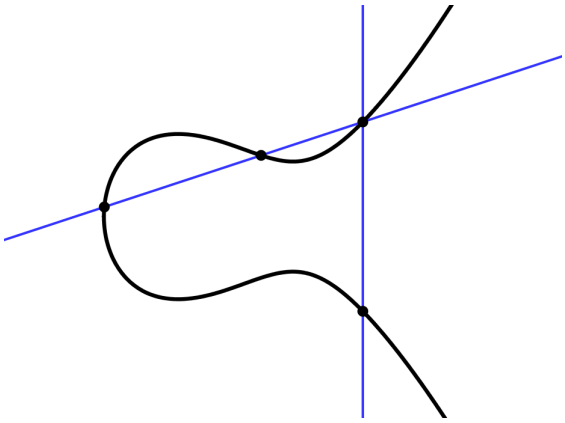
The Mordell-Weil theorem and the Birch and Swinnerton-Dyer conjecture

Gilles Felber

Semester paper

supervised by Prof. Dr. E. Kowalski

July 2, 2019



D-Math, ETH Zürich

Contents

1 Introduction

The theory of Diophantine equations is the study of rational or integer solutions of polynomial equations. As its name indicates, its history goes back to the ancient Greeks, especially Diophantus of Alexandria. The simplest case of a Diophantine equation is

$$aX + bY = c$$

with $a, b, c \in \mathbb{Z}$ and a, b not both 0. Using Euclid algorithm, we easily (and explicitly) find if the equation has solutions and, in the later case, describe the infinite amount of solutions. Indeed, the Euclid algorithm tells us that there exist $r, s \in \mathbb{Z}$ such that

$$ar + bs = (a, b),$$

where (a, b) is the biggest common divisor of a and b. If $(a, b) \mid c$, then we can take X and Y as multiple of r and s to get $aX + bY = c$. All the other solutions are given by the changes of variables

$$X \mapsto X + k\frac{b}{(a, b)},\ Y \mapsto Y - k\frac{a}{(a, b)}$$

for $k \in \mathbb{Z}$. If $(a, b) \nmid c$, there is clearly no solution since $(a, b) \mid a$ and $(a, b) \mid b$, so (a, b) divides the left of the equation, hence it must divide c.

The second simplest case is the case of a quadratic equation

$$aX^2 + bXY + cY^2 + dX + eY + f = 0,$$

$a, b, c, d, e, f \in \mathbb{Z}$. This case can be completely solved using the famous Hasse Principle.

Theorem 1.1 (Hasse-Minkowski, [2], xvii)**.** *Let $f \in \mathbb{Q}[X, Y]$ a quadratic polynomial. Then $f(X, Y) = 0$ has solution in $\mathbb{Q}^2$ if and only if $f(X, Y) = 0$ has a solution over all completions of $\mathbb{Q}$, i.e. in $\mathbb{R}^2$ and in $\mathbb{Q}_p^2$.*

Using the law of quadratic reciprocity, we can test if $f(X, Y) = 0$ has a solution modulo p for all p prime. From Hensel's lemma, we then deduce the existence or not of a solution in $\mathbb{Q}_p$ (we must also solve the equation in $\mathbb{R}$) and deduce, using Hasse principle, if the equation has a solution in $\mathbb{Q}$.

After having done the degree one and the degree two cases, one would like to study integer solutions of equations of degree 3. Unfortunately, the Hasse principle is known to fail in this case. Ernst S. Selmer provided in 1951 the classical example

$$3x^3 + 4y^3 + 5z^3 = 0,$$

which has a solution in every completion of $\mathbb{Q}$ but none in the rational numbers (see [1]). Nevertheless, the study of Diophantine equations of degree 3 leads to a incredibly rich theory named the arithmetic of elliptic curves. This paper is an introduction and overview of some topics in this theory. After a general but not thorough introduction to the theory of elliptic curves, we will give a mostly self-contained proof of the Mordell-Weil theorem, following the treatment of Silverman [2]. We will finish by stating the famous Birch and Swinnerton-Dyer conjecture, one of the greatest open problem in mathematics nowadays.

2 Elliptic curves

There are two ways to define elliptic curves and their group law. One using purely the language of algebraic geometry and the other using more concrete equations and geometric meaning. We will prove that theses definitions give the same object in the end and for this, the famous Riemann-Roch theorem will be of central use.

Definition 2.1. An **elliptic curve** is a pair (E, O), where E is a smooth projective algebraic curve of genus one and $O \in E$. We generally omit O and denote the elliptic curve only by E.

Definition 2.2. A **Weierstrass equation** is a homogenenous cubic equation of the form

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

with $a_1, a_2, a_3, a_4, a_6 \in \bar{K}$, an algebraic closure of a fixed field K. We fix $O = [0 : 1 : 0] \in \mathbb{P}^2(K)$ and denote by $E \subseteq \mathbb{P}^2(K)$ the vanishing locus of the equation over K. If $char(K) \neq 2, 3$, we can rewrite the equation using affine variable changes to get

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

with $y = \frac{Y}{Z}$ and $x = \frac{X}{Z}$. We say that E is **non-singular** or **smooth** if the **discriminant**,

$$\Delta = -16(4A^3 + 27B^2) = -b_2^2b_8 - 8b_4^3 - 27b_6^2 + 9b_2b_4b_6$$

is non-zero, where

$$b_2 = a_1^2 + 4a_4,\ b_4 = 2a_4 + a_1a_3,\ b_6 = a_3^2 + 4a_6,\ b_8 = a_1^2a_6 + 4a_2a_6 - a_1a_3a_5 + a_2a_3^2 - a_4^2.$$

If $\Delta = 0$, then E has a unique singular point (see [2], chapter III, 1.4 (a)).

Proposition 2.3. *If E is a curve given by a singular Weierstrass equation, then it is birational to $\mathbb{P}^1$.*

Proof. Without loss of generality, we can suppose that E has a unique singular point at $(x, y) = (0, 0)$. By definition of smoothness, both partial derivative of the Weierstrass equation must be 0. This implies that the equation has the form

$$y^2 + a_1xy = x^3 + a_2x^2.$$

Thus, the map

$$E \to \mathbb{P}^1,\ (x, y) \mapsto [x, y]$$

has degree one because its inverse can be explicitly given by (thinking as $t = \frac{y}{x}$)

$$\mathbb{P}^1 \to E,\ [t, 1] \mapsto (t^2 + a_1t - a_2, t^3 + a_1t^2 - a_2t).$$

□

Theorem 2.4 (Riemann-Roch)**.** *Let C a smooth curve, K_C a canonical divisor on C. There exists an integer $g \geq 0$, the **genus** of C, such that for any divisor $D \in Div(C)$,*

$$\ell(D) - \ell(K_C - D) = \deg(D) - g + 1.$$

Proof. See [2], chapter II, theorem 5.4. □

Corollary 2.5. *(a)* $\ell(K_C) = g$.

(b) $\deg(K_C) = 2g - 2$.

(c) If $\deg(D) > 2g - 2$, *then* $\ell(D) = \deg(D) - g + 1$.

Proof. 1. This is immediate using the Riemann-Roch formula for $D = 0$.

2. This follow taking $D = K_C$ and using (a).

3. From (b), we have that $\deg(K_C - D) < 0$. But, for any $f \in \bar{K}(C)^*$, $\deg \operatorname{div}(f) = 0$, so there is no $f \neq 0$ such that $\operatorname{div}(f) \geq -D$. So $\ell(K_C - D) = 0$.

