/\mathfrak{p}$ is finite since it is a finite extension of $\mathbb{Z}/(p) \cong \mathbb{F}_p$ where $(p) = \mathfrak{p} \cap \mathbb{Z}$. In particular, the results from Proposition 6.26 and Corollary 6.27 apply.

7 Appendix C - Completion

We present briefly completion of fields with respect absolute values and discuss the particular case of non-archimedean absolute values. We then focus on completion of number fields.

7.1 Completion of fields with respect to an absolute value

We briefly sketch the construction of the completion of a field with an absolute value and discuss the case of fields with discrete valuations. We conclude this section with a particular case of Hensel's Lemma.

Definition 7.1. Let K be a field. An absolute value on K is a function $|\,| : K \longrightarrow \mathbb{R}$ that satisfies the following properties: for all a, b in K,

(i) $|a| \geq 0$ and $|a| = 0$ if and only if $a = 0$.

(ii) $|ab| = |a||b|$.

(iii) $|a + b| \leq |a| + |b|$ (triangle inequality).

An absolute value that satisfies these properties is called archimedean. If in addition it satisfies the following property:

(iv) $|a + b| \leq \max\{|a|, |b|\}$ (strong triangle inequality),

then it is called a non-archimedean absolute value.

Definition 7.2. Let $(K, |\,|)$ be a field with an absolute. A completion of K with respect to $|\,|$ is a field $(\hat{K}, \|\,\|)$ with an absolute value that is complete with respect to this absolute value and there exists an embedding $\iota : K \longrightarrow \hat{K}$ that respects absolute values in the sense that $\|\iota(k)\| = |k|$ and such that $\iota(K)$ is dense in $\hat{K}$.

Proposition 7.3. *Let $(K, |\,|)$ be a field with an absolute value. There exists a unique completion of K with respect to $|\,|$.*

Sketch of proof. We start by proving uniqueness. Suppose $(\hat{K}, \|\cdot\|)$ and $(\hat{K}', \|\cdot\|')$ are two completions of $(K, |\cdot|)$. Let x be an element of $\hat{K}$. By density, there exists a sequence (x_n) in K such that $x = \lim_{n\to\infty}\{(\iota(x_n))\}$. Define $\lambda(x)$ to be $\lim_{n\to\infty}\{\iota'(x_n)\}$. One sees that $\|\iota'(x_m) - \iota'(x_n)\|' = \|\iota(x_m) - \iota(x_n)\|$ so that the sequence $\{\iota'(x_n)\}$ is a Cauchy sequence in

$\hat{K}'$ and hence converges in $\hat{K}'$ by completeness. One checks easily that $\lambda(x)$ does not depend on the chosen sequence x_n. We therefore have a well-defined morphism $\lambda : \hat{K} \longrightarrow \hat{K}'$. Repeating the argument, we construct another morphism $\lambda' : \hat{K}' \longrightarrow \hat{K}$ and we see that it is the inverse of λ and vice-versa.

We now sketch the proof of the existence part. Let K be a field and v and absolute value on K. Consider the set $\mathcal{C}$ of Cauchy sequences in K. This is a ring and the set $\mathcal{M}$ of null sequences, meaning the sequences that converge to zero, forms a maximal ideal. Thus the residue class ring $\mathcal{C}/\mathcal{M}$ is a field that we shall denote K_v. There is a natural embedding ι of K into K_v by sending and element x of K to the constant sequence $(x \mod \mathcal{M})$ in K_v. We define the norm of an element (x_n) of K_v by setting $\|(x_n)\|_v = \lim_{n\to\infty} |x_n|_v$ which makes sense since $\mathbb{R}$ is complete. One checks that this is well-defined and that $(K_v, \| \ \|_v)$ is a completion of K. □

Remark 7.4. Consider a field K with a discrete valuation v. Choosing any real number c in the interval $]0,1[$ gives rise to an absolute value on K by defining $|x|_v = c^{v(x)}$ for any x in K and one can check that this absolute value is non-archimedean. Note that different choices of c give rise to different absolute values which all yield the same topology on K, so the choice does not really matter. Conversely, one can show that all non-archimedean absolute values are constructed like this from a discrete valuation. We shall therefore indifferently talk about the completion of field with respect to a discrete valuation or with respect to a non-archimedean absolute value.

Let K be a field with a discrete valuation v. In this case, one can construct the completion of K with respect to v by first completing the discrete valuation ring A of v. There exists a theory of completion for abitrary rings, but we restrict ourselves to the case of discrete valuation rings since this is the only case we need and it requires less work. For general completion theory, see for example [AM] ch. 10, p. 100-105.

Definition 7.5. Let R be any ring and let I be a proper ideal of R. We define the completion of R with respect to I to be the ring

$$\hat{R}_I = \varprojlim R/I^n.$$

There is a canonical homomorphism of rings from R to $\hat{R}_I$ which sends and element r to the sequence $(r \mod I^n)_{n\geq 1}$. The kernel of this map is $\bigcap_{n\geq 1} I^n$.

Why is the case of a discrete valuation ring A simpler than the general case ? There are mainly two reasons. The first one is that we know the proper ideals of a discrete valuation ring. If π is a uniformizing element, then the proper ideals of A are the ideals (π^n) where n is a natural number (see proof of Proposition 6.5). The second reason is that discrete valuation rings are special cases of Noetherian domains and therefore Krull's Intersection Theorem ([AM] ch. 10 Corollary 10.18, p. 110) applies.