□

Lemma 2.6. *Let* C, D *curves,* $\phi : C \to D$ *a rational map. Then*

(a) If C *is smooth, then* ϕ *is a morphism.*

(b) If ϕ *is a morphism, then* ϕ *is either surjective or constant.*

(c) If C *and* D *are smooth and* ϕ *is of degree one, then* ϕ *is a isomorphism.*

Proof. See [2], chapter II, 2.1, 2.3 and 2.4.1 respectively. □

Proposition 2.7. *1. Let* E/K *an elliptic curve. Then* E *is given by a Weierstrass equation. More precisely, there exist functions* $x, y \in K(E)$, $a_1, \ldots, a_6 \in K$ *and a morphism*

$$\phi : E \to \mathbb{P}^2, \ P \mapsto [x(P), y(P), 1]$$

satisfying $\phi(O) = [0, 1, 0]$ *and that is an isomorphism between* E *and its image, which is a curve given by a Weierstrass equation.*

2. Every smooth curve given by a Weierstrass equation over a field K *is an elliptic curve over* K *with base point* $[0, 1, 0]$.

Proof. 1. We look at $\mathcal{L}(n(O))$ for $n \geq 1$. Using corollary 2.5 (c) of Riemann-Roch, we get

$$\dim(\mathcal{L}(n(O)) = \ell(n(O)) = n.$$

So, we can find functions $x, y \in K(E)$ such that $\{1, x\}$ is a basis of $\mathcal{L}(2(O))$ and such that $\{1, x, y\}$ is a basis of $\mathcal{L}(3(O))$. Moreover, by definition, x must have a pole of order exactly 2 at O and similarly y must have a pole of order 3 at O. Now, if we look for $n = 6$, we see that $\mathcal{L}(6(O))$ has dimension 6 but contains the following 7 elements:

$$1, x, x^2, x^3, y, y^2, xy.$$

Therefore, there exists a non-trivial linear relation

$$A_1 + A_2 x + A_3 x^2 + A_4 x^3 + A_5 y + A_6 y^2 + A_7 xy = 0$$

with $A_1, \ldots A_7 \in K$. Note that $A_4 \neq 0$ and $A_6 \neq 0$ because all the other terms have a pole at O of different order and so would vanish otherwise. Replacing x by $-A_4 A_6 x$ and y by $A_4^2 A_6$

and then dividing by $A_4^4 A_6^3$ gives us an Weierstrass equation in the desired form. Looking at the map

$$\phi : E \to \mathbb{P}^2,\ P \mapsto [x(P), y(P), 1],$$

we see that its image C satisfies the Weierstrass equation. It is a surjective morphism by lemma 2.6 (a) and (b). Furthermore, $\phi(O) = [0, 1, 0]$ since y has an pole of higher order than x.
Now, we prove that ϕ has degree one, or equivalently, that $K(E) = K(x, y)$. First, the map $[x, 1] : E \to \mathbb{P}^1$ has a double pole at O and no other pole so it must be of degree 2. Similarly, $[y, 1] : E \to \mathbb{P}^1$ has degree 3. Hence, the degree of ϕ divides 2 and 3, so it can only be one. So if C is smooth, then ϕ is an isomorphism by lemma 2.6 (c).
It remains to show that C is smooth. If C is singular, there exists a rational map $\psi : C \to \mathbb{P}^1$ of degree one by 2.3. So $\psi \circ \phi$ is a degree-one map between smooth curve, therefore an isomorphism, which contradicts the fact that E has genus 1. Hence, C is smooth and ϕ is an isomorphism.

2. Let E be given by a non-singular equation. The differential

$$\omega = \frac{dx}{2y + a_1 x + a_3} \in \Omega_E$$

has no zero or pole (see [2], III.1.5), so its associated divisor must be 0. Applying corollary 2.5 (b) of Riemann-Roch theorem on it, we get for the genus g of E

$$2g - 2 = \deg \operatorname{div}(\omega) = 0,$$

so $g = 1$.

□

Using these two definitions, we can define two addition laws on elliptic curves.

Given 2 points on the curve, the Riemann-Roch theorem tells us that there exist an unique function $f \in K(E)$ up to constant and a unique point on E such that f has poles exactly at the two first points and zeros at O and this last point. This defines a commutative group law on E which is induced by the one on the degree-zero part of the Picard group as we will see in more detail.

Given two points $P, Q \in E$, there is a unique third point (not necessary distinct from the two first) intersecting the line passing through the first points by Bézout's theorem. Similary, we take the line through this point and O to get a fourth point that we denote by $P + Q$. This gives a commutative group law on E.

Lemma 2.8. *Let C a curve of genus 1 and $P, Q \in C$. Then*

$$(P) \sim (Q) \Leftrightarrow P = Q.$$

Proof. Let $f \in \bar{K}(C)$ such that $(P) - (Q) = \operatorname{div}(f)$ so $f \in \mathcal{L}((Q))$. Riemann-Roch theorem 2.5 (c) tells us that $\ell((Q)) = 1$. But $\mathcal{L}((Q))$ must contain the constant functions, so $f \in \bar{K}$ and $P = Q$. □

Using this lemma, we get a map from E to $Pic^0(E)$.

Proposition 2.9. *Let E an elliptic curve. The map*

$$\kappa : E \to Pic^0(E),\ P \mapsto [(P) - (O)]$$

is an isomorphism of groups. Its inverse is given by

$$\sigma : Pic^0(E) \to E,\ [D] \mapsto P,$$

where P is the unique point in E such that $D \sim (P) - (O)$.

Proof. First, the map σ is well-defined: since E has genus one, the Riemann-Roch theorem 2.5 (c) tells us that $\ell(D + (O)) = 1$, so it has a basis given by a single non-zero element $f \in \bar{K}(E)$. Since $\mathrm{div}(f) \geq -D - (O)$ and $\deg \mathrm{div}(f) = 0$, it follows that

$$\mathrm{div}(f) = -D - (O) + (P)$$

for some $P \in E$. Hence $D \sim (P) - (O)$. If $P' \in E$ is such that $D \sim (P') - (O)$, then $(P) \sim (P')$, and by lemma 2.8, $P = P'$. Furthermore, if D and D' are two divisors in $Div^0(E)$, let $P, P' \in E$ be the points given by $D \sim (P) - (O)$ and $D' \sim (P') - (O)$. Then $(P) - (P') \sim (D) - (D')$. So $P = P'$ if and only if $D \sim D'$ by lemma 2.8. Hence, σ is well-defined.

Now, clearly κ and σ are inverse of each other. Hence, we just need to prove that κ is a morphism, i.e. that

$$\kappa(P + Q) = \kappa(P) + \kappa(Q).$$

Let

$$f(X, Y, Z) = aX + bY + cZ = 0$$

be the line in $\mathbb{P}^2$ through P, Q and a third point $R \in E$ (not necessary distinct) and

$$g(X, Y, Z) = 0$$

the line through R, O and $P + Q$. Then, since $Z = 0$ intersect O with multiplicity 3, we have

$$\mathrm{div}(f/Z) = (P) + (Q) + (R) - 3(O),$$
$$\mathrm{div}(g/Z) = (R) + (P + Q) - 2(O).$$

So we get

$$0 \sim \mathrm{div}(g/f) = (P) + (Q) - (P + Q) - (O),$$

which is the same as

$$\kappa(P + Q) - \kappa(P) - \kappa(Q) = 0.$$

□

3 The Mordell-Weil Theorem

The goal of this chapter is to prove the following theorem.

Theorem 3.1 (Mordell-Weil Theorem)**.** *Let E/K be an elliptic curve over a number field K. Then $E(K)$ is finitely generated.*

To prove this, we will first prove the so called weak Mordell-Weil theorem and then use height functions to prove the general case. We fix the following notations : K is a number field, E/K is an elliptic curve, M_K is the set of valuations v of K, with M_K^∞ the archimedean ones and M_K^0 the normalized nonarchimidean ones.

For a Galois field extension L/K, we denote $G_{L/K}$ the corresponding Galois group. We denote $|\cdot|_v$ the absolute value associated to v, K_v is the completion of K at v. if $v \in M_K^0$, we denote R_v the corresponding ring of integer, k_v the residue field and $\tilde{E}(k_v)$ the reduced curve.

3.1 The weak Mordell-Weil Theorem

Theorem 3.2 (Weak Mordell-Weil Theorem)**.** *Let $m \geq 2$ an integer. Then $E(K)/mE(K)$ is finite.*

Lemma 3.3. *Let L/K a finite Galois extension. If $E(L)/mE(L)$ is finite, then so is $E(K)/mE(K)$.*

Proof. We consider the inclusion $E(K) \to E(L)$. It induces a morphism $E(K)/mE(K) \to E(L)/mE(L)$ with kernel $N := \frac{mE(L)\cap E(K)}{mE(K)}$. We only need to prove that this kernel is finite. For each $P \in N$, we fix a $Q_P \in E(L)$ such that $[m]Q_P = P \pmod{mE(K)}$. Let

$$\lambda_P : G_{L/K} \to E[m],\ \sigma \mapsto Q_P^\sigma - Q_P.$$

It is well defined since $G_{L/K}$ leave $mE(K)$ invariant so $[m](Q_P^\sigma - Q_P) = P^\sigma - P = O$. Moreover, if there exists P' such that $\lambda_P = \lambda_{P'}$, then

$$(Q_P - Q_{P'})^\sigma = Q_P - Q_{P'},$$

for all $\sigma \in G_{L/K}$, hence $Q_P - Q_{P'} \in E(K)$ and $P - P' \in mE(K)$. Therefore λ_P gives an injection $\lambda : N \to \mathrm{Map}(G_{L/K}, E[m])$. Since both sets on the right hand side are finite, so must be N. □

Using the lemma, we immediately deduce that we can reduce to the case where

$$E[m] \subseteq E(K),$$

by adding the corresponding elements to the field. To prove this case, we introduce an analogue to the theory of Kummer extensions.

Definition 3.4. The **Kummer pairing** is

$$\kappa : E(K) \times G_{\bar{K}/K} \to E[m],$$

$$(P, \sigma) \mapsto Q^\sigma - Q,$$

where $Q \in E(\bar{K})$ is such that $[m]Q = P$.

Proposition 3.5. *1. The Kummer pairing is well-defined.*

2. The Kummer pairing is bilinear.

3. The kernel on the left is $mE(K)$.

4. The kernel on the right is $G_{\bar{K}/L}$ where $L := K([m]^{-1}E(K))$ is the composition of all extensions $K(Q)$ for $Q \in E(\bar{K})$ with $mQ \in E(K)$.

Hence the Kummer Pairing gives a perfect bilinear pairing

$$E(K)/mE(K) \times G_{L/K} \to E[m].$$

Proof. 1. As before, since $[m]Q = P \in E(K)$, we have

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

Moreover, if $[m]Q' = P$, then there exists $R \in E[m]$ such that $Q' = Q + R$. Then

$$(Q')^\sigma - Q' = Q^\sigma + R^\sigma - (Q + R) = Q^\sigma - Q,$$

because $E[m] \subseteq E(K)$ and σ leaves $E(K)$ invariant.

2. Linearity on the left is clear. On the right, let $\sigma, \tau \in G_{\bar{K}/K}$, $P \in E(K)$ and $Q \in E(\bar{K})$ such that $[m]Q = P$. Then

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

since $Q^\sigma - Q \in E(K)$.

3. If $P \in E(K)$ is such that $Q^\sigma = Q$ for all $\sigma \in G_{\bar{K}/K}$, then $Q \in E(K)$, so $P \in mE(K)$.

4. If $\sigma \in G_{\bar{K}/L}$, then by definition $Q^\sigma = Q$ so $\kappa(P, \sigma) = O$ for all $P \in E(K)$. Reciprocally, if $\kappa(P, \sigma) = O$ for all $P \in E(K)$, then σ fixes L by definition.

For the final statement, we just have to verify that L is Galois, which is true since all elements of $G_{\bar{K}/K}$ map L to itself. □

Remark. Hence, to prove the finiteness of $E(K)/mE(K)$, it suffices to prove that L/K is a finite extension. Indeed, the Kummer pairing gives an injection

$$E(K)/mE(K) \to \mathrm{Hom}(G_{L/K}, E[m]),\ P \mapsto \kappa(P, \cdot).$$

Definition 3.6. Let $v \in M_K^0$ a discrete valuation. A **minimal Weierstrass** equation for E over K_v is a Weierstrass equation such that the valuation of all coefficients is non-negative and $v(\Delta)$ is minimal. $E(K)$ has **good reduction** at v if the reduced curve $\tilde{E}$ over k_v is non-singular. Otherwise, $E(K)$ has **bad reduction** at v.

Proposition 3.7. *Let $v \in M_K^0$ such that $v(m) = 0$ and such that E has good reduction at v. Then the reduction map*

$$E(K)[m] \to \tilde{E}_v(k_v)$$

is injective.

Proof. See [2], chapter VII, proposition 3.1 (b). □

Proposition 3.8. 1. *The extension L/K is abelian of exponent m, i.e. all elements have order dividing m.*

2. *The valuations $v \in M_K^0$ such that E has good reduction at v and $v(m) \neq 0$ are unramified. In particular, there are only finitely many ramified valuations.*

Proof. 1. It is clear since the Kummer pairing gives us an injection into an abelian group

$$G_{L/K} \to \mathrm{Hom}(E(K), E[m]).$$

Moreover, this group is of exponent m, by definition of $E[m]$.

2. Let v as in the proposition. We only need to look at $K' = K(Q)$ for $Q \in L$. Let w a valuation above v. We prove that the ramification group $I_{w/v} \subseteq G_{\bar{K}/K}$ is trivial. Let $\sigma \in I_{w/v}$, so $\sigma = id \pmod{R_w}$. Hence

$$Q^\sigma - Q = O \pmod{R_w}.$$

Moreover, $[m]Q \in E(K)$ implies that $[m](Q^\sigma - Q) = O$. By proposition 3.7, we get that $Q^\sigma - Q = O$ so Q is fixed by all elements of the inertia group $I_{w/v}$ hence it must be trivial and K' is unramified.

□

To conclude the proof of the weak Mordell-Weil theorem, we show that any extension satisfying the properties of proposition 3.8 is finite. For this, we will use the following theorem.

Theorem 3.9 (Main Theorem of Kummer Theory)**.** *Let K a field of characteristic 0 containing all the m-th roots of unity. The maximal abelian extension of exponent m of K is given by adding the m-roots of all the elements of K^*, i.e. it is the extension*

$$K(\sqrt[m]{a} : a \in K^*).$$

Proof. See [2], chapter VIII, proof of proposition 1.6. □

Lemma 3.10. *Let R a Dedekind domain with finite class number, $K = Frac(R)$, $I_1 = R, \dots, I_r$ representatives of the ideal class group. If S is a finite subset of the valuations of R containing all infinite places and all prime ideals dividing $I_1, \dots, I_r$, then*

$$R_S := \{a \in K | v(a) \geq 0 \ \forall v \notin S\}$$

is a principal ideal domain. Moreover, its prime ideals are exactly the one corresponding to valuations not in S.