Proposition 7.6. *Let K be a field with a discrete normalized valuation v. Let A be the discrete valuation ring of v. The completion $\hat{A}$ of A with respect to its maximal ideal is also a discrete valuation ring. Denote by $\hat{v}$ the normalized valuation of $\hat{A}$. The ring A is embedded in $\hat{A}$ via the canonical homomorphism and $\hat{v}$ extends v with index* 1 *and A and $\hat{A}$ share the same uniformizing parameter. Moreover, $\hat{A}$ is complete with respect to the topology induced by $\hat{v}$. Let $\hat{K}$ be the fraction field of $\hat{A}$ and extend $\hat{v}$ to $\hat{K}$. Then $(\hat{K}, \hat{v})$ is the completion of (K, v).*

Sketch of proof. By Krull's Intersection Theorem ([AM] ch. 10 Corollary 10.18, p. 110)

$$\bigcap_{n\geq 1} \pi^n A = \{0\}. \tag{7.6.1}$$

This is the kernel of the canonical homomorphism from A to $\hat{A}$ and therefore this homomorphism is injective and A is embedded in $\hat{A}$.

Define a function $\hat{v} : \hat{A} \longrightarrow \mathbb{Z} \cup \{\infty\}$ as follows. Let x be an element of $\hat{A}$. Then x is an infinite sequence (x_n) of elements in A such that $x_m \equiv x_n \mod \pi^n$ for all n and m with $m \geq n$. Define

$$\hat{v}(x) = \begin{cases} \inf\{n \in \mathbb{N} \mid x_n \neq 0\} - 1, & \text{if } x \neq 0 \\ \infty, & \text{if } x = 0. \end{cases}$$

It is not to difficult to check that this does indeed define a discrete valuation. Note that (7.6.1) guarantees that $\hat{v}(x) = 0$ if and only if $x = 0$. All the remaining results can now be proved without to much effort. □

Proposition 7.7. *Let A be a discrete valuation ring with uniformizing parameter π. Let $\hat{A}$ be the completion of A with respect to its maximal ideal. Then for all n, we have an isomorphism of rings: $A/\pi^n A \cong \hat{A}/\pi^n\hat{A}$.*

Proof. In order to prove this, consider the natural projection maps $\rho_n : \hat{A} \longrightarrow A/\pi^n A$ that send an element (a_m) of $\hat{A}$ to its n^{th} component a_n. These projection maps are obviously surjective homomorphisms. We have the following exact sequence :

$$0 \longrightarrow \hat{A} \xrightarrow{.\pi^n} \hat{A} \xrightarrow{\rho_n} A/\pi^n A \longrightarrow 0.$$

The multiplication-by-π^n is clearly injective and the image of $\hat{A}$ via this map is $\pi^n\hat{A}$ which is clearly contained in the kernel of ρ_n. Let (a_m) be an element of this kernel. Then a_n is zero and by coherence of the sequence, a_m is zero for all $m \leq n$. Thus $\hat{v}((a_m))$ is greater or equal to n. Since π is a uniformizing element of $\hat{A}$, this implies that (a_m) belongs to $\pi^n\hat{A}$. Thus $\pi^n\hat{A} = \ker(\rho_n)$ and the sequence is exact. We conclude that $\hat{A}/\pi^n\hat{A} \cong A/\pi^n A$ for all n. □

Lemma 7.8. *Let A be a discrete valuation ring with uniformizing parameter π and let $\hat{A}$ be the completion of A with respect to its maximal ideal πA. Let u be a unit of $\hat{A}$ and let α be an element of $\hat{A}$ such that $\alpha = u + i$ for some element i in (π). Then α is a unit in $\hat{A}$.*

Proof. Consider the sequence in $\hat{A}$ defined by

$$v_n = \sum_{k=0}^{n} \left(-\frac{i}{u}\right)^k .$$

The difference $v_{n+1} - v_n$ is equal to $(-\frac{i}{u})^{n+1}$ which belongs to (π^{n+1}). So the sequence (v_n) defines an element of $\hat{A}$ which we call v. For each natural number n, v is equal to v_n modulo (π^n). Now,

$$\alpha\frac{v}{u} = (u+i)\frac{v}{u} \equiv v_n - \frac{i}{u}v_n \mod (\pi^n) \equiv 1 \mod (\pi^n)$$

and this holds for all n. Thus $\alpha\frac{v}{u} = 1$ in $\hat{A}$ and consequently α is invertible. □

We provide the proof of the well known lifting lemma by Hensel but restrict ourselves to the case of a discrete valuation ring.

Lemma 7.9 (Hensel). *Let A be a discrete valuation ring with uniformizing parameter π and let $\hat{A}$ be the completion of A with respect to its maximal ideal πA. Let F be a polynomial in one variable with coefficients in $\hat{A}$ and suppose that there exist an element a in $\hat{A}$ and a positive integer n such that $F(a)$ belongs to $\pi\hat{A}$ and $F'(a)$ belongs to $\hat{A}^*$. Then, for any element α in $\hat{A}$ that is equal to $F'(a)$ modulo (π), the sequence defined by*

$$w_0 = a, \qquad w_{m+1} = w_m - \frac{F(w_m)}{\alpha}$$

converges to a unique element b of $\hat{A}$ which satisfies

$$F(b) = 0 \quad \textit{and} \quad b \equiv a \mod (\pi^n).$$

Proof. We first proceed to a simplification: replacing the polynomial $F(X)$ by $F(X+a)/\alpha$, the assumptions of the lemma become

$$F(0) \in (\pi^n), \quad F'(0) \in \hat{A}^*, \quad F'(0) \equiv 1 \mod (\pi)$$

and the sequence is simply defined by $w_0 = 0$ and $w_{n+1} = w_n - F(w_n)$.

First, remark that if w_m belongs to (π^n), then it is also true for w_{m+1}. Indeed, by definition w_{n+1} is equal to $w_m - F(w_n)$ which in turn equals $-F(0)$ modulo (π^n). But $F(0)$ belongs to (π^n) by assumption which concludes the argument. Now, since the initial term of the sequence equals 0 which is certainly an element of (π^n), we may conclude that w_m belongs to (π^n) for any non-negative integer m.