Proof. We denote $I(R)$ the set of ideals of R and $I(R_S)$ the ideals of R_S. We also denote $I_v \subseteq R$ the ideal corresponding to the valuation v. We consider the function

$$\phi : I(R) \to I(R_S), \ I \mapsto IR_S.$$

It is a surjective function. To prove this, let

$$\psi : I(R_S) \to I(R), \ J \mapsto J \cap R_S.$$

If we have $\phi \circ \psi = id$, then ϕ must be surjective, because it send $\psi(J)$ to J for any $J \in I(R_S)$. Clearly,

$$\phi \circ \psi(J) = (J \cap R)R_S \subseteq JR_S = J.$$

Now, let $\frac{a}{b} \in J$, with $a, b \in R$, so $v(a) - v(b) \geq 0$ for all $v \notin S$. We consider the two ideals

$$A = \prod_{v \notin S} I_v^{v(b)},$$

$$B = \prod_{w \in S} I_w^{w(b)}.$$

They are obviously coprime so $A + B = R$ and there exist $c \in A$, $d \in B$ such that

$$c + d = 1$$

and, by definition of A and B, $v(c) \geq v(b)$ for all $v \notin S$ and $w(d) \geq w(b)$ for all $w \in S$. We rewrite

$$\frac{a}{b} = (c + d)\frac{a}{b} = a\frac{c}{b} + \frac{da}{b}.$$

For all $v \notin S$, we have $v(c) - v(b) \geq 0$, so $\frac{c}{b} \in R_S$ and clearly, $a \in I \cap R$. Also, $w(d) - w(b) \geq 0$ for all $w \in S$, so $\frac{da}{b} \in R$. Since J is an ideal, $\frac{da}{b} \in J \cap R$. We conclude that

$$\frac{a}{b} = a\frac{c}{d} + \frac{da}{b} \in (J \cap R)R_S,$$

so $J = (J \cap R)R_S$.

So ϕ is a surjective map. Now consider a maximal ideal $I_v \subseteq R$ corresponding to a valuation $v \notin S$, i.e. $I_v = \{a \in R | v(a) \geq 1\}$. Since $v \notin S$, we have the inclusion

$$\phi(I_v) = I_v R_S \subseteq \{a \in R_S | v(a) \geq 1\},$$

so $\phi(I_v) \neq R_S$. Conversely, if $I_w \subseteq R$ correspond to a valuation $w \in S$, then clearly $I_w^{-1} \subseteq R_S$ so

$$\phi(I_w) = I_w R_S \supseteq I_w I_w^{-1} = R.$$

Hence, $\phi(I_w) = R_S$ since it is an ideal of R_S containing 1.

We can now deduce that R_S is principal. Let $J \subseteq R_S$ an ideal. $I := J \cap R$ is in the same class as a representative K_i, so there exists $a \in K^*$ such that $I = aI_i$. If $i = 1$, $I_i = R$ so $J = IR_S = aRR_S = aR_S$ is principal. If $i \neq 1$, then

$$I_i = \prod_{v \in S} I_v^{v(I)}$$

and so

$$I_i R_S = \prod_{v \in S} (I_v R_S)^{v(I)} = \prod_{v \in S} R_S^{v(I)} = R_S.$$

Hence, we also have $J = IR_S = aR_S$. Hence, in all cases, J is principal.

Finally, if $J \subseteq R_S$ is a prime ideal, then clearly $J \cap R$ must be prime so it correspond to a valuation v (that can not be in S). Therefore $J = (I \cap R)R_S \subseteq \{a \in R_S | v(a) \geq 1\}$ but this last set has the same restriction to R so they are equal. Conversely, if $v \notin S$, then $I_v R_S \subseteq \{a \in R_S | v(a) \geq 1\}$ and this last set is clearly a prime ideal with restriction equal to I_v, hence they are equal.

□

Proposition 3.11. *Let K be a number field, $S \subseteq M_K$ be a finite set of valuations containing M_K^∞ and L/K the maximal abelian extension of K with exponent m that is unramified outside of S. Then L/K is finite.*

Proof. Note first that without loss of generality, we can suppose that all the m-th roots of unit are in K since if we take a finite extension K'/K, then the fact that LK'/K' is finite implies that L/K is so. Moreover, since the class number is finite, we can suppose that the ring of S-integers

$$R_S := \{a \in K | v(a) \geq 0 \ \forall v \notin S\}$$

is a principal ideal domain by adding to S the valuations corresponding to all prime ideal dividing a set of representatives of the ideal class group by lemma 3.10. We can do this since adding valuations to S will only grow the extension L. So we can also suppose that $v(m) = 0$ for all $v \notin S$.

The main theorem of Kummer theory tells us that L is the largest subfield of $K(\sqrt[m]{a} : a \in K)$ which is unramified outside S. Let $v \notin S$ and $a \in K$ and consider the equation $X^m - a = 0$ over the local field K_v. Since $v(m) = 0$ and the discriminant of the polynomial is $\pm m^m a^{m-1}$, $K_v(\sqrt[m]{a})/K_v$ is unramified if and only if $\mathrm{ord}_v(a) = 0 \pmod m$. Note that we only need to take elements in a set of representatives of $K^*/(K^*)^m$. Hence, we only need to prove that

$$T_S := \{a \in K^*/(K^*)^m : \mathrm{ord}_v(a) = 0 \pmod m \ \forall v \notin S\}$$

is finite. We are going to find a surjective map $R_S^* \to T_S$. Since R_S^* is finite generated, expliciting the kernel will conclude the proof.

Consider $R_S^* \to K^*/(K^*)^m$. If $a \in R_S^*$, then $v(a) = 0$ for all $v \notin S$ by definition. Hence, it goes in T_S. To prove surjectivity, let $a \in K^*$ representing an element of T_S. aR_S is the m-th power of an ideal of R_S, so there exists $b \in K^*$ such that $aR_S = b^m R_S$, so $a = ub^m$ for some $u \in R_S^*$. Hence, $a = u \pmod{(K^*)^m}$. Finally, the kernel of this map clearly contains $(R_S^*)^m$ so we are done. □

3.2 The Descent Theorem

To prove the general Mordell-Weil theorem, we will use the theory of height functions over abelian groups.

Theorem 3.12 (Descent Theorem)**.** *Let A an abelian group, m an integer and $h : A \to \mathbb{R}$ a* ***height function*** *satisfying the following properties:*

1. *For all $Q \in A$, there exists a constant C_1, depending on Q, such that*
$$h(P + Q) \leq 2h(P) + C_1$$
for all $P \in A$.

2. *There exists a constant C_2 such that*
$$h(mP) \geq m^2 h(P) - C_2$$
for all $P \in A$.

3. *For any constant C, the set*
$$\{P \in A | h(P) \leq C\}$$
is finite.

If A/mA is finite, then A is finitely generated.

Proof. Let $Q_1, \ldots, Q_r \in A$ representatives of A/mA and P a fixed element of A. Then there exist $1 \leq i_1 \leq r$ and $P_1 \in A$ such that $P = mP_1 + Q_{i_1}$. Doing the same reasoning for P_1 and iterating, we get two sequences $P_0 = P, P_1, P_2, \ldots$ and $i_1, i_2, i_3, \ldots$ such that
$$P_n = mP_{n+1} + Q_{i_{n+1}}$$
for any integer n. Note that
$$m^2 h(P_{n+1}) \leq h(P_n - Q_{i_{n+1}}) + C_2 \leq 2h(P_n) + C_1' + C_2$$
where C_1' is the maximum of the constant C_1 given by the first property for $Q \in \{-Q_1, \ldots, -Q_r\}$. By iterating, we conclude that
$$h(P_n) \leq \left(\frac{2}{m^2}\right)^n h(P) + (C_1' + C_2)\left(\frac{1}{m^2} + \frac{2}{m^4} + \frac{2^2}{m^6} + \cdots + \frac{2^{n-1}}{m^{2n}}\right) \leq \frac{1}{m^n} h(P) + C_1 + C_2,$$
since $m \geq 2$. So if we fix a constant $C > C_1' + C_2$, we have that $h(P_n) \leq C$ for n big enough. Since the set $B := \{P \in A | h(P) \leq C\}$ is finite,
$$B \cup \{Q_1, Q_2, \ldots, Q_r\}$$
is a finite generating set, hence A is finitely generated. □

3.3 Heights on projective space

To prove the Mordell-Weil theorem we need to define a height function on $E(K)$. For this, we begin by defining it on a projective space.