Next, we prove that the sequence (w_m) is defines an element of $\hat{A}$. In other words, we prove by induction that the difference $w_{m+1} - w_m$ belongs to (π^{n+m}) for any non-negative integer m. If m equals 0, the same goes for w_0 and since w_1 is equal to $-F(0)$ which, by assumption, belongs to (π^n), there is nothing to prove. Now, let m be greater than 1 and suppose the relation is true for $m-1$. Using the definition of the sequence we obtain an equality between $w_{m+1} - w_m$ and $(w_m - F(w_m)) - (w_{m-1} - F(w_{m-1}))$. Rearranging the terms this is $(w_m - w_{m-1}) - (F(w_m) - F(w_{m-1}))$. We perform a trick by letting X and Y be two independent variables and decomposing $F(X) - F(Y)$ as follows :

$$F(X) - F(Y) = (X - Y)(F'(0) + XG(X,Y) + YH(X,Y))$$

where G and H are elements of $\hat{A}[X,Y]$. By linearity, we need only to check this decomposition on polynomials of the form $X^k - Y^k$ and for the latter it is clear. Using this we obtain an equality with

$$(w_m - w_{m-1})(1 - F'(0) - w_m G(w_m, w_{m-1}) - w_{m-1} H(w_m, w_{m-1})).$$

From our induction hypothesis we know that $w_m - w_{m-1}$ belongs to (π^{n+m-1}). Furthermore, by assumption $F'(0) - 1$ belongs to (π) and from our observation above, $w_j \in (\pi^n)$ for any j. As a consequence, $w_{m+1} - w_m$ belongs to (π^{n+m}).

We have proved that the sequence (w_m) defines an element of $\hat{A}$. We will call this element b. Since w_m belongs to (π^n) for all non-negative m, we must have that b is also an element of (π^n). Using the recursive definition of the sequence and taking the limit as m goes to infinity, we see that b equals $b - F(b)$ so that $F(b)$ is necessarily 0.

We now prove the uniqueness assertion. Let c be another element of (π^n) such that $F(c)$ is 0. Then

$$0 = F(b) - F(c) = (b - c)(F'(0) + bG(b,c) + c(Hb,c)).$$

Since $\hat{A}$ is a domain, we must have that $F'(0) + bG(b,c) + c(Hb,c)0$ equals 0. Whence the equality between $F'(0)$ and $-bG(b,c) - cH(b,c)$ which is an element of (π^n). But this contradicts the fact that $F'(0) - 1$ belongs to (π). This proves uniqueness. □

Remark 7.10. In the case where K is a number field, with the same notations as above, the valuation ring A of v is the localization $\mathcal{O}_{K,\mathfrak{p}}$ of the ring of algebraic integers $\mathcal{O}_K$ away from the prime ideal corresponding to v. We have a natural isomorphism between $\mathcal{O}_K/\mathfrak{p}\mathcal{O}_K$ and $\mathcal{O}_{K,\mathfrak{p}}/\mathfrak{p}\mathcal{O}_{K,\mathfrak{p}}$. In fact, both these quotient rings are fields. The first one because $\mathcal{O}_K$ is a Dedekind domain and the second one because $\mathcal{O}_{K,\mathfrak{p}}$ is a local ring with maximal ideal $\mathfrak{p}\mathcal{O}_{K,\mathfrak{p}}$. The isomorphism is given by

$$\begin{array}{ccc} \mathcal{O}_K/\mathfrak{p}\mathcal{O}_K & \longrightarrow & \mathcal{O}_{K,\mathfrak{p}}/\mathfrak{p}\mathcal{O}_{K,\mathfrak{p}} \\ a & \longrightarrow & a \\ ab^{-1} & \longleftarrow & a/b. \end{array}$$

Using Proposition 7.7 with $n = 1$, we get that

$$\hat{\mathcal{O}}_{K,\mathfrak{p}}/\mathfrak{p}\hat{\mathcal{O}}_{K,\mathfrak{p}} \cong \mathcal{O}_{K,\mathfrak{p}}/\mathfrak{p}\mathcal{O}_{K,\mathfrak{p}} \cong \mathcal{O}_K/\mathfrak{p}\mathcal{O}_K.$$

7.2 Completion of number fields

Let K denote a number field of degree n. By Ostrowski's Theorem ([AM] ch 1 § 4 prop. 4.3 p. 23, § 3 Theorem 3.4, 3.5 p. 15-17), all non-archimedean absolute values on $\mathbb{Q}$ arise from the p-adic absolute values. For any prime p, we let $|\cdot|_p$ denote the absolute value defined by

$$\begin{array}{ccc} \mathbb{Q} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & |x|_p = p^{-v_p(x)} \end{array}$$

where $v_p(x)$ denotes the p-adic valuation of x, that is the order of x at p. Moreover we let $|\cdot|_\infty$ denote the standard archimedean absolute value on $\mathbb{Q}$ which is defined via the inclusion of $\mathbb{Q}$ in $\mathbb{R}$.

Definition 7.11. The set $M_{\mathbb{Q}} = \{|\cdot|_p \ : \ p \text{ prime}\} \cup \{|\cdot|_\infty\}$ is called the set of standard absolute values on $\mathbb{Q}$.

Let $\mathfrak{p}$ be a prime ideal of the ring of algebraic integers $\mathcal{O}_K$ and let π be a uniformizing parameter for the localization $\mathcal{O}_{K,\mathfrak{p}}$ which by Corollary 6.9 is a discrete valuation ring. Let p be the prime such that $\mathfrak{p} \cap \mathbb{Z} = p\mathbb{Z}$. Then $p = \pi^{e(\mathfrak{p}/p)}u$ where u is a unit of $\mathcal{O}_{K,\mathfrak{p}}$ and $e(\mathfrak{p}/p)$ is the ramification index of $\mathfrak{p}$ at p. We define an absolute value that we shall denote $|\cdot|_\mathfrak{p}$ by

$$\begin{array}{ccc} K & \longrightarrow & \mathbb{R} \\ x & \longmapsto & |x|_\mathfrak{p} = (p^{-1/e(\mathfrak{p}/p)})^{v_\mathfrak{p}(x)} \end{array}$$

where $v_\mathfrak{p}(x)$ denotes the $\mathfrak{p}$-adic valuation of x, that is the order of x at π. As we have already discussed, all non-archimedean absolute values on K arise from prime ideals of $\mathcal{O}_K$.

Moreover, we let M_K^∞ denote the archimedean absolute values on K. These absolute values are obtained through embedding of K into $\mathbb{C}$ in the following way. Let $\sigma : K \longrightarrow \mathbb{C}$ be a homomorphism of fields that fixes $\mathbb{Q}$. Define the following function

$$\begin{array}{rccc} |\,|_\sigma : & K & \longrightarrow & \mathbb{R} \\ & x & \longmapsto & |\sigma(x)|_\infty \end{array}$$

where $|\,|_\infty$ denotes the standard absolute value on $\mathbb{C}$ ($|a+bi|_\infty = \sqrt{a^2+b^2}$). One can check that this defined an archimedean absolute value on K and that all archimedan absolute value on K is equivalent to an absolute value defined from such an embedding σ. If the image $\sigma(K)$ is contained in $\mathbb{R}$, we say that σ is a real embedding. Otherwise, we say that it is a complex embedding. Different real embeddings give rise to inequivalent absolute values on K, whereas complex embeddings always come in pairs through complex conjugations and two conjugated embeddings define the same absolute value. If r_1 and r_2 respectively are the number of real embeddings of K and the number of conjugate pairs of complex embeddings of K, then $n = r_1 + 2r_2$ and the size of M_K^∞ is $r_1 + r_2$.

Definition 7.12. The set $M_K = \{|\cdot|_\mathfrak{p} \ : \ \mathfrak{p} \text{ prime ideal of } \mathcal{O}_K\} \cup M_K^\infty$ is called the set of standard absolute values on K. These are the absolute values on K whose restriction to $\mathbb{Q}$ is one of the absolute values in $M_\mathbb{Q}$. We denote by M_K^0 the set of non-archimedean absolute values in M_K. We will indifferently talk about the absolute value $|\,|_\mathfrak{p}$ in M_K^0 or the valuation $v_\mathfrak{p}$ and we will write $v_\mathfrak{p} \in M_K^0$.

If L is a finite extension of K, then using a similar construction as above, one defines M_L to be the absolute values on L whose restriction to $\mathbb{Q}$ is one of the absolute values in $M_\mathbb{Q}$. This construction being consistent by multiplicativity of the ramification index, this is also those absolute values on L whose restriction to K is one of the absolute values in M_K.

Let v be in M_K^0 and let L be a finite extension of K. We would like to understand the relation between K_v and the different completions L_w where $w \in M_L$ extends v. In order to do this, we first need to know more about extensions of complete fields.

Proposition 7.13. *Let K be a field of characteristic zero with a discrete valuation v, let A be its valuation ring and suppose that K is complete with respect to v. Let L/K be a finite extension of K of degree n and let B denote the integral closure of A in B. Then B is a discrete valuation ring, it is a free A-module of rank n and L is complete with respect to the topology defined by B. In particular, v extends uniquely to a place $[w]$ of L.*

Proof. Since A is the valuation ring of v, it is a discrete valuation ring by Proposition 6.5. By Proposition 6.8, it is integrally closed and therefore A is a Dedekind domain. Thus B is also a Dedekind domain. Moreover we know that B is a finitely generated A-module. But A is principal and B is torsion free, so by the Structure Theorem for finitely generated modules over a principal ring ([Sam] ch. 1 § 5 Corollary 1 and 2 p.22), B is a free A-module of rank n.

Consider the prime ideals $\mathfrak{q}_i$ of B and the corresponding normalized valuations w_i. By Proposition 6.17, these valuations extend ord_v with index $e_{\mathfrak{q}_i}$ so in particular they extend v. Each of the w_i's defines an absolute value on L making L into a normed K-vector space of dimension n. But K is complete, so all norms on L are equivalent. In particular, $(L, |\,|_{w_i})$ is isomorphic to $(K^n, |\,|_{\sup})$ which is complete. Thus L is complete with respect to any norm on L. By Corollary 6.6, $[w_i]$ is uniquely determined by its valuation ring. But the valuation ring is entirely determined by the topology defined by $[w_i]$: it is the set of elements x of L for which the sequence (x^n) converges to zero when n goes to infinity. Since the norms defined by the w_i's are all equivalent, they define the same topology on L and therefore have the same valuation ring. Thus all the w_i's are equivalent and determine the same place $[w]$. So there is only one place $[w]$ of L that extends v. By Remark 6.18, the Dedekind domain B has a unique non-zero prime ideal and by Proposition 6.8 it is therefore a discrete valuation ring. □

Remark 7.14. With the same notations as in Proposition 7.13, by Proposition 6.14, we have $n = ef$ where e is the ramification index of B over A and f is the residual degree. There is no ambiguity in the notations when talking about ramification since both A and B are discrete valuation rings and therefore only have one prime ideal each.

Corollary 7.15. *Let K be a field of characteristic zero with a discrete valuation v and suppose that K is complete with respect to v. Let L/K be a finite extension of K of degree n and let $[w]$ be the unique place of L that extends v. Then two conjugate elements of L have the same normalized valuation.*

Proof. It suffices to prove this for a finite extension L' of L since, by Proposition 7.13, L is complete and the place $[w]$ extends uniquely to a place of L'. We may therefore suppose that L is a Galois extension of K. Let σ be an element of $\mathrm{Gal}(L|K)$. The composition $w \circ \sigma$ defines another discrete valuation on L which extends v since K is fixed by σ. By uniqueness, we must have that $[w \circ \sigma] = [w]$. In particular, $\mathrm{ord}_{w\circ\sigma}(x) = \mathrm{ord}_w(x)$ for all x in L. But $\mathrm{ord}_{w\circ\sigma} = \mathrm{ord}_w \circ \sigma$ so that $\mathrm{ord}_w(\sigma(x)) = \mathrm{ord}_w(x)$ for all x in L. This equality holds for all σ and since the conjugates of an element x are exactly the elements $\sigma(x)$ where σ is a K-homomorphism of L, we have proved the claim. □