Definition 3.13. Let $P \in \mathbb{P}^N(K)$ with homogeneous coordinates $P = [x_0, \dots, x_N]$, $x_0, \dots, x_N \in K$. The **height** of P (relative to K) is

$$H_K(P) := \prod_{v \in M_{\mathbb{K}}} \max\{|x_0|_v, \dots, |x_r|_v\}^{n_v},$$

where $n_v := [K_v : \mathbb{Q}_v]$ is the **local degree** at v.

Proposition 3.14. *Let $P \in \mathbb{P}^N(K)$.*

1. *H_K is well defined;*
2. *$H_K(P) \geq 1$;*
3. *For a finite extension L/K, $H_L(P)^{[L:K]} = H_K(P)$.*

To prove these, we recall two basic facts of algebraic number theory.

Proposition 3.15 (Extension formula)**.** *Let $L/K/\mathbb{Q}$ a tower of number fields, $v \in M_K$. Then*

$$\sum_{M_L \ni w | v} n_w = [L : K] n_v.$$

Proposition 3.16 (Product formula)**.** *Let $x \in K^*$. Then*

$$\prod_{v \in M_K} |x|_v^{n_v} = 1.$$

Proof. See, for example, [3], chapter II, corollary 8.4 and chapter III, proposition 1.3. □

Proof of 3.14.

1. Using product formula, we immediately get that if $[y_0, \dots, y_N] = \lambda[x_0, \dots, x_N]$ for some $\lambda \in K^*$, then $H_K([y_0, \dots, y_N]) = H_K([x_0, \dots, x_N])$.
2. Using the first point, we can assume without loss of generality that at least one of the coordinate is one. Then $H_K(P) \geq 1$ since $|1|_v = 1$ for all $v \in M_K$.
3. Using the extension formula, we get

$$H_L(P) = \prod_{w \in M_L} \max\{|x_0|_w, \dots, |x_r|_w\}^{n_w} = \prod_{v \in M_L} \prod_{M_L \ni w | v} \max\{|x_0|_w, \dots, |x_r|_w\}^{n_w}$$

$$= \prod_{v \in M_L} \max\{|x_0|_v, \dots, |x_r|_v\}^{\sum_{M_L \ni w | v} n_w} = H_K(P)^{[L:K]}.$$

□

Definition 3.17. Let $P \in \mathbb{P}^N(\bar{\mathbb{Q}})$. The **absolute height function** of P is

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

For any number field K such that $P \in \mathbb{P}^N(K)$ is well-defined.

Example 3.18. The set

$$S := \{P \in \mathbb{P}^N(\mathbb{Q}) | H(P) \leq C\}$$

is finite:

Notice first that we can suppose, by cleaning denominators, that $P = [x_0, \ldots, x_N]$ with all $x_i \in \mathbb{Z}$ and such that they have no common divisors. Then for all non-archimedean valuations $v \in M_K^0$, at least one of the coefficient is one and none can be bigger than one, so we get the inequality:

$$C \geq H(P) = H_{\mathbb{Q}}(P) = \prod_{v \in M_K} \max\{|x_0|_v, \ldots, |x_N|_v\} = \max\{|x_0|, \ldots, |x_N|\},$$

where $|\cdot|$ is the usual absolute value. So our set S is included into

$$\{(x_0, \ldots, x_N) \in \mathbb{Z}^N | \max\{|x_0|, \ldots, |x_N|\} \leq C\},$$

which is clearly finite.

Theorem 3.19. *Let $P \in \mathbb{P}^N(\bar{\mathbb{Q}})$ and $\sigma \in G_{\bar{\mathbb{Q}}/\mathbb{Q}}$. Then*

$$H(P^\sigma) = H(P).$$

Proof. Let $P = [x_0, \ldots, x_N]$ and $K = \mathbb{Q}(P)$. σ gives an isomorphism $K \xrightarrow{\sim} K^\sigma$ which identifies the absolute values, so we also have $n_{v^\sigma} = n_v$ and $[K : \mathbb{Q}] = [K^\sigma : \mathbb{Q}]$. We compute:

$$H_{K^\sigma}(P^\sigma) = \prod_{v^\sigma \in M_{K^\sigma}} \max_{j=0,\ldots,n} \{|x_j^\sigma|_{v^\sigma}\}^{n_{v^\sigma}} = \prod_{v \in M_K} \max_{j=0,\ldots,n} \{|x_j|_v\}^{n_v} = H_K(P).$$

□

Notation. We fix the following notations: for any $P = [x_0, \ldots, x_N] \in \mathbb{P}^N(\bar{\mathbb{Q}})$ and $v \in M_K$, we denote:

$$|P|_v := \max\{|x_0|_v, \ldots, |x_N|_v\};$$
$$|F(P)|_v := \max\{|f_0(P)|_v, \ldots, |f_M(P)|_v\};$$
$$|F|_v := \max\{|a|_v : a \text{ coefficient of some } f_j\}.$$

We have the identities

$$H_K(P) = \prod_{v \in M_K} |P|_v^{n_v};$$
$$H_K(F(P)) = \prod_{v \in M_K} |F(P)|_v^{n_v}.$$

We also define

$$H_K(F) := \prod_{v \in M_K} |F|_v^{n_v} = H_K([a : a \text{ coefficient of some } f_j]),$$

$$\epsilon(v) := \begin{cases} 1 & \text{if } v \text{ is archimedean,} \\ 0 & \text{else.} \end{cases}$$

Hence, for all $v \in M_K$, we have the identity

$$|x_1 + \cdots + x_n|_v \leq n^{\epsilon(v)} \max\{|x_1|_v, \ldots, |x_n|_v\}.$$

Theorem 3.20. *Let $F : \mathbb{P}^N \to \mathbb{P}^M$ a morphism of degree d, i.e. $F = [f_0, \dots, f_M]$ with $f_0, \dots, f_M \in \bar{\mathbb{Q}}[X_0, \dots, X_N]$ homogeneous polynomials of degree d without other common zero than $[0, \dots, 0]$. Then there exist constants $C_1, C_2 > 0$, depending on F, such that for all $P \in \mathbb{P}^N(\bar{\mathbb{Q}})$,*

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

Proof. Using the notations above, we see that

$$|f_j(P)|_v \leq C_1^{\epsilon(v)} |F|_v |P|_v^d,$$

For a constant C_1 only depending on F, N and M. Since $\sum_{v \in M_K} \epsilon(v) n_v = [K : \mathbb{Q}]$, by taking product over all $v \in M_K$, we get the upper bound

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

Now, using that $f_1, \dots, f_M$ have no common non-trivial zero, we know that the radical of the ideal generated by these polynomials in $\bar{\mathbb{Q}}[X_0, \dots, X_N]$ is $(X_0, \dots, X_N)$ by the Nullstellensatz. Therefore it must contains some power of each X_j, i.e. there exists a positive integer e such that we have

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

Without loss of generality, we can suppose that $g_{ij} \in K[X_0, \dots, X_N]$ by replacing K by a finite extension, and also that all terms on the right are homogeneous of degree e, hence all g_{ij} are of degree $d - e$. We denote

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

$$H_K(G) = \prod_{v \in M_K} |G|_v.$$

With these notations, we get, for some constant C_2, C_3,

$$|x_i|_v^e \leq C_2^{\epsilon(v)} |G(P)|_v |F(P)|_v \leq (C_2 C_3)^{\epsilon(v)} |G|_v |F(P)|_v |P|_v^{e-d}.$$

Hence,

$$|P|_v^d \leq (C_2 C_3)^{\epsilon(v)} |G|_v |F(P)|_v,$$

and we conclude

$$H(P)^d \leq C_2 C_3 H(G) H(F(P)).$$

□

Notation. For $x \in \bar{\mathbb{Q}}$, we write $H(x) := H([x, 1])$ and if $x \in K$, we also write $H_K(x) := H_K([x, 1])$.