Corollary 7.16. *Let K be a field of characteristic zero with a discrete valuation v and suppose that K is complete with respect to v. Let L/K be a finite extension of K of degree n and let $[w]$ be the unique place of L that extends v. We have $ord_w(x) = \frac{1}{f} ord_v(N_{L/K}(x))$ for all x in L. Here f is the residual degree of L/K.*

Proof. Let Σ be the set of K-homomorphisms of L. Using the previous corollary, we have

$$\begin{aligned}\mathrm{ord}_v(N_{L/K}(x)) &= \mathrm{ord}_v\left(\prod_{\sigma\in\Sigma}\sigma(x)\right) = \frac{1}{e}\mathrm{ord}_w\left(\prod_{\sigma\in\Sigma}\sigma(x)\right)\\ &= \frac{1}{e}\sum_{\sigma\in\Sigma}\mathrm{ord}_w(\sigma(x)) = \frac{1}{e}\sum_{\sigma\in\Sigma}\mathrm{ord}_w(x) = f\mathrm{ord}_w(x)\end{aligned}$$

since the size of Σ is n and $n = ef$ by Remark 7.14. □

Proposition 7.17. *Let L/K be an extension of number fields of degree n. Consider v a discrete valuation on K, let w_i be the different extensions of v to L and let e_i and f_i be respectively the ramification index and the residual degree of w_i at v (meaning of the corresponding prime ideals). Let K_v and L_{w_i} be the completions of K and L with respect to v and w_i and denote by $\hat{v}$ and $\hat{w}_i$ the normalized valuations on the completions that extend v and w_i.*

(i) We have : $e_i = e_{\hat{w}_i/\hat{v}}$ and $f_i = f_{\hat{w}_i/\hat{v}}$.

(ii) The field L_{w_i} is an extension of K_v of degree $n_i = e_i f_i$.

(iii) We have an isomorphism : $L \otimes_K K_v \cong \prod_i L_{w_i}$.

Proof. Fix i and to simplify notation we write w instead of w_i. We need to understand the ramification index and residual degree of the extension L_w/K_v. Let A and B be the valuation rings of v and w and let A_v and B_v be the respective valuation rings of K_v and L_w. They are all Dedekind domains since they are discrete valuation rings and B_w is the integral closure of A_v in L_w just as B is the integral closure of A in L. The equality of the ramification degrees is a consequence of the fact that A and A_v (resp. B and B_w) share the same uniformizing parameter.

Concerning the residual degrees, we have

$$f = [B/\mathfrak{q} : A/\mathfrak{p}] \quad \text{and} \quad \hat{f} = [B_w/\hat{\mathfrak{q}} : A_v/\hat{\mathfrak{p}}].$$

The equality between f and $\hat{f}$ is a direct consequence of Remark 7.10. We have proved (i).

Property (ii) is a consequence of (i). By Remark 7.14, the degree n_i of L_{w_i}/K_v is $e_{w_i/v}f_{w_i/v}$. By (i), this is exactly $e_i f_i$. Note that by Proposition 6.14, if n is the degree of L/K, then

$$n = \sum_i e_i f_i = \sum_i n_i.$$

We now prove (iii). By the Primitive Element Theorem, there exists α an element of L such that $L = K(\alpha) = K[X]/(\min(\alpha, K))$, where $\min(\alpha, K)$ denotes the minimal ploynomial of α over K. This is an irreducible polynomial of degree n. Let F be the polynomial $\min(\alpha, K)$ viewed as a polynomial with coefficients in K_v via the embedding of K into K_v. Decompose F into irreducible components F_i with i ranging between 1 and say d. Since the characteristic of K_v is zero, all roots of F are distinct and the tensor product then becomes a product of fields

$$L \otimes_K K_v \cong K_v[X]/(F) \cong K_v[X]/(F_1) \times \ldots \times K_v[X]/(F_d).$$

The fields in the product are exactly the different compositums of L and K_v. Let w be a valuation on L that extends v. We now prove that L_w is a compositum of L and K_v. Clearly, K_v is contained in L_w and the same is true for L via the embedding given by the completion. Thus the compositum LK_v is contained in L_w. The compositum is uniquely defined in this context since both fields are contained in the larger field L_w. The compositum LK_v is a finite degree extension of the complete field K_v and is therefore complete by Proposition 7.13, hence closed in L_w. But it contains L and we know that L is dense in L_w. Thus $LK_v = L_w$. Consequently, all the L_{w_i}'s appear in the above product of fields.

We now prove that the L_{w_i}'s are all distinct. By definition of the completion, $\hat{w}_i$ extends the valuation w_i. Suppose that $L_{w_i} \cong L_{w_j}$. Then $[\hat{w}_i|_L] = [w_i] = [w_j]$ and thus $i = j$ since $[w_i] = [w_j]$ if and only if $i = j$ by assumption. We conclude that the L_{w_i}'s are indeed distinct.

The tensor product $L \otimes_K K_v$ is a K_v-algebra of dimension n and by the remark in our proof of (ii) the product $\prod_i L_{w_i}$ is a K_v-algebra of dimension $\sum_i n_i = n$. We conclude that the above product of fields is exactly the product of the L_{w_i}'s, hence the result. □

Corollary 7.18. *With the same hypothesis as in the proposition and the same notation,* $N_{L/K}(x) = \prod_i N_{L_{w_i}/K_v}(x)$ *for all* x *in* L.

Proof. Let x be an element of L and denote by F its characteristic polynomial over K. Consider F as a polynomial with coefficients in $L \otimes_K K_v$. This the characteristic polynomial of x viewed in this tensor product. By the above isomorphism, F factorizes as a product of polynomials F_i who are the characteristic polynomials of x in L_{w_i}/K_v. In particular, since we can read off the norm of x on the constant coefficient of its characteristic polynomial, the result follows. □

Notation If K is a number field, we shall denote by n_v the degree $[K_v : \mathbb{Q}_{v|_{\mathbb{Q}}}]$ where $v \in M_K$.

Corollary 7.19 (Extension formula). *Let L/K be an extension of number fields of finite degree and v a valuation in M_K. Then*

$$\sum_{\substack{w \in M_L \\ w|v}} n_w = [L : K] n_v.$$

Proof. Suppose first that v belongs to M_K^0. By Formula 6.14 and Proposition 7.17 (ii), we have the following formula :

$$n = \sum_{w|v} e_{\mathfrak{b}/\mathfrak{p}} f_{\mathfrak{b}/\mathfrak{p}} = \sum_{w|v} [L_w : K_v].$$

Multiplying both sides of this equality by $[K_v : \mathbb{Q}_p]$ we obtain the desired result.

Now, let v be an archimedean absolute value on K. We distinguish two cases. Suppose that v comes from a real embedding σ of K into $\mathbb{C}$. If r_1 and r_2 respectively are the number of real embeddings of L into $\mathbb{C}$ and the number of conjugate pairs of complex embeddings of L whose restriction to K is σ, then $n = r_1 + 2r_2$. Thus,

$$\sum_{w \in M_L} n_w = \sum_{w \in M_L^\infty} n_w = r_1 + 2r_2 = n = [L : K] n_v$$

since $n_w = 2$ if w results from a pair of complex embeddings of K and $n_w = 1$ if w results from a real embedding. Suppose that v comes from a pair of embeddings σ and $\bar{\sigma}$. If r_1 and r_2 (resp. $\bar{r}_1$ and $\bar{r}_2$) respectively are the number of real embeddings of L into $\mathbb{C}$ and the number of conjugate pairs of complex embeddings of L whose restriction to K is σ (resp. $\bar{\sigma}$), then $n = r_1 + 2r_2 = \bar{r}_1 + 2\bar{r}_2$. Then

$$\sum_{w \in M_L} n_w = \sum_{w \in M_L^\infty} n_w = r_1 + 2r_2 + \bar{r}_1 + 2\bar{r}_2 = 2n = [L : K] n_v.$$

□

Proposition 7.20 (Product Formula). *Let K be a number field. For every non-zero element x of K, we have the following formula:*

$$\prod_{v \in M_K} |x|_v^{n_v} = 1.$$

Proof. We first prove it in the case $K = \mathbb{Q}$. The general case will follow. Recall that $M_{\mathbb{Q}} = \{|\ |_p \mid p \text{ prime}\} \cup \{|\ |_\infty\}$. Let x be a non-zero element of $\mathbb{Q}$. We decompose it into its prime factorization:

$$x = \pm \prod_p p^{v_p(x)}.$$

Then $|x|_\infty = \prod_p p^{v_p(x)}$ and $|x|_p = p^{-v_p(x)}$. Thus $\prod_{v \in M_{\mathbb{Q}}} |x|_v = |x|_\infty \prod_p |x|_p = 1$.

Let K be any number field. By Remark 7.14, for $v \in M_K^0$, we have $n_v = [K_v : \mathbb{Q}_v] = e_v f_v$ where e_v and f_v are the ramification and residual indexes of $K_v/\mathbb{Q}_v$. Let $\hat{v}$ be the discrete valuation on K_v. By Corollary 7.16, $\hat{v}(x) = \hat{v}(N_{K_v/\mathbb{Q}_v}(x))/f_v$. Remembering our convention for the absolute value defined by v,

$$\prod_{\substack{v \in M_K \\ v|p}} |x|_v^{n_v} = \prod_{\substack{v \in M_K \\ v|p}} p^{-n_v v(x)/e_v} = \prod_{\substack{v \in M_K \\ v|p}} p^{-\hat{v}_p(N_{K_v/\mathbb{Q}_p}(x))} = p^{-\hat{v}_p(\prod N_{K_v/\mathbb{Q}_p}(x))} = |N_{K/\mathbb{Q}(x)}|_p$$

where the last equality is the application of Corollary 7.18.

Now consider $v \in M_K^\infty$. Either there is $s : K \longrightarrow \mathbb{R}$ an embedding such that $|x|_v = |s(x)|_\infty$ or there are two conjugate embeddings $s, \bar{s} : K \longrightarrow \mathbb{C}$ such that $|x|_v = |s(x)|_\infty = |\bar{s}(x)|_\infty$. In the first case, $n_v = 1$ whereas in the second case, $n_v = 2$. Thus

$$\begin{aligned} |N_{K/\mathbb{Q}}(x)|_\infty = \left| \prod_s s(x) \right|_\infty &= \prod_{s \text{ real}} |s(x)|_\infty \prod_{s \text{ complex}} |s(x)|_\infty |\bar{s}(x)|_\infty \\ &= \prod_{s \text{ real}} |s(x)|_\infty \prod_{s \text{ complex}} |s(x)|_\infty^2 = \prod_{v \in M_K^\infty} |x|_v^{n_v}. \end{aligned}$$

Since x is non-zero, its norm is also non-zero. By the case $K = \mathbb{Q}$, we have

$$1 = \prod_{v \in M_{\mathbb{Q}}} |N_{K/\mathbb{Q}}(x)|_v = \prod_{v \in M_{\mathbb{Q}}^0} \prod_{\substack{w \in M_K \\ w|v}} |x|_w^{n_w} \prod_{w \in M_K^\infty} |x|_w^{n_w} = \prod_{w \in M_K} |x|_w^{n_w}.$$

□

8 Appendix D - Formal Groups

Definition 8.1 (Formal group)**.** Let A be a ring. A formal group defined over A is a formal power series in two variables $F \in A[[X,Y]]$ that satisfies the following conditions :

(i) $F(X,Y) = X + Y +$(higher order terms) ;

(ii) $F(X, F(Y,Z)) = F(F(X,Y),Z)$;

(iii) $F(X,Y) = F(Y,X)$;

(iv) $\exists \iota(T) \in A[[T]]$ such that $F(\iota(T), T) = 0$;

(v) $F(X,0) = X$ and $F(0,Y) = Y$.