Theorem 3.21. *Let*

$$f(T) = a_d T^d + a_{d-1} T^{d-1} + \cdots + a_1 T + a_0 = a_d (T - \alpha_1)(T - \alpha_2) \dots (T - \alpha_d) \in \bar{\mathbb{Q}}[T]$$

a polynomial of degree d. Then

$$2^{-d} \prod_{j=1}^{d} H(\alpha_j) \leq H([a_0, \dots, a_n]) \leq 2^{d-1} \prod_{j=1}^{n} H(\alpha_j).$$

Proof. Note first that we can suppose that f is monic, since multiplying it by a constant does not change the inequalities. Let $K := \mathbb{Q}(\alpha_1, \dots, \alpha_d)$. For a valuation $v \in M_K$, we denote

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

The triangle inequality reads now

$$|x + y|_v \leq \epsilon(v) \max\{|x|_v, |y|_v\}$$

for all $v \in M_K$, $x, y \in K$. We will prove that

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

Rising to the power n_v, taking product, and taking root will then give the result. We will do the proof by induction on d. For $d = 1$, $f(T) = T - \alpha_1$ and the result is clear. If $d > 1$, we choose an index k such that $|\alpha_k|_v \geq |\alpha_j|_v$ for all $j = 0, \dots, d$. We define

$$g(T) = (T - \alpha_1) \dots (T - \alpha_{k-1})(T - \alpha_{k+1})(T - \alpha_d) = T^{d-1} + b_{d-2}T^{d-2} + \dots + b_1 T + b_0,$$

so that $f(T) = (T - \alpha_k)g(T)$, hence each coefficient is given by $a_j = b_{j-1} - \alpha_k b_j$ (we set $b_{-1} = b_d = 0$). For the second inequality, we compute

$$\max_{j=0,\dots,d}\{|a_j|_v\} \leq \epsilon(v) \max_{j=0,\dots,d}\{\max\{|b_{j-1}|_v, |\alpha_k b_j|_v\}\} = \epsilon(v) \max_{j=0,\dots,d-1}\{|b_j|_v\} \max\{|\alpha_k|_v, 1\}.$$

The induction hypothesis gives us

$$\max_{j=0,\dots,d}\{|a_j|_v\} \leq \epsilon(v)^{d-1} \prod_{j=1}^{d} \max\{|\alpha_k|_v, 1\}.$$

For the other inequality, first if $|\alpha_k|_v \leq \epsilon(v)$, then

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

and since f is monic, we conclude in this case. In the other case, we note first that for $v \in M_K^\infty$, $\epsilon(v) = 2$ and

$$\max_{j=0,\dots,d}\{|b_{j-1} - \alpha_k b_j|_v\} \geq (|\alpha_k|_v - 1) \max_{j=0,\dots,d}\{|b_j|_v\} \geq \epsilon(v)^{-1} |\alpha_k|_v \max_{j=0\dots,d}\{|b_j|\}.$$

So we get (for $v \in M_K^0$, this is an equality):

$$\max_{j=0,\dots,d}\{|a_j|_v\} = \max_{j=0,\dots,d}\{|b_{j-1} - \alpha_k b_j|_v\} \geq \epsilon(v)^{-1} \max\{|\alpha_k|_v, 1\} \max_{j=0,\dots,d}\{|b_j|_v\}.$$

Applying the induction hypothesis, we conclude that case. □

Theorem 3.22. *Let $C > 0$, $d \in \mathbb{N}$ constant. Then the set*

$$\{P \in \mathbb{P}^N(\bar{\mathbb{Q}}) | H(P) \leq C \text{ and } [\mathbb{Q}(P) : \mathbb{Q}] \leq d\}$$

is finite. In particular, for a number field K, the set

$$\{P \in \mathbb{P}^N(K) | H(P) \leq C\}$$

is finite.

Proof. We first reduce to a one-dimensional case. Let $P = [x_0, \ldots, x_n] \in \mathbb{P}^N(\bar{\mathbb{Q}})$ such that $H(P) \leq C$ and $[\mathbb{Q}(P) : \mathbb{Q}] \leq d$ and let $K = \mathbb{Q}(P)$. We also suppose that at least one of the $x_j = 1$. For any $i = 0, \ldots, n$, we have

$$H_K(P) = \prod_{v \in M_K} \max_{j=0,\ldots,n} \{|x_j|_v\}^{n_v} \geq \prod_{v \in M_K} \max\{|x_i|_v, 1\}^{n_v} = H_K(x_i).$$

Hence, if $H(P) \leq C$ and $[\mathbb{Q}(P) : \mathbb{Q}] \leq d$, then $H(x_i) \leq C$ and $[\mathbb{Q}(x_i) : \mathbb{Q}] \leq d$ for all $i = 0, \ldots, n$. We are reduced to proving that the following set is finite:

$$X_{C,d} := \{x \in \bar{\mathbb{Q}} | H(x) \leq C \text{ and } [\mathbb{Q}(x) : \mathbb{Q}] \leq d\}.$$

Now, let $x \in \bar{\mathbb{Q}}$ such that $H(x) \leq C$ and $[\mathbb{Q}(x) : \mathbb{Q}] \leq d$ and $f(T) = T^n + \cdots + a_1 T + a_0 \in \mathbb{Q}[T]$ its minimal polynomial with $n \leq d$. Theorems 3.19 and 3.21 tell us

$$H([a_0, \ldots, a_n]) \leq 2^{n-1} \prod_{\sigma \in G_{\mathbb{Q}(x)/\mathbb{Q}}} H(x^\sigma) \leq 2^{d-1} H(x).$$

Hence the degree and the height of the coefficients of the minimal polynomial of x are bounded. Example 3.18 tells us that there are finitely many choices for the coefficients of the polynomials, so there is a finite amount of such polynomials, hence a finite amount of such x. □

3.4 Height on elliptic curves

We want now to study the behaviour of height functions of projective space under the addition law of elliptic. As saw in theorem 3.20, they tend to be multiplicative. Therefore, we introduce the following definitions.

Definition 3.23. The **absolute logarithmic height** on projective space is given by

$$h : \mathbb{P}^N(\bar{\mathbb{Q}}) \to \mathbb{R}, \ P \mapsto \log H(P).$$

Let E/K an elliptic curve and $f \in \bar{K}(E)$ a function. If f is non-constant, it defines a surjective morphism

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

The **height on E relative to f** is

$$h_f : E(\bar{K}) \to \mathbb{R}, \ P \to h(f(P)).$$

The finiteness result 3.22 is still valid for h_f, since f gives a finite to one correspondence between $E(K)$ and $\mathbb{P}^1(K)$.

Proposition 3.24. *For any constant C, the set*

$$\{P \in E(K) | h_f(P) \leq C\}$$

is finite.

To understand the relationship between height functions and the addition law, we first prove a sort of parallelogram identity.

Theorem 3.25. *Let $f \in K(E)$ an even non-constant function, i.e. $f \circ [-1] = f$, and $P, Q \in E(\bar{K})$. Then*

$$h_f(P+Q) + h_f(P-Q) = 2h_f(P) + 2h_f(Q) + O(1),$$

where the constant in $O(1)$ only depends on E and f.