Notation For notational ease (to be understood later), we shall often denote the formal group by $(\mathcal{F}, F)$.

Remark 8.2. In other words, a formal group over A is an operation that imitates the law of an abelian group but that has no underlying elements : we have associativity (ii), commutativity (iii), existence of an inverse (iv) and a neutral element (v). So these conditions seem natural to impose if one wants to create an abstract notion of group. Only condition (i) remains a mystery and seems in some way non-intuitive.

Definition 8.3 (Morphism of formal groups)**.** A morphism of formal groups $f : (\mathcal{F}, F) \longrightarrow (\mathcal{G}, G)$ where both are defined over a ring A is a formal power series $f \in A[[T]]$ such that $f(F(X,Y)) = G(f(X), f(Y))$. The two formal groups are said to be isomorphic if there exists a morphism $g : (\mathcal{G}, G) \longrightarrow (\mathcal{F}, F)$ such that $f(g(T)) = g(f(T)) = T$.

Definition 8.4. Let m be an integer and $(\mathcal{F}, F)$ be a formal group over a ring. We define an element $[m]$ in $A[[T]]$ inductively by :

$$[0](T) = 0\,; \quad [m+1](T) = F([m](T), T)\,; \quad [m-1](T) = F([m](T), \iota(T)).$$

Proposition 8.5. *For any integer m, the power series $[m]$ is a morphism of formal groups from $(\mathcal{F}, F)$ to itself.*

Proof. Case $m = 0 :$ $[0](F(X,Y)) = 0$ and $F([0](X),[0](Y)) = F(0,0) = 0$.
Case $m > 0 :$ By induction, suppose that $[m]$ is a morphism. We prove that $[m+1]$ is also a morphism. Using the definition of the multiplication-by-m map yields an equality between $[m+1](F(X,Y))$ and

$$F([m](F(X,Y)), F(X,Y)).$$

Applying the induction hypothesis, this is in turn equal to

$$F(F([m](X),[m](Y)), F(X,Y)).$$

Using the associativity and then the commutativity of F, this is

$$F([m](X), F(F(X,Y),[m](Y))).$$

Another use of associativity yields $F([m](X), F(X, F(Y,[m](Y))))$. Commutativity and the definition of the map give $F([m](X), F(X,[m+1](Y)))$ which is equal to

$$F(F([m](X),X),[m+1](Y))$$

by associativity. Finally, by definition of the map we obtain $F([m+1](X),[m+1](Y))$ which ends the proof.
Case $m < 0 :$ Also by induction and analogous to the previous case. □

Proposition 8.6. *The multiplication-by-m map is of the form*

$$[m](T) = mT + (\textit{higher order terms}).$$

From now on we shall write h.o.t. for higher order terms.

Proof. Case $m = 0 :$ Nothing to prove.
Case $m > 0 :$ Suppose $[m]$ has the required form. We prove that it is also the case for $m+1$. Observe that by definition of the map, $[m+1](T) = F([m](T),T)$. By the induction hypothesis, this is $F(mT + (h.o.t.), T)$. Using property (i) of Definition 8.1, we see that it is equal to

$$(mT + (h.o.t.)) + T + (h.o.t.)$$

which is $(m+1)T + (h.o.t.)$.
Case $m < 0 :$ Also by induction and analogous to the previous case. □

Lemma 8.7. *Let A be a ring and f be an element of $A[[T]]$ of the form*

$$f(T) = aT + (h.o.t.)$$

where a is a unit of A. Then there exists a unique element g of $A[[T]]$ such that $f(g(T)) = T$. We also have $g(f(T)) = T$.

Proof. We construct a sequence of polynomials $(g_n) \in A[T]$ enjoying the following properties:

(i) $g_{n+1}(T) \equiv g_n(T) \mod T^{n+1}$

(ii) $f(g_n(T)) \equiv T \mod T^{n+1}$.

This is a quite natural idea, since (i) will ensure that the sequence (g_n) belongs to the inverse limit $\varprojlim A[[T]]/(T^m)$ which is isomorphic to $A[[T]]$ since the latter is complete. Thus the limit $g(T)$ of the sequence is a well-defined element of $A[[T]]$. The condition (ii) ensures that $f(g(T))$ equals $f(g_n(T))$ modulo T^{n+1} which is equal to T modulo T^{n+1} for any positive integer n and therefore we obtain equality between $f(g(T))$ and T.

Proceeding to the construction we define $g_1(T)$ to be $a^{-1}T$. By the assumption made on f, we see that

$$f(g_1(T)) = a(a^{-1}T) + (h.o.t.) \equiv T \mod T^2$$

so g_1 satisfies (ii).

Now, suppose that we have constructed g_{n-1} satisfying the above conditions. In view of the first condition, we must find some element λ of A so that

$$g_n(T) = g_{n-1}(T) + \lambda T^n.$$

By our assumption on g_{n-1}, there is some b belonging to A such that $f(g_{n-1}(T)) = T + bT^n$. Now, we must determine λ so that $f(g_n(T))$ equals T modulo T^{n+1}. We compute

$$\begin{aligned} f(g_n(T)) &= f(g_{n-1}(T) + \lambda(T)T^n) \\ &= a(g_{n-1}(T) + \lambda(T)T^n) + \alpha(g_{n-1}(T) + \lambda(T)T^n)^2 + (h.o.t.) \end{aligned}$$

for some α that belongs to A. By construction, the degree of g_{n-1} is $n-1$, so by examining degrees when opening the square, we see that

$$\alpha(g_{n-1}(T) + \lambda(T)T^n)^2 \equiv \alpha(g_{n-1}(T))^2 \mod T^{n+1}.$$

Therefore,

$$\begin{aligned} f(g_n(T)) &\equiv ag_{n-1}(T) + \alpha(g_{n-1}(T))^2 + a\lambda T^n \mod T^{n+1} \\ &\equiv f(g_{n-1}(T)) + a\lambda T^n \mod T^{n+1} \\ &\equiv T + (b + a\lambda)T^n \mod T^{n+1}. \end{aligned}$$

Choosing λ to be $-a^{-1}b$, we get the desired result.

Since $g(T) = a^{-1}T + (h.o.t.)$, applying the same reasoning as above we may find an element h of $A[[T]]$ such that $g(h(T))$ equals T. Then

$$g(f(T)) = g(f(g(h(T)))) = g \circ (f \circ g)(h(T)) = g(h(T)) = T.$$

To prove uniqueness, let $\hat{g}$ be in $A[[T]]$ and suppose that $f(\hat{g}(T))$ equals T. Then

$$g(T) = g(f(\hat{g}(T))) = \hat{g}(T)$$

and this completes the proof. □

Groups associated to formal groups Let A be a complete local ring, $\mathcal{M}$ its maximal ideal and $k = A/\mathcal{M}$ the residue field. Let $(\mathcal{F}, F)$ be a formal group over A. As already noted, $(\mathcal{F}, F)$ is a "group operation" with no underlying group. In general, there is no obvious way to assign values to a formal power series. But evaluating F in elements x and y of $\mathcal{M}$ yields an element $F(x, y)$ which belongs to A. To see this, define polynomials F_n in $A[X, Y]$ by cutting off terms of order greater or equal to $n+1$. Then each $F_n(x, y)$ belongs to $\mathcal{M}$ and it follows immediately from the definition of this sequence that it is coherent. It is therefore an element of the inverse limit $\varprojlim A/\mathcal{M}^m$, so it converges in A by completeness. It is clear that the limit is $F(x, y)$ and as a consequence $F(x, y)$ belongs to A. This makes $\mathcal{M}$ into a group under the operation F. This is made precise in the following definition.

Definition 8.8. The group associated to $(\mathcal{F}, F)$, denoted by $\mathcal{F}(\mathcal{M})$, is the set $\mathcal{M}$ with the group law defined by :

$$x \oplus_{\mathcal{F}} y = F(x, y) \qquad \ominus_{\mathcal{F}} x = \iota(x)$$

for elements x and y in $\mathcal{M}$.

Proposition 8.9. *Let m be a positive integer that is relatively prime to char(k)= p (we may have $p = 0$). Then the group $\mathcal{F}(\mathcal{M})$ has no non-trivial points of order m.*

Proof. We know already that the multiplication-by-m map

$$[m] : \mathcal{F}(\mathcal{M}) \longrightarrow \mathcal{F}(\mathcal{M})$$

is a homomorphism of abelian groups. We have also seen already that $[m]$ is of the form

$$[m](T) = mT + (h.o.t.).$$

The fact that m and p are relatively prime implies that p does not divide m. So $\tilde{m}$ is not zero in the residue field k. Thus m does not belong to $\mathcal{M}$. But A is a local ring so this implies that m is a unit in A. Lemma 8.7 then implies that $[m]$ is an isomorphism. So, if x is an element of $\mathcal{F}(\mathcal{M})$ such that $[m](x)$ is zero, then x must also be zero. □

References

[AM] M.F. Atiyah, I.G. Macdonald, *Introduction to Commutative Algebra.* Addison-Wesley Publishing Company, 1969.

[Coh] P.M. Cohn, *Algebraic Numbers and Algebraic Functions.* Springer, 1991.

[Fal] G. Faltings, *Endlichkeitssätze für abelsche Varietäten über Zahlkörpern.* Inventiones Mathematicae 73, 1983.

[Ful] W. Fulton, *Algebraic curves: an introduction to Algebraic Geometry.* http://www.math.lsa.umich.edu/ wfulton/CurveBook.pdf.

[GZ] B.H. Gross, D.B. Zagier, *Heegner points and derivatives of L-series.* Inventiones Mathematicae 84, 1986.

[Har] R. Hartshorne, *Algebraic Geometry.* Graduate Texts in Mathematics, Vol. 52, 2006.

[HH] D. Hilbert, A. Hurwitz, *Über die diophantischen Gleichungen vom Geschlecht Null.* Acta Math. 14, 1891.

[Hil] D. Hilbert, *The Theory of Algebraic Number Fields.* Springer, 1998.

[Lan] S. Lang, *Algebraic Number Theory.* Graduate Texts in Mathematics, Vol. 110, 1986.

[Maz] B. Mazur, *Modular curves and the Eisenstein ideal.* Inst. Hautes Études Sci. Publ. Math. 47, 1977.

[Mer] L. Merel, *Bornes sur la torsion des courbes elliptiques sur les corps des nombres.* Inventiones Mathematicae 124, 1996.

[Mor] L. J. Mordell, *On the rational solutions of the indeterminate equations of the third and fourth degrees.* Proc. Cambridge Phil. Soc. 21, 1922.

[Poi] H. Poincaré, *Sur les propriétés arithmétiques des courbes algébriques.* J. de Liouville 7, 1901.

[Sam] P. Samuel, *Algebraic Theory of Numbers.* Dover, 2008.

[Se1] J. P. Serre, *Lectures on the Mordell-Weil Theorem.* Aspects of Mathematics, 3rd edition, 1997.

[Se2] J. P. Serre, *Corps locaux.* Actualités Scientifiques et Industrielles, Hermann, 1968.

[Sha] I. R. Shafarevich, *Basic Algebraic Geometry.* Springer, 1977.

[Sil] J. H. Silverman, *The Arithmetic of Elliptic Curves.* Graduate Texts in Mathematics, Vol. 106, 2nd edition, 2009.