Lemma 3.26. *Let $f, g \in K(E)$ be even non-constant functions. Then*

$$\deg(g) h_f = \deg(f) h_g + O(1)$$

Proof. We take a Weierstrass equation for E:

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

First, since f and g are even, they factor through x, i.e. we can write $f = r \circ x$ for some $r \in K(X)$. Indeed, looking at addition formula [2] III.2.3, we see that if $P = (x_0, y_0) \in E$, then

$$-P = (x_0, -y_0 - a_1x_0 - a_2).$$

Using the Weierstrass equation, we can write

$$f(x, y) = f_1(x) + f_2(x)y.$$

Hence, the evenness of f implies that

$$f_1(x_0) + f_2(x_0)y = f_1(x_0) - f_2(x_0)(y_0 + a_1x_0 + a_2).$$

So $(2y_0 + a_1x_0 + a_2)f_2(x_0) = 0$ for all $(x_0, y_0) \in E$. If $y = a_1 = a_2 = 0$, then the discriminant of E is also zero, so the only possibility is that f_2 is identically zero. Now, consider theorem 3.20 for r. If we take logarithm on the inequalities, we get

$$\deg(r) h_x(P) + O(1) \leq h_f(P) \leq \deg(r) h_x(P) + O(1).$$

Hence, $h_f(P) = \deg(r) h_x(P) + O(1)$. We also know that $\deg(f) = \deg(x)\deg(r) = 2\deg(r)$. We can rewrite this

$$\frac{2h_f(P)}{\deg(f)} = h_x(P) + O(1).$$

Moreover, the same is valid for g. Combining both equalities, we conclude that

$$\frac{2h_f(P)}{\deg(f)} = \frac{2h_g(P)}{\deg(g)} + O(1).$$

□

Proof of theorem 3.25. We fix a Weierstrass equation for E/K

$$E : y^2 = x^3 + Ax + B.$$

We first reduce to the case $f = x$ for this equation. From the lemma, since the degree of x is 2, we deduce that

$$h_f = \frac{1}{2}\deg(f)h_x + O(1).$$

So the general case immediately follows from $f = x$ by multiplying the equation by $\frac{\deg(f)}{2}$. The result is also clear if $P = O$ or $Q = O$ since $h_x(O) = 0$ and $h_x(-Q) = h_x(Q)$. We write $x(P) = [x_1 : 1]$, $x(Q) = [x_2, 1]$, $x(P+Q) = [x_3 : 1]$ and $x(P-Q) = [x_4 : 1]$ (x_3 or x_4 are infinite if $P = \pm Q$). Looking at addition formulas [2] III.2.3, we get

$$x_3 + x_4 = \frac{2(x_1 + x_2)(A + x_1x_2) + 4B}{(x_1 + x_2)^2 - 4x_1x_2},$$

$$x_3x_4 = \frac{(x_1x_2 - A)^2 - 4B(x_1 + x_2)}{(x_1 + x_2)^2 - 4x_1x_2}.$$

We define a map $g : \mathbb{P}^2 \to \mathbb{P}^2$ by

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

And two other maps:

$$G : E \times E \to E \times E,\ (P, Q) \mapsto (P + Q, P - Q),$$

$$\sigma : E \times E \to \mathbb{P}^2,\ (P, Q) \mapsto (x(P), x(Q)) = ([a_1, b_1], [a_2, b_2]) \mapsto [b_1b_2, a_1b_2 + a_2b_1, a_1a_2].$$

Looking at the formulas for x_3 and x_4, we see that $\sigma \circ G = g \circ \sigma$. The idea is that t represents 1, u represents $x_1 + x_2$ and v, x_1x_2, so $g([t, u, v])$ represents $[1, x_3 + x_4, x_3x_4]$. We want to show that g is a morphism, i.e. the polynomials defining g have no common zeros except $t = u = v = 0$.

If $t = 0$, we immediately get that $u = 0$ and $v = 0$. If $t \neq 0$, there exists a solution of the equation only if $\Delta = 0$, which is excluded since E is an elliptic curve. Now, since g is a morphism of degree 2,

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

To conclude, we just need to prove that

$$h(\sigma(P, Q)) = h_x(P) + h_x(Q) + O(1).$$

If $P = O$ or $Q = O$, the result is clear. Otherwise, taking the same notations for x_1 and x_2 as before, we get

$$h(\sigma(P, Q)) = h([1, x_1 + x_2, x_1x_2]),$$

which are the coefficient of the polynomial $(T - x_1)(T - x_2)$ and

$$h_x(P) + h_x(Q) = h(x_1) + h(x_2),$$

which are the roots. Applying theorem 3.21, we see that

$$h(x_1) + h(x_2) - \log 4 \leq h([1, x_1 + x_2, x_1x_2]) \leq h(x_1) + h(x_2) + \log 2,$$

which conclude the proof. □

Corollary 3.27. *Let $P, Q \in E(\bar{K})$, $f \in K(E)$ an even non-constant function and $m \in \mathbb{Z}$. Then:*

1. $$h_f(P+Q) \leq 2h_f(P) + O(1),$$
   where the constant only depends on E, f and Q;

2. $$h_f([m]P) = m^2 h_f(P) + O(1),$$
   where the constant only depends on E, f and m.

Proof. 1. Since $h_f(P-Q) \geq 0$, the formula follows immediately.

2. Multiplication by m is a morphism of degree m^2, hence last lemma gives us
$$\deg(f)h([m]P) = m^2 h_f(P) + O(1).$$
In the proof of the lemma, we also saw that $h_f(P) = \deg(f)h(P) + O(1)$, so we conclude. □

We can now conclude the proof of the Mordell-Weil theorem, applying theorem 3.12 with any non-constant even $f \in K(E)$, like $f = x$.

4 The conjecture of Birch and Swinnerton-Dyer

Let K a number field and E an elliptic curve over K. The Mordell-Weil theorem tells us that $E(K)$ is finitely generated, i.e. $E(K)$ is isomorphic to the product of its torsion part and a free abelian group of finite rank:
$$E(K) \cong E(K)_{tors} \times \mathbb{Z}^r.$$
This group and more precisely the constant r are closely related to one of the *Millennium Prize Problems* of the *Clay Mathematical Institute*, namely the Birch and Swinnerton-Dyer conjecture. To state it, we define an L-function associated to E.

Let $v \in M_K^0$ a finite place where E has good reduction, k_v the residue field at v of cardinality $q_v := |k_v|$ and $k_{v,n}$ the unique extension of degree n. The **zeta function** of $\tilde{E}_v/k_v$ is defined as
$$Z(\tilde{E}_v/k_v, T) := \exp\left(\sum_{n=1}^{\infty} \frac{|\tilde{E}_v(k_{v,n})|}{n} T^n\right).$$

Notation. For $\psi \in End(E)$, we denote ψ_l the corresponding endomorphism on the Tate module of E (see [2], III.7 for basic definitions).

Lemma 4.1. *Let $\psi \in End(E)$. Then*
$$\det(\psi_l) = \deg(\psi)$$
and
$$tr(\psi_l) = 1 + \det(\psi) - \det(1-\psi).$$
Moreover, if ψ is separable, then
$$\deg(\psi) = |\ker(\psi)|$$
.

Proof. See [2], chapter V, proposition 2.3 and chapter III, theorem 4.10(c). □

Proposition 4.2. *$Z(\tilde{E}_v/k_v, T)$ is a rational function, more precisely*

$$Z(\tilde{E}/k_v, T) = \frac{L_v(T)}{(1-T)(1-q_vT)},$$

with $L_v(T) = 1 - a_vT + q_vT^2$, where $a_v = q_v + 1 - |\tilde{E}_v(k_v)|$. Moreover, it satisfies the functional equation

$$Z(\tilde{E}_v/k_v, \frac{1}{q_vT}) = Z(\tilde{E}_v/k_v, T)$$

and the roots of $L_v(T)$ have both modulus $\sqrt{q}$.

Proof. Let

$$\phi_v : \tilde{E}_v \to \tilde{E}_v,\ (x, y) \mapsto (x^{q_v}, y^{q_v})$$

the Frobenius automorphism at v, which has degree q_v. Since the Galois group $G_{\bar{k}_v/k_v}$ is topologically generated by ϕ_v, we have that $P \in E_v(k_v)$ if and only if $\phi_v(P) = P$. Hence, since $1 - \phi_v$ is separable (see [2], chapter III, corollary 5.5),

$$|\tilde{E}_v(k_v)| = |\ker(1-\phi_v)| = \deg(1-\phi_v).$$

Moreover, lemma 4.1 tells us that

$$\det(\phi_{v,l}) = \deg(\phi_v) = q_v,$$

$$\operatorname{tr}(\phi_{v,l}) = 1 + \deg(\phi_v) - \deg(1-\phi_v) = 1 + q_v - |\tilde{E}_v(k_v)| = a_v.$$

Hence, the characteristic polynomial of $\phi_{v,l}$ is

$$\det(T - \phi_{v,l}) = T^2 - \operatorname{tr}(\phi_{v,l})T + \det(\phi_{v,l}) = T^2 - a_vT + q.$$

Factorising over $\mathbb{C}$, we denote

$$(T-\alpha_v)(T-\beta_v) = \det(T - \phi_{v,l}).$$

For every rational number $\frac{m}{n} \in \mathbb{Q}$, we have

$$\det\left(\frac{m}{n} - \phi_{v,l}\right) = \frac{\det(m - n\phi_{v,l})}{n^2} = \frac{\deg(m - n\phi_v)}{n^2} \geq 0.$$

So the characteristic polynomial is always non-negative on $\mathbb{R}$. Thus, it has a double real root or two complex conjugate roots. In either case, they have the same modulus and $|\alpha_v| = |\beta_v| = \sqrt{q_v}$ since the constant term of $\det(T - \phi_{l,v})$ is q_v. Similarly,

$$|\tilde{E}_v(k_{v,n})| = \deg(1 - \phi_v^n)$$

and

$$\det(T - \phi_{v,l}^n) = (T - \alpha_v^n)(T - \beta_v^n),$$

which is immediate if we put $\phi_{v,l}$ in Jordan normal form, since its diagonal will have the two values α and β. In particular,

$$|\tilde{E}_v(k_{v,n})| = \deg(1 - \phi_v^n) = \det(1 - \phi_{v,l}^n) = 1 - \alpha_v^n - \beta_v^n + q_v^n.$$

We conclude that

$$\log(Z(\tilde{E}_v/k_v,T)) = \sum_{n=1}^{\infty} \frac{|\tilde{E}_v(k_{v.n})|}{n} T^n = \sum_{n=1}^{\infty} \frac{1-\alpha_v^n-\beta_v^n+q_v^n}{n} T^n$$

$$= -\log(1-T) + \log(1-\alpha_v T) + \log(1-\beta_v T) - \log(1-q_v T).$$

Hence, we have the rational form of the zeta function. Moreover, using $\alpha_v\beta_v = q_v$, we compute

$$Z(\tilde{E}_v/k_v, \frac{1}{q_v T}) = \frac{(1-\frac{1}{\beta_v T})(1-\frac{1}{\alpha_v T})}{(1-\frac{1}{q_v T})(1-\frac{1}{T})} = \frac{(\beta_v T-1)(\alpha_v T-1)}{(q_v T-1)(T-1)} = Z(\tilde{E}_v/k_v, T)$$

□

If E has bad reduction at v, we define

$$L_v(T) = \begin{cases} 1-T & \text{if } E \text{ has split multiplicative reduction at } v, \\ 1+T & \text{if } E \text{ has nonsplit multiplicative reduction at } v, \\ 1 & \text{if } E \text{ has additive reduction at } v. \end{cases}$$

So in all cases, we have ([2], VII.5.1)

$$L_v(q_v^{-1}) = \frac{|\tilde{E}_{ns}(k_v)|}{q_v}.$$

The L-**series** of E/K is then defined as the inverse product of these functions at all places $v \in M_K^0$:

$$L(E/K,s) := \prod_{v\in M_K^0} L_v(q_v^{-s})^{-1}.$$

For almost all v, its factor in the product has modulus

$$|L_v(q_v^{-s})|^{-1} = \frac{1}{|1-\alpha_v q_v^{-s}|}\frac{1}{|1-\beta_v q_v^{-s}|} = \left|\sum_{n=0}^{\infty}(\alpha_v q_v^{-s})^n\right| \left|\sum_{n=0}^{\infty}(\beta_v q_v^{-s})^n\right| \leq \left(\sum_{n=0}^{\infty}\left(q_v^{-s+\frac{1}{2}}\right)^n\right)^2 = \left(1-q_v^{-s+\frac{1}{2}}\right)^{-2}$$

So we have (omitting the finite number of v with bad reduction)

$$|L(E/K,s)|^{-2} \leq \prod_{v\in M_K^0}(1-q_v^{-s+\frac{1}{2}}).$$

This Euler product converges if and only if the sum $\sum_{v\in M_K^0} q_v^{-s+\frac{1}{2}}$ converges. Since over each prime number, there is a bounded number of prime ideals (corresponding to discrete valuation), the later sum can be bounded as follow:

$$\sum_{v\in M_K^0} q_v^{-s+\frac{1}{2}} \leq \sum_{p \text{ prime}} [K:\mathbb{Q}]p^{-s+\frac{1}{2}} \leq [K:\mathbb{Q}]\sum_{n=1}^{\infty} n^{-s+\frac{1}{2}}.$$

So in particular, it converges to an holomorphic function for all $Re(s) > \frac{3}{2}$.

Conjecture 4.3. *$L(E/K, s)$ has an analytic continuation to $\mathbb{C}$ and satisfies a functional equation relating the values at s and $2 - s$.*

This has been only proved in some cases. In particular, generalisations of the work of Wiles on Fermat Great Theorem implies that the conjecture is true for $K = \mathbb{Q}$.

We can now state the weak form of the Birch and Swinnerton-Dyer conjecture.

Conjecture 4.4 (Birch and Swinnerton-Dyer)**.** *The degree of annulation of $L(E/\mathbb{Q}, s)$ at 1 is equal to r.*

There exists a more refined version specifying the first non-zero term of the Taylor series of $L(E/\mathbb{Q}, s)$ at 1 which involve various invariant of the curve E, like the number of torsion element of $E(\mathbb{Q})$ or the size of the Shafarevich-Tate group which is not even known to be finite. This conjecture and the refined version can be seen as rather surprising, considering the fail of the Hasse principle for equations of degree 3.

A lot of evidences supporting the conjecture have been gathered. Birch and Swinnerton-Dyer originally formulated it after having explored numerically some curves (see [4] and [5]). It was then numerically checked in a lot of other cases. Coates and Wiles proved that if $E(\mathbb{Q})$ is infinite and $E(\mathbb{Q})$ has complex multiplication, then $L(E/\mathbb{Q}, 1) = 0$. Reciprocally, Gross and Zagier proved that if $L(/\mathbb{Q}, s)$ has a zero of rank one, then $E(\mathbb{Q})$ has a point of infinite order. A lot of other evidences and a survey of the conjecture is presented in [2], appendix C.16.

References

[1] Ernst S. Selmer. The diophantine equation ax 3 + by 3 + cz 3 =0. *Acta Math.*, 85:203–362, 1951.

[2] Joseph H. Silverman. *The arithmetic of elliptic curves*, volume 106 of *Graduate Texts in Mathematics*. Springer, Dordrecht, second edition, 2009.

[3] Jürgen Neukirch. *Algebraic number theory*, volume 322 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1999. Translated from the 1992 German original and with a note by Norbert Schappacher, With a foreword by G. Harder.

[4] B. J. Birch and H. P. F. Swinnerton-Dyer. Notes on elliptic curves. I. *J. Reine Angew. Math.*, 212:7–25, 1963.

[5] B. J. Birch and H. P. F. Swinnerton-Dyer. Notes on elliptic curves. II. *J. Reine Angew. Math.*, 218:79–108, 1965